PROCEEDINGS —

Twentieth Annual Meeting

Theme:

"Transportation Alternatives in A Changing Environment"

October 29-30-31, 1979
Drake Hotel
Chicago, Illinois

Volume XX • Number 1 1979

T R F

TRANSPORTATION RESEARCH FORUM

The Efficient Financing of Highways

*by Douglass Lee**

INTRODUCTION

ALTHOUGH efficiency has not explicitly guided highway pricing and investment in the past, there seems to be some interest in viewing transportation systems as economic enterprises rather than as public services. Perhaps this is because the direct benefits of transportation are largely or entirely captured by the users of transportation facilities, and the allocation of resources in transportation seems similar to resource allocation problems in other sectors of the economy.[1] This paper attempts to set forth the principles of allocative efficiency as they might be applied to pricing and investment in the highway system.

In seeking to illuminate the public enterprise perspective on highways, we will begin by breaking the discussion into two major subject areas. The first of these is directed at questions concerning the utilization of the existing system, emphasizing the short run. The existence of the present highway system will be taken as given. In the second subject area, the long run status of the highway enterprise will be evaluated, to assess how well the system is performing as a major capital investment. These two subject areas are interrelated in important ways, so the final section will integrate the short and long run perspectives and offer guidelines for efficient highway financing.

UTILIZATION OF THE EXISTING HIGHWAY SYSTEM

In the short run, the rational enterprise will charge a price to consumers that is at least sufficient to cover the immediate costs of providing the service; otherwise, the firm is losing money on each customer it serves. Given its fixed capital facilities, the firm will seek to get as much revenue from using them as it can, subject to competitive pressures that prevent the firm (ideally) from exploiting any monopoly power. It will stay in business as long as it can cover out-of-pocket costs in the short run, but it is also trying to earn enough to pay for the replacement or expansion of its capital facilities over the course of the lifetimes of the fixed assets.

Theory of Short Run Pricing

From the standpoint of social efficiency, a public enterprise should follow the above principles where they are suitable. Referring to the diagram in Figure 1, average variable costs of additional vehicle trips on a given highway link are represented by the AVC curve. Above some level of traffic, marginal costs continue to increase because the fixed capacity of the highway leads to increasing costs resulting from congestion. This curve can be used to derive a marginal cost curve, showing the additional cost of each vehicle added to the traffic stream. Eventually, of course, no more vehicles can be accommodated and the marginal cost becomes infinite.

It is marginal cost which must be covered by the firm in the short run. If the marginal social benefits of each trip are reflected in the demand curve, then q is the optimal number of trips and p is the correct unit price. For trip volumes above q the additional costs are greater than the additional benefits (the MC curve lies above the MSB curve), while for volumes below q the benefits of additional trips outweigh the costs.

Type of Costs

For evaluating short run utilization, only variable costs are relevant, i.e., only those costs that vary directly with the volume of traffic should be considered in comparing short run costs and short run revenues.

OPTIMUM SHORT RUN UTILIZATION AND PRICE

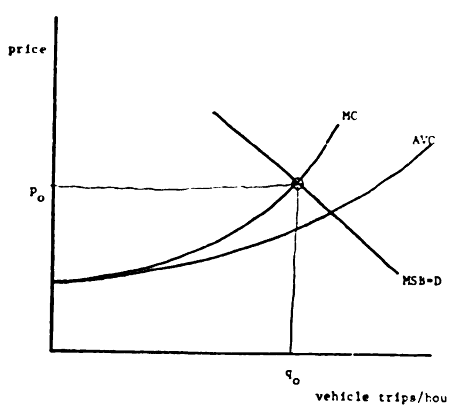

FIGURE 1

*Transportation System Center.

Maintenance and Administration of the Highway System. Not all maintenance costs result from traffic, since weather and soil conditions can result in pavement deterioration independently of the volume of traffic. For the variable portion of maintenance cost, vehicle mix will have a major impact: heavy vehicles do many times the damage of light vehicles, per vehicle mile, and the difference is magnified on light duty roads compared to high-type construction. Administration includes traffic control, spot vehicle inspections, accident liability adjudication, and similar efforts that result directly from the amount of traffic on the highways.

Vehicle Operation. These costs are borne entirely by users of the highway system. The capital cost of the vehicle itself can be thought of as having two components: one—wear—represents the amount of travel the vehicle will perform (e.g., total lifetime mileage), and the other—opportunity cost—represents the time value of the sources tied up in the vehicle and therefore not available for other uses (e.g., the interest on the remaining value of the vehicle).

Other variable costs include tires, parts, accessories, insurance, accidents, fuel, oil, parking (excluding storage at place of residence or other primary location), and repair labor. Some of these categories may overlap, e.g., insurance covers part of the cost of accidents. User charges, such as gasoline taxes, should not be included.

Non-Monetary User Costs. In addition to the out-of-pocket costs described above, highway users also must supply time and labor in order to use the system. This component of costs includes a basic travel time and its associated value, plus additional factors for delay, congestion, interference from other vehicles, aggravation, fatigue, errors, accidents, and so forth, that result from the interaction between the users of the system and the quality (including capacity) of the highways. While measured delay is a plausible surrogate or functional argument for many of these costs, the relationship between units of delay and units of value is not well understood. Markets in which these values might be estimated are complicated, and most research to date has not sought to relate the micro-characteristics of the traffic stream to the social costs of vehicle interaction and delay.

Negative Externalities. Because highway users degrade the natural and human environment in the vicinity of the highway, costs are not limited to the highway and its users. Air pollutants for which highways are an important source include carbon monoxide, oxides of nitrogen, lead, asbestos, and suspended particulate. Water quality is affected by these pollutants, as well as by petroleum distillates. Neighborhood amenities are adversely affected by noise, danger, visual intrusion, and barriers to social interaction and physical movement, among other things. Some effort has been directed at estimating values for these impacts, but they are hard to quantify.

THE HIGHWAY SYSTEM AS AN INVESTMENT

Besides making sure that resources are allocated efficiently in the short run and that the stock of capital available is optimally utilized, the highway enterprise should be checking to determine whether previous investment has proved to be worthwhile and whether or not additional investment is justified. This requires a long run perspective, in which all costs are variable and the major question is whether revenues are sufficient to cover long run costs. If short run prices are correctly set, and significant economies or diseconomies of scale are not present, a revenue surplus suggests that expansion is warranted, while a revenue shortfall implies that the system should be contracted.

Theory of Long Run Equilibrium

In the short run, price is variable and the capital stock is fixed; price is set to achieve optimum utilization of the fixed resources. In the long run, the level of investment is variable and is set so as to generate the optimum plant size. Long run and short run are in joint equilibrium when price equals both LRMC and SRMC. This is shown in Figure 2 under constant returns to scale, with the marginal benefits curve labelled D.

Long run benefits represented by D would result in a disequilibrium for the curves shown. The theoretically correct response is to price at SRMC (reducing consumption to q), and disinvest until the SRMC and SRAC curves have shifted far enough to the left to gain the equilibrium described above.[2] During this shift, price will move up the D curve until it equals LRMC, and output will drop to q.

Decreasing Long Run Costs. The disequilibrium case from Figure 2 is repeated in Figure 3, with the change of increasing instead of constant returns to scale. The cost curves imply a long run output of q, while equilibrium requires an output of q. Long run price would be announced as equaling LRMC at q, with the transitional short run prices set in accordance with the rate of disin-

LONG RUN EQUILIBRIUM AND DISEQUILIBRIUM UNDER CRTS

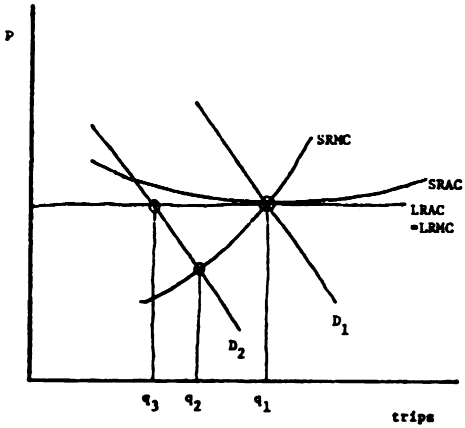

FIGURE 2

vestment. The difference between the resulting equilibrium and the one under constant returns to scale is that price will not cover long run costs (the difference between LRAC and LRMC at q); the shortfall must be made up from general revenues, multipart tariffs that do not affect long run demand, or deviations from short run marginal costs that have the least effect on efficiency. Since total efficiency urges that the consumers of the service fully bear the costs in the long run, some mechanism which places the burden on users is preferable.[3]

Alternative Design Characteristics. Up to this point it has been implicitly assumed that output could be measured by a single homogeneous variable. Because the attributes of almost any good or service can be richly varied, the demand and cost curves illustrated above imply that the mix of characteristics has already been optimized and is known. The particulars of this process bear some study.

Suppose we look at the optimization of a single design attribute such as surface quality, using Figure 4. The unit of output is an increment in surface quality, normalized to a trip volume; the horizontal axis does not measure trip volume either directly or indirectly, but assumes the volume is given. Demand reflects the willingness-to-pay for additional quality by users, on a per-trip basis. In this case, demand may be derived from fuel and wear savings, travel time savings, and improved comfort, but demand for surface quality is separate from the demand for trips. The supply curve reflects the cost of providing various levels of quality, on a per-trip basis for a given trip volume. Because some of these costs are fixed, the location of the cost curve (up or down) is greatly affected by the number of trips upon which it is based. The increasing steepness of the cost curve is meant to suggest that maintenance of higher levels of surface quality is increasingly costly. No congestion or other user costs are included in the curve, and no short run curves are pertinent because there are no variable costs. Suppliers of surface quality are highway contractors and highway department maintenance crews, while demand must be inferred or constructed indirectly from user benefits.

Multiple User Classes. A major complexity is introduced by allowing for more than one class of users. Two fundamental questions are (1) How are design compromises arrived at when user classes have different preferences for design attributes; and (2) How should classes be charged for attributes that may be valued by only a subset of classes? For example, how thick should pavements be for heavy trucks, and how much should they be charged in comparison to autos?

Going back over the previous discussion of long run equilibrium, the existence of multiple user classes does not affect the results as long as demand can be aggregated into a single function and output can be measured along a single dimension. The common unit for both is the standardized vehicle trip or passenger car equivalent (PCE). As long as truck costs and truck demand can be stated in PCE units, the same analysis

LONG RUN EQUILIBRIUM UNDER IRTS

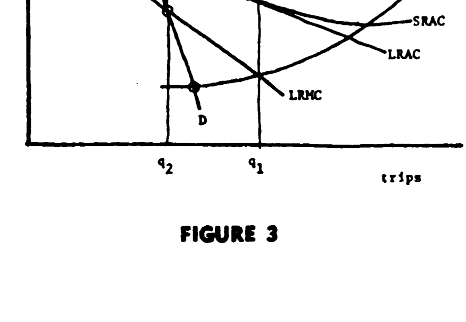

FIGURE 3

EQUILIBRIUM FOR SINGLE ATTRIBUTE

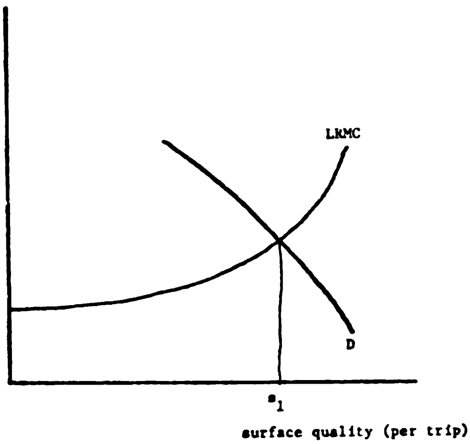

FIGURE 4

holds throughout and the horizontal axis is PCE vehicle trips per unit of time.

For example, take the case of two user classes, one called autos and the other called trucks. Each class presents its demand schedule, and the two are aggregated horizontally: at a given price, the quantity demanded equals the number of auto vehicle trips demanded at the price plus the number of truck trips demanded. The number of truck trips must be stated in terms of PCE's to be consistent with the output unit for auto trip demand. The combined demand serves as the single demand schedule shown in Figures 1-3.

This demand aggregation assumes that attributes such as pavement quality have already been determined exogenously, through a mechanism such as that shown in Figure 4. For this purpose, demand must be aggregated vertically: the willingness-to-pay for improved surface quality on the part of the auto user class is added to the willingness-to-pay for the same attribute on the part of the truck user class, and the combined demand resolved against the supply curve as in Figure 4.

Demand aggregation for two attributes —surface quality and pavement depth— is shown in Figure 5. Operationally, the incremental costs and benefits of higher quality levels would be compared until the maximum net benefit value of the attribute were identified. Benefits would be estimated by adding up user savings in vehicle wear, time, fuel, and perhaps driver fatigue and the like.

We have addressed the question of long run versus short run marginal cost pricing with respect to output of travel, but we have not dealt with the variable attributes of highway quality as to their pricing. From Figures 4 and 5, it might be inferred that both the level of output (the correct surface quality, pavement depth, etc.) as well as the price per trip had been determined, but this is incorrect; the prices are long run prices and not relevant to user charges. The fact that light vehicles would not be willing to clear the market for pavement depth at a price of zero (in Figure 5) does not mean that light vehicles should be absolved from paying for pavement depth.

To see this, we should scrutinize carefully the implications of the long run equilibrium condition that equates short run marginal cost with long run marginal cost. When long run average costs equal short run average costs, the components are identical because long run average costs are nothing more than a selection of short run average costs; they are one and the same. With marginal costs, however, the equilibrium condition requires that short run rents exactly cover the long run marginal costs of capacity. Figure 6 illustrates the costs that are repeated in both short and long run MC, as well as the distinct elements. Vehicle wear, fuel, travel time, accidents, etc., under conditions of long run optimization will be the same for both short and long run. The same is true of pavement wear and other variable cost components. On the short run side, however, the deviation between marginal and average variable costs of delay, fuel costs, wear, etc., from congestion have no counterpart on the long run side, while the marginal capacity costs for lane width, pavement depth, gradient reduction, right-of-way, base and subbase, engineering, landscaping and so on have no counterparts on the short run side. An equilibrium is obtained when these two sets of cost components total to the same amount. Short run efficiency is served by setting prices equal to short run marginal costs, while long run efficiency is achieved by adjusting investment until the short run price equals the long run marginal cost. These principles were presented at the beginning of this section, and they still apply.

Correct pricing, then, ignores the incremental fixed costs of differentially serving various user classes. Instead, only the costs which vary with volume of usage in the short run should be assigned to user classes; costs associated with design attributes that are fixed in the short run are relevant to design and investment decisions, but only indirectly to pricing. No direct connection should

DEMAND AGGREGATION FOR QUALITY ATTRIBUTES

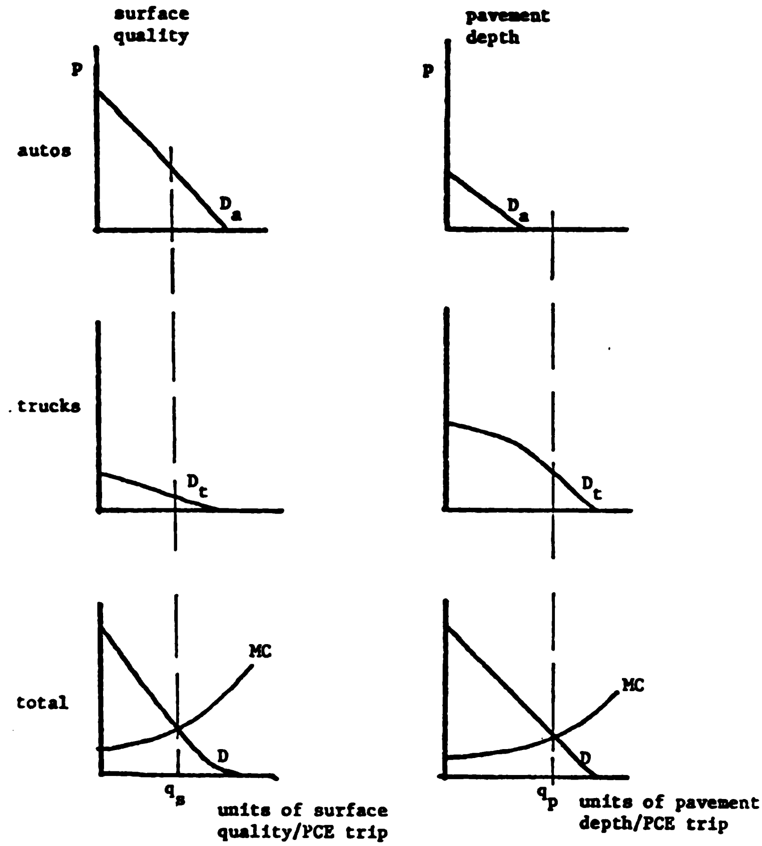

FIGURE 5

be made, for example, between the extra lane width or pavement thickness suitable for heavy vehicles and the prices (net of variable costs, e.g., pavement damage) charged to those vehicles. Once the decision is made to provide a heavy duty facility, any vehicle which uses the facility appropriates the full lane width and pavement depth at the time of usage. The extra pavement depth cannot be reassigned to another user just because a light vehicle passing over it does not need the extra depth.

Types of Costs

For the long run, all of the costs listed as variable in the short run plus all those costs that are fixed in the short run are included. Only the additional ones are described below.

Construction of the Highway System. Right-of-way (ROW), grading, drainage, base and pavement construction, shoulders, landscaping, and administration are components of the capital costs of highways.

Fixed Maintenance and Administration. Some costs of highway maintenance and administration must be incurred whether or not any traffic uses the road, and some are necessary to keep the road in operation but are unrelated to vehicle volumes. Traffic signal maintenance and operation, street lighting, vehicle and driver registration, and general administration of state and Federal departments of transportation are some examples.

General Overhead Obligation. A cost item almost universally overlooked but one which is real to railroads and other

COMPARISON OF SHORT AND LONG RUN MARGINAL COST COMPONENTS

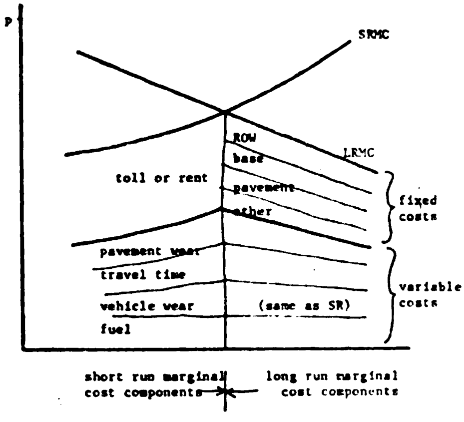

FIGURE 6

private firms is the cost of providing government services. For private firms, this obligation is met through payment of property, sales, and corporate income taxes. Following through on the private enterprise analogy implicit in the public enterprise perspective, an equitable share of general government cost should be assigned to the highway system. This suggests that the road enterprise should collect from road users, in the same way that commercial enterprises collect from their customers, enough revenues to cover property, sales, and income tax payments as well as capital and operating costs.

HIGHWAY EFFICIENCY ANALYSIS AND POLICY ISSUES

On the short run side, the empirical question of whether or not marginal costs are matched by variable user charges is obscure because of both the paucity of information about marginal cost and the crude relationship between existing user charges and variable costs. On the long run side, most of the cost components have never been estimated. Nonetheless, a few limited conclusions can be proffered.

(a) The primary instrument (the fuel tax raises about two-thirds of all user revenues) does not distinguish well between the pavement damage done by heavy vehicles versus light ones, or between underutilized roads and heavily congested ones.

(b) Given a highway system that appears to have a generous supply of capacity in most areas, the marginal costs of travel on newer heavy duty roads is probably low.

(c) To the extent that congestion and interference are important costs, the difference between average cost and marginal cost is ignored because there is no user charge instrument in place that even approximately reflects this cost.

(d) Negative externalities are difficult to put prices on even crudely, but they appear to be large in the aggregate and substantial in some areas. Only the fuel tax relates in any way at all to these impacts, and the relationship is so poor that the negative externalities are treated by highway users as essentially free.

(e) The damage done by heavy vehicles to light duty roads is very large, relative to the cost of the road, and no user charge responds to this.

(f) National expenditure and revenue accounts show that at least thirty percent of total expenditures are derived from non-user revenue sources. Since expenditures are undoubtedly below long run costs, revenues clearly fall short of full cost by some large but unknown amount.

(g) Several costs of significant magnitude have not been included (capital replacement value at current prices, general overhead obligation) on the total cost side, while revenue sources recognized as user charges are declining (excise taxes fixed in dollar terms while costs inflate, excise taxes preempting general sales taxes), hinting that the apparent revenue shortfall is further biased toward the low side.

Because the theory presented above is highly distilled and hence abstract, it is difficult to relate actual circumstances directly to the theory. An attempt has been made to do this in Figure 7, but it should be regarded as illustrative of plausible hypotheses rather than as based on current data. As already indicated, the data have not been collected. What the diagram shows, however, is that (a) average private costs (APC) are generally below both marginal and average social costs; (b) highways users base their decision about highway consumption on APC; and (c) average total costs are much higher than average user costs. While there may be portions of the highway system for which user charges exceed short run marginal cost,

APPROXIMATE CURRENT EQUILIBRIUM

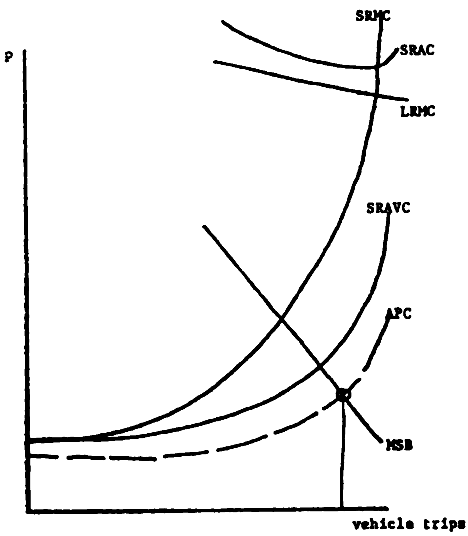

FIGURE 7

it is likely that these segments are the exception rather than the rule. Overall, then, short run user charges are at least misapplied if not too low in the aggregate, and total costs are far from matched by total revenues.

The conclusion, then, is that analysis of the long run efficiency of the highway system should concern itself with

(a) the direct attribution of variable costs (not fixed costs, but only those costs which vary with traffic volume in the short run) to user classes;

(b) the degree to which the scale of the highway system is in long run equilibrium; and

(c) the optimality of design attributes such as maintenance standards, pavement thickness, lane widths, and the like.

Such analysis requires information pertaining to the full long run costs of the national highway system (not simply current expenditures) as well as volume-related costs such as vehicle interference, pavement damage, and externalities. Many pieces of the analysis are already well developed, but there are critical gaps and much effort is wasted in tangential endeavors. Both the research side as well as the policy side would be vastly improved by adopting an efficiency framework along the lines described above, rather than searching for "an equitable distribution of the tax burden"[4] as a solution to the problems of highway financing.

FOOTNOTES

1 See, for example, Herbert Mohring, Transportation Economics, Cambridge, Mass: Ballinger, 1976.

2 The argument presented here parallels the discussion in such references as Richard H. Leftwich, The Price System and Resource Allocation, 7th ed., Hinsdale, Illinois: Holt, Rinehart, and Winston, 1979, with the exception that the distinction between the firm and the industry is suppressed.

3 There is strong empirical and theoretical evidence that the long run average cost curve is either flat or declining for most industries throughout the range of possible interest (see A. Koutsoyiannis, Modern Microeconomics, New York: Johy Wiley, 1975, Chapter 4). General Motors represents an example (Fortune, May 7, 1979), in that scale economies for very large units of production still outweigh management and other potential diseconomies. The significance of this is due to the fact that despite pervasive decreasing costs and monopoly or oligopoly market strength, private firms forced to cover full costs out of revenues are reasonably if not highly efficient. The presence of decreasing costs provides very little real justification for public subsidy.

4 Section 506, Surface Transportation Act of 978, U.S. Congress, in requesting a new highway cost allocation study.