PROCEEDINGS —

Twenty-third Annual Meeting

Volume XXIII • Number 1 1982

T R F

TRANSPORTATION RESEARCH FORUM

PROCEEDINGS —
Twenty-third Annual Meeting

Theme:

"Developing Concinnity in Transportation"

October 28-30, 1982
Fairmont Hotel
New Orleans, LA

Volume XXIII • Number 1 1982

TRANSPORTATION RESEARCH FORUM

Toward A Theory of Rational Road Pricing

*by John R. G. Brander**

THIS PAPER explores one aspect of a broader, on-going research effort focusing on the questions of the pricing of and investment in transportation infrastructure. In the broader context, the questions of user pay, congestion pricing and investment expenditure are all considered. Research to date indicates that the three questions must be answered simultaneously, and prior to the installation of the facilities. The objective of the present paper is much more limited, focusing only on the issue of user pay. The objective is to develop a theoretical framework for the user pay component which can be integrated into the broader analysis.

The first part of the paper deals with the overall analytical framework, and discusses some differences between the present approach and that in the received literature. The following section is devoted to a consideration of the problem of excess capacity in the highway system. With this background in place, attention is turned to the question of nonuser cost responsibility for infrastructure costs. It will be argued that the separation of cost responsibility between users and nonusers must be the first step in the analysis. The next section of the paper deals with the allocation of the remaining costs among classes of user. With this issue resolved, the actual mechanism for determining the level of user charges is explored. A final section of the paper presents some conclusions and implications of the analysis.

BACKGROUND TO THE DISCUSSION

In his writings on the theory of subjective value, Carl Menger distinguished goods which could be consumed directly from those which could be consumed only indirectly. The former he regarded as first order goods. The latter, which in reality are inputs into the production of the former, included several categories of higher order goods.[1] Menger's framework provides a useful schema for considering the whole question of highway pricing. In this instance, the directly consumable first order goods would be the actual transportation. Its production, in turn, requires several higher order goods as inputs. On the capital input side, the most important of these would be the infrastructure and some type of vehicle. Current inputs would include, among other things, labour, fuel and time. Each of these inputs has a price, and the total cost of transportation is a function of the prices and the amounts of each input required to produce the final output.

Menger's analytical structure demonstrates the need to deal with the question of infrastructure pricing separately from the other elements of transportation cost. Where the infrastructure itself is concerned, this implies a two part pricing policy. The first component would be the user fee. It must be set in such a fashion as to cover the costs of owning and operating the infrastructure. The other component is a congestion toll which must be both time and location specific. The latter, of course, would be designed to ration available capacity. It also suggests that it is necessary to carry out the analysis at the level of individual vehicle classes, ideally by class of road.

The objective of full user pay as an element of long-run transportation policy is generally accepted. In the interests of optimum resource allocation, users of a transportation facility should pay their "fair" share of the costs incurred in its provision. Were they to pay less than this, users would be receiving a subsidy. Were they to pay more, general taxation would be carried out in the guise of user charges. Either would result in a misallocation of resources through a misallocation of traffic among competing modes. This much is generally accepted. Controversy arises when attempts are made to translate the objective into an operational framework. The intermingling of user pay and congestion pricing, the use of aggregate cost and revenue comparisons, and the failure to properly address the nature of the debate at hand are all sources of disagreement.

Despite the research efforts of the past decade, there remains in the literature some uncertainty about the nature of the infrastructure input. The pay-as-you-go approach treats highway capital investment as a current expenditure, implying that user charge revenues in

**Professor and Chairman, Department of Economics, University of New Brunswick.*

any year should be sufficient to cover these capital costs as well as current operating and maintenance outlays. The alternative is to place the highway system on commercial principles and thus view any year's investment as generating a stream of economic costs in the future.[2] The Menger-type framework noted above follows the latter approach, an approach which is on sounder footing from the point of view of economic theory. This is the approach adopted in this paper.

A second difficulty is the treatment of the time variable in the analysis. Once the pay-as-you-go approach has been rejected in favor of the economic analysis approach, it is necessary to decide upon the time framework in which the analysis will be carried out. In most of the studies of infrastructure pricing, time is included only in historical perspective.[3] Generally, the analyses end with the present (or the very recent past). Admittedly, this is useful in explaining how any given system evolved. It does not, however, provide any indication as to where it might be going. A complete understanding of the forces at work cannot be had without some reference to the future as well.

Most serious, from the theoretical perspective at least, is the potential incompatibility between pricing decisions and investment analysis in this time framework. In an ideal world, the pricing and investment decisions for a particular highway would be determined simultaneously. This is the case because the price determined usage which, in turn, is one of the elements defining the magnitude of benefits arising from the project. Clearly where the price determined ex post is different from that used in the investment analysis, the project's benefits will differ as well. This suggests that the price used in the investment analysis must be the one which will actually be charged, and that future traffic volumes and hence benefit streams must be estimated on the basis of this price. At the level of the individual project, it is clear that future time must be included. Obviously, the same thing holds at the level of the highway system. In the case of the latter, it is also necessary to incorporate historical time into the analysis as well, for the numerous investment decisions made in the past have, in total, resulted in the creation of the present highway system. The investigator must, therefore, stand in the present and look into both the past and the future in carrying out the pricing analysis.

This perspective of time raises serious questions about the possibility of engaging in a retrospective analysis to determine the appropriate level of user charges by class of vehicle. It is admitted that one or more allocative mechanisms could be found to apportion costs (however defined) among the various classes of vehicle using a highway system. Railroads, after all, have been doing this sort of analysis for decades among their various traffic elements. Whether this results in anything more than an arbitrary, albeit consistent, allocation of cost is another matter. Average, rather than specific costs emerge from this analysis, but whether these averages represent the "true" costs is another question. This question will be returned to later in the paper.

On balance, while the objective of full user pay is accepted, it is clear that there are problems with earlier approaches to the analysis which must be overcome. In the present paper, an attempt is made to overcome these difficulties using a different perspective of time. Before this is done, however, there is one other matter deserving of attention. That is the question of excess capacity.

THE PROBLEM OF EXCESS CAPACITY

One problem which has not been given sufficient attention in the user pay debate is the issue of excess capacity and its impact on the cost structure of the highway infrastructure. While some attention has been given to pricing schemes applicable to lightly used transportation infrastructure,[4] the value of service prices which result are inconsistent with the desired fiscal neutrality that is a part of the user pay debate. As the first steps in finding an alternative, it is necessary to inquire as to the causes of this excess capacity and its impacts on highway costs.

Three sets of factors interact to produce considerable amounts of excess capacity in the highway network. The first of these is the transportation planning process itself. The second factor is the lumpiness of highway investments, a difficulty which creates indivisabilities on the cost side. The third factor results from the fact there are a variety of nontransportation factors involved both in the decision to construct and the design standard of particular highways. Each aspect must be discussed briefly.

The transportation planning process is concerned almost exclusively with supply side adjustments to capacity problems. In practice, at least, the manipulation of demand through the price mechanism is almost unknown. Faced with a congested highway, the typical highway planner will recommend an expansion of the existing facility of the construction of a new and larger one. The capacity of this larger facility is determined by the notion of a "design hourly volume." The design hour is the thirtieth highest hourly volume in the design year. The design year is typically twenty years away. In the present context, it might be useful to turn the issue around. If the traffic forecasts are completely accurate, one would expect the design hourly volume to be exceeded for only 29 hours in the design year. This is clearly an exaggeration, but it does emphasize the point that even the busiest highways are busy only part of the time. The balance of the time finds them in an excess capacity situation. Given its nature, it should be clear that the transportation planning process results in the installation of a considerable amount of excess capacity in the highway network. What needs to be recognized is that this cannot be done without imposing severe ownership penalties on the operators of the infrastructure.

This difficulty is compounded by a second problem, the lumpiness of highway investment alternatives. Under certain circumstances, at least, the lumpiness also contributes to the existence of excess capacity. It causes a second difficulty as well, namely discontinuities in the cost structure. Of equal importance is the fact that the lumpiness tends to bias cost-revenue comparisons. Large systems will tend to produce better financial results simply because they are larger. Put another way, the burdens of excess capacity weigh more heavily upon small systems than on large ones. Account must be taken of these factors as well in the setting of the level of user charges.

Much of the received literature assumes implicitly that only user demands were considered in the decision to construct a particular highway. Were this the case, and were the two problems discussed above adequately incorporated into the analysis, the issues at hand might well be resolved quickly. This implicit assumption is, however, incorrect. A variety of non-used demands have historically had an impact on highway investment, and as a result, the system is not a user-optimized one. Some of these demands are political rather than economic in origin. A case in point would be the desire for a uniform standard on a particular class of highway regardless of the traffic volumes. The minimum design standards of the Interstate Highway System would be an example. A variety of other non-user demands have had an impact as well. Regional economic development motives, the desire for social integration of isolated areas, and even military preparedness all have had roles to play. All of these motives create cost streams in the future, and increase the amount of excess capacity in the present. Given these impacts, the failure to incorporate them into the analysis in a full user pay system amounts to the use of that system as an instrument of general taxation. Accordingly, they must be incorporated into the analysis.

USER PAY IN GENERAL

The difficulties discussed above can best be overcome by remembering that there are three closely related questions in the debate, two related to the question of allocating cost, the other relating to the determination of the user fee to be charged. Only by considering the non-user cost responsibility first, then focusing on the cost responsibility of various classes of user, and finally setting the price for the use of the facilities is it possible to avoid the pitfalls of earlier approaches. This section of the paper presents a general discussion of these topics. Later sections will consider them in greater detail.

Highways, roads and streets are multipurpose capital assets, built and maintained to serve a variety of functions. Only one of these, actual vehicle usage of the facilities, is transportation related. It follows, therefore, that there are non-transportation considerations involved in the provision of this capital stock, and to these, there are costs attached. A cost separation between the two elements is obviously necessary, and this separation must be the first step in the analysis. While this seems straightforward, it is not a subject which has received any consistent treatment in the received literature. For example, Haritos[5] recognizes the existence of this problem in the abstract, but not in his empirical analysis. On the other hand, McGillivray et al.[6] bypass the issue entirely.

There have, of course, been attempts made to incorporate this non-user component into the analysis. At one time or another, such techniques as the relative use method and the predominant use method have experienced popularity as a means of allocating costs to non-users.[7] There is a serious theoretical flaw in both of these techniques. Both employ a measure of vehicle usage to determine a cost responsibility which would presumably exist even in the absence of such vehicles. It may well have been the problems stemming from this shortcoming which led Harbeson[8] to note that the user-non-user allocation of costs appears to differ depending upon whether or not the system is earning a "profit." Serious difficulties would arise from the adoption of the past attempts to make this cost separation.

What is required is an approach which is independent of the motor vehicle and its usage of a particular facility. In turn, this suggests that there are not likely to exist general allocative rules, so that each increment to the existing stock of highway capital must be considered on its own merits. The focus must be on the externalities generated by individual projects. The cost allocation between users and non-users must then be related to these externalities. In the extreme case where none exist, all of the costs would be attributable to the motor vehicle. Where they do exist, both their nature and extent must be explored before any allocation is possible. Such an approach would obviate the need to consider the motor vehicle at all in determining this cost separation, and would clearly be independent of the current "profitability" of the system. Such an approach is developed in a later section of the present paper.

In passing, it might be noted that this approach precludes any retrospective cost separation. It reinforces the earlier contention that full user pay is an objective which can be attained only in the long run.

The second question relates to the allocation of the remaining costs among the various classes of user. It is generally agreed that this phase of the analysis must be carried out at the level of vehicle class. The rationale here is straightforward. The fundamental question in the user pay debate involves the allocation of traffic among competing modes. Interesting as a macro comparison between highway costs and revenues may be, it masks the relative performance of vehicle classes, and might result in the acceptance of cross subsidization among those classes. Such cross subsidization is incompatible with the objective of improving resource allocation in transportation. Consequently, the focus must be on individual vehicle classes.

The implications of this position are considerably broader than the present paper. This principle has implications for other modes of transportation as well. In the case of rail, for example, concern for intermodal competition and the allocation of resources dictates that there must be a separation of national policy from national transportation policy. If, in the interests of the former, railroads are forced to operate uneconomic branch lines, then they should receive a compensatory subsidy, including an allowance for profit, in performing this function. The focused use of internat cross subsidization, whether by government or by one of its regulatory agencies, is incompatible with intermodal competition based on inherent advantage.

One further point needs to be made with respect to the allocation of costs among classes of vehicle. It must be remembered that there also exist various classes of highway, and that one should expect that the interclass allocation of costs will differ with highway type. With this in mind, it can safely be said that general allocation rules will not exist. Each case will have to be considered individually. In this regard, it should be noted that if the earlier difficulties have not ruled out retrospective cost allocation, this one does.

Only after these allocations have been successfully completed is it possible to consider the third question, the determination of the appropriate level of user charge. This price must be set subject to two constraints. The first is that, for any class of vehicle, the revenues generated must equal those costs attributed to that class. The other constraint is that the nature of the pricing mechanism not be such as to alter the inherent advantages of competing modes. It should be obvious that the pricing scheme adopted must satisfy both of these constraints simultaneously. Most of the received literature suggests one of two alternatives for price setting. The first of these involves setting price equal to marginal cost. The second sets price equal to the product of marginal cost and the reciprocal of the coefficient of price elasticity of demand. Both must be considered from the perspective of satisfying the constraints above.

In evaluating the first alternative, the crucial question is the relationship between marginal cost and average total cost, for equating price and marginal cost will not alter the inherent advantages of competing modes. In the Menger-type framework utilized in this paper, only the costs of owning and operating the infrastructure are relevant to infrastructure pricing. The social costs of congestion are excluded. In this situation, for any given capital stock, average variable costs will constantly decline until the capacity of the facility is achieved. Marginal costs are therefore less than average variable costs which, in turn, are less than average total costs. In this situation, setting price equal to marginal costs will result in revenues insufficient to cover total costs. This result will hold for all classes of user, and as a result, all classes of users will receive a hidden subsidy from government in their utilisation of the highway infrastructure. This approach must therefore be rejected as a pricing mechanism for it fails to satisfy the first of the two constraints above.

The second usual approach, making price a function of marginal cost and the elasticity of demand, can be employed to overcome the revenue shortfall inherent in strict marginal cost pricing. What emerges, however, is value-of-service pricing, an approach which results in the extraction of either producer or consumer surplus (depending on the class and other characteristics of the user). Furthermore, given that price is now dependent on something more than cost, the approach impinges on the inherent advantages of competing modes and thus violates the second constraint above. As a consequence, this pricing scheme must be rejected as well.

What alternative remains? The one which comes most readily to mind is an average cost approach. Under this sort of arrangement, price would be set equal to average total cost. On the surface, at least, such an approach would seem to satisfy both of the constraints. However, before the matter can be considered in depth, the determination of this average total cost must be considered in greater detail. This will be done in the next two sections of the paper.

Before proceeding, however, it should be noted that the cost allocations must be made on the basis of capital cost before the installation of the infrastructure. There are two reasons for this. The first is the need to integrate the pricing decision with the investment decision. The benefits of a particular investment in highways depend in part on traffic volumes. These, in turn, depend partly on the price charged for the use of the infrastructure. Only where the pricing and investment decisions are taken at the same time is an accurate estimate of benefits possible. To this must be added the fact that it is only at the time the investment decision is taken is it likely that there will exist sufficient familiarity with the data to permit the cost separations. The attainment of the objective of full user pay is therefore only attainable in the long run, for a retrospective allocation of capital costs appears impossible. It is therefore of crucial importance to ensure that each new project meets the test of full user pay. In this fashion, the goal will be attained gradually, and eventually, each part of the transportation system can become financially self-sufficient.

ALLOCATION OF COST: USERS AND NON-USERS

Highways, streets and roads are multi-purpose capital assets, built and maintained to perform a variety of functions. One of these, clearly the most important involves the motor vehicle. Other motives for investment in highways would include land access, military preparedness, social integration and economic development. The impact of these other factors is to increase the size of the physical plant beyond that which would exist in a user-optimized network. Needless to say, they increase associated costs as well. On the revenue side, however, they do not generate an equivalent increase in traffic volumes. Without some allocation of costs to these non-transportation motives for highway construction, the imposition of user charges as a general means of taxation tends to creep into the analysis. In turn, this would result in a misallocation of traffic among competing modes—the very evil that full user pay seeks to avoid. Some discussion of the nature of these non-user costs is therefore appropriate.

The question of the impact of improved transportation upon land values is straightforward once the basic interrelationships are understood. Highway investments increase the accessibility to certain areas. This improved access generates increased land rents, particularly in the peripheral areas of an economy. When this increase in rents is capitalized, it becomes an unearned increment to those fortunate enough to possess property rights in the affected areas. This process tends to occur most frequently in area adjacent to large urban agglomerations, and comes about because of the typical reaction of transportation planners to the process of urban growth and sprawl. As an urban area grows, congestion develops and worsens. Instead of imposing congestion tolls to restrict demand, the adjustment is made on the supply side. Highway investments are made to increase capacity. In turn, this response fuels the forces driving urban and suburban expansion, and so the cycle is repeated. The process is, in other words, a cumulative one. The essential point is that without this reaction on the part of the highway planner, the process would be considerably weakened. It is the reaction, therefore, which generates part of the increase in land rents and the concomitant increase in capital values. It is apparent that some fraction of this unearned increment should be captured by government to assist in the financing of the highway project which made it possible. The location rents so captured would then be deducted from the total capital cost of the project as the first step in the analysis. A collection mechanism would have to be developed, but this would present no real difficulty once the principle was accepted.

The impact of the economic development motive for highway investment is to advance the construction dates of particular projects. Put another way, the physical plant is improved sooner than it would otherwise be in the absence of the development effort. In addition, it might be noted that there is at least the possibility that standard of construction will be higher. Both factors impose costs on current users of the network beyond the costs which would exist in their absence. The solution to the problems created by the first development motive is best dealt with by means of a simple example. Suppose that, as a result of the economic development motive, the construction of a particular highway project is advanced by five years. If the project has a capital cost of $10 million, and the real rate of time preference is 4 percent, the present value of the project to current users is $8.22 million. It is this cost which should be allocated to users, while the balance, $1.78 million, should be allocated to non-users. An adjustment is also necessary in the other case. The cost of the higher quality facility will already have been estimated. What remains to be done is to estimate the cost of the facility which would be installed in a user-optimized system. Only the cost of the latter would be allocated to users, while the difference between the former cost and the latter would be funded from general revenues, or perhaps from the economic development budget.

The military preparedness motive for highway construction is best handled in essentially the same fashion as the quality impact of the economic development motive. The cost of the project, in the absence of the military motive, would be simulated, and this portion of total capital costs would be allocated to transportation users. The balance would then be financed from general revenues or from the defense budget.

The impact of the social integration motive for highway construction is somewhat more difficult to assess. Part of the difficulty can be overcome by remembering that there are three levels of transportation policy. The first of these is to fulfill a basic social need, to connect an area with adjacent areas. The second level relates to the laying of the groundwork for economic development. The third level of policy involves the construction of specific projects in specific areas, generally to meet specific needs. It is the first which is of interest here. Some portion of the highway capital stock in any area can be regarded as fulfilling this basic social need. Some, perhaps small, adjustment must therefore be made to highway capital stock to incorporate this motive into the analysis. In the case of new construction, the question need only be asked as to whether this motive is still present. If the answer is affirmative, some portion of the cost of the project must be allocated to non-users. A negative answer, of course, means no such allocation is necessary.

In passing, it might be noted that it is highly unlikely that general allocation rules for performing these analyses will exist. Most probably, each case will have to be treated on the basis of its own merits. While this complicates matters for the analyst, the outcome should be a more accurate determination of the costs attributable to the various non-user motives for highway construction.

Clearly, this approach to the allocation of costs between non-users and users is different from those which have been used or suggested in the past. It offers a number of advantages. The first of these is that the motor vehicle is not a consideration in making the non-user-user cost allocation. A second advantage is that the allocative mechanisms function independently of whether or not the components of the highway system are being operated at a "profit." In both of these respects, it overcomes deficiencies of past approaches. In addition, since it considers each project on the basis of its own merits, it appears to force a more disaggregated approach than do the earlier methods. This must be considered an advantage as well.

ALLOCATION OF COST: CLASSES OF USER

The capital costs which remain after the allocation to non-users form the basis for the next stage in the analysis—the allocation of capital costs among the various classes of user. This, of course, is the second question which must be considered in the determination of the optimum level of user charges.

Utilization of capital cost as the basis for the allocation permits the construction, via a perpetual inventory method, of a highway capital stock for which a given class of vehicle is responsible. A major advantage of this approach is that it permits an annual updating of the capital stock figures at this disaggregative level. This would, in turn, forestall the necessity of engaging in a succession of cost allocation studies as in the past. It need hardly be mentioned that the analysis must be carried out in real terms, and the user fee which ultimately emerges must then be converted into nominal terms through the application of an appropriate price index.

Broadly speaking, there would seem to be two approaches which could be employed to allocate capital cost responsibility among classes of user. The first of these is the benefit principle under which capital cost would be allocated on the basis of the proportion of user benefits received from a particular project. The alternative would be to employ some type of incremental cost approach. In this case, the capital cost to be borne by a particular class of vehicle would be determined on the basis of the costs incurred by government in providing for the infrastructural needs of that class vehicle. Both approaches have the common advantage of focusing on individual vehicle classes, the only theoretically valid approach.

The benefit principle is attractive because of the ease with which it could be utilized. Most, if not all of the required analysis will have already been carried out in the investment analysis stage. What would remain would be the calculation of ratios of benefit by class to total benefit and the allocation of capital costs on the basis of those ratios. This advantage would, of course, be lost in those circumstances where no investment analysis is carried out.

Critics of this approach tend to dismiss it for two reasons. The first of these is that benefits cannot really be measured with any degree of accuracy. While there is a substantial amount of truth in this statement, it must be pointed out that estimates of the volumes of benefits are employed in determining the viability of investments, and would have equal validity here. The second objection frequently voiced is that its use precludes the role of price as a rationing device. This objection would hold only in those cases where a single user charge/congestion toll was being contemplated. However, as noted above, these two aspects of the price finally charged to users need to be separated. A properly structured congestion toll must, at the minimum, be both time and location specific. Given this point, the second objection loses its validity.

The incremental cost approach begins by defining the cost of a basic road, one which is capable of carrying a small volume of light traffic. The tendency, one which appears theoretically unsound, is then to pro rate the costs associated with this basic road across all classes of traffic. It appears that this step is taken because the basic road accounts for an overwhelming proportion (as much as two-thirds) of the capital cost of a highway. This point aside, the approach does recognize that different highway users do place differing demands on the facility. It recognizes, furthermore, that there are differing costs associated with these differing demands. Costs above and beyond those of the basic road are then allocated to the particular vehicle class which necessitated their incurrance. Such an approach is intuitively attractive. Yet in practice, there is at least one major pitfall.

In evaluating the incremental cost approach, it is important to recognize that capital costs rise with not one, but rather with two factors. Capital costs increase with **vehicle size and weight.** The interrelationship has received much attention in the literature. Capital costs also increase with **rising traffic volumes.** This volume parameter seems to be frequently neglected in the literature. Yet, it is of overwhelming importance in urban areas. However, its importance as an escalator of highway capital costs is clearly shown in some early work completed by Ross.[9] He found that with a traffic volume of 2000 vehicles per day, the capital cost per mile of a facility designed for heavy vehicles (defined as 14,001 to 18,000 pounds per axle) was only about 1.5 times the capital cost of a facility designed to accommodate basic vehicles (defined as those with an axle weight of 6000 pounds or under). On the other hand, with only the basic vehicle included, capital cost increased by almost 2.7 times when traffic volumes rose from the range 300-750 per day to the 2000 per day volume. Given that capital costs increase with both increasing vehicle weight and with increasing traffic volumes, there arises an apparently insurmountable difficulty of separating the one force from the other. It does no good, as the incremental cost approach usually does, to circumvent the difficulty by focusing only on the relationship between weight and capital cost.

Given the problems with the incremental cost approach, one is left with the benefit approach of cost allocation.

However, there are problems with its use which make it less than ideal. To begin with, it brings the analysis uncomfortably close to the notion of a 'just price.' On the other side of the coin, this may not be as serious a problem as appears at first glance, for much of the debate about user pay is couched in terms of 'fair shares.' More serious are objections such as those offered by Lee.[10]

Lee suggests that the use of the benefit approach implies a compensatory payment among vehicle classes for each and every change in vehicle mix. Were the analysis carried out only periodically, and at the macroeconomic level, this would be a valid criticism. What seems to be forgotten in all of this is the fact that it is a specific vehicle mix which leads to the investment taking place. Furthermore, given that the vehicle mix will change slowly through time, that the present approach involves an annual updating and re-estimation of the user fee on a vehicle class basis, and that specific investments frequently are the result of the needs of particular classes, the argument is weakened considerably. On balance, the weight of evidence lies on the side of the benefit principle where the alternative is the incremental cost approach.

It will be noted that nothing has been said about the allocation of cost responsibility within classes of vehicle. For the purpose of establishing the basic user fee, it appears best to treat individual vehicle classes as a unit, and to make payment proportional to facility usage. Intra-class concerns relate primarily to questions concerning the time and/or location of vehicle usage, and the social costs thus imposed. In turn, this, once again, suggests some confusion concerning the role of a congestion toll. Congestion pricing must be dealt with separately from the question of the basic user charge.

In summary, the proposed approach invloves the allocation of user capital cost responsibility among classes of vehicle through the use of the benefit principle for each project. The results of these cost separations can then, through the use of a perpetual inventory model, be combined to generate a capital stock responsibility for each vehicle class. In turn, as discussed in the following section, this capital cost becomes the basis for the estimation of the optimum user fee in each time period. Attention must now be turned to the mechanisms for determining the charge for the basic use of the highway infrastructure.

SETTING THE BASIC USER CHARGE

This section of the paper utilizes the capital stocks generated above to determine the user charge for the various classes of vehicle using the facility. The initial step is to convert the capital costs into an annual figure. Once this has been done, some attention is devoted to the question of traffic flows. Finally, these two elements are combined to determine the optimum level of user charges. Three alternatives are considered. The first involves the recovery of annual costs each year. The second examines the consequences of recovering annual costs only at capacity levels of utilization of the system. The final alternative changes the focus and considers costs and revenues only over the life cycle of a facility.

In estimating the annual costs, four elements of cost must be considered. These are: administrative costs, maintenance costs, depreciation and the opportunity cost of capital. Of these, the first, third and fourth are independent of the volume of traffic in the aggregate, though not (obviously) independent of the cost per vehicle mile. Maintenance costs are partially a function of usage, but are partly independent of traffic volume as well. Each of these must be discussed in turn.

Initially, one might be tempted to simply incorporate historical administrative and maintenance costs into the analysis. However, given the time perspective of the paper, where future time as well as past time is included, some other means must be found. One alternative would be to forecast the historical data. However, there is a better way. Rules of thumb which relate future administrative and maintenance costs to capital costs are frequently employed in investment analysis. This is a more fruitful approach to adopt here. Clearly, historical costs would be a useful aid in determining the relevant ratios. Data from elsewhere would be useful as well. Administration costs can be dealt with quickly. They are held to be a constant proportion of the gross capital stock attributed to each class of vehicle, both expressed in real terms. Maintenance expenditures are somewhat more complex and require further discussion.

Maintenance expenditures can be subdivided into three categories—those which are attributable to the non-user, those which are independent of the motor vehicle for other reasons, and those which can be technically attributable to specific vehicle classes. With the exception of controlled access highways, some component of the annual highway maintenance budget is expended to provide access to land. Driveway culverts would be the most important example. Much of the remaining maintenance is time associated rather than use associated. The costs of ditching, striping, maintenance of signs and lighting and the painting of structures are all examples. Snow removal expenditures, while not fitting this category, is not really vehicle associated either. Only the maintenance of the surface course can, strictly speaking, be attributed to specific vehicle classes. Even here, the problem is complicated by the fact that there exists a trade-off between initial capital cost of a facility and the annual maintenance it requires. The first category of maintenance costs can be simply deleted and charged to general taxation. The second category, plus snow removal, is largely a function of the size of the highway network and can be dealt with by determining a ratio to capital stock by class of vehicle. The final category is best treated incrementally by vehicle class, though once this is done, it can be ratioed to capital stock as well.

Depreciation expenses are a straightforward matter. Annual depreciation for a single facility considered in isolation is simply its initial capital cost less its scrap value divided by its useful life. The system, being an aggregate of such projects is simply a weighted expense total. It should be obvious that this amount can be reduced to a vehicle class basis without difficulty. The opportunity cost of capital differs from depreciation only in that it is based on capital stock net of depreciation rather than gross capital stock. It should be noted in passing that since the analysis is carried out in real terms, a real opportunity cost of capital must be utilized.

The total annual cost of the highway system for any given class of vehicle is then simply the sum of these four cost elements.

With respect to the revenue side, user charge revenue is simply the product of two things: the level of the user charge (a parameter to be determined) and the volume of traffic. Volume depends on the user fee and a host of exogenous factors. Among the latter would be real incomes, the prices of competing modes, and the cost of energy. Their impact will be through an impact of the rate of growth in traffic. In the present paper, for the sake of convenience, traffic will be assumed to grow at the constant proportionate rate v.

The first method of setting the level of user charges, full cost recovery every year, does meet the test of full cost recovery. For a single project considered in isolation, at least, the approach creates problems. In this case, total annual cost will decline through time, largely because of the decline in the opportunity cost of capital. With increasing volumes, this implies a declining user fee over time. Failure to adjust the user fee would be to employ it as a means of general taxation. At the system level, the same result holds where the real rate of growth of revenues exceeds the real rate of growth of annual costs. If the former expands at 7 percent, the latter expands at 5 percent, and initially revenues are 70 percent of costs, it will take between 19 and 20 years for revenues to catch up.

The fundamental problem is that the approach does not adequately take into account the impact of excess capacity in the early years of a facility's life. Users are presented with a physical plant, and hence a cost, beyond their present demands. The approach would overcharge users in the early years, undercharging them in the later years. Finally, the approach is incompatible with the desire to ration the highways through congestion pricing. This system, in fact, increases congestion problems through diverting traffic by decreasing price. For all of these reasons, the approach must be rejected.

The second approach, recovering full costs only at capacity levels of utilization, can be rejected out of hand. It fails to meet the test of full cost recovery, and consequently involves a hidden subsidy to the users of the infrastructure. It does have one positive aspect. Since it does involve a user charge which in real terms is constant through time, it does not result in the congestion complication of the first approach.

The third approach considers the relationship between costs and revenues from the perspective of the life cycle. For a single project, annual costs for each year in its lifespan are estimated, discounted to the present and summed. The same procedure is followed for traffic volumes. The former is then divided by the latter to determine a user fee which is held constant over the life of the facility. All costs, including the interest payments on the early year deficits, are recovered in that period. The congestion problem noted above is avoided, for the real fee is constant through time. It should be noted that an annual cost-revenue comparison is not made for any year. In this framework, such a comparison would be irrelevant. At the level of the highway system, the same approach is followed with a single change. At the system level, the user fee must be recomputed annually to reflect the fact that the highway plant is expanding.

A practical example for a single facility may serve to clarify some of the issues. Suppose there is a single class of user whose capital cost responsibility is $1.0 million for a facility with a useful life of 10 years. Initial traffic volumes are set at 3.5 million vehicle miles, and the growth rate is 5 percent. The real discount rate is also 5 percent. Administrative costs might be 2 percent of capital costs, maintenance might be 5 percent, depreciation would be 10 percent given the ten-year useful life, and the opportunity cost of capital 5 percent. The total annual costs would then be 22 percent of capital costs each year.

If the first method of setting the user fee were adopted, the initial level of user fee would be 6.3 cents per vehicle mile. In the final year, it would be 3.1 cents per vehicle mile. Over the life of the facility, the total costs of $1.975 million are recovered, so there is no hidden subsidy to users. Traffic would, however, be diverted from other roads or modes increasing congestion difficulties.

In the second case, the user fee would be a constant 3.1 cents per vehicle mile. However, the accumulated deficit at the end of the period would be $5.438 million, or 27.5 percent of the total annual costs incurred. This, of course, represents a substantial hidden subsidy to users. It is this hidden subsidy which causes the rejection of the approach.

The third method of setting the fee results in a user charge of 4.4 cents per vehicle mile, a charge constant over the facility's life. Total costs, including the imputed interest on the early year deficits, are recovered so there is no hidden subsidy to users. Because the fee is constant, there is no tendency to increase the pressures of congestion as the facility moves toward the end of its useful life. On balance, this approach is the best one to employ in establishing the level of user charges. There are early year deficits, but these reflect the existence of indivisibilities and excess capacity early on. They do not imply a hidden subsidy to the user. There are late year surpluses, and again, it must be noted that these simply reflect the indivisibilities and excess capacity problems above. They do not imply the application of user charges as a means of general taxation. This third approach has clear advantages over the others and should be adopted.

Two other points need to be made. The first of these is that for each level of capital stock, traffic volume and mix, and traffic growth, there exists an optimum set of user charges. Where expansion occurs, as it inevitably does, the user fee must be recalculated. The process is straightforward. For each year's highway investment projects, capital cost must be allocated. The first allocation is between users and non-users. The balance of capital costs is then allocated among classes of user. For each class of user, this capital cost is added to the existing gross capital stock, and there is an equivalent dropping of the worn out capital stock. The result is a new estimate of capital stock, and this forms the basis for the calculation of the revised user fee. The current validity of the four cost ratios must be checked, and revisions made where necessary. Once this is done, the total annual cost streams are calculated, discounted and summed. The same process is followed on the traffic volume side. The user fee, which is held constant within the year, is then estimated. Following this procedure, the groundwork is laid for full cost recovery in the future. The second point to be remembered is that the analysis is to be carried out in real terms. The resulting user fee would then have to be converted into nominal terms through the use of an appropriate price index.

SUMMARY AND CONCLUSIONS

This paper has considered the question of user pay in isolation from other, closely related aspects of transportation policy. It has argued that, in dealing with this question, it is not sufficient to consider time only in historical perspective, but that future time must be included as well. The inherent tendency toward excess capacity in the highway network was discussed, and the need for taking account of this problem in setting the level of user fees stressed.

It was argued that the initial step in the determination of the level of user fees must be the allocation of capital costs between users and non-users. Several means of performing this allocation, all of which were independent of the motor vehicle and related to the non-user requirements, were discussed. It was concluded that general allocation rules do not exist and that each case must be treated on its own merits.

The next question considered was the allocation of costs among classes of vehicle. Despite its numerous short-comings, it was argued that the best allocator to employ here was the ratio of benefits received. After these allocations have been made, they are accumulated, using a perpetual inventory model, into capital stock estimates by class vehicle. In turn, these capital stock estimates form the basis for the estimation of total annual costs.

After a brief discussion of the revenue side of the question, the paper explored three methods that could be employed in setting the level of user fee. Both the annual equating of annual costs and revenues and the equating of annual costs and revenues at capacity levels of utilization were rejected because of some problems they created. It was argued that a lifecycle approach, in which total costs are recovered only over the lifespan of a facility, offered overwhelming advantages and should be adopted.

Finally, it was pointed out that, in the real world where highway capital stock is expanding, it is necessary to recalculate the user fee annually. It was also noted that, since the analysis must be carried out in real terms, some price index must be found to permit the conversion of the user fee into nominal terms. Specific price indexes, however, were not reviewed.

FOOTNOTES

1 For a discussion of Menger, see: George J. Stigler, **Production and Distribution Theories,** (New York, MacMillan, 1941), pages 136-147.

2 K. Bhatt et al., "Review of Road Expenditures and Payments by Vehicle Class: 1956 to 1975," **Transportation Research Record 680,** (Washington, TRB, 1978)

3 For example, see: R. McGillivray et al., "Toward Rational Road-User Charges," Transportation Research Record 680, (Washington, TRB, 1978)

4 See for example: Steven A. Morrison, "The Structure of Landing Fees at Uncongested Airports," a paper presented to the Canadian Economics Association, June, 1982.

5 Z. Haritos, **Rational Road Pricing Policies in Canada,** (Ottawa, Canadian Transport Commission, 1973).

6 R. McGillivray et al., **Op. Cit.**

7 See, for example, A. D. LeBaron, "The Theory of Highway Finance: Roots, Aims and Accomplishments," **National Tax Journal,** 1963.

8 R. W. Harbeson, "Some Unsettled Issues in Highway Cost Allocation," in Kleinsorge (Ed.), **Public Finance and Welfare.**

9 W. B. Ross, **Financing Highway Improvements in Louisiana** (Baton Rouge, 1955) as reported in J. R. Meyer et al., **The Economics of Competition in the Transportation Industries,** (Cambridge, Harvard University Press, 1964)

10 Douglass B. Lee, **Methods for Allocating Highway Costs,** U.S.D.O.T. Staff Study SS-24-U.3-181, (Cambridge, U.S.D.O.T. Transportation Systems Center, 1981)