Reaching an Economic Balance Between Mass Transit and Provision for Individual Automobile Traffic

by William Vickrey

ABSTRACT

In 1959, William Vickrey presented a statement, along with exhibits and appendices, to the Joint Committee on Washington Metropolitan Problems. Judging the material to be of considerable historical interest in the context of urban economic transport theory and policy, and to rescue it from obscurity, we recommended to the Editor that part of it be published in the *Review*. The Editor graciously assented.

The statement provides an overview of Vickrey's thinking on the technological implementation of efficient urban transport policy. The appendix to the statement presents "A Preliminary Sketch of Possible Schemes for Automatic Toll Assessment with Reference to the Washington, D.C., Metropolitan Area." Exhibit 51 is an essay, "Reaching an Economic Balance between Mass Transit and Provision for Individual Automobile Traffic." The appendix to that exhibit contains a derivation of Ramsey-Boiteux pricing in the context of urban transportation. Exhibit 52 deals with the "Construction of Tentative Fare Schedules for Washington, D.C., Regional Rapid Transit System." And Exhibit 53 is an essay on "The Economizing of Curb Parking Space - A Suggestion for a New Approach to Parking Meters."

The total length is some 100 typewritten pages. To make the material suitable for journal publication, we excluded Exhibit 52, divided the remaining material into two coherent papers, and did some editing. This paper includes Exhibit 51 and its appendix, very lightly-edited. The statement, its appendix, and Exhibit 53 are forthcoming in the *Journal of Urban Economics* under the title "Statement to the Joint Committee on Washington, D.C., Metropolitan Problems."

This paper not only demonstrates Vickrey's creativity, economic intelligence, breadth of vision, and farsightedness, but also provides further evidence that he is the first to have thought deeply of pricing solutions to urban transport problems. With the tremendous advances in technology over the last thirty-five years, his proposals are even more compelling now than when they were originally formulated.

- Richard Arnott and Marvin Kraus, Boston College

I. Introduction

Incremental Costs of Rush-Hour Trips

No comparison of methods for providing local transportation facilities for a metropolitan area would be complete without some indication of the relative incremental costs of caring for rush-hour traffic by various methods. To come to intelligent decisions it is necessary to be aware of the consequences for costs of shifts in the division of the total traffic between one medium and another. In the light of these cost consequences, measures may appear appropriate that might be dismissed as too extreme or inappropriate in the

* William Vickrey is Professor Emeritus, Columbia University.

absence of a full appreciation of what these costs mean when reduced to terms of cost per trip.

The following table, while subject to a wide margin of error and to some very serious qualifications as to possible bias in the underlying data, can be considered to give at least a general idea of the order of magnitudes involved, for the purposes of a preliminary discussion.

TABLE 1
Incremental Costs of Rush Hour Travel by Various Modes

		Investment cost			
	Operating cost per round trip	Investment per round trip per year	Rate of capital charge (%)	Cost per round trip	Total cost per round trip
Express bus	$0.35	$2.70	15	$0.40	$0.75
Rail:					
At constant frequency	.12	4.20	10	.42	.54
At constant train length	.19	4.20	10	.42	.61
Private automobile, 1 person:					
Expressways and streets		63.00	5	3.15	3.15
Out-of-pocket operating expenses	.30				.30
Total, without parking					3.45
Parking, central area	.20	10.00	8	.80	1.00
Total, with parking					4.45
Persons per car (average)	1	1.5	1.8	2	6
Cost per person, without parking	$3.45	$2.30	$1.95	$1.75	$0.60
Cost per person, with parking	$4.45	$2.95	$2.50	$2.25	$0.75

The salient figure of $63 of highway investment per round trip per year perhaps deserves some comment, as this is what accounts for most of the striking variation in costs. The difference between the highway outlay called for under the recommended plan and under the auto dominant plan is $590 million. Corresponding to this difference in costs is a difference of about 45,000 persons travelling by auto during the morning peak hour to the zero sector. In addition there is some traffic with other destinations that forms, over part of its total journey, part of the controlling volumes of total traffic upon which the lane requirements in the various plans are based; with due allowance for the smaller average mileage over which the traffic not destined to zero sector will be a part of the dominant flow, a division of 65,000 persons would appear appropriate for determining the degree to which each of 45,000 rush hour trips is responsible for the difference in cost. Allowing 1.8 persons per car, this is equivalent to 36,000 car trips, during the peak hour, and the cost differential works out at $16,300 per daily car trip, or on the basis of 260 heavy traffic days per year, this becomes $63 per rush hour vehicle round trip per year. The percentages used for capital charges are admittedly somewhat arbitrary; they are intended to cover interest and amortization or depreciation, plus possibly some allowance for taxes foregone on taxable property taken for highway or parking use; if anything they probably somewhat understate the relative cost of providing for private automobile transportation.

These figures, rough as they are, serve to point up the relatively high cost of providing facilities for large scale rush hour movement by private automobile in urbanized areas. Some of the implied questions raised by these figures are:

a. Is it not excessively wasteful to insist on providing free facilities on a scale which which may greatly exceed that which the users would require if they were asked to pay a price individually for the service they use commensurate with its cost?
b. Are there possibilities for the diversion of traffic from private automobiles to mass transit which have not been fully explored?
c. Are there possibilities for inducing a reduction in the intensity of the peak movement which have not been exploited?
d. Are all of these trips really necessary, in the sense of being of sufficient importance, even including all of the indirect benefits and consequences, to warrant the indicated heavy expenditures?

The Free Highway Tradition and Urban Congestion

Highway planning and finance is strongly conditioned by a long tradition of planning in terms of predominantly free highways, financed either from general tax revenues or from fuel and similar taxes representing a relatively modest addition to other costs of operating automobiles. As applied to nonurban roads and highways, this has worked fairly well. On low traffic secondary roads, the incremental costs in the form of wear and tear or increased congestion caused by incremental traffic has been generally negligible, and once the decision has been made at the governmental level to construct such roads in such a manner as to provide a given standard of alinement, grade, and paving, little purpose would be served by imposing any constraints on their use, aside from weight restrictions designed to prevent breakdown from overloading. On heavily traveled roads, construction costs in nonurbanized areas have been such that the cost of providing for additional traffic while maintaining a constant or even improved level of service has not been much if any in excess of the user charges imposed through fuel and other taxes. It is therefore only rarely, in these areas, that the use of the facility is carried to the point where the costs imposed on fellow users, taxpayers, and others by the last increment of use has outweighed to any significant extent the benefit derived by the user. Most such occasions involve accidents, emergencies, or contingencies of various kinds, or not readily controlled occasions such as holidays and special events.

Urban highway conditions, however, are so radically different that the application of this philosophy in this area is likely to produce very wasteful consequences. It becomes impossible, or at least prohibitively expensive, to preserve a reasonably free-flowing condition on streets and expressways during rush hours. At least for the Washington area, under the assumptions upon which the traffic assignments have thus far been based, even with the most attractive transit system feasible, if the assumption is made that competing with this there is a free-flowing highway system without tolls or other constraints on its use, a volume of automobile traffic is projected that calls for a highway system that appears out of the question on technical grounds alone, even aside from cost considerations. In the absence of other constraints on highway use, there is then implicit in all of the plans proposed a degree of rush hour highway congestion sufficiently severe to cause a diversion of a substantial number of potential private car users to transit, or perhaps the complete abandonment of a substantial number of trips that would otherwise be made. Under urban conditions we cannot have both free flowing rush hour traffic and the absence of user charges or other constraints on highway use. One or the other of these desiderata must yield.

Congestion and Express Bus Service

Permitting such a degree of congestion to develop raises additional problems in connection with plans for express bus service. If the bus service is nonsegregated and uses the same lanes and facilities as other traffic, then it becomes subject to the same delays that affect private cars, so that the attractiveness of the bus over the private car in terms of time tends to vanish. If this is to be avoided, it then becomes essential to provide for segregated bus lanes on the routes that are subject to this congestion and possibly, in addition, for some means of insulating the buses from the effects of congestion in the central business district during the distribution portion of their run. The additional cost of providing this segregation is likely to be considerable, and it is not at all clear that the estimates presented for express bus operation include an adequate allowance on this score. The extra ramps and roadway separations required may not be a major item, but in many cases it will not be possible to have bus traffic closely geared to an integral number of lanes, so that lanes devoted exclusively to bus traffic will be only partially utilized. The estimates for bus operation appear to have been based on the assumption that in the majority of the cases buses would be accommodated in the lanes planned for the general flow of traffic, and where bus traffic alone is less than the full capacity of a lane, the cost of the full lane that would be required for segregated operation has apparently not been charged.

More important, in the absence of constraints on highway use it is not at all clear how the congestion likely to develop on the local CBD streets is to be prevented from slowing the bus distribution service down to the point where it fails to compete timewise with private automobiles.

The Advantages of Constraints on Rush Hour Auto Traffic

The economics of constraints on rush hour highway use can be illustrated at the planning level as follows. Under the recommended plan, total peak hour traffic to the zero sector is put at 145,000 persons, of whom 100,000 are assumed to use transit, and 45,000 to use automobiles. If some means could be found of diverting an additional 15,000 persons to transit from automobiles, and if we assume that the diversion is most effective for solo automobile traffic so that we have an average of 1.5 persons per car for this diverted traffic, then costs of $44,500 per day would be avoided by reason of the reduced automobile traffic, assuming the highway construction program to be cut back so as to preserve the same degree of congestion; the increased cost of providing transit service of the same quality as before would be only $10,000 per day, a saving of $34,500 per day, or $8,800,000 per year on the basis of 255 heavy traffic days per year.

Conversely, if for some reason the attractiveness of transit turns out to be less than anticipated, and if the traffic divides 85,000 on transit and 60,000 in automobiles, then if congestion is to be limited to the same level as is assumed in the traffic assignment as presented, additional costs of $8,800,000 per year will be required.

In the above comparisons we have assumed that 36 percent of transit passengers used rail, the change in traffic being taken care of on a constant frequency basis by adjusting the length of the trains. If the adjustment is on a constant train length basis, the figure would be $10,500, and if the system is an all-bus system, the figure would be $11,200; as a partial offset to these higher costs in the latter two cases, there would be some added benefit through increased frequency of service and possibly, in the case of bus service, a greater variety of routes. But the differential is, in any case, of the same order of magnitude.

On a short run basis, the same underlying relationship reveals itself in a slightly different form. A lane of expressway can be considered to have a nominal capacity of 1,500 cars per hour, at which rate of flow traffic can be considered to be generally free-flowing and subject only to delays of a minor or exceptional sort. At 1,800 cars per hour, we may assume that traffic still flows at a good speed as long as weather conditions are favorable and no traffic incidents occur, but any unfavorable developments tend to cause delays of considerable magnitude, and driving is in any case a somewhat trying experience for the average driver. Beyond 1,800 cars per hour we may consider that conditions deteriorate rapidly; traffic is slowed and access is difficult; queues lengthen, and attempts by more cars to pass over a given stretch may merely result in the forcible spreading of the traffic over a longer period by holding up the excess traffic until the peak demand subsides. In the absence of constraints on traffic, and in the absence of highway facilities approaching the impossible levels called for by the auto-dominant program, we can expect highway traffic to be somewhere in the 1,800 to 2,000 cars per lane range, let us say 1,900 cars per lane per hour.

It would not be possible to remedy this situation merely by increasing the number of lanes by 25 percent; the result is not the division of the former traffic among more lanes, but the attraction of more traffic as conditions get better, so that the final result will be not the division of the former traffic among more lanes, which would have given a very tolerable level of 1,520 cars per lane per hour, but rather say 1,850 cars per lane per hour. The lanes really required to provide free flow when additional traffic is allowed to add itself to the flow without limit, would be excessive.

Suppose, however, that sufficient inducement is given to say 20 percent of the traffic to cause them to shift to transit, or possibly to abandon their trips entirely, and to stay off the highway with their private cars in spite of the subsequent improvement in traffic conditions. It would then be possible for the remaining 1,520 cars per hour to move freely over the highway system, which they could not do under any previous highway program; on the other hand the traffic that is shifted is enjoying a service which, while in some respects inferior to that which those who remain on the highways are now enjoying, is only slightly inferior to the congested highway travel which is what they would get if they all went back to the highway. Particularly if there is a substantial number of automobile users who have only a very slight margin of preference for this mode of transportation over the alternatives, a very great gain to the other highway users can be secured at only very slight inconvenience to those on the margin.

Still another way of looking at the matter is to consider that there are always in a community individual highway trips to be made for which any conceivable form of transit is a very inferior substitute. It is likely to be of considerable value to at least some of those making such trips, to be able to make them, even during rush hours, under reasonably free-flowing traffic conditions. Without constraints it may become prohibitively expensive to provide free-flowing conditions for such traffic. With constraints on the use of the highway system, however, conditions can be created such that every person for whom it is adequately important to use his private automobile for a given trip will be able to do so under reasonably satisfactory conditions, an opportunity that would be completely closed to him if the free-highway tradition is adhered to.

Methods of Reducing Rush Hour Traffic

It is accordingly of the utmost importance, in evaluating plans for traffic facilities, to consider the various ways by which their use may be suitably controlled. Such control may be either direct constraints related directly to the use made of the highways themselves, or constraints applied to complementary facilities such as parking, or inducements attached to the use of substitute facilities such as mass transit.

Reduced Transit Fares

One possibility is to reduce transit fares, even though this may mean operating the transit system at a deficit. Suppose, for example, that transit fares are eliminated altogether, or perhaps merely reduced to a nominal level of 5 cents from an average of, say, 25 cents for a one-way trip. Suppose that rush hour transit riding thereupon increases by 30,000, of which 15,000 comes from reduced automobile traffic, the remainder being additional trips not previously made. Then if we assume the highway construction program to be cut back to the point where the same degree of congestion occurs as would have occurred under the original program, the total cost of providing for auto traffic would again be reduced by $44,500 per day, while the cost of providing transit would be increased by $20,000 per day, a net decrease in the cost of transit facilities of $24,500 per day, or $6,200,000 per year. With this saving we have provided for 15,000 new round trips valued at 30 cents on the average (since the riders in question were formerly not willing to pay the round trip fare of 50 cents, but are now willing to pay 10 cents) or an additional service value of $4,500 per day. On the other hand the service used by the riders who gave up auto riding may be said to be on the average less valuable than their former auto ride by 60 cents, since if we assume that their out-of-pocket expense, including parking, averaged 90 cents per passenger, their previous cost differential was 90 cents minus 50 cents or 40 cents for the round trip, while the new cost differential is 90 cents minus 10 cents, or 80 cents:

those who shift will be indicating that the value differential between the two modes of transportation lies for them between these two cost differentials, and we can assume that the average value differential is at the midpoint of the interval. Thus we have a reduction in cost of $24,500, a gain in service value to new riders of $4,500, a loss in service value to shifting riders of $9,000, a net gain from this change of $20,000 per day.

Unfortunately it would not be possible, in general, to offer such low fares to rush-hour riders without offering equally low fares to others, partly because of the political objections to discrimination of this sort, but more practically because of the tendency which this would have to cause a shift in riding from the non-rush to the rush hours; such a shift would tend to produce a reduction in the efficiency of transit operation which might very largely wipe out the indicated gains. In any case, heavy deficits would be incurred from transit operations, and though it would in principle be possible to meet these deficits by appropriate tax increases in the various jurisdictions concerned, the difficulties of doing this may well be considered to be such as to preclude this as a practical possibility, or at least such as to offset to a considerable extent the potential gains evaluated above. Thus while this analysis does serve to underline the importance of keeping transit fares at a low level, it does not provide a major contribution to the solution of the problem of diverting an adequate proportion of rush-hour traffic to transit.

Parking Restrictions

The second major alternative is to impose restrictions on supplementary services which in this case means parking. If it could be assumed that all all-day parkers in the CBD are invariably drivers who make their trips to and from the CBD during the rush hours and over equally lengthy and congested routes, then a restriction in the number of available all-day parking spaces in the downtown area to the number for which there would be a demand at a charge of $4 or $5 per day would go far towards solving the problem.

To enforce such a restriction without undesirable results of various kinds, however, is by no means easy. Merely to hold the number of spaces down by restricted issue of construction permits would mean substantial windfall profits to the favored operators, and would also tend to produce a creeping expansion of parking in various alleyways, service stations, service areas, and the like, in a manner which would be far more expensive in terms of interference with other activities than if the equivalent space had been provided in properly designed structures. If an attempt is made to hold down the excess profits of the parking lot operators by public regulation of parking charges to a rate yielding only a normal return, parking space is likely to be preempted by the firstcomers in one sense or another, leaving late arrivals and those with irregular needs with no place to go, regardless of how urgent or legitimate their need for the use of a private automobile may be.

The ideal method, perhaps, would be to levy a tax of say $3 per day on all day parking. But even this would in many ways fail to meet the need. There would remain the tendency of those with access to private or free parking of various kinds and those whose trips, though forming part of the congested flow, take them to places where parking is less expensive or where the tax would not apply, to add their bit to the congestion of the highways without adequate motive. On the other hand it would be difficult to segregate the all-day parker in such a way as to avoid charging excessively heavy parking rates to the short-term parker whose trip involves only part of the day and who does not become part of any of the controlling traffic flows that determine the magnitude of the highway structure needed. Similarly it may be difficult to discriminate between the 9 to 5 all-day parker who is the person to be singled out for the tax, and other all-day parkers, such as theater employees who may park say from 2 to 11, thus making no contribution to rush-hour congestion or highway lane requirements. The closest approach would be a special charge levied specifically on those who place their car in parking spaces between say 8 and 9:30, and on those who remove their cars between 4:30 and 6; but again this would improperly catch those who merely move their cars at 5 p.m. from a space near their office to a space near a downtown bar or restaurant.

Thus while the control of traffic through control over parking spaces and charges could make a considerable contribution to the problem, this is awkward and at best only partially and unevenly effective.

Direct User Charges for Rush Hour Highway Use

The direct approach would of course call for the assessment of charges specifically on those who actually use the highways in such a manner as to contribute to congestion or to require the construction of additional facilities in order to avert congestion. Usually such suggestions are discounted, not only on the arbitrary and emotional ground that this is contrary to the tradition of the free use of highways, but also on the more practical ground that toll collection by any of the techniques now practiced would be inordinately expensive and would also hamper the free flow of traffic and reduce the capacity of the highway system. There is also the possibility that if toll assessment is concentrated, for economy reasons or otherwise, at a small number of barriers, this may lead either to inefficient distribution of traffic around the toll barriers, or to restrictions on the number or location of entrance and exit ramps, in a manner not conducive to the most direct flow of traffic, or perhaps both.

The importance of this issue, however, warrants a careful examination of the possibilities as they may be affected by technological developments. It seems entirely within the realm of possibility, using only elements that are already at an advanced stage of development, to devise a plan for the assessment of tolls in a completely flexible manner, using an insignificant amount of manpower, and causing no disturbance whatever to the flow of traffic.

Electronic Toll Assessment

Electronic apparatus has already been devised which will sense the passage, at moderate distances of up to several feet, of simple tuned drone circuits, distinguishing among these drone circuits according to the frequency to which they are tuned. The tuned drone circuits themselves are simple, inexpensive combinations of a coil or loop of wire and a fixed condenser, requiring no power supply. This principle has already been applied by the Chicago Transit Authority to the identification of trains and the automatic setting of routes.

There would seem to be nothing inherently difficult or inordinately expensive about encasing say 10 of these circuits, tuned to different frequencies, in a plastic block, and having each regular user of the highway system fasten such a block, with a combination of frequencies that would be unique for him, under the bumper of his car, just below the license plate, for example. Or the circuits might even be incorporated in the license plate itself. Each entrance and exit lane to the freeway system would then be provided with an electronic scanning device, which would sense the drone circuit combination of each equipped car, test the combination for conformity with certain simple error-detecting criteria, and record this combination on a magnetic tape, which would also be made to contain a record of the date and time at suitable intervals. Each month (oftener if desired) the tapes would be collected and processed in a general purpose computer. The computer would collate the various entries pertaining to a given car with data from a master tape containing the name and address of the owner of the car for each frequency combination, compute the charges due according to the indicated trips made by the car, and print out a bill, which could, if desired, include a detailed list of the trips for which the charge is being made.

To take care of the occasional user of the highway system, as well as cases of malfunction or tampering, if the electronic sensing unit detects the passage of a car which either does not contain a code block, or for which the code response obtained is improper, it could be made to trigger a camera aimed to show the license plates of the car. This film could be developed and projected, with the license number being transcribed manually from the film to magnetic tape, along with indications of the place and time, after which processing could be carried out as for the regular users, assuming that a magnetic tape file of names and addresses for the various license numbers has been compiled.

In either case, ultimate payment of the charges assessed could be enforced in much the same way as is done for parking tickets. In order to encourage motorists to take the trouble to install the code block properly, an extra 5 or 10 cents could be charged for each trip for which the charge is assessed by photograph, in order to cover the extra cost. To minimize billing costs associated with infrequent use, perhaps one free rush hour round trip per month could be allowed each nonlocal car. As a check against fraud, a small random sample of cars could be photographed as well as recorded electronically, the two records being checked against each other.

Pattern of Charges Suitable for Traffic Control

Such a system would be capable of assessing charges in a wide variety of patterns. Since the highway system proposed is considered desirable even if financed from sources other than specific user charges, it is clear that the purpose of the user charges is not primarily to finance the highway system or even to raise any designated amount of revenue; the primary purpose of the charge is to control the use of the highway system so as to permit the maintenance of free-flowing conditions and preserve the availability of rapid private automobile transportation to those requiring it, without wasteful expenditure on excessively extensive facilities.

Assuming a correct forecast of traffic developments, and that the period of peak traffic continues to be substantially 1 hour, a charge during the rush hour of $1.70 for a typical one-way trip to the zero sector would be in order, and if capacity has been correctly programmed, then with the charge at this level the capacity of the system should be just about comfortably utilized during the peak hour. A person making 40 such rush-hour trips during the month would be billed for about $68 at the end of the month. The amounts are thus substantial, even in relation to the presumably higher incomes and standards of living to be expected in 1980.

On the other hand as long as we keep to the principle of using the charges primarily as a means of promoting the efficient use of the highways, there is little justification for any charge at all during off-peak periods, considering the highway system strictly by itself. Considering the possibilities of substitution between off-peak transit riding and off-peak auto travel, however, there would be something to be said for keeping a small charge of 10 to 20 cents in effect to match the excess of the transit fare over the cost of the off-peak ride, if there is no corresponding off-peak reduction in the transit fare. Something can also be said for such a charge as a relatively innocuous and cheap way of raising additional revenue for public purposes, given that the apparatus is already in existence for the purpose of assessing the rush-hour charge.

If, however, charges for peak-hour trips at the level of $1.70 per trip, coupled with zero or nominal charges in off-peak periods, have a substantial effect in the direction of shifting travel to the periods before and after the peak, the appropriate pattern of charges may shift slightly, since we do not want merely to create new peaks and congested periods at the time just before the peak-hour rate goes into effect, or after the peak hour. One might emerge with a pattern of charges calling for $1 during the peak half hour, 60 cents for the half hour on either side of this peak, and 30 cents for the hours preceding and following these. Or the steps could be even more frequent and gradual. The charge would, of course, also be graduated according to the distance traveled over routes carrying capacity traffic. In general, no charge or a purely nominal charge would be made for travel in the direction opposed to the dominant flow, though where reversible lanes are used this may not hold in all cases.

The essential principle is to so graduate the charge that as long as the facility involved is being used at or near capacity, the charge is high enough to prevent overloading and congestion under normal circumstances, and that as long as the facility is not being used to capacity there is only a nominal charge, if any. Whether the revenues collected reach any preassigned level is for this purpose entirely irrelevant. It might even be argued that the charges should vary with the weather or other adventitious circumstances affecting traffic and highway capacity, even though such variations would probably be less effective in influencing the volume of traffic than variations of a more regular and predictable sort.

User Charges and the Staging of New Construction

A flexible system of user charges of this sort can greatly facilitate the staging of the highway construction program. It can greatly ease the transition from one stage to the next, as by helping to route traffic away from congestion caused by work in progress; by permitting a more economical timing of additions to the highway system, by providing a continuously up-to-date record of traffic flows by origins and destinations in whatever detail needed; by minimizing losses resulting from the inevitable errors in planning, through permitting full or at least unconstrained utilization of facilities that have been overbuilt, and preventing the usefulness of underbuilt facilities from being made still more inadequate by congestion.

Thus, for example, if a particular artery happens to be for the moment overbuilt, either through an error in forecasting or through the necessity for constructing in one piece a unit with a large traffic-carrying capacity, the excess capacity need not lie idle, but can be put to use immediately to the maximum extent possible, pending the growth of demand, by correspondingly reducing the user charges associated with it for the time being. On the other hand, if a particular artery is underbuilt, again either through error of forecasting or because of the difficulty of scheduling an appropriate addition at the precise time it is needed, this need not result in undue congestion, since traffic can be held appropriately in check by a corresponding temporary increase in user charges, to be abated when the new facilities become available. It will no longer be necessary to alternate between conditions of undue congestion just before new facilities are completed and substantial unused capacity immediately following the completion of a new facility, nor will it be necessary to build so much in advance of actual traffic growth in order to avoid congestion.

Implications of Highway User Charges for Choice of Transit Medium

If highway user charges of this sort are considered as a possible element in the solution of the traffic problem, the balance of considerations bearing on the choice of rail or bus as the appropriate means of mass transit for given arteries is considerably affected, though not all of the considerations bear in the same direction. On the one hand, the existence of a method of limiting the flow of automobile traffic to that which can be accommodated without undue congestion makes it less necessary to insulate buses from the effects of possible congestion, thus obviating the need for specially reserved bus lanes, and in many cases making it possible to accommodate buses without as many additional lanes as would otherwise be necessary.

Moreover, given the possibility of adjusting user charges to effect an appropriate distribution of traffic between bus and auto, the maintenance of the superiority of buses over automobiles with respect to speed will no longer be crucial, and a moderate amount of congestion can be tolerated much more readily than if such congestion would be likely to have the effect of diverting passengers from transit to private automobiles, further increasing congestion in a vicious circle. The advantage of rail transit in being largely unaffected as to speed by the development of highway congestion would be to this extent set aside.

On the other hand the staging problems presented by the necessity of adding rail transit in units of very large capacity may be greatly mitigated by the possibility of using the user charges to funnel large volumes of traffic to rail transit immediately after the completion of the facility. This could be done, for example, in such a manner as to keep the traffic volume to the then capacity of highways, so as to obviate the need for any further highway construction for some time thereafter. Moreover, the uncertainty connected with possible errors in traffic projections would be reduced if there exists a means at hand of influencing the choice of travel mode so as to assure the division of the total traffic in the proportions planned for, or nearly so.

There is also the further consideration that a program involving highway user charges is likely to rely more heavily on mass transit than one that does not, and that of course the greater the volume of riding on mass transit, the more favorably does the use of rail transit show up in terms of cost. Still another point along this line is that if highway user charges are used primarily with respect to rush hour traffic, transit is likely to be more than ever a rush-hour facility, non-rush hour traffic being more than ever likely to prefer private auto travel as long as there is neither a charge levied nor serious congestion to contend with. Rail transit is likely to show itself better capable of dealing with such a heavy peak and light off-peak situation than bus, in part because of the lower crew requirements per passenger in the longer rush-hour trains than in the shorter slack-hour trains, so that the problem of fitting crew schedules to union requirements as to full working days and limitations on split shifts will be less severe.

Highway User Charges and the Financing of Transportation Facilities

The imposition of user charges along these lines of course will require the reconsideration of plans for financing the various facilities. Thus far the rule has generally been that except for specially costly facilities such as bridges and tunnels, Federal funds have not been available for highways for which a toll has been charged. It could reasonably be argued that in terms of the higher cost of urban expressway construction the situation is more nearly akin to the bridge than to the highway in open country. If this argument cannot be made to prevail, then the system of user charges faces the serious disadvantage from a purely local viewpoint of tending to involve the forfeiture of a considerable share of Federal highway funds.

Aside from the problem of avoiding the forfeiture of Federal aid, there is the problem of organization for the determination, collection, and distribution or expenditure of the user charge revenues by the District and the two States involved. Ideally some kind of joint interstate agency would seem indicated; such an agency might or might not also concern itself with mass transit, preferably it should. But technically, at least, there would be no particular difficulty in assessing the user charges separately for the portions of the highway system lying in each jurisdiction; this could be done either on a completely independent basis, with additional traffic detectors located at the boundaries, or the records for the system as a whole could be analyzed to show what charges are related to the portions of each trip within the several jurisdictions. If the several jurisdictions determine their respective charges independently of one another, there is the difficulty of so adjusting the separate charges that the overall effect on traffic would be as desired, while at the same time an equitable distribution of the proceeds results. Some kind of coordination would thus be essential, although it could, in a pinch, be merely in the form of consultation and informal cooperation rather than a rigid interstate organization, if the barriers to such an interstate organization are too great to overcome. With independent action possible, however, there would be the ever- present temptation for one jurisdiction to raise its charges unilaterally with the aim of securing a larger share of the revenues for itself, a temptation that if yielded to would result in the breakdown of the system of charges as a means of inducing the efficient utilization of the facilities.

APPENDIX

Derivation of a Mathematical Formula for the Determination of the Optimum Transit Fare Structure Under Conditions of Automobile Substitution and a Limit to the Amount of Transportation Subsidy[1]

Let the various types of traffic to be carried between various points at various periods be designated by a subscript i ($i=1,2,\ldots,n$); let Q_i be the total amount of traffic of type i to be carried, this total being fixed for each type i of traffic and subject only to a division between the amounts q_i, that will travel by mass transit and the amounts $a_i = Q_i - q_i$ that will travel by private car. Let the total cost of providing transit services of a given level of quality be $T = T(q_1, q_2, \ldots, q_n)$, and $m_i = \frac{\partial T}{\partial q_i}$ be the marginal cost of the ith category of traffic. Similarly let $A = A(a_1, a_2, \ldots, a_n)$ be the total cost of providing automobile services of a given quality, depending on the pattern of use, and let $b_i = \frac{\partial A}{\partial a_i}$ be the inclusive incremental cost of carrying traffic of type i by automobile.

Let p_i be the fare charged for a transit ride of type i, and assume that each q_i depends on the corresponding p_i, and not on the other p_j (we thus ignore the possibility that a ride may be shifted to an off-peak period because of a fare differential). Let c_i be the out-of-pocket cost to the motorist of an automobile trip of type i. We will measure the difference between the value of a motor trip and that of the corresponding transit ride to a given individual by the difference between this out-of-pocket cost c_i and the specific fare p_i at which this individual is just induced to switch from auto transit (or vice versa). Suppose we begin with a situation where transit service, though theoretically available, is priced prohibitively high so that all travel is by automobile; let us call the value of the service rendered V_1. Now let us lower the price from the prohibitive level to the optimum level p_i^o by successive small increments; for each increment in traffic that shifts from automobile to transit as the corresponding transit fare is reduced, the increment to the total value of the service will be $dV = (p_i - c_i)\,dq_i$, and the total value of the service becomes

$$V = V_1 + \sum_i \int_{q_i=0}^{q_i=q_i^o} (p_i - c_i)\,dq_i$$

We wish to maximize the amount by which this value of the service V exceeds the total cost of providing it, $T + A$, subject to the condition that the total subsidy to transportation be held fixed at a given value. Let the fixed contribution of motorists to their costs, i.e., that part of their automobile expense that does not figure in out-of-pocket costs, be designated F. The subsidy will then be $T + A - F - \sum_i p_i q_i - \sum_i c_i a_i$. Using a Lagrangian multiplier k, we can state that the necessary conditions for the problem to have a solution correspond to the necessary conditions for $W = [V - T - A] - k\,[T + A - F - \sum_i p_i q_i - \sum_i c_i a_i]$ to be a maximum for a suitable value of k which will depend upon the stringency of the restriction on the subsidy. These conditions include $\frac{\partial W}{\partial q_i} = 0,\ i = 1, \ldots, n$. Performing this partial differentiation, noting that the c_i are constants, that $a_i = Q_i - q_i$ with Q_i constant, so that $\frac{da_i}{dq_i} = -1$, and that p_i is functionally dependent only on q_i, we have:

$$\frac{\partial W}{\partial q_i} = p_i - c_i - m_i + b_i - k[m_i - b_i - p_i - q_i\frac{dp_i}{dq_i} + c_i] = 0$$

If we put $e_i = -\frac{p_i dq_i}{q_i dp_i}$ for the elasticity of the ith type of traffic, we can write this as

$$[(p_i - m_i) - (c_i - bBi)]\,(1+k) = k\frac{p_i}{e_i}.$$

From this general formula we can derive three interesting special cases. First, if the limit on the subsidy is either removed entirely or made so liberal that it is not a constraint, then we can put $k = 0$, and have $p_i - m_i = c_i - b_i$; that is, the excess of the fare over marginal cost must for each type of traffic be equal to the cost which the corresponding motorist shifts to others. Second, if we assume that the motorist is made to bear in all cases the entire marginal cost of his trip so that we can put $c_i = b_i$, then we can write $\frac{p_i - m_i}{p_i} = \frac{k}{(1+k)}\frac{1}{e_i}$; that is, the relative excess of the fare for each type of traffic over its marginal cost must be inversely proportional to the elasticity of that traffic. Third, we may suppose that two types of traffic, i and j, have equal elasticity, and ask under what circumstances would a lack of fare differential be justified. Putting $p_i = p_j$ and $e_i = e_j$, we have

$$-m_i - c_i + b_i = -p_i + \frac{k}{(1+k)}\frac{1}{e_i} = -p_j + \frac{k}{(1+k)}\frac{1}{e_j} = -m_j - c_j + b_j$$

or $(b_i - c_i) - (b_j - c_j) = m_i - m_j$; that is, differences in costs shifted to others by motorists must correspond to differences in marginal cost of transit rides.

ENDNOTES

1. M. Boiteux, in "Sur la Gestion des Monopoles Astreints à l'Equilibre Budgétaire," *Econometrica*, 24, 1956, 22-40, was the first to pose and solve the Ramsey-Boiteux problem of second-best efficient pricing in the presence of decreasing costs and a deficit constraint. Vickrey's derivation is independent, and is the first formal application of the theory of the second best to urban transportation.