CONGESTION CHARGES AND WELFARE

Some Answers to Sharp's Doubts

By William Vickrey

In his examination of the case for a "congestion tax" as presented in the Smeed report,1 Clifford Sharp2 manages to convey the impression that, in view of its many complexities, the case for the use of road pricing as a crucial element in the control of congestion is not all that strong, and that, if some such scheme were to be adopted, it could only be a "very partial solution". Moreover, he concludes that putting price equal to a marginal social cost defined in terms of congestion will not produce a satisfactory approach to the maximum net yield in social benefit. But many of the considerations adduced by Sharp in support of his position are open to serious question; and some of them, on examination, turn out to be capable of being used as arguments for more, rather than less, reliance on congestion charges than would be indicated by the marginal cost criterion alone.

market processes in the usual sense. If for purposes of analysis we suppose that roadway space is sufficiently divisible to permit the establishment of a perfectly competitive market, we could expect to find a multiplicity of roads with a variety of toll rates and a variety of degrees of congestion to suit individual tastes, and in equilibrium the toll on each road would at one and the same time equal, in the absence of fluctuations in demand, both the long-run average cost per vehicle of providing the road and the short-run marginal quality deterioration costs associated with an increment of traffic on the road. The price so arrived at would fully represent the quality deterioration aspect of the cost.

This may be seen as follows. The revenues for each road would have to just cover total costs, otherwise the profits or losses would induce changes in the pattern. Thus 
\[ pq = cw, \]
where \( p \) is price per vehicle, \( q \) is the number of vehicles per hour, \( w \) is the capacity or width of the roadway, and \( c \) is the constant average cost of the roadway per unit of capacity. This can be written 
\[ p = cw/q = c/h, \]
where \( h = q/w \) is the density of traffic relative to capacity. Thus the price on the various roadways will vary inversely with the density of traffic. To evaluate the marginal cost in terms of the deterioration in the value of the service resulting from an increment in traffic, we can first put \( c_i(h) \) for the value placed on the \( i^{th} \) use by the user when that use is at a density \( h \), this value being defined as the price that the user would be willing to pay to travel at this congestion level rather than not make the trip at all. Then 
\[ (dq_i/dh) \]
the change in the value of the \( i^{th} \) use resulting from an increment in traffic density \( dq \) produced by an increment of traffic \( dq = w \, dh \) on the roadway used by this \( i^{th} \) use. Now for any use \( i \) of a particular road priced at \( p_i \) we must have 
\[ dq_i/dh = dp_i/dh, \]
othewise the user would find it to his advantage to move either to a higher priced and less congested road or to a more congested but lower priced road, depending on the sign of the inequality. In the one case he would value the gain in quality more than the increase in price, and in the other he would consider that the reduction in price more then compensated for the loss in quality. Thus, if the variety of roadways available provides a continuous spectrum, then, although different uses of the same roadway may have a different absolute value, the differential between value of each use at density \( h \) and what it would be for adjacent densities, expressed by \( dq_i/dh \), must be the same for all users of the same roadway, and equal to \( dp/dh \).

Accordingly, the marginal cost \( m \, dq \) of an increment of traffic \( dq \) as measured by the total value of the impairment in the quality of the service enjoyed by other users of the roadway, associated with the resulting increase in traffic density \( dh = dq/w \), is given by 
\[ m \, dq = (c/h) \frac{dq}{dh} = (c/h)(1/w) dq = (c/w)(-d(c/k)(dh)/dq) = hck^{-2} dq \]
\[ = (c/h) dq = p \, dq; \text{ thus } m = p. \]

Thus, in a long-run competitive equilibrium, short-run marginal congestion cost and long-run average cost are equal, and it makes no difference whether the charge is viewed as representing the average cost of providing the roadway or the marginal cost inflicted on other users. Of course, the two measures of cost are alternatives, not to be added together.

Moreover, in a short-run competitive equilibrium in which total roadway capacity has not been adjusted to the demand, price on the different roads will still vary inversely according to intensity of use, at a level higher or lower than the long-run equilibrium according as capacity is below or above the long-run equilibrium; the proprietors of the roads will be receiving Marshallian "quasi-rents". Any deviation
from the relation \( p = k/h \) would be symptomatic of a failure to maximize the quasi-rent \( k \). Traffic, being ignorant of the cost of the roads, will again be distributing itself over the various roads in such a way that \( dp/dh = dv/dh \); the previous argument that price will represent marginal congestion costs still holds. Thus, if "the normal processes of the market" is interpreted to mean a perfectly competitive market, then indeed the congestion-determined charge is identical with the market-determined price, and has much greater claim to be termed a price than has a charge based on average historical cost.

If there were diminishing returns, occasioned possibly by a shortage of available space or favourable alignments, quasi-rents would remain above costs in the long run, the excess representing land rent for the superior locations. Even in this case, the congestion charges, being inclusive of a rent element, would be more nearly analogous to the price of, say, corn than would a "cost"-based charge that included no rent element. (A congestion charge could be considered to include a "tax" element in the strict sense only if we adopt the convention that what is rent when received by a private person becomes a tax when it is received by a government. This would, however, be a confusing usage.)

Where, as in some less populous areas, there are substantial economies of scale and increasing returns, marginal cost lies below average cost and optimal congestion tolls would, given optimal levels of capacity, be below the "average cost". If we adopt the Lerner terminology, the amount by which a price set equal to average cost exceeds the marginal cost (which at optimal equilibrium would be the same whether based on outlays for construction or on congestion effects) could properly be termed a quasi-tax, the revenue from which would be a counterpart of the "negative rent" arising from the economies of scale.

Where the traffic has a periodic fluctuation pattern and prices can fluctuate accordingly, a competitive market would generate prices which cannot be calculated from any cost-of-roadway considerations, but would correspond to the congestion costs at each moment of time, assuming the pattern to be regular and predictable. Such prices could be interpreted as covering a rental value for the roadway at each moment of time, the rentals over a cycle being in the aggregate just sufficient, in long-run equilibrium, to cover the costs. There would, however, be no way of calculating the price pattern from the roadway cost without reference not only to the time pattern of demand, but also to the impact of demand on congestion and the value of different levels of service as revealed by the users in their shifting among routes of varying degrees of congestion. In such circumstances, congestion costs become the measure "par excellence" of what a price structure arrived at by the normal processes of the market would tend to be.

Of course, where economies of scale make it uneconomical to offer more than one level (or, in some cases, a relatively small number of levels) of service, it becomes difficult to find an objective test by which the relative value of different qualities of service to the road users can be measured. But this practical difficulty stems merely from the presence of economies of scale; the conceptual mensurandum itself is just as definite and tangible as it would be if the variety of facilities that would test this value could actually be provided without excessive cost. In other contexts, a variety of service can be arranged, even if only temporarily or experimentally: the value of increasing the frequency of a bus service, for example, might be tested by designating
alternate bus stops as red and green stops, and similarly designating alternate buses as red and green buses, charging a lower fare if a passenger boards a bus of a colour corresponding to that of the stop and a higher fare if he boards a green bus from a red stop or vice versa, and observing how many would wait for the next bus rather than pay the higher fare, for different levels of the differential. Utilization would of course be non-optimal for the duration of the test, so this procedure could be only a temporary experimental expedient and would be subject to biases due to its temporary nature. Comparable experiments with road service would be more expensive to arrange, but some measurements can be teased out of studies of traffic splits between toll and free routes, and of traffic changes in response to changes in toll rates. Indeed, if a general system of congestion tolls were adopted, much information on this point could be gleaned from a study of the reactions of traffic to toll changes as these were adjusted in a tâtonnement process in the search for an approximation to the equilibrium pattern.

Of course, in the competitive case where a continuous variety of service qualities is offered, the patrons of routes with a given level of congestion are a self-selected group, placing at the margin a uniform value on an increment in the quality of service. In practice the choice of quality offered is more limited, and this value will have a more or less wide range. Nevertheless, so far as congestion cost is concerned, the relevant figure is a simple average of these values over the units of traffic concurrently using the facility, and this same average becomes the basis for a uniform charge to all users regardless of their individual tastes (assuming, of course, the single user to be a negligibly small unit compared with the total traffic to which a specific element in the charge structure is to be applied).

II. THE USE OF THE REVENUES

Further on, Mr. Sharp apparently considers the funds yielded by the congestion charge to be an embarrassment rather than an advantage (pages 809-811), partly because there might be considerable political pressure for using the revenues to reduce other motor vehicle taxes, and if this is done the beneficial effects of the congestion charges could, he feels, be seriously impaired by motorists using the funds thus released to “buy back” congested road space. But this represents a rather peculiar picture of the motorist as a person who has a fixed budget allocated to motor transport and will insist on pouring back into further expenditure in this same category any savings realised as a result of a reduction in a particular charge. A more conventional form of analysis would be in terms of income elasticities: assuming that total outlay on motor travel in congested circumstances amounts to 5 per cent of the incomes of those affected by the motor vehicle tax reductions, that the income elasticity of this category of expenditure is 3, and that the congestion tax amounts to half the total expenditure on congested travel – all of which seem wildly liberal assumptions – the “buy back” effect will be 7.5 per cent, rather than the 34.8 per cent suggested by Sharp. If, indeed, congestion charges result in relatively little curtailment of demand, this is better regarded as inherent price inelasticity of demand rather than as the effect of a greater initial response of the demand to the charge offset by a “buy back”, or income effect.
More relevant would be the idea that the other motor vehicle taxes are to some degree complementary to the congestion charge, and that the effective price for the overall trip is a composite of various elements of cost, so that some of the congestion charge would be directly offset by the reduction in other motor vehicle taxes. But this is basically merely a caution that to achieve a given level of effectiveness in the face of these offsetting adjustments the congestion charge might have to be made a little higher than if the motor vehicle taxes were left unchanged and the revenue used in some other way.

Sharp considers that the purchase tax, and presumably, by extension, the annual licence fees, would be illogical targets for reduction by use of revenues from a congestion tax (page 809). Actually the justification for substantial annual licence fees, as against higher fuel taxes, is that cars driven relatively few miles are likely to contribute more to congestion in proportion to miles run than cars driven more extensively. If congestion charges become reasonably widespread, so as to cover the bulk of cases where significant congestion occurs, there is very little case for licence fees, and the fees can appropriately be reduced to a level representing administrative and other similar costs associated with registration.

More important is Sharp's criticism of the Smeed report for disregarding the yield of the congestion tax and treating it as a pure transfer payment which represents no real cost to the community. Sharp appears to feel that this transfer involves a positive cost (page 811). Actually, aside from the explicit costs of the collection procedure, transfer of funds to the State is likely on balance to have positive benefits, because in general it enables excess burdens attached to other sources of revenue to be abated. Indeed, on this ground, once the explicit collection cost has been covered, it would become desirable to push the congestion charges above the level representing marginal congestion cost to the point where the "marginal cost of public funds" derived from the congestion charges, resulting from traffic being depressed to slightly below the optimum level, becomes equal to the marginal cost of public funds derived from other sources. How far the charges should be pushed on this criterion would of course vary according to the jurisdiction involved. Under conditions prevailing in the United States, for example, with core cities generally extremely hard pressed for tax revenues which will not have undesirable repercussions, revenues from a congestion tax, if made available to cities, could provide an extremely welcome means of abating the baneful effects of local sales taxes, high property taxes on improvements, and a number of miscellaneous imposts on business activity which are often far from neutral in their impact. In Britain, if the revenues were devoted to a reduction in fuel tax, this would probably be beneficial on the whole, as it would reduce the under-utilization of the relatively uncongested streets and roads. In effect, the price for the use of congested streets should be pushed above marginal cost by an amount similar to the corresponding excess, if any, that would persist in the price charged for the use of uncongested roads.

III. DISCONTINUITIES OF DEMAND

Sharp claims further that "the case for a congestion tax assumes that demand varies continuously with price, or that each motorist puts a different valuation on travelling
time" (page 812), and also that it "is based on the value judgement that some people's time is more important than others" (page 813). While variations in the value of motorists' time increase the potential benefits derivable from congestion charges, the case would remain adequate even if motorists were completely homogeneous in income and in the preciousness of their time. Nearly always there are wide variations between different users of a facility in the closeness of the alternatives available and in the amount of time that they will save by using the congested facility rather than an alternative. Congestion charges are necessary to achieve optimum use of a roadway network, even if this is defined merely as minimizing the total time needed to complete a given transportation task.

Sharp further suggests that, if demand for a given type of trip is completely inelastic over certain price ranges, the administrators might be faced with the dilemma of removing either too few or too many vehicles from the road (page 812). Actually, what would create this dilemma would be a demand that was perfectly elastic at a given price over a range of quantities; that is a different matter, and a situation much less likely to arise in practice. It is conceivable that there might be a situation in a relatively small community where there is only one feasible route for a given class of trips, no public transport, and a dispersal of origins, destinations and times of travel such as to make car pooling unfeasible, so that changes in congestion charges over a fairly wide range would have little or no effect on traffic. It is somewhat more difficult to imagine that as the charge is pushed up past a critical point a large bloc of traffic vanishes, though this might occur if a dominant employer were simply to move out because of the toll; even here, a more likely reaction for a dominant employer would be to stagger his working hours, which presumably need not be done on an all-or-nothing basis. But, in any case, the chief field for the application of congestion tolls is likely to be in the major metropolitan centres or on elements of a network where reasonably good alternatives are available. Even in the isolated case where alternative routes are not available, many trips will have at least some flexibility as to time.

Even where the dilemma put by Sharp exists, the optimum resolution will usually be not to abandon pricing altogether, but to select the better of the two alternative levels of charge. Suppose, for example, there are 2,000 persons having to make trips involving a given congested segment during a given period, and that if they all make the trip in their individual automobiles the trip will take 40 minutes and cost 4s. per person in operating cost. Suppose further that two persons could get together and double up in a car pool to save 2s. each in operating cost, but if the inconvenience of joint scheduling and the time lost in the pick-up process is valued at say 3s. per person they will not do this, since there will be no time saved for them on the congested segment. Moreover this pattern would not be changed by any toll less than 2s. A toll of 2s. 1d. per car, however, would push the savings from pooling above 3s. per person and make it worthwhile for everybody, so that only 1,000 cars would make the trip, congestion would be reduced, and the trip would take only 30 minutes. If the time saved on the trip is worth as much as 7s. 1d. per hour, everyone is better off with a toll of 2s. 1d. and pooling 2 persons per car than with a toll of 1s. 11d. and everyone riding singly; if the time savings were worth as much as 18s. 7d. per hour, everyone would be better off with the 2s. 1d. toll than with no toll, and the entire revenue would be derived at no sacrifice to the contributors.
CONGESTION CHARGES AND WELFARE

William Vickrey

The only difficulty is that an attempt to make the toll correspond to the marginal congestion at a given time might result in an oscillation between tolls of 2s. 1d. and congestion costs somewhat lower than 2s. 1d. and tolls of 1s. 11d. and congestion costs much higher than 1s. 11d., whereas the ideal solution would be to have a toll of 2s. and manage to have just enough car pooling to bring the congestion cost to that level. In practice one could probably find a sufficient amount of variation between the case with which various individuals could form car pools, so that something approaching this could be achieved. The result might then be that half the travellers would double up, resulting in a traffic level of 1,500 cars making the trip in 31 minutes. In any case, the problem is no different from that arising whenever a commodity or service with an inelastic supply meets a demand curve that is perfectly elastic over a range that includes the equilibrium quantity.

The opposite case, of inelastic demand, represents an opportunity rather than a problem, because on second best principles it will then be possible to raise the charge to a point near the upper end of the range of inelasticity and use the revenues thus derived to reduce the excess burden of taxation in other areas without significantly adding to the excess burden in this area.

IV. DISCONTINUITIES OF SUPPLY, GROWTH AND PLANNING

The more important discontinuity, indeed, is on the supply side, and here again the congestion charge may have an important contribution to make. In many instances there will be, in the absence of congestion tolls, a dilemma between providing a facility so limited that serious congestion results and providing one with substantial excess capacity. And in a dynamic situation it is almost impossible to programme additions to capacity, at any cost, in such a way as to avoid undesirable degrees of congestion without shoudering the burden of having at any given time a substantial amount of capacity that is grossly in excess of current needs. One potential benefit from congestion tolls is to keep congestion on an existing facility within tolerable limits for a longer time, so that, when an addition to capacity becomes justified, tolls can be so adjusted that the new facility will be more nearly fully utilized from the time of its completion.

Sharp also fears that the levy of congestion charges would provide revenues of a kind that would generate a political demand for additional highway construction which might or might not be justified. There is indeed a danger here, but surely the answer is not to resign oneself to obvious inefficiency in the use of what one has, but to work for a more rational and integrated approach to the whole problem of urban transport and finance. Possibly this would be an added reason for devoting the revenues to reductions in other specifically urban taxes, such as rates, on the ground that to do so would provide a closer offset to the new burden for individuals than would the reduction in other vehicular levies, which might go to a large extent to the rural population. The fear that congestion levies might overburden the available public transport facilities (page 815) seems somewhat fanciful. Surely the reduction in congestion should make bus service more attractive in comparison with rail, at the same time releasing the road space to permit added bus service to be offered; the skill required to operate a bus is surely not as scarce as all that.
The availability of a well-designed system of congestion tolls may also make an important contribution to the efficiency with which expansion of the highway system is designed, a point largely overlooked in the Smed report. Without tolls, expressways penetrating close to or through city centres are likely, if uncongested, to attract a considerable amount of through traffic that would be more economically accommodated on circumferentials; the planner may be faced with the dilemma of building the central artery in such a way that it becomes undesirably congested, or providing at inordinate expense for a central artery that will accommodate through traffic. Or, without tolls, a circumferential that is to attract traffic away from the central district may have to be built close in, at much greater expense than if tolls could be applied and traffic thereby induced to use a cheaper circumferential built at a greater distance from the centre. It is the overlooking of this kind of possibility that seems to be responsible for Sharp’s statement that “There cannot be very much scope for through traffic to find alternative non-congested routes in Britain” (page 813). If true, this would seem to indicate that the construction of circumferential routes has lagged, possibly in part because in the absence of tolls on the congested routes it is more difficult to attract enough traffic to them to make their construction sufficiently worth while.

Indeed, in the short run it is precisely where a road without close alternatives is overburdened that a congestion charge may be salutary by inducing more travellers to share in car pools or use buses: more people may then be moved faster in fewer vehicles, and this result is likely to be advantageous even if all travellers value their time at a like figure. Sharp’s claim that if time has a constant value “it is no longer true that it is better for fewer people to be travelling faster than for more people to be travelling more slowly” (page 816) is patently dubious. If 1,900 persons make a given trip which has a uniform value to them of £2 and takes one hour at a value per hour of 10s., the net gain is £2,850; if 2,000 go and the trip thereupon is lengthened to 75 minutes, the value of the trip is £4,000 less the travel cost of £2,000 (1) (1) = £1,250 or a net gain of £2,750. The 2,000 persons wishing to go would do better to draw lots to see which 100 persons should be kept from going and compensated to the extent of £1 8s. 6d. each, which is then the gain each gets; if they all insisted on going each would gain only £1 7s. 6d.

V. REACTIONS OF GOODS TRAFFIC

In many cases, however, it is not so much a matter of alternative routing as of alternative timing. I have heard it said that even under existing conditions the cost of night deliveries in Manhattan, after paying premium wage rates, is still less than half that of daytime deliveries, the difference being largely due to congestion costs. Since marginal congestion costs tend, under these conditions, to run from three to five times the average congestion cost, congestion charges reflecting costs at these levels might well be sufficient to overcome the reluctance of consignors to provide facilities to receive and despatch outside congested hours. Sharp’s doubt that “the existing haulage industry is extraordinarily inefficient” (page 813) may perhaps be correct if interpreted from the point of view of the individual firm reacting to the costs and prices as it finds them. It is certainly not correct in the aggregate, given the strong
CONGESTION CHARGES AND WELFARE

William Vickrey

externalities in the situation, and there is room for a great deal of congestion-charge-induced economizing, not only by utilization of off-peak hours but also by consolidation of deliveries, greater use of off-street loading bays, and minimization of delays in loading and unloading at the curbside. This is true at least in Manhattan, and I suspect elsewhere.

The most surprising of Sharp’s objections to congestion charges, as applied to commercial vehicles, is that they would have “fairly widespread inflationary consequences” (page 813). Conceivably this would be true in the short run if the funds collected were spent on additional road construction or used to reduce income tax. But if the revenues are used to reduce other taxes, whether specifically motor vehicle taxes or even other taxes impinging on business activity, such as purchase tax in general or corporation tax (or even rates, if the yield can be used as a local revenue), it would seem that costs in these directions would in general go down by at least as much as transport costs went up, and that on balance costs would have to go down if the objective of increasing the efficiency of use of road space was realized at all. And in the long run even the transitory effects of income tax reduction or road construction would be likely to have a net deflationary effect from increased efficiency. To be sure, effects on different industries would vary according to their use of congested road facilities. Home deliveries of bread and milk, with the possible exception of deliveries to West End flats, seem to be one of the activities less likely to succumb (page 817): residential streets are seldom congested to a degree that would justify any substantial levy of congestion charge. Home deliveries of furniture and domestic removals might incur significant charges, since the use of urban arterials is a larger part of the total movement; but here the demand is unlikely to show much elasticity. More important, the rationalization of the pattern of business activity might be stimulated, activities generating relatively large amounts of traffic (such as warehouses) being induced by the tolls to move to less congested areas, leaving room for activities better suited to metropolitan cores. But it is hard to think of such changes as being on balance baneful, except on the basis of an overpowering enchantment with the status quo.

VI. INCOME REDISTRIBUTION, QUEUEING AND PRICE RATIONING

Perhaps the most appealing objection to congestion charges is that they would tend to exclude from the highways those who can better afford to pay for this privilege in terms of time, inconvenience or discomfort than in terms of coin of the realm. It does indeed seem somewhat harsh to deny to those with more time than money the opportunity to buy road space with that coin of which they have a relative abundance. But while payment in money gives to the recipient, even if he is merely a representative taxpayer, what is taken away from the payer, time that is given up in queueing does not become available to anyone else and is an economic waste. One could as well offer priority in the use of roads to those who would be prepared to travel in springless and unupholstered wagons. The analogy of the surgery queue (page 813) is not quite apposite, for in the provision of health services there are strong
elements of both the "merit want" concept (it is thought that individuals systematically underestimate the value of the service to themselves) and of the external economies or positive neighbourhood effects represented by the general desire both to reduce the risk of infection through contagion and to avoid the generation of burdens on public welfare institutions through neglect of potentially disabling conditions. Moreover there is usually the possibility, for those who can ill afford to wait in the queue, of arranging for treatment outside the National Health Service. While it might in principle be desirable to have alternative highways available with different levels of congestion, offering the traveller a similar choice between paying in time and paying in money, the economies of scale involved make this impractical as a general solution. Even in the surgery, one could probably find a better way of rationing the service than the deterrent effect of the queue (at least in principle, as by a schedule of fees related to income). (Of course, we are talking here about extended and persistent queues, and not about that minor degree of queueing associated with smoothing out random fluctuations and justified as a necessary cost of assuring a high level of utilization of the available services.)

There is very little either of the merit want element or of positive externalities to justify a special concern for those who might find it especially burdensome to pay for highway use. If there is a felt injustice in the withdrawal of the established privilege of the free access to the highway and street, this is a short-run phenomenon stemming from a change from a given status quo; any long-run injustice is better expressed in terms of given groups having unduly low real incomes, rather than having specific privileges denied them. The situation is not unlike that attending land enclosures, as expressed in the old ditty:

The law condemns that man or woman
Who steals the goose from off the common,
But lets the greater felon loose
Who steals the common from the goose!

If these sentiments can find their echo today, at least the benefits of modern enclosure of streets will not accrue to upper-class individuals but to taxpayers generally. And, as with land enclosure, the charging of rent for the use of city streets is a necessity if they are to be efficiently utilized.

Nor is it really correct to say that congestion tolls would bring about a situation where "driving in cities would then be limited to the rich or the privileged" (pages 812-813). The aim of imposing congestion charges is not to make a drastic reduction in the number of trips actually accomplished, but to make a relatively modest reduction, of the order of 10 or 20 per cent, in the number of trips attempted, with the consequence that increased speeds would decrease by 20 to 50 per cent the number of vehicles attempting to move at any one time, and traffic congestion would be decreased to an acceptable level. The picture conjured up of Regent Street cleared of all but Bentleys and Jaguars is quite inappropriate. If we look at the 10 or 20 per cent of the trips that will be "trolled off", a fair number are likely to be the less essential trips of the moderately well-to-do; a substantial bloc is likely to consist of commuting trips of suburbanites who are induced to shift to rail, bus, or car pool, often giving up a second car. Those for whom the crocodile tears are shed turn out, on examination, to be on the whole a quite prosperous class.
CONGESTION CHARGES AND WELFARE

William Vickrey

To some extent, indeed, city driving may become concentrated among the rich and the privileged, to much the same degree, with or without congestion tolls. Passenger automobile traffic in lower Manhattan is already noticeably the special province of the taxi, the doctor, the diplomat, the chauffeur-driven car, and the car with parking privileges somewhere. The high cost of parking already makes a car in lower Manhattan something of a white elephant for the driver of modest means.

In any case, the claim that congestion charges might burden the middle class for the benefit of the rich overlooks the much more important fact that there would be important benefits to the really poor in terms of better bus service, reductions in the less progressive forms of taxation, and reductions in the cost of living resulting from more efficient transportation of goods. There would seem to be no overwhelming difficulty, indeed, in coupling the institution of congestion charges with adjustments in income tax rates which would leave every income group as a class better off.

Where the transitional injustice is felt to be sufficiently acute to require a specific remedy, one can always ask whether, in lieu of giving certain individuals or groups an exemption or a concessional rate, it would not be better to charge them the full rates and give them a cash sum equivalent to the difference between what they would pay at the reduced rates and what they would pay for the same travel at the full rates, leaving them to choose whether they would use this money to finance this same amount of travel or would prefer to divert it to other uses. Only where it was found that the administrative difficulty involved in making these compensatory payments was excessive, and the felt injustices were severe enough to warrant the inefficiency, would the concessional rate be justified. Put to this test, few of the special exemptions often proposed for the clergy, the disabled, current and former members of the armed services, large families or public officials would survive. Not even the Post Office, and perhaps not even fire and police vehicles, should be exempt from the pressure to avoid, as far as possible, undue contribution to the costs of congestion. (The treatment of Coronation parades had perhaps best be regarded as beyond the scope of this paper.)

VII. SCOPE

Sharp seems to feel that, while congestion charges may have a role to play in a few critical areas, the scope for their application is fairly limited. Actually, for most of the schemes that have been proposed, a large share of the total collection cost is represented by the equipment to be carried on the vehicle and by the establishment of collection procedures; therefore, once a congestion charge scheme has been adopted in a given area, the extension of the system of charges to cover contiguous areas can be accomplished at relatively little additional expense. Accordingly there seems to be very little reason for not gradually extending the scheme to cover nearly all areas where significant congestion occurs, with the possible exception of minor amounts of congestion in relatively isolated areas or occurring primarily on a seasonal basis, where inclusion would be likely to bring a relatively large number of additional vehicles within the scheme.

All told, therefore, there seems to be relatively little reason for hesitating to use the best available estimate of marginal congestion cost as a basis for setting the con-
gestion charges, perhaps adding a surcharge, in accordance with the general theory
of second best, to reflect the desirability of obtaining additional revenues so as to allow a reduction in the excess burden imposed by other forms of taxation. Admittedly estimates of congestion costs will be rough, but it is to be doubted whether one can come much closer to an optimum than by this uniform rule. There will, of course, be important and often unpredicted side effects stemming from the institution of such an important change in the pattern of transport costs; in the absence of specific indications to the contrary, however, there is surely a presumption that reactions to a system of charges which comes closer to representing real costs will in the long run be more often desirable than otherwise, and those who would use such unpredicted reactions as a basis for keeping congestion charges low must sustain the burden of proving that the benefits of the higher charges would not outweigh the injuries. It may of course be desirable to initiate the charges at a level somewhat below the estimated optimum level and increase charges gradually until the optimum balance is achieved, to allow reactions to the charges to be observed and estimates of the ultimate optimum pattern to be revised and to avoid excessive disturbance from possibly overshooting the mark on the basis of initial uncertainties. After the initial transition period, however, wherever there are charges at all, they should preferably be set above, rather than below, the level reflecting estimated marginal cost.

There will be, no doubt, problems involved in the application of any system of congestion charges. It must be decided how detailed the schedule of charges should be; what means should be used to inform the motorist of the charges he is likely to incur; whether the charges should be fixed by schedule in advance or should be made to vary with conditions actually experienced; whether they should be based on mileage or on time, or both; how to integrate charges for movement in congested areas with charges for parking and standing; to what extent weather conditions should be taken into account; whether areas favoured with new facilities representing a substantial increment in capacity should be favoured with lower congestion charges as well. There is also, to be sure, the question how to measure congestion cost, but this is not a problem raised only by the intention to levy a congestion charge; measurement of congestion cost is an essential part of any cost benefit analysis used as a basis for intelligent planning of new facilities and the maintenance of existing ones, though perhaps congestion charges require more detailed investigations into the variation of costs between peak and off-peak periods. But, however these questions are resolved, road pricing on the basis of marginal costs is likely to prove the crucial key to the solution of the problem of efficient use of road space wherever significant congestion occurs.

*Columbia University, New York*
I am grateful to the Editors of this Journal for giving me space to reply to Professor Vickrey's criticisms of my views on the proposals for the introduction of a congestion tax. In restating my arguments I have taken the opportunity to range a little more widely than the specific points made by Professor Vickrey.

There have been some important contributions to the discussion on a congestion tax since I wrote the article which appeared in the Economic Journal in December 1966. These include the report of a Ministry of Transport study group on the means of restraining traffic on urban roads, an important study of the possible effects of the imposition of a special licence charge on private cars entering an area of Central London (and a comparison with the effects of a parking tax) by J. M. Thomson, and a comprehensive and impressive survey of the whole subject by Professor Beesley at the European Conference of Ministers of Transport. I restrict my present discussion of Professor Beesley's paper mainly to his specific criticisms of my article.

My basic position is that of a critic, but not an opponent, of the congestion tax proposals. I believe that there is a very strong case for restraint of the use of private cars in towns, and that in some areas a congestion tax may be the best form of restraint available. But there are a number of points in the theoretical case for a congestion tax made in the Smedd Report and elsewhere with which I disagree, and some aspects of the practical proposals for the actual application of the tax which I would question. My original aim was to criticise what I considered to be the oversimplified and uncritical approach of the Smedd Report. I was most interested to read the recent comment by C. B. Winsten, himself a member of the Smedd Panel, that the Panel "seemed to me to suffer from the difficulty that it was too much composed of road-pricing enthusiasts and they were less critical of the complicated and perhaps in some cases unworkable devices, mostly electronic, which it was considering". Better Use of Town Roads, although still favouring the use of a congestion tax, is noticeably less enthusiastic than the Smedd Report, particularly in the conclusion that "Road pricing cannot, therefore, be a quick solution and it would probably be unrealistic to expect that it could take much less than six or seven years to bring a system into full operation". The Minister of Transport's Foreword is even more non-committal. "It is not certain that road pricing can offer a solution; it will

---

7Op cit., page 35.
in any case not be an immediate one”\textsuperscript{8}. The pioneering study by J. M. Thomson of the possible effects of a limited form of “pricing” in a specific area is the kind of development which, by asking questions about actual results rather than relying on the achievement of a theoretical optimum, would remove most of my objections to the introduction of a congestion tax.

I tried to argue in the Economic Journal article that, if the congestion tax is a price, it is in some ways an unusual one. In following up this point I wrote, probably unwisely, that a profit-maximising road-and-vehicle-providing undertaking would not arrive at the “optimum” price suggested by congestion tax theory. Professor Vickrey agrees that this is true for a monopolistic corporation, but shows that, if one could imagine a number of competing road corporations providing “a multiplicity of toll rates and a variety of degrees of congestion”, that price would reflect the deterioration of quality associated with an increased traffic flow. The optimization of the flow over a road network which was described in the pioneering work of Beckmann\textsuperscript{9} is, however, a different case from the single road with no alternatives to which my arguments applied. Merely imposing a congestion tax would still leave us a long way from the conditions of perfect competition. Professor Beesley argues that a monopolistic road corporation would differentiate in price between roads with different demand and supply conditions. This is, of course, true, but does not mean that the charge on any one road would be equal to the congestion tax “price”. What the corporation could quite clearly not do is to discriminate between the users of a single road, since it could not provide more and less congested road space simultaneously. I was probably unwise to follow this argument at all, since one could reject a congestion tax even if a road corporation could conceptually arrive at the same “price” or accept it despite proof that no real world corporation would ever impose an equivalent charge. But the original point of departure is not unimportant. A congestion tax would reflect a special kind of cost. Alternatively, if, as Professor Beesley argues,\textsuperscript{10} the costs involved in congestion tax pricing are costs just like any other costs, then so are a considerable number of other costs which the original congestion tax case ignored. Motorists who turn right, learner drivers, cars pulling caravans, underpowered lorries and pedestrians seeking to cross the road all impose costs which are just as real as those caused by vehicles which by their mere presence reduce the speed of a traffic flow. If there are \(n\) costs arising from a flow of traffic, we shall not necessarily come nearer to an optimum resource allocation by raising the number of costs reflected in price from \(x\) to \(x + 1\) when \(n > x + 1\). If travelling time has a cost, so also have road accidents. The basic problem of measuring the results of the 70 m.p.h. speed limit is to evaluate the trade-off between increasing time costs and reducing accident costs. The increased speeds resulting from a congestion tax could lead to more accidents, which might give a net decrease in benefits. This may be unlikely in practice,\textsuperscript{11} but it is an example of a question which needs to be asked about the results of a congestion tax. It illustrates one way in which the achievement of a sub-optimum in the allocation of road space could result in a net decrease in

\textsuperscript{8}Op cit., page vii.

\textsuperscript{9}Beckmann, McGuire and Winsten: Studies in the Economics of Transportation.

\textsuperscript{10}Beesley, op. cit., page 58.

welfare. Much the most important welfare problem, which neither Professor Beechley nor Professor Vickrey solves, arises (as I argued in the Economic Journal article) because we do not start with an optimum distribution of income. If income is not optimally distributed, there is a widely accepted case for public action to attempt to correct some of the malallocation of resources which would result from allowing unrestricted consumer sovereignty. In practice there are areas of the economy where we are prepared to allow the pricing mechanism alone to allocate resources and others, such as the provision of housing, medical care and library books, where we are not, and road space for the journey to work may well fall into the latter class.

The consumer surplus criterion will be inefficient in distinguishing between different situations in relation to the alternative time costs of those who may be tolled off the road, because it multiplies the actual time losses by weights which have suspect values. The congestion tax proposals would in effect weight the 1 hour extra travelling time which might result for a tolled-off £3,000 a year man by 3 (or more) and the 2 hours’ extra travelling time of a man with a salary of £1,000 a year by 1. Because the price mechanism is the best method of allocation we have over a large section of the economy, it does not follow that we must accept unquestioningly the result of this system of weighting where the influence of income swamps that of the relative actual (non-monetary) time costs involved in the reallocation. I agree with Professor Vickrey that there is no element of merit want involved, but would suggest that there is an important external diseconomy (or an external economy if we think of moving away from a full consumer sovereignty “road pricing” allocation) in the time losses of the tolled-off. Once measurement of benefit as “what the consumer will pay” is questioned, then, of course, the whole theoretical basis of the congestion tax case comes into doubt, since the time gains of the continuing “tolled-on” road users may be valued at less than the time losses of the “tolled-off”.

The proposition that income-determined time evaluation is not a wholly acceptable criterion for allocating road space for the journey to work can be made purely on the basis of a value judgement. It can also be supported by the economic argument that, since the levels of wages and salaries cannot be said to represent any economic optimum, they are unlikely to lead to an optimum valuation of time. The output of travellers to work may be interdependent, and resources will not necessarily be more efficiently allocated if the imposition of a congestion tax allows the managing director of a firm to get to work half an hour earlier but also results in the boiler maintenance men being half an hour late, and if the director finds that he cannot make his decisions in the cold.

If it has been established that an “optimum” congestion tax does not automatically optimise anything, we can proceed to the practical question of where such a tax would seem likely to yield a net benefit, and where it would not. We may in this case want to distinguish between roads which differ in the number and quality of transport alternatives available to the average traveller;¹² between roads having

¹²Some of the arguments for a congestion tax in Britain seem to suffer from a “Central London bias”, as a much wider choice between private and other forms of transport is assumed than exists in many provincial cities. There is also the added provincial problem that more traffic may be destined for workplaces outside the city centre, and that the journey time difference between public and private transport for reaching such places may then be very great indeed.
users with different patterns of travel purposes; and also between different classes of road users.

On any given road a congestion tax could yield benefits by persuading those who are free to do so to alter the times of their journeys or not to travel at all; by directing through travellers to alternative routes; and by leading to a greater use of public transport and a more economic use of road space. The opportunities to postpone or cancel journeys must occur almost exclusively amongst “non-business” travellers, who made up only 9 per cent of morning peak traffic in Central London in 1962, and whose complete removal from the roads “could not be expected to make much difference to traffic congestion”.

As I have suggested before in this Journal, one criterion for measuring the effects of redistributing existing travellers between alternative routes, and between public and private transport, would be the effect on average door-to-door journey time. Where it can be calculated that average journey times would be reduced there would seem to be a prima facie case for introducing some form of restraint on private motoring. The practical choice for a means of restricting motorists lies between parking charges or restrictions and a form of congestion tax. The evidence of Mr. Thomson’s Central London Survey suggests that a congestion tax would be more efficient. In those circumstances I would accept the case for such a tax (subject to all the practical difficulties discussed very fully by Professor Beesley in the European Conference article).

There might still, of course, be cases of hardship among poorer paid workers which might need to be remedied, but the fulfilment of the “shorter-average-journey-time” criterion would result in a clear gain to the community in time savings. A tax resulting in longer average journey times would not necessarily be rejected, but the situation would have to be examined very carefully to decide whether the consumer surplus weighting effects were acceptable. Where traffic restraint would be justified by the shorter-average-journey-time criterion, it would almost certainly also be justified by the consumer surplus measurements of the standard case for a congestion tax.

My position on the inclusion of commercial vehicles in the congestion tax system is again pragmatic. Professor Vickrey’s expectations about the results of a congestion tax on commercial vehicle operation are different from mine, and this issue could not be definitely resolved unless a trial tax was actually imposed. There are already penalties, in the form of increased costs, for hauliers who operate with poor load factors or use congested routes when less congested ones are available. A congestion tax would merely increase these penalties. Since there is already very widespread competition in the road haulage industry, it would be surprising if there was much further scope for driving out the inefficient firm. Similarly, all the possible economies which Professor Vickrey mentions (the utilization of off-peak hours, consolidation of deliveries, greater use of off-street loading bays and minimization of curbside loading and unloading delays) are ones which hauliers would already like to achieve. Off-peak work and the use of off-street loading bays can only take place after major policy changes by the consumers rather than by the providers of transport. Professor


Vickrey is quite right in pointing out that the overall consequences of a rise in road haulage costs need not be inflationary if the tax revenue were used to reduce costs elsewhere in the economy. But this is a very important qualification. I would still argue that where a congestion tax was imposed it might be applied to retail delivery vehicles (24 per cent of total goods vehicles mileage in London in 1962), but that the benefits resulting from including general haulage vehicles are very uncertain. Most general haulage vehicles operating in urban areas must use congested roads or go out of business. Their demand will therefore be highly inelastic and would probably remain unchanged by the imposition of a congestion tax; the only short-term effects of the tax, apart from the unlikely achievement of significantly improved operating economies, would then be to transfer money from road hauliers (and presumably from transport users) to the government. This would bring no advantages unless it were to result in a more favourable distribution of industry or a more nearly optimum allocation of road investment. These are both important objectives, but they both involve the solution of complex problems which will certainly not be resolved merely by the introduction of a congestion tax.

If there is a case for excluding general haulage goods vehicles from the tax, there would seem to be an even stronger case for excluding buses. One of the main advantages of the congestion tax would be to bring about a greater use of public transport, and raising bus costs and prices would obviously tend to reduce the scale of a car-to-bus transfer. In theory, of course, one can argue that all road users should be treated alike, and that the pricing system would reflect the economies in using road space which can be achieved by buses. But these economies depend crucially on the relative load factors of cars and buses, which would be altered in unpredictable ways by the congestion tax. In order to determine their post-tax price structure bus undertakings would need to predict load factors which would themselves be partly dependent on fares. There would also be interdependencies between the speed and regularity of the bus services and the number of people choosing to transfer from car to bus. It would seem very unwise to add the uncertain reactions of bus undertakings when faced by a very sharp increase in costs to the many other problems which would be created by a congestion tax. Professor Beesley has estimated that the application of a congestion tax to buses might result in a charge amounting to 60d. per vehicle mile. This is rather more than the average total operating costs of bus undertakings (including all overheads except interest payments), which in 1964/65 averaged 42.4d. per vehicle mile for municipal undertakings (and 49.8d. for London Transport). Only on the most optimistic assumptions of improvements in load factor, reductions in lost time and improvements in the relation between peak and off-peak running costs could such a cost increase be met without a considerable increase in fares. It is, incidentally, an interesting illustration of the difficulties which can arise from treating congestion costs as identical with other costs that Professor Beesley continues his analysis of the effect of a congestion tax on bus operation with the statement that, on certain assumptions, "the cost savings... would give savings of the order of 53d. a mile". But these cost savings do not represent a drop in the paid-out costs of bus undertakings (often equivalent to their total disappearance). They are the passengers' time savings, which, if all the calculations of demand elasticities and time evaluation work out right, may be translated into potential increased revenue for the bus undertakings, but which the undertakings
would doubtless regard as a rather more uncertain and distant benefit than a fall in operating costs.

My arguments about the "buy back" effect were not intended to be a reason for rejecting a congestion tax, but were an attempt to show that the Smeed Report suggestion that the total burden of tax on motorists should be kept constant would make it more difficult to calculate "optimum" tax levels. Neither Professor Vickrey nor I can be certain how motorists would behave if they were met by an increase in the price of road space and a simultaneous rise in income; but my hypothesis that the vital "journey to work" travellers would be likely to spend most or all of their increased income on road space still seems to me to be likely to hold good. The same point can be put in another way by arguing that recent research supports the contention that income is the major determinant of the value put on travelling time, and that a rise in income will, therefore, increase the average value of leisure time and with it the demand for relatively high-speed private motoring. Since not all the increase in income would accrue to urban road users, successive adjustments to the tax would take the form of a convergent series, and a new equilibrium could be reached; but the process would be more complex than if there were no directly related income changes. J. C. Tanner has shown that the present taxes on rural roads may (if we regard them as a price) be too high and that some reduction might therefore be desirable. But the Smeed Report constraint of keeping the total tax burden constant could lead to an actual subsidy to rural road users, which it would be hard to justify. Unless the present tax level is considered to be optimum, the only reason for requiring the total tax yield to be kept constant must be a realisation of the political difficulties, which may well mean that the congestion tax discussion will remain academic to the end. There is no income weighting of the votes which interest politicians, and the political party which introduces a congestion tax may well lose more from the tolled-off than it will gain from the tolled-on. And, as I pointed out in the Economic Journal article, even the "tolled-on" may consider themselves worse off unless they feel a direct benefit from the way in which the government uses the revenue from the tax.

Professor Beesley has argued, no doubt correctly, that a congestion tax would provide valuable information for cost-benefit studies of alternative road investments; but the introduction of a tax can hardly be justified on this ground alone. The much more sweeping claims of G. Roth, in which a full "road pricing" system would be linked with investment determined through the price system and profit maximization as in a competitive private sector of industry, cannot be accepted. There would be insuperable difficulties in reproducing competitive conditions and in allowing for the social costs which a congestion tax would not reflect. Urban roads might be extended to the point where the city was finally destroyed as a place where people

---


CONGESTION CHARGES AND WELFARE

C. H. Sharp

could live and be happy, and, as Professor Beesley points out in a full analysis of these proposals, some rural roads might be allowed to deteriorate or be closed.

Professor Vickrey is right in criticising the wording of my point that if all travellers valued time equally "it is no longer true that it is better for fewer people to be travelling faster than for more people to be travelling more slowly". I should have said that it is not necessarily true, and that in these circumstances the congestion tax solution would not work.

Some of the differences between my view of the congestion tax proposals and those of both Professor Beesley and Professor Vickrey centre round the long-term results of a tax and its effects on town planning, the distribution of industry and the location of residential areas. All I would ask is that we should be unbiased in assessing these long-term effects, and not assume that net benefits must result merely out of a desire to justify a congestion tax: it is evident that this would be only one factor amongst many others influencing these complex issues. Some of my differences with Professor Vickrey over the immediate results of a tax arise because he appears to be thinking mostly of metropolitan-type road networks where alternative routes and alternative modes of transport are available. I have been arguing about the applicability of the tax to some of the highly congested roads into cities like Birmingham, where for very many people the only short-term alternative to car travel would be very much slower bus journeys, and where the time losses of the tolled-off may be greater than the gains of the tolled-on. There are almost certainly some urban roads in Britain for which the sensible short-term solution is to operate them to maximum capacity (measured by traffic flow per hour), which could mean accepting much slower speeds than in the congestion tax solution.

I can accept the conclusion of Professor Vickrey's chapter "Pricing as a Tool in Coordination of Local Transportation" that a congestion tax can be useful when it can provide "a direct incentive to the individual driver to economise in the use of high cost facilities during periods of peak demand and potential congestion". But I would like to add that there are some roads in Britain where there may be no scope for such economies, and where it would be difficult to justify forcing some people to undertake very much slower work journeys so that others might travel faster. Where a congestion tax does not produce some immediately recognisable economies in the allocation of road space, the appeal to its very uncertain long-term effects cannot carry much weight. A congestion tax must be judged by its results, which may differ from road to road. I cannot accept the argument that we are necessarily moving towards any kind of allocative optimum if we charge road users the costs which they impose on each other while other similar costs are ignored, while the tax structure on which the congestion tax would be imposed is determined by the general revenue requirements of the government rather than by the average or marginal cost of road provision, and while the distribution of income is not optimal.

University of Leicester

19M. Beesley: European Conference paper, op. cit., page 32.