PRICING AND CONGESTION: ECONOMIC PRINCIPLES RELEVANT TO PRICING ROADS

DAVID M NEWBERY
Department of Applied Economics, University of Cambridge

I. INTRODUCTION

The road network is a costly and increasingly scarce resource. For the UK the Department of Transport (1989a) calculates that total road expenditures (capital and current) or 'road costs' averaged £4.34 billion per year at 1989/90 prices for the period 1987/8-1989/90. Public expenditure on roads has been fairly stable recently, increasing by about 6 per cent in real terms between 1982/3 and 1988/9, but with no strong trend. (Department of Transport, 1989b, Tables 1.18, 1.22, 1.23.) From the 24.6 million vehicles registered, road taxes of £12.7 billion were collected (including the £1.4 billion car tax), or 2.9 times the Department's figures for 'road costs'. In 1987 15.1% of consumers' expenditure was on transport and vehicles, and 11.3% was on motor vehicles alone. Clearly, road transport is of major economic significance. Car ownership per 1000 population in the UK appears to be catching up on the rates in the larger European countries and is now about 83% of French and Italian levels, 73% of West German levels. Over the decade 1979-1989 the number of private cars increased from 14.3 to 18.5 million, or by 29%. From 1978-1988, the number of total vehicle-km driven rose from 256 to 363 billion or by 42%. As the length of the road network increased rather less, the average daily traffic on each km of road rose by 34% over the same decade on all roads and by 52% on motorways. Traffic on major roads in built up areas (ie those with a speed limit of 40 mph or less) increased by 13%. (Department of Transport, 1989b, Tables 2.1, 2.3.)

As road space is a valuable and scarce resource, it is natural that economists should argue that it should be rationed by price - road users should pay the marginal social cost of using the road network if they are to be induced to make the right decisions about whether (and by which means) to take a particular journey, and, more generally, to ensure that they make the correct allocative decisions between transport and other activities. If road users paid the true social cost of transport, perhaps urban geography, commuting patterns, and even the sizes of towns would be radically different from the present. The modest aim here is to identify these social costs, provide rough estimates of their magnitude for Britain, and hence identify the major policy issues.

One way to focus the discussion is to ask how to design a system of charges for road use. The problem of designing road charges can be broken down into various subproblems. First, what is the marginal social cost (that is, the extra cost to society) of allowing a particular vehicle to make a particular trip? Part will be the direct cost of using the vehicle (fuel, wear and tear, driver's time, and so forth) and will be paid for by the owner. This is the private cost of road use. Other costs are social: some will be borne by other road users (delays, for example); some by the highway authority (extra road maintenance); and some by the society at large (pollution and risk of accidents). These are called the *road use cost* - the social costs (excluding the private costs) arising from vehicles using roads. It seems logical to attempt to charge vehicles for these road use costs, so as to discourage them from making journeys where the benefits are less than the total social costs (private costs plus road use costs). The first task, therefore, is to measure these road use costs.

The second question is whether road users should pay additional taxes above these road use costs. One argument is that road users should pay the whole cost of the highway system, not just the extra cost of road use, either to be 'fair' in an absolute sense or to achieve parity or equity with, say, rail users (in those rare countries where the railway is required to cover its total costs without subsidy). Another argument is that the government needs to raise revenues and some part of this revenue should be collected from road users, since to exempt them would be to give them an unreasonable advantage over the rest of the population. Both arguments appeal either to the desire for equity or fairness, or to the need for efficiency in the allocation of resources (road versus rail), or both.

(i) Relevant Principles of Taxation

The modern theory of public finance provides a powerful organizing principle for taxing and pricing. Under certain assumptions policies should be designed to achieve production efficiency, with all distortionary taxes falling on final consumers. Broadly, the conditions for this result, set out formally in Diamond and Mirrlees (1971), are (a) that production efficiency is feasible and (b) that any resulting private profits are either negligible or can be taxed away. The feasibility condition would be satisfied if the economy were competitive and externalities could be corrected or internalized.

The theory has immediate implications for road charges and taxes. Road users can be divided into two groups: those who transport freight, which is an intermediate service used in production, and those who drive their own cars or transport passengers, who enjoy final consumption. Freight transport, which is roughly and conveniently synonymous with diesel using vehicles, should pay the road use costs to correct externalities and to pay for the marginal costs of maintenance. Additional taxes (comprising the *pure tax element)* on (largely gasoline using) passenger transport can be set, using the same principles that guide the design of other indirect taxes. We shall show below that one would expect a close relationship between road use costs and total road expenditures. There is no logical reason to attribute the taxation of passenger transport to the highway budget, since it is a component of general tax revenue. But if all road taxes and charges are taken together, there are good reasons to expect that they will exceed total highway expenditure. In short, in a well-run country no conflict need arise between the goals of designing an equitable and efficient system of road use charges and taxes and the desire to cover the highway system's costs.

The theory provides a useful framework for the study of road user charges. The first step is to identify the road use costs. The second is to see what methods are available for levying charges and how finely they can be adjusted to match these costs. The third step is to examine how far these methods have repercussions outside the transport sector and, where these occur, how to take them into account. These three steps will suffice for freight transport. For passenger transport, one other step is needed: to determine the appropriate level (and method of levying) the pure tax element.

II QUANTIFYING THE SOCIAL COSTS OF ROAD USE

Vehicles impose four main costs on the rest of society - accident externalities, environmental pollution, road damage, and congestion. Accident externalities arise whenever extra vehicles on the road increase the probability that other road users will be involved in an accident. To the extent that accidents depend on distance driven and other traffic these accident costs can be treated rather like congestion costs. Newbery (1988a) argued that accident externalities could be as large as all other externality costs taken together, and are thus possibly of first order importance. There are two reasons for this high estimate, both disputed. The first is that the figure critically depends on the value of a life saved or the cost of a life lost. If one bases this on apparent willingness to pay to reduce risks, then the cost per life saved might be between £650,000 and £2 million at 1989 prices, based on the survey results of Jones-Lee, reported in this issue. The lower figure is over double that originally used by the Department of Transport, who based their earlier estimates on the expected loss of future earnings of a representative victim. Apparently the Department of Transport has been persuaded of the logic behind the willingness-to-pay approach, and now uses a figure of £500,000 (Jones-Lee, 1990).

The second reason is that in the absence of convincing evidence, the estimate assumed that the number of accidents increased with the traffic flow as the 1.25 power of that flow. (That is, if the traffic is twice as heavy, the risk of an accident happening to each car is increased by 19 per cent. Compare this with the number of pairwise encounters between vehicles, which rises as the square of the flow.) This in turn means that one-quarter of the cost of mutually caused accidents is an uncharged externality, even if each driver pays the full cost of the accident to him. (To the extent that society pays through the NHS, these individual costs are borne by society and attributable as part of 'road costs'. The Department of Transport includes direct costs. It might argue that their earlier valuation of life was based on the loss of earnings which might have to be made good through the social security system to survivors.) Note that it is important to relate the accident rate to traffic levels in order to identify the size of the externality. Indeed, one might argue from the fact that the accident rate has fallen as traffic has increased that this 1.25 power law is invalid, and that at best there is no relationship between traffic and the accident rate. If so, then there would be no externality between motor vehicles (other than that already counted in the cost falling on the NHS). This would be the case if one took seriously the explanation of 'risk compensation', according to which road users choose a desired level of perceived risk with which they are comfortable - too little risk is boring, too much is frightening. Improvements in road safety then induce compensating increases in risk taking, while deteriorating road conditions (ice, snow, heavier traffic) induce more caution. Of course, one should be wary of using time series information about accident rates as road improvements are continuously undertaken to improve road safety. The relationship between accident rate and traffic should be derived from a properly estimated cross-section analysis.

Jones-Lee, in his article in this issue, does indeed assume that the accident *rate* (ie the risk of an accident per km driven) is independent of the traffic flow, from which it follows that there is no externality between motor vehicles (except those caused by the system of social and health insurance). He also assumes that the probability of any vehicle having an accident involving a pedestrian or cyclist is constant per km driven, in which case it follows that the accident rate experienced by cyclists and pedestrians is proportional to the number of vehicle km driven. If this is the case, then motor vehicles do impose an externality on non-motorised road users (though not on other motorists), which Jones-Lee calculates to be quite large--perhaps 10-20 per cent of total road costs. Of course, we remain relatively uncertain about the relationship between accidents to other road users and traffic. It has been remarked that not so long ago children were allowed to play in the street, and cycle or walk unaccompanied to school. The number of accidents to such children was quite high. Now it is so obviously insane to allow such activities that the number of accidents may have fallen with the increase in traffic. Of course, the accident externality is still there, though hidden in the form of the extra costs of ferrying children to school, and not allowing them to play or cycle unsupervised.

The main problem therefore lies in identifying the relationship between traffic and accidents - in the words of the US Federal Highway Cost Allocation Study 'Quantitative estimation of accident cost and vehicle volume relationships, however, has not yet proved satisfactory..' (U.S. Federal Highway Authority, 1982). Given the huge costs involved and the potential gains from lowering accident rates, identifying such relationships should have overwhelming research priority.

Similarly, pollution costs share many of the same features as congestion costs (and tend to occur in the same places). Where they have been quantified (for the US) the appear to contribute less than 10 percent of total road costs. They are normally dealt with by mandating emission standards, and by differential taxes on the more polluting fuels (for example, by having higher taxes on leaded petrol). A new European Directive on vehicle emissions, known as the Luxembourg Agreement, will be implemented in the UK. This mandates NO_x levels of half the current limit, and reductions in hydrocarbon releases of three-quarters, at an estimated cost to the motorist of £800 million or about 4 per cent of motoring costs (Department of the Environment, 1989). Provided the pollution costs are reflected in fuel taxes, and the requirement to meet emissions standards, these costs will have been satisfactorily internalised. One should, however, be rather cautious about mandating stringent emissions standards without a careful cost-benefit analysis. Crandall *et al* (1986, p114-5) estimate that the programme costs for the US of the more stringent 1984 emissions standards might be about $20 billion per year with a replacement rate of 10.5 million cars, which is several times rather optimistic estimates of the potential benefits of reducing pollution. (Safety regulations in contrast, though expensive, seem to have been justified on cost-benefit criteria.) Some of these issues are discussed further in Newbery (1990).

(i) Road damage costs

These are the costs borne by the highway authority of repairing roads damaged by the passage of vehicles, and the extra vehicle operating costs caused by this road damage. The damage a vehicle does to the road pavement increases as the fourth power of the axle load, which means that almost all damage is done by heavy vehicles such as trucks. Increasing the number of axles is a potent method of reducing the damaging effect of a vehicle - doubling the number of equally loaded axles reduces the damage to one-eighth its previous level. Consequently most highway authorities closely regulate the axle configuration and maximum legal axle loads. Increasing the road thickness dramatically increases the number of vehicles that can be carried before major repairs are required - doubling the thickness increases this number by 2 to the power 6.5 or so (Paterson, 1987). Consequently the most damaging and therefore costly combination is a heavy vehicle on a thin road.

The theory which allows these costs to be quantified is set out in Newbery (1988a,b). The road damage costs of a vehicle will be proportional to its damaging power (and will be measured in terms of Equivalent Standard Axles, or ESAs). Britain, in common with most advanced countries, follows a condition-responsive maintenance strategy in which the road is repaired when its condition reaches a pre-determined state. In such cases, the road damage costs will be equal to the average annual costs of maintaining the road network in a stable state, multiplied by the fraction of the road deterioration caused by vehicles as opposed to weather, allocated in proportion to ESA-miles driven. The fraction of total costs allocated to vehicles will depend on the climate, the strength of the road, and the interval between major repairs, and the formula is given in (Newbery 1988a, b). In hot dry climates it will be between 60 and 80 percent, while in freezing temperate climates the proportion will be between 20 and 60 percent, the lower figures corresponding to more stringent maintenance criteria or lower traffic volumes. For Britain, Newbery (1988a) argued that the appropriate fraction was 40 percent. If maintenance is condition-responsive then it is not necessary to charge vehicles for the damage they indirectly do to subsequent vehicles which experience increased operating costs on the damaged pavement - on average the condition of the pavement will remain unchanged.

It is simple to update the road damage costs given in Newbery (1988a) using the latest estimates of road track costs provided in Department of Transport (1989a, Table 5). The total cost identified is £785 million, or 8.7 p/ESAkm. The allocable fraction is 0.4, giving £314 million or 3.5 p/ESAkm. As such, road damage costs are a small fraction of total road costs. To provide a quick estimate of how large a fraction, Table 1 below updates the results for 1986 to 1990 from Newbery (1988a, Table 2). There the value of the road network was estimated at £50 billion excluding land. Updated to 1990 prices this is £62 b, to which annual capital expenditure of £2 b brings it to £70 b. The cost of land is estimated at 14 per cent of this, to give a total of £80 billion. In Newbery (1988a) the rate of interest on this capital value was taken to be the then Test Discount Rate of 5% real, and if this figure is again used, then interest on the value of the road network would be £4,000 million, compared to the actual capital expenditure of only £1,921 million. Recently the Test Discount Rate has been revised upward to 8% real (presumably reflecting the perceived higher real rate of return in the rest of the economy), and at this rate the interest costs would rise to £6,400 million.

There is little logic in combining current and capital expenditures as the Department of Transport does in estimating ‘road costs’, and Table 1 only includes imputed interest at the two different rates of 5% and 8%. It will be seen that allocable road damage costs amount to 3.5 - 5% of total road costs, and are thus essentially negligible (which is not to deny that it is important to charge them appropriately to heavy goods vehicles.

This estimate is quite close to that for 1986 of 3.5% given in Newbery (1988a). Even if repair costs currently allocated by the UK Department of Transport in proportion to Gross Vehicle Weight are included (and the theoretical justification for so doing is rather unclear) the figure only rises to 9 - 13 per cent (depending on the choice of the TDR). Small et al (1988) estimate that pavement costs, including construction and periodic resurfacing, are less than 16 per cent of road costs in their simulations of an optimised US road system, and road damage charges would only account for 2 per cent of total charges. Far and away the largest element (again, ignoring accident costs, which might also be very large) are the congestion costs.

(ii) Congestion costs

These arise because additional vehicles reduce the speed of other vehicles, and hence increase their journey time. The standard way of calculating the short-run marginal congestion cost (MCC) of an extra vehicle in the traffic stream starts by postulating a relationship between speed (v kph) and flow (q vehicles or PCU/h) where PCU are passenger car units, a measure of the congestive effect of different vehicles in different circumstances (eg higher for heavy lorries on steep hills than on the level). If the travel cost per km of a representative vehicle is

$$c = a + b/v, \qquad \textbf{(1)}$$

where b is the cost per vehicle hour, including the opportunity cost of the driver and occupants, then the total cost of a flow of q vehicles per hour is $C = cq$. If an additional vehicle is added to the flow, the total social cost is increased by

$$\frac{dC}{dq} = c + q.\frac{dc}{dq} \qquad \textbf{(2)}$$

The first term is the private cost borne by the vehicle and the second is the marginal externality cost borne by other road users.

The next step is to establish the speed-flow relationship, $v = v(q)$ and here one must be careful to pose the right question. Engineers designing particular sections of the road network are concerned with flow at each point, and most of the relationships estimated are of this form. They show that traffic flow is heavily influenced by junctions, where additional traffic enters or disrupts the smooth flow, and it is possible for the speed flow relationship to be backward bending, as in Figure 1.

The curve is to be interpreted as follows. As traffic increases above q the speed is given by points such as A, B. As traffic nears the capacity of the link, k, at the point C, the flow changes to a condition of stop-start, and traffic flow through the bottleneck drops, to a point such as D, associated with a lower speed. This is an unstable situation, and as flow falls, so the traffic leaving the bottleneck will accelerate, and eventually clear the blockage. At that time, the speed will jump back up to point A. (Further details are given in Newbery (1987) and Hall et al (1986).)

Table 1
Road costs at 1989/90 prices

£ million

Cost category	*Annual average* 5% TDR	8% TDR
Interest on capital	4000	6400
(Capital expenditure)	(1921)	
Maintenance *less* costs attrib to pedestrians	2093	2093
Policing and traffic wardens	327	327
Total road costs	6420	8820
of which attributable to		
road damage costs	314	314
Gross vehicle mass	495	495
VKT	225	225
Balance attributable to PCU	4884	7284
PCU km (billion)	375 billion km	
Cost per PCUkm pence/km	(*1.30p/km*)	(*1.94p/km*)

Source: Department of Transport (1989a).
Notes: Figures are annual averages for the years 1987/88 to 1989/90
TDR: Test Discount Rate; VKT: vehicle km travelled. Costs attributable to Gross vehicle mass and VKT taken from Department of Transport (1989a), adjusted in the same way as those given in Newbery (1988a, Table 5).

Useful though this relation is for road design, it is not what is wanted for estimating the cost of congestion, where we need a measure of the total extra time taken by the remaining traffic to complete their planned journeys, not their speed at a particular point on the road network. The Department of Transport, when planning roads to alleviate congestion, uses formulas estimated by the Transport and Road Research Laboratory, and reported in Department of Transport (1987). These are based on 'floating car' methods in which the observing vehicle remains in the traffic stream for a period of time, and hence gives a better estimate of the average relationship between speed and low. The find a reasonably stable linear relationship of the form

$$v = v_0 - \beta q \qquad \textbf{(3)}$$

where q is measured in PCU/lane/hr. The estimated value of β for urban traffic is 0.035. This agrees closely with a careful study of traffic flows within zones of Hong Kong, reported in Harrison et al (1986), itself commissioned as part of Hong Kong's road pricing experiment.

Figure 1
Speed flow relationship for a link

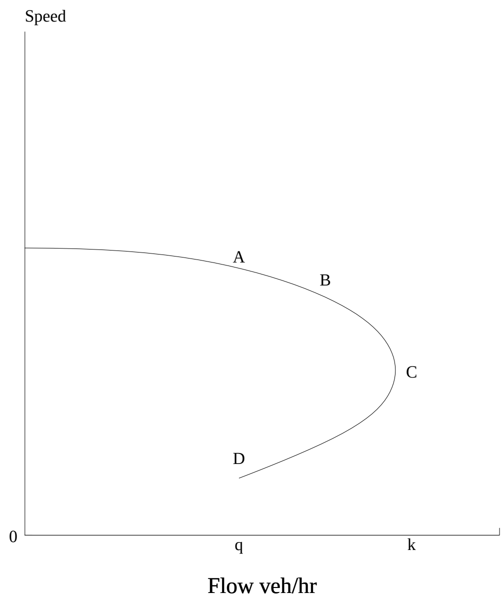
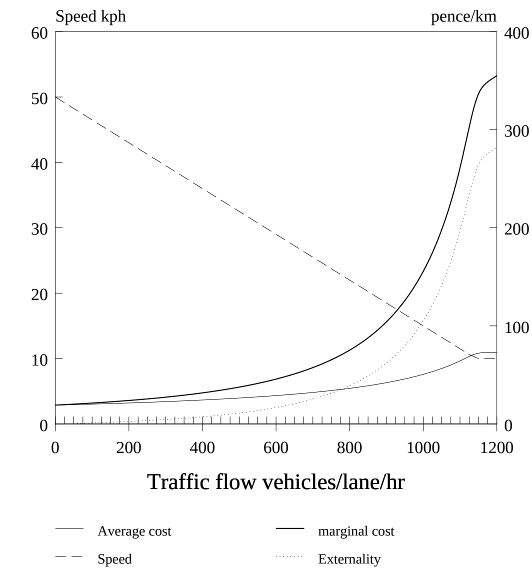


Figure 2
Average and Marginal Cost of Trips



Urban non-central roads, 1990 prices

This linear relationship can be used to quantify the average and marginal costs of traffic and hence to determine the MCC. Figure 2 below gives this relationship for suburban roads at 1990 prices, based on the estimated COBA 9 formula. The left hand scale gives the speed associated with the traffic flow, and on such roads average speeds rarely fall below 25 kph, so the relevant part of the diagram is to the left of that speed level. Another, quite useful way of representing this limitation is to suppose that the demand for using suburban roads becomes highly elastic at this level, an idea pursued further below.

In some ways a better measure of the congestion relationship is given by the marginal time cost (MTC) in vehicle-hours per veh.km, which can then be multiplied by the current value of the time use of the vehicle, b in (1). From (3), the MTC is just $\beta q/v^2$. Given data for q and v, the MTC can be estimated. Newbery (1988a, Table 1) estimated these costs for Britain for the year 1985. Rather than repeat the rather time-consuming calculations reported there, the following short-cut has been adopted. If m is the MCC as a function of q, and if Δq is the increase in traffic over some period, then the revised estimate of the MCC is $m + dm/dq.\Delta q$. The factor by which to scale up the original estimates of MCC can be found from the above equations and is

$$\left(1+\frac{2\beta q}{v}\right)\cdot\frac{\Delta q}{q} \qquad \textbf{(4)}$$

The results of this updating procedure is given in Table 2. If anything, the estimate will be on the low side, as the relationship is very non-linear. If some roads have an above average increase in traffic while others have a below average increase, then taking the average increase $\Delta q/q$ will underestimate the average of the costs on each road.

Table 2
Marginal Time Costs of Congestion in Great Britain, 1990

	MTC (veh h/ 100 PCUkm)	*VKT fraction*	*MCC p/PCUkm*	*Index of MCC*
Motorway	0.05	0.17	0.26	8
Urban central peak	5.41	0.01	36.37	1070
Urban central off-peak	4.35	0.03	29.23	860
Non-central peak	2.36	0.03	15.86	466
Non-central off-peak	1.30	0.10	8.74	257
Small town peak	1.03	0.03	6.89	203
Small town off-peak	0.63	0.07	4.20	124
Other urban	0.01	0.14	0.08	2
Rural dual carriageway	0.01	0.12	0.07	2
Other trunk and principal	0.04	0.18	0.19	6
Other rural	0.01	0.11	0.05	1
Weighted average			3.40	100

Source: Updated from Newbery (1988a, Table 1)

Table 2 shows in a vivid way the great variation in marginal congestion costs by time of day and location. Urban central areas at the peak have an *average* congestion cost of 10 times the average over all roads, and more than 100 times that the average motorway or rural road.

The table shows that the average congestion cost is 3.4 p/PCUkm, and, given the 375 billion PCUkm driven from Table 1, if road users were to be charged for congestion, the revenue collected would have been £12,750 million. If we add to this sum the road damage costs (£314 m), the amounts allocated according to Gross Vehicle Mass (£495 m), and VKT (£225 m), all taken from Table 1, then the appropriate level of road charges should yield £13,784 million. Total road taxes were £12,700 million, or 92 percent the required level. Congestion charges would amount to 92 percent of the total appropriate road charge.

It is interesting to compare these estimates with those given in Newbery, (1988a). The estimated congestion charges for Britain for 1986 were £6,203 million out of the total appropriate charge (excluding accident costs) of £7,033 million, or 88 percent. The figures are high as the amount of time wasted is so high, though the costs are frequently ignored by highway authorities, as they are entirely borne by highway users. The Confederation of British Industries has calculated that traffic congestion costs the nation £15 billion a year, or £10 a week on each household's shopping bill (*The Times*, May 19, 1989). This figure is comparable to the £13 billion for the appropriate congestion charge calculated above, though the question of how best to measure the true social cost of congestion is discussed more fully below.

Small et al (1988) cite evidence that suggests congestion costs are also high in the US. Thus average peak-hour delays cross the Hudson River to Manhattan have roughly doubled in the past ten years, while congestion delays in the San Francisco Bay Area grew by more than 50 percent in just two years. In 1987 63 percent of vehicle-miles were driven on Interstate Highways at volume to capacity ratios exceeding 0.8, and 42 percent on other arterials. Although their study does not quantify the congestion costs they suggest figures of some tens of billions of dollars annually - a figure which squares with the evidence in the next paragraph.

The correct charge to levy on vehicles is equal to the congestion costs they cause (in addition to other social costs like damage costs). If roads experience constant returns to scale, in that doubling the capital expenditure on the road creates a road able to carry twice the number of vehicles at the same speed, and if roads are optimally designed and adjusted to the traffic, then it can be shown that the optimal congestion charge would recover all of the non-damage road costs. (Newbery, 1989). These include interest on the original capital stock, as well as the weather-induced road damage costs not directly attributable to vehicles, and other maintenance expenditures, collectively identified as 'road costs attributable to PCU' in Table 1. The available evidence supports constant returns to scale or possibly rather mildly increasing returns, of the order of 1.03-1.19. (This always surprises highway engineers, who *know* that it does not cost twice as much to double the capacity of a highway, as many costs - embankments, etc - are fixed, and the capacity of a two-lane divided highway is considerably greater than of a one-lane divided highway. But most capacity increases are needed in congested urban areas where the costs of land acquisition can be extremely high. The econometric estimates pick this factor up. See Keeler and Small, 1977; Kraus, 1981). If we take the estimates as supporting constant returns, then the result is directly applicable, and provides a useful benchmark against which to judge the estimated congestion charges. If we assume increasing returns to scale, then road charges would not cover road costs of an optimal road network.

In 1986 the estimated average congestion charge was 1.42 times as high as the road costs attributable to PCU, suggesting either that the road network was inadequate for the level of traffic, or that the correct rate of interest to charge was higher than 5 per cent real. (At 8 per cent the ratio was only 1.0.) The corresponding ratio for 1990 is 2.82 (or 1.75 at 8 per cent interest). In short, if roads were undersupplied in 1986, they are becoming critically scarce as traffic volumes increase faster than road space is supplied. Notice that assuming that there are increasing returns to capacity expansion strengthens the conclusion that roads are undersupplied.

III. CHARGING FOR ROAD USE

Ideally, vehicles should be charged for the road use cost of each trip, so that only cost-justified trips are undertaken. In practice it is not too difficult to charge for road damage, which is largely a function of the type of vehicle and the extent to which it is loaded. Ton-mile taxes as charged by some of the States of the US can approximate the damage charge quite closely, provided they are made specific to the type of vehicle. Vehicle specific distance taxes would be almost as good, provided axle loading restrictions were enforced. Fuel taxes are moderately good in that they charge trucks in proportion to ton-miles. As they charge different types of vehicles at (almost) the same ton-mile rate, they must be supplemented by vehicle specific purchase taxes or license fees and combined with careful regulation of allowable axle configurations and loadings (Newbery et al 1988; Newbery 1988c).

The more difficult task is to charge for congestion, which varies enormously depending on the level of traffic, which in turn varies across roads and with the time of day, as Table 2 shows dramatically. The most direct way is to charge an amount specific to the road and time of day, using an ‘electronic numberplate’ which signals to the recording computer the presence of the vehicle. The computer then acts like a telephone exchange billing system, recording the user, the time and place and hence the appropriate charge, and issuing monthly bills. The cost per numberplate is of the order of $100, or perhaps one quarter of the cost of catalytic converters which are now mandatory for pollution control in many countries. Such systems have been successfully tested in Hong Kong (Dawson and Catling, 1986) but there was initially some pessimism at the political likelihood that they would ever be introduced (Borins 1988). The stated objection was that the electronic detectors could monitor the location of vehicles and hence would violate the right to privacy. This objection may be valid in a society suspicious of (and not represented by) its government, though evidence suggests that this is not likely to be much of a problem in Europe (ECMT, 1989). The objection could be levied against telephone subscribers once itemised bills are introduced, and the objection can be overcome in much the same way by the use of ‘smart cards’, rather like magnetic telephone cards. The electronic licence plate would be loaded with the smart card and would debit payments until exhausted. Only thereafter would the central computer monitor and bill for road use.

A more plausible explanation for the lack of success in Hong Kong may have been that it was not clear to car owners that the new charges (which were quite high, of the order of $2-3 per day) would replace the existing and very high annual licence fee. Faced with a doubling of the cost of road use commuters understandably objected. But the whole point of charging for road use by electronic licence plates is to replace less well designed road charging schemes, such as fuel taxes and licence fees. The proposition that needs to be put to the public is that in exchange for the entire system of current road taxes (fuel taxes in excess of the rate of VAT, the special car purchase tax, and the licence fee), road users will be charged according to their use of congested road space, at a rate which for the average road user will be roughly the same. (For the UK this is still just about feasible on the figures given above, at least if the pure tax element on private motorists is allowed to fall to zero.) As more than half the road-using population drive less than the average number of miles in congested areas, this should command majority support.

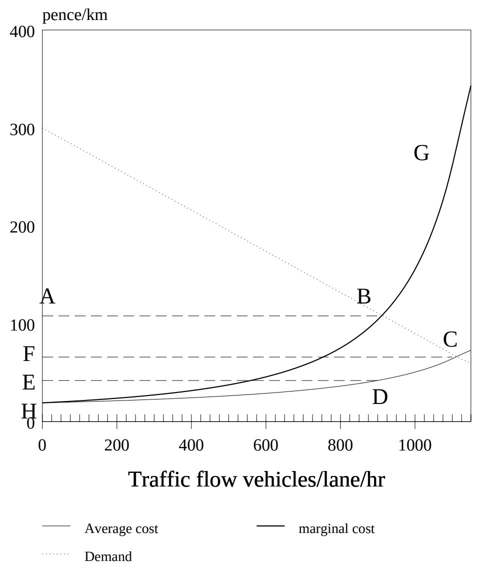
An alternative method of selling the use of electronic licence plates might be to offer rebates for tax on fuel used (at the estimated rate per km, or on production of receipts) and to waive the licence fee for those installing the licence plates. This might be necessary as an interim measure when their use is confined to major urban centres, notably London. It is noticeable that despite the claim by Channon as Minister of Transport that the Government had no plans to introduce road pricing because of the perceived public hostility, that hostility seems to be diminishing, at least in London. A recent opinion survey conducted by the Metropolitan Transport Research Unit showed 87 per cent in favour of some form of traffic restraint, and 53 per cent in favour of a fixed charge to drive into Central London. Charging per mile was widely supported, and 48 per cent said they would use public transport if such charges were introduced (*The Independent*, January 26, 1990).

Until road pricing is introduced, alternative and less satisfactory methods of charging are necessary. One such is selling area licenses which grant access to congested zones such as city centres during rush hours. This solution has been used in Singapore for over a decade, with considerable success. Heavy parking charges and restricted access may also be effective to varying degrees (World Bank, 1986). At the moment, however, the only way to charge vehicles for the congestion they cause is in proportion to the distance they drive, by fuel taxes and/or vehicle purchase taxes (which approximate reasonably well to distance charges for heavily used vehicles, but less well for automobiles). Such taxes, combined with access charges (area licenses, or even annual licenses) achieve the desired effect of charging road users on average for congestion, but do little to encourage them to drive on less congested roads or at less busy times of the day (or, indeed, to take public transport instead). In Britain, the estimates above suggest that road taxes might usefully be increased somewhat. It is worth remarking that there are clear advantages in raising such corrective taxes to their efficient level, as they allow other distortionary taxes, which incur deadweight losses, to be reduced.

IV MEASURING THE COSTS OF CONGESTION

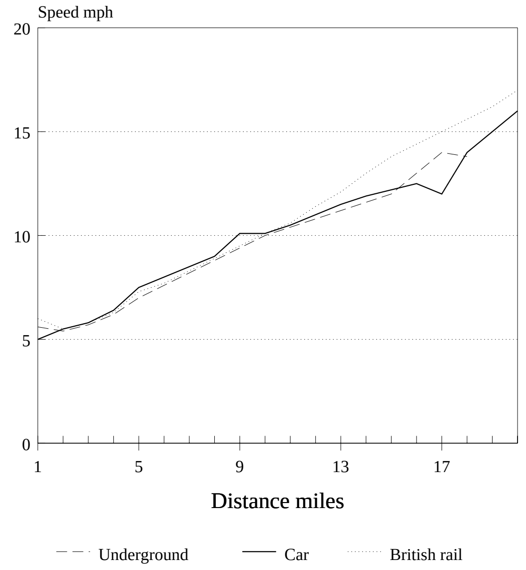
So far we have avoided discussing the actual cost of congestion in Britain, and instead calculated the revenue that would be generated if vehicles were to be charged for the congestion they caused. Figure 3 shows how the equilibrium demand for trips is established in the absence of such charges. If road users pay only the private costs of the trip, their costs are given by the average cost schedule, which meets the demand schedule at point C. At that point the willingness to pay by the marginal road user is equal to the cost of the trip. The efficient congestion charge would be an amount BD, which, if levied, would cause demand to fall to the level associated with point B.

Figure 3
Cost of Congestion



Urban non-central roads, 1990 prices

Figure 4
Equilibrium trip speed

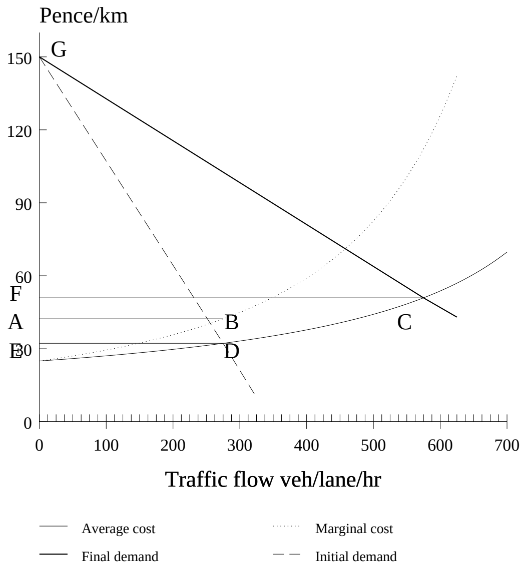


Average door-to-door speed, Cent. London

The revenue then raised would be ABDE, and this is the amount referred to above as the revenue attributable to the congestion charge. In the figure it would amount to £600 per lane hour. But is this the correct measure of the congestion cost? Consider various alternative measures. One measure frequently cited in newspaper accounts of the cost of congestion is the extra costs involved in travelling on congested rather than uncongested roads. This might be measure by FC times FH, the excess of the average actual cost over the cost on a road with zero traffic. On the figure this would amount to £530 (per lane hour). But this is an unrealistic comparison, as it would be uneconomic to build roads to be totally uncongested. Instead one might compare the extra costs of the excessive congestion at point C above that which would be efficient, at point D. These extra costs might be measured as CF times EC, or £270. This is not satisfactory either, as fewer trips would be taken at the efficient level of charges. A better alternative is the loss in social surplus associated with excessive road use. The efficient total social surplus associated with point B is the consumer surplus triangle KBA, plus the tax revenue ABDE. The social surplus associated with point C is just the triangle KCF, and the difference is the rectangle FJDE *less* the triangle BCJ, or £198. This is also equal to the area of the standard deadweight loss triangle BCG. In this case the deadweight loss is equal to one third of the revenue measure.

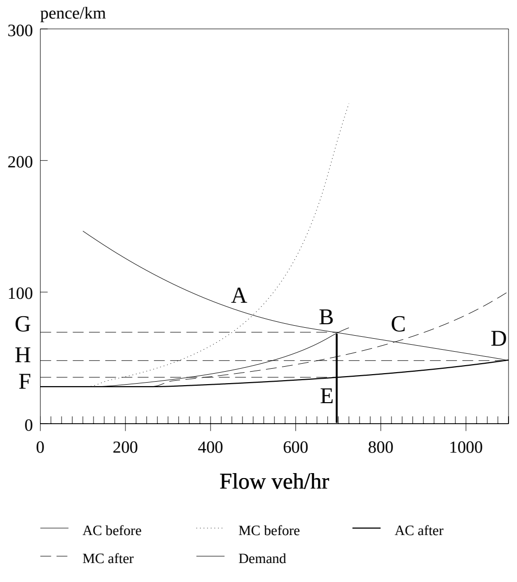
There is one important case in which the revenue measure accurately measures the deadweight loss, and that is where the demand for trips becomes perfectly elastic beyond its intersection with the marginal social cost schedule at point B. In this case there is no gain in consumer surplus in moving from the efficient level of traffic to the equilibrium level, and hence the loss is just equal to the foregone tax revenue. Put another way, by not charging the efficient toll in this perfectly elastic case, the government forgoes revenue but the consumer makes no gain. In crowded urban areas this is a plausible situation. The equilibrium level of traffic is found where the marginal road user is indifferent between using a car or some alternative - public transport, or walking. Thus traffic speeds in London are about the same as they were in the nineteenth century, and the time taken to get to work (or, more properly, the total perceived cost) for many commuters is no better by road than alternatives. Figure 4, which is taken from a graph in *The Times*, Dec 5, 1988, itself based on work by Martin Mogridge, illustrates this graphically for central London. The average door-to-door time taken between points in central London is rremarkably similar for all three modes of transport, so car speeds in equilibrium are determined by the speeds of alternative public transport.

Figure 5
Congestion with public transport



Urban central areas, 1990 prices

Figure 6
Cost-benefit of road improvements



In such circumstances, the social costs of congestion may even be understated by the revenue measure. Consider the situation portrayed in Figure 5. Initially demand is HC and equilibrium is established at C with commuting costs of 0F. But if road users were charged an amount BD the some road users would switch to public transport, reducing demand for private road use. This increased demand for public transport would, after appropriate investment and expansion, lead to an improved frequency of service, while the reduced traffic would lead to a faster public transport service, lowering the costs of travel. Given this lower public transport cost, the demand for private transport would fall from GC to GD, and the willingness to pay for private commuting would fall, from the level 0F to 0A for the marginal commuter. In this case charging the congestion tax would yield tax revenue (a social gain) while reducing the cost of commuting (an additional consumer gain). The tax foregone thus understates the cost of the congestion.

V. COST-BENEFIT ANALYSIS OF ROAD IMPROVEMENTS

The average road tax paid by vehicles per km is now somewhat below the average efficient level. On roads of below average congestion, vehicles may be overcharged, but in urban areas they are certainly undercharged, in many cases by a large margin. Faced with growing congestion, one natural response is to increase road capacity. Figure 6 illustrates the pitfalls in simple-minded cost-benefit analysis. If the road capacity is doubled, then the average and marginal cost schedules will move from ‘before’ to ‘after’. The initial equilibrium will be at B, where demand is equal to the initial average cost (AC), and traffic will be BG. If the AC is lowered to E, then the apparent cost-saving is GBEF, to be compared with the cost of the improvement. But the lower AC will induce increased traffic and the new equilibrium will be at point D, not E. The benefit will be GBDH, which may be significantly lower.

Figure 7 shows that where traffic increases come from previous users of public transport, as in Figure 4, the effect of the road ‘improvement’ may be to increase traffic volumes but to raise average costs, making everyone worse off, and adding negative value. It is hard to escape the conclusion that as journey speeds in London are now below their level at the turn of the century, before the introduction of the car, much of the road investment has been self-defeating.

Figure 7
Perverse road improvement

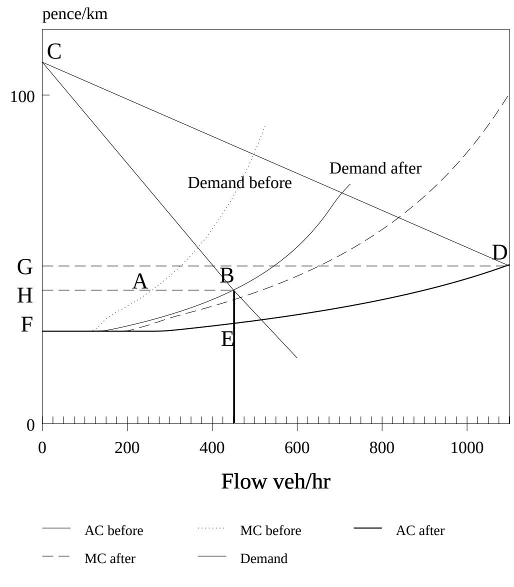
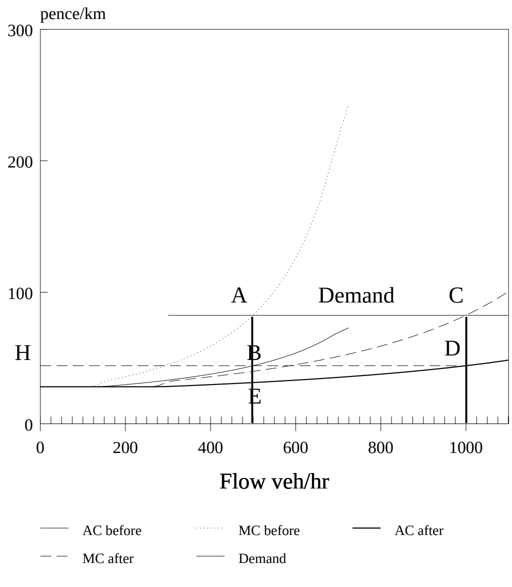


Figure 8
Road improvements with charges



The situation is radically changed when road users pay the efficient congestion charge. Consider Figure 8, in which the demand for trips is perfectly elastic along AC, and congestion charges at the rate AB (marginal cost *less* average cost) are levied by means of an electronic numberplate. If road capacity is expanded, but the charge maintained, then the social gain is the increase in the revenue collected, ACDE, which can be compared with the cost. If there are constant returns to capacity expansion and the road improvement is self-financing, it is justified. This is another application of the proposition that with constant returns to road expansion, congestion charges will recover the costs of the road investment. If there are increasing returns to expansion, then expansion may still be justified even if it is not self-financing - one should compare the marginal expansion costs with the marginal benefit measured by the increased revenue.

The implications of this are clear - without road pricing, road improvements may yield low or even negative returns, but with efficient pricing, not only are improvements easy to evaluate, they should also be self-financing, at least if there are constant or diminishing returns to expansion, as is likely in congested areas.

VI PRICING PUBLIC TRANSPORT

If private road users are significantly undercharged for using the road in urban areas, then commuters face the wrong relative prices when choosing between public and private transport. If it is hard to raise the price of private driving, why not lower the price of public transport to improve the relative price ratios? There are a number of problems with this proposal. First, it is hard to compute the second-best public transport subsidies given the great variation of congestion costs by time of day and location. Second, the subsidies have to be financed by taxes which have distortionary costs elsewhere. (Note that congestion charges would reduce the distortionary costs of the tax system.) Third, subsidies to public transport appear to be rather cost-ineffective. Few motorists are attracted off the road into private transport, with much of the increase in public transport use coming from those previously not using either mode.

There are also political economic problems with operating public bus companies at a loss - there is a temptation to ration them and lower the quality, thus defeating the purpose of making them more attractive to commuters. It becomes more difficult to gain the benefits of deregulation and privatisation. It may just lead to rent-dissipation by the supplier of public transport - as seems to have happened to some extent when the London Underground was heavily subsidised. The same applies to subsidising rail travel - it becomes unattractive to expand capacity if this just increases the size of the loss.

A more promising approach is to make private transport relatively less attractive and public transport more attractive, by improving the quality of the latter, possibly at the expense of the former. Bus-lanes which reduce road space to other road users have this effect, as would electronic signalling to give priority to public transport at traffic lights. Banning private cars from congested streets during the working day has a similar effect. Arguably the greatest obstacle to overcome is that of making private road users aware of the true social costs of their road use. Table 2 reveals that *average* congestion charges of 36 pence/km would be appropriate in urban centres at the peak, and 30 pence/km during the off-peak, with higher charges appropriate on roads with higher volume and/or lower speeds. Figure 2, which is plotted for 1990 costs for suburban areas, shows that when traffic speeds have fallen to 20 kph, then congestion charges of 50 pence/km are in order. It is hard to make public transport 50 pence/km cheaper than its unsubsidized level, and thus more productive to think of ways of raising the cost of private transport at congested periods, or directly reducing its level.

It was argued above that road pricing gives rise to sensible cost-benefit rules for road improvements. The same is also true for other transport investments, especially in public transport and that most capital intensive from, rail transport. If road users paid the full social cost of road use, then there would be every reason to charge users of public transport the full social marginal cost, including congestion. It is unlikely that there are economies of scale in peak-period road or rail use in London, or in peak-period bus use in other cities, and this would lead to fares that would cover the full operating costs, including interest on capital. It would therefore again be easier to apply commercial criteria to investment in public transport, and, indeed, there would no longer be a strong case for keeping these services in the public sector.

VII. CONCLUSION

Road pricing is the best method of dealing with congestion, and would have far-reaching implications for the viability and quality of public transport, for the finance of urban infrastructure, and ultimately for the quality of life. The average road charges might not need to increase above current levels, for if roads were correctly priced, demand for the use of the most congested streets would fall, and with it the efficient charge to levy. Current road taxes are heavy and yield revenue greater than the average cost of the current road system. In equilibrium, with efficient road pricing and an adequate road network, one would expect road charges to be roughly equal to road costs, so either road charges would fall below current tax levels, or substantial road investment would be justified, and possibly both.

A shift to road pricing would cause a fall in the cost of driving in non-urban areas, and an increase in the cost of urban driving. In the medium run the quality of urban public transport would improve, and the average cost of urban travel by both public and private transport might fall below current levels. Energy consumption would increase, as there would no longer be any reason to have heavy fuel taxes (other than those needed to reflect the costs of pollution, and this problem is arguably better addressed by mandatory emissions standards). A shift to road pricing matched by offsetting adjustments to other taxes to raise the same total tax revenue is unlikely to be inflationary, both because the average road tax/charge would be almost unchanged, and because any increase in charge leading to higher travel costs could be matched by lower taxes on other goods, leading to a fall in those elements of consumer expenditure.

Similarly, the impact on the distribution of income is likely to be slight and probably favourable. Urban private travel costs would rise, at least in the short run, and urban public transport quality-adjusted costs may fall (depending on how rapidly any subsidies were phased out, and how quickly the increased demand for public transport translated into more supply at higher average quality). Rural transport costs would fall. As urban car owners are richer than average, and users of public transport are poorer, the redistribution should be favourable. The Government would have to decide on the fate of the car tax (which raised £1.4 billion in 1989/90), and which is a fairly progressive form of indirect tax, being levied at an *ad valorem* rate on the purchase price. Its logic as a method of rationing access to road space would disappear, but it could be retained as a pure consumer tax if it were thought to be justifiable on redistributive grounds.

REFERENCES

Borins, Sandford F. (1988), 'Electronic Road Pricing: An Idea whose Time may never come', *Transpn Res., -A* 22A(1), 37-44.

Dawson, J.A.L., and Catling, I. (1986), 'Electronic Road Pricing in Hong Kong', *Transportation Research*, 20A (March), 129-34.

Crandall, R.W., H.K. Gruenspecht, T.E. Keeler and L.B. Lave (1986), *Regulating the Automobile*, Brookings, Washington DC

Department of the Environment (1989), *Environment in Trust: Air Quality*.

Department of Transport (1987), *COBA 9*, Department of Transport, London.

Department of Transport (1989a), *The allocation of road track costs 1989/90*, Department of Transport, London.

Department of Transport (1989b), *Transport Statistic Great Britain 1978-88*, Department of Transport, London.

Diamond, P.A., and Mirrlees, J.A. (1971), 'Optimal taxation and public production, I: productive efficiency', *American Economic Review*, 61, 8-27.

Hall, F.L., B.L. Allen, and M.A. Gunter (1986), 'Empirical Analysis of Freeway Flow-Density Relationships', *Transportation Research* 20A, 197-210.

Harrison, W.J., C. Pell, P.M. Jones and H. Ashton (1986), 'Some advances in model design developed for the practical assessment of road pricing in Hong Kong', *Transportation Research* 20A, 135-44.

Jones-Lee, M.W., (1990), 'The Value of Transport Safety', *Oxford Review of Economic Policy*, this issue.

Keeler and Small, (1977), 'Optimal Peak-Load Pricing, Investment and Service Levels on Urban Expressways', *Journal of Political Economy*, 85(1), 1-25.

Kraus, Marvin (1981), 'Scale Economies Analysis for Urban Highway Networks', *Journal of Urban Economics*, 9(1), 1-22.

Newbery, D. M. G. (1987), Road User Charges and the Taxation of Road Transport. *I M F Working Paper WP/87/5* International Monetary Fund, Washington DC.

Newbery, D.M. (1988a), 'Road User Charges in Britain', *The Economic Journal*, 98 (Conference 1988), 161-76.

Newbery, D.M. (1988b), 'Road Damage Externalities and Road User Charges', *Econometrica*, 56(2), 295-316.

Newbery, D. M. (1988c), 'Charging for roads', *Research Observer*, 3 (2), 119-38.

Newbery, D. M., (1989), 'Cost recovery from optimally designed roads', *Economica*, 56, 165-85.

Newbery, D. M., (1990), 'Acid Rain', *Economic Policy*, 11, October.

Newbery, D.M., Hughes, G. A., Paterson, W. D. O. and Bennathan, E. (1988), *Road Transport Taxation in Developing Countries: The Design of User Charges and Taxes for Tunisia*. World Bank Discussion Papers, 26, Washington, D.C.

ECMT (1989), European Conference of Ministers of Transport Round Table 80, *Systems of Infrastructure Cost Coverage*, Economic Research Centre, OECD, Paris.

Paterson, W. D. O. (1987), *Road Deterioration and Maintenance Effects: Models for Planning and Management*, Baltimore: Johns Hopkins University Press for World Bank.

Small, K. A., Winston C. and Evans, C. A. (1988), *Road Work: A New Highway Policy*, Washington DC Brookings. (In Press).

U.S. Federal Highway Authority (1982), *Final Report to the Federal Highway Cost Allocation Study*. Washington, D.C.:U.S. Government Printing Office.

World Bank (1986), *Urban Transport*, Washington D.C.

Appendix: Update to 1999

This article was first published in 1990 and updated in 1995. Since then, road taxes have substantially increased as a result of the fuel tax escalator. This was originally introduced by the Conservative government and increased the rate of duty on motor fuels by 5% above the rate of inflation each year. The original defence was that increasing fuel taxes were required to control vehicle emissions, particularly carbon dioxide, in order to meet the countries' targets for total carbon emissions. The labour government committed itself to raise the fuel tax escalator to 6% in real terms each year while at the same time freezing or reducing expenditure on new roads. Expenditure on road maintenance has only increased slightly, despite the substantial increase in traffic, and as a result the road quality has deteriorated (as measured by the defects index of road conditions published in the DETR Transport Statistics Great Britain).

By 1997/98, road tax revenue was £24 billion, compared with total road costs of £11 billion, measuring the interest on the infrastructure at 6%). With road traffic growing at just over 2% per annum, and the real duty rising at 6% per annum, real tax revenue will rise at about 8% per annum, rapidly raising the excess of road tax revenue over road costs.

Given the emphasis placed by the government on sustainable development, the obvious questions to ask are what does one mean by sustainability for transport, and what implications does that have for the taxation of transport. In its document on *Sustainable Development* (HMSO, 1994, p.169), sustainability is taken to mean that "users pay the full social and environmental cost of their transport decisions, so improving the overall efficiency of these decisions for the economy as a whole and bringing environmental benefits." Newbery (1998a) reviewed the evidence on social and environmental costs to assess what sum should be added to road costs to give the total costs of road transport. His estimates were somewhat lower than some other sources, at £2-8 billion per year, mainly because he attached a lower cost to the mortality effects of air pollution. Some commentators have valued the mortality from air pollution at the same rate per person as accidents (£2 million per person), but the overwhelming medical evidence is that pollution shortens lives by weeks or possibly months rather than the expected value of forty years for a road accident. If one adjusts for life years lost as a result of pollution, the figures are dramatically reduced. Nevertheless, the degree of uncertainty about these social and environmental costs remains high, as indicated by the factor of 4 to 1 in the range from the low to the high figure.

Even if the high figure is added to the road costs, and even if road maintenance currently understates the depreciation of the network (as the quality is deteriorating), making the necessary adjustments to give the total social cost of road use results in costs substantially less than current road taxes.

The present debate on road pricing has two strands. The first and more fundamental question is whether the charges that road users pay should be related to the costs of the road system, as with other public utilities. The case for separately identifying these road costs is that they could be levied as road charges with the revenue allocated to one or more agencies charged to provide road services. These agencies would be the road counterpart of Railtrack, the privatised owner of the railway infrastructure, and would be subject to regulation to ensure that road charges were set at the efficient level and to ensure that the agencies delivered an acceptable quality of service. Environmental ("green") taxes could then be levied on road fuel as on other fuels to reflect the environmental costs of various emissions. Both road charges and environmental charges would be subject to VAT, just as the current road fuel excise tax is subject to VAT.

The second issue is how best to levy the road charges. Road pricing implies that different roads charge different prices depending on their scarcity and congestion. Sophisticated road pricing would require electronic equipment, and would only be justified if the improvement in the efficiency of road use were greater than the costs of installing the system. Until that time, the road charge could continue to be collected by fuel charges and vehicle excise duties. When the technology is sufficiently mature and the cost-benefit analysis indicates that the time is ripe, these road charges can be reduced and replaced by road prices yielding the same total revenue. The regulator of Roadtrack would ensure that total revenues did not exceed the regulated level of costs. The regulator would therefore give road users the assurance that road pricing was not intended as an additional charge but as a more efficient way of collecting the amount deemed appropriate.

A decade after the original article, there is growing acceptance that transport policy is in disarray, that road taxes can no longer be defended as reflecting road costs including environmental costs, and that the system of financing transport investments is deeply unsatisfactory. Newbery (1998b) discusses various solutions to these problems, but other countries such as Australia and New Zealand are further advanced in contemplating the proper financing of the road system.

Economic principles continue to be relevant for the design of transport policy and these principles have not changed. What has changed are the magnitudes of the costs, and the more interesting problems of the design of institutions to manage investment and maintenance and mechanisms to improve the efficiency of road use. No doubt the next decade will provide more evidence to guide practical policy making.

References

HMSO (1994) *Sustainable Development: the UK Strategy*, London: HMSO Cm 2426

Newbery, D.M. (1998a) *Fair payment from road-users: A review of the evidence on social and environmental costs*, Report published by Automobile Association, February

Newbery, D.M. (1998b) ‘Fair and efficient pricing and the finance of roads’, (The 53rd Henry Spurrier Memorial Lecture), *The Proceedings of the Chartered Institute of Transport in the UK*, **7**(3), October, 3-19