ISSN 0265-8003

THE FUNDAMENTAL THEOREM OF ROAD USER CHARGES

David M Newbery

Discussion Paper No. 59
April 1985

Centre for Economic Policy Research
6 Duke of York Street
London SW1Y 6LA

Tel: 01 930 2963

CEPR Discussion Paper No. 59
April 1985

The Fundamental Theorem of Road User Charges*

ABSTRACT

Highway Authorities are divided as to whether to charge road users the average cost of road maintenance or the social marginal cost of road use, which includes not only the extra maintenance costs, but also the increased vehicle operating costs attributable to the road damage caused by vehicle passage. The Fundamental Theorem states that in a steady state with a consistent (but not necessarily optimal) maintenance policy the two concepts are identical for zero traffic growth on uncongested roads for a wide class of road damage and vehicle operating cost functions, and approximately equal in other cases.

JEL classification: 615

Keywords: road user charges, road damage, average cost, social marginal cost

David M Newbery
Churchill College
Cambridge
(0223) 61200
On leave until 5 June 1985

Woodrow Wilson School
Princeton University
Princeton, NJ 08544
USA
(609) 452 4825

*The major part of this November 1984 paper was completed whilst the author was on leave at the World Bank, as part of a research project on Taxing and Pricing Transport Fuels. He is greatly indebted to W.D.O. Paterson for continuous help and enlightenment on all matters to do with roads, without implicating him in the interpretation presented here. The World Bank does not accept responsibility for the views expressed herein which are those of the author and should not be attributed to the World Bank or its affiliated organisations. The findings, interpretations, and conclusions are the results of research supported by the Bank; they do not necessarily reflect official policy of the Bank.

NON-TECHNICAL SUMMARY

In 1979 $37.5 billion was allocated for highway purposes in the US, of which about 60 per cent was collected by taxes and charges levied on highway users. Given the sums involved, governments are everywhere concerned with how road user charges should be levied, and in particular, whether they should be <u>equitable</u> or <u>efficient</u>. The Federal Highway Administration interprets equity to mean that "users should pay for the highway costs they occasion" where costs are defined as highway agency expenditures. The efficient charge, on the other hand, is equal to the marginal social cost incurred by the vehicle, and includes not only the extra damage done to the highway, which will require maintenance expenditure by the highway agency, but also the increase in operating costs of other road users. On uncongested roads, the source of these extra vehicle operating costs will be the damage done to the pavement by earlier vehicles. This damage, measured by the increased roughness of the pavement surface, is now known to increase vehicle operating costs appreciably, and is a major component of the social marginal cost of road use. Indeed, these vehicle operating costs typically exceed highway maintenance costs by a factor of ten or more, and so it would seem to be of major importance to measure these costs and correctly allocate them.

Governments are tempted to choose the equitable charge approach, as it covers government costs, sounds acceptable, and avoids the very considerable problem of calculating the efficient costs, which require a knowledge of the road deterioration and vehicle cost relationships, both hard to estimate accurately.

The Fundamental Theorem states that in a steady state with consistent (but not necessarily optimal) maintenance policy the two concepts are identical for zero traffic growth on uncongested roads for a wide class of road damage and vehicle operating cost functions. The two concepts are approximately equal in other cases.

This result thus greatly simplifies the problem of calculating road user charges, and solves the potential conflict between the two criteria. It does not, however, deal with two other important components of road user charges - recovering the original capital costs, and internalising congestion costs. It does strengthen the case for allocating maintenance costs in proportion to the damaging power of different vehicles (i.e. essentially all to heavy lorries).

The Fundamental Theorem of Road User Charges

by David M. Newbery
Churchill College
Cambridge

November 1984

In 1979 $37.5 billion was allocated for highway purposes in the U.S., of which about $22.8 billion was collected by taxes and charges levied on highway users. Given the large sums involved, it is economically and politically important to be able to defend any proposed cost allocation between different classes of road users, and as a result a whole series of cost allocation studies have been commissioned by federal and state agencies. The latest in this line has recently been published by the Federal Highway Administration (1982). The United States is not alone in experiencing a rapid increase in road maintenance expenditure as large sections of its road network approaches the end of its design life, and other countries, especially developing countries, are now attempting to decide how best to allocate road costs to the various classes of road users.

Any cost allocation exercise involves answering two questions. First, what is the relationship between the vehicles using the road and the costs they cause, and second, what principles should guide the allocation of these costs to vehicles. The first question has been the subject of extensive (and expensive) research by engineers, and we now have a much clearer understanding of the complex processes involved, though there remains substantial uncertainty about the exact form and parameter values of the various cost relationships. The second question is pre-eminently one which economists ought to have a comparative advantage in addressing. The key issue which concerns Congress and other Governments is whether road user charges should be <u>equitable</u> or <u>efficient</u>. The Federal Highway Administration interprets equity to mean that "users should pay for the highway costs they occasion" where costs are defined as highway agency expenditures. The efficient charge, on the other hand, is equal to the marginal social cost incurred by the vehicle, and includes not only the extra damage done to the highway, which will require maintenance expenditure by the highway agency, but also the increase in operating costs of other road users. On uncongested roads, the source of these extra vehicle operating costs will be the damage done to the pavement by earlier vehicles. This damage, measured by the increased roughness of the pavement surface, is now known to increase vehicle operating costs appreciably, and is a major component of the social marginal cost of road use. Indeed, these vehicle operating costs typically exceed highway maintenance costs by a factor of ten or more, and so it would seem to be of major importance to measure these costs and correctly allocate them.

Governments everywhere are typically more concerned with expenditures borne by the Government than those borne by individuals, and when it is also deemed equitable to charge road users just for these expenditures, then the economist is likely to have some difficulty in suggesting his alternative approach. This difficulty is compounded by the far greater complexity of his calculation. The equitable charge is readily calculable and transparent - take highway expenditures, divide them by the number of damaging units (equivalent standard axle miles) and hence obtain the cost per ESA mile. Different classes of vehicle contribute different numbers of ESA miles p.a., and hence the costs can readily be calculated. In contrast, the efficient charge requires a knowledge of the damage done by a given vehicle to the pavement (which will depend on the characteristics of the pavement), and a knowledge of the effect this damage has on vehicle operating costs. Typically this calculation is performed by a large computer simulation model (see, for example, Markow and Wong, 1983). The resulting estimates are often hard to interpret, and their derivation anything but transparent. They frequently raise but leave unanswered such questions as - are the estimated costs high because the roads were poorly designed, poorly maintained, or because the vehicle was excessively laden? In short, whose fault is it that the costs are what they are, and, by implication, is it fair to penalise road users for the mistakes of the highway authority? If the reasons for the estimate are opaque, then emotive arguments of this sort are likely to cloud the issue.

It would therefore seem very desirable to know what is the relationship between equitable and efficient charges. If they can be shown to be close, then it is relatively unimportant to argue the difficult case for the efficient pricing of roads, but it they are very different, then the costs of inappropriate road charges might be quite substantial. Since there remains some uncertainty about the various damage relationships, it would be particularly useful to be able to make statements about equitable and efficient charges which were true of a wide class of damage relationship. The aim of this paper is to specify the circumstances in which the two charges are <u>equal</u>, and then explore their relationship when these circumstances are not satisfied. The main result, which is quite remarkable, and which is termed the Fundamental Theorem of Road User Charges, is that if all highway damage is attributable to traffic, and if highways are maintained when this damage reaches a predetermined (but not necessarily optimally determined) state, then, in the absence of traffic growth, the equitable and efficient road user charges are identical regardless of the exact form of the damage relationships.

This result is immediately applicable to unpaved roads which are of considerable importance in developing countries. It requires two modifications to be applicable to paved roads. First, traffic growth over the maintenance cycle (which may be of the order of 15 years) must be allowed for. Second, the effect of weather which also contributes to road damage, implies that not all damage is attributable to vehicles. This effect also needs to be quantified.

These modifications qualitatively reduce the efficient charge below the equitable charge, and quantitatively their importance will depend on the particular damage relationships chosen. Fortunately it is possible to derive analytical expressions for the efficient charge using the present best estimate of these damage relationships, and it is therefore relatively easy to calculate numerically the value of the efficient charge, and to test its dependence on particular features of interest (such as the road strength, the maintenance strategy, the traffic composition, etc.). Whilst large computer simulation models will remain useful for capturing more subtle effects, the approach proposed here can be applied quickly in a given country once the basic information is available, and will not require the often expensive software development needed to adapt these simulation models.

Formulating the Problem

When a vehicle drives on a road, it causes some damage to the surface and advances the date at which the road should be repaired. We suppose that the Highway Authority repairs roads when they reach some predetermined level of distress, and this seems not only sensible, but a reasonable description of reality. In general the criterion for maintenance will depend on the level of traffic, possibly the political clout of the road users, and the stringency of the Authority's budget constraint, and there is no reason to suppose that the criterion will be socially optimal, but we assume that however it is determined, the maintenance criterion is applied consistently to any given road over time, and does not alter.

The most important type of damage done to a road is best measured by the increased roughness of the surface, which can be quantified by a variety of different instruments, such as the Bump Integrator, or assessed subjectively (though with reasonable precision). For a well-designed road, its initial roughness will be low, and will steadily increase with the passage of traffic until, after 10-20 years, it reaches a level at which major maintenance, such as an asphalt overlay, is required to restore the surface to its initial low level of roughness. For an unpaved road the process of deterioration is very much faster, and the road may need to be bladed every few months, or, for more durable unpaved roads, graded every few years, to reduce the level of roughness.

Paved roads often exhibit distress modes such as cracking, rutting, ravelling, potholing etc., over shorter periods (5-15 years), and require minor maintenance (filling potholes, possible resealing, etc.). The timeliness of this maintenance will affect the rate at which the road subsequently deteriorates, as measured by roughness, and again we suppose that this minor maintenance policy is constant over time, though not necessarily optimal.

Vehicle operating costs depend on the quality of the road surface as measured by its roughness (and also on road geometry, which, however, remains constant). The damage vehicles do to roads depends on the characteristics of the vehicle and the type of road, and, for any given type of road, the damaging power of a vehicle can be expressed as some fraction or multiple of a standard damaging unit. Thus on paved roads, the damaging power of a single axle is considered to be proportional to the fourth power of its load, and the standard is taken as a load of 18000 lbs or 80 Kilo Newtons, described as an Equivalent Standard Axle Load (ESAL). Trucks may vary from less than 0.1 ESALs to more than 50 ESALs for (illegally) overloaded vehicles.[1] Private cars have completely insignificant damaging factors. On unpaved roads, however, axle load is far less important, and cars may have almost as much damaging effect as trucks (between 30% and 100% of the damaging effect).

---

[1]Jones [1977] reports that about 5% of 3 and 5 axle vehicles (the worst offenders) had damaging factors of greater than 50 ESALs in Kenya in 1974 when the legal limit was less than 5 ESALs. Only 2% of heavy vehicles in the U.K. on the M1 had damaging factors greater than 12 in 1964.

Determining the damaging power of any given vehicle on a particular surface is an extremely complex problem of first order importance, since small differences in the exponent of the damage load can dramatically shift the responsibility for road damage between different classes of road users, and alter the social desirability of heavily laden large trucks. More research on this aspect of the problem is undoubtedly needed (see, for example, OECD, 1982, Chapter 6), but is not central to our present concern, where we assume that the damaging power of each vehicle is known, and hence also that of the whole traffic stream, or of the representative vehicle. For brevity, suppose that this damaging power is measured by ESALs, with the appropriate interpretation for unpaved roads.

We shall ignore all costs that are independent of road or vehicle damage, such as verge maintenance, lighting, police, pollution, etc. and also ignore the costs of congestion and accidents. As such the model developed below is more suitable for uncongested inter-urban roads, though the other externality costs can of course be separately computed and added in where necessary to more accurately measure the social marginal costs incurred by particular vehicles on specified roads.[2]

Efficient pricing of roads

The efficient price to charge a vehicle for using a given road is the extra social cost which the passage of the vehicle causes. Were the vehicle to be charged this price, it would then make the correct (efficient) decision as to whether the journey was justified. In practice, it is unrealistic to charge separately for each road, and we shall be interested in the average cost of a given type of vehicle using a given strength of road. (This can be further

[2]There is a more subtle form of congestion which probably should be included. Road maintenance typically disrupts traffic and increases vehicle costs (if only through time loss), and hence these additional costs should be included in the calculation of efficiency charges. They will typically be small for lightly trafficked inter-urban roads, but may be large on already congested roads.

averaged over all roads if necessary). Charging vehicles of a given type this average efficient price should then lead to the right choice of vehicles and the right degree of utilization on average, though it may mean that some roads are over or under used relative to the first best system of pricing. In particular, we shall be primarily concerned with the relationship between the efficient price to charge a vehicle, and the average cost of maintaining the road, allocating these costs in proportion to the damage done to the road. This cost allocation procedure has been advocated as the equitable charge to levy for road use.

The Fundamental Theorem

Suppose that traffic is constant at N vehicles per annum, and that the average vehicle inflicts E equivalent standard axles of damage to the road. If up to date z the number of cumulative standard axle transits has been X, then the cumulative total by date f will be

$$X(t) = X - NE(t-z), \quad t > z. \tag{1}$$

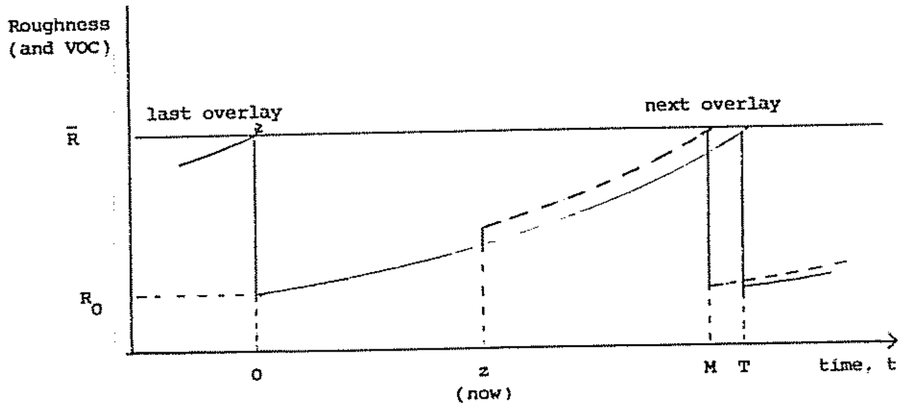
The road damage relationship gives roughness R at date t as some function of cumulative ESALs:

$$R = R\{X(t)\}. \tag{2}$$

The average vehicle operating cost, v, depends on roughness, R:

$$v = v(R). \tag{3}$$

The maintenance strategy is to restore the road to roughness R_o whenever the roughness reaches a predetermined critical level $\bar{R}$, at a cost C. At constant traffic levels, this will recur every T years, where T depends on the strength of the road, the traffic, and $\bar{R}$. Figure 1 shows the time path of roughness and, on a different scale, of vehicle operating costs (VOC).



<u>Fig 1</u> <u>The effect of an extra standard axle on roughness and costs</u>

The present discounted value of all future vehicle operating and road maintenance costs at date z can be written as

$$P = e^{-r(M-z)} \frac{C}{1-e^{-rT}} + N \int_z^M v(R(t))e^{-r(t-z)}dt$$

$$+ e^{-r(M-z)} \frac{N \int_0^T ve^{-ru}\, du}{1-e^{-rT}}, \tag{4}$$

In equation (4) we distinguish between the life of the present road, M, and the design life of the road after overlay, T. These will coincide if traffic was as predicted. The first term is the cost of repeatedly resurfacing the road every T years, the second term is the operating cost of all vehicles from now until the next overlay at date M, and the last term is the operating cost of all subsequent vehicles.

The marginal social cost of an extra standard axle transit at z can be found by varying X in (1), which will affect R and hence M and v in (4). Differentiate (1) to obtain

$$\frac{dF}{dX} = -re^{-r(M-z)} \frac{C}{1 - e^{-rt}} \frac{\partial M}{\partial X} + N \int_z^M \frac{dv}{dR} \frac{\partial R}{\partial X} e^{-r(t-z)} dt$$

$$+ N\, v(\bar{R}) e^{-r(M-z)} \frac{\partial M}{\partial X} - re^{-r(M-z)} N \frac{\partial M}{\partial X} \frac{\int_0^T ve^{-ru} du}{1-e^{-rT}}, \tag{5}$$

The date of overlay, M, is implicitly defined by the total number of ESALs required to cause the road surface to reach its critical level of roughness, $\bar{R}$:

$$\bar{R} = R(\bar{X}), \qquad \bar{X} = X + NE(M - z). \tag{6}$$

The dependence of the date of overlay, M, on X is found by differentiating (6):

$$\frac{\partial M}{\partial X} = -\frac{1}{NE}, \tag{7}$$

whilst roughness at date t, R(t), depends on X(t), and hence, from (1) on both X and t. From (1)

$$\frac{\partial R(t)}{\partial X} = \frac{\partial R(t)}{\partial X(T)} = \frac{1}{NE} \frac{dR(t)}{dt}$$

This, together with the expression for $\partial M/\partial X$ from (7), and $M = T$, can be substituted in (5) to yield

$$\frac{dF(z)}{dX} = re^{-r(T-z)} \frac{C/NE}{1-e^{-rT}} + \frac{1}{E} \int_z^T \frac{dv}{dR} \frac{dR}{dt} e^{-r(t-z)} dt$$
$$- \frac{v(\bar{R})}{E} e^{-r(T-z)} + re^{-r(T-z)} \frac{\frac{1}{E} \int_0^T ve^{-ru} du}{1-e^{-rT}}. \tag{8}$$

The efficient charge as calculated depends on the age of the road since the last overlay, z. and the average efficient charge averages this over all roads of ages 0 to T:

$$\overline{\frac{dF}{dX}} = \frac{1}{T} \int_0^T \frac{dF(z)}{dX} dz \tag{9}$$

It is instructive to consider the maintenance costs (the first term in (8) separately, for which the average charge is, from (8) and (9)

$$\frac{r}{T} \int_0^T e^{-r(T-z)} dz \frac{C/NE}{1-e^{-rT}} = \frac{C}{NTE}. \tag{10}$$

which is exactly the average maintenance cost, or the equitable road user charge. The remaining terms cancel, as the following argument shows. Consider the second term in (8), and let

$$\int_z^T \frac{dv}{dt} e^{-rt} dt = \Phi(z),$$

so that

$$\frac{d\Phi}{dz} = \frac{dv(z)}{dt} e^{-rz}.$$

The time average of the second term in (8) is then, integrating by parts:

$$\frac{1}{TE} \int_0^T dz\, e^{rz}\Phi(z) = \frac{1}{rTE} \{[e^{rz}\Phi(z)]_0^T - \int_0^T e^{rz} d\Phi\}$$

$$= \frac{1}{rTE} \{ -\int_0^T \frac{dv}{dt} e^{-rt} dt + \int_0^T dv\}. \qquad (11)$$

The time average of the last two terms in (8) is

$$\frac{1}{TE} \{-v(T) \int_0^T e^{-r(T-z)} dz + \int_0^T v(u) e^{-ru} du\} = \frac{1}{TE} \int_0^T \{v(u) - v(T)\} e^{-ru} du. \qquad (12)$$

Again, this can be integrated by parts to yield

$$\frac{1}{rTE} \left[\{v(u) - v(T)\} e^{-ru}\right]_T^0 + \frac{1}{rTE} \int_0^T \frac{dv}{du} e^{-ru} du.$$

$$= -\frac{1}{rTE} \{v(T) - v(0)\} + \frac{1}{rTE} \int_0^T \frac{dv}{du} e^{-ru} du \qquad (13)$$

The average externality costs are the time averages of the last three terms in (8), or the sum of (11) and (13), and hence identically zero.

The intuitive reason why the externality costs (effect on vehicle operating costs) can be ignored is that the effect of an extra vehicle is to advance slightly the whole saw tooth pattern of costs in Figure 1. The effect is to advance both higher costs (the later part of the cost cycle) and lower costs (the earlier part), so that on average they cancel.

This is a remarkable result, for it is independent of the exact form of the road damage relationship, $R\{X(t)\}$, of the vehicle damage relationship, $v(R)$, and of the road maintenance criterion, $\bar{R}$. All that is required is that R depend only on $X(t)$, that v depends only on R, that the maintenance strategy

If the maintenance frequency T is optimally chosen, $\partial P/\partial T = 0$, and the marginal social cost will then be equal to the marginal externality costs. Thus the efficient charge will again coincide with the equitable charge when the maintenance frequency is fixed, independent of the state of the road, provided this frequency is optimally chosen. If the frequency is not optimal, then the efficient charge may be above or below the equitable charge, depending on whether maintenance is too infrequent or too frequent. Whilst one should not be too surprised that optimal maintenance leads to equality between the efficient and equitable charge, it should be stressed again that if maintenance is triggered by the road condition, and not at a set interval, any constant maintenance strategy leads to the same equality.

Extensions to the Fundamental Theorem

The road damage relationship relating roughness to traffic has been the subject of intense empirical investigation since the pioneering model estimated by AASHTO (Highway Research Board, 1962) from accelerated wear studies performed in Illinois in 1957-8. Paterson (1984) summarizes four road distress prediction models which were estimated from four major empirical studies, in Brazil, Kenya, Arizona and Texas. In all of these studies except the latest Brazilian study and the Arizona study, the damage done to the road depends on the number of cumulative ESAs, since construction (or reconstruction), $X(t)$, and not separately on the time which has elapsed, t. That is, if damage done is measured by the roughness of the road, R, then

$$R = R(X(t), Z) \qquad (15)$$

where Z is a vector of road and environmental characteristics. For example, the Kenyan model postulates that

$$R = R_0 + m(SNC).X(t) \qquad (16)$$

where SNC is the modified or corrected structural number of the road, and measures its strength.[3] The Texas model is more complex, but has the general form

$$R = f(\exp[-\rho/X(t)] \tag{17}$$

where f is a function relating roughness to the loss in serviceability since construction, and ρ is a parameter which depends on structural and climatic factors (pavement thickness, plasticity of subgrade soil, a measure of moisture, the number of annual freeze-thaw cycles, etc.)

The Arizona model is at the other polar extreme in supposing that all damage is caused by time and weather:

$$R = (R_0 + k)e^{mt} - k \tag{18}$$

where R_0 is initial roughness, and both m and k depend on elevation, average rainfall, number of freeze-thaw cycles, temperature, etc., but not on traffic.

The Arizona model in effect presupposes that roads are optimally designed for expected traffic, with more heavily trafficked roads being of stronger construction. Variations in rates of road deterioration then do not appear to depend on differences in traffic, leaving weather and time as the major apparent causes of damage. Whilst this is clearly an unsatisfactory feature of the model for calculating road user charges, it points to an important phenomenon not captured by the earlier relationships, namely that roads deteriorate both as a result of the passage of vehicles, and the elapse of time and weathering. Thus the road damage relationship should be written, not as in (2) or (15), but as

$$R(t) = R(X(t), t). \tag{19}$$

---

[3] AASHTO (1972) gives $m = k(1 + SN)^{-5.25}$, where k is a constant and SN is the unmodified structural number (a linear function of the layer thickness). Paterson (1984) gives more elaborate forms for the parameter m.

Whereas the earlier analysis of the data collected in the Brazilian study estimated a relationship similar to (15) a later reworking of the data found that the fit was improved by making incremental deterioration depend on the current state of the road, and incremental traffic. The final equation estimated can be integrated to give the following relationship:

$$R = e^{mt}\left[R_0^{1-\alpha} + (1-\alpha)kX(t)\right]^{\frac{1}{(1-\alpha)}} \tag{20}$$

where k depends on the road strength (measured by SNC), and both α and m are constants.[4] Two modifications must be made to the previous analysis. First $\partial M/\partial X$ in equation (4) is found by totally differentiating

$$\bar{R} = R(X + NE(M-z), M) \tag{21}$$

whence

$$\frac{\partial M}{\partial X} = \left.\frac{-1}{NE + \dfrac{\partial R/\partial t}{\partial R/\partial X(t)}}\right|_{t=T} \equiv \frac{1}{NE^*} \tag{22}$$

where E^* measures the <u>combined</u> effect of traffic and time. Second, equations (9) and (1) give

$$\frac{\partial R(T)}{\partial X} = \frac{\partial R}{\partial X(t)} = \frac{1}{NE}\left(\frac{dR}{dt} - \frac{\partial R}{\partial t}\right). \tag{23}$$

With these changes equation (6) becomes

$$\frac{dP(z)}{dX} = re^{-r(T-z)}\frac{C/NE^*}{1-e^{-rT}} + \frac{1}{E}\int_z^T \frac{dv}{dt}e^{-r(t-z)}dt - \frac{1}{E^*}v(\bar{R})e^{-r(M-z)}$$

$$+ re^{-r(T-z)}\frac{\frac{1}{E^*}\int_0^T ve^{-ru}du}{1-e^{-rT}} - \frac{1}{E^*}\int_z^T \frac{dv}{dR}\frac{\partial R}{\partial t}e^{-r(t-z)}dt. \tag{24}$$

---

[4]The form given in Paterson (1984, p.2) is

$$dR = kR^{\alpha}dX + mRdt$$

but this cannot be integrated to give a path-independent relation for R. However, a slight modification to

$$dR = kR^{\alpha}e^{(1-\alpha)mt}dX + mRdt$$

can be integrated to give (20), and behaves very similarly to the estimated relationship.

remains constant, and that there is no traffic growth. Traffic growth means that either the interval between overlays will steadily decrease, or that the road will need to be strengthened, and so the simplicity of the constant maintenance cycle is lost. Its effect will be considered below.

The importance of the maintenance strategy

There is good empirical evidence that Highway Authorities respond to the state of paved roads when deciding on major maintenance such as overlays, but unpaved roads have a much shorter life between resurfacings (or regrading) and the Highway Authority may therefore choose to resurface at fixed time intervals, rather than in response to the state of the surface. Thus in Tunisia, unpaved roads are bladed after the wet season and just before the harvest, to ensure that the harvest can be efficiently evacuated. This maintenance stragegy significantly alters the calculation of the efficient road user charge, for a vehicle transit at date z in Figure 1 now has no effect on the date of maintenance, T, but it does affect the roughness at T, R(T), and, for unpaved roads, this affects the initial roughness immediately after blading, $R(T+\epsilon)$, and hence thereafter. Road maintenance costs are independent of vehicle traffic, and hence the first term in equation (4) is zero. Consequently, the efficient charge will be just equal to the externality terms (the increase in subsequent vehicle operating costs), independent of the road maintenance costs, and thus the link between efficient and equitable charges has been broken. However, if the frequency of maintenance is <u>optimally</u> chosen, this link is restored, for the following reason. The marginal social cost of an extra ESAL is found by totally differentiating P in equation (4). It can also be written, replacing M by T, as

$$\frac{dF}{dX} = \frac{\partial F}{\partial X}\bigg|_{T\ \mathrm{const}} + \frac{\partial F}{\partial T}\frac{\partial T}{\partial X}, \tag{14}$$

The time average efficient charge is then

$$\overline{\frac{dP}{dX}} = \frac{C}{NTE^*} + \frac{1}{ET}\,{}_0\!\int^T {}_z\!\int^T \frac{dv}{dR}\left[(1 - \frac{E}{E^*})\,\frac{dR}{dt} - \frac{\partial R}{\partial t}\right] e^{-rt}dt\; e^{rz}dz. \qquad (25)$$

The first point to make is that if traffic does not contribute to road deterioration, then dP/dX will be zero, and the efficient road charge will also be zero. This extreme is implausible, and the first point to check is that the effect of time and weather does reduce the efficient charge below the equitable charge.

This is readily done. The first term in (25) is the average attributable maintenance cost, which is a fraction E/E^* of the average total cost. The Appendix demonstrates that for the Brazilian model of equation (20)

$$\frac{E^*}{E} = 1 + \frac{m(\alpha-1)T}{\left[\frac{\bar{R}}{R_0}\,e^{-mT}\right]^{\alpha-1} - 1} \qquad (26)$$

and typical numerical values for the parameters are shown in Table 1 below. Note that E/E^* is an increasing function of $(\alpha-1)mT$, and that T, the age at

Table 1 Values of Parameters for Brazilian Model

Parameters	Value	
α	1.42	
m	0.011 - 0.025	p.a.
T	15 - 25	years
R_O	25 - 35	QI units
R	75 - 120	"
E/E^*	0.80 - 0.90	

which the road is restored, and $\bar{R}/R_0$, the ratio of final to initial roughness, will typically be correlated.

Two points emerge from the calculations. First, E/E^* is a fraction, less than one, so the attributable maintenance costs are indeed reduced by weather. Second, the fraction is quite large and reasonably stable for different climates and maintenance strategies (at least, those observed in Brazil and Tunisia).

The second set of terms in equation (25), which is the vehicle externality cost (VEC), is shown to be negative in the Appendix, and so the efficient charge is reduced below the attributable maintenance cost.

Quantitative Importance

The Brazilian model was calibrated for Tunisian roads which experience a drier climate, and hence have a lower rate of time degradation, m, of 1.1% p.a.. A wide range of parameter values describing roads of different strengths, design lives, maintenance criteria, traffic volume, and damaging power per vehicle, were used in the sensitivity analysis reported in Table 2. The model takes the average daily traffic (ADT of col.4; total, both directions), the overlay criterion, $\bar{R}$, the design life T, and the damaging power of the average vehicle, E (col.5, in ESAs per 100 vehicles), and computes the design strength as the corrected structural number (SNC of col. 2) consistent with these parameters. The ratio E/E^* gives the fraction of maintenance costs to allocate to traffic, whilst the ratio of the vehicle externality cost to the average maintenance cost is given as a per cent in col. (6). Col. 8 gives the efficient charge per ESA per km to levy for the assumed parameters. Cols. 9 and 10 give the efficient charge on two different assumptions about the effect of weather. When m = 0 all costs are attributable to traffic, whilst when m = 2.45% p.a. (as in Brazil) then typically E/E^* will be 50-70% and the attributable costs will be appreciably lower.

The main conclusion is that the vehicle externality cost is quantitatively small and remarkably constant across alternative specifications, varying from 0.04 US¢/ESA km to 0.08 ¢/ESA km as traffic varies by a factor of 30 and m varies by a factor of 2. As a proportion of average costs it is small except for heavily trafficked strong roads, for which the maintenance cost per vehicle is so small.

Table 2 Sensitivity Analysis for Tunisian Data

	R QI units	SNC	T yrs	ADT veh/day	E per 100 veh	VEC/AC %	E/E^* %	Marginal Social Cost US¢/ESA km[c] m=0.011	m=0.0	0.0245
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A	80	1.28	15	500	37	-1.0	88	3.4	3.9	2.8
B	80	1.50	25			-1.8	80	1.8	2.3	1.1
C	100	1.43	25			-2.7	84	1.8	2.3	1.3
D	80	1.41	25		30	-2.3	80	2.3	2.9	1.4
E	80	1.53	15	1000	37	-1.6	88	2.1	2.4	1.7
F		1.75	25			-2.9	80	1.1	1.5	0.7
G	80	1.7	15	1500	37	-1.9	88	1.7	2.0	1.4
H		2.0	25			-3.6	80	0.9	1.2	0.6
I	80	3.1	15	15,000	37	-16.7	88	0.2	0.2	0.1
J		3.25	25			-31.2	80	0.1	0.1	0.0

Notes (a) blanks indicate previous values unchanged

(b) r = 8%, overlay costs C vary from $20,000 per km for the weak roads to $35,000/km for strong roads. R_0 = 25.

(c) Cols. 1-8 assume m = 0.011, col.9 m = 0.0, col.10 m = 0.0245.

Thus the main conclusion seems to be that in the absence of traffic growth the main effect of weather is to reduce the fraction of the maintenance costs which can be attributed to vehicles, and the interaction between weather and vehicle externality costs are small and can be essentially ignored.

The Effects of Traffic Growth with Weathering

The maintenance cycle for paved roads may be 15-25 years, during which time traffic may grow significantly. Equation (3) must now be modified in two ways. First, the terms involving traffic must reflect this growth, and second, the road will need additional strengthening to deal with the increase in traffic volume. The simplest maintenance strategy would be to increase the strength to preserve its design life over the next overlay cycle. If the design life is T years, the initial and final roughness are R_0 and $\bar{R}$, then equation (20) gives a relation between k (which depends on road strength) and X(T):

$$\left[\bar{R}\, e^{-mT}\right]^{1-\alpha} = R_0^{1-\alpha} + (1-\alpha)kX(T) \qquad (27)$$

$$X(T) \propto \frac{1}{k} = 0.926\ SN^{1/\sigma} \quad , \sigma = 0.2595 \qquad (28)$$

where SN is the corrected structural number of the pavement, and $1/\sigma$ measures the increasing returns to scale in road strengthening. Thus, if traffic grows at an annual rate g, and if the structural number of the road is increased by a factor $e^{\sigma gT}$ each T years, the road will be capable of carrying the increased traffic volume for a further T years before deteriorating to a roughness $\bar{R}$. Rolt (1981 p.7) gives the following equation for the costs of overlay

$$C_L = C_0 + \beta\ SN = C_0 + \beta SN_0 e^{n\sigma gt} \qquad (29)$$

for the nth overlay cycle. Vehicle operating costs will follow the same time profile on all overlay cycles, and are hence simple function of the time since the last overlay, as is roughness. The present discounted cost of all future maintenance is then

$$e^{-r(M-z)}\left[\frac{C_0}{1-e^{-rT}} + \frac{k\ SN_0 e^{g\sigma T}}{1-e^{-(r-g\sigma)T}}\right] = e^{-r(M-z)}C \qquad (30)$$

where, as in the first model, z is the current date, M is the date of the next overlay, and (M-z) is the time before the next overlay. Total social costs will be

$$F(X) = e^{-r(M-z)}C + \int_z^M N_0 e^{gT} v(R(t)) e^{-r(t-z)} dt$$

$$+ e^{-r(M-z)} \left[N_0 e^{gM} \frac{\int_0^T v(R(t)) e^{-(r-g)u} du}{1 - e^{-(r-g)T}} \right], \tag{31}$$

where N_0 was the traffic flow p.a. at the date of the last overlay and N_0e^{gt} the traffic flow at date t. The effect of an extra ESAL now (at date $t = z$) is to raise social costs, and the marginal social cost is

$$\frac{dF}{dX} = -re^{-r(M-z)}C \frac{\partial M}{\partial X} + N_0 \left[e^{gM} v(\bar{R}) e^{-r(M-z)} \frac{\partial M}{\partial X} + e^{rz} \int_z^M \frac{dv}{dR} \frac{\partial R}{\partial X} e^{-(r-g)t} dt \right.$$

$$\left. - (r-g) e^{rz} e^{-(r-g)M} \frac{\int_0^T ve^{-(r-g)u} du}{1 - e^{-(r-g)T}} \frac{\partial M}{\partial X} \right]. \tag{32}$$

The relationship between M and X can be deduced by totally differentiating

$$\bar{R} = R \left[X + \int_z^M N_0 E \, e^{gt} dt \, , M \right]$$

$$\frac{\partial M}{\partial X} = - \frac{1}{N_1 E + R_t / R_X} \bigg|_T \equiv - \frac{1}{N_1 E'}, \tag{33}$$

where $N_1 = N_0 e^{gT}$ is the traffic volume at next overlay, and E^* again measures the combined effect of traffic and time (cf equation 22). Similarly

$$\frac{\partial R(t)}{\partial X} = \frac{\partial R}{\partial X(t)} = \frac{1}{N(t)E} \left[\frac{dR}{dt} - \frac{\partial R}{\partial t} \right]. \tag{34}$$

If the road is overlaid after T years, then equation (32) can be written as

$$\frac{dF(z)}{dX} = \frac{e^{rz}}{E^*} \left[\frac{rCe^{-rT}}{N_1} + [v^* - \bar{v}] \, e^{-rT} + \frac{E^*}{E} \int_z^T \frac{dv}{dR} \left[\frac{dR}{dt} - \frac{\partial R}{\partial t} \right] e^{-rt} dt \right] \tag{35}$$

where C is defined by (30) and

$$v^* = \frac{\int_0^T ve^{-(r-g)u}du}{\int_0^T e^{-(r-g)u}\,du} \quad , \qquad \bar{v} = v(\bar{R}).$$

The time averages can again be computed, and Table 3 shows the effect of traffic growth on the ratio of the efficient charge to the average total maintenance cost.

As the importance of weathering increases (as m increases) so the ratio decreases, as one would expect. For low and moderate traffic volumes the effect of traffic growth is to further reduce the ratio, but for high volume strong roads, traffic growth raise the ratio, in some cases above 100%, but weathering has a larger proportionate effect on the ratio, and can even make the marginal social cost negative.[5] It should be remembered that the high volume roads have very low average costs anyway, and these are swamped by the increasingly important vehicle externality costs.

Table 3 Effect of Growth on Allocatable Costs

ADT veh/day	T years	MSC/AC per cent g=0% p.a. m=0.011	g=5% m=0.011	g=5% m=0.0	g=0% m=0.0245	g=5% m=0.0245
(1)	(2)	(3)	(4)	(5)	(6)	(7)
500	15	87	78	87	71	65
	25	78	61	76	49	40
1500	15	86	58	69	69	46
	25	76	39	63	47	16
15000	15	72	85	141	47	30
	25	48	60	258	22	-98

Notes $R_0 = 25$, $\bar{R} = 80$, $r = 8\%$, $E = 37/100$ vehicles.

[5] A negative MSC implies that the maintenance period T is too long, and that it pays to damage the road in order to advance the date of overlay, and provide a better surface for the higher volume of later traffic.

Conclusions

The Fundamental Theorem states that if road damage is solely caused by traffic and if there is no traffic growth, then provided the road is restored whenever its roughness reaches a predetermined level, the average efficient road user charge is equal to the equitable charge, or the average cost of the repair. In other words, the externality caused by vehicles damaging roads, which raises the operating cost of subsequent vehicles, exactly cancels out when averaging over roads of differing age.

This result is modified if weather is a significant source of road damage, in which case only a fraction of the road costs can be attributed to vehicles. In dry climates the effect is to reduce the efficient charge to about 75-90% of the equitable charge, except on very heavily trafficked roads for which the equitable charge is in any case very low. In more severe climates the attributable fraction is lower, but again, the externality costs can be essentially ignored.

The result is further modified if traffic grows steadily, and traffic growth can make the efficient charge either larger or smaller than the equitable charge, depending on the maintenance frequency and design strength of the road. Again, for strong low volume roads for which the equitable charge is high, the effect of growth is to further lower the ratio of efficient to equitable charges. For high volume weak roads, with a low equitable charge, the relationship is very sensitive to weathering, traffic growth, and road strength, but the cost levels are very low so this sensitivity is not very important. The main conclusion is that efficient road pricing will in general not recover maintenance costs, let alone other, non-traffic related costs (such as policing, lighting, and, most important, interest on the original investment in the road). Congestion charges, on the other hand, will more than cover the costs of congestion borne by the Highway Authority, and would go some way to meeting the deficit on the efficient pricing of uncongested roads.

References

Federal Highway Administration (1982), Final Report on the Federal Highway Cost Allocation Study, Department of Transportation, Washington DC.

Highway Research Board (1962), The AASHO Road Test: Pavement Research, HRB Special Report 61E, Washington DC.

Jones, T.E. (1977), Axle-loads on paved roads in Kenya, TRRL Laboratory Report 763, Transport and Road Research Laboratory, Crowthorne, Berks.

Markow, Michael J. and Wong, Thomas K. (1984), 'Life-Cycle Pavement Cost Allocation'. Transportation Research Record 900, Washington DC.

OECD (1983), Impacts of Heavy Freight Vehicles, Paris.

Paterson, W.D.O. (1984), 'Summary of Distress Prediction Models from Major Studies: Brazil, Kenya, Arizona and Texas', World Bank Transport Department, Feb. 6.

Rolt, J. (1981), Optimum Axle Loads of Commercial Vehicles in Developing Countries, TRRL Laboratory Report 1002, Transport and Road Research Laboratory, Crowthorne, Berks.

Appendix The relative importance of time and traffic

The fraction of maintenance costs attributable to vehicles is E/E, where E is defined by equation (22):

$$\frac{E^*}{E} = 1 + \left.\frac{\partial R/\partial t}{NE\partial R/\partial X(t)}\right|_{t=T},$$

If R(X(t),t) is given by (20), then

$$\frac{1}{NE}\,\frac{R_t}{R_X} = \frac{m\left[\bar{R}e^{-mT}\right]^{1-\alpha}}{NEk},$$

Also from (20)

$$\left[\bar{R}e^{-mT}\right]^{1-\alpha} = R_0^{1-\alpha} + (1-\alpha)kX(T)$$

and X(t) = NET, so

$$NEk = \frac{R_0^{1-\alpha} - \left[\bar{R}e^{-mT}\right]^{1-\alpha}}{(\alpha - 1)T}.$$

Hence

$$\frac{E^*}{E} = 1 + \frac{m(\alpha - 1)T}{\left[\frac{\bar{R}e^{-mT}}{R_0}\right]^{1-\alpha} - 1}.$$

The sign of the term in square brackets in equation (25) is found as follows.

Define $\Phi(t) = \frac{\partial R/dt}{NE\partial R/\partial X(t)}$

so that

$$\frac{E^*}{E} = 1 + \Phi(T)$$

and

$$\frac{dR}{dt} = \frac{\partial R}{\partial t}\left[\frac{1 + \Phi(t)}{\Phi(t)}\right].$$

Then

$$\left[1 - \frac{E}{E^*}\right]\frac{dR}{dt} - \frac{\partial R}{\partial t} = \frac{\Phi(T) - \Phi(t)}{\Phi(t)\left[1 + \Phi(T)\right]}.$$

But

$$\Phi(t) = \frac{m\left[Re^{-mt}\right]^{1-\alpha}}{NEk} = \frac{m}{NEk}\left[R_0^{1-\alpha} + (1-\alpha)kNEt \right]$$

so

$$\frac{d\Phi}{dt} = (1-\alpha)m < 0 \quad , \quad \Phi(t) > 0,$$

and the expression is negative.