Road Damage Externalities and Road User Charges

Author(s): David M. Newbery
Source: *Econometrica*, Vol. 56, No. 2 (Mar., 1988), pp. 295-316
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1911073
Accessed: 10/11/2010 11:39

Econometrica, Vol. 56, No. 2 (March, 1988), 295–316

ROAD DAMAGE EXTERNALITIES AND ROAD USER CHARGES[1]

BY DAVID M. NEWBERY

Vehicles on uncongested roads damage the pavement and advance the date at which repairs are necessary. Recent empirical work has identified another potentially more important social cost, for the damage to the pavement raises the operating costs of subsequent vehicles, and these operating costs may be an order of magnitude larger than the road maintenance costs. The paper develops a theory to handle this newly identified externality and establishes a remarkable result. If roads are repaired when they reach a predetermined critical condition (not necessarily optimally set) and if road damage is fully attributable to traffic, then in steady state with zero traffic growth the average road damage externality is zero, and the average marginal social cost of road use is equal to the average road maintenance cost. The result holds for arbitrary road damage and vehicle damage functions. Recent empirical studies reveal that weather accounts for a significant fraction of road deterioration, and in this case the road damage externality is no longer exactly zero, but is quantitatively negligible. The appropriate road user charge now only recovers the fraction of damage attributable to traffic.

KEYWORDS: Externalities, efficient pricing, road user charges.

IT HAS LONG BEEN RECOGNIZED that road users create externalities, notably congestion, but their study falls into four quite widely separate phases. The first phase was the theoretical exploration of optimal pricing and investment rules for congested and uncongested roads, which were developed by Ellet (1840), Dupuit (1844), Pigou (1912), and Knight (1924). The second stage really began with Walters' (1961) attempt to quantify the congestion externalities and the implied optimal tolls. This had to wait for the underlying congestion relationships to be empirically estimated, mainly by traffic engineers in the first instance (see the references in Walters (1961)). The magnitude of the congestion costs appeared impressively high for congested urban streets, and the fact that the problem had been quantified and shown to be important stimulated a great deal of subsequent research (see the admirable survey by Winston (1985)).

Walters had been primarily concerned with pricing issues, and hence with the short run marginal social cost of extra traffic on a given road system. Other writers were quick to take up the related theme of the optimal investment rule, and Mohring and Harwitz (1962) and Mohring (1976) pointed out that if road capacity experienced constant returns and could be continuously adjusted, then optimal congestion tolls would exactly recover the costs of providing the optimal

[1] The major part of this paper was completed while the author was on leave at the World Bank, as part of a research project on taxing and pricing transport fuels. An earlier version of this paper was issued in Newbery (1985). 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. I am also indebted to A. Deaton, G. Dixon, M. Katz, P. Swan, and two referees for their comments and mathematical advice. 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 organizations. The finding, interpretations, and conclusions are the results of research supported by the Bank; they do not necessarily reflect official policy of the Bank.

degree of capacity. This qualitative result is attractive to economists, who can now argue that if traffic engineers have indeed chosen the correct capacity (perhaps on average) *and* if there are constant returns to scale, then the (average) congestion charge should be the average cost of capacity. This in turn has shifted attention from the calculation of the congestion function (which, as Dewees (1979) has shown, is remarkably difficult) to the question of whether there are constant, increasing, or decreasing returns to capacity expansion (see Mohring (1976); Keeler and Small (1977)). Once the degree of returns to scale has been identified, then the relationship between the congestion charges and the average capacity cost can be calculated (if necessary, taking account of distortions elsewhere in the economy, as reviewed in Winston (1985, pp. 79–81)).

Useful though these relationships are, they require optimal capacity expansion and need modification if there are indivisibilities in road capacity (going from two lanes to three, for example). Starkie (1982) argues that such indivisibilities are of little practical importance. But even if roads can be *designed* for a smoothly varying range of capacities, if traffic volumes grow, the process of *adjustment* of capacity is likely to be intermittent and lumpy, if only because of the set-up and disruption costs.

During the 1960's the evidence available suggested that congestion (of various types (see Vickery (1969)) and including accidents) was the only interactive externality affecting road users. (Noise and pollution primarily affect non-road users.) It was known that vehicle use damaged the pavement and it was therefore argued that the consequent maintenance costs should be charged to vehicles in proportion to their damaging effect. The sums involved are large. In 1979 the U.S. allocated $37.5 billion for highway purposes, and collected about $22.8 billion in taxes and charges on highway users. (In the U.K. and Europe the sums collected exceed those spent on highways by a factor of two or more.) Consequently a great deal of effort has been devised to determining the damaging power of different vehicles, and devising methods of allocating and charging for these costs. The Highway Research Board (1962) reports the results of the AASTHO road test in which variously laden vehicles were driven millions of miles on a wide range of surfaces to determine their rate of deterioration in an accelerated wear test—an experiment which it is estimated would cost over $300 million at 1980 prices to replicate. This test provided good evidence that the damaging power of a vehicle axle on paved roads was approximately proportionate to the fourth power of its loading, and the Highway Research Board proposed measuring the damaging power of vehicles by the number of Equivalent Standard Axle Loads (ESALs), where 1 ESAL is taken as a load of 18000 lbs. or 80 kilo Newtons. On this scale, trucks may vary from less than 0.1 ESALs to more than 50 ESALs for (illegally) overloaded vehicles. Typically, the average for the truck fleet is about 2–3, with a typical legal limit of 5. Only 2 percent of heavy vehicles in the U.K. on the main traffic road (the M1) had damaging factors greater than 12 in 1964 (Jones, 1977). Private cars have totally insignificant damaging factors.

Given the sums involved, and the wide disparity in damaging costs of different road users, the issue of cost allocation and recovery becomes economically and politically important, 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 approach the end of its design life, and other countries, especially developing countries, are now attempting to decide how best to charge for road use.

By and large the study of the allocation of maintenance costs has not attracted so much mainstream economic interest as the study of congestion costs, perhaps because it appears to be relatively simple. Even on the simple approach of just allocating costs in proportion to vehicle damage, though, the degree of distortion in the current U.S. system of road user charges may be large, as Small and Winston (1984) argue. However, the past decade has produced new evidence on the consequence of road damage which suggests that road damage creates a new type of externality which is quantitatively very significant, and qualitatively quite different from congestion. It is the aim of the present paper to develop the theory to deal with this kind of externality, characterize the optimal pricing rules, and in turn relate them to current methods of cost allocation.

1. ROAD DAMAGE EXTERNALITIES

When a vehicle travels on a road it causes some damage to the surface and advances the date at which the road needs to be repaired. 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 a typically period of 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.

The damage vehicles do to roads depends on the characteristics of the vehicle and the type of road (paved, unpaved) 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, which, for paved roads, is measured by the Equivalent Standard Axle Load. 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). Consequently, when a vehicle damages the road surface and increases its roughness, it thereby increases the vehicle operating costs of subsequent vehicles, and creates a *road damage externality*. On well-trafficked interurban roads these vehicle operating costs are between 10 and 100 times as large as maintenance costs, while the average *increase* in vehicle operating costs are comparable to the direct cost of repairing the damage, and therefore potentially of the first importance (Newbery, 1986a). At first sight it would seem to be very misleading to ignore these externalities when calculating and allocating highway costs.

If every link in the road network had a tollgate, then vehicles could be charged at a vehicle and road specific rate (per ESAL km, at different rates depending on the characteristics of the road, or the road *type*, measured by strength, traffic flow, maintenance strategy). In practice, of course, such fine tuning of the system of road user charges is impractical with current technology, and the best that is likely is a vehicle-specific distance-related charge, which will have to be equated to some average of the road specific charges. Even if the authorities could charge for each road link separately, the marginal social cost (MSC) (and the appropriate road user charge) would vary with the age and state of the road, and hence would need continuous adjustment. With the more realistic vehicle-specific distance charge, the average would have to be taken over roads of different types as well as different ages.

The best strategy for calculating the appropriate road user charge is to calculate the MSC of an ESAL km on each type of road (distinguishing as finely as possible given the available data from the Road Traffic Census and information about the road network—lengths, width, strength, maintenance strategy) and then to calculate the optimum uniform charge as a weighted average of these MSC's, as shown below. The problem then reduces to finding the MSC of an ESAL km on a road of a particular type, but varying age. If the network is in static equilibrium with constant traffic flow on each link, the natural assumption is that roads of each type have a uniform age distribution (measured equivalently by calendar time or cumulative traffic flow). The appropriate measure of the MSC of an extra vehicle on a particular type of road is then the expected MSC averaged over the age distribution of the roads. (This can either be thought of as the MSC of a vehicle km on a road of unknown age, or as the average MSC per km of an extra vehicle which travels over a representative fraction of roads of this type, and hence experiences roads of all ages.)

The main result of this paper is to show that there is an important case in which the road damage externality is identically zero, and in which the efficient road user charge, equal to the marginal social cost of highway use, is exactly equal to the average maintenance cost, allocated in proportion to the road damaging power of the vehicles. This holds for any damage sensitive maintenance strategy, and not just for optimal maintenance strategies. It holds despite the long elapse of time between major maintenance (typically 10–20 years), and despite the presence of discounting, and it holds for a whole class of road damage functions and vehicle operating cost functions. Since these functions are difficult to estimate to a desirable level of precision, it offers an extremely useful short-cut for the calculation of road user charges.

Nevertheless, for the theorem to hold two important conditions need to be satisfied. The first is that traffic is the sole source of road damage, and the second is that traffic flow is constant over time. Until quite recently, the available evidence on the road damage relationship was largely limited to the accelerated wear tests reported in Highway Research Board (1962), and in these tests it did appear that all damage was attributable to vehicles. Road engineers remained concerned that weather also affected the road surface and its rate of deterioration, but until recently this effect could not be disentangled from the influence of traffic. The two are obviously collinear, the time periods are long, and roads are designed to the expected traffic and conditions, and hence all tend to behave similarly, despite variations in the separate factors. Statistical identification is therefore very difficult.

Recent research, supported by the World Bank and the Brazilian government, and reported in Paterson (1984), has gone a considerable way towards disentangling the two effects, and the best current model of road damage makes the change in roughness a function of both the passage of traffic and the elapse of time (which is a proxy for the effect of weathering). Section 3 below explores the consequence of this recent development and shows that it lowers the fraction of maintenance costs attributable to traffic, since some is now attributable to the weather. The road damage externality is no longer identically zero, but it is quantitatively insignificant.

Finally, the common experience of road networks everywhere is that traffic volumes are growing. Of course, beyond some point, the traffic *per lane* reaches optimal capacity (defined jointly by congestion effects and the magnitude of the road damage externality) and thereafter growth is accommodated by increases in capacity. As the model considers maintenance cost per unit of road width (or per lane), if capacity could be steadily adjusted, then traffic growth could be ignored. However, if capacity is not so adjusted, then traffic growth becomes important, and is analyzed in Section 4. Even if all damage is attributable to traffic (as on unpaved roads with a short maintenance cycle, for example), traffic growth prevents the equality between efficient charges and average maintenance costs. Depending on the maintenance strategy adopted, the marginal social cost may be above or below the average cost. Again, however, the road damage externality is negligible.

It is therefore important to recognize the strengths and limitations of the Fundamental Theorem of Section 2. First, the sense in which the road damage externality is zero is *on average*, taken over roads of uniformly varying ages. Since it is currently impractical to imagine charging vehicles differently depending on the age of the road surface they use, any charge will perforce have to be an average, and the uniform age distribution is consistent with the assumption of static flow. Some readers have claimed that it is then not surprising that the average of a set of short run marginal costs should be the average cost, but this is not at all obvious. First, the claim is that the average externality cost is *zero*. Second, the costs are discounted, so arguments which might look plausible for undiscounted flows are not obviously true for discounted flows. This is seen most clearly when traffic growth affects the time pattern of costs, in which case it ceases to be true that the average externality cost is zero. Finally, the theorem requires that all damage be attributable to vehicles, and if weather also damages roads, the average road damage externality is no longer identically zero (though it is approximately zero). So simple arguments are most suspect, and the theorem remains remarkable.

The second point to stress is that it assumes a particular type of maintenance strategy, in which the Highway Authority decides to resurface any particular road when it reaches a predetermined level of roughness. This appears to describe actual maintenance behavior in developed countries well, and is to be contrasted with the rule that ordains resurfacing of roads, at fixed intervals, or in a fixed sequence. The optimal moment to resurface will involve balancing the extra costs of advancing the date of resurfacing with the lowering of vehicle operating costs, and will depend sensitively on the nature and magnitude of the road damage externality. However, the theorem does not assume that maintenance is optimally timed, only that the maintenance criterion is unchanged over time. (Optimal maintenance is a specific type of consistent maintenance strategy for which the theorem continues to apply.) At the end of Section 2 the relationship between the two maintenance strategies (those triggered by the state of the road and those by the elapse of time) is briefly discussed.

The practical importance of those results is substantial for two reasons. First, it is statistically difficult to identify the various road damage relationships with any precision, and it is reassuring to note that the design of the optimal road user charge is insensitive to the exact form of the relationship (and, in special cases, quite independent of it). Second, the data requirements to calculate an approximately optimal road user charge are minimal—road maintenance costs per ESAL km under a consistent maintenance policy. The figures can be refined with data on traffic growth rates, traffic census data, and estimates of road strengths, surface lives, etc., but is likely to be reasonably well approximated by the minimal data of average maintenance costs, corrected for the effect of weathering, which will depend mainly on the type of climate.

2. 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 t will be

$$(1) \qquad Y(t) = X + NE(t-z), \qquad t \geqslant z.$$

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

$$(2) \qquad R = R\{Y(t)\}.$$

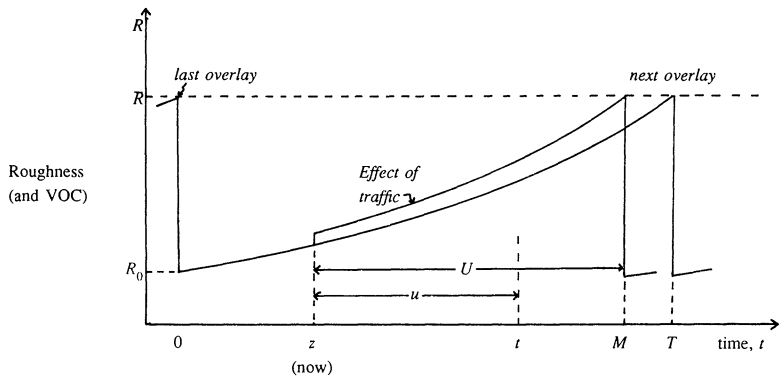

FIGURE 1.—The effect of an extra standard axle on roughness and costs.

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

$$(3) \qquad v = v(R).$$

The maintenance strategy is to restore the road to roughness R_0 whenever the roughness reaches a predetermined critical level $\bar{R}$, at a cost C.[2] 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).

The present discounted value of all future vehicle operating and road maintenance costs on a road which was restored z years ago and has since experienced X ESALs, can be written as

$$(4) \qquad F(X,z) = e^{-r(M-z)}\frac{C}{1-e^{-rT}} + N\int_z^M v[R\{Y(t)\}]e^{-r(t-z)}\,dt + e^{-r(M-z)}\frac{N\int_0^T ve^{-ru}\,du}{1-e^{-rT}}.$$

In equation (4) the life of the present road, M, is distinguished from the design life of the road after overlay, T. These will coincide if traffic was as predicted or if $X = NEz$. The first term is the (present discounted) 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 date of the next overlay, M, is implicitly defined by the total number of ESALs required to cause the road surface to reach its critical level of roughness,

[2] It appears to be the case that resurfaced roads can be restored to essentially the condition of a new road (though with traffic growth they are likely to be strengthened: see Section 3 below). The maintenance strategy can thus be described as an $[R_0, \bar{R}]$ policy, analogously to the (S, s) inventory policy. Given the large fixed costs involved in overlaying it is easy to show that an $[R_0, \bar{R}]$ policy is optimal for static traffic flow. There are further analogies in that both policies give rise to a modulo arithmetic with powerful uniformity implications that give the Fundamental Theorem as a direct implication. See Caplin (1985) for the $[S, s]$ counterpart.

$\bar{R}$:

$$(5) \qquad \bar{R} = R\{\bar{Y}\}, \quad \bar{Y} = X + NE(M - z), \quad M = z + (\bar{Y} - X)/(NE).$$

(The life of the surface is thus measured by the constant number of ESAL's, $\bar{Y}$, which it can sustain before reaching the critical level of roughness.)

The MSC of an extra ESAL at date z can be found by varying X in (4), which, from (1) and (2) will affect $R(t)$ and M, and the MSC will be equal to $\partial F/\partial X$. The *expected* MSC can be found by averaging over roads of all ages, z, from 0 to T years:

$$(6) \qquad \frac{\bar{\partial}\bar{F}}{\partial X} = \frac{1}{T}\int_0^T \frac{\partial F}{\partial X}(X, z)\, dz.$$

The Fundamental Theorem claims that $\bar{\partial}\bar{F}/\partial X = C/(NET)$, that is, that the expected *marginal* social cost is exactly equal to the *average* maintenance cost, and the road damage externality is exactly zero. There are (at least) three proof strategies available: (i) Evaluate (6) directly by differentiating (4), and demonstrating that the road damage externality terms cancel, as in Newbery (1985). (ii) Transform variables to simplify the problem. (iii) Calculate the optimal road user charge for roads of uniform age distribution directly.

The first method is the only practical one for the more complicated problems of Sections 3 and 4, but it does not explain *why* the externality terms cancel, as the second method does. The last method (which, strictly speaking, proves the more embracing result that it is also optimal to set the flat charge equal to the expected MSC) provides a remarkably simple insight into the theorem. I shall therefore provide the second and third methods of proof.

The first point to note in equation (4) is that M, z, and X are linked by equation (5), and that if $M - z$ is replaced by U, the time before the next overlay, then the present discounted social cost, (4), can be rewritten as

$$(7) \qquad G(X) = e^{-rU}\frac{C}{1 - e^{-rT}} + N\int_0^U v\{R(X + NEu)\}e^{-ru}\, du + e^{-rU}\frac{N\int_0^T ve^{-ru}\, du}{1 - e^{-rT}},$$

where the variable of integration has been changed from $t - z$ to u. U is a function of X from (5), for $U = (\bar{Y} - X)/(NE)$, and so G is a function solely of X. The age of a road can be described as well by U as by z, and hence

$$(8) \qquad \frac{\bar{\partial}\bar{F}}{\partial X} = \frac{1}{T}\int_T^0 \frac{dG}{dX}\, dU.$$

The effect of increasing X is exactly the same as increasing the age of the road, and the effect of averaging is thus simply to undo the effect of the integration. This can be seen formally as follows. Since U is a function of X, G could as readily be differentiated with respect to U, for

$$(9) \qquad \frac{dG}{dX}=\frac{dG}{dU}\cdot\frac{dU}{dX}=-\frac{1}{NE}\frac{dG}{dU}.$$

Hence equation (8) can be written

$$\begin{aligned}\frac{\partial \bar{F}}{\partial X}&=\frac{-1}{NET}\int_0^T\frac{dG}{dU}\,dU=\frac{1}{NET}\left[G(\bar{Y}-NEU)\right]_T^0\\&=\frac{1}{NET}\left[[1-e^{-rT}]\frac{C}{1-e^{-rT}}-N\int_0^T ve^{-ru}\,du\right.\\&\qquad\qquad\left.+[1-e^{-rT}]\frac{N\int_0^T ve^{-ru}\,du}{1-e^{-rT}}\right],\\&=C/(NET).\end{aligned}$$

This is a remarkable result, for it is independent of the exact form of the road damage relationship, $R\{Y(t)\}$, 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 $Y(t)$, the v depend only on R, that the maintenance strategy remain constant, and that there be no traffic growth.

The other method of proving the Fundamental Theorem (or its extension to road charging) is to find the optimal road user charge directly. Suppose that the road network consists of a single road of length 1 km,[3] and uniform age distribution between 0 and T years, where T years is the initial period between overlays. Suppose also, without loss of generality, that at date $t=0$, the age of the road x km from one end is xT years. Each road user i $(i=1,2,\ldots,I)$ demands q_i trips per annum, and imposes a damaging effect equivalent to e_i ESALs, so that the total annual traffic is $Q=\Sigma q_i e_i$ ESALs. Until date $t=0$, total traffic has been constant at $NE=\bar{Q}=\Sigma\bar{q}_i e_i$, so that NET is the carrying capacity of the road before it requires overlay. At date t and distance x the age of the road surface since last overlay, measured by cumulative ESALs, is

$$X(x,t)=NETx+\sum_i q_i e_i t \pmod{NET},$$

as shown in Figure 2 below. The crucial point to note is that the age distribution of the road is completely independent of the current volume of traffic, provided only that the road is overlaid whenever it reaches the same level of roughness, $\bar{R}$, corresponding to NET cumulative ESALs, and is restored to the same initial roughness, R_0. The vehicle operating cost (VOC) of traversing the road is

$$\bar{v}_i=\int_0^1 v_i[R\{X(x,t)\}]\,dx$$

and since the road has length 1 km, this is the average VOC per km. If vehicle i is charged a fee f_i per km, the total private cost will be $p_i=\bar{v}_i+f_i$ per km, and this

[3] Normalizing the length to 1 km makes it unnecessary to distinguish between total costs and costs per km, but is otherwise inessential.

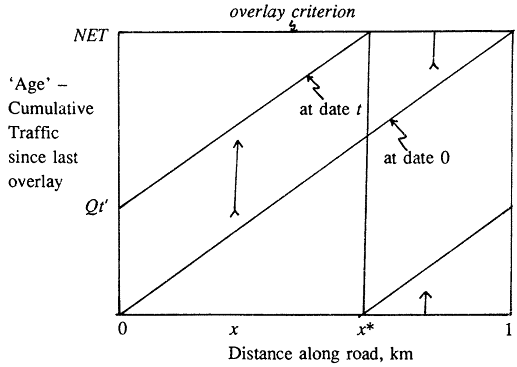

FIGURE 2.—'Age' of road as a function of distance and time.

will determine i's demand for trips, q_i, as follows. If $V_i(p_i, b_i)$ is the indirect utility function, where b_i is lump sum income, then $q_i = -\partial V_i/\partial p_i/\beta_i$, where $\beta_i = \partial V_i/\partial b_i$ is the private marginal utility of income.

The annual cost of operating the road system is the cost of restoring sections of the road which reach the critical level of roughness each year. The number of km reaching this critical state each year is $\Sigma q_i e_i/(NET)$, and if the cost of overlay is C per km then the total annual cost is $C\Sigma q_i e_i/(NET)$ (also per km). The surplus of fee income over maintenance costs to be redistributed to vehicle users is $\Sigma f_i q_i - C\Sigma q_i e_i/(NET)$, and vehicle owner i receives a share α_i, $\Sigma \alpha_i = 1$. The Government's problem is to choose the road user charges, f_i, to achieve efficiency:

$$\max_{f_i} W\left[V_i\left[\bar{v}_i + f_i, \alpha_i\left(\sum_j f_j q_j - C\sum_j q_j e_j/NET\right)\right]\right] \tag{10}$$

when $(\partial W/\partial V_i)\cdot(\partial V_i/\partial b_i) = 1$ (i.e. the social welfare function respects the current income distribution). The key point to note is that although traffic, Q, depends on f_i, and roughness, R, depends on Q, $\bar{v}_i$ is independent of f_j for all i, j, because the age distribution of the road is independent of traffic. Formally, let $Qt \pmod{NET} = Qt'$, $x^* \equiv 1 - Qt'/(NET)$ (both shown on Figure 2).

$$\bar{v}_i = \int_0^{x^*} v_i[R\{Qt' + NETx\}]\,dx + \int_{x^*}^1 v_i[R\{Qt' + NETx - NET\}]\,dx,$$

so

$$\frac{d\bar{v}_i}{dQ} = \frac{dx^*}{dQ}\left[v_i(\bar{R}) - v_i(R_0)\right] + \int_0^1 \frac{dv_i}{dR}\frac{dR}{dX}\frac{dX}{dQ}\,dx.$$

But $dx^*/dQ = -t'/(NET)$, and $dX/dQ = t' = (t'/NET)\,dX/dx$, so

$$\frac{d\bar{v}_i}{dQ} = \frac{-t'}{NET}\left[v_i(R)\right]_{R_0}^{\bar{R}} + \frac{t'}{NET}\int_0^1 \frac{dv_i}{dR}\frac{dR}{dX}\frac{dX}{dx}\,dx \equiv 0.$$

The first order condition for equation (10) is thus

$$(11)\qquad \frac{\partial W}{\partial V_i}\frac{\partial V_i}{\partial b_i}\left[-q_i+\alpha_i q_i+\alpha_i\frac{\partial q_i}{\partial f_i}\left[f_i-\frac{Ce_i}{NET}\right]\right]+\sum_{j\neq i}\frac{\partial W}{\partial V_j}\frac{\partial V_j}{\partial b_j}\alpha_j q_i$$

$$=q_i\left(1-\sum_j\alpha_j\right)+\alpha_i\frac{\partial q_i}{\partial f_i}\left(f_i-\frac{Ce_i}{NET}\right)=0.$$

Since $\Sigma\alpha_j=1$ and $\partial q_i/\partial f_i<0$, the optimal fee is $f_i=Ce_i/(NET)$, or the charge per ESAL km (f_i/e_i) is exactly equal to the average maintenance cost per ESAL km, $C/(NET)$.

This last method can be directly extended to address the question of choosing the best vehicle-specific tax for a network of varying types of roads, each of 1 km length, $k=1,\ldots,K$. Let q_{ik} be the number of trips by vehicle i on road segment k, $\bar{v}_{ik}$ be the VOC of vehicle i on road segment k, $c_k=C_k/(N_k E_k T_k)$ be the *average* maintenance cost per ESAL km on road segment k; then the indirect utility function can be written $V_i\ [\bar{\boldsymbol{v}}_i+f_i\boldsymbol{u},\ \{\alpha_i\Sigma_k(\Sigma_j f_j q_{jk}-c_k\Sigma q_{jk}e_j)\}]$, where $\bar{\boldsymbol{v}}_i$ is the vector of VOC across roads, and $\boldsymbol{u}$ is a k-vector of 1's. The same analysis then gives the optimal road user charge as

$$(12)\qquad f_i=\frac{e_i\Sigma_k c_k\,\partial q_{ik}/\partial f_i}{\Sigma_k\partial q_{ik}/\partial f_i}.$$

If the average maintenance cost per ESAL km is high on lightly trafficked low strength roads (economies of scale in strengthening roads) and if these roads are allowed (sensibly) to have a higher average roughness, then c_k and v_{ik} will be positively correlated. If the demand for trips is an isoelastic function of the total private transport cost, with the same demand elasticity across roads, then the optimal fee per ESAL km, f_i/e_i, will be somewhat *less* than the average maintenance charge, since c_k and $-\partial q_{ik}/\partial f_i$ will be negatively correlated. If on the other hand, demand is linear with uniform slope coefficient: $q_{ik}=A_{ik}-B_i(\bar{v}_{ik}+f_i)$, then the optimal fee will be the (unweighted) average cost. We can summarize the results as follows.

PROPOSITION 1: *If the age distribution of roads of a given type is constant, and the traffic flow is constant, and all road damage is attributable to traffic, then the average road damage cost of a vehicle is identically equal to the average maintenance cost allocated in proportion to its number of equivalent standard axles. The road damage externality is zero.*

PROPOSITION 2: *With the assumptions of Proposition* 1, *the optimal flat charge per ESAL km will exactly recover the road maintenance costs of the network if the demand for trips is linear with a slope coefficient uniform across vehicles. If not, the optimal damage charge will be a weighted average of the road maintenance costs.*

2.1. *The Importance of the Maintenance Strategy*

There is good empirical evidence[4] 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 strategy 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, $R(T+\varepsilon)$, and hence thereafter. Road maintenance costs are independent of vehicle traffic and hence the derivative of 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 the marginal social cost and the average maintenance cost has been broken. However, if the frequency of maintenance is *optimally* chosen, this link is restored, for the following reason. The marginal social cost of an extra ESAL is found by totally differentiating F in equation (4). It can also be written, replacing M by T, as

$$\frac{dF}{dX} = \left.\frac{\partial F}{\partial X}\right|_{T\,\text{const}} + \frac{\partial F}{\partial T}\frac{\partial T}{\partial X}.$$

If the maintenance frequency T is optimally chosen, $\partial F/\partial T = 0$, and the marginal social cost will then be equal to the marginal externality cost. Thus the efficient charge will again coincide with the average cost when the maintenance frequency is fixed, independent of the state of the road, provided this frequency is optimally chosen. (A formal proof, which provides yet another proof of the Fundamental Theorem, is available in Newbery, 1986b.) If the frequency is not optimal, then the efficient charge may be above or below the average cost, 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 charge and average cost, 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.

This is perhaps the appropriate place to mention that on multi-lane roads, standard practice is to build all lanes to the same strength, and overlay them at the same moment, even though heavy vehicles are concentrated on the outer lanes. This does not affect the Fundamental Theorem, but it does alter the

[4] Thus the U.S. Federal Highway Authority publishes criteria ("minimum tolerable standards") for various repairs and reconstruction in the Technical Manual of the U.S. Highway Performance Monitoring System (USFHWA, 1983) and each year determines the required expenditure given the state of the highways—e.g. USFHWA (1984). The U.K. has a similar system.

calculation of the maintenance cost, which is now that of the whole *road* allocated to total ESALs. The number of ESALs required to precipitate overlay will be determined by the time taken for the number of ESALs in the outer *lane* to reach the critical number. See Newbery (1986b) for details and empirical implications.

3. EXTENSIONS TO THE FUNDAMENTAL THEOREM

The road damage relationship relating 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 the late 1950's. 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 ESALs since construction (or reconstruction), $Y(t)$, and not separately on the time which has elapsed, t, as in (2).

The Arizona model is at the other polar extreme in supposing that all damage is caused by time and weather and none by traffic. The Arizona model in effect presupposes the 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 may be useful for predicting maintenance intervals, it is clearly an unsatisfactory feature of the model for calculating road user charges. Nevertheless, 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), but as

$$R(t) = R\{Y(t), t\}. \tag{13}$$

Whereas the earlier analysis of the data collected in the Brazilian study estimated a relationship of the form of (2), 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:

$$R(t) = e^{mt}[R_0 + kY(t)], \tag{14}$$

where k depends on the road strength (measured by the modified structural number, SNC), and m is a constant which depends on the climate.[5]

It is clear that the Fundamental Theorem will no longer hold for damage relationships like (13) or (14), because the state of the road can no longer be

[5] The form actually estimated was $\Delta R = K\Delta Y + mR\Delta t$, which is Markovian, but this cannot be integrated to give a path-independent relation for R. However, a slight modification to $dR = ke^{mt}dY + mRdt$ can be integrated to give (14), and behaves very similarly to the estimated relationship. See Newbery (1986d). The parameter k is $9450/(1 + SNC)^{6.45}$ when Y is measured in millions of ESALs.

unambiguously measured by cumulative ESALs. Looking at Figure 2, changes in the volume of traffic will no longer leave the roughness distribution of the road unchanged, and will therefore have an effect on subsequent traffic—the damage externality will no longer vanish. The two practical questions are, first, what is the sign of the road damage externality, and second, quantitatively how important is the externality. The first question can be answered analytically for equation (14), but the second question requires numerical calculations. The answers are that the road damage externality is *positive* for the case of zero traffic growth, but quantitatively small.

The proof that the externality is positive and likely to be small can be constructed by extending the geometrical proof of the Fundamental Theorem. Consider a road of uniform age distribution in steady state with traffic NE ESALs per year, which is overlaid whenever it reaches $\overline{R}$, restored to R_0. Suppose the overlay cycle is T years, where $T = T(N)$ solves:

$$\overline{R} = e^{mT(N)}\{R_0 + kNET(N)\}. \tag{15}$$

The fraction of road, z, of roughness less than R, can be found as follows. If $R(t) = R$ for some age t, then $z = t/T$ because all roads younger than t will be less rough than $R(t)$. Thus z solves

$$R = R(t) = R(zT) = e^{mzT}(R_0 + kNETz). \tag{16}$$

$R(zT)$ can thus be treated as a function of z for given $T = T(N)$, i.e. for given traffic, and the shape of the graph $R(zT)$ will vary with N. As $N \to 0$, (15) and (16) give

$$R(zT) \to e^{mzT}R_0 = \overline{R}^z R_0^{1-z} \tag{17}$$

which is graphed in Figure 3 as the case of no traffic. As $N \to \infty$, $T \to 0$ (for example, in an accelerated wear test taking place over a very short period of

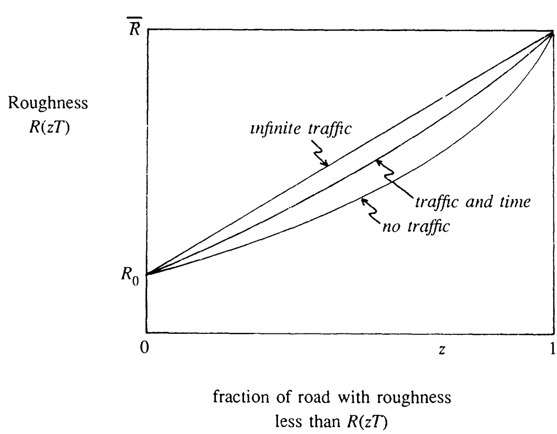

FIGURE 3.—Distribution of roughness as a function of traffic flow.

time), then (15) and (16) give

$$(18)\qquad R(zT) \to R_0(1-z) + z\bar{R}.$$

(This corresponds to the earlier case in which weathering has no effect on wear, as in the Fundamental Theorem.) The effect of holding the maintenance criteria constant but varying traffic is thus to alter the shape of the roughness profile as shown in Figure 3. As traffic increases, the roughness profile becomes less convex and more nearly uniform (linear), thereby raising the average roughness and also raising the average vehicle operating cost. As extra traffic raises operating costs it creates a positive externality. That it is likely to be small can be seen intuitively from the very small increase in average vehicle operating costs represented by a very large increase in traffic flows, which has only a slight effect on the curvature of the roughness profile, but it is easy to check its magnitude by substituting values into the analytical formula presented in the Appendix.[6]

The expression for the average MSC can be found by differentiating (4), replacing (2) by (13) or (14):

$$(19)\qquad \frac{\partial F}{\partial X}(X,z) = -\frac{re^{-r(M-z)}}{1-e^{-rT}}C\frac{\partial M}{\partial X} + \frac{\partial D(X,z)}{\partial X},$$

where $\partial D(X,z)/\partial X$ is the road damage externality (and is the derivative of the last two terms in (4)). The general expression for $\partial M/\partial X$ is found by totally differentiating equation (13) at the date of overlay, $t = M(=T)$, where $\bar{R} = R\{X + NE(M-z), M\}$. This gives

$$(20)\qquad \frac{-\partial M}{\partial X} = \frac{1}{NE + R_t/R_Y|_T} \equiv \frac{\mu}{NE},$$

where μ is the fraction of maintenance costs attributable to traffic. This interpretation of μ can be seen by calculating the expected MSC as before:

$$(21)\qquad \frac{\partial \bar{F}}{\partial X} = \frac{1}{T}\int_0^T \frac{\partial F}{\partial X}(X,z)\,dz = \frac{\mu C}{NET} + \frac{1}{T}\int_0^T \frac{\partial D(X,z)}{\partial X}\,dz.$$

The second term in (21) is the road damage externality, which has been argued to be positive, and this can be confirmed using the formulas given in the Appendix, which also derives the expression for the allocable fraction, μ, in terms of readily observed parameters R_0, $\bar{R}$, m, T (and not the harder to observe road strength):

$$(22)\qquad \mu = \left[1 + \frac{mT}{1 - R_0 e^{mT}/\bar{R}}\right]^{-1},$$

which lies in the range (0, 1), with $\mu \to 0$ as $N \to 0$; $\mu \to 1$ as $N \to \infty$. Since roads

[6] The original road damage function estimated from the same data and presented in Newbery (1985) had $R(t) = e^{mt}[R_0^{1-\alpha} + (1-\alpha)kY(t)]^{1/1-\alpha}$, $\alpha > 1$, for which similar operations to (17) and (18) *reverse* the positions of the two extreme profiles, and lead to a negative externality. This qualitative ambiguity is also indicative of its quantitative unimportance, since both equations fit the same (large) data set almost equally well (and give almost the same estimates for the road damage costs).

TABLE I

SENSITIVITY OF ALLOCABLE FRACTION OF ROAD DAMAGE

Climate	$\bar{R}/R_0$	T years	Allocable fraction μ
Humid Subtropical	4	20	0.54
or nonfreezing	$\underline{3}$	20	0.47
temperate	3	20	0.47
$m = 0.025$	3	$\underline{15}$	0.58
	$\underline{2}$	$\underline{10}$	0.59
Arid subtropical	2	$\underline{15}$	0.74
	2	$\underline{25}$	0.59
$m = 0.01$	$\underline{3}$	$\underline{15}$	0.80
	$\underline{4}$	$\underline{20}$	0.78
Freezing temperate	2	$\underline{10}$	0.26
	$\underline{4}$	10	0.54
$m = 0.05$	4	$\underline{15}$	0.38

Note: Underlines indicate parameter changes.

are designed to a strength appropriate to the expected traffic, the observed range of μ is usually quite narrow, and typical numerical values for the parameters and for μ are shown in Table I. Note that $\bar{R}/R_0$ (the ratio of final to initial roughness) and mT will be positively correlated. The coefficient m depends on the prevailing climate, and is higher for Brazil (which is wetter) at 0.0245 than for Tunisia (which is drier, and for which $m = 0.011$).

The main point to emerge from the calculation is that the fraction μ is quite large and reasonably stable for climates and maintenance strategies likely to be encountered in developing countries such as Brazil and Tunisia but lower and more sensitive to maintenance criteria in developed countries with a freezing climate.

4. 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 (4) 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. If there were no strengthening, and all wear were attributable to traffic, then the lifetime of roads would steadily shorten as traffic grew, but it would still be true that provided roads of each type initially had a uniform age distribution, then the roughness distribution would be unaltered by traffic growth (as should be clear from Figure 2). In this case the Fundamental Theorem will continue to hold, and the efficient charge is again the average maintenance cost.

This is, however, unrealistic, as Highway Departments will take the opportunity of the overlay to strengthen the road to cope with the increased traffic. In addition, weathering is important and it is necessary to check for possible interactions between traffic growth and weathering. The simplest (and empirically plausible) maintenance strategy would be to increase the strength to preserve its design life, T, and the initial and final roughness levels, R_0 and $\bar{R}$, over the next overlay cycle.

The next problem to be faced is how to average the road damage costs. Two options suggest themselves. The first is to average over all roads of the same initial strength and initial traffic but varying ages (and hence varying levels of current traffic). The problem with this approach is that the available data typically classify roads by current traffic, not traffic at date of last overlay, so this first option is difficult to implement. The alternative is to average over roads of the same current traffic, N_0, assuming a uniform age distribution. If the present age of the road is z, then its traffic at last overlay was $N_0 e^{-gz}$, and if it had been designed to last T years, its original strength would have been $S(z)$, whose precise formula is given in the Appendix. If this road will be next overlaid at date $M(=T)$, as before, and as in Figure 1, and if the present discounted cost of the next and all subsequent overlays at the overlay date is $C(z)$ (for a road of current age z, and traffic N_0), then the total social cost corresponding to (4) now becomes

$$F(X,z) = e^{-r(M-z)}C(z) + N_0 \int_z^M e^{g(t-z)} v\{R(t)\} e^{-r(t-z)}\,dt + N_0 e^{(g-r)(M-z)} \left[\frac{\int_0^T v\{R(t)\} e^{-(r-g)u}\,du}{1-e^{-(r-g)T}} \right], \tag{23}$$

where N_0 is the current traffic flow. The effect of an extra ESAL now (at date $t = z$) is to raise social costs, and the marginal social cost (evaluated at $M = T$) is

$$\frac{\partial F}{\partial X} = -\frac{\partial M}{\partial X}\left\{ re^{-r(T-z)}C(z) + e^{-(r-g)(T-z)}N_0(v^* - \bar{v}) \right\} + N_0 e^{(r-g)z} \int_z^M \frac{dv}{dR}\frac{\partial R}{\partial X} e^{-(r-g)t}\,dt, \tag{24}$$

where $C(z)$ is given in the Appendix and

$$v^* = \int_0^T v e^{-(r-g)u}\,du \Big/ \int_0^T e^{-(r-g)u}\,du, \qquad \bar{v} = (\bar{R}). \tag{25}$$

The time averages can be analytically derived and are given in the Appendix.

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 percent 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, representative results of which are reported in Table II. The model takes the average daily traffic (*ADT* of col. 1; total, both directions), the maintenance strategy, R_0, $\bar{R}$, the design life T, and the damaging power of the average vehicle, E (col. 5, in ESALs per 100 vehicles), which span

TABLE II

SENSITIVITY ANALYSIS

	ADT veh/day	R_0 IRI units	$\overline{R}$ IRI units	*T* Years	*E* ESALs/ 100 veh	*r*	*g* % p.a.	*m* % p.a.	Road Damage Extern U.S. ¢/ESSAL km	Marg. Soc. Cost U.S. ¢/ESSAL km	MSC/AC %
A	500	3.5	6.0	15	38	12	0	1.11	0.03	2.58	66
B				20					0.03	1.63	56
C			6.5	20					0.03	1.69	61
D				20	30				0.03	1.97	56
E				20	38	8			0.03	1.65	57
F		3.31	6.6	15		12		2.45	0.02	1.41	36
G		3.5	6.0	20		8	5	0.0	0.03	1.57	54
H				20		12	10	1.11	0.03	1.58	50
I				20		8	5	0.0	0.14	2.47	85
J	1000	3.0	5.5	15	38	8	0	1.11	0.04	1.70	70
K				20			0		0.04	1.12	61
L				20			5		0.05	1.05	58
M				20			5	2.45	−0.03	0.35	19
N		3.0	6.0	15			5	2.45	0.01	1.01	44
O		3.0	5.5	20			5	0.0	0.14	1.59	87
P	1500	3.0	5.5	20	38	8	5	1.11	0.05	0.85	58
Q	15,000	1.8	3.7	15	38	8	0	1.11	0.03	0.31	80
R							5	1.11	0.06	0.32	82
S							5	2.45	0.02	0.19	50
T		1.23	3.7				5	5.00	0.01	0.13	32
U		1.23	3.7				0	5.00	0.03	0.14	36
V		1.8	3.7				5	0.0	0.09	0.42	107

Notes: (1) Blanks indicate previous values unchanged, underlines indicate change. (2) Overlay costs range from $20,000/km for weak roads to $60,000/km for strong roads. (3) MSC, AC are marginal social cost and average maintenance cost/ESAL km. (4) *ADT*: Average Daily Traffic; other parameters defined in text.

Source: Computations by author.

the range encountered in Tunisia. The rate of interest, r, the rate of traffic growth, g, and the rate of weathering, m, were then varied systematically, and for each set of parameters the strength of the road consistent with these assumptions was calculated. The final three columns of Table II show the Road Damage Externality and the (Average) Marginal Social Cost (i.e. $\overline{\partial F/\partial X}$) in U.S. ¢/ESAL km at 1983 prices for Tunisian overlay and vehicle operating costs (which are not very different from those elsewhere). The final column gives the ratio of the MSC to the average overlay cost as a percentage.

The rate of weathering, m, was varied across an extreme range (i.e. from zero, to 5 percent p.a., consistent with freezing climates). Across such a climate range the design life of the road would also vary (inversely). Within a given climate zone the variation would of course be very much less.

The main conclusion to draw from Table II is that the vehicle externality cost is quantitatively small and remarkably constant across alternative specifications, and is typically about 0.03 U.S. ¢/ESA km, varying little as traffic varies by a factor of 30.

The only noticeable variation is for negative values (though small) in row M (corresponding to growth and a high rate of weathering), and for rather high values in rows I, O, and V (each corresponding to positive traffic growth but zero and therefore implausible weathering). Traffic growth and weathering appear to have opposite effects on the road damage externality.

The road damage externality is insensitive to variations in traffic growth, rates of interest, the damaging power of traffic, the maintenance criterion, and traffic volumes, provided roads are appropriately designed and consistently maintained.

The second major conclusion is that the ratio of the marginal social cost to the average overlay cost (i.e. the ratio of the efficient road user charge and the costs of maintaining the network) are primarily determined by the climate, or the rate of weathering, m. For Tunisian conditions (semi-arid); the ratio is high and fairly stable, varying from 50 percent (row H, *very* high traffic growth) to 82 percent (row R, very high traffic volume). Since the ratio tends to be higher on more heavily trafficked roads, (partly because road strengths and maintenance standards are relatively higher on such roads), the overall network average will be nearer the upper end of the range. In very adverse conditions (freezing climates) the fraction of maintenance costs recovered by efficient charges is lower, possibly less than half. (The very low fractions in rows M, F, T, and U derive in large part from the choice of inappropriate initial roughness and design life—in adverse climates the initial roughness is likely to be lower as higher construction standards are normal.)

The interaction of growth and weathering can produce potentially surprising results. With high weathering, modest growth and long maintenance intervals (requiring very strong roads) traffic can generate negative externalities, which may dramatically reduce the MSC. Very low weathering combined with traffic growth can result in MSC *above* the average cost (Row V).

5. CONCLUSIONS

The Fundamental Theorem states that if road damage is solely caused by traffic, if there is no traffic growth, and if the age distribution of the road network is uniform, then provided the road is restored to its original state and strength whenever its roughness reaches a predetermined level, the road damage externality is identically zero when averaged over roads of different ages, and the efficient road user charge is equal to 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 ages.

For all other cases, where weather damage and traffic growth may be quantitatively important, the road damage externality, though no longer identically zero, remains quantitatively negligible (less than 0.1 U.S. ¢/ESAL km) provided the highway authority follows a consistent condition-responsive maintenance strategy. The main effect of weather damage is to lower the fraction of maintenance costs which can be efficiently charged to traffic, possibly by a large amount for severe climates and weak roads. Within any climatic zone for roads appropriately designed for expected traffic volumes, the fraction of costs efficiently chargeable to traffic is reasonably stable, most of the outlying values being associated with lightly trafficked roads involving a small fraction of total traffic. The formula for the allocable fraction of road maintenance costs, μ, provides a reasonably good approximation for the fraction of maintenance charges which can be efficiently recovered (and is increasingly accurate for low traffic growth rates), and can be very easily calculated.

The main conclusion is that this element of efficient road pricing will in general not recover maintenance costs, let alone other, non-traffic related costs (such as policing, lighting, etc.) On the other hand, congestion charges on an optimally adjusted highway system will cover (marginal) capacity costs, and if, as is likely in cities, there are increasing expansion costs, then it is possible that congestion charges will generate a surplus which may go some way to offset the likely deficit on maintenance costs.

There is another important practical consequence of considering both road damage costs and congestion costs together. The efficient road damage charge is proportional to the damaging power of the vehicle, and is thus predominantly borne by lorries. Given the fourth power law involved, increases in lorry loading, which tend to reduce private transport costs, greatly increase road damage costs, but tend to increase operating costs roughly linearly with vehicle mass. It is therefore hard to find a tax system based on fuel, tires, or vehicle depreciation, which successfully charges for distance driven and damage done. On the other hand the congestion effects of large lorries, while greater than cars and small lorries, tend to rise less than proportionately with load carried and weight. Taking both congestion and road damage costs together, it is considerably easier to design a system of vehicle taxation which varies with the total costs imposed. (See Newbery (1986c) for further discussion.)

Department of Agricultural and Resource Economics, University of California-Berkeley, Berkeley, CA 94720, U.S.A.

Manuscript received October, 1985; final revision received July, 1987.

APPENDIX

Analytical Derivation of the Road Damage Cost

The general case has traffic growing at steady rate g, and the case of zero traffic growth can be derived by setting $g = 0$. Current traffic is N_0 at date z, and the equation for roughness at date $t \geqslant z$ is

$$\text{(A1)} \qquad R(t) = e^{mt}\left\{ R_0 + k\left(X + N_0 E\int_z^t e^{g(u-z)}\,du \right)\right\},$$

(where X is cumulative traffic up to the current date). At $t = M$, roughness reaches its critical level $\bar{R}$:

$$\text{(A2)} \qquad \bar{R}e^{-mM} = R_0 + k\left\{ X + N_0 E\int_z^M e^{g(t-z)}\,dt \right\}.$$

Differentiate (A2) totally with respect to X and evaluate at $M = T$:

$$\text{(A3)} \qquad -m\bar{R}e^{-mT}\partial M/\partial X = k + kN_0 Ee^{g(T-z)}\partial M/\partial X$$

which can be solved for $\partial M/\partial X$ in terms of k. It is convenient to substitute for k from (A2) (replacing M by T) in which case

$$\text{(A4)} \qquad -\frac{\partial M}{\partial X} = \frac{\mu e^{-g(T-z)}}{N_0 E}, \qquad \mu = \left\{ 1 + \frac{m\phi(g)}{1 - R_0 e^{mT}/\bar{R}} \right\}^{-1}$$

where

$$\text{(A5)} \qquad \phi(\alpha) \equiv \int_0^T e^{-\alpha t}\,dt = (1 - e^{-\alpha T})/\alpha.$$

Thus at $g = 0$ (no traffic growth), $\phi(0) = T$, giving the value of μ shown in (22).

The cost of overlays depends on the required road strength after overlay, which can be found as follows. From (14) $kY(T) = \bar{R}e^{-mT} - R_0$, so that if R_0, $\bar{R}$, and T are to remain constant, $1/k$ must be proportional to $Y(T)$. Empirically $k = 9450\ S^{-1/\sigma}$ when Y is measured in millions of ESALs, where S is a measure of road strength related to the corrected structural number of the roads, SNC (itself a function of layer thicknesses and sub-grade strength), by the equation $S = 1 + SNC$. The power $1/\sigma$ measures the extent of increasing returns to road strengthening ($1/\sigma = 6.65$, or $\sigma = .1504$). If traffic grows at an annual rate g, the $Y(T)$ will increase by e^{gT} between overlay cycles, and if S increases by a factor $e^{g\sigma T}$ each T years, then R_0, $\bar{R}$ and T can remain constant. If S_0 is suitable for initial traffic N_0, then $S(z) = S_0 e^{-g\sigma z}$ is the strength of a road of age z now carrying traffic N_0.

The cost of an overlay of corrected structure number SNC, strength S, is given by Rolt (1981, p. 7) as

$$\text{(A6)} \qquad C_L = C_0 + \beta SNC = C_0 - \beta + \beta S.$$

For a road current age z, strength at next overlay $S_0 e^{g\sigma(T-z)}$, the present discounted cost of all future overlays is

$$\text{(A7)} \qquad e^{-r(M-z)} \left\{ \frac{C_0 - \beta}{1 - e^{-rT}} + \frac{\beta(SNC_0 + 1)e^{g\sigma(T-z)}}{1 - e^{-(r-g\sigma)T}} \right\} \equiv e^{-r(M-z)} C(z).$$

Empirically the equation for vehicle operating cost (3) is linear in roughness to a high degree of approximation:

$$\text{(A8)} \qquad v(R) = v_0 + b(R - R_0)$$

so that dv/dR in (25) is b, while $\partial R/\partial X$ is, from (A1), ke^{mt}. The last term in (25) is thus $N_0 e^{(r-g)z} bk[e^{(-r-g-m)z} - e^{-(r-g-m)T}]/(r-g-m)$. The value of k, which is a function of z, is given by (A2) at $M = T$; $\bar{R}e^{-mT} = R_0 + kEN_0 e^{g(T-z)}\phi(g)$ (ϕ is defined in (A5)).

The time average of equation (24) is thus

$$\text{(A9)} \qquad \frac{\partial \bar{F}}{\partial X} = \frac{1}{T}\int_0^T \frac{\partial F}{\partial X}\,dz = \frac{\mu}{N_0 ET} \left\{ \int_0^T re^{-(g+r)(T-z)} C(z) + e^{-r(T-z)} N_0 (v^* - \bar{v}) \right\}$$
$$+ \frac{b\left(\bar{R}e^{-mT} - R_0\right)}{TE\phi(g)} e^{-gt} \int_0^T \left\{ e^{(g+m)z} - e^{-(r-g-m)T} e^{rz} \right\} dz$$

where μ is given by (A4) and v^* is given in (25). The integrals can be evaluated to give

$$\text{(A10)} \qquad \frac{\partial \bar{F}}{\partial X} = \frac{\mu}{N_0 ET} \left\{ \frac{(C_0 - \beta)\phi(r+g)}{\phi(r)} + \frac{\beta(SNC_0 + 1) r\phi(r + g - g\sigma)}{(r - g\sigma)\phi(r - g\sigma)} + N_0 (v^* - \bar{v})\phi(r) \right\}$$
$$+ \frac{b\left(\bar{R}e^{-mT} - R_0\right) e^{mT}}{ET\phi(g)} \left\{ \frac{\phi(g+m) - \phi(r)}{r - g - m} \right\}.$$

The expression $(v^* - \bar{v})$ can be evaluated using (A8) and expressed in terms of b, R_0, $\bar{R}$, and the other parameters m, r, g, and T. The current rate of expenditure per ESAL km on overlays will be $C_0/(N_0 ET)$ since a fraction $1/T$ roads will have just reached the end of their overlay cycle and will require resurfacing to a strength suitable for initial traffic N_0, i.e. to a strength SNC_0 at a cost C_0.

The road damage externality is the sum of the last two terms in (A10), the first of which, involving $v^* - \bar{v}$, is negative, the second positive. Both are proportional to b (the rate of increase in VOC with roughness) and, since the RDE is small, the road damage costs are insensitive to b and to VOC.

REFERENCES

CAPLIN, A. S. (1985): "The Variability of Aggregate Demand with (S, s) Inventory Policies," *Econometrica*, 53, 1395–1410.

DEWEES, D. (1979): "Estimating the Time Costs of Highway Congestion," *Econometrica*, 47, 1499–1512.

DUPUIT, J. (1844): "On the Measurement of the Utility of Public Works," reprinted in *Transport*, ed. by D. Munby, pp. 19–57.

ELLET, C. JR. (1840): "A Popular Exposition of the Incorrectness of the Tariffs on Tolls in Use on the Public Improvements of the United States," *J. Franklin Institute*, 29, 225–232.

FEDERAL HIGHWAY ADMINISTRATION (1982): *Final Report on the Federal Highway Cost Allocation Study*. Washington, D.C.: Department of Transportation.

HIGHWAY RESEARCH BOARD (1962): *The AASHTO Road Test: Pavement Research*. Washington, D.C.: HRB Special Report 61E.

JONES, T. E. (1977): *Axle-loads on Paved Roads in Kenya*, TRRL Laboratory Report 763. Crowthorne, Berks: Transport and Road Research Laboratory.

KEELER, T. E., AND K. SMALL (1977): "Optimal Peak-load Pricing, Investment, and Service Levels on Urban Expressways," *Journal of Political Economy*, 85, 1–25.

KNIGHT, F. (1924): "Some Fallacies in the Interpretation of Social Costs," *Quarterly Journal of Economics*, 38, 582–606.

MOHRING, H. (1970): "The Peakload Problem with Increasing Returns and Pricing Constraints," *American Economic Review*, 60, 693–705.

——— (1976): *Transportation Economics*. Cambridge, Mass.: Ballinger.

MOHRING, H., AND M. HARWITZ (1962): *Highway Benefits: An Analytical Framework*. Evanston, IL: Northwestern U. Press.

NEWBERY, D. M. G. (1985): "The Fundamental Theorem of Road User Charges," CEPR Discussion Paper No. 59, Centre for Economic Policy Research, London.

——— (1986a): "Efficient Pricing of U.S. Interstate Highways," mimeo, Fiscal Affairs Department, International Monetary Fund, July 21.

——— (1986b): "On Cost Recovery from Optimally Designed Roads," mimeo, Fiscal Affairs Department, International Monetary Fund, Washington, D.C., Aug. 21.

——— (1986c): *The Design of Road User Charges and Transport Taxation in Tunisia*. Washington, D.C.: World Bank Transportation Department.

——— (1986d): "Markovian and Integrable Road Damage Functions," mimeo, Cambridge, October.

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

——— (1985): "Study on Pricing and Taxing of Transport Fuels, Part II: Technical Issues," Mimeo, Transportation Department, World Bank, May 22.

PIGOU, A. C. (1912): *Wealth and Welfare*. London: MacMillan.

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

SMALL, K. A., AND C. WINSTON (1984): "Welfare Effects of Marginal Cost Taxation of Motor Freight Transportation: A Study of Infrastructure Pricing," *NBER Conference on State and Local Public Finance*.

STARKIE, D. N. M. (1982): "Road Indivisibilities: Some Observations," *Journal of Transportation Economics and Policy*, 259–266.

U.S. FEDERAL HIGHWAY ADMINISTRATION (USFHWA) (1983): *Highway Performance Monitoring System Technical Manual*. Washington, D.C.: FHWA.

——— (1984): *1983 HPMS Needs Analysis*, HHP-12 of 11/08/84. Washington D.C.: FHWA.

VICKERY, W. (1969): "Congestion Theory and Transport Investment," *American Economic Review*, 59, 251–260.

WALTERS, A. A. (1961): "The Theory and Measurement of Private and Social Cost of Highway Congestion," *Econometrica*, 29, 676–699.

WINSTON, C. (1985): "Conceptual Developments in the Economics of Transportation: An Interpretive Survey," *Journal of Economic Literature*, 23, 57–94.