TRAVEL DEMAND FORECASTS AND THE EVALUATION OF HIGHWAY SCHEMES UNDER CONGESTED CONDITIONS

By Huw C. W. L. Williams and Yaeko Yamashita*

1. INTRODUCTION

It has long been argued, particularly by those opposed to the level of resources committed to the highway sector, that travel demand forecasts which are used to justify investments carry with them a self-fulfilling prophecy. Highway construction, it is suggested, is used as a means of accommodating large anticipated increases of traffic, thereby allowing that growth to occur. Furthermore, through the use of growth factors based on national projections or local "trip-end" estimates which treat incompletely or ignore the effects of capacity limitations and of generated traffic, significant distortion may arise in both local traffic forecasts and investment benefits. Indeed, as a result of such projections the benefits of improving speeds in heavily congested areas may often be such that many road schemes can be justified simply by placing a value on the associated travel time savings.

These views, which have not been subject to a detailed quantitative study, exist as a broad-based critique of institutionalised methods and national policies, and their possible inter-relationship. They remain of considerable interest and importance not only because of the need to adopt widely-accepted appraisal techniques, but also because of the influence which road traffic forecasts have in setting national strategies for transport investments. In the UK, for example, road traffic is projected to increase in the range 83-142 per cent by 2025, and these figures are cited by the government as the basis for a substantial highway construction programme (DTp, 1989; HMSO, 1989). The methods of major road appraisal currently adopted by the UK Department of Transport are subject to increasing criticism for the reasons noted above and in particular for assuming that the elasticity of demand with respect to travel time is zero.

Although such critiques do not always represent accurately the methods applied in practice, as we discuss later in the paper, nevertheless we believe that the forecasting approach and its implications for benefit assessment merit considerable additional study. The effects of local capacity limitations and changes in travel costs on estimates of demand and benefit are cases in point. Specifically, the assumptions relating to inelastic demand estimations at national and local levels and of 'growth cut-off' procedures require, and are currently receiving, detailed scrutiny.

In this paper we offer some views on the development of a consistent framework for travel demand forecasting and economic appraisal of highway schemes, and explore the consequences of a failure to achieve an appropriate specification of transport models.

* Department of City and Regional Planning, University of Wales, College of Cardiff.
We draw on an equilibrium approach, which has its origins in the work of Beckmann et al. (1956) and Thomson (1970), and seek to extend the work of Williams and Moore (1990), Williams and Lam (1991), Williams and Lai (1991), Williams et al. (1991) and Emmerson (1991), who considered some aspects of appraisal under elastic demand. Here we shall be explicitly concerned with the mutual interaction between the demand for and cost of travel over the lifetime of a project, and thus seek to introduce consistently the effects of changes in exogenous variables, such as car ownership and economic activity, together with the influence of travel cost changes and capacity limitations into the forecasting and evaluation methodology. Using this equilibrium approach, we shall address a key question:

What are the implications of forecasting traffic and evaluating user benefits derived from a highway scheme under the assumption that demand is insensitive to travel cost changes (zero elasticity) when the true elasticity is in fact non-zero? This mis-specification study and the range of sensitivity tests proposed will shed some light on the ranges of errors which might result from conventional appraisal assumptions and thereby allow the significance of the critiques noted above to be assessed.

In Section 2, we specify the equilibrium model of a simplified transport system within a temporal setting and identify the forecasts of traffic and user benefit measures which emerge from different model assumptions. In Section 3, we assume linear forms for the demand and user cost functions which comprise the model and derive analytic and numerical results for the traffic forecasts and benefit measures. While the linear analysis may have some practical relevance in non-urban contexts, its inclusion here is primarily to interpret the method and to provide insights into the effect of many variables and parameters which are at the heart of a highway appraisal. The linear analysis is thus used to guide the tests involving demand and cost functions more closely akin to those used in practical studies. The non-linear equilibrium model is specified in Section 4, and the numerical results derived from a range of assumptions and different parameter values are presented in Section 5. In a final section, we consider the implications of the analysis for both theoretical research and practical analysis.

2. THE EQUILIBRIUM FRAMEWORK

2.1 Equilibrium Forecasts
The equilibrium framework we adopt is an extension of that considered by Williams and Moore (1990). In a simplified model of a highway system, journeys are made by private transport between an origin and destination on a link of capacity $K$ and length $L$. The demand for travel $V(t)$ in year $t$ is expressed as a function of a set of exogenous variables $X(t)$, which includes car ownership and measures of economic activity, and variables endogenous to the transport system which are subsumed within the (generalised) cost of travel $c(t)$. This is represented by the demand function

$$V(t) = D[c(t), X(t), \sigma(t)]$$

in which $\sigma(t)$ is a set of (possibly time-varying) parameters.

In turn, the supply function representing the variation of user cost with volume is expressed generally as

$$c(t) = S[V(t), \phi]$$

262
TRAVEL DEMAND FORECASTS

H. C. W. L. Williams and Y. Yamashita

with the set \( \phi \) parameterising the functional form.

Interest focuses on the solution of these simultaneous equations which generate comparative static equilibrium states \([V(t), c(t)]\) over time. The effect of policies will be simulated by changes to the cost function from an initial form \(S_1\) to a final form \(S_2\), representing an investment and giving rise to modifications to the highway link characteristics such as an increase in its capacity.

In the following analysis, we shall adopt a demand model in which the effects of exogenous and endogenous factors are separable and of the form

\[
V(t) = \bar{V}(t_0) \ W[X(t), \gamma(t)] \ F[c(t), \theta(t)]
\]  

(3)

in which \(\bar{V}(t_0)\) is the volume of travel per unit time at time \(t=0\), the time of the investment, \(W\) is a function of exogenous variables alone and \(F(c)\) incorporates the effects of cost changes. We shall be particularly interested in the influence of the latter in the temporal development of the system with and without the investment.

The consequences of ignoring the effects of costs by assuming that \(F\) takes the value unity will be explored in inelastic demand (ID) calculations, in contrast to those derived from elastic demand (ED) functions in which \(F\) exhibits a cost dependency. We avoid the common terminology of 'fixed' and 'variable' demand calculations in the present case as changes will also occur under the influence of exogenous growth processes.

2.2 Notation

We introduce the following notation for volumes and generalised costs at an arbitrary time \(t\):

- \(V_1(t), V_2(t)\) equilibrium volumes before and after the investment, respectively
- \(c_1(t), c_2(t)\) equilibrium generalised costs before and after the investment.

Equilibrium states at time \(t\) are thus identified by the vectors

- \([V_1(t), c_1(t)]\) without the investment (reference system)
- \([V_2(t), c_2(t)]\) with the investment (policy system)

in which \(t\) takes the discrete yearly values from \(t_0\) up to \(t_N\), the time period over which the project is to be evaluated.

A bar over a variable will signify its calculation when the elasticity of demand is taken as zero. That is, the ID variables are defined by the following forms

- \([\bar{V}(t), \bar{c}_1(t)]\) without the policy (reference system under inelastic demand)
- \([\bar{V}(t), \bar{c}_2(t)]\) with the policy (under inelastic demand)

In ID calculations, the same value \(\bar{V}(t)\) is associated with both the reference and policy states — no additional traffic will result at time \(t\) as a consequence of a highway investment.

The central features of the equilibrium calculations are illustrated in Figure 1. The outward shift in the demand curve, a direct result of the time development in the exogenous variables, gives rise to the temporal succession of equilibrium states, as intersections of the demand and cost curves.

2.3 Forms of the demand function

For inelastic demand forecasts the volume of travel along the link will be given by

\[
V(t) = \bar{V}(t_0) \ W[X(t)]
\]  

(4)

which will also be expressed in terms of a cumulative growth factor \(G(t)\) up to time \(t\)

\[
V(t) = \bar{V}(t_0)[1+G(t)]
\]  

(5)
FIGURE 1
Succession of Equilibrium States over Time With and Without an Investment

FIGURE 2
The Benefit of an Investment under Inelastic and Elastic Demand
and decomposed into growth contributions through rates $r(t')$ in prior time periods $t_0 \leq t' \leq t$, by

$$G(\theta) = \prod_{t' = t_0}^{t} [1 + r(t')] - 1$$

(6)

In elastic demand calculations, the demand function $F(c)$ will take a difference or ratio form with respect to the costs at time $t_0$. An example of the former

$$F(c) = F[c(t) - c(t_0), \theta]$$

(7)

is the negative exponential function

$$F(c) = \exp\{-\beta[c(t) - c(t_0)]\}$$

(8)

and an example of the latter

$$F(c) = F[c(t)/c(t_0), \theta]$$

(9)

is the power function

$$F(c) = \left[\frac{c(t)}{c(t_0)}\right]^{-\alpha}$$

(10)

We would stress that changes in costs will result in a modelled response, from a wide variety of sources including alternative modes, time periods, locations, frequencies, and so on, and we shall not enquire of the specific source of any traffic generated by an investment. It is well known that the above functions (8) and (10) result in demand elasticities (with respect to generalised cost) of the form $-\beta c(t)$ and $-\alpha$, respectively. We shall consider the implications of the difference in cost dependence of the elasticity for demand forecasts and benefit estimation in Section 5.

In general discussion of the demand model and its response properties we shall refer to $\theta$ as an elasticity parameter, while $\alpha$ and $\beta$ will imply the specific forms in equations (10) and (8), respectively.

The function $F(c)$ has the following properties for both forms

$$F(c) = 1 \text{ for all values of } \theta \text{ when } c(t) = c(t_0)$$

(11)

and

$$F(c) = 1 \text{ if } \theta = 0$$

(12)

This implies that at any time $t$ the demand function characterised by a non-zero elasticity $\theta \neq 0$ passes through the point of intersection of the inelastic demand function and the line of constant cost $c(t) = c(t_0)$. That is,

$$V[X(t), c(t) = c(t_0), \theta \neq 0] = V[X(t), \theta = 0]$$

(13)

This condition is illustrated in Figure 2. Under inelastic demand the equilibria before and after the introduction of the policy are, at time $t_0$, at $A$ and $C$, respectively. At time $t$, under the influence of changes in the exogenous variables, the equilibria are formed at $E$ and $H$. For non-zero elasticities the equilibria at $t_0$, $A$ and $A'$ move to $D$ and $D'$, respectively, at time $t$. As expressed formally in equation (13), at a cost $c(t_0)$ the elastic and inelastic demand curves $DD'$ and $EH$ intersect (at point $F$).

We shall be particularly interested in the relationship between travel volumes computed under ED and ID assumptions, which will be expressed through the following ratios:

$$\xi(t) = \frac{V_1(t)}{V(t)} \text{ and } \xi_2(t) = \frac{V_2(t)}{V(t)}$$

(14)

In both cases the denominator is the volume computed under inelastic demand $(\theta = 0)$,
as in equation (5), while the numerators are respectively the equilibrium volumes before and after the investment, computed under non-zero values of θ.

We would expect $\xi_1(t)$, at any time t, to be less than or equal to unity. In Figure 2, the equilibrium volume at $D$ is less than that at $E$ because of the influence of cost increases on demand. At $t_0$, due to the generation of traffic by the investment under non-zero elasticity, the ratio $\xi_2(t_0)$ will be greater than unity. It is however difficult to say, a priori, whether $\xi_2(t)$ at an arbitrary future time will be less or greater than unity. In the figure the volume at the equilibrium point $D'$, corresponding to the new transport system, is shown to be greater than the volume at $H$ computed under the inelastic demand $\bar{V}(t)$, yielding $\xi_2(t) > 1$. However, this condition clearly depends on the time period and the nature of the curve $S_2$ for, if this were shifted upwards such that the point $H$ lay above $F$, then the ratio $\xi_2(t)$ would be less than unity. The functional dependence of $\xi_2(t)$ and $\xi_2(t)$ on the parameters of the demand and cost functions will be investigated in Section 3 and an explicit condition governing the sign of $\xi_2(t)$ will be established.

We now turn to the quantification of user benefit measures under ED and ID assumptions.

2.4 Benefit Measures
We define the following components of benefit arising from an investment at time $t_0$.

- $BED(t)$, $BID(t)$ user benefit in period t, computed under elastic and inelastic demand, respectively.
- $BED$, $BID$ the present value of the benefit streams computed under elastic and inelastic demand.

Adopting a social discount factor $d$, these quantities are related in the usual way by

$$BED = \sum_{t_0}^{t_f} \frac{BED(t)}{(1 + d)^t}, \quad BID = \sum_{t_0}^{t_f} \frac{BID(t)}{(1 + d)^t}$$  \hspace{1cm} (15)

In later sections we shall present analytic and numerical results for the benefit streams and present values of benefits, and for the proportional difference

$$\Delta = \frac{BID - BED}{BID}$$  \hspace{1cm} (16)

which represents the error in benefit assessment accompanying inelastic demand calculations in cases for which a non-zero elasticity is appropriate. From equations (15) and (16) we may achieve the following decomposition

$$\Delta = \sum_{t} \Delta(t) \frac{BID(t)}{BID} (1 + d)^t$$  \hspace{1cm} (17)

in which $\Delta(t)$ is the proportional difference between benefit measures computed in time t, as in equation (16). The variation of the quantities $\Delta(t)$ and $\Delta$ will provide useful information on the implications of mis-specifying the transport models and will be considered in the sensitivity tests described below.

With reference to Figure 2, the values of $BID(t)$ and $BED(t)$ are represented as follows

$$BID(t) = \text{Area } UEHT = \bar{V}(t)[\xi_1(t) - \xi_2(t)]$$  \hspace{1cm} (18)

$$BED(t) = \text{Area } VDDY = \frac{1}{2}[V_1(t) + V_2(t)][c_1(t) - c_2(t)]$$  \hspace{1cm} (19)

The proportional difference $\Delta(t)$ may thus be expressed in terms of the areas
\[ \Delta(t) = \frac{\text{Area } UEHT - \text{Area } VDDY}{\text{Area } UEHT} \]

\[ = \frac{\text{Area } UEFDV + \text{Area } YTHI - \text{Area } FID'}{\text{Area } UEHT} \]

the numerator being the difference between the light shaded and chequered areas.

The inelastic demand measurement of benefit may overestimate the benefit derived from elastic demand for two reasons: firstly, an overestimation of the 'true' cost \( c_1(t) \) will imply an excessive cost reduction achieved under the investment; and secondly, the trips generated by a policy, although receiving a benefit themselves, undermine the investment benefit accruing to existing travellers as \( c_2(t) \) is greater than (or equal to) \( \bar{c}_2(t) \). We shall refer to these as Type I and Type II errors, respectively.

From the discussion of sections 2.3 and 2.4, it is clear that the effect of generated traffic on investment benefits is inextricably bound to the cost dependency of traffic forecasts, and the consequences of inappropriate assumptions in the specification of the demand function — and in particular the elasticity parameter value — will be reflected in errors in both.

It is readily demonstrated that the quantity \( \Delta(t) \) reduces at time \( t = t_0 \) to the value given by Williams and Moore (1990, p.67), in which the source of any overestimation of benefits due to inelastic demand calculations is solely due to a Type II error as noted above.

3. TEMPORAL ANALYSIS UNDER LINEAR COST AND DEMAND FUNCTIONS

We shall compute the co-ordinates of the equilibrium states identified in Figure 2 in terms of parameters governing (locally) linear cost and demand functions, and from these derive measures of benefit accompanying the introduction of an investment. The process involves the adoption of linear approximations to the demand and cost curves \( DD' \) and \( S_1,S_2 \), and seeking their intersection.

3.1 Linear Expansions and the Computation of Equilibrium States

By expanding to first order the demand function (equation (1)) about the equilibrium state in the reference system \( A \{ \overline{V}(t_0), \overline{c}(t_0) \} \), we obtain an expression for the demand at time \( t \) in terms of changes in the exogenous and endogenous variables

\[ V_1(t) = \overline{V}(t_0) + \sum_{\mu} \frac{\partial D}{\partial X_\mu}[X_\mu(t) - X_\mu(t_0)] + \frac{\partial D}{\partial c}[c(t) - \overline{c}_1(t_0)] , \quad (20) \]

the derivatives with respect to the components \( X_\mu \) and \( c \) being evaluated at \( A \). If the demand function (1) is linear in \( X \) and \( c \), this would specify the function exactly, otherwise it represents a locally linear approximation. Equation (20) will be written in terms of the growth proportion \( G(t) \) as follows

\[ V_1(t) = \overline{V}(t_0) [1 + G(t)] + D' [c(t) - \overline{c}_1(t_0)] \quad (21) \]

with \( D' \), denoting the derivative with respect to travel cost.

Similarly, expanding the cost function \( S_1 \) about \( A \) yields
\[ c(t) = \bar{c}_1(t_0) + \bar{S}_1 [V(t) - \bar{V}(t_0)] \]  
(22)

with \( \bar{S}_1 \) denoting the derivative \( \partial \bar{S}_1 / \partial \bar{V} \), evaluated at \( A \). The simultaneous equations (21)-(22) may now be solved to identify the equilibrium state \( D \), the components of which are given by

\[ V_1(t) = \bar{V}(t_0) [1 + G(t) \psi] \]  
(23)

\[ c_1(t) = \bar{c}_1(t_0) [1 + E^s \psi G(t)] \]  
(24)

We have introduced the absolute values of the demand and cost elasticities

\[ E^D = \frac{\bar{c}(t_0)}{\bar{V}(t_0)} \left( \frac{\partial D}{\partial c} \right) \]  
(25)

\[ E^s = \frac{\bar{V}(t_0)}{\bar{c}(t_0)} \left( \frac{\partial \bar{S}_1}{\partial \bar{V}} \right) \]  
(26)

and the factor

\[ Z_1 = (1 + E^D E^s)^{-1}. \]  
(27)

In a similar way, the above expansion of the demand function about \( A \) and that of the cost function \( S_2 \) about point \( C \) in Figure 2 yield

\[ V(t) = \bar{V}(t_0) [1 + G(t)] + D^c [c(t) - \bar{c}(t_0)] \]  
(equation (21))

\[ c(t) = \bar{c}_2(t_0) + S^2 [V(t) - \bar{V}(t_0)] \]  
(28)

from which the co-ordinates of the equilibrium state \( D' \) may be established

\[ V_2(t) = \bar{V}(t_0) [1 + Z_2 \{ \psi E^D + G(t) \}] \]  
(29)

\[ c_2(t) = \bar{c}_2(t_0) [(1 - \psi) + E^s Z_2 \{ \psi E^D + G(t) \}] \]  
(30)

in which we define the following

\[ \psi = \frac{\bar{c}_1(t_0) - \bar{c}_2(t_0)}{\bar{c}_1(t_0)} \]  
(31)

\[ E^s = \frac{\bar{V}(t_0)}{\bar{c}(t_0)} S^2 \]  
(32)

and

\[ Z_2 = (1 + E^D E^s)^{-1}. \]  
(33)

\( \psi \) is a measure of the fall in travel cost achieved by the investment at the initial volume \( \bar{V}(t_0) \), while \( E^s \), is proportional to the cost elasticity at point \( C \), the constant of proportionality being \( \bar{c}_2(t_0) / \bar{c}_1(t_0) \).

By putting \( E^D = 0 \) the co-ordinates of the equilibrium state computed under inelastic demand are obtained

\[ \bar{V}(t) = \bar{V}(t_0) [1 + G(t)] \]  
(as in equation (5))

\[ \bar{c}_1(t) = \bar{c}_1(t_0) [1 + E^s G(t)] \]  
(34)

\[ \bar{c}_2(t) = \bar{c}_2(t_0) [(1 - \psi) + E^s G(t)] \]  
(35)

The equilibrium states, defined by equations (23), (24), (29) and (30), together with (5), (34) and (35), provide information about the effect of cost increases, arising from exogenous sources, on traffic levels in the reference system, and the extra traffic generated as a consequence of the investment policy.

268
TRAVEL DEMAND FORECASTS

H. C. W. L. Williams and Y. Yamashita

From equations (23), (29) and (5) the ratios $\xi_1(t)$ and $\xi_2(t)$, introduced in section 2.3, can now be determined for the linear model:

$$\xi_1(t) = \frac{V_1(t)}{\tilde{V}(t)} = 1 - \left[ \frac{G(t) E_1^D}{1 + G(t)} \right] \cdot \frac{E_1^D E_2^\psi}{1 + E_1^D E_2^\psi}$$  \hspace{1cm} (36)

The second term in (36) expresses quantitatively the effect of a Type I error, the extent to which the inelastic demand calculations overestimate the equilibrium volume derived with $E^D \neq 0$. Indeed, from equation (23) it can be seen that the factor $Z_2$ plays the role of establishing an 'effective' growth proportion $G(t)Z_1$

In turn, $\xi_2(t)$ is given by

$$\xi_2(t) = \frac{V_2(t)}{\tilde{V}(t)} = 1 - \frac{[G(t)(1-Z_2)-Z_2 E^D \psi]}{1 + G(t)}$$  \hspace{1cm} (37)

In the presence of the investment the equilibrium volume $V_2(t)$ will be greater or less than $\tilde{V}(t)$, that estimated under inelastic demand, according to the inequality

$$\xi_2(t) \geq 1 \text{ when } \frac{G(t) E_2^\psi}{\psi} \geq 1$$  \hspace{1cm} (38)

which, as expected, does not depend on the elasticity of demand, but only on the growth factor $G(t)$, the characteristics of the new cost function $S_2$, and the cost reduction $\tilde{c}_1(t_0) - \tilde{c}_2(t_0)$, achieved under the volume $\tilde{V}(t_0)$, which is measured by $\psi$. In the early stages of the project lifetime, and in particular at $t = t_0$ when $G = 0$, $\xi_2$ will exceed unity. Thereafter the direction of the inequality will depend crucially on $G(t)$.

We note from equations (23) and (29) that the extra traffic at time $t$ generated by the investment is given by

$$\delta(t) = V_2(t) - V_1(t) = \tilde{V}(t_0) E_1^D Z_2 \{ \psi + Z_1 G(t) (E_1^s - E_2^s) \}$$  \hspace{1cm} (39)

which, in the opening period $t_0$, is

$$\delta(t = 0) = E_1^D \tilde{V}(t_0) Z_2 \psi$$

a result established by Williams and Moore (1990). In that paper the quantity $B$ is equal to $\psi/2$ as defined in equation (31) above.

3.3 User Benefits Derived from the Investment

From equation (19), applied in conjunction with (23), (24), (29) and (30), we can express the perceived user benefits arising under elastic demand as

$$BED(t) = \tilde{V}(t_0) \tilde{c}_1(t_0) \left[ 1 + \frac{G(t)}{2} (Z_1 + Z_2) + \frac{Z_2 E_1^D \psi}{2} \right]$$  \hspace{1cm} (40)

which reduces under $E^D = 0$ to

$$BID(t) = \tilde{V}(t_0) \tilde{c}_1(t_0) [1 + G(t)] \{ \psi + G(t) (E_1^s - E_2^s) \}$$  \hspace{1cm} (41)
The proportional reduction of benefit at \( t \) measured by

\[
\Delta(t) = \frac{BID(t) - BED(t)}{BID(t)}
\]

will be written as

\[
\Delta(t) = 1 - \Pi(t)
\]  

with

\[
\Pi(t) = \frac{1 + G(t) (Z_1 + Z_2) + \frac{Z_2 E^D}{2} \frac{\Psi}{2} [\Psi Z_2 + G(t) (Z_4 E^s_1 - Z_2 E^s_2)]}{[1 + G(t)] [\Psi + G(t) (E^s_1 - E^s_2)]}
\]  

(42)

When \( E^D = 0 \), \( \Pi(t) = 1 \) and \( \Delta(t) = 0 \), as required, and at \( t = t_0 \)

\[
\Pi(t_0) = \left(1 + \frac{Z_2 E^D}{2}\right) Z_2,
\]

\[
\Delta(t_0) = (1 - Z_2) \left[ \frac{(Z_2)^2 E^D}{2} \frac{\Psi}{2} \right]
\]  

(43)

(44)

The quantity \( \Delta(t_0) \) will be negative when

\[
E^s_2 (1 + E^D E^s_2) < \frac{\Psi}{2}
\]

(45)

in which case the benefit to the traffic generated by an investment is greater than the disbenefit inflicted on other road users. If this condition is satisfied at \( t = t_0 \), then \( \Delta(t) \) may continue to be negative in the early years of the project, depending on the sizes of the elasticity parameters, a condition observed in some of the numerical experiments described below.

By expanding the function \( \Delta(t) \), defined by (42) and (43) to first order in the demand elasticity, it may be shown that

\[
\Delta(t) = \lambda(t) E^D
\]

(46)

with

\[
\lambda(t) = \frac{G(t) [(E^s_1)^2 - (E^s_2)^2] + \Psi E^s_2}{G(t) (E^s_1 - E^s_2) + \Psi} + \frac{1 + G(t)^{-1}}{2} [G(t) (E^s_1 + E^s_2) - \Psi]
\]

(47)

The error of applying an inelastic model is thus proportional to \( E^D \), if the latter is small (compared with \( 2/E^s_1 \)). At \( t = t_0 \) we have

\[
\lambda(t_0) = E^s_2 - \frac{\Psi}{2}
\]

(48)

consistent with equation (45), and at a large elapsed time

\[
\lim_{t \to \infty} \lambda(t) = \frac{3 (E^s_1 + E^s_2)}{2} \geq 0
\]

(49)

At small times, \( t = t_0 \), Type II errors predominate and these will be strongly influenced by the size of \( E^s_2 \). Thereafter Type I errors become increasingly important (notwithstanding the discounting factor) and their size is primarily determined by the magnitude of \( E^s_1 \). This may be seen from Figure 2 as the demand curve shifts outwards from its initial state through \( A \).
TRAVEL DEMAND FORECASTS

H. C. W. L. Williams and Y. Yamashita

FIGURE 3
Variation of $\xi_{s}(t)$ and $\xi_{d}(t)$ with Time for Growth Rates $r = 2\%$ and $r = 4\%$

FIGURE 4
Variation of $\Delta$ with Demand Elasticity $E^0$ for Different Values of $E^s_1$
3.4 Numerical Results
In Figures 3 and 4 we illustrate some of the key features of investment appraisal under elastic demand. The variation of $\xi_1(t)$ with time shows the extent to which a traffic forecast made on the basis of changes in exogenous factors alone overstates that which allows for cost increases for different values of the elasticity parameter. For example, under the conditions $E_D^D = 0.5, E_S^S = 1.0$ and a constant annual growth rate of 4 per cent, the overestimation after 20 years is given by $1 - \xi_1(t = 20)$ which is approximately 20 per cent. Figure 3 also shows for different growth rates the behaviour of $\xi_2(t)$, and the extent to which $V_2(t)$ exceeds $V(t)$ in the early years of a project before the inequality reverses in later years.

While Figure 3 exhibits the characteristics of the reference equilibrium state evolving under the influence of changes in the factors $X$ and $c$, Figure 4 illustrates the effects of cost changes on investment benefits as a function of the demand elasticity $E_D^D$ for different values of $E_S^S$, which is proportional to the slope of the cost curve $S$ at the initial equilibrium state $A$. Through the incorporation of Type I and Type II effects, the behaviour in Figure 4 subsumes that in Figure 3 — errors in benefit estimation and demand forecasts are, under elastic demand, inextricably linked as we noted in Section 2.

4. TRAVEL FORECASTS AND BENEFIT CALCULATIONS WITH NON-LINEAR DEMAND AND COST FUNCTIONS

4.1 User Costs
The generalised cost of travel at time $t$ will be expressed in terms of the travel time $\tau$ and operating cost components through the standard form

$$c(t) = \phi_1(t) \tau(t) + \phi_2(t) L,$$

in which the unit value parameters $\phi_1$ and $\phi_2$, together with their temporal variation, are taken from the COBA Manual (Department of Transport, 1980). Travel time is dependent on the volume of travel per unit time through the relation

$$\tau(t) = \frac{L}{u[V(t)]}$$

with $u(V)$ taken as a standard speed-flow relation, the form of which is shown in Figure 5 for the reference system $P_0$ and three policies $P_1, P_2$ and $P_3$. Each curve is characterised by three regions: $R_1$, a free flow regime in which velocity is independent of flow up to a 'free flow limit'; $R_2$, a regime below the nominal capacity $K$, in which velocity decreases linearly with flow; and $R_3$, a regime above capacity which accounts for the influence of queueing. The variation of generalised cost with volume and more specifically the slope of the function

$$c(t) = S[V(t)]$$

determined by equations (50) and (51), will depend on the detailed specification of these regimes for each policy, and the latter has been selected to allow the effects of different values of $E_S$ in the neighbourhood of capacity to be examined.

4.2 The Demand Function
The two forms of demand function have already been described in Section 2 and will be summarised here as follows:

272
TRAVEL DEMAND FORECASTS

H. C. W. L. Williams and Y. Yamashita

FIGURE 5
Speed-Flow Curves for the Reference System (P₀) and for Three Policies (P₁, P₂, P₃)

Negative Exponential Form:
\[ V(t) = \bar{V}(t₀)(1 + r)^t \exp \left\{ -\beta [c(t) - \bar{c}(t₀)] \right\} \tag{52} \]

Constant Elasticity Form:
\[ V(t) = \bar{V}(t₀)(1 + r)^t \frac{c(t)}{\bar{c}(t₀)} \tag{53} \]

In both forms a constant rate of growth \( r \) is assumed.

When implementing the negative exponential form (52), a value of \( \beta \) is selected to yield a required (absolute) value of elasticity \( E^D(t₀) \) at the initial time \( t₀ \), that is
\[ \beta = \frac{E^D(t₀)}{\bar{c}(t₀)} \tag{54} \]

The elasticity of demand in any future period \( E^D(t) \) will depend on the increase in the equilibrium cost over time through the relation
\[ E^D(t) = E^D(t₀) \frac{c(t)}{\bar{c}(t₀)} \tag{55} \]

4.3 Equilibration and the Computation of Benefit Measures

Although 'exact' measures of perceived user benefits are available in the present case through analytic manipulation of the standard consumers' surplus integral at any time period, here we adopt the usual rule-of-a-half approximation (19) for the elastic demand case, and the change in total travel cost (18) under inelastic demand, using the appropriate values of equilibrium costs in each case.
As in the single period analysis considered by Williams and Moore (1990) the computation of these equilibrium states and resultant benefit measures is achieved by straightforward numerical methods. Under inelastic demand calculations the volume
\[ \bar{V}(t) = \bar{V}(t_0)(1 + r)^t \]
is determined for each time period, and the costs \( \bar{c}_1(t) \) and \( \bar{c}_2(t) \) are then computed directly from the user costs functions \( S_1 \) and \( S_2 \). For non-zero elasticities a self-consistent solution of the simultaneous equations (50)-(53) is achieved to give \( V_1(t) \) and \( c_1(t) \) through a simple iterative process, with a damping mechanism employed to eliminate the possibility of divergent oscillations between successive estimates of volumes and costs. The new cost function under investment \( S_2 \) is then substituted to derive \( V_2(t) \) and \( c_2(t) \).

### 4.4 The Implementation of ‘Growth Cut-offs’ in the ID Approach

If, in practical tests, projected traffic volumes are such that the corresponding speed falls to an ‘unrealistically low value’, a procedure is invoked to ‘cut off’ the traffic accommodated by a link or part of a network. This procedure, which will be applied only in conjunction with the inelastic demand assumption, is introduced to prevent a rapid increase in travel costs and excessive values of benefits arising from an investment.

The introduction of ‘cut-offs’ has been much criticised in the past for their somewhat arbitrary nature and the failure adequately to reflect behavioural responses to cost changes. We view the approach as a practical expedient which represents a simple but imprecise and ‘discontinuous’ model of the behaviour of a transport system in the neighbourhood of capacity. We propose to investigate the effect of ‘cut-offs’ in a set of ID tests in two forms, involving:

(i) a single cut-off \( T \) common to both the reference and policy states;
(ii) two cut-offs \( T_1 \) and \( T_2 \) applied before and after the investment respectively, to allow for any increase in capacity to generate additional traffic.

Above the cut-off no further growth is assumed to occur.

### 5. NUMERICAL RESULTS FOR THE NON-LINEAR CASE

#### 5.1 Introduction

The equilibrium volumes and costs before and after the investment and the benefit measures computed from them are dependent on the parameters and forms which characterise the demand and cost functions. We have examined the sensitivity of these measures to the following: the specification of the demand function (negative exponential (NE) or constant elasticity (CE) forms and the elasticity parameters \( \beta \) and \( \alpha \) embedded in each); \( \kappa \), the ratio of the volume \( \bar{V}(t_0) \) to capacity \( K \) of the initial highway; the policy \( P \), expressing the transition between two speed-volume relations; and \( r \), the annual rate of growth of traffic due to exogenous sources. We also seek the dependence of the results determined under inelastic demand to any growth ‘cut-offs’ imposed on the traffic volume. In all benefit calculations, the values accruing over time will be discounted at a constant rate of 8 per cent per annum.

No attempt is made to cover comprehensively the full interdependence of the above variables but we focus on what we perceive as the more significant characteristics of the problem.
TRAVEL DEMAND FORECASTS

H. C. W. L. Williams and Y. Yamashita

FIGURE 6
Variation of $\xi_j$ with Demand Elasticity for Different $\kappa$

FIGURE 7
Variation of $\Delta$ with Elasticity of Demand for Different $\kappa$
5.2 The Variation of $\xi_1$ and $\Delta$ with Elasticity and Initial Volume

In Figure 6 we display the variation of $\xi_1$ at the time $t = 30$, the life of the project, with the elasticity of demand $E^D$ for different values of $\kappa$, the initial volume to capacity ratio. A negative exponential demand function is assumed and a growth rate of 4 per cent adopted.

At a low value of $\kappa$, say 0.25, when the initial volume is well within the free-flow regime $R_1$, the diagram indicates a relatively low overestimation of the equilibrium volume $V_1(t = 30)$ by the inelastic value $\bar{V}_1(t = 30)$. For example, when $E^D = 0.5$, the ratio $\xi_1(t = 30) = 0.85$, that is, the discrepancy between the forecasts $V_1$ and $\bar{V}_1$, is only 15 per cent. When $E^D = 1$ this increases to 25 per cent.

In contrast, a higher value of $\kappa$, in or near the regime $R_2$ of the initial speed-flow curve, results in a rapid rise of costs and a correspondingly greater response to cost changes at a given elasticity value. For $\kappa = 0.75$, for example, an elasticity value as low as 0.25 results in an estimate of the equilibrium volume of just over a half (0.55) that derived from inelastic demand calculations. As $\kappa$ increases thereafter, and for elasticities greater than 0.3, the ratio $\xi_1$ is relatively insensitive to either $\kappa$ or to $E^D$. For low values of $E^D$, however, $1 - \xi_1$ is approximately proportional to $E^D$, a form suggested by the linear analysis.

The implications of this relationship are clear. In congested regimes, forecasts of demand which ignore the potential travel response to cost increases may be subject to very large errors.

This variation, characterising Type I effects, is reflected in benefit calculation and the discrepancy between $BID$ and $BED$, and their temporal components. In Figure 7 we display the variation of $\Delta$ with elasticity value, for different ratios $\kappa$. All curves pass through the origin and asymptotically approach the line $\Delta = 100$ per cent at high values of elasticity as Type I and Type II effects erode the benefits which would otherwise have accrued under inelastic demand. For the low value $\kappa = 0.25$ the discrepancy between $BID$ and $BED$ for an elasticity as high as unity is only 15 per cent. However, for $\kappa$ equal to or greater than 0.5, inelastic demand estimates of benefits overstate the equilibrium values derived from an elasticity of $E^D = 0.25$ by at least 50 per cent.

In Table 1 we exhibit the variation of $\Delta$ with elasticity for different policies and for the NE and CE forms for the demand function. Differences arise because of the variation of elasticity with costs. At the initial time the elasticities are equal, but towards the end of the 30-year period they will differ according to the rise in the equilibrium cost. The NE function is characterised by a higher average elasticity than the CE form over the period of the project appraisal, and this results in a relatively higher value of $\Delta$.

The important feature of this table is that the effects are very large even for 'small' elasticities, and the differences, of the order of 10 per cent, arising from policy and model form do not compromise this conclusion. Errors in forecasts and benefits due to a non-zero elasticity are potentially very large.

5.3 Variation of $\xi_1$ and $\Delta$ with Demand Elasticity and Growth Rates

In Figure 8 we display the variation of $\xi_1$ with respect to time for different growth rates, for a constant elasticity demand curve with $\alpha = 0.5$, and with $\kappa = 0.75$. For a rate of 4 per cent inelastic forecasts will overstate equilibrium values by a little over 50 per cent after year 30, the temporal relationship being of approximately linear form. The variation of $\Delta$ with respect to growth rate also indicates a strong functional dependency,
TRAVEL DEMAND FORECASTS
H. C. W. L. Williams and Y. Yamashita

TABLE 1

<table>
<thead>
<tr>
<th>ED Policy</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50.43</td>
<td>66.63</td>
<td>74.55</td>
<td>79.05</td>
</tr>
<tr>
<td></td>
<td>61.92</td>
<td>73.18</td>
<td>78.32</td>
<td>81.39</td>
</tr>
<tr>
<td>2</td>
<td>53.57</td>
<td>69.56</td>
<td>77.91</td>
<td>82.41</td>
</tr>
<tr>
<td></td>
<td>65.27</td>
<td>76.77</td>
<td>81.98</td>
<td>84.98</td>
</tr>
<tr>
<td>3</td>
<td>58.73</td>
<td>75.70</td>
<td>84.46</td>
<td>89.14</td>
</tr>
<tr>
<td></td>
<td>71.39</td>
<td>83.61</td>
<td>89.12</td>
<td>92.03</td>
</tr>
</tbody>
</table>

as shown in Figure 9. This behaviour was suggested in the linear analysis, expressed in equations (36) and (43).

5.4 Temporal Effects
In Figure 10 we display the temporal variation of benefits in both undiscounted and discounted forms for both CE and NE demand models. Values of ED = 0.25, 0.75 and r=4 per cent are assumed. The larger values of BED(t) for large t in the constant elasticity form are a result of the lower effective elasticity over the period.

The figures display a feature not uncommon in the estimation of benefits, namely the rapid rise in undiscounted yearly benefits in the inelastic demand calculations. Even after discounting at 8 per cent, BID(t) is rising at year t = 30. In contrast, traveller response to cost changes in the neighbourhood of capacity is such that the equilibrium values of the discounted benefits BED(t) decrease with time. In aggregate terms the value of Δ is 50 per cent for the constant elasticity demand form.

5.5 The Effect of Growth ‘Cut-offs’
It is precisely to prevent this increase of benefits in the later years of a project that the growth cut-offs are introduced in ID calculations, as discussed in Section 4. In Figure 11 we illustrate the effects of applying cut-offs in different forms. The case with no ‘cut-off’, which is used for reference purposes, is shown in Figure 11a. In all cases the elastic demand calculations are applied without any cut-off, the reduction in demand compared with the value which would have resulted from exogenous growth alone, is achieved through the response to cost changes.

In Figure 11b, the effect of applying a cut-off of 2400 vehicles per hour (the nominal capacity of the new road) in both S1 and S2, is shown. The suppression of higher values in BID(t) draws the distributions closer together. Figure 11c involves the calculation with a lower cut-off at the capacity of the initial highway, 2200 vehicles per hour. We note here the similarity in the benefit streams derived from ID calculations with a cut-off and the ED results derived in this case with an elasticity of 0.25.

277
FIGURE 8
Variation of $z(t)$ with Time for Different Growth Rates $r$: $E^D = 0.5, \kappa = 0.75$, Policy $P_1$

FIGURE 9
Variation of $\Delta$ with Growth Rate, $r$
TRAVEL DEMAND FORECASTS

H. C. W. L. Williams and Y. Yamashita

In the above ID tests an expansion of capacity will not generate additional traffic, the benefit will simply arise from a reduction of cost which will be maintained after the threshold is passed. It seems unrealistic to suppose that, in congested conditions, a major increase in capacity will not generate some additional travel. We have allowed for this in a further numerical test by applying different thresholds in the initial and final states, indeed, cutting off growth after their respective nominal capacities. This allows additional traffic to grow and receive a benefit in the process until the new capacity is reached. In so doing, the equilibrium travel cost is raised and the benefit of the investment reduced in later years (due to Type II effects) as the comparison between 11c and 11d indicates.

These figures indicate both the sensitivity of the inelastic demand calculations to the growth cut-off convention adopted, and the possibility that their effects may be used to approximate the results of a theoretically more appropriate but implementationally more problematic elastic equilibrium method.

6. CONCLUSION

In all countries road traffic is growing rapidly due to the increase of car ownership and economic activity. In aggregate terms, and particularly at the national level, it will continue to do so for the foreseeable future. At a local level, however, and in particular time periods, this growth will be modified by capacity limitations. We know of no practically applied models which treat this effect convincingly.

We have considered the theoretical and numerical implications of an equilibrium framework which allows self-consistency in traffic forecasting and the evaluation of transport policies. Both the equilibrium demand forecasts and the benefit measures have been shown to be sensitive to a range of parameters and in particular the elasticity of demand. We have demonstrated that application of inelastic, or ‘fixed demand’, methods which are commonly applied in highway appraisal might result in a significant overestimation both of road traffic and the benefits of policies if there is even a small propensity to respond to cost changes. The framework presented allows some of the critical issues in the debate on the validity of forecasting and appraisal methods to be subject to scrutiny and numerical (sensitivity) analysis.

Our results generalise those of Williams and Moore (1990) who considered the influence of generated traffic on scheme benefits when an investment is first introduced. Here, the mutual effect of travel costs on equilibrium volumes results in an even greater difference between benefits computed under elastic and inelastic demand when these are considered over the lifetime of a project.

It is important to extend this basic approach to consider:
(a) applications to network problems and the refinement of the link-based representation to consider detailed junction characteristics;
(b) extension of the single mode model, which embraces different forms of response, to consider specific substitutes, such as time period, location, mode, and so on. In a general spatial setting the specification of the demand function will have important implications for the variation of elasticity with trip length.

We would expect a qualitative similarity between the results of the single link and network problems as in the corresponding single period studies reported by Williams.
and Lam (1991), Williams and Lai (1991), and Williams et al. (1991). Such extensions are currently being considered by the authors and by the UK Department of Transport.

The development of methods and techniques in the realm of public policy appraisal does of course seek a compromise between what is theoretically desirable and practically achievable. The adoption of several parameters to characterise the demand functions places a great onus on estimation and verification. In the present context very little empirical information is available to estimate the elasticity of demand for highway travel with a great degree of confidence. Identifying the effects of capacity or level-of-service variables on demand, and disentangling causes and effects (are capacity additions the cause or consequence of changes in traffic?) in longitudinal or cross-sectional studies will continue to prove a significant empirical challenge.

National or regional level traffic models which omit capacity effects generate 'unrestrained' forecasts which should be applied with great caution at the local level. How traffic increases will be accommodated over space and in different time periods is subject to much uncertainty. In order to avoid biases in forecasting and the appraisal of specific schemes, we would argue that it is important to take into account the influence of changes in natural growth processes and travel costs on the volume of traffic using generalisations of the model framework described in this paper. Some form of sensitivity analysis of the results within the equilibrium framework would appear to be justified. Our results give some support to the practical expedient of applying growth cut-off methods in forecasting and evaluation, and we would advocate further research into the use of these methods to assess the extent to which they may approximate the results derived from equilibrium models. We believe that the effects described will become increasingly important as congestion on local and major roads increases and the problem of reconciling national and regional traffic forecasts with those appropriate to local schemes becomes even more evident.

Date of receipt of final typescript: October 1991

REFERENCES