# RAND EUROPE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

Jump down to document ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

## Support RAND

Browse Books & Publications

Make a charitable contribution

## For More Information

Visit RAND at www.rand.org

Explore RAND Europe

View document details

## Limited Electronic Distribution Rights

# Cost Damping in Travel Demand Models

Report of a study for the
Department for Transport

Andrew Daly

RAND EUROPE

The research described in this report was prepared for the UK Department for Transport.

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND®** is a registered trademark.

# Preface

This report has been prepared under contract for WSP and PB, joint managers of the Framework, under the terms of the Department for Transport (DfT) Technical Research Framework PPRO 4/45/4. Under Lot 4 of this Framework, DfT issued a brief to make a study of a number of questions relating to Cost Damping and Realism Testing in travel demand models. The consortium led by WSP and PB proposed a team and the work has been completed by that team.

The present report, which has been prepared by RAND Europe, deals with three of the questions that specifically relate to Cost Damping, which refers to the feature in some models that the impact of cost and/or time is reduced for longer journeys. This report is intended as input to the final recommendations of the study and revised draft Guidance which DfT may issue to local planners. However, it may also be of general interest to transport analysts.

Any interpretations or opinions expressed in this report are those of the author and do not necessarily reflect the views of the Department or the other agencies involved in the study.

RAND Europe is an independent not-for-profit policy research organisation, which aims to improve policy and decision making in the public interest, through research and analysis. RAND Europe's clients include European governments, institutions, non-governmental organisations and firms with a need for rigorous, independent, multidisciplinary analysis. This report has been peer reviewed in accordance with RAND's quality assurance standards.

For more information about RAND Europe or this document, please contact:

Charlene Rohr
Director of Modelling

RAND Europe
Westbrook Centre
Milton Road
Cambridge CB4 1YG
United Kingdom
Tel. +44 (1223) 353 329
Email: crohr@rand.org

# Contents

# Table of Tables

# Acknowledgements

# Summary

## Introduction

The Department for Transport maintains Guidance for those concerned with predicting traffic flows, on the internet site www.webtag.org.uk. In particular Unit 3.10 (on Variable Demand Modelling) describes the modelling procedures that are recommended. Unit 3.10 and this report focus on personal travel – that is, freight movements are not considered.

To summarise, the modelling procedures represent travel as being movements between pairs of zones defined in the area of study and are based on the representation of the impedance to travel between two zones. Impedance or generalised time – WebTAG calls it 'generalised cost' – is measured in minutes of travel time, with cost converted to time units by dividing by a 'value of time' and additional weighting given to travel components such as walking, waiting or time spent in crowded circumstances. When used in the models, generalised time is multiplied by a parameter (called lambda), which represents the marginal sensitivity of behaviour to each minute of travel.

The aim of the present study is to help the Department to improve its Guidance concerning the specification of these lambda parameters. The study is organised in two Tasks:

Task A:      to answer eight Questions relating to key issues which were previously unanswered concerning the Guidance that ought to be given on the lambdas (see text box); and

Task B:      to prepare new Guidance on the basis of the findings of Task A.

**Text Box: The eight DfT questions for this study**

---

The Department wishes to know:

1.  How the requirements for realism testing in WebTAG are addressed in practice;

2.  Whether there are any cost damping mechanisms in use in the UK or elsewhere other than those listed in the Brief;

3.  How widely each mechanism is used a) in the UK and b) elsewhere;

4.  How effective are these mechanisms in addressing requirements for realism testing?

5.  What are the other practical advantages and disadvantages of each mechanism?

6.  What are the theoretical basis and empirical evidence for each of these mechanisms?

7.  If the scope for realism testing should be broadened, and, if so in what way?

8.  What methods are available for implementing realism testing in models used for major scheme appraisal and how does this compare with the methods that have been adopted?

---

Task A is itself broken down into three Work Packages (WP):

WP1:    to answer the Questions (1, 4, 5 and 8) which relate to issues concerning the testing of models and the adjustment of lambdas to give more realistic models;

WP2:    to answer the Questions (2, 3 and 6) which relate to the ways in which the lambdas should vary within a model – in particular how they should be adjusted by 'Cost Damping'; and

WP3:    to answer Question 7 which asks how realism testing should be broadened and how the tests should be specified.

The present document forms the final report of WP2 of Task A; that is, it answers Questions 2, 3 and 6 which relate specifically to Cost Damping.

The context of the present report is given by the following definition of Cost Damping, slightly amended from the Department's original Brief for this work.

> *The 'standard' modelling approach, as recommended in WebTAG Unit 3.10, assumes a model in which a) generalised cost is linear in its components (time, money costs and so on) and b) lambda parameters are constant throughout. However, the Department is aware that some models include additional elements or different structures that have the effect of reducing the model sensitivity with increasing distance. These mechanisms are what we mean by 'cost damping'.*[1]

---

[1] By 'constant' here is meant that the parameters do not vary as a function of time or cost. Different lambda values would generally be used for different travel purposes and for different aspects of behaviour, such as mode or destination choice. When mode choice is 'above' destination choice in the hierarchy, it is also possible to define separate lambda values by mode for modelling destination choice.

The following sections of the Summary give brief answers to the three Cost Damping questions and draw conclusions and recommendations which are intended to support the development of new Guidance. The technical discussions on which these answers and conclusions and recommendations are based are presented in the following chapters of the report.

- Chapter 1 reports correspondence with international experts which has supported the parts of the answers to Questions 2 and 3 relating to the use of cost damping outside the UK. (The corresponding information within the UK was collected for Work Package 1.)

- Chapter 2 is a revised version of a paper presented at the 2008 European Transport Conference in The Netherlands. This Chapter gives the basis for answering Question 6.

- Chapter 3 summarises the microeconomic theory against which recommendations for cost damping mechanisms might be tested

## Answers to the Cost Damping questions

The context in which these questions are asked is that models are being used to forecast matrices of trips or tours (origin-destination or production-attraction) by predicting some or all of travellers' choices of frequency, mode, destination and time of travel. The models are expected to make their predictions by calculating an impedance of travel for each of the alternatives which can be expressed in minutes of travel time; this measure can probably best be called generalised time. The models are of the nested logit type.

In fact, this approach seems to be almost universally used, both in the UK and elsewhere, although some models would require to be slightly reformulated to fit this paradigm. This uniformity means that information from a wide range of sources can be used to answer the questions.

### QUESTION 2:   What are the cost damping mechanisms in use in the UK and elsewhere?

In the UK, the following cost damping mechanisms are used[2]:

    A.      damping generalised time by a power of distance;

    B.      increasing the value of time by a function of distance;

    C.      log cost formulation;

    D.      addition of constants that depend on distance;

    F.      origin-destination-specific scaling of generalised time;

    G.      power function of generalised time;

---

2      The labelling of the mechanisms A, B etc. relates to that used in Chapter 2, specifically in Table 1. More detailed explanations of each of the mechanisms are given in that chapter.

H.      detailed segmentation by purpose, e.g. distinguishing 10-20 purposes.

We cannot claim that our survey is exhaustive, but it is unlikely that there are other mechanisms that are used in a significant number of studies.

Outside the UK, our survey has necessarily been much more limited in scope and it is very likely that mechanisms have been missed. It is notable that most of the mechanisms used in the UK are *not* used in other countries; in fact, only mechanism C is known to be used in exactly comparable form both inside and outside the UK. The following mechanisms have been found in foreign studies:

C.      log cost formulation;

E(a)    Box-Tukey[3] transformations applied to the generalised time;

E(b)    Box-Tukey transformations applied to components of the generalised time.

It should be noted that mechanism C is actually a special case of mechanism E(b) and G is a special case of E(a). Mechanisms A and F are closely related in practice. Moreover mechanisms D and H are not strictly cost damping mechanisms at all, although they have some features that relate to cost damping; they are not discussed further in any detail.

## QUESTION 3:   How widely is each mechanism used a) in the UK; and b) elsewhere?

By far the most common mechanism in use in the UK is A from the list above, which is used by different consultants in several large-scale studies. Mechanisms B and C appear to be used in a small number of studies each; B is sometimes used in conjunction with A. Mechanism D is, to our knowledge, used only in the National Transport Model. Mechanism F is provided in the DIADEM software, but we know of only one application that has used it, and that a rather limited one. Mechanism G is used in only one study. Mechanism H has been reported as in use, but we do not know where.

RAND Europe has used mechanism C in a number of studies outside the UK. Mechanism E(b) appears to be used rather seldom, primarily in work by Marc Gaudry, and E(a) has been seen in only one study (Swiss National Model). In general it appears that cost damping is more prevalent in the UK than elsewhere.

## QUESTION 6:   What are the theoretical basis and empirical evidence for each of these mechanisms?

The *empirical evidence* for requiring cost damping is very strong. It may be summarised as follows.

- Modelling studies, in the UK and elsewhere that test for cost damping at the time that the models are estimated almost always find that specifications with cost damping are significantly and substantially better than undamped formulations.

- Practical modelling studies in the UK often find that they are not able to obtain reasonable forecasting behaviour in the models – in particular, reasonable

---

[3]      The better-known Box-Cox transformation is actually a special form of the more general Box-Tukey transformation. Both of these transformations are explained in detail in Table 1, Chapter 2.

elasticities – without cost damping, but that the use of cost damping allows reasonable models to be obtained.

- Value-of-time studies in the transport sector most often find that the value of travel time in money terms increases with the length of the trip, usually concluding that this effect is due to a decline in the marginal impact of money cost, rather than an increase in the marginal impact of time.

The strength of the evidence comes from the large number and overwhelming proportion of studies that make these findings.

It should be noted that almost all of the tests that have been made assume that the modelling scale does not vary with trip length – that is, that the impact of one unit of impedance on choice proportions is constant; this restriction is inherent in the logit formulation used for practical modelling. In more sophisticated work, when that assumption is abandoned and models other than nested logit may be used, different results can be obtained. However it should be noted that current models that allow varying model scale are both experimental and very computer-intensive, so that their use in practical studies would not currently be feasible.

While the question is posed to require empirical evidence for each of the tests separately, we have found no evidence that would allow discrimination between the power of the mechanisms. Study teams typically select one mechanism *a priori*, show that it is better than having no mechanism, and continue with their work. There is strong empirical evidence[4] that the value of time increases with trip length, so that cost damping mechanisms that permit this change are to be preferred to those that do not. This effect apart, the large body of empirical evidence justifies the use of cost damping in general but does not indicate that one mechanism is better than another.

The *theoretical basis* is a more complicated issue.

If the models are viewed as microeconomic constructs, there exists a body of theory against which they may be judged. A summary of that theory is given in Chapter 3, which concludes that there is almost no economic theory to support or reject cost damping.

If however the models are viewed more generally as derived from an unspecified theory of behaviour, then several points may be made which allow us to test the mechanisms for their general reasonableness.

- Mechanisms involving a distance variable, such as A and B (and possibly D and F) are difficult to justify because the role of the distance variable and its impacts on cost and time sensitivity are not clear; therefore the assumption that its impact in the future will remain unchanged, not subject to influence by policy or exogenous changes, is questionable.
- The 'kilometrage test', which requires that a model should predict a decline in the total kilometrage driven when the cost per kilometre increases uniformly, can be used to test the reasonableness of models. This test allows the rejection of several model forms, including (in principle) those that are not smooth.

---

[4]     See the literature reviewed in Chapter 2.

- The evidence that value of time increases with trip length is very strong, so mechanisms that do not allow that effect (A, D, E(a), F and G) are inferior in this respect to mechanisms that do allow this effect (B, C, E(b) and possibly H). In the latter group of mechanisms the sensitivity with respect to cost should decline more rapidly than the sensitivity with respect to time, which may not decline at all (and cannot in B and C).

- Taking all the preceding considerations together, the only mechanism of those listed that meets all the requirements is the Box-Tukey transformation of components, E(b). Mechanism C does not allow variation in the influence of time and allows only the most extreme decline in the influence of cost.

- However, models that use 'mixed' mechanisms – that is, linear combinations of different functional forms, in particular mixed log and linear forms, can also meet all the requirements while not imposing the complications of a power transformation such as the Box-Tukey.

- It is better from the viewpoint of explaining behaviour and of justifying the model that is being used, that cost damping should be included in model estimation, not added afterwards in an attempt to obtain realistic elasticities but abandoning fit to the original data – for example, to observed trip lengths.

These tests and conclusions are explained in more detail in Chapter 2.

It is clear that cost damping is an intrinsic aspect of how people respond to cost and time in making travel choices. Therefore, calibrating models that exclude cost damping will simply obtain incorrect parameters – for example, for sensitivity. Applying cost damping subsequently cannot give a correct result, because the parameters used initially will not take account of the subsequent adjustment. The correct approach is to calibrate the models *with* cost damping in the first place. If parameters are being taken from other regions – e.g. via WebTAG – those imported parameters should be estimated with account taken of cost damping; for example, the parameters in WebTAG should be estimated taking account of the impact of cost damping.

## Conclusions and Recommendations

The conclusions and recommendations emerging from this work may be stated as follows.

**Cost Damping is used more widely in the UK than abroad**. In the UK, damping is primarily performed by scaling generalised time by a power of distance, sometimes also adjusting the value of time; a number of other mechanisms including log cost are also used. In other countries, log cost and Box-Tukey transformations are used; it is quite likely that other forms are used also but these have not been revealed by our limited search. Foreign governments take differing views to non-linear scaling such as cost damping, some (e.g. US) being suspicious, others (e.g. France) being apparently strongly in favour.

**Empirically, the evidence for cost damping is very strong**, given the nested logit framework used in most practical models. This evidence comes from

– modelling studies where the effect has been tested,

– the need to introduce cost damping to achieve realistic elasticities and (

- c) value-of-time studies, which show that damping operates more strongly on cost than on time.

**There is little empirical evidence that indicates that one cost damping mechanism should be preferred to another**, except that mechanisms which permit values of time to increase with trip length are to be preferred to those that do not.

**Microeconomic theory can be used to give a basis for discrete choice models** of the type used for travel demand modelling. However, no indications can be drawn from this theory concerning the propriety of cost damping.

**While microeconomic theory does not rule out the use of model parameters that vary with distance, there is a risk that the effects of time and money variables could be biased** if distance is included as an explicit variable. However, the methods for assessing these risks fall outside the scope of the current work. Moreover, the empirical evidence on the following issues does not rule out

- generalised cost parameters that vary with distance;

- non-linear functions of some components, such as log costs;

- linear combinations of functions of this nature.

**Other theoretical considerations than microeconomics can be used to support various forms of cost damping**. These include the following.

- The use of a distance variable is questionable, since there is no behavioural basis and consequently no case that the effect of distance will remain constant over time.

- The use of transformations of individual variables, rather than of the overall generalised time function, is to be preferred, as this allows the value of time to increase with trip length.

- Box-Tukey or power functions can be applied to the variables comprising generalised time or to the whole function; in either case these functions should have exponents less than 1 to give the required damping effect. The exponent used on time should be greater than that on cost to allow the value of time to increase with trip length.

- A *kilometrage test* has been devised, which requires that the total kilometrage driven should decrease when the marginal cost per kilometre increases uniformly. This test can be used to derive limits on the degree of cost damping that is acceptable:

  o power functions should have exponents not less than zero;

  o a pure log cost function does not give a change in the kilometrage when the marginal cost changes and therefore forms the limit of acceptability; however, the small addition to cost that is often made before the log is taken means that the function gives a small decrease in kilometrage when cost increases;

  o exponential functions are not acceptable;

- o 'kinks' in functions are in principle not acceptable, but this may not be important in practice.
- – 'Mixtures', i.e. linear combinations with positive coefficients, of functions that are acceptable with respect to these tests are themselves acceptable and these functions can be used to simplify practical estimation work, when power functions are awkward.

**When nested logit models are used, cost damping is part of our current best understanding of traveller behaviour** and it would be expected that it would be incorporated in models.

**Cost damping should be part of the initial set-up of a model to be estimated** and should not be introduced solely as a means of achieving realistic elasticities.

These recommendations represent the current state of knowledge as it has been assessed for the present study. It is to be hoped that further research and practical work will allow continual improvement to modelling Guidance and practice.

**International expertise**

As set out in the main text, the answers to Questions 2 and 3 posed by the Department required in part that a limited international survey was made to determine the mechanisms in use in other parts of the world and the extent to which they are used. This chapter reports the contacts that were made and what was learned.

The findings of this survey are summarised in the main text of the report, along with the findings of UK applications which are based on the work of WP1.

A limited search of the internet was made, but this proved to be largely fruitless. The survey is therefore based on our own work and that of contacts around the world.

## 1.1 The work of Marc Gaudry

The most prominent advocate of non-linear utility functions over the last 20 years or so has been Marc Gaudry, working in Canada, France, Sweden and Germany. He has consistently proposed and used Box-Cox transformations applied to individual variables as the basis for improving the explanation of traveller behaviour. He has developed software which can estimate models of this form.

Correspondence with him for this project revealed a number of papers, both published and posted on his institution's web-site (Agora Jules Dupuit, www.e-ajd.org). A number of these papers are reviewed in Chapter 2 of this report.

In email correspondence, Gaudry added a number of useful points. In particular, he tells us that "the French Ministry of Transport now *requires* Box-Cox tests in the Terms of Reference of its major calls" (emphasis added). He then listed five "principal questions" in the application of Box-Cox methods (paraphrased here):

1. difficulty of estimation, because of the form of the log likelihood function, together with the influence of scale on the results;
2. plausibility of the function curvature obtained;
3. impact of Box-Cox on the size and sign of coefficients;
4. impacts on forecasting;
5. consumer surplus calculations, which are made more difficult because of the changing marginal value of money.

This list of difficulties suggests that the widespread adoption of Box-Cox (or other power functions) would be challenging. However, Gaudry presents plenty of evidence to show

that functions of this type give a better explanation of behaviour, in terms of likelihood, and give significantly different results in forecasting, in both cases confirming UK and other foreign findings.

## 1.2    The work of RAND Europe

RAND Europe and its predecessor Hague Consulting Group have developed large-scale travel demand models by estimation from local data over a period of more than 25 years. Initially, estimations were made sequentially and with the assumption that destination choice was higher in the hierarchy than mode choice; poor results were often obtained in the initial mode choice stage. However, with the emergence of software capable of making simultaneous estimations in the early 1980's, it was found that much better estimates could be made – i.e. the addition of the destination dimension added useful variance – and that alternative structures could also be tested and sometimes performed better.

The work up to about 1985 did not employ cost damping. However, in a study conducted in that period, which later formed the basis for the Netherlands National Model, data was available from both home interview and intercept surveys (highway and train) and the range of tour lengths thus obtained allowed tests of log and linear forms, leading to the conclusion that the log form gave a much better fit to the data. The reasons for this were explored in a paper (Ben-Akiva *et al.*, 1987) which concluded that there were several possible explanations for non-linear utility functions, concluding that these alternative forms deserved to be tested.

Practical work by Hague Consulting Group and (from 2001) RAND Europe continued over the next 15 years, with work in The Netherlands (both national and regional models), Stockholm, Norway, Paris, Copenhagen and Sydney (Fox *et al.*, 2003). The majority of these systems featured log transformations of cost, which was shown to give better results than a linear cost formulation. The basis of these models was explored once again in a paper (Daly and Carrasco, 2009) which concluded that the performance of the log cost transformation could be equalled by models in which heteroskedasticity with respect to trip length was introduced and that this heteroskedasticity seemed to be best parametrised proportional to the cost of the journey.

Value-of-time studies in several countries by this group also consistently found that the value of time increased with trip length and that this seemed to be due to a declining sensitivity to cost.

The most significant recent work by RAND Europe has been done in the West Midlands (PRISM model, see www.prism-wm.com) and in Manchester (Fox *et al.*, 2008). The findings in those UK studies that cost sensitivity declines with increasing journey cost is clearly supported by a wide range of earlier foreign work.

The RAND Europe work is the most substantial body of evidence we have found for models estimated on local data supporting the need for cost damping.

## 1.3    Other foreign researchers

A number of other foreign researchers were also contacted for this study, though their relevant work is not on the scale of that of RAND Europe or Gaudry. The Zurich Eidgenössische Technische Hochschule (ETH) in Switzerland appears to have done the largest amount of work in this area.

**Kay Axhausen** at ETH in Zürich contributed a number of papers, which indicate improvements in models obtained by Box-Cox transformation. The ETH researchers are also (probably with others) responsible for the Swiss National Model, which uses Box-Tukey transformations to improve the fit of the models.

**Eric Petersen** at Cambridge Systematics in Chicago responded, also referring to work with his collaborator **Peter Vovsha** at Parsons Brinckerhoff in New York. They apply non-linear transformations of distance to destination and/or mode choice modelling. The use of these functions raises issues with the Federal overseers, but these difficulties have been overcome in their work.

**Mark Bradley**, independent consultant in California, responded:

> Non-linear cost and travel time functions are very rare in US transport planning, mainly because they are not accepted by the Federal funding agencies. In fact, LOS coefficients [i.e. the coefficients applied to level-of-service variables] are restricted to such a tight range and specification that they are usually constrained instead of estimated. I have often looked at detailed price functions in marketing SP [stated preference] applications, but those are outside of transport.

**Staffan Algers**, at Kungliga Tekniske Högskolan (KTH) in Stockholm, referred to work he had done with Marc Gaudry which compares non-linear transformations with segmentation (reviewed in Chapter 2). He also contributed a note on Swedish long-distance models, which showed that Box-Cox transformations, which include the log function as a special case (see Table 1) gave improvements comparable with those of allowing extensive segmentation or heteroskedasticity when applied to the 1984/85 National Travel Survey. When applied to the 1994–2000 Survey, substantial improvements were also obtained from Box-Cox transformations, but problems were encountered in finding global optima.

**David Hensher**, at the University of Sydney, said that he was working on non-linear utility functions, particularly involving thresholds, but that his work was not yet ready. Much of his recent work is concerned with *changes* in time and cost relative to current levels, mainly in Stated Choice experiments – very interesting but not for the present study.

**Juan de Dios Ortúzar**, at the Catholic University of Chile in Santiago, referred to his work on thresholds, which (like Hensher's work) mainly arises in the context of *changes* from current levels. He also told us of a debate in Santiago concerning the appropriate level of exponents in Box-Cox transformations and correspondence with Gaudry indicating that values greater than 1 (i.e. cost amplification) might be appropriate.

**The relationship of cost sensitivity and trip length**

## 2.1 Introduction

A recurring feature of recent travel demand modelling studies is the focus on variation of the sensitivity of travellers to changes in travel cost or travel time. Specifically, it has consistently been found that sensitivity appears to decline as trip length increases. A wide range of theories have been advanced to explain the effect and a similarly wide range of functions have been tried in practice to incorporate this variation into the models used in planning studies. The term 'cost damping' has frequently been used to describe the various mechanisms used to implement the required variation in scaling.

The questions that naturally arise in these circumstances are whether such variation is justified theoretically or empirically and what are the implications of the choice of particular functional forms. A paper presented at the 2006 European Transport Conference (Daly and Hyman, 2006) showed that pure microeconomic theory was quite weak in discriminating between potential forms. However, there are other considerations, based on empirical or forecasting considerations, that can be utilised to indicate which functional forms can be used to represent time or cost sensitivity.

The empirical considerations are based primarily on disaggregate, mainly cross-sectional, model estimation results, since it is difficult to imagine that aggregate data would be adequate to distinguish differential impacts by trip length. However, in model estimation, the role of trip length variables, such as the road distance, is open to different interpretations. It is possible that such variables represent a 'pure' separation effect, for example that they are reflecting traveller's knowledge about the existence of opportunities. However, it is also possible that a distance variable that appears to be significant in model estimation is actually reflecting a travel time or cost for which distance is a proxy.

Forecasting considerations are also relevant, as the assumption of particular functional forms can be important in determining how models behave when forecasting for scenarios substantially different from the present. Even though the economic theory is not very strong, it does rule out some models that have been used in practice, for example, where utility increases with increasing cost(!). Sharper tests can be formulated, such as requiring that the total distance driven would decrease with an increase in fuel price, which are not strictly economic in nature but which may be considered as formulations of common

sense. These tests also allow us to identify unsatisfactory features of other models that have been used in practice.

### 2.1.1 Random utility maximisation

It will be assumed that the behaviour of travellers results from their maximisation of utility. We recognise that the concept of utility presents a number of difficulties and that there exist concepts, e.g. state-dependent preferences, that go beyond a strict interpretation of utility maximisation. However, the models used in even quite advanced travel demand modelling practice do not require a more sophisticated basis than utility maximisation, while the framework of reference given by utility allows a clear interpretation of many of the most relevant effects arising in the discussion. For practical implementation for a population of varying tastes, Random Utility Models (RUM) represent the application of the utility approach.

Two points may be made immediately, starting from the RUM concept.

First, the utility that appears in the models is *indirect* utility, i.e. the utility that can be obtained subject to the traveller's budgets (of both money and time) and the alternatives available. This means that non-linearity apparent in the models does not necessarily relate to ideas such as satiation, which would apply to the direct utility, but may also, possibly more importantly, relate to budget constraints and to the opportunities available. That is, a 'bend' in the function is not necessarily caused by a change in the marginal utility of time or cost *per se*, but may be due to the increasing consumption of scarce resources, which are then not available for other purchases and activities. Different travellers may then have different budgets and hence different non-linearities. In general, budget effects might reasonably be expected to cause indirect utility functions to become steeper as time and cost increase. An effect of *increasing* sensitivity with trip length can be called 'cost amplification.

Second, any interpretation we are able to make of the measured part of utility appearing in models has to be seen in the context of the random terms of those models. It is quite possible that the random terms do not have constant variance (i.e. heteroskedasticity exists) and this effect may easily be confused with non-linearity in the measured components of utility. It may not be possible to identify whether a given finding is 'really' due to heteroskedasticity or to non-linearity. Importantly, if utility variance increases for longer trips, as would be the case under many plausible hypotheses, this increased variance would imply a lower model sensitivity (Daly, 2008), which could be represented in a simple model by one of the cost damping mechanisms. In general, if not incorporated in the model, heteroskedastic effects would cause indirect utility functions to appear less steep as time and cost increase.

However, as we shall see, a proper application of microeconomic theory indicates that the description of the function as having a 'bend' is entirely arbitrary and depends entirely on the scale on which utility is measured. That is, the shape of the function is not determined by theoretical considerations but by practical and empirical ones.

### 2.1.2 Structure of the Chapter

The following section of the chapter describes the range of cost damping mechanisms that have been considered in practical models, while the third section reviews some of the quite

extensive literature that has addressed this issue. Section 2.4 discusses the ambiguous role of the distance variable which has been used in a number of these models. Section 2.5 presents a number of tests that can be applied to models to determine whether they implement cost damping in a way that is likely to be satisfactory in practice.

The Summary of the chapter puts forward a restricted range of model forms which would satisfy the tests that are discussed and shows that these models can indeed represent the range of varying cost sensitivities that have been found in practice.

## 2.2 Potential Forms of Cost Damping

The general context in which we find cost damping is in models of the RUM form, in which the attractiveness of each alternative is represented by a utility function, part of which is measured and part of which is random. In practice nearly all practical models take the form of logit or tree logit (i.e. nested logit without cross-nesting). Cost damping is then understood as the use of non-linear functions in the measured part of the utility functions; it is generally expected that the non-linearity has the effect that the marginal disutility of time and/or cost declines as trip length increases.

A review of literature and practice in the UK and more widely has identified that at least nine mechanisms for cost damping have been applied. The nine explicitly identified are listed in Table 1.

**Table 1: Cost damping mechanisms found in practice**

|   | Definition | Formula | Notes |
|---|---|---|---|
| A | damping utility by a power of distance | $g = (d/k)^{-\alpha}.(t + c/v)$ | commonly used in UK practice |
| B | increasing the value of time by a function of distance | $g = t + c/v(d)$ | also used in UK practice |
| C | log cost formulation | $g = t + \beta \log(c + \delta)$ | see section 1.2 |
| D | addition of constants that depend on distance | $g = t + c/v + k(d)$ | used in UK NTM |
| E | Box-Tukey or Box-Cox transformations<br><br>a<br><br>b | $g^{(\alpha,\delta)} = \dfrac{(g+\delta)^{\alpha} - 1}{\alpha}$ when $\alpha \neq 0$; $g^{(\alpha,\delta)} = \log(g+\delta)$ when $\alpha = 0$.<br><br>$x^{(\alpha,\delta)} = \dfrac{(x+\delta)^{\alpha} - 1}{\alpha}$ when $\alpha \neq 0$; $x^{(\alpha,\delta)} = \log(x+\delta)$ when $\alpha = 0$. | see section 1.1 |
| F | origin–destination–specific scaling | $g = \lambda_{OD}(t + c/v)$ | in standard UK software, DIADEM |
| G | power function of utility | $g^* = g^{0.9}$ | used in one UK study |
| H | detailed segmentation by purpose | $g^* = \lambda_p g$ | used in some UK studies |
| I | heteroskedasticity with respect to trip length | $g$ unchanged in principle | not known to be used in practice |

General notation for Table 1:

$g$      refers to the overall disutility, generalised cost or generalised time;

$d$      is the trip or tour length;

$t, c$      are the trip or tour time and cost, respectively;

$v$      is the value of time;

$\lambda > 0, \; \alpha > 0$ are parameters that need to be estimated or supplied.


Specific notes and notation for mechanisms in Table 1:

A. $k$ is generally an average trip length, e.g. 30 km and sometimes the transformation is applied only for trips longer than $k$. $\alpha$ is often 0.5, but has been 1 or greater than 2 in some cases.

B. The issue is what the specific function $v$ should be – often a function like $v_0 (d/k)^{\alpha}$ is used, where $v_0$ is a base value of time.

C. $\delta$ is a small constant (e.g. 1p or €0.01) and $\beta$ is a coefficient to be estimated. A recent variant, suggested by Hyman (2007), is to include both log and linear cost functions in the utility, an approach that can be called a 'mixed' function.

D. NTM is the UK National Transport Model; $k(d)$ represent distance-specific constants.

E. Box-Tukey is a simple generalisation of the Box-Cox transformation, introducing $\delta > 0$, which allows $g$ to be zero even when $\alpha \leq 0$. Negative $g$ is not covered.

The two variants shown here apply the Box-Tukey transformation either (a) to the entire utility (or generalised time) or (b) to selected variables $x$, such as cost or time. Note that transformation C is a special case of Box-Tukey applied to cost with $\alpha = 0$. It may be noted that when $(g + \delta) = 1$ the transformation gives the value 0 and has derivative 1, irrespective of $\alpha$, thus effectively normalising the output variable.

F. This is not very different from mechanism A.

G. This can be seen as a form of Box-Cox transformation applied to the whole function, i.e. a special case of E(a).

H. The intention is to separate regular (short) and specialised (long) trips, which can then be modelled with separate $\lambda$'s.

The model scale effectively diminishes, as a function of increasing time and/or cost, as the variance of the error term increases.

It is possible to use mechanisms in combination. In particular, mechanisms A and B are used together in some UK studies.

The key issues here are whether the cost damping is fixed (i.e. does not change with policy) and whether it applies to the entire generalised time function or just to components. The various mechanisms above can be classified as in Table 2.

**Table 2 Classification of Cost-Damping Mechanisms**

| operating on: | Entire function | Components |
|---|---|---|
| Fixed function 'differential scaling' | A, F | B |
| Depends on costs 'transformed variables' | E(a), G | C, E(b) |

Mechanisms D, H and I should be considered as alternatives which offer some of the advantages or address the same problems as cost damping but are not cost damping in themselves, although they are sometimes seen as damping functions. It suffices to

investigate mechanisms A, B, E(a) and E(b) as the other mechanisms shown in the table (C, F and G) are special cases of these.

The 'differential scaling' mechanisms A and B generally depend on the use of the distance variable, which is largely unchanged in forecasting (although route changes may affect distance). As such, they suppose that distance has some unspecified impact on behaviour which will remain constant through time. A detailed discussion of the use of the distance variable is given in Section 4.

## 2.3 Evidence for Cost Damping in Theory and Practice

A number of papers have appeared in the literature giving evidence of significant improvements in model form and accuracy offered by cost damping. It should be noted, however, that many large-scale model applications are not reported in the international literature.

### 2.3.1 Theoretical papers

Blayac and Causse (2001) apply a general time–money optimisation framework of the type proposed by de Serpa. They argue that functions non-linear in time and cost can be permitted in such a framework and illustrate this in a small empirical study. However, the direction of curvature of the non-linear functions is not clear in their paper.

Orro *et al.* (2005) investigate the feasibility of using Box-Cox transformations along with taste heterogeneity. Interestingly, they assert that power parameters greater than 1 are "contrary to theory and should not be used". They find that Box-Cox transformations are quite feasible to apply and give an important improvement in the model when non-linearity is present in reality. They also note that "nonlinearity in the influence of the variable may be confused with heterogeneity in tastes".

Jara-Díaz has written extensively on the appropriate functions for modelling time and cost trading within overall budgets. The time available for this study precludes a full review of his work, but a flavour is given in Jara-Díaz (2006) in which indirect utility functions are indicated that may be non-linear in either time or cost or both. Indications are not given of the sign of the second derivative.

Daly and Hyman (2006) investigate the conditions that microeconomic theory might place on the form of travel demand models, which can be formulated to be rigorously consistent with utility theory using the RUM approach. They argue that relatively weak conditions on the form of the distribution of the error term are generally acceptable, but that models should not be restricted by either

- the assumption of uniform marginal disutility of expenditure or

- the assumption that tastes do not vary in the population.

Analysing the form of utility functions, they find that the curvatures of the direct or indirect utility functions are of little ultimate impact in the demand function. Even the requirement for quasi-convexity does not impose a restriction on the sign of the second derivative. Microeconomic theory would not preclude any form of cost damping that retained a negative first derivative of indirect utility with respect to price.

Hyman (2007) investigates the elasticities of travel demand, expressed as kilometrage, for a series of very simple models. When the main change in demand is either pure mode switching or pure destination switching, elasticity may be proportional to cost in a linear model, or independent of cost in a log cost model. He suggests the use of a model with *both* linear and log cost (called a 'Tanner' function in a subsequent note), suggesting that this could give more reasonable intermediate results than either simple log or simple linear formulations.

### 2.3.2 Forecasting studies

Vrtic *et al.* (2007) report the development of a national model for Switzerland. The focus of the paper is very much on the Erzeugung–Verteilung–Aufteilung (EVA) algorithm for equilibrating a demand model with assignment and on the activity-pair approach for generating trips, but for present purposes the interest lies in their use of Box-Tukey transformations applied to the generalised time functions, which were estimated using maximum likelihood methods but transformed subsequently because "the introduction of the marginal constraints in the EVA approach requires adjusting some of the variables to obtain the observed distance distributions". In the Box-Tukey transformations, the shift parameter was consistently set to 1 and the power function "adjusted by hand". All of the values of the power parameter lie between 0 and 1, i.e. they show a cost damping effect.

In the re-estimation of the UK National Transport Model, Gunn and Petersen (2007) developed models to explain travellers' choices of mode, trip length and destination area type. They showed in preliminary models that a formulation with the log of the cost of each alternative gave a very much better fit to the data than a formulation with a linear cost function. They went on to estimate very detailed models (more than 100 parameters) with linear cost and distance–band–specific constants. These constants indicated a downward 'bending' of the total disutility function, presumably corresponding to the effect detected by the log cost function in the simpler models, but (unlike the log cost function) the constants do not give a damping of the effect of cost *changes* at longer distances.

Mandel *et al.* (1994; the 1997 paper reports the same models) use the Box-Cox transformation applied to separate variables, which substantially improves the fit of the model to intercity data. They applied the same $\alpha$ for the separate variables, as significant improvements were not obtained with separate parameters. A value of $\alpha$ of 0.521 was obtained in the best model, which represents a moderate damping effect; because both time and cost are damped, the value of time would not increase consistently with distance in this model.

Gaudry has written extensively (in addition to the Mandel papers) on non-linearity in travel demand models, largely exploiting the Box-Cox transformation. A recent paper (Gaudry, 2008) summarises some of his earlier work and is a convenient reference. Work using disaggregate Revealed Preference data in the Québec–Windsor corridor shows standard Box-Cox models with parameters generally in the range 0–1, though more complicated models can have values outside that range. Further Stated Preference data from the same corridor gave similar results. Tests on Swedish data showed that Box-Cox transformations and detailed segmentation could be substitutes for each other (a justification for transformations of type H in Table 1) and that heteroskedasticity was also a possible substitute; more detail of this study can be found in Algers and Gaudry (2001).

Gaudry (2008) goes on to make the important point that the function that is taken for utility components can be very important in forecasting: in the case of interest for that paper – high-speed rail – a non-linear response can mean higher forecasts for long trips but lower forecasts for the more numerous shorter trips and hence a more accurate, lower, total revenue forecast when non-linearity is taken into account.

A further study by Gaudry *et al.* (2007) for trans-Pyrenean freight, concludes that Box-Cox transformations give a far superior explanation of observed behaviour than linear logit. The elasticities obtained from the two models also appear substantially different, but the elasticities appear to be calculated only at the mean value of the variables, so that this point is of little value.

Daly and Carrasco (2009), continuing the work of Ben-Akiva *et al.* (1987), cite large-scale modelling studies in several countries (Fox *et al.*, 2003, RAND Europe, 2004) as showing the superiority of the formulation of cost as a logarithmic variable over the linear formulation. Daly and Carrasco set out a list of hypotheses which might explain the observation of increasing value of time with trip length, which is a key feature of the log cost formulation. They note that several of the most plausible hypotheses rely on heteroskedasticity and proceed to test for heteroskedasticity in two large-scale datasets from Paris and Sydney. They find that heteroskedasticity is significant as an influence on behaviour and gives an explanation at least as good as that given by non-linear functions applied to cost and/or time.

Work forecasting railway passenger demand in the UK (e.g. Wardman and Whelan, 2004a, summarised, 2004b) typically uses a constant–elasticity model. Indeed, in the work referenced, variation of elasticity with distance was tested and found not to be important. This formulation implies a utility formulation with a log transformation, i.e. a considerable degree of cost damping.

Informal contacts with international researchers indicate that academic analysts are convinced of the significance of non-linear effects in the influence of time and cost on behaviour, but that government authorities often insist on the use of linear models for forecasting and, particularly, for the appraisal of transport proposals. It appears that practice is lagging well behind research on this point.

### 2.3.3 Studies of the value of time

While the main interest of the current study is in forecasting models, some information is also available from value-of-time studies, some of which have considered the relationship of the marginal disutility of time and cost for trips of different length.

In addition to the large-scale studies discussed in the previous section, Daly and Carrasco (2009) also cite several western European value-of-time studies indicating an increase in the value of time with trip length. Again, hypotheses indicating heteroskedasticity are advanced to explain this apparent effect and an analysis is made of Dutch data. This again showed significant heteroskedasticity, so that the observed increase in values of time could be caused by self-selection, so that travellers with high values of time consistently chose more distant destinations.

Bates and Whelan (2002) find, somewhat to their surprise, that the marginal impacts of both time and cost diminish in value-of-time Stated Preference data collected in the UK.

The cost effect predominated, which enabled them to conclude that the tendency of values of time to increase with trip length was primarily due to a reduced sensitivity to cost changes for more expensive journeys. More complicated issues arise when considering cost and time changes relative to the current situation, but these issues are not immediately relevant for the current study.

In a recent study, aimed more particularly at modelling demand for the M6 Toll motorway, but also giving results on the value of time, Wardman *et al.* (2007) found that values of time increased with trip duration. Unfortunately for our present purposes, the analysis primarily considered only increases in the time coefficient, rather than the more plausible and more widely tested decreases in the cost coefficient, so they give little insight into cost damping; rather time amplification would be implied by their main model. However, they also report a separate test indicating minor reductions in the toll coefficient with increases in journey duration.

### 2.3.4   Summary of the evidence

There is no reason in microeconomic or other theory not to accept nonlinearities in indirect utility functions such as those used in travel choice models to represent cost damping. The finding of Orro *et al.* that cost amplification is not acceptable is neither justified in their paper nor supported by other research we have found.

The possible interchangeability of heteroskedasticity and non-linearity has a long pedigree. Gaudry and Dagenais (1979) discuss this issue and present methods for dealing with *both* effects in modelling – in that case regression modelling. They also refer to still earlier work by Zarembka (1974, not available for review) in this area. The work of Daly and Carrasco (2009) confirms this interchangeability in the context of travel choice models.

In models intended for forecasting, there is a consistent and quite strong finding that the marginal disutility of cost declines with trip length. The marginal disutility of time may also decline, but the evidence is less clear.

From studies of the value of travel time, there is a consistent finding that the value of time increases with trip length. Partly this is due to a preference for longer trips among travellers with high incomes, but even in models where income is a covariate the same finding is made within income groups. These findings imply a non-linear damping of the cost coefficient and/or a non-linear amplification of the time coefficient, although there is evidence that the former is more likely.

We conclude that the existence of cost damping as an influence on behaviour is consistent with economic and behavioural theory and is strongly supported by the evidence of large-scale forecasting studies and studies of the value of travel time.

### 2.3.5   Reference dependence

An important advance in thinking about choice behaviour is given by Prospect Theory, in particular the work of Tversky and Kahnemann. This work draws attention to the importance of reference conditions in consumers' formulations of value. Key features are

- the asymmetry of gains and losses,

- diminishing marginal utility of changes (both gains and losses) and

— risk aversion with respect to gains but risk-seeking with respect to losses.

Behaviour of this nature is reasonably well established with respect to monetary gains and losses, but less so with respect to time gains and losses. Tapley *et al.* (2006) give an introduction for value-of-time estimation.

However, for the purposes of the present discussion, the impact of these features is less important, as we are generally considering long-term forecasts, in which travellers are subject to a series of changes in differing directions, as well as to changes in their personal circumstances (home, job etc.). Divergences from a utility explanation of behaviour will occur in both directions and in a wide range of ways for different travellers. It does not seem important to try to incorporate reference-dependent effects in models of the type being considered.

## 2.4 The Role of Distance

A number of cost damping mechanisms used in the UK employ a distance variable to implement the reduction in model sensitivity, although this mechanism has not been found in other countries. The functional form is, with increasing distance, to scale down the entire utility function and/or to adjust the value of time upwards, in either case reducing the marginal impact of cost for longer distances (mechanisms A and B in Table 1).

It seems reasonable to ask whether the procedures using distance are used simply as mechanical means of adjusting the function to obtain model performance that meets *a priori* expectations better or whether it is believed that distance *per se* has some explicit impact on traveller behaviour, distinct from the impacts of time and cost.

In the former case, where a mechanical adjustment is made that is not understood, there is an urgent need for research to clarify the role of this adjustment. In particular, since the assumption has to be made that the parameters associated with distance will remain constant for forecasting, it is essential that the effect of distance can be associated with behavioural mechanisms that are indeed constant. Without validating this assumption, the use of models containing such variables is seriously undermined.

In the latter case, where a behavioural basis is claimed for the distance variable, distinct from time and cost, hypotheses can be advanced to explain the role of distance:

— that it serves to correct measurement errors in time and/or cost;

— that it represents an information effect, connected with the radius of travellers' other trips, (e.g. they pass opportunities they would otherwise not be aware of), the distribution areas for Yellow Pages, etc.;

— that there is some psychic influence of being far from home, which reduces the marginal impact of time and cost, perhaps because of heteroskedasticity[5];

---

[5] But heteroskedasticity defined as a function of distance is less plausible than heteroskedasticity defined as a function of time and/or cost, which arises naturally if there are heterogeneous preferences for marginal time and/or cost changes in the population.

    –    that there are other effects we do not know about.

In each case, to justify the adjustment to the model, the behavioural impact of distance must be to diminish the marginal impact of cost, or of time and cost, but in such a way that the effect cannot be represented by cost itself.

If distance does have some impact on behaviour, a concern is that we are often modelling travellers' choices of destination, so that cost damping involving distance will imply that the marginal impact of cost differs between destinations at different distances. While varying marginal cost utilities are acceptable in theory, they imply a degree of complication beyond the simple analyses usually used in practice.

Whichever basis is used to justify the inclusion of distance-based effects, these mechanisms can be adopted only after the analyst has conducted careful tests using non-linear functions of time and cost. The use of distance to explain behaviour in the model raises the suspicion that the impacts of time and/or cost may be biased and this suspicion must be dispelled for the model to be fully credible.

The use of distance–band–specific constants in the model is similar in its impact to the use of a continuous distance variable, except that there may be arbitrary jumps in the impact of the distance effect between one band and another. Again, the suspicion exists that cost and/or time impacts may be distorted by the use of these variables.

## 2.5    Tests on Cost Damping Mechanisms

In Section 2.2 four broad groups of cost damping mechanisms were set out in Table 2. This section considers how each of these four groups can be tested in each of a series of ways.

### 2.5.1   Classical microeconomic theory

The clearest requirement of classical microeconomic theory, effectively the 'Law of Demand', is that the utility should increase (generalised time should decrease) when the cost of an alternative decreases. To this may be added a similar requirement with respect to time, although some authors think that the utility of an alternative may remain constant under some circumstances when the time required for it changes. These requirements place clear conditions on the coefficients in the (indirect) utility function.

There seem to be few other requirements that can clearly be stated from classical theory, in particular there is no requirement that the conditional indirect utility function should be linear. It is shown in the main part of this report that, if that utility function is differentiable with respect to the own cost of an alternative, a strictly monotonic transformation cannot change the sign of the first derivative, but *may well change the sign of the second derivative*. Since any such transformation may always be made, without losing consistency with utility theory, it is not possible for microeconomic theory to give information about the curvature of the utility function. A requirement of quasi-convexity is not relevant in a discrete choice context, since the costs of other alternatives do not enter the conditional utility functions, so that a requirement for quasi-convexity reduces to a requirement for monotonicity, which we would have in any case.

An issue that has not been considered in practical transport planning contexts is that if the indirect utility function is non-linear – or indeed if it is linear with different slopes for different alternatives – this implies that simple methods of appraisal such as the 'rule of a half' may not be applicable. This issue requires further attention (see, for example, de Jong *et al.*, 2008).

### 2.5.2 Practical Tests

In practice, it has been found in many studies that cost damping is required to fit the base data and/or to give reasonable elasticity values. We are not aware of any studies where cost amplification has been implemented, although the impact of cost is sometimes increased *proportionately* to obtain increased elasticity.

An important issue is to achieve a balance between local evidence and generally accepted expectations of model performance. If a model is estimated to give an optimal explanation of behaviour in a specific area, and subsequently adjusted (e.g. by applying cost damping to meet generally expected values of elasticity) then of course the model will no longer give an optimal explanation of local behaviour and may indeed be far from it. Thus if cost damping is believed to be necessary, it should be introduced at the model estimation stage and not subsequently.

Similarly, when models are set up by importing parameters from other areas, as is common in UK practice when nationally-defined parameters are used, then it is necessary that those parameters are defined for the model forms, i.e. with or without cost damping, that are being used.

### 2.5.3 The Kilometrage Test

The kilometrage test rejects models that predict an increase in the total kilometrage driven when the cost of driving is increased. The test is not a requirement of economic theory and is primarily a practical test of the reasonableness of a model.

This apparently simple and commonsense test turns out to be quite useful in indicating which models are acceptable and which are not. Details are given in the Appendix – in summary the findings are that, to achieve cost damping and pass the test:

- piece-wise linear models are *unacceptable*, in the strictest interpretation of the kilometrage test;

- power functions and Box-Tukey transformations with power parameters between 0 and 1 inclusive, i.e. including log and linear functions, are *acceptable*;

- negative power functions are *unacceptable*;

- if value of time is expressed as a negative power function of distance[6], and this function is used to reduce the impact of cost increasing with distance, the power *must* lie between 0 and 1 inclusive;

- exponential functions are *unacceptable*.

- positive mixtures of acceptable functions are *acceptable* and may be quite useful.

---

[6]     See also Section 2.4.

The kilometrage test applies to all four types of damping transformation in Table 2. Strictly, the test is an approximation and exactly correct only when the per-kilometre cost of driving is constant. Moreover, the exact conditions have been worked out for mechanisms A and B in Table 2 only when the cost damping factor is viewed as a constant fixed by trip length, rather than as one of the 'moving parts' of the model. But the test applies equally whether damping is applied to the entire generalised time function or to the cost component within the generalised time.

In principle the kilometrage test can be applied equally well for a combination of damping mechanisms, e.g. mechanisms A and B which are sometimes applied together. In that case it seems that the power functions applied must both be positive and *sum to 1 at most*. The analysis of combinations of Box-Tukey functions is quite complicated and it appears that a non-analytic approach, e.g. a graphical method, might be preferable.

### 2.5.4 Change of Elasticity with Trip Length

Investigating the variation in elasticity with respect to trip length was not part of the objectives of this study, but information is available from the analyses that have been done and this section contains that information.

The elasticity of demand for an alternative $j$ with respect to the cost of driving per kilometre $f$ can be calculated by

$$\eta_{jf} = \frac{\partial p_j}{\partial f} \cdot \frac{f}{p_j}$$

where $p_j$ is the probability of choosing that alternative.

If demand is predicted by a cost-damped model using a generalised time formulation

$$g = t + \frac{fd + r}{v}, \qquad \text{i.e.} \quad \frac{\partial g}{\partial f} = \frac{d}{v}$$

with    $t$ being the time components of generalised time;

   $v$ the (fixed) value of time;

   $d$ the trip length, so that $fd$ is the cost of driving; and

   $r$ all the other cost components, such as parking and tolls;

then, by the chain rule for differentiation,

$$\eta_{jf} = \frac{\partial p_j}{\partial g^{(\alpha)}} \cdot \frac{\partial g^{(\alpha)}}{\partial g} \cdot \frac{d_j}{v} \cdot \frac{f}{p_j}$$

where    $g^{(\alpha)}$ represents the damped generalised time.

If, for example, cost damping was applied using Case E(a), a Box-Tukey transformation applied to the whole generalised time, we obtain

$$\frac{\partial g^{(\alpha)}}{\partial g} = (g + \delta)^{\alpha - 1}$$

where $\delta$ is the (small) positive offset applied in the transformation; and

$0 \le \alpha \le 1$ to meet the theoretical requirements of the situation.

If demand is predicted by a model of the logit form, then considering only the response to the 'own' cost

$$\frac{\partial p_j}{\partial g_j^{(\alpha)}} = \lambda p_j (1 - p_j)$$

and applying the two last equations to the elasticity gives an 'own' elasticity

$$\eta_{jf}^{own} = \lambda(1 - p_j).(g_j + \delta)^{\alpha-1}.\frac{d_j f}{v} = \frac{\lambda f}{v}.(1 - p_j).(t_j + \frac{fd_j + r_j}{v} + \delta)^{\alpha-1} d_j$$

Considering the three components of the right hand side of this equation:

- $\frac{\lambda f}{v}$ is constant;

- $(1 - p_j)$ will always be close to 1, since there are many alternatives, and will generally tend to increase with trip length, as more distant destinations are more rarely chosen;

- $(t_j + \frac{fd_j + r_j}{v} + \delta)^{\alpha-1} d_j$ will also increase with trip length, as $\alpha \ge 0$, so that distance appears with a net non-negative power; the presence of $r_j/v$ and $\delta$ in the bracketed term dilutes the negative impact; the presence of $t_j$ in the bracketed term will also dilute the negative impact of that term, as although time increases with distance it will usually not do so proportionately as faster roads can be used for longer trips.

Thus it is reasonable to expect that $\eta_{jf}^{own}$ increases with distance.

If the model is of logit form, the cross-effects will simply be proportional, since this is a well-known property of the logit model. Nesting with all the destinations at the same level of the hierarchy will not affect the property of proportionality between destinations.

In the Appendix it is shown that similar conditions affecting $\alpha$ apply to the four generic Cases of cost damping, so that similar conclusions would apply to the dependence of elasticity on trip length.

In conclusion, for the types of models generally used in transport planning, the elasticity of response to fuel price changes must increase with trip length. This point is not directly relevant to the topic of central interest, but is a useful by-product of the analysis that has been done.

## 2.6    Conclusions and Recommendations

Cost damping has been found to give an improved explanation of behaviour in both large-scale forecasting studies and value-of-time studies. A range of mechanisms have been used,

which can be classified into four broad groups on the basis of whether (a) they use a 'fixed' damping based on distance or apply a functional transformation and (b) whether the damping applies to the entire utility or generalised time or just to the monetary cost component.

Evidence from value-of-time studies and from large-scale forecasting studies is that the value of time increases with trip length, which would imply that cost damping needs to be applied more strongly, or only, to the monetary cost component. While heteroskedasticity can also explain this effect, and the limited evidence suggests that the explanation is at least as good as obtained by cost damping, heteroskedastic models are not yet practical for large-scale application.

The use of distance to modify values of time or the scale of generalised time, risks obscuring the impact of cost and/or time in the model, since distance, cost and time are so highly correlated. Where possible, the impact of cost damping should be brought into models through transformations to the policy variables (i.e. time and cost) to improve the defensibility of the models.

These considerations would lead to a recommendation of transformations of individual variables without using distance as an auxiliary, i.e. transformations of the type E(b) in Table 2. Transformations could be applied to both time and cost, but a stronger transformation should be applied to cost than to time.

A power function between 0 and 1 should be used, where it is understood that the value 0 implies using the log function. Mixtures of acceptable functions can be used, but 'kinks' in the function should be avoided. Mixtures (e.g. log and linear functions) can be useful in practical modelling.

In any case, cost damping should be included as part of the model development process, as applying damping subsequently undermines any claim that the model is based on evidence on behaviour.

# CHAPTER 3 The impact of microeconomic theory

Microeconomics is the study of consumer choices under monetary constraint. It presents a theory of behaviour which can be applied to the travellers in a transport system and therefore it is not unreasonable to enquire whether travel demand models are consistent with microeconomic theory. In particular, the role of travel demand models in providing input to an economic appraisal procedure, which is an essential part of the DfT approach, suggests that a microeconomic approach to demand modelling would be natural.

The classical microeconomic theory of consumer demand is presented in textbooks such as Deaton and Muellbauer (1980, see Ch. 2). This theory explains the choice of a consumer choosing a bundle of quantities of continuously divisible goods, each of which has a unit price, subject to a budget constraint. Within this theory, the concept of indirect utility is clearly defined: indirect utility is the maximum utility that can be obtained with given income and goods prices and is a function of income and prices.

For travel demand modelling, the most appropriate representation of many aspects is that of discrete choice. Within this framework, joint choices of mode, destination, time of travel and even of travel frequency can be represented by models of varying degrees of sophistication. Almost all travel demand models can be set in the discrete choice framework, with the (negative) generalised time function of the travel demand model corresponding with the attractiveness or 'utility' function used in discrete choice models.

Discrete choice can be presented within the framework of microeconomics: some of the issues are discussed by Small and Rosen (1981) and McFadden (1981). A discrete choice model is seen as reflecting part of a consumer's behaviour, while the consumption of other goods is modelled using classical methods. The resources available for the present study do not permit a thorough investigation to be made of the full correspondence between the two bodies of theory. However, it is clear that there is a broad agreement in the principles. For example, the first three Axioms posited by Deaton and Muellbauer for their presentation of classical utility theory (reflexivity, completeness and transitivity) are also necessary for discrete choice modelling, while their fourth Axiom (continuity) is not necessary when, as in discrete choice modelling, the choice set is finite.

The correspondence of theory between microeconomics and discrete choice is in any case sufficiently strong to allow an identity to be drawn between the negative generalised time or attractiveness function used in discrete choice models and the kernel of the indirect utility function of classical microeconomics, *conditional* on the discrete choice that is made. In this context, the 'kernel' of the indirect utility means that the discrete choice considers only those aspects of utility that vary between the alternatives considered in that choice.

The indirect utility function conditional on each discrete choice alternative would naturally have among its arguments all the variables of quality pertaining to that alternative, such as the travel time[7] and comfort levels, the price of the alternative and, in principle, the traveller's income. The prices of other discrete choice alternatives would not appear, though general price levels in the economy would in principle be present.

Resources available for the present study also preclude a full investigation of the way in which income and general prices should appear in the indirect utility function. Relevant issues include the following.

- The impact of general prices in the economy should impact in the same way on the prices of discrete choice alternatives and on income. In travel demand models, prices and incomes are usually expressed in base-year values and this appears to be adequate to protect our models from 'money illusion' effects.
- Homogeneity of specific prices and incomes can be assured by including in the indirect utility function only functions of $(c/Y)$, where c is the price of the alternative and Y is the traveller's income. We note that if the function is the logarithm, this becomes $(\log c - \log Y)$, at which point $\log Y$ is constant across the alternatives and falls out of the kernel of the utility function.
- There is an argument that the proper form in which c and Y should appear in the indirect utility function is as a function of $(Y-c)$ only, since the only impact of the cost of this alternative on utility is through the remaining income (also called 'free cash') available to spend on other goods. We note that if the function is $\log(Y-c)$, then this can be approximated[8] as $(\log Y - c/Y)$, with $\log Y$ dropping out as before and leaving a homogenous form.
- There appears to be some tension between the requirements of the two previous points and it is not clear that they can be resolved simultaneously with functions sufficiently flexible to represent real-world behaviour. The 'polar form' due to Gorman (1961, not available for review) may be relevant here.

In most cases, however, this discussion is entirely moot, as travel demand models do not often contain any information about income.

Returning to the Axioms of classical demand modelling in Deaton and Muellbauer, Axiom 5 requires non-satiation, in at least one good. This Axiom is required so that the consumer's budget is exhausted. In discrete choice modelling it would imply that the indirect utility of an alternative was always decreasing (not just non-increasing) with respect to the price of the alternative. This monotonicity is a relevant and useful property to require in travel demand models.

Deaton and Muellbauer continue with convexity: "one further 'axiom', that we shall use from time to time but that will not be generally assumed to hold". Convexity of preferences implies quasi-convexity of the utility function with respect to prices. However, in the discrete choice context, with only the price of the specific alternative appearing in

---

[7]   Some presentations consider time also to be a constrained resource with a budget similar in its operation to the financial constraint, but these issues go beyond the scope of the present work.

[8]   $\log(Y-c) = \log Y + \log(1-c/Y)$ and if $c<<Y$ this can be approximated by the first term of the Taylor expansion: $\log(1-x) = x + x^2/2 + \dots$ .

the indirect utility, convexity implies no more than monotonicity and this is already implied by the non-satiation Axiom discussed in the preceding paragraph.

If the indirect utility function $V$ is differentiable with respect to cost $c$, monotonicity obviously implies that the first derivative is always negative:

$$\frac{\partial V}{\partial c} < 0 .$$

It is useful to note that this property survives a strictly monotonic transformation $g$, i.e.

$$\text{if } W = g(V), \text{ then } \frac{\partial W}{\partial c} = g'\frac{\partial V}{\partial c} < 0,$$

because $g' > 0$ follows from the assumed monotonicity of $g$.

Such a strictly monotonic transformation can always be applied to utility, which is fundamentally an ordinal rather than a cardinal concept. If we consider the *second* derivative of transformed utility $W$, however, we find that its sign properties are not necessarily consistent with those of $V$:

$$\frac{\partial^2 W}{\partial c^2} = g''\frac{\partial V}{\partial c} + g'\frac{\partial^2 V}{\partial c^2} .$$

Here we see that the second term does preserve the sign of the equivalent derivative of $V$, but the first term may be positive or negative and can easily dominate the second term, as no constraint is placed on the magnitude or sign of $g''$.

What this simple analysis shows is that economic theory cannot give any indication for the sign of the second derivative of indirect utility with respect to price. More generally, for example for non-differentiable functions, we can get no indication of the direction of curvature. Since it is exactly the curvature of the utility function that defines cost damping, we see that economic theory is not able to give any indication for the existence or validity of cost damping.

Effectively, what happens with a monotonic transformation of utility such as $g$ is that the *scale* of utility is changed. In practical models, however, we have no such freedom to change the utility scale, as this is defined by the form of the model. The issue of cost damping, therefore, is to determine what is the requirement for the second derivative of the utility function with respect to price, *given the specific form of the model*. From the discussion above, there is no requirement from economic theory for the function to have any specific curvature, but there is clear empirical evidence that, given the specification of scale implicit in a logit model, that the impact of cost declines as trip lengths increase, i.e. that the second derivative of indirect utility with respect to price is positive. This is however an empirical rather than a theoretical requirement.

# References

Algers, S. and Gaudry, M. (2001) Are non-linearity and segmentation substitutes, complements or unrelated in logit models?  The Swedish intercity passenger case, Agora Jules Dupuit, Publication AJD-28, www.e-ajd.org.

Bates, J. and Whelan, G. (2002), private communication.

Ben-Akiva, M., Daly, A. J. and Gunn, H. F. (1987), Destination Choice Models: Design and Appraisal, presented to *P.T.R.C. Summer Annual Meeting*.

Blayac, T. and Causse, A. (2001), Value of travel time: a theoretical legitimization of some nonlinear representative in discrete choice models, *Trans. Res.* B, 35, pp. 391-400.

Daly, A. (2008), Elasticity, model scale and error, presented to European Transport Conference, Noordwijkerhout.

Daly, A. and Carrasco, J. (2009), The influence of trip length on marginal time and money values,  in *The Expanding Sphere of Travel Behaviour Research*: Selected papers from the proceedings of the 11th Conference on Travel Behaviour Research, Kitamura, R. *et al.* (eds.), Emerald Books.

Daly, A. and Hyman, G. (2006), The rules of the game, presented to European Transport Conference, Strasbourg.

Daly, A., Fox, J. and Tuinenga, J. G. (2005) Pivot-point procedures in practical travel demand forecasting, presented at conference of European Regional Science Association; to be submitted to Traffic Engineering & Control.

Deaton, A. and Muellbauer, J. (1980) *Economics and consumer behaviour*, Cambridge University Press.

Fox, J., Rohr, C., Daly, A., Patruni, B., Hyman, G. (2008) Manchester motorway box: post survey research of induced traffic effects, presented to European Transport Conference, Noordwijkerhout, Netherlands.

Fox, J., Daly, A. and Gunn, H. (2003), Review of RAND Europe's Travel Demand Model Systems, RAND Monograph Report 1694, http://www.rand.org/pubs/monograph_reports/MR1694/.

Gaudry, M. (2008) Non linear logit modelling developments and high speed rail profitability, Agora Jules Dupuit, Publication AJD-127, www.e-ajd.org.

Gaudry, M. and Dagenais, M. (1979), Heteroskedasticity and the use of Box-Cox transformations, Economics Letters, 2, pp. 225-229.

Gaudry, M., Briand, A., Paulmyer, I. and Tran, C-L. (2007) Choix modal transpyrenéen ferroviaire, intermodal et routier: un model logit universel de forme Box-Cox généralisée, Agora Jules Dupuit, Publication AJD-88, www.e-ajd.org.

Gorman, W. M. (1961) On a class of preference fields, Metroeconomica, 13, August, pp. 53-56.

Gunn, H., and Petersen, E. (2007) Alternative strategies to estimating large-scale travel demand models: the UK national transport model, presented to European Transport Conference, Noordwijkerhout.

Hyman, G. (2007), Three propositions relating to the realism of car traffic elasticities, Department for Transport, unpublished.

Jara-Díaz, S. (2006) The complete goods-activities framework to model the assignment, consumption and discrete choices, presented to European Transport Conference, Strasbourg.

Jong, G. de, Daly, A., Ibáñez, J.N., Batley, R. and Bok, M. de (2008) Welfare measures from discrete choice models in the presence of income effect, presented to European Transport Conference, Noordwijkerhout.

Mandel, B., Gaudry, M. and Rothengatter, W. (1994), Linear or nonlinear utility functions in logit models? The impact on German high-speed rail demand forecasts, *Trans. Res.* 28B, pp. 91-101.

Mandel, B., Gaudry, M. and Rothengatter, W. (1997), A disaggregate Box-Cox Logit mode choice model of intercity passenger travel in Germany and its implications for high-speed rail demand forecasts, *Ann. Reg. Sci.*, 31 pp. 99-120.

McFadden, D. (1981) Econometric models of probabilistic choice. In: *Structural Analysis of Discrete Data: With Econometric Applications*, Manski, C. and McFadden, D., eds., The MIT Press, Cambridge, Massachusetts.

Orro, A., Novales, M., and Benitez, F. G. (2005), Nonlinearity and taste heterogeneity influence on discrete choice models, presented to European Transport Conference, Strasbourg.

RAND Europe (2004), Tour based mode destination modelling, Available from www.prism-wm.com.

Small, K.A. and Rosen, H.S. (1981) Applied welfare economics with discrete choice models, *Econometrica*, 49, 1, pp. 105-130.

Tapley, N., Wardman, M. and Whelan, G. (2006) Nonlinearities in discrete choice attribute valuations, presented to European Transport Conference, Strasbourg.

Vrtic, M., Fröhlich, P., Schüssler, N., Axhausen, K. W., Lohse, D., Schiller, C and Teichert, H. (2007) Two-dimensionally constrained disaggregate trip generation, distribution and mode choice model: Theory and application for a Swiss national model, Trans. Res. A, 41, pp. 857-873.

Wardman, M. and Whelan, G. (2004a) Generalised journey time reformulation, Institute for Transport Studies Report.

Wardman, M. and Whelan, G. (2004b) Estimating the effects of service quality changes on the demand for rail travel, presented to European Transport Conference, Strasbourg.

Wardman, M., Whelan, G., Vaughan, B., Murphy, P., and Hyman, G. (2007), Modelling the demand for toll roads in the UK, presented to European Transport Conferences, Noordwiijkerhout.

Zarembka, P. (1974), Transformation of variables in econometrics, in Zarembka, P. (ed.), Frontiers in Econometrics, Ch. 3.

# Appendix: The Kilometrage Test

This Appendix presents the detailed mathematics on which the conclusions in the main text with respect to the kilometrage test are based.

For a given travel segment (i.e. purpose, origin and person type), define the total car kilometrage as $D$

$$D = G.\sum_j d_j p_j$$

where    $G$  is the total demand (number of tours or trips) for this segment;

$d_j$  is the car distance from the origin to alternative $j$  (0 if the alternative relates to a  mode other than car driver);

$p_j$  is the probability of choosing alternative $j$  for this segment.

It would be required by commonsense that under all circumstances this *kilometrage would not increase with an increasing marginal cost of driving*: this is the *kilometrage test*[9]. This is not an absolute requirement of economic theory, since it can happen that consumption of goods increases with increasing price, but it would be very strange if it were so in this case. That is, assuming $p$  is differentiable[10], we formulate the kilometrage test as

$$0 \geq \frac{\partial D}{\partial f} = G\sum_j d_j \frac{\partial p_j}{\partial f}$$

where    $f$  is the marginal cost/km. of driving.

To ensure that the kilometrage test is always passed, this should apply for all combinations of destinations and other alternatives. In particular, suppose there are just two alternatives, with car distances $d$  and $d + \delta$  respectively ($\delta > 0$), then we obtain

---

[9]    Fuel consumption functions that vary with speed (and therefore are dependent to some extent on trip length) mean that the discussion has to be conducted in terms of marginal driving cost, rather than fuel cost, but the assumption that fuel consumption per kilometre is constant is reasonable, at least as an approximation and for discussion purposes. Hence the notation $f$ .

[10]    It is probably reasonable to expect that the demand functions $p$  are differentiable with respect to marginal driving cost, i.e. there are no threshold effects or kinks. However, kinks with respect to distance are more common.

$$0 \geq G\left\{d\left(-\frac{\partial p}{\partial f}\right) + (d + \delta)\left(\frac{\partial p}{\partial f}\right)\right\} = G\delta\frac{\partial p}{\partial f}$$

where    $p$  is the probability of choosing the alternative at  $d + \delta$  (the other alternative has of course probability  $(1 - p)$ ).

That is, we require the demand derivative  $\partial p / \partial f$  for the more distant destination to be negative (or at most zero), since  $G$  and  $\delta$  are positive.

Suppose we are modelling demand using some kind of generalised time function  $g$ , as in the four generic cases of cost damping set out in Table 2 of Chapter 2, so that in the binary case the demand for the two alternatives depends only on the difference in generalised time, e.g. in some kind of logit model. We then obtain

$$\frac{\partial p}{\partial f} = -K\lambda\frac{\partial(\Delta g)}{\partial f} \leq 0$$

where    $K$  is a positive function which depends on the nature of the demand model (e.g. in a logit model it would be equal to  $p(1 - p)$ );

$\lambda > 0$  is a sensitivity parameter and is not dependent on  $f$ ;

$\Delta g$  is the difference in generalised time  $g$  between the alternatives (more distant minus less distant).

The negative sign here expresses the fact that the demand for the more distant alternative will decline when the generalised time difference increases. The kilometrage test is then equivalent to the requirement that  $\partial(\Delta g) / \partial f$  should always be non-negative. *A marginal cost increase must not reduce the difference in generalised time between car driver alternatives.*

To be more specific, we can write the generalised time for car driver alternatives as a function of time, distance-related cost and other costs

$$g = g(t, fd, r)$$

where    $t$  represents all the time elements of travel and

$r$  represents the non-driving cost elements, such as tolls and parking.

Then if  $c = fd$  is the driving cost, we can differentiate  $g$  with respect to  $d$ [11]

$$g' = \frac{\partial g}{\partial d} = f\frac{\partial g}{\partial c} = \frac{f}{d}\frac{\partial g}{\partial f}$$

It seems reasonable to expect that generalised time should be differentiable with respect to distance, except perhaps at a limited number of points or 'kinks'; this differential should never be negative.

---

[11]        In this differentiation, we focus on the 'cost-bearing' impact of distance, i.e. the term  $fd$ , and the impacts of distance that modify behaviour indirectly, i.e. through value of time or through an overall scaling on generalised time, are not considered at this stage.

We can now express the generalised time difference as a function of distance

$$\Delta g = g(d + \delta) - g(d)$$

Hence

$$\frac{\partial(\Delta g)}{\partial f} = \frac{1}{f}\{(d + \delta)g'(d + \delta) - dg'(d)\} = \frac{1}{f}\{d(g'(d + \delta)) - g'(d)) + \delta g'(d + \delta)\}$$

For $\partial(\Delta g)/\partial f$ to be non-negative, we must have

$$-\frac{1}{\delta}(g'(d + \delta) - g'(d)) \le \frac{1}{d}g'(d + \delta) \qquad (1)$$

Considering this inequality, the left-hand side represents the rate of decline in the slope of the generalised time from $d$ to $d + \delta$. We see that to pass the kilometrage test *the slope of the cost function may decline, but the rate of decline cannot be more than proportionate to distance*. The extent of cost damping must be limited.

If the generalised time is twice differentiable with respect to distance, then we can take the limit of (1) as $\delta \to 0$ and obtain the second derivative, which must be negative to achieve cost damping[12]

$$-g''(d) \le \frac{1}{d}g'(d) \qquad (2)$$

However, in some models the generalised time function is specified to have one form up to a specified distance and another form thereafter. Usually, the function would be continuous, but the slope might not be continuous, i.e. there might be a 'kink' in the curve. In this case, if we choose $d$ and $d + \delta$ to lie on opposite sides of the kink, we see from equation (1) that, when $\delta$ is very small, either the slopes of the function must be the same (no kink) or that the bend in the generalised time function must be upwards, which of course is not what is required for cost damping. *Kinks are not a theoretically acceptable means of achieving cost damping*. While the practical implications of this finding may be limited, the theoretical restriction on the functional forms that can be used needs to be noted.

Assuming now that the function $g$ is twice differentiable with respect to distance, we can consider the implications of the kilometrage test in the form of equation (2) for the four generic forms of cost damping set out in Table 2 (see Chapter 2).

Case A: $g = \left(\frac{d}{k}\right)^{-\alpha}\left(t + \frac{fd + r}{v}\right)$, with constant value of time $v$

Case B: $g = t + \frac{fd + r}{v(d)}$

---

[12] At this stage we need to consider *all* the impacts of distance, both cost-bearing and modifying.

In these cases, the cost damping modifiers, $\left(\dfrac{d}{k}\right)^{-\alpha}$ in Case A and $\dfrac{1}{v(d)}$ in Case B, are usually not considered to be cost-bearing or 'movable' parts of the model, e.g. it would not be expected that a change in driving cost would impact on behaviour directly through these terms, rather that they would modify the impact of the direct effects. Comments on this interpretation of the behavioural impact of distance are given in Section 2.4. If we accept the validity of this approach, then to obtain the terms for equation (1) we need to calculate the first derivatives considering only the directly cost-bearing distance components $fd$, obtaining

Case A: $g' = \left(\dfrac{d}{k}\right)^{-\alpha}\dfrac{f}{v}$

Case B: $g' = \dfrac{f}{v(d)}$

For the second differentiation, moving from equation (1) to equation (2), however, the form of the function with respect to distance needs to be taken into account and here we do need to consider the modifying effect of the cost damping terms.

In Case A, we obtain

$$g'' = -\alpha\left(\dfrac{d}{k}\right)^{-\alpha-1}\dfrac{f}{kv} = -\alpha\dfrac{g'}{d}$$

which satisfies inequality (2) providing $\alpha \leq 1$ and offers cost damping providing $\alpha > 0$.

In Case B, we get

$$g'' = -\dfrac{fv'(d)}{\left(v(d)\right)^2} = -g'\cdot\dfrac{v'(d)}{v(d)}$$

so that inequality (2) is satisfied providing $v'(d) \leq v(d)/d$. Usually $v$ would be given as a power function and the discussion under cases E below shows that the power in this function needs to lie between 0 and 1 inclusive.

Thus in Cases A and B, the impact of distance is *directly* through a linear relationship with driving cost and *indirectly* through a cost damping which must have a power between –1 and 0. The net impact is therefore that cost depends on distance with a power relationship between 0 and 1.

If the view is taken that the damping factors are cost-bearing, then the calculations for Cases A and B become considerably more complicated and have not been worked out for this study.

Case E(a):      $g = \left(t + \dfrac{fd + r}{v}\right)^{(\alpha,\delta)}$ , a Box-Tukey transform of the generalised time

Case E(b):    $g = t + \left( \dfrac{fd + r}{v} \right)^{(\alpha, \delta)}$ , a Box-Tukey transform of the money cost

In both cases E(a) and E(b) we find

$$g' = \left( \frac{f}{v} \right)(g* + \delta)^{\alpha - 1} \quad \text{and} \quad g'' = \left( \frac{f}{v} \right)^2 (\alpha - 1)(g* + \delta)^{\alpha - 2}$$

where $g*$ is the argument of the Box-Tukey function. Hence

$$\frac{dg''}{g'} = \frac{fd(\alpha - 1)}{v(g* + \delta)}$$

In the slightly simpler case E(b), this gives

$$\frac{dg''}{g'} = \frac{fd(\alpha - 1)}{fd + r + \delta v}$$

which will always be negative since $r$ and $\delta v$ are always non-negative and $\alpha < 1$ to obtain cost damping rather than cost amplification. If $\alpha < 0$, then for high values of $d$ we may find the fraction less than –1, whatever the values of $r$ and $\delta v$, so we require $\alpha \geq 0$.

In case E(a), we obtain

$$\frac{dg''}{g'} = \frac{fd(\alpha - 1)}{fd + vt + r + \delta v}$$

and the argument follows the same lines as for E(b), except that the presence of the term $vt$ in the denominator may mean that satisfactory results can be obtained for a greater range of distance; however we still require $\alpha \geq 0$ to be certain of compliance with the kilometrage test.

We conclude that for both the Box-Tukey formulations it is necessary that $0 \leq \alpha < 1$ to meet the kilometrage test and offer cost damping.

Three further points are also useful, concerning exponential functions 'mixed' functions and the presence of constants in the model.

*Exponential functions* $g = \lambda \exp \mu(t + fd + r)$

To give disutility increasing with increasing cost we must have $\lambda \mu > 0$. Then

$$g' = g.\mu f$$

$$g'' = g.(\mu f)^2$$

The kilometrage test (2) requires

$$- g.(\mu f)^2 \leq g.\frac{\mu f}{d}$$

If $\lambda$ and (therefore) $g$ are positive, the kilometrage test is always satisfied, but in this case $g''$ is positive and there is no cost damping. But if $\lambda$ and (therefore) $g$ and $\mu$ are

negative, the kilometrage test is equivalent to (dividing through by the *positive* number $g\mu f$ )

$$-\mu f \leq \frac{1}{d}, \qquad \text{i.e.} \, d \leq \frac{1}{-\mu f}$$

That is, there is a maximum trip cost length after which the model is not acceptable, again a disquieting feature but which may not be serious in specific instances.

*Mixtures of functions*, i.e. $g = \lambda_1 g_1 + \lambda_2 g_2$ , with both $\lambda_1, \lambda_2$ positive

Because of the form of tests (1) and (2), if $g_1$ and $g_2$ separately pass the test then $g$ will pass. In an important example, a mixture of log and linear functions (Box-Tukey with $\alpha = 0$ and $\alpha = 1$ respectively) is acceptable without further testing, providing the mixing factors are both positive. This can be useful in practical modelling since it allows linear-in-parameters specifications to be estimated. However, if either or both $g_1$ and $g_2$ fail at some points, then $g$ may or may not pass and little can be said. Special tests would be required in any specific case.

*Constants in the model*

The kilometrage test, whether in form (1) or form (2), depends only on the differentials of the generalised time function, not on the absolute value of generalised time. Therefore, the conclusions relating to the test are unaffected by the presence or otherwise of constants in the model, such as 'alternative-specific constants', 'mode handicaps etc..

Of particular importance in this context is that models are very often applied either 'incrementally' or by pivoting from a base matrix. It can be shown (e.g. Daly *et al.*, 2005) that 'incremental logit' and pivoting by taking the ratio of synthetic estimates (and then correcting the row totals), which are the most common methods used in the UK, are both equivalent to the use of a full matrix of K factors. Since K factors can be represented as constants in the model, the use of any of these methods for a model that passed the kilometrage test would leave forecasts which also passed the test.

When synthetic forecasts for two or more segments, e.g. person types or travel purposes, are summed before factoring, a similar argument can be applied. In this case a K factor exists that could be applied to all of the segments to obtain the exact fit to the base matrix and the addition of this constant would not affect whether or not the model passes the kilometrage test