Issues in the use of choice modelling to support transport policy

Andrew Daly
RAND Europe, ITS Leeds and CTS Stockholm

Seminar: Advances in Discrete Choice Models

University of Cergy-Pontoise
18 December 2015

Introduction

limited aspects of transport policy have been supported by behavioural analysis for some time (Dupuit 1844)

the application of choice and utility was extended in the 3rd quarter of the 20th century (Wardrop, 1954; Beesley, 1965; Wilson, 1969)

but most transport policy was supported only by models based on physical analogues (gravity model, entropy maximisation..)

- travellers as continuously divisible incompressible homogenous commodity

so that "...there still does not exist a solid foundation in behavioural theory for demand forecasting practices" (McFadden, 1974)

The BART project gave an opportunity to develop and apply a demand modelling framework

BART project

McFadden and his team:

- set out a consistent framework for forecasting based on a plausible behavioural hypothesis (utility maximisation)
- collected data and estimated models
- applied these models to forecast BART patronage
- showed that the forecasts were rather close to actual turnout demand

- demonstrated a proof of concept for models based on disaggregate utility maximisation

this demonstration was crucial to the funding of further work in using discrete choice to predict travel demand

Utility-based modelling was also being developed in other areas

contemporary pieces of work included

- Ben-Akiva thesis, MIT, completed 1972
- work by LGORU, England, 1970-73
- and several other studies in US and UK
  - linking to UK government work using generalised cost

- by the mid 70s we were able to take on broader studies
  - e.g. in England, Huddersfield Bus Study (1974-76), Value-of-Time study (1975) etc.
  - also with important associated theoretical developments
- and in the US, the MTC study

Two key model developments in 70s and 80s

MTC model developed by Ben-Akiva & associates

- it aimed to cover all travel in the metropolitan area
- set up a comprehensive sequence of decisions
- the basis was to view travel as trips
- followed BART in estimating model from local data using maximum likelihood

➢the decision sequence idea led into activity-based models

Netherlands National Model

- based on earlier Dutch studies, themselves drawing strongly from MTC but viewing travel as tours
- and focussed more on European issues
  - e.g. licence holding, to get more precise mode choice
- the model has now been used continuously for 30 years (of course with multiple updates and improvements)

➢the effectiveness of this model led to series of national and conurbation models in Europe and elsewhere

These models required some further advances to work well

- the socio-economic detail in the models is one of their key advantages
  - but population forecasts need to be made in compatible detail
  - different methods are available for this task and research is still needed to determine which works better in given circumstances
  - forecasting income presents particular challenges
- spatial detail is another important development
  - but this increased the computational burden at a time when computing was still very expensive
  - sampling of alternatives McFadden 1978, extended by Guevara/Ben-Akiva 2013
  - there are still issues in using these advances for planning work
- non-linear functions were found to be necessary
  - to some extent interchangeable with heteroskedastic modelling
  - but speed of execution demanded logit-based models

The rational basis meant that the models could be extended

- beyond the classical 4 stages...
  - frequency, mode, destination, route
- ...to include at least time period choice (5 stages)
  - also car ownership, toll roads, pt passes, sub-mode choice...
  - these extensions are often crucial to policy appraisal (e.g. road user charging)

- how much behaviour needs to be modelled?
  - activity-based models are widely used in the USA
- when is this a good idea?
  - in practical work, the benefit needs to be shown convincingly to the funding agency
  - so if European analysts want to develop these models, they need to prepare evidence of benefits for European funders
    - some progress is being made, but we're not there yet

WTP analyses could be made on the same theoretical basis

- this comes of course from the utility paradigm
- WTP methods, based on stated choice, are widely used to estimate the value of travel time
  - which in turn is a key input to transport policy appraisal
- there are however some differences in practice...
- governments impose simplifications on VOT
  - to keep it simple
  - to impose ‘equity’ across regions, modes and income groups

- it’s difficult (impossible?) to devise realistic consumer surplus for models with alternative behavioural paradigms

Governments are convinced by these models

- tour-based models are currently under development in numerous countries, regions, agglomerations and cities
  - also for the whole of Europe (though with data problems!)
- and activity-based models in numerous cities in the US, Israel and Denmark

- funding is justified by the usefulness of the models in supporting policy

- there are small groups in several countries who are capable of doing the work

..but there are still several issues

e.g. two issues from my current work..

- can we estimate models using informal data along with specific RP or SP records?
  - we reject models when unacceptable properties are found
    e.g. inappropriate parameter ratios
  - how can this be formalised or at least used with justification
- how can we handle interactions of choice models with assignment systems?
  - assignments give effective treatment of route choice under congestion
  - but are generally poor with respect to interpersonal variation

1. Likelihood is not enough!

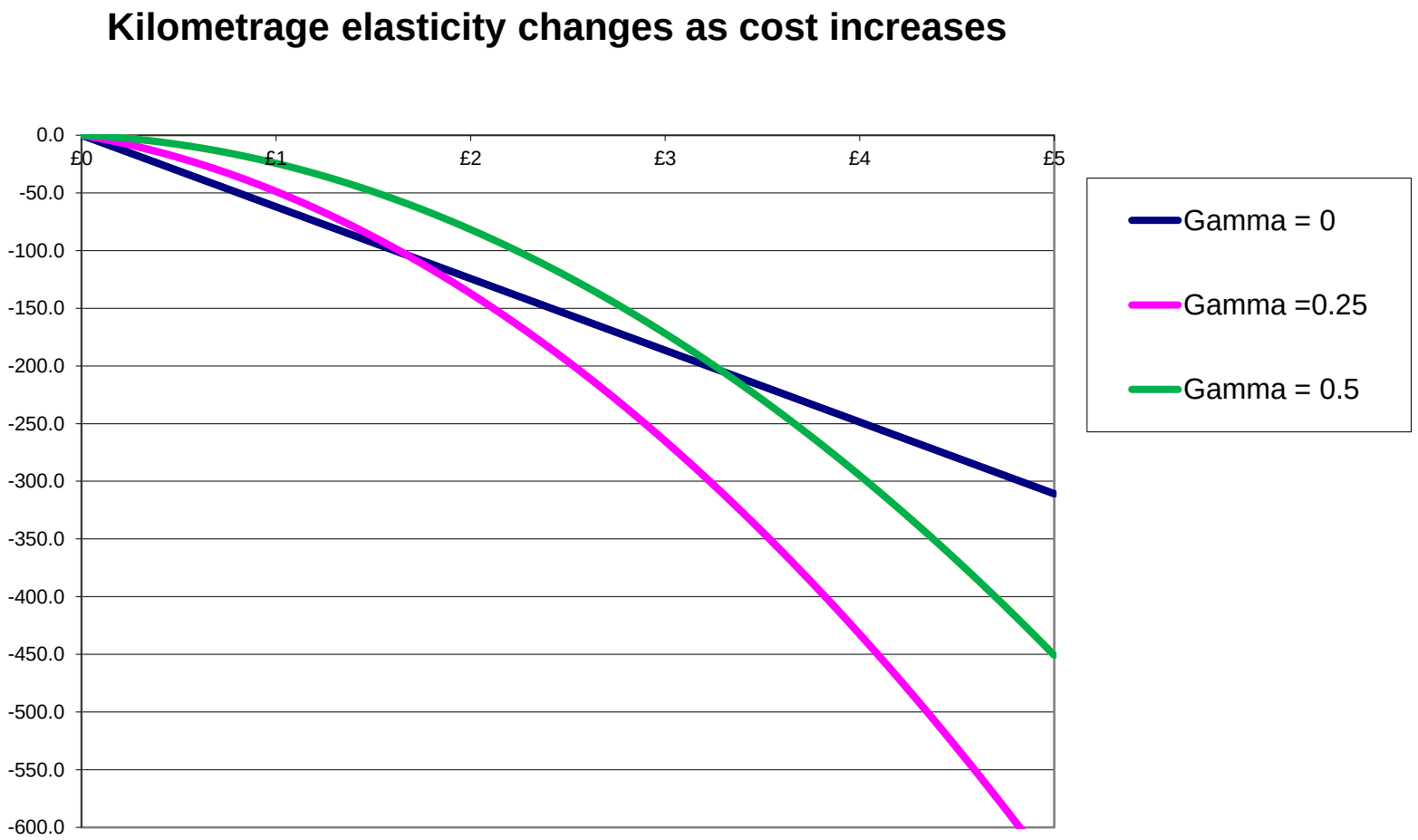
- models can be estimated to maximise fit to local data
  - by maximum likelihood
- but funding agencies expect good output performance
  - in the UK, this is a requirement to access national funds
- and local data is not always ideal
  - sometimes quite poor!
- this suggests the use of background information
  e.g. VOT and elasticity from national/international sources

- example of a large-scale model estimated by RAND Europe for West Midlands County in England
  - Birmingham and surrounding area

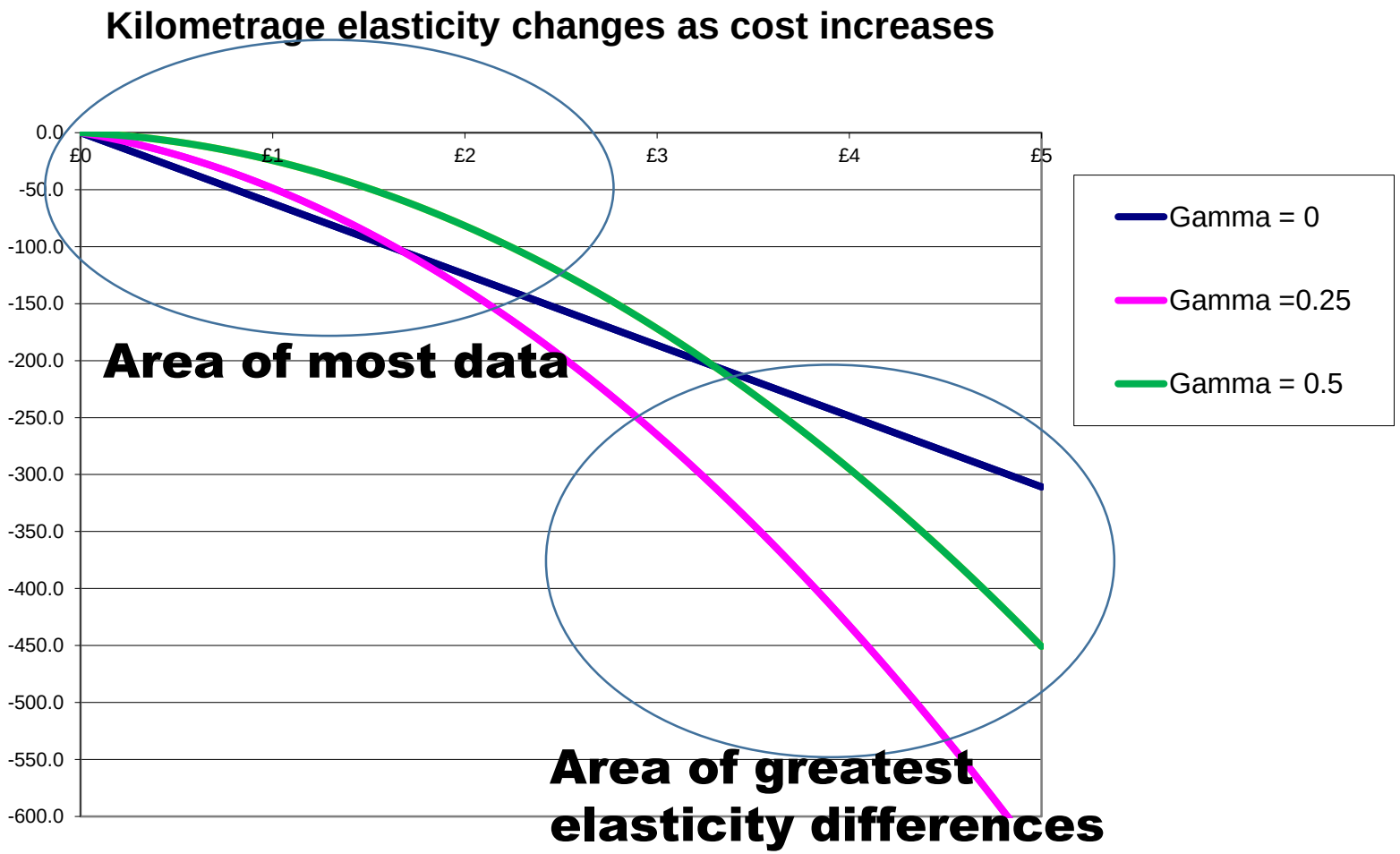
The West Midlands (PRISM) saga

- PRISM was estimated based on local data (of moderate quality)
  - with non-linear functions when justified by that data
- but the elasticity of kms driven with respect to fuel cost was not acceptable to DfT
  - we got –0.18 and it should have been –0.3000
- this led to a long (and animated) debate
  - there were some reasons why local behaviour could be inelastic
  - but the evidence for a higher national value was quite strong
  - we were able to persuade DfT that there should be more flexibility, so current advice requires elasticity –0.25 to –0.35
- but we realised we were too rigid in our approach
  - elasticity was not very well determined
  - and why should we not include other information sources?

Most data is for short trips, but long trips are more important for elasticity



Most data is for short trips, but long trips are more important for elasticity



Resolution of the PRISM saga

- DfT near-approval of the model was the key to multiple applications
  - infrastructure plans, both roads and metro
  - major highway management project
  - city centre planning
  - redevelopment strategy
  - etc.
- £ tens of millions of national money mobilised for West Midlands projects

- clear benefit to local area, but we need to be clear that our approach is professionally justified
- can the use of professional wisdom be formalised?

2. Can we combine assignment and choice models?

- there is substantial investment in assignment algorithms and software
  - predicting route choice in large networks with congestion
  - multi-modal for public transport
  - iteration is needed, so efficiency is essential
- the problem is how to use these models in combination with choice models of (e.g.) mode and destination
- issues concern the calculation of choice probability and of utility

- consistency can be achieved in a sampling application, but we are not yet (if ever) ready to abandon our software
  - agent-based applications are attractive, but the investment in aggregate assignment software and training is huge

Issues are more apparent in public transport models

- ..but they also exist for highway models

- the problem is essentially that some decisions are taken by the traveller and some by the system
  - and some that are joint efforts!
- the traveller decides, as in a stochastic model
  - preferences for sub-mode
    (train vs. bus, high-speed vs. classical rail,..)
  - timing preference and whether to learn timetable
- the system decides, as in an assignment approach
  - timing of services at same/alternative stations
  - geographical alignment of services

‘Who decides’ affects the choice function and resulting utility

- if traveller decides, we usually need a continuous probability model

  ...and a logsum-type utility measure
- if the system decides, we may well need an all-or-nothing assignment

  ...and an average utility
- of course logsum > average
  (sometimes logsum >> average !)
- so it makes a difference which approach is taken

- the difference can be large, particularly in long-distance models
  - because the scale parameter in the model is small

HS2 is the proposal for high-speed rail from London to N. England

Cost is around £60 bn
Benefit is uncertain, both for travellers and environment

Is it a new train route or a new mode?

- affects the degree to which heterogeneity influences benefit

But the focus shifted to economic benefit

- ultimately most HS2 demand will come from existing train users
- with a substantial element of newly generated traffic
  - depends on projected income growth
- impact on the air and car markets is small
  - the air market is very small and the car market is very diverse
  - so little environmental benefit to set against the construction impacts

- the main issue is economic impacts
- with time gains also important
- both depend on projected income growth

Concluding

- the relationship between transport analysis and choice modelling has been fruitful
  - transport has generated funding for choice modelling
  - choice modelling has provided effective tools for transport analysis
- the relationship has been productive over 40 years
- there have been repeated advances over this period
  - but not all of the problems have been solved
- and implementation lags far behind research

..and looking forward

- the issues I have outlined are not the only ones
- e.g. we need to know how the model should be applied
  - preparing inputs that are reliable and/or allow for alternative scenarios
- it would be good to get more university involvement in solving the problems
  - very limited in Europe, seems that US is better
  - clients also have a role here, but they are resistant to enlightenment
- one concern for academics is the challenge of publication
- another is that developing practical models is hard work

My message

- there are important transport planning issues that require demand forecasts and discrete choice is an excellent way to provide these
- in the past, leading academics have contributed to development
- but this is not currently happening in Europe, despite the need for further developments
  - there are interesting developments, e.g. in Singapore, which we could perhaps exploit
- there is an opportunity here and some urgent problems to solve