

Introduction

Choice modelling is becoming an indispensable tool in many sectors of economic activity, from the marketing of clothing, through the design of neighbourhoods to forecasting demand for public transport in the developing world. But the approaches and techniques that researchers in different sectors are developing are increasingly converging. For this reason, multi-sector conferences offer the research community the opportunity to present and compare methodology and for users of choice modelling to learn what is available and what can be done. To meet these needs, the second International Choice Modelling Conference was organised by the Institute for Transport Studies at the University of Leeds and was held at Oulton Hall near Leeds from 4 to 6 July 2011.

The conference brought together leading researchers and practitioners from across the many different areas in which choice modelling is a key technique for understanding behaviour and evaluating policy. This included many of the delegates who had attended the inaugural International Choice Modelling Conference held in Harrogate in Spring 2009. The presence of numerous new young delegates illustrated the growing international popularity of choice modelling as a topic of research.

The diversity of the field was reflected in presentations by both academics and practitioners, coming from five continents and various different fields, with a similar mix in non-speaking delegates. The papers presented at the conference gave a broad overview of the state of the art, encompassing a number of presentations by leading academics in the field, and of the state of practice, including works by people prominent in the application of the latest techniques to practical issues such as predicting demand or evaluating policy options.

This edited volume contains a selection of the best work presented at the conference, giving an overview of theoretical as well as applied progress in the field. The first two papers in the book correspond to keynote addresses made to the conference by researchers of very long standing in the field. To facilitate the rapid interchange and reporting of ideas the conference included four workshop sessions, whose reports form the following four chapters of the book, thus giving insight into

four important and active research areas. The remaining seven chapters comprise key papers presented at the conference, focussing on a diverse set of topics.

The first chapter is by Professor Jordan Louviere from the Centre for the Study of Choice at the University of Technology, Sydney, and focusses on the estimation of individual level models. The vast majority of choice models are estimated at a sample population level and thus give a representation of the choice behaviour of groups of respondents, albeit while potentially allowing for heterogeneity across decision makers, using one of a number of different approaches. The chapter by Louviere highlights the potential benefits of estimation at the individual level, i.e. the use of respondent specific models, showing how this gives greater insights into behavioural diversity and more flexibility in representing such diversity.

A key focus of Louviere's chapter is on the data requirements of such models, where he looks at length at different types of best worst scaling data. Louviere also highlights that individual level models are not new and in fact go back to early work by psychologists in the 1920s, an interest that continues in behavioural economics and mathematical psychology to this day. Louviere goes on to show how relatively simple modelling approaches can be used to analyse choices at the individual level, and also discusses issues relating to differences across respondents in error variances, and the stability of preferences over time.

The wide set of possible uses of these models is highlighted in a number of case studies, looking at topics as diverse as the choice of wine for a dinner party, residents' preference for different public transport improvements in Sydney, emission trading plans, and cross-country flights. This broad set of topics is also a reflection of the inherent widespread nature of real world uses for choice modelling techniques.

The second chapter of the book, by Professor Andrew Daly of the Institute for Transport Studies at the University of Leeds, describes work from the 1960s onwards that analysed travel demand and the valuation of travel time from the perspective of a choice process. Initially the leading work seems to have been done in the UK, but in the 1970s the baton passed to the US, in particular with McFadden's work, which laid the foundations for modern utility theory as the basis for choice modelling. All of this work was inspired and funded to meet the needs of transport studies.

Another long-standing feature of transport work was the reliance on revealed preference data, which began to change only in the 1980s. While the subsequent achievements in transport and other fields of stated preference analyses, in particular using stated choice, should not be neglected or discarded, Daly argues that revealed preference data still has much to offer

and can form an important substitute, supplement or real-world validation of stated preference information.

Finally Daly points out that the issue of forecasting using choice models is in need of further research attention. He notes in particular the need for forecasting the population whose behaviour is modelled, the characteristics of the choices they will face and the choice function itself. These issues then form an important part of the research he calls for in the coming years.

The following four chapters present the findings from four interactive workshops at the conference. The key benefit to readers of this book who may not have been present at the conference, or may have participated in a different workshop, is that they give overviews of the cutting edge of research and practice in four key areas of choice modelling.

The first of these chapters, by Pinjari, Bhat and Bunch, is a report on recent developments in modelling multiple discrete and continuous choices. While choice modelling to date has primarily focussed on the issue of a consumer choosing a single product from a set of available options, many consumer choice situations involve not just a choice from among a set of competing alternatives, but also a decision on the quantity of the given product that is bought or consumed. Additionally, there are numerous examples where consumers choose not just one product but choose a number of different products, with different quantities selected for each product.

Pinjari et al. list a number of examples of such multiple discrete continuous choices, including decisions to engage in different types of activities and time allocation to each activity, decisions relating to investment portfolios, and brand choice and purchase quantity. Approaches to represent such choices are being more widely used with increases in computer power and a number of theoretical developments have recently been made, many by the workshop participants themselves. This chapter provides a timely overview of these developments.

The chapter highlights the fact that the naïve approach of enumerating all possible combinations of products (and quantities) as individual alternatives within a traditional discrete choice framework quickly becomes intractable as the number of combinations rises. It also discusses how treating the choice as a number of correlated univariate discrete choice model equations fails to provide a unified, underlying theoretical framework. The chapter then focusses on more appropriate and elegant methods derived from classical microeconomic theory of constrained utility maximisation, contrasts the direct utility maximization approach (also called the Kuhn-Tucker approach) and the indirect utility approach, and presents a number of recent advances for the former.

The interest in the modelling of multiple discrete continuous choices is only likely to increase in coming years, across numerous fields of study, and the chapter by Pinjari and colleagues will serve as reference point for such work, while also listing an important set of priorities for future research in the area.

The second workshop report describes the two alternative approaches that have emerged for dealing with multiple-response data: one making explicit assumptions about the distribution of preferences across the population and leading to the development of sample-level models and the other, following Louviere's keynote address, making no such assumption and developing individual-level models.

The key issues identified in the survey of sample-level models are the treatment of fatigue or learning during an interview, the ways in which heterogeneity of preferences can be modelled and the methods available for the correction of errors to account for correlated responses. Individual-level models encounter these difficulties much less, but do have their own specific problem of the sparsity of information available for each respondent. The promising approach of using a Newton-Raphson iteration step is discussed, followed by an alternative procedure based on the consistency of the choices made by an individual. The results of simulation studies are described.

The strengths of the approaches are found to be complementary, the statistical method developed for sample-level models contrasting with the insights given by the individual-level modelling, so that collaborations between the two groups of researchers should be fruitful. However, because of the time taken by detailed discussion of the two approaches, the workshop was not able to discuss future research needs in any detail.

The third of the workshop chapters, by Dellaert, Arentze, Chorus, Oppewal and Wets looks at an emerging area of research, namely the importance of mental representations in making choices. The key argument brought forward by the authors is that when an individual makes a choice, the mental representation of the choice problem at hand plays a crucial role in that respondent's formation of preferences.

The authors argue that this in turn implies that modelling the formation of mental representations has the potential to increase model fit and – perhaps even more importantly – to generate new behavioural insights and associated policy implications. The authors of this chapter highlight the key issue of how to measure the formation of such mental representations and how to model their role in choice behaviour. They discuss a relatively new interview approach that allows them to capture the *decision network* of a respondent which is then represented in the models.

However, the chapter also makes clear that the incorporation of mental

representations in choice models comes with a number of associated conceptual and operational modelling challenges: the authors provide an extensive discussion of these challenges and of ways to address them.

Finally, the fourth workshop chapter, by Sheldon, Dix, Flynn and Metcalfe focussed on the more applied but no less important issue of how to collect appropriate data for the estimation of choice models. This issue is becoming ever more important with the diversification in the types of choice that are modelled as well as reduced funding for data collection.

While a large amount of recent academic research on surveys has focussed on improvements to experimental design, the chapter by Sheldon et al also highlights the importance of the qualitative process, reminding especially academic audiences of the potential benefits of cognitive testing of their surveys, and the role of diagnostic questions.

At a time when face-to-face interviews are in strong decline and there is more and more focus on online surveys, the authors also stress the potential issues analysts face in such cases, notably linked to the growing reliance on *online panels* and the risk of misleading results produced by what the authors term “professional respondents”.

The chapter also goes on to provide examples of good practice in a number of different directions, and highlights the crucial need for further qualitative and quantitative advances in the field of data collection, in the face of the growing complexity of the model structures that rely on data from these surveys.

The remaining chapters are formed by key papers presented at the conference.

First, Batley and Ibáñez discuss theoretical issues in the construction of consumer surplus measures from choice models, an important application of the models. Starting from the influential 1981 paper by Small and Rosen, Batley and Ibáñez show that the derivation of a consumer surplus measure from a discrete choice model requires quite rigorous conditions to be imposed on the choice situation and the choice model. However, if demand can be represented as varying continuously as well as involving discrete choice, then the requirements can be substantially less stringent. They conclude that consumer surplus should be measured relative to discrete-continuous demand rather than strictly discrete demand, i.e. the types of models discussed by Pinjari et al. in their workshop report. This paper, which questions some current practice, exemplifies the rigorous theoretical basis on which choice models can be constructed.

Enam and Choudhury, in contrast, present a very practical application of choice modelling to predict demand for public transport in the context of poor data availability in Dhaka (Bangladesh). They describe a sub-model to reflect the bias and noise in the data. Further, the wide diversity

of incomes and accessibility to modes in this very large city means that it is not certain what choices are available to travellers, so that a further submodel was needed to describe the alternatives from which choices were made. Enam and Choudhury base these submodels and the model of mode choice itself on both Stated Preference and Revealed Preference. They suggest that the methods they have developed could be applicable in other developing countries. Here we have a demonstration that in applications for poor (particularly data-poor) areas, the use of sophisticated models – essentially, modelling the problems – is of particular advantage in overcoming the handicaps that exist and in representing the wide diversity of the population.

Iglesias, Greene and Ortúzar show, in a very interesting application, how choice modelling can be used to help in the design of neighbourhoods in which residents can feel safe. This chapter addresses a very important issue in terms of the use of stated choice surveys. Indeed, while such surveys have been used very widely to study respondents' trade-offs between easy-to-measure attributes such as time and cost, their uptake has been much slower when it comes to qualitative attributes. A key reason for this has been the difficulty in presenting such attributes in the context of a hypothetical survey.

Iglesias and colleagues address this problem through an innovative survey that presents respondents with digital images of public spaces and allows them to choose between them. To further increase realism, the images are customised to different neighbourhoods, allowing better identification with their spatial context. The key finding of the empirical work is that, in spite of low average income levels, there is a relatively high willingness by respondents to pay for improving the perceived safety of areas they use every day. Among the most valued attributes are street maintenance and natural vigilance, in particular the natural flow of people and the presence of elements associated with potential vigilance (e.g. windows and access points to buildings).

La Paix, Bierlaire, Cherchi and Monzón also look at the features of neighbourhoods, but in their case the focus is on the influence that neighbourhood structure has on the travel of residents. In particular, they study not just the types of tours that travellers make but also their actual underlying propensity to travel. With this propensity being unobserved, La Paix and her colleagues use a flexible hybrid choice model which treats the propensity as a latent variable, which is affected by neighbourhood characteristics and which in turn affects the actual choice of tours. The empirical results confirm the underlying hypotheses, showing that neighbourhood attributes do indeed have a significant impact on the choice of the type of tours either directly and through the latent propensity to travel.

Chorus and Rose, in perhaps the most surprising application, discuss the use of the recently proposed Random Regret Minimisation (RRM) model to study the selection of potential dating partners online. This constitutes the first application of RRM outside the areas of transportation and environmental economics.

The chapter starts by arguing, building on results obtained in the field of (social) psychology, how date selection forms a decision type that is particularly likely to trigger regret minimisation-based decision making. Subsequently the RRM model and a more conventional linear-in-parameters Random Utility Maximization (RUM) model are estimated on data obtained from a Stated Choice experiment administered among visitors of a dating website.

The authors find that the RRM model fits the data better than the equally parsimonious linear-in-parameters RUM model (the model fit difference being small, but statistically significant). The chapter also argues – supported by empirical analysis based on the estimated models – that the RRM model captures preferences for so-called compromise alternatives: ‘dates’ with an intermediate performance on all attributes (such as job type, looks) are somewhat preferred over ‘dates’ with a strong performance on some attributes and a poor performance on other attributes. This so-called compromise effect is ignored by linear-in-parameter RUM models.

The chapter concludes with a brief segmentation study, which shows that the RRM model does a better job than RUM in terms of explaining choices made by males, while RRM and RUM perform equally well in terms of explaining date selection by females.

Mokhtarian and Tang show the extent of interaction between online and in-store choices on the behaviour of those shopping for clothes in California. The key link is between the choice of purchase channel, i.e. online or in-store, and choices of pre-purchase channel, which could be either in-store or online, or neither, or both. Strong correlations are found between the pre-purchase and purchase channels and insights are given into the variables influencing the choice of channels in each case. Here we see the commercial value of choice modelling in this case indicating to retailers how they should best approach the marketing of clothes to coordinate their online and store channels of sales.

Finally, Vovsha discusses advances being made in the forecasting of activity participation and the travel generated by those activities. To model this behaviour, which is increasingly being used to support government decision-making in the transport sector, models from the Generalised Extreme Value family are generally used. Vovsha discusses the ways in which these model types can be exploited to represent the

additional choice dimensions and household interactions that activity-based modelling introduces. In particular, he gives attention to the need to parametrise the structural coefficients of the model. Another major issue in this work is the large number of alternatives that can arise because of the combinations of different choice dimensions that need to be treated; treating the choices as sequential is behaviourally unappealing. Vovsha suggests some ways forward in this context.

This paper, by a leading practitioner, shows how choice modelling structures developed in academic contexts can be applied on a large scale to address important problems faced by society at large.

Together, these chapters show the richness of the field and how it is expanding to help policy makers to deal with the infinitely varied aspects of life throughout the world. This book is intended to provide a resource for practising modellers and an insight for those new to the field into the varied and innovative choice modelling techniques and applications that are now being developed and used.

It remains to us as conference organisers to once again thank our six conference sponsors, MVA Consultancy, Peter Davidson Consultancy, PTV, RAND Europe, Resource Systems Group Inc. and Significance. Further thanks need to go to Dan Johnson, Jeremy Shires, Julián Arellana and Qian (Mike) Fu, who worked tirelessly behind the scenes. Finally, all delegates will remember the wonderful musical performance by Supertram, joined on stage by the Red Hot Chilean Professors and John Bates Piano Services.

In closing, we hope to see many of the authors and readers of this volume at the third International Choice Modelling Conference, to be held in Sydney in July 2013.

Stephane Hess and Andrew Daly
Leeds, October 2012