

# 19

## METHODOLOGICAL DEVELOPMENTS IN ACTIVITY-TRAVEL BEHAVIOUR ANALYSIS

***Juan de Dios Ortuzar***

*Pontificia Universidad Católica de Chile, Santiago, CHILE*

***Naveen Eluru***

*McGill University, Montreal, CANADA*

***Karthik K. Srinivasan***

*Indian Institute of Technology, Chennai, INDIA*

### 1. INTRODUCTION

The main focus of developing advanced methodologies in a field such as activity-travel behaviour analysis is precisely to enhance understanding of, in this case, traveller's behaviour. Brownstone (2012) and Cherchi (2012) have provided in-depth reviews and analyses of some of the most important advances and remaining challenges in this vast field. Considering these and the preliminary conclusions of an open and frank discussion about these issues at the recent *International Conference of the Association for Travel Behaviour Research* in Jaipur, India, we were able to identify some important directions towards this objective.

First, to what extent is it really possible to identify real underlying behaviour? Even with the most gloriously sophisticated models at our disposal it is practically impossible to disentangle all the various effects that could be underpinning a given observed or declared set of choices by individuals. Closely tied to this, there seems to be a growing conviction that it is more worthwhile to put increased emphasis on systematic or observed utility, rather than on blindly trying to “extract more juice” from stochastic error components. Third, and probably obvious or inescapable, there is no denial that data is crucial, in particular, there is the need to know what data and in what form is required to settle the previous two considerations. Finally, there seems to be also some consensus on

the need to review the goals of modelling as an important component of social and private project evaluation.

The rest of the paper is organised along these four directions, and an attempt is made to give useful insights and recommendations about these issues.

## **2. TO WHAT EXTENT CAN WE IDENTIFY REAL UNDERLYING BEHAVIOUR?**

There is growing recognition within the travel behaviour community that representing behaviour in a random maximization based framework might not always be acceptable. A number of behavioural approaches to understand choice making have been developed in the literature. These include non-compensatory and semi-compensatory behaviour, bounded rationality, learning and adaptation, local searches, and evolutionary processes. There is growing evidence of non-rational and imperfect rational behaviour in practice, and methods and theory to integrate these into analytical techniques must be developed appropriately.

We wish to stress the need to apply decision rules based on the choice context under consideration. The examination of the choice context needs to incorporate the decision process at play, in addition to analysing choice outcomes. Considering the decision process along with the choice outcome, should enable us to consider potential changes to decision rules during the choice process. For example, it might be helpful to study how people respond to different environments and policy interventions, such as road pricing, by incorporating psychological factors into modelling. To what extent can we identify the true behavioural phenomena underlying observed choice outcomes, remains an important research issue.

The increasing interest to make further use of alternative behavioural paradigms, such as regret theory and prospect theory (in general, non-utility maximization frameworks), for examining choice making is bound to benefit choice behaviour analysis. The consistency and deviation from micro economic foundations and theory need to be assessed for newer generation/emerging methods. Further, some caution is appropriate at this juncture. First, the analyst must recognize that the choice framework should be relevant to the choice context under consideration. Specifically, the applied framework needs to allow for the evaluation of different policy contexts effectively. Secondly, it is not easy to detect (in particular from the prevailing cross-sectional choice data), which of a large variety of decision-making paradigms is the best suited to a particular situation (and even if there is a single one). Indeed work on mental processes, albeit extraordinarily demanding in terms of data, reveals that decision making approaches can vary extraordinarily among members of homogenous samples; moreover, a given individual can change approach depending on certain features of the case in hand (Denstadli, *et al.*, 2012). Thus, the challenges of identifying the underlying true behavioural phenomena are compounded by the possibility of multiple and simultaneous heuristics/choice rules.

To enable more behavioural realism a number of newer frameworks are being considered in addition to random utility maximization. These include game theoretic, constrained, multi-objective, and evolutionary agent based frameworks (Chen and Ben-Akiva, 1998; Laporte, *et al*, 2010; Schmöcker, *et al*, 2008). An open question is how well can these alternative frameworks represent and replicate choice processes and outcomes relative to the more traditional framework. In general, two types of modelling frameworks can be adopted: i) a bottom up approach and ii) a top-down approach. Most current activity-based models seem to follow the bottom up approach, where the overall travel patterns are obtained as a result of individual choice decisions (Salvini and Miller, 2005; Pinjari *et al*, 2008). Alternatively, in the top-down approach the primary focus is on the role of macroscopic factors as the major driving factors of society and the economy, and of their influence on travel demand decisions, with corrections and adaptations based on local information. The advantage of this approach may be not only data and model parsimony, but that it may be directly policy sensitive. In addition, it may enable a more natural representation of regulatory and policy instruments and “governmental” decision agents.

These approaches have been applied in system dynamics and ecological system models (Moffatt and Kohler, 2008). Not enough of the latter approach has been adopted in applied travel demand models. The potential for determining whether and when such frameworks may be applied to forecasting gainfully remain to be examined and investigated systematically.

An important aspect in studying choice behaviour is also to understand that various choices tend to be interconnected. The scope and focus of investigations have expanded substantially from individual and cross sectional discrete choice models to collective and interdependent/group decisions at varying levels of aggregation. Typically, decision makers’ short term choices are constrained by their long term choices. For instance, individuals’ mode to work is contingent on the residential location choice of their households (Pinjari *et al*, 2008). The analyst has to explicitly incorporate the interplay across choices for capturing realism in choice behaviour. Neglecting this interplay could result in models that either under-predict or over-predict the decision maker’s sensitivity to choice behaviour. Further, capturing spatial and temporal interdependencies is becoming more important from an applications perspective (emissions, time-of-day, etc.). Not only choices are of interest, but the underlying latent and observed processes (e.g., learning, adaptation, etc.) are also gaining attention (Arentze and Timmermans, 2003). The possibility of using panel data to address dynamics and changes in choice patterns over time is also of interest. This is now increasingly possible with the greater availability of dynamic/panel data and repeated measurements data (Spissu, *et al*, 2009).

Context seems to be very important in many empirical choice modelling settings. A natural question is how best to bring in the role of context in models. However, there is little discussion within the field today to incorporate the influence of choice sets, constraints and context dependence in complex choices explicitly. There is also little research on how to compile choice sets, and particularly so in terms of their spatial context. Work in the area of time-space geography has highlighted the practical importance of context. Explicit modelling of how context affects choice sets and constraints influencing choice could provide a more realistic representation of choice

process (Waddell, *et al*, 2007). Attempts towards mapping models, behaviours and contexts are desirable. Identification of interactions which are appropriate under different contexts may be essential. Evidence of such contextual influence has been observed in shopping/land-use work. Frameworks and methods that illuminate how interpersonal interactions and budgetary allocations shape the context and ultimately influence activity-travel decisions would be valuable (Castro, *et al*, 2011). Further, past research has not clearly examined how to distinguish between what can be labelled a *constraint* and what is a *component* of the utility function; finally, another cloudy area is how to identify constraints in the decision process and capture them in the data collection process. Neglecting these important aspects of a choice process could lead to incorrect model results, and applying such incorrect models results for policy evaluation would lead in erroneous advice to policy makers.

Overall, there seems to be agreement in the research community about the need to explore further all the aforementioned issues. However, in addition to the need for developing every day more advanced models, we face challenges on different fronts including data sources. These challenges are briefly discussed in the next two sections below.

### **3. ISSUES OF SPECIFICATION, ESTIMATION AND INCREASED EMPHASIS ON SYSTEMATIC OR OBSERVED UTILITY**

In recent years, research on activity travel behaviour has concentrated excessively on examining the error component of the utility associated with the traditional compensatory model, particularly since the advent of simulation based methods such as the Mixed Logit model (Train, 2009). Newer tools and models enable richer specification of empirical choice-related phenomena including heterogeneity, state dependence, inertia, shocks, cohort-dynamics, etc. A number of issues relating to richer specification, notably of the error term, are being investigated with these enhanced tools (Bhat and Castelar, 2002). Specification issues which require attention include: distribution of random components, dependence versus correlation, interactions and confounding effects, the masking of genuine effects and appearance of spurious effects. Further the issue of asymmetric dependence and the limitation of the concept of correlation to capture such features must also be understood.

The broader research context and the growing push for richer specifications entail a number of challenges for model estimation. The spatial context and interdependencies of some choice situations require that the combinatorial nature and dimensionality problem be addressed explicitly. The issue of choice set specification and heterogeneity is important in this regard. Often, with richer models and empirical specifications, problems of numerical convergence, ill-conditioning, practical identification, multiple and non-unique optima are encountered (Bhat, 2011). While some of these issues may be an artefact of the estimation methods (typically simulated maximum likelihood) others may be due to data or model structure limitations. To circumvent or mitigate these difficulties, approaches such as alternative re-parameterization (e.g., value of time), use of robust

standard errors, and dimensionality reduction can be considered (Fosgerau, 2006). Emerging methodological attempts that show promise in addressing these issues include Bayesian estimation, semi and non-parametric techniques, composite maximum likelihood approaches and copula based methods (Bhat and Eluru 2009; Ferdous, *et al*, 2010). Regardless of the approach used, some key concerns remain in relation to practical and theoretical identification, irretrievability of parameters and generating process, and inferential power. Some degree of endogeneity between choice process and choice outcomes may be expected, and needs to be accounted for by appropriate modelling techniques. With increasing data and ranges of choice phenomena, the role of measurement errors and explicit recognition of uncertainty in choice process might be properly captured.

Notwithstanding, there is now a renewed call (see the early insights of Ortúzar and Garrido, 2002; and the strong cautionary note in Ortúzar, 2006) for increased emphasis on improving the systematic (or observed) component of the choice process. The difficulties experienced in recent years, with confounding effects and the like in our more complex and powerful model structures (Walker, 2002; Walker, *et al*, 2007; Cherchi and Ortúzar, 2008), have made it clear that the more we capture of the systematic component, the better the developed models will be for forecasting, and the easier it will be to interpret results. After all, as Train (2009) most famously declared ...”there is a limit about what can we learn from things that we cannot see”.

#### **4. DATA ISSUES TO SUPPORT THE ABOVE**

Data issues are closely intertwined with some of the methodological questions and challenges posed above. In particular, many choice decisions are made under imperfect information. Thus, the role of imperfect information to the decision maker and modeller needs to be separately incorporated. Advanced information technology and tools such as trip planners/ GPS / ITS, and the like, can support decision makers and modellers with better data and to some extent reduce these issues (Bricka and Bhat, 2006). Understanding the extent to which models and methods are constrained by existing data collection tools and survey instruments is an essential pre-requisite for inference and forecasting. Advances in survey methods to circumvent these limitations through innovative data collection techniques and stated choice experiments with efficient/optimal experimental designs offer promise and need to be better integrated with the methodological main-stream. Data quality is certainly affected by self-selection and measurement errors. Attempts to resolve these and identify their presence in empirical contexts will be beneficial in practice.

On the other hand, a significant impediment to incorporating more substance in the specification of observed utility is lack of quality data. In particular, there is lack of panel data, focus group and in-depth interviewing efforts. Enriching the quality of available data should enable the analysts to develop more behaviourally rich models but this is costly in comparison to just collecting the classical cross-sectional information. A cost-effective alternative to long-term data collection is retrospective surveys. Such surveys

may yield biased estimates for short-term travel decisions but can provide reliable estimates for longer term choices. Another alternative is to collect data as a series of cross-sections while ensuring a small sample is continuously interviewed (see the review by Ortúzar, *et al*, 2010).

It is also necessary to invest in focus groups and in-depth interviewing efforts to clearly understand the choice process at hand. In this area of enquiry transport research is clearly behind other fields as noted by Cherchi (2012), see for example Payne, *et al*, (1992); Gilbride and Allenby (2004); Denstadli and Lines (2007). Of course, one of the main reasons behind the lack of this type of data is the prohibitive costs involved in collecting it. Also, the data collected through focus groups and in-depth interviews is not generally directly applicable for forecasting.

There is need to encourage panel data collection efforts to study decision maker level variations (such as inertia and shock effects). The usefulness of such panels in characterizing and understanding changes in behavioural phenomena is well acknowledged. Despite their value addition, some key challenges have hampered the deployment of models based on such data, such as dimensionality, attrition, and problems of initial conditions in contexts ranging from vehicle ownership and transactions to synthetic population dynamics. Also, there is little research on examining how panel data can be employed for forecasting purposes. In this sense, the efforts done in the small (about 300 respondents) and relatively short-lived (only four waves) *Santiago Panel* is noteworthy (Yañez, *et al*, 2011; Yañez, *et al*, 2010b), but it would be desirable to see more and larger efforts elsewhere. From a forecasting perspective, the stability of parameters (or behaviour) over time as well as the stability of model is a cause of concern. In some contexts parameters may exhibit significant dynamics, and predictions about the future may become challenging.

To address some of the data issues, focus groups, one-on-one in-depth interviews and other such direct elicitation techniques are being adopted in the environmental economics and marketing research communities. These methods may yield rich and useful insight into understanding the behavioural processes underlying activity-travel decisions. However, sample size and costs remain a barrier for wider adoption. Therefore, simultaneously, it is important for the research community to consciously make an effort towards developing methodologies suited to analysing panel data, focus group and in-depth interviewing data.

## 5. REVIEW THE GOALS OF MODELLING

Forecasting ability and policy sensitivity should always occupy an important place when considering the goals of modelling. Many models tend to focus more on the estimation process, with greater emphasis on allowing for a complex error structure than specifying appropriate systematic components. More insights are needed on the stability of models and their ability to forecast over time and across spatial units. The value and prospects of using descriptive models and qualitative research can be explored in order to obtain new

insights regarding information acquisition and choice process, with possible implications for improved forecasts.

On the other hand, concern has been expressed about incorporating confounding effects in the models, and about the numerical stability of complex models in the context of simulation and higher dimensionality (Cherchi, 2012). There is a strong need to clarify or develop forecasting procedures for complex models. If the objective of the model becomes forecasting, then the framework for analysis might be entirely different. The Bayesian approaches might offer some advantages in this context (Brownstone, 2012).

Another important aspect often neglected within the travel behaviour research community is the need to have hold out samples for validation. It is important to examine the performance of any model developed for prediction using a—hopefully independent—hold out sample. This should enable the analyst to iron out deficiencies within the developed models. Although methods for selecting a hold out sample from cross-sectional data commonly available have been on offer for some time (Ortúzar and Willumsen, 2011), it is surprising to find so little in the literature about experience with hold out samples. Yañez, *et al*, (2010a) provide an interesting example of such use, in the context of longitudinal data, by examining the performance of panel data models estimated using the first three waves of their panel, when applied to predict behaviour in the fourth wave.

The use of latent variables in forecasting also needs to be developed further for practical applications. Policy analysis and planning objectives today are increasingly aligned towards sustainable development. Traditional models have largely evolved to evaluate infrastructure development objectives. In this regard, it is worth pondering about how to incorporate sustainability objectives for forecasting purposes in existing models.

## ACKNOWLEDGEMENT

We sincerely thank the resource paper authors and all of the workshop participants at the Jaipur Conference for their valuable insights and intellectually stimulating perspectives that greatly contributed to the success of the workshop and made this report possible. Thanks are also due to FONDECYT (Project 1100483), and to the Millennium Institute on Complex Engineering Systems (ICM: P-05-004-F; CONICYT: FBO16) for having supported this work.

## REFERENCES

- Arentze, T. and Timmermans, H., 2003. Modeling learning and adaptation processes in activity-travel choice: a framework and numerical experiment. *Transportation*, 30(1), pp.37-62.

- Bhat, C.R. and Castelar, S., 2002. A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay Area. *Transportation Research Part B*, 36(7), pp.593-616.
- Bhat, C.R. and Eluru, N., 2009. A copula-based approach to accommodate residential self-selection effects in travel behaviour modelling. *Transportation Research Part B*, 43(7), pp.749-765.
- Bhat, C.R., 2011. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B*, 45(7), pp.923-939.
- Bricka, S., and Bhat, C.R., 2006. A comparative analysis of GPS-based and travel survey-based data. *Transportation Research Record*, 1972, pp.9-20.
- Brownstone, D., 2012. Methodological developments in activity-travel behaviour analysis. In: R.M. Pendyala and C.R. Bhat, eds. *Travel Behaviour Research in an Evolving World*. Jaipur: IATBR. Ch.11.
- Castro, M., Eluru, N., Bhat, C.R. and Pendyala, R.M., 2011. A joint model of participation in non-work activities and time-of-day choice set formation for workers. *Transportation Research Record*, 2254, pp.140-150.
- Chen, O.J. and Ben-Akiva, M.E., 1998. Game-theoretic formulations of interaction between dynamic traffic control and dynamic traffic assignment. *Transportation Research Record*, 1617, pp.179-188.
- Cherchi, E., 2012. Modelling individual preferences: state of the art, recent advances and future directions. In: R.M. Pendyala and C.R. Bhat, eds. *Travel Behaviour Research in an Evolving World*. Jaipur: IATBR. Ch.10.
- Denstadli, J.M. and Lines, R., 2007. Conjoint respondents as adaptive decision makers. *International Journal of Market Research*, 49(1), pp.117-133.
- Denstadli, J.M., Lines, R. and Ortúzar, J. de D., 2012. Information processing in choice-based conjoint experiments: a process-tracing study. *European Journal of Marketing*, 46(3), pp.422-446.
- Ferdous, N., Eluru, N., Bhat, C.R. and Meloni, I., 2010. A multivariate ordered response model system for adults' weekday activity episode generation by activity purpose and social context. *Transportation Research Part B*, 44(8), pp.922-943.
- Fosgerau, M., 2006. Investigating the distribution of the value of travel time savings. *Transportation Research Part B*, 40(8), pp.688-707.
- Gilbride, T.J. and Allenby, G.M., 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, 23(3), pp.391-406.
- Laporte, G., Mesa, J.A. and Perea, F., 2010. A game theoretic framework for the robust railway transit network design problem. *Transportation Research Part B*, 44(4), pp.447-459.
- Moffatt, S. and Kohler, N., 2008. Conceptualizing the built environment as a social-ecological system. *Building Research and Information*, 36(3), pp.248-268.
- Ortúzar, J. de D. and Garrido, R.A., 2002. Methodological developments: workshop report. In: H.S. Mahmassani, ed. *In Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges*. Oxford: Pergamon, pp.471-477.
- Ortúzar, J. de D., 2006. Defining the state of the art in discrete choice modelling. Presented at: *Double Workshop on Transportation and Sustainable Cities*. Universidad de Chile, Santiago, Chile 7-11 August 2006.

- Ortúzar, J. de D., Armoogum, J., Madre, J.L. and Potier, F., 2010. Continuous mobility surveys: the state of practice. *Transport Reviews*, 31(3), pp.293-312.
- Ortúzar, J. de D. and Willumsen, L.G., 2011. *Modelling Transport*. 4<sup>th</sup> ed. Chichester: John Wiley & Sons. pp. 284-287.
- Payne, J.W., Bettman, J.R. and Johnson, E.J., 1992. Behavioural decision research: a constructive processing perspective. *Annual Review of Psychology*, 43, pp.87-131.
- Pinjari, A.R., Eluru, N., Srinivasan, S., Guo, J.Y., Copperman, R.B., Sener, I.N. and Bhat, C.R., 2008. CEMDAP: Modelling and micro-simulation frameworks, software development, and verification. Compendium of papers CD-ROM for: *Transportation Research Board 87<sup>th</sup> Annual Meeting*. Washington, DC, USA 13-17 January 2008.
- Salvini, P. and Miller, E.J., 2005. ILUTE: an operational prototype of a comprehensive micro-simulation model of urban systems. *Networks and Spatial Economics*, 5(2), pp.217-234.
- Schmöcker, J.D., Ahuja, S. and Bell, M.G.H., 2008. Multi-objective signal control of urban junctions: framework and a London case study. *Transportation Research Part C*, 16(4), pp.454-470.
- Spissu, E., Pinjari, A.R., Bhat, C.R., Pendyala, R.M. and Axhausen, K.W., 2009. An analysis of weekly out-of-home discretionary activity participation and time-use behaviour. *Transportation*, 36(5), pp.483-451.
- Train, K.E., 2009. *Discrete Choice Methods with Simulation*. 2<sup>nd</sup> ed. Cambridge: Cambridge University Press.
- Waddell, P., Bhat, C., Eluru, N., Wang, L. and Pendyala, R., 2007. Modelling the interdependence in household residence and workplace choices. *Transportation Research Record*, 2003, pp.84-92.
- Walker, J., 2002. The mixed logit (or logit kernel) model: dispelling misconceptions of identification. *Transportation Research Record*, 1805, pp.86-98.
- Walker, J., Ben-Akiva M.E. and Bolduc, D., 2007. Identification of parameters in Normal error components logit mixture (NECLM) models. *Journal of Applied Econometrics*, 22(6), pp.1095-1125.
- Yañez, M.F., Cherchi, E. and Ortúzar, J. de D., 2010a. Forecasting mode choice in the presence of inertia and shock effects: the case of the Transantiago implementation. In: *12<sup>th</sup> World Conference on Transport Research*. Lisbon, Portugal 11-15 July 2010.
- Yañez, M.F., Mansilla, P. and Ortúzar, J. de D., 2010b. The Santiago Panel: measuring the effects of implementing Transantiago. *Transportation*, 37(1), pp.125-149.
- Yañez, M.F., Cherchi, E., Heydecker, B.G. and Ortúzar, J. de D., 2011. On the treatment of repeated observations in panel data: efficiency of mixed logit parameter estimates. *Networks and Spatial Economics*, 11(3), pp.393-418.