Stated preference surveys and experimental design: an audit of the journey so far and future research perspectives

Elisabetta Cherchi and David A. Hensher

Version: June 2 2015 (revised 27 August 2015)

Abstract

This paper is a synthesis of the discussions and ideas that were generated during the workshop on "Stated preference surveys and experimental design" at the 2014 Travel Survey Methods Conference in Leura (Australia). The workshop addressed the challenges related to the design and implementation of stated preference surveys as a way to capture richer behavioural information on the preferences of individuals and groups. The discussion began by reviewing the current state of stated preference surveys and whether and what we have been doing correctly. We then analysed the areas where improvements are still needed, how we can achieve them, and some pros and cons of each improvement.

Introduction

Stated preference (SP) or stated choice (SC) surveys have been extensively used in the last decades in many different fields such as marketing, transport, health economics, agricultural and environmental economics. There now exists a very rich literature showing the ability of SP studies to elicit behavioural responses and to allow identification through estimation of the preferences of a single individual or a group. The ability of SP surveys to support identification of individual preferences in a context that efficiently reduces the cognitive effort and fatigue of the respondent has made them a dominant data paradigm to study individuals' behavioural market decisions.

One the foremost problems generally recognised in SP surveys is that although they have the capability through careful design to mimic the real world, this does not ensure that such experiments represent it (lack of realism). A similar comment could be made about revealed preference (RP) surveys which have a number of challenges associated with non-chosen alternatives and measurement error more generally. The issue of realism has been predominantly related to the use of SP data in prediction, where the alternative specific constants and scale need to be adjusted to the real market shares or some assumption needs to be made to forecast new alternatives, not currently available in the market (Cherchi and Ortúzar, 2006; Glerum et al., 2013, Hensher et al., 2015). When the focus of SP studies is on the derivation of marginal rates of substitution between attributes (i.e., willingness to pay measures) their validity is generally supported although the matter of hypothetical bias still remains; the concern exists however when researchers use SP data almost unconditionally. The use of such data in deriving (direct and cross) demand elasticities of the alternatives is somewhat more problematic since it is calculated using knowledge of predicted choice probabilities, and unless the SP model has been calibrated to reproduce true market shares, the elasticities may be unreliable.

However, the question to what extent the preferences elicited in SP surveys reflect real market observed preferences of the respondent, or to what extent are we manipulating these results, has been raised on a number of occasions, igniting renewed interest in the role of RP data. Problems of reality (or far from reality) are indeed present in both RP and SP data, for different reasons (at least in transport and some other contexts). Some shrewdness to improve realism applies equally to both types of data. For example, in both types of data, realism can be jeopardised by contacting respondents who are not the right persons to interview (a very real issue in, for example, freight studies where the driver of a truck may have some ownership of the travel time but not the travel cost), by not precisely defining the context or by using attributes not measured correctly, such as for example disregarding trip frequency when measuring travel costs. These are all well-known problems, and there are several studies where these issues have been carefully accounted for (e.g., Hensher and Bradley, 1993; Hensher and Raimond, 1995; Cherchi and Ortúzar, 2002; Iragüen and Ortúzar, 2004; Ehreke et al. 2014).

The issue of realism (linked to hypothetical bias, i.e. "the potential error induced by not confronting the individual with an actual situation" Schulze et al. (1981); see also Hensher (2010)) is especially relevant because the SP survey is based on constructed hypothetical scenarios designed to elicit individuals' preferences for specific attributes. The scenarios consist of a set of alternatives defined as a combination of a number of attributes at specific levels. Alternatives, attributes and attributes levels, as well as their combinations presented to the respondents, are defined in advance by the analyst. Keeping the balance between realism (i.e., relevance) and complexity represents probably the major challenge in building a stated preference survey. This is not a new problem. A "to do" list to achieve realism containing complexity is reported in Ortúzar and Willumsen (2011, page 114). However, this is a serious problem that still affects many SP experiments. During the workshop, several points included in that list and new points were discussed in the light of the most recent experiences and advances in the field.

Tasks complexity and respondents engagement

In order to ensure realism and reduce potential hypothetical bias, analysts may need to build rather complex survey tasks which respondents are asked to process in a short time, potentially exerting high burden and risking damaging the quality of response. Individuals have limitations in their capacity to process information, and are not always willing to invest the required degree of effort in evaluating alternatives. When presented with a complex task, it is then likely that they show disengagement, adopting simplifying strategies to reduce the mental effort required to solve the problem. On the other hand, simplifying the survey tasks to reduce the cognitive burden for respondents is also risky. Simplified survey tasks are often too simplistic and they can be seemingly perceived as unrealistic by the respondents, leading to problems with respondents' engagement, or respondents choosing based on other attributes not included in the design. The risk is that a design with too little alternatives or attributes or levels might produce more errors than more "complex" designs.

There are many different experiences and recommendations from the literature on how complex an SP survey should be (Arentze et al., 2003; Hensher, 2004, 2006; Caussade et al., 2005).

Hensher (2004) promotes a position that 'relevance' is what matters and this is likely to result in a large range (i.e., heterogeneous) of 'complexity' as perceived by a survey analyst. The risk that the analyst will impose their own (biased) views on what is complex is a concern, resulting in a study bias that is artificially created and has nothing to do with preference revelation (see also the discussion in Hensher, 2015). The acceptable level of complexity varies among contexts and can also depend on cultural aspects (Rose et al., 2009). Some authors have reported a consistently positive experience with complex designs that include at least 15 attributes, and in some cases more than 40 choice sets (Brazell and Louviere, 1997; Stopher and Hensher, 2000; Hensher, 2001), whereas some studies recommend a number not higher than six (Ortúzar and Willumsen (2011, page 101), because there are some evidence that the number of choice situations to be evaluated led to an increase in the error term variance because the fatigue (Bradley and Daly, 1994; Ortúzar et al., 2000). Hess et al. (2012) suggest otherwise and conclude that "Using a comprehensive testing framework employing both multinomial logit and mixed logit structures, we provide strong evidence that the concerns about fatigue in the literature are possibly overstated, with no clear decreasing trend in scale across choice tasks in any of our studies." The appropriate answer has to be context dependent, and it is something that should always be investigated and carefully tested prior to starting any new SP study. Swait and Adamowicz (1996), for example, found that choice task complexity and cumulative burden affect mainly reliability. More recently, Petrik et al. (2014) used a control scenario to detect potentially disengaged respondents. The control scenario, which consists of replicating, at the end of the survey, one of the scenarios previously presented to each individual, increases the complexity of the survey and can be used only for relatively simple surveys. Moreover, other effects, such as learning, can explain differences in the control scenario. Mazaheri, et al. (2014) grouped the alternatives according to related attributes and select one alternative as a representative alternative for each group. Rose and Bliemer (2014) implement an alternative stated choice survey strategy where choice tasks are constructed such that the first alternative presented in the previous task is removed from the set shown, with the remaining alternatives shifted to fill the resulting gap, and a replacement alternative drawn from the candidate set inserted. This technique allows for the inclusion of a large number of attributes, whilst at the same time potentially reducing the cognitive load placed on respondents undertaking the experiment.

A typical way to minimise the cognitive effort required by respondents is the use of blocks to reduce the number of choice scenarios. However, recently Rose et al. (2014) questioned the use of blocking techniques because the dispersion of survey questions will be perfectly confounded with any detected response heterogeneity, hence making it impossible to determine whether respondents truly have different preferences or error variances, or whether the heterogeneity is an artefact of the survey questions asked. They propose an alternative method of constructing experimental designs that involves repeating a set of questions across individuals, which allow testing for possible demand artefacts in stated choice experiments.

Unfeasible, implausible or dominated alternatives

The potential lack of realism is one of the most challenging elements of SP experiments, induced in part by poor experimental design, such as designs that include unfeasible, implausible or dominated alternatives. Using designs where the attributes are pivoted around the real attribute levels experienced by respondents (as typical, for example, in studies where joint models with revealed and stated preference data are estimated; see e.g., Ben-Akiva and Morikawa, 1990; Swait et al., 1994; Bradley and Daly, 1997; Louviere et al., 2000, chapter 8; Brownstone et al., 2000; Cherchi and Ortúzar. 2002; Bhat and Castelar, 2002; Börjesson, 2008) can reduce the risk of alternatives as packages of attributes that are confusing and lacking meaning.

It is also important to ensure that the combination of attribute levels presented is realistic. With the traditional orthogonal designs, this task was mainly undertaken *a posterior*, checking the tasks after the design was built. In the class of efficient designs, it is possible to incorporate behavioural constraints *a priori*, for example, imposing a condition that an attribute can take a given level only if another attribute is greater than a certain value (these are sometimes referred to as nested designs). Research by Collins et al. (2014) has developed ways to account for constraints when building experimental designs, which involve some amount of trading against statistical efficiency. But this is a small price to pay for increased behavioural relevance. They analysed various rule structures for specifying the constraints and proposed two new algorithms that can handle these constraints effectively. Which algorithm performs best depends on the type of constraints specified; however a convergence solution is not guaranteed. The general message is that it is better to lose some statistical efficiency if it results in greater experimental realism.

Realism is also improved by avoiding dominant alternatives. Since most experimental designs are computer-generated, it is easy to overlook choice tasks that contain a dominant alternative (Bliemer et al. 2014). There are currently no techniques that allow avoiding *a priori* for dominant alternatives, with the solution methods usually involving eliminating problem design features *a posterior*, after the design is built. Any *a posterior* variation of the SP design violates the design's statistical properties, but this is appropriate if it increases realism. Importantly we know that the more attributes and levels there are in a choice experiment design, the less likely that dominant alternatives will exist, something that does tend to pervade relatively simple designs. Using an illustrative simple case with two attributes, Bliemer et al. (2014) show that data sets that contain strictly dominant alternatives (i.e., alternatives always being chosen by a respondent in every offered choice scenario) are problematic, as they may lead to substantially biased parameter estimates. They proposed two different ways to mitigate the problem, namely (i) automatic detection and removal of problematic choice tasks, and (ii) adapting the discrete choice model to ensure that the parameter estimates are less sensitive to scaling issues.

Another potential problem related to the definition of attributes in the SP design is the inability to differentiate between the impacts of the attributes *per se* (i.e., the average impact of an attribute in the evaluation of an alternative), and the impact of the attribute levels shown in the experiment (i.e., the variation in impact due to the range of the attribute in an experimental setting). Recent literature has suggested that the best-worst experiment form nicely allows for this differentiation (e.g., Finn and Louviere, 1992; Marley and Flynn, 2015; Balbontìn et al., 2015). Best-Worst experiments can be also used to avoid redundancy in the offered set of 'statistically designed' alternatives. For example, Hensher and Ho (2014) found empirical evidence to support a simple best-worst binary form, in contrast to other ways of preserving the (constructed) full rank choice set, concluding that this appears to represent the most appropriate processing rule in choice making which can, *ex post*, be used to jointly estimate process and outcome choices.

Use of pictures

The use of pictures has also been discussed as a way to improve realism in SP experiments. "Image theory" suggests that using images instead of words allows a study to better capture the psychological or distorted representation of objective reality residing and existing in the mind of the individual (Myers, 1968). From a theoretical point of view, the use of images is then recommended, but its use in practice can be risky because pictures can convey too much information that, if not controlled through some very explicit conditions, might generate responses that cannot be confidently associated with the attributes that are being assessed. Confounding outcomes are thus quite likely with a risk of ambiguity in the interpretation of the choice process. Hence if images are to be used, they need to be controlled; for example, real images modified on the computer in order to leave only the elements of relevance for the SP experiment, eliminating all the elements that can distract from the main focus avoiding confounding the attributes of interest with features we do not want to measure. Good examples of this can be found in Hensher and Mulley (2015), Galilea and Ortúzar (2005), Sillano and Ortúzar (2005), Strazzera et al. (2010), Iglesias et al. (2012), Kamargianni and Polydoropoulou (2013), Sottile et al. (2014). The majority of these studies report a positive experience. However, there is also evidence where no differences were found between using images or textual information or even worse where using images induced mistakes in the responses (Stopher et al., 2003; Alsnih et al., 2004).

The use of images requires more research since there is an intuitive sense that visualisation is value adding in respect to behavioural relevance and reducing hypothetical bias. A context where (a sequence of) images (or possibly videos) can be very useful is that involving public transport and the response to crowding. The perception of crowding is a tricky notion to capture, since it has both a subjective and an objective dimension. One can show respondents pictures of crowding on buses and trains that depict how many people are standing and how many are sitting; however this may only represent a single time point in the journey (typically defined as when one enters the train or bus), and yet this can vary along the trip and will influence a respondent's perception and assessment of the relevance of this attribute in preference revelation and consequent choice making. We also do not know why some people prefer to stand (possibly to alight quicker). What would be interesting is to show how the crowding (however defined) varies as the trip progresses so one gets a better sense of its true influence, unless it can be shown that decisions are only based on crowding levels upon entering a vehicle or carriage. Results from images can also be used to infer whether attributes are attended to or not. i.e., which attributes people are evaluating, though it is important in this case to avoid confirmation bias.

It has been suggested by a number of studies that the use of the available technology, such as eyes tracking, virtual reality and simulators, represents an important direction to improve the use of images in SP experiments (e.g., Meißner and Decker, 2010; Campbell et al., 2013). .

Qualitative work

There was substantial discussion on ways in which the realism of SP surveys could be improved using qualitative research as an *ex ante* approach to teasing out relevant attributes, processing capabilities, and strategies of candidate respondents. Traditionally, qualitative work was carried out before designing an SP experiment, in the form of focus groups or other forms of group or in-depth interviews (Louviere et al., 2000, page 257). Such methods can be used at the preliminary design stage using (excel) mock ups of an SP instrument as a way to gather information from a small sample of individuals on matters such as the comprehensive and comprehendable nature of the survey instrument, including the relevance of attributes and their levels as well as the length and overall complexity of the tasks.

Qualitative work is also of importance to understand how people evaluate images. This work can be carried out during the focus group to help defining the best way to present the images or during (or immediately after) the SP experiment in the form of supplementary questions to better understand a posterior how people have used and evaluated the information conveyed in the images. If asking qualitative questions during the SP experiment might affect the evaluation of the subsequent tasks, it is then preferable to ask questions at the end of the SP survey.

Supplementary questions have often been used to better understand how respondents process the information in choice experiments. The goal of the SP survey is to elicit the behavioural responses; however, a growing number of studies and authors have recognised the need to understand and account for the underling processing rules (or heuristics) that respondents use in assessing choice experiments and making a choice (including ranking the offered alternatives). A growing body of literature in different fields, including transport, (see Hensher, 2014 for a broader review) has now consistently argued and demonstrated that individuals use different rules (or heuristics) to assist them in making choices, and it is important to identify the right choice response mechanism.

The great majority of choice modelling research has taken, as a maintained assumption, the behavioural position that individuals are fully compensatory in the way that they assess and trade-off attributes in choice making and that in circumstances where the analyst imposes a set of attributes to evaluate, as is common in stated choice experiments, it is commonly assumed that all attributes are relevant in choice making. One consequence of these assumptions is the resulting view that studies that require individuals to assess an increasing number of attributes impose growing complexity and cognitive burden that risks the loss of identifiable comprehendible settings of choice making.

One criticism of this position is that individual choosers are very heterogeneous in the way that they make choices in real and hypothetical markets, and that they draw on rules (or heuristics) to assist them in making choices. These rules may vary by contexts typified, for example, by habit or variety seeking behaviour. They may also reflect the fact that individuals bring to the choice making setting their views on what attributes are the key drivers of specific choice outcomes, and that these attributes may or may not be included in the set defined by an analyst. There is a real risk that the analyst may self-impose their own views (or prejudices) or even those of a client funding a study, on the number of attributes and alternatives that are deemed comprehendible to a sample of respondents in a survey.

These presumptions have been questioned in the broader literature on heuristics and decision making that has evolved in a number of literatures, notably, psychology, economics and marketing; however the migration of ideas from this literature, which we refer to as process heuristics, has been slow to influence the way that discrete choice modelling has been represented. This is changing now, with a growing number of studies questioning the standard fully compensatory choice paradigm.

An example of a process heuristic that have gained a lot of traction in a number of disciplines (especially transport and environmental economics) is attribute non-attendance (ANA), which recognises that individuals often ignore one or more attributes of the alternatives when making their choices (see Collins and Hensher, 2015 for an updated review). The reasons for non-attending to an attribute are many including the absence of relevancy of the attributes and disengagement, as discussed previously, to true indifference to the attribute. Alemu et al. (2013) provide an interesting example of supplementary questions used to detect the nature of attributes non-attendance. It was suggested that supplementary qualitative questions can also be used to reveal if people are telling the 'truth', and hence test the level of engagement and believability of the answers (Morrison et al., 2002).

Whatever is the scope of supplementary qualitative questions, particular attention needs to be pay to the way questions are asked in order to be able to correctly elicit what you would like to measure.

Interaction within groups

In the majority of stated choice surveys, the respondent (a single agent) is asked to evaluate the scenarios presented. However, for many type of choices the preferences underlying decisions may be affected/depend upon the preferences of other people (typically, but not only, other members of the household). It is then important to understand to what extent the preferences elicited in SP surveys reflect the real preferences of the respondent. This is particularly important for choices, such as vehicle ownership or residential choice, but might be true for simpler everyday mode choices (e.g., the use of the family car is often the result of a bargaining process among the activity schedules of different household members). Some studies have included a simple question to measure if the choices stated in the experiment were taken alone or involved a level of bargaining with other people. Although this is useful information that supports maintaining the simplicity of a design, in order to identify individual preferences correctly, group decisions may need to be measured explicitly. There are few examples in the literature where the interaction within groups is taken into account when eliciting individual's preferences. Some studies have attempted to do this by asking individuals to answer the SP experiment twice: first alone and then with the group (usually the partner) (Beharry-Borg et al., 2009), or by building a two-step SP experiment, where in the first step the scenarios imply a change only for one member of the family, while in the second they imply a change for several members of the family (Swärdh and Algers, 2009), or by simulating a bargaining situation (Dosman and Adamowitz, 2005; Hensher et al., 2008). Other studies have included an attribute in the design to measure the effect of friends, relatives or peers' choices on individual preferences (Rasouli and Timmermans, 2013). The problem in this case is to understand if the market share is endogenous or exogenous, and then modelling this properly.

Most advanced studies in this line of research use interactive SP experiments, where each individual provides the evaluation of a set of alternatives several times, every time after being informed about the choice of others (e.g., Beck et al. 2013, 2014). The process continues until both parties agree (or not). These approaches are important to measure individual preferences correctly, but they are also extremely time and cost consuming. Research along these lines is still rare, though crucial. Simulators can be particularly useful to increase realism in testing competition among respondents.

Adaptive/dynamic SP

Another important direction where major research is needed is the use of adaptive and dynamic SP surveys. Adaptive Stated Preference consists in adjusting the levels of the attributes presented in each choice task based on the answers provided in the previous tasks. The adjustment can be across respondents, if the attributes levels are adjusted based on the preferences elicited in previous respondents, or within respondent, if the attributes levels are adjusted based on the preferences elicited in the previous tasks presented to the same respondent. Adaptive designs are particularly suitable for small samples, but might generate endogeneity issues.

The lack of knowledge of the product presented represents another problem that affects the ability of the SP experiment to reflect the real preferences of the respondent. Some SP surveys are used to identify individual preferences for products that do not currently exist in the market. However, individuals might have a misconception about the relevance of the new product and its characteristics and this might affect the measure of the individual preference that we obtain from SP data. Even if the product is not completely new, individuals are adaptive decision makers who learn through trial and error. Their preferences are then affected by (and hence might change with) the experience obtained from using or consuming a certain product. It is important, then, to study the dynamic evolution of individual preferences in order to try and measure correctly individual preferences. Jensen et al. (2013 and 2014) set up a "long panel" survey, where an SP experiment was repeated twice, before and after individuals experienced an electric vehicle in real life during a three-month period. They found that individual preferences indeed change significantly after a real experience and that willingness to pay and demand elasticity for some characteristics change significantly, whether measured before or after the real experience. The use of panel SP before and after experience or real interventions represent an important direction of research to properly account dynamic evolution of individual preferences. This reinforces the findings of Hensher et al. (1997) that overt experience changes preferences.

Hensher (2105) indicates that when the analyst assumes that all attributes offered in a choice experiment are relevant, there is an implicit sense that the respondent has complete familiarity or awareness of both the alternatives on offer and the context within which a choice is being made. A classic example where this fails is the debate on road pricing reform, which is substantively not understood by the majority of respondents, certainly in Australia. The value of identifying respondent awareness and/or familiarity seems appealing to behavioural study as a way of establishing the extent to which better information will improve awareness and familiarity in a way that might engender greater support for a reform package. This is a sensible position for studies interested in stakeholder buy in.

Improving the efficiency of an experimental design

Based on the work of Rose and Bliemer (Bliemer and Rose; 2006; Rose and Bliemer, 2008; 2012), in the last few years there has been a substantial amount of discussion in the literature on strategies to generate experimental designs. Although, in principle, any strategy can be used to build a design, some strategies are statistically more efficient than others, allowing for smaller choice sets and hence smaller sample sizes and/or less cognitive effort and possible fatigue for the respondents. This motivated the use of efficient design versus orthogonal fractional factorial designs, which were originally the most used strategy within the transport literature. As highlighted in recent literature (Louviere et al., 2008; Bliemer and Rose, 2011), statistical efficiency has an impact not only on the standard errors of the parameter estimates but it can also influence the scale of the model. Some concerns about the use of efficiency criteria are related with the consistency between the phenomenon we are able to measure with SP data (i.e., the prior assumptions required about model specification and parameters) and the phenomenon we are able to estimate using such data. Since, in SP experiments, the behaviour underlying the data is controlled by the modeller via the experimental design, an additional challenge is to build the experimental design such that the behaviour underlying the data is consistent with the phenomenon we attempt to estimate with the discrete choice model. Concerns related to the statistical efficiency measure adopted do not matter. Bliemer and Rose (2014) show that although the design principles in different methods may seem very different, all methods follow exactly the same theory, but merely make different assumptions on the efficiency criterion, model specification, assumed parameter priors, and attribute level constraints.

Conclusions

The workshop on "Stated preference surveys and experimental design" generated a lively discussion and many fruitful ideas. As highlighted in this synthesis, researchers in the field have made significant advances in key directions such as a better understanding of behavioural response and improved design efficiency. The availability of modern technology (like the internet or programs that optimise the design of an experiment) have facilitated the use of stated preferences, with more complex experiments, although at the potential risk to the quality of the survey in terms of behavioural representation. At the same time, the advances in econometric models have enabled us to account *a posteriori* for behaviour that cannot be elicited directly from the survey. The research needs that were identified during the workshop, and are discussed in this paper, can be summarised as follows:

- Tasks complexity and respondents engagement
- Unfeasible, implausible or dominated alternatives
- Use of pictures
- Qualitative work
- Interaction within groups

- Adaptive/dynamic SP
- Improving the efficiency of an experimental design

A key recommendation is that future research should aim at improving realism in surveys, revisiting known techniques (such as for example, qualitative surveys), exploring the adoption of the new technologies available (such as eye tracking, simulators, smart phones) and exploiting the knowledge available from other disciplines (such as psychology, neuro-science, and marketing). The message from the workshop was that "better to have excellent data with simple models, rather than poor data with excellent econometric models". Finally, it was also pointed out that it is not always necessary to collect new data; it might be relevant to revisit existing data sets and to re-analyse them as we gain better insights into how data might be further re-analysed and interpreted.

Aknowledgment
The authors are grateful to Kay Axhausen, who co-chaired the workshop with Elisabetta Cherchi, and to all the participants at the workshop for their valuable contribution in discussing the future direction of the stated preference methods. In particular we thank: Matthew Back, Camila Balbontin, Michiel Bliemer, Lidia Bonet, Kelly Clifton, Andrew Collins, Ilka Ehreke, Sergio Garcia, Leighton Haworth, Ana Herrera, Nodoka Kasahara, Kouros Mohammadian, Dean Ranche Manel, John Rose, Olga Petrik, and Simon Washington.

References

Alemu, M.H, Mørbak, M.R., Olsen, S. and Jensen, C. (2013) Attending to the Reasons for Attribute Non-attendance in Choice Experiments. *Environmental and Resource Economics* 54(3), 333-359.

Alsnih, R., Rose, J.M. and Stopher, P. (2004) Dynamic Travel Demand for Emergency Evacuation: The Case of Bushfires, 27th *Australasian Transport Research Forum*, Adelaide.

Arentze, T., Borgers, A., Timmermans, H. and del Mistro, R. (2003) Transport stated choice responses: effects of task complexity, presentation format and literacy. *Transportation Research* 39E, 229‑244.

Balbontín, C., Ortúzar, J. de D. and Swait, J.D. (2015) A Joint Best-Worse Scaling/Stated Choice Model Considering Observed and Unobserved Heterogeneity: an application to residential location choice. *Working Paper* Pontificia Universidad Católica de Chile.

Beharry-Borg, N., Hensher, D.A. and Scarpa, R. (2009) An Analytical Framework for Joint vs Separate Decisions by Couples in Choice Experiments: The Case of Coastal Water Quality in Tobago. *Environmental Resource Economics* 43, 95–117.

Beck, M.J., C.G. Chorus, J.M. Rose and D.A. Hensher (2013) Vehicle purchasing behaviour of individuals and groups: regret or reward? *Journal of Transport Economics and Policy*, 47(3), 475-92.

Beck, M.J., Rose, J.M. and Zhang, J. (2014) Collecting instant and simultaneous group responses in stated preference experiments. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Ben-Akiva, M.E. and Morikawa, T. (1990) Estimation of travel demand models from multiple data sources. *Proceedings 11th International Symposium on Transportation and Traffic Theory*. Yokohama, Japan

Bhat, C. and Castelar, S. (2002) A unified mixed logit framework for modelling revealed and stated preferences: Formulation and application to congestion pricing analysis in the San Francisco Bay Area. *Transportation Research Part B: Methodological* 36(7), 593-616.

Bliemer, M.C.J., and Rose, J.M. (2011) Experimental design influences on stated choice outputs: an empirical study in air travel choice. *Transportation Research 45A*, 63-79.

Bliemer, M.C., Rose, J.M. and Chorus, C. (2014) Dominancy in stated choice surveys and its impact on scale in discrete choice models. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Bliemer, M.C. and Rose, J.M. (2014) A unified theory of experimental design for stated choice studies. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Bliemer, M.C. and Rose, J.M. (2006) Designing stated choice experiments: The state of the art. 11th *International Conference on Travel Behaviour Research*, Kyoto, Japan, 20th August 2006

Börjesson, M. (2008) Joint RP/SP data in a mixed logit analysis of trip timing decisions. *Transportation Research Part E: Logistics and Transportation Review*, 44(6), 1025-1038.

Bradley, M. and Daly, A.J. (1997) Estimation of logit models using mixed stated preference and revealed preference information. In P. Stopher and M. Lee-Gosselin (eds.), *Understanding Travel Behaviour in an Era of Change*. Pergamon Press, Oxford.

Brazell, J. And Louviere, J.J. (1997) Respondents' help, learning and fatigue. *INFORMS Marketing Science Conference*, University of California at Berkeley.

Brownstone, D., Bunch, D. and Train, K. (2000) Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research* 34B, 315-338.

Campbell, B., Behe, B.K., Khachatryan, H.; Hall, C.R., Dennis, J.H., Huddleston, P.T. and Fernandez, R.T. (2013) Incorporating Eye Tracking Technology and Conjoint Analysis to Better Understand the Green Industry Consumer. *Agricultural & Applied Economics Association's AAEA & CAES Joint Annual Meeting*, Washington.

Caussade, S., Ortúzar, J.de D., Rizzi, L.I. and Hensher, D.A. (2005) Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research* 39B, 621-640.

Cherchi, E. and Ortúzar, J. de D. (2002) Mixed RP/SP models incorporating interaction effects: modelling new suburban train services in Cagliari. *Transportation* 29, 371-395.

Cherchi, E. and Ortúzar J. de D. (2006) On fitting mode specific constants in the presence of new options in RP/SP models, *Transportation Research Part A: Policy and Practice,* 40(1): 1-18.

Collins, A. T. Bliemer, M.C. and Rose, J.M. (2014) Constrained stated choice experimental designs. 10th *International Conference on Transport Survey Methods,* Leura, Australia.

Collins, A. and Hensher, D.A. (2015) The influence of varying information load on inferred attribute non-attendance. In *Bounded Rational Choice Behaviour: Applications in Transport.* S. Rasouli and H. Timmermans (Eds.), Emerald. Chapter 4, 73-94.

Dosman, D. and Adamowicz, W. (2006) Combining stated and revealed preference data to construct an empirical examination of intrahousehold bargaining. *Review of Economics of the Household* 4, 15–34

Ehreke, I., Axhausen, K. and Weis, C. (2014) The German value of travel time savings (VTTS) and value of reliability (VOR) study. 10th *International Conference on Transport Survey Methods*, Leura, Australia..

Galilea, P. and Ortúzar, J. de D. (2005) Valuing noise level reductions in a residential location context. *Transportation Research, D,* 10, 305-322

Glerum A., Stankovikj L., Thémans M. and Bierlaire, M. (2013) Forecasting the demand for electric vehicles: accounting for attitudes and perceptions, *Transportation Science,* published in advance Online: December 12 (2013) Doi:10.1287/trsc.2013.0487

Finn, A. and Louviere, J.J. (1992) Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy and Marketing* 11, 12–25.

Hensher, D.A. (2001) Measurement of the valuation of travel time savings. *Journal of Transport Economics and Policy* 35, 71‑98.

Hensher, D.A., (2004) Identifying the influence of stated choice design dimensionality on willingness-to-pay for travel time savings. *Journal of Transport Economics and Policy,* 38(3), 425-446.

Hensher, D.A. (2006) How do respondents process stated choice experiments? – Attribute consideration under varying information load, *Journal of Applied Econometrics,* 21(6), 861-878.

Hensher, D.A. (2010) Hypothetical bias, choice experiments and willingness to pay, *Transportation Research Part B*, 44, 735-752.

Hensher, D.A. (2014) Attribute processing as a behavioural strategy in choice making. In S. Hess, A.J. Daly (Eds.), *Handbook of Discrete Choice Modelling,* Edward Elgar, UK.

Hensher, D.A. (2015) Data challenges: more behavioural and (relatively) less statistical – a think piece, invited plenary paper for Elsevier Procedia, edited by Marcela Munizaga, *Selected proceedings of the 2014 International Conference on Travel Survey Methods.*

Hensher, D.A. and Bradley, M. (1993) Using stated responses choice data to enrich revealed preference discrete choice models. *Marketing Letter* 4(2), 139-151.

Hensher, D.A., Louviere, J.J. and Swait, J. (1999) Combining sources of preference data. *Journal of Econometrics* 89, 197–221.

Hensher, D.A., and Ho, C. (2014) Identifying a behaviourally relevant choice set from stated choice data. *Transportation* DOI 10.1007/s11116-014-9572-z.

Hensher, D.A. and Mulley, C.M. (2015) Modal image: candidate drivers of preference differences for BRT and LRT, *Transportation* 42 (1), 7-23.

Hensher, D.A. and Raimond, T. (1995) Evaluation of fare elasticities for the Sydney region. *Report* prepared for the Institute of Transport Studies for NSW Government Pricing Tribunal, Sydney,

Hensher, D.A., Battellino, H.C. and Gee, J.L. (1997) The role of stated preferences and discrete choice models in identifying community preferences for traffic management devices, in

P.R. Stopher and M. Lee-Gosselin (ed.) *Understanding Travel Behaviour in an Era of Change*, Proceedings of 6th International Conference on Travel Demand Modelling, Pergamon Press, Oxford, 233-256.

Hensher, D.A., Rose, J.M. and Black, I. (2008) Interactive agency choice in automobile purchase decisions: the role of negotiation in determining equilibrium choice outcomes. *Journal of Transport Economics and Policy* 42(2), 269–296.

Hensher, D.A., Rose, J.M. and Greene, W.H, (2015) *Applied Choice Analysis*, Second Edition, Cambridge University Press, Cambridge.

Hess, S., Hensher, D.A., and Daly, A.J. (2012) Not bored yet – revisiting respondent fatigue in stated choice experiments, *Transportation Research Part A*, 46, 626-644.

Iglesias, P. Green, M. and Ortùzar, J. de D. (2012) On the perception of safety in low income neighbourhoods: using digital images in a stated choice experiment. In Choice Modelling: the state of the art and the state of practice. S. hess and A. Daly (eds). Chapter 9, 193-210.

Iragüen, P, and Ortúzar, J. de D. (2004) Willingness-to-pay for reducing fatal accident risk in urban areas: an internet-based web page stated preference survey. *Journal of Accident and Prevention* 36(4), 513-524.

Kamargianni, M., and A. Polydoropoulou (2013) Hybrid choice model to investigate effects of teenagers' attitudes towards walking and cycling on mode choice behavior. *Transportation Research Record* 2382(1), 151-161.

Jensen, A., Cherchi, E. and Mabit, S. (2013) On the stability of preferences and attitudes before and after experiencing an electric vehicle. *Transportation Research* 25 D, 24-32.

Jensen, A., Cherchi, E., and Ortúzar, J. de D. (2014) A long panel survey to elicit variation in preferences and attitudes in the choice of electric vehicles. *Transportation* 41(5), 973-993.

Louviere, J.J. and Hensher, D.A. (1983). Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. *Journal of Consumer Research* 10 (3), 348–361.

Louviere, J.J., Hensher, D.A. and Swait, J. (2000) *Stated Choice Methods*, Cambridge University Press, Cambridge

Louviere, J.J., Islam, T., Wasi, N., Street, D., and Burgess, L. (2008) Designing discrete choice experiments: do optimal designs come at a price? *Journal of Consumer Research* 35(2),360–375.

Meißner, M. and Decker, R. (2010) Eye-tracking information processing in choice-based conjoint analysis. *International Journal of Market Research* 52(5), 593–612.

Marley, A.A.J., and Flynn, T.N. (2015) Best and worst scaling: theory methods and application. In J.D. Wright (Editor-in-Chief), *International Encyclopedia of the Social and Behavioral Sciences*, 2nd Edition. Elsevier. 548-552.

Mazaheri, A, Zabihi, M., Seyedabrishami, S. and Mahmud Saffarzadeh, M. (2014) Behavioral Analysis of Iranian International Air Passengers in Selecting Low Cost Carriers: A Stated Preference Approach. 10^{th} *International Conference on Transport Survey Methods*, Leura, Australia.

Morrison, M.A, Haley, E., Sheehan, K.B. and Taylor, R.E. (2002) *Using Qualitative Research in Advertising*, Thousand Oaks, CA: SAGE Publications.

Myers, J.G. (1968). Consumer Image and Attitude. *Institute of Business and Economic Research*. University of California, Berkeley.

Ortúzar, J. de D., Roncagliolo, D.A. and Velarde, U.C. (2000) Interactions and independence in stated preference modelling.In: Ortúzar, J. de D. (Ed.), *Stated Preference Modelling Techniques.Perspectives* 4, PTRC, London.

Petrik, O., de Abreu e Silva, J. and Moura, F. (2014) Revealing Disengagement of Respondents in Travel Stated Preference Surveys. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Rasouli, S. and Timmermans, H. (2013) The Effect of Social Adoption on the Intention to Purchase Electric Cars: A Stated Choice Approach. *Transportation Research Record*, 2344, 10-19.

Rose, J.M. and Bliemer, M.C. (2008) Stated preference experimental design strategies. In *Handbook of Transport Modelling*, ed. D.A Hensher and K.J. Button, Elsevier, Oxford, United Kingdom, pp. 151-179

Rose, J.M. and Bliemer, M.C. (2012) Stated preference experimental design strategies. In *Transport Economics: Critical Concepts in Economics (Volume 1)*, ed. David A. Hensher, Routledge, Oxford, United Kingdom, pp. 304-32

Rose, J.M. and Bliemer, M.C. (2014) Sliding doors: A transitional representation of presenting stated choice survey questions. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Rose, J.M., Bliemer, M.C., Goos, P. and Louviere, J. (2014) Survey artefacts in stated choice experiments. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Rose, J., Hensher, D.A, Caussade, S., Ortúzar, J.de D. and Jou, R.C. (2009) Identifying differences in willingness to pay due to dimensionality in stated choice experiments: a cross country analysis. *Journal of Transport Geography* 17, 21-29.

Schulze, W.D: d'Arge, R.C. and Brookshire, D.S. (1981) Valuing environmental commodities: Some recent experiments. *Land Economics*, 57(2), 151-172.

Sillano, M. and Ortúzar J. de D. (2005) Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning* 37A, 525-550.

Sottile, E. Cherchi, E, and Meloni, I. (2014) Measuring soft measures within a stated preference survey: the effect of pollution and stress from traffic in the mode choice. 10th *International Conference on Transport Survey Methods*, Leura, Australia.

Stopher, P.R. and Hensher, D.A. (2000) Are more profiles better than fewer. Searching for parsimony and relevance in stated choice experiments. *Transportation Research Record* 1719, 165–174.

Stopher, P.S., Alsnih, R. and Rose, J.M. (2003) Emergency Bushfire Evacuations (Reports I & 2), report prepared for Emergency Management Australia, Institute of Transport Studies, The University of Sydney.

Strazzera, E., Cherchi, E. and Ferrini, S. (2010) Assessment of regeneration projects in urban areas of environmental interest: a stated choice approach to estimate use and quasi-option values. *Environmental and Planning A, 42(2*), 452-468.

Swait, J., Louviere, J. and Williams, M. (1994) A sequential approach to exploiting the combined strengths of SP and RP data: Application to freight shipper choice. *Transportation*, 21(2), 135–152

Swärdh, J.E. and Algers, S. (2009) Willingness to accept commuting time for yourself and for your spouse: Empirical evidence from Swedish stated preference data. *Working Paper*, Royal Institute of Technology, Sweden.