ITLS

WORKING PAPER

ITLS-WP-08-09

Should reference alternatives in pivot design SC surveys be treated differently?

By

Stephane Hess* & John M. Rose

* ITLS Sydney and ITS Leeds

April 2008

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-08-09

TITLE: **Should reference alternatives in pivot design SC surveys be treated differently?**

ABSTRACT: Analysts are increasingly making use of pivot style Stated Choice (SC) data in the estimation of choice models. These datasets often contain a reference alternative whose attributes remain invariant across replications for the same respondent. This paper presents some evidence to suggest that the standard specification used for such data may not be appropriate. As such, our analysis shows differences not only in the specification of the observed part of utility between the reference alternative and hypothetical SC alternatives, but also suggests differences in the error terms.

KEY WORDS: ***Stated choice; pivot designs; discrete choice***

AUTHORS: Stephane Hess & John M. Rose

Institute of Transport and Logistics Studies **(C37)**

The Australian Key Centre in Transport Management

The University of Sydney NSW 2006 Australia

CONTACT:

Telephone:	+61 9351 0071
Facsimile:	+61 9351 0088
E-mail:	itlsinfo@itls.usyd.edu.au
Internet:	http://www.itls.usyd.edu.au

DATE: April 2008

1 Introduction

Stated Choice (SC) theory and methods are not new. The first SC experiment is thought to have been conducted by Thurstone (1931) who, using a crude form of experimental design, attempted to estimate indifference curves by asking respondents to make choices between different combinations of coats, hats and shoes. Whilst Thurstone reported some success using this method, economists at the time derided the methodology on the basis of a lack of realism which they argued was likely to give rise to spurious results. In particular Wallis and Friedman (1942) stated:

> *"It is questionable whether a subject in so artificial an experimental situation could know what choices he would make in an economic situation; not knowing, it is almost inevitable that he would, in entire good faith, systematise his answers in such a way as to produce plausible but spurious results."*

and

> *"...for a satisfactory experiment it is essential that the subject give actual reactions to actual stimuli... Questionnaires or other devices based on conjectural responses to hypothetical stimuli do not satisfy this requirement. The responses are valueless because the subject cannot know how he would react."*

Such arguments against the use of SC data have not abated over time (e.g. Camerer and Hogarth, 1999; Diamond and Hausman, 1994; List, 2001). Yet despite these criticisms, SC techniques have progressed to become the dominant method for collecting and analysing data related to individual choice behaviour. In part, this is due to the limitations of revealed preference (RP) data, such as restrictions on the alternatives available for modelling purposes as well as issues related to capturing data on non-chosen alternatives. The success in making SC methods more realistic may also partly account for the increased use of SC data. Studies showing that SC results may be similar to those obtained from RP models up to a scale also strengthen the case for using SC data (for a review of such literature, see Louviere et al., 2000).

Attempts to make SC choice tasks more realistic have taken many forms over the years. Making SC choice tasks more incentive-compatible (e.g. Ding, 2007; Rousseas and Hart, 1951) and individual customisation of SC choice tasks to respondent specific experiences (e.g. Rose et al., 2008; Train and Wilson, 2007) represent just two approaches researchers have taken to improve the realism of SC choice tasks. It is with the latter approach that this paper is concerned.

Within the literature, there has been a growing trend towards the promotion of SC experiments where the attributes of the alternatives are pivoted around the knowledge base of sampled respondents (for applications see for example Hensher and Greene, 2003; Hensher, 2004). The use of a respondent's knowledge base to derive the attribute levels of the SC experiments is supported by a number of theories derived in behavioural and cognitive psychology, economics, case-based decisions theory and minimum-regret theory (cf. Starmer, 2000; Kahnemann and Tversky, 1979; Gilboa et al., 2002). In total, these theories support approaches that psychologically relate experiments to individual specific experiences and perceptions. For example, prospect theory (Kahnemann and Tversky, 1979) argues that individuals use decision heuristics when making choices and promotes the idea that the context in which a decision is made is an important determinant of the choice process, supporting the use of reference alternatives in SC choice tasks.

Typically, the relating of SC experiments to individual specific experiences can take one of three forms. Firstly, such experiments may involve the presence of a status quo alternative which is represented as a null alternative with the attributes and attribute levels of the alternative not shown as part of the experiment. A second form of these experiments involves respondents being shown alternatives with attribute levels based on their own experiences but not the exact levels as described (e.g. Hensher and Rose, 2007; Train and Wilson, 2007). A final form of these experiments involves the inclusion of one or more alternatives in the choice task being described with the exact levels representing each respondent's recent experiences (e.g. Hensher and Greene, 2003; Hensher, 2004). It is this last representation that we wish to focus on in this research.

When faced with an SC choice task where one alternative is represented as an individual specific reference alternative, there exists the possibility that respondents may treat that alternative systematically different to other 'hypothetical' alternatives that they are presented with. This systematic difference may arise either as a result of an experienced reference alternative being compared with constructed, non-experienced, hypothetical alternatives, or from the fact that the reference alternative is held constant across choice tasks whereas the remaining alternatives are forced to vary by way of the experimental design. Theoretically, it is also possible for the hypothetical alternatives of SC experiments involving reference alternatives to be more highly correlated with each other than with the reference alternative (e.g. Train and Wilson, 2007).

The purpose of this paper is to examine different approaches to modelling SC experiments involving the inclusion of individual specific reference alternatives. An important reference in this context is the work of Scarpa et al. (2007), who look at the inclusion of a constant for the *status quo* alternative, as well as dealing with the differences between this alternative and the hypothetical SC alternatives through a Nested Logit or Error Components specification. In the present paper, we expand the investigation into potential differences in treatment across alternatives. In particular, we examine whether the parameter estimates of hypothetical choice tasks should be treated the same as those of reference alternatives. We further examine the error structures of these hypothetical versus real alternatives in SC tasks. Of particular interest are differences in willingness to pay estimates as well as the possible presence of non-linearity of preferences as attribute levels of the hypothetical

alternatives in the experiment are either improved or made worse when compared to the reference alternative.
The remainder of this paper is organised as follows. The next section looks at the various methodological issues arising with pivot design surveys. This is followed in Section 3 by an application illustrating some of these issues with the help of a typical pivot design dataset. Finally, Section 4 summarises the work and provides some directions for future research.

2 Methodological issues

We will base our discussion of the methodological and conceptual issues on the case where the choice set contains three alternatives, of which one is a reference alternative corresponding to a recently chosen *real world* option. Consequently, the attributes of this alternative are kept constant across choice situations, while those of the remaining two alternatives vary across replications, with attribute levels being framed around those of the reference alternative. This is consistent with the data used in the application in Section 3 which in turn is representative of a general type of pivot design dataset.
Let $U_{i,n,t}$ give the utility of alternative i in choice situation t for respondent n. Then in our three alternative example:

$$\begin{aligned} U_{1,n,t} &= V_{1,n} + \epsilon_{1,n,t} \\ U_{2,n,t} &= V_{2,n,t} + \epsilon_{2,n,t} \\ U_{3,n,t} &= V_{3,n,t} + \epsilon_{3,n,t} \end{aligned} \quad (1)$$

where $V_{i,n,t}$ and $\epsilon_{i,n,t}$ refer to the observed and unobserved utility components respectively. The fact that the attributes of the reference alternative are kept constant across replications leads to the use of $V_{1,n}$ rather than $V_{1,n,t}$.
We will first focus on the observed part of utility $V_{i,n,t}$, before looking at the potential implications that the data type has on the unobserved part of utility $\epsilon_{i,n,t}$. In the observed part of utility, an attribute $x_{i,k}$ of alternative i is associated with a marginal utility coefficient $\beta_{i,k}$, where we have that:

$$V_{i,n,t} = \sum_{k=1}^{K} f_{i,k} \left(\beta_{i,k,n}, x_{i,n,t,k}, \Omega_{i,k} \right), \quad (2)$$

where $x_{i,n,t,k}$ gives the value of attribute k for alternative i as faced in choice situation t by respondent n. The function $f_{i,k}()$ interacts this attribute with the associate marginal utility coefficient $\beta_{i,k,n}$ and potentially also additional shape parameters contained in the vector $\Omega_{i,k}$.
In the formulation given in Equation 2 we made very few assumptions about the shape of the utility function. In most applications, a linear in parameters formulation is used, such

that:

$$f_{i,k}\left(\beta_{i,k,n}, x_{i,n,t,k}, \Omega_{i,k}\right) = \beta_{i,k,n}\, x_{i,n,t,k}. \tag{3}$$

Two further simplifications generally apply. The estimation of respondent specific marginal utility coefficients is not practical, such that $\beta_{i,k,n}$ is replaced by $\beta_{i,k}$ where taste heterogeneity is potentially included through interaction with socio-demographic attributes[1] or through a random coefficients formulation such as Mixed Multinomial Logit (MMNL) (cf. Train, 2003).

Moving away from a notation with respondent specific coefficients does not cause any more issues here than in a case using data without a reference alternative. However, the situation is less straightforward with the second simplification. Thus far, we have assumed that all marginal utility coefficients are alternative specific. In practice however, at least some of the attributes will be shared across alternatives, and generic coefficients will be used. As such, while, in a mode choice model, separate travel time coefficients may be used for different modal alternatives, a common travel time coefficient will be used for alternatives sharing the same mode. Such a situation for example arises when using data from an unlabelled route choice experiment.

In most situations, there would be no reason to suppose that the marginal utility for a shared attribute is not identical across alternatives. However, in the case of pivot style data, this may no longer be the case. Two main reasons exist for why the marginal utility valuations might differ between the reference alternative and the SC alternatives. As a first example, the fact that one alternative is a *real world* alternative while the other alternatives are hypothetical may play a role. As such, respondents may react differently to the attributes of an alternative that they have actually chosen in real life, potentially repeatedly so[2]. An additional reason for potential differences in the marginal utility coefficients lies in the fact that the attributes of the reference alternative stay constant across replications while those of the SC alternatives vary. Here, the variation in attributes for the SC alternatives may potentially lead to differences in response.

On the basis of the above, it seems important for researchers to test for the possibility of alternative specific responses, by estimating separate coefficients at least for the reference alternative and for any remaining hypothetical SC alternatives[3]. As such, for our three

[1] This can be accommodated in $f_{i,k}\left(\right)$.

[2] As an example, in the case of commuters, the reference alternative may be the result of experience collected over many years.

[3] There is little reason to expect differences in the marginal sensitivities across the hypothetical SC alternatives.

alternative example, the base situation from Equation 1 would be adapted as follows:

$$\begin{aligned}
U_{1,n,t} &= \sum_{k=1}^{K} f_{R,k}\left(\beta_{R,k,n}, x_{1,n,k}, \Omega_{R,k}\right) + \epsilon_{1,n,t} \\
U_{2,n,t} &= \sum_{k=1}^{K} f_{S,k}\left(\beta_{S,k,n}, x_{2,n,t,k}, \Omega_{S,k}\right) + \epsilon_{2,n,t} \\
U_{3,n,t} &= \sum_{k=1}^{K} f_{S,k}\left(\beta_{S,k,n}, x_{3,n,t,k}, \Omega_{S,k}\right) + \epsilon_{3,n,t},
\end{aligned} \tag{4}$$

where the functional forms as well as the marginal utility and shape coefficients are specific to a given group, either the reference alternative (R) or the two hypothetical SC alternatives (S). In practice, such an approach is not used, and coefficients relating to the same attribute are generic across alternatives. Testing the validity of this assumption is one of the objectives of this study. Here, it should also be noted that the difference in coefficients across alternatives may extent from the fixed coefficient scenario to a random coefficients case, a point addressed below. Furthermore, there is a question as to whether the same functional form can be used for the utility functions across the two types of alternatives.
Another issue that should be kept in mind when dealing with pivot style data is the possibility that some of the attributes may take on a zero value for the reference alternative. This can for example be the case in toll road studies where the respondent was observed to choose an untolled option on the reference trip. Here, the issue is not so much one of alternative specific coefficients, but a division of the sample population into two groups, depending on whether or not a non zero value was observed for the reference trip.
In modelling terms, and working with a linear formulation, this would equate to replacing $\beta_k x_k$ by $\beta_{k,A} x_k I(A) + \beta_{k,B} x_k I(B)$, where $I(A)$ is set to 1 only if the given respondent falls into the first group, with a corresponding formulation for $I(B)$. In the present context, the two groups would correspond to respondents who did or did not face a non-zero value for a given attribute in the reference trip. The grouping becomes increasingly complex with more attributes potentially taking on zero values for the reference alternative.
A final important point is the possibility that respondents may in fact evaluate the attributes of the SC alternatives relative to those of the reference alternative, rather than working with absolute values. In this context, increases and decreases (gains and losses) may be evaluated asymmetrically, as discussed in detail by Hess et al. (2007) and also applied by Hensher (2007). With this specification, any attributes shared across alternatives will disappear from the observed utility function for the reference alternative, where this will generally reduce to a constant. In the observed utility functions for the hypothetical SC alternatives, and working with a linear formulation, the term $\beta_k x_{i,k}$ will be replaced by $\beta_k^{inc} \max\left(x_{i,k} - x_{R,k}, 0\right) + \beta_k^{dec} \max\left(x_{R,k} - x_{i,k}, 0\right)$, where $x_{R,k}$ and $x_{i,k}$ give the value for attribute k for the reference and concerned SC alternative respectively, and where β_k^{inc}

and β_k^{dec} are separate coefficients for increases and decreases in attribute x_k relative to the reference alternative.

We now move to a discussion of how the nature of pivot style data may affect the unobserved part of utility of choice models estimated on such data.

The unobserved part of utility $\epsilon_{i,n,t}$ may be made up of several individual components reflecting a number of different processes that cannot be represented adequately in the formulation of $V_{i,n,t}$. In addition to the usual extreme value terms, this may include terms related to the representation of random taste heterogeneity, inter-alternative correlation and heteroscedasticity across alternatives and/or respondents.

A first problem arises due to the presence of the extreme value terms. Here, we limit the discussion to the Multinomial Logit (MNL) case, extension to other structures is straightforward. In the MNL model, $\epsilon_{i,n,t}$ is given by a type I extreme value term, say $\varepsilon_{i,n,t}$ distributed identically and independently across alternatives and observations. The extreme value term captures a host of phenomena that cannot be explained in the observed part of utility. However, whether the underlying model is MNL, Nested Logit (NL) or Cross-Nested Logit (CNL), the assumption is made that these error terms are distributed independently across observations. It is not clear whether this assumption is appropriate in the case of an alternative whose attributes do not change across observations. Indeed, some of the unobserved components of the utility for this alternative can clearly be assumed to be shared across observations. In fact, it could be argued that this should apply to the majority of the components in the unobserved part of utility. Possible exclusions include learning and fatigue effects, but these are generally accommodated through terms other than the extreme value terms.

It is not clear a priori what approach can be used to address this issue, or what the effects of not addressing it are on model estimates. If the underlying extreme value structure is to be retained[4], the correlation in the errors across alternatives needs to be represented through additional error terms. This is possible with the use of an Error Components Logit (ECL) formulation in which normally distributed random variates are added to the utility functions. However, to maintain homoscedasticity, error components with the same variance need to be added to the utility functions of all alternatives. However, with the correlation only applying for the first alternative, the formulation in Equation 1 would have to change to:

$$\begin{aligned} U_{1,n,t} &= V_{1,n} + \sigma\,\xi_{1,n} + \varepsilon_{1,n,t} \\ U_{2,n,t} &= V_{2,n,t} + \sigma\,\xi_{2,n,t} + \varepsilon_{2,n,t} \\ U_{3,n,t} &= V_{3,n,t} + \sigma\,\xi_{3,n,t} + \varepsilon_{3,n,t}, \end{aligned} \tag{5}$$

where $\xi_{1,n}$, $\xi_{2,n,t}$ and $\xi_{3,n,t}$ are normal variates with a mean of zero and a standard deviation of 1, where the multiplication by a common σ ensures homoscedasticity. Here, $\xi_{1,n}$ varies

[4] I.e., short of moving away from a model structure based on extreme value distributions.

only across respondents, while $\xi_{2,n,t}$ and $\xi_{3,n,t}$ are also distributed within respondents. This results in the required correlation structure, but leads to significant estimation problems due to the placement of integrals at different locations in the likelihood function, as discussed by Hess and Rose (2007).

For two reasons, the assumption of identical error terms across alternatives may also be violated in the case of data with a reference alternative. Firstly, the variation in the attributes for some of the alternatives but not for the reference alternative may lead to a greater error for the former. Secondly, the difference between a *known* reference alternative and *unknown* SC alternatives may similarly lead to a greater error for the latter.

Such heteroscedasticity across alternatives can be accommodated either in a heteroscedastic extreme value (HEV) model, as discussed by Bhat (1995) or through an appropriate ECL specification of the MMNL model, as discussed by Walker (2001). Both models can be difficult to estimate, and the identification issues for the ECL specification are non-trivial.

As already mentioned above, the potential for differences in marginal sensitivities for the two sets of alternatives are not constrained to the fixed coefficients case but extend to a random coefficients scenario, in which case the random component would enter the unobserved part of utility, again leading to differences in error terms across alternatives.

Another possible reason for differences in the distribution of the error terms, this time across respondents, is the case where some respondents have a zero value for one or more of the attributes in the reference alternative. This could potentially lead to a different relative weight for the observed part of utility for these respondents. These differences can be accommodated rather straightforwardly with most estimation packages, through identification of group specific scale parameters (with appropriate normalisation).

Finally, the possibility of correlation in the error terms between alternatives should not be ignored, and here there is clearly a case to be made for a structure that groups together the SC alternatives, nesting them against the reference alternative.

In summary, the following issues potentially need to be addressed when using pivot style data that contains a reference alternative with invariant attributes across observations:

- different marginal sensitivities across reference and hypothetical SC alternatives, including differences in any random distributions
- different structural forms for the observed utility function
- separate coefficients for respondents with a zero value for a given attribute for the reference alternative
- asymmetrical preference formation relative to attribute values for the reference alternative
- correlation across replications in the error term for the first alternative

- heteroscedasticity between the reference and hypothetical SC alternatives
- different scale for respondents with a zero value for some attributes
- correlation in the error terms between hypothetical SC alternatives

The complexity in testing for these various phenomena varies significantly. In the following section, we will illustrate the effects of at least some of these phenomena in an empirical application.

3 Empirical example

In this section, we present the results for an empirical application testing various hypotheses set out in Section 2. We first provide an introduction to the data used in the analysis before turning our attention to the actual results.

3.1 Data

The data used in this analysis were collected in Sydney in 2004 as part of a wider study to obtain estimates of Value of Travel Time Savings (VTTS) of car drivers in the Sydney metropolitan area. For this paper, we use only data collected for respondents undertaking non-commuting trips.
As part of the initial study, a sampling strategy was employed whereby only respondents who had recently taken a trip within a particular corridor where a new toll road is proposed to be built were eligible to be surveyed. Recruitment took place using a computer aided telephone interview (CATI) employing a stratified geographical sampling frame drawn from a wide catchment area.
As part of the survey task, respondents were asked information about a recent trip that they had undertaken and which could potentially have used the proposed toll road had it been in existence. This information was then used to frame the context of the SC experiment. Based on the actual trip attribute levels reported, respondents were given 16 choice scenarios, each with three alternative routes described by time spent in free flow (FF) and slowed down time (SDT) travel conditions, travel time variability (VAR), running (petrol) costs (RC) and toll costs (TOLL). In all cases, the first alternative shown presented the respondent with the attribute levels provided as part of their recent trip as reported (the reference alternative). The remaining two alternatives represented competing hypothetical routes. As such, the reference alternative remained invariant across the 16 choice situations with only the levels of the hypothetical SC alternatives varying. Before commencing, respondents were given an example game to practice with, which was explained to them

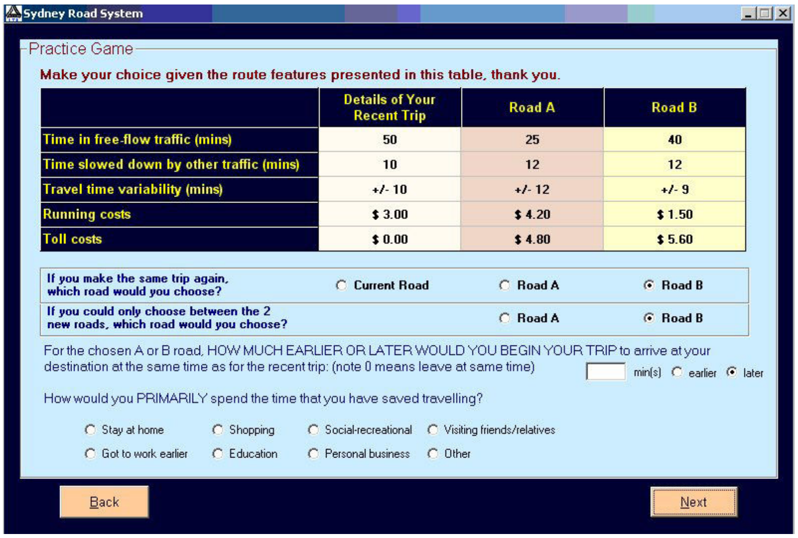

Figure 1: An example of a stated choice screen

by the interviewer. An example choice situation (taken from a practice game) is shown in Figure 1.

The SC experiment was constructed using efficient experimental design methods. For a review of efficient SC design methods, see Bliemer and Rose (2006) or Ferrini and Scarpa (2007). The final sample consisted of 205 effective interviews. For modelling purposes, this equates to 3,280 choice observations. For other recent applications using this dataset, see Hess et al. (2007) and Hess and Rose (2007).

3.2 Estimation results

We will now look in turn at the effects of allowing for the various phenomena discussed in Section 2. A mixture of Biogeme (Bierlaire, 2003), NLogit (Econometric Software, 2007) and purpose written Ox code (Doornik, 2001) was used for the estimation of the models.

3.2.1 Base model

During the specification search for the base model, several non-linear transformations of the attributes were attempted (mainly with the help of Box-Cox transforms), but no significant levels of non-linearity were retrieved. To this extent, a purely linear specification of the utility function was used. Specifically, we have that:

$$\begin{aligned}
V_{R,n} &= \delta_{\mathrm{R}} + \beta_{\mathrm{FF}}\mathrm{FF}_{\mathrm{R},n} + \beta_{\mathrm{SDT}}\mathrm{SDT}_{\mathrm{R},n} + \beta_{\mathrm{RC}}\mathrm{RC}_{\mathrm{R},n} \\
&+ \beta_{\mathrm{TOLL}}\mathrm{TOLL}_{\mathrm{R},n} + \beta_{\mathrm{VAR}}\mathrm{VAR}_{\mathrm{R},n} \\
V_{\mathrm{S}_1,n,t} &= \delta_{\mathrm{S}_1} + \beta_{\mathrm{FF}}\mathrm{FF}_{\mathrm{S}_1,n,t} + \beta_{\mathrm{SDT}}\mathrm{SDT}_{\mathrm{S}_1,n,t} + \beta_{\mathrm{RC}}\mathrm{RC}_{\mathrm{S}_1,n,t} \\
&+ +\beta_{\mathrm{TOLL}}\mathrm{TOLL}_{\mathrm{S}_1,n,t} + \beta_{\mathrm{VAR}}\mathrm{VAR}_{\mathrm{S}_1,n} \\
V_{\mathrm{S}_2,n,t} &= \beta_{\mathrm{FF}}\mathrm{FF}_{\mathrm{S}_2,n,t} + \beta_{\mathrm{SDT}}\mathrm{SDT}_{\mathrm{S}_2,n,t} + \beta_{\mathrm{RC}}\mathrm{RC}_{\mathrm{S}_2,n,t} \\
&+ +\beta_{\mathrm{TOLL}}\mathrm{TOLL}_{\mathrm{S}_2,n,t} + \beta_{\mathrm{VAR}}\mathrm{VAR}_{\mathrm{S}_2,n}. \qquad (6)
\end{aligned}$$

Here, the observed utility for the reference alternative (R) is again kept constant across replications for the same respondent, while the utilities for the two hypothetical SC alternatives (S_1 and S_2) vary over choice situations. Alternative specific constants (ASC) are included for the first two alternatives (δ_{R} and δ_{S_1}), and all marginal utility coefficients are generic across alternatives, with the associated attributes labelled as in Section 3.1.
The estimation results for this model are summarised in Table 1. All estimated parameters are statistically significant, including the two alternative specific constants, where these indicate, all else being equal, choice inertia as well as a reading from left to right effect. The implied trade-offs from these models show higher willingness to pay (WTP) for slowed down time than for free flow time, with the valuations being lower relative to toll than relative to running cost[5]. However, the results also show that the former of these two differences is only significant at the 91% level.

3.2.2 Model with differences in marginal sensitivities

Our next set of models move away from the assumption of equal response to attributes across all alternatives by estimating alternative specific taste coefficients. Here, no differences were identified between the two hypothetical SC alternatives, so that only two sets of coefficients were estimated, one set linked to the reference alternative (subscript R) and one set linked to the two hypothetical SC alternatives (subscript S). The estimation results for this model are summarised in Table 2.
In this model, all estimated coefficients remain statistically significant, with the exception of $\beta_{\mathrm{VAR,S}}$, the coefficient for variability in the travel time for hypothetical SC alternatives. With 5 additional parameters, the model obtains an improvement in log-likelihood (LL) by 13.59 units, which is significant above the 99% level of confidence.
Looking at the actual differences between coefficients, we observe that the difference between the coefficients for the two types of alternatives is significant above the 95% level for the two cost coefficients, while it is significant above the 90% level for the two travel time coefficients. The pattern of differences is bi-directional. For free flow time and tolls,

[5]Standard errors of the ratios were obtained using simulation methods taking into account the asymptotic normal distribution of estimators (cf. Armstrong et al., 2001).

Table 1: Estimation results for base model

Observations	3,280	
Parameters	7	
LL($\hat{\beta}$)	-2,395.88	
adj. $\rho^2(0)$	0.3332	
	est.	asy. t-rat.
δ_{R}	0.2860	3.60
δ_{S_1}	0.1480	2.51
β_{FF}	-0.0813	-17.87
β_{SDT}	-0.0926	-17.12
β_{RC}	-0.3640	-12.82
β_{TOLL}	-0.4430	-28.22
β_{VAR}	-0.0087	-2.50
WTP indicators (AUD/hr)	pt. est.	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	13.40	11.14
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	15.26	10.72
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	11.01	16.64
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	12.54	15.97
Diff. between WTP indicators	asy. t-rat.	
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$ *vs.* $\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	-1.71	
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$ *vs.* $\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	-1.73	
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$ *vs.* $\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	2.30	
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$ *vs.* $\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	2.28	

the sensitivity is lower for the reference alternative than for the two hypothetical SC alternatives, with the converse applying for slowed down time and running costs. Finally, no significant differences are observed between $\beta_{\mathrm{VAR,R}}$ and $\beta_{\mathrm{VAR,S}}$.

Some interesting differences also arise when looking at the implied WTP indicators. Here, the valuations of slowed down time remain higher than the valuations of free flow time, just as in the base model in Section 3.2.1. However, while for the hypothetical SC alternatives, the valuations relative to toll are still lower than those relative to running cost, the converse is now the case for the reference alternative. This comes as a result of the sharp drop in the sensitivity to tolls for the reference alternative.

Looking at the asymptotic t-ratios for the differences between WTP indicators, we can see that the WTP for reductions in free flow time relative to running cost is significantly higher for the hypothetical SC alternatives than for the reference alternative (the difference

Table 2: Estimation results for model with differences in marginal sensitivities

Observations	3,280		
Parameters	12		
LL($\hat{\beta}$)	-2,382.292		
adj. $\rho^2(0)$	0.3356		
	est.	asy. t-rat.	asy. t-rat. (Δ)
δ_{R}	0.3210	2.50	
δ_{S_1}	0.1510	2.54	
$\beta_{\text{FF,R}}$	-0.0734	-10.16	1.66
$\beta_{\text{FF,S}}$	-0.0833	-17.83	
$\beta_{\text{SDT,R}}$	-0.0993	-14.17	-1.75
$\beta_{\text{SDT,S}}$	-0.0915	-16.38	
$\beta_{\text{RC,R}}$	-0.4640	-8.30	-2.08
$\beta_{\text{RC,S}}$	-0.3620	-12.60	
$\beta_{\text{TOLL,R}}$	-0.3680	-12.65	2.82
$\beta_{\text{TOLL,S}}$	-0.4500	-26.81	
$\beta_{\text{VAR,R}}$	-0.0079	-2.19	0.55
$\beta_{\text{VAR,S}}$	-0.0150	-1.22	
WTP indicators (AUD/hr)	pt. est.	asy. t-rat.	
$\beta_{\text{FF,R}}/\beta_{\text{RC,R}}$	9.49	5.08	
$\beta_{\text{SDT,R}}/\beta_{\text{RC,R}}$	12.84	6.08	
$\beta_{\text{FF,R}}/\beta_{\text{TOLL,R}}$	11.97	8.06	
$\beta_{\text{SDT,R}}/\beta_{\text{TOLL,R}}$	16.19	9.64	
$\beta_{\text{FF,S}}/\beta_{\text{RC,S}}$	13.81	11.16	
$\beta_{\text{SDT,S}}/\beta_{\text{RC,S}}$	15.17	10.48	
$\beta_{\text{FF,S}}/\beta_{\text{TOLL,S}}$	11.11	16.69	
$\beta_{\text{SDT,S}}/\beta_{\text{TOLL,S}}$	12.2	15.35	
Diff. between WTP indicators	asy. t-rat.		
$\beta_{\text{FF,R}}/\beta_{\text{RC,R}}$ *vs.* $\beta_{\text{FF,S}}/\beta_{\text{RC,S}}$	-2.26		
$\beta_{\text{SDT,R}}/\beta_{\text{RC,R}}$ *vs.* $\beta_{\text{SDT,S}}/\beta_{\text{RC,S}}$	-1.12		
$\beta_{\text{FF,R}}/\beta_{\text{TOLL,R}}$ *vs.* $\beta_{\text{FF,S}}/\beta_{\text{TOLL,S}}$	0.68		
$\beta_{\text{SDT,R}}/\beta_{\text{TOLL,R}}$ *vs.* $\beta_{\text{SDT,S}}/\beta_{\text{TOLL,S}}$	2.83		

of AUD4.32 is significant at the 97% level), while the converse is the case when looking at the valuation of slowed down time relative to road tolls (the reference alternative valuation is AUD3.99 higher, with a confidence level of 99%). The differences in the remaining two WTP indicators are of lower absolute value, and are not significant at the usual levels of confidence.

3.2.3 Models with different structural forms for observed utility function

Various non-linear specifications of the utility functions were attempted for a model with group specific coefficients, based on the model from Section 3.2.2. However, just as for the base model in Section 3.2.1, no evidence of non-linear responses was found.

3.2.4 Models with separate groups for respondents with zero values for reference alternative

In pivot designs, the possibility exists that some of the respondents are always presented with a zero value for one of the attributes of the reference alternative. In the present context, this would for example arise in the case of respondents who did not pay a toll on their reference trip.

In the data used here, 1 respondent had a zero value for the free flow time attribute (i.e. a fully congested trip), while 16 respondents had a zero value for the slowed down time attribute (i.e. a congestion-free trip). For either of these two attributes, no differences were observed between the two segments with a zero or non-zero value, in the observed or unobserved part of utility.

More promising however is the fact that 72 respondents, i.e. just over a third of the sample, did not pay a toll on their reference trip. It is a major assumption to suggest that these respondents behave identically to respondents who did pay a toll on their reference trip.

Two approaches were used to explore the differences between these two groups. In the first approach, generic coefficients were used across the two groups, but the utility functions for respondents who paid a toll were multiplied by a scale parameter (α). If the estimate of this scale parameter is larger than 1, it indicates a lower relative weight for the unobserved part of utility for these respondents. In the second approach, we move away from the assumption that the scale differences are uniform by estimating separate models in the two groups.

The estimation results for the resulting three models are summarised in Table 3. The first model compares directly to the base model in Table 1. Here, the addition of one additional parameter (α) leads to a highly significant increase in log-likelihood by 8.94 units. The actual value of α suggests that the variance of the unobserved part of utility is larger for respondents with a non-zero toll for the reference trip than for respondents with a zero toll. A very easy explanation arises for this in that respondents with a zero toll are far more likely to choose their reference trip than are respondents who did pay such a toll. While, in the overall sample, the market share for the reference alternative is 37.16%, it is much higher for respondents with a zero reference toll, at 58.42%, but is only 25.66% for respondents with a non-zero reference toll. This suggests that the behaviour of respondents

Table 3: Estimation results for models with grouping depending on whether or not a toll was paid on reference trip

	All		**Zero ref. toll**		**Non-zero ref. toll**	
Observations	3,280		1,152		2,128	
Parameters	8		7		7	
LL($\hat{\beta}$)	-2,386.94		-710.57		-1,669.83	
adj. $\rho^2(0)$	0.3350		0.4330		0.2830	
	est.	asy. t-rat.	est.	asy. t-rat.	est.	asy. t-rat.
δ_{R}	0.2590	2.74	0.1270	0.87	0.3290	3.33
δ_{S_1}	0.1810	2.52	0.2440	1.93	0.1180	1.76
β_{FF}	-0.0995	-11.31	-0.1000	-9.63	-0.0765	-15.40
β_{SDT}	-0.1130	-11.59	-0.1020	-8.6	-0.0894	-15.04
β_{RC}	-0.4480	-9.56	-0.5510	-6.12	-0.3370	-11.48
β_{TOLL}	-0.5170	-16.02	-0.5480	-20	-0.4020	-16.71
β_{VAR}	-0.0098	-2.34	-0.0127	-1.81	-0.0091	-2.27
$\alpha^{\dagger}$	0.7620	-3.18	-	-	-	-
WTP ind. (AUD/hr)	pt. est.	asy. t-rat.	pt. est.	asy. t-rat.	pt. est.	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	13.33	11.20	10.89	4.65	13.62	9.94
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	15.13	10.95	11.11	4.87	15.92	9.56
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	11.55	16.31	10.95	9.93	11.42	12.24
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	13.11	16.02	11.17	9.04	13.34	11.95

†: asymptotic t-ratio for α calculated with respect to 1

with a zero reference toll is easier to model by being more deterministic (captive to reference alternative). The actual WTP indicators from this model are comparable to those from the base model.

The results from the two separate models confirm the findings of the joint model. The model performance in terms of the adjusted ρ^2 measure is far superior in the model for respondents with a zero reference toll than in the model for respondents with a non-zero reference toll[6]. We obtain a combined LL of $-2,380.40$ for the two models, meaning that, at the cost of 7 additional parameters compared to the base model, we obtain a statistically significant gain in LL by 15.48 units. As expected, the WTP indicators relative to toll are lower for respondents with a zero reference toll; these respondents have a higher aversion to road tolls. However, there is also a reduction in the WTP indicators relative to running costs, which suggests that respondents who chose an untolled option on their reference

[6]The adjusted ρ^2 measure for the combined model is 0.3355.

journey have a higher overall cost sensitivity, independent of the actual cost component.

3.2.5 Model with asymmetrical preference formation

The next model structure we explore allows for the possibility that respondents treat the attribute levels of hypothetical SC alternatives not as absolute values, but relative to the values of the reference alternative. In this case, there is a possibility of asymmetrical preference formation, with respondents reacting differently to gains and losses relative to the attribute values for this reference trip. For an in-depth discussion of this topic, see Hess et al. (2007).

The utility function for the reference alternative now contains only a constant, while the utility functions for the two hypothetical SC alternatives contain two coefficients for each attribute, one associated with increases relative to the reference trip (e.g. $\beta_{\text{FF,inc}}$) and one associated with decreases (e.g. $\beta_{\text{FF,dec}}$). The estimation results for this model are summarised in Table 4.

With 5 additional parameters compared to the base model, we obtain a statistically significant increase in LL by 19.42 units. Coefficients associated with reductions in attributes take on positive signs, with the converse being the case for increases. The only exception is the positive value for $\beta_{\text{VAR,inc}}$, where this is however not statistically significant. All other estimated parameters are significant at high levels of confidence.

Looking in detail at the response to increases and decreases, we observe a perfectly symmetrical response in the case of running costs, i.e. an increase in running cost by AUD1 is valued as negatively as a reduction in running cost by AUD1 is valued positively. For free flow time, the difference between increases and decreases (after taking into account the sign differences) is only significant at the 86% level but suggests that increases in free flow time are valued more negatively than reductions are valued positively.

For slowed down time and toll costs, the differences between increases and decreases are highly significant, where, by coincidence, the value of the asymptotic t-ratio is identical. For toll costs, there is clear evidence that increases are valued more negatively than decreases, which is consistent with intuition, and which is also reflected in the high aversion by respondents with a zero reference toll to move to a tolled SC alternative. However, for slowed down time, the surprising observation is made that decreases are valued far more positively than increases are valued negatively. A potential explanation for this could be in the way respondents trade off free flow time and slowed down time, in that an increase in slowed down time was in general associated with a reduction in free flow time, where the magnitude of the latter was greater. So respondents can be seen to accept increases in slowed down time in return for reductions in free flow time.

It should also be said that the asymmetries are at least in part a result of the design, where gains and losses were not presented in a symmetrical fashion. As such, very few increases in variability were presented, explaining the low significance level of $\beta_{\text{VAR,inc}}$. Furthermore,

Table 4: Estimation results for model with asymmetrical preference formation

Observations	3,280
Parameters	12
LL($\hat{\beta}$)	-2,376.464
adj. $\rho^2(0)$	0.3372

	est.	asy. t-rat.	asy. t-rat. $(\Delta)^\dagger$
δ_{R}	0.2670	2.34	
$\delta_{\mathrm{S_1}}$	0.1590	2.72	
$\beta_{\mathrm{FF,dec}}$	0.0736	10.13	1.48
$\beta_{\mathrm{FF,inc}}$	-0.1050	-6.57	
$\beta_{\mathrm{SDT,dec}}$	0.1180	12.92	-3.99
$\beta_{\mathrm{SDT,inc}}$	-0.0313	-2.02	
$\beta_{\mathrm{RC,dec}}$	0.3740	7.24	0.09
$\beta_{\mathrm{RC,inc}}$	-0.3640	-4.88	
$\beta_{\mathrm{TOLL,dec}}$	0.3320	10.56	3.99
$\beta_{\mathrm{TOLL,inc}}$	-0.5060	-22.10	
$\beta_{\mathrm{VAR,dec}}$	0.0077	2.18	-0.18
$\beta_{\mathrm{VAR,inc}}$	0.0045	0.07	

WTP indicators (AUD/hr)	pt. est.	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	12.13	4.34
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	19.45	4.34
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	8.73	9.61
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	13.99	11.93

WTA indicators (min/AUD)	pt. est.	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	3.56	5.57
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	11.95	1.98
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	3.16	5.83
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	10.61	2.02

Diff. between WTP and WTA	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	-0.88
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	2.61
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	-2.69
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	2.31

†: sign differences taken into account in calculation of standard errors for differences for road tolls, any reductions presented were always a 100% reduction, in which case the marginal gains of a reduction in tolls (per AUD) could indeed be expected to be lower, as observed. Independently of the reasons, the analysis presents clear evidence of asymmetries and suggests that a symmetrical specification is not appropriate.

With separate coefficients estimated for increases and decreases, two different trade-offs can be calculated, namely the willingness to pay (WTP) for reductions in travel time, and the willingness to accept (WTA) increases in travel time in return for reductions in travel cost. Here, the asymmetry for the marginal utility coefficients translates into similar asymmetry patterns for the actual trade-offs, where these are significant with the exception of the trade-offs between free flow time and running costs.

3.2.6 Model with cross-replication correlation in error term for reference alternative

Our next model allows for correlation across replications in the error terms for the reference alternative, with identically but cross-sectionally distributed error terms for the remaining two alternatives ensuring homoscedasticity.

The estimation results for this model are summarised in Table 5. Compared to the base model, we obtain an increase in LL by 70.84 units, which, at the cost of just one additional parameter is highly significant. These gains are at least in part due to the fact that this model recognises the panel nature of the data, which also leads to an upwards correction of the standard errors, with the actual WTP measures remaining largely unaffected. The estimate for σ is highly significant, and suggests a correlation of 0.26 in the error terms for the reference alternative across replications for the same respondent. In a hypothetical panel MMNL model estimated in parallel[7], this correlation measure drops to 0.21. This would suggest that the correlation picked up in the model in Table 5 is primarily due to the fact that we have multiple observations per respondent, although it is not possible to completely discount the possibility that some of the correlation is caused by the invariant attributes nature of the reference alternative.

3.2.7 Model with heteroscedasticity across alternatives

To allow for heteroscedasticity across alternatives, an ECL formulation of the MMNL model was used, with independently distributed Normal variates with a mean of zero included in the utility functions (cf. Walker, 2001). The initial identification search showed that the most appropriate normalisation is to set the standard deviation for the error component for the reference alternative to zero. As such, only two error components were used, one for each of the two hypothetical SC alternative. With no observable difference between the two error components, the two standard deviations were constrained to the same value.

[7] I.e. allowing for the same correlation effect for all alternatives, not just the reference alternative.

Table 5: Estimation results for model with cross-replication correlation in error term for reference alternative

Observations	3,280	
Parameters	8	
$LL(\hat{\beta})$	-2325.04	
adj. $\rho^2(0)$	0.3530	
	est.	asy. t-rat.
δ_{R}	0.3140	2.38
δ_{S_1}	0.1710	2.52
β_{RC}	-0.4150	-9.12
β_{FF}	-0.0961	-11.33
β_{SDT}	-0.1040	-11.38
β_{TOLL}	-0.5140	-13.37
β_{VAR}	-0.0108	-1.92
σ	0.7570	12.56
WTP indicators (AUD/hr)	pt. est.	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	13.89	7.10
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	15.04	7.54
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	11.22	9.94
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	12.14	9.67

The results for this heteroscedastic model are summarised in Table 6. With one additional parameter compared to the base model, we obtain a statistically significant improvement in LL by 3.43 units. All estimated parameters are statistically significant, and the WTP indicators differ only marginally from the base model. With the variance of the unobserved part of utility for the reference alternative being fixed to $\frac{\pi^2}{6}$, the variance for the hypothetical SC alternative, at $\frac{\pi^2}{6} + 1.5^2$, is higher by over 136%. This strongly suggests the presence of heteroscedasticity[8].

3.2.8 Model with cross-alternative correlation in error terms

Attempts were made to estimate various NL structures on the data. All models collapsed back to a MNL specification with the exception of a model nesting together the two hypothetical SC alternatives. However, in this model, the structural parameter took on a value larger than 1, with the fit being very similar to the heteroscedastic ECL model. It is not

[8]These results were confirmed through the estimation of a HEV model, with detailed results available on request.

Table 6: Estimation results for model with heteroscedasticity across alternatives

Observations	3,280	
Parameters	8	
$\mathrm{LL}(\hat{\beta})$	-2,392.45	
adj. $\rho^2(0)$	0.3340	
	est.	asy. t-rat.
δ_{R}	0.5340	2.58
δ_{S_1}	0.2480	2.21
β_{RC}	-0.5010	-4.76
β_{FF}	-0.1130	-4.92
β_{SDT}	-0.1260	-4.93
β_{TOLL}	-0.6010	-5.35
β_{VAR}	-0.0104	-2.23
σ_{R}	0	-
σ_{S}	1.5	2.36
WTP indicators (AUD/hr)	pt. est.	asy. t-rat.
$\beta_{\mathrm{FF}}/\beta_{\mathrm{RC}}$	13.53	10.30
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{RC}}$	15.09	10.26
$\beta_{\mathrm{FF}}/\beta_{\mathrm{TOLL}}$	11.28	15.15
$\beta_{\mathrm{SDT}}/\beta_{\mathrm{TOLL}}$	12.58	15.03

immediately clear how the results from this model should be interpreted although there is a possibility that they also point towards a heteroscedastic specification.

4 Summary, conclusions and directions for future research

This paper has discussed the issue of what special precautions might need to be taken when estimating models on pivot style data, where the choice sets include a reference alternative with invariant attributes across replications. The theoretical discussions in the paper identify a number of possible phenomena that may act on the observed or unobserved utility components.

The application conducted on a representative pivot style dataset collected in Sydney confirms the theoretical suspicions. As such, different advanced model specifications offer statistically significant gains in model fit when compared to the base model. In particular, the models suggest:

- the presence of differences in the marginal sensitivities between the reference alternative and the hypothetical SC alternatives
- the presence of differences in the marginal sensitivities for respondents who did not pay a toll on the reference trip
- the presence of differences in the variance of unobserved utility components for respondents who did not pay a toll on the reference trip
- the prevalence of asymmetrical preference formation around the attribute values of the reference alternative
- the existence of correlation in the error terms across replications for the reference trip
- the existence of differences in the variance of error terms between the reference alternative and the hypothetical SC alternatives

There is clearly a distinct possibility of some confounding across models, such that not all of the above factors may act to the same degree on the data. Further work is required to establish the extent of confounding, through specification of an appropriate joint model. Furthermore, even though the evidence in this paper is limited to one dataset, the findings do clearly suggest that special care is required when estimating models on pivot style data. With the increasing popularity of such data, these warnings are very timely indeed.

5 Acknowledgements

The majority of this work was carried out during a visit by the first author to the Institute of Transport and Logistics Studies at the University of Sydney, funded by a Faculty of Economics and Business Visiting Scholar Grant.

References

Armstrong, P., Garrido, R. A., Ortúzar, J. de D., 2001. Confidence interval to bound the value of time. Transportation Research Part E: Logistics and Transportation Review 37 (1), 143–161.

Bhat, C. R., 1995. A heteroscedastic extreme value model of intercity travel mode choice. Transportation Research Part B: Methodological 29 (6), 471–483.

Bierlaire, M., 2003. BIOGEME: a free package for the estimation of discrete choice models. Proceedings of the 3^{rd} Swiss Transport Research Conference, Monte Verità, Ascona.

Bliemer, M. C. J., Rose, J. M., 2006. Designing Stated Choice Experiments: State-of-the-art. paper presented at the 11^{th} International Conference on Travel Behaviour Research, Kyoto, Japan.

Camerer, C. F., Hogarth, R. M., 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. Journal of Risk and Uncertainty 19 (1-3), 7–42.

Diamond, P. A., Hausman, J. A., 1994. Contingent valuation: Is some number better than no number? Journal of Economic Perspectives 8 (4), 45–64.

Ding, M., 2007. An incentive aligned mechanism for conjoint analysis. Journal of Marketing Research XLIV, 214–223.

Doornik, J. A., 2001. Ox: An Object-Oriented Matrix Language. Timberlake Consultants Press, London.

Econometric Software, 2007. Nlogit 4.0. Econometric Software, New York and Sydney.

Ferrini, S., Scarpa, R., 2007. Designs with a-priori information for nonmarket valuation with choice-experiments: a monte carlo study. Journal of Environmental Economics and Management, 53, 342–363.

Gilboa, I., Schmeidler, D., Wakker, P., 2002. Utility in case-based decision theory. Journal of Economic Theory 105, 483–502.

Hensher, D. A., 2004. Accounting for stated choice design dimensionality in willingness to pay for travel time savings. Journal of Transport Economics and Policy 38 (2), 425–446.

Hensher, D. A., 2007. Joint estimation of process and outcome in choice experiments and implications for willingness to pay. Journal of Transport Economics and Policy, forthcoming.

Hensher, D. A., Greene, W. H., 2003. The Mixed Logit Model: The State of Practice. Transportation 30 (2), 133–176.

Hensher, D. A., Rose, J. M., 2007. Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study. Transportation Research Part A: Policy and Practice 41 (5), 428–443.

Hess, S., Rose, J. M., 2007. Intra-respondent taste heterogeneity in instantaneous panel surveys. paper to be presented at the 11^{th} triennial WCTR conference, July 2007, Berkeley, CA.

Hess, S., Rose, J. M., Hensher, D. A., 2007. Asymmetrical Preference Formation in Willingness to Pay Estimates in Discrete Choice Models. Transportation Research E, forthcoming.

Kahnemann, D., Tversky, A., 1979. Prospect theory: an analysis of decisions under risk. Econometrica 47 (2), 263–291.

List, J. A., 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? evidence from field auctions for sportscards. American Economic Review 91 (5), 1498–1507.

Louviere, J. J., Hensher, D. A., Swait, J., 2000. Stated Choice Models: Analysis and Application. Cambridge University Press, Cambridge.

Rose, J. M., Bliemer, M. C. J., Hensher, D. A., Collins, A. C., 2008. Designing Efficient Stated Choice Experiments Involving Respondent Based Reference Alternatives. Transportation Research Part B: Methodological, forthcoming.

Rousseas, S. W., Hart, A. G., 1951. Experimental verification of a composite indifference map. Journal of Political Economy 59, 288–318.

Scarpa, R., Willis, K. G., Acutt, M., 2007. Valuing externalities from water supply: Status quo, choice complexity and individual random effects in panel kernel logit analysis of choice experiments. Journal of Environmental Planning and Management 50 (4), 449–466.

Starmer, C., 2000. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. Journal of Economic Literature XXXVIII, 332–382.

Thurstone, L., 1931. The indifference function. Journal of Social Psychology 2, 139–167.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, MA.

Train, K., Wilson, W. W., 2007. Estimation on stated-preference experiments constructed from revealed-preference choices. Transportation Research Part B: Methodological, forthcoming.

Walker, J., 2001. Extended discrete choice models: Integrated framework, flexible error structures, and latent variables. Ph.D. thesis, MIT, Cambridge, MA.

Wallis, W. A., Friedman, M., 1942. The empirical derivation of indifference functions. In: Lange, O., McIntyre, F., Yntema, T. O. (Eds.), Studies in Mathematical Economics and Econometrics in memory of Henry Schultz. University of Chicago Press, Chicago, IL, pp. 175–189.