EJTIR

Issue 14(4), 2014
pp. 349-361
ISSN: 1567-7141
tlo.tbm.tudelft.nl/ejtir

Impact of unimportant attributes in stated choice surveys

Stephane Hess[1]
Institute for Transport Studies, University of Leeds, United Kingdom.

Despite growing interest in the notion that respondents in stated choice surveys may make their decisions on the basis of only a subset of the presented attributes, the impact of any *unimportant* attributes on the estimates of other valuations is somewhat unclear. This paper presents evidence from a two stage survey where the second stage eliminates attributes deemed unimportant in the first stage. Our analysis shows no evidence of systematic differences between the results of the two stages. This leads to the conclusion that, up to a point where respondent burden may become an issue, analysts should include all attributes that *may* be relevant, and allow the respondent to filter out those that play no role.

Keywords: *information processing, attribute ignoring, non-attendance, respondent burden, attribute relevance, stated choice.*

1. Introduction

Over the last few years, a substantial amount of research effort has gone into investigating the possibility of individual respondents using different strategies in processing the information describing the scenarios they face in stated choice (SC) surveys. A comprehensive overview of this work is given in Hensher (2010). The work is especially important in the context of monetary valuation studies, such as for example in the appraisal for new infrastructure of policy schemes. The main emphasis has been on the notion that some respondents may ignore certain attributes, often described as attribute non-attendance. While the origins of this work are in the transport field, there are now applications across numerous different fields, with some examples being the work of Hensher et al. (2005), Hensher (2006), Hensher et al. (2007), Hole (2011), Mariel et al. (2013), Balcombe et al. (2011), and Scarpa et al. (2011). There has also been some interest in looking at whether specific individuals may process *similar* attributes jointly rather than separately (see e.g. Layton and Hensher, 2010).

Stated choice surveys now routinely include questions asking respondents whether they ignored a given attribute. Early work in this context deterministically imposed the processing rule on the basis of such information, but it is fairly straightforward to see that this leads to issues with endogeneity, given the likely correlation between respondent reported processing information and other unmodelled components. Additionally, the question arises as to how reliable this information is, with work repeatedly showing non-zero coefficients for such respondents (cf. Hess and Hensher, 2010; Alemu et al., 2013; Carlsson et al., 2010). Later work made use of more robust approaches that treat the processing strategies as latent components (see e.g. Hess and Rose, 2007; Hensher, 2008; Dumont et al., 2011), and/or offer more flexible approaches to

[1] A: University Road 36-40, Leeds LS2 9JT, United Kingdom T: +31 152 788 546 F: +31 152 788 546 E: s.hess@its.leeds.ac.uk

capturing the possible confounding between non-attendance and low sensitivity (e.g. Hess et al., 2013; Collins et al., 2013).[2]

As is evident from the above, research on how to adequately capture differences in processing strategies in data, and how to accommodate them in our models, has made substantial headway in recent years. Two quite different views exist as to why attribute non-attendance may arise. First, it is seen by some as a sign that SC surveys are too complex, with overburdening leading to some of the information being ignored. Unfortunately, such arguments still drive the widespread use of very simple SC scenarios, often based on just two attributes (and two alternatives), especially in applied work; such surveys are clearly not in line with the real world complexity of human decisions. In contrast with this is the view that relevance matters more than respondent burden; individuals can be trusted to determine which attributes matter to them, and the impact of presenting attributes that are *unimportant* for some respondents is less severe than if attributes that are only *important* for some respondents are excluded from the choice sets for all respondents (cf. Hensher, 2006). The view here is thus that it is important to present all attributes that could possibly be important and let respondent make a choice as in real life, possibly disregarding some attributes. This is reflected in the extensive use by respondents of data processing tools in the work of Collins et al. (2012).

Independently of the precise cause of such attribute non-attendance, the impact of any *unimportant* attributes on the estimates of other valuations is somewhat unclear, despite some evidence to the contrary in Alemu et al. (2013). It is however of crucial importance. Indeed, analysts who are concerned about respondent burden will not want to include additional attributes if evidence suggests that such attributes may have an undue influence on core valuations. Conversely, those concerned with presenting all attributes that may be relevant[3] will be reassured if there is evidence of minimal impact on those attributes actually used by the respondent. While the main interest in the literature has thus been the impact on sample level estimates of some respondents in the sample ignoring specific attributes, the emphasis in this paper is the impact on individual behaviour of presenting attributes that do not matter to this individual, or at least are of low importance.

It is possible to come up with a number of different reasons why the presence of *unimportant* attributes may have an impact on the valuation of the *important* attributes. The first of these relates respondent burden. It is entirely reasonable to hypothesise that, when respondents become too cognitively burdened, as may be the case in the presence of a large number of attributes, they may be more likely to use simplifying heuristics for making their choices, leading to reduced response quality. A similar reasoning, which would also lead to reduced data quality, is that the presence of several unimportant attributes may reduce the attention that respondents pay to the important ones, again adding more noise to the decision process. Finally, it is also possible to imagine impacts on the substantive model outputs, i.e. not just error variance. Indeed, the presence of a larger number of attributes may reduce the marginal willingness-to-pay for individual components of an alternative - this may in fact be in line with real world behaviour, where the options we choose are made up of many more components than in simplified stated choice settings.

The detailed checking of these hypotheses, and the many other ones that are possible, is beyond the scope of the present paper and would require further information at the respondent level, including detailed post-survey questioning. It remains however an important area for future work. The specific question the paper seeks to address is thus *whether* including attributes that

[2] It should be noted that the early work by Train and Sonnier (2005) is also relevant in this context.

[3] It should be acknowledged that it will never be possible to include all attributes that may be relevant to any given person in a population. What we refer to here is the inclusion of all attributes that are of interest to the study at hand - any work is then still based on the typical stated choice assumption that anything that is not included is equal across all of the alternatives.

may not be relevant to a given respondent unduly affects that respondent's behaviour, not *why* that may be the case. As such, rather than focussing on how to accommodate non-attendance at the modelling end, this paper aims to add some insights to the discussion about whether surveys should focus on core attributes and give a simplified representation of reality, or whether they should include everything that may be relevant.

Specifically, we compare the results from two separate SC components, one including the full set of attributes considered in the study, the other being limited to the subset of attributes deemed to be of interest to a given respondent. Reassuringly, no conclusive evidence is found that presenting *unimportant* attributes unduly affects behaviour. This suggests that the impact of such attributes on other valuations is minimal at best and leads to the conclusion that, up to a point where respondent burden may become an issue, analysts should include all attributes that may be relevant, and allow the respondent to filter out those that play no role.

Before proceeding with the remainder of the paper, a final question to address is whether non-attendance is restricted to the SC context or reflects a real world characteristic. The view of this author is that non-attendance in real world scenarios is potentially less likely for continuous attributes, where insufficient ranges may be the cause for any non-attendance in a hypothetical context, a point reinforced by Alemu et al. (2013). On the other hand, a different picture arises for quality of service attributes with simple present/absent levels; here, it is conceivable that subsets of respondents are indifferent to such attributes independently of other attributes. This would then clearly also affect continuous attributes linked to discrete attributes (e.g. price for wifi would be ignored by those ignoring wifi provision). It should also be acknowledged that there is some evidence from real world data (cf. Scarpa et al., 2012; Morkbak et al., 2013).

The remainder of this paper is organised as follows. The next section presents the data collected for this study and the empirical framework used for the analysis. This is followed in Section 3 by the results from the empirical analysis. Finally, Section 4 presents the conclusions from the work.

2. Survey work and empirical framework

The empirical framework for this study consisted of conducting a two stage SC survey. In the first stage, respondents are faced with a set of SC scenarios where the alternatives are described by the full set of attributes considered in this study. After completion of this initial stage, respondents are given the option of specifying whether they ignored any attributes in the first stage. The second stage then presents each respondent with choice scenarios from a new design, using only those attributes specified as *relevant* by the respondent. The actual analysis looks at differences between the two stages, and in particular whether the inclusion of *unimportant* attributes in the first stage led to a different valuation for those attributes that matter to the respondent.

At this stage, an important question arises. As mentioned in the introduction, doubts have been expressed as to the validity of respondent stated information in relation to attribute non-attendance, and the importance of recognising that non-attendance may only be partial rather than complete for some respondents, i.e. apply to some tasks but not all (cf. Hess and Hensher, 2010; Alemu et al., 2013; Carlsson et al., 2010). Additionally, the use of such information in models arguably puts the analyst at risk of endogeneity bias, given the likely correlation between the answers to such questions and other unobserved components. It should first be noted that the analysis in this paper is purely exploratory with no interest in producing forecasts or unbiased monetary valuations. Secondly, the aim was to collect data from both stages in a single sitting, so as to avoid any impacts on results by a gap between the two surveys, and also to counter the resulting large drop-out of respondents that is likely in two-stage surveys. Nevertheless, we acknowledge that this on the other hand creates more scope for carry-over effect.

To avoid the use of stated non-attendance strategies with the present data would have meant relying on retrieved processing strategies, as in the work of Hess and Hensher (2010). This would however have led to the requirement for model estimation work in between the two stages, meaning that a single sitting would no longer have been an option. Furthermore, it is important to recognise that it is impossible to retrieve the actual processing strategy used by a given respondent with certainty, and we can only state that a respondent ignored a specific attribute up to a probability. This would thus also mean assigning respondents to the different versions of the second stage on a probabilistic basis. From this perspective, making use of the stated non-attendance information is preferable, even unavoidable, despite the obvious caveats.

The data for the present paper was collected using an internet based SC survey conducted in January 2010. The survey was framed in a rail travel context, with only respondents who had completed a journey of at least one hour in the last year being eligible to participate. In each scenario, a respondent was faced with a choice between three alternatives, described by five attributes, namely fare (£), travel time (minutes), the guarantee of a reserved seat (yes/no), the provision of free wifi (yes/no), and whether the given option allows for ticket flexibility (yes/no), for example in terms of rebooking on a different train. For fare and travel time, the attribute values were pivoted around respondent reported reference values, with variations between -20% and +20% for time, and -15% and +15% for fare. The survey was based on a D-efficient design generated in NGene (Choicemetrics, 2010), where in this first stage, each respondent was faced with eight choice tasks, with the attributes presented being pivoted around the levels of a recent trip for the given respondent.

After completion of this first stage, respondents were asked whether they had ignored any of the attributes across the eight tasks. Here, a decision was taken to limit this to the three qualitative attributes, working on the reasonable assumption that time and costs are core attributes that matter to some extent to all respondents. Sufficiently wide ranges were used in the data to ensure that this was the case. For the purposes of the present study, these questions were asked at the end of the stage, thus relating to all choice tasks, rather than using a choice task specific questioning approach. This is partly motivated on the grounds that non-attendance for qualitative attributes is less likely to be choice task specific than would be the case for continuous attributes, given that, for the latter, there is scope for smaller differences between alternatives depending on the levels used in a given task.[4]

By focussing on potential non-attendance for the three qualitative attributes only, eight possible classifications of respondents arise, as summarised in Table 1. This shows that while just under a quarter of respondents stated that they based their choices on all five attributes, almost 16% stated that they had ignored all three qualitative attributes, where the highest rate of stated non-attendance applies to the provision of free wifi.

Table 1. Classification of respondents on the basis of stated non-attendance information

Group	Strategy	Share of sample
1	All attributes considered	24.13%
2	Ignore only seat reservation	2.51%
3	Ignore only wifi	29.80%
4	Ignore only ticket flexibility	5.79%
5	Ignore seat reservation & wifi	9.39%
6	Ignore seat reservation & ticket flexibility	1.97%
7	Ignore wifi & ticket flexibility	10.59%
8	Ignore seat reservation & wifi & ticket flexibility	15.82%

[4] See also Carlsson et al. (2010) for a discussion on non-attendance at the choice task level *vs.* overall non-attendance.

For the second stage, eight separate designs were generated, in line with the eight groups listed in Table 1. Each respondent was then assigned to the appropriate stage and was faced with six choice tasks from this new design - the lower number was chosen so as to reduce possible respondent burden but also given the lower number of combinations possible in some of the stage 2 experiments. A final sample of 916 respondents was obtained, giving 7,328 observations for the first stage and 5,496 observations across the eight different versions of the second stage.

3. Empirical results

3.1 Analysis of data from stage 1

As a first step, models were estimated only on the data from the first stage, i.e. the eight choice scenarios per respondent in which all five attributes were included. A preliminary analysis indicated the presence of decreasing marginal time and cost sensitivities, leading to the use of a natural logarithm transform for the fare and travel time attributes. The actual analysis of the data is based on simple Multinomial Logit (MNL) models, using a panel specification of the sandwich estimator to account for the repeated choice nature of the data when calculating the covariance matrix (cf. Daly and Hess, 2011). The use of simple MNL models is justified in the context of a study looking at overall effects, but is also in part motivated by the small sample sizes in some of the subgroups for stage 2.

Table 2. Estimation results for stage 1 models

	Generic		Non-attendance Group separate		Non-attendance Group zero	
Obs.	7,328		7,328		7,328	
Resp.	916		916		916	
LL	-5,685.90		-5465.60		-5491.00	
Par.	7		10		7	
Adj. ρ^2	0.2929		0.3199		0.3171	
	Est.	t-rat.	Est.	t-rat.	Est.	t-rat.
$B_{\text{L-FARE}}$	-6.835	27.2	-70.070	27.1	-65.960	30.1
$B_{\text{L-time}}$	-4.949	24.9	-50.830	25.0	-47.220	27.7
$B_{\text{flex, attend}}$	0.6555	7.9	0.8628	9.3	0.7101	9.4
$B_{\text{seat, attend}}$	0.8306	15.2	11.350	16.0	12.100	19.0
$B_{\text{wifi, attend}}$	0.4778	9.5	10.050	12.2	0.9201	12.3
$B_{\text{flex, ignore}}$	-		0.2718	2.7	0	
$B_{\text{seat, ignore}}$	-		0.2214	3.5	0	
$B_{\text{wifi ,ignore}}$	-		0.2003	3.9	0	
WTP time (£/hr)	8.69	28.3	8.70	28.2	8.59	26.3
WTP flex, attend (£)	3.36	9.1	4.31	10.6	3.77	10.0
WTP seat, attend (£)	4.25	11.7	5.67	13.3	6.42	16.3
WTP wifi, attend (£)	2.45	10.7	5.02	13.2	4.88	12.3
WTP flex, ignore (£)	-		1.36	2.8	0	
WTP seat, ignore (£)	-		1.11	3.3	0	
WTP wifi, ignore (£)	-		1.00	4.2	0	

The estimation results for the base model for stage 1 are summarised in the first column of Table 2. This shows the expected negative effects of increases in (the logarithm of) fare β_{L-fare} and time β_{L-time}, with positive effects for ticket flexibility, guaranteed seat reservation and the provision of free wifi. Table 2 also shows willingness-to-pay (WTP) indicators for this model, where these are calculated at the average chosen fare in the data, which is £35, while, for the value of time, we used the average ratio of £0.2/minute, as observed in the sample data.

As a next step, we make use of respondent stated non-attendance strategies, with a view to testing their empirical correctness, notwithstanding the earlier comments about risk of endogeneity. Given the doubts expressed by Hess and Hensher (2010), Carlsson et al. (2010) and Alemu et al. (2013) as to the validity of such stated non-attendance information, we estimate two separate models. In one of the models, we estimate separate coefficients for the three qualitative attributes depending on whether a respondent stated that they had ignored the attribute, while, in the other model, we impose a value of zero on such coefficients, thus assuming that the stated non-attendance strategies are in fact valid.

Both models that take into account the stated non-attendance information obtain improvements in fit over the base models. We see gains by 220.3 units in log-likelihood (LL) for the model using separate coefficients in the non-attendance group, at the cost of three additional parameters. On the other hand, the gains for the model setting the coefficients to zero in the non-attendance group are slightly smaller. This is directly due to the fact that the true coefficient values in the non-attendance groups are not in fact equal to zero. Indeed, in line with the results by Hess and Hensher (2010), Carlsson et al. (2010) and Alemu et al. (2013), we see that the three coefficients in question ($\beta_{flex,ignore}$, $\beta_{seat,ignore}$ and $\beta_{wifi,ignore}$) are still significantly different from zero, albeit that their values are much smaller than in the non-ignoring groups. This is consistent with the earlier comment that respondents who state that they had ignored a certain attribute may simply have assigned it a lower value. Similarly, it is clearly also possible that full non-attendance still applied to some respondents within this group. The differences between the groups are also reflected in the WTP measures, which are much lower in the non-attendance group, while those in the non-ignoring group are visibly higher than what was observed in the base model. Interestingly, and somewhat reassuringly, accounting for this heterogeneity has no impact on the retrieved WTP for travel time, with the only impact on the time and fare coefficients being a minor increase in scale. This suggests no cross-attribute bias by not accounting for non-attendance.

While these results have reinforced the a priori expectations that attributes allegedly ignored by a given respondent may still have been given some weight in the decision making, the question still remains whether the presence of any such attributes that were unimportant, or had significantly lower importance, had an impact on the valuations of other attributes. This is the motivation of the remainder of this study.

3.2 *Analysis of data from stage 2*

As a first step, we estimate a model only on the data for stage 2, with results summarised in the first column of Table 3. All parameters are of the expected sign and statistically significant, with the same applying to all four WTP indicators. There is obviously no need for the parameters relating to ignored attributes as respondents in stage 2 were not presented with attributes for which they had stated non-attendance in stage 1. The results from this model can also be used in a comparison with those from stage 1, notably in the forms of changes in the four WTP indicators. As shown at the bottom of the table, we observe reductions in three WTP measures, where the drop in the WTP for wifi is statistically significant, and indicates a reduction by more than half compared to stage 1. There is also a drop in the WTP for travel time changes, albeit that this is not significant at the usual levels of confidence.

We next look at models estimated jointly on the data from the two stages. Table 3 shows the results for three such models, using the three different specifications from Section 3.1. These again show the gains in fit made by allowing for differences between the various groups in the data (i.e. different patterns in terms of stated non-attendance), in line with the earlier observations, and following the same patterns as in Table 2. Before proceeding with a detailed analysis of these results, it is worth remembering that the log-likelihood for the final model for stage 1 (using separate coefficients for the non-attendance part of the sample) was -5,465.60, using 10 parameters. The fit for the model estimated on stage 2 is -4,443.70, with 7 parameters. Estimating a simple joint model without accounting for possible scale differences gives a log-likelihood of -9,920.07, with 10 parameters. The likelihood-ratio test comparing the two separate models with the joint model has a value of 21.55 units; with 7 degrees of freedom, we can thus firmly reject the assumption of homogeneity between the two stages.

Table 3. Estimation results for stage 2 and joint models

Model	Stage 2		Generic		Joint Non-attendance group separate		Non-attendance group zero	
Obs.	5496		12824		12824		12824	
Resp.	916		1832		1832		1832	
LL	-4,443.70		-10,165.90		-9,918.90		-9,944.80	
Par.	7		8		11		8	
Adj. ρ^2	0.2629		0.2779		0.2952		0.2936	
	Est.	t-rat	Est.	t-rat	Est.	t-rat	Est.	t-rat
$B_{\text{L-FARE}}$	-6.8030	22.80	-6.7100	29.40	-7.0230	29.60	-6.7180	31.60
$B_{\text{L-time}}$	-4.5330	19.30	-4.6480	25.70	-4.8920	26.10	-4.6370	28.20
$B_{\text{flex, attend}}$	0.7755	7.20	0.6836	9.80	0.8350	10.40	0.7302	10.50
$B_{\text{seat, attend}}$	1.1340	15.20	0.9196	18.60	1.1620	18.20	1.2060	19.40
$B_{\text{wifi, attend}}$	0.7073	7.90	0.5039	11.20	0.8866	12.20	0.8279	12.20
$B_{\text{flex, ignore}}$	-		-		0.2274	2.40	-	
$B_{\text{seat, ignore}}$	-		-		0.2443	4.00	-	
$B_{\text{wifi ,ignore}}$	-		-		0.1817	3.90	-	
$\mu_{\text{stage 2}}$			1.0610	1.85*	0.9556	1.51*	0.9662	1.13*
WTP time (£/hr)	8.00	23.00	8.31	28.90	8.36	28.30	8.28	27.10
WTP flex (£)	3.99	8.30	3.57	11.10	4.16	11.60	3.80	11.00
WTP seat (£)	5.83	11.50	4.80	14.40	5.79	15.20	6.28	17.10
WTP wifi (£)	3.64	8.20	2.63	11.90	4.42	12.80	4.31	12.20
WTP flex ignore (£)	-		-		1.13	2.50	-	
WTP seat ignore (£)	-		-		1.22	3.90	-	
WTP wifi ignore (£)	-		-		0.91	4.10	-	
Δ WTP time (£/hr)	-0.7080	1.52						
Δ WTP flex not ignore (£)	-0.3185	0.51						
Δ WTP seat not ignore (£)	0.1610	0.24						
Δ WTP wifi not ignore (£)	-1.3790	2.36						

* Calculated against a base value of 1

In the three models shown in Table 3, we allow for scale differences between the two stages. While the model not incorporating the stated non-attendance information shows higher scale for stage 2, this is a reflection of the generic coefficients being scaled down due to the presence of some respondents with low/zero sensitivities. The remaining two models show slightly lower scale for stage 2, but the differences are not statistically significant. With this in mind, it becomes clear that the differences between the two stages are in the relative sensitivities, as reflected in the significant changes in the WTP for wifi, and the WTP for travel time, where this is significant at the 87% level.

The important question in the context of the present paper is now as follows. Are the differences between the two stages caused by the inclusion in stage 1 of unimportant attributes, or are they the result of more general changes in sensitivities as respondents progress through the survey?

To allow us to answer that question, separate models were estimated for the eight groups identified in Table 1. We estimated separate models for stage 1 and stage 2, where coefficients were estimated for all attributes presented in the survey (i.e. disregarding the stated non-attendance strategies). The results for these models are summarised in Table 4. The table first shows the result of a likelihood ratio (LR) test between separate models for the two stages and a joint model, to allow us to investigate possible differences between stages. For completeness, we also show the adjusted ρ^2 measures for the separate models. Next, the table shows the scale parameter for the second stage from a joint model, allowing us to test differences in error variance between stages. Finally, the table shows the WTP measures for the attributes included in each stage, along with the differences in these WTP measures between the two stages, allowing us to test whether there are significant changes in WTP measures.

Starting with group 1, i.e. respondents who state that they did not ignore any of the attributes, we see that the LR test cannot reject the assumption of homogeneity between the two stages. We also observe no significant differences in scale between the two stages. There are some differences between the two stages in the WTP measures, but only the WTP for wifi comes close to significance. Overall, these results suggest consistency in behaviour between the two stages, which is reassuring for a group where the same attributes were used in both stages, albeit with choice scenarios coming from a different experimental design.

The second group contains respondents who stated that they had ignored the seat reservation attribute in stage 1. This is a very small sample of respondents, meaning that the results are not very reliable. We note that the LR test once again cannot reject the homogeneity assumption, and while there is some evidence of scale reductions in stage 2, this is not overly significant, with the same applying to the increases in the WTP measures.

The third group contains respondents who stated that they had ignored the wifi attribute. We once again cannot reject the assumption of homogeneity between the two stages, and there is no evidence of scale differences either, while there is also no significant change in the WTP measures for those attributes used in both stages. These results are highly interesting given that the WTP for wifi was in fact significant (albeit low) in stage 1, despite the respondents stating that they had ignored it.

The fourth group contains respondents who stated that they had ignored the flexibility attribute in stage 1. This is once again a relatively small sample, potentially meaning unreliable results; the coefficient for the *ignored* flexibility attribute was negative, but not significant. In this group, the LR test rejects the assumption of homogeneity between the two stages, and while there are no significant scale differences, there is a clear drop in the WTP for wifi provision. This is a surprising result, although it could potentially indicate that after the exclusion of the unimportant flexibility attribute, respondents place more value on time and seat reservation, and reduced value on the provision of wifi.

For respondents who stated that they had ignored both seat reservation and wifi provision (group 5), we again cannot reject the assumption of homogeneity. We observe no significant differences in either scale or WTP measures between the two stages. This group is however another example of respondents stating that they had ignored a specific attribute (seat reservation) when the estimate is in fact significant (albeit low).

For respondents who stated that they had ignored both ticket flexibility and seat reservation (group 6), we again cannot reject the assumption of homogeneity, but observe a drop in scale that is significant at the 93% level. In this group, the WTP for flexibility was significant in stage 1 even though apparently *ignored* by respondents, while the coefficient for seat reservation was in fact negative. However, this is also a very small sample, so results are not reliable, and there is no evidence of significant changes in the two WTP measures used in both stages.

Table 4. Differences between stage 1 and stage 2 by group

Group	1		2		3		4	
Respondents	221		23		273		53	
LR p-value (diff. between games)	0.44		0.14		0.06		0.00	
Adj. ρ^2 game 1	0.22		0.26		0.35		0.39	
Adj. ρ^2 game 2	0.24		0.21		0.36		0.29	
	Est.	t-rat	Est.	t-rat	Est.	t-rat	Est.	t-rat
μ_2	1.05	0.62*	0.85	1.53*	1.03	0.50*	0.90	1.18*
WTP time (£/hr) – game 1	9.89	12.90	14.33	4.60	8.79	14.20	5.79	3.50
WTP flex (£) – game 1	4.94	5.80	2.52	0.90	4.31	6.10	-4.77	1.40
WTP seat (£) – game 1	4.47	5.10	0.87	0.50	6.95	7.90	13.29	3.50
WTP wifi (£) – game 1	4.97	9.20	5.73	3.00	1.13	2.50	8.19	4.30
WTP time (£/hr) – game 2	10.04	12.70	14.40	2.80	8.19	15.40	5.60	4.30
WTP flex (£) – game 2	3.88	4.20	3.97	1.00	4.75	5.90	-	
WTP seat (£) – game 2	5.24	5.50	-		5.78	6.30	8.84	3.10
WTP wifi (£) – game 2	3.72	6.40	11.09	2.30	-		2.78	2.10
Δ VTTS	0.15	0.14	0.07	0.01	-0.60	0.73	-0.19	0.09
Δ WTP flex	-1.05	0.84	1.45	0.30	0.44	0.42	-	
Δ WTP seat	0.77	0.59	-		-1.18	0.93	-4.45	0.94
Δ WTP wifi	-1.25	1.58	5.36	1.03	-		-5.42	2.34

Group	5		6		7		8	
Respondents	221		18		97		145	
LR p-value (diff. between games)	0.44		0.14		0.06		0.00	
Adj. ρ^2 game 1	0.31		0.29		0.44		0.44	
Adj. ρ^2 game 2	0.21		0.06		0.52		0.07	
	Est.	t-rat	Est.	t-rat	Est.	t-rat	Est.	t-rat
μ_2	0.93	0.79*	0.77	1.80*	0.96	0.54*	0.70	3.91*
WTP time (£/hr) – game 1	6.70	9.10	10.71	4.40	10.39	9.10	7.86	15.70
WTP flex (£) – game 1	3.00	2.90	4.00	2.50	1.45	1.10	1.81	2.70
WTP seat (£) – game 1	1.20	2.00	-2.27	1.90	9.95	5.40	0.68	1.60
WTP wifi (£) – game 1	0.02	0.00	6.81	3.30	0.29	0.30	1.07	3.00
WTP time (£/hr) – game 2	5.79	7.00	8.63	4.30	9.44	8.70	7.48	10.80
WTP flex (£) – game 2	3.62	4.00	-		-		-	
WTP seat (£) – game 2	-		-		6.00	3.50	-	
WTP wifi (£) – game 2	-		5.86	3.70	-		-	
Δ VTTS	-0.91	0.82	-2.09	0.66	-0.95	0.60	-0.38	0.45
Δ WTP flex	0.62	0.45	-		-		-	
Δ WTP seat	-		-		-3.94	-1.57	-	
Δ WTP wifi	-		-0.95	0.36	-		-	

* Calculated against a base value of 1

For respondents who stated that they had ignored both ticket flexibility and provision of wifi (group 7), we again cannot reject the assumption of homogeneity, and the scale differences between the two stages are not significant. We note a drop in the WTP for reserved seats, but this is only significant at the 88% level.

Finally, for respondents who stated that they had ignored all three qualitative attributes (group 8), the LR test rejects the assumption for homogeneity between the two stages. We observe a significant drop in scale between the two stages, but the change in the WTP for travel time is not significant. The question arises whether the drop in scale is the result of fatigue, but no such effects were observed in the other groups, notwithstanding the possibility that this group contains individuals who paid less attention to the survey. Another explanation is that the new stage may be too simplistic by focussing on only two attributes, potentially leading to a drop in respondent engagement with the survey, and resulting low data quality. Finally, by looking at the results for stage 1, it also seems that the *ignored* attributes were once again not really ignored but just given lower valuations.

Before proceeding to the conclusions, it is worth noting the high level of consistency across the eight groups. We note overall a lower level of sensitivity in the stated non-attendance group, where this however remains different from zero. Additionally, and crucially for the present paper, we observe little evidence of significant differences between stages, in terms of error variance or relative sensitivities.

4. Conclusions

The aim of the present paper was to contribute to current knowledge in the field of attribute processing in stated choice surveys, and in particular attribute non-attendance. It is now well established that some respondents will make their decisions in such surveys based on a subset of attributes. Separate strands of research have looked at appropriate ways of identifying these respondents, understanding the causes for non-attendance, studying their impacts on overall results, and making appropriate provisions for the presence of such respondents at the modelling end.

The present paper has looked at a distinct issue, namely what impact any ignored attributes may have on the remaining parameter estimates. This is in contrast with existing work, which has principally been concerned with the impact that ignoring a specific attribute may have on the sample level estimates for the associated coefficient. The results from this paper should provide some support to analysts facing the difficult trade-off between relevance and respondent burden; should all attributes that may matter be included, or should surveys focus only on core attributes likely to be of importance to all respondents?

The specific approach used in this paper was to first collect responses from a stated choice component involving a full set of five attributes. On the basis of respondent reported information on attribute non-attendance, each respondent was then presented with a second experiment, excluding any attributes that this given respondent stated to have ignored in the first stage. Such an approach would clearly not be applicable in practical research, and there are also arguably concerns about endogeneity bias by making use of respondent stated information on attribute non-attendance, in our models as well as in the selection of attributes for designing the second stage. The aim in the present paper was simply to give a first indication on the likely impact of ignored attributes on the remaining valuations. Another possible shortcoming relates to the possibility that the first stage scenarios allow respondents to *learn* their preferences and that these then carry over into the second stage. In this context, it is important to note that, during the first stage, respondents did not know of the existence of the simpler second stage. Also, the sensitivities to the *unimportant* attributes were not zero in the first stage, so it is not the case that respondents focussed solely on learning their sensitivities to those attributes which would then also be carried over into the second stage.

As a first observation, the analysis has once again shown that respondent reported information on non-attendance may not be completely reliable, i.e. there is evidence that in some cases, respondents who state that they had ignored a given attribute simply assigned it lower importance. This is reflected in statistically significant estimates for the concerned coefficients. However, the estimates in the non-attendance groups are invariably lower than those for the remaining respondents, confirming that they did indeed treat the concerned attributes in a different manner. It should be acknowledged that different respondents potentially have different reasons for indicating non-attendance (cf. Alemu et al., 2013), and that this could explain why the mean values in the group are still different from zero. This is also in line with the discussions in Hess et al. (2013) and a random treatment of sensitivities within the stated non-attendance group can provide further insights. This is however beyond the scope of the present analysis, and also difficult given the small sample sizes in some of the subgroups.

Turning to the issue of main interest in the present paper, we can observe that assigning a second stage with a reduced set of attributes produces slightly different results in some cases, but no overall trends. Crucially, in the present study, there is no conclusive evidence that presenting *unimportant* attributes unduly affects behaviour, in terms of scale or relative valuations.

Like many other studies in the field, the results from this paper are based on just a single dataset, and further corroboration would be useful. Nevertheless, they thus give some support to the notion that respondents are able to focus on those attributes that do matter to them, and that including attributes that may be irrelevant to some respondents does not have any detrimental impact on their overall behaviour. The risk of overburdening respondents thus seems relatively small in the present context. It should also be noted that our data showed clear heterogeneity across respondents in terms of which attributes are considered important, meaning that it would indeed be difficult for an analyst to specify a subset of attributes that would be relevant to all respondents while ensuring that all relevant attributes are included for each respondent.

It should be acknowledged that the present study made use of a relatively simple survey with a maximum of five attributes, all of which are relatively familiar to most people, and the situation may well be different in surveys with larger numbers and/or unfamiliar attributes, where respondent burden may become more of an issue. Indeed, it may then not be possible to present all respondents with every single attribute. What is too many attributes is a question that is probably survey and context dependent. Nevertheless, the question then arises how the choice scenarios can be customised to each respondent, potentially based on prior information. Once again, there are likely to be important important issues with endogeneity, and this thus remains another area for further research.

While this analysis has allowed us to test the impact of including unimportant attributes, it is more difficult (though similarly important) to look at the impact that not including relevant attributes may have on results. Analysts routinely produce monetary valuations on the basis of surveys with only two or three attributes, thus assuming that the valuations in such trade-offs are consistent with those from a real world setting where numerous other attributes play a role. Here, a risk for example arises that respondents may infer the values of such missing attributes, taking heed for example of the warnings in Islam et al. (2007). If evidence could be produced that including only a subset of relevant attributes has an impact on the estimates produced by these models, this would raise concerns as to the continued widespread use of such surveys. As one example, the question could be asked whether the focus in value of time research on simple time-money trade-offs is potentially misguided and if different valuations would be obtained when including other relevant attributes.

With hindsight, it would also have been interesting to include an additional third stage in which respondents are presented with random subsets of attributes to investigate what happens if important attributes are not included. This is another area for future work, as is a setting in which unimportant attributes are added in a later stage, following a stage which already has a higher level of burden than what was used here. Finally, explicit testing of hypotheses relating to the processing of *unimportant* attributes and their impact on other components remains of interest.

Acknowledgements

The author acknowledges the financial support by the Leverhulme Trust, in the form of a *Leverhulme Early Career Fellowship*. The author would also like to thank Andrew Daly, Amanda Stathopoulos and Kelvin Balcombe for comments on an earlier version of this paper.

References

Alemu, M. H., Mørbak, M. R., Olsen, S. B., Jensen, C. L., 2013. Attending to the reasons for attribute non-attendance in choice experiments. *Environmental and Resource Economics*, 54(3), 333-359.

Balcombe, K., Burton, M., Rigby, D., 2011. Skew and attribute non-attendance within the bayesian mixed logit model. *Journal of Environmental Economics and Management*, 62 (3), 446–461.

Carlsson, F., Kataria, M., Lampi, E., 2010. Dealing with Ignored Attributes in Choice Experiments on Valuation of Sweden's Environmental Quality Objectives. *Environmental and Resource Economics*, 47(1), 65-89.

Choicemetrics, 2010. NGene 1.1. http://www.choice-metrics.com/.

Collins, A. C., Rose, J. M., Hess, S., 2012. Interactive stated choice surveys: a study of air travel behaviour. *Transportation,* 39 (1), 55–79.

Collins, A. T., Rose, J. M., Hensher, D. A., 2013. Specification issues in a generalised random parameters attribute nonattendance model. *Transportation Research Part B*, 56, 234-253.

Daly, A., Hess, S., 2011. Simple methods for panel data analysis. paper presented at the 90th Annual Meeting of the *Transportation Research Board*, Washington, D.C.

Dumont, J., Adler, T. J., Hess, S., Neafsey, W. C., 2011. Using individual-level parameter estimates to investigate evidence of lexicographic behavior and heterogeneity in information processing strategies. paper presented at the *Second International Choice Modelling Conference*, Oulton Hall, Leeds.

Hensher, D. A., 2006. How do respondents handle stated choice experiments? – attribute processing strategies under varying information load. *Journal of Applied Econometrics,* 21, 861–878.

Hensher, D. A., 2008. Joint estimation of process and outcome in choice experiments and implications for willingness to pay. *Journal of Transport Economics and Policy*, 42 (2), 297-322.

Hensher, D. A., 2010. Attribute processing, heuristics and preference construction in choice analysis. In: Hess, S., Daly, A. J. (Eds.), *State-of Art and State-of Practice in Choice Modelling: Proceedings from the Inaugural International Choice Modelling Conference*. Emerald, Bingley, UK, Ch. 3, 35–70.

Hensher, D. A., Rose, J. M., Bertoia, T., 2007. The implications on willingness to pay of a stochastic treatment of attribute processing in stated choice studies. *Transportation Research Part E*, 43 (2), 73–89.

Hensher, D. A., Rose, J. M., Greene, W. H., 2005. The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, 32 (3), 203–222.

Hess, S., Hensher, D. A., 2010. Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B*, 44 (6), 781–790.

Hess, S., Rose, J. M., 2007. A latent class approach to recognising respondents' information processing strategies in SP studies. paper presented at the *Oslo Workshop on Valuation Methods in Transport Planning*, Oslo.

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., Caussade, S., 2013. It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity. *Transportation*, 40 (3), 583–607.

Hole, A. R., 2011. Attribute non-attendance in patients' choice of general practitioner appointment. paper presented at the *Second International Choice Modelling Conference*, Oulton Hall, Leeds.

Islam, T., Louviere, J., Burke, P., 2007. Modeling the effects of including/excluding attributes in choice experiments on systematic and random components. *International Journal of Research in Marketing*, 24, 289–300.

Layton, D., Hensher, D. A., 2010. Aggregation of common-metric attributes in preference revelation in choice experiments and implications for willingness to pay. *Transportation Research Part D*, 15 (7), 394–404.

Mariel, P., Hoyos, D., Meyerhoff, J., 2013. Stated or inferred attribute non-attendance? A simulation approach. *Economa Agraria y Recursos Naturales*, 13, 51–67.

Morkbak, M., Olsen, S., Denver, S., Campbell, D., 2013. Providing real economic incentives in choice experiments what are the behavioral implications? paper presented at the *Third International Choice Modelling Conference*, Sydney.

Scarpa, R., Raffaelli, R., Notaro, S., Louviere, J. J., 2011. Modelling the effects of stated attribute non-attendance on its inference: an application to visitors benefits from the alpine grazing commons. paper presented at the *Second International Choice Modelling Conference*, Oulton Hall, Leeds.

Scarpa, R., Zanoli, R., Bruschi, V., Naspetti, S., 2012. Inferred and stated attribute nonattendance in food choice experiments. *American Journal of Agricultural Economics*, 95 (1), 165–180.

Train, K., Sonnier, G., 2005. Mixed logit with bounded distributions of correlated partworths. In: Scarpa, R., Alberini, A. (Eds.), *Applications of Simulation Methods in Environmental and Resource Economics*. Springer Publisher, Dordrecht, The Netherlands, Ch. 7, 117–134.