ISSN 1835-9728

Environmental Economics Research Hub

Research Reports

Inducing Strategic Bias: and its implications for Choice Modelling design

Michael Burton

Research Report No. 61

May 2010

About the authors

Michael Burton is a Professor at the School of Agricultural and Resource Economics, University of Western Australia

E-mail: michael.burton@uwa.edu.au

* This research is part of the Environmental Economics Research Hub Project 8: *Divergence between community and expert valuations of ecosystems*, which is funded by the Australian Commonwealth Environmental Research Facilities (CERF) programme. Valuable comments from Jonelle Cleland, Ben White and Dan Rigby are gratefully acknowledged

Environmental Economics Research Hub Research Reports are published by The Crawford School of Economics and Government, Australian National University, Canberra 0200 Australia.

These Reports present work in progress being undertaken by project teams within the Environmental Economics Research Hub (EERH). The EERH is funded by the Department of Environment and Water Heritage and the Arts under the Commonwealth Environment Research Facility.

Crawford School of Economics and Government
THE AUSTRALIAN NATIONAL UNIVERSITY

http://www.crawford.anu.edu.au

Table of Contents

Abstract		4
1.	Introduction	5
2.	Conceptual Focus	5
3.	Study Design	7
4.	Treatments	9
5.	Descriptive Analysis	13
6.	Summary of Descriptive Analysis	19
7.	Estimation	20
8.	Identification of Bias	25
9.	Conclusions	27
References		29
Appendix 1		30

Abstract:

It has been suggested that the nature of the task within a multi-attribute multi-alternative choice experiment may be sufficiently complex to make it difficult for individuals to develop response strategies to strategically bias their answers. This experiment tested that hypothesis by setting experimental conditions that provide incentives for strategic bias. By changing design parameters one can investigate whether the strategic bias can be reduced. The answer is effectively no: under most circumstances, respondents could find a strategy that achieved significant bias in inferred preferences. The circumstances where this did not occur (involving ranking alternatives, rather than selecting a single preferred alternative) the inferred preferences reflected neither the intended bias, nor their original preferences, making the answers useless to both respondent and researcher.

Keywords: Strategic bias, choice modeling, complexity.

JEL classification: Q51, C91

.

1. Introduction

Valuation of non-market environmental goods has relied upon hypothetical, stated preference methods. One issue with the contingent valuation method is its exposure to strategic bias, whereby someone with a particular interest can inflate the perceived value they place upon an outcome, beyond the level they would give if they actually had to pay for that outcome.

There is an unsupported assertion made in some of the environmental valuation literature that the “choice modelling” (CM) approach to experimental design can reduce the ability of respondents to systematically select choices that are strategically biased, as compared to the more traditional dichotomous choice techniques used in contingent valuation. Surprisingly, in the context of designing choice experiments, technical guidelines for minimising strategic bias are lacking, and there appears to be little empirical research exploring this issue. This is a particular issue for the CERF project, as it will be working with environmental scientists from a range of disciplines, and there is a concern that their responses may be biased by their disciplinary perspective.

The objective of the proposed experiment is to investigate the implications of choice experiment design on the possibility of a respondent achieving a strategic bias in the outcome. It uses students as the respondents, and gives them a strong financial incentive to simulate strategic bias.

2. Conceptual focus

The issue of strategic bias is linked to the concept of incentive compatibility, which is receiving increasing attention within the choice modelling literature. For incentive compatibility, a survey question needs to fulfil two criteria (Carson and Groves, 2007). Firstly it has to be consequential in so far as the respondents answer has to have an influence on some decision, and that the respondent cares about the outcome of that decision. Without consequentiality it is claimed that it not possible for economic theory to predict respondents behaviour (Carson and Groves, 2007), although Mazar et al (2007) suggest that truth telling may be the outcome as a result of individuals preference to see themselves as honest. Assuming that an experiment is consequential does not necessarily imply that answers will be truthful if the design of the question is such that there are incentives to misrepresent one’s preferences, if that increases one’s personal benefits. An experiment is incentive compatible if it is designed so that there are no incentives to give anything but answers that reflect the true preferences of the respondent. The requirements for incentive compatibility within the context of a typical choice modelling experiment are quite severe: apart from the issue of consequentiality, single shot binary-choice questions are the only format that can ensure incentive compatibility within the context of public good provision.

This is not the format that is currently most typically used (multiple questions, with multiple alternatives within each) and not the context within which the arguments about the benefits of choice experiments are made. In fact, arguments about the benefits of choice experiments (CE) have been based only informally around the issue of incentive compatibility, to the extent that the format may make it harder for respondents to work out how to behave strategically. We now turn to consider these arguments.

Although the suggestion is made, the literature is often vague on the exact reasons as to why the choice experiment[1] should be less prone to strategic bias. Thus: "choice modelling generally avoids an explicit elicitation of respondents' willingness to pay by relying instead on ratings, rankings or choices amongst a series of alternative packages of characteristics from where willingness to pay can be indirectly inferred. As such, CE may minimise some of the response difficulties found in CVM… (protest bids, strategic behaviour, yeah saying). But this point has yet to be demonstrated." Hanley et all (2001) p448. Similarly Adamowicz et al , (1999) state "Strategic Behaviour should be minimal in Stated Preference tasks since the choices are made from descriptions of attributes and it is not clear which choice will over- or under represent a valuation" (p467). Lu et al (2008) suggest there may be two processes at work within the CM environment: "By adding complexity to the SP task, respondents may exhibit less bias. This may be partly occur because of the extra effort required to complete the exercise with bias, but it is more likely to occur because of respondents failing to see any clear single purpose to the exercise" (p128).

Two possible issues emerge from these perspectives. The first is that the benefit of the CE approach is through masking intent. The implication is that respondents may hold values for all attributes, but would be prepared to overstate the value held for any of them, on the assumption that it was the policy target. Faced with a number of possible attributes, and unable to identify which of these is the true target of interest to the policy maker, they resort to revealing the truth. However, this argument does not preclude the possibility of respondents overstating the value of all attributes, presumably by downplaying the cost attribute.

The second possibility is that the respondent wants to overstate the value placed on an attribute relative to others i.e. they genuinely hold a preference for one attribute, and wish to inflate this value compared to others in the design. The increased complexity of the CE structure is hypothesised to make it difficult to identify which choices will lead to this result. However, if this argument is true, then it undermines the efficacy of using the CE structure at all. Indeed, one could argue that strategic bias is simply an expression of a well-behaved but constructed utility function, and it is unclear as to why the CE structure should make it difficult to reveal this constructed utility function: and if it does, why it allows the revelation of a 'normal' utility function. As such, the assertions that CE will lead to reductions in strategic bias seem to be speculative at best. The study implemented is focused on the second of these hypotheses: that task complexity leads to a reduced ability to strategically influence outcomes.

[1] This working paper assumes that the basic structure and analytics of choice experiments is known: see Hensher et al (2005) for an overview.

3. Study Design

The objective of this study is to investigate whether complexity in the repeated choice experiment task is sufficient to prevent strategic behaviour. It does not address the issues of incentive compatibility of the design itself, or issues of provision mechanisms: indeed it controls for incentive compatibility by attempting to deliberately induce incentives for strategic behaviour.

The vehicle for the valuation study was the desirable characteristics of rented accommodation. This was chosen as it is a familiar concept to students, and one could easily identify a number of relevant characteristics. The fact that this is a private good rather than a public good (such as an environmental outcome) is not seen as an issue, as the objective of the study is to identify the ability of respondents to influence the outcome of the design. Five attributes were selected (signs in parentheses indicate the anticipated impact of the attribute on utility) : Rent per week (-ve), Distance from UWA (-ve), Total number of people sharing (?), Furnished/unfurnished (+ve) and North/South of the river (+ve). Furnished state is 1 if furnished, 0 otherwise, and North/South of the Swan River: 1 if North, 0 if South. The latter attribute needs some explanation with respect to the geography of Perth: the Swan River splits the city in two, with limited vehicle access via 2 bridges from one side to the other within the city. The traffic bottle neck this imposes, and the need for changing on public transport if one travels from the South means that access to UWA is likely to be seen as less easy if one lives a similar geographical distance from UWA, but South of the river. However, it may also act as a proxy for other characteristics of accommodation, such as access to beaches or other infrastructure.

Respondents were initially asked to complete a set of 8 choice sets, having been given the information that the study was interested in the issue of preferences for student housing, given its shortage in Perth. Having completed those initial choice sets, the true purpose of the study, the investigation of bias, was revealed. Respondents were told that the type of survey they had just completed may be subject to strategic behaviour and that in the remaining parts of the study they were going to be asked to deliberately behave strategically, and attempt to bias the outcomes of the research. The incentive mechanism for doing so was the reward system of the survey itself: all participants were entered into a lottery for a significant ($400) reward but those who managed to bias their responses the most, in the way required, would get a greater number of entries into the lottery, increasing their chances of winning. The full text of the instructions given is in Box 1 below.

Box 1: Instructions to respondents, explaining the incentive mechanism to induce strategic behaviour

The type of questions that you have just completed are very commonly used in valuing new products or environmental assets.
From the choices made, and the levels of the attributes that are included in the alternatives, it is possible to identify how, on average, the respondents to the survey are trading off the different attributes of the accommodation.

One issue is the extent to which they are open to manipulation: that people will not give their true answers to the question because they want to try and influence the outcomes in a particular way. It is unlikely that you were doing this, but it may be the case where people try and overstate the importance of some feature, in an effort to change public policy, or change the type of product provided.

In the next section of this survey, you will be asked to deliberately change the way that you answer the questions, to mimic this type of biased response".

"Please read the following information carefully. Understanding it will have a strong impact on your chances of winning the $400.

In the following section, you will be presented with an additional set of questions. When answering these questions you should behave as if you want to give the impression that

Being CLOSE to UWA as the most important attribute to you,
You want this to be identified as more important than the other attributes of the accommodation.

A group of 40 of you have been given the task of trying to influence the value attached to being CLOSE to UWA

Other groups of respondents have been given the task of biasing the importance placed on other attributes.

The extent to which you, as a group, are more successful in manipulating the outcomes of the study will change the likelihood that you will win the $400.

Everyone who completes the survey will be awarded 1 entry into the draw. Each person in the group of 40 who manage to bias their designated attribute the most will be awarded 3 entries each into the draw, increasing their chance of winning the $400.

The winner of the $400 will be drawn at random, and the draw will be monitored, to ensure that the process outlined above is followed

The first paragraph explains the true issue being investigated, while the second establishes the conditions for incentive compatibility for achieving bias. The survey was designed to prevent a return to the earlier choice sets to change answers. The second set of choice questions was an exact replication of the initial set. Evaluation of the extent of bias was made relative to the initial set in each case. One criticism of this approach is that there may have been learning or fatigue effects, and an alternative may have been to compare the bias outcomes with a control group who completed 16 choice sets, without incentives to bias. This was not undertaken because it was anticipated that the inducement to bias would overwhelm any other learning or fatigue effects that may be present. No further information was given on the protocols that would be used to define the “most successful” bias outcome. Analysing the responses to the choice questions then allows one to identify the ability of the respondent to influence responses, and the impact of question design on that ability. The statistical design of the survey and the exact instructions to respondents was varied, allowing some judgement to be made on the extent to which respondents were able to bias outcomes relative to treatments. To understand that process, the different treatments need to be outlined.

4. Treatments

As noted in the introduction, one hypothesis is that the complexity of the repeated choice experiment may make it difficult to identify appropriate strategic behaviour. Therefore in the design a number of different aspects of complexity were included. Work on the impacts of complexity on behaviour in choice experiments has identified a number of potential aspects that may impact on behaviour (e.g. Hensher, 2006). Two of those are selected here: a number of alternatives within a choice set and the number of levels within attributes. Whether respondents were asked to choose the best alternative or to rank all was also used. It was unclear *ex ante* how this may affect behaviour: ranking is clearly a more complex process, and hence makes it harder to behave strategically, but it may appear to offer more subtle options for influencing outcomes. It is important to remember that no information was given on how the choices made would be analysed, but it is perhaps reasonable to assume that respondents would infer that the full ranking information would be used, given it was being requested (but see the statistical analysis below). The final issue is on the number of attributes that the respondents were asked to reveal bias towards. A question that has not received explicit attention within the incentive compatibility literature is whether there are specific attributes within an alternative that respondents are concerned about and trying to influence the valuation of. Usually the discussion is phrased in terms of the alternatives as a whole (i.e. of three candidates, who do you vote for: of four possible environmental interventions, which is preferred). However, within most choice experiments, there can be a diverse set of attributes, ranging over a variety of outcome types (for example, designs may include attributes that reflect pure existence values for the resource, recreational use values and social values such as the impact of any policy change on employment levels or rural populations). Strategic behaviour may be focussed on particular attributes. Of course, its attribute levels that lead to aggregate utility from each alternative, and that will translate into choices over alternatives (strategic or otherwise). As such, the results from the incentive compatibility literature remain: faced with more than a binary choice, there may be incentives to select an alternative other than the truly most preferred. However, the nature of the rules that deliver optimal strategic outcomes are likely to be more complex if one is attempting to influence the provision of more than one attribute. To investigate this, some samples are asked to bias one attribute, while others are asked to bias 2.

In summary, the meta-design consists of 4 variables, each with 2 modes:

Number of alternatives per choice set (i.e. 3 or 6)

The number of levels for cost and distance attribute (i.e. 4 or 6)

Whether respondents had to select the most preferred alternative or rank all alternatives

Whether respondents were attempting to bias the effect of one attribute, or two.

This gives a total of 16 potential treatments (2^4) in the meta-design. However, due to limited student numbers (see below on recruitment methods) only, only 11 treatments were implemented . Given the objective of the study, the focus was on the more complex designs. Table 1 below shows the full set of 16 possible designs, with the 11 implemented designs shaded in grey.

Table 1 Experimental design showing 16 potential treatments and the 11 implemented treatments shaded in grey.

			Attributes being influenced			
			1		2	
			Attribute levels			
			4	6	4	6
Number of alternatives	3	Rank All	D1		D6	D9
		Select Best	D2			
	6	Rank All	D3	D4	D7	D10
		Select Best		D5	D8	D11

The attributes and their levels are reported in Table 1 below.

Table 2 The attribute levels

Attribute	Levels
Rent per week ($)	4 levels: 75,125,150,200 6 levels: 75,100,125,150,175,200
Distance from UWA (km)	4 levels: 5,10,15,20 6 levels: 8,10,12,14,16,20
Total number of people sharing	1,2,3,4
Furnished/unfurnished	1/0
North/South of river	1/0

With regards to the issue of whether they have to influence one or two attributes, in the case of the latter, the instructions on the target for bias were adjusted accordingly:

Box 2: Alternative wording for respondents required to influence 2 attributes.

> *"In the following section, you will be presented with an additional set of questions. When answering these questions you should behave as if you want to give the impression that*
>
> *Being CLOSE to UWA and being NORTH of the river are the most important attributes to you,*
> *You want these to be identified as more important than the other attributes of the accommodation.*
>
> *A group of 40 of you have been given the task of trying to influence the value attached to being CLOSE to UWA and being NORTH of the river."*

Although the meta-design contains 16 potential treatments, there are only 4 statistical designs required (as ranking versus selecting best and influencing 1 or 2 attributes operates independently of the formal CE design). All designs were generated using *Ngene* (Rose et al, 2009), assuming a linear utility function in the 5 attributes, using 8 choice sets per design. The relative simplicity of the attribute specification means that designs could be completed with 8 choice sets, and hence there was no requirement to block the design.

An adaptive design process was employed. A pre-test to check the functionality of the on-line survey allowed an initial sample to be collected based on a 3 alternative - 4 attribute level design, designed to minimise D error. This convenience sample was used to estimate priors for the parameters of the utility function. These priors were then used to re-generate a revised design, but this time using an S efficiency criterion (i.e. minimising sample size required to estimate the models: see Scarpa and Rose, 2008). Given the limited number of respondents and the large number of potential treatments, achieving efficiency on this criterion was seen as the most valuable.

This revised design was then used as the basis for collecting a full set of data for treatments D1 and D2 (see Table 2 above), and this data used to estimate a further revised set of priors. These priors were then used to generate designs for the three more complex cases (6 alternatives-4 attribute levels; 3 alternatives-6 attribute levels; 6 alternatives-6 attribute levels). In all cases S efficiency was used as the design criteria. The minimum number of respondents needed was estimated at 9 for all designs, given the revised priors and assuming that each student would complete the full set of 8 choice questions.

A general call was made to students at the University of Western Australia (UWA), Perth, to participate in a survey. This required an initial 'sign-up'. Students who expressed an interest in completing the study were randomly allocated to a group of approximately 40 students. Each group completed a unique survey, termed a "treatment". It was initially intended that all 6 treatments would be covered but there were only sufficient students to complete 11. Each treatment differed only in the structure of the choice experiments presented to the respondent.

A total of 40 invitations were sent out to students for each version of the survey. Although they had volunteered to complete the survey, not all invitations were taken up. Table 3 gives sample numbers for those who completed the entire survey by relevant treatments (and hence were eligible for reward).

Table 3: Number of completed surveys by treatment

			Attributes being influenced			
			1		2	
			Attribute levels			
			4	6	4	6
Number of alternatives	3	Rank All	D1 27		D6 26	D9 21
		Select 1	D2 29			
	6	Rank All	D3 24	D4 27	D7 21	D10 29
		Select 1		D5 30	D8 24	D11 24

5. Descriptive analysis

Before considering a more formal statistical analysis, it is informative to look at the actual choices made, by question. Although there are 11 different treatments considered within the overall experiment, in terms of the attribute structures within the choice sets there are only four unique designs (using 3 or 6 alternatives, and 4 or 6 attribute levels): the other elements of the experiment (selecting 1 or ranking, and biasing 1 or 2 attributes) are independent of the choice set design. Each of the four will be considered in turn. Note that for those who had to rank (as opposed to select a single best alternative) have been re-scored so that their first ranked alternative only is considered. In the four tables that follow, the percentage of respondents who select an alternative are reported, for both the pre-bias and post-bias condition. Note that these are the same people within each treatment. The attribute value for Distance and North/South are reported for each question and alternative. For those who are attempting to bias 1 attribute only, then that is distance, for those influencing 2, it is both. The alternative which would appear to be the dominating alternative, given the bias condition is indicated with two stars (**). This will be the case where there is only one alternative with the minimum distance, or only one alternative where there is both a minimum distance and North=1. In the case where there are several alternatives with the same minimal distance, or where it is not possible to find a single alternative that achieves both minimal distance and North, all possible alternatives are marked with one star (*). The star system gives a simplistic assessment of the options one might expect the bias-condition to select. However, note that it does not consider any of the other attribute levels (i.e. cost) and in the case of rankings, it does not consider what was being selected as 2nd or 3rd options.

Table 4 **Percentage of sample selecting an alternative, pre-bias and post-bias: Design with 3 alternatives and 4 levels (Treatments D1,D2,D6), by question.**

		Alternative 1		Alternative 2		Alternative 3	
		% sample Pre-bias	Post-bias	% sample Pre-bias	Post- bias	% sample Pre-bias	Post-bias
Q1		Dist=15,North=0		Dist=20,North=0		Dist=10,North=0	
	D2 select: bias 1	4	**4**	38	**0**	58	**96 ****
	D1 rank: bias 1	28	**8**	28	**88**	44	**4 ****
	D6 rank 2	28	**16**	36	**76**	36	**8 ****
Q2		Dist=10,North=0		Dist=10,North=1		Dist=20,North=0	
	D2 select: bias 1	38	**48 ***	35	**36 ***	27	**0**
	D1 rank: bias 1	36	**56 ***	28	**27 ***	36	**0**
	D6 rank: bias 2	50	**12**	23	**88 ****	27	**0**
Q3		Dist=20,North=0		Dist=8, North=0		Dist=8, North=1	
	D2 select: bias 1	42	**0**	35	**48 ***	23	**52 ***
	D1 rank: bias 1	48	**4**	16	**0 ***	36	**96 ***
	D6 rank: bias 2	58	**4**	19	**8**	23	**88 ****
Q4		Dist=20,North=0		Dist=8, North=1		Dist=8, North=1	
	D2 select: bias 1	8	**8**	42	**56 ***	50	**36 ***
	D1 rank: bias 1	8	**4**	4	**0 ***	88	**96 ***
	D6 rank: bias 2	8	**4**	0	**0 ***	92	**96 ***
Q5		Dist=8, North=1		Dist=10, North=1		Dist=15, North=1	
	D2 select: bias 1	19	**88 ****	8	**8**	73	**4**
	D1 rank: bias 1	28	**84 ****	24	**12**	48	**4**
	D6 rank: bias 2	12	**84 ****	46	**16**	42	**0**
Q6		Dist=8, North=1		Dist=15, North=1		Dist=20, North=1	
	D2 select: bias 1	35	**100 ****	62	**0**	4	**0**
	D1 rank: bias 1	36	**88 ****	40	**12**	24	**0**
	D6 rank: bias 2	19	**80 ****	54	**20**	27	**0**
Q7		Dist=15, North=1		Dist=20, North=1		Dist=10, North=0	
	D2 select: bias 1	81	**13**	4	**0**	15	**87 ****
	D1 rank: bias 1	88	**20**	4	**76**	8	**4 ****
	D6 rank: bias 2	65	**92 ***	23	**8**	11	**0 ***
Q8		Dist=10, North=1		Dist=15, North=1		Dist=15, North=0	
	D2 select: bias 1	31	**88 ****	62	**8**	8	**4**
	D1 rank: bias 1	28	**84 ****	16	**8**	56	**8**
	D6 rank: bias 2	23	**84 ****	23	**8**	54	**8**

Treatment codes, whether they select or rank, and whether they bias 1 or 2 attributes in Column 2.
Attribute levels for the potential target attributes - Distance and North/South (1,0) are reported.
In ranking designs, Alternatives ranked 1 recoded to selected.
% of sample selecting each Alternative reported. Values in **bold** are the post-bias treatment
** indicates a single alternative is identified as dominant for that condition, * indicates one of several alternatives that may be considered dominant.

In general, these results are consistent with what one would expect: respondents select the Alternative with the smallest distance, or make some tradeoff when required to bias both distance and N/S. The anomalies are: in Q1, the samples that rank both select Alternative 2, even though it has the highest distance. Similarly, in Q7, those who rank, and bias distance alone, have a very low probability of selecting Alternative 3, which has the lowest distance. However, in Q8 this effect is gone, and it was not present in Q2. In Q7, those who in d6 (bias 2, full ranking) appear to be trading off the increased distance in Alternative 1 for the presence of the North attribute, which is plausible. In summary, the anomalies seem to occur (selectively) within those who are required to rank

Table 5. Percentage of sample selecting an alternative, pre-bias and post-bias: Design with 3 alternatives and 6 levels, by question.

		Alternative 1		Alternative 2		Alternative 3	
		% sample		% sample		% sample	
		Pre-bias	Post-bias	Pre-bias	Post-bias	Pre-bias	Post-bias
Q1		Dist=10,North=0		Dist=20,North=0		Dist=8,North=1	
	D9 select: bias 2	9	**5**	19	**94**	71	**5 ****
Q2		Dist=12,North=0		Dist=10,North=1		Dist=10,North=0	
	D9 select: bias 2	9	**5**	14	**86 ****	76	**10**
Q3		Dist=16,North=1		Dist=8,North=0		Dist=12,North=1	
	D9 select: bias 2	43	**5**	52	**43 ***	5	**52 ***
Q4		Dist=8,North=0		Dist=12,North=0		Dist=16,North=1	
	D9 select: bias 2	71	**52 ***	19	**48**	10	**0 ***
Q5		Dist=20,North=1		Dist=16,North=1		Dist=8,North=0	
	D9 select: bias 2	10	**5**	10	**10 ***	81	**86 ***
Q6		Dist=10,North=1		Dist=14,North=0		Dist=14,North=0	
	D9 select: bias 2	76	**100 ****	24	**0**	0	**0**
Q7		Dist=8,North=0		Dist=8,North=1		Dist=20,North=0	
	D9 select: bias 2	57	**5**	38	**95 ****	5	**0**
Q8		Dist=14,North=0		Dist=10,North=1		Dist=10,North=1	
	D9 select: bias 2	5	**5**	62	**0 ***	33	**95 ***

Treatment codes, whether they select or rank, and whether they bias 1 or 2 attributes in Column 2.
Attribute levels for the potential target attributes - Distance and North/South (1,0) are reported.
In ranking designs, Alternatives ranked 1 recoded to selected.
% of sample selecting each Alternative reported. Values in **bold** are the post-bias treatment
** indicates a single alternative is identified as dominant for that condition, * indicates one of several alternatives that may be considered dominant.

In general the results are consistent with expectations, given that this sample is attempting to bias 2 attributes: the only anomaly appears to be in question 1, where Alternative 3 is the dominant Alternative, and yet does not get selected, compared to strategy.

Table 6. Percentage of sample selecting an alternative, pre-bias and post-bias: Design with 6 alternatives and 4 levels, by question.

		Alternative 1		Alternative 2		Alternative 3		Alternative 4		Alternative 5		Alternative 6	
		% sample		% sample		% sample		% sample		% sample		% sample	
		Pre-bias	Post-bias	Pre-bias	Post-bias	Pre-bias	Post-bias	Pre-bias	Post-bias	Pre-bias	Post-bias	Pre-bias	Post-bias
Q1	Treatment	Dist=8,N =0		Dist=10,N=0		Dist=15, N=1		Dist=10, N=1		Dist=8,N =0		Dist=10, N=1	
	D3 rank: bias 1	7	**76 ***	16	**24**	18	**0**	19	**0**	20	**0 ***	20	**0**
	D7 rank: bias 2	52	**12 ***	24	**6**	19	**18**	0	**18 ***	0	**41 ***	5	**6 ***
	D8 select: bias 2	33	**5 ***	8	**0**	4	**5**	8	**9 ***	8	**0 ***	28	**82 ***
Q2		Dist=20, N=0		Dist=15, N=0		Dist=10,N=0		Dist=8,N =1		Dist=20, N=1		Dist=20, N=1	
	D3 rank: bias 1	23	**10**	14	**0**	9	**0**	45	**81 ****	9	**0**	0	**10**
	D7 rank: bias 2	17	**12**	11	**0**	6	**0**	28	**0 ****	22	**6**	17	**81**
	D8 select: bias 2	25	**9**	4	**5**	17	**0**	50	**86 ****	4	**0**	0	**0**
Q3		Dist=10,N=0		Dist=8,N =1		Dist=8,N =1		Dist=15, N=1		Dist=10,N=0		Dist=15, N=0	
	D3 rank: bias 1	18	**10**	36	**5 ***	9	**14 ***	18	**67**	5	**0**	14	**5**
	D7 rank: bias 2	32	**6**	5	**0 ***	16	**0 ***	11	**44**	16	**31**	21	**19**
	D8 select: bias 2	17	**0**	17	**55 ***	30	**41 ***	9	**5**	22	**0**	4	**0**
Q4		Dist=15, N=1		Dist=20, N=0		Dist=8,N =1		Dist=20, N=0		Dist=20, N=1		Dist=10,N=0	
	D3 rank: bias 1	5	**5**	0	**0**	32	**86 ****	37	**0**	27	**5**	0	**5**
	D7 rank: bias 2	16	**6**	16	**69**	26	**19 ****	11	**0**	21	**0**	11	**6**
	D8 select: bias 2	4	**0**	26	**9**	39	**82 ****	13	**5**	13	**0**	4	**5**
Q5		Dist=20, N=1		Dist=15, N=0		Dist=10,N=0		Dist=8,N =0		Dist=15, N=1		Dist=8, N=1	
	D3 rank: bias 1	5	**5**	0	**0**	18	**0**	14	**0 ***	9	**5**	55	**90 ***
	D7 rank: bias 2	5	**6**	0	**0**	0	**38**	11	**6**	21	**19**	63	**31 ****
	D8 select: bias 2	0	**0**	13	**0**	9	**5**	35	**9**	35	**5**	9	**82 ****
Q6		Dist=10,N=0		Dist=20, N=1		Dist=15, N=0		Dist=15, N=1		Dist=8, N=0		Dist=15, N=1	
	D3 rank: bias 1	27	**14**	18	**81**	9	**0**	36	**5**	9	**0 ****	0	**0**
	D7 rank: bias 2	16	**19**	0	**12**	0	**0**	58	**19 ***	5	**44 ***	21	**6 ***
	D8 select: bias 2	23	**9**	5	**5**	18	**0**	5	**27 ***	0	**9 ***	50	**50 ***
Q7		Dist=15, N=1		Dist=10,N=1		Dist=20, N=0		Dist=20, N=0		Dist=10,N=1		Dist=8, N=0	
	D3 rank: bias 1	5	**5**	0	**0**	9	**5**	45	**86**	32	**0**	9	**5 ****
	D7 rank: bias 2	5	**13**	16	**0 ***	0	**44**	42	**44**	32	**0 ***	5	**0 ***
	D8 select: bias 2	5	**0**	55	**59 ***	0	**5**	14	**0**	9	**27 ***	18	**9 ***
Q8		Dist=8, N=1		Dist=8, N=1		Dist=20, N=1		Dist=10,N=0		Dist=15, N=0		Dist=20, N=0	
	D3 rank: bias 1	18	**81 ***	14	**14 ***	9	**0**	23	**5**	36	**0**	0	**0**
	D7 rank: bias 2	37	**88 ***	26	**6 ***	5	**0**	11	**0**	21	**0**	0	**6**
	D8 select: bias 2	23	**68 ***	5	**23 ***	27	**9**	32	**0**	9	**0**	5	**0**

Treatment codes, whether they select or rank, and whether they bias 1 or 2 attributes in Column 2.
Attribute levels for the potential target attributes - Distance and North/South (1,0) are reported.
In ranking designs, Alternatives ranked 1 recoded to selected.
% of sample selecting each Alternative reported. Values in **bold** are the post-bias treatment
** indicates a single alternative is identified as dominant for that condition, * indicates one of several alternatives that may be considered dominant.

In Q1, samples D3 and D8 perform as expected, with a concentration on the lower distances, and D8 making a tradeoff between 8km and 10km based on the North attribute. However, D7 shows a considerable spread, with a significant proportion selecting Alternative 3, even though it is dominated by Alternative 4 and 6. In Q2, D7 again generates an anomaly selecting Alternative 6, which appears to be inferior to Alternative 4.In Q3, again D7 selects Alternatives (4 & 5) which appear inferior to either 2 or 3. D3 also appears to be selecting dominated Alternative 4. In Q4, Q5 and Q7 D7 selects Alternatives 2, 3 and 3-4 respectively with greater frequency than one might expect given the objectives set. However, in question 8, all three align, selecting Alternatives that minimise distance or reward North. Although not present everywhere, it would appear from this analysis that when asked to rank 6 Alternatives, and attempt to achieve bias in 2 outcomes, respondents have difficulty. This issue will be returned to in the statistical analysis later.

In Table 7 below, in Q1, both D4 and D10 seem to over select Alternative 5, which has the highest distance. In Q2, both ranking samples again seem to favour high distances (one would have expected that Alternatives 1, 5 and 6 would be selected). In Q3, 93% of D4 (who rank) select Alternative 6, which has a higher distance, while 25% of D10 select Alternative3, which would appear to be dominated by Alternative 2. In Q4, again D4, who rank, select Alternative 6, which with a distance of 10, is greater than Alternative 4. Question 5 seems to be consistent, but in Q6 88% of D10 select Alternative 5, which has the smallest distance, but South, while zero respondents select Alternative 2, which has both the smallest distance plus North and would hence appear to dominate 5. In Q7, both groups that rank select Alternative 2, which has the highest distance, and is clearly dominated by Alternative 4. In Q8, D4 again selects Alternative 2, which would appear to be dominated by either 1 or 3.

Table 7. Percentage of sample selecting an alternative, pre-bias and post-bias. Design with 6 alternatives and 6 levels, by question.

		Alternative 1		Alternative 2		Alternative 3		Alternative 4		Alternative 5		Alternative 6	
		% sample Pre-bias	Post-bias	% sample Pre-bias	Post-bias	% sample Pre-bias	Post-bias	% sample Pre-bias	Post-bias	% sample Pre-bias	Post-bias	% sample Pre-bias	Post-bias
Q1	Treatment	Dist=14, N=0		Dist=10,N=1		Dist=10,N=1		Dist=12, N=1		Dist=20, N=0		Dist=8, N=0	
	D4 rank: bias 1	4	4	15	0	11	0	15	0	33	96	22	0 **
	D5 selct: bias 1	4	4	32	0	14	0	0	4	4	0	46	93 **
	D10 rank: bias 2	4	0	0	0 *	29	8 *	22	0	33	84	11	8 *
	D11 selct: bias 2	8	9	22	48 *	9	39 *	0	0	9	0	52	4 *
Q2		Dist=8, N=0		Dist=14, N=0		Dist=10,N=0		Dist=20, N=1		Dist=8, N=1		Dist=8, N=1	
	D4 rank: bias 1	19	35 *	19	46	19	15	26	0	15	0 *	4	4 *
	D5 select: bias 1	18	29 *	32	14	4	0	4	0	36	32 *	7	25 *
	D10 rank: bias 2	36	0	44	8	4	40	16	48	0	4 *	0	0 *
	D11 selct: bias 2	17	4	26	4	0	0	4	0	39	61 *	13	30 *
Q3		Dist=16, N=1		Dist=8, N=1		Dist=8, N=0		Dist=14, N=0		Dist=12, N=0		Dist=14, N=0	
	D4 rank: bias 1	15	7	35	0 *	12	0 *	8	0	23	0	8	93
	D5 select: bias 1	11	4	14	36 *	68	57 *	4	0	0	0	4	4
	D10 rank: bias 2	4	0	23	48 **	19	24	12	16	15	4	27	8
	D11 selct: bias 2	9	4	9	87 **	74	9	9	0	0	0	0	0
Q4		Dist=20, N=1		Dist=16, N=0		Dist=12, N=1		Dist=8, N=0		Dist=10,N=1		Dist=10,N=0	
	D4 rank: bias 1	12	8	12	0	15	0	38	0 *	12	0	12	92
	D5 select: bias 1	21	0	4	11	4	0	4	71 *	36	14	32	4
	D10 rank: bias 2	15	0	12	4	8	36	35	4 *	12	44 *	19	12
	D11 selct: bias 2	4	9	9	4	0	4	4	0 *	52	83 *	30	0
Q5		D=10,N=1		Dist=10,N=0		Dist=16, N=0		Dist=10,N=0		Dist=16, N=1		Dist=20, N=1	
	D4 rank: bias 1	27	15 *	23	65 *	31	15	15	0 *	4	4	0	0
	D5 select: bias 1	14	21 *	7	39 *	0	0	61	39 *	14	0	4	0
	D10 rank: bias 2	15	88 **	31	8	8	4	23	0	23	0	0	0
	D11 selct: bias 2	17	91 **	9	0	9	0	43	4	22	4	0	0
Q6		Dist=12, N=0		Dist=8, N=1		Dist=14, N=0		D=10,N=1		Dist=8, N=0		D=10,N=1	
	D4 rank: bias 1	4	4	12	0 *	19	4	23	12	38	76 *	4	4
	D5 select: bias 1	0	0	75	71 *	14	4	4	0	7	25 *	0	0
	D10 rank: bias 2	4	0	12	0 **	15	4	27	4	31	88	12	4
	D11 selct: bias 2	0	4	70	83 **	22	4	9	4	0	0	0	4
Q7		D=10,N=0		Dist=20, N=1		Dist=20, N=1		Dist=8, N=1		Dist=14, N=0		Dist=16, N=0	
	D4 rank: bias 1	4	4	12	92	12	0	27	0 **	42	0	4	4
	D5 select: bias 1	4	7	0	4	21	4	57	86 **	0	0	18	0
	D10 rank: bias 2	8	0	12	48	8	0	15	44 **	54	8	4	0
	D11 selct: bias 2	0	4	4	0	9	0	70	96 **	0	0	17	0
Q8		Dist=8, N=1		Dist=12, N=0		Dist=8, N=0		Dist=16, N=0		D=10,N=1		Dist=12, N=1	
	D4 rank: bias 1	27	58 *	42	42	12	0 *	12	0	4	0	4	0
	D5 select: bias 1	32	32 *	4	4	46	57 *	18	0	0	4	0	4
	D10 rank: bias 2	32	76 **	32	20	24	0	8	0	0	4	4	0
	D11 selct: bias 2	17	57 **	9	0	48	39	22	0	4	4	0	0

Treatment codes, whether they select or rank, and whether they bias 1 or 2 attributes in Column 2.
Attribute levels for the potential target attributes - Distance and North/South (1,0) are reported.
In ranking designs, Alternatives ranked 1 recoded to selected.
% of sample selecting each Alternative reported. Values in **bold** are the post-bias treatment
** indicates a single alternative is identified as dominant for that condition, * indicates one of several alternatives that may be considered dominant.

6. Summary of descriptive analysis

This analysis is based on the assumption that respondents would have responded to the request to bias their answers by selecting alternatives that contain the required outcomes: the shortest distance and/or North of the river. Where they are required to bias both, and it is not possible to optimise on both simultaneously one would expect to see some dispersion across alternatives, which one does see. However, there are some cases, especially in the samples who are asked to rank all Alternatives where they appear to not be following this strategy. It is possible that the ability to rank introduced a more subtle strategy in the case where 2 attributes are being optimised for: iteratively 'favouring' one alternative in the first selection, and then the other in the second, but its not clear that this will rationalise all anomalies. Alternatively they may be considering other attributes jointly i.e. avoiding alternatives with low cost (even if they have the lowest distance and North) so as to avoid appearing to select 'rationally' on that basis. Unfortunately the designs are not full factorials, so that revealed behaviours cannot be compared with all possible strategies.

Table 8 provides a summary of the previous descriptive analysis, by identifying the percentage of respondents who selected what has been identified as a "dominated" alternative as their preferred option within each choice set question i.e. who make a 'wrong' choice. They are grouped by statistical design, as per the previous 4 Tables. Overall, the results indicate that when asked to select a single best alternative, the choices conform to expectations, or if they rank over 3 alternatives. When ranking, with 6 alternatives, they do not. The only counter example to this is D8, which has a reasonable good performance.

In the following section we will consider comparisons of econometric models.

Table 8. Percentage of respondents selecting a 'dominated' alternative within a choice set, by treatment and question.

		q1	q2	q3	q4	q5	q6	q7	q8
3 alternatives and 4 levels									
D2	Select: Bias 1	4	0	0	8	12	0	13	12
D1	Rank: Bias 1	96	0	4	4	16	12	96	16
D6	Rank: Bias 2	92	12	12	4	16	20	8	16
3 alternatives 6 levels									
D9	Select: Bias 2	99	15	5	48	5	0	5	5
6 alternatives and 4 levels									
D3	Select: Bias 1	24	20	82	15	10	100	96	5
D7	Rank: Bias 1	24	99	100	81	69	31	101	6
D8	Rank: Bias 2	5	14	5	19	19	14	5	9
6 alternatives and 6 levels									
D4	Rank: Bias 1	100	61	100	100	19	24	100	42
D5	Select: Bias 1	8	14	8	29	0	4	15	12
D10	Rank: Bias 2	84	96	52	52	12	100	56	24
D11	Select: Bias 2	22	8	13	17	8	16	4	43

7. Estimation

An issue with the analysis of the data is how to identify the extent of any change in behaviour between the first and second rounds, i.e. without and with an explicit incentive to bias the outcome. The instructions simply said (for the case of influencing a single attribute):

> *In the following section, you will be presented with an additional set of questions. When answering these questions you should behave as if you want to give the impression that*
>
> > *Being CLOSE to UWA as the most important attribute to you,*
> > *You want this to be identified as more important than the other attributes of the accommodation.*

There was no guidance given to the respondent as to how the data would be analysed, or how the measure of ‘more importance’ will be calculated. This is probably consistent with the process in most choice experiments: respondents are asked to make choices, and the subsequent manipulation of the data is not made explicit (the exception to this would be “one shot” CEs where an individual only answers one choice question, and which explicitly deal with incentive compatibility by stating that the alternative that is selected most frequently will be implemented). As such, this lack of clarity in process may be considered as a contributing factor towards the ability of CEs to thwart strategic bias, but in the current context it makes modelling the strategies that the respondents used when completing the questions under the ‘bias’ incentive clear.

An obvious technical interpretation of “preference” is that an attribute is preferred if the partworth of the attribute is largest, compared to the other attributes. Thus bias in preferences would imply bias in partworths. However, the partworth depends on a measure of the marginal utility of cost (it’s the (negative) ratio of attribute parameter to cost parameter). If cost is ignored by the respondent, the estimated coefficient will tend to zero, and hence the partworth gets inflated. Three issues then arise:

all partworths will be inflated, including those for non target attributes,

all partworths will tend to become insignificant

if the marginal utility of cost becomes positive (which in the estimation it may) then the implication will be that the attribute is NOT preferred at all, if one takes the conventional interpretation of the partworth.

Given that a feasible response is to lexicographically place weight upon only the target attributes and ignore all others, the possibility of a zero (or insignificant) coefficient on cost is high.

Below we present a method of evaluating effectiveness of the biasing activity, based on the probability of selecting an alternative.

The initial task is to estimate models over the initial 8 questions. Table 9 reports conditional logit results: in those cases where the treatment asked for rankings, the alternative with the highest rank is used as the selected alternative, so that the estimation can use a common basis. In addition, Table 10 reports the Rank Ordered Logit (Hair et al., 2010, Hausman & Ruud, 1987) results for those samples that were required to rank alternatives. All estimation has been undertaken with Stata Version 11 (StataCorp. 2009)

In general, the parameters are well defined, with expected signs on cost and distance to UWA. The results with 6 alternatives which require ranking have relatively few significant, suggesting that at this level of complexity some cognitive dissonance is setting in.

The second column of estimates in each cell represent the estimates of the post-bias data for the same models.

Table 9. Results for both pre-bias and post-bias data, using conditional logit models, by design.

Number of alternatives		Selection method	Distance, 4 attribute levels	Distance, 6 attribute levels	Distance and NS, 4 attribute levels	Distance and NS, 6 attribute levels
	3	Rank	D1		D6	D9
	3	Select	D2			
	6	Rank	D3	D4	D7	D10
	6	Select		D5	D8	D11

D1

	Pre-bias	Post-bias
cost	-0.013***	0.006**
dist	-0.069***	-0.052**
num	-0.187**	0.061
furn	-0.123	-0.255
Ns	0.084	1.414***

D6

	Pre-bias	Post-bias
cost	-0.012***	-0.014***
dist	-0.039**	-0.150***
num	-0.161**	0.031
furn	-0.168	-0.716***
Ns	-0.252	1.827***

D9

	Pre-bias	Post-bias
cost	-0.009***	-0.026***
dist	-0.169***	-0.241***
num	-0.112	0.385***
furn	-0.252	2.469***
ns	0.133	0.667**

D2

	Pre-bias	Post-bias
cost	-0.009***	0.013***
dist	-0.076***	-0.319***
num	-0.498***	0.002
furn	0.339*	-0.305
Ns	0.456	0.259

D3

	Pre-bias	Post-bias
cost	-0.001	-0.013***
dist	-0.051***	-0.099***
num	-0.197**	0.141
furn	0.037	0.382**
ns	0.025	1.080***

D4

	Pre-bias	Post-bias
cost	0.003	-.002
dist	-0.019	0.091***
num	-0.019	0.772***
furn	-0.271*	0.928***
Ns	-0.313**	-1.641***

D7

	Pre-bias	Post-bias
cost	-0.001	0.009***
dist	-0.044**	0.067***
num	-0.113	0.261**
furn	-0.101	-0.837***
Ns	0.374**	-0.397**

D10

	Pre-bias	Post-bias
cost	0.001	-0.006**
dist	-0.024	-0.20
num	0.070	0.539***
furn	0.026	0.316
ns	-0.486***	-0.047

D5

	Pre-bias	Post-bias
cost	-0.022***	-0.006**
dist	-0.183***	-0.548***
num	-0.451***	-0.127
furn	0.374**	0.412***
ns	0.516***	-0.167

D8

	Pre-bias	Post-bias
cost	-0.009***	-0.001
dist	-0.112***	-0.380***
num	-0.665***	-0.412***
furn	0.214	0.048
ns	0.360**	2.597***

D11

	Pre-bias	Post-bias
cost	-0.022***	-0.018***
dist	-0.183***	-0.367***
num	-0.316***	-0.011
furn	0.682***	0.025
ns	0.345**	2.871***

***= p<0.01, **= p<0.05, *= p<0.1

Table 10. Results for both pre-bias and post-bias data, using Rank Ordered Logit models, by design.

		Attributes being influenced											
		Distance						Distance and NS					
		Attribute levels											
		4			6			4			6		
Number of alternatives	Selection method	D1	Pre-bias	Post-bias				D6	Pre-bias	Post-bias	D9	Pre-bias	Post-bias
3	Rank	cost	-0.013***	0.005**				cost	-0.010***	-0.006***	cost	-0.011***	-0.016***
		dist	-0.063***	-0.063***				dist	-0.036***	-0.126***	dist	-0.122***	-0.107***
		num	-0.113**	0.098*				num	-0.093*	-0.151***	num	-0.192**	0.307***
		furn	-0.436***	0.081				furn	-0.418***	-0.537***	furn	-0.097	0.875***
		Ns	0.055	0.226				Ns	-0.261**	1.356***	ns	0.231	1.223***
		D3	Pre-bias	Post-bias	D4	Pre-bias	Post-bias	D7	Pre-bias	Post-bias	D10	Pre-bias	Post-bias
6	Rank	cost	-0.009***	-0.011***	cost	-0.001	0.006***	cost	-0.002**	0.002**	cost	-0.002**	-0.000
		dist	-0.028***	-0.114***	dist	-0.001	0.002	dist	-0.008	-0.046***	dist	-0.012	-0.030***
		num	-0.016	0.054	num	-0.001	-0.237***	num	0.089**	0.079	num	-0.004	0.071*
		furn	-0.028	0.164**	furn	0.235***	-0.210***	furn	-0.127	-0.353***	furn	0.214***	-0.009
		ns	0.014	0.684***	Ns	0.069	0.144*	Ns	0.112	0.346***	ns	0.020	0.440***

8. *Identification of Bias*

Ideally one would like to identify a measure of bias that is holistic, and does not require a dependence on a significant (and negative) cost parameter for interpretation. Here we propose a measure based on the simulated probability of choosing an alternative. We do this by simulating a 2 alternative choice situation based on the parameters estimated from the pre-bias condition and the post-bias condition.

Consider the case where only distance is being targeted. Assume there are two alternatives (i=1,2), and that the parameters of the utility function are those obtained using the “unbiased” responses within treatment t. Assume that the attribute levels for alternatives 1 and 2 are given by:

	Alternative 1	Alternative 2
North/South	0	0
Furnished	0	0
Distance	8	12
Cost	100	100+ PW_{d4}

Then it is possible to identify a value for PW_{d4} such that[2]

$$P(y=1) = f(A_{t1}, A_{t2}, \hat{\beta}_t) = 0.5$$

where A_{ti} are the attribute levels, and $\hat{\beta}_t$ the pre-bias parameter estimates for treatment t.
Thus, the value of PW_{d4} is identified that results in the respondent being indifferent between the two alternatives.

The measure of success in achieving bias is then identified by calculating:

$$P_b(y=1) = f(A_{t1}, A_{t2}, \hat{\beta}_{tb})$$

where $\hat{\beta}_{tb}$ is the parameter vector estimated for treatment t in the incentive to bias phase. If the design is such that it has not been possible to induce bias, then $P_b(y=1) = 0.5$ i.e. predicted choice is unchanged. A successful bias outcome would increase the value beyond 0.5. It is necessary to calibrate the initial probability at 0.5 (as opposed to comparing the probabilities associated with the actual design attributes), because it would be possible

[2] Note that the value of PW_{d4} needed to achieve equality in the probability of selection will, in fact, be the partworth associated with a 4 unit change in distance.

for some configuration of attributes and parameters to lead to a high probability of selecting the first alternative without bias. Because of the non-linear nature of the logit probability function, that would then reduce the extent to which one could identify biasing behaviour.

In the case where the 2 attributes are targets for bias, in addition the North/South variable is set to 1 in the attribute set for Alternative 1, and zero for Alternative 2, and the calibrating level of cost is determined by $PW_{d4\text{-}ns}$ i.e. the partworth associated with a 4 unit change in distance **and** a change in location from South to North.

There are two alternative vectors of parameters for those treatments where respondents have to rank alternatives. They can either be estimated using a Ranked Ordered Logit (ROL) model to take advantage of the full ranking data, or the alternative ranked 1 can be used within a conventional CL model. As a result one can identify two alternative measures of strategic bias, based on which of these approaches is used. Results for both are reported in Table 11 below (the value in parenthesis in each cell is the ROL result).

Table 11: Implied probability of selecting the 'target' alternative

			Attributes being influenced			
			1		2	
			Attribute levels			
			4	6	4	6
Number of alternatives	3	Rank All	D1 0.73 (0.74)		D6 0.97 (0.94)	D9 0.08 (0.53)
		Select 1	D2 0.99			
	6	Rank All	D3 0.01 (0.53)	D4 0.28 (0.52)	D7 0.99 (0.75)	D10 0.21 (0.69)
		Select 1		D5 0.99	D8 0.99	D11 0.99

Thus, the results for treatment D1 show that the simulated probability of selecting alternative 1 (the alternative which has the desirable level of distance) is increased from 0.5 under the pre-bias condition, up to 0.73 under the incentive to bias condition. Thus, one could conclude that this group have achieved a reasonable degree of success in conveying their strategic preferences. The conclusion is very similar (0.74) if the parameters estimates obtained from the ranked ordered logit results are used. However, this success is dwarfed by the success of those in treatment D2, who shift the probability of selecting the target alternative from 0.5 up to 0.99. In treatment D3, on the other hand, where respondents had to rank across a 6-alternative choice set, the conclusions from the simulation are quite different. Using just the first ranked as the selected option as the basis for identifying parameters, the probability of selecting the target alternative drops from 0.5 to 0.01 i.e. their observed behaviour in the incentive-to-bias setting gives absolutely no indication of the preferred attribute they were targeting. Although the results improve if the parameters from the ROL are used (the probability rises to 0.53) this is effectively no different to the baseline probability of 0.5.

To the extent that there is any consistent pattern emerging across these results, it would appear to be that when asked to select just one alternative, consistently high levels of bias can be induced in the results, irrespective of the number of attributes levels or alternatives, or the number of attributes targeted. When asked to rank all alternatives this capacity is reduced (the exception is when ranking with 4 attribute levels and influencing one attribute: its not clear why that is the case). It is possible that the possibility of ranking alternatives offers the opportunity for a more complex heuristic that is then not being captured in the simple linear utility function representation i.e. that selection strategies vary across the sequence of choices, while the ROL model assumes the same process is being used in each step.

9. Conclusions

This study is the first attempt we are aware of that attempts to actively induce strategic bias within a choice experiment setting. The results from this simple exercise suggest that a motivated respondent can influence the outcomes from a Choice Experiment to some considerable degree. This is particularly true within the context of selecting a single, most preferred, alternative which is the dominate elicitation method within the CE literature. These results are not affected by other elements of the design. The extent of the bias may be reduced if the respondents are required to rank the full set of alternatives. Whether the latter conclusion arises because the act of ranking makes the decision harder, or whether it is because the statistical model that has been estimated does not capture the heuristic the individuals have adopted is open to question.

To link these results to the ideas raised in the literature on CM and bias:

> "As such, CE may minimise some of the response difficulties found in CVM… (protest bids, strategic behaviour, yeah saying). But this point has yet to be demonstrated." Hanley et all (2001) p448.

If anything, the results from this study demonstrate the opposite.

> "Strategic Behaviour should be minimal in Stated Preference tasks since the choices are made from descriptions of attributes and it is not clear which choice will over- or under represent a valuation" Adamowicz et al , (1999) (p467).

Again, if the intention is to exhibit a strategic bias in favour of one (or more) of the attributes, then the results would appear to contradict this.

> "By adding complexity to the SP task, respondents may exhibit less bias. This may be partly occur because of the extra effort required to complete the exercise with bias, but it is more likely to occur because of respondents failing to see any clear single purpose to the exercise" Lu et al (2008) (p128).

The first of these (the impact of effort) would appear to be confounded: even with relatively complex structures, a well motivated respondent seems to be able to apply the effort required to induce bias. The second is less clear, given the respondents were given very explicit information about the purpose of the exercise: they were asked to overstate values for specific attributes. But whether an individual with a particular motivation towards an attribute would fail to identify the appropriate response without prompting would seem to be unlikely.

Whether in a more complex structure, such as a ranking, the respondent struggles to identify how to influence the (to them) unknown statistical process of inferring values, would seem to be an open question. However, of more importance is not whether the respondent has failed to bias the response in the way that they required, but whether they have generated choices which reveal useful information about their true preferences. It would be a pyrrhic victory if the complexity of the design that prevents strategic bias leads to choices that reveal little about true preferences.

References

Adamowicz, W. Boxall,P.C., Louviere,J.J..,and Swait,J. (1999) Stated preference methods for evaluating environmental amenities, in Bateman, IJ. and Willis, KG (eds) *Valuing environmental Preferences: Theory and Practice of the Contingent Valuation Methods in the US, EU and Developing Countries*: 460-482

Carson, R.T. & Groves,T. (2007) Incentive and informational properties of preference questions *Environmental and Resource Economics* 37: 181-210

Hair Jr., J. F., W. C. Black, B. J. Babin, and R. E. Anderson. (2010). *Multivariate Data Analysis*. 7th ed. Upper Saddle River, NJ: Pearson.

Hanley,N. Mourato,S. and Wright, R. (2001) Choice modelling approaches: a superior approach for environmental valuation? *Journal of Economic Surveys* 15(3): 435-462.

Hausman, J. A., and P. A. Ruud. 1987. Specifying and testing econometric models for rank-ordered data. *Journal of Econometrics* 34: 83–104.

Hensher, DA (2006) How do respondents process stated choice experiments? Attribute consideration under varying information load *Journal of Applied Econometrics* 21:6 pp861-878

Hensher, D.A, Rose,J. and Greene,W.H. (2005) *Applied Choice Analysis: a primer* Cambridge University Press, New York.

Lu,H., Fowkes,T. and Wardman,M. (2008) Amending the incentive for strategic bias in stated preference studies *Transportation Research Record* 2049 :128-135.

Mazar,N. Amir,O., Ariely,D. (2008) The Dishonesty of Honest People: A Theory of Self-Concept Maintenance *Journal of Marketing Research*, Volume 45, Issue 6, p. 633-644.

Ngene v.1.0.0 (2009) Rose, J. Collins, A. Bliemer, M, Hensher, D.

Scarpa, R. and Rose,J. (2008). Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why *Australian Journal of Agricultural and Resource Economics*, vol. 52(3), pp 253-282

StataCorp. 2009. Stata: Release 11. Statistical Software. College Station, TX: StataCorp LP.

Appendix 1

Text version of survey. Separators identify page splits in the web-based survey

---

Thank you for agreeing to take part in this survey.
Its purpose is to help understand how people answer a particular type of survey question, where one of several options has to be selected. The context that is being used is rented accommodation.

As a reward for completing the survey you will be offered an entry into a prize draw for $400.
To be eligible for the draw you will have to complete it fully, to the very end. It should only take about 5-10 minutes.

It is assumed that completing the survey implies you are giving consent to participate in this study.
It will be possible to link your responses to your email address, as this is needed to allow us to enter you into the draw, but once the draw for the prize has been completed, all links between the survey and your personal details will be deleted. Because of this, once submitted to us, it may not be possible to withdraw your data from the study.   If you have any questions please feel free to contact me at the email address below.

Yours truly,

Dr. Michael Burton

michael.burton@UWA.edu.au

The Human Research Ethics Committee at the University of Western Australia requires that all participants are informed that, if they have any complaint regarding the manner, in which a research project is conducted, it may be given to the researcher or, alternatively to the Secretary, Human Research Ethics Committee, Registrar's Office, University of Western Australia, 35 Stirling Highway, Crawley, Wa 6009 (telephone number 6488-3703).

There is currently a considerable shortage of rented accommodation in Perth, making it difficult for people to identify somewhere to live which has the right characteristics, at an affordable price.

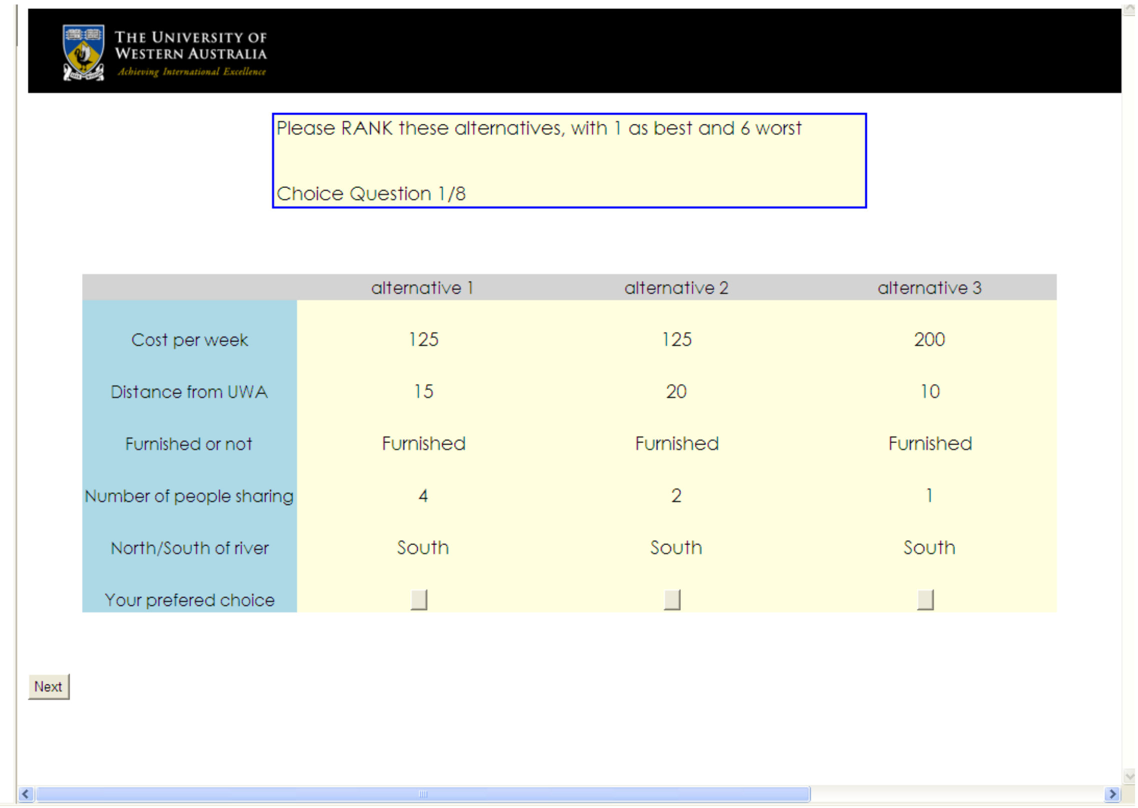
In this survey you will be presented with a number of descriptions of hypothetical rented accommodation, and asked to rank them in order of preference.

The attributes of the accommodation will include:

The rent per week (bills included)
The total number of people sharing the flat/house
Whether it is furnished/unfurnished
Whether its North or South of the river
The distance from UWA

When looking at the options, and selecting the one that you prefer, you should imagine that you are in the position of having no other alternative than to accept one of these, or be homeless.

There then followed 8 choice questions of the form:



The University of Western Australia
Achieving International Excellence

Please RANK these alternatives, with 1 as best and 6 worst

Choice Question 1/8

	alternative 1	alternative 2	alternative 3
Cost per week	125	125	200
Distance from UWA	15	20	10
Furnished or not	Furnished	Furnished	Furnished
Number of people sharing	4	2	1
North/South of river	South	South	South
Your prefered choice			

Next

The type of questions that you have just completed are very commonly used in valuing new products or environmental assets.
From the choices made, and the levels of the attributes that are included in the alternatives, it is possible to identify how, on average, the respondents to the survey are trading off the different attributes of the accommodation.

One issue is the extent to which they are open to manipulation: that people will not give their true answers to the question because they want to try and influence the outcomes in a particular way. It is unlikely that you were doing this, but it may be the case where people try and overstate the importance of some feature, in an effort to change public policy, or change the type of product provided.

In the next section of this survey, you will be asked to deliberately change the way that you answer the questions, to mimic this type of biased response.

Please read the following information carefully. Understanding it will have a strong impact on your chances of winning the $400.

In the following section, you will be presented with an additional set of questions. When answering these questions you should behave as if you want to give the impression that

> Being CLOSE to UWA as the most important attribute to you,
> You want this to be identified as more important than the other attributes of the accommodation.

a group of 40 of you have been given the task of trying to influence the value attached to being CLOSE to UWA

Other groups of respondents have been given the task of biasing the importance placed on other attributes.

The extent to which you, as a group, are more successful in manipulating the outcomes of the study will change the likelihood that you will win the $400.

Everyone who completes the survey will be awarded 1 entry into the draw. Each person in the group of 40 who manage to bias their designated attribute the most will be awarded 3 entries each into the draw, increasing their chance of winning the $400.

The winner of the $400 will be drawn at random, and the draw will be monitored, to ensure that the process outlined above is followed.

There then followed the same 8 choice questions

What strategy did you use for trying to achieve the most bias in your answers?

____

Finally, some questions about yourself:

Do you live in rented accommodation?	Yes No
What is your postcode?	____
What is your age?	____

Thank you for completing the survey. In order to ensure that I can contact you if you win the prize draw, please enter your UWA email address. ONLY valid UWA student addresses will be eligible for the prize draw

____

repeat for confirmation

____

Thank you very much for completing this survey.

The study is currently undersubscribed. If you know of any UWA students who might be willing to complete this survey, please direct them to the original sign-up page:

http://surveys.webservices.uwa.edu.au/sw/wchost.asp?st=cerfintro

If you have any questions about the study, please feel free to contact me on:

Michael.burton@uwa.edu.au

Yours,

Dr. Michael Burton

The Human Research Ethics Committee at the University of Western australia requires that all participants are informed that, if they have any complaint regarding the manner, in which a research project is conducted, it may be given to the researcher or, alternatively to the Secretary, Human Research Ethics Committee, Registrar's Office,
University of Western Australia, 35 Stirling Highway, Crawley, Wa 6009 (telephone number 6488-3703).