

Comparing Responses from Internet and Paper-Based Collection Methods in more Complex Stated Preference Environmental Valuation Surveys

Jill Windle

*Centre for Environmental Management
Central Queensland University
North Rockhampton QLD Australia 4702
(Email: j.windle@cqu.edu.au)*

and

John Rolfe

*Centre for Environmental Management
Central Queensland University
North Rockhampton QLD Australia 4702
(Email: j.rolfe@cqu.edu.au)*

Abstract: Internet surveys are becoming an increasing popular survey collection method because collection times are quicker and survey costs are lower than other collection techniques. Many studies have been conducted overseas to compare the effects of survey collection modes with results still remaining inconsistent. Fewer studies have compared collection methods for nonmarket valuation surveys, particularly for the more complex stated preference, choice modelling surveys. In this study, a comparison of internet and paper-based surveys is made to determine if the results for overseas studies can be replicated in Australia. The valuation exercise elicited values from Brisbane respondents for future improvements in the environmental condition of the Great Barrier Reef. The results indicate that there were some socio-demographic and attitudinal differences between the two samples and the models developed to explain the influence on choice selection were also significantly different. However, no differences in value estimates were found in the final results; household willingness-to-pay for an improvement in the condition of the GBR was equivalent across collection methods.

I. INTRODUCTION

Determining an appropriate survey collection method is an important consideration in the design of all questionnaire surveys. However, as survey complexity increases, such as in the collection of data in stated preference surveys, the choice of collection mode requires more careful consideration, in case biases are generated (Mitchell and Carson 1989, Bennett and

Blamey 2001, Bateman *et al.* 2002, Champ *et al.* 2003, Alberini and Khan 2006). Stated preference surveys have traditionally been collected using mail-outs, face-to-face interviews and telephone interviews (Bateman *et al.* 2002, Champ 2003, Alberini and Khan 2006). The most complex of these valuation surveys is the choice modelling (CM) technique. Respondents are required to make a series of choices where they have to make tradeoffs between different levels of different attributes. The added complexity of CM surveys, and the need to show choice tasks to respondents, means that mail-out and face-to-face interviews have been the primary collection modes. Limitations of these collection modes are the high costs involved, the difficulties of generating high response rates and representative community samples, and the intensity of effort and time involved. In Australia, the use of a drop-off/pick-up collection technique, as applied in the Australian Bureau of Statistics main Census survey, is commonly adopted to collect choice experiments (Bennett and Blamey 2001).

An alternative method of survey collection is to use internet surveys. This collection technique is becoming more popular as there is growing familiarity with internet usage across most sectors of the community. Key advantages of internet surveys are low collection costs, rapid collection times, increased flexibility of tailoring questionnaires to respondent groups, and increased automation of data recording and coding (Berrens 2003, Marta-Pedroso 2007, Fleming and Bowden 2009, Maguire 2009, Olsen 2009). In addition, internet formats are able to incorporate new and innovative design features and information provision.

The most commonly cited disadvantages of internet surveys are potential sample frame bias (non-random exclusion of individuals who do not use the internet) and response bias (responses of those who respond may be different from those who do not) (Bateman *et al.* 2002, Champ 2003, Marta-Pedroso 2007, Fleming and Bowden 2009, Olsen 2009). However, other survey collection methods may also be associated with sample frame bias and response bias as access to, and involvement with, different groups in society can be expected to vary across collection techniques. As access and familiarity with internet communication increases in society, there may be a convergence between the total sampling frame biases associated with this survey collection method compared to other methods.

There is a growing pool of studies that examine whether internet collection elicits a different pattern of responses than other collection methods in stated preference surveys, but there is limited evidence about whether the results can be replicated in Australia. There is also little examination of the use of pre-recruited internet panels for survey collection. Unlike the general public, these respondents complete surveys on a regular basis and are survey savvy, and may respond differently to the completion of more complex surveys, such as CM surveys.

In this paper, a comparison is made of the results from a CM survey using two collection methods. Identical surveys were collected in 2009 in a paper-based format using a drop-off/pick-up collection technique and in an internet format using a pre-recruited internet panel. The CM surveys were focused on eliciting values for improvements in the environmental condition of the Great Barrier Reef (GBR). The total cost of the paper-based survey was approximately \$70 per survey and took three months to complete. In contrast, the internet survey cost approximately \$15 per survey and was completed in two weeks. This differential is likely to widen in the future as it is increasingly difficult to collect surveys in urban areas as access to residents becomes more difficult with increases in housing density and security

gates surrounding blocks of flats and units. On the other hand, increased demand for internet panels is leading to increased supply and more competition is driving down costs.

Tests for differences generated by the two collection methods are focused on whether there were sampling differences and whether the estimated protection values varied across the two sample groups. The results indicate that there are some demographic and attitudinal differences between the sample groups and the models developed to explain the influence on choice selection were also significantly different. However, no differences in values were found in the final results; household willingness-to-pay (WTP) for an improvement in the condition of the GBR was equivalent across collection methods.

This paper makes an important contribution to the literature on the influence of internet collection in complex surveys generally in Australia, and in stated preference surveys in particular. It also provides new insights into the use of survey savvy pre-recruited internet panellists. The paper is outlined as follows. In the next section, an overview is given of collection comparisons that have been made for stated preference surveys. In Section III, details are provided about the choice modelling case study and the results are outlined in Section IV. The implications are discussed and conclusions drawn in the final section.

II. COLLECTION MODE AND STATED PREFERENCE SURVEYS

The literature on collection method comparisons for stated preference surveys has generally focused on three main areas; sample frame and response bias, differences in model performance, and survey results (WTP estimates).

The wide availability and coverage of pre-recruited internet panellists has been a relatively recent development as many earlier internet surveys were administered by researchers themselves and hosted on their organisational websites. This led to difficulties in recruiting respondents, with several studies reporting low response rates when using email or internet surveys. Shih and Fan (2009) examined 35 studies and although individual studies reported inconsistent findings about response rate differences, their meta-analysis suggests that the response rates from internet surveys were on average about 20% lower than mail surveys. However, the response rates from pre-recruited internet panels are different as they are not only related to the nature of survey being completed, but also to the terms and conditions associated with panel membership and the nature of the incentive that participants receive.

Several studies have compared the socio-demographic differences between respondents in stated preference surveys using internet and other collection modes, and all report some sample differences (Berrens *et al.* 2003, Canavari *et al.* 2005, Marta-Pedroso *et al.* 2007, Olsen 2009, Hatton Macdonald *et al.* 2010, Nielsen 2011). However, there are no consistent socio-demographic differences across studies. Where socio-demographic comparisons have been made between samples and the population, the inconsistencies continue with no collection method consistently providing a more representative sample (Berrens *et al.* 2003, Olsen 2009).

Comparisons of WTP estimates from stated preference surveys also reveal a similar pattern of inconsistency. Nielsen (2011) recently reported the WTP estimates from collection comparisons conducted in seven stated preference studies, including their own. In five studies, there was no significant difference between WTP estimates from internet collection models and face-to-face

(two), postal (two), and telephone collection (one). In contrast, WTP from internet surveys was found to be either higher than face-to-face interviews (Canavari *et al.* 2005) or lower than face-to-face interviews (Marta-Pedroso *et al.* 2007). Given the potential impact of interviewer bias in face-to-face interviews, the most consistent result from these methodological comparisons is that WTP estimates do not vary between internet surveys and other collection modes. In the only other CM internet collection comparison in Australia known to the authors, Hatton MacDonald *et al.* (2010) find survey collection effects in their comparison of internet and mail surveys in a CM valuation of WTP for improving water quality in the River Murray and the Coorong. They report lower WTP estimates for the internet survey sample compared with a mail survey.

The evidence presented above suggests that there is no reason not to use internet surveys, particularly in light of the practical benefits associated with the use of internet panels. While socio-demographic and WTP related factors (model performance) may vary across collection techniques, there is no clear evidence that internet samples are any more or less biased than other collection samples. However, there is little evidence in the literature that compares the attitudes of different respondent groups to a complex survey completion, particularly in Australia.

In the CM study outlined in this paper, a comparison is made between internet and paper-based surveys. Based on the evidence outline above the *a priori* expectation was that there would be collection differences in the socio-demographic characteristics of respondents and in model performance but that the WTP estimates would be the same.

Attitudinal information was also collected to help identify any underlying differences between the two test groups. A series of follow-up questions were included after the choice tasks to determine how respondents reacted to the complexity of the choice scenarios. There was some *a priori* expectation that the internet panel respondents, who regularly complete surveys, would find the choice tasks easier to complete than the paper-based respondents. It was unclear how these attitudes may impact on choice selection.

In a CM survey, respondents always have the choice to select a do-nothing option (also called a status quo or no-cost option). This option is designed to be selected by respondents who cannot afford one of the improvement options with an associated cost. There are a range of other reasons why this option might be selected, even when the respondent might really prefer one of the improvement options. For example, it might also be selected by respondents who object to paying extra (above what they already pay in taxes) for environmental improvements. If the status quo is selected in all choice tasks, it might be an indication of some form of protest (von Haefen *et al.* 2005, Boxall *et al.* 2009). Respondents in the paper survey have the option of not completing the survey as a form of protest, but internet panellists are provided with an incentive (for completed surveys) and therefore might be more likely to register their protest through non-participation by always selecting the status quo option. There was also *a priori* expectation that serial selection of the status quo option would be higher in the internet group.

III. CASE STUDY DETAILS

In 2009, a CM survey was conducted to estimate the non-market values of state capital (Brisbane) residents for improvements in the condition of the GBR. Two experiments were designed to explore respondents' preferences and how they may vary across different levels of geographic

scope. One experiment referred to the whole GBR and the other referred only to the Mackay-Whitsunday regional section of the GBR. Data for both experiments was gathered across two collection modes. The surveys were designed as a part of a much larger research project, and many issues associated with design and performance are reported in other research papers. The main focus of this study is to examine the differences associated with the collection method. The experiment was designed to keep the surveys equivalent across collection modes, with the whole and regional surveys the same, apart from the absolute levels for the area of GBR in good condition.

The CM technique requires respondents in a survey format to choose a single preferred option from a set of a number of resource use options (Bennett and Blamey 2001). The economic theory underlying CM assumes that the most preferred option yields the highest utility for the respondent (Louviere *et al.* 2000; Bennett and Blamey 2001). Respondents are presented with a number of similar choice tasks. Each choice task contains the same number of options which are described in terms of a common set of underlying attributes that vary across a set number of levels. The variation in the levels of attributes differentiates the options to respondents. By offering the combinations of attributes and levels in a systematic way through the use of an experimental design (Louviere *et al.* 2000), the key influences on choice can be identified (Rolfe 2006).

In this study, the choice scenario was framed in terms of a 25 year time period and was described in terms of three primary attributes:

- GBR CONDITION – area of the GBR in good condition.
- CERTAINTY – different levels of certainty were associated with the predicted outcomes to help frame the variability surrounding any predictions about current and future condition of the GBR.
- COST – an annual payment for a five year period. A shorter time frame than the 25 year valuation scenario was applied to match a more realistic policy investment scenario.

Each choice task included four options. The first option (referred to as the status quo) described the predicted level of GBR in good condition in 25 years time if no additional funding was allocated. It had no associated cost. This option remained the same in each choice task. Three other options were available where the level of the different attributes varied across options and across choice tasks. Each of the options was labelled in terms of the management option that would be applied to achieve the predicted benefits. The three management options: Improve water quality; Increase conservation zones; Reduce Greenhouse gases were designed to address the three main pressures impacting on the condition of the GBR. These labels remained constant across choice tasks. Example choice tasks for the whole and regional surveys are presented in *Figure 1*.

The choice tasks were the same for the whole and regional surveys and the levels for both the COST and CERTAINTY attributes remained the same. The levels for the GBR CONDITION attribute were described in both percentage and absolute levels and only the absolute levels (sq kms) varied across survey scope. A D-efficient experimental design was created to allocate attribute levels across different options. A 12 choice task design was generated which was used for both the whole and regional surveys. Each of the surveys was divided into two versions so that each respondent was required to complete six choice tasks.

Figure 1: Example Choice Tasks

Whole GBR					
	Management	Amount of GBR in good condition	Will it happen?	Cost	Your choice
	 Option for particular focus	 Current condition: 90% in good condition (311,000 sq km) Condition in 25 years time	 Level of certainty	 How much you pay each year (5 years)	 Select one option only
Option A	Current trends	65% in good condition (225,000 sq km)	80%	\$0	<input type="checkbox"/>
Option B	Improve water quality	68% (235,000 sq km) = 3% improvement	60%	\$100	<input type="checkbox"/>
Option C	Increase conservation zones	66% (228,000 sq km) = 1% improvement	75%	\$50	<input type="checkbox"/>
Option D	Reduce greenhouse gases*	85% (294,000 sq km) = 20% improvement	40%	\$100	<input type="checkbox"/>

Regional section of the GBR					
Central section: Tully to Airlie Beach (Whitsundays)					
	Management	Amount of GBR in good condition	Will it happen?	Cost	Your choice
	 Option for particular focus	 Current condition: 90% in good condition (77,750 sq km) Condition in 25 years time	 Level of certainty	 How much you pay each year (5 years)	 Select one option only
Option A	Current trends	65% in good condition (56,250 sq km)	80%	\$0	<input type="checkbox"/>
Option B	Improve water quality	76% (65,750 sq km) = 11% improvement	50%	\$50	<input type="checkbox"/>
Option C	Increase conservation zones	68% (58,750 sq km) = 3% improvement	85%	\$50	<input type="checkbox"/>
Option D	Reduce greenhouse gases*	80% (66,000 sq km) = 15% improvement	10%	\$500	<input type="checkbox"/>

3.1 Survey Collection Details

Two collection methods were used to collect responses for identical surveys from Brisbane residents. The first was a drop-off/pick-up, paper-based collection technique¹ and the second was an internet survey with respondents from a pre-recruited internet panel. The paper-based collection was conducted in a three month period from June to September 2009, with a high response rate of 91% recorded². A private organisation was contracted to host the internet survey and provide access to an internet panel over a two-week period in August 2009. Two segmentation criteria were implemented to ensure a 50:50 split between males and females and between respondents aged between 18-45 years and 46-88 years. However, a large number of surveys were collected and these segmentation requirements were not specific to the subset of surveys reported in this study. An accurate response rate for the internet survey was not obtained. Emails were sent to nearly 11,000 panellists and it is unclear what proportion responded before the required sample size was achieved and the survey closed. The use of segmentation quotas further confounded the issue.

¹ The authors have experienced very low responses rates (<20%) with the use of mail surveys even with pre-emptive and follow-up reminders.

² This appears high and was repeatedly checked with the collectors. Collection was organised through the Lions Club. The coordinator was an experienced survey collector and had been involved in the collection of the National Census in 2006. However, there may have been some non-random selection of Brisbane suburbs and some avoidance of certain areas.

The whole GBR and regional GBR versions of the paper-based survey yielded 93 and 86 responses respectively. As each respondent completed six choice tasks, there were 558 and 516 observations for each survey. In the internet survey, 80 responses were collected for both the whole and regional surveys, yielding 480 observations for each survey.

IV. RESULTS

The aim of this paper is to examine how the two different survey collection modes may impact on the results of a choice modelling valuation experiment. The potential differences are explored in four key ways. The first is to identify if the two collection modes captured equivalent samples of the population. The second is to see if respondents' attitudes and usage of the GBR were the same and if there was any variation in the influence these attitudes may have had on choice selection. The third is to identify if the samples from the two collection modes generated similar valuation models and the fourth is to identify if the final results or value estimates are the same.

4.1 Comparing Socio-Demographic Characteristics of Sample Respondents

The results of this study indicate little difference between collection modes in generating an accurate random sample of the population. There was no significant difference³ in the education and income levels of respondents in the two collection samples, or whether or not they had children (*Table 1*). As expected, there was a difference in the ages of the two groups with the online group having a higher proportion of younger respondents and fewer older people. There was also a lower proportion of females in the online sample.

Table 1: Respondent Characteristics for Survey Modes

		Internet	Paper	ABS 2006 census
Gender*	Female	48%	58%	50%
Children	Have children	67%	73%	n/a
Age***	18-29	23%	12%	24%
	30-45	32%	32%	31%
	46-65	34%	33%	30%
	66-89	11%	23%	16%
Education	post school qualification	59%	57%	56%
	tertiary degree	33%	31%	24%
Income ¹	less than \$499 per week	11%	14%	17%
	\$500 – \$799 per week	16%	21%	18%
	\$800 – \$1199 per week	27%	20%	21%
	\$1200 – \$1999 per week	29%	28%	24%
	\$2000 or more per week	16%	17%	21%

³ All reference to tests of significance relate to Pearson's chi squared crosstabs at 5%, unless otherwise stated.

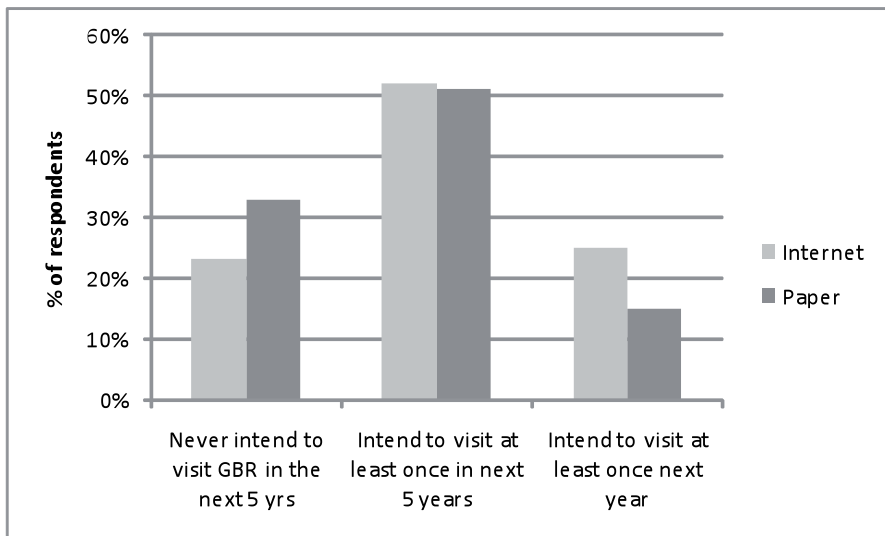
*** significant difference between collection modes at 1% level of confidence; * at 10% level of confidence
¹ 21% of respondents in the paper-based survey did not complete the question

The socio-demographic characteristics of both groups can also be compared with the census data for the Brisbane population (*Table 1*). It appears that the gender composition of the internet sample was better aligned with the population. There were some age biases in both samples. The oldest age category was under-represented in the online sample and over-represented in paper sample. In the youngest age category, the internet sample was well aligned with the population, while the paper sample was underrepresented. The education levels were the same across samples and both were well matched with the population in terms of post school qualifications. Both samples had higher levels of tertiary education compared with the population. Income levels were similar across samples, with both samples underrepresented in the lowest income category (more so in the internet sample) and highest categories compared with the population. However, a large proportion (21%) of the paper sample did not report their income.

4.2 Comparing Attitudes and GBR Usage

Information was gathered about both past use of the GBR and expected future use. Due to the high number of missing values in the information about past use in the paper survey, only information about expected future use is presented here. However, because the valuation scenario was framed in terms of changes in the future condition of the GBR, expected future use was an appropriate indicator. There was no significant difference in the five category responses about the frequency of expected future use across collection modes, although more

Figure 2: Expected Future Use of the GBR in the Next Five Years

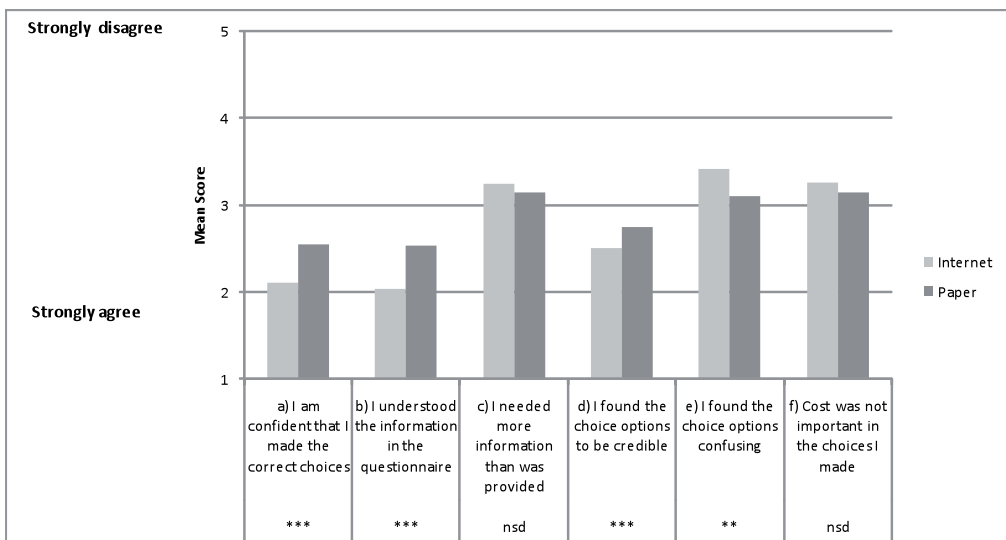


paper respondents thought they would never visit the GBR in the next five years and more internet respondents thought they would visit the GBR at least once in the next year (*Figure 2*).

There was also no significant difference in respondents’ attitudes about how the condition of the GBR had changed over the past 10 years. In both samples, only 1% of respondents thought it had improved, while 66% and 59% in the internet and paper samples respectively thought it had declined. The rest of respondents were unsure.

A series of six follow-up questions were included after the choice tasks to determine how respondents reacted to the choice tasks. Respondents were asked to score the extent to which they agreed with each statement on a scale from 1 = strongly agree to 5 = strongly disagree. The mean scores and significant sample differences are presented in *Figure 3*. The results confirm that the internet respondents had fewer problems with the survey, most likely because of their experience in survey completion. Internet respondents were more agreed that (a) they were confident they had made the right choices; (b) that they understood the information in the questionnaire; (d) that they found the choice options credible and (e) were more in disagreement that they found the choice options confusing.

Figure 3: Respondent’s Attitudes to the Choice Tasks¹



¹ Mean scores on a rating from 1= strongly agree to 5 = strongly disagree
 *** significant difference between collection modes at 1% level; ** at 5%; nsd = no significant difference

All the attitudinal questions were included in the choice models outlined in the next section to determine their potential influence on choice selection.

4.3 Comparing the Performance of the Choice Models

The information from choice models is usually analysed using logistical regression models. The probability that a respondent would choose a particular option can be related to the levels of the different attributes in all the choice tasks, the socio-demographic characteristics of

the respondents, and other factors such as attitudinal information collected in the surveys. A constant term (ASC) is included in the models to account for any unexplained effects (ie. not accounted for in the specified model variables) (Bennet and Blamey 2001).

In this case, random parameter logit (RPL) models were developed which meant a random draw could be taken (from a specified distribution) for each variable that was selected as a random parameter. It was expected that most heterogeneity across respondents' choices would be in their preferences for the different management options, and the three ASCs associated with the management options were randomised with a specified normal distribution. Choice responses were modelled as panel data to account for the fact each respondents made six choices. The socio-demographic variables and other use and attitudinal variables were all modelled to explain the choice of the status quo option.

Details of the model variables are presented in *Table 2*. Separate models were developed for the whole and regional levels of surveys scope for both the internet and paper samples. To provide a more robust sample size for comparative purposes, the responses of the whole and regionally scoped surveys were then pooled for each collection sample, and an additional dummy variable was included to account for the scope of the two surveys. The results are presented in *Table 3*.

All six models are significant (high chi-squared values) and the two main attributes, COST and GBR CONDITION, are highly significant and signed as expected. Higher levels of COST are not preferred, and higher levels for GBR CONDITION are preferred. The standard deviations for the randomised ASC distributions are all highly significant which signifies the heterogeneity in preferences for the three management options. However the significance of the ASCs does vary across the models indicating there are differences in the unexplained influences in the selection of the different management options which vary across both collection modes and survey scope. In the pooled models, the ASCs are consistently significant, indicating there are significant unexplained influences in both collection modes.

In the first two comparisons (whole and regional GBR models), there is little consistency in the significance of the other variables across collection samples, although more consistency does occur in the pooled models. For example, AGE is significant in the whole internet model (1) (the negative sign indicating older people are less likely to select the status quo option/ more likely to select an improvement option) but not in the whole paper model (2). The situation is reversed in the regional model comparisons (3-4), and not significant in either sample in the pooled models (5-6).

Most of the attitudinal responses were not a significant influence on choice selection so that the underlying differences in the respondents' attitudes of the two groups did not significantly affect their choices. While some inconsistency appeared in the whole and regional GBR comparisons, more consistency exists in the pooled models. In both samples, respondents seem to have been cost conscious and the influence of cost question (COST_NOTIMPORTANT) is significant and positive indicating that those who disagreed (thought cost was important) were more likely to select the status quo option.

There are more underlying similarities between the two collection samples in the pooled models neutralising some of inconsistencies of the small samples sizes in the separate models. However, some differences in collection mode still remain as confirmed by a log likelihood ratio test of the two pooled models.

The other comparison of interest in the larger samples of the pooled models is the extent to which respondents always selected the status quo option. While many may have selected this

Table 2: Model Variables and Descriptions

Main attributes	Description
COST	Annual payment for a 5 year period
GBR_CONDITION	Area of GBR in good health (Sq kms)
CERTAINTY	Will it happen? Level of certainty
ASC	Alternative specific constant
SQ...	Prefix to denote status quo (current situation) alternative
WQ...	Prefix to denote management option: Improve water quality
CZ...	Prefix to denote management option: Increase conservation zones
GG...	Prefix to denote management option: Reduce greenhouse gases
AGE	Categories (1-4) – see <i>Table 1</i> for details
GENDER	Male = 0; Female = 1
CHILDREN	Children = 1; no children = 2
EDUCATION	Coded from 1= primary to 5 = tertiary degree or higher
INCOME	Categories (1-5) – see <i>Table 1</i> for details. The mid point of each category was used for analysis with an additional 25% added to the last category
FUTUSE	Categories (1-5) from 1= no plan to visit the GBR in the future to 5= plan to visit more than one in the next year
GBR_COND_CHANGE	Change in condition of the GBR in the last 10 years effects coded: -1 = declined; 0 = unsure; 1= improved
CONFIDENT	I am confident that I made the correct choices (1 = strongly agree to 5= strongly disagree)
UNDERSTOOD	I understood the information in the questionnaire (1 = strongly agree to 5= strongly disagree)
MOREINFO	I needed more information than was provided (1 = strongly agree to 5= strongly disagree)
CREDIBLE	I found the choice options to be credible (1 = strongly agree to 5= strongly disagree)
CONFUSED	I found the choice options confusing (1 = strongly agree to 5= strongly disagree)
COST_NOTIMPORT	Cost was not important in the choices I made (1 = strongly agree to 5= strongly disagree)
SURVEY SCOPE	Survey scale: Whole = 0; Regional = 1

Table 3: RPL Models for Different Sample Responses to GBR Scope Surveys

	Whole GBR		Regional GBR		Whole + Reg GBR	
	Internet	Paper	Internet	Paper	Internet	Paper
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Random parameters in utility functions</i>						
ASC_WQ	-14.262***	-0.904	-0.546	-9.510***	-7.631***	-4.431***
ASC_CZ	-14.006***	-0.487	-0.501	-10.468***	-8.378***	-5.607***
ASC_GG	-15.743***	-3.550	-1.150	-10.022***	-6.668***	-4.325***
<i>Derived standard deviations of parameter distributions</i>						
ASC_WQ	2.297***	1.568***	1.464***	1.438***	2.165***	1.713***
ASC_CZ	2.638***	1.854***	1.775***	2.231***	2.176***	2.088***
ASC_GG	3.471***	2.363***	2.678***	2.602***	3.086***	2.734***
<i>Non Random parameters in utility functions</i>						
COST	-0.005***	-0.006***	-0.007***	-0.006***	-0.007***	-0.007***
GBR CONDITION	0.004***	0.005***	0.016***	0.013***	0.002***	0.002***
CERTAINTY	0.011	0.012	0.029**	0.044***	0.040***	0.047***
AGE	-0.700**	0.009	0.206	-0.590**	-0.089	-0.202
GENDER	-2.477***	-0.846*	-1.382***	-1.736***	-1.552***	-0.869***
CHILDREN	-2.229***	-0.178	-0.557	-1.060**	-1.046***	-0.265
EDUCATION	-0.254	-0.393*	-0.066	0.256	-0.230	-0.266*
INCOME	-2.5E-05**	6.0E-06	-1.7E-05***	-1.6E-05**	-2.3E-05***	-1.0E-05**
FUTUSE	-0.054	-0.553**	0.188	-0.283	0.140	-0.319**
GBR_CONDIT_CHANGE	2.847***	1.532**	0.192	1.807***	1.180***	1.704***
CONFIDENT	-0.330	-0.074	-0.384	-0.569**	-0.282	-0.251
UNDERSTOOD	-0.464	0.635***	0.055	-0.059	-0.242	0.237
MOREINFO	0.295	-0.161	-0.272	0.123	-0.059	0.076
CREDIBLE	0.047	-0.177	0.311	-0.512**	0.274	-0.282*
CONFUSED	-0.778**	0.300	0.075	-1.191***	-0.271*	-0.448***
COST_NOTIMPORTANT SURVEY SCOPE	0.915***	0.554***	0.755***	0.648***	0.704***	0.399***
					-1.882***	-0.348
<i>Model statistics</i>						
No of Observations	480	558	480	516	960	1074
Log Likelihood	-437	552	-492	-481	-961	-1080
AIC	1.92	2.06	2.14	1.95	2.05	2.05
Chi Sqrd	457	442	346	468	740	818
McFadden Pseudo R-sqrd	0.34	0.29	0.26	0.33	0.28	0.28

*** significant at 1% level of confidence; ** significant at 5%; * significant at 10%;

option because they could not afford to pay for one of the improvement alternatives, others may have used this as some form of protest vote. However, there is no significant difference between samples with 18% always selecting this option in the internet sample and 12% in the paper sample.

4.4 Comparing WTP Estimates

The WTP estimates are generated from the models by dividing the attribute coefficient by the cost coefficient (Table 3) and the models are designed to provide estimates of average annual (for a five-year period) household WTP for a 100 sq km improvement in the condition of the GBR in 25 years time. These estimates are presented in *Table 4* and demonstrate the similarity across collection samples and survey scope. In the pooled models (5-6), the WTP for a 100 sq km improvement is exactly the same (\$0.34 per household per year for five years) for both the internet and paper samples.

Table 4: Household WTP Estimates for a 100 sq km Improvement in GBR Condition

Survey mode	Whole GBR	Regional GBR	Whole + Reg GBR
Internet (mean WTP)	\$0.87***	\$2.27***	\$0.34***
Confidence intervals (CIs)	\$0.31 – \$2.07	\$0.95 – \$4.37	\$0.07 – \$0.64
Paper (mean WTP)	\$0.88***	\$2.12***	\$0.34***
Confidence intervals (CIs)	\$0.44 – \$1.62	\$0.64 – \$4.61	\$0.11 – \$0.63

*** significant at 1% level of confidence

V. DISCUSSION AND CONCLUSIONS

The key findings of this study are that internet panels do not exactly replicate sample selection or the statistical modelling of the drop-off/ collect surveys, but that there is no difference in the subsequent value estimates. However, some of the inconsistencies found in model performance are less evident in the pooled models with larger sample sizes.

In this study, the main socio-demographic difference between the two samples was in age, with some differences also appearing in gender. The use of segmentation quotas in the internet sample meant the internet sample was more closely aligned with the population than the paper-based sample. Some alignment differences occurred in the paper survey even though there was a very high response rate. These results suggest the issue of sampling bias should no longer be a concern in internet surveys when pre-recruited internet panels are used and specific segmentation criteria are applied. Although socio-demographic variables had an inconsistent affect on choice selection in the separate models (1-4), there was little difference in the pooled models.

There were some differences between the two samples in respondents’ intended use of the GBR (which might be expected to influence their preferences for improvement in GBR condition); in attitudes of respondents to the complexity of the survey as well as in attitudes

to changes in the condition of the GBR. However, these differences did not have a consistent influence on respondents' preferences and choice selection in the separate models. Again, some inconsistencies were resolved in the pooled model. In addition, there was no evidence of more protest responses in the internet sample, even though they had more inducement to complete the survey compared with the paper-based respondents.

Even though the internet respondents were more survey savvy, this did not ultimately impact on the underlying focus of the survey to elicit respondents WTP. Values remained equivalent across collection modes in all three comparisons. This confirms the parity found in two previous studies comparing WTP estimates for internet and postal surveys (Nielsen 2011), but contradicts the findings of the other Australian comparison (Hatton MacDonald *et al.* 2010).

It has often been argued by policy makers that the process of acquiring information about nonmarket value estimates for environmental goods takes too long and is too costly. The use of internet panels really negates the traction of this argument. The number of different panels and the size of panels are growing rapidly, extending coverage to a larger section of the population. Given the substantial cost and logistical efficiency associated with their use, this is a collection mode that is going to increase in popularity. The results of this study indicate that the WTP estimates from a nonmarket valuation survey are not compromised by such choice of collection methodology.

ACKNOWLEDGEMENT

This research has been supported through the Environmental Economics Research Hub, funded by the Australian Department of Environment and Water Heritage and the Arts under the Commonwealth Environment Research Facility.

REFERENCES

- Alberini, A. and R. Kahn (2006). *Handbook on contingent valuation*. Cheltenham: Edward Elgar.
- Bateman, I.J., R.T. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, M. Jones-Lee, G. Loomes, S. Mourato, E. Ozdemiroglu, D.W. Pearce, R. Sugden, and J. Swanson (2002). *Environmental valuation with stated preference techniques*. Cheltenham: Edward Elgar.
- Bennett, J.W. and R.K. Blamey (2001). *The choice modelling approach to environmental valuation*, Cheltenham: Edward Elgar.
- Berrens, R., A. Bohara, H. Jenkins-Smith, C. Silva, and D. Weimer (2003). The advent of internet surveys for political research: A comparison of telephone and internet samples, *Political Analysis*, 11: 1-22.
- Boxall, P., W.L. Adamowicz, and A. Moon (2009). Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement, *Australian Journal of Agricultural and Resource Economics*, 53: 503-519.
- Canavari, M., G. Nocella, and R. Scarpa (2005). Stated willingness-to-pay for organic fruit and pesticide ban, *Journal of Food Products Marketing*, 11: 107-134.
- Champ, P. (2003). Collecting survey data for non-market valuation, in: P. Champ, K.J. Boyle, and T.C. Brown (eds.), *A primer on nonmarket valuation*. Dordrecht: Kluwer Academic Publishers.
- Fleming, C. and M. Bowden (2009). Web-based surveys as an alternative to traditional mail methods, *Journal of Environmental Management*, 90: 284-292.
- Hatton MacDonald, D., M. Morrison, J. Rose, and K. Boyle (2010). Untangling differences in values from internet and mail stated preference studies, Paper presented at the Fourth World Congress of Environmental and Resource Economists, Montreal.

- Louviere, J., D. Hensher, and J. Swait (2000). *Stated choice models – analysis and application*, Cambridge: Cambridge University Press.
- Marta-Pedroso, C., H. Freitas, and T. Domingos (2007). Testing the survey mode effect on contingent valuation data quality: A case study of web based versus in-person interviews, *Ecological Economics*. 62:388-398.
- Maguire, K. (2009). Does mode matter? A comparison of telephone, mail, and in-person treatments in contingent valuation surveys, *Journal of Environmental Management*. 90:3528-3533.
- Mitchell, R.C. and R.T. Carson (1989). *Using surveys to value public goods: the contingent valuation method*, Washington DC: Resources for the Future.
- Nielsen, J.S. (2011). Use of the Internet for willingness-to-pay surveys: A comparison of face-to-face and web-based interviews, *Resource and Energy Economics*. 33:119-129.
- Olsen, S.B. (2009). Choosing between internet and mail survey modes for choice experiments surveys considering non-market goods, *Environmental and Resource Economics*. 44: 591-610.
- Rolfe, J. (2006). Theoretical issues in using choice modelling data for benefit transfer, in: J. Rolfe and J. Bennett (eds.): *Choice modelling and the transfer of environmental values*. Cheltenham: Edward Elgar.
- Shih, T-H and X. Fan (2009). Comparing response rates in e-mail and paper surveys: A meta-analysis, *Educational Research Review*. 4:26-40.
- Von Haefen, R.H., D.M. Massey and W.I. Adamowicz (2005). Serial nonparticipation in repeated discrete choice models, *American Journal of Agricultural Economics*. 87:1061-1076.

