ISSN 1835-9728

Environmental Economics Research Hub

Research Reports

Testing construct validity of verbal versus numerical measures of preference uncertainty in contingent valuation

Sonia Akter and Jeff Bennett

Research Report No.46

January 2010

About the authors

Sonia Akter[1] is a PhD student in the Crawford School of Economics and Government at the Australian National University.

Jeff Bennett is Professor in the Crawford School of Economics and Government at the Australian National University.

[1]Crawford School of Economics and Government, The Australian National University, Canberra, ACT 0200, Australia, Tel: +61 2 6125 6556, Fax: +61 2 6125 5570

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

The numerical certainty scale (NCS) and polychotomous choice (PC) methods are two widely used techniques for measuring preference uncertainty in contingent valuation (CV) studies. The NCS follows a numerical scale and the PC is based on a verbal scale. This paper presents results of two experiments that use these preference uncertainty measurement techniques. The first experiment was designed to compare and contrast the uncertainty scores obtained from the NCS and the PC method. The second experiment was conducted to test a preference uncertainty measurement scale which combines verbal expressions with numerical and graphical interpretations: a composite certainty scale (CCS). The construct validity of the certainty scores obtained from these three techniques was tested by estimating three separate ordered probit regression models. The results of the study can be summarized in three key findings. First, the PC method generates a higher proportion of 'Yes' responses than the conventional dichotomous choice elicitation format. Second, the CCS method generates a significantly higher proportion of certain responses than the NCS and the PC methods. Finally, the NCS method performs poorly in terms of construct validity. We conclude that, overall, the verbal measures perform better than the numerical measure. Furthermore, the CCS method is promising in measuring preference uncertainty in CV studies. However, further empirical applications are required to develop a better understanding of its strengths and the weaknesses.

Keywords: Preference uncertainty, contingent valuation, numerical certainty scale, polychotomous choice method, composite certainty scale, climate change, Australia

JEL CODE: Q51, Q54,

1. Introduction

The NOAA Blue Ribbon Panel on contingent valuation (CV) advocated the use of response formats that allow for expressions of uncertainty in respondents' preferences (NOAA, 1993). In general terms, 'preference' denotes an individual's taste for a specific good and 'uncertainty' denotes a lack of knowledge about the probability of an outcome (Knight, 1921). 'Preference Uncertainty' in the context of market transactions, therefore, refers to an individual's lack of knowledge about their choices of purchases given the price and other attributes of the good in question. Hanemann et al. (1995) first proposed a welfare model that incorporated an element of uncertainty about individual preferences in the context of stated preference (SP) studies. They argued that individuals do not necessarily know their true valuations (a) of a good with certainty. Rather they perceive the value of the good to lie within an interval {a-h, a+h}, where 'h' refers to the amount of unknown component in preference (h>0).

Subsequent to the NOAA Panel report and the Hanemann et al. (1995) study, researchers have developed and applied a variety of methods to address preference uncertainty in CV studies. The numerical certainty scale (NCS) method and the polychotomous choice (PC) method are two widely used techniques to measure the extent of preference uncertainty. Li and Mattson (1995) first devised the NCS method in which the standard 'Yes/No' dichotomous choice (DC) valuation question is followed by a numerical certainty scale ranging from 1 to 10. Ready et al. (1995) introduced and applied the PC format in which respondents express their uncertainty by choosing from a set of responses: 'Definitely Yes', 'Probably Yes', 'Maybe Yes', 'Maybe No', 'Probably No', 'Definitely No'.

Empirical findings in the psychology literature suggest that there can be important consequences of measuring psychological uncertainty with verbal versus numerical measures. In general, numerical measures are considered less appropriate for measuring psychological uncertainty especially in situations that involve probabilistic information (Gigerenzer, 2003). A large body of research dealing with peoples' statistical competency demonstrates that information about probabilistic events tends to be difficult to understand, even for highly educated people (Burkell, 2004). Furthermore, psychologists argue that most people in every day life use words rather than numbers when describing their own uncertainty (Windschitl and Wells, 1996; Renooij and Witteman, 1999). For example, a question like “Are you going to the gym today?” will commonly generate answers like ‘Definitely’, ‘Probably’, ‘Maybe Not’.

However, numerical measures have some advantages over verbal measures. Numbers are precise, allow calculations and have a fixed rank-order (Renooij and Witteman, 1999). Verbal measures are imprecise, cannot be used in calculations and more subjectively interpretable (Weber and Hilton, 1990). For decision making situations that require precise information of psychological uncertainty, verbal measures are less informative and useful than numerical measures (Budescu and Wallesten, 1990). Yet, the numerical measure adds complexity as it requires the respondents to think about their uncertainty in a deliberate, controlled and rule-based manner (Windschitl and Wells, 1996). Numerical measure, therefore, may lead to incorrect assessments of psychological uncertainty in situations where people’s decisions and behaviours are products of intuitive and associative processing. In a series of experiments, Windschitl and Wells (1996) showed that verbal measure is more advantageous than numerical measure to assess psychological uncertainty that people experience in their everyday life.

Debate persists in the CV literature about whether the NCS or the PC method provides the better uncertainty measure (Akter et al., 2008). However, to the best of our knowledge, no study has been conducted to date to compare the performance of these two methods. The choice of a preference uncertainty measurement technique in CV studies is subjective and ad-hoc. Some empirical studies have compared the treatment effects of certainty calibrated WTP against the conventional DC CV WTP estimates in terms of accuracy and efficiency (see Samnaliev et al., 2006; Chang et al., 2007). These studies provide no conclusion regarding the superiority of one technique over the other. In this paper, we applied the NCS and the PC methods and one other that we develop: the composite certainty scale (CCS), using split sample treatments. The self-reported certainty scores obtained from these sample splits were compared and contrasted. The scores, furthermore, were modelled using ordered probit regression model to test if their variations could be explained by theoretically and intuitively expected explanatory variables.

The rest of the paper is organized as follows. Section 2 presents a discussion of the NCS, PC and CCS method followed by a description of the survey and case study in Section 3. In Section 4, the distribution of the self-reported uncertainty scores are compared across the different certainty measurement techniques. Section 5 presents the statistical model and regression results followed by the results of a calibration exercise in Section 6. Section 7 presents discussion of the results and concludes.

2. The NCS, PC and CCS method

Proponents of the NCS method argue that it provides more precise information about the level of certainty as the respondent is able to specify a numerical certainty value in a 1 to 10 point scale. However, the NCS method is based on two assumptions (Loomis and Ekstrand, 1998). First, it assumes that respondents are able to assess accurately their own degree of certainty when answering the willingness to pay (WTP) question. Second, it is assumed that all respondents interpret the certainty scale equivalently. The main argument for measuring preference uncertainty in CV studies is that respondents are uncertain about their valuation of the good. Hence, the first assumption, by implying that respondents are certain about their levels of confidence in their WTP choice, appears to be contradictory (Akter et al., 2008). Furthermore, the second assumption of comparable rating responses across individuals is dubious as it has been observed that respondents show 'scale preference' in which some individuals tend to be low raters or high raters (Roe et al., 1996).

The performance of the PC scale has also been debated. The incentive compatibility property of a CV study is considered to be diminished when this format is used to elicit preference uncertainty. Carson and Groves (2007) argue that questions that contain more than two alternatives are not incentive compatible as they provide opportunities for respondents to respond strategically. Accordingly, Ready et al. (1995) found that the PC format generates higher rates of 'Yes' responses than the standard DC question because it allows the respondent to give an affirmative response, without making a strong commitment. Alberini et al. (2003) argued that PC format may cause false uncertainty to arise in the SP framework because it provides respondents with an inducement to leave unresolved their lack of confidence in answering the valuation question. One other pitfall of the PC format is the subjectivity of the words that are generally used to elicit respondent uncertainty (Hanley et al., 2009). For example, when a respondent is given the choice between ‘Probably Yes’, and ‘Maybe Yes’, unless the distinction between the terms ‘Probably’ and ‘Maybe’ are explicitly demonstrated, the interpretation of these two responses could be highly subjective and may lead to a potential measurement bias.

The CCS method was designed to overcome the shortcomings of the PC method. The method is based on the word-graphic rating scale, first developed by Tesler et al. (1991). The method uses verbal expression of uncertainty which is similar to the PC method but the scale is introduced followed by a DC WTP question. Thus, it allows the incentive compatibility property of the CV study to be maintained. Furthermore, the method includes two steps. In the first step, respondents are asked whether they are certain about their answers to the WTP question. Only respondents who were not fully certain about their decisions are shown the CCS and asked to indicate their levels of certainty. This decomposition exercise is based on the domain-range questionnaire structure suggested by Beatty et al. (1999). They suggest decomposing questions into major components so that the first component serves as the domain (which tells the respondent what the question is about) and the second component serves as the range (which tells the respondent what they are expected to give back).

In the CCS method, the domain is whether or not respondents are uncertain about their decisions. Having identified the presence of any level of uncertainty, the range is a list of different certainty levels. Five categories of responses are added to the scale, namely ‘Extremely Unsure’, ‘Highly Unsure’, ‘Fairly Unsure’, ‘Highly Sure’, and ‘Extremely Sure’. To overcome any subjective interpretation of the verbal scale (the second pitfall of the PC method), each of these response categories is associated with a numerical interpretation and a graphical expression (see Figure 1)[1]. A pie diagram was added to help respondents visualise the information. The literature in health risk communication indicates that pictures accompanied by clear text help communication through higher attention, comprehension, recall, and adherence by respondents (Houts et al., 2006). The psychology literature suggests that graphical representation of data may improve judgment and decision-making because it can facilitate information processing and analogue reasoning by providing a holistic view of the information (Stone et al., 1997; Lipkus and Hollands, 1999; Larkin and Simon, 1987). In addition, it can minimise mental effort and be used either as an alternative to numbers or in addition to them in order to aid further understanding of numerical risks.

INSERT FIGURE 1 HERE

3. Survey and data collection

3.1. Set-up of the survey

The context of these experiments was Australian households' preferences towards the occurrence and mitigation of anthropocentric climate change. As part of fulfillment of its Kyoto Protocol obligations, the Australian Government has recently proposed a national emissions trading scheme known as the Carbon Pollution Reduction Scheme (CPRS). The aims of the CPRS are to reduce emissions by 60 per cent of the 2000 level by 2050 and to encourage the development and use of reduced emission technologies (Department of Climate Change, 2008). The

[1] One limitation of attaching numerical percentage points to the five point verbal scale was that the whole continuum between 0 and 100 could not be included in the scale.

implementation of the CPRS will affect Australian households as the prices of a wide range of emission-intensive goods and services are expected to rise. The case study aimed to explore Australian households' willingness to bear extra expenses to support the CPRS.

A web-based CV survey was conducted with about 300 respondents in each of three split samples in Sydney from the third week of November 2008 until the first week of December 2008. Respondents were asked if their household would be willing to bear specified extra expenses per month resulting from the CPRS. Eight different extra expenses "bids" ranging from AUD 20 to AUD 400 per month per household were randomly assigned across the respondents. These amounts were based on responses obtained from an open-ended WTP question asked during the first focus group. The bid amounts were tested in the second focus groups and the pilot survey.

In the first split sample (S_{NCS}), respondents were shown a ten point numerical certainty scale followed by a DC WTP question. Respondents in the second split sample (S_{PC}) answered the PC question that included six response categories ('Definitely Yes/No', 'Probably Yes/No' and "Maybe Yes/No'). In the third split sample (S_{CCS}), respondents answered the CCS questions. The development process of the CCS method involved peer discussions and two focus group sessions. Five university staffs and graduate students were invited to participate in the testing of the first version of the CCS method. Based on their feedback, the scale was revised and tested in a focus group discussion session with 12 members of the general public. Participants were asked to provide feedback on the level of comprehensibility and appropriateness of the word choices of the CCS. The design of the scale was then revised further and tested again in a second focus group. Before the final survey, a small-scale pilot test was conducted where twenty respondents participated.

In addition to the valuation question and question(s) related to preference uncertainty, the survey questionnaire included a number of questions regarding respondents' perceptions of the extent of climate change and effectiveness of climate change policy. Respondents were asked to indicate their best guess of temperature change in 2100 relative to the current year. They were subsequently asked to indicate a range around their best guess of average change in temperature in the form of high and low guesses. A numerical probability scale was used to elicit respondents' perceptions of their 'best guess' of policy effectiveness followed by their 'high guess' and 'low guess' of the policy being effective in slowing down climate change. The high guess and low guess of expected temperature change and climate policy effectiveness provide a measure of respondents' perceptions of uncertainty of the future scenario and policy in question.

These two factors (uncertainty in scenario and policy) are expected to be of particular interest as psychology theory suggests that environmental factors are important source of psychological uncertainty (Downey et al., 1975). The term 'environment' does not refer natural environment. Instead it refers to the basic conditioning of the decision making framework within which the individual decision maker operates from. More specifically, environmental uncertainty occurs whenever the outcome of an event in future and/or the probability of an outcome to occur cannot be precisely determined (Downey et al., 1975; Duncan, 1972). In a situation like this, decision maker try to rate the outcome and/or probability of the external events as a prominent coping strategy (Kahneman and Tversky, 1982). According to Downey et al. (1975), these external factors or decision making environment provides direct input into individual's cognitive mapping process and thus determines the level of psychological uncertainty.

Respondents were asked a set of socio-demographic (e.g. sex, age, education, occupation, income) and attitudinal questions (level of concern regarding climate change, relative importance of a climate change mitigation policy, belief if climate change is caused by human action). Finally, a set of questions that measured respondent's knowledge and level of familiarity with the scenario and policy context (respondents' awareness of the CPRS, Kyoto protocol, Intergovernmental Panel of Climate Change (IPCC)) was included in the questionnaire.

2.2. General survey results

Table 1 compares the socio-economic characteristics of the sampled households with the regional and national population statistics. A chi-square test of proportions revealed that the differences in the three sample splits with respect to sex ratio, age, education and household income are not statistically significant at the ten percent level. When each of the split sample statistics were compared with the Sydney population and the Australian population, the differences in sex ratio are not found to be statistically significant. However, although the educational attainments of the sample were not found to be significantly different than the Sydney population, they were significantly different than the educational attainments of the Australian population. Finally, Z tests revealed that the sample respondents' age and weekly household income are not significantly different than the median age and weekly average income of the Sydney population and the national population. These test results demonstrate that the sample is representative of the Sydney population as well as the Australian population at least with respect to sex ratio, age and household income.

Although over eighty percent of the respondents had heard of the Kyoto Protocol and one third of them knew the Protocol's objectives, a majority (82 percent) of the respondents had not heard of the IPCC. While more than half of the respondents (57 percent) had heard of the CPRS prior to the survey, a majority (83 percent) did not know when the CPRS would be implemented. Around two thirds of those who claimed to know when the CPRS would be implemented (five percent of the total sample) could correctly indicate the proposed implementation year of the CPRS. Respondents' knowledge of the Kyoto Protocol and the CPRS were positively correlated ($r=0.221$, $p<0.001$) implying that respondents who were informed about the Kyoto Protocol were also aware of the CPRS. Likewise, a low but statistically significant positive correlation was observed between respondents' knowledge of the CPRS and carbon offset programs ($r=0.118$, $p<0.001$). This implies that respondents who were familiar with carbon offsets were also familiar with the CPRS.

Respondents' mean best guess about the change in average temperature in 2100 relative to the current year was 3.75 degrees centigrade. The median was three degrees centigrade with a maximum of 10.5 degrees and minimum of minus 4.5 degrees. The average range (the difference between high guess and low guess) around stated best guess temperature change was about three degrees centigrade. The range of expected temperature change varied within the range of 15.5 degrees and zero degrees centigrade. The mean of respondents' best guess of the CPRS being effective in slowing down climate change was 25 percent. The average range around this best guess was also 25 percent.

4. Results concerning preference uncertainty

In this section the self-reported certainty scores obtained from the three sample splits are compared and contrasted. First, we examine if the PC method generates a higher proportion of 'Yes' responses compared to the DC WTP elicitation format. Second, the distributions of the stated certainty scores across the three measurement techniques are determined. Finally, the distributions of certainty scores across 'Yes/No' WTP responses are displayed for all three sample splits.

Recoding 'Definitely Yes', 'Probably Yes' and 'Maybe Yes' responses to 'Yes', and 'Definitely No', 'Probably No' and 'Maybe No' responses to 'No', 54 percent of 'Yes' responses and 46 percent 'No' responses were observed in the S_{PC}. In Table 2, we compared this distribution of 'Yes/No' responses with the DC WTP responses obtained from the S_{NCS} and the S_{CCS}. In both the S_{NCS} and S_{CCS}, about two thirds of the respondents replied 'No' to the WTP question and the rest said 'Yes'. The proportions of 'Yes' responses in the S_{NCS} and the S_{CCS} were about 63 percent lower than that of the S_{PC}. No statistically significant difference was observed between the 'Yes/No' WTP responses across the S_{NCS} and the S_{CCS} (Chi square=0.155, $p<0.7$). The difference in distribution of 'Yes/No' responses in the S_{PC} from the other two samples was found to be statistically significant (S_{PC} & S_{NCS}: chi square=7.053, $p<0.01$; S_{PC} & S_{CCS}: chi square=9.692, $p<0.01$). These results imply that the PC method, as previously claimed by Ready et al. (1995)

and Alberini et al. (2003), generates a higher proportion of ‘Yes’ responses than the conventional DC WTP question.

INSERT TABLE 2 HERE

Figure 2 presents the distribution of the highest and the lowest certainty scores across the measurement methods. A third of the respondents in the S_{PC} indicated the highest level of certainty (Definitely Yes/No) about their preferences of paying (or not paying) for the CPRS while over a third of the respondents indicated the lowest level of certainty (Maybe Yes/No). Likewise, about a third of respondents in the S_{NCS} indicated the highest level of certainty (certainty score of 10) about their ‘Yes/No’ responses to the DC WTP question whereas less than three percent of the respondents indicated the lowest level of certainty about their decisions (certainty score of 1). In the S_{CCS}, an overwhelming majority of 87 percent respondents indicated the highest level of certainty (that they were completely certain) about their voting decisions while only less than a percent of the respondents indicated that they were extremely unsure[2]. The differences in the proportion of the highest and the lowest level of certainty score between the S_{CCS} and other two sample splits were statistically significant (S_{CCS} and S_{NCS}: Chi square=177.59, $p<0.001$; S_{CCS} and S_{PC}: Chi square=182.52, $p<0.001$). These results imply that the S_{CCS} method generates the highest level of certain responses compared to the two other methods.

INSERT FIGURE 2 HERE

[2] Note that none of the respondents selected the ‘Extremely Sure’ response and only one respondent selected the ‘Extremely Unsure’ response.

The distribution of the self-reported certainty scores in three sample splits were examined for differences across 'Yes/No' WTP responses. In the S_{PC} , only six percent of the 'Yes' respondents indicated the highest level of certainty ('Definitely Yes') whereas over a quarter of the total respondents reported the highest level of certainty about their decisions of not to pay ('Definitely No'). The difference in distribution of the highest certainty scores between 'Yes' and 'No' responses was statistically significant at the one percent level (chi square= 54.102, $p<0.001$). Over 40 percent of the S_{NCS} respondents, who said 'No' to the WTP question, were very certain (certainty score = 10) about their decisions as opposed to less than 20 percent of the 'Yes' respondents who were very certain. Hence, the respondents who replied 'No' to the WTP question stated significantly (Chi square=28.64, $p<0.001$) higher certainty scores than the respondents who replied 'Yes'. In S_{CCS}, 91 percent of all 'No' respondents said that they were absolutely sure about their decisions whereas 78 percent of all 'Yes' responses were absolutely sure about their decisions. This difference between certainty scores associated with 'Yes' and 'No' answers was statistically significant (Chi square=17.28, $p<0.001$). These results imply that, the 'No' responses tend to be held with greater certainty scores than 'Yes' responses regardless of the measurement method. Therefore, it could be argued that the level of self-reported certainty score is, in part, determined by the responses to the WTP question.

5. Construct validity of preference uncertainty scores

5.1. Statistical model of preference uncertainty

Construct validity refers to how well the estimated coefficient signs and values of the explanatory variables used to explain the variation in the dependent variable fit the theoretical expectation on which the model is based (Mitchell and Carson, 1989). Although no explicit theoretical model to explain variations in preference uncertainty scores has yet been developed, there is a general agreement about some explanatory hypotheses that have emerged after Loomis and Ekstrand (1998). Three other studies (Champ and Bishop, 2001; Samneliev et al., 2006; Akter et al., 2009) have estimated preference uncertainty models. The explanatory variables used in those models are more ‘intuitive’ than ‘theoretical’. We use the results of these studies and psychology theories as a foundation on which to test the construct validity of the certainty scores.

Loomis and Ekstrand (1998) estimated an ordinary least square regression model on pooled (both ‘Yes' and ‘No' responses) data. Champ and Bishop (2001) and Akter et al. (2009) estimated ordered probit regression models on certainty scores for ‘Yes’ responses whereas Samneliev et al. (2006) estimated separate logistic regression models for ‘Yes’ and ‘No’ responses. Loomis and Ekstrand (1998) used a follow-up DC certainty scale varying between 1 (=very uncertain) and 10 (=very certain) to elicit respondents’ levels of certainty regarding their responses to the DC WTP question. They found a quadratic relationship between the self-reported certainty levels and bid levels. This implies that respondents experience the highest level of uncertainty at the middle bid and relatively lower levels of uncertainty at the high and low bids. Furthermore, the relationships between certainty scores and respondents’ prior knowledge about the particular endangered species and their visiting the area proposed for protection, were positive and statistically significant.

Like Loomis and Ekstrand (1998), Champ and Bishop (2001) and Samnaliev et al. (2006) applied a follow-up DC certainty scale. Champ and Bishop (2001) estimated an ordered logit regression model of the ‘Yes’ certainty scores. Respondents’ perceptions of and attitudes towards the proposed program were found to be responsible for the observed variation in the self-reported ‘Yes’ certainty scores. Respondents in favour of the program and willing to pay the extra cost expressed higher certainty levels than other respondents. Samnaliev et al. (2006) found similar results to those of Champ and Bishop (2001). Respondents who objected to the imposed user fees in principle were more certain in rejecting the bid than others. This reflects respondents’ general attitudes to the hypothetical market, usually referred to as protest responses in CV.

Akter et al. (2009) used a five category PC question format (Extremely Unlikely, Fairly Unlikely, Not Sure, Fairly Likely, Extremely Likely) to ask respondents if they would pay the stated WTP value under a voluntary payment provision. The authors found a significant negative relationship between start bid and the stated likelihood of making a voluntary payment. The study also provides evidence that supports the relationship between respondent attitudes, perceptions and stated likelihood of paying, consistent with the findings of Champ and Bishop (2001) and Samnaliev et al. (2006). A respondent’s perceived individual responsibility for contributing to climate change, attitude towards paying to protect the environment and belief in the effectiveness of the proposed tree plantation program on climate change mitigation were found to be the main sources of stated uncertainty.

We used Loomis and Ekstrand (1998) s’ model as the basis of our statistical model. This model is the most relevant because it is the only study where a pooled regression model on both ‘Yes’ and ‘No’ certainty scores has been estimated. Loomis and Ekstrand (1998) suggest that the self-reported certainty score (C) is a quadratic function of bid level (BID and BIDSQ) and prior knowledge and familiarity with the good in question (KNOW). In addition to these variables, we hypothesize that an affirmative or negative response to the WTP question (A_WTP) influences the self-reported certainty score, i.e. respondents who say ‘Yes’ to the WTP question tend to state a lower certainty score and vice versa. Furthermore, as psychology theory predicts, uncertainty associated with the decision making context may contribute to psychological uncertainty. We identify two different forms of uncertainty in the current context, namely, scenario uncertainty (S_UNCERT) and policy uncertainty (P_UNCERT). S_UNCERT is related to respondents’ uncertainty regarding the extent of future climate change measured through respondents’ perception about the range of temperature rise in future. P_UNCERT refers to the perceived uncertainty associated with the proposed CPRS being effective in slowing down climate change. This variable was measured by asking respondents to indicate their high guess and low guess probability that the CPRS will be successful. The higher the perceived scenario and policy uncertainty, the lower would be the self-reported certainty score.

Finally, the certainty scores are expected to vary across respondents’ age (AGE). The psychology literature presents evidence of negative effects of aging on performance in cognitive tasks mainly due to slower information processing capacity (Salthouse, 1996; Hartley 2006). However, others argue that relatively older people bring knowledge and experience which may partially or completely offset any decrease in cognitive functioning that may have occurred with age (Park, 1998; Marsiske and Margrett, 2006). Therefore, the net effect of age on psychological uncertainty can be either positive or negative depending on the relative magnitude of the decline of cognitive processing effect and the wisdom effect due to higher level of knowledge and experience.

The statistical model to be tested takes the following form:

$C_i = \beta_1 BID_i + \beta_2 BIDSQ_i + \beta_3 KNOW_i + \beta_4\ A\_WTP + \beta_5\ S\_UNCERT + \beta_6\ P\_UNCERT + \beta_7$
$AGE + \varepsilon_i$

5.2. Results

Three ordered probit regression models were estimated using the certainty scores of both 'Yes' and 'No' responses. In the $S_{NCS,}$ the stated certainty scores within the range of 1 (absolutely uncertain) to 10 (absolutely certain) were used as the dependent variable. In the $S_{PC,}$ 'Definitely Yes/No' responses were recoded as three, 'Probably Yes/No' responses were recoded as two and 'Maybe Yes/No' responses were recoded as one. Finally, in the $S_{CCS,}$ certainty scores of the respondents who were absolutely certain about their decisions were recoded to six. The rest of the five categories (Extremely Sure, Highly Sure, Fairly Sure, Highly Unsure and Extremely Unsure) were recoded from five (Extremely Sure) to one (Extremely Unsure).

Tables 3 presents the regression results. Model 1 is the regression result obtained from the NCS certainty scores. No statistically significant effect could be detected for any of the explanatory variables used in Model 1.

INSERT TABLE 3 HERE

Model 2 presents the results obtained by regressing certainty scores of the PC responses. . The coefficient of the variable A_WTP (answer to the WTP question) is statistically significant in this model. The coefficient of A_WTP is negative which implies that, other things remaining the same, respondents who said 'No' to the WTP question stated higher certainty score than respondents who said 'Yes' . The coefficient of the variable 'CPRS' (if respondents have heard about the CPRS before the survey) is positive and statistically significant at the ten percent level. This implies that, respondents who had heard of the CPRS were more certain about their

'Yes/No' WTP decisions. The coefficient of the variable AGE was found to be positive and statistically significant which means that the self-reported certainty levels were higher for respondents who belonged to a higher age group. Finally, scenario uncertainty (S_UNCERT) was negative and statistically significant at the five percent level. The negative sign of the coefficient of S_UNCERT implies that more uncertain the respondents were about the future increase of temperature, the less certain they were about their decisions to support or not to support the policy. The coefficients of the variables BID, BIDSQ, and P_UNCERT were not statistically significant in Model 2.

In Model 3, the coefficients of the variables BID and BIDSQ were statistically significant at the one percent level. The signs of the coefficients of BID and BIDSQ are negative and positive respectively. This implies that the self-reported certainty scores are convex in bid level. In other words, at extremely low and high bids respondents were more certain of their decisions and less certain at intermediate bid levels. These results correspond to the theory and empirical evidence provided by Loomis and Ekstrand (1998). The coefficient of the variable P_UNCERT is negative and significant at one percent level. The negative sign of the coefficient was expected. This implies that respondents, who were more uncertain about the proposed CPRS to be effective in slowing down climate change, stated lower certainty scores about their decisions of paying or not paying for the CPRS. As in Model 2, respondents' age is found to have a positive influence on self-reported certainty score. The coefficients of the variables S_UNCERT and CPRS were not statistically significant in Model 3.

6. Results of certainty calibration

Table 4 presents mean WTP and their ninety five percent confidence interval estimated from the S_{NCS} and the S_{CCS} (hereafter WTP_{NCS} and WTP_{CCS}). Krinsky and Robb confidence intervals for the point estimates of mean WTP were estimated (Cooper 1999). Note that the WTP_{CCS} is estimated from a truncated WTP distribution. This is because a 'fat-tail' pattern was observed in the probability distribution function (PDF) of the stated WTP in the S_{CCS}. A common and relatively simple statistical approach to mitigate the fat-tail problem is to truncate the distribution of WTP at some upper limit, usually the second largest bid (Bishop and Heberlein 1979). Although the truncated WTP_{CCS} is ten percent lower than the WTP_{NCS}, their 95 percent confidence intervals overlap with each other. This implies that there is no statistical significance between WTP_{CCS} and WTP_{NCS}. The efficiency[3] score of the WTP_{NCS} is higher (0.37) than the WTP_{CCS} (0.56). Note that respondents answered a DC WTP question in both of these sample splits. Therefore, the difference in the PDF of the WTP and their respective efficiency scores are independent of the certainty measurement techniques.

INSERT TABLE 4 HERE

In Table 5, we present the results of certainty calibration in each split samples. Certainty calibration refers to the exercise of recoding original 'Yes' responses to 'No' responses based on some certainty scale cut-off points. The certainty calibrated mean WTP estimates were compared with the original DC WTP estimates to test for accuracy and loss (or gain) in efficiency. Note that for the $S_{PC,}$ a DC WTP estimate was not available[4]. The average of the S_{NCS} and the S_{CCS} WTP estimates and the average of the upper bounds and lower bounds of their 95 percent confidence intervals were used as the baseline to compare the certainty calibrated mean WTP estimates for the $S_{PC.}$

[3] Efficiency scores of the WTP estimates were calculated using the following formula: Efficiency = Difference between upper and lower confidence intervals over the mean WTP (Loomis and Ekstrand, 1998). The lower the efficiency score, the higher the efficiency of the estimator.
[4] In a PC method, respondents express their uncertainty directly through their WTP answers. This method does not include a DC WTP question.

INSERT TABLE 5 HERE

The calibration rules of the certainty scores are subjective and ad-hoc in CV studies (Akter et al., 2008). The most commonly used technique of incorporating the NCS measure is to recode the original 'Yes' DC responses to 'No' based on certainty scale cut-off points varying from seven to ten (Champ et al., 1997; Loomis and Ekstrand, 1998; Champ and Bishop, 2001; Samnaliev et al., 2006). In a multiple choice PC format, recoding is applied in a variety of combinations, e.g. calibrating all 'Yes' responses ('Definitely Yes', 'Probably Yes', 'Maybe Yes') as 1 and the rest as 0, or calibrating only 'Definitely Yes' responses as 1 and the rest as 0 (Johannesson et al., 1998; Blumenschein et al., 1998; Whitehead et al., 1998; Blumenschein et al., 2001; Chang et al., 2007).

In the S_{NCS} , the original 'Yes' responses were recoded based on four different certainty scale cut-off points (original 'Yes' responses were recoded to 'No' if certainty scores were 7, 8, 9 and 10), a calibration technique first used by Champ et al. (1997). Three different certainty calibration techniques were applied in the S_{PC} (Maybe Yes, Probably Yes and Definitely Yes = Yes; Probably Yes and Definitely Yes = Yes; Definitely Yes=Yes) following the calibration techniques applied by Chang et al. (2007). Finally, in the $S_{CCS,}$ two calibration techniques were applies (Fairly sure, Highly Sure and Absolutely Sure Yes= Yes; Absolutely Sure = Yes). The reason for applying only two calibration treatments in the S_{CCS} was that there were no uncertain 'Yes' responses below or above 'Fairly Sure'.

A calibration exercise reduces the probability of 'Yes' responses to each bid level as 'Yes' responses are recoded to 'No' responses. Therefore, it is expected that the certainty calibrated mean WTP values would be lower than the original DC WTP estimates. Accordingly, results presented in Table 5 show that all calibrated mean WTP values are lower than the mean WTP without certainty calibration except in one case. In the S_{PC}, the certainty calibrated mean WTP estimate is 117 percent higher than the DC WTP estimate when ‘Maybe Yes’, ‘Probably Yes’ and ‘Definitely Yes’ responses were recoded as ‘Yes’ responses. This result is expected. A recoding of ‘Maybe Yes’ responses to ‘Yes’ responses increased the proportion of ‘Yes’ responses by 40 percent in the S_{PC} than the proportion of ‘Yes’ responses in the DC elicitation format. Except for this one case, inclusion of the different certainty scale cut-off points yields mean WTP values that are 3 percent to 83 percent lower than the original DC mean WTP estimate. Comparing the value changes of calibrated mean WTP estimates across certainty measurement techniques, it appears that the lowest change of mean WTP value occurred in the S_{CCS} (3 to 32 percent). The reason for this small value change in the S_{CCS} was that the scale of calibration was relatively small in the S_{CCS} compared to the two other sample splits. Only 21 of the all ‘Yes’ responses were recoded to ‘No’ responses at the first stage (Fairly Sure=19 and Extremely Sure=2) and only one ‘Yes’ response was recoded to ‘No’ response (Highly Unsure=1) at the second stage of the calibration exercise.

In all the cases presented in Table 5, the certainty calibrated mean WTP were less efficient than the DC WTP estimate. The range of efficiency losses as a consequence of certainty calibration (7 percent to 276 percent) was larger than the range observed in existing empirical studies (22 percent to 149 percent) (Akter et al., 2008). Comparing the range of efficiency losses across different measurement techniques, the highest level of efficiency loss occurred in the S_{NCS} (100 percent to 276 percent). The range of efficiency loss was the lowest in the S_{CCS} (10 percent to 7 percent). The efficiency loss range in the S_{PC} lies between the ranges between of 65 percent and 100 percent which clearly implies that the calibrated WTP_{PC} is relatively less efficient than the calibrated WTP_{CCS}.

7. Discussions and conclusions

The main aim of the study was to test the construct validity of the self-reported certainty scores obtained from three preference uncertainty measurement techniques. We applied the NCS and PC methods using split sample treatments. The CCS method, a new method for measuring preference uncertainty, was applied in a third split sample.

The distribution of the certainty scores varied across the measurement techniques. In general, 'No' responses were found to be accompanied by relatively higher level of certainty scores than 'Yes' responses. However, this pattern was consistent across all the sample splits. The relative proportion of the highest certainty score for both 'Yes/No' responses was the lowest in the PC method whereas the CCS method generated the highest proportion of absolute certain responses. This systematic and significant variation in the distribution of the certainty scores in different sample split raise the question if preference uncertainty is an outcome of the technique used to measure them. Put simply, a specific preference uncertainty measurement technique may be one of the sources of uncertainty in respondents' preferences.

The regression results of the study provide new information about the sources of variation in the certainty scores. The 'Yes/No' response to the WTP question was found to be a significant determinant of the certainty level in two of the certainty measurement techniques. This implies that respondents who said ‘No’ to the WTP question consistently expressed significantly higher levels of certainty than those who said ‘Yes’. In addition, respondents’ perceived uncertainty about the future scenario and effectiveness of the proposed policy were found to be influencing their expressed certainty scores in two of the sample splits. As theoretically expected, the stated certainty scores were negatively affected by these attributes of environmental uncertainty. These results imply that the level of preference uncertainty in CV studies can be reduced if the environmental uncertainties associated with the valuation framework are resolved. Finally, respondents’ age group was found to be a significant, positive determinant of the certainty sores in two of the sample splits. In other words, older respondents were more certain about their responses to the valuation question than younger respondents. This result implies that part of the experienced uncertainty in preferences is purely exogenous and, therefore, can never be completely eliminated.

Respondents’ familiarity to the good being valued was found to have statistically significant positive impact on experienced certainty in the S_{PC}. The variable was not significant in the S_{CCS}. This result may be plausible because our questionnaire included an information section containing key descriptions of the good being valued and the policy measures under consideration. As a result, respondents may be equally informed about the good and the policy when they answer the WTP question. It is important to note that none of the other CV studies except Loomis and Ekstrand (1998) found prior knowledge or experience to have statistically significant impacts on the self-reported certainty scores.

While testing the construct validity of the self-reported certainty scores, verbal expression based techniques were found to be more appropriate than purely numerical scale based techniques. This result is consistent with the empirical evidence of the psychology literature that suggest verbal measure is more appropriate to measure psychological uncertainty that lay people experience in their everyday decision making. The NCS method showed poor construct validity. The variation of the self-reported numerical certainty scores could not be explained by the variations of any of the explanatory variables added in the model. This finding corresponds to the conclusion drawn by Windschitl and Wells (1996) in their paper. They asserted that researchers who use numerical measures may either underestimate or even fail to identify the effect of relevant variables on the psychological uncertainty experienced by the decision maker (Windschitl and Wells, 1996 pp 360).

Between the two verbal expression based techniques, both the PC and CCS performed equally well on the grounds of construct validity. Therefore, no conclusion can be drawn regarding the superiority of one technique over the other based on the validity criteria. However, the PC method generated significantly higher numbers of 'Yes' responses than the other two sample splits. From this perspective, the CCS method offers an improvement over the PC method as it allows the researcher to keep the conventional DC valuation question format unchanged. Also, the CCS method outperformed the PC method on efficiency grounds. However, there are some caveats to the CCS method. About 90 percent of the S_{CCS} respondents indicated complete certainty about their responses to the valuation question. Such a high proportion of complete confidence in WTP responses is unprecedented in the CV literature. This finding raises more questions than it answers. One plausible explanation of this result could be respondents' attempts to avoid or deny self-contradiction as pointed out by Samneliev et al. (2006). One might argue that this attempt should have been equally reflected in the two other sample splits. Note that unlike the S_{NCS} and the S_{PC}, a question decomposition exercise was undertaken in the $S_{CCS.}$ Before introducing the certainty scale, respondents were first asked if they were completely certain about their responses to the WTP question (the domain question). It might be the case that this domain question had sparked respondents' urges to stand by their choices at an extraordinary level compared to the NCS and PC method. This explanation is a speculation than conclusion. However, this is a hypothesis which can be tested in future research.

Finally, this study is one of the first attempts to compare the performance of the existing certainty measurement methods and to develop a new method that overcomes the deficiencies in existing measures. The results of our experiments provide two clear messages. First, the suitability of the numerical scale based method in capturing preference uncertainty in CV studies is questionable. Second, the CCS method holds promise as a useful measure of preference uncertainty. It helps to overcome some problems associated with the PC method. However, the method can be improved further particularly by adding more response options such as very (un)sure, quite (un)sure, somewhat (un)sure and so on. Furthermore, the pie diagrams could be replaced by facial expressions that reflect uncertainty. It might be challenging to develop such pictograms that explicitly distinguish the level of uncertainty through facial expressions. The 'Pain Rating Scales' developed by medical researchers could be a useful starting point.

References

Akter, S., Bennett, J. and Akhter, S. 2008. Preference uncertainty in contingent valuation. Ecol. Econ. 67: 345–351

Akter, S., Brouwer, R., Brander, L. and Beukering, P. V. 2009. Respondent uncertainty in a contingent market for carbon offsets, Ecol. Econ., 68: 1858-1863

Alberini, A., Boyle, K. and Welsh, M. 2003. Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. J. Env. Econ. and Mang., 45:40–62.

Beatty, P., Fowler, F.J., and Fitzgerald, G. 1999. Construction Strategies for Complex Survey Questions. Proceedings of the Survey Research Methods Section, American Statistical Association, pp. 973-978.

Bishop, R.C. and Heberlein, T.A. 1979. Measuring values of extra-market goods: are indirect measures biased? Am. J. Agr. Econ. 61: 926–930.

Blumenschein, K., Johannesson, M., Yokoyama, K.K. and Freeman, P.R. 2001. Hypothetical versus real willingness to pay in the health care sector: results from a field experiment, Journal of Health Economics, 20:441–57.

Blumenschein, K., Johannesson, M., Blomquist, G.C., Liljas, B. and O'Conor R.M. 1998. Experimental results on expressed certainty and hypothetical bias in contingent valuation, Southern Economic Journal, 65: 169–77.

Budescu, D.V. and Wallesten, T.S. 1990. Dyadic decisions with numerical and verbal probabilities, Organ. Behav. Hum. Dec., 46: 240-63.

Burkell, J. 2004. What are the chances? Evaluating risk and benefit information in consumer health materials. J. Med. Libr. Assoc., 92: 200-208.

Carson, R. and Groves, T., 2007. Incentive and informational properties of preference questions, Env. and Res. Econ. 37: 181-210.

Champ, P. A. and Bishop, R. C. 2001. Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. J. Env. Econ. and Manag. 19: 383-402.

Chang, J.I., Yoo, S.H. and Kwak, S.J., 2007. An investigation of preference uncertainty in the contingent valuation study. App. Econ. Lett., 14:691-695.

Cooper, J.C. 1999. Referendum CVM programs. Economic Research Service, USDA, Washington, DC.

Department of Climate Change. 2008. Carbon Pollution reduction Scheme Green Paper, url:http://www.climatechange.gov.au/greenpaper/summary/pubs/greenpaper-summary.pdf

Downey, H. K., Hellriegel, D., and Slocum, J. W. 1975. Environmental uncertainty: The construct and its applications. Adm. Sci. Q., 20: 613-629.

Duncan, R. 1972. Characteristics of organizational environment and perceived environment uncertainty, Adm. Sci. Q., 17:313 - 327.

Gigerenzer, G. 2003. Why does framing influence judgment? J. Gen. Intern. Med., 18: 960–961.

Hanley, N., Kristrom, B. and Shogren, J. 2009. Coherent arbitrariness: on preference uncertainty for environment goods, Land Econ. 85:41-50.

Hanemann, W. M., Kriström, B. and Li, C. 1995. Nonmarket valuation under preference uncertainty: econometric models and estimation, paper presented at the Annual

Meeting of the European Association of Environmental & Resource Economists, Umea, Sweden, June 17-20.

Hartley, A. 2006. Changing role of the speed of processing construct in the cognitive psychology of human aging. In Handbook of the Psychology of Aging, ed. James E. Birren and K. Warner Schaie. New York: Elsevier, 183–207.

Houts, P.S., Doak, C.C., Doak, L. G., Loscalzo, I.V.I.J. 2006. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ. Couns., 61:173-90.

Johannesson, M., Liljas, B. and Johansson, P-O. 1998. An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions, App. Econ., 30: 643–7.

Knight, F. H. 1921. Risk, uncertainty and profit. Houghton Mifflin, New York.

Kahneman, D., and Tversky, A. 1982. On the study of statistical intuitions. In D. Kahneman, P. Slovic & A. Tversky (Eds.). Judgment under uncertainty: Heuristics and biases (pp.493–508). New York: Cambridge University Press.

Larkin, J. H. and Simon, H. A. 1987. Why a diagram is (sometimes) worth ten thousand words. Cognitive Sci., 11: 65 – 100 .

Li, C. and Mattson, L. 1995. Discrete choice under preference uncertainty: an improved structural model for contingent valuation. J. Env. Econ. and Manag. 28:256–269.

Lipkus, I.M. and Hollands, J. 2000. The visual communication of risk. Monographs of the National Cancer Institute. , National Institutes of Health, Bethesda, MD.

Loomis, J. and Ekstrand, E. 1998. Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican Spotted Owl. Ecol. Econ. 27:29–41.

Marsiske, M. and Margrett, J. A. 2006. Everyday problem solving and decision making. In Handbook of the Psychology of Aging, ed. James E. Birren and K. Warner Schaie. New York: Elsevier, 315–42.

Mitchell, R. C. and Carson, R. T. 1989, Using Surveys to Value Public Goods. Washington, D.C.: Resources for the Future.

NOAA Blue Ribbon Panel on Contingent Valuation, 1993. Natural resource damage assessments under the oil Pollution Act of 1990, Federal Reg. 58:460–4614.

Park, D. C. 1998. Cognitive aging, processing resources, and self-report. In N. Schwarz, D. C. Park, B. Knauper, & S. Sudman. (Eds.). *Aging, cognition, and self-report,* (pp. 35-68). Philadelphia: Psychology Press.

Ready, R., Whitehead, J. and Blomquist, G. 1995. Contingent valuation when respondents are ambivalent. Journal of Env. Econ. and Manag. 29:181–197.

Renooij, S. and Witteman, C. 1999. Talking probabilities: communicating probabilistic information with words and numbers. Int. J. Approx. Reason., 22: 169-194

Roe, B., Boyle, K. and Teisl, M. 1996. Using conjoint analysis to derive estimates of compensating variation, J. Env. Econ. and Mang., 31: 145–159.

Salthouse, T. A. 1996. The processing-speed theory of adult age differences in cognition. Psychological Review, 103: 403–28.

Samnaliev, M., Stevens, T.H. and More, T. 2006. A comparison of alternative certainty calibration techniques in contingent valuation. Ecol. Econ. 57:507-519.

Stone, E.R., Yates, J.F. and Parker, A.M. 1997. Effects of numerical and graphical displays on professed risk-taking behaviour. J. Exp. Psychol-Appl, 3: 243–256

Tesler, M.D., Savedra, M.C., Holzemer, W.L. 1991. The word graphic rating scale as a measure of children's and adolescent's pain intensity. Research in Nursing and Health. 14: 1361-1371.

Weber, E. U. and Hilton, D. J. 1990. Contextual effects in the interpretations of probability words: Perceived base-rate and severity of events. J. Exp. Psychol. Human., 16, 781-789.

Whitehead, J.C., Huang, J.C., Blomquist, G.C. and Ready, R. C. 1998. Construct validity of dichotomous and polychotomous choice contingent valuation questions. Env. and Res. Econ., 11:107-116.

Windschitl, P. D. and Wells, G. 1996. Measuring psychological uncertainty: Verbal versus numeric. J. Exp. Psychol-Appl, 2:343-364.

Figure 1: The composite certainty scale

19. Are you sure about your answer in the previous question?

☐ Yes

☐ No

20. How do you feel about your answer to the question no 18? (TICK ONE BOX)

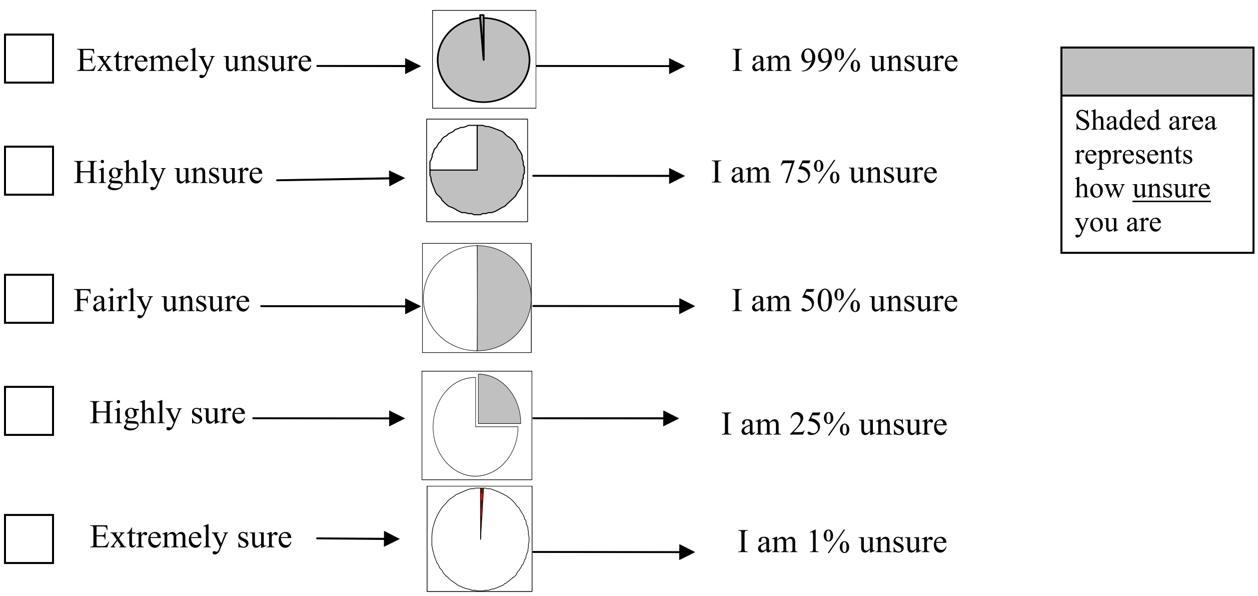

Figure 2: Distribution of the highest and the lowest certainty scores across measurement methods.

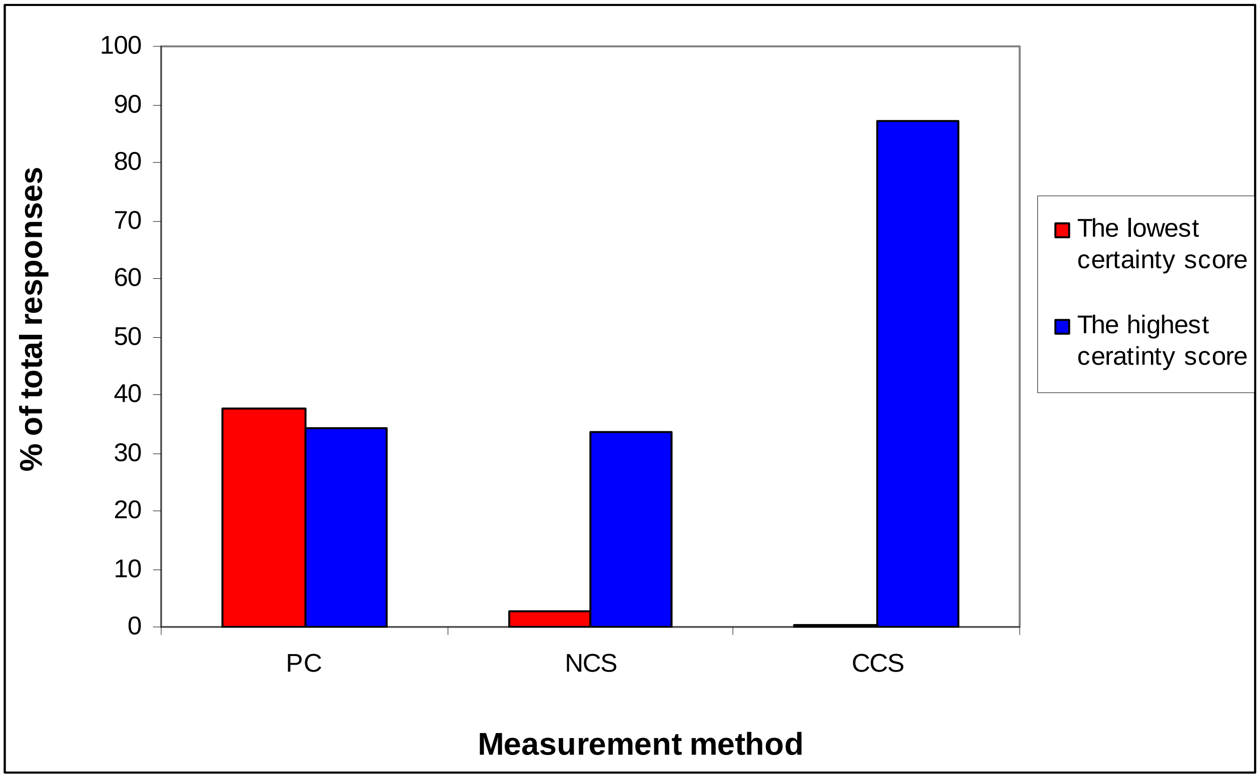

Table 1 Summary statistics of respondents' socio-economic characteristics.

Respondent characteristic		$^{a}S_{NCS}$	$^{b}S_{PC}$	$^{c}S_{CCS}$	*[d]Sydney average*	*[d]National average*
Sex ratio (male/female)		*0.93*	*1.07*	*0.84*	*1.16*	*0.99*
Respondent mean age (years)		*35*	*35*	*34*	*35*	*37*
Highest level of education (%)	*Year 12 or below*	*31*	*33*	*33*	*36*	*51*
	Certificate	*29*	*35*	*31*	*21*	*16*
	Bachelor's degree or above	*40*	*32*	*36*	*44*	*22*
Gross average household income (AUS$/week)		*1384*	*1372*	*1395*	*1360*	*1305*

Note:

[a]Sample split: Numerical certainty scale.

[b]Sample split: Polychotomous choice method.

[c]Sample split: Composite-certainty scale.

[d]Source: Australian Bureau of Statistics (2008)

Table 2 Comparison of 'Yes' and 'No' responses across different sample splits.

	Yes	No
[a]S_{NCS} (%)	33	67
[b]S_{CCS} (%)	32	68
[c]S_{PC} (%)	54	46

Note:

[a]Sample split: Numerical certainty scale.

[b]Sample split: Polychotomous choice method.

[c]Sample split: Composite-certainty scale.

Table 3 Ordered probit regression results for stated certainty scores of both 'Yes' and 'No' responses.

Variable	Description (Value range)	S_{NCS} : Model 1	S_{PC} : Model 2	S_{CCS} : Model 3
BID	Bid level (20, 50, 100, 150, 200, 250, 300, 400AUS$/month)	0.002 (0.002)	-0.098 (0.134)	-0.008*** (0.003)
BIDSQ	Square of bid level	-2.85e-06 (4.48e-06)	0.004 (0.15)	1.95E-05*** (7.50E-06)
A_WTP	Answer of the WTP question (Yes=1, No=0)	-0.199* (0.141)	-1.01*** (0.140)	-0.628*** (0.204)
CPRS	Respondents have heard of the CPRS (Yes=1, No=0)	0.044 (0.121)	0.256* (0.133)	0.103 (0.198)
S_UNCERT	Uncertainty about climate change (0 to 14.5)	-0.031 (0.022)	-0.067*** (0.025)	0.008 (0.040)
P_UNCERT	Policy uncertainty (0 to 100)	-0.002 (0.003)	-0.002 (0.004)	-0.010** (0.005)
AGE	Respondents' age group (1=18 to 24 years and 6= 65 years and above)	-0.008 (0.048)	0.091* (0.054)	0.272*** (0.079)
Model fit statistics				
Log likelihood		-595.76	-308.89	-136.01
LR chi square		9.43 (df=7, p<0.22)	78.73 (df=7, p<0.01)	33.05 (df=7, p<0.01)
Pseudo R^2		0.01	0.11	0.11
N		306	319	308

Explanatory notes:

Standard errors of the parameter estimates between brackets.

***: p<0.01; **: p<0.05; *: p<0.10.

Table 4 Mean WTP for the CPRS and 95% confidence interval (Krinsky and Robb 5000 repetitions).

Sub-sample	Mean WTP (AUS$)	95% confidence interval
S_{NCS}	133	112 – 162
S_{CCS}	119	92 – 159

Table 5 Mean WTP for the CPRS and 95% confidence interval (Krinsky and Robb 1000 repetitions).

Calibration Technique	% change in WTP estimate relative to Baseline	% change in efficiency[5] score of the WTP estimate relative to baseline
S_{NCS}		
YES7 (WTP Yes=Yes only for certainty ≥ 7)	-41	-100
YES8 (WTP Yes=Yes only for certainty ≥ 8)	-61	-284
YES9 (WTP Yes=Yes only for certainty ≥ 9)	-78	-216
YES10 (WTP Yes=Yes only for certainty ≥ 10)	-83	-276
S_{PC}		
MBYES (Maybe, Probably and Definitely Yes= Yes)	117	-145
PRYES (Probably and Definitely Yes =Yes)	-38	-33
DFYES (Definitely Yes =Yes)	-77	-115
S_{CCS}		
FSYES (Fairly Sure, Highly Sure and	-3	-10

[5] Efficiency score was calculated using the following formula: Efficiency =Difference between upper and lower CI over the Mean WTP.

Absolutely Sure Yes =Yes)		
ASYES (Absolutely Sure Yes=Yes)	-32	-7