Preference uncertainty in stated preference studies: facts and artefacts

Sonia Akter[a*] & Jeff Bennett[b]

[a]Department of Economics, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany

[b]Crawford School of Economics and Government, The Australian National University, Canberra, Australia

"*This is an Author's Original Manuscript of an article whose final and definitive form, the Version of Record, has been published in Applied Economics [05 April 2012] [copyright Taylor & Francis], available online at:* *http://www.tandfonline.com/*[10.1080/00036846.2012.654914]."

Abstract

The ordinal scale and polychotomous choice methods are two widely used techniques for estimating preference uncertainty in stated preference studies. This paper presents the results of two experiments that apply these estimation techniques. The first experiment was designed to compare and contrast the scores of the ordinal scale and polychotomous choice method. The second experiment was conducted to test a scale that combines verbal expressions with numerical and graphical interpretations: a composite scale. The results of the study can be summarized in three key findings. First, the polychotomous choice method generates a higher proportion of 'yes' responses than the conventional dichotomous choice elicitation format. Second, the composite scale generates a significantly higher proportion of certain responses. Finally, the ordinal scale performs poorly on the ground of construct validity.

Keywords: Preference uncertainty, contingent valuation, ordinal scale, polychotomous choice method, composite scale, climate change

1. Introduction

The concept of 'preference (or respondent) uncertainty' has gained significant attention in the stated preference (SP) literature over the past 20 years. Preference uncertainty, in the context of a hypothetical non-market valuation exercise, refers to an individual's lack of knowledge or confidence about their true value of a good in question (Li and Mattsson 1995). Hanemann *et al.* (1995) argued that individuals do not necessarily know their true valuations of a good with certainty. Rather they perceive their values of the good to lie within an interval. This knowledge gap gives rise to preference uncertainty which is considered to be partly responsible for hypothetical bias in SP studies (Champ and Bishop 2001). A number of empirical studies have estimated the magnitude of preference uncertainty to mitigate the disparity between hypothetical and actual behaviour (Champ *et al.* 1997; Champ and Bishop 2001; Ready *et al.* 2010). Others have used preference uncertainty information to understand the discrepancy between the dichotomous choice (DC) and open ended willingness to pay (WTP) contingent valuation (CV) elicitation formats (Welsh and Poe 1998; Ready *et al.* 2001).

Researchers have developed and applied a variety of scales to estimate preference uncertainty in SP studies. These scales can be either numerical (cardinal or ordinal) or verbal. Li and Mattson (1995) were the first to devise a cardinal scale. They asked respondents to express their uncertainty by selecting a number from a probabilistic scale ranging from 0 to 100 percent. Champ *et al.* (1997) introduced an ordinal scale in which respondents were asked to state their level of uncertainty on a 1 (very uncertain) to 10 (very certain) point scale. Ready *et al.* (1995) first introduced a verbal scale which is commonly known as the polychotomous choice method. Under this method, respondents are provided with the opportunity to express their certainty directly by choosing from a set of responses, e.g. 'definitely yes (no)', 'probably yes (no)' and 'maybe yes (no)'. Later, Blamey *et al.* (1999) proposed an enhanced version of this format, the dissonance-minimizing format, which allows respondents to select one response from a series of theoretically informed response categories that most closely resembles their view about the valuation scenario.

Practitioners do not all agree about which method is more appropriate to estimating preference uncertainty in SP studies (Akter *et al.* 2008). The choice of an uncertainty estimation technique is primarily based on an analyst's subjective judgment. The survey method literature suggests that the choice of an estimation scale may have important implications for determining response validity in social science research (Stevens 1946; DeVellis 2003). In particular, the estimation of latent variables (e.g. satisfaction, pain, happiness) can exhibit methodological artefacts (Peterson and Wilson 1992). Methodological artefact refers to systematic non-substantive survey responses often invoked by the questionnaire frame, order of alternatives and context of the question (Billiet and McClendon 2000; Horan *et al.* 2003). This issue has been a matter of long-standing interest in the psychology and survey method literatures. Some studies have concluded that people do not reveal their true preferences or pre-existing attitudes during surveys; rather, they use the questionnaire to determine their preferences and attitudes (Feldman 1990; Zaller and Feldman 1992).

These findings question if respondents reveal their true uncertainty about their WTP or if the level of preference uncertainty is determined by the estimation scales. The aim of this study is to investigate this issue by examining the validity of alternative uncertainty estimation scales.

Validity is defined as the degree to which the scale delivers what it is intended to estimate (Anastasi and Urbina 1997). This can be assessed by analyzing concurrent and construct validity. Concurrent validity refers to similarities in patterns among values estimated by different scales (McDowell and Newell 1987). Construct validity compares results against the theoretical construct of the notion being estimated. One of the most commonly used approaches to testing construct validity is to test hypotheses indicating correlation between the estimated variable and other variables of interest. Another indicator is evidence of the ability of the scale to discriminate among different groups of respondents (McDowell and Newell 1987).

To the best of our knowledge, no previous studies have tested the validity of preference uncertainty scales in SP experiments. Some empirical studies have compared the treatment effects of uncertainty adjusted WTP against the conventional DC WTP estimates in terms of accuracy and efficiency (Samnaliev *et al.* 2006; Chang *et al.* 2007). These studies provide no conclusions regarding the validity of estimation scales. In this paper, we apply the ordinal, polychotomous choice and composite scale methods using split sample treatments. The concurrent validity of these scales was examined by comparing (1) the distributions of the self-reported certainty scores and (2) the certainty adjusted mean WTP estimates. The construct validity test involved the assessment of regression results obtained by regressing the certainty scores against a set of theoretically and intuitively expected explanatory variables.

The rest of the paper is organized as follows. A discussion of the preference uncertainty estimation scales is presented in Section 2. The analytical model to be used as a framework for testing construct validity is developed in Section 3. A description of the survey and case study is provided in Section 4. The concurrent validity test results are presented in Section 5 followed by the regression results in Section 6. A discussion of the results and concluding remarks are presented in Section 7.

2. Scales to measuring preference uncertainty

Proponents of the probabilistic scale argue that it provides more precise information about the level of preference uncertainty as the respondent is able to specify a numerical probability on a 0 to 100 percent scale (Loomis and Ekstrand 1998). These values can be directly used as estimates of probabilities of paying (or not paying) the offered bid level. However, the appropriateness of this method has been criticized for its assumption that respondents are able to assess accurately the probability of their decisions to pay for an environmental good. This assumption is unlikely to hold true in practice as lay people tend to demonstrate poor understanding of numerical probability (Burkell 2004; Gigerenzer *et al.* 2005). Research dealing with peoples' statistical competencies reveals that information about probabilistic events tends to be difficult to comprehend, even for highly educated people (Burkell 2004).

In contrast, on an ordinal scale, preference uncertainty is assigned as numbers in a way that the order of the numbers reflects the level of uncertainty across respondents. For example, if respondent A selects 7 on the ordinal scale and respondent B selects 8, then respondent B is considered to be more certain about the decision of paying (not paying) than respondent A. However, the ordinal scale may be subject to distortions for two reasons. First, it has been observed that individuals interpret the levels of an ordinal scale differently (Brady 1985). Respondents show 'scale preference' in which some individuals tend to be low raters or high raters (Roe *et al.* 1996). This is commonly known as the ‘differential item functioning’ problem in the survey method literature which poses a serious threat to interpersonal comparison of ordinal responses. Second, the performance of an ordinal scale with regards to response reliability, validity and consistency has been found to be varying based on the number of categories included in a scale (see Preston and Coleman 2000 for a review).

Psychologists argue that verbal scale-based techniques are more appropriate to estimating uncertainty that people experience in everyday life (Windschitl and Wells 1996). This is because most people in every day life use words rather than numbers when describing their own uncertainty (Windschitl and Wells 1996). However, debate persists about the performance of the polychotomous choice method to estimating preference uncertainty in SP studies. The scale is considered to involve two perverse incentives. First, the scale allows respondents to assess the alternatives less carefully and give an affirmative response, without making a strong commitment (Ready *et al.* 1995). Thus, it provides respondents with an inducement to leave unresolved their lack of confidence in answering the valuation question (Ready *et al.* 1995; Alberini *et al.* 2003). As a consequence, the format generates higher rates of ‘yes’ responses (including ‘definitely yes’, ‘probably yes’ and ‘maybe yes’) than the standard DC question format. Second, the incentive compatibility property of a SP study is considered to be diminished when this format is used. Carson and Groves (2007) argue that the DC response format is a necessary but not sufficient condition to ensure incentive compatibility of a SP survey. Questions that involve more than two choice alternatives may provide opportunities for respondents to respond strategically.

One other pitfall of the polychotomous choice format is the subjectivity of the words used to elicit preference uncertainty (Hanley *et al.* 2009). For example, when a respondent is given the choice between 'probably yes' and 'maybe yes', unless the distinction between 'probably' and 'maybe' are objectively defined, the interpretation of these two responses could be highly subjective.

The composite scale was designed to overcome the shortcomings of the polychotomous choice method. The scale is based on the word-graphic rating scale (Tesler *et al.* 1991). The scale was developed using peer discussions, two focus group sessions and a small-scale pilot test. It used a verbal expression of certainty which is similar to the polychotomous choice method but the scale was introduced after the DC WTP question. Thus, it satisfies the necessary condition for incentive compatibility. The scale was decomposed into two major components so that the first component served as the domain (which tells the respondent what the question is about) and the second component served as the range (which tells respondents what they are expected to give back). This decomposition exercise was based on the domain-range questionnaire structure suggested by Beatty *et al.* (1999).

In the first step, respondents were asked whether they were certain about their answers to the WTP question (the domain question). Those who said 'no' to this question were asked the range question. Five categories of responses comprised the range question: 'extremely unsure', 'highly unsure', 'fairly unsure', 'highly sure', and 'extremely sure'. To overcome subjective interpretation of the verbal scale, each of these categories was associated with a numerical probability and a graphical expression (see Figure 1). Pie diagrams were added to help respondents visualize the information. The literature in health risk communication suggests that pictures accompanied by clear text help communication through higher attention, comprehension, recall and adherence by respondents (Houts *et al.* 2006). Furthermore, graphical representation of data improves judgment and decision-making by providing a holistic view of information (Lipkus and Hollands 1999).

INSERT FIGURE 1 HERE

3. Analytical framework for testing construct validity

Although no explicit theoretical model has been developed to explain variations in preference certainty scores, there is general agreement about some hypotheses that have emerged after Loomis and Ekstrand (1998). They suggested that the self-reported certainty score (C) is a quadratic function of bid level (Bid and BidSQ). Respondents' prior knowledge and familiarity with the good in question were also found to be important determinants of the self stated certainty scores. Psychology theory suggests that environmental uncertainty is an important source of psychological uncertainty (Downey *et al.* 1975). Here the term 'environment' refers to the basic conditioning of the decision making framework within which the individual decision maker operates from. 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). They provide direct input into an individual's cognitive mapping process and thus help determine the level of psychological uncertainty (Downey *et al.* 1975).

Two different forms of environmental uncertainty can be identified in a SP context, namely, subjective scenario and policy uncertainty. Subjective scenario uncertainty refers to an individual's lack of knowledge or information about the scale of future damage to an environmental good. Subjective policy uncertainty refers to an individual's lack of knowledge about the probability of a policy intervention being effective in remeding an environmental problem. A number of SP studies (Cameron 2005; Burghart *et al.* 2007; Akter and Bennett 2009) have investigated the role of these uncertainties on individual's WTP for non-marketed goods. They show that the magnitude of subjective scenario and policy uncertainties negatively influence individuals' valuations, i.e. the higher the subjective uncertainty, the lower the WTP. Since preference uncertainty arises from value uncertainty (i.e. individual not knowing the value of the good with certainty), it can be argued that the magnitude of subjective scenario and policy uncertainty may intensify the uncertainty experienced for the value of the good in question. This analogy suggests that subjective scenario and policy uncertainties are likely to affect preference uncertainty negatively. In other words, the higher the subjective uncertainty about scenario and policy, the lower the stated certainty score and vice versa.

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 (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 (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:

$$C_i = \alpha + \beta_1 \text{Bid}_i + \beta_2 \text{BidSQ}_i + \beta_3 \text{Knowledge}_i + \beta_4 \text{ Subjective Scenario Uncertainty}_i + \beta_5 \text{ Subjective Policy Uncertainty}_i + \beta_6 \text{ Age}_i (25) + \beta_7 \text{ Age}_i (35) + \beta_8 \text{ Age}_i (45) + \beta_9 \text{ Age}_i(55) + \beta_{10} \text{ Age}_i(65) + \varepsilon_i \quad (1)$$

Table 1 presents a description of the independent variables and the expected signs of their coefficients. In Equation 1, the variables Age(25), Age(35), Age(45), Age(55) and Age(65) are dummy variables to represent respondents from five age groups between 25 and over 65 years. The age group 18-24 years is the base level. α is a constant term and ε stands for error term.

INSERT TABLE 1 HERE

4. Survey and data

The context of the experiment was Australian households' preferences for the occurrence and mitigation of anthropocentric climate change. As part of fulfillment of its Kyoto Protocol obligations, the Australian Government proposed a national emissions trading scheme known as the Carbon Pollution Reduction Scheme (CPRS) in 2008. The aims of the CPRS were to reduce emissions by 60 percent of 2000 levels by 2050 and to encourage the development and use of reduced emission technologies (Department of Climate Change and Energy Efficiency 2008).

The implementation of the CPRS was expected to increase the prices of a wide range of emission-intensive goods and services and therefore, increase household expenditure. The case study aimed to explore Australian households' willingness to bear extra expenses to support the CPRS.

A web-based CV survey was conducted in Sydney in November-December 2008. The sample was stratified with respect to gender (50% male and 50% female), age group (50% lower than 35 and 50% higher) and education (33% high school or below, 33% certificate or graduate diploma and 33% university). These quotas were chosen based on the information collected from the Australian Bureau of Statistics (ABS 2008). Respondents were asked if their household would be willing to pay extra expenses (higher prices for electricity, fuel, public transport, food in restaurants and groceries) per month resulting from the CPRS. Eight different 'bids' ranging from A$20 to A$400 per month per household were randomly assigned across the respondents.

The survey was conducted in three split samples: (1) $SS_{Ordinal}$, (2) $SS_{Composite}$ and (3) SS_{PC}. Respondents were asked a DC CV question in the $SS_{Ordinal}$ and $SS_{Composite}$. In the $SS_{Ordinal}$, respondents were asked about their preference uncertainty using a 10 point numerical scale (1=highly uncertain, 10=highly certain) after the DC CV question. In the $SS_{Composite}$, respondents answered the composite scale preference uncertainty questions after the DC CV question. Respondents in the SS_{PC} answered the polychotomous choice CV question that included six preference uncertainty response categories ('definitely yes (no)', 'probably yes (no)' and 'maybe yes (no)').

The survey questionnaires included four questions to measure the two key variables included in the analytical model (Equation 1) – scenario and policy uncertainty. These questions were framed following the approach adopted by Cameron (2005) and Viscusi and Zeckhauser (2006). It was assumed that respondents do not know the exact change of future temperature (ΔT_i) nor the true probability of the policy being effective in mitigating the temperature change (P_i). However, respondents have distributions of ΔT and P in their minds. These distributions vary across individuals with respect to their mean ($\mu_{\Delta T}$, μ_P) and variance ($\sigma^2_{\Delta T}$, σ^2_P). Subjective scenario and policy uncertainties are reflected by the variances of these distributions.

Respondents were first shown a figure displaying average annual temperature in Australia for the period of 1910 to 2007. They were then presented with a series of 32 different levels of possible changes in annual average temperature ranging from minus 5° to 10°C and asked to indicate their perceptions of high and low guesses of temperature change in 2100 relative to the current year (see Figure 2). The difference between high and low guess temperatures was used as a measure of variance ($\sigma^2_{\Delta T}$). A similar approach was applied to estimate the variance of the distribution of the likelihood of policy effectiveness (σ^2_P). A numerical probability scale was used to elicit respondents' perceptions of their 'high guess' and 'low guess' of policy effectiveness. A verbal probability classification, consistent with the IPCC likelihood scale, was attached to the numerical scale (see Figure 3). The difference between high and low guess of the subjective probability of policy effectiveness was used to estimate subjective policy uncertainty (σ^2_P)[1].

[1] Note that all responses regarding the ranges of temperature change and likelihood of policy effectiveness were included considered in the analysis. In other words, the stated ranges were not assessed for validity by comparing them with official estimates.

INSERT FIGURE 2 HERE

INSERT FIGURE 3 HERE

No statistically significant differences were observed among respondents' socio-demographic characteristics across the three sample splits. Over 50 percent of the respondents were female and their average age was about 34 years. On average, the households consisted of three members. Three-quarters of the sample respondents were employed and two-thirds of them were working full time. Average weekly household income was A$1,442, which was significantly higher than the Sydney (A$1,360) and national (A$1,305) household income per week (ABS 2008). Trimming off the five percent lowest and highest values, the average weekly household income equalled to A$1,346. The differences between the trimmed sample mean weekly income and Sydney population and national population income were not statistically significant at the 10 percent level.

5. Concurrent validity results

5.1. Distributions of certainty scores

In this sub-section, the distributions of self-reported certainty scores obtained from the three sample splits are compared on the basis of three criteria. First, we examine whether the polychotomous choice method generates a higher proportion of 'yes' responses compared to the DC elicitation format. Second, we investigate whether the three estimation techniques generate equal proportions of low-end, mid-scale and high-end certainty scores. Finally, we examine whether the distributions of certainty scores across 'yes/no' WTP responses show similar patterns.

Recoding 'definitely yes', 'probably yes' and 'maybe yes' responses to 'yes' and 'definitely no', 'probably no' and 'maybe no' responses to 'no', 54 (46) percent of 'yes' ('no') responses were observed in the SS_{PC}[2]. The proportion of 'yes' responses in the S_{PC} is 63 percent higher than that of $SS_{Ordinal}$ and $SS_{Composite.}$ This difference in the distribution of 'yes/no' responses in the S_{PC} was significantly different from the other two sub-samples (SS_{PC} & $SS_{Ordinal}$: Chi square=7, p<0.01; SS_{PC} & $SS_{Composite}$: Chi square=10, p<0.01). No statistically significant difference was observed between the 'yes/no' WTP responses across the $SS_{Ordinal}$ and $SS_{Composite}$ (Chi square=0.16, *p*<0.7). The proportions of 'yes' responses across bid levels were compared in the three sample splits (Figure 3). At bid levels A$200, A$250 and A$300, a significantly higher proportion of respondents in the S_{PC} said 'yes' than the other two sample splits (A$200: Chi square=6, p<0.05; A$250: Chi square=3, p<0.10; A$300: Chi square=10, p<0.01). These results provide evidence in support of Ready *et al.*'s (1995) proposition that the polychotomous choice format induces a tendency to give affirmative responses, particularly at high bid levels, without any strong commitment.

INSERT FIGURE 4 HERE

A third (34%) of the respondents in the SS_{PC} stated the high-end certainty level (definitely yes/no) about their preferences for paying (or not paying) for the CPRS while less than a third (28%) chose the mid-scale response ('probably yes') and over a third (38%) stated the low-end certainty level (maybe yes/no). In the $SS_{Ordinal}$, about half (47%) of the respondents stated the high-end certainty scores (8, 9 and10), over 40 percent stated the mid-scale certainty scores (5, 6

[2] It might be argued that 'probably yes (no)' and 'maybe yes (no)' responses are not the same as the DC 'yes (no)' responses. However, this treatment approach was applied following Ready *et al.* (1995) and Whitehead *et al.* (1998).

and 7) while the rest (11%) stated the low-end certainty scores (1, 2, 3 and 4). The difference in the distribution of stated certainty scores in the polychotomous choice and ordinal scale was statistically significant at the one percent level (Chi square=54, $p<0.001$). This result implies that respondents answering the ordinal scale tend to state relatively higher certainty about their WTP decisions than the polychotomous choice format.

In the $SS_{Composite}$, 87 percent of respondents said 'yes' to the domain question (are you sure about your answer to the WTP question?). The remaining 13 percent proceeded to the range question. None of the respondents who answered the range question selected the 'extremely sure' option and one respondent selected the 'extremely unsure' option. About 10 percent said they were 50 percent sure and three percent indicated they were 75 percent sure. The domain question was meant to detect respondents who had absolutely no doubts about their decisions. Therefore, respondents who said 'yes' to the domain question (hereafter domain-yes) were exempted from answering the range question to avoid repetition[3]. However, the large number of 'domain-yes' responses raises concern about how the domain question was interpreted by different respondents. The term 'sure' in the domain question may not have been interpreted as 100 percent certainty by all respondents who answered 'yes' to this question.

Two alternative approaches – symmetric and asymmetric – were adopted to address the potential ambiguity associated with the domain-yes responses. The symmetric approach is widely used to allocate numerical probabilities to mirror-image verbal probability phrases (e.g. 'sure-unsure') (Clarke *et al.* 1992). This approach assumes perfect symmetry between positive and negative

[3] Participants in the focus groups and peer discussions reported that they found the range question repetitive once they had said 'yes' to the domain question.

phrases and therefore assigns 50 percent probability to each. According to this approach, the domain-yes responses were assumed to be associated with at least 50 percent certainty. More specifically, the symmetric approach assumes that the certainty scores of those respondents who answered ‘yes’ to the domain question lie between 50 and 100 percent.

The validity of the symmetric approach has been questioned by research that demonstrates mirror-image probability phrases are more likely to be asymmetric with positive expressions having greater probabilities than negative expressions (Lichtenstein and Newman 1967; Reagan *et al.* 1989). Empirical studies have found that the positive phrases (e.g. probable, probably, likely, sure) are associated with over 70 percent numerical probability (Clarke *et al.* 1992; Lau and Ranyard 1999). Therefore, according to the asymmetric approach, it can be assumed that the domain-yes responses are associated with at least 70 percent certainty. In other words, the certainty scores of domain-yes respondents lie between 70 and 100 percent.

The asymmetric approach generates 90 percent high-end certainty responses in the composite scale. This implies that about 90 percent of respondents in the $SS_{Composite}$ stated 70 percent or higher certainty about their WTP answers. This proportion was significantly higher than the high-end certainty responses observed in the two other scales ($SS_{Composite}$ and SS_{NCS}: Chi square=178, p<0.001; $SS_{Composite}$ and SS_{PC}: Chi square=183, p<0.001). The symmetric probability allocation approach generated over 95 percent mid-scale and high-end certainty response. This implies that over 95 percent of respondents in the $SS_{Composite}$ were at least 50 percent certain about their WTP answers. This proportion is significantly higher than the ordinal scale (85%) and polychotomous choice format (62%) ($SS_{Composite}$ and SS_{NCS}: Chi square=85,

$p<0.001$; $SS_{Composite}$ and SS_{PC}: Chi square=89, $p<0.001$). These results imply that the composite scale generates the highest proportion of certain responses compared to the two other scales regardless of the treatment applied to the domain-yes responses.

The distributions of the self-reported certainty scores in the three sample splits were examined for differences across ‘yes/no’ WTP responses. In the SS_{PC} , a significantly higher proportion (49%) of the respondents who said ‘no’ to the WTP question were very certain (‘definitely no’) about their decisions as opposed to only 15 percent of the ‘yes’ respondents who were very certain (‘definitely yes’) (Chi square=54.102, $p<0.001$). Likewise, a significantly higher proportion (40%) of ‘no’ respondents in the $SS_{Ordinal}$ were very certain (certainty score =10) about their decisions as opposed to less than 20 percent of the ‘yes’ respondents who were very certain (Chi square=29, $p<0.001$). In the $SS_{Composite}$, a significantly higher proportion (91%) of ‘no’ respondents said ‘yes’ to the domain question compared to 78 percent of ‘yes’ respondents who said ‘yes’ to the domain question (Chi square=17, $p<0.001$). These results imply that the ‘no’ responses tend to be held with greater certainty scores than ‘yes’ responses regardless of the estimation method.

5.2. Certainty adjustment results

Table 2 presents certainty adjusted mean WTP and their 95 percent confidence intervals. Certainty adjustment refers to the exercise of recoding original ‘yes’ responses to ‘no’ based on some cut-off points, e.g. recoding the DC ‘yes’ responses to ‘no’ if ordinal scores are greater than seven. In the polychotomous choice method, adjustments are made by recoding ‘definitely yes’ and ‘probably yes’ as ‘yes’ and the rest as ‘no’, or recoding only ‘definitely yes’ responses as ‘yes’ and the rest as ‘no’ (see Akter *et al.* 2008 for a review). We applied two recoding principles to allow inter-scale comparison of uncertainty adjusted mean WTP estimates: high-end and mid-scale adjustment. Under the high-end adjustment principle, the ‘yes’ responses accompanied by the high-end certainty scores (only definitely yes=yes in the SS_{PC}; 8, 9 and 10 in the $SS_{Ordinal}$; certainty score ≥ 70% in the $SS_{Composite}$) were considered as true ‘yes’ and the rest were coded as ‘no’. The mid-scale adjustment principle refers to the recoding rule where ‘yes’ respondents who selected a certainty score located at the middle of the scale (‘probably yes’ and ‘definitely yes’ in the SS_{PC}, 5 and above in the $SS_{Ordinal}$ and ‘fairly sure’, ‘highly sure’, ‘extremely sure’ and domain-yes responses in the $SS_{Composite}$). These certainty-adjusted mean WTP estimates were compared with the original DC WTP estimates.

INSERT TABLE 2 HERE

A non-parametric approach, suggested by Kriström (1990) was applied to estimate the mean WTP vales. The main advantage of nonparametric estimators is that they are robust against functional misspecification (Kerr 2000). Krinsky and Robb (1986) confidence intervals for the point estimates of mean WTP were estimated using the referendum CVM programs (in GAUSS) written by Cooper (1999). The mean WTP estimate for $SS_{Ordinal}$ (A\$143) and $SS_{Composite}$ (A\$153) (the DC WTP format) were not statistically[4] different from each other. The polychotomous choice WTP estimate[5] (A\$230) was about 50 percent higher than the DC WTP estimates. This difference was statistically significant at the five percent level.

[4] The convolution based Poe *et al.* (1994) test was used to test statistical differences between mean WTP estimates.
[5] This was obtained by recoding ‘maybe yes’, ‘probably yes’ and ‘definitely yes’ responses to ‘yes’ and the rest as ‘no’.

The mid-scale adjustment principle generated a slightly lower mean WTP estimate for $SS_{Ordinal}$ (A$123) while the same principle caused mean WTP estimate for $SS_{Composite}$ (A$154) to increase. However, these estimates were not statistically different from each other. The mid-scale adjusted WTP estimated for SS_{PC} (A$70) was significantly lower than the estimates obtained for $SS_{Ordinal}$ and $SS_{Composite}$. The composite scale generated the highest high-end adjusted mean WTP estimate (A$136). However, note that this estimate was obtained by applying the asymmetric probability principle (domain-yes ≥ 70% certainty). The high-end adjusted mean WTP estimated for $SS_{Ordinal}$ (A$52) was significantly higher than the high-end adjusted mean WTP estimate obtained from SS_{PC} (A$27).

6. Construct validity results

This section presents the regression results obtained from the estimation of Equation 1 presented in Section 3. The stated certainty scores obtained from the ordinal scale and polychotomous choice method are ordinal. The ordered probit model, first introduced by McKelvey and Zavoina (1975), serves as an appropriate framework for statistical analysis in this case. It uses the information that one response category is higher than the other by ignoring the magnitude of the differences. Therefore, two ordered probit regression models were estimated using the certainty scores of both ‘yes’ and ‘no’ responses. The interval (or grouped data) regression approach, similar to an ordered probit model, was applied to analyze the scores from the composite scale. The interval regression approach is applicable when the dependent variable is limited to a certain number of categories, but the ranges of the underlying variable to which each category refers to are known (Wooldridge 2007) [6].

The results obtained for the ordinal scores are presented in Model 1 (Table 3). No statistically significant effects could be detected for any of the explanatory variables used in this model. Model 2 presents the polychotomous choice results. The coefficient of Knowledge (if respondents have heard about the CPRS before the survey) was positive and statistically significant at less than 10 percent level. This implies that, respondents who had heard of the CPRS were more certain about their 'yes/no' WTP decisions. The coefficient of Subjective Scenario Uncertainty ($\sigma^2_{\Delta T}$) was negative and statistically significant at the five percent level, implying that the more uncertain respondents were about the future increase of temperatures, the less certain they were about their decisions to support or not to support the policy. The coefficients of Age(45) and Age(55) were positive and statistically significant at the five and 10 percent level respectively. This implies that respondents of these two age groups (45-54 years and 55-64 years) were significantly more certain about their WTP decisions than other respondents. Although the coefficient of Age(55) was higher than the coefficient of Age(45), the difference was not statistically significant at the 10 percent level (Chi square=0, p<0.96).

---

[6] In the $SS_{Ordinal}$ and SS_{PC}, preference certainty is a latent variable. In the $SS_{Composite}$, the stated certainty scores have some quantitative meaning and so in that case preference certainty is not a latent variable. However, the exact level of certainty is still not observable. We only observe whether certainty falls within a specific range. Theoretically, an interval regression approach is more efficient than an ordered probit approach to model this latter type of variable since the estimation procedure utilizes information provided by the thresholds values to produce an estimate of the standard deviation rather than requiring that this be normalized to one (Horowitz 1994).

INSERT TABLE 3 HERE

Models 3 and 4 in Table 3 present the results for the composite certainty scores[7]. Note that the dependent variables in these two models were coded as an interval with an upper and lower limit. These two models vary in their assumptions about the interpretation of the domain-yes responses. Model 3 is estimated based on the asymmetric probability allocation principle to the domain-yes responses. It assumes that the certainty scores of respondents who answered 'yes' to the domain question lie between 70 (lower limit) and 100 (upper limit) percent. Model 4 assumes symmetric probability allocation to the domain-yes responses. The certainty scores of respondents who answered 'yes' to the domain question were assumed to lie between 50 (lower limit) and 100 (upper limit) percent. The certainty scores of the rest of the respondents (who answered the range question) were coded according to the values included in the scale. In both Models 3 and 4, the coefficients of the variables Bid and BidSQ were statistically significant. As expected, the signs of the coefficients of Bid and BidSQ were negative and positive respectively. The coefficient of Subjective Policy Uncertainty (σ^2_P) was negative and significant at the five percent level in Model 3. However, the coefficient of this variable was not statistically significant in Model 4. As in Model 2, respondents' age was found to have a positive influence on self-reported certainty scores in both Model 3 and 4. The coefficients of Age(35), Age(45) and Age(55) were positive and statistically significant at the 10, one and five percent level respectively. This implies that respondents of these three age groups were significantly more certain about their decisions than respondents belonging to the other age groups. However, none

[7] The full set of certainty responses (n=308) was included in the analysis.

these age coefficients were significantly different from each other implying the absence of non-linear relationships between age and stated certainty scores.

7. Discussion and conclusion

The aim of the study was to investigate the concurrent and construct validity of the preference uncertainty estimation techniques. We applied the ordinal scale, polychotomous choice method and composite scale using split sample treatments. The concurrent validity test results show evidence of methodological artefact. It was observed that the polychotomous choice format generates higher proportion of 'yes' responses, particularly at higher bid levels. The distributions of low-end, mid-scale and high-end scores were found to be significantly different across preference uncertainty estimation scales. The composite and ordinary scales generated significantly higher proportion of mid-scale and high-end scores than the polychotomous choice format.

Differences were observed in the certainty adjusted mean WTP estimates obtained from the three sample splits. The polychotomous choice format produced a significantly higher mean WTP estimate than the conventional DC WTP estimates. The mid-scale and high-end adjusted mean WTP estimated for the polychotomous choice sample was significantly lower than the mid-scale and high-end adjusted mean WTP estimated for ordinal and composite scales. Furthermore, the high-end adjusted mean WTP for the composite scale was significantly higher than the high-end adjusted mean WTP of the ordinal scale. These findings imply that the choice of a preference uncertainty estimation scale may generate significantly different welfare estimates and therefore, may lead to different conclusions with regards to the magnitude of the deviation between hypothetical and actual behaviour or the DC WTP and open-ended WTP estimates.

In the absence of a theoretical model of preference uncertainty in SP literature, a number of hypotheses were formed based on theoretical and intuitive reasoning to test the construct validity of the self-reported certainty scores. The variations of the polychotomous choice responses were explained by variations in respondents' familiarity, scenario uncertainty and age. The construct validity of the composite scores was found to be sensitive to the treatment of the domain-yes responses. When an asymmetric numerical probability assignment principle was applied, variations in the composite scores were explained by bid levels, policy uncertainty and age. Under the symmetric probability allocation principle, policy uncertainty did not have any statistically significant influence on the composite scores. The ordinal scale, the most widely used scale to estimate preference uncertainty in SP studies, showed poor construct validity. The variation of the ordinal scores could not be explained by variations in any of the explanatory variables in the regression model.

The construct validity test is not conclusive given that the hypotheses were not drawn from a theoretical model of preference uncertainty. However, the results are indicative of inadequacy of the ordinal scale in estimating preference uncertainty. Although the polychotomous choice method showed better construct validity than the ordinal and composite scale, the concurrent validity test results provided evidence in support of the widely-held belief that this format induces false uncertainty (Ready *et al.* 1995; Alberini *et al.* 2003). The composite scale performed better than the ordinal scale on construct validity grounds. Furthermore, the composite scale offers an improvement over the polychotomous choice format. It maintains the conventional DC valuation question format, allows expression of certainty on a verbal scale and provides numerical and visual interpretation of the verbal scales to avoid subjective interpretation and to ensure better comprehensibility. However, there are some caveats to the use of the composite scale. The scale generated an unprecedented proportion of certain responses. This could be due to the structure of the domain-range of the scale. One possible remedy for this problem would be to ask all respondents both the domain and range questions. Another option could be to add a measure of numerical probability to the domain question, e.g. are you 100% sure about your response?

Finally, this study is one of the first attempts to compare preference uncertainty scales and to develop a new scale that overcomes some deficiencies in existing scales. The results of our experiments provide three conclusions. First, the choice of an estimation scale influences the level of preference uncertainty in CV studies. Second, the suitability of the ordinal scale in estimating preference uncertainty in CV studies is questionable. Finally, the composite scale holds promise as a useful estimation technique. However, further research is necessary to explore the full potential of this method.

References

ABS (Australian Bureau of Statistics) (2008). *Year Book of Australia*, Australian Bureau of Statistics, Canberra.

Akter, S., Bennett, J. and Akhter, S. (2008). Preference uncertainty in contingent valuation, *Ecological Economics* 67, 345–351

Akter, S. and Bennett, J. (2009). Household perceptions of climate change and preferences for mitigation action: the case of the Carbon Pollution Reduction Scheme in Australia, EERH Research Report No.19

Alberini, A., Boyle, K. and Welsh, M. (2003). Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty, *Journal of Environmental Economics and Management* 45, 40–62.

Anastasi, A. and Urbina, S. (1997). *Psychological Testing*. 7th edn. Upper Saddle River, NJ: Prentice Hall.

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.

Billiet, J. B. and McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items, *Structural Equation Modeling* 7, 608–628.

Blamey, R.K., Bennett, J.W. and Morrison, M.D. (1999). Yea-saying in contingent valuation surveys, *Land Economics* 75(1), 126–141

Brady, H. (1985). The perils of survey research: Inter-personally incomparable responses, *Political Methodology* 11(3–4), 269–90.

Burghart, D.R., Cameron, T.A. and Gerdes, G.R. (2007). Valuing publicly sponsored research projects: risks, scenario adjustments, and inattention, *Journal of Risk and Uncertainty* 35(1), 77–105.

Burkell, J. (2004). What are the chances? Evaluating risk and benefit information in consumer health materials, *Journal of the Medical Library Association* 92, 200-208.

Cameron, T.A. (2005). Individual option prices for climate change mitigation, *Journal of Public Economics* 89(2–3), 283–301.

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

Champ, P.A., Bishop, R.C., Brown, T.C. and McCollum, D.W. (1997). Using donation mechanisms to value non-use benefits from public goods, *Journal of Environmental Economics and Management* 33, 151–162.

Champ, P. A. and Bishop, R. C. (2001). Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias, *Journal of Environmental Economics and Management* 19, 383-402.

Chang, J.I., Yoo, S.H. and Kwak, S.J. (2007). An investigation of preference uncertainty in the contingent valuation study, *Applied Economics Letter* 14, 691-695.

Clarke, V. A., Ruffin, C. and Hill, D. (1992). Ratings of orally presented verbal expressions of probability by a heterogeneous sample**,** *Journal of Applied Social Psychology* 22, 630-656.

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

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

DeVellis, R.F. (2003). *Scale Development* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Downey, H. K., Hellriegel, D. and Slocum, J. W. (1975). Environmental uncertainty: The construct and its applications, *Administrative Science Quarterly* 20, 613-629.

Duncan, R. (1972). Characteristics of organizational environment and subjective environment uncertainty, *Administrative Science Quarterly* 17, 313 - 327.

Feldman, S. (1990). Theory, measurement, and the survey response, Presented at the annual meeting of the American Political Science Association, Atlanta.

Gigerenzer, G., Hertwig, R., Broek, E. van den, Fasolo, B. and Katsikopoulos, K. V. (2005). "A 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts? *Risk Analysis* 25, 623-629.

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.

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

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.

Horan, P.M., DiStefano, C. and Motl, R.W. (2003). Wording effects in self esteem scales: methodological artifact or response style?, *Structural Equation Modeling* 10, 444–455..

Horowitz, J.L. (1994). Bootstrap-based critical values for the information matrix test. *Journal of Econometrics*, 61, 395-411.

Houts, P.S., Doak, C.C., Doak, L. G. and 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 Education and Counselling* 61, 173–90.

Kerr, G.N. (2000). Dichotomous choice contingent valuation probability distributions, *AustralianJournal of Agricultural and Resource Economics* 44, 233–252.

Krinsky, I. and Robb, A. L. (1986). On approximating the statistical properties of elasticities, *Review of Economics and Statistics* 68(4), 715–719.

Kriström, B. (1990). A non-parametric approach to the estimation of welfare measures in discrete response valuation studies, *Land Economics* 66: 135–139

Lau, L. and Ranyard, R. (1999). Chinese and English speakers' linguistic expression of probability and probabilistic thinking, *Journal of Cross-Cultural Psychology* 30, 411–421.

Li, C. and Mattson, L. (1995). Discrete choice under preference uncertainty: an improved structural model for contingent valuation, *Journal of Environmental Economics and Management* 28, 256–269.

Lichtenstein, S. and Newman, J. R. (1967). Empirical scaling of common verbal phrases associated with numerical probabilities, *Psychonomic Science* 9(10), 563-564.

Lipkus, I.M. and Hollands, J. (1999). The visual communication of risk, *Journal of the National Cancer Institute Monographs* 25, 149-163.

Loomis, J. and Ekstrand, E. (1998). Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican Spotted Owl, *Ecological Economics* 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.

McDowell I. and Newell C. (1987). *Measuring health: a guide to rating scales and questionnaire*, Toronto: Oxford University Press.

McKelvey, R.D. and Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables, *Journal of Mathematical Sociology* 4, 103–120.

Peterson, R. A. and Wilson R. W. (1992). Measuring customer satisfaction: fact and artefact, *Journal of the Academy of Marketing Science* 20, 61-71

Poe, G.L., Severance-Lossin, E.K. and Welsh, M.P. (1994). Measuring the difference (X - Y) of simulated distributions: a convolutions approach, *American Journal of Agricultural Economics* 76(4), 904–915.

Preston, C.C. and Colman, A.M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences, *Acta Psychologica* 104, 1–15.

Ready, R.C., Champ, P.A. and Lawton, J. L. (2010). Using respondent uncertainty to mitigate hypothetical bias in a stated choice experiment**,** *Land Economics* 86(2), 363-381.

Ready, R.C., Navrud, S. and Dubourg, W.R. (2001). How do respondents with uncertain willingness to pay answer contingent valuation questions?, *Land Economics* 77 (3), 315– 326.

Ready, R.C., Whitehead, J. and Blomquist, G. (1995). Contingent valuation when respondents are ambivalent, *Journal of Environmental Economics and Management* 29, 181–197.

Reagan, R. T., Mosteller, F. and Youtz, C. (1989). Quantitative meanings of verbal probability expressions, *Journal of Applied Psychology* 74, 433-442.

Roe, B., Boyle, K. and Teisl, M. (1996). Using conjoint analysis to derive estimates of compensating variation, *Journal of Environmental Economics and Management* 31, 145–159.

Samnaliev, M., Stevens, T.H. and More, T. (2006). A comparison of alternative certainty calibration techniques in contingent valuation, *Ecological Economics* 57, 507-519.

Stevens, S.S. (1946). On the theory of scales of measurement, *Science* 103, 677–680.

Tesler, M.D., Savedra, M.C. and 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.

Viscusi, W.K. and Zeckhauser, R.J. (2006). The perception and valuation of the risks of climate change: a rational and behavioral blend, *Climatic Change* 77(1–2), 151–177.

Welsh, M.P. and Poe, G.L. (1998). Elicitation effects in contingent valuation: comparisons to a multiple bounded discrete choice approach, *Journal of Environmental Economics and Management* 36(2), 170-85.

Whitehead, J.C., Huang, J.C., Blomquist, G.C. and Ready, R. C. (1998). Construct validity of dichotomous and polychotomous choice contingent valuation questions, *Environmental and Resource Economics* 11, 107-116.

Windschitl, P. D. and Wells, G. (1996). Measuring psychological uncertainty: Verbal versus numeric, *Journal of Experimental Psychology: Applied* 2, 343-364.

Wooldridge, J. 2007. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Zaller, J. and Feldman, S. (1992). A simple theory of the survey response: answering questions versus revealing preferences, *American Journal of Political Science* 36(3), 579-616.

Figure 1: Valuation question and composite scale

18. Would you be willing to spend an extra <u>$ X</u> per month <u>starting from 2010</u> on your current household spending to pay for the 'Carbon Pollution Reduction Scheme'?

[Please note that this is not a tax or levy. The extra expense is due to rise in prices of necessary goods and services.]

[When answering this question, keep in mind your views on the likelihood of other polluting countries introducing a similar scheme.]

☐ Yes

☐ No

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)

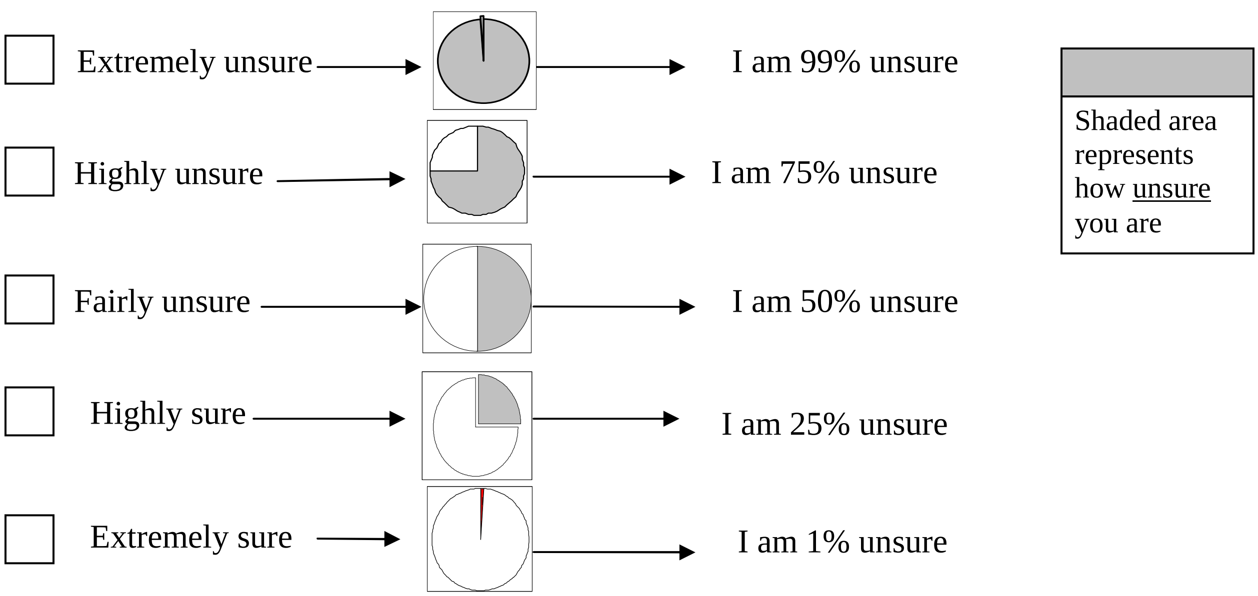
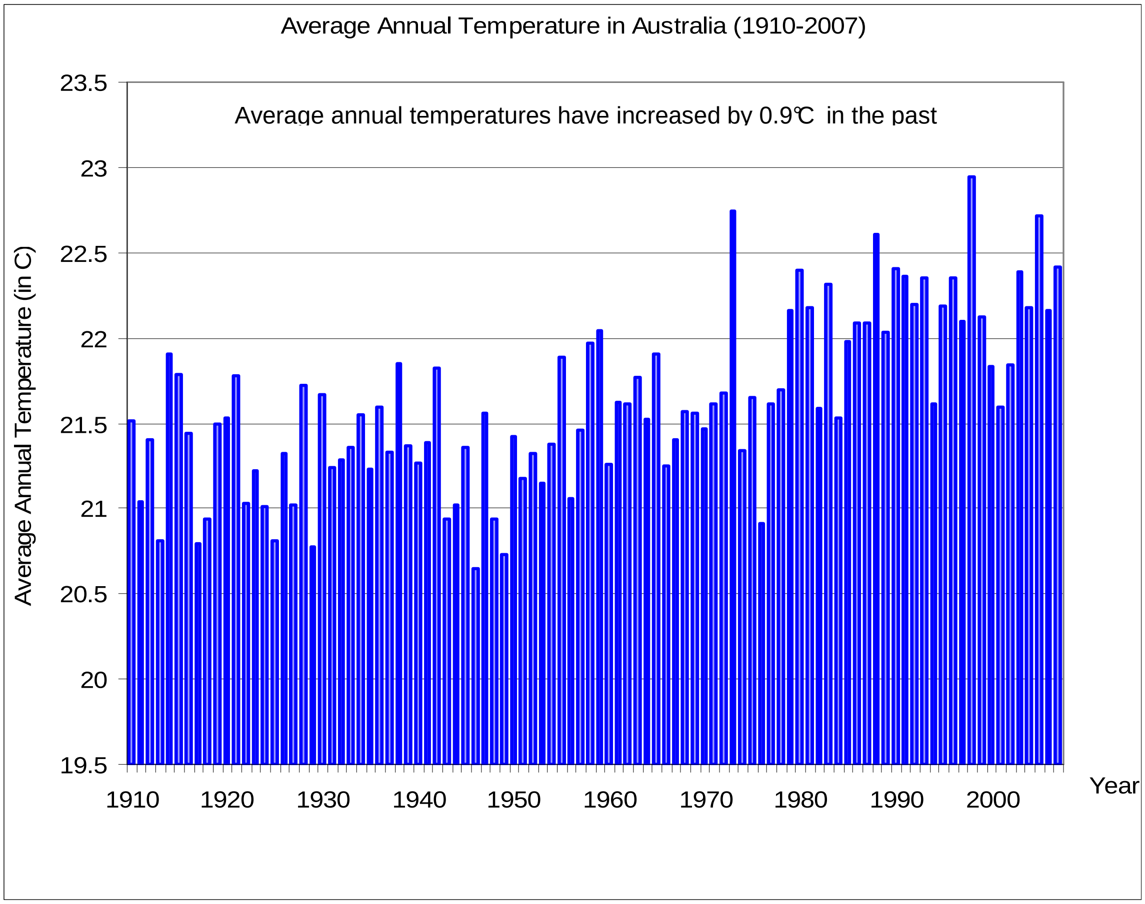


Figure 2 **Questions asked to estimate subjective scenario uncertainty**

<u>Section A:</u> We would like to know your perception about future temperature change. The following graph shows the average annual temperatures in Australia over the past 100 years.

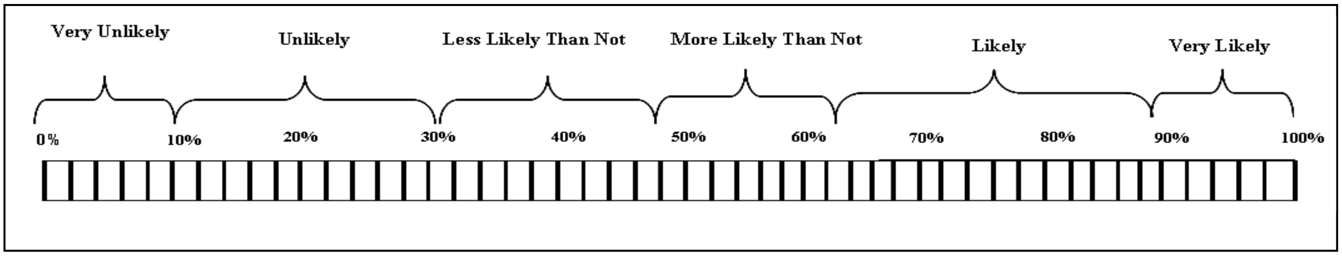


Source: The Bureau of Meteorology, Australia (2008).

1.a. What is the most you think average temperatures might change in 100 years time if nothing is done to prevent climate change? _____°C (a drop down box appeared here with 32 different levels of temperature)

1.b. What is the least you think average temperatures might change in 100 years time if nothing is done to prevent climate change? _____°C (a drop down box appeared here with 32 different levels of temperature)

Figure 3 **Questions asked to estimate subjective policy uncertainty**



3. What do you think are the <u>highest</u> and the <u>lowest</u> chances of the 'Carbon Pollution Reduction Scheme' being successful in slowing down the rise in temperature in Australia if major polluting countries like USA, China and India <u>INTRODUCE</u> a similar scheme?

Highest chance _____%

Lowest Chance _____%

Figure 4 **Distribution of certainty scores across bid levels**

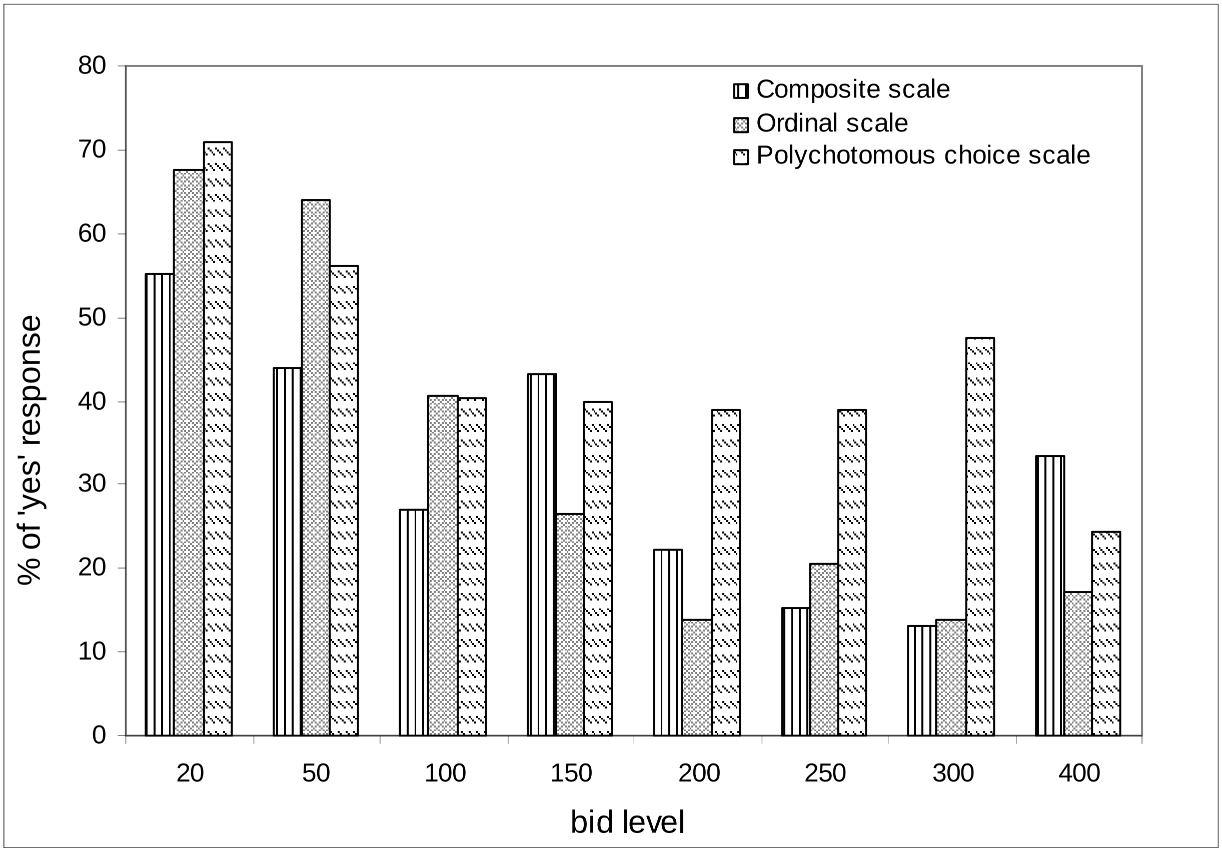


Table 1 Explanatory variables and their expected signs

Variable	Description	Expected sign
Bid	Bid level	–
BidSQ	Square of bid level	+
Knowledge	Knowledge/familiarity with the good in question	+
Subjective Scenario Uncertainty	Uncertainty about future scenario measured by estimating the variance of individual's subjective distribution of future temperature change	–
Subjective Policy Uncertainty	Uncertainty about the effectiveness of the policy measured by estimating the variance of individual's subjective distribution of likelihood of policy success	–
Age(25)	Respondents aged between 25 and 34 years=1, otherwise=0	+/–
Age(35)	Respondents aged between 35 and 44 years=1, otherwise=0	+/–
Age(45)	Respondents aged between 45 and 54 years=1, otherwise=0	+/–
Age(55)	Respondents aged between 55 and 64 years=1, otherwise=0	+/–
Age(65)	Respondents aged 65 years=1, otherwise=0	+/–

Table 2 Non-parametric mean WTP (AUD/Month) for the CPRS and 95% confidence interval (2000 repetitions).

Ordinal Scale	*DC response*	*Mid-scale adjustment*[a]	*High-end adjustment*[b]
Mean WTP (95% confidence intervals)	143 (126 – 160)	123 (107 – 138)	52 (34 – 66)
Composite scale	*DC response*	*Mid-scale adjustment*[c]	*High-end adjustment*[d]
Mean WTP (95% confidence intervals)	153 (114 – 227)	154 (114 – 218)	136 (87 – 233)
Polychotomous Choice format	*PC response*[e]	*Mid-scale adjustment*[f]	*High-end adjustment*[g]
Mean WTP (95% confidence intervals)	230 (171 – 343)	71 (55 – 87)	27 (16 – 38)

<u>Explanatory notes:</u>

[a]DC WTP yes=yes only for certainty ≥5; otherwise=no.
[b]DC WTP yes=yes only for certainty ≥8; otherwise=no.
[c]Fairly sure, highly sure and extremely sure yes =yes; yes to the domain question=yes; otherwise=no.
[d]Highly sure and extremely sure yes =yes; yes to the domain question=yes; otherwise=no.
[e]Maybe, probably and definitely yes (no)= yes (no)
[f]Probably and definitely yes =yes; otherwise=no.
[g]Definitely yes =yes; otherwise=no.

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

Variable	Description	$^{a}S_{Ordinal}$: Model 1	$^{b}S_{PC}$: Model 2	$^{c}S_{Composite}$: Model 3	$^{d}S_{Composite}$: Model 4
Constant [e]		–	–	80*** (5.4)	64*** (6.54)
Bid	Bid level (20, 50, 100, 150, 200, 250, 300, 400A$/month)	0.002 (0.001)	-0.01 (0.07)	-0.06** (0.003)	-0.06* (0.04)
BidSQ	Square of bid level	-3.27e-06 (4.40e-06)	-2.04e-08 (3.30e-06)	1.50E-04** (7.08E-05)	1.63E-04** (8.69E-05)
Knowledge	Respondents have heard of the CPRS (Yes=1, No=0)	0.033 (0.12)	0.26* (0.13)	0.75 (2.03)	1.13 (2.42)
Subjective Scenario Uncertainty	Uncertainty about future scenario measured by estimating the variance of individual's subjective distribution of future temperature change	-0.032 (0.023)	-0.08*** (0.02)	0.12 (0.43)	0.08 (0.51)
Subjective Policy Uncertainty	Uncertainty about the effectiveness of the policy measured by estimating the variance of individual's subjective distribution of likelihood of policy success	-0.002 (0.003)	-0.0002 (0.004)	-0.13** (0.06)	-0.11 (0.07)
Age(25)	Respondents aged between 25 and 34 years=1, otherwise=0	-0.07 (0.19)	0.24 (0.19)	3.09 (3.02)	2.48 (3.29)
Age(35)	Respondents aged between 35 and 44 years=1, otherwise=0	0.19 (0.19)	0.31 (0.18)	6.23** (3.15)	6.90* (3.57)
Age(45)	Respondents aged between 45 and 54 years=1, otherwise=0	0.15 (0.22)	0.48** (0.23)	10.75*** (3.24)	12.53*** (3.96)
Age(55)	Respondents aged between 55 and 64 years=1, otherwise=0	-0.09 (0.29)	0.49* (0.07)	9.88** (4.04)	11.52** (5.14)
Age(65)	Respondents aged 65 years=1, otherwise=0	-0.37 (0.33)	0.14 (0.54)	6.88 (6.32)	6.78 (7.55)
α_1 [e]		-1.8*** (0.28)	0.003 (0.34)	–	–
α_2 [e]		-1.5*** (0.27)	0.76** (0.35)	–	–
α_3 [e]		-1.2*** (0.26)	–	–	–
α_4 [e]		-1.1*** (0.25)	–	–	–

α_5^e	-0.5** (0.25)	–	–	–
α_6^e	-0.11 (0.25)	–	–	–
α_7^e	0.23 (0.25)	–	–	–
α_8^e	0.52** (0.25)	–	–	–
α_9^e	0.63** (0.25)	–	–	–
Model fit statistics				
Log likelihood	-594	-335	-321	-132
LR chi square	13 (df=10, *p*<0.22)	26 (df=10, *p*<0.01)	23 (df=10, *p*<0.05)	19 (df=10, *p*<0.05)
N	306	319	308	308

Explanatory notes:

[a] Dependent variable varies between 1 (absolutely uncertain) and 10 (absolutely certain).
[b] Dependent variable varies between 1 and 3 ('maybe yes/no'=1; 'probably yes/no'=2 and 'definitely yes/no'=3).
[c] Interval dependent variable (Yes to question 19 in Figure1=70–100%; extremely sure=76–99%; highly sure=51–75%; fairly sure=24–50%; highly unsure=2–25%; extremely unsure=0–1%)
[d] Interval dependent variable (Yes to question 19 in Figure1=50–100%; extremely sure=76–99%; highly sure=51–75%; fairly sure=24–50%; highly unsure=2–25%; extremely unsure=0–1%)
[e]Models 1 and 2 do not include a constant term because they include the full set of cut-off points (nine cut-off points in Model 1 and two cut-off points in Model 2). It is not possible to identify both the constant term and all the cut-off points.
[e]$\alpha_{i\ s}$ are threshold values associated with response categories.
– Standard errors of the parameter estimates between brackets.
– ***: p<0.01; **: p<0.05; *: p<0.10.