Framing for incentive compatibility in choice modelling

54th Annual Conference of the Australian Agricultural and Resource Economics Society, Adelaide, Australia, 10 – 12 February 2010

Kasia Mazur[1] and Jeff Bennett[2]

Key words: Choice modelling, Incentive comparability, Provision rule, Non-market valuation, Environment

---

[1] Kasia Mazur is a PhD student in the Crawford School of Economics and Government at the Australian National University and an economist at the Australian Bureau of Agricultural and Resource Economics (ABARE).

[2] Jeff Bennett is a Professor at the Crawford School of Economics and Government at the Australian National University and the Director of the Environmental Economics Research Hub

Abstract

The incentives that motivate respondents to reveal their preferences truthfully have been a long-standing area of research in the non-market valuation literature. A number of studies have been undertaken to investigate incentive compatibility in non-market valuation. Most of these used laboratory environments rather than field surveys (e.g. Carson and Burton, 2008, Harrison, 2007, Lusk and Schroeder, 2004, Racevskis and Lupi, 2008). Only a few studies investigating incentive compatibility have considered multi-attribute public goods with an explicit provision rule in a choice experiment (Carson and Groves, 2007, Collins and Vossler, 2009, Carson and Burton, 2008).

The design of a choice modelling study that avoids strategic behaviour has proven particularly difficult because of multiple choices and difficulties in developing a majority voting provision rule.

This study investigates the impact of the inclusion of a framing statement for incentive compatibility in a field survey choice modelling study. An incentive compatible statement (provision rule) that sets out to respondents the rule relating to when the good under consideration will be provided was employed. The impact of a provision rule across three alternative choice modelling multiple choice questionnaires was tested by comparing results between split samples with and without a provision rule. Four split samples were used to test the impact of a provision rule on preferences across different communities including local/rural residents and distant/urban residents. A choice modelling analysis that involved a conditional logit model and a random parameter model was used to elicit household willingness to pay for improvements in environmental quality in the Hawkesbury-Nepean catchment.

The results of the study show that the inclusion of a provision rule had an effect on preferences in the distant/urban communities. However, the impact of a provision rule in the local/rural community sub-samples was negligible. This study suggests that the impact of a provision rule should be analysed in the context of different community characteristics.

1. Introduction

Social choice theory concentrates on the analysis of collective decision-making processes with the goal of maximising social welfare. Social welfare is a state of human perception that is based on community preferences (Georgiou et al., 2000) where the community is defined as the aggregation of the individuals who constitute that social grouping. Private goods markets work to reveal peoples' preferences. However, preferences for public goods can only indirectly be identified, if at all, through market mechanisms. Therefore, in order for such goods to be supplied at the socially optimal level by governments, peoples' preferences need to be estimated using non-market methods. This can involve people being asked to reveal their preferences.

Economic theory predicts that individuals will make choices to maximise their own utility given their constraints, their knowledge and the possible actions of others. This may involve individuals misrepresenting their preferences when asked about their willingness to pay (WTP) for public goods. Because of the non-rival and non-exclusive[3] characteristics of public goods the dominant Nash equilibrium[4] behaviour of individuals is to "free ride" (Poe et al., 2002). This means that some people may seek to pay less than their WTP for the provision of the good, leading to inefficiency in market provision of a public good. In making stated preference choices, respondents may have a private incentive to behave strategically. This means that respondents intentionally reveal a WTP amount that is different from their true WTP amount. However, the incentive for strategic behaviour could be different from the case of market provision. For example, respondents may behave strategically to influence the provision of the good by overstating their WTP.

Free riding occurs in a situation when respondents state a lower value or do not agree to pay at all for the provision of public goods in the expectation that others will provide enough to cover the cost of provision of this good (Venkatachalam, 2004). According to Samuelson (1954) "it is in the selfish interest of each person to give false signals, to

[3] Individuals receive these goods regardless of their level of contribution and nobody can be excluded from using public goods.
[4] At a Nash equilibrium, expression of preferences of each individual is a best response to the equilibrium strategies of others.

pretend to have less interest in a given collective activity than he really has…" (p 388). Evidence from experiments involving real money shows that individuals usually contribute 40 to 60 percent of the Pareto optimal level (Davis and Holt, 1993). Roberts (1976) found that the larger the number of consumers of a public good the greater is the incentive to free ride due to the smaller contribution made by each individual.

Over-stating of WTP may arise in situations when respondents to stated preference questions think that they don't have to actually pay for a good. By expressing the higher value they have to influence the decision of the provision of that good. While some studies have found evidence of over-statement (e.g. Posavac, 1998, Cummings et al., 1995, Blamey et al., 1999, Champ et al., 1997). Bohm (1971) in one of the first Contingent Valuation (CV) studies of incentive compatibility (IC), used different treatments in eliciting WTP to see a preview of a television showed found no evidence of free-riding or over-pledging.

The understatement and overstatement issue was also tested using direct WTP questioning and the Smith auction format by Bennett (1987). In the Smith auction respondents respond to an iterative sequence of WTP questions. Significant over-statements were observed under direct questioning whilst under-statements were observed in the Smith auctions. Bennett concluded that under-statement was balanced by over-statement behaviour (Bennett, 1987).

The misrepresentation of preferences embodied in strategic behaviour can lead to inefficiency in the allocation of resources. Therefore there is a need for preference revelation techniques to be incentive compatible (IC). Incentive compatibility (IC) in non-market valuation concentrates on the incentives that motivate respondents to reveal their preferences truthfully which means avoiding potential question misinterpretation and strategic behaviours including free riding or over-stating. The design of IC stated preference techniques has been a long-standing area of research in the non-market valuation literature (Hammond, 1979). Most of the IC studies in the non-market valuation have used laboratory environments rather than field surveys (e.g. Carson and Burton, 2008, Harrison, 2007, Lusk and Schroeder, 2004, Racevskis and Lupi, 2008). Only a few studies investigating IC have considered multi-attribute public goods with an explicit provision rule in a choice experiment (Carson and Groves, 2007, Collins and Vossler, 2009, Carson and Burton, 2008).

This study investigates the impact of framing for IC in a field survey choice modelling study. A key component of IC in dichotomous choice CV questions is the specification of a rule relating to when the good under consideration will be provided. In a Choice Modelling (CM) application involving multiple choice tasks, using a 'majority provision rule' is particularly infeasible. Because the single dichotomous (DC) choice format with majority provision rule is the only IC format for stated preference (SP) questioning, the IC of CM applications with multiple alternatives and multiple choice sets is in doubt. Testing the impacts on stated preferences of including a modified form of the majority provision rule is therefore the goal of this paper. A split sample approach is used where the only difference between the sub-sample treatments is the inclusion of a provision rule.

Natural resource management in the Hawkesbury-Nepean catchment was used as the context for the case study. A CM survey was conducted with respondents from two different locations, Hawkesbury-Nepean catchment and Sydney to test for differences in responsiveness to a provision rule between local/rural and distant/urban communities.

This paper is constructed as follows. Section 2 describes the theoretical basis of IC. Section 3 describes the study design and sets out two research hypotheses. Section 4 details the case study catchment. Section 5 sets out the questionnaire design procedure. Section 6 describes the survey logistics. The sample characteristics are set out in Section 7. Section 8 provides an analysis of the results to test the hypothesis. The last section (9) presents some concluding comments.

2 Theoretical background of incentive compatibility

2.1 The theoretical foundation

The theoretical foundation of IC lies in neoclassical social choice theory and mechanism design theory (Hurwicz, 1986, Groves et al., 1987, Varian, 1992). Analysis of collective decision-making processes that maximise social welfare is based on social choice theory pioneered by Kenneth Arrow (1951). In social choice theory, Arrow's impossibility theorem, or Arrow's paradox, demonstrates that it is impossible to design a set of non-dictatorial rules that can convert the preferences of individual members of a group into a consistent set of preferences for the group as a whole. Arrow (1951 p.7) indicated that "once a machinery for making social choices from individual tastes is established, individuals will find it profitable, from a rational point of view, to misrepresent their tastes by their actions, either because such misrepresentation is somehow directly profitable or, more usually, because some other individual will be made so much better off by the first individual's misrepresentation that he could compensate the first individual in such a way that both are better off than if everyone really acted in direct accordance with his tastes."

The Arrow theorem provided a basis to the development of the Gibbard (1973) and Satterthwaite (1975) theorem on the manipulability of voting schemes. The Gibbard-Satterthwaite theorem states that if there are at least three candidates and at least two voters, there is no non-dictatorial voting scheme in which the revelation of true preferences is the dominant strategy. This is because individuals act strategically to avoid wasting their votes on their most preferred candidate if they have a low chance of winning. Hence, they vote for their second best option. In such a case IC preferences are not revealed, therefore only the single DC choice format is IC.

Mechanism design theory provides a framework for analysing institutions, or "allocation mechanisms", with a focus on the incentives that motivate individuals to reveal or not reveal their private information about preferred resources allocations (RSAS, 2007). Hurwicz (1972) first introduced the IC concept into mechanism design theory. He defined a mechanism to be IC if the respondent's dominant strategy is to reveal truthfully his preferences.

Carson and Groves (2007) suggest that people respond truthfully if the outcome of the survey could not affect the respondent in any way. However, the lack of interest in the outcome may lead to careless and meaningless responses – the so-called 'hypothetical bias'. Also if respondents are informed that their answers will not be used in the decision, they may ignore this information and answer the questions consistently with how they think this information may be used because of the effort made in collecting the data (Carson and Groves, 2007). Andreoni (1989) explains strategic behaviour from the 'warm glow' that comes from giving. Also, evidence from experimental economics shows that in some circumstances, people do not always behave in a way that maximises their economic self-interest (Reeson and Nolles, 2009). For example, in some cases agents may have a strategic interest to cooperate voluntarily and tell the truth about their demand for public goods (Mitchell and Carson, 1989). McMillan (1979) found that if the gain from dynamic cooperative planning is higher than the static gains from free riding then people have an incentive to tell the truth. Also Hammond (1979), Bowen (1943) and Dorfman (1969) argue that agents have an incentive to reveal their true preferences if the costs and benefits of producing a public good is shared equally between agents.

The literature shows that even after 50 years of IC research, the problem of strategic behaviour is still an important issue in modern preference reviling techniques such as CM. The design of a CM study that avoids strategic behaviour has proven particularly difficult because of its multiple choices and the difficulties of developing a majority voting provision rule. Therefore, this study investigates the impact of the inclusion of a framing statement for IC in a field survey choice modelling study.

2.2 Addressing incentive compatibility in non-market valuation

A number of studies have investigated different designs of non-market valuation mechanisms to encourage participants to reveal their true preferences.

Survey design

While in some cases respondents try to manipulate outcomes in their own interest by sending false signals, some of the reasons of IC lie in poor CV survey design (Carson and Groves, 2007). Carson and Groves (2007) provide a list of survey design characteristics that can improve the IC of the non-market valuation studies. These include: ensuring a consequential survey design, transparency of task, credibility of policy, relevance of the issue to the respondent, inclusion of information on how the results will be used, plausibility of scenario and that the study and respondents contribution is perceived to have a positive impact.

Carson and Groves (2007) argue that in the case of an inconsequential survey design, economic theory makes no predictions about the accuracy of respondents' stated preferences. A consequential design, however, may lead to strategic behaviour. Nevertheless, most surveys conducted by government or businesses are consequential. Cummings et al. (1997) and Burton et al. (2007) argue that a hypothetical survey design generates unreliable results that are not consistent with economic theory. According to Scott (1965) asking a hypothetical question would lead to a hypothetical answer. Therefore, the respondents should view their choices as potentially influencing the policy decisions (Carson and Groves, 2007). These authors also suggest reminding the respondents about their budget constraint which would reduce the incentive to overstate their true WTP. The understatement of true values could be reduced by explaining that lower values could result in a situation that the good may not be provided.

The transparency of the task is also recognised as an important factor for improving the IC of the non-market valuation studies. In general, researchers assume that respondents answer the question being asked. This assumption may be dubious if respondents do not understand the question (Sudman et al., 1996). The poor design of a stated preference questionnaire can result in interpretation of the question in a number of different ways. Therefore, the clarity and type of language used in questionnaires is very important. Moreover, the issue of preference uncertainty can arise if respondents are presented with unfamiliar goods. “This can lead to a high variance in WTP estimates or systematically biased estimates” (Bateman et al., 2008a p.128). Carson and Groves (2007) agree that while familiarity with a good can influence WTP, it does not influence the IC properties of the question. Another problem is when respondents are not sensitive to the question asked, because critical details are missing or too many details are presented (Fischhoff et al., 1993).

The IC of a survey may also be affected by a lack of credibility. The credibility of the policy, organisation and researchers involved in the study need to be established in the questionnaire. The respondent needs to be provided with true and consistent information.

The scenario presented to respondents needs to be believable. In a situation where too high or too low a cost for the provision of the good is presented, respondents may substitute the presented costs with one that they regard as being more realistic (Carson and Groves, 2007). Another example of lack of plausibility is when an improvement is presented in the questionnaire that does not seem to be fully achievable in the eyes of the respondents. They may discount the stated improvement to one that they selected as being more realistic (Fischhoff et al., 1993, Bennett et al., 1998, Carson and Mitchell, 1995, Smith and Osborne, 1996). Carson and Mitchell (1995) give the example of Kahneman and Knetsch’s (1992) study where the description of the goods is argued to be inadequate and their provision implausible, making it difficult for respondents to identify the relevant scope. Similar criticism was also directed at studies by Smith (1992) and Harrison (1992).

Mitchell and Carson (1989) argue that a larger number of people be involved in a CV survey to give the impression that their individual preferences would not influence the overall outcome.

Provision rule

Attention has been given to the inclusion of an explicit provision rule in the CV non-market valuation literature. A provision rule provides a connection between survey choices and actual outcomes (Hoehn and Randall, 1987). A lack of a provision rule can create ambiguity relating to the outcomes of respondents' choices (Harrison, 2007). This is because respondents do not know how their votes are taken into account to decide whether the good should be provided. Moreover, respondents can have different perceptions of what the actual provision rule is and, by not knowing how their utilities are affected by the outcome of the vote, it could be difficult for respondents to understand why they should answer truthfully (Polomé, 2003). If the respondents are not given with an explicit provision rule, an assumption may be that the option receiving the greatest support would be implemented which is a plurality voting rule (Taylor et al., 2007). This may not be the case. Therefore to avoid misguided assumptions, the inclusion of a provision rule can add greater certainty to the outcome and credibility to the study.

The most common provision rules used are plurality and majority voting rules (Arrow et al., 1993). In plurality voting rule applications, the option that receives the greatest number of votes is implemented for the whole group. In the majority voting rule the option chosen by more than half of the participants is implemented. Quite often these two terms are used interchangeably without a clear distinction between majority and plurality provision rules (Lin et al., 2003).

Despite the reduced ambiguity from the inclusion of a provision rule, the IC of the selection process of choosing one option from more than two alternatives is affected by the strategic behaviour problem specified by the Gibbard-Satterthwaite theorem. Only the application of a provision rule to the two-option choice format or to a voting process where there are two wining options out of three could potentially yield an IC outcome. In such formats there is no possibility to behave strategically. Moreover voters think that their behaviour will have some impact on the outcome, and they know exactly how their choices will have an impact on the decision regarding which alternative to implement. Therefore, respondents cannot do better than to vote honestly (Harrison, 2007).

Elicitation format

As predicted by the Gibbard-Satterthwaite theorem, a number of studies have found that only the single binary choice elicitation format of CV is a potential IC mechanism. First Clarke (1971) and Groves (1973) showed that a binary decision is a mechanism where a dominant strategy is to reveal truthfully people's preferences (assuming no income effects on the demand for public goods) (RSAS, 2007). Following these findings, the single binary choice (DC) elicitation format was introduced by Bishop and Heberlein (1979) to the CV method and was recommended by the NOAA Panel (Arrow et al., 1993). In this format, each respondent is asked a single question whether they are willing to pay a pre-specified amount for a specific good or not (known as a "take it or leave it" offer). In order to estimate the distribution of WTP the amount of money presented to the respondents varies between questionnaires (Boyle et al., 1996). In this format respondents cannot do better than express their real preferences.

Despite the IC character of the single binary elicitation format, it has been criticised for being statistically inefficient and prone to starting point bias (Ready et al., 1996). To address the inefficiency of the single bounded choice format Hanemann et al., (1991) proposed the 'double-bound' (DB) elicitation method. However, even though this format yields some efficiency gains, the answers to the second question are not IC (McFadden, 1994, Cameron and Quiggin, 1994, Bateman et al., 2001). This is because answers to the second question can be influenced by the first choice (Carson et al., 1994). Moreover, respondents may assume that the actual cost could be the weighted average of the two prices from both questions (Carson and Groves, 2007). There are other formats: the multiple binary choice format that involves a sequence of paired comparisons of the *status quo* and the alternative option, single multiple choice format; and the repeated multiple choice elicitation format. Even though these formats are statistically efficient they are potentially prone to strategic bias.

2.3 Choice Modelling and incentive compatibility

CM is a relatively recently emerged non-market valuation method. The main advantages of this method over CV are: the estimation of marginal values of a number of attributes and policy options, the facilitation of benefit transfer, the reduction in some of the biases (e.g. “yea-saying”) and the possibility of testing for internal consistency (Collins and Vossler, 2009). However, CM, unlike the single binary discrete choice format of CV, does not have IC properties: the elicitation format of CM in which choices are made over more than two alternatives across a series of questions potentially suffers from strategic bias.

In CM, respondents are presented with a sequence of choice sets comprising a number of alternatives. The CM choice sets usually include three to four alternatives (usually including the *status quo*) described by different levels of attributes including a cost attribute. If respondents are presented with more than two options they may choose the second best option rather then their most preferred if they expect that their first choice does not have a chance of winning.

The other property of CM that reduces IC is the multiple question format. A sequence of questions creates uncertainty as to how respondents to the survey treated the information across the different grouped alternatives (i.e. independently or not). In the multiple choice format, respondents are presented with different costs for the same good or the same cost for different goods at varying points in the questionnaire. The variability of the costs between similar options and the variability of the good for the same cost can confuse respondents and affect the credibility of the study (Carson and Groves, 2007). When respondents are presented with a range of different prices for the same good they may assume that the true cost is somewhere in the middle. Respondents may also believe that there is no risk of not getting the good. Then they may have a strategic interest to always choose the minimum cost option even if their real WTP is higher (Carson and Groves, 2007, Bateman et al., 2008b). If there is a risk of not getting the good, the respondent may first select (from all choices) the best change option at a lower cost and reject other options even if they would prefer the alternative change options over the *status quo* in other choice sets. That is, in a multiple choice set format, respondents are aware of available combinations of goods and prices on offer. They may therefore look for relatively ‘good deals’ compared to other options on offer over the whole questionnaire rather than in one choice set (Bateman et al., 2008b).

Some authors argue that multiple response formats are preferable because of repetition and learning experience, which are important in obtaining consistent and stable preferences (Bateman et al., 2008a). Mitchell and Carson (1989) argued that the indirect non-market valuation methods which are based on observing how respondents make trade-offs between different resource allocations are less prone to strategic behaviour than the methods that directly ask the respondents to state their true preferences. Therefore, the CM method, because it indirectly asks respondents about their WTP for different resource allocations and due to its choice complexity, is considered by Mitchell and Carson (1989) to be more difficult for respondents to develop strategic behaviour than in the open-ended CV method. Carson and Groves (2007) argue that through the complexity of CM, with a large sample size and careful questionnaire design, strategic behaviour problems in CM can be managed.

2.4 Testing for incentive compatibility in choice modelling

Some laboratory experiments have tested for IC of CM by investigating different elicitation formats and provision rules. However, field testing for IC in CM is largely unexplored (Carson and Burton, 2008).

Carson and Burton (2008) conducted a laboratory experiment that investigated the IC of different elicitation formats in CM. The study compared alternative questioning formats such as a single binary choice, six binary choices, single multiple choice (a single choice set with more than two alternatives) and repeated multiple choices (more than one choice set with more than two alternatives). The study found that the single binary choice format is demand revealing. The comparison of the results across these four different elicitation formats showed that the proportion of respondents choosing each option in the repeated binary choice format was not significantly different from the single binary choice format. There were also no differences in the distributions of choices between the single multiple choice format and repeated multiple choice treatment. However, the total rate of non-demand revelation (e.g. *status quo* bias or the respondent does not choose the utility maximising change option) in the repeated single binary choice treatment was significantly lower than in the multiple choice treatment. The main problem identified with the repeated choice format was an increased rate of *status quo* bias. Similarly, Racevskis and Lupi (2008) found that while the multiple choice response format improved statistical efficiency, the WTP estimates were not consistent with the potentially IC single binary choice format.

Explicit provision rules applied to public goods have had limited applications in CM (e.g. Carson and Groves 2007). Some studies (e.g. Collins and Vossler, 2009, Scheufele and Bennett, 2010) used provision rules in multiple choice formats. For example, a recent CM field survey study conducted by Scheufele and Bennett, (2010) employed a single binary elicitation format with a majority vote provision rule as the baseline to investigate the effects of sequential binary DC elicitation formats. The study presented evidence for effects caused by institutional learning and either strategic behaviour or value learning in repeated choice questions. However, the study did not find any evidence of strategic behaviour caused by sole awareness of presence of multiple choices in a CM questionnaire.

Collins and Vossler (2009) in their laboratory CM study found that there was no statistically significant difference between the standard DC format questions and the three-option trichotomous choice (TC) elicitation format, under the plurality voting rule. The authors also used a random selection rule for which the percent in favour of some alternative was assumed to be the probability of the implementation of the alternative.

After all votes were collected, one ballot was selected randomly to determine the option for implementation. In this treatment, DC and TC were significantly different.

Lusk and Schroeder (2004) tested for IC in CM by comparing hypothetical and real treatments using a private good (different quality beef steaks). The result of the study showed that respondents to the hypothetical treatment choose the change option more frequently. Moreover, the WTP from this treatment for the five rib-eye steaks versus not having the steak at all was also higher than from the non-hypothetical setting. However it needs to be remembered that the real treatment may have experienced understatement.

3. *Hypotheses and study design*

In the research reported here, the impact of a provision rule in a three alternative CM multi choice questionnaire is explored in the context of an improvement in the environmental quality of the Hawkesbury-Nepean catchment. Half of the questionnaires used for this test included the framing statement: "Only options that are chosen by more than 50 precent of the people surveyed will be considered further for implementation by the Catchment Management Authority". This provision rule refers to each choice set in the questionnaire. Unlike in previous studies, the outcome are determined based on people's choices and further analysis of the management scenarios, rather than being based on random selection from the set of most preferred options. It is argued that the inclusion of a statement that provides more information about further consequences of peoples' decisions should make a difference to WTP estimates as it provides a clearer and therefore stronger impression of respondents' contributions to the overall outcome of the policy. With additional information about how the outcome of the survey will be used in the decision process, comes greater realism of potential actions to take place. This consequently should generate different results.

The results from questionnaires with and without a provision rule were compared based on responses from the local/rural community (Hawkesbury-Nepean) and a distant/urban

(Sydney) community. To test for the impact of the provision rule on the responses the following hypotheses were formulated:

***HA**: In Sydney and the Hawkesbury-Nepean sub-samples there are no differences between attribute parameter (β) estimates obtained from the questionnaires with a provision rule (β_{PR}) and without a provision rule (β_{NPR}).*

The null hypothesis:

HA_0: $\beta_{NPR} = \beta_{PR}$

The alternative hypothesis:

HA_1: $\beta_{NPR} \neq \beta_{PR}$

The null hypothesis (HA_0) implies that the β estimates obtained from the questionnaires with and without a provision rule are the same. The alternative hypothesis (HA_1) states that the βs differ significantly. Our prior expectation is that the HA_1 will not be rejected.

***HB**: In both Sydney and the Hawkesbury-Nepean sub-samples there are no differences between WTP estimates obtained from the questionnaires with a provision rule (WTP_{PR}) and without a provision rule (WTP_{NPR}).*

The null hypothesis:

HB_0: $WTP_{NPR} = WTP_{PR}$

The alternative hypothesis:

HB_1: $WTP_{NPR} \neq WTP_{PR}$

The null hypothesis (HB_0) implies that WTP estimates for improvements in environmental quality are the same with (WTP_{PR}) and without (WTP_{NPR}) a provision rule in a standard three option CM questionnaire design. The alternative hypothesis (HB_1) states that the WTP_{PR} and WTP_{NPR} differ significantly. Our prior expectation is that the HB_1 will not be rejected.

4. Case study

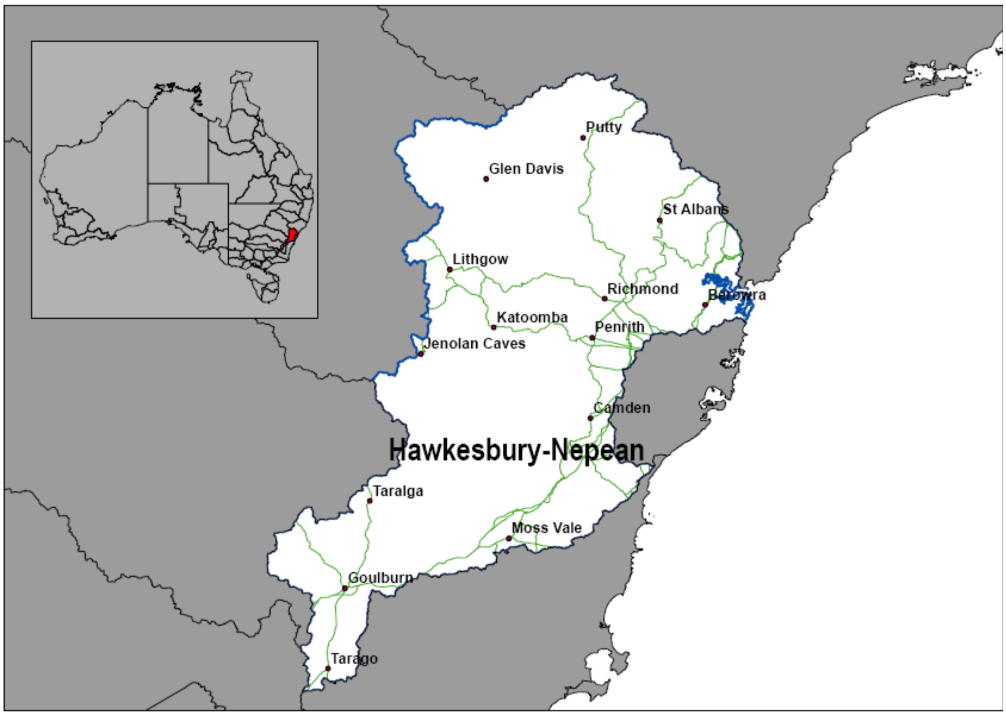
The Hawkesbury-Nepean catchment (see Figure 1) was chosen as a case study for this study. The Hawkesbury-Nepean catchment covers 22,000km^2 and 1,000,000 people live in this catchment. More than 50 percent of the catchment area is in the National Parks. Agriculture occupies about 30 percent with more than half of this agricultural area used for grazing. Native vegetation covers about 70 percent of the catchment area. About 20 percent is urbanised.

The main environmental issues of the Hawkesbury-Nepean catchment include declining biodiversity, loss of native vegetation and reduced water quality. The greatest area of native vegetation in good quality in the Hawkesbury-Nepean (of the total) area of native vegetation in good quality covers 50 percent of the catchment.

Water quality has declined in 75 percent of the total waterways in the catchment. Currently only 15 percent of the waterways in the Hawkesbury-Nepean catchment are of good enough quality for drinking, swimming and fishing. The amount of NS in the Hawkesbury-Nepean catchment is 3,000 from which 233 is endangered and vulnerable.

NRM actions such as planting more trees, protecting existing vegetation, fencing and revegetating river banks and wetlands, pest and weed control are just some of the actions that can improve environmental quality in the catchments. More information about each catchment's characteristics is included in Mazur and Bennett (2009).

Figure 1. The Hawkesbury-Nepean catchment.



5. Questionnaire development

The questionnaire was designed to ensure respondents considered it to be consequential. Respondents were ensured of the importance of the study and its further use in the policy making process. The credibility of the organisation conducting the study was clearly displayed. The plausibility of the scenarios and transparency of the task were tested during focus group discussions and verified with the specialists in the local area. The impact of inclusion of a provision rule was tested using two split samples in each of the areas (the Hawkesbury-Nepean catchment and Sydney). These two different communities were presented with questionnaires with and without a provision rule. Table 1 presents the research design.

Table 1. Research design and the study sub-samples

Questionnaires / Sub-sample location	With provision rule PR	Without provision rule NPR
Hawkesbury-Nepean	Local /rural	Local /rural
Sydney	Distant/ urban	Distant/ urban

The attributes and their current and potential future levels used in the questionnaire were developed in consultation with local specialists and during focus group discussions (see Mazur and Bennett, 2008b).

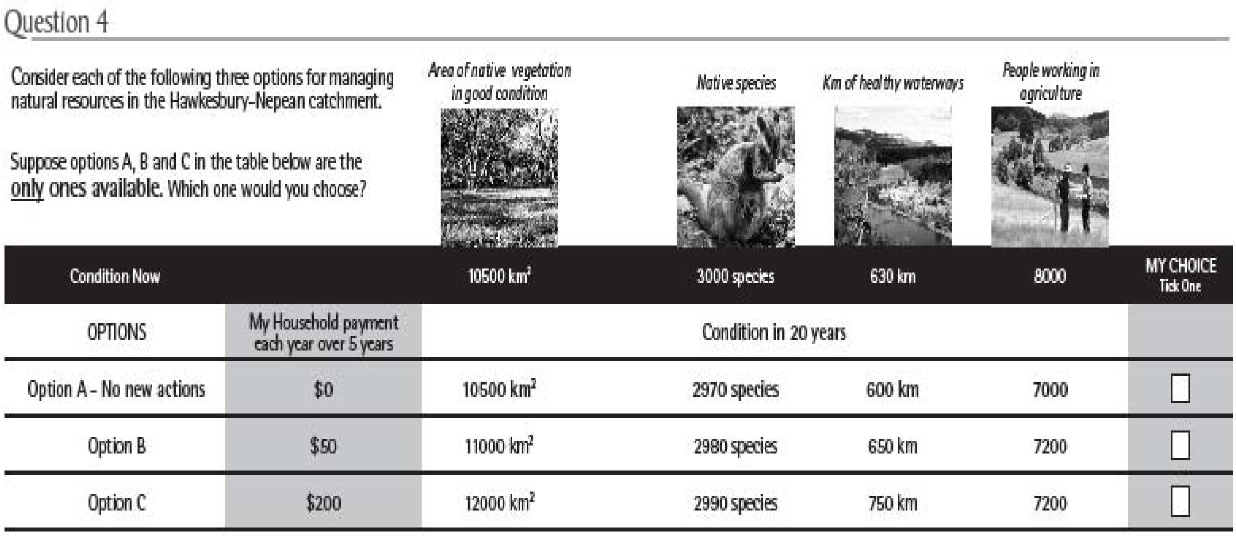
Three attributes that represent the main potential environmental benefits derived from NRM investments in the Hawkesbury-Nepean catchment were used: area of native vegetation in good condition (NV), kilometres of healthy waterways (HW), and number of native species (NS). One additional attribute - people working in agriculture (PA) - was chosen to capture the social consequences of changes in NRM actions. The fifth attribute was a monetary cost. The annual payment to be made by respondents from new NRM actions was specified to continue for five years. The payment vehicle was described as a mixture of increased taxes, council rates, prices for goods and services and recreational charges. Three different levels of each attribute in each questionnaire type were determined and used in an experimental design to structure the choice set used in the questionnaires. The ranges of the attribute levels are included in Table 2.

Table 2. Attributes and their levels

	Cost	NV	NS	HW	PA
Condition Now		**10500**	**3000**	**630**	**800**
Status Quo	**$0**	**10500**	**2970**	**600**	**7000**
Outcomes in 20 years time	$50	11000	2980	650	7100
	$200	11500	2990	700	7200
	$300	12000	3000	750	7300

The levels of each attribute across the predicted range were used in an orthogonal design that produced 25 alternative NRM options. These alternatives were randomly blocked into five different versions, each with five choice sets. This resulted in ten different versions of the questionnaire. Two change options and a *status quo* option were included in each choice set. An example choice set is presented in Figure 2.

Figure 2. Example of a choice set for the Hawkesbury-Nepean catchment questionnaire


Question 4

Consider each of the following three options for managing natural resources in the Hawkesbury-Nepean catchment.

Suppose options A, B and C in the table below are the **only ones available.** Which one would you choose?

		Area of native vegetation in good condition	Native species	Km of healthy waterways	People working in agriculture	
Condition Now		10500 km^2	3000 species	630 km	8000	MY CHOICE Tick One
OPTIONS	My Household payment each year over 5 years	Condition in 20 years				
Option A - No new actions	$0	10500 km^2	2970 species	600 km	7000	☐
Option B	$50	11000 km^2	2980 species	650 km	7200	☐
Option C	$200	12000 km^2	2990 species	750 km	7200	☐

6. Survey Logistics

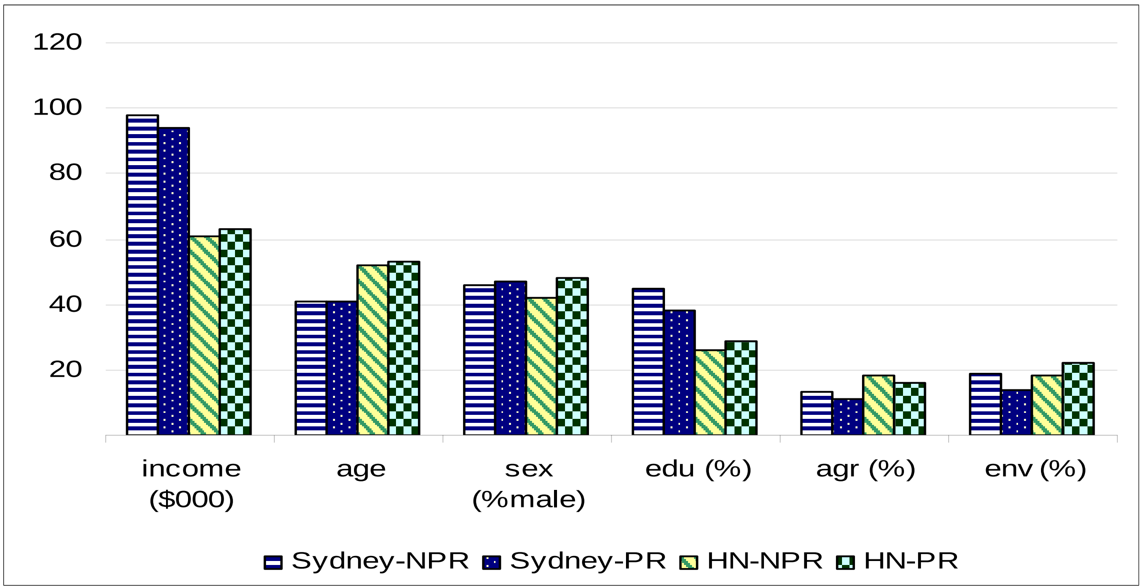
A drop-off/pick-up approach was used for the survey. Questionnaires were distributed in two main towns in the Hawkesbury-Nepean catchment (Goulburn and Moss Vale) and in Sydney. Geographically stratified random sampling was applied to choose the households to ensure a representation of the NSW population in terms of gender, age, income etc.[5]

[5] A more detailed description about the sampling procedure is included in Mazur and Bennett MAZUR, K. & BENNETT, J. (2009) A Choice Modelling Survey of Community Attitudes for Improvements in Environmental Quality in NSW Catchments. *EERH Report No 10.* Canberra, Australian National University..

7. Sample characteristics

The socio-economic characteristics of the sub-samples are presented in Figure 3.

Figure 3. Descriptive statistics: Hawkesbury-Nepean catchment and Sydney sub-samples.



Note: income- annual household income ($000), edu – represents respondents with a tertiary degree and above, agr-represents association with agricultural industry of the respondents and their close family, env-represents association with environmental organisations of the respondents and their close family. Sydney S-NPR - the questionnaire without a provision rule tested in Sydney, Sydney S-PR - the questionnaire with a provision rule tested in Sydney, HN-NPR - the questionnaire without a provision rule tested in the Hawkesbury-Nepean catchment, HN-PR - the questionnaire with a provision rule tested in the Hawkesbury-Nepean catchment.

A comparison of the socio-economic characteristics of the sub-samples with ABS (2006) Census data was undertaken. The χ^2 test was used to compare the distribution of age, income and education level between the sub-samples against the Census data. There were no significant differences in age between the ABS Census data and all the sub-samples (Sydney: NPR χ^2 =8.97, and PR χ^2=12.21, Hawkesbury-Nepean: NPR χ^2 =23.05, and PR χ^2=22.09). There were also no significant differences in age distribution between the sub-samples within each area (Sydney χ^2=11.68, Hawkesbury-Nepean χ^2 =21.28) and the Census.

No significant differences in household size between the samples and the ABS census data were found. However, the distribution of educational level was significantly different from the Census for all the sub-samples[6] (Sydney: NPR χ^2 =80.99, and PR χ^2=39.24, Hawkesbury-Nepean: NPR χ^2 =150.92, and PR χ^2=195.25). The proportion of people with a tertiary degree was higher in the study sub-samples than recorded by the 2006 Census. However, there were no significant differences in education level between the sub-samples within each area (Sydney χ^2=9.59 Hawkesbury-Nepean χ^2 =3.44).

The income ranges presented in the questionnaire were consistent with ABS household ranges presented in the 2006 Census. Significant differences[7] between the sub-samples and Census income were recorded in the Sydney sub-samples (NPR: χ^2=22.41 and PR: χ^2=24.15). Also significant differences from Census were observed for the data from the Hawkesbury-Nepean sub-sample with the provision rule (: χ^2=22.65) but not for the sub-sample without a provision rule (: χ^2=18.50). There were no significant differences between the sub-samples in each of the sampled area (Sydney χ^2=9.59 Hawkesbury-Nepean χ^2 =14.23).

8. The econometric models

Conditional logit (CL) and Random Parameters (RPL) with panel specification models were estimated using Limdep (version 4.0) software. [8] The CL model provides the probability of an individual i choosing alternative j as a function of attributes that describe each alternative:

[6] The critical χ^2 = 12.59 at 0.05 level d.f. 6

[7] The critical χ^2 = 21.06 at 0.05 level d.f. 12

[8] Random Parameters model without panel specification was also conducted. As no significant improvement in model fit and no significant differences in WTP in comparison to the CL model were identified the results of this model were not recorded in the following paper.

$$P_{ij} = \frac{\exp\left(\mu\beta_i x_{ij}\right)}{\sum_{q=1}^{J}\exp\left(\mu\beta_i x_{iq}\right)} \tag{1}$$

where x_{ij} is a vector of attributes j and individual characteristics i, β is a vector of parameters, j and q are the vectors of attributes describing different options, and μ is a scale parameter, which is usually normalized to one. Due to the fact that the scale factor μ and parameter vector β are confounded in the MNL model it is difficult to observe differences in the estimated parameters and scale factors. In order to compare β parameters across the two data sets (with and without a provision rule) the scale factor μ needs to be isolated. The scale parameter is inversely proportional to the standard deviation of the error distribution $V(e_{ij})= \pi^2/6\sigma^2$ where σ is the standard deviation of error distribution (Ben-Akiva and Lerman, 1985). In order to be able to compare the parameters between two data sets the Swait and Louviere (1993) test is conducted. This test involves two stages. The first stage tests the null hypothesis that the parameters are equal while permitting the scale factors to vary between the data sets.

HA1: $\beta_{NPR} = \beta_{PR} = \beta$

This test uses the likelihood ratio (LR) test:

$$LR = -2[LL_{pooled} - (LL_{NPR} + LL_{PR}) \tag{2}$$

where LL_{NPR} and LL_{PR} are the log-likelihoods corresponding to each model. LL_{pooled} is the log-likelihood value of the combined data set of NPR and PR. The scale parameter for one of these data sets should be rescaled (in this case NPR). The correct value of the relative scale parameter is found by conducting a grid search using different values of the scale parameter. The scale parameter for which the log-likelihood of the pooled model (PR+NPR) is optimised is chosen. The test statistic is χ^2 distributed with (K+1) degrees of freedom, where K is the number of common parameters in each of the models (pooled, NPR and PR) and the additional degree of freedom occurs because μ varies under the alternative hypothesis (Swait and Louviere, 1993). If the first hypothesis cannot be rejected the second stage of the test needs to be conducted. This involves a test of the null hypothesis of equal scale factors.

HA2: $\mu_{NPR} = \mu_{PR} = \mu$

The test statistic is:

$$LR = -2*(LL - LL_{pooled}) \quad (3)$$

Where LL is the log likelihood value for the model using the combined data set in which the scale factors of the two sub-samples are assumed to be equal, and LL_{pooled} is as previously defined. The test statistic is χ^2 distributed with one degree of freedom. Only if both hypothesis HA1 and HA2 are not rejected at a given confidence level can we retain the hypothesis that:

HA3: $\beta_{NPR} = \beta_{PR}$ and $\mu_{NPR} = \mu_{PR}$

If only HA1 is not rejected μ_{PR} can be interpreted as a measure of the heterogeneity or homogeneity of the error variance of the two data sets (Swait and Louviere, 1993). "If HA1 is rejected the estimated value is simply an average multiplier that optimally scales the data of sample PR to offset the imposition of the β parameter equality assumption" (Swait and Louviere, 1993 p 309).

The standard assumption of the CL model is that the ε term is an independently and identically distributed (IID) Gumbel random variable (McFadden, 1974). The irrelevance of independent alternatives (IIA) assumption is derived from the IID. According to the IIA assumption, the inclusion of an irrelevant alternative in a choice set has no impact on the probability of the selection of a particular alternative by the respondent. This assumption can be violated and in such cases a different assumption regarding the stochastic term needs to be made, necessitating the use of alternative models including random parameter logit (RPL).

RPL accounts for observed and unobserved preference heterogeneity across respondents and relaxes the IIA "despite the presence of the IID assumption for the random components ε_{ij} of the alternatives" (Louviere et al., 2000). This means that the RPL model separates IIA from IID and allows cross-correlation amongst alternatives in the estimated models (Hensher and Reyes, 2000). The form of RPL is described below drawing from (Train, 1998) and Greene (2007).

A specification for the RPL model is the same as for the conditional logit model (see equation 1, this time including α_{ij} and multiple choice situation *t*) except that coefficient β_i varies in the population.

$$\mathrm{L_{ijt}}\left(\beta\right)=\frac{\exp\left(\alpha_{ij}+\beta_{\mathrm{i}}\ \mathrm{x_{jit}}\right)}{\sum_{\mathrm{q=1}}^{\mathrm{J}}\exp(\alpha_{iq}+\beta_{\mathrm{q}}\ \mathrm{x_{qit}})} \qquad (4)$$

The variance in β_i induces correlation in utility. Therefore the coefficient vector β_i of each respondent *i* can be expressed as:

$$\beta_{\mathrm{j}} = \beta+\sigma v_{\mathrm{i}} \qquad (5)$$

Where β_i is the population mean β, v_i is the individual specific heterogeneity, with mean zero and standard deviation one. σ is the standard deviation of β_i around the mean β (Greene, 2007). σ accommodates the presence of unobservable preference heterogeneity in the sampled pollution (Hensher et al., 2005) .

In the RPL model it is assumed that individual preferences β_i vary across the population with density $\mathrm{f}\,(\beta/\theta)$ where θ are the parameters of this distribution (representing the mean and standard deviation of preferences) (Train, 1998). Hence, the probability that the individual *i* chooses the alternative *j* can be expressed as the integral of the conditional probability (equation 6) over all possible values of β weighted by the density of β (Train, 1998).

$$P_{ijt}\left(\theta\right)=\int L_{ijt}\left(\beta\right) f\,(\beta/\theta)\, d\beta \qquad (6)$$

In the RPL model, choice probability cannot be calculated exactly. Instead, the probability is approximated through simulation. "For a given value of the parameter θ a value of β is drawn from its distribution" (Train, 1998 p.5).

In CM, respondents are usually presented with a sequence of choices. However, standard CL and RPL models treat each choice set as a separate observation and do not account for error correlation between choices made by one individual. The panel specification of RPL model calculates the conditional probability at the level of each individual respondent by accounting for error correlations between repeated choices of each individual. Therefore, the assumption that the choices made by the same respondent are independent no longer holds. Therefore the log likelihood becomes:

$$LL_{ijt}(\beta)=\prod_{t}\int L_{ijt}\left(\beta\right) f\,(\beta/\theta)\, d\beta \qquad (7)$$

9. Results

In total, 1121 responses producing 56,055 choice observations were collected from the four sub-samples. In about 32 percent of the choice sets, the *status quo* option was chosen. In both the Sydney sub-samples (with and without a provision rule) 35 percent of choices were the *status quo* option. In the Hawkesbury-Nepean sub-samples the *status quo* option was chosen in 30 percent of choice stets in the questionnaires with a provision rule and in 26 percent of choice sets in the questionnaires without a provision rule. There is however no significant difference (at the 5 percent level) between the sub-samples in this regard.

The CL model used in this choice modelling analysis was estimated using Limdep (version 4.0) software. The equations for this model are:

$$U(A)=\beta_1 costs+\beta_2 NV+\beta_3 NS+\beta_4 HW+\beta_5 PA$$

$$U(B)=ASC+\beta_1 costs+\beta_2 NV+\beta_3 NS+\beta_4 HW+\beta_5 PA+ASC*AGE+ASC*EDU+ ASC*INC+ASC*GEN+ASC*CHIL+ASC*ENV+ASC*AGR \quad (8)$$

$$U(C)=ASC+\beta_1 costs+\beta_2 NV+\beta_3 NS+\beta_4 HW+\beta_5 PA+ASC*AGE+ASC*EDU+ ASC*INC+ASC*GEN+ASC*CHIL+ASC*ENV+ASC*AGR$$

where:

A - *Status quo* option

B and C - change options

β - estimated coefficients

ASC - alternative specific constant

The attributes are described in Table 3.

Table 3. Variables used in the Choice Models

ASC	alternative specific constant
NV	km^2 of native vegetation in good condition
NS	number of native species
HW	km of healthy waterways
PA	number of people working in agriculture
COST	cost of choice alternative ($ pa per household over 5 years)
ASCAGE	respondent age x ASC
ASCEDU	respondent education status (1=with tertiary degree) x ASC
ASCINCOME	respondent household income ($000) x ASC
ASCGENDER	respondent gender (1= female) x ASC
ASCCHILDERN	respondent children (1= with children) x ASC
ASCENV	respondent association with environmental organisation (1=associated) x ASC
ASCAGR	respondent association with agricultural industry (1=associated) x ASC

The *status quo* level was treated as the constant base for each attribute. Therefore, the differences in choice probabilities between the *status quo* and a specific option with different attribute levels were expressed in the estimated model parameters. All parameters used in the models are generic. In order to account for preference heterogeneity, models with socio-economic and attitudinal variables were estimated. Socio-economic characteristics such as age, education, income, gender, number of children, association with agricultural industry and association with environmental organisation were included in the CL by interacting them with the ASC.

A Hausman test was conducted in order to test for any violation of the IIA. This test showed that there was no breach of the IIA assumption (at the 5% significance level) in all of the CL models. However to account for observed and unobserved preference heterogeneity RPL models were estimated.

In order to estimate the RPL model, simulations were undertaken to determine the appropriate distributions for the random variables. Normal distributions were used for the final models. The cost attribute coefficient was treated as a fixed parameter whilst other coefficients were allowed to vary. Estimates for the RPL models were derived using 500 Halton draws (Train, 2000). The attributes that consistently showed an insignificant standard deviation were treated as non-random and the model was re-estimated. The best model in terms of model fit and significance of the attributes was chosen.

In order to control for unobserved heterogeneity across the choices from the same individual, a model with panel specifications was used. The results from the choice models for each sample are presented in Tables 4 to 7. The results indicate a good overall model performance. A better model fit (higher pseudo-R^2) was obtained in the RPL by accounting for error correlations between repeated choices of each individual. The pseudo R^2 for most of the CL models was around the ten percent level but for RPL was twenty percent.

The ASC (coded as 1 for the change options) was negative and significant for all the models. The results show that for all the split samples, the signs of the model parameters are in accordance with *a priori* expectations. All the significant environmental attribute parameter coefficients have positive values implying that those NRM scenarios which result in higher levels of any single attribute are preferred. The cost coefficient was negative and significant for all the models. The significance of the attributes varied between different community types and models.

The significance of the attributes obtained from the RPL models were the same as obtained from the CL models in most of the sub-samples. In both Hawkesbury-Nepean sub-samples (NPR and PR) the NS, HW and PA attributes were significant at the ten percent level and NV was insignificant. NS attribute even significant in the CL model for PR sub-sample it become insignificant in the RPL model. In both Sydney sub-samples (NPR and PR) HW and NS were significant, NV was only significant in the NPR sub-sample and PA was insignificant in both sub-samples. Similarly for the Hawkesbury-Nepean PR sub-sample the NS attribute become insignificant in the RPL model.

The respondents with a higher education level were more likely to choose the change option in all the sub-samples. Income also had a positive and significant effect on peoples' choices but only in the sub-samples without a provision rule. Also people from Sydney who were associated with an environmental organisation were more likely to choose the change options in both sub-samples (PR and NPR).

Table 4. Choice models without provision rule (NPR): Sydney sub-sample

	CL	RPL
Random parameters		
HW		.0071*** (.0020)
Random parameter standard deviations		
HW		.0162*** (.0016)
ASC	-3.955*** (.7050)	-3.9550*** (.9024)
COST	-.0052*** (.0005)	-.0081*** (.0008)
NV	.0003*** (.0001)	.0017*** (.0002)
NS	.0269*** (.0067)	.0272*** (.0075)
HW	.0057*** (.0013)	
PA	.0009 (.0007)	.0020** (.0009)
ASCAGE	.0189*** (.0061)	.0195** (.0067)
ASCEDU	.0777* (.0418)	.0779 (.0516)
ASCNCOME	.0119*** (.0018)	.7811D-05*** (.2061D-05)
ASCGENDER	.5283*** (.1701)	.5283** (.1997)
ASCCHILDREN	-.1987 (.1990)	-.1987 (.2322)
ASCENV	.5477*** (.2320)	.5477** (.2568)
ASCAGR	.4834** (.2529)	.4834 (.3208)
AIC	1.95713	1.76165
BIC	1.95761	1.83777
HQIC	1.98416	1.79076
Pseudo R2	0.12079	.2127340
D.F.O	13	14
Log likelihood	-847.1601	-760.2472
Chi^2	232.78460	41.8660
Observations	879	879

Notes: Significance levels indicated by: * 0.1, **0.05, ***0.01, standard errors in brackets .

Table 5. Choice models with provision rule (PR): Sydney sub-sample

	CL	RPL
Random parameters		
NS		*-.0239 (.0150)*
Random parameter standard deviations		
NS		*.1440 (.0148)*
Non-random parameters		
ASC	-1.9670*** (.5689)	-2.1681* (1.188)
COST	-.0055*** (.0005)	-.0082*** (.0007)
NV	.9464D-04 (.0001)	.0002 (.0002)
NS	.0124* (.0067)	
HW	.0038*** (.0013)	.0048*** (.0016)
PA	.0003 (.0007)	.0006 (.0008)
ASCAGE	.0168*** (.0058)	.0320** (.0134)
ASCEDU	.1134*** (.0318)	.1704** (.0760)
ASCINCOME	-.0013 (.0015)	.2192D-06 (.3339D-05)
ASCGENDER	.0814 (.1501)	.1669 (.3351)
ASCCHILDREN	-.2582 (.1720)	-.4746 (.3873)
ASCENV	.5484*** (.2183)	.4643 (.4545)
ASCAGR	.0726 (.2279)	-.4647 (.5003)
AIC	2.05123	1.71307
BIC	1.12255	1.78987
HQIC	2.07852	1.74246
Pseudo R2	0.0762	.2350127
D.F.O	13	14
Log likelihood	-878.2585	-730.3288
Chi^2	144.8780	448.7305
Observations	869	869

Notes: Significance levels indicated by: * 0.1, **0.05, ***0.01, standard errors in brackets

Table 6. Choice models without provision rule (NPR): Hawkesbury-Nepean sub-sample

	CL	RPL
Random parameters		
NS		.0093 (.0149)
Random parameter standard deviations		
NS		*.1592*** (.0151)*
Non-random parameters		
ASC	-3.8937*** (.6445)	-6.0505*** (1.4503)
COST	-.0047*** (.0005)	-.0077*** (.0007)
NV	.0001 (.0001)	.0002 (.0001)
NS	.0325*** (.0061)	
HW	.0042*** (.0012)	.0051*** (.0015)
PA	.0011* (.0006)	.0020*** (.0007)
ASCAGE	.0223*** (.0049)	.0468*** (.0121)
ASCEDU	.1257*** (.0368)	.2446*** (.0851)
ASCINCOME	.0085*** (.0023)	.1272D-04*** (.5027D-05)
ASCGENDER	.5930*** (.1508)	1.2081*** (.3646)
ASCCHILDREN	-.2922 (.2016)	-.4751 (.4817)
ASCENV	.3475 (.2174)	-.6295 (.5039)
ASCAGR	-.0238 (.2077)	.8301* (.5039)
AIC	2.01965	1.63562
BIC	2.08325	1.70411
HQIC	2.04382	1.66164
Pseudo R2	0.08862	.2682919
D.F.O	13	14
Log likelihood	-1000.866	-807.0790
Chi^2	194.6540	591.8555
Observations	1004	1004

Notes: Significance levels indicated by: * 0.1, **0.05, ***0.01, standard errors in brackets

Table 7. Choice models with provision rule (PR): Hawkesbury-Nepean sub-sample

	CL	RPL
Random parameters		
NS		*.0260* (.0147)*
Random parameter standard deviations		
NS		.1632*** (.0168)
Non-random parameters		
ASC	-1.5332** (.6733)	-.7948 (1.4519)
COST	-.0058*** (.0005)	-.0089*** (.0007)
NV	.0001 (.0001)	.0002 (.0002)
NS	.0337*** (.0064)	
HW	.0046*** (.0012)	.0064*** (.0015)
PA	.0011* (.0006)	.0020*** (.0008)
ASCAGE	.0002 (.0055)	.0116 (.0125)
ASCEDU	.1030*** (.0373)	.1345 (.0859)
ASCINCOME	.0014 (.0020)	.3579D-05 (.4666D-05)
ASCGENDER	.3127** (.1642)	.3683 (.3819)
ASCCHILDREN	-.3232 (.2457)	-.9180 (.5781)
ASCENV	.2682 (.2719)	-.5640 (.5900)
ASCAGR	-.1447 (.1929)	.1103 (.4537)
AIC	1.99819	1.693929
BIC	2.06527	1.71152
HQIC	2.02377	1.66682
Pseudo R2	0.0886	.2675001
D.F.O	13	14
Log likelihood	-925.1525	-755.6446
Chi^2	178.7100	551.9046
Observations	939	939

Notes: Significance levels indicated by: * 0.1, **0.05, ***0.01, standard errors in brack

8.2 The implicit prices

The WTPs for changes in each attribute were estimated as implicit prices (IP). The marginal value of a change in a single attribute was calculated by dividing the β coefficients of the attributes (NV, NS, HW, and PA) by the β coefficient of the cost parameter and multiplying by -1.

$$\text{IP}=-1\left(\frac{\beta_{\text{attribute}}}{\beta_{\text{cost}}}\right) \qquad \textbf{\textit{(9)}}$$

The 95 precent confidence intervals (CI) for the WTPs were obtained by using a bootstrapping procedure from the unconditional parameter estimates (Krinsky and Robb, 1986). A vector of 1000 sets of parameters was drawn for each attribute from the covariance matrix for each sub-sample. The WTP estimates obtained from CL and RPL models are presented in Table 8.

The results show that the IPs for the environmental attributes are positive implying that respondents have positive WTPs for improvements in environmental quality. The IPs for environmental improvements in the Hawkesbury-Nepean catchment obtained from the questionnaires with and without a provision rule from both Sydney and the Hawkesbury-Nepean sub-samples are presented in Table 8.

Table 8. The mean annual households WTP (for 5 years)

	CL			
	Sydney		HN	
Attributes	NPR	PR	NPR	PR
NV	\$0.06** (0.01 ~ 0.11)	\$0.02 (-0.03 ~ 0.06)	\$0.03 (-0.02 ~ 0.08)	\$0.02 (-0.02 ~ 0.06)
NS	\$5.25*** (2.61 ~ 7.90)	\$2.30* (-0.13 ~ 4.56)	\$6.97*** (4.21 ~ 9.74)	\$5.89*** (3.69 ~ 6.15)
HW	\$1.10*** (0.56 ~ 1.64)	\$0.70*** (0.24 ~ 1.20)	\$0.90*** (0.37 ~ 1.42)	\$0.79*** (0.39 ~ 1.24)
PA	\$0.17 (-0.08 ~ 0.42)	\$0.06 (-0.19 ~ 0.30)	\$0.23* (-0.02 ~ 0.48)	\$0.20* (-0.01 ~ 0.41)
	RPL			
NV	\$0.21*** (0.18 ~ 0.25)	\$0.03 (-0.01 ~ 0.06)	\$0.02 (-0.01 ~ 0.06)	\$0.02 (-0.02 ~ 0.05)
NS	\$3.31*** (1.54 ~ 5.04)	\$-2.94 (-6.80 ~ 0.49)	\$1.18*** (-2.65 ~ 4.88)	\$2.91 (0.24 ~ 6.01)
HW	\$0.87*** (0.41 ~ 1.37)	\$0.58*** (0.20 ~ 0.97)	\$0.66*** (0.27 ~ 1.05)	\$0.72*** (0.37 ~ 1.08)
PA	\$0.24** (0.02 ~ 0.46)	\$0.07 (-0.12 ~ 0.26)	\$0.25** (0.05 ~ 0.45)	\$0.22** (0.04 ~ 0.39)

8.3 Hypothesis testing

HA*: In Sydney and the Hawkesbury-Nepean sub-samples there are no differences between attribute parameter (β) estimates obtained from the questionnaires with a provision rule (β_{PR}) and without a provision rule (β_{NPR}).*

To compare the models from two different sub-samples the coefficient vectors of these models need to contain the same elements. In order to identify the best RPL model, the model specification in terms of the type and number of random parameters is varied between the sub-samples.[9] Therefore, to be able to test for equivalence of the preferences across the two data sets CL models that have the same specification were used. To test the null hypothesis of equivalence of the preferences across the two data sets (with and without a provision rule) a grid search was conducted to estimate the ratio of scale parameters (Swait and Louviere, 1993) where μ_{NPR} for the data set without a provision rule was constrained to one and μ_{PR} for the data with a provision rule become the relative scale factor.

Following Swait and Louviere, (1993) the following steps were undertaken. First the log likelihood values (LL_{NPR} and LL_{PR}) for two separate models PR and NPR were estimated. Second, the log likelihood value (LL_{pooled}) was estimated for the pooled model data (PR+NPR). In the pooled model data sets PR and NPR were assumed to be independent. The third step involved combining the two data sets but rescaling one of the data sets (NPR) by conducting a grid search to identify the scale factor that maximises the log-likelihood of the pooled model (Earnhart, 2002). The results of the pooled models are presented in Table 9.

[9] The consistent RPL models in terms of number and type of random parameters between sub-samples were determined but the standard deviation of some of the random parameters were insignificant and the models were not superior the CL model.

Table 9. Pooled model results (CL)

	Pooled – Sydney (no scaling)	Pooled – Sydney (with scaling)	Pooled – HN (no scaling)	Pooled – HN (with scaling)
ASC	-2.7997*** (.4293)	-2.7566***(.4244)	-2.8182***(.4593)	-2.8109***(.4312)
COST	-.0053***(.0004)	-.0052***(.0004)	-.0052***(.0004)	-.0055***(.0004)
NV	.0002**(.9214D-04)	.0002**(.9101D-04)	.0001(.8515D-04)	.0001(.892D-04)
NS	.0193***(.0047)	.0191***(.0046)	.0328***(.0044)	.0346***(.0046)
HW	.0048***(.0009)	.0047***(.0009)	.0043***(.0008)	.0046***(.0009)
PA	.0006(.0005)	.0006(.0005)	.0011***(.0004)	.0012***(.0005)
ASCAGE	.0162***(.0040)	.0162***(.0040)	.0149***(.0040)	.01482***(.0040)
ASCEDU	.0959***(.0246)	.0916***(.0241)	.1133***(.0260)	.1204***(.0263)
ASCINCOME	.0046***(.0011)	.0046***(.0017)	.0049***(.0015)	.0049***(0015)
ASCGENDER	.2200**(.1095)	.2162**(.1082)	.4729***(.1098)	.5091***(.1160)
ASCCHILDREN	-.2324*(.1273)	-.2305*(.1260)	-.2615*(.1532)	-.2949*(.1604)
ASCENV	.5314***(.1562)	.5253***(.1544)	.2987*(.1652)	.3412**(.1758)
ASCAGR	.2681*(.1670)	.2678*(.1651)	-.0475 (.1391)	-.0691(.1459)
Pseudo R2	.08677	.08683	.08439	.08569
Log likelihood	-1748. 578	-1748. 461	-1936. 484	-1933. 548
D.F.O	13	13	13	13
Chi^2(critical Chi^2 in brackets)	332.262(19.675)	332.496(19.675)	356.944(19.675)	362.416(19.675)
Observations	1748	1748	1943	1943

Notes: Significance levels indicated by: * 0.1, **0.05, ***0.01, standard errors in brackets.

The log-likelihood ratio test was then conducted to test for similarities between the two data sets (NPR and PR) for the local/rural (Hawkesbury-Nepean) and distant/urban (Sydney) community sub-samples. The results of these tests are presented in Table 10.

Table 10. Test results for simulated data for the local/rural (Hawkesbury-Nepean) and distant/urban (Sydney) communities sub-samples.

Sub-sample	μ	L_{HN}	L_{HI}	LL_{pooled}	λ_{H1}	Reject HA1	LL	λ_{H2}	Reject HA2
Sydney	1.024	-847.16	-878.26	-1748.46	46.08	YES	-	-	-
Hawkesbury -Nepean	0.904	-1000.87	-925.15	-1933.55	15.06	NO	-1936.48	5.87	YES

HA1: Chi-squared statistic for 14 d.f. and 95% confidence level = 23.68479

HA2: Chi-squared statistic for 1 d.f. and 95% confidence level = 3.84146

For the Sydney sub-samples the null hypothesis of equal parameters (HA1) is rejected at the 95% confidence level. For the Hawkesbury-Nepean sub-samples the HA1 hypothesis could not be rejected at the 95% confidence level. However, the rejection of HA2 for the Hawkesbury-Nepean sub-samples implies that the scale factor and the error variance of the two sub-samples are different at the 95% confidence level. Thus, we can conclude that the inclusion of a provision rule in the CM questionnaire had an effect on the preference parameters for the Sydney sub-samples but it only had an effect on the scale parameters for the Hawkesbury-Nepean sub-samples. Therefore, at the 5 percent level of significance we cannot reject the hypothesis (HA_0) of equal parameters between the sub-samples with a provision rule (PR) and without (NPR) for both locations (Sydney and Hawkesbury-Nepean).

***HB**: In both Sydney and the Hawkesbury-Nepean sub-samples there are no differences between WTP estimates obtained from the questionnaires with a provision rule (WTP_{PR}) and without a provision rule (WTP_{NPR}).*

In order to perform the IC hypotheses tests it is necessary to identify whether the differences between the estimated IPs of the attributes across the different sub-samples are statistically significant. The Poe et al. (1994) test was used to compare IPs between different sub-samples. The Krinsky and Robb (1981) bootstrapping procedure was used to simulate the distribution of each WTP by using 1000 random draws. Using these random draws, the distributions of WTP differences between model pairs were compared.

This process was repeated 100 times for each pair of WTP estimates in order to generate the average proportion of differences where the differences are greater than zero.

The results of the Poe et al. (1994) test show that there are no significant differences in WTP between the values obtained from the CL and RPL models for most of the attributes[10]. Only two significant differences in WTP for NS in the Hawkesbury-Nepean NPR sub-sample and in the Sydney PR sub-sample were identified. As the RPL model specifications are different between the sub-samples the CL model that has the same specification across sub-samples is more suitable for the comparisons of the WTP estimates. Therefore, due to the similarity of outcomes in the IC test for both models and for clarity and consistency only the Poe et al. (1994) test for the CL models are discussed.

The WTP estimates obtained from the questionnaires with the provision rule (PR) are consistently greater than the marginal WTP estimates obtained from the questionnaires without a provision rule (NPR) in both Sydney and the Hawkesbury-Nepean sub-samples (see Table 8). The Poe et al. (1994) test however, showed that there are no significant differences (at the 10 percent level) between values from the two different treatments (PR and NPR) in the Hawkesbury-Nepean sub-samples.

In the Sydney sub-sample the values obtained from the questionnaires without a provision rule were significantly higher (at the 10 percent level) for NS. No significant differences were observed for the NV, HW and PA attributes between the two Sydney sub-samples (PR and NPR). [11] Therefore, HB_0 cannot be rejected for any attribute in the Hawkesbury-Nepean sub-sample. HB_0 however is rejected for NS (at the 5 percent level) in the Sydney sub-sample.

[10] There are no significant differences in any of the attributes' MWTP between the CL and RPL model without the panel specification.
[11] A significant difference (at the 10 percent level) was observed for NV in RPL model.

10. Conclusion

Due to increasing interest in non-market valuation and its reliability, a lot of work has been devoted to finding mechanisms that give stated preference respondents incentives to reveal their true preferences (Dasgupta et al., 1979). The design of a multiple choice format CM study that avoids strategic behaviour has proven particularly difficult because of the multiple choices included and difficulties in developing a majority voting provision rule.

The tests for IC in CM have been investigated in just a few non-market valuation laboratory experiments. This study investigates the impact of including a provision rule aimed at improving the IC of a field survey choice modelling study. The test was conducted by investigating the impact of a provision rule on the stated preferences for a common catchment management strategy of two different communities (local/rural and distant/urban). Careful design of the survey and large sample sizes were additional steps undertaken to improve IC.

The test for equivalence between parameters from the two sub-samples (with and without a provision rule) showed that there were significant (at the 5% level) differences between the parameters of the two data sets (PR and NPR) in the Sydney sub-samples. This indicates that the inclusion of a provision rule had an impact on respondent preferences. However, the lower scale parameter for the Sydney sub-sample with a provision rule indicates that choice behaviours are more random when a provision rule is used. On the other hand, in the Hawkesbury-Nepean sub-samples, significant (at the 5% level) differences were observed in the variances but not in the actual coefficients. However, in this case a higher scale parameter for the sub-samples with a provision rule indicates a reduction in variance. Therefore there is a lack of consistency in the response to the provision rule across the two sub-samples.

Further results showed that the WTP estimates for the questionnaires with and without a provision rule were not significantly different for the Hawkesbury-Nepean sub-samples.

However a significant difference (at the 10 percent level) between the Sydney sub-samples (PR and NPR) was observed for the NS attributes and for NV in the RPL model. The values for this attributes were significantly higher (at 10 percent level) when the provision rule was not included in the questionnaire. In other words the inclusion of a provision rule reduced WTP estimates of respondents in Sydney only for NS and NV attribute.

These results suggest that the inclusion of a provision rule in the local sub-sample had a negligible effect on the results of the CM study. However, the provision rule had a significant (at 10 percent level) effect on NS and NV attributes in the distant/urban community sub-sample. This implies that the distant/urban community was more sensitive to the inclusion of a provision rule than the local/rural community sub-sample. Hence, the local/rural community sub-sample with its closer association with the environment under consideration valued the change consistently across the two treatments. The distant urban communities' values for the changes in NRM in the Hawkesbury-Nepean catchment where a provision rule was included were significantly (at the 10 percent level) lower than the values obtained from the questionnaires without a provision rule. The impact of a provision rule should thus be analysed in the context of different community characteristics.

11. References

ANDREONI, J. (1989) Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy,* 97**,** 1447-1458

ARROW, K. (1951) *Social Choice and Individual Values,* New York., Wiley.

ARROW, K., SOLOW, R., PORTNEY, P. R., LEAMER, E. E., RADNER, R. & SCHUMAN, H. (1993) Report of the NOAA Panel on Contingent Valuation. *Federal Register,* 58**,** 4601-4614.

BATEMAN, I. J., BURGESSB, D., HUTCHINSONC, G. W. & MATTHEWS, D. I. (2008a) Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. *Journal of Environmental Economics and Management* 55**,** 127-141.

BATEMAN, I. J., CARSON, R. T., DAY, B., DUPONT, D. P., LOUVIERE, J., MORIMOTO, S., SCARPA, R. & WANG, P. (2008b) Choice Set Awareness and

Ordering Effects in Discrete Choice Experiments. CSERGE Working Paper EDM.
BEN-AKIVA, M. & LERMAN, S. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand,* Cambridge, MIT Press.
BENNETT, J. (1987) Strategic Behaviour Some Experimental Evidence. *Journal of Public Economics,* 32**,** 355-368.
BISHOP, R. C. & HEBERLEIN, T. A. (1979) Measuring Values of Extra-Market Goods: Are Indirect Measures Biased? *American Journal of Agricultural Economics* 61**,** 926-930.
BLAMEY, R. K., BENNETT, J. W. & MORRISON, M. D. (1999) Yea-Saying in Contingent Valuation Surveys. *Land Economics* 75**,** 126-141.
BOHM, P. (1971) Estimating Demand for Public Goods: An Experiment. *European Economic Review,* 3**,** 111-130.
BOWEN, H. R. (1943) The Interpretation of Voting in the Allocation of Economic Resources. *The Quarterly Journal of Economics,* 58**,** 27-48.
BOYLE, K., JOHNSON, F., MCCOLLUM, D., DESVOUSGES, W., DUNFORD, R. & HUDSON, S. (1996) Valuing public goods: discrete versus continuous contingent-valuation responses. *Land Economics* 72 381- 96.
BURTON, A. C., CARSON, K. S., CHILTON, S. M. & HUTCHINSON, W. G. (2007) Resolving questions about bias in real and hypothetical referenda. *Environmental Resource Economics,* 38**,** 513-525.
CARSON, K. S. & BURTON, A. C. (2008) Can Choice Experiment Surveys be Incentive Compatible? An Economic Experiment on Repeated Choice Formats with Financial Incentives in Place. *Environmental Economic Research Conference.* Gothenburg, Sweden.
CARSON, R. & GROVES, T. (2007) Incentive and informational properties of preference questions. *Environmental and Resource Economics,* 37**,** 181-210.
CARSON, R. & MITCHELL, R. C. (1995) Sequencing and nesting in contingent valuation survey. *Journal of Environmental Economics and Management,* 28**,** 155-173.
CARSON, R. T., MITCHELL, R. C., HANEMANN, W. M., KOPP, R. J., PRESSER, S. & RUUD, P. A. (1994) Contingent valuation and lost passive use: damages from the Exxon Valdez, Discussion Paper 94-18. *Resources for the Future.* Washington, D.C.
CHAMP, P. A., BISHOP, R. C., BROWN, T. C. & MCCOLLUM, D. W. (1997) Using Donation Mechanisms to Value Non-use Benefits from Public Goods. *Journal of Environmental Economics and Management,* 33**,** 151-162.
CLARKE, E. H. (1971) Multipart pricing of public goods *Public Choice***,** 17-33.
COLLINS, J. P. & VOSSLER, C. A. (2009) Incentive compatibility tests of choice experiment value elicitation questions. *Journal of Environmental Economics and Management,* 58**,** 226-235.
CUMMINGS, R. G., ELLIOT, S., HARRISON, G. W. & MURPHY, J. (1997) Are hypothetical referenda incentive compatible? *Journal of Political Economy,* 105**,** 609-621.

CUMMINGS, R. G., HARRISON, G. W. & RUTSTROM, E. E. (1995) Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive-Compatible? *American Economic Review,* 85**,** 260-66.

DASGUPTA, P., HAMMOND, P. & MASKIN, E. (1979) The implementation of Social Choice Rules: Some General Rules on Incentive Compatibility. *The Review of Economics and Statistics,* 46**,** 185-216.

DAVIS, D. D. & HOLT, C. A. (1993) *Experimental Economics.* , Princeton: Princeton University Press.

DORFMAN, R. (1969) General Equilibrium with Public Goods. IN MARGOLIS, J. & GUITTON, H. (Eds.) *Public Economics* Macmillan.

EARNHART, D. (2002) Combining Revealed and Stated Data to Examine Housing Decisions Using Discrete Choice Analysis. *Journal of Urban Economics,* 51**,** 143-169.

FISCHHOFF, B., QUADREL, M. J., KAMLET, M., LOEWENSTEIN, G., DAWES, R., FISCHBECK, P., KLEPPER, S., LELAND, J. & STROH, P. (1993) Embedding effects: stimulus representation and response mode. *Insurance: Mathematics and Economics,* 13**,** 161.

GEORGIOU, S., BATEMAN, I., COLE, M. & HADLEY, D. (2000) Contingent ranking and valuation of river water quality improvements: Testing for scope sensitivity, ordering and distance decay effects. . CSERGE Working Paper GC 2000-18.

GIBBARD, A. (1973) Manipulation of voting schemes: a general result. *Econometrica,* 41**,** 587-601.

GROVES, T. (1973) Incentives in teams. *Econometrica* 41.

GROVES, T., RADNER, R. & REITER, S. (1987) *information, Incentives, and Economic Mechanisms: Essays in Honour of Leonid Hurwicz,* Minneapolis, University of Minnesota Press.

HAMMOND, P. (1979) Straightforward Individual Incentive Compatibility in Large Economies. *The Review of Economic Studies,* 46**,** 263-282.

HANEMANN, W. M., LOOMIS, J. & KANNINEN, B. (1991) Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics* 73 1255-1263.

HARRISON, G. (2007) Making Choice Studies Incentive Compatible. *Valuing Environmental Amenities Using Stated Choice Studies.*

HARRISON, G. W. (1992) Valuing public goods with the contingent valuation method: a critique of Kahneman and Knestch. *Journal of Environmental Economics and Management,* 23**,** 248-57.

HENSHER, D. & REYES, A. (2000) *Trip chaining as a barrier to the propensity to use public transport,* Netherlands, Kulwer Academic Publishers.

HENSHER, D. A., ROSE, J. A. & GREENE, W. H. (2005) *Applied Choice Analysis,* Cambridge, Cambridge University Press.

HOEHN, J. P. & RANDALL, A. (1987) A Satisfactory Benefit-Cost Indicator from Contingent Valuation. *Journal of Environmental Economics and Management,* 14**,** 226-246.

HURWICZ, L. (1986) Incentive Aspects of Decentralization. IN ARROW, K. J. & INTTRILLIGATOR, M. D. (Eds.) *Handbook of Mathematical Economics.* Amsterdam.

KAHNEMAN, D. & KNETSCH, J. L. (1992) Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management,* 22**,** 57-70.

KRINSKY, I. & ROBB, A. L. (1986) On approximating the statistical properties of elasticises. *Rev. Econ. Stat.,* 68**,** 715-719.

LIN, X., YACOUB, S., BURNS, J. & SIMSKE, S. (2003) Performance analysis of pattern classifier combination by plurality votin. *Pattern Recognition Letters* 24 1959-1969.

LOUVIERE, J., HENSHER, D. A. & SWAIT, J. (2000) *Stated Choice Methods, analysis and application*, Cambridge University Press, U.K.

LUSK, J. L. & SCHROEDER, T. C. (2004) Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks. *American Journal of Agricultural Economics,* 86**,** 467-482.

MAZUR, K. & BENNETT, J. (2009) A Choice Modelling Survey of Community Attitudes for Improvements in Environmental Quality in NSW Catchments. *EERH Report No 10.* Canberra, Australian National University.

MAZUR, K. & BENNETT, J. (2008) Using focus groups to design a choice modelling questionnaire for estimating NRM benefits in NSW. *EERH Report No 2.* Canberra, Australian National University.

MCFADDEN, D. (1974) Conditional logit analysis of qualitative choice behaviour. IN ZAREMBKA, P. (Ed.) *Frontiers of Econometrics.* London, Academic Press.

MCMILLAN, J. (1979) Individual Incentives in the Supply of Public Inputs. *Journal of Public Economics,* 12**,** 87-98.

MITCHELL, R. C. & CARSON, R. T. (1989) *Using surveys to value public goods: the contingent valuation method. Resource for the Future,* Washington, DC.

POE, G. L., CLARK, J. E., RONDEAU, D. & SCHULZE, W. D. (2002) Provision Point Mechanisms and Field Validity Tests of Contingent Valuation. *Environmental and Resource Economics,* 23**,** 105-131.

POE, G. L., SEVERANCE-LOSSIN, E. & WELSH, M. P. (1994) Measuring the Difference (X-Y) of Simulated Distribution: A Convolution Approach. *American Journal of Agricultural Economics,* 76**,** 904-914.

POLOMÉ, P. (2003) Experimental evidence on deliberate misrepresentation in referendum contingent valuation. *Journal of Economic Behavior & Organization,* 52 387-401.

POSAVAC, S. (1998) Strategic overbidding in contingent valuation: stated economic value of public goods varies according to consumers expectations of funding source *Journal of Economic Psychology,* 19**,** 205-214.

RACEVSKIS, L. A. & LUPI, F. (2008) Incentive Compatibility in an Attribute-Based Referendum Model. *The American Agricultural Economics Association Annual Meeting.* Orlando.

READY, R., BUZBY, J. & HU, D. (1996) Differences between continuous and discrete contingent value estimates

Land Economics

72**,** 397- 411.

REESON, A. & NOLLES, K. (2009) Experimental economics applications to environmental policy *Working Paper Series 2009-03.* Canberra, CSRIO.

ROBERTS, J. (1976) Incentives for the Correct Revelation of Preferences and the Number of Consumers. *Journal of Public Economics,* 6**,** 359-374.

RSAS (2007) Mechanism Design Theory. *Scientific background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2007.* Stockholm, Sweden, The Royal Swedish Academy of Science.

SAMUELSON, W. (1954) The Pure Theory of Public Expenditure. *The Review of Economics and Statistics,* 36**,** 387-389.

SATTERTHWAITE, M. A. (1975) Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory,* 10**,** 187-217.

SCHEUFELE, G. & BENNETT, J. (2010) Effects of alternative elicitation formats in discrete choice experiments. *54st Annual Conference of the Australian Agricultural and Resource Economics Society.* Adelaide, SA.

SCOTT, A. (1965) The valuation of game resources: some theoretical aspects. *Can Fish Rep,* 4**,** 27-47.

SMITH, V. K. (1992) Arbitrary values, good causes, and premature verdicts. *Journal of Environmental Economics and Management,* 22**,** 71-89.

SUDMAN, S., BRADBURN, N. & SCHWARZ, N. (1996) *Thinking about answers: the application of cognitive processes to survey methodology,* San Francisco, Jossey-Bass Publishers,.

SWAIT, J. & LOUVIERE, J. J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Market. Res,* 30**,** 305-314.

TAYLOR, L., O, MORRISON, M., D & BOYLE, K. J. (2007) Exchange Rules and the Incentive Compatibility of Choice Experiments.

TRAIN, K. (1998) Recreation demand models with taste differences over people. *Land Economics,* 74**,** 230-239.

TRAIN, K., E (2000) Halton Sequences for Mixed Logit. *Paper E00-278.* University of California, Berkeley, Institute of Business and Economics.

VARIAN, H. (1992) *Microeconomic Analysis,* New York.

VENKATACHALAM, L. (2004) The contingent valuation method: a review. *Environmental Impact Assessment Review,* 24**,** 89-124.