The Australian Journal of
Agricultural and Resource Economics

AARES
AUSTRALASIAN AGRICULTURAL & RESOURCE ECONOMICS SOCIETY

Check for updates

Australian Journal of Agricultural and Resource Economics, 66, pp. 280–308

Reducing bias in preference elicitation for environmental public goods*

Daniel A. Brent, Lata Gangadharan, Anke D. Leroux and Paul A. Raschky†

The recent stated preference literature emphasises the importance of incentive compatible elicitation methods, which depend on respondent beliefs that payment can be collected if provision occurs. We investigate this condition in a randomised field experiment where stated choices are incentivised financially. The objective of the treatment was to make choices salient by making each decision financially relevant and to increase the respondents' beliefs that future payments will be enforced. Our results show that the treatment increases estimates of the marginal utility of income, with the effect being economically and statistically significant for low-income respondents. We develop a stylised theoretical framework that allows us to quantify the bias that is implied by the observed differences between the treated and control groups. We find that failure to account for respondents' doubts about payment coercion in an otherwise well-designed survey inflates the marginal willingness to pay among low-income respondents by a factor of at least 1.72.

Key words: consequentiality, field experiment, non-market goods, payment coercion, quasi-public goods, stated preference.

JEL classifications: Q51, C93

* The authors are grateful for helpful comments from Jeff Bennett, Mike Burton, Thijs Dekker, Denzil Fiebig, Glenn Harrison, Robert Johnston, Andrea Leiter-Scheiring, John List, Jim Murphy, Christian Vossler and participants at AARES 2016, WCERE 2014 and the Behavioral and Experimental Workshop at Monash University, 2 anonymous referees, the associate editor and the editor that greatly improved the paper. Funding from the Cooperative Research Centre for Water Sensitive Cities (CRC grant number 20110044) and the USDA National Institute of Food and Agriculture and Hatch/Multi-State Appropriations under Project #PEN04631 and Accession #1014400 is acknowledged. This project has been approved by the Monash University Human Research Ethics Committee; MUHREC#: CF12/2511 2912001358.

† Daniel A. Brent is Assistant Professor at the Department of Agricultural Economics, Sociology, & Education, Pennsylvania State University, State College, Pennsylvania, USA. Lata Gangadharan (email: Lata.Gangadharan@monash.edu) is Professor at the Department of Economics, Monash University, Clayton, Victoria, Australia. Anke Leroux is Associate Professor at the Department of Economics, Monash University, Clayton, Victoria, Australia. Paul A. Raschky is Professor at the Department of Economics and Principal Investigator at SoDa Laboratories at Monash University, Clayton, Victoria, Australia.

doi: 10.1111/1467-8489.12463

1. Introduction

Stated preference surveys remain the most commonly used technique for estimating monetary values of non-marketed goods and services. Throughout its history of use, there has been significant debate over the ability of stated preference surveys to elicit truthful responses (Diamond & Hausman, 1994; Hanemann, 1994; Hausman, 2012; Kling et al., 2012). In this paper, we examine the role financial incentives may play in such survey settings so as to identify and reduce bias in the non-market valuation of projects that yield both private and public benefits.

In recent years, efforts to identify the sources and extent of potential bias have shifted their focus from a discussion about the hypothetical nature of surveys[1] to understanding how surveys can be designed to be consequential to respondents and motivate truthful answers. According to Carson and Groves (2007) and Vossler et al. (2012), the following conditions are required for consequentiality to hold. First, the respondents care about the survey outcome and believe that their answers influence whether a proposed policy or project is acted upon (policy consequentiality). Second, the respondents must also believe that the policymaker can enforce payment for the public good from the respondents once the policy is implemented (payment consequentiality). A number of empirical studies have highlighted the importance of consequentiality in survey design, by either randomising the inclusion of a consequentiality statement in the survey (Bulte et al., 2005), varying the probability of the outcome to be binding (Carson et al., 2014; Mitani & Flores, 2014) or checking the interviewees' beliefs about the consequentiality of the survey in follow-up questions (Vossler et al., 2012). More recent empirical studies by Czajkowski et al. (2017) and Lloyd-Smith et al. (2019), however, raise some doubts that the inclusion of simple consequentiality scripts indeed yields more truthful preference revelations.

Even in situations in which the survey is policy consequential, respondents may still be doubtful as to whether the payment for the public good will ever be coerced. Groothuis et al. (2017) show that respondents' beliefs in the survey's consequentiality decrease with higher project costs, while Champ et al. (2002) found that a large portion of respondents believe that the actual project costs are higher than the costs used in the survey. If respondents believe that they will never have to pay for the project in the form of added taxes, council rates or service fees, they may have an incentive to overstate their willingness to pay. Research by Zawojska et al. (2019) and Borger et al.

[1] While some studies in this early literature find no evidence of hypothetical bias (Carson et al., 1996; Carlsson and Martinsson 2001), other laboratory (Harrison, 2006a; Harrison, 2006; Andersen et al., 2006) and field experiments (Cummings et al., 1997; List et al., 2006) find evidence of hypothetical bias in many common valuation methods with meta-analyses showing that hypothetical willingness to pay typically exceeds the actual value by a factor of two to three (List and Gallet, 2001; Loomis, 2011).

(2021) suggests that self-indications of consequentiality can be used to model and control for consequentiality in willingness to pay (WTP) studies.

The present paper proposes an alternative approach to increase payment consequentiality in a survey. Incentivising choices financially, using experimental methods, means that the gains and losses that arise for individuals from their decisions are actually experienced and thereby become salient to them. We thus include a treatment in our stated preference study, designed to increase perceptions that payment for the public good will be enforced. The objective is to make choices salient and investigate whether this increase in the probability of payment coercion has an impact on respondents' choices. In particular, we conduct personal interviews with a randomised sample of almost 1000 individuals and elicit their preferences for the non-market benefits of local water management. A randomly determined subset of the respondents is incentivised with earned or endowed cash before they choose among alternative water management projects that vary in the provision of public and private benefits and in their costs. Prior to the choice task, respondents in the treatment group are informed that one of their choices will be randomly selected at the end of the experiment and the cost for the associated project will be deducted from their earnings for the benefit of a specified water management pilot project in their local community. We label this treatment as the *salient treatment* as it connects financial incentives to the decisions of the respondents (Smith, 1982). Respondents in the control group are presented with decisions that follow the non-incentivised protocol. The specific purpose of the salient treatment is to link the survey instrument to a payment requirement that reflects the stated choices, thereby increasing the credibility of the coercive payment vehicle.[2]

Our first hypothesis is that the estimates of the marginal utility of income, (the negative of the coefficient on cost from our econometric model), are the same across both treatment and control groups. We focus on the marginal utility of income since it measures how responsive respondents are to hypothetical costs. This allows us to examine whether the salient treatment makes respondents more sensitive to project costs, thereby decreasing their WTP. Our second hypothesis tests if the treatment differentially affects the preferences for the public and private benefits of the good. Provided the existing evidence for private goods is transferable to quasi-public goods, one would expect to see no systematic difference in preferences across the treatment and control group for the public and private benefits of the good.

With respect to the first hypothesis, we find mixed evidence. We find no statistically significant treatment effect on aggregate. However, our results show that the effect of our treatment is highly heterogeneous; low-income

[2] The randomisation of the salient treatment across the survey respondents gives rise to a between-subject design, eliminating concerns about respondents' desire to be consistent in their preferences when presented with both, the control and treatment protocols (Johansson-Stenman and Svedsäater, 2008).

households are more sensitive to the treatment with the marginal utility of income increasing by 84 percent for the treated sample compared to the control group. Regarding our second hypothesis, we do not find systematic treatment effects. Preferences for the benefits of the quasi-public good are not statistically significantly different across the treatment and control groups. Importantly, this result holds for both public and private benefits of the quasi-public good, thereby extending the existing results on the marginal utility of private benefits (List et al., 2006) to public good benefits as well.

We develop a simple theoretical framework of the effect of increasing payment consequentiality in stated preference settings to formalise the interpretation of our empirical results. In this framework, the salient treatment affects marginal WTP by weakly increasing the subjective probability that payments will be coerced. In addition, we explicitly allow for any discrepancies in perceived benefits that arise from the payment supporting a pre-specified water management pilot project instead of the chosen alternative. The theoretical framework allows us to determine the extent of bias under different assumptions regarding treatment efficacy as well as the degree of potential benefit distortions during the implementation stage of the project. For the low-income group, for which this effect is more precisely estimated, we show that marginal WTP measures in the control group are at least 72 percent higher than in the treatment group.

Our study makes several contributions to the literature on non-market valuation. First, we complement the above-mentioned literature on consequentiality in stated preference methods by investigating the importance of respondents' beliefs in payment coercion. Respondents believing that payments for the selected policy or project will be coerced is a condition of truthful preference revelation that is often assumed to hold rather than being put to the test. Our results suggest that payment consequentiality is a critical feature of stated preference surveys. We find lower WTP among some respondent groups for whom payments were coerced. Second, our investigation takes place in a setting that is typical for non-market valuation studies in the field, whereby the objective is to elicit the value of non-market goods to inform their provision—that is values are elicited for goods that are non-existent at the time of the survey. This is a challenging setting for testing truthful preference revelation and we design a novel experiment, involving a proxy good. Third, we develop a theoretical framework that allows us to interpret the treatment effect under different assumptions of treatment efficacy and in the light of the proxy good being an imperfect substitute for the good under consideration. Fourth, the experimental treatment we designed represents an easy to implement tool that practitioners can use to address bias when they expect that payment coercion may not be widely believed in by survey respondents.

The paper is organised as follows: Section 2 describes the design of the field experiment and the survey. Section 3 outlines the theoretical framework, and Sections 4 and 5 introduce the data and present differences in the raw choice data. Section 6 describes the empirical framework, presents the results and discusses the observed treatment effects. Section 7 concludes.

2. Design of the field experiment and the survey

2.1 Field experiment

The sequence of the experiment and survey was as follows. Interviewers went to randomly selected homes, introduced themselves and asked the householder whether he/she would be willing to participate in a survey about local water management.[3] After confirming the eligibility requirements (older than 18 years and owner–occupier status), the interviewer started the survey on an iPad. At this stage, the software randomly assigned the interviewee into the treatment ("Salient") or the control group.

The control group immediately started with the choice task, while the treatment group was allocated randomly by the software to two equal sized groups called "Earned Salient" and "Endowed Salient." The "Endowed Salient" group received 1 out of 4 potential endowments (each with a probability of 0.25): A$30.60, A$39.60, A$42.00, A$53.10.[4] The "Earned Salient" group received an initial endowment of A$30.00 and, before commencing the choice task, participated in a risk elicitation task based on Holt and Laury (2002). An example of the decision problem can be found in Appendix S1: Figure A3. The earnings from this game ranged between A$0.60 and A$23.10 and were added to the respondent's initial endowment. The two salient treatments were designed such that the distribution of total earnings would a priori be comparable in both subsamples. Throughout the remainder of the survey, the respondent's money balance was shown on the upper right corner of the screen. The two salient treatments allow us to examine if project choice differs according to the source of the income obtained (endowed versus earned).[5] We do not find any differences across "Earned Salient" and "Endowed Salient," so in our analysis we pool the data for the two salient treatments.[6]

At the beginning of the choice task, the interviewer carefully explained the choice situation as well as the procedure of the choice task to the respondent

[3] A copy of the introduction letter can be found in Appendix S1; Figure A1. The list of households to be visited resulted from a random draw from the council's homeowner database.

[4] At the time the experiment was conducted, 1 Australian dollar was about 0.96 of the US dollar.

[5] Ideally, we would have asked participants in the treatment group to pay for the cost of their choice without first receiving a cash endowment, but the field implementation of such a design is problematic. Section 5.4 discusses potential implications of the initial cash endowment for the interpretation of our results.

[6] The test statistic of a non-parametric Mann–Whitney test for average cost of selected alternatives across the earned versus endowed salient treatments is 0.319 with a p-value of 0.75. In several other contexts, particularly in the laboratory, researchers have found differences in decisions between endowed and earned treatments. For examples, see Cherry et al. (2002); Hoffman et al. (1994) among others.

(see Appendix S1: Figure A4). It was explicitly mentioned that we were interested in their truthful valuation of the benefits. Each individual was asked to select their preferred option in 10 subsequent choice sets. In addition, respondents in the treatment group were informed that at the end of the interview they would be asked to draw a number between 1 and 10, and this would determine which choice set was selected for payment. They were also informed that the annual cost of their selected option would be deducted from their interview earnings and transferred to an existing pilot water management project in their local area, which could be scaled up and modified depending on the choices made by the subjects.[7] Thus, treated respondents' selections in the choice experiments were directly tied to financial incentives.

Each choice set contained a status quo that was a scenario with no changes in the attribute levels with a cost of A$0, as well as two options (options A and B) that provide improvements in at least one of five attributes (discussed in the next section) and always had costs >A$0. Our trichotomous format mimics many recent field applications (see, for example, Rogers et al., 2020), where policymakers are interested in using stated preference methods as an instrument to identify the preferred scope and features of a multi-dimensional public good to be provided in the future. As such, our research expands on related work by Vossler et al. (2012), who conduct a field test of truthful preference revelation in a referendum for a public project of much narrower scope.

Moreover, opting for this choice format allows us to investigate the impact of our salient treatment at the intensive margin, whereby treated respondents, if given the option, may choose a non-status quo project with similar probability to individuals in the control group, but then opt for the lower-cost alternative. We acknowledge that earlier arguments in favour of trichotomous choice formats (Rolfe & Bennett, 2009)[8] have not been adopted in the most recent guidelines for stated preference studies (Johnston et al., 2017) due to concerns over incentive incompatibility and status quo bias (Carson et al., 2020). However, all design features are constant across all respondents and treatments so that any observed differences between the control and treatment groups can be attributed to the treatment effect.

The treated subjects received their final payout at the end of the interview. The payout was always positive and ranged between A$0.60 and A$53.10. As per information provided to the salient groups, the total amount paid by the

[7] Depending on the pilot project, what respondents paid for, matched the attributes of the selected option in most cases, but not all. Our theoretical framework considers the potential effects of a mismatch.

[8] The argument is that including more than two alternatives (at least two options in addition to the status quo) may also provide better value elicitation since a dichotomous choice masks much of the variation in specific alternatives and transforms the decision to being primarily pro-project or anti-project.

survey participants was transferred to the respective water management pilot projects and published in the councils' newsletters.

2.2 Survey

The survey and the discrete choice experiment were designed to elicit stated preferences for urban water management in Australia. A random sample of 981 Australian individuals from four councils in Melbourne, Victoria (VIC), and Sydney, New South Wales (NSW), metropolitan areas were personally interviewed using iPads. The four councils (Fairfield [NSW], Manningham [VIC], Moonee Valley [VIC] and Warringah [NSW]) were initially chosen from a list of 29 Cooperative Research Centre (CRC) partner communities.[9] Having access to the CRC partner councils helped implement the salient treatment as these councils were inclined towards setting up pilot stormwater management projects in the near future. Similarly, residents in these councils may be more familiar with local water management initiatives, aiding the plausibility of our survey. Hence, running this survey in the credible setting of a CRC partner council undertaking the proposed activity is the best attempt to provide robust, reliable and consequential estimates even in the control group.

Among the list of partner councils, we examined different data sources to select councils that were similar along several important dimensions. First, we selected councils that were similar in the local precipitation patterns since we expect climatic factors to affect preferences for water management.[10] Next, we accessed data from the HILDA database and compared the different councils along a list of demographic characteristics (income, age composition, percentage of homeowners) as well as by responses to questions about environmental preferences.[11] We selected a subset of four councils, based on their similarity in the most relevant demographic characteristics. As the objective of the study is to investigate valuation in stated preference methods, we make no claims of sample representativeness beyond seasonality, age and home ownership status of the target population.

The survey was conducted by a professional survey company between March and August 2013.[12] Opting for personal interviews as the

[9] The set of partner communities was established by the Cooperative Research Centre (CRC) for Water Sensitive Cities, an Australian research initiative funded by the federal government.

[10] We accessed the daily rainfall statistics for all Australian councils from January 1890 to February 2013 from the Australian Bureau of Meteorology. We then compared long-term mean and variance in daily, weekly and monthly precipitation between the councils.

[11] The Household, Income and Labour Dynamics in Australia (HILDA) is a government-funded Australian household panel study.

[12] This extended data collection period means that the results are not driven by the seasonality in rainfall and therefore ensures greater representativeness of the value estimates for rain-dependent attributes. Most importantly, the length of the data collection period does not adversely affect the field experiment as the selection of treatments is equally distributed over the sampling period.

methodology, instead of phone, mail or Internet surveys, was important because we needed to ensure that the respondents understood the information and the alternative scenarios presented to them in the choice experiment. Moreover, the use of an iPad, with its clear visual images of the choice sets and user-friendly interface aided the respondents' understanding of the available options, thereby helping to smooth the effects of varying cognitive abilities on choices.

The survey consisted of three parts: First, an introduction to the study, providing some explanation and motivation for the survey (see Appendix S1: Figure A1 for the introduction letter). Second, the choice task to elicit individuals' preferences for attributes associated with local water management projects.[13] The third part of the survey was a demographic questionnaire, comprising questions on socioeconomic characteristics and attitudes towards environmental goods and services.

The attributes were selected in the following way: The CRC holds quarterly meetings of the key stakeholders involved in water management projects in Australia. These include representatives of local councils, water authorities and providers, as well as researchers from various disciplines (engineering, hydrology, climate science, urban studies, economics, law, sociology and political science). The audience was divided into small groups, each containing at least one representative from each stakeholder group.

The groups were then asked to list the 10 most important benefits associated with stormwater management. From these lists, the following final set of five attributes was agreed upon in a plenary forum: Reduction in Water Restrictions, Reduction in Flash Flooding, Improvements in Stream Health, Improvements in Recreational and Amenity Benefits, and Cooler Summer Temperatures. In the next step, the levels for each individual attribute were defined in collaboration with researchers from the respective disciplines. For example, the attribute levels for reduction in flash flooding were defined by a group of hydrologists, engineers and climate scientists, while levels in the attribute Improvements in Stream Health were defined by a group of hydrologists, biologists and ecologists.

In the context of water management, many attributes are subject to risk in the sense that whether or not a promised outcome is achieved also depends on exogenous factors. For example, while stormwater harvesting may go some way to reduce the need for compulsory water restrictions, it may not be sufficient to achieve this outcome during a severe drought. Similarly, investment in stormwater harvesting infrastructure may improve the level of biodiversity in the local stream, but the final outcome is subject to a variety

[13] The study had two choice tasks and the treatment only incentivised payments for the first choice task; the treatment did not apply to the second choice task in any manner. Analysing the effect of treatment on the second choice task serves as a placebo test that treatment has no impact on outcomes not targeted by the treatment. Dorner et al. (2019) study the second choice task and find no impact of the treatment on the second choice task. This strengthens the argument that any observed treatment effects stem from the intended mechanism.

of other ecological factors. In contrast, the costs of investing in stormwater infrastructure are more certain. Therefore, we allow two attributes, the removal of water restrictions and improvements in stream health, to be achieved subject to some probability. We frame risk as the probability of success rather than the risk of failure. The five attributes were presented to the participants as the benefits from local water management and were defined as follows:[14]

Reductions in Water Restrictions range from a status quo scenario with no change (attribute level 1), to the exemption from less invasive restrictions (level 2), to the exemption from the most austere restrictions in the local area (level 3). The second attribute relates to the *Reduction in Flash Flooding*. Under the status quo (level 1), the average number of flash floods over a five year period remains the same. Smaller water management projects (level 2) are able to reduce the frequency of flash floods by half, while larger water management projects (level 3) are capable of reducing the number of flash floods to almost none. *Improvements in Stream Health* account for the fact that urban water management can have a direct impact on the health of local waterways. The status quo (level 1) is an unhealthy stream characterised by littered and eroded banks and low species diversity. Moderate improvements (level 2) are comprised of reduced erosion, no litter and improved species diversity, whereas large improvements (level 3) involve the return to a diverse stream community with few nuisance species. *Improvements in Recreational and Amenity Benefits* include, for example, recreational use benefits associated with local water ways such as paddling and swimming or the use of water for irrigation of local sports grounds and parks. The status quo (level 1) is characterised by rivers that are only fit to paddle, sports grounds and parks that are dry during extended periods without rain, and street line vegetation (i.e. trees) that is not watered. Moderate (level 2) recreational and amenity benefits include greener sports grounds and parks during extended dry periods and permit watering of street line vegetation. High-level benefits (level 3) improve the local waterway quality to being fit for swimming and increase the amount of street line vegetation. *Cooler Summer Temperatures* involve either no change in local summer temperatures under the status quo (level 1) or hot summer days being 2 degrees C cooler on average as a result of shading from additional trees being planted and evaporative cooling from artificial water bodies (level 2).

Finally, the *Costs* for the different projects are presented as additions to the household's annual water bill and range from A\$0 to A\$30.[15] Given the current legal framework in Australia, this payment vehicle is the most likely mechanism to fund stormwater management projects at the communal level.

[14] A more detailed description of the attributes is presented in Appendix S1. Brent et al (2017) estimates households' willingness to pay (WTP) for these attributes.

[15] The explanation of the attribute *Costs* reads as follows: "These are the costs per household per year of providing the water management option. These costs would be added to your annual water bill."

As a result, only individuals, who are owner–occupiers and therefore responsible for paying the water bill, were interviewed. The selected cost levels were also endorsed by legal and policy experts as we were interested in presenting realistic scenarios that respondents would take credibly. The experiment was designed such that the respondents in the salient treatment would always earn more than the highest cost in any choice set. Each respondent was presented with 10 different choice sets that represented water management projects that varied along five attributes as well as costs.[16] The choice sets were generated using the NGene software package, where the D-efficiency criterion was applied to a 4x10 block design.

The questionnaire in the third part was designed to collect additional information about the respondent. A set of questions about the respondent's experience with the different attributes, the use of environmental goods, as well as experience with natural hazards, is followed by a number of questions that allow respondents to be categorised into different types depending on their attitude towards water management (i.e. their concerns for water quality and biodiversity in and around the waterways). A set of demographic controls concludes the questionnaire. Among these are an income variable, collected in intervals (based on the HILDA classification) as well as a self-reported categorisation into low, medium and high income. Both, the sequence and the content of the choice task and the questionnaire were the same across all 981 participants.

Before going into the field, interviewers were carefully briefed and trained by the authors. The fieldwork commenced with two rounds of pilot studies. The first round was conducted with a group of 10 employees from Manningham and Mooney Valley City Councils (VIC) who volunteered for the study. Of the ten volunteers, one had professional experience in local water management. The pilot was supervised by one of the authors as well as a trained social psychologist, who interviewed the volunteers before and after they completed all survey components to evaluate the overall survey design (i.e. wording, length, information content) as well as the cognitive demands of the survey. The revised version of the survey was field tested with randomly selected homeowners living in Warringah council (NSW), before the final version was rolled out.

3. Theoretical framework

Our starting point is a random utility model (McFadden, 1973) of householders' choices over a set J of local water management options, including the status quo and its alternatives. The utility U of individual i from choosing a water management option j in choice occasion t is given by

[16] Figure A6 in the Appendix S1 provides an example choice set within the explanation document.

$$U_{ijt} = V_{ijt} + \epsilon_{ijt}, \tag{1}$$

where V_{ijt} is typically assumed to be a deterministic function of the observable characteristics of water management option j and ϵ_{ijt} is a random component. In this framework, individual i chooses the water management option that yields the highest level of utility with probability

$$\begin{aligned} \pi_{ijt} &= \Pr(Y_{it} = j) = \Pr(U_{ijt} > U_{iht}) : \forall : h \neq j \\ &= \Pr(V_{ijt} + \epsilon_{ijt} > V_{iht} + \epsilon_{iht}) : \forall : h \neq j \end{aligned} \tag{2}$$

To illustrate the channels through which we expect our salient treatment to work, we augment the standard theoretical framework by explicitly allowing for payment uncertainty in the context of a well-designed survey, where choices are consequential in terms of respondents believing that their responses matter to them and for policy decisions (Herriges et al., 2010). In this setting, bias may still arise from the respondent's belief that she may never have to pay for the implemented water management alternative. Formally, the subjective probability of payment being coerced is given by $0 < p \leq 1$ and so the expected utility from selecting a particular water management alternative is $E[U] = E[V] + \epsilon$, where the subscripts i, j and t have been dropped for ease of exposition and where

$$E[V] = \beta\mathbf{X} - p\beta_c C. \tag{3}$$

In Equation (3), $\boldsymbol{\beta}$ is a vector of preference parameters associated with the different levels of each attribute in $\mathbf{X}$ and C is the cost of the selected alternative, evaluated in perpetuity. It follows that the annual cost of the selected alternative is given by δC, where $\delta > 0$ is the rate of discount.[17] The term $p\beta_c$ is the marginal utility of income under payment uncertainty. The higher is the probability that payments for the implemented alternative will be coerced the higher is the marginal utility of income, whereby certainty about payment coercion $p = 1$ yields the deterministic marginal utility of income β_c.

The marginal WTP for an improvement in attribute x is given by

$$MWTP_x = \frac{\beta_x}{p\beta_c}, \tag{4}$$

where β_x is the preference parameter associated with attribute $x \in \mathbf{X}$. Equations (3) and (4) show that bias in this framework manifests itself in

[17] This specification assumes positive discounting. Alternatively, one could think of δ as $1/N$ where N is the number of years a respondent expects to pay for the selected alternative.

the form of a lower marginal utility of income, causing the marginal willingness to pay for each attribute to be inflated by $\frac{1}{p}$.

We envisage the salient treatment to operate via two channels. Firstly, requiring the respondents to pay for the selected choice in the survey weakly increases the credibility of the coercive payment vehicle: the subjective probability of having to pay for an implemented water management alternative is $p \leq p_s \leq 1$ for respondents in the salient treatment. The second channel is linked to our use of a proxy good. Our non-market valuation study is typical for many studies of this kind in that its purpose is to inform the future provision of a currently non-existent public good. Introducing a salient treatment in this setting requires the identification of a proxy public good with closely matched attributes that can be scaled as a result of the one-off payments made by the treated respondents. In our case, the proxy good is a local pilot water management project. Introducing such a proxy, however, may lead to distortions to the extent that the benefits from the pilot project may not always match perfectly the attributes in the selected choice set. Hence, a treated respondent who selects a water management option other than the status quo incurs the disutility from having to pay immediately δC and receives in return the increase in the pilot project benefits that result from the payment made. Expressed formally, $E[V]$ for respondents in the treatment group is equal to

$$E[V] = \beta \mathbf{X} - p_s \beta_c C + p_s \delta (\beta_z z - \beta_c) C, \tag{5}$$

where β_z is the marginal utility of provisions through the pilot project and where it is assumed that the production of the proxy good takes the simple form $z \times C$. Here, z, can be thought of as the constant marginal provision per dollar spent. The last term in Equation (5) allows for this potential discrepancy between the disutility from immediate payment and the increased benefit from the pilot project evaluated over the course of one year. Equation (5) gives rise to a marginal utility of income for the treated respondent equal to

$$MU_I = p_s \beta_c + p_s \delta (\beta_c - \beta_z z). \tag{6}$$

Taking the difference between the marginal utility of income of the treated respondent and that of a respondent in the control group and dividing by the marginal utility of income of the latter yields the relative treatment effect.

$$\theta = \frac{p_s}{p} - 1 + \frac{p_s \delta}{p} \left(1 - \frac{\beta_z z}{\beta_c} \right). \tag{7}$$

Equation (7) shows that the magnitude of the effect of our salient treatment is a function of the extent to which the treatment reduces bias, $\frac{p_s}{p} - 1$, the marginal benefit–cost ratio from the pilot project, $\frac{\beta_z z}{\beta_c}$, and the discount rate, δ.

Solving (7) for $\frac{1}{p}$, yields

$$\frac{1}{p} = \frac{\theta + 1}{p_s\left(1 + \delta\left(1 - \frac{\beta_z z}{\beta_c}\right)\right)}, \tag{8}$$

which is the bias that is implied by a given treatment effect, θ, and by given values of p_s, δ and $\frac{\beta_z z}{\beta_c}$. To be precise, $\frac{1}{p}$ is the factor by which willingness to pay estimates in the control group are inflated due to the respondents' belief that payments will be coerced with probability less than one.

Closer inspection of (8) reveals that the greater is the observed treatment effect, the greater is the implied bias, $\frac{\partial(1/p)}{\partial\theta} > 0$, *ceteris paribus*. Equation (8) also shows that the implied bias is increasing in the benefit–cost ratio of the proxy project, $\frac{\partial(1/p)}{\partial(\beta_z z/\beta_c)} > 1$.

With regard to the relationship between the bias, $\frac{1}{p}$, and p_s, we expect the salient treatment to at least weakly increase the respondent's belief that payment for the selected alternative will be coerced, that is $p_s \geq p$. Using Equation (8), it can be shown that this condition holds as long as $\theta \geq \delta\left(1 - \frac{\beta_z z}{\beta c}\right)$. That is, irrespective of the assumed benefit–cost ratio from the pilot project, p_s is always greater than p as long as θ is greater than the discount rate. Moreover, greater effectiveness of the salient treatment, represented by a higher coercive payment probability p_s, reduces the bias that is implied by a given treatment effect, $\frac{\partial(1/p)}{\partial p_s} < 0$. Hence, the lower bound of the bias that is implied by the salient treatment is given by $\frac{1}{p} = \frac{\theta+1}{\delta+1}$. It represents the case where the salient treatment mitigates all payment uncertainty, $p_s = 1$, and the one-off payment to the pilot project yields no benefits for the respondent, $\beta_z z = 0$.[18]

Finally, in the ideal case where there is no distortion from the pilot project and our treatment is fully effective, that is $\frac{\beta_z z}{\beta_c} = 1$ and $p_s = 1$, the bias that is implied by an observed treatment effect is given by $\frac{1}{p} = \theta + 1$.

4. Data

We commence with the presentation of the descriptive statistics of our data set. Panel (a) of Figure 1 shows the income distribution in the sample. Many individuals refused to provide detailed information about their income, but did provide information on their general income category as seen in panel (b) of Figure 1. Since income is an important driver in our main results, we focus

[18] The theoretical framework presented here assumes that the subjective probability of payment coercion that results from the salient treatment, p_s, applies in the same way to the one-off payment as to future payments. It can be shown that relaxing this assumption has no effect on the lower bound of the implied bias.

on the general income categories to avoid losing a considerable proportion of the data. Moreover, the respondents' perception of which income category they fall into is perhaps a more appropriate determinant of their WTP for an improvement in environmental quality. This captures their subjective income and likely incorporates their gross income relative to household expenses. Bertrand and Mullainathan (2001) find subjective data are useful at explaining behaviour across individuals, as it is being used in our context.

In addition to standard demographic data, we ask questions about environmental preferences and activities that are likely to affect the

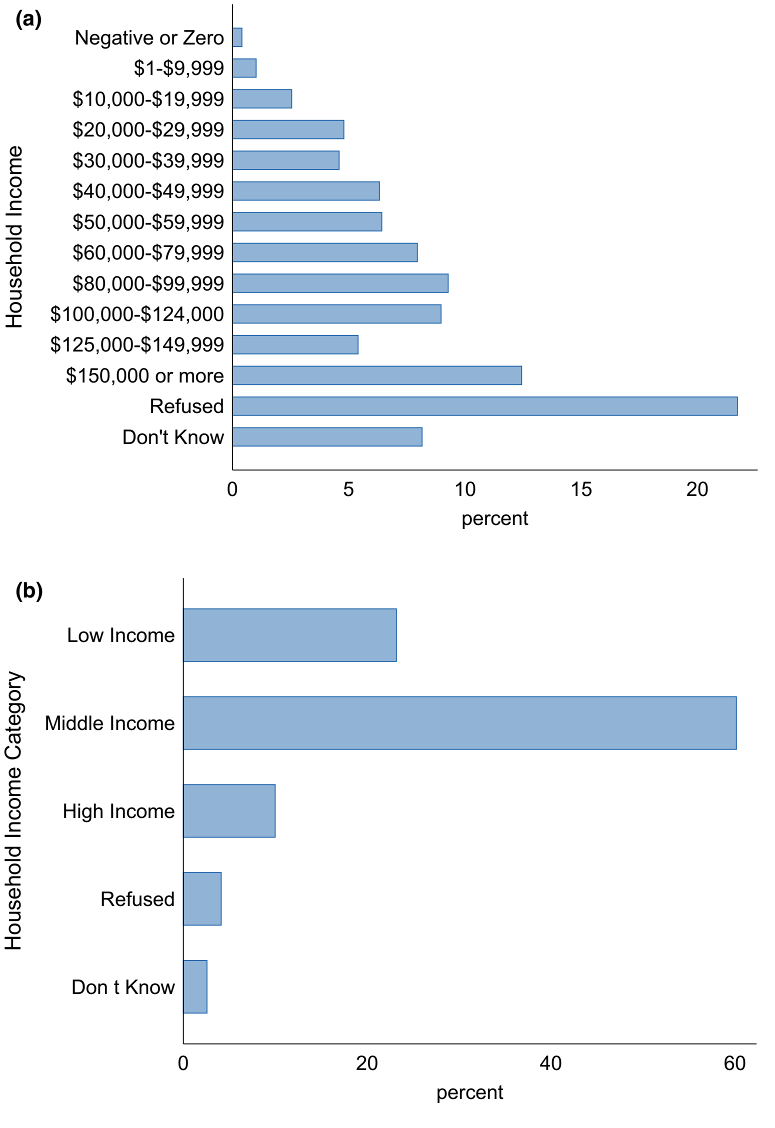

Figure 1 Income in the sample. [Colour figure can be viewed at wileyonlinelibrary.com]

Table 1 Balance on observables

	Mean_C	N_C	Mean_T	N_T	Difference	Std. Difference
Low Income	0.25	597	0.25	309	−0.00	−0.00
Medium Income	0.66	597	0.62	309	0.04	0.07
High Income	0.10	597	0.13	309	−0.03	−0.11
Age	55.08	647	53.33	332	1.75	0.11
Female	0.46	647	0.48	334	−0.03	−0.05
Nature	0.38	647	0.36	334	0.02	0.04
Restrict	0.24	647	0.21	334	0.03	0.08
Water Quality	0.35	647	0.36	334	−0.01	−0.03
Flood	0.31	630	0.33	323	−0.02	−0.05
Summer Heat	0.50	640	0.59	333	−0.09	−0.17

Notes: The columns show the means and samples sizes for relevant demographic and attitudinal variables for both the salient group and the non-salient group, as well as the difference in means and the standardised difference in means. All variables except age are indicator variables and the means are sample proportions, and age is measured in years.

willingness to contribute to a water management project. These questions include whether individuals engage in nature activities (*Nature*), if they are currently facing watering restrictions (*Restrictions*), their concern for water quality (*Water Quality*), if they think a flash flood is likely or if they have experienced a flood (*Flood*) recently and whether they are concerned about increasing summer temperatures (*Summer Heat*). Some of these variables have multiple levels that we collapse into binary indicators that represent a natural division of the variable of interest.[19] This saves degrees of freedom in the estimation while still incorporating important information into the regressions. Table 1 displays the means and sample sizes of demographic and attitudinal variables for both the treatment and control group. Following the advice of Deaton and Cartwright (2018), we do not report statistical tests of balance, but rather show absolute and standardised differences in means. The raw and standardised differences are all low.

5. Nonparametric choice analysis

We first examine differences in the cost of selected alternatives between the treatment and control groups for subsections of the sample. By examining the raw choice data, we can observe if there are differences in behaviour between the treatment and control groups. In particular, we focus on the cost of alternatives that respondents select since treatment primarily impacts the sensitivity to cost. The decisions for each respondent are likely to be highly correlated so we average the costs of selected alternatives across all choice sets for each respondent to examine differences in average costs across treatment

[19] For example, an indicator for concern over water quality takes on a value of one if the answer to the question of whether there is a need to be concerned over water quality was “very much reason” or “quite a lot of reason” and is set equal to zero if the responses are “not very much reason” or “no reason at all.”

Table 2 Cost of selected alternatives across treatment status

	Control	N_C	Treatment	N_T	Difference	*p*-value
(a) Income: Average selected cost						
All	13.76	647	13.13	334	0.63	0.0939
Low Income	12.71	148	10.97	77	1.74	0.0411
Medium Income	14.24	395	13.90	195	0.34	0.3616
High Income	15.14	57	16.01	40	−0.87	0.7468
Low & Med Income	13.63	590	12.74	294	0.89	0.0498
(b) Income: Probability of choosing status quo						
All	0.22	647	0.25	334	−0.03	0.0924
Low Income	0.28	148	0.36	77	−0.08	0.0724
Medium Income	0.19	395	0.21	195	−0.01	0.7206
High Income	0.13	57	0.10	40	0.03	0.3137
Low & Med Income	0.23	590	0.27	294	−0.04	0.1034
(c) Income: Intensive margin contribution						
All	16.72	505	16.36	253	0.36	0.1259
Low Income	16.60	109	15.35	48	1.24	0.0826
Medium Income	16.85	316	16.61	156	0.24	0.2464
High Income	16.85	48	16.50	38	0.35	0.5535
Low & Med Income	16.70	457	16.34	215	0.37	0.1711
(d) Demographic variables						
Nature	14.80	244	14.75	119	0.05	0.8033
Children	13.83	204	13.83	121	−0.00	0.7542
Female	14.21	295	12.65	161	1.56	0.0471
Water Quality	15.74	224	15.15	120	0.58	0.1478

Notes: The columns show the average cost of selected alternatives for the Treatment group and the Control group as well as the difference in means and the *p*-value from a non-parametric Mann–Whitney test. The rows designate different subsections of the sample across key demographic variables.

assignment with the respondent as the unit of observation.[20] This allows us to exploit the randomisation without imposing distributional assumptions on the preference parameters in a discrete choice model. Since the treatment and control groups receive the same choice sets, variation in the attributes is differenced out.

The panels of Table 2 present several specifications of the average choice decisions. Each panel shows the means for the variable of interest across treatment status, samples sizes for each group, the difference in means and the *p*-value from a non-parametric Mann–Whitney test.[21] Panel (a) shows the

[20] We perform two robustness checks for the analysis of the average costs of selected alternatives. First, we use median costs instead of average costs. Second we regress the costs of selected alternatives on treatment, which captures the panel structure of the data. The regressions also control for differences in the costs of projects presented to the respondents, though the average presented project costs do not vary substantially, taking on values of A$17, A$17.75 and A$18.25 in the different choice set blocks. The results are consistent in terms of both magnitudes and statistical significance so we focus on average costs of selected alternatives.

[21] The distribution of costs is bi-modal due to a mass at zero which represents the status quo, and therefore, a t-test that assumes normality is not appropriate.

effect of treatment on the average selected cost by income level. For the whole sample, the treatment group chose projects that cost A$0.63 less than the control, representing approximately 5 percent of the average cost, and the Mann–Whitney test rejects equality of the two distributions at the 10 percent level. This shows a noticeable treatment effect even at a highly aggregated level.

We further analyse the average selected cost by respondents' income to assess heterogeneity in the treatment effect. The salient treatment primarily affects decisions for low-income respondents whose average selected alternative costs A$1.74 lower compared to the control, a 14 percent difference that is statistically significant at the 5 percent level. Medium-income households receiving the treatment select slightly cheaper alternatives and the high-income households actually choose more expensive alternatives, although neither difference is statistically significant. Since the treatment actually induces the high-income group to select more expensive projects we also restrict the sample to low and medium income; and the Mann–Whitney test rejects the null of equal distributions at the 5 percent level.

We envision the salient treatment working on two margins: the extensive margin represents the probability of choosing some positive payment over the status quo, which has zero cost. The intensive margin reflects the cost of a selected project conditional on paying some non-zero amount. Panel (b) shows the effect of treatment on the probability of selecting the status quo option separated by income group. The results show small differences in the probability of choosing the status quo across treatment for all income levels; the whole sample and low-income subsample are significant at the 10 percent level while tests for the other subsamples cannot reject the null. The signs are consistent with the results for the pooling both margins as presented in panel (a).

To investigate the effect on the intensive margin, we restrict the sample to decisions where the respondent did not choose the status quo. Conditional on selecting a project that offers some improvement, we test for the impact of treatment on the *cost* of the selected alternative. The results, displayed in Table panel (c), show that all groups decrease the conditional size of the project, which is the expected result. In particular, the high-income subsample, conditional on non-zero contributions, selects cheaper projects if the resulting costs are immediately paid by the respondent.[22]

Next, we examine other demographic and attitudinal variables that we expect to drive differences in selected costs between the treatment and control groups. Panel (d) shows the differences in average selected costs by treatment across several demographic variables. Treated respondents who engage in nature and have children do not reduce their average selected costs. These groups have stronger preferences for water management as evidenced in

[22] This finding is consistent with the interviewer effect that respondents want to please the interviewer and not appear stingy. So when respondents from high-income households know, they are paying with their own money they are more likely to contribute, but at lower levels. Only 9% of the sample are high income so we lose substantial statistical power for hypothesis tests of this subgroup.

exploratory analysis where both of these variables reduce the probability of selecting the status quo. This is consistent with the hypothesis that these groups are already predisposed to value the benefits of water management and the monetary treatment does not affect their average selections. Of all the demographic variables, the treatment has the largest difference across gender, with women reducing their average selected cost by A$1.49. This result is intriguing because on average female respondents are less likely to choose the status quo option. So while the treatment magnifies the effect of low-income respondents choosing cheaper alternatives, the treatment also mitigates the effect of women choosing more expensive alternatives. Concern for water quality has a smaller and insignificant reduction in average selected cost for those in the treatment group.

We also test for endowment effects by comparing the average selected cost in the treatment group among different levels of the initial endowment.[23] None of the differences in means are statistically different from each other, and there is no monotonic relationship between the endowment and the difference in selected costs. In the next section, we estimate a discrete choice model to investigate the impact of treatment on preference parameters.

6. Regression framework

6.1 Econometric model

In the random utility model described by Equations 1 and 2, if it is assumed that ϵ_{ijt} follows a type I extreme value distribution then the choice probabilities can be modelled in the logit specification shown in Equation 9.[24]

$$pr\left(Y_{ij}=j\right)=\frac{\exp\left(V_{ijt}\right)}{\sum_{h\in J}\exp\left(V_{iht}\right)} \tag{9}$$

In our setting, the respondents select one of three alternatives from each choice set, requiring a model that accommodates multiple categories. Based on the results of a Hausman test (Hausman & McFadden, 1984), we reject that the IIA assumption on restrictions of substitution patterns holds in our setting and therefore eliminate the standard multinomial logit as a valid econometric model. Our preferred specification is the mixed logit (MXL), which McFadden and Train (2000) show can accommodate any set of substitution patterns.[25] Additionally, the MXL model is popular in the

[23] Table A2 in the Appendix S1 shows the difference in average selected cost for the salient group separated by the initial endowment.

[24] As is common practice in empirical analyses, the cost of the alternative, C_{jt}, is preceded by a positive sign in the function to be estimated; the marginal utility of income is therefore the parameter defined by $-1 \times \beta_c$.

[25] We also estimate a nested logit with the nests as the status quo and the two non-status quo options that produces similar results.

applied literature estimating WTP from discrete choice experiments; see among others Revelt and Train (1998); Train (1998); Greene and Hensher (2003); Hensher et al. (2005); Balcombe et al. (2011). The MXL also allows for individual level heterogeneity by estimating a distribution of parameters across the individuals in the sample. The mixed logit has random coefficients and the probability that respondent i selects alternative j for choice t is

$$P_{ijt} = \int_{\beta} \prod_{t=1}^{t=10} \frac{\exp\left(\beta_n' X_{\mathrm{int}}\right)}{\exp \sum_j \exp\left(\beta_n' X_{jnt}\right)} f\left(\mu_\beta \Omega_\beta\right) dt. \tag{10}$$

The joint unconditional choice probability of a panel of observed choices is therefore the weighted average of the product of the conditional probabilities of choice, where the weights are the densities of the random draws. The integral representing the probability of the sequence of each respondent's choices does not have a closed form and therefore the estimates are approximated through numerical simulation (Train, 2009).

6.2 Regression results

The results from the base regression can be found in column (1) of Table 3. The level of each attribute is modelled as a dummy variable equal to one if that attribute level is present for a given alternative within a choice set. We pool flood-never and flood-half as well as recreation-high and recreation-medium. Each regression model has two columns for the mean (mean) and standard deviation (SD) of the random parameters. Fixed parameters are reported in the mean column and do not have standard deviations. Standard errors, clustered at the respondent level, are reported in parentheses below the parameter.

The attributes are all modelled as random parameters and the mean of the distribution for each attribute has the expected sign with the exception of flood protection.[26] It is important to note the substantial heterogeneity in preferences for both the attributes and cost, as evidenced by the large standard deviations. Respondents prefer a water management alternative to the status quo, all else being equal.

We model the cost parameter as normally distributed, which is not standard practice since it allows for probability mass in the positive region. While this is not ideal in general we believe, it is appropriate in our setting. We tested other specifications such as fixed costs, interacting cost with income and modelling cost as log-normally distributed. Each of these provided

[26] One explanation why flood protection may not be desirable for respondents is that this refers to flash floods as opposed to large-scale flooding. Many residents are likely to undertake averting behaviour through the purchase of private goods such as flood insurance, elevated housing and sandbags. Thus, they may not see a role for their local council in reducing the likelihood of flash floods.

Table 3 Cost and treatment interactions

	None		Salient Interactions					
	(1) Base		(2) Cost		(3) Cost*Income		(4) Attributes	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Status Quo	−0.9119***		−0.9114***		−0.9258***		−0.9135***	
	(0.1237)		(0.1237)		(0.1285)		(0.1237)	
Cost	−0.0221***	0.0929***	−0.0186***	0.0928***	−0.0115**	0.0872***	−0.0221***	0.0927***
	(0.0041)	(0.0057)	(0.0047)	(0.0057)	(0.0054)	(0.0055)	(0.0041)	(0.0057)
Restrictions 3,4	0.3530***	0.8336***	0.3530***	0.8337***	0.3732***	0.8403***	0.3344***	0.8338***
	(0.0658)	(0.0775)	(0.0658)	(0.0774)	(0.0680)	(0.0804)	(0.0786)	(0.0776)
No Restrictions	0.2794***	0.7290***	0.2794***	0.7284***	0.2882***	0.7407***	0.3161***	0.7258***
	(0.0613)	(0.0782)	(0.0613)	(0.0783)	(0.0632)	(0.0800)	(0.0732)	(0.0789)
Flood Protection (Both)	−0.1897***	0.5464***	−0.1899***	0.5465***	−0.2159***	0.5378***	−0.1801**	0.5485***
	(0.0587)	(0.0874)	(0.0587)	(0.0876)	(0.0605)	(0.0900)	(0.0700)	(0.0872)
Stream High	0.3016***	0.5251***	0.3015***	0.5257***	0.2761***	0.5048***	0.3532***	0.5333***
	(0.0685)	(0.0775)	(0.0685)	(0.0773)	(0.0692)	(0.0814)	(0.0813)	(0.0768)
Stream Medium	0.2748***	0.6989***	0.2751***	0.6968***	0.2524***	0.6649***	0.3791***	0.6897***
	(0.0738)	(0.0724)	(0.0737)	(0.0726)	(0.0749)	(0.0762)	(0.0886)	(0.0735)
Recreation (Both)	0.0644	1.2209***	0.0645	1.2211***	0.1205*	1.2119***	0.0531	1.2241***
	(0.0611)	(0.0571)	(0.0611)	(0.0571)	(0.0630)	(0.0599)	(0.0749)	(0.0572)
Temp −2	0.0772*	0.7757***	0.0774*	0.7753***	0.0865**	0.7683***	0.0860*	0.7772***
	(0.0418)	(0.0572)	(0.0418)	(0.0572)	(0.0429)	(0.0587)	(0.0509)	(0.0573)
Cost*Salient			−0.0101					
			(0.0073)					
Low Income*Cost					−0.0218*			
					(0.0112)			
High Income*Cost					0.0137			
					(0.0128)			
Cost*Salient*Low Income					−0.0279*			
					(0.0167)			
Cost*Salient*Med Income					−0.0053			
					(0.0088)			

Table 3 (*Continued*)

	None		Salient Interactions					
	(1) Base		(2) Cost		(3) Cost*Income		(4) Attributes	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Cost*Salient*High Income					0.0130 (0.0148)			
Flood (Both)*Salient							−0.0276 (0.1165)	
Restrictions 3,4*Salient							0.0541 (0.1246)	
Restrictions None*Salient							−0.1081 (0.1217)	
Stream medium*Salient							−0.3037** (0.1348)	
Stream High*Salient							−0.1535 (0.1284)	
Recreation (Both)*Salient							0.0330 (0.1130)	
Temp −2*Salient							−0.0259 (0.0867)	
BIC/N	18		18		18		18	
AIC/N	18		18		18		18	
Observations	9,774		9,774		9,060		9,774	
Individuals	981		981		906		981	

Notes: All regressions are mixed logit model with random coefficients. All random coefficients are normally distributed. Each regression has two columns: Mean and SD that refer to the mean and standard deviation of the random parameters. Fixed coefficients have no standard deviation. Significance levels are based on standard errors clustered at the respondent level that are reported in parentheses below the parameter estimate. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

anomalous results such as positive and or insignificant cost parameters. This is likely because the survey was designed in an interdisciplinary team and the cost levels were based on realistic levels informed by policymakers. These cost levels were likely too low or did not have high enough levels to more precisely estimate the cost parameter. We provide the specification tests for the base model (column (1) of Table 3) in Table A4 in the Appendix S1.

We model the random variables as uncorrelated. While off diagonal elements of the covariance matrix are individually and jointly significant, the overall model fit is only slightly better as measured by the AIC and BIC per respondent. Additionally, none of the attributes are statistically significant in the correlated model. For these reasons, we prefer the uncorrelated model. The correlated model is presented in column (2) of Table A4.

Columns (2)–(4) in Table 3 include interaction terms with the treatment modelled as fixed coefficients. The interaction of cost and treatment shows how the treatment impacts the respondents' sensitivity to the project cost, also interpreted as the marginal utility of income. We only present the interaction with cost, but models that replace the interaction terms with status quo yield similar results. Additionally, the interaction of treatment and cost captures changes on both the extensive and intensive margin because the status quo option always has a zero cost. The interaction term in column (2) is negative but not statistically significant. Despite the lack of precision, the treatment effect is over half the magnitude of the mean of the cost coefficient, indicating that the estimated marginal utility of income increases by approximately 50 percent in the treatment group. Since willingness to pay is inversely scaled by the marginal utility of income, higher estimates for the marginal utility of income reduce the willingness to pay in the treated group relative to the control.

We test the heterogeneity in treatment impact on the estimates of marginal utility by income level to build on our results in Section 5. Similar to the raw choice data, we find that in magnitude the treatment effect is largest for the low-income group, and the point estimate for the high-income respondents is positive. The low-income group is statistically significant at the 10 percent level and also statistically different from the treatment effect for the income group at the 10 percent level ($p = 0.07$). One reason why it is difficult to investigate treatment heterogeneity with respect to cost in a mixed logit model is that we need to include substantially more interactions with the cost attribute; consequently, the standard errors on the treatment interactions with the low- and high-income respondents are roughly twice as large as the base treatment effect.

In our setting, there was no specific project to be built, and the primary focus was estimating values for the benefits that various stormwater management projects could provide. The econometric analysis thus far (columns [1] - [3] of Table 3) has implicitly assumed that the response to the treatment is independent of the attributes in the choice set and only impacts the decisions related to program cost. Our second hypothesis tests this assumption by examining if the treatment changes preferences for the attributes.

We test for treatment effects within the attributes of the choice set by interacting the salient treatment dummy with each attribute. We find that most of the signs are insignificant except for medium stream health, which is significant at the 10 percent level. A Wald test for joint significance of the interaction terms fails to reject the null at the 10 percent level. With the exception of stream health, our result that the treatment does not affect most attributes is in line with our second hypothesis and also consistent with the finding of List et al. (2006) that marginal willingness to pay for attributes is not susceptible to hypothetical bias. While List et al. (2006) only examined a private good, our setting must be viewed in the context of valuing attributes of a quasi-public good that provides different types of benefits. The fact that most preferences for attributes are statistically indistinguishable across treatment status provides evidence that the treated respondents' focus is on the attributes of the proposed water management alternatives and not on the attributes of the pilot project.

We consider several robustness checks to consider the possibility that the optimisation achieved a local, as opposed to global, maxima. The first check simply replicates Table 3 using different seeds for the random number generator. As shown in Appendix S1 in Table A5, the results are identical. We also consider the role of starting values in global convergence. By default, the mixed logit regression uses starting values from the conditional logit. We use our base treatment effect regression in column (2) of Table 3 as our starting values and add random disturbance terms. We draw the disturbance terms from uniform distributions with limits of $\left[-\overline{\beta}_k\kappa,\ \overline{\beta}_k\kappa\right]$, where $\kappa \in 0.25,\ 0.33,\ 0.5,\ 1$. This allows the disturbance term to be proportional to the mean of the original parameter value, with larger values of κ allowing greater deviations from the original starting values. The results, presented in Table A6, are consistent with original results in Table 3.

6.3 Interpretation of treatment effects

Equation (8) in Section 3 allows us to interpret the differences in the estimated marginal utilities of income between the control and treatment groups for given assumptions about the efficacy of our treatment and the distortion that is introduced by the proxy good.

Model 3 in Table 2 shows that our treatment has a significant effect on low-income respondents. Taking into account the interaction terms, the observed treatment effect for low-income respondents is $\theta = \frac{-0.0279}{-0.0115-0.0218} = 0.84$ relative to non-treated low-income respondents. The standard error for θ is 0.68 and the 95 percent confidence interval is [−0.5, 2.18]. The bias that is implied by the observed treatment effect is shown graphically in Figure 2, where the implied bias, $\frac{1}{p}$, is plotted against the benefit–cost ratio of the pilot project, $\frac{\beta_z z}{\beta_c}$, for a discount rate of $\delta = 0.07$. The solid, dashed and dotted lines

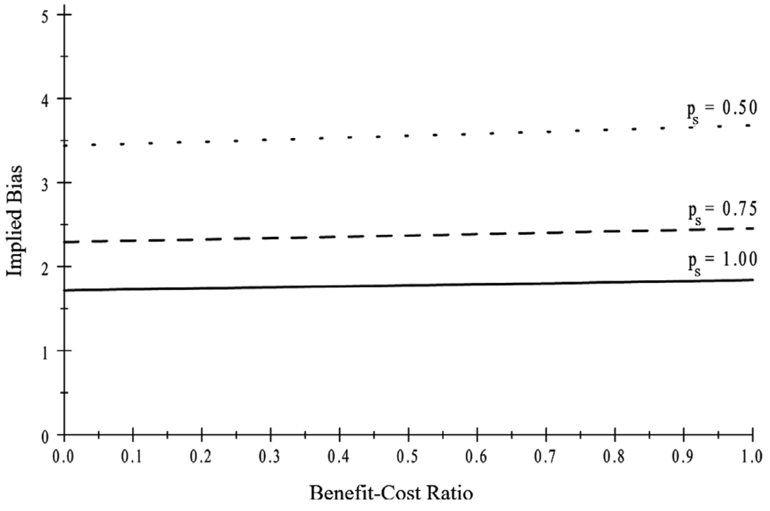

Figure 2 Implied bias.
Notes: Magnitude of bias, defined as 1 = p in Equation (8) that is implied by an observed treatment effect of $\theta = 0.84$ for subjective treatment-induced coercive payment probabilities of 0.5, 0.75 and 1, over the range of values of the benefit-cost ratio for funding the pilot project ($\beta_z z/\beta_c$) between 0 to 1 and a discount rate of $\delta = 0{:}07$.

assume, respectively, decreasing levels of treatment effectiveness as characterised by average subjective payment probabilities in the treatment group of $p_s = 1.0$, $p_s = 0.75$ and $p_s = 0.5$.

Our best-case scenario, where the salient treatment eradicates all payment uncertainty, $p_s = 1$, and there is no distortion from the pilot project, $\frac{\beta_z z}{\beta_c} = 1$, implies that willingness to pay estimates of low-income respondents in the control group are inflated by a factor of 1.84. This bias increases if we allow for the treatment to work less than perfectly. For example, assuming that treated respondents believe that there is 50–50 chance of having to pay for the selected alternative, implies a bias of 3.68. Under these assumptions, willingness to pay would be 3.68 times higher among the non-treated low-income respondents than their treated counterparts.

In contrast, the substitutability of the pilot project for the selected alternative has very little impact on the bias that is implied by our observed treatment effect. Assuming $p_s = 1$ (solid line in Figure 2), the implied bias ranges from 1.72 when $\frac{\beta_z z}{\beta_c} = 0$ to 1.84 when $\frac{\beta_z z}{\beta_c} = 1$. Hence, under the assumptions of the model, the minimum bias that is implied by our observed treatment effect is 1.72.

Overall, Figure 2 shows that assumptions regarding the effectiveness of the salient treatment in terms of p_s have a bigger impact on the implied bias than any distortion that was introduced by the proxy good.

In the interest of completeness, we now turn to other features of our experimental design that could potentially lead to differences in the control and treatment groups. One of these relates to the potential effects of the initial cash endowment on the respondents' choices. First, respondents could consider the initial endowment as additional income, affecting the respondents' budget constraints. This effect would be relatively more important for low-income households. However, easing the budget constraint should lead to a greater WTP among treated lower income households and not, as we observe, a reduction in WTP. Second, respondents might consider the cash endowment as house money, which they value less than their personal income. If respondents value endowed money less than earned money, we should see a difference between the "Endowed Salient" and the "Earned Salient" groups. However, there are no statistically significant differences in the contributions between those sub-groups. More importantly, a house money effect in our setting would manifest itself in the form of a lower aggregate cost sensitivity of respondents in the treatment group, while we find the opposite effect. In this context, our results can be interpreted as a lower bound.

The treatment could be considered a donation to the pilot project and may induce an incentive to free ride. This effect would be captured in our theoretical framework as free riding respondents would have a low value for the marginal benefits from their contribution to the pilot project. However, as shown in Figure 2, the majority of the bias that is implied by the treatment effect persists even when no benefits are derived from the pilot project.

Alternatively, the treatment could affect the incentive compatibility of the survey design. For example, wealthier respondents in the treatment group might perceive that relatively poorer respondents in the treatment group are now less likely to opt for a non-status quo option due to the relatively higher costs induced by the treatment. Therefore, they might act strategically and opt for a less costly option to influence the aggregate outcome. We do not find this in our data as treatment does not increase sensitivity to cost for high-income respondents. However, it is difficult, if not impossible, to account for all the potential effects on strategic behaviour.

We find these alternative explanations less plausible than those that are explicitly accounted for in our theoretical framework. Although we cannot fully discount that some of the treatment effect may be due to channels outside this stylised framework, the evidence of the existence of bias among certain respondents is compelling and should be considered by policymakers who are concerned with distributional effects of projects with non-market benefits. In particular, if low-income households, who generally have lower WTP, overstate their WTP, a uniform tax increase based on average WTP will exacerbate the negative welfare effects on the low-income population.

7. Conclusion

Stated preference surveys are used to estimate the values the public places on non-market goods and often provide the basis for the design, scope and scale of public and quasi-public good provision. For surveys to be informative, individuals need to reveal their true preferences, which has been a topic of much debate in the context of valuing non-market goods. We introduce a method that involves financially incentivising survey respondents' choices and illustrate this method through a discrete choice experiment designed to estimate the benefits of water management policies in Australia. In a door-to-door survey with 981 individuals, a group of respondents was randomly assigned a salient treatment. The treatment group received a monetary endowment prior to the actual choice task. One of their choices was then randomly selected and the cost associated with this choice was deducted from their initial monetary endowment. The money collected from the respondents was used to implement a water management project in their community. The treatment therefore ensures that the payment for the good is realised, hence improving salience in the cost of stated choices.

We find an economically and statistically significant treatment effect among low-income respondents. This effect is significant in the non-parametric and econometric analyses. In particular, the treatment increases estimates of the average marginal utility of income of low-income respondents by 84 percent. We examine the channels through which we expect the treatment to work within a stylised utility framework in a stated preference setting, which allows us to interpret the treatment effect. We find evidence of bias, inflating the marginal willingness to pay for low-income respondents by a factor of at least 1.72. While other mechanisms, not explicitly accounted for in the theoretical framework, may also contribute to differences across treatment status, we find these explanations less compelling.

As we are interested in understanding the effect of the treatment on both the extensive and intensive margin, the mechanism we designed involves three options. This is useful when utilising the survey results for scoping or refining features of a quasi-public good project. We show that the saliency method can help improve our understanding of the willingness to pay for such goods.

Our findings suggest several avenues for future research. For instance, methodologically, it would be useful to examine the robustness of our treatment effects using a survey that involves a single binary choice question, as this elicitation format is incentive compatible under weaker assumptions relative to the format employed in this study. It would also be interesting to compare different approaches to mitigate non-truthful preference revelation using field experiments. In particular, the saliency method could be evaluated against a binding referendum, expert advice or an intervention using a cheap talk script. Field experiments of the kind discussed in this paper can be often challenging to implement as they are placed within a natural environment that typically imposes many constraints (Brent et al, 2016; List 2011). Some of the bottlenecks we faced led to broadly defined projects and overall low sensitivity to costs. Therefore, while our study identifies the role played by economic incentives when eliciting values for the general benefits of a water management policies, it would be important to also examine whether they play a similar role when projects and benefits are more precisely defined.

Conflict of interest

The authors have no conflict of interest to declare.

Data availability statement

The data and code underlying this paper can be accessed at https://github.com/praschky/ajaere_2021.

References

Andersen, S., Harrison, G.W., Lau, M.I. & Rutström, E.E. (2006) Elicitation using multiple price list formats. *Experimental Economics*, 9(4), 383–405.

Balcombe, K., Burton, M. & Rigby, D. (2011) Skew and attribute nonattendance within the Bayesian mixed logit model. *Journal of Environmental Economics and Management*, 62(3), 446–461.

Bertrand, M. & Mullainathan, S. (2001) Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91(2), 67–72.

Börger, T., Abate, T.G., Aanesen, M. & Zawojska, E. (2021) Payment and policy consequentiality in dichotomous choice contingent valuation: Experimental design effects on self-reported perceptions. *Land Economics*, 97(2), 407–424.

Brent, D.A., Friesen, L., Gangadharan, L. & Leibbrandt, A. (2017) Behavioral Insights from Field Experiments in Environmental Economics. *International Review of Environmental and Resource Economics*, 10(2), 95–143. http://doi.org/10.1561/101.00000084

Brent, D.A., Gangadharan, L., Lassiter, A., Leroux, A. & Raschky, P.A. (2017) Valuing environmental services provided by local stormwater management. *Water Resources Research*, 53(6), 4907–4921.

Bulte, E., Gerking, S., List, J.A. & De Zeeuw, A. (2005) The effect of varying the causes of environmental problems on stated WTP values: evidence from a field study. *Journal of Environmental Economics and Management*, 49(2), 330–342.

Carlsson, F. & Martinsson, P. (2001) Do hypothetical and actual marginal willingness to pay differ in choice experiments?: Application to the valuation of the environment. *Journal of Environmental Economics and Management*, 41(2), 179–192.

Carson, K.S., Chilton, S.M., George Hutchinson, W. & Scarpa, R. (2020) Public resource allocation, strategic behavior, and status quo bias in choice experiments. *Public Choice*, 185 (1), 1–19.

Carson, R.T., Flores, N.E., Martin, K.M. & Wright, J.L. (1996) Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. *Land Economics*, 72(1), 80–99.

Carson, R.T. & Groves, T. (2007) Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37(1), 181–210.

Carson, R.T., Groves, T. & List, J.A. (2014) Consequentiality: a theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, 1(1), 171–207.

Champ, P.A., Flores, N.E., Brown, T.C. & Chivers, J. (2002) Contingent valuation and incentives. *Land Economics*, 78(4), 591–604.

Cherry, T.L., Frykblom, P. & Shogren, J.F. (2002) Hardnose the dictator. *American Economic Review*, 92(4), 1218–1221.

Cummings, R.G., Elliott, S., Harrison, G.W. & Murphy, J.J. (1997) Are hypothetical referenda incentive compatible? *Journal of Political Economy*, 105(3), 609–621.

Czajkowski, M., Vossler, C.A., Budziński, W., Wiśniewska, A. & Zawojska, E. (2017) Addressing empirical challenges related to the incentive compatibility of stated preferences methods. *Journal of Economic Behavior & Organization*, 142, 47–63.
Deaton, A. & Cartwright, N. (2018) Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2–21.
Diamond, P.A. & Hausman, J.A. (1994) Contingent valuation: is some number better than no number? *The Journal of Economic Perspectives*, 8(4), 45–64.
Greene, W.H. & Hensher, D.A. (2003) A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), 681–698.
Groothuis, P.A., Mohr, T.M., Whitehead, J.C. & Cockerill, K. (2017) Endogenous consequentiality in stated preference referendum data: the influence of the randomly assigned tax amount. *Land Economics*, 93(2), 258–268.
Hanemann, W.M. (1994) Valuing the environment through contingent valuation. *The Journal of Economic Perspectives*, 8(4), 19–43.
Harrison, G.W. (2006) Experimental evidence on alternative environmental valuation methods. *Environmental and Resource Economics*, 34(1), 125–162.
Harrison, G.W. (2006) Hypothetical Bias Over Uncertain Outcomes. In List, J.A. (ed.), *Using Experimental Methods in Environmental and Resource Economics, Chapter 3*, Edward Elgar Publishing, 41–69.
Hausman, J. (2012) Contingent valuation: from dubious to hopeless. *Journal of Economic Perspectives*, 26(4), 43–56.
Hausman, J. & McFadden, D. (1984) Specification Tests for the Multinomial Logit Model. *Econometrica*, 52(5), 1219–1240.
Hensher, D.A., Rose, J.M. & Greene, W.H. (2005) *Applied choice analysis: a primer*. Cambridge, UK: Cambridge University Press.
Herriges, J., Kling, C., Liu, C.-C. & Tobias, J. (2010) What are the consequences of consequentiality? *Journal of Environmental Economics and Management*, 59(1), 67–81.
Hoffman, E., McCabe, K., Shachat, K. & Smith, V. (1994) Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7(3), 346–380.
Holt, C.A. & Laury, S.K. (2002) Risk aversion and incentive effects. *The American Economic Review*, 92(5), 1644–1655.
Johansson-Stenman, O. & Svedsäter, H. (2008) Measuring hypothetical bias in choice experiments: The importance of cognitive consistency. *The B.E. Journal of Economic Analysis & Policy*, 8(1), 1–10. https://doi.org/10.2202/1935-1682.1898
Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R. & Cameron, T.A. et al. (2017) Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), 319–405.
Kling, C.L., Phaneuf, D.J. & Zhao, J. (2012) From Exxon to BP: has some number become better than no number? *Journal of Economic Perspectives*, 26(4), 3–26.
List, J.A. & Gallet, C.A. (2001) What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics*, 20, 241–254.
List, J.A. (2011) Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off. *Journal of Economic Perspectives*, 25(3), 3–16.
List, J., Sinha, P. & Taylor, M. (2006) Using choice experiments to value nonmarket goods and services: evidence from field experiments. *The B.E. Journal of Economic Analysis & Policy*, 5 (2), 1–39. https://doi.org/10.2202/1538-0637.1132
Lloyd-Smith, P., Adamowicz, W. & Dupont, D. (2019) Incorporating stated consequentiality questions in stated preference research. *Land Economics*, 95(3), 293–306.
Loomis, J. (2011) What's to know about hypothetical bias in stated preference valuation studies? *Journal of Economic Surveys*, 25(2), 363–370.
McFadden, D. (1973) Conditional logit analysis of qualitative choice be. In Zarembka, P. (ed.), *Frontiers in Econometrics*. New York: Academic Press, 105–142.

McFadden, D. & Train, K. (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5), 447–470.

Mitani, Y. & Flores, N. (2014) Hypothetical bias reconsidered: payment and provision uncertainties in a threshold provision mechanism. *Environmental & Resource Economics*, 59 (3), 433–454.

Revelt, D. & Train, K. (1998) Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, 80(4), 647–657.

Rogers, A.A., Burton, M.P., Cleland, J.A., Rolfe, J.C., Meeuwig, J.J. & Pannell, D.J. (2020) Expert judgements and community values: preference heterogeneity for protecting river ecology in Western Australia. *Australian Journal of Agricultural and Resource Economics*, 64 (2), 266–293.

Rolfe, J. & Bennett, J. (2009) The impact of offering two versus three alternatives in choice modelling experiments. *Ecological Economics*, 68(4), 1140–1148.

Smith, V.L. (1982) Microeconomic systems as an experimental science. *The American Economic Review*, 72(5), 923–955.

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

Train, K. (2009) *Discrete choice methods with simulation*. Cambridge University Press.

Vossler, C.A., Doyon, M. & Rondeau, D. (2012) Truth in consequentiality: theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, 4(4), 145–171.

Zawojska, E., Bartczak, A. & Czajkowski, M. (2019) Disentangling the effects of policy and payment consequentiality and risk attitudes on stated preferences. *Journal of Environmental Economics and Management*, 93, 63–84.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure A1. Introduction letter.
Figure A2. Explanation document.
Figure A3. Holt and Laury Lottery – example of a decision problem.
Figure A4. Long explanation sheet.
Table A1. Difference in probability of choosing status quo.
Table A2. Cost of selected alternatives by endowment.
Table A3. Balance on observables: exclude high income.
Table A4. Specification tests.
Table A5. Robustness to seed.
Table A6. Robustness to starting values.