Sources of Hypothetical Bias in Public Goods Experiments: A Disaggregated Approach

Julie I. Santos[1], Emi Uchida, Christopher Anderson and Stephen Swallow

Department of Environmental and Natural Resource Economics

University of Rhode Island

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's

2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

***** Do Not Cite *******

Sources of Hypothetical Bias in Public Goods Experiments: A Disaggregated Approach

June 4, 2013

1 Introduction

Research in environmental and natural resource valuation rely heavily on hypothetical survey data to estimate values for public goods. In many cases, the hypothetical survey is the only way to assign monetary values for such goods. However, when doing so, researchers run into complications involving free-riding behavior as well as hypothetical bias. Free-riding occurs because an individual is likely to conclude that by not paying for the provision of the goods and services provided by nature she will not be precluded from benefiting if enough external support exists to provide the good. In such cases, the individual has a strong incentive to ride free on the contributions of others. Freeriding occurs as a consequence of the fact that public goods lack well-defined property rights and as such are non-rival and non-excludable. Overcoming the tendency to free-ride has been a focus in experimental economics and involves designing incentive-compatible elicitation mechanisms.

Hypothetical bias is not restricted to public goods per se, but is a problem associated with hypothetical choice scenarios. Essentially, any difference between stated and revealed valuations (Chang et al 2009[4], Hensher 2010[5]) is termed hypothetical bias. While a number of causes have been identified, a comprehensive theory and definition of hypothetical bias has eluded researchers (Murphy et al 2005[9]). Opaluch and Segerson ([10]) hypothesize that when presented with unfamiliar goods, individuals may not know their willingness to pay precisely but may be able to place the value within an "ambivalence" range. They suggest that when real payments are not required, less cognitive effort is spent exploring the valuation question. Because public goods lack well-defined property rights, they are not directly traded in markets. Therefore, assessing the societal value for such goods has been the motivation of many stated preference studies which utilize hypothetical scenarios to generate a snapshot of public preferences.

There exists a substantial literature on measuring the extent of the hypothetical bias and on investigating potential means of mitigating the problem. Within the environmental economics literature in particular, interest in measuring the severity of hypothetical bias has been waxing. Johnston ([7]) compared a CV survey with subsequent results of a referendum to provide a public water supply in the town of North Scituate, Rhode Island. Aadland and Caplan ([1]) compare responses to a phone survey of communities that did and did not have curbside recycling programs using a cheap-talk script. While most studies find significant evidence of differences between stated and revealed valuations, Carlsson and Martinsson ([3]) found no such evidence in scenarios involving donations for environmental projects. Common culprits include: subject pool variety, differences in information provided across experiments, social norms and whether willingness-to-pay (WTP) or willingness-to-accept (WTA) is being measured (Hensher[5]). However, almost without exception, these studies involve comparisons of aggregated measures of values across experimental settings.

The objective of this study is to further understand the sources of hypothetical bias as pertains to the unique valuation issues that surround public goods provision. Our interest is in tackling this issue at the level of the individual. In this way, we can identify response strategies and test their impacts on behavior in situations that involve real payments for public goods. This is a timely approach as choice modelling research has of late emphasized the importance of accounting for preference heterogeneity. As a by-product, these models permit calculation of individual-level marginal values. A non-parametric latent-class approach to modeling taste variation is used. The choice of models will inform an exploration of the extent to which random utility maximization or some other heuristic is being employed in our hypothetical choice data set. The model yields information about the individual respondent that is hence utilized in the analysis of revealed behavior that manifests in the market experiment. By estimating WTP values for each individual in our sample, we are able to identify strategies in responses such as yeah-saying (Hensher[5]) and attribute non-attendance (Campbell et al.[2]) in order to make inferences about how these dynamics affect measurements at the aggregate level. In addition, we explore how demographic characteristics influence hypothetical bias. Finally, we address free-riding as a separate issue by analyzing differences in response to elicitation mechanisms.

We utilize revealed preference (RP) and stated preference (SP) data from a novel experiment involving the design and implementation of a local market-like process for ecosystem services. The first phase of the experiment involved eliciting preferences for farmland ecosystem services via a mail-in hypothetical choice experiment. The ecosystem service being offered to the respondents was a contract to protect grassland-nesting habitat for fledgling bird populations on hay fields in the community. Each contract contained a variety of related attributes, with both public and private goods characters. In the second phase, the community members were subsequently asked to participate in a market-like process to pay the local farmers to provide the farmland amenity. This process had characteristics of a non-hypothetical choice experiment but served as a real market for the species protection service that the farmers offered for sale.

This particular experiment was uniquely conducive to analysis of hypothetical bias for several reasons. First, preference elicitation for public goods has been found to exhibit more hypothetical bias than private goods because they are less familiar to the respondent (List and Gallet 2001 [8]). However, in contrast to other survey formats, the choice experiment approach adds an element of realism to the choice task and is widely considered the state of the art among survey researchers (Hensher 2010 [5]). Second, both SP and market experiments were administered to the same sample of individuals in a small community in rural Rhode Island. Thus, differences in subject pool (i.e. university students vs. other types of respondents) can be ruled out. We have a substantial amount of demographic and attitudinal information regarding our respondents and can rule out differences in subject pool. Third, in contrast to most studies using SP and RP data, our study conducted an SP experiment before the market experiment. Finally, care was taken to engineer the market good and the market process to be as closely consistent with the hypothetical choices as possible. Moreover, nearly identical elicitation mechanisms including the pivotal point mechanism and provision point mechanisms were administered in order to reduce freeriding. This way, we are able to examine the performance of different elicitation mechanisms to address freeridership as a separate issue.

The paper proceeds as follows. Section 2 describes the ecosystem service under analysis. Sections 3 and 4 review the stated preference and revealed preference experiments. Section 4 describes the method used to construct the willingness-to-pay estimates. Section 5 presents the data. Section 6 describes the method used to combine the data. Section 7 presents results and section 8 provides a summary and conclusion.

2 Constructing a local market for wildlife preservation

The primary objective of the experiment described herein was to provide a means by which community members can exercise their preferences for farmland wildlife protection on a local scale. Often, provision of public goods is left to legislative bodies which act on a regional scale and may be influenced by special interests to act in a way that is not in line with maximizing social welfare. Placing these decisions in the hands of the community members may prove a more efficient means of providing or protecting the goods and services produced by the natural environment.

Bobolinks Every spring, hay farms on Jamestown serve as nesting grounds for a species of ground nesting birds with a large migratory range and charismatic song called the Bobolink (*Dolichonyx oryzivorus*). Historically, hay fields in many U.S. states have been in decline as preference is given to other crops. In addition, crops are cut 2-3 weeks earlier than has been historically [in the 1940s and 1950s] (Cornell BNA[11]. This shift in cropping practices has led to serious mortality for Bobolink fledglings. Consequently, Bobolinks are now protected under the Migratory Bird Treaty Act and are listed as a Species of Special Concern in some states. The Bobolink experiment sought to transfer compensatory payments from community members to farmers to delay harvesting of hay crops thereby permitting Bobolink offspring to fledge and avoid devastation at the hands of the plow.

The two phases of the experiment was administered as a mail-in survey and solicitation to the inhabitants of Jamestown, Rhode Island and spanned a three-year period from 2006 to 2008.

The SP survey was designed as a multi-question contingent valuation experiment mailed to the residents of Jamestown from October to December of 2006. There were 5 questions comparing two potential contracts and a sixth question with one potential contract. The sixth response was not utilized in this analysis, but may be useful in determining predictive validity in the future. Each contract was described by a list of attributes (see Table 1). There were six attributes described: 1. acreage under contract to delay harvest (Acres), 2. the number of acreage to restore to active farmland (Restore), 3. whether the acreage was found to have a high or low concentration of Bobolink (HighBobolink), 4. whether the contracted acreage is viewable from the road, 5. whether or not a birdwalk is offered, and 6. the cost of implementing the contract. Respondents were presented with two competing contracts presented side by side. Individuals were then asked whether they would choose contract A, contract B, both, or neither. A full description of the survey design and implementation can be found in Uchida et al. 2007[13]. We shall describe the three mechanisms that were common to both experiments. The pivotal mechanism (PM) is designed such that the respondent pays only if her bid is pivotal; that is, her contribution makes the difference between the good being provided and not. This mechanism was included because it has been demonstrated to be incentive-compatible in mitigating freeriding. The second mechanism, the proportional rebate mechanism (PR), collects bids from all respondents and if more money than is needed is collected, returns the balance as a proportion of bid amount. The third mechanism, uniform price auction (UPA), is designed so that a uniform price is ultimately administered to all bidders whose willingness to pay exceeds a reserve price. That all participants ultimately pay the same price has two effects. First, it mirrors the law of one price aspect of conventional markets and, second, there is a notion of equity not present in the other mechanisms. That is, all participants ultimately have the same financial responsibility toward provision of the good. Uchida et al. 2007 explored the effects of the various elicitation mechanisms on the marginal utilities of the attributes of the contracts and the marginal utility of income. They found that elicitation mechanisms had the most significant effect on the marginal utility of income. The mechanisms were not found to jointly affect the marginal utility of the contract attributes. Base utility level was also not found to be affected by mechanisms. In addition, the authors computed marginal willingness-to-pay estimates and mean WTP for a contract with typical attributes and compared them by mechanism. They concluded that, for a typical farm-wildlife contract, the pivotal mechanism had the lowest estimated WTP.

The second experiment was a revealed-choice experiment whereby acceptable contracts were drawn up between the mediators and farmers in the same community of Jamestown, RI. Community members were solicited for payment toward provision of the contracts, again via the different types of elicitation mechanisms. In addition to the application of competing mechanisms, individuals were randomly assigned either open- ended or binary choice questions. This experimental market was open to the residents of Jamestown in early 2007 and again in early 2008. The markets successfully provided five of ten potential field contracts. Similarly, it was determined that elicitation mechanism had an impact on valuation estimates. By comparing the incentive-compatible mechanisms between the two experiments, we may, in theory, eliminate the effect of free-ridership and determine the effects of hypothetical bias on the prediction of willingness to pay in the market. A comparison of the attributes across both experiments is listed in table 2. The market experiment did not offer participants a choice to restore fallow land to active cultivation or a birdwalk. In addition, concentrations of Bobolink activity in a particular plot was represented differently across treatments. In the choice experiment, the attribute "Low/High Bobolink Concentration" was mildly correlated with plot size and, in all specifications, was found to be insignificant. The market experiment included two separate representations of Bobolink concentration. The number of Bobolink territories observed in 2006 was used in the 2007 treatment and the number of Fledglings expected to be supported by the field. Both choice experiment and market treatments described contracts as having a view or no view, with the addition of a 'partial view' option in the market treatments. The field size attribute had a broader range in the stated preference survey and costs were comparable across years.

A wealth of demographic and attitudinal information was available to the researchers for individuals who participated in both experiments. In addition to mailing addresses, resident demographics obtained from a market research firm included age, income, education level, donation history, and mail-order history. In addition, choice experiment respondents were asked to provide information about community members' preferences for farmland amenities.

A summary of market experiment participation by major group is listed in table 3. The table is divided into four groups: all individuals who returned the market mailing, those who returned both the market and choice experiment surveys, the subset of CE respondents who chose the "Both" option for all questions, and all other individuals who were mailed a market experiment survey. The individuals who chose the "Both" option consistently are singled out as "yeahsayers". There are two specific issues with this group. First, choosing the Both option for each question offers no information about one's relative preferences for the attributes of the contracts. Specifically, no trade-off between cost and attribute levels is observed. Thus, a model that assigns either arbitrarily low values for the marginal utility of income or arbitrarily high values for the marginal valuation of all attributes is likely to result. Any model that aggregates these respondents into the mix will produce mean values that are too high.

The second issue with this subsection of the SP respondents is the actual signal that is being sent by adopting a ”yeah-saying” strategy. One might imagine that some of the respondents are truly more interested in and willing to pay for ecosystem preservation and that some individuals may simply be expressing interest but not expending the mental effort to assess whether they would actually be willing to pay the stated amount.

From the table, we can see that, over all, individuals who returned the CE survey were more likely to return the market experiment solicitation, make an offer, and offer larger bids. The subset of yeah-sayers, in fact, have even higher participation and offers (with the exception of the second year).

3 Ways to address heterogeneity in response

Our objective is to identify strategic behaviors that might violate the assumptions of neoclassical utility maximization and examine hypothetical bias given these behaviors. The Latent Class Logit (LCL) model is particularly useful for this kind of analysis. The LCL model has been used both to explore patterns of attribute non-attendance and other violations of continuous preference ordering as well as modeling non-parametric preference heterogeneity.

There is by now a substantial literature which uses the LCL model to identify AN-A behavior, especially in the absence of self-reported non-attendance. It is widely exploited in marketing and transportation studies but has recently been used in cases of public goods valuation. The unique feature of the CV survey was the presence of the Both option. Coupled with the fact that the mail-in survey format is bound to select out individuals who have strong feelings for or against the contracts and their features. Scarpa et al. ([12]) used the Latent Class approach with a stated preference survey designed to elicit preferences for landscape improvements in Ireland. They stress that, if an attribute is ignored, then relative trade-offs that involve that attribute are not meaningful. That is, no increase/decrease in the ignored attribute compensates for a change in an attended attribute. Most of these studies assume preference homogeneity across classes except in the case that a particular attribute in a class is not attended to. That is, a parameter on an attribute is either restricted to zero if it is not attended to in a class or estimated and fixed in the sample population. In this respect, it is called a fixed parameter latent class model and the class membership probability determines an individual's placement into a non-attending class.

One of the prevailing debates among researchers using this approach has been whether a parameter restricted to equal zero truly represents ANA on the respondent's part. Certainly within the context of a fixed parameter latent class structure, it is likely that such classes capture individuals whose marginal utility is very close to zero. Hess et al. ([6]) point out this consideration and relax the assumption of fixed values for fully attended attributes. Our model builds on this and Train ([**?**]) who treats the LC model as a "non-parametric" representation of preference heterogeneity, where the number of support points, defined as number of classes, characterizes the distribution of heterogeneity within the population and can approximate arbitrarily complex distributions.

Modelling response Heuristics

Non-parametric estimation: EM algorithm

4 Methods

4.1 The LCL Model

Stage 1: The LCL model The model presented herein combines Train's (2008) Expectation Maximization algorithm for nonparametric estimation of the random parameter latent class logit model with Hess et al.'s (2011) expansion of attribute non-attendance for heterogeneous taste variation to account for respondents who may have ignored key attributes omitted in the market experiment but for whom varying taste intensities were permitted for attributes that were attended to. In particular, our aim is to catch and contain the yeah-sayers in our sample, whose insensitivity to contract price would otherwise inflate marginal values for all other attributes, and to identify respondents who were perhaps only mildly sensitive to the two attributes that were omitted from the revealed preference survey: acres of restored farmland and invitation to a bird walk. Attribute non-attendance is expected to be a significant problem with this particular type of choice task since respondents are not likely to be familiar with the ecosystem service for offer and thus may make unforeseen assessments of the true meaning of the attributes of the contract, or to decide that a particular attribute is too cryptic to assess a value for.

Lay out the model The Expectation-Maximization algorithm has been utilized by Train (2008) as a form of non-parametric estimation of underlying taste heterogeneity whereby a discrete distribution whose accuracy in approximating the true underlying distribution rises with the number of parameters. This is an extension of Bhat ([**?**]) where increasing the number of classes allows for better approximation of taste heterogeneity.

Given the standard choice modeling scenario, N agents choose among J alternatives in each of T choice occasions, let y_{njt} be an indicator variable equal to 1 if agent n chooses alternative j in choice situation t. Each alternative is defined by a set of attributes with varying levels and the choice of attributes is assumed to result from standard neoclassical utility maximization. Further assume that there are C distinct sets of taste parameters in the population, $\beta = \{\beta_1, ..., \beta_C\}$. In this framework, there are two sets of unknowns: the β_cs, that is, the estimates of taste parameters for each class, and the class membership status of the agents.

Model selection is generally achieved based on minimizing an information criterion such as AIC, BIC, or CAIC. Train (2008) compares AIC against CAIC which penalizes additional parameters more. If the information criteria do not agree on which model is preferred, the researcher must choose based on examination of standard errors and feasibility of parameter signs. It is important to test several different starting points to ensure that a global maximum has been obtained.

Given the class parameter estimates, we estimated the conditional probabilities of each individual belonging to each of the classes. We use these conditional probabilities as weights on the class parameters to construct the marginal utilities for each individual. We combined the estimated parameters with the RP contract attributes. We were only able to use the acres attribute and the view attribute, since high bobolink was not estimable in the first stage. So, we constructed choice-instance expectations of WTP by combining the individual constructed marginal utilities of acres and view, subtracting the status quo estimate, and dividing by the marginal utility of income. An individual is assigned to a class based on their highest conditional probability.

We first explore determinants of the decision to participate in the market experiment. We consider an individual to participate in the market if she returned the market mailing, even if her offer amount was zero. We model the decision to participate using a random parameter probit model in which the dependent variable is whether or not the individual returned the market mailing. We are interested in determining whether individuals who fell into a particular class were found more or less likely to participate. In addition, we'd like to quantify the change in likelihood of participating based on the fact that an individual returned the SP survey. We test whether the elicitation mechanisms affect the decision to participate. Finally, we test demographic variables for inclusion in the model.

We then turn our interest in determining if we can quantify differences in hypothetical bias based on class membership, relevant characteristics of the market mailing, and demographic characteristics. We model the amount offered using interval regression because we have a mix of payment card and binary choice data.

We choose to focus on non-attendance of attributes from the SP experiment for which there is no RP analogue. That is, we model attribute non-attendance to the restore and tour attributes. We hypothesize that these estimates of stated wtp are in a sense more reliable because they assume that the decision process is more aligned with the RP scenario.

We tested this model with 50 starting points so as to ensure a global rather than local maximum as per the recommendations of Train (2008).

5 Results

Upon examination of the choice experiment data, there appeared a few prominent strategies. For instance, a significant proportion of respondents answered that they would purchase both contracts for all questions. There was also evidence that many individuals were simply choosing the lowest cost option. Based on the choice experiment survey, many individuals indicated that they would not be interested or would not be able to attend a birdwalk. Thus, ignoring this attribute seemed to be a plausible option based on first section of survey.

The final latent class model was arrived at after extensive testing. Table 4 lists the results for several candidate models. The first model tested assumed that there was one class of individuals who ignored the Restore attribute of the contracts, one class who ignored (or had marginal valuation of) the Tour attribute, one class of individuals who ignored the cost attribute, a class who ignored all attributes, and a full attribute attendance class. The final model chosen, based on BIC and highest average maximum conditional class membership probability, had seven classes: four restricted classes and three full attendance classes.

A description of the classes can be found in Table 6. The yeah-sayers fell into the cost ANA class. We were unable to calculate WTP estimates for this group, given the absence of the cost parameter estimate. This is as it should be since no clear valuations can be deduced from such behavior and only a lower bound on WTP can be determined.

Following some authors, we included a class that represented full attribute non-attendance to capture idiosyncratic behavior. There were 116 respondents who fell into this category. A portion of these respondents were found to have extreme reactions to a change from the referendum to one of the elicitation mechanisms. Individuals who consistently chose the cheapest alternative were relegated to the Tour ANA class.

Of the full attendance classes, a few trends emerge. Class 5 includes individuals whose preferences are best described by large valuations for both a view and tour, but, little particular interest in the number of acres under contract. These individuals seem to be driven by use values. Class 6 individuals have generally high values for all attributes, with particular interest in a tour. The final class has the highest valuations, and seem drawn to the Restored acreage attribute as well as the tour. Average willingness to pay for ten acres of protected farmland with a view from the road is listed in the table. It is clear that there are significant differences in valuation across classes.

The results of the participation equation (table 7) reveal some interesting findings. Age and having a history of donating both increase participation. Mail ordering through children's catalogues decreases the likelihood of participation.

Unsurprisingly, having returned the choice experiment survey increases the likelihood of returning the market mailing. Class 6 individuals have the highest likelihood of participation.

Table 10 Interpretation The interval regression results indicate that individuals who were administered a discrete choice survey were found to contribute significantly more than those who received payment card solicitations. Class 3 participants were found to pay less than individuals who returned only the market mailing. All other classes were willing to pay more. From table 9, we can see that class 4 individuals (the "yeah-sayers") were likely to contribute highly. In addition, class 6 individuals were found to have the most truthful revelations of willingness to pay. That said, there is a significant amount of hypothetical bias in these results. The coefficient on Class 6 estimated willingness to pay is 0.386 indicating that actual payment for this class was 38% of what was estimated in the first stage Latent Class model.

Summary and Conclusions There are several interesting findings that emerge from these results. From a disaggregated perspective, it is clear to see that some sources of hypothetical bias indeed arise from certain attribute processing strategies that preclude a reliable estimation of marginal valuations.

Table 1: Summary of Contingent Valuation contract attributes

Attribute	Description	Levels
Acres	Number of acres to be placed under contract upon which farmer will delay mowing and harvesting	10, 25, 40, 55
Restore	Number of acres to be restored to active hay fields, not restricted to delayed mowing/harvesting	0, 10, 20, 30
High Bobolink	Level of expected fledglings saved (correlated with acreage)	Low, High
View	Whether the parcel will be viewable from the road	View, No View
Tour	Whether individuals paying into the contract are invited to an expert-led birdwalk	Tour, No tour

Table 2: Comparison of Attributes and Levels across SP and RP treatments

Attribute	Choice Experiment	Market 2007	Market 2008
Field size to delay harvesting	10 – 55	10 – 18	10
Acres of farmland restored to cultivation	0 – 30		
Bobolink Concentration	Low, High		
# of 2006 territories		1 – 4	
# of Fledglings			6—10, 10—14,14—18
View from road	View, No view	View, Partial, No view	View, No view
Mechanism	SP, VCM:PPMBG, PPPR, PM, UPA	PM, PR, UPA	PM, PR, UPA
Cost	\$10 — \$85	\$10 — \$120	\$10 — \$120

Table 9: Testing Payment Specifications With Demographic and Contract Attributes

VARIABLES	Eq. 3	Eq. 4	Eq. 5
dc	25.56***	25.60***	20.90***
	(5.336)	(7.453)	(7.464)
oehigh		2.770	1.104
		(3.415)	(3.557)
spwtpdc			-0.0606
			(0.201)
spwtp	0.173*	0.187*	0.178*
	(0.0943)	(0.109)	(0.107)
Class 1 WTP	0.143	0.162	0.119
	(0.492)	(0.505)	(0.492)
Class 1 WTP*dc		0.277	0.353

Continued on next page

	Eq. 3	Eq. 4	Eq. 5
VARIABLES			
		(1.655)	(1.560)
Class 2 WTP	0.202	-0.00634	-0.0223
	(0.181)	(0.222)	(0.216)
Class 2 WTP*dc		0.420	0.592*
		(0.310)	(0.340)
Class 3 WTP	-2.504***	-1.732	-1.481
	(0.924)	(1.179)	(1.157)
Class 3 WTP*dc		-1.733	-1.983
		(1.622)	(1.628)
Class 4	26.83***	18.88*	23.40**
	(9.582)	(11.04)	(11.10)
Class 5 WTP	0.130	0.380	0.316
	(0.599)	(0.626)	(0.617)
Class 5 WTP*dc		-15.26	-13.66
		(557,414)	(107,420)
Class 6 WTP	0.386***	0.330**	0.314**
	(0.126)	(0.142)	(0.139)
Class 6 WTP*dc		0.0466	0.168
		(0.198)	(0.266)
Mail Order (kids)	-26.84***		-27.99***
	(7.243)		(7.526)
UPA			-14.43
			(8.818)
PR			3.080
			(7.612)
PM			-1.200

Continued on next page

Table 9 – *Continued from previous page*

VARIABLES	Eq. 3	Eq. 4	Eq. 5
			(7.941)
Missing Income	9.080		1.780
	(15.16)		(20.92)
Donation History (any)			-1.908
			(6.780)
Donation History (environment)			5.777
			(8.850)
Age			-0.215
			(0.259)
Same Mechanism		0.179	-0.500
		(5.684)	(6.007)
Class 7 WTP*dc		-0.0736	
		(0.207)	
2008 Indicator	4.196		
	(4.334)		
Constant	21.58***	18.34***	39.32**
	(5.146)	(5.709)	(17.55)
sigma_u	47.13***	48.17***	45.74***
	(3.482)	(3.522)	(3.477)
sigma_e	23.89***	24.29***	24.64***
	(1.996)	(1.436)	(1.899)
Observations	576	576	576
Number of Individuals	471	471	471

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Summary of Market Response

	Total (2007/2008)	SP respondents (2007/2008)	SP Yeah-sayers (2007/2008)	RP only (2007/2008)
N	2791 / 2680	764 / 713	105 / 101	2027 / 1967
Returned Market mailing	365 (13%) / 211 (7.9%)	220 (28.8%) / 112 (15.7%)	36 (34.3%) / 14 (13.9%)	145 (7.2%) / 99 (5%)
Made Offer	198 (7%) / 141 (5.3%)	141 (18.5%) / 81 (11.4%)	28 (26.7%) / 11 (10.9%)	57 (2.8%) / 60 (3.1%)
Average Value of Offer	$47.94 / $46.49	$50.30 / $46.98	$60.54 / $41.82	$42.11 / $45.83

Table 4: Results of Specification Search

Classes	Restore ANA	Tour ANA	Cost ANA	All ANA	Full A	LL	AIC	BIC	Avg. max. P(membership)
5	1	1	1	1	1	-3448	6950	7076	0.86
6	1	1	1	1	2	-3420	6908	7068	0.85
7	1	1	1	1	3	-3397	6876	**7067**	0.8
8	2	2	1	1	1	-3385	6862	7077	0.8
11	3	3	1	1	3	-3355	**6840**	7144	0.75
12	3	3	1	1	4	-3349	6842	7178	0.67
14	4	4	1	1	4	-3336	6842	7239	0.71

Note: Bold-faced numbers indicate the best values for AIC/BIC.

Table 5: Latent Class Model Results

Class	Cost	Statquo	Acres	Restore	View	Tour	share
Class1	-0.02	-0.5	-0.02	0	0.88	-0.42	0.070*
Class2	-0.05***	-1.577***	0.036***	0.024***	0.1	0	0.341***
Class3	-0.09***	-1.144***	0	0	0	0	0.13
Class4	0	-14.27	0.435*	5.64	11.14	0.19	0.13
Class5	-0.021**	0.14	-0.03*	0.04	0.958**	0.16	0.034*
Class6	-0.016***	-1.09***	0.037***	0.066***	0.225*	0.295**	0.16
Class7	-0.008*	-1.768***	0.01	0.034***	-0.07	0.35***	0.13
BIC				7067			
AIC				6876			
LL				-3397			

Notes:*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Summary of Classes

Class	Definition	Unique Features	Avg. WTP for 10 acres with view	N
1	Restore ANA	Cares highly about view	$54.68	22
2	Tour ANA	Includes all individuals who chose the cheaper alt. all or most of the time	$40.70	302
3	All Attributes ANA		$12.71	116
4	Cost ANA	All yeah-sayers	NA	115
5	Full Attendance	Cares most about having view and tour, not interested in acreage	$24.76	16
6	Full Attendance	High Values for all attributes, esp. tour	$105.31	122
7	Full Attendance	Cares most about restore and tour	$218.63	98

Note: N is the number of individuals for whom the maximum conditional probability places them in the class

Table 7: Participation Equation

VARIABLES	Participation Equation
ln(Purchasing Power)	0.0320
	(0.0567)
Age	0.0136***
	(0.00362)
Donation History (any)	0.247***
	(0.0834)
Mail Order (kids)	-0.321***
	(0.0986)
lnMinAmt of Offer	-0.148***
	(0.0472)
SP Respondents	0.741***
	(0.197)
Class 1	0.521
	(0.413)
Class 2	0.372*
	(0.213)
Class 3	0.220
	(0.249)
Class 4	0.506**
	(0.246)
Class 5	0.399
	(0.473)
Class 6	0.708***
	(0.243)
Constant	-2.680***
	(0.394)
lnsig2u	0.123
	(0.171)
Observations	5,239
Number of nameid	3,025

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Payment Equation

VARIABLES	Eq. 1 lboe	Eq. 2 lboe
Discrete Choice		27.25***
		(5.386)
Class 1		24.11
		(21.94)
Class 2		15.40**
		(7.632)
Class 3		-34.46***
		(13.14)
Class 4	24.54**	24.70**
	(9.622)	(9.734)
Class 5		36.35
		(25.36)
Class 6		58.03***
		(10.16)
Class 7		28.44**
		(13.13)
First Stage $\widehat{WTP}$	0.359***	
	(0.0727)	
Constant	26.44***	17.49***
	(3.922)	(4.938)
sigma_u	48.65***	48.33***
	(3.505)	(3.463)
sigma_e	25.32***	23.40***
	(1.053)	(2.229)
Observations	576	576
Number of Individuals	471	471

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

References

[1] David Aadland and Arthur J Caplan. Willingness to pay for curbside recycling with detection and mitigation of hypothetical bias. *American Journal of Agricultural Economics*, 85(2):492–502, 2003.

[2] Danny Campbell, David A Hensher, and Riccardo Scarpa. Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of environmental planning and management*, 54(8):1061–1076, 2011.

[3] Fredrik Carlsson and Peter Martinsson. 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, 2001.

[4] Jae Bong Chang, Jayson L Lusk, and F Bailey Norwood. How closely do hypothetical surveys and laboratory experiments predict field behavior? *American Journal of Agricultural Economics*, 91(2):518–534, 2009.

[5] David A Hensher. Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological*, 44(6):735–752, 2010.

[6] Stephane Hess, Moshe Ben-Akiva, Dinesh Gopinath, and Joan Walker. Advantages of latent class over continuous mixture of logit models. Technical report, working paper, 2011. URL http://www. stephanehess. me. uk/papers/Hess_Ben-Akiva_Gopinath_Walker_May_2011. pdf, 2011.

[7] Robert J Johnston. Is hypothetical bias universal? validating contingent valuation responses using a binding public referendum. *Journal of Environmental Economics and Management*, 52(1):469–481, 2006.

[8] John A List and Craig A Gallet. What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics*, 20(3):241–254, 2001.

[9] James J Murphy, P Geoffrey Allen, Thomas H Stevens, and Darryl Weatherhead. A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30(3):313–325, 2005.

[10] James J Opaluch and Kathleen Segerson. Rational roots of irrationalbehavior: new theories of economic decision-making. *Northeastern Journal of Agricultural and Resource Economics*, 18(2):81–95, 1989.

[11] A. Poole. The birds of north america online, August 2005.

[12] Riccardo Scarpa, Timothy J Gilbride, Danny Campbell, and David A Hensher. Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European review of agricultural economics*, 36(2):151–174, 2009.

[13] Emi Uchida, Christopher M. Anderson, and Stephen K. Swallow. Marketing ecosystem services from agricultural land: Stated preferences over payment mechanisms and actual sales of farm-wildlife contracts. 2007 Annual Meeting, July 29-August 1, 2007, Portland, Oregon TN 9955, American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association), 2007.