8. Inferences from stated preference surveys when some respondents do not compare costs and benefits[1]

Edward Leamer and Josh Lustig[2]

INTRODUCTION

Stated preference surveys are often used to estimate willingness to pay (WTP) for environmental improvements. These surveys typically ask respondents to choose between the status quo and one or more environmental improvements at hypothetical costs. The feature of these survey data that determines the estimate of the WTP is the declining fraction of respondents who choose an environmental improvement as the hypothetical cost increases. But data alone are not enough. A formal choice model is also needed to turn these data into WTP estimates. The traditional model presumes that all respondents know the gain in utility that they would experience if the environmental improvement were enacted, and they also know how much utility would be lost if they were compelled to pay the hypothetical cost. They then are assumed to choose the option with the greatest hypothetical net benefit, or choose the status quo if all the hypothetical net benefits are negative.

The WTP estimates derived from this traditional analysis are valid only if respondents are actually behaving in a way consistent with the utility maximization assumption, and in particular are making the kind of thoughtful trade-offs between costs and benefits that would have them choosing the environmental improvement if the cost is low and rejecting it if the cost is high. Rather than presupposing this ideal behavior, this

chapter offers models that allow for the possibility that some respondents may ignore costs while others may ignore benefits.

The fraction of respondents who behave according to the traditional utility maximization model is an estimated parameter in our model, as are the fractions of respondents whose decisions are better described by one of the heuristic decision rules. This statistical model can be thought to be a way of purging from the data suspicious responses, thus providing a formal basis for the common practice of excluding "protestors" who would oppose any environmental improvement regardless of cost (e.g., Meyerhoff et al., 2012). Protestors who ignore the benefits inappropriately drag down the estimated WTP if they are included in the data set, but there may also be respondents who want to improve the environment but do not weigh the costs and benefits of doing so. For example, some respondents may ignore the costs of environmental improvements and others may make choices only on the basis of costs. Therefore, it takes a model with at least three heuristics to identify protestors and respondents whose preferences for environmental improvements do not reflect trade-offs between costs and benefits, and to do statistically valid and conceptually unbiased two-sided trimming of responses. Inevitably, as we allow more heuristic rules into the model, the data trimming becomes more substantial and the estimated fraction of respondents whose choices reflect cost–benefit trade-offs declines but the estimated WTP of these respondents can go either up or down.

We apply our methodology to data collected by a National Oceanic and Atmospheric Administration ("NOAA") survey designed to value eight threatened and endangered marine species. We discover that many of the NOAA survey respondents' choices are more consistent with one or more heuristic rules than with a model of utility maximization. Using a model that includes a mixture of five heuristic decision rules competing with utility maximization, we estimate that only 23.4% of survey respondents answer the survey in a manner consistent with utility maximization. This estimated model offers a substantially improved fit of the data over the traditional mixed logit model that assumes all respondents maximize utility. We also discover that as the list of included heuristics is varied, there is considerable variability in the estimate of the fraction of utility maximizers and the estimate of their WTP. This variability of conclusions is troubling because we have not attempted to identify a full set of heuristics and if more heuristics are included in the model the range of alternative estimates would inevitably increase. The operative concluding words are thus: credibility and fragility. A credible analysis of data from stated preference surveys needs to allow formally for aberrant decision-makers, but attempts to increase credibility are likely to uncover an uncomfortable amount of fragility.

The rest of this chapter proceeds as follows. The next section briefly reviews related literature. The third section describes the stated preference survey we exploit. In sections four and five, we present our model and describe our estimation strategy. The sixth section presents our results and the seventh concludes.

RELATED LITERATURE

Latent class models allow survey respondents to use different rules or strategies when responding to surveys. For example, one class we assume below will absorb survey respondents who weigh the costs and benefits of marine species improvements when responding to the survey. What class a particular respondent belongs to is hidden to the researcher. But the latent class model infers probabilistically which class each respondent likely belongs to. Below, we use latent class models to infer which NOAA stated preference survey respondents likely make trade-offs between species improvement costs and benefits and which respondents likely use alternative choice rules. Other researchers have also used latent class models to make this distinction. In this section, we briefly review this literature.

First is a recent literature that uses latent class models to study attribute non-attendance.[3] Attribute non-attendance is present when some respondents ignore one or more attributes of the good or service when making choices. To accommodate multiple classes of respondents, these papers modify the multinomial and mixed logit models and allow for separate utility specifications for each class of respondents.[4] The literature on attribute non-attendance allows some respondents to have zero coefficients for the neglected attributes, but those respondents are otherwise identical to utility maximizers. We depart from this literature by allowing respondents who use heuristic decision rules to be entirely different, with no parameters in common with the utility maximizers who trade off costs and benefits.[5]

The attribute non-attendance literature comes to the same conclusion

---

[3]   See, for example, Hensher et al. (2005, 2012), Scarpa et al. (2009), Hensher (2010), Hensher and Greene (2010), McNair et al. (2010), Campbell et al. (2011), and Greene and Hensher (2013).

[4]   All but one of the papers cited above estimate multinomial logit models with latent classes. Hensher et al. (2005) estimate a multinomial *mixed* logit model with latent classes.

[5]   Another difference between our chapter and the non-attendance literature is that all but one of the heuristic rules in our model are deterministic in the sense that the choices made by respondents following the heuristic rules can be perfectly predicted based on the observed characteristics of the options.

we do. Many individuals do not weigh the costs and benefits of improvement options when responding to stated preference surveys. For example, in a survey intended to elicit WTP for rural landscape improvements in Ireland, Campbell et al. (2011) find that 65.2% of survey respondents ignore costs. Using a model that does not allow non-attendance, they estimate WTP for landscape improvements between from $163 to $221. After accounting for attribute non-attendance, estimated WTP ranged from $49 to $109. In other words, WTP falls by more than 50% after adjusting the model to allow for the possibility that not all respondents are attentive to the costs of landscape improvements. Below, we will find a similar downward adjustment to WTP when the model allows a set of heuristics as alternatives to utility maximization.

A second related literature uses latent class models to identify "protestors" based on serial non-participation.[6] See, for example, Von Haefen et al. (2005), Burton and Rigby (2009) or Cunha-e-Sa et al. (2012). In these analyses, a class of respondents is assumed to reject environmental improvements (or other improvements) regardless of their benefits and costs. Inferring "protest" responses using latent class models complements the standard approach that identifies "protestors" based on their responses to follow-up questions. For example, Lew and Wallmo (2011) and Wallmo and Lew (2011, 2012)[7] also rely on the NOAA survey and define protestors as (1) respondents who choose the status quo in all three questions and (2) indicate they are not confident in their responses or their answers to other follow-up questions indicate the respondent is not making cost–benefit trade-offs. For example, respondents who distrust the government or are unwilling to pay higher taxes for any reason are classified as protestors if they choose the status quo in all three questions.

Other applications of latent class models do not neatly fall into the two categories described above. For example, McNair et al. (2012) use a latent class model to identify survey respondents who learn their preferences while completing the survey and respondents who behave strategically (i.e., misrepresent their preferences to manipulate the survey outcome). Similarly, Hess et al. (2012) use a latent class model to distinguish between respondents who use reference points when completing stated preference surveys and respondents whose choices reflect lexicographic preferences.

---

[6] This literature also uses the term "hurdle" model to refer to latent class models.

[7] Throughout the text, we refer to the definition of protestors used by Lew and Wallmo. In each of these instances we are referring to the definition of protestor adapted by Lew and Wallmo (2011), Wallmo and Lew (2011), and Wallmo and Lew (2012).

NOAA SURVEY DATA

Our analysis relies on survey data from Phase I of the Protected Species Valuation Survey undertaken by the National Marine Fisheries Service of the National Oceanic and Atmospheric Administration ('NOAA survey'). The purpose of the NOAA survey was to value potential improvements in the Endangered Species Act (ESA) status of eight threatened and endangered ('T&E') marine species – the North Pacific right whale, the North Atlantic right whale, the loggerhead sea turtle, the leatherback sea turtle, the Hawaiian monk seal, the wild Upper Willamette River Chinook salmon, the wild Puget Sound Chinook salmon, and the smalltooth sawfish. Each version of the survey offered respondents the opportunity to improve three of these eight species.

Before answering choice questions about species improvements, respondents were first shown information about the ESA and the three species in their version, as well as what actions are currently being done to protect them and what additional actions could be undertaken. They were then asked to select their most preferred option in three choice questions. Each question offered respondents three alternatives to choose from: a status quo option and two alternative options offering additional protection actions for at least one of the three T&E species. Respondents were asked to select the option they would most prefer. Figure 1 shows an example choice screen. Each option is described by the ESA status of each species (endangered, threatened, or recovered) before and after the option is implemented and the amount of added household cost per year over a period of ten years. The three options are labeled A, B, and C from left to right, with Option A always being the status quo option, with no added household cost.

The NOAA survey was conducted by Knowledge Networks using a random sample of their Internet panel of US households.[8] A pretest including only three of the eight T&E species was fielded in December 2008 and January 2009. The main survey was fielded in June and July of 2009, yielding 13,684 completed surveys with a completion rate of 70.8%. There are 44 versions of the main survey, differing by species combination, species order, which cost scale was used, and whether a "cheap talk" script was given to the respondent. Each version is further divided into 16 sub-versions with different levels of ESA status and costs.[9]

---

[8]   We obtained the NOAA survey data through a Freedom of Information request.
[9]   We chose not to weight the Knowledge Networks (2009) survey data for several reasons. First, Wallmo and Lew did not use weighted data, and we wanted to do our analysis using the

As in the previous question, please compare Options A, B, and C in this table and select the option you most prefer.

Remember that any money you spend on these options is money that could spent on other things.

Expected result in 50 years for each option

	Option A *No additional protection actions*	Option B *Additional protection actions*	Option C *Additional protection actions*
Loggerhead sea turtle ESA status	Threatened	Recovered	Recovered
North Pacific right whale ESA status	Endangered	Threatened	Endangered
Leatherback sea turtle ESA status	Endangered	Threatened	Recovered
Cost per year Added cost to your household each year for 10 years	$0	$100	$60
Which option do you prefer?	○	○	○

Figure 1   Example of choice experiment

We restrict our analysis to survey respondents who provide answers to each of the three choice tasks assigned to them. Imposing this restriction reduces the number of respondents from 13,684 to 11,459 and the number of observed choices from 41,052 to 34,377.

same data they used to allow direct comparison of our results. Second, the data sample used in the survey was enormous, so any distortions caused by not weighting should be minimal. Third, our focus in our analysis was to show that estimated WTPs vary depending on the inclusion of different heuristics classes in the estimation methodology; weighting the data was not required for this purpose.

MODEL

In this section we describe the model we use to explain respondents' answers to the NOAA stated preference survey. Our model allows a number of alternative choice rules that respondents could have used to solve this task. These include choice rules in which survey respondents make rational trade-offs between the costs and benefits of environmental improvements and choice rules where choices do not reflect such trade-offs. This distinction is important because survey responses are informative about willingness to pay only if they reflect rational cost–benefit trade-offs.

We note where the choice rules we include in the model are similar to choice rules used in previous analyses of stated preference surveys. Although our set of alternative rules spans a wide range of behaviors, we have not attempted to include all likely choice rules, and in particular we do not study context effects here.[10,11]

Trade-off Respondents

We assume a fraction π_{TO} of respondents make rational trade-offs between the costs and benefits of species improvements. A "trade-off" respondent i is assumed to choose option j that maximizes utility v_{ij}:

$$v_{ij} = \sum_s \beta_{is} * d_{js} - \alpha Cost_j + \varepsilon_{ij}$$

Utility v_{ij} includes the benefits from species improvements, the costs of species improvements, and an idiosyncratic zero mean error term ε_{ij}. The error term represents either unobserved utility or personal indecision (wavering). If ε_{ij} is unobserved utility with mean zero it does not affect mean WTP and if ε_{ij} is wavering it does not affect individual WTP. The average WTP calculations we perform below that exclude the error terms are correct in either case.

The binary indicator d_{js} turns on when a species improvement is offered and the parameter β_{is} is the "utility" that respondent i would experience given a particular species improvement. The coefficient on cost α measures the utility of income, which is implicitly assumed to be constant over the chosen cost scale and the same for all individuals. Since option A (the

---

[10]   We study the context effects in the NOAA survey in a companion paper.
[11]   We also chose not to use the NOAA stated preference survey to study learning (e.g., Plott, 1996). Researchers typically identify learning by looking for changes in respondents' behavior as they proceed through the survey. However, the NOAA survey only asks respondents three questions. Learning studies typically use surveys with more than three questions.

status quo option) offers no species improvements and imposes no costs, we use the normalization $v_{iA} = 0$.

WTP is the level of the cost that perfectly offsets the benefits and makes the respondent indifferent between paying for a species improvement and the status quo. This cost is the solution to $0 = \beta_{is} - \alpha Cost$, namely $WTP_{is} = \beta_{is}/\alpha$.

We model utility maximizers' behavior with a mixed logit specification similar to that used by Wallmo and Lew (2012) to analyze the same survey data.[12] The mixed logit model has been widely adopted to analyze stated preference survey data in the recent literature.[13] Similar to other studies, we assume that each level of marginal utility for species improvements, β_{is}, is drawn from a normal distribution with mean $\bar{\beta}_s$ and standard deviation σ_s. We assume the marginal utility of income, α, is fixed across respondents, leading to WTP that is also normally distributed.[14] Finally, the mixed logit model carries the assumption that ε_{ij} is drawn from an *iid* extreme value distribution.

Alternatives to Benefit–Cost Trade-offs: Heuristic Decision Rules

We consider several heuristic decision rules that capture three broad categories of respondents who do not compare and make trade-offs between costs and benefits – respondents who consider costs but not benefits, respondents who consider benefits but not costs, and respondents who consider neither costs nor benefits. We make no claims that the set of heuristics we consider is comprehensive, and the heuristic rules that we describe below represent only a subset of all the heuristics that respondents may be using.[15] However, even a small set of heuristics is sufficient to demonstrate that there are a large number of responses that are more consistent with heuristic decision rules than with utility maximization:

---

[12] Wallmo and Lew also used a similar model in the analysis of the NOAA survey's pre-test data (Lew and Wallmo, 2011; Wallmo and Lew, 2011).

[13] Examples in the environmental literature include studies on global climate change (Layton and Brown, 2000), biodiversity (Cerdaa et al., 2013), river ecology (Hanley et al., 2006), coral reef ecosystems (Parsons and Thur, 2008), landscape (Olsen, 2009), endangered species (Lew et al., 2010), and wetlands (Kaffashi et al., 2012).

[14] When estimating mixed logit models, researchers do not typically assume random coefficients for all product characteristics. All contingent evaluation studies cited above as examples in the environmental literature assume a fixed cost or price coefficient, while allowing the coefficients of all other attributes to be random.

[15] The rules we use are intended as examples of heuristic rules in which respondents are not attentive to costs or benefits. There are alternative heuristic rules we could have tested in which respondents are also not attentive to costs and/or benefits. However, identifying the heuristic rules that best explain respondents' choices is beyond the scope of this chapter.

1   Status quo only

A fraction π_{SQ} of respondents are assumed to be protestors who always choose the status quo option no matter what other alternatives are offered in the survey question. These respondents are not attentive to costs or benefits when responding to the survey. Thus, the probability an individual i chooses an option k in question q conditional on following the protestor heuristic is given by:

$$Prob(y_{iq} = k|Status\,Quo) = \begin{array}{l} 1\ if\ k = A \\ 0\ if\ k \neq A \end{array}$$

In a secondary specification, we remove survey respondents identified by Wallmo and Lew (2012) as protestors before estimating the model. These include the choices made by 2,800 respondents who chose the status quo in all three questions and whose responses to follow-up questions suggest protest behavior. While this secondary specification yields results that are similar to our main specification, we prefer a latent class approach to identifying protest behavior because it does not require a priori exclusion of respondents based on ad hoc protestor definitions. This view is supported in the literature. See, for example, Meyerhoff and Liebe (2006, 2008) and Meyerhoff et al. (2012).

2   Attentive to environmental improvements only ("steps only" respondents)

We assume that a fraction of the population π_{Steps} is composed of respondents who follow the "improvement steps only" heuristic. These respondents support environmental improvements but do not make benefit–cost trade-offs. Instead, they first identify how many steps of improvement are offered by each option in their choice set. For example, an option that improves the North Pacific right whale from endangered to recovered and leatherback turtle from endangered to threatened offers three steps of improvement. These respondents then choose the option that offers the greatest total number of steps of species improvements, irrespective of costs. If options B and C offer the same number of steps of improvements, "max steps" respondents are assumed to randomly choose between the two options with probability 0.5 on both.

   "Steps only" respondents' choice probabilities are as follows. Since the status quo option offers no species improvements, for all questions:

$$Prob(y_{iq} = A|Steps\ Only) = 0$$

Instead, respondent i chooses the option that offers the most steps of improvement. If options B and C offer the same number of steps

of improvement, respondents' choices are determined by a coin flip. Therefore, the probability that individual i chooses B is given by:

$$Prob(y_{iq} = B|Steps\ Only) = \begin{array}{l} 0 \text{ if } Steps_{qB} < Steps_{qC} \\ .5 \text{ if } Steps_{qB} = Steps_{qC} \\ 1 \text{ if } Steps_{qB} > Steps_{qC} \end{array}$$

The probability that i chooses C in any question is defined analogously.

3   Attentive to environmental costs only ("costs only" respondents)

We include a heuristic that captures respondents who want to improve the status of marine species but at the lowest cost possible. We assume these respondents choose the environmental improvement option with the lower cost, and randomly choose between the two improvement options if they share the same cost. Since the status quo option offers no species improvements, for all questions:

$$Prob(y_{iq} = A|Costs\ Only) = 0$$

Instead, i chooses the option that offers some form of improvement at the lowest cost. Like the steps only heuristic, if options B and C are equally costly, respondents' choices are determined by a coin flip. Therefore, the probability that individual i chooses B is given by:

$$Prob(y_{iq} = B|Costs\ Only) = \begin{array}{l} 0 \text{ if } Cost_{qB} > Cost_{qC} \\ .5 \text{ if } Cost_{qB} = Cost_{qC} \\ 1 \text{ if } Cost_{qB} < Cost_{qC} \end{array}$$

The probability that i chooses C in any question is defined analogously. This heuristic represents a fraction π_{Costs} of the population.[16] For example, some respondents might have "attribute non-attendance" with a zero coefficient on cost in their utility function but might have

---

[16]   We have also experimented with a "high cost only heuristic." High cost only respondents always choose the species improvement with the highest cost. If options B and C have the same cost, high cost only respondents are assumed to randomly choose between the two options with probability 0.5 on both. We do not include this heuristic in our main specification because the set of respondents compatible with the high cost only heuristic overlaps the set compatible with the improvement steps only heuristic. More than twice as many respondents make choices that are consistent with the steps only heuristic than the high cost only heuristic and there are very few respondents who make choices that are consistent with the high cost only heuristic but inconsistent with the steps only heuristic. In a sensitivity analysis below, we include the high cost only heuristic in order to demonstrate how our results are affected by the inclusion of an unnecessary heuristic.

all the species improvement coefficients in common with the trade-off respondents.

4 Choose environmental improvements ignoring costs and benefits

Some survey respondents may want to indicate support for the environment and get the survey over as rapidly as possible. These respondents may always choose the first improvement, option B, and others may always choose the second (and last) improvement, option C. These heuristics represent π_B and π_C of the population. The probability an individual i chooses an option k in question q conditional on following one of these two heuristics is given by:

$$Prob(y_{iq} = k|B\ Only) = \begin{array}{l} 1\ if\ k = B \\ 0\ if\ k \neq B \end{array}$$

$$Prob(y_{iq} = k|C\ Only) = \begin{array}{l} 1\ if\ k = C \\ 0\ if\ k \neq C \end{array}$$

5 Randomizers ignoring costs and benefits

Last is a set of respondents who ignore the costs and the proposed improvements and act as if they were randomly choosing a response. This heuristic represents a fraction π_{Random} of the population. The other heuristics predict behavior with probability either one or zero depending on whether the decisions conform or not with the rule. Both the utility maximization rule and the random rule predict observed behavior with a probability between zero and one. An individual is judged by the estimation routine likely to be a randomizer if his or her decisions are incompatible with the other deterministic rules and also if the random model predicts better than utility maximization. Under this heuristic individuals choose options A, B, and C with probabilities π_{RandA}, π_{RandB}, and π_{RandC}, respectively.[17] Thus, the probability an individual i makes a sequence of choices Y_i is given by:

$$Prob(Y_i|Randomizer) = \pi_{RandA}^{3\ \Sigma=1\ (y_{iq}=A)} \pi_{RandB}^{3\ \Sigma=1\ (y_{iq}=B)} \pi_{RandC}^{3\ \Sigma=1\ (y_{iq}=C)}$$

[17] An alternative would be to model randomizers as having an equal probability of choosing A, B, or C. We chose to use the less restrictive rule of allowing the data to determine the probabilities of choosing each option.

ESTIMATION

A critical modeling assumption that we make is that individuals use the same decision rule for all three choice sets.[18] Therefore, the unconditional probability of observing individual *i* make a sequence of choices Y_i is a probabilistic mixture of the shares (π) and conditional probabilities (*Prob* ($Y_i|Choice\ Rule$)) described in the fourth section:

$Prob(Y_i)$ = $\pi_{TO}Prob(Y_i|Trade\text{-}offs)$ ++$\pi_{SQ}Prob(Y_i|Status\ Quo)$ +$\pi_{Steps}Prob(Y_i|Steps\ Only)$ + $\pi_{Costs}Prob(Y_i|Costs\ Only)$ + $\pi_B Prob(Y_i|B\ Only)$ + $\pi_C Prob(Y_i|C\ Only)$+$\pi_{Random}Prob(Y_i|Randomizer)$.

We multiply these probabilities across respondents and take the log to form the likelihood function:

$$Ln(Prob(Y_1, Y_2, \ldots, Y_n|\pi, \theta)) = \sum_i \log\left[\sum_{s=1}^{S} \pi_s Prob_s(Y_i|\theta_S)\right].$$

We estimate the model by via maximum likelihood in Matlab.[19,20] Maximum likelihood estimation of this model has the same estimating equations for respondents' preferences and WTP (α and β) as the standard mixed logit model, except that respondents' choices are weighted on the basis of whether they are likely to be utility maximizers. Below we demonstrate this result and provide additional intuition underlying identification of the model's parameters.

Let $\Theta \equiv (\alpha, \beta, \sigma)$ represent the parameters entering the mixed logit component of our model. Recall that these are the parameters determining WTP. Below, we compare the first-order conditions that characterize

---

[18] Mariel et al. (2011) allow choice rules to vary across questions for the same respondent. The authors allow respondents to attend to attributes in some questions but not others. In our analysis we make the assumption that individuals use the same decision rule in each choice set because of the data limitations of having only three choice sets per individual. In addition, the focus of our study was on showing the importance of including in the estimation methodology latent classes of individuals using heuristic decision rules not based on cost–benefit rules and not on identifying the optimal set of heuristics. We demonstrate that under the assumption that each individual's decision rules are fixed that our estimates are statistically superior to standard mixed logit results. A fertile area for future research would be to investigate whether respondents' decision rules change across choice sets.

[19] To determine the reliability of our estimation routine, we applied it to simulated data. Our estimation routine was able to recover the parameters used to simulate the data (subject to sampling error).

[20] We based our Matlab estimation routine on code provided by Professor Kenneth Train (accessed December 7, 2016 at http://eml.berkeley.edu/Software/abstracts/train1006mxlmsl. html). To incorporate heterogeneity into the mixed logit component of the model, we used shuffled Halton draws.

Θ when heuristics are included in the model to the first-order conditions when heuristics are excluded. Below $L_{MixLogit + Heuristics}$ represents the log-likelihood function when heuristics are added to the mixed logit model and $L_{MixLogit}$ represents the log-likelihood function when heuristics are not included.

In our model and the mixed logit model, Θ is chosen to maximize the probability of respondents' observed choices generated by the model. Both first-order conditions are satisfied by a Θ such that a weighted sum of $\frac{\partial Prob(Y_i|Trade\text{-}offs)}{\partial\Theta}$ across respondents equals zero. The first-order conditions associated with our model and the mixed logit model differ only in how respondents' choices are weighted:

$$\frac{\partial L_{MixLogit + Heuristics}}{\partial\Theta} = \sum_{i=1}^{N} w_{MixLogit + Heuristics,i} \frac{\partial Prob(Y_i|Trade\text{-}offs)}{\partial\Theta} = 0$$

$$\frac{\partial L_{MixLogit}}{\partial\Theta} = \sum_{i=1}^{N} w_{MixLogit,i} \frac{\partial Prob(Y_i|Trade\text{-}offs)}{\partial\Theta} = 0$$

Thus, the nature of variation in the data used to identify Θ (and, WTP) in the mixed logit model and a model merging the mixed logit with heuristics is the same. The parameters are chosen so the observed propensity of respondents to choose species improvements at stated costs over the status quo is best explained by the model.

However, the measures of WTP emerging from the mixed logit and our model will differ because of differences in how the models weigh respondents. A careful examination of the weights in the first-order conditions illustrates differences in identification across the two models. For individual i:

$$w_{MixLogit,i + Heuristics,i} = \frac{\pi_{TO}}{Prob(Y_i)}$$

$$= \frac{\pi_{TO}}{\pi_{TO}Prob(Y_i|Trade\text{-}offs) + \cdots + \pi_{Random}Prob(Y_i|Randomizer)}$$

and

$$w_{MixLogit,i} = \frac{1}{Prob(Y_i|Trade\text{-}off Rule)}$$

Dividing $w_{MixLogit + Heuristics,\ i}$ by $w_{MixLogit,\ i}$ and applying Bayes' rule yields:

$$\frac{w_{MixLogit + Heuristics,i}}{w_{MixLogit,i}} = Prob(i\ is\ a\ Trade\text{-}off\ Respondent|Y_i)$$

Given *i*'s choices, *Prob*(*i is a Trade-off Respondent* $|Y_i$) expresses the probability that *i* is maximizing a neoclassical utility function. When a respondent's choices are easily rationalized by the mixed logit model, this probability is high. When *i*'s choices cannot be easily rationalized by the mixed logit model, and it appears that *i* is following a heuristic rule, this probability is low. In contrast to the mixed logit model, our estimation routine places greater weight on respondents who appear to be more likely to be maximizing utility. Our routine appropriately places less weight on respondents whose choices appear less likely to reflect trade-offs between species improvements and costs.[21]

Next, we discuss identification of the heuristic shares. Let π_h denote the share of respondents following an arbitrary heuristic. The first-order condition associated with $\pi_{heuristic\ h}$ is:[22]

$$\sum_{i=1}^{N} \frac{Prob(Y_i|heuristic\ h)}{Prob(Y_i)} = \sum_{i=1}^{N} \frac{Prob(Y_i|Trade\text{-}offs)}{Prob(Y_i)}$$

For any individual *i*, $Prob(Y_i)$ is equal to the portion of *i*'s choice explained by the estimated model and $\frac{Prob(Y_i|heuristic\ h)}{Prob(Y_i)}$ equals the fraction of the model's fit attributable to heuristic *h*. The heuristic shares are chosen so the marginal contribution of each heuristic in explaining respondents' behavior (averaged across respondents) is equal across choice rules.

Intuitively, the estimation routine evaluates each heuristic's marginal contributions using variations in choice sets across respondents. Recall that each respondent is given one of 44 versions and 16 sub-versions of the NOAA survey so that there is extensive variation in the species status improvements and costs offered to the respondents. Each choice rule makes unique predictions about how patterns in respondents' choices should vary across choice sets. We identify heuristic shares on the basis of whether these unique predictions are confirmed in the data. For example, we identify the share of respondents following the status quo heuristic (π_{SQ}) by examining changes in respondents' choice patterns when choice sets' species status improvements are fixed but their costs vary. Respondents following the status quo heuristic will indiscriminately choose option A regardless of whether species improvements are offered at low or high costs. Trade-off respondents, however,

---

[21] When choosing the parameters that determine WTP, the estimation routine does place weight on all respondents. This is because all respondents' choices are potentially explained by the trade-off rule. No respondents are assigned to a decision rule with certainty.

[22] This formulation of the first-order conditions relies on the identity $\pi_{TO} = 1 - \pi_{Steps} - \pi_{Costs} - \cdots - \pi_{Random}$

will choose the status quo more often when they are presented choice sets with relatively high costs. Our estimation routine chooses π_{SQ} so that the estimated model best explains changes in respondents' observed behavior related to this choice set variation. If the share of respondents choosing option A in all three questions varies little with options B and C's costs, the model will infer that a large fraction of respondents consistently choose option A because they follow the status quo heuristic. Similar intuition explains identification of the other heuristic shares.

RESULTS

In this section we first describe the estimates implied by a model that includes all the heuristics, and then we show how much these estimates change if the set of alternative heuristics is varied. We argue that the estimates lack credibility if the set of alternative heuristic decision rules is too narrow (e.g., none), and we show that the estimates are fragile (very dependent on the particular mix of heuristics that is allowed) when the set of heuristics is wide enough to be credible.

Estimated Population Shares of Choice Rules

Our main specification is a probabilistic mixture of all the heuristics described above plus a mixed logit group of trade-off respondents. The first column of Table 1 reports estimates of the population shares of the seven choice rules of this specification and also the choice probabilities of the maximizers.[23]

The six simple heuristic rules absorb a large majority of respondents and only 22.4% are better explained by the trade-off rule.[24] A larger fraction is estimated to follow the status quo and randomizer heuristics (24.6% and 27.8% respectively) and large fractions of respondents (12.4% and 8.7%) are estimated to behave according to the steps only and costs only heuristics. According to the estimated model, few

----

[23]   The results in Table 1 and all subsequent tables are based on a sample of 11,459 respondents who answer all three survey questions. The choices of 1,822 respondents who did not answer all three survey questions were excluded from the analysis. Of these respondents 1,478 failed to answer one question. The remaining 344 respondents failed to answer two of the three questions.

[24]   We expect that incorporating a more comprehensive set of heuristic rules will reduce this share further; in addition, using heuristic rules that assume errors in respondents' choices would likely increase the estimated heuristics shares and reduce the estimated trade-off share.

Table 1   Estimated choice rule shares

	Heuristic Shares (%)	St. Err. (%)
Trade-off	22.40	0.80
Status quo only (Always A)	24.60	0.90
Steps only	12.40	0.40
Costs only	8.70	0.40
B or C only	4.20	0.30
Randomizers	27.80	0.80
Probability of A	4.70	0.30
Probability of B	44.50	0.60
Probability of C	50.80	0.60

respondents systematically choose options B and C and belong to the B and C only heuristics (4.2%). Finally, the estimates imply that respondents following the randomizer heuristic tend to ignore option A. They randomly choose between the two improvement options, with a higher tendency to choosing the middle option B (50.8% of respondents following the randomizer heuristic) rather than option C (44.5% of respondents following the randomizer heuristic). In other words, this turns out to be describing random environmentalists. If we sum the shares of the steps only heuristic (12.4%), the costs only heuristic (8.7%), and the randomizer heuristic (27.8%), we find that 48.9% of survey respondents reject the status quo regardless of how expensive the two environmental options may be.

These population shares are estimated with high precision, per the standard errors reported in column two. These standard errors imply, for example, that there is a 90% chance that the share of respondents following a trade-off decision rule is between 21% and 24.5%.

A deterministic heuristic rule predicts behavior with probability one or zero. For example, the status quo rule perfectly predicts the choices if the status quo was chosen for all three choice sets, but otherwise is incompatible with these choices. This allows us to separate the observations into those that are perfectly compatible with the rule and those that are incompatible. Table 2 compares the estimated heuristic shares with the % of respondents whose responses were perfectly compatible with the rule.

Responses that are compatible with more than one decision rule are counted more than once in the percentage of respondents column and these have to be allocated probabilistically across the alternatives including the utility maximization rule to compute the estimated population shares. Consequently, in all cases the estimated population share is

Table 2   *Percentage of respondents whose behavior is consistent with heuristic rules*

	% Respondents	(%) Estimated Heuristic Share
Status quo only (Always A)	27.0	24.6
Steps only	24.4	12.4
Cost only	18.2	8.7
B or C only	16.7	4.2

smaller than the fraction of compatible respondents. This has the smallest effect on the protestors (who always choose the status quo) because most of the cases in which the status quo was chosen all three times are not well explained by any of the other rules, which is a partial justification for the common procedure of dropping these "protestors."

While the data cannot reveal exactly which of the decision rules each respondent followed, the estimated model and data do allow us to compute an estimated probability that a respondent was following one of the decision rules.[25] Estimates of the parameters of the trade-off rule with maximum likelihood place more weight on respondents who were probably using the trade-off rule. Since respondents who chose the status quo in all three questions are unlikely to have done so while following the trade-off decision rule, the estimation routine discounts these respondents' choices when determining the trade-off rule parameters and WTP.

Next, we turn to our measures of WTP for various species status improvements that are calculated from estimates of species valuations ($\bar{\beta}_{sk}$ and σ_{sk}) and cost sensitivity (α) parameters of the mixed logit model. It is important to emphasize these are the WTPs of trade-off respondents, who constitute fewer than one-quarter (22.4%) of respondents.[26]

Estimated Willingness to Pay for Species Status Improvements

Table 3 presents estimates of WTP for species status improvements. These measures reflect the choices of respondents whose patterns of

---

[25]   Intuitively, the more favorable the improvement options are in the respondent's choice set (large improvements at low costs), the less likely it is that the respondent picked the status quo because the respondent was rationally weighing costs and benefits.

[26]   Moreover, these WTP estimates would change if more heuristic rules were added and if we took into account the types of context effects we analyze in a companion paper, Prelec et al. (forthcoming).

Table 3 Estimated mean willingness to pay for species improvements

	Trade-off Respondents		Population	
	Mean WTP ($)	Standard error ($)	Mean WTP ($)	Standard error ($)
One step from endangered to threatened				
Smalltooth sawfish	10.30	2.30	2.41	0.54
Leatherback turtle	11.08	3.12	2.59	0.73
Hawaiian monk seal	15.60	2.02	3.65	0.47
North Pacific right whale	14.13	2.04	3.31	0.48
North Atlantic right whale	12.15	2.47	2.84	0.58
One step from threatened to recovered				
Upper Willamette River Chinook salmon	15.78	2.53	3.69	0.59
Puget Sound Chinook salmon	17.92	2.37	4.19	0.55
Loggerhead	18.55	1.61	4.34	0.38
Two steps from endangered to recovered				
Smalltooth sawfish	16.98	2.03	3.97	0.71
Hawaiian monk seal	23.93	2.63	5.60	0.62
Leatherback turtle	27.00	3.00	6.32	0.70
North Pacific right whale	23.52	4.32	5.50	1.01
North Atlantic right whale	23.92	4.12	5.60	0.96
Average one-step	14.44		3.38	
Average two-step	23.07		5.40	

Note: The population WTP is equal to the trade-off respondents' WTP multiplied by the share of respondents following the trade-off rule. The population WTP estimates assume that the non-trade-off respondents have zero WTP because there is no evidence in the data to support non-zero WTPs for these respondents.

responses indicate that they are likely to use a trade-off decision rule.[27] While the mixed logit model generates a distribution of WTP across individuals (reflecting the distributions of β), we only present the estimated means of these distributions below and the standard errors of the means.[28]

[27]   Since all respondents' choice rules are unobserved, technically all respondents' choices enter the WTP calculation. However, the estimation routine places much more weight on respondents likely to be maximizing utility. For example, a respondent following the trade-off choice rule with 75 percent probability will receive 15 times as much weight as a respondent following the trade-off rule with 5 percent probability.

[28]   For each species improvement, we report $\frac{\hat{\bar{\beta}}_s}{\hat{\alpha}}$ where $\bar{\beta}_s$ and $\hat{\alpha}$ are point estimates of $\bar{\beta}_s$ and α. To calculate the standard error of mean willingness to pay, we perform a bootstrap

The first set of columns reports mean WTP estimates of the trade-off respondents. The mean WTP estimates range from $10.30 (WTP for the smalltooth sawfish improvement from endangered to threatened) to $27.00 (WTP for the leatherback turtle improvement from endangered to recovered). The second set of columns presents average WTP across all respondents. This overall WTP estimate is obtained by multiplying the estimated WTP of trade-off respondents by the estimated share of respondents who are utilizing the trade-off rule, thus assigning zero WTP to all non-trade-off respondents. While some of these non-trade-off respondents may have positive WTP, the model and the database we are using do not allow us to infer their WTP from their responses.

Additional Specifications: Sensitivity Analysis

It seems clear that analyses of stated preference surveys have to allow for protestors and other types of non-conforming respondents. But what other types should be considered? That is not at all clear. While the utility maximization hypothesis does not dictate all the features of the logit or mixed logit models, the extent of model ambiguity in a utility maximization model seems very small compared with the model ambiguity that afflicts studies that allow for non-conforming respondents. Inferential fragility thus becomes a very serious issue. Do minor changes in the model lead to major changes in the inferences? To explore the fragility of estimates of WTP, we present in this section results from models with different sets of included heuristics, beginning with the traditional model that includes no heuristics.

Tables 4 and 5 contain results of ten models that have different lists of included heuristics. Table 4 reports the estimated heuristic shares and measures of mean WTP (averaged across species improvements and respondents) implied by each model. Table 5 reports measures of fit (likelihood value and Bayesian information criterion [BIC]) for each model. The mixed logit trade-off rule is included in all ten models and the estimated fraction of the population using this rule is reported in the first column of Table 4 labeled "Trade-off." The intent of Table 4 is to help discover how much the estimates of WTP depend on the choice of heuristics and which heuristics matter most. Table 5 supplements Table 4 with information on measures of fit.

The first two rows of Table 4 contain the results of traditional mixed logit models with and without the Lew and Wallmo "protestors" included.

---

using the estimated distributions of $\hat{\bar{\beta}}_S$ and $\hat{\alpha}$. These standard errors are measures of how precisely mean willingness to pay is estimated and not the amount of heterogeneity across respondents.

The third row refers to a model that includes only the mixed logit model and the status quo heuristic. Row (4) includes all of the heuristics plus a high cost only heuristic.[29] Rows (5) through (10) refer to models that include all the heuristics but one.

The first seven columns of numbers are the estimated fractions of each of the heuristic types, and columns (8) to (10) report the estimated choice probabilities for the randomizers. The last two columns report average WTP, first for the trade-off respondents and then for the population overall, assigning zero WTP to non-trade-off respondents.

The purpose of rows (1) through (3) is to evaluate the implications of different treatments of protestors. In row (1), we report estimates using the choices of all 11,459 respondents.[30] In row (2), we exclude from the data set the 24.4%[31] of respondents labeled "protestors" by Lew and Wallmo. These Lew and Wallmo "protestors" are respondents who both chose the status quo in all three choice sets and also revealed in follow-up questions confusion regarding the task they had performed. In row (3), we include all 11,459 respondents in the data set but use a model that adds the status quo heuristic to the mixed logit model. Notice how similar are rows (2) and (3), the former using the Lew and Wallmo definition of protestors and the latter using the protestor heuristic to drive the trimming. The estimated fractions of trade-off respondents are 75.6% and 73.5%, and the estimates of mean willingness to pay of the trade-off respondents are $45.80 and $51.20 compared with $21.34 for the row (1) case with all the data included. Also note that the population WTP estimates in row (2) and row (3) are both substantially larger than the row (1) estimate based on all the data, because the increase in the estimated WTP for the trade-off respondents in rows (2) and (3) more than offsets the reduced fraction of respondents to which the estimated WTP applies.

Although rows (2) and (3) yield very similar conclusions, it is important to understand that the model-driven results in row (3) do not exclude respondents who always chose the status quo since respondents who in all three choice sets saw high prices for the environmental improvements may be well explained by the utility maximization model. The model-driven estimates of WTP put lower weights but not zero weights on the respondents who are reasonably well explained by the utility maximization model. To put it another way, row (3) uses the choice model to infer the likelihood a respondent's status quo choices result from cost–benefit trade-offs and

[29] See footnote 18 for a description of the high cost heuristic.
[30] The NOAA data actually contain 13,684 respondents. However, we exclude respondents who do not answer each of the three questions posed to them.
[31] 2,800.

Table 4 Sensitivity analyses

Model Specifications	Estimated Heuristic Shares (%)							Randomizers' choice probabilities			WTP ($)	
	Trade-off	Status quo only (always A)	Steps only	B or C only	Costs only	High cost option only	Randomizer	Status quo	Option B	Option C	Trade-off respondents	Population
(1) Traditional mixed logit with all data	100										21.34	21.34
(2) Traditional mixed logit excluding protestors	76										45.80	34.61
(3) Include status quo heuristic only	73.5	26.5									51.20	37.62
(4) All heuristics including high cost only heuristic	22.4	24.6	12.4	4.2	8.7	0.0	27.8	4.7	44.5	50.8	17.28	3.87
(5)	41.7		10.8	10.0	3.1	1.8	30.5	96.1	2.1	1.8	39.38	16.41
(6)	33.9	26.2		3.5	9.8	3.5	23.0	32.9	32.8	34.3	59.03	20.00
(7) Include all heuristics but one	25.0	24.8	12.8		8.7	0.2	28.6	3.2	48.9	47.9	18.70	4.67
(8)	29.1	24.9	10.9	4.5		0.0	30.7	3.1	45.6	51.4	20.48	5.95
(9)	22.4	24.6	12.4	4.2	8.7		27.8	4.7	44.5	50.8	17.28	3.86
(10)	44.6	25.9	12.1	5.1	10.2	2.0					33.44	14.90

Notes:
Mean WTP is averaged across both species improvements and respondents.
Non-conforming respondents assigned zero WTP.
The row labeled Always B or C combines the Always B and Always C heuristics.
The population WTP estimates assume that the non-trade-off respondents have 0 WTP because there is no evidence in the data to support non-zero WTPs for these respondents.

244

Table 5   Sensitivity analyses measures of fit

Model Specifications	Measures of Fit	
	Mean likelihood value	BIC criterion
(1) Traditional mixed logit with all data	−2.934	5.886
(2) Traditional mixed logit excluding protestors	−2.828	5.678
(3) Include status quo heuristic	−2.566	5.151
(4) All heuristics including high cost only heuristic	−2.377	4.777
(5)	−2.415	4.840
(6)	−2.433	4.889
(7) Include all heuristics but one	−2.395	4.811
(8)	−2.405	4.833
(9)	−2.377	4.777
(10)	−2.446	4.914

Notes:
Mean WTP is averaged across both species improvements and respondents.
Non-conforming respondents assigned zero WTP.
The row labeled Always B or C combines the Always B and Always C heuristics.

"trims" the data by placing lower weight on respondents whose status quo choices cannot be rationalized by utility maximization, while row (2) uses responses to ad hoc follow-up questions to exclude status quo choices that may be the result of utility maximization.

It is noteworthy that we obtain similar results in (2) and (3) using alternative strategies to deal with protestors. However, comparing rows (1) with either (2) or (3) reveals the large impact of protestor treatment, either by excluding them (row 2) or by absorbing them into a protestor heuristic (row 3), in both cases substantially increasing the estimated WTP of the trade-off respondents.

The model represented in row (4) contains the same heuristics as our main specification and an additional heuristic for respondents who systematically choose the species improvement alternative with the highest cost instead of making cost–benefit trade-offs. Survey respondents might adopt this choice strategy if they believe the highest cost option always does the most to improve the environment. Including the high cost only heuristic doesn't really matter much if it is accompanied by other heuristics. This reminds us that adding additional heuristics won't matter if the new behavior can be mimicked by one of the heuristics already in the

model.[32] Because of this, the high cost only heuristic has a near-zero share in the all-heuristics model reported in row (4).

In rows (5) through (10) of Tables 4 and 5, we remove heuristics from the model one at a time. There are two purposes behind this exercise. First, we evaluate the contribution from each heuristic to goodness of fit. Second, we evaluate whether measures of willingness to pay are reliable when an incomplete set of heuristics is used.

Table 5 indicates that models with heuristics better explain respondents' choices than models without heuristics.[33] We have used maximum likelihood estimation, and the likelihood value at the maximum is a standard measure of the fit of the model. The model that includes all heuristics (row 4) has a maximum likelihood value equal to –2.377, better than traditional mixed logit models or models using sets of heuristics known to be incomplete. This, however, is a simple consequence of the fact that all the other models are special cases of this general model. The issue is not whether this model allows a greater likelihood value; the issue is whether the increment is great enough when adjusted for the number of parameters included. To account for the size of the models we report the Bayesian information criterion (BIC) (or, Schwarz criterion). The BIC is one of several measures of fit that adjust for the number of parameters with more parameters.[34] Models with small BIC measures better fit the data than models with large BIC measures. Here, the smallest BIC equal to 4.77 applies to the general model in row (4) and the model without the high cost only heuristic in row (9). On this basis, we conclude it is appropriate to include each of the heuristics used in our main specification.

The measures of willingness to pay we report in Table 3 are reliable only if one of two conditions hold. First, our measured willingness to pay is reliable if our underlying model is properly specified and we have included in our model all relevant heuristics survey respondents rely upon when not making cost–benefit trade-offs. As this set of potential heuristics is very large, we believe it is very unlikely our model includes every possible heuristic.[35] Unfortunately, it is not practical to test our first condition

---

[32]   Of the 11,459 respondents, 2,795 respondents' choices are consistent with the steps only environmentalist heuristic and 1,353 respondents' choices are consistent with the high cost heuristic. Of these 1,353 respondents, only 172 respondents make choices that are inconsistent with the steps only environmentalist heuristic.

[33]   The WTPs reported in Table 5 are averages across both species improvements and respondents.

[34]   Greene (2002), at p. 160.

[35]   Indeed, we do not include heuristics that absorb respondents who behave strategically or who have poorly formed preferences as in McNair et al. (2012). Nor do we include heuristics to absorb respondents whose trade-offs are affected by context as in Prelec et al. (forthcoming). Many other heuristics are described in Gilovich et al. (2002).

because it is not possible to list all possible heuristics, much less incorporate them into an econometric model of choice. If our first condition is not satisfied or untestable, however, it might be the case that biases in willingness to pay that emerge from models with incomplete sets of heuristics are small in practice. If this second condition is satisfied, the measures of willingness to pay reported in Table 3 are reliable.

In Table 4, we investigate the second condition listed in the preceding paragraph by estimating models with sets of heuristics known to be incomplete. Using this top-down approach, except for the case of the high cost only heuristic, which has a near zero estimated share in row (4), omitting any one of the heuristics increases both the WTP of the trade-off respondents and the WTP of the population. A reduction in the list of included heuristics inevitably increases the estimated fraction of trade-off respondents because there are fewer competitors for the trade-off rule but the effect on the estimated WTP is not clear-cut. The estimated trade-off fraction in the complete model in row (4) is 22.4%, in contrast to the range of estimates from 22.4%, to 44.6% in rows (5) to (10).

When heuristics are removed from the model one at a time, estimated WTP increases from $17.28 to values between $18.70 and $59.03 depending on which heuristic is removed from the model.[36] The biggest effect on WTP occurs in row (6) when the steps only heuristic is omitted and mean willingness to pay increases from $17.28 to $59.03. This increase occurs because the model needs to explain the observed B and C choices without the benefit of the 12.4% of the respondents who are steps only respondents in the all-heuristics model and it does this by both increasing the percentage of respondents who are trade-off respondents and by reducing the cost sensitivity of those respondents (resulting in more B and C choices by that class). The exclusion of the status quo only heuristic in row (5) also substantially increases WTP, from $17.28 to $39.38. To understand the source of this increase, consider the randomizers' choice probabilities before and after the status quo only heuristic is excluded. Before the exclusion, randomizers accept species improvements (options B and C) in 95.3% of questions. After the exclusion, randomizers choose the status quo in 96.1% of questions. In other words, the randomizer heuristic mimics the status quo only heuristic in row (5) and respondents who were classified as randomizers in row (4) are classified as trade-off respondents in row (5). The model now needs to explain the observed B and C choices, which it

---

[36]   We do not include the model when we exclude the high cost only heuristic (row 9) in this range since the estimated share on the high cost only heuristic in row (4) is zero. So it is no surprise that WTP in rows (4) and (9) are identical.

does by increasing the share of trade-off respondents and reducing their cost-sensitivity by lowering the absolute value of the cost coefficient.

Table 4 indicates that reliance on incomplete sets of heuristic potentially introduces substantial biases into measures of willingness to pay. These results demonstrate the challenge of using stated preference surveys to measure WTP. The set of heuristics included in row (4) of Table 4 is incomplete and to some extent arbitrary and accidental. Although this initial foray into alterations of the model to allow non-conforming respondents offers explains choices better than traditional mixed logit models, it has an "accidental" feel to it, meaning that a different set of analysts would likely propose a different list of heuristics. Unfortunately, no one can say with confidence what estimates would result if different or more complete sets of heuristics are employed. Perhaps the bias in row (4) from omitted heuristics is small and relying on a more complete set of heuristics would decrease WTP by 7.5% or less, which occurred when add the B only or C only heuristics to the model (i.e., we go from row 7 to row 4). But the bias from relying on an incomplete set of heuristics could be much larger, as was the case in rows (5) and (6) when the status quo only and steps only heuristics were excluded from the model. On the other hand, there is a downside to including too many heuristic rules. With only three choices per respondent some utility maximizing individuals will match the choice profile of one of the heuristics by chance. Including more heuristic rules increases this probability, reducing the fraction of the population estimated to be utility maximizers. Exploring the issues how to determine the "best" heuristic rules and the optimal number of heuristic rules are beyond the scope of this chapter. For these and other reasons, we regard the results reported in Table 4 to be suggestive but they are not definitive estimates of WTP.

CONCLUSIONS

Wallmo and Lew (2011) have offered a traditional mixed logit analysis of the stated preference data collected by a NOAA endangered species survey. After excluding what they regard to be protestors, they find an estimated WTP averaged across respondents and species improvements equal to $45.80. We have analyzed the same data using models that allow various types of non-conforming respondents – those who do not behave according the traditional utility maximization model. Our most general model has a mixture of six heuristic rules and also the traditional mixed logit utility maximization rule. This model estimates that only 23.4% of respondents were conforming with the utility maximization model. The randomizers have the largest estimated population share equal to 24.8%.

These respondents choose the status quo in 3.7% of questions, the species improvement in column B in 53.4% of questions, and column C in 43.0% of questions – basically flipping a coin to decide which of the species options to select.

This model offers a substantially improved fit of the data relative to the mixed logit model and is the best fitting model according to the BIC criterion of any considered here. It supports an estimate of WTP equal to $17.28 for the 22.4% of respondents who are maximizing utility and $3.87 for the population as a whole after assigning $0 WTP to the other respondents.

We emphasize that our model is not intended as a realistic and comprehensive description of the survey respondents' decisions. The heuristic decision rules we have included are stylized and limited in number. Together with our companion paper, Leamer (2016), which extends the heuristic decision rules to allow individual wavering, what we have accomplished is a demonstration of the inadequacy of the traditional utility maximization model for estimation of environmental valuation. But this isn't news to analysts who omit subsets of aberrant respondents, including Lew and Wallmo who omit the non-conforming respondents they call protestors. To those who omit protestors, we offer the rhetorical question: "Why not more?" and we demonstrate the very large impact when more heuristics are included in the model and when the effective trimming of the sample is two-sided, not just protestors.

REFERENCES

Burton, M. and D. Rigby (2009), "Hurdle and latent class approaches to serial non-participation in choice models," *Environmental and Resource Economics*, **42**(2), 211–26.

Campbell, D., D.A. Hensher, and R. Scarpa (2011), "Non-attendance to attributes in environmental choice analysis: A latent class specification," *Journal of Environmental Planning and Management*, **54**(8), 1061–76.

Cerdaa, C., A. Ponceb, and M. Zappic (2013), "Using choice experiments to understand public demand for the conservation of nature: A case study in a protected area of Chile," *Journal of Nature Conservation*, **21**(3), 143–53.

Cunha-e-Sa, M., L. Madureira, L. Nunes, and V. Otrachshenko (2012), "Protesting and justifying: A latent class model for contingent valuation with attitudinal data," *Environmental and Resource Economics*, **52**(4), 531–48.

Gilovich, T., D. Griffin, and D. Kahneman (2002), *Heuristics and Biases: The Psychology of Intuitive Judgement*, Cambridge, UK: Cambridge University Press.

Greene, W. (2002), *Econometric Analysis*, 5th edition, New York: New York University Press.

Greene, W.H. and D.A. Hensher (2013), "Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model," *Applied Economics*, **45**(14), 1897–902.

Hanley, N., R.E. Wright, and B. Alvarez-Farizo (2006), "Estimating the economic value of improvements in river ecology using choice experiments: An application to the water framework directive," *Journal of Environmental Management*, **78**(2), 183–93.

Hensher, D.A. (2010), "Attribute processing, heuristics and preference construction in choice analysis," in S. Hess and A. Daly (eds), *Choice Modelling: The State-of-the-Art and the State-of-Practice – Proceedings from the Inaugural International Choice Modelling Conference*, Bingley, UK: Emerald Group Publishing, pp. 35–70.

Hensher, D.A. and W.H. Greene (2010), "Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification," *Empirical Economics*, **39**(2), 413–26.

Hensher, D.A., J.M. Rose, and W.H. Greene (2005), "The implications on willingness to pay of respondents ignoring specific attributes," *Transportation*, **32**(3), 203–22.

Hensher, D.A., J.M. Rose, and W.H. Greene (2012), "Inferring attribute non-attendance from stated choice data: Implications for willingness to pay estimates and a warning for stated choice experiment design," *Transportation*, **39**(2), 235–46.

Hess, S., D.A. Hensher, and A. Daly (2012), "Not bored yet – revisiting respondent fatigue in stated choice experiments," *Transportation Research*, **46**(A), 626–44.

Hess, S., A. Stathopoulos, D. Campbell, V. O'Neill, and S. Caussade (2012), "It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity," *Transportation*, **40**(3), 583–607.

Kaffashi, S., M.N. Shamsudin, A. Radam, M.R. Yacob, K.A. Rahim, and M. Yazid (2012), "Economic valuation and conservation: Do people vote for better preservation of Shadegan International Wetland?," *Biological Conservation*, **150**(1), 150–58.

Knowledge Networks (2009), *Field Report, Protected Species Valuation Survey Conducted Under Contract to Pacific States Marine Fisheries Commission and OAK Management, Inc., Submitted to Dr. Kristy Wallmo, Ph.D., July 31, 2009*.

Layton, D.F. and G. Brown (2000), "Heterogeneous preferences regarding global climate change," *The Review of Economics and Statistics*, **82**(4), 616–24.

Leamer, E.E. (2016), "Tradeoffs between costs and benefits in choice experiments, or not," working paper.

Lew, D.K. and K. Wallmo (2011), "External tests of scope and embedding in stated preference choice experiments: An application to endangered species valuation," *Environmental and Resource Economics*, **48**(1), 1–23.

Lew, D.K., D.F. Layton, and R.D. Rowe (2010), "Valuing enhancements to endangered species protection under alternative baseline futures: The case of the Steller sea lion," *Marine Resource Economics*, **25**(2), 133–54.

Mariel, P., J. Meyerhoff, and D. Hoyos (2011), "Stated or inferred attribute non-attendance? A simulation approach," working paper presented at International Choice Modelling Conference, Leeds, UK.

McNair, B.J., D.A. Hensher, and J. Bennett (2010), "Modelling heterogeneity in response behaviour towards a sequence of discrete choice questions: A latent

class approach," *Environmental Management & Development Occasional Paper, No. 16*, Crawford School of Economics and Government, The Australian National University.

McNair, B., D.A. Hensher, and J. Bennett (2012), "Modelling heterogeneity in response behaviour towards a sequence of discrete choice questions: A probabilistic decision process model," *Environmental and Resource Economics*, **51**(4), 599–616.

Meyerhoff, J. and U. Liebe (2006), "Protest beliefs in contingent valuation: explaining their motivation," *Ecological Economics*, **57**(4), 583–94.

Meyerhoff, J. and U. Liebe (2008), "Do protest responses to a contingent valuation question and a choice experiment differ?," *Environmental and Resource Economics*, **39**(4), 433–46.

Meyerhoff, J., A. Bartczak, and U. Liebe (2012), "Protester or non-protester: A binary state? On the use (and non-use) of latent class models to analyse protesting in economic valuation," *The Australian Journal of Agricultural and Resource Economics*, **56**(3), 438–54.

Olsen, S.B. (2009), "Choosing between Internet and mail survey modes for choice experiment surveys considering non-market goods," *Environmental and Resource Economics*, **44**(4), 591–610.

Parsons, G.R. and S.M. Thur (2008), "Valuing changes in the quality of coral reef ecosystems: A stated preference study of SCUBA diving in the Bonaire National Marine Park," *Environmental and Resource Economics*, **40**(4), 593–608.

Plott, C.R. (1996), "Rational individual behavior in markets and social choice processes: The discovered preference hypothesis," in K. Arrow, E. Columbatto, M. Perlman, and C. Schmidt (eds), *The Rational Foundations of Economic Behavior*, London: Macmillan.

Prelec, D., J. Burrows, and P. Dixon (forthcoming), "Context sensitivity in stated preference experiments," available from the paper's authors on request.

Scarpa, R., T.J. Gilbride, D. Campbell, and D.A. Hensher (2009), "Modelling attribute non-attendance in choice experiments for rural landscape valuation," *European Review of Agricultural Economics*, **36**(2), 151–74.

Wallmo, K. and D.K. Lew (2011), "Valuing improvements to threatened and endangered marine species: An application of stated preference choice experiments," *Journal of Environmental Management*, **92**(7), 1793–801.

Wallmo, K. and D.K. Lew (2012), "Public willingness to pay for recovering and downlisting threatened and endangered marine Species," *Conservation Biology*, **26**(5), 830–39.