1. Response to cost prompts in stated preference valuation of environmental goods[1]

James Burrows, Powell Dixon, and Hiu Man Chan[2]

INTRODUCTION

The stated preference discrete choice experiment, also known as conjoint analysis, is now a standard method for estimating non-use values of natural resources from respondents' answers to survey questions.[3] A choice experiment (CE) consists of a sequence of choices among several options, each offering various combinations of features together with costs (often expressed as taxes imposed on each household over some number of years). Through their choices respondents are presumed to reveal whether they would accept a given cost in exchange for a better level of a natural resource. These hypothetical "votes" are then fed into econometric models to produce willingness-to-pay (WTP) estimates, both individual and aggregate, which are the ultimate objects of interest in most applications.[4] In the standard random utility model (RUM) widely used for inferring WTP from data from choice experiments, the underlying utility functions are

[1] The authors gratefully acknowledge the essential contributions to this chapter by Drazen Prelec, Ed Leamer, Renee Miller-Mizia, Jerome Genser, and Stamatia Kostakis.

[2] Respectively, Vice Chairman, Charles River Associates; Associate Principal, Charles River Associates; Vice President, Analysis Group.

[3] Boxall et al. (1996) commented: "For approximately 30 years contingent valuation (CVM) methods have been employed by economists to value environmental goods and services. . . Other types of [stated preference] approaches capable of eliciting environmental preferences have not been widely used in environmental valuation." According to Hanley et al. (1998), the choice experiment ("CE") technique had not been applied to environmental management problems until 1994 by Adamowicz et al., and the first application of the technique to estimating non-use or passive use values was as late as 1998 by Adamowicz et al. By 2008, however, Bateman et al. indicated that CE "has become the most popular approach for valuing a range of multi-attribute public goods."

[4] See Train (2009) and Hensher et al. (2015), commonly used reference texts for discrete choice modeling.

assumed to be logit and uniform throughout the cost distribution. Under these assumptions, the estimated WTPs should be independent of the cost scale, as long as the cost scale is in a range in which some respondents select a do-nothing option and some select a do-something option.

Because of the unfamiliar nature of many environmental improvements, and the unfamiliar task of evaluating a non-market good, survey respondents may not know their WTP for environmental amenities or even how to think about their WTP. To perform the choice tasks, the respondents may look for clues in the survey to assist in determining what they think is an appropriate or reasonable WTP. In particular, the costs that are offered in the survey might affect what respondents think they are, or should be, willing to pay. If the effect is small, then it can perhaps be ignored, but if two studies that present different costs but are otherwise identical lead to large differences in willingness-to-pay estimates, it would be inappropriate to assume – without additional analysis – that either set reveals respondents' prior valuations of the natural resource.

A number of studies have documented cost-anchoring effects (or starting-point bias) in double-bounded or multiple-bounded dichotomous choice CV surveys, in which respondents are presented with follow-up valuation questions after the first valuation question – examples include Silverman and Klock (1989), Herriges and Shogren (1996), Green et al. (1998), Frykblom and Shogren (2000), and Flachaire et al. (2007). Other studies, including most notably Ariely et al. (2003), have shown anchoring effects of costs that respondents see outside of the survey itself (for example, costs seen in a different context in advance of the survey or in test questions for the survey).

Several studies have examined the influence of cost scales on WTPs estimated from CV surveys; these differences generally arise not from the effects of cost anchoring but on the effects of truncating responses in the tails of the distribution, which can affect estimated WTPs if the missing observations do not accord with the functional form of the assumed distribution function. Cameron and Huppert (1989) examine the impact of truncating higher tails of the cost vector and find that this can change the estimated WTPs. They conclude that:

> [. . .]it seems that an unscrupulous researcher could readily influence the estimated total value of the resource by appropriately tailoring the upper intervals of the payment card, making a judicious choice of the arbitrary "midpoint" for that interval, and then selecting either the medians or the means in order to achieve the desired effect. (Cameron and Huppert, 1989, p. 241)

Duffield and Patterson (1991) focus on the selection of the measure of central tendency and the question of sample allocation. They find through simulation that the optimal sample allocation is sensitive to the measure chosen by the analyst. Cooper and Loomis (1992) evaluate the sensitivity of WTP estimates to the bid range and the size of the bid intervals. They report fluctuations in their WTP measures with respect to both the range of bids and coarseness of bid intervals ranging from a 63% decrease in WTP to a 37% increase in WTP. Cooper and Loomis (1992) find their preferred WTP measure is particularly sensitive to the higher bid levels. With respect to the bid intervals, Cooper and Loomis (1992) find that while the effect on WTP can be large in magnitude, its affect is generally unpredictable.

Only a handful of studies have examined the effects of cost scales on estimated WTPs in CE surveys (see Table 1). With respect to use amenities, Ryan and Wordsworth (2000) assess the sensitivity of estimates of WTP for different attributes of cervical screening, including the cost attribute. The two split samples vary both the cost scale and the scales of two of the attributes (time to receive results and chance of dying from cervical cancer), considerably complicating the interpretation of their results. They find that the WTPs for four of the five attributes they measured are different between the low-cost sample and the high-cost sample, but two of these are not in the expected direction. They also find that when "the overall estimated WTP for a hypothetical policy change was considered, WTP estimates were shown not to differ substantially across the two estimates." Carlsson and Martinsson (2008) analyze the WTP for reducing power outages, using a base cost scale of 150, 200, 275, and 375 kroner and a high-cost scale that is 200 kroner higher at each point (an average increase in cost of 80%), and report an average WTP increase of 105%. Mørkbak et al. (2010) analyze the effects of increasing just the maximum price (from 65 kroner to 80 or 120 kroner) in a cost scale on the WTP for minced pork, and find increases in WTPs of 21.7–68.31%. They observe that at the higher maximum prices fewer respondents say yes.

With respect to non-use amenities, in a study of WTP to protect nature areas from motorized roadway development, Ladenburg and Olsen (2008) find that an increase in the cost vector from 100 and 200 kroner to 400 and 1,100 kroner in an instructional pre-test question has a significant effect on their WTP measure. However, their finding of a significant effect is limited to females only. In a split sample survey valuing river health improvements in which the cost attribute varies by a factor of 3 (£0.67, 1.67, 3.67, 5, 8 vs 2, 5, 11, 15, 24), Hanley et al. (2005) estimate mean WTP to be 73–113% greater for all attributes in the high-cost survey version. However, the WTP differences are not statistically significant. The authors report that a larger fraction of respondents (25% vs 17%) reject any option but the status quo

Table 1 Summary of cost scale studies

Author	Year	Amenity	Fielded Surveys	Completed Surveys	Cost Scales (Low) (High)	Use/ Non-use	Effects
Ryan and Wordsworth	2000	Cervical screening	2,000	641	(2, 8, 29, 35) (7, 29, 40, 60)	Use	Mixed results
Carlsson and Martinsson	2008	Power outages	2,000	791	(150, 200, 275, 375) (350, 400, 475, 575)	Use	Avg. increase in WTP of 105% for 82% increase in cost
Mørkbak et al.	2010	Minced pork		3,345	Vary max price (65 | 80 | 120)	Use	Cost scales the same except for maximum price; avg. increase in WTPs 22–68%; fewer say yes at maximum price
Hanley et al.	2005	Water quality		330	(0.67, 1.67, 3.67, 5, 8) (2, 5, 11, 15, 24)	Non-use	WTPs lower in low scale, but not significantly different In high-cost sample 25% reject any option but the status quo; this drops to 17% in low-cost sample, with the difference significant at $p = 0.05$
Kragt	2013	Catchment area	772	523	(30, 60, 200, 400) (50, 100, 300, 600)	Non-use	WTPs higher in high cost, but not significantly different No significant differences in proportions of choices at any of the cost scales

in the high cost scale vs the low cost scale. In a study analyzing the valuation of improved river catchment areas, Kragt (2013) finds that WTPs using a high-cost scale of A$50, 100, 300, and 600 are higher than WTPs estimated using a low-cost scale of A$30, 60, 200, and 400, but that the differences in WTPs are not significant. She reports that there are no significant differences in the proportions of choices at any of the cost levels.

One would expect that cost scales would have a larger anchoring effect for non-use amenities than for use amenities. In the case of use amenities, where respondents have more experience valuing the amenities, respondents may enter the choice situation with more information and thus be less reliant on the survey to supply missing information. However, two of the three studies that examined the effects of cost scales in CEs for use amenities found substantial and significant cost scale effects, while the two studies of the effects of cost scales in CEs for non-use amenities found substantial directional effects that had the right sign but were not significant. The study of the effects of cost scale in an instructional choice set found a significant effect, but only for females.

The study reported here contributes to the literature on cost scale effect in CEs by examining the effects of cost scale on an important non-use amenity (threatened and endangered species status) using a data set larger than any of the prior CE studies of the effects of cost scale on CE estimates of WTPs for non-use amenities. We extend the literature by examining the pattern of choices for the status quo, the do-something options, and the choices made at the highest costs seen by respondents and find evidence that supports the conclusion that respondents are making choices based on relative, not absolute costs. We also extend the literature by analyzing responses by choice set and show that absolute costs seem to have some effect on choices in the first choice set, while choices in the second and third choice sets seem to be based on relative costs.

NOAA SURVEY DATA

The data used for this study come 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").[5] The survey objective was to value potential improvements in the Endangered Species Act ("ESA") status of eight threatened and endangered ("T&E") marine species under the stewardship of NOAA – the

[5] The data and related documents were obtained through a Freedom of Information Act request.

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.[6] Each version of the survey presented only three of these eight species.

Before answering choice questions, respondents are provided information about the ESA and the three species in their version, as well as information on what is currently being done to protect the species and what additional protection actions could be undertaken. They are then asked to select the most preferred option in each of a sequence of three choice questions. In each choice set respondents are queried on the same three (of the eight) T&E species; both species improvements and costs are varied across options and across questions. Figure 1 presents a sample choice screen. Within each question, respondents compare the status quo with two alternative options, each of which offers additional protection actions for at least one of the three T&E species, and select the option that they would most prefer. 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.

Knowledge Networks conducted the survey with a random sample of their Internet panel of US households. 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, cost scale, and "cheap talk" treatment. Each version is further divided into 16 sub-versions with different levels of ESA status and costs.[7] Two version pairs (four of the 44 survey designs) are identical except that the cost scales differ by a factor of 4.[8] In our analyses

[6] Wallmo and Lew (2012) use a substantial portion of these survey data as the basis for their analysis, and their paper describes the portion of the data set upon which they relied.

[7] According to Wallmo and Lew (2012, p. 833), "Attribute levels were determined with an experimental design that accounted for main effects and maximized a D-efficiency criterion (i.e., a measure of the goodness of a design relative to an optimal orthogonal design that may be impossible to attain) (Louviere et al. 2000)."

[8] The following "cheap talk" script is used in all but three of the 44 survey versions, and in the four versions used in this study: "For hypothetical questions like these, studies have shown that many people say they are willing to pay more for protecting threatened and endangered species than they actually would pay out of their pockets. We believe this happens because people do not really consider how big an impact an extra cost actually has to their family's budget when answering these types of questions. It is easy to be generous when you do not really need to open your wallet. To avoid this, as you consider each question, please imagine your household actually paying the cost of the choice you select out of your household's budget."

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 be 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 Sample choice question screen

we follow the common practice of excluding so-called "protesters."[9] In addition to the Wallmo and Lew exclusions, we also eliminate data for respondents who failed to respond to all three choice problems.

[9] We use the same exclusion criteria used by Wallmo and Lew (2012). Protesters are (a) respondents who indicated that they were not confident at all about their answers, or (b) who chose status quo Option A in all three choice questions, and indicated any of the following as the reason for choosing Option A in the first choice questions: "Protecting threatened and endangered species places too many restrictions on industries or private landowners"; "I did not feel it is my responsibility to pay for protecting these species"; "I don't trust the government to run the program"; "I should not have to pay more taxes for any reason"; "I need more information to make a choice"; "I am too unsure about how I feel about threatened and

The Impact of a Fourfold Increase in Costs on Willingness to Pay

The four survey versions (two version pairs) testing cost scale are identical except that the dollar amounts in one pair (high-cost versions, with costs ranging from $20 to $200) are exactly four times the dollar amounts in the other pair (low-cost versions, with costs ranging from $5 to $50). Three of the species are endangered with two possible steps of improvements. Two are threatened species with one possible step of improvement. Hence, the two split-sample experiments evaluate a total of eight possible status improvements. A total of 916 completed questionnaires (completion rate of 74.7%) were obtained for the four split-sample versions, resulting in 2,572 completed choice responses.[10]

If respondent utility functions are rational and well behaved, the RUM estimation methodology should yield the same expected WTPs regardless of cost scale, as long as the cost scale includes costs that result in some respondents making trade-offs between the status quo option (A) and the do-something options (B and C). A RUM model assumes that each respondent chooses the option that maximizes utility, and that the measured utility from each option is a function only of the attributes or cost of the good in each choice option and not a function of the attributes or costs of other options seen by the respondent (either in the current question or in earlier questions). An implication of these assumptions is that respondents' choices in the higher-cost scale should result in more status quo choices and fewer do-something choices at each point in the two cost scales. For example, if the low-cost scale has costs ranging from $5 to $50 and the high-cost scale has costs ranging from $20 to $200, the percentage of status quo choices at $20 in the high-cost scale should be higher than at $5 in the low-cost scale. As shown in Table 2, the proportion of choice sets in which an improvement option is chosen at corresponding costs does not differ significantly between respondents facing different cost scales in any of the data sets (the entire sample, the sample after excluding non-responses, and the sample after excluding both non-responses and protestors). In fact, contrary to expectation, the incidence of choosing an

endangered species"; "I do not think the programs will be effective"; and "More research to understand one or more species needs to be done before I would pay for additional protection options."

[10] The Group A version includes the North Atlantic right whale, the Upper Willamette River Chinook salmon, and the smalltooth sawfish (survey versions 41 and 43 in the original NOAA survey, respectively; each version includes all three species, with version 41 having the low-cost scale and version 43 having the high-cost scale). The Group B version includes the North Pacific right whale, the loggerhead sea turtle, and the smalltooth sawfish (survey versions 42 and 44, respectively; each version uses all three species, with version 42 using the low-cost scale and version 44 the high-cost scale).

Table 2 Frequency of choosing an improvement option by survey cost scale and average cost

Average Cost in Low-cost/ High-cost versions	# Choice Sets		% B/C Chosen		(%) Difference in Proportion (Low–High)	*P*-value for Test of Difference in Proportion
	Low-cost Versions	High-cost Versions	Low-cost Versions	High-cost Versions		
No data exclusions						
<$15/<$60	232	227	61.2	58.1	3.1	0.504
$15–17.5/$60–70	386	382	58.5	59.2	−0.6	0.863
$17.5–20/$70–80	269	269	63.2	57.2	5.9	0.159
$20–27.5/$80–110	309	306	53.1	65.0	−12.0	0.003
$27.5–30/$110–120	348	343	56.3	60.3	−4.0	0.283
≥$30/≥$120	307	303	58.6	58.4	0.2	0.957
Total	1,851	1,830	58.2	59.8	−1.6	0.324
Exclude non-responses[a]						
<$15/<$60	214	208	66.4	63.5	2.9	0.533
$15–17.5/$60–70	354	350	63.8	64.6	−0.7	0.840
$17.5–20/$70–80	245	234	69.4	65.8	3.6	0.403
$20–27.5/$80–110	286	281	57.3	70.8	−13.5	0.001
$27.5–30/$110–120	315	322	62.2	64.3	−2.1	0.589
≥$30/≥$120	272	282	66.2	62.8	3.4	0.402
Total	1,686	1,677	63.9	65.3	−1.4	0.411
Exclude non-responses & protesters[b]						
<$15/<$60	171	152	81.9	86.2	−4.3	0.292
$15–17.5/$60–70	270	272	83.0	81.6	1.3	0.682
$17.5–20/$70–80	199	180	83.4	83.9	−0.5	0.901
$20–27.5/$80–110	208	212	77.9	90.1	−12.2	0.001
$27.5–30/$110–120	241	241	80.1	84.6	−4.6	0.189
≥$30/≥$120	212	214	84.0	80.8	3.1	0.398
Total	1,301	1,271	81.7	84.3	−2.6	0.075

Notes:

a. "Non-responses" are choice sets for which the respondent did not select any of the three options.

b. "Protesters" are defined as respondents who always chose the status quo (Option A), and indicated their protest in the debriefing questions. See footnote 9.

improvement option is slightly (but not statistically significantly) higher among respondents receiving the high-cost versions. The difference is statistically significant only for one of the six subgroups, and the direction is unexpected, with the frequency of choosing an improvement option significantly higher for respondents receiving high-cost surveys. These results are similar to those reported by Kragt (2013), but are at variance with those reported by Hanley et al. (2005) (status quo rejection higher in high-cost scale).

Table 3 presents a similar analysis by choice set. We find that in the first choice set a higher percentage of respondents in the high-cost scale choose the status quo and a lower percentage choose a do-something option, but that these relationships are reversed in the second and third choice sets. The differences in the choice percentages are small across all choice sets. These results are consistent with an interpretation that some respondents are responsive to absolute costs in the first choice set but that as a result of anchoring, more respondents are making choices in the second and third choice sets on the basis of relative costs. Using the same data, a related paper (Prelec et al., forthcoming) develops and tests formal utility models to explain the results reported here.

We also review evidence that respondents are constrained by the smaller scale by looking at the proportion of respondents that choose the improvement option with the highest cost seen. If respondents are considering only absolute costs, a higher fraction of respondents in the low-cost scale

Table 3 Status quo and corner choices by cost scale

	Status Quo (%)		
	Low scale	High scale	Difference
All choice sets	18.3	15.7	2.6
Choice set 1	12.9	20.0	−7.1
Choice set 2	19.4	11.8	7.6
Choice set 3	22.2	15.5	6.7
	At Corner (%)		
	Low scale	High scale	Difference
All choice sets	44.2	44.9	−0.7
Choice set 1	38.2	28.3	9.9
Choice set 2	48.1	50.8	−2.7
Choice set 3	45.9	54.6	−8.7

Note: Protesters are excluded as described above. "At Corner" describes respondents who choose the improvement option with the highest cost seen.

versions should choose the maximum cost. As shown in Table 3, this relationship is found only in the first choice set. In the second and third choice sets, fewer respondents appear to be "at the corner" in the low-cost scale. Mørkbak et al. (2010) test the effects of increasing the highest cost in the cost scale; they find fewer respondents say yes at the maximum price in the cost scales with higher maximum prices, but they do not examine this effect across choice sets. If the responses to all three choice sets are pooled, there is virtually no difference between the two cost scales. These results indicate that the choices of most respondents depend on the cost of the option relative to other costs seen in the survey, rather than on the absolute cost of the option itself.

Table 4 presents estimates of the conditional logit models for the low-cost and high-cost data samples.[11] WTPs based on these models are summarized in Table 5, which shows that mean WTP estimates from the high-cost versions vary from 1.60 to 5.22 times higher than the WTP estimates from the low-cost versions, and average 2.97 times higher. The differences in WTPs between the two versions are statistically significant for seven of the eight species' improvements.

Our hypothesis, explored in detail in Prelec et al. (forthcoming), is that our results with respect to cost scale effects are driven by respondents making choices based on relative, not absolute, costs. It is possible that our results could be explained by preference learning, in which respondents "discover" their preferences through completing multiple choices. If the impact of earlier costs on later choices disappeared over the course of the survey, that finding would constitute evidence in favor of the learning hypothesis. We do not see this pattern in our data. However, the three choices presented in the NOAA design may not be a sufficiently long series to give the learning hypothesis a fair test (Dekker et al., 2014).

CONCLUSION

It is implicitly assumed in choice experiments that the choice of cost scale has no effect on the utility function that is revealed by the survey. We find that this is not true in the case of the NOAA endangered and threatened species survey: a fourfold increase in the cost scale leads to about a

[11] The models presented here are estimated using conditional logit. This modeling choice derives from the small samples utilized in the cost scale designs. Using all survey versions, models estimated with mixed logit estimation methods are very similar to models estimated with the conditional logit methodology.

Table 4 Estimation results for conditional logit model by cost scale

	Low-cost Scale ($5–50)			High-cost Scale ($20–200)		
	Estimate	Std. Err.	Z-value	Estimate	Std. Err.	Z-value
One-step improvement from endangered to threatened						
Sawfish = T	0.3453	0.0865	3.99	0.4734	0.0906	5.22
N. Atl. right whale = T	0.3166	0.1165	2.72	0.2831	0.1199	2.36
N. Pac. right whale = T	0.2728	0.1132	2.41	0.7963	0.1224	6.51
One-step improvement from threatened to recovered						
UWR Chinook salmon = R	0.6746	0.1020	6.62	0.9242	0.1074	8.61
Loggerhead = R	0.6452	0.0999	6.46	0.8588	0.1075	7.99
Two-step improvement from endangered to recovered						
Sawfish = R	0.6523	0.1030	6.33	1.0851	0.1086	10.00
N. Atl. right whale = R	0.8544	0.1385	6.17	1.2049	0.1414	8.52
N. Pac. right whale = R	0.6448	0.1247	5.17	1.5103	0.1376	10.98
Cost	−0.0160	0.0034	−4.69	−0.0090	0.0009	−9.58
Log likelihood	−1,337			−1,225		
Observations	1,301			1,271		

Table 5 Effect of quadrupling costs on mean willingness-to-pay estimates derived from the conditional logit model

Species Improvements	Low-cost Scale			High-cost Scale			Ratio
		95% C.I.			95% C.I.		
	Estimate	Lower	Upper	Estimate	Lower	Upper	
N. Pac. right whale (E) = T	$17.01	$4.96	$35.04	$88.82	$62.91	$121.54	5.22
N. Pac. right whale (E) = R	$40.20	$25.01	$64.64	$168.46	$137.52	$209.56	4.19
Sawfish (E) = R	$40.67	$29.25	$62.23	$121.04	$101.38	$147.68	2.98
N. Atl. right whale (E) = R	$53.27	$35.80	$83.75	$134.40	$106.89	$169.56	2.52
Sawfish (E) = T	$21.53	$12.20	$35.41	$52.81	$35.35	$73.89	2.45
UWR Chinook salmon (T) = R	$42.06	$28.49	$67.61	$103.09	$81.32	$129.49	2.45
Loggerhead (T) = R	$40.23	$27.87	$64.98	$95.80	$73.96	$124.30	2.38
N. Atl. right whale (E) = T	$19.74	$6.07	$38.05	$31.58	$5.23	$57.21	1.60

Note: The initial ESA status for each improvement is indicated in parentheses: E denotes endangered, T denotes threatened. The distribution of WTP is simulated using the Krinsky-Robb method. The difference in WTP distributions, evaluated through the method of convolutions, is significant at the 5% level for all improvements, except N. Atl. right whale improving to the threatened ESA status.

threefold increase in the average estimated WTPs of the species surveyed in the study. Our analysis of the data also reveals that the high-cost scale results in fewer, not more, choices of the status quo option and in about the same number of responses in which the highest cost offered is chosen. It appears from our results that after the first choice set, respondents base their choices on relative, not absolute costs.

The cost scale lift we have found is for certain non-use environmental goods (marine species status). While the cost scale effect we document here may be present in many stated preference experiments, such effects are likely to be more pronounced in stated preference studies involving non-use environmental goods because respondents have no experience valuing such goods and may be more receptive to contextual information (such as cost prompts) to assist in deriving values.

The type of cost scale effect we observe may also occur in consumer goods, as occurs in the minced pork study we review (Mørkbak et al., 2010). With consumer products in stable markets, the selection task by respondents is more straightforward than for non-use environmental goods, in that the prices displayed in the study would not generally be dramatically different from the prices that the consumer has experienced in the marketplace. If these prices determine what will be regarded as expensive or cheap relative to other products in the market, then that fact is precisely what a preference measurement method should uncover. When no market exists, as is the case with public goods, the realism criterion is replaced by fidelity to a thought experiment in which public goods are secured through an imaginary private transaction. Whether the individualistic model of consumer preferences does justice to the issues raised by this thought experiment is a separate question – one arguably more critical for the CV method of public good valuation than the contextual anomalies discussed here.

Our results raise two important issues with respect to both use and non-use studies and to our understanding of actual choice behavior. First, since the effect of cost scale is found in both use and non-use studies, it is important that the influence of cost scale be measured. Even in the case of use amenities, if the respondent interprets the prices presented as suggesting options as cheap or expensive relative to the alternatives, this effect should be separated from other aspects of preferences. Second, we note that the effects measured in this chapter and elsewhere may not be a result of the measurement technique, but may reflect the realities of the psychology of choice. On both fronts, the evidence suggests that survey design is an important determinant of measured values, and therefore survey design must be considered in analyzing and interpreting survey results. Further research into the effect of survey design on measurement is warranted.

REFERENCES

Adamowicz, W., J. Louviere, and M. Williams (1994), "Combining revealed and stated preference methods for valuing environmental amenities," *Journal of Environmental Economics and Management*, **26**(3), 271–92.

Adamowicz, W., P. Boxall, M. Williams, and J. Louviere (1998), "Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation," *American Journal of Agricultural Economics*, **80**(1), 64–75.

Ariely, D., G. Lowenstein, and D. Prelec (2003), "Coherent arbitrariness: Stable demand curves without stable preferences," *The Quarterly Journal of Economics*, **18**(1), 73–105.

Bateman, I.J., R.T. Carson, B. Day, D. Dupont, J.J. Louviere, S. Morimoto, R. Scarpa, and P. Wang (2008), "Choice set awareness and ordering effects in discrete choice experiments," *CSERGE Working Paper: EDM*, No. 08-01.

Boxall, P.C., W.L. Adamowicz, J. Swait, M. Williams, and J. Louviere (1996), "A comparison of stated preference methods for environmental valuation," *Ecological Economics*, **18**(3), 243–53.

Cameron, T.A. and D.D. Huppert (1989), "OLS versus ML estimation of non-market resource values with payment card interval data," *Journal of Environmental Economics and Management*, **17**(3), 230–46.

Carlsson, F. and P. Martinsson (2008), "How much is too much? An investigation of the effect of the number of choice sets, context dependence and the choice of bid vectors in choice experiments," *Environmental and Resource Economics*, **40**(2), 165–76.

Cooper, J. and J. Loomis (1992), "Sensitivity of willingness to pay to bid design in dichotomous choice contingent valuation models," *Land Economics*, **68**(2), 211–22.

Dekker, T., P. Koster, and R. Brouwer (2014), "Changing with the tide: Semi-parametric estimation of preference dynamics," *Land Economics*, **90**(4), 717–45.

Duffield, J.W. and D.A. Patterson (1991), "Inference and optimal design for a welfare measure in dichotomous choice contingent valuation," *Land Economics*, **67**(2), 225–39.

Flachaire, E., G. Hollard, and S. Luchini (2007), "Heterogeneous anchoring in dichotomous choice valuation framework," *Louvain Economic Review*, **73**(4), 369–85.

Frykblom, P. and J.F. Shogren (2000), "An experimental testing of anchoring effects in discrete choice questions," *Environmental and Resource Economics*, **16**(3), 329–41.

Green, D., K.E. Jacowitz, D. Kahneman, and D. McFadden (1998), "Referendum contingent valuation, anchoring, and willingness to pay for public goods," *Resource and Energy Economics*, **20**(2), 85–116.

Hanley, N., W. Adamowicz, and R.E. Wright (2005), "Price vector effects in choice experiments: An empirical test," *Resource and Energy Economics*, **27**(3), 227–34.

Hanley, N., R.E. Wright, and V. Adamowicz (1998), "Using choice experiments to value the environment," *Environmental and Resource Economics*, **11**(3–4), 413–28.

Hensher, D.A., J.M. Rose, and W.H. Greene (2015), *Applied Choice Modeling*, Cambridge, UK: Cambridge University Press.

Herriges, J.A. and J.F. Shogren (1996), "Starting point bias in dichotomous choice valuation with follow-up questioning," *Journal of Environmental Economics and Management*, **30**(1), 121–31.
Kragt, M.E. (2013), "The effects of changing cost vectors on choices and scale heterogeneity," *Environmental and Resource Economics*, **54**(2), 201–21.
Ladenburg, J. and S.B. Olsen (2008), "Gender-specific starting point bias in choice experiments: Evidence from an empirical study," *Journal of Environmental Economics and Management*, **56**(3), 275–85.
Louviere, J.J., D.A. Hensher, and J.D. Swait (2000), *Stated Choice Methods: Analysis and Applications*, Cambridge, UK: Cambridge University Press.
Mørkbak, M.R., T. Christensen, and D. Gyrd-Hansen (2010), "Choke price bias in choice experiments," *Environmental and Resource Economics*, **45**(4), 537–51.
Prelec, D., J. Burrows, and P. Dixon (forthcoming), "Context sensitivity in stated preference experiments," available from the authors on request.
Ryan, M. and S. Wordsworth (2000), "Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments," *Scottish Journal of Political Economy*, **47**(5), 504–24.
Silverman, J. and M. Klock (1989), "The behavior of respondents in contingent valuation: Evidence on starting bids," *Journal of Behavioral Economics*, **18**(1), 51–60.
Train, K. (2009), *Discrete Choice Methods with Simulation*, Cambridge, UK: Cambridge University Press.
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.