Journal of Agricultural and Resource Economics 44(2):439–455 ISSN 1068-5502

The Effect of Forced Choice with Constant Choice Experiment Complexity

Jerrod M. Penn, Wuyang Hu, and Linda J. Cox

In a choice experiment, when respondents are not given the opportunity to choose none of the options offered in a choice set, the choices can be considered forced. In this study of visits to Hawaiian beaches, we adopt a dual-response choice experiment that allows a comparison between forced and unforced choices while avoiding the possible confounding effect of choice set complexity found in previous research. The results suggest that individual willingness to pay is different in forced and unforced choice sets. Joint tests for parameter equality provide evidence to support the use of unforced choice designs.

Key words: beach, dual response, Hawaii, opt-out

Introduction

Continuous examination and refinement of choice experiments (CEs) have made them a staple tool for nonmarket valuation. Researchers make a number of decisions about how to implement a CE, each of which can potentially affect welfare estimates. One choice is whether to make respondents' choices forced or unforced by including what is often called an "opt-out," "status quo," or "choose none" alternative, referred to as "opt-out" hereafter. Choices are considered forced when the opt-out alternative is not included in the CE and respondents must select their most preferred among costly alternatives, with no opportunity to indicate that the truly preferred alternative may be the opt-out option. Welfare measurements without an opt-out alternative are inconsistent with demand theory (Hanley, Mourato, and Wright, 2001) and may produce biased results.

Considerable work has focused on the format of opt-out options (Barreiro-Hurle et al., 2018; Kontoleon and Yabe, 2003) and their welfare implications on respondents (Pedersen et al., 2011, 2012). However, given that some CE studies do not include an opt-out option in their design (e.g., Hasund, Kataria, and Lagerkvist, 2011), surprisingly little research has empirically examined the effect of including or excluding an opt-out. Using private goods as the target, both Dhar and Simonson (2003) and Kallas and Gil (2012) find significant differences in choice behavior when choices were given as forced versus unforced, and Kallas and Gil further demonstrate significant differences in willingness to pay (WTP) estimates based on whether choices were given as forced. Carlsson, Frykblom, and Lagerkvist (2007) and Veldwijk et al. (2014) directly compare forced and unforced choices in the context of a CE. Carlsson, Frykblom, and Lagerkvist find that including an opt-out alternative does not have a significant effect on marginal WTP but does affect unobserved heterogeneity (i.e., the number of attribute standard deviations that are estimated to be significant).

Jerrod M. Penn is an assistant professor in the Department of Agricultural Economics and Agribusiness at Louisiana State University and the LSU AgCenter. Wuyang Hu (corresponding author) is an invited courtesy professor in the College of Economics and Management, Huazhong Agricultural University, and a professor in the Department of Agricultural, Environmental, and Development Economics at The Ohio State University. Linda J. Cox is a professor in the Department of Natural Resources and Environmental Management at the University of Hawaii at Mānoa.
The authors acknowledge Lara Kozloff for her contribution collecting the data. The authors are grateful for the support the first two authors received while they both were in the Department of Agricultural Economics at the University of Kentucky during the process of completing this article.

Review coordinated by Jeffrey Peterson.

Conversely, Veldwijk et al. (2014) find significant variations in mean WTP between forced and unforced choices. Because both studies used a design in which respondents were either given an opt-out alternative in each presented choice set or forced choice set without an opt-out alterative, it creates a limitation. By introducing an additional alternative, even when this additional alternative is the opt-out option, choice complexity increases, thus confounding the effect of including/excluding the opt-out option with the choice set complexity effect. This study re-evaluates the effect of the opt-out alternative while holding the cognitive complexity of the CE constant using an application designed to investigate recreational beach quality.

Literature Review

Opt-out alternatives are recommended in CE (Hensher, 2010), specifically since opt-outs reflect a realistic choice that respondents can make (Carson et al., 1994). Kallas and Gil (2012) provide a comprehensive overview of the theory and implications of including an opt-out alternative and the form in which the opt-out alternative should be introduced. Boxall, Adamowicz, and Moon (2009) review the potential biases from not including an opt-out alternative.

However, recent studies without opt-out alternatives exist (Rigby, Alcon, and Burton, 2010; Hasund, Kataria, and Lagerkvist, 2011; Kallas, Escobar, and Gil, 2012; Milte et al., 2018; Patterson, Holdford, and Harpe, 2018). One reason not to include an opt-out alternative may be to maintain the incentive compatibility of the CE design. Incentive compatibility ensures that respondents can only maximize their utility by revealing their true preferences (Vossler, Doyon, and Rondeau, 2012). The standard approach to CE is to create an efficient CE design with respect to some criteria, most commonly D-optimality. Often the design uses two or three alternatives per choice set, and the opt-out alternative is simply added on afterward. Although an opt-out alternative is not an actual alternative with attributes, it still constitutes an additional alternative. Including an opt-out, while appropriate from a utility theoretical and practical perspective, makes the design incentive incompatible. Although we focus on the impact of including or excluding an opt-out in a choice set featuring multiple alternatives, the most commonly implemented form used in CE studies, incentive compatibility remains an important design element and area of research.

Most work has focused on the impact on choice behavior from how the opt-out alternative is described—such as a "choose none" or a "status quo" option—or on the general respondent perception of the opt-out alternative. This work appears across a number of fields such as environment (Scarpa, Willis, and Acutt, 2007; Marsh, Mkwara, and Scarpa, 2011; Ahtiainen, Pouta, and Artell, 2015), food (Kontoleon and Yabe, 2003; Alemu and Olsen, 2018), transportation (Hensher, 2010; Le Pira et al., 2017), and health (Vass, Gray, and Payne, 2016). Some have studied the opt-out alternative's effect on different population segments (de Blaeij, Nunes, and van den Bergh, 2007; Milte et al., 2018; Pedersen et al., 2011). Rose and Hess (2009) study the effect of a respondent-constructed reference alternative, demonstrating its usefulness in capturing more information about preferences yet minimizing the respondent's cognitive burden.

Only a limited number of studies have specifically examined the impact of including or excluding an opt-out alternative. All of this work analyzes the effect by introducing two treatments. In one treatment, choice sets in a CE appear without an opt-out alternative; the other treatment uses the same choice sets with the addition of an opt-out alternative (Dhar and Simonson, 2003; Carlsson, Frykblom, and Lagerkvist, 2007; Kallas and Gil, 2012; Veldwijk et al., 2014). However, by adding an additional opt-out alternative to a choice set, this approach changes choice set complexity. Several dimensions related to choice set complexity exist in the context of CE design. DeShazo and Fermo (2002) and Rolfe and Bennett (2009) define choice set or CE complexity as the number of alternatives or attributes under consideration in a CE. Researchers have also found that the number of choice sets presented affects results from learning or fatigue (Savage and Waldman, 2008; Czajkowski, Giergiczny, and Greene, 2014). Carlsson and Martinsson (2001) demonstrate significantly different preferences comparing choices made in the first and second halves of a series

Table 1. Choice Experiment Beach Attributes and Levels

Attribute	Level
Sand quality	Poor – Brown/gray; composed of 75% foreign materials and 25% sand (reference category)
	Average – Dark tan/light brown; composed of 50% sand and 50% foreign materials
	Good – Light tan; composed of 75% sand and 25% foreign materials
	Excellent – White; all sand
Water quality	Poor –Murky water, brownish in color; probability of illness from wading occurs in 0.025 of healthy adults (reference category)
	Average – Cloudier water affecting visibility, green in color; probability of illness from wading occurs in 0.019 of healthy adults
	Good – Visible particles floating in otherwise clear water, blue in color; probability of illness from wading occurs in 0.012 of healthy adults
	Excellent – Clear, aqua-colored water; probability of illness from wading occurs in 0.005 of healthy adults
Congestion	Poor – Overcrowded and extremely noisy (reference category)
	Average – Beach congestion and noise are present but do not hamper the experience
	Good – Ample open space and little noise
Water entry/swimming safety	Unsafe – Conditions not safe for any recreationists
	Safe – Conditions safe for experienced beach recreationists
	Very Safe – Conditions safe for the majority of beach recreationists (reference category)
Fuel cost (roundtrip)	$0, $5, $10, $15, $20

of choice scenarios, evidence of a learning effect. Hu (2008) discovers similar learning effects and concludes that given a constant number of alternatives per choice set and the total number of choice sets, more complex choice set content (i.e., more variation in attributes across alternatives) affected choices. Boxall, Adamowicz, and Moon (2009) show that the number of choice alternatives and Oehlmann et al. (2017) show that the number of changes in attributes across choice alternatives affect the choice probability of the opt-out alternative and subsequent welfare implications.

In this study, we refer to choice complexity as the number of alternatives a choice set contains. Following this, previous studies dealing with the effect of forced versus unforced choices do not explicitly control for choice set complexity. In this study, we consider the effect of the opt-out alternative while holding choice set complexity constant.

Application and Methods

We examine the opt-out effect within the context of a valuation for recreational beaches among tourists and residents in Oahu, Hawaii. Beach attributes and levels were selected and developed primarily to understand recreational beach choice for swimming and wading based on a number of sources, including correspondence with the Clean Water Branch of the Hawaii Department of Health, the city and county of Honolulu Ocean Safety and Lifeguard Services, previous scholarly work (Mak and Moncur, 1998; Murray, Sohngen, and Pendleton, 2001; Mourato et al., 2003), and focus group feedback.

Table 1 describes the attributes of the beach valuation CE: water quality (four levels), sand quality (four levels), congestion (three levels), safety (three levels), and round-trip fuel costs (five levels).

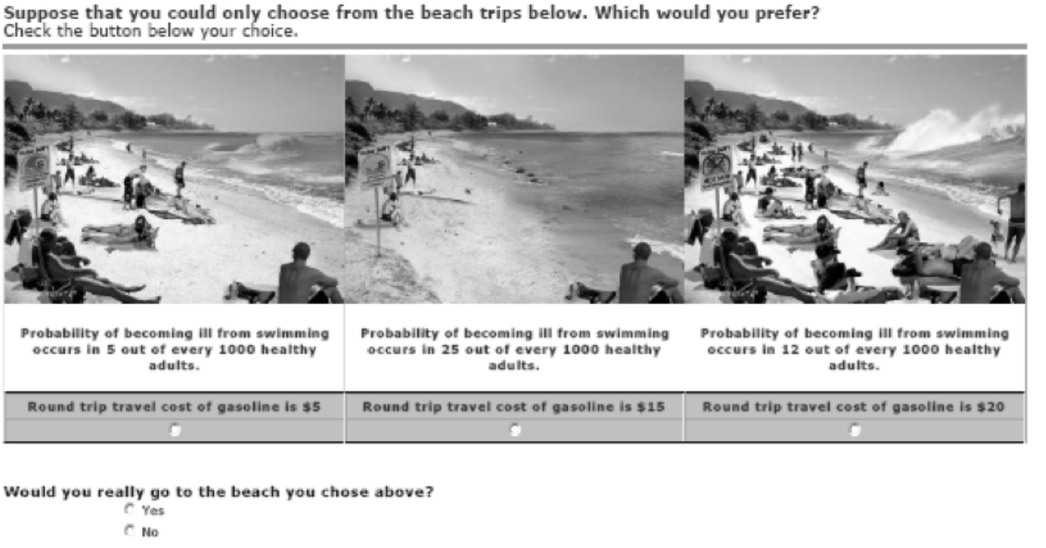

Figure 1. Example Choice Experiment Scenario with Dual Response

This study examines the effects of the opt-out option on choices and WTP estimates while holding choice complexity constant. To do so, the CE used the dual-response method formulated by Brazell et al. (2006). In this study, dual response is akin a follow-up certainty question in that immediately after a respondent provides their preferred beach in a forced setting, they are given the opportunity to opt out of their selected option. The technique has been used previously by Kallas and Gil (2012) and Mentzakis, Ryan, and McNamee (2011) with no attention given to the implications on choice complexity. Figure 1 provides an example CE scenario.

As seen above, the CE communicated attribute levels via computer-augmented pictures, except for the risk of illness and round-trip cost of fuel, which were presented as text. We used a complete enumeration strategy, which considers all possible choice alternatives and chooses a combination of alternatives and choice sets that produce the greatest similarity to an orthogonal design for each respondent, in terms of main effects. We chose this design because (i) orthogonal designs do not exist for all numbers of attributes and levels in our study and (ii) the efficiency loss that can occur with the complete enumeration design is offset in this study by greater flexibility and the elimination of potential attribute level position and order bias (Carlsson, Mørkbak, and Olsen, 2012; Day et al., 2012). Using this technique, D-efficiency exceeds 99%. Each respondent completed ten scenarios (choice sets), each containing three alternatives as well as the dual question in which the respondents can elect to opt out. The survey was fielded in five locations around Oahu by a professional survey firm from late September to mid-October 2009. Participants had to be at least 18 years old and U.S. citizens.[1] The survey was administered on a computer tablet device.

In the CE, every respondent must provide two sequential answers per choice set, answered in two steps, which we refer to as the dual response. In the first stage, each respondent initially makes a choice from among three beach alternatives per scenario. Immediately after the first stage, a dual-response question was presented, asking the respondent, "Would you really go to the beach you chose above?" The respondent could then opt out by saying "no" to the dual-response question, which means that their answer in the first step represents a forced choice. Those who selected "yes" in the dual response are indicating that the beach they selected in the first stage represents an unforced choice. By utilizing the dual-response question format in the CE, the second answer (the "dual") determines whether the choice made in the first stage was a true forced choice. Since all choice scenarios in the first stage have the same number of alternatives and whether the choices

[1] Many of Hawaii's tourists are from abroad, so welfare results are only representative of the United States.

were forced is determined after the choices were made, our approach can examine the forced choice effect while maintaining choice complexity.[2]

In this study, we did not adopt the conventional approach of comparing choice behavior by excluding or including the opt-out alternative. In other words, we did not implement two treatments in which respondents are made to see choice scenarios without the opt-out alternative (i.e., forced choices) or scenarios with the opt-out alternative included for each choice set (i.e., unforced choices). As a result, the goal of this study is not to compare whether our approach is superior to the conventional approach in modeling forced and unforced choices. Rather, our goal is to show the impact of forced choices while holding choice complexity constant.[3]

To compare forced and unforced choices, we partition the sample scenario responses into two sets: (i) scenarios in which the respondent opted out by saying "no" in the dual (i.e., forced choices) and (ii) scenarios in which the respondent did not opt out by saying "yes" in the dual (i.e., unforced choices in the first step). This data segmentation allows us to test for potential effects of including an opt-out using the random utility theory (RUT) that discrete choice experiments rely upon (McFadden, 1973), as in equation (1):

$$U_{ijt} = \alpha C_{ijt} + \boldsymbol{X_{ijt}\beta} + e_{ijt}. \tag{1}$$

RUT asserts that person i's utility for alternative j in scenario t is composed of a systematic component, C and $\boldsymbol{X}$, which represent the payment vehicle and the nonpayment attributes, respectively, and an unobserved component, e. By assuming the difference of the error between two alternatives has a type I largest extreme value distribution, equation (1) can be modeled as a conditional logit model using maximum likelihood estimation. Because of scaling inherent to the unobservable component associated with a logit model, the scale is also reflected in the utility coefficients, α and $\boldsymbol{\beta}$, such that direct comparison of the parameters across models is prohibited (Swait and Louviere, 1993).

In this study, since the attributes are qualitative, we convert each attribute's levels into separate dummies. For example, three dummy variable indicators exist for excellent, good, and average water quality, with poor water quality as the omitted reference category. Each indicator variable is included in the model in equation (2):

$$\begin{aligned} U_{ijt} = & \alpha Cost_{njt} + \beta_1 AverageWater + \beta_2 GoodWater + \beta_3 ExcellentWater + \\ & \beta_4 AverageSand + \beta_5 GoodSand + \beta_6 ExcellentSand + \\ & \beta_7 AverageCongestion + \beta_8 LittleCongestion + \beta_9 Safe + \beta_{10} VerySafe + e_{ijt} \end{aligned} \tag{2}$$

Equation (2) represents the standard practice of modeling CEs, in which the researcher makes distributional assumptions about parameters and the error term and subsequently derives an estimable equation and the WTP estimates by dividing an attribute's coefficient, β, by the negative of α, the coefficient of the payment vehicle, which is the marginal utility of income.

[2] This two-step decision-making process may be less intuitive to respondents relative to a traditional decision in which the opt-out choice is jointly listed alongside the other alternatives. Some respondents may need to learn to become more familiar with the tasks as the CE continues. To reduce the impact learning may have on choice behavior, each respondent saw an example choice set to demonstrate the two-step decision process prior to beginning the CE, which has been shown to mitigate this effect (Day et al., 2012). We analyzed additional models controlling for choice set order effect, but these did not alter the effect of forced choices described in this article.

[3] In addition to our dual question approach, a forced choice situation can also be created by simply not giving respondents the option to opt out. A maintained hypothesis in our dual-response design is that if a respondent prefers an alternative other than the opt-out, they will not change their preference after the dual-response question is asked. We believe this to be a reasonable assumption. However, we cannot maintain the hypothesis that when respondents do not prefer any of the alternatives offered in the first stage, their choices would be the same whether they know there will be a dual question or not. Robustness checks later in the result section offer support for the hypothesis that respondents do not change their preferences.

Table 2. Sample Summary Statistics

Characteristics	Hawaii Average[a]	Resident Sample	Tourist Sample	U.S. Average[a]
Median household income[b]	\$67, 116	\$56,930	\$75,932	\$52,762
Female	49.9%	52.6%	48.3%	50.8%
Associate degree or more	39.1%	59.1%	68.8%	36.3%
Age 18–25	10.99%	30.7%	27.8%	11.23%
Age 55 or older	27.8%	12.4%	16.9%	25.5%
Days on Oahu[b]			8.00	7.37
No. of CE respondents		350	351	

Notes: [a] Based on information from the Hawaii Tourism Authority and the U.S. Census.
[b] Based on midpoint of each available response.

We can test the effect of forced choice versus unforced choice using a likelihood ratio test for equality based on Swait and Louviere (1993)[4] with steps outlined in Louviere, Hensher, and Swait (2000, p. 364). The test statistic is

$$-2(LL_r - \sum LL_u), \tag{3}$$

which is distributed χ^2 with $K(M-1)$ degrees of freedom, where K is the number of restrictions and M is the number of treatments. LL_r is the log-likelihood of the pooled data, and $\sum LL_u$ is the summation of the log-likelihood for each of the individual treatments. The null hypothesis of the test is that parameters are not significantly different between treatments.

In this study's case, LL_r is the logit model of forced and unforced scenarios combined, allowing the scale parameter to be different between these two treatments. $\sum LL_u$ is the sum of the log-likelihood function of the forced and unforced choice models estimated separately and is distributed χ^2 with 12 degrees of freedom, since the number of treatments (M) for forced and unforced equals 2, and the number of parameter restrictions (K) equals 12.[5] If the test statistic exceeds the critical value, then evidence suggests that the treatments sufficiently contribute to explaining the model and reject the null hypothesis of no difference between the forced and unforced datasets.

We can also test for the effect of forced choices using the pooled data of unforced and forced choice sets and adding an interaction for each attribute multiplied by an indicator for whether the response represented a forced choice. If significant, the interactions represent systematic differences in the effect of an attribute in a forced response situation.

To capture consumer taste heterogeneity, we use a mixed logit model with normally distributed nonpayment coefficients for comparison across models. Following Train (2009), the probability of individual i choosing alternative j in the tth scenario can be written as in equation (3), where h is the probability density function of the normally distributed nonpayment coefficients:

$$Prob_{ijt} = \int \frac{\exp\left(\alpha C_{ijt} + \boldsymbol{X_{ijt}\beta}\right)}{\sum_{k=1}^{J} \exp\left(\alpha C_{ijt} + \boldsymbol{X_{ikt}\beta}\right)} h(\boldsymbol{\beta})\, d(\boldsymbol{\beta}). \tag{4}$$

Finally, we also considered scale multinomial logit specifications, which have been previously overlooked in similar queries on the effect of forced choice. This is important because these specifications allow for heteroskedasticity in e_{ijt}, sometimes referred to as scale heterogeneity, in the absence or presence of parameter heterogeneity. Models of this type have been shown to be statistically appropriate in multiple settings (Fiebig et al., 2010; Greene and Hensher, 2010). Results are discussed in the next section.

[4] For applications of the pooled LR test used in a similar manner, see de Magistris, Gracia, and Nayga (2013) and Tonsor and Shupp (2011).

[5] It is important to stress that forced and unforced are not treatments in the sense of a traditional split-sample design. The forced or unforced treatments are determined by how the same individual answers the dual question.

Results

Table 2 reports summary statistics for tourists and residents in the sample, which is generally more well-educated and younger relative to population estimates for both groups. Tourist respondents tend to have larger household incomes compared to the U.S. average. This is not surprising since vacationing in Hawaii is relatively expensive compared to alternative recreational trips within the contiguous United States. Resident respondents tend to have incomes smaller than the Hawaii norms. Lastly, tourists appeared to be representative in terms of the length of time spent on Hawaii. All model analyses henceforth combine the tourist and resident CE experiment data unless specified otherwise.

A total of 701 respondents participated in the CE and completed 6,823 choice sets, with the remaining choice sets left unanswered. Of the completed choice sets, 6,023 (88.3%) chose to keep their selection in the dual question, representing unforced choices. The remaining 800 (11.7%) chose to opt out in the dual, representing forced choices in the first stage of choices. About 63% of the respondents did not exhibit any forced choices and about 5% of the respondents showed forced choices in at least six of the ten choices with which they were presented. The econometric results are presented next, with the mixed logit model results in Tables 3a–3e and corresponding WTP in Table 4.

Prior to examining mixed logit model results, we investigate the appropriateness of pooling/differences in the forced and unforced responses based on the likelihood ratio test from equation (3). The test statistics is equal to $[5,631.273 - (733.598 + 4,860.239\}] \times 2 = 74.654$. The number of restrictions (K) is equal to 12 since the number of parameters in the forced and unforced models must be equal after allowing for different scales in the pooled model. two treatments exist (i.e., $M = 2$). The test statistic exceeds critical value for $\chi^2_{(12\}}$, with a p-value of <0.001. The evidence suggests that a pooled model can be rejected and that separating models and parameters is appropriate.

Given this, we proceed to the mixed logit results. To begin, the results for a standard mixed logit model pooling all data and treating the respondents' answers in the dual question as if they were made simultaneously with the first-stage question (the standard CE approach in which the opt-out appears alongside the other alternatives), appears as Model I in Table 3a. These results match the usual expectations. Given that the omitted reference attribute levels are poor water quality, poor sand quality, very congested beaches, and unsafe swimming conditions, we would expect and observe that all non-opt-out attribute coefficients are positive. Furthermore, we would expect and find that the magnitude of coefficients for higher quality levels within an attribute increase (e.g., excellent sand and water are preferred to good sand and water, and good sand and water are preferred to average sand and water). Thus, the data and basic model behave well under a standard approach.

We also consider possible differences in preferences between tourists and Hawaii residents. For instance, a resident may be more likely to opt out given the lower opportunity cost of missing a day at the beach, while tourists have fewer chances to visit an Oahu beach so they may discount the opt-out alternative more heavily. Model I includes the interaction terms between a dummy variable representing tourists and all beach attribute variables. This leads to only a few significant coefficients; in particular, the interaction with the opt-out alternative is insignificant. As a result, we focus on interpreting the main effects, while maintaining these interactions in the model.

Although misspecified, the pooled model that ignores the answers from the dual opt-out question, representing responses only from the first-stage question can be analyzed, as shown in Model II (Table 3b), which is equivalent to the specification in Carlsson and Martinsson (2001). Each attribute is significant with the anticipated magnitude and direction. However, since the opt-out constant is significant in Model I, this model at least suffers from omitted variable bias. This finding is also consistent with previous literature (Carlsson, Frykblom, and Lagerkvist, 2007). In comparing the coefficient estimates across Model I and Model II, we see that the value of each attribute level has decreased relative to the reference level in Model I. With respect to understanding preferences, this

Table 3a. Mixed Logit Estimation Result: Standard Opt-Out Model (Model I)

	Coefficient	Std. Err. of Coefficient	Std. Dev.	Std. Err. of Std. Dev.
Fuel cost	−0.040***	0.004		
Opt-out	0.643**	0.260	2.963***	0.185
Average sand	0.557***	0.085	0.106	0.195
Good sand	0.770***	0.084	0.011	0.131
Excellent sand	1.052***	0.086	0.359***	0.104
Avgerage water	0.945***	0.099	0.307**	0.143
Good water	1.470***	0.093	0.074	0.167
Excellent water	2.308***	0.117	1.231***	0.071
Average congestion	0.623***	0.071	0.042	0.093
Little congestion	0.901***	0.084	0.835***	0.058
Safe	0.798***	0.078	0.505***	0.079
Very safe	1.109***	0.089	0.939***	0.060
Tourist interaction				
Opt-out	−0.001	0.006		
Fuel cost	−0.220	0.340		
Average sand	0.156	0.121		
Good sand	0.170	0.118		
Excellent sand	0.140	0.120		
Average water	0.287**	0.143		
Good water	0.448***	0.134		
Excellent water	0.540***	0.165		
Average congestion	−0.197**	0.098		
Little congestion	−0.195*	0.115		
Safe	−0.049	0.108		
Very safe	0.036	0.122		
No. of choice sets	6,823			
No of obs.	27,292			
Log likelihood	−6,334.6			

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

indicates that a forced choice model that ignores the opportunity to opt out underestimates the utility gained from providing higher quality levels for each attribute.

We next use the information from the dual opt-out question to separate the pooled model into unforced choices (those who chose to keep their option given the chance to opt out) and forced choices (those who elected not to visit the beach they previously choose), Models II (Table 3b) and III (Table 3c), respectively. The results vary dramatically between the two groups. Considering the unforced choice model, the results are nearly equivalent to Model I results in terms of statistical significance and magnitude of the coefficient estimates, their standard deviations, and WTP per attribute. Conversely, the forced choice model results are different from Model I and unforced results. Average and good sand quality now have negative signs, while good and excellent sand quality are no longer statistically significant. WTP for good and excellent water quality also increased much more relative to the difference between pooled and unforced models. Given the unrealistic signs of some of these parameter estimates, the results suggest that "forcing" responses without offering the opt-out option may generate biased or unrealistic estimates.

Table 3b. Mixed Logit Estimation Result: Pooled (Model II)

	Coefficient	Std. Err. of Coefficient	Std. Dev.	Std. Err. of Std. Dev.
Fuel cost	−0.040***	0.004		
Average sand	0.386***	0.078	0.107	0.178
Good sand	0.585***	0.077	0.112	0.117
Excellent sand	0.832***	0.078	0.291**	0.115
Avgerage water	0.839***	0.090	0.189	0.219
Good water	1.354***	0.085	0.191	0.13
Excellent water	2.282***	0.115	1.308***	0.072
Average congestion	0.533***	0.066	0.002	0.085
Little congestion	0.776***	0.079	0.828***	0.055
Safe	0.708***	0.070	0.394***	0.084
Very safe	0.973***	0.083	0.914***	0.057
Tourist interaction				
Fuel cost	0.003	0.006		
Average sand	0.187*	0.112		
Good sand	0.181*	0.109		
Excellent sand	0.218**	0.111		
Average water	0.343***	0.130		
Good water	0.515***	0.124		
Excellent water	0.479***	0.160		
Average congestion	−0.171*	0.091		
Little congestion	−0.159	0.109		
Safe	−0.068	0.098		
Very safe	0.036	0.115		
No. of choice sets	6,823			
No of obs.	20,469			
Log likelihood	−5,279.9			

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

In Model V (Table 3e), the forced and unforced data were again pooled, but with interactions for each attribute level multiplied by an indicator if it was in a forced choice.[6] The implicit model-defined reference alternative under Model I is the opt-out option, while Model V allows for flexibility in an unobserved, self-identified next-best alternative if respondents do not need the opt-out alternative offered in the second stage to indicate their preferred option. The forced interactions represent potential differences in the effect of an attribute in a forced choice relative to an unforced choice. Ten of the eleven forced choice attribute interactions are significant. Once again, some produce counterintuitive results, such as forced average sand ($0.751 - 0.916 = -0.165$) and forced good sand ($0.552 - 0.998 = -0.446$), demonstrating that these levels are less preferable than poor sand quality, coinciding with the results of Model III. A likelihood ratio test of Model III versus Model V—which tests whether all 11 forced choice interactions are jointly equally 0—generated a χ^2 test statistic of 91.45, corresponding to a p-value of <0.001.

We now proceed to discuss the WTP measures (see Table 4); the confidence interval calculation relied on the delta method. The results reinforce the oddities of the previous parameter estimates. First, we observe that some negative WTP measures are produced under the forced response model (Model IV; Table 3d) and that the confidence intervals associated with these WTP measures are much

[6] A separate model that also included a series of joint dummy interactions of forced and tourist (versus Hawaii residents) choices did not generate any statistical significance.

Table 3c. Mixed Logit Estimation Result: Unforced Responses (Model III)

	Coefficient	Std. Err. of Coefficient	Std. Dev.	Std. Err. of Std. Dev.
Fuel cost	−0.039***	0.005		
Average sand	0.558***	0.089	0.104	0.265
Good sand	0.742***	0.087	0.022	0.135
Excellent sand	1.018***	0.090	0.338***	0.115
Avgerage water	0.927***	0.102	0.316**	0.151
Good water	1.454***	0.096	0.211*	0.110
Excellent water	2.352***	0.124	1.247***	0.075
Average congestion	0.580***	0.074	0.002	0.095
Little congestion	0.881***	0.088	0.876***	0.060
Safe	0.795***	0.079	0.464***	0.087
Very safe	1.083***	0.093	0.977***	0.063
Tourist interaction				
Fuel cost	−0.001	0.006		
Average sand	0.157	0.125		
Good sand	0.166	0.121		
Excellent sand	0.185	0.125		
Average water	0.304**	0.146		
Good water	0.448***	0.138		
Excellent water	0.489***	0.170		
Average congestion	−0.191*	0.101		
Little congestion	−0.220*	0.120		
Safe	−0.076	0.109		
Very safe	0.034	0.127		
No. of choice sets	6,023			
No of obs.	18,069			
Log likelihood	−4,549.4			

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

larger than those based on unforced responses (Model III). Forced versus unforced WTP suggested under Model V also confirms the same conclusion. We find that WTP is insignificantly different for average, good, and excellent sand quality under forced choices in Model V. In comparing these results to the pooled results of Model II, we see that deviations in WTP are modest, with starker changes in WTP for sand quality, which coincides with Model V's forced WTP. Furthermore, comparing the WTP from Model I's standard approach versus Model II, which ignores respondents' opt-out choice, we see that WTP decreases for 10 of the 11 attributes. In 2015, 5.3 million visitors spent 36.4 million visitor days on Oahu. Using the forced WTP values to approximate even a small percentage of visitor days would lead to large biases in the aggregate estimated economic benefit.

To further investigate the potential differences among tourists, we also considered a model of only tourists that includes an interaction for the length of stay on Oahu with each beach attribute level. This specification also leads to large variations in WTP estimates of forced and unforced choice sets, significantly so with respect to good and excellent sand quality. This again would lead to biased predictions of aggregate effects to Hawaii tourists.

Table 3d. Mixed Logit Estimation Result: Forced Responses (Model IV)

	Coefficient	Std. Err. of Coefficient	Std. Dev.	Std. Err. of Std. Dev.
Fuel cost	−0.032***	0.011		
Average sand	−0.442**	0.200	0.007	0.418
Good sand	−0.113	0.190	0.014	0.404
Excellent sand	−0.007	0.198	0.034	0.495
Avgerage water	0.309	0.206	0.008	0.374
Good water	0.865***	0.199	0.077	0.617
Excellent water	2.141***	0.326	2.107***	0.314
Average congestion	0.336**	0.168	0.226	0.441
Little congestion	0.360*	0.198	0.952***	0.220
Safe	0.303*	0.167	0.003	0.318
Very safe	0.321*	0.182	0.621**	0.259
Tourist interaction				
Fuel cost	0.021	0.016		
Average sand	0.241	0.304		
Good sand	0.056	0.302		
Excellent sand	0.248	0.297		
Average water	0.542*	0.326		
Good water	0.784**	0.316		
Excellent water	0.174	0.476		
Average congestion	−0.069	0.257		
Little congestion	0.017	0.293		
Safe	−0.066	0.255		
Very safe	0.310	0.273		
No. of choice sets	800			
No of obs.	2,400			
Log likelihood	−693.3			

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

As a robustness check, we also modeled the data as a scale multinomial logit model.[7] In this case, rather than assume individual heterogeneity for the preference of attribute levels, we allow heteroskedasticity in the error term of equation (1) across respondents. We again find that forced choices are drastically different to unforced choices. In the model of forced choices, three of the attributes are no longer statistically significant. Similar to the mixed logit results, the odd result shows that average sand among forced respondents has a negative sign (−0.06), which indicates that average sand quality is less likely to be chosen compared to poor sand quality, holding all other factors constant.

The scale multinomial logit model may also help identify peculiarities across choice scenarios induced by the design in this analysis. It is possible that the using the dual opt-out design unduly influences respondent choice compared to the commonly used approach of including the opt-out as an additional alternative. For example, suppose a respondent finds all three alternatives in a choice scenario to be relatively unappealing. Because the respondent knows that they will have the option to opt out immediately after the initial choice, compared to a situation in which they are not given the opt-out alternative at all, the respondent may not take the time to select their most preferred of

[7] We attempted generalized multinomial logit models (Fiebig et al., 2010) that allow for scale as well as attribute heterogeneity across respondents but failed to achieve model convergence in the forced choice model. Consequently, we choose to test the data separately as both a mixed logit and a scale multinomial logit.

Table 3e. Mixed Logit Estimation Result: Pooled Responses with Fored Interactions (Model V)

	Coefficient	Std. Err. of Coefficient	Std. Dev.	Std. Err. of Std. Dev.	Coefficient Interacted with Forced Choice Indicator	Std. Err. of Coefficient Interacted with Forced Choice Indicator
Fuel cost	−0.041***	0.004				
Average sand	0.751***	0.083	0.064	0.203	−0.916***	0.155
Good sand	0.552***	0.084	0.108	0.116	−0.998***	0.156
Excellent sand	0.996***	0.084	0.321***	0.109	−0.885***	0.153
Avgerage water	0.919***	0.095	0.170	0.268	−0.500***	0.174
Good water	1.410***	0.090	0.119	0.215	−0.380**	0.165
Excellent water	2.345***	0.120	1.304***	0.072	−0.404**	0.190
Average congestion	0.560***	0.070	0.001	0.084	−0.204	0.130
Little congestion	0.845***	0.083	0.825***	0.055	−0.475***	0.140
Safe	0.790***	0.074	0.391***	0.087	−0.472***	0.132
Very safe	1.079***	0.088	0.934***	0.057	−0.690***	0.142
Tourist interaction						
Fuel cost	0.004	0.006				
Average sand	0.139	0.112				
Good sand	0.133	0.110				
Excellent sand	0.176	0.112				
Average water	0.319**	0.130				
Good water	0.487***	0.124				
Excellent water	0.441***	0.161				
Average congestion	−0.172*	0.092				
Little congestion	−0.178	0.110				
Safe	−0.089	0.098				
Very safe	0.009	0.117				
No. of choice sets	6,823					
No of obs.	20,469					
Log likelihood	−5,234.2					

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

Table 4. Willingness to Pay Estimates

	I Standard Opt-Out		II Pooled		III Unforced Responses		IV Forced Responses		V Pooled Responses with Forced Interactions				
Attribute	WTP	95% CI	WTP	95% CI	WTP	95% CI	WTP	95% CI	WTP	95% CI	WTP	95% CI	*p*-Value
Average sand	13.85***	8.82, 18.89	10.01***	5.59, 14.44	14.28***	8.84, 19.72	−13.70*	−28.72, 1.33	13.49***	8.68, 18.31	−14.64	−29.72, 0 .44	<0.001
Good sand	19.15***	13.67, 24.63	15.18***	10.43, 19.93	19.00***	13.17, 24.82	−3.50	−15.45, 8.45	18.37***	13.17, 23.56	−5.39	−17.73, 6.95	<0.001
Excellent sand	26.18***	19.30, 33.06	21.58***	15.67, 27.49	26.06***	18.69, 33.43	−0.20	−12.22, 11.81	24.35***	17.99, 30.72	−0.07	−11.96, 11.83	<0.001
Average water	23.50***	16.66, 30.33	21.78***	15.46, 28.10	23.73***	16.34, 31.12	9.59	−4.07, 23.24	22.47***	16.04, 28.91	16.87*	0.26, 33.48	0.505
Good water	36.57***	27.81, 45.34	35.12***	26.87, 43.37	37.20***	27.62, 46.79	26.81***	6.85, 46.76	34.47***	26.33, 42.62	41.40*	13.45, 69.36	0.620
Excellent water	57.41***	44.36, 70.46	59.19***	45.97, 72.41	60.21***	45.47, 74.94	66.36***	21.69, 111.03	57.33***	44.48, 70.19	78.02*	28.09, 127.95	0.404
Avg. congestion	15.48***	10.75, 20.22	13.82***	9.46, 18.18	14.84***	9.86, 19.83	10.42*	−1.65, 22.48	13.69***	9.35, 18.03	14.32*	1.22, 27.41	0.924
Little congestion	22.41***	16.16, 28.67	20.14***	14.37, 25.92	22.55***	15.79, 29.3	11.15	−2.51, 24.81	20.65***	14.83, 26.47	14.87*	0.63, 29.12	0.414
Safe conditions	19.85***	14.26, 25.43	18.37***	13.26, 23.48	20.35***	14.34, 26.36	9.41	−2.21, 21.02	19.32***	14.06, 24.57	12.80	−0.02, 25.62	0.316
Very safe conditions	27.59***	20.48, 34.70	25.26***	18.72, 31.79	27.71***	20.03, 35.39	9.96	−2.41, 22.33	26.37***	19.67, 33.07	15.63*	1.14, 30.13	0.142

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

the three alternatives, leading to more noise in choice in the forced choice situation. This is akin to the respondent's perceived credibility to the choice set described by Carson and Groves (2007). If a particular good or bundle of attributes is offered at an implausibly low or high price, the respondent may infer additional information about the product beyond what is presented. Similarly, Czajkowski, Giergiczny, and Greene (2014) generalized this as the choice order effect.

Although we have no benchmark on which to base a conclusion about whether our dual opt-out design may generate choice order effects too large to be acceptable, the possibility of choice order effects is tested in two ways: (i) by including interactions between attribute levels and the order in which the choice scenarios was presented to the respondents and (ii) in the scale multinomial logit model with a variable indicating the order of choice scenarios to explain the scale. None of these specifications generates consistently significant results that suggest evidence of any clear choice order effects.

Conclusions

Holding choice set complexity constant, we investigate the effect of forced choice in a choice experiment of recreational beach sites and find significant differences. Model results suggest that the data from forced and unforced choices are not from an equivalent data-generating process and should not be pooled. Furthermore, some attribute-specific differences in forced versus unforced results exist. Forcing a choice in a scenario of undesirable alternatives produces different welfare estimates and, in our case, some attributes with the theoretically incorrect sign, thus violating construct validity. These differences are also economically significant given the large number of visitors to Oahu (5.34 million visitors and 36.4 million visitor days in 2015). Even if a small fraction of these visitor days is spent at the beach, the \$0 estimate of the value of sand quality under forced choice as compared to the \$26 estimated otherwise leads to large biases when generating the aggregate value of Hawaii beaches.

These results are somewhat different from those of Carlsson, Frykblom, and Lagerkvist (2007), who found no significant differences in attribute-level WTP but did find significantly greater unobservable heterogeneity for their mixed logit models of unforced choice. Given that Carlsson, Frykblom, and Lagerkvist did not allow choice set task complexity to be constant between different treatment groups in their research, discerning what may have led to the insignificant differences between treatment effects is difficult. Furthermore, the context of their research on cattle and broiler processing, a credence attribute, is different to the use values associated with beach recreation choices faced here, which may explain some of the differences in the results. Similarly, the context of forced choice for food may cause a different level of decision conflict relative to recreation.

Our results coincide somewhat with the slightly different models of Kallas and Gil (2012) and Veldwijk et al. (2014), with significantly different WTP for the forced versus unforced model. Our results for significantly different parameters and WTP persist using a scale multinomial logit model. Allowing scale parameter differences in testing the difference between forced and unforced choices is absent in previous works. It also shows our findings are robust across model specifications.

While this work shows the importance of including an opt-out alternative, even after maintaining task complexity, factors that attenuate the decision-making process among respondents may exist. For example, credibility of the computer-augmented choice alternatives presented can seem implausible and may induce protest or seemingly inconsistent preferences (Carson and Groves, 2007). Another potential shortcoming of the dual design is the circumstance of all alternatives being relatively unappealing to a respondent. In this case, the respondent may be more prone to randomly select from among the unappealing alternatives with the knowledge that they will opt out rather than select the most preferred among them. Despite the significant implication on aggregated value discussed previously, the proportion of forced choices is relatively small, such that the number of random choices made in the first stage that are caused by our dual question design might also be small. Our robustness check did not find different choice order effects, specifically behavior between early in the CE when respondents are less familiar with the dual response and later in the CE when their familiarity with the dual-response question has increased. Other applications in which the number of forced choices is much greater, the impact on WTP and the overall parameter estimation result may also be much larger. Finally, comparing our dual-response question approach to the more conventional approach to measure the effect of forced choices—in which the opt-out alternative is simply excluded or included in different choice scenarios—remains an interesting future question.

[First submitted July 2016; accepted for publication September 2018.]

References

Ahtiainen, H., E. Pouta, and J. Artell. "Modelling Asymmetric Preferences for Water Quality in Choice Experiments with Individual-Specific Status Quo Alternatives." *Water Resources and Economics* 12(2015):1–13. doi: 10.1016/j.wre.2015.10.003.

Alemu, M. H., and S. B. Olsen. "Can a Repeated Opt-Out Reminder Mitigate Hypothetical Bias in Discrete Choice Experiments? An Application to Consumer Valuation of Novel Food Products." *European Review of Agricultural Economics* 45(2018):749–782. doi: 10.1093/erae/jby009.

Barreiro-Hurle, J., M. Espinosa-Goded, J. M. Martinez-Paz, and A. Perni. "Choosing Not to Choose: A Meta-Analysis of Status Quo Effects in Environmental Valuations Using Choice Experiments." *Economía Agraria y Recursos Naturales* 18(2018):79–109.

Boxall, P., W. L. Adamowicz, and A. Moon. "Complexity in Choice Experiments: Choice of the Status Quo Alternative and Implications for Welfare Measurement." *Australian Journal of Agricultural and Resource Economics* 53(2009):503–519.

Brazell, J. D., C. G. Diener, E. Karniouchina, W. L. Moore, V. Séverin, and P.-F. Uldry. "The No-Choice Option and Dual Response Choice Designs." *Marketing Letters* 17(2006):255–268. doi: 10.1007/s11002-006-7943-8.

Carlsson, F., P. Frykblom, and C. J. Lagerkvist. "Consumer Willingness to Pay for Farm Animal Welfare: Mobile Abattoirs versus Transportation to Slaughter." *European Review of Agricultural Economics* 34(2007):321–344. doi: 10.1093/erae/jbm025.

Carlsson, F., and P. Martinsson. "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments?" *Journal of Environmental Economics and Management* 41(2001):179–192. doi: 10.1006/jeem.2000.1138.

Carlsson, F., M. R. Mørkbak, and S. B. Olsen. "The First Time Is the Hardest: A Test of Ordering Effects in Choice Experiments." *Journal of Choice Modelling* 5(2012):19–37.

Carson, R. T., and T. Groves. "Incentive and Informational Properties of Preference Questions." *Environmental and Resource Economics* 37(2007):181–210. doi: 10.1007/s10640-007-9124-5.

Carson, R. T., J. J. Louviere, D. A. Anderson, P. Arabie, D. S. Bunch, D. A. Hensher, R. M. Johnson, W. F. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, and J. B. Wiley. "Experimental Analysis of Choice." *Marketing Letters* 5(1994):351–367. doi: 10.1007/BF00999210.

Czajkowski, M., M. Giergiczny, and W. H. Greene. "Learning and Fatigue Effects Revisited: Investigating the Effects of Accounting for Unobservable Preference and Scale Heterogeneity." *Land Economics* 90(2014):324–351. doi: 10.3368/le.90.2.324.

Day, B., I. J. Bateman, R. T. Carson, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa, and P. Wang. "Ordering Effects and Choice Set Awareness in Repeat-Response Stated Preference Studies." *Journal of Environmental Economics and Management* 63(2012):73–91.

de Blaeij, A. T., P. A. L. D. Nunes, and J. C. J. M. van den Bergh. "'No-Choice' Options within a Nested Logit Model: One Model Is Insufficient." *Applied Economics* 39(2007):1245–1252. doi: 10.1080/00036840600852955.

de Magistris, T., A. Gracia, and R. M. Nayga. "On the Use of Honesty Priming Tasks to Mitigate Hypothetical Bias in Choice Experiments." *American Journal of Agricultural Economics* 95(2013):1136–1154. doi: 10.1093/ajae/aat052.

DeShazo, J., and G. Fermo. “Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency.” *Journal of Environmental Economics and Management* 44(2002):123–143. doi: 10.1006/jeem.2001.1199.

Dhar, R., and I. Simonson. “The Effect of Forced Choice on Choice.” *Journal of Marketing Research* 40(2003):146–160. doi: 10.1509/jmkr.40.2.146.19229.

Fiebig, D. G., M. P. Keane, J. Louviere, and N. Wasi. “The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity.” *Marketing Science* 29(2010):393–421.

Greene, W. H., and D. A. Hensher. “Does Scale Heterogeneity across Individuals Matter? An Empirical Assessment of Alternative Logit Models.” *Transportation* 37(2010):413–428.

Hanley, N., S. Mourato, and R. E. Wright. “Choice Modelling Approaches: A Superior Alternative for Environmental Valuatioin?” *Journal of Economic Surveys* 15(2001):435–462.

Hasund, K. P., M. Kataria, and C. J. Lagerkvist. “Valuing Public Goods of the Agricultural Landscape: A Choice Experiment Using Reference Points to Capture Observable Heterogeneity.” *Journal of Environmental Planning and Management* 54(2011):31–53.

Hensher, D. A. “Hypothetical Bias, Choice Experiments and Willingness to Pay.” *Transportation Research Part B: Methodological* 44(2010):735–752. doi: 10.1016/j.trb.2009.12.012.

Hu, W. “My Favourite Is Not on the Shelf: Effects of the Revealed-Preference Bundle on Consumers’ Stated Choice.” *European Review of Agricultural Economics* 35(2008):29–50. doi: 10.1093/erae/jbn002.

Kallas, Z., C. Escobar, and J. M. Gil. “Assessing the Impacts of Alternative ‘Opt-Out’ Formats in Choice Experiment Studies: Consumer Preferences for Genetically Modified Content and Production Information in Food.” *Appetite* 58(2012):285–298. doi: 10.1016/j.appet.2011.09.017.

Kallas, Z., and J. M. Gil. “A Dual Response Choice Experiments (DRCE) Design to Assess Rabbit Meat Preference in Catalonia: A Heteroscedastic Extreme-Value Model.” *British Food Journal* 114(2012):1394–1413. doi: 10.1108/00070701211262984.

Kontoleon, A., and M. Yabe. “Assessing the Impacts of Alternative ‘Opt-Out’ Formats in Choice Experiment Studies: Consumer Preferences for Genetically Modified Content and Production Information in Food.” *Journal of Agricultural Policy and Resources* 5(2003):1–53.

Le Pira, M., E. Marcucci, V. Gatta, G. Inturri, M. Ignaccolo, and A. Pluchino. “Integrating Discrete Choice Models and Agent-Based Models for Ex-Ante Evaluation of Stakeholder Policy Acceptability in Urban Freight Transport.” *Research in Transportation Economics* 64(2017):13–25. doi: 10.1016/j.retrec.2017.08.002.

Louviere, J. J., D. A. Hensher, and J. D. Swait. *Stated Choice Methods: Analysis and Applications*. Cambridge, UK: Cambridge University Press, 2000.

Mak, J., and J. Moncur. “Political Economy of Protecting Unique Recreational Resources: Hanauma Bay, Hawaii.” *Ambio* 27(1998):217–223.

Marsh, D., L. Mkwara, and R. Scarpa. “Do Respondents’ Perceptions of the Status Quo Matter in Non-Market Valuation with Choice Experiments? An Application to New Zealand Freshwater Streams.” *Sustainability* 3(2011):1593–1615. doi: 10.3390/su3091593.

McFadden, D. “Conditional Logit Analysis of Qualitative Choice Behavior.” In P. Zarembka, ed., *Frontiers in Econometrics,* Economic Theory and Mathematical Economics. New York, NY: Academic Press, 1973, 105–142.

Mentzakis, E., M. Ryan, and P. McNamee. “Using Discrete Choice Experiments to Value Informal Care Tasks: Exploring Preference Heterogeneity.” *Health Economics* 20(2011):930–944.

Milte, R., J. Ratcliffe, G. Chen, and M. Crotty. “What Characteristics of Nursing Homes Are Most Valued by Consumers? A Discrete Choice Experiment with Residents and Family Members.” *Value in Health* 21(2018):843–849. doi: 10.1016/j.jval.2017.11.004.

Mourato, S., S. Georgiou, E. Ozdemiroglu, J. Newcombe, and A. Howarth. “Bathing Water Directive Revisions: What are the Benefits to England and Wales? A Stated Preference Study.” CSERGE Working Paper ECM 03-12, University of East Anglia, Centre for Social and Economic Research on the Global Environment, Norwich, UK, 2003.

Murray, C., B. Sohngen, and L. Pendleton. “Valuing Water Quality Advisories and Beach Amenities in the Great Lakes.” *Water Resources Research* 37(2001):2583–2590.

Oehlmann, M., J. Meyerhoff, P. Mariel, and P. Weller. “Uncovering Context-Induced Status Quo Effects in Choice Experiments.” *Journal of Environmental Economics and Management* 81(2017):59–73. doi: 10.1016/j.jeem.2016.09.002.

Patterson, J. A., D. A. Holdford, and S. E. Harpe. “Patient Preferences for Objective Quality Metrics during Community Pharmacy Selection: A Discrete Choice Experiment.” *Research in Social and Administrative Pharmacy* (2018). Status: forthcoming. doi: 10.1016/j.sapharm.2018.08.002.

Pedersen, L. B., T. Kjær, J. Kragstrup, and D. Gyrd-Hansen. “Does the Inclusion of a Cost Attribute in Forced and Unforced Choices Matter?” *Journal of Choice Modelling* 4(2011):88–109. doi: 10.1016/S1755-5345(13)70044-7.

Pedersen, L. B., T. KjÃ̧er, J. Kragstrup, and D. Gyrd-Hansen. “Do General Practitioners Know Patients’ Preferences? An Empirical Study on the Agency Relationship at an Aggregate Level Using a Discrete Choice Experiment.” *Value in Health* 15(2012):514–523. doi: 10.1016/j.jval.2012.01.002.

Rigby, D., F. Alcon, and M. Burton. “Supply Uncertainty and the Economic Value of Irrigation Water.” *European Review of Agricultural Economics* 37(2010):97–117. doi: 10.1093/erae/jbq001.

Rolfe, J., and J. Bennett. “The Impact of Offering Two versus Three Alternatives in Choice Modelling Experiments.” *Ecological Economics* 68(2009):1140–1148. doi: 10.1016/j.ecolecon.2008.08.007.

Rose, J. M., and S. Hess. “Dual-Response Choices in Pivoted Stated Choice Experiments.” *Transportation Research Record: Journal of the Transportation Research Board* 2135(2009):25–33. doi: 10.3141/2135-04.

Savage, S. J., and D. M. Waldman. “Learning and Fatigue during Choice Experiments: A Comparison of Online and Mail Survey Modes.” *Journal of Applied Econometrics* 23(2008):351–371. doi: 10.1002/jae.984.

Scarpa, R., K. G. Willis, and M. Acutt. “Valuing Externalities from Water Supply: Status Quo, Choice Complexity and Individual Random Effects in Panel Kernel Logit Analysis of Choice Experiments.” *Journal of Environmental Planning and Management* 50(2007):449–466. doi: 10.1080/09640560701401861.

Swait, J., and J. Louviere. “The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models.” *Journal of Marketing Research* 30(1993):305–314. doi: 10.2307/3172883.

Tonsor, G. T., and R. Shupp. “Cheap Talk Scripts: Effectiveness in Online Choice Experiments and Private Good, Niche Market Assessments.” *American Journal of Agricultural Economics* 93(2011):1015–1031. doi: 10.1093/ajae/aar036.

Train, K. E. *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press, 2009, 2nd ed.

Vass, C., E. Gray, and K. Payne. “Discrete Choice Experiments of Pharmacy Services: A Systematic Review.” *International Journal of Clinical Pharmacy* 38(2016):620–630. doi: 10.1007/s11096-015-0221-1.

Veldwijk, J., M. S. Lambooij, E. W. de Bekker-Grob, H. A. Smit, and G. A. de Wit. “The Effect of Including an Opt-Out Option in Discrete Choice Experiments.” *PLoS ONE* 9(2014):e111,805. doi: 10.1371/journal.pone.0111805.

Vossler, C. A., M. Doyon, and D. Rondeau. “Truth in Consequentiality: Theory and Field Evidence on Discrete Choice Experiments.” *American Economic Journal: Microeconomics* 4(2012):145–171. doi: 10.1257/mic.4.4.145.