Are Choice Experiment Treatments of Outcome Uncertainty Sufficient? An Application to Climate Risk Reductions

Christos Makriyannis, Robert J. Johnston, and Adam W. Whelchel

Choice experiments addressing outcome uncertainty (OU) typically reframe continuous probability densities for each risky outcome into two discrete categories, each with a single probability of occurrence. The implications of this simplification for welfare estimation are unknown. This article evaluates the convergent validity of willingness-to-pay (WTP) estimates from a more accurate multiple-outcome treatment of OU, compared to the two-outcome approach. Results for a case study of coastal flood adaptation in Connecticut, United States, suggest that higher-resolution OU treatments increase choice complexity but can provide additional information on risk preferences and WTP. This tradeoff highlights challenges facing the valuation of uncertain outcomes.

Key Words: choice scenario, climate change adaptation, discrete choice experiments, generalized multinomial logit model, outcome uncertainty

Discrete choice experiments (DCEs) in environmental economics often include choice attributes that are subject to outcome uncertainty (OU), defined as uncertainty about whether attribute levels indicated in valuation scenarios will occur when policies are implemented. OU can be attributed to multiple factors, including uncertainty in scientific predictions related to environmental phenomena and the efficacy of policy interventions. OU is often associated with policy changes that directly influence the probability of a particular outcome, such as continued survival of a threatened species. However, it may

Christos Makriyannis is a Postdoctoral Associate within the Institute for Data, Systems, and Society, Massachusetts Institute of Technology, 50 Ames St., Building E18-407A, Cambridge, MA 02142, USA. Robert J. Johnston is Director, George Perkins Marsh Institute and Professor, Department of Economics, Clark University, 950 Main St., Worcester, MA 01610, USA. Adam W. Whelchel is Director of Science, The Nature Conservancy Connecticut Chapter, 55 Church St., New Haven, CT 06510, USA. *Correspondence: Christos Makriyannis ▪ Institute for Data, Systems, and Society ▪ Massachusetts Institute of Technology ▪ 50 Ames St., Cambridge, MA 02142 ▪ Phone 617.458.0165 ▪ Email chrismak@mit.edu*

also occur if the effect of a policy change depends on an uncertain future event. The latter body of research often focuses on reducing the risk of loss associated with uncertain future events such as fire, storms, or disease. Although the underlying cause of the uncertainty differs, and some researchers consider the two strands of literature to be distinct,[1] the implications for DCE design are similar—at least one attribute has a level that is uncertain.

Recent guidance for stated preference methods indicates that "when risk or uncertainty is an important aspect of the baseline or change being valued, scenarios should communicate this information in terms that are readily understood by respondents" (Johnston et al. 2017, 329). However, most DCEs include no formal communication of OU (Lundhede et al. 2015). Underlying this treatment is an often-unstated assumption that outcomes are certain (Wielgus et al. 2009, Glenk and Colombo 2013) or that individuals' utilities depend only on attributes' final states (Roberts, Boyer, and Lusk 2008). Concerns that these assumptions may have implications for the validity of DCE results have sparked a small but growing body of DCE studies that account for OU (Roberts, Boyer, and Lusk 2008, Rolfe and Windle 2009, 2010, 2015, Ivanova, Rolfe, and Tucker 2010, Shaw and Baker 2010, Glenk and Colombo 2011, 2013, Phillips 2011, Reynaud, Aubert, and Nguyen 2013, Lundhede et al. 2015). OU may be particularly important in contexts such as climate change, where uncertainty regarding future conditions is central to decision-making (Glenk and Colombo 2011, 2013).

DCEs in environmental and health economics were among the first to provide risk information and include attributes subject to uncertainty.[2] Within these and other DCEs, OU is typically framed as a situation of risk. As such, potential outcomes and probabilities are assumed to be known, and it is possible to describe either a continuous or discrete probability distribution of outcomes. Most DCEs that address OU in this way allow for only two possible outcomes of each risk-related attribute, with each outcome distinguished by an exogenously varying numerical probability embedded within the choice set. We refer to this as a binary treatment of OU. This practice is common both for studies addressing the success of policy interventions and those estimating the value of risk reductions.

For example, within studies framed in terms of policy success (versus failure), Roberts, Boyer, and Lusk (2008) include the probability of an algal bloom (versus no bloom). In a DCE on land-based climate change mitigation, Glenk and Colombo (2011, 2013) include the probability of policy failure (versus success) as a standalone attribute. Ivanova, Rolfe, and Tucker (2010) include

---

[1]   For example, some researchers frame the former case as OU and the latter case as risk valuation. However, as both involve the valuation of outcome(s) subject to some type of risk or uncertainty, we combine both under a broader heading of OU for purposes of this analysis.

[2]   In the health literature, marketing survey-based methods eliciting risk attitudes go as far back as 1987 (e.g., Viscusi, Magat, and Huber 1987), while Macmillan, Hanley, and Buckland (1996) is one of the first environmental stated preference studies to include explicit treatment of OU.

the probability that a policy will contribute to target GHG levels. Rolfe and Windle (2010) include the probability that Great Barrier Reef conservation will achieve specified outcomes.

Similar practices are common within the risk valuation literature, which also may be framed in terms of OU. For example, Reynaud, Aubert, and Nguyen (2013) include the probability of flood damage to agricultural land, damage to the respondent's home, and death in the family. Rolfe and Windle (2009) value the benefits of controlling invasive fire ants in Brisbane, Australia, and embed verbal risk descriptors in choice option labels indicating "high" and "low" certainty of policy success (versus failure). Although many different forms of risk communication and outcomes are reflected within such DCEs, most consider some form of binary OU.[3] For example, Table 1 lists a sample of recent environmental DCEs addressing OU using numerical probabilities, including a summary of the methods used to accommodate uncertainty; all apply some form of binary OU.

Because few actual policies or environmental phenomena are characterized accurately by only two possible outcomes (e.g., bloom/no bloom; success/ failure), characterizing scenarios in this way requires analysts to reframe or discretize actual conditions. This is done under the typically unstated assumption that more accurate treatments of OU (i.e., >2 possible outcomes) would result in overly complex choice scenarios, or that the most relevant policy outcomes are discrete rather than continuous. For example, harmful algal blooms can occur over different areas and durations, each with different probabilities of occurrence; this is reframed to a simpler case of bloom versus no bloom, presumably to reduce the complexity of choice scenarios and/or because the primary, welfare-relevant distinction is assumed to be whether a bloom (of any size or duration) occurs (cf. Roberts, Boyer, and Lusk 2008).

In some cases, this reframing may be intuitive or flow naturally from the policy context. For example, an endangered species context may lend itself to a binary survival versus non-survival framing, and a fire-risk context may naturally lead to a probability of damage to a respondent's home. In other cases, a binary framing may be less intuitive.[4] Regardless, the implications of

---

3   Other examples include Phillips (2011), who investigate beach visitors' preferences for coastal management options where the relative risk of flood damage to public and private property is a primary attribute that takes on the values "low (1 in 20 years)", "medium (1 in 10 years)" and "high (1 in 3 years)." Shaw and Baker (2010) estimate Hurricane Katrina evacuees' willingness to pay (WTP) to avoid hurricane risks and also use the "low," "medium," and "high" terms as independent attributes. Still other studies account for OU by providing framing statements prior to the choice questions (e.g., Macmillan, Hanley, and Buckland 1996, Wielgus et al. 2009). To value bird population changes while accounting for policy-related outcome uncertainty, Lundhede et al. (2015) include verbal risk descriptors such as "very certain" and "rather uncertain" as standalone attributes.

4   Even intuitive cases require assumptions, however. For example, in the fire-risk example, the binary damage versus no damage discretization is a simplification of the actual circumstance in

Table 1. Recent DCE Studies that Use Objective Numerical Probabilities to Address OU Related to Climate Change and other Environmental Outcomes

Authors	Risk-Related Attributes	Numerical Probability Communication Format	Standard Two-Outcome Model
Torres, Faccioli, and Font (2017)	Number of bird species	Numerical probabilities vary exogenously across—but not within—choice sets (i.e., probabilities are held fixed across alternative policies and embedded in choice scenarios as a standalone attribute)	Yes
Torres, Bujosa, and Font (2017)	Temperature increase due to climate change (indicated as a framing variable)	Three numerical probabilities varying exogenously across choice sets not embedded in valuation scenarios (i.e., framing variables)	Yes
Veronesi et al. (2014)	Wastewater overflowing in rivers and lakes	One exogenously varying numerical probability embedded in valuation scenario as a standalone attribute	Yes
Wibbenmeyer et al. (2013)	Wildfire management strategy	One exogenously varying numerical probability embedded in valuation scenario as a standalone attribute	Yes
Reynaud, Aubert, and Nguyen (2013)	Flooding/damage to respondent's home, flooding/damage to paddy field and agricultural land	Two exogenously varying numerical probabilities embedded in valuation scenario as standalone attributes	Yes
Glenk and Colombo (2011, 2013)	GHG emissions reductions	One exogenously varying numerical probability embedded in valuation scenario as a standalone attribute	Yes
Ivanova, Rolfe, and Tucker (2010)	Emissions reductions of technology options (green power, efficient technology, and carbon capture)	Three exogenously varying numerical probability ranges embedded in valuation scenario as three standalone attributes	Yes
Rolfe and Windle (2010)	Water quality improvement, increase in conservation zones, GHG emissions reductions	Three exogenously varying numerical probabilities embedded in valuation scenario as standalone attributes	Yes
Wielgus et al. (2009)	Number, weight, and type of fish caught	One fixed-objective numerical probability stated prior to choice questions	Yes
Roberts, Boyer, and Lusk (2008)	Algal bloom, water levels	Three exogenously varying numerical probabilities associated with attributes and embedded in valuation scenarios	Yes

OU reframing for welfare estimates are almost universally unknown. For example, it is possible that willingness to pay (WTP) for reframed outcomes might differ from WTP for the original uncertain outcomes—that is, OU reframing may not be welfare-neutral. Despite this possibility, the literature provides no systematic guidance regarding the potential effects of OU reframing on DCE results, and what this implies for welfare analysis when policy outcomes or environmental phenomena are uncertain.

Addressing this gap in the literature, this article presents a systematic evaluation of whether and how a more accurate treatment of OU influences DCE results, compared to a traditional two-outcome approach. Methods are illustrated using an application to coastal flood adaptation in Connecticut, United States, with OU related to the number of homes expected to flood during coastal storms of varying intensities and probabilities. We systematically compare models estimated using two otherwise-identical DCEs that incorporate alternative treatments of OU while presenting the same policy outcomes. The first (version one) allows for three different possible outcomes of a risky event—here a coastal storm—with associated damage to exposed homes. Specifically, it presents variations in home flooding that will occur conditional on storm events of two different intensities and likelihoods of occurrence (Category 2 versus Category 3 storms), compared to zero flooding with no storm. The second (version two) maintains the standard two-outcome OU treatment of the same risky event, presenting homes expected to flood during a Category 3 storm event with a given probability, compared to zero flooding with no storm.

Systematic comparison of these otherwise-identical DCEs is used to test the convergent validity[5] of welfare estimates derived through otherwise-identical DCEs that include alternative framing of OU. That is, we evaluate whether or not welfare estimates for the same policy outcomes (e.g., a flood of a specific size that affects a given number of homes) are robust between a simplified two-outcome representation of OU and the higher resolution (and more accurate) multiple-outcome representation. We also compare evidence of DCE complexity, including symptoms of decision heuristics.

Results identify tradeoffs associated with the design of choice scenarios that present higher-resolution OU. We reject convergent validity between the two treatments for multiple implicit prices, suggesting that welfare estimates can be sensitive to the accuracy with which OU is presented. Higher resolution treatments of OU also provide additional, potentially relevant information on respondents' preferences for risk-related attributes. At the same time, we find

---

which a fire may cause varying degrees of damage to the home, from minor damage to total destruction. The endangered survival versus no survival case is a simplification of survival in different regions and/or at different population levels.

[5] Convergent validity holds if different measurement techniques provide statistically indistinguishable estimates of the same theoretical construct, here WTP (Carmines and Zeller 1979, Boyle and Özdemir 2009).

evidence that higher-resolution OU treatments increase symptoms of choice complexity and decision-making heuristics, suggesting that even modest increases in the accuracy of risk communication can present cognitive challenges to respondents. Hence, while standard treatments may be inadequate to fully reflect welfare change when outcomes are uncertain, more accurate treatments may introduce challenges related to the complexity. This conundrum is a specific illustration of a more general concern discussed by Johnston and Swallow (1999) and Mazzotta and Opaluch (1995), among others, in which survey designs intended to increase the information content of stated preference questions lead to complexity that at least partially negates the advantages for welfare estimation.

Treatment of Outcome Uncertainty in Choice Scenarios

A mature literature considers issues related to whether or not and how individuals process risk information when making choices (e.g., Slovic, Fischhoff, and Lichtenstein 1985, Slovic 1987, Fischhoff, Bostrom, and Quadrei 1993, Kahneman 2003, Slovic et al. 2005, Peters et al. 2006, Gilboa, Postlewaite, and Schmeidler 2008, Baker et al. 2009), along with the presentation and impact of this information within DCEs (e.g., Corso, Hammitt, and Graham 2001, Patt and Schrag, 2003, Hanley, Kriström, and Shogren 2009, Shaw and Baker, 2010, Cameron, DeShazo, and Johnson 2011, Akter, Bennett, and Ward 2012). However, the impact of the two-outcome reframing of OU in DCEs has received little explicit attention, particularly with regard to implications for welfare estimation.

The DCE literature demonstrates that communicating OU influences welfare estimates compared to otherwise-identical scenarios that do not communicate OU (Rolfe and Windle 2015). As a result, omitting OU information (i.e., treating outcomes as certain) may lead to WTP estimates that do not provide accurate expressions of ex ante welfare change under uncertainty. Results of these analyses also demonstrate that the stated probabilities, as well as the presentation of probabilistic outcomes, influence welfare estimates (e.g., Roberts, Boyer, and Lusk 2008, Glenk and Colombo 2011, 2013, Rolfe and Windle 2015). That is, they demonstrate that OU—when presented to respondents—is relevant to welfare estimation. These empirical findings are consistent with theory demonstrating that ex ante WTP for risky outcomes (before risk is resolved) is not generally the same as WTP for the expected value of these outcomes—in the general case (unless risk neutrality holds), expected utility is not equivalent to the utility of the expected value (Graham 1981).[6]

---

[6] For a review of these concepts and the related literature, see Freeman, Herriges, and Kling (2014) (chapter 5) and Holland, Sanchirico, and Johnston (2010) (chapter 5).

Hence, when DCE outcomes are subject to substantial OU, accurate welfare elicitation generally requires this OU to be communicated to respondents (Johnston et al. 2017). In the general case, uncertain outcomes are most accurately characterized using a continuous probability density function (PDF). For management, modeling, or communication purposes, however, these continuous distributions are often discretized into a small number of intervals, each with an associated probability. In general, the accuracy of these discrete approximations (in terms of representing second and higher moments of the underlying distribution) is negatively related to the number of intervals used to approximate the continuous PDF (Miller and Rice 1983). For example, discretizing a normal PDF into three intervals more accurately represents variance (the second moment) and kurtosis (the fourth moment) than an otherwise-identical discretizing using two intervals (Miller and Rice 1983), where the latter is the standard approach within the DCE literature.

Although past DCE studies often recognize the difference between binary representations of OU and the underlying continuous PDFs (e.g., Glenk and Colombo 2013), they do not evaluate whether or not and how welfare estimates respond to this simplification. Yet if WTP estimates are conditional on OU (re)framing, then there may be a divergence between WTP elicited by the survey and that which would occur under a more accurate understanding of OU. The divergence between reframed DCE scenarios and actual conditions may constrain the relevance of results for policy analysis (Starmer 2000, Gilboa, Sochen, and Zeevi 2002, Fifer et al. 2011, Glenk and Colombo 2013). At the same time, it is unclear whether binary treatments of OU are required to ensure choice scenarios of manageable complexity; respondents' ability to process more accurate presentations of OU[7] remains largely untested by the DCE literature. These dual concerns apply whether the goal of the study is to evaluate WTP for policies with a certain probability of success, for policies whose impacts depend on an uncertain exogenous event (e.g., flood or fire), or for policies that explicitly reduce the risk of damage due to an uncertain future event.

The Theoretical Model

To evaluate these dual concerns (the accuracy of OU framing and the resulting complexity of choice scenarios), we develop a coupled theoretical and empirical model for a case study of coastal climate change adaptation in the town of Old Saybrook, Connecticut, United States. The analysis focuses on OU related to the protection of homes vulnerable to flooding during storms of different intensities, where these storms have different probabilities of occurrence that

---

7   For example, one can discretize continuous outcome pdfs using > 2 intervals, such that the resulting discrete distributions more accurately represent moments of the underlying continuous distributions.

cannot be influenced by local-level policy interventions. The underlying distribution of storm event intensity is continuous, with intensities typically measured using sustained wind speed. To communicate storm intensities, this continuous PDF is almost always discretized into an interval scale. A common example is the Saffir-Simpson Hurricane Wind Scale (http://www. nhc.noaa.gov/aboutsshws.php, accessed April 15, 2016), which distinguishes five categories of hurricane intensity. These range from Category 1 (least intense, with sustained winds of 74–95 mph) to Category 5 (most intense, with sustained winds of 157 mph or higher). This categorization may be further simplified into various single binary PDFs (e.g., Category 3 storm versus no Category 3 storm).

To formalize the alternative treatments of OU, we specify the probability of a less intense Category 2 coastal storm as P_L. The number of homes expected to flood conditional on a Category 2 storm is given by X_{1A}. The probability of a more intense (and less likely) Category 3 storm is given by P_H, leading to the flooding of an additional X_{1B} homes, beyond those that would flood in a Category 2 storm. Hence, the total number of homes expected to flood in a Category 2 storm is given by X_{1A}, and the total number of homes expected to flood in a Category 3 storm is given by $X_1 = X_{1A} + X_{1B}$. This combination of attributes and probabilities allows for OU associated with the risky attribute (homes flooded) to be represented using alternative treatments.

Grounded in this framework, we develop a split-sample DCE to evaluate the implications of representing OU using a more accurate distribution of risk-related outcomes. Survey version one, which we denote as S1, presents the number of homes that would flood conditional on a Category 2 storm (X_{1A}) and the number of additional homes that would be flooded conditional on a Category 3 (or higher) storm (X_{1B}). This is shown in Figure 1.[8] We refer to S1 as the *multiple-outcome* treatment of OU, as it allows for more than two possible storm outcomes (i.e., no storm, Category 2 storm, Category 3 storm). The otherwise-identical second version, which we denote as S2, communicates the same policy outcomes, but the risky attribute (homes flooded) is categorized into only two possible outcomes. This version presents only the total number of homes expected to flood in a Category 3 storm (given by $X_1 = X_{1A} + X_{1B}$). This is shown in Figure 2. We refer to S2 as the *two-outcome* treatment. Both versions S1 and S2 present a simplified reframing of OU, but S1 presents a more accurate discrete distribution.[9]

---

[8] In the study area, the distinction between a Category 3 storm and Category 3 or higher storm is effectively trivial, as no Category 4 or 5 storms have been recorded. Hence, the probability of storms of this intensity, based on historical frequency, is zero.

[9] Attribute X_{1B} in version S1 indicates homes to flood in a "Category 3+" storm while attribute X_1 in version S2 indicates homes to flood in a "Category 3" storm. Despite this minor textual difference, focus group results indicate that respondents interpreted these two descriptions as equivalent, as they were intended to be.

PLEASE VOTE

Question 4.

PROTECTION OPTION A and **PROTECTION OPTION B** are possible protection options for Old Saybrook. **NO NEW ACTION** shows what is expected to occur with no additional protection. All plans involve hard and soft defenses in different areas. Given a choice between the three, how would you vote?

Methods and Effects of Protection	Result in 2020s with NO NEW ACTION	Result in 2020s with PROTECTION OPTION A	Result in 2020s with PROTECTION OPTION B
	No Change in Existing Defenses	More Emphasis on SOFT Defenses	More Emphasis on SOFT Defenses
Homes Flooded in Category 2 Storm	**28%** 1,411 of 5,034 homes expected to flood in a Category 2 storm	**32%** 1,611 of 5,034 homes expected to flood in a Category 2 storm	**28%** 1,411 of 5,034 homes expected to flood in a Category 2 storm
Homes Flooded Only in Category 3+ Storm	**23%** 1,174 of 5,034 homes expected to flood only in a Category 3+ storm	**23%** 1,174 of 5,034 homes expected to flood only in a Category 3+ storm	**27%** 1,359 of 5,034 homes expected to flood only in a Category 3+ storm
Wetlands Lost	**5%** 25 of 497 wetland acres expected to be lost	**2%** 10 of 497 wetland acres expected to be lost	**2%** 10 of 497 wetland acres expected to be lost
Beaches and Dunes Lost	**10%** 3 of 30 beach acres expected to be lost	**16%** 5 of 30 beach acres expected to be lost	**16%** 5 of 30 beach acres expected to be lost
Seawalls and Coastal Armoring	**24%** 12 of 50 miles of coast armored	**24%** 12 of 50 miles of coast armored	**24%** 12 of 50 miles of coast armored
$ Cost to Your Household per Year	**$0** Increase in annual taxes or fees	**$125** Increase in annual taxes or fees	**$95** Increase in annual taxes or fees
HOW WOULD YOU VOTE? (CHOOSE ONLY ONE) I vote for	☐ I vote for NO NEW ACTION	☐ I vote for PROTECTION OPTION A	☐ I vote for PROTECTION OPTION B

Figure 1. Sample Choice Question from Survey Version S1

Beyond the inclusion of OU, the model is specified as a traditional random utility model in which household h chooses among three options ($k = A,B,N$), including two multi-attribute adaptation options (A,B) and a status quo (N),

PLEASE VOTE

Question 4.

PROTECTION OPTION A and **PROTECTION OPTION B** are possible protection options for Old Saybrook. **NO NEW ACTION** shows what is expected to occur with no additional protection. All plans involve hard and soft defenses in different areas. Given a choice between the three, how would you vote?

Methods and Effects of Protection	Result in 2020s with NO NEW ACTION	Result in 2020s with PROTECTION OPTION A	Result in 2020s with PROTECTION OPTION B
	No Change in Existing Defenses	More Emphasis on HARD Defenses	More Emphasis on HARD Defenses
Homes Flooded	**51%** 2,585 of 5,034 homes expected to flood in a Category 3 storm	**51%** 2,585 of 5,034 homes expected to flood in a Category 3 storm	**51%** 2,585 of 5,034 homes expected to flood in a Category 3 storm
Wetlands Lost	**5%** 25 of 497 wetland acres expected to be lost	**5%** 25 of 497 wetland acres expected to be lost	**2%** 10 of 497 wetland acres expected to be lost
Beaches and Dunes Lost	**10%** 3 of 30 beach acres expected to be lost	**4%** 1 of 30 beach acres expected to be lost	**4%** 1 of 30 beach acres expected to be lost
Seawalls and Coastal Armoring	**24%** 12 of 50 miles of coast armored	**24%** 12 of 50 miles of coast armored	**24%** 12 of 50 miles of coast armored
$ Cost to Your Household per Year	**$0** Increase in annual taxes or fees	**$95** Increase in annual taxes or fees	**$125** Increase in annual taxes or fees
HOW WOULD YOU VOTE? (CHOOSE ONLY ONE) I vote for	☐ I vote for NO NEW ACTION	☐ I vote for PROTECTION OPTION A	☐ I vote for PROTECTION OPTION B

Figure 2. Sample Choice Question from Survey Version S2

with no adaptation action and zero household cost. Each option is characterized by a vector of variables, $\boldsymbol{X} = [X_{1A}, X_{1B}...X_J]$ representing coastal adaptation outcomes. Variables $X_{1A}, X_{1B}...X_{J-1}$ represent non-monetary adaptation outcomes, including the number of homes expected to flood in storms of different intensity, and X_J represents unavoidable household cost. The vector P_{x1} represents a set of probabilities associated with home flooding attributes X_{1Ak} and X_{1Bk}. We represent the utility of household h from option k as

(1) $$U_{hk}(X_{1Ak}, X_{1Bk} \ldots X_{Jk}, P_{x1}) = u_{hk}(X_{1Ak}, X_{1Bk} \ldots X_{Jk}, P_{x1}) + \varepsilon_{hk},$$

where U_{hk} is the utility derived by household h from option k; U_{hk} (.) is a function representing the empirically measurable component of utility, and ε_{hk} is the unobservable component of utility and modeled as econometric error.[10]

If one assumes that outcomes are certain or that respondents maximize linear expected utility (e.g., are risk neutral), equation 1 may be specified using an additively separable, linear functional form. The resulting model may be estimated using maximum likelihood models for discrete outcomes (e.g., conditional, mixed, or generalized mixed logit), with likelihood functions determined by assumptions regarding such factors as the unobservable component of utility, scale heterogeneity, and preference heterogeneity (Hensher and Greene 2003, Train and Weeks 2005, Train 2009, Fiebig et al. 2010).

Models such as (1) may also be estimated using non-linear (in the parameters) functional forms that allow for behavior beyond risk neutral expected utility maximization (Glenk and Colombo 2013). Here, such models cannot be estimated because (as noted above), the probabilities of Category 2 and 3 storms are exogenous and fixed within our data, such that P_L and P_H are scalars (respondents were provided information on actual storm probability, which does not vary across the case study area). Hence, we proceed with estimation using a linear expected utility framework, while noting that models allowing for non-linear risk preferences may be more appropriate when probabilities vary across respondents or scenarios.

Given the structure of storm probabilities and home flooding specified above, a linear expected utility model for version S1 reduces to the general form

(2) $$U_{hk} = \alpha_{1Ah}(P_L + P_H)X_{1Ak} + \alpha_{1Bh}P_H X_{1Bk} + \sum_{j=2}^{J} \alpha_{jh}X_{jk} + \varepsilon_{hk},$$

because homes of type X_{1A} are expected to flood in either a Category 2 or 3 storm, whereas homes of type X_{1B} are only expected to flood in a Category 3 storm. However, given that P_L and P_H are fixed, equation 2 may be further simplified to

(3) $$U_{hk} = \beta_{1Ah}X_{1Ak} + \beta_{1Bh}X_{1Bk} + \sum_{j=2}^{J} \beta_{jh}X_{jk} + \varepsilon_{hk},$$

where $\beta_{1Ah} = \alpha_{1Ah}(P_L + P_H)$, $\beta_{1Bh} = \alpha_{1Bh}P_H$, and $\alpha_{jh} = \beta_{jh}$. Equations 2 and 3 are structurally and statistically equivalent when P_L and P_H are constant; the only difference is the interpretation of parameter estimates.

---

[10]   A parallel theoretical model applies to version S2, the only difference being that attributes X_{1A} and X_{1B} are represented by the single attribute X_1.

Grounded in utility of the form shown in (3), our primary hypotheses relate to whether or not and how a more continuous treatment of OU in DCEs influences results. The model leads to three primary hypotheses. First, if estimated marginal utilities (and associated implicit prices) are robust to OU reframing (assuming representative samples from the same population and that utility may be modeled using a linear approximation), then the marginal utility for X_{1B} in S1 should be equal to that for X_1 in S2; both reflect WTP to protect a marginal home that would otherwise be flooded in a Category 3 storm. This is the first hypothesis to be considered. That is, do the different treatments of OU lead to different inferences regarding marginal WTP for home protection? One may also consider similar comparisons for other model attributes, including whether variations in OU treatment influence other, non-risky attributes within the DCE.

If marginal utilities for X_{1A} and X_{1B} differ significantly within version S1, it would further suggest that there are potentially policy-relevant welfare differences that are obscured by the version that aggregates these two types of homes into a single category. This is the second hypothesis to be tested. That is, does the marginal WTP to prevent home flooding depend on the relative degree of flood risk facing those homes? Note that there are intuitive reasons for the implicit prices for X_{1A} and X_{1B} to differ. For example, in addition to the greater probability of a Category 2 storm relative to a Category 3 storm (which influences utility as shown in [2] and [3] above), respondents might have fundamentally different preferences for protecting higher- versus lower-risk homes.

Note that these and other hypothesis tests involve implicit prices defined as ex ante compensating surplus measures; they reflect WTP before uncertainty regarding future storm events is resolved (Just, Hueth, and Schmitz 2005). For example, an implicit price for X_1 would represent ex ante WTP for a policy that would protect a marginal home from flooding if a Category 3 storm occurs, where a Category 3 storm has a given probability of occurrence. Here, this is implied by the structure of (3), in which $\beta_{1Ah} = \alpha_{1Ah}(P_L + P_H)$ and $\beta_{1Bh} = \alpha_{1Bh}P_H$, such that β_{1Ah} and β_{1Bh} embed the fixed probability of storm occurrence. These welfare measures are not equal to ex post measures of compensating surplus, which would reflect WTP for a *guarantee* that the same homes will be prevented from flooding with 100 percent certainty. Results should be interpreted accordingly.

The third and final hypothesis to be tested concerns the presence of symptoms of greater complexity within version S1. Unlike the first and second hypotheses, the third hypothesis is not evaluated using a single test, but rather is evaluated using a weight of evidence that considers multiple symptoms of task complexity across the two models, such as measures of model fit, symptoms of decision heuristics, and the prevalence of related behaviors such as extreme and/or random preferences (e.g., as captured by variations in scale). These are discussed as part of the description of the empirical model below.

Note that the focus of this paper is distinct from most DCE studies that address OU: Its goal is not to estimate the effect of *varying* numerical and/or verbal risk information on choices; this issue has been addressed by numerous papers in the literature (e.g., Roberts, Boyer, and Lusk 2008, Rolfe and Windle 2009, 2015, Shaw and Baker 2010, Phillips 2011, Glenk and Colombo 2013). Here, the same underlying probabilities (i.e., the actual probabilities of Category 2 and 3 storms) are relevant to utility change under any adaptation option, and these probabilities cannot vary across choice scenarios. For this reason, the DCE holds probabilities (i.e., the probabilities of different intensity storms) constant across all surveys and scenarios and describes these probabilities explicitly on a separate page prior to choice questions.[11] This format is fundamental to the objective of the analysis. Studies that allow for only two possible outcomes of the OU-related attribute while embedding exogenously varying probabilities in valuation scenarios cannot easily isolate the effect of OU reframing from the effect of varying numerical risk information. The risk communication format adopted in this study is designed to isolate the effect of OU reframing, independent of whether and how probabilities vary across scenarios.[12]

Empirical Application

We implement the model using a DCE, addressing preferences for coastal flood adaptation, including the protection of homes and natural systems such as beaches and salt marshes. The DCE was developed over two years in a process involving economists and natural scientists; meetings and interviews with town planners and community officials, engineers and stakeholder groups; and 13 focus groups with residents. Issues considered during focus groups included but were not limited to: (a) respondents' understanding of the coastal adaptation policy context, (b) adaptation outcomes most relevant to respondents, (c) the interpretation of valuation scenarios, including any symptoms of scenario adjustment (Cameron, DeShazo, and Johnson 2011), (d) respondents' understanding of information presented by the questionnaire, including risk information, and (e) cognitive processes used to answer DCE questions. The first version of the survey instrument was developed after input from four initial focus groups that helped define attributes and scenarios to be considered. Later focus groups led to extensive questionnaire

---

[11]   The omission of numerical probabilities from choice scenarios does not imply that indicated outcomes are certain. Indicated outcomes of homes expected to flood are subject to OU because they are contingent on storm events whose occurrence is not known with certainty, and this was explicitly communicated to respondents prior to the choice questions. Omitting numerical probabilities from choice scenarios constitutes merely an alternative risk communication format.

[12]   Further, there are ethical concerns with purposefully communicating incorrect storm probabilities to respondents (e.g., that they might take real actions based on incorrect information).

improvements, including updates to information provision, DCE attribute definitions, the set of attributes included, and the communication of OU.[13] Additional details of pretesting conducted in focus groups are described below. Particular attention was given to the provision and understanding of information on storm probability and associated flooding.

Prior to administration of choice questions, both survey versions provided identical text describing the historical frequency of Category 2 and 3 storms in Old Saybrook and the numerical probability that a storm of each intensity would occur at least once by the mid-2020s (based on historical frequencies). These probabilities were calculated from historical patterns of Category 2 and 3 storms over the previous 75 years, leading to an approximate 20 percent probability that at least one Category 3 storm would occur in the region by 2025, and a parallel 55 percent probability of at least one Category 2 storm. To ensure that respondents understood relationships between storms of different intensities and home flooding, questionnaire version S1 (with two risk-related attributes X_{1A} and X_{1B}) included Figure 3, a simple graphic illustrating the risk-related differences between these two groups of homes. This figure also clarified the definition of these two groups of homes, for example highlighting that they are independent and non-overlapping groups. Respondents' understanding of this information was tested in focus groups.

Both survey versions also provided identical information describing tradeoffs associated with alternative approaches to coastal adaptation, along with projected inundation scenarios in the mid-2020s and baseline (status quo) effects with no new adaptation actions. The data used to inform choice scenarios and inundation models (i.e., future storm and inundation scenarios) were provided by the Center for Climate Systems Research at Columbia University, NASA's Goddard Institute for Space Studies, the National Oceanic and Atmospheric Administration (NOAA) Services Center and The Nature Conservancy (TNC), as reflected in coastal flooding scenarios for TNC's Coastal Resilience platform (http://www.coastalresilience.org) (Beck et al. 2013).

This and other information was conveyed via a combination of text, custom-designed graphics, geographic information system (GIS) maps, and photographs. Detailed instructions were also provided, including reminders to consider budget constraints and statements highlighting consequentiality (Carson and Groves 2007). Survey language and graphics were subject to extensive pretesting in focus groups and cognitive interviews (Johnston et al. 1995, Kaplowitz, Lupi, and Hoehn 2004), including the use of verbal protocols to gain insight into respondents' comprehension and decision processes (Schkade and Payne 1994). This helped ensure that language, graphics and format were understood, that respondents and researchers shared interpretations of

---

[13]   For example, focus groups motivated the design of a figure illustrating the differences in flood risk facing different types of homes (Figure 3).
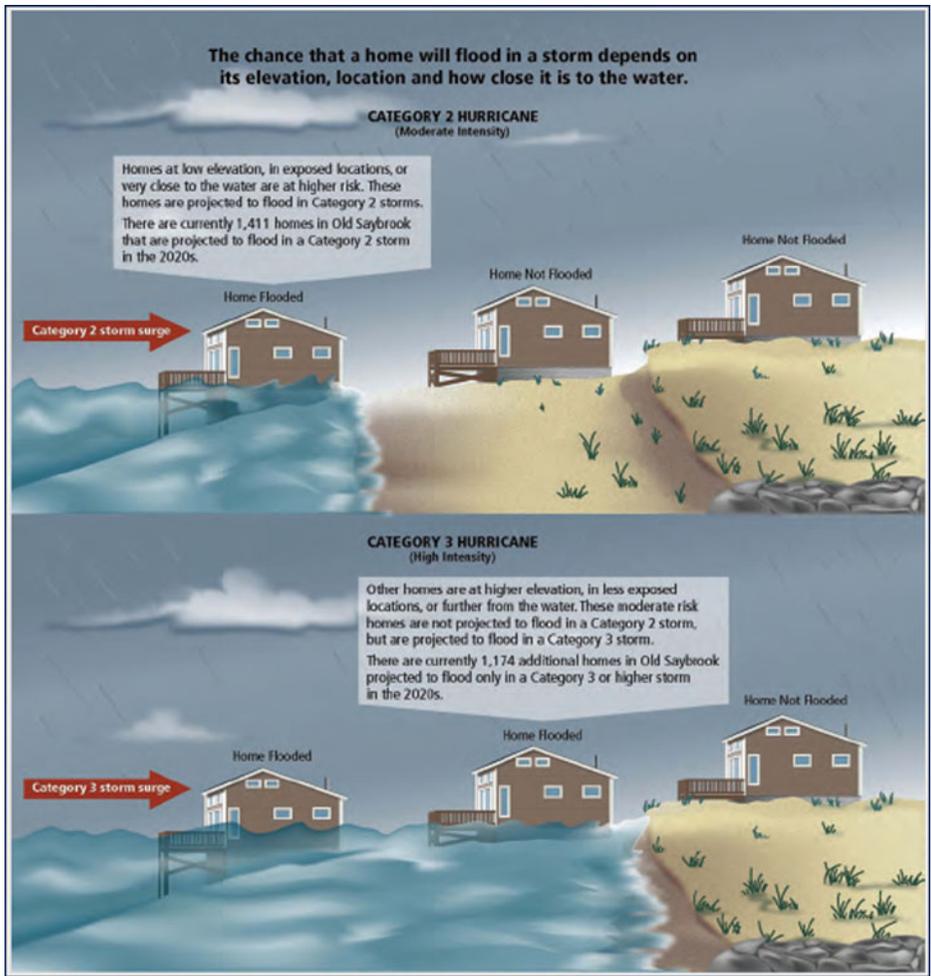
Figure 3. Illustrates the Difference between the Two Home Groups in Survey Version S1

terminology and scenarios, and that scenarios captured outcomes viewed as relevant and accurate.

In questionnaire version S1, choice options are characterized by seven attributes: the percentage and number of homes expected to flood only in a Category 3 storm, the percentage and number of homes expected to flood in a Category 2 storm, wetland acreage expected to be lost, beach and dune acreage to be lost, the length of coastline to be hard-armored, whether or not the overall plan emphasizes hardened coastal defenses, and unavoidable household cost (Table 2). All outcomes are forecast as of the mid-2020s.

Table 2. Choice Experiment Attributes and Descriptive Statistics[14]

Variable Name	Definition	S1 Survey (More Accurate OU Scenario) Mean (Std. Dev.)	S2 Survey (Two-Outcome OU Scenario) Mean (Std. Dev.)
Homes	The total percentage of Old Saybrook homes expected to flood in a high-intensity (Category 3) storm in the mid-2020s	—	48.95 (7.28)
Homes2	The percentage of Old Saybrook homes expected to flood in a Category 2 storm in the mid-2020s	26.18 (3.90)	—
Homes3	The percentage of Old Saybrook homes expected to flood *only* in a Category 3 or higher storm in the mid-2020s (i.e., not expected to flood in a Category 2 storm)	21.97 (3.46)	—
Wetlands	The percentage of Old Saybrook's coastal marshes expected to be lost by the mid-2020s due to flooding or erosion	5.05 (2.63)	5.06 (2.61)
Beaches	The percentage of Old Saybrook's beaches and dunes expected to be lost by the mid-2020s due to flooding or erosion	9.53 (4.28)	9.40 (4.32)
Seawalls	The percentage of Old Saybrook's coast shielded by hard armoring by the mid-2020s	24.94 (6.29)	24.77 (6.39)
Cost	Cost to the household per year, in unavoidable taxes and fees. These funds are legally guaranteed to be spent only on the selected coastal protection option	62.46 (56.49)	62.45 (56.14)
Neither	Alternative specific constant (ASC) associated with the status quo, or a choice of neither Option A or Option B	0.33 (0.47)	0.33 (0.47)
Hard	Binary variable indicating whether the adaptation policy emphasizes hard or engineered defenses; a value of 1 indicates an emphasis on hard defenses. Range 0–1	0.19 (0.39)	0.20 (0.40)

[14]Means and standard deviations include status quo option of no adaptation action.

Questionnaire version S2 is identical except that it only included one attribute characterizing the percentage and number of homes expected to flood in a Category 3 storm. That is, version S1 allows for different degrees of home flooding conditional on storms of different intensity, with each storm intensity level associated with a given probability. In contrast, version S2 mirrors standard treatments by communicating homes flooded under a single binary event (storm/no storm), with one associated probability.

Following the general approach of Johnston et al. (2011, 2012), attributes represent each adaptation method and effect in relative (percentage) terms with regard to upper and lower reference conditions (i.e., best and worst possible in Old Saybrook) as defined in survey informational materials. Scenarios also present the cardinal basis for relative levels where applicable. With the exception of hard armoring levels and household cost, relative attribute levels represent losses or flooding approaching the upper reference condition (100 percent loss or flooding), starting from the lower reference condition (0 percent loss or flooding). For example, the attribute representing the number of homes expected to flood in a Category 2 storm (*homes2*) is presented both as a cardinal number of homes and as a percentage relative to the total number of homes in Old Saybrook. To ensure comparability, all home flooding attributes (*homes*, *homes2* and *homes3*) are denominated as a percentage of all homes in Old Saybrook, where 1 percent = 50.34 homes.

Grounded in these attribute levels, a fractional factorial experimental design was generated using a D-efficiency criterion (Sándor and Wedel 2001, 2002, Ferrini and Scarpa 2007, Rose and Bliemer 2009) for main effects and selected two-way interactions, yielding 72 profiles blocked in 24 booklets. Other measures of efficiency were also reviewed for each candidate design, e.g., S-efficiency, to evaluate potential sample sizes required to estimate preference parameters, for assumed utility specifications (Scarpa and Rose 2008, Bliemer, Rose, and Hensher 2009, Rose and Bliemer 2009). Design efficiency was also reevaluated using alternative assumptions for preference parameters and utility structure.[15] Each respondent was provided with three choice questions and instructed to consider each as independent and non-additive. Table 2 provides attribute summary statistics and definitions. Table 3 illustrates attribute levels. Sample choice questions from Surveys S1 and S2 are shown in Figures 1 and 2, respectively.

Surveys were sent to 1,152 randomly selected Old Saybrook households via U.S. postal mail during May through July 2014, with repeated mailings

---

[15]   Independent D-efficient designs were developed for each choice experiment, reflecting the distinct characteristics of each design and associated utility function. These designs were tested using a variety of approaches, including robustness to alternative parameter priors. Given the established capacity of these types of designs to enable unbiased parameter estimates (Sándor and Wedel 2001, 2002, Ferrini and Scarpa 2007, Rose and Bliemer, 2009), and the fact that independent designs were optimized for each DCE, there is no evidence that results of the present analysis are driven by the experimental design.

Table 3. Attribute Levels in Choice Experiment Designs (Identical for S1 and S2 Surveys)

S1 and S2 Survey Versions Variable Noted	Attribute Levels Across Choice Options
Homes (relevant only to S2 survey)	36%; 43%; 51%; 59%
Homes2	20%; 24%; 28%; 32%
Homes3	16%; 19%; 23%; 27%
Wetlands	2%; 5%; 10%
Beaches	3%; 5%; 7%; 10%
Seawalls	15%; 24%; 35%
Cost	$0; $35; $65; $95; $125; $155
Hard	0 (Emphasis on natural defenses) 1 (Emphasis on hard defenses)

following Dillman et al. (2009) to increase response rates. The analysis is based on 167 returns from 493 deliverable S1 surveys, for a 33.87 percent response rate, and 163 returns from the 488 deliverable S2 surveys, for a 33.40 percent response rate.[16]

Model Estimation

The models are estimated using generalized multinomial logit (GMNL; Fiebig et al. 2010) in WTP-space (cf. Cameron and James 1987, Scarpa, Thiene, and Train 2008, Thiene and Scarpa 2009, Hensher and Greene 2011). WTP-space estimation circumvents challenges associated with ambiguously defined welfare estimates from certain types of preference-space mixed logit specifications. These are caused by the presence of a randomly specified cost coefficient in the denominator of the WTP expression, leading to behaviorally and statistically implausible WTP values, e.g., due to infinite WTP moments (Hensher and Greene 2003, Train and Weeks 2005, Scarpa, Thiene and Train 2008, Thiene and Scarpa 2009, Daly, Hess, and Train 2012). Cost coefficient distributions that lead to finite WTP moments can lead to additional

---

[16]   The two samples are similar in age, education, and household income. For example, the median age of version S1 and S2 respondents is 58 and 61, a 4-year college or higher degree is held by 61.2 percent of version S1 respondents and 61.7 percent of version S2 respondents, and both samples' median household income category is $100,000 to $249,000. Both samples were thus somewhat older and wealthier than the median Old Saybrook resident who is 50 years of age and has a household income of $71,800. Also, the percentage of Old Saybrook residents holding a 4-year college or higher degree is 44.79 percent, whereas in both samples it is about 61 percent.

challenges, such as unrealistic welfare estimates associated with the tails of the lognormal distribution (Hensher and Greene 2003).

The model is specified following Hensher and Greene (2011), based on the WTP-space model of Scarpa, Thiene, and Train (2008) and Thiene and Scarpa (2009). Grounded in (3), for household h facing choice alternatives $k = A$, B, N and $t = 1...T$ choice sets, the model in preference space may be restated in streamlined form as

$$(4) \qquad U_{hkt} = \boldsymbol{B'}_h \boldsymbol{X}_{hkt} + \varepsilon_{hkt},$$

where $\boldsymbol{B}_h$ is a vector of random parameters underlying marginal utilities of attributes $\boldsymbol{X}_{hkt}$ and ε_{hkt} is a traditional i.i.d. type-one extreme-value error with constant variance. Within a random parameter preference-space model, $\boldsymbol{B}_h$ may be specified as

$$(5) \qquad \boldsymbol{B}_h = \boldsymbol{B} + \boldsymbol{\Gamma v}_h,$$

where $\boldsymbol{\Gamma}$ is a lower triangular matrix that provides the standard deviations and covariances of $\boldsymbol{B}_h$ and $\boldsymbol{v}_h$ is a vector of random variables that provide the stochastic element of preference weights.

Fiebig et al. (2010) extend this model to distinguish the possibility that preference heterogeneity is either independent of, or proportional to, scale heterogeneity by specifying

$$(6) \qquad \boldsymbol{B}_h = \sigma_h \boldsymbol{B} + \sigma_h[\gamma + (1 - \gamma)]\boldsymbol{\Gamma v}_h$$

$$(7) \qquad \sigma_h = \exp(-\tau^2/2 + \tau w_h).$$

In this model, $\boldsymbol{\Gamma}$ and $\boldsymbol{v}_h$ are as defined as above, σ_h is the logit scale parameter, τ is the primary parameter that determines scale heterogeneity (with $w_h \sim N(0,1)$), and γ is a weighting parameter that lies between zero and one, governing how the variance of residual preference heterogeneity varies with scale (Hensher and Greene 2011). The traditional random parameters model in (4) is a special case of the GMNL model with $\gamma = 0$ and $\tau = 0$ (the multinomial logit model arises when $\gamma = 0$, $\boldsymbol{\Gamma} = 0$, and $\tau = 0$ so that $\sigma_h = \sigma = 1$).[17]

---

[17]   As noted by Hess and Train (2017), the scale parameter in this model captures *all* correlation among utility coefficients, including that related to preferences rather than scale. Results must be interpreted accordingly.

To derive the GMNL in WTP-space, we assume that utility is separable in price, p_{hkt},[18] and other attributes, $\boldsymbol{x_{hkt}}$, and rewrite equation 4 as

$$(8) \qquad\qquad U_{hkt} = \lambda_h p_{hkt} + \boldsymbol{\delta}'_{\boldsymbol{h}}\boldsymbol{x_{hkt}} + \varepsilon_{hkt}$$

where $\lambda_h = c_h/\sigma_h$ is the preference-space coefficient on program cost, c_h is the marginal utility of income, and $\boldsymbol{\delta}'_{\boldsymbol{h}} = (\boldsymbol{\beta_h}/\sigma_h)'$ is a conforming vector of coefficients on non-price attributes $\boldsymbol{x_{hkt}}$. From (8), WTP may be calculated as the ratio of the coefficient on any non-price attribute and the coefficient on cost, $\boldsymbol{\omega_h} = (\boldsymbol{\delta_h}/\lambda_h) = (\boldsymbol{\beta_h}/c_h)$. Given this, we can rewrite (8) to derive the parallel WTP-space specification (Train and Weeks 2005),

$$(9) \qquad\qquad \begin{aligned} U_{hkt} &= \lambda_h[p_{hkt} + (1/\lambda_h)\boldsymbol{\delta}'_{\boldsymbol{h}}\boldsymbol{x_{hkt}}] + \varepsilon_{hkt}, \\ &= \lambda_h[p_{hkt} + \boldsymbol{\omega}'_{\boldsymbol{h}}\boldsymbol{x_{hkt}}] + \varepsilon_{ihk}, \end{aligned}$$

where (9) is behaviorally equivalent to (8), its preference-space analog. By setting $\gamma = 0$, the elements of the resulting vector $\boldsymbol{\omega}'_{\boldsymbol{h}}$ are random coefficients representing direct estimates of WTP, assumed to be normally distributed. The parameter λ_h becomes the normalizing constant (Hensher and Greene 2011), and the WTP-space model thus circumvents the need to estimate the ratio of two parameters via a direct estimate of the ratio. This specification enables the analyst to specify the assumed distribution of WTP directly (Train and Weeks 2005, Scarpa, Thiene, and Train 2008). To ensure the positive marginal utility of income, we specify $\lambda_h = -e^{\varphi_h}$, where $\varphi_h = (-\tau^2/2 + \tau w_h)$ is the latent normal factor defining the lognormally distributed cost coefficient (Hensher and Greene 2011).

From equation 9, additional multiplicative interactions are included between a set of socioeconomic variables and *neither*, the alternative specific constant for the status quo (Morrison et al. 2002, Columbo, Calatrava-Requena, and Hanley 2007, Johnston and Duke 2010). These variables included the respondent's age in years (*age*), and dummy variables identifying households earning more than $60,000 per year (*hi_income*) and respondents with at least a 4-year college degree (*college*). An identical specification was applied to each model, with parameters on *neither* and homes expected to flood (*homes2* and *homes3* in S1 and *homes* in S2) specified random, assuming normal and independent distributions. Other parameters are specified as non-random.[19] The sign of the cost variable was reversed prior to estimation. Models were estimated with different starting values and numbers of Halton draws to evaluate robustness and local versus global convergence. Parallel preference-space

---

[18]   Note that p_{hkt} is equivalent to X_{jk} in equation (1).
[19]   This was necessary to ensure convergence of identical specifications for both DCE versions.

models were estimated to further evaluate the robustness of results; these lead to similar conclusions as models reported in the main text below (see appendix). The final model is estimated using 500 Halton draws.

Results

Results are reported in Table 4. Two models are illustrated, corresponding to versions S1 and S2. Estimated WTP coefficients from versions S1 and S2 are jointly significant at $p \leq 0.0001$ ($\chi^2 = 163.52$, df $= 16$ and $\chi^2 = 186.94$, df $= 14$). Signs of estimated coefficients are identical in both versions, with the exception of coefficient estimates on *seawalls* (length of coastline to be hard-armored by the mid-2020s). The sign of all coefficient estimates match prior expectations, where prior expectations exist.[20] All main effect coefficient estimates are statistically significant in version S2. In version S1, all are significant except for the coefficients on *neither* and *hard*. In version S1, coefficient estimates on all socioeconomic interaction terms with the ASC are statistically significant, except for that on $ne \times hi\_income$. In version S2, only the coefficient estimate on $ne \times college$ is statistically significant.

For continuous attributes (*homes, homes2, homes3, beaches, wetlands, seawalls*), coefficients reflect mean implicit prices (marginal WTP) per percentage point increase, along with the standard deviations of these implicit prices (for random parameters). For example, within version S1, annual per-household WTP per percentage point increase in *homes2* is $-\$6.75$, implying that the average household is willing to pay \$6.75 to avoid a one-percentage-point increase in homes expected to flood in a Category 2 storm. Similarly, in version S2, the average household is willing to pay \$7.36 to avoid an identical magnitude increase in *homes* expected to flood in a Category 3 storm.[21] The coefficient estimate on binary attributes such as *hard* and the ASC *neither* reflect the mean implicit price for the presence of these attributes relative to their absence. In version S2, for example, the coefficient estimate of -76.56 on *hard* suggests that the average household is willing to pay \$76.56 annually to avoid adaptation action that emphasizes hard, engineered defenses.

---

[20] The coefficient on *seawalls* has an ambiguous expected sign, given the potentially positive and negative impacts of seawalls on utility (e.g., due to the protective services of seawalls, but also their generally undesired aesthetic properties). These dual positive and negative aspects were highlighted in focus groups.

[21] There is no necessary relationship between WTP estimates for homes and flood insurance premiums. Insurance premiums reflect private values—i.e., what people are willing to pay to protect/insure their own home. The WTP values generated here reflect public values—i.e., what people are willing to pay to protect homes community-wide (primarily other people's homes). For example, the latter estimates of WTP may include altruistic and other non-use values that are not incorporated into insurance premiums.

Table 4. Results of WTP-Space Generalized Multinomial Logit Models

Choice Attribute	Version S1 (Multiple Outcome Scenario) (Std. Error)	Version S2 (Two-Outcome Scenario) (Std. Error)	Difference WTP(S1)–WTP(S2)[a] (Std. Deviation)
Random parameters			
Neither (ne)	−137.46 (106.93)	−148.34** (65.25)	10.96 (126.05)
ne × age	−6.09*** (2.22)	0.44 (0.81)	−6.24*** (2.34)
ne × hi_income	−104.73 (87.22)	−19.21 (26.05)	−75.13 (93.60)
ne × college	−388.29*** (134.43)	−85.11*** (25.19)	−288.37** (147.41)
Homes	— (—)	−7.36*** (1.02)	— (—)
Homes2	−6.75*** (1.56)	— (—)	0.58[a] (1.84)
Homes3	−11.74*** (2.67)	— (—)	−4.26*[a] (2.90)
Wetlands	−6.67** (2.92)	−4.43*** (1.71)	−2.23 (3.38)
Beaches	−6.49*** (1.56)	−6.27*** (1.06)	−0.26 (1.83)
Hard	−21.35 (18.78)	−76.56*** (23.57)	55.50*** (23.95)
Seawalls	−2.27* (1.26)	2.09** (0.83)	−4.27*** (1.47)
Standard deviation of random parameters			
Neither	2499.22*** (834.42)	333.49*** (61.34)	2054.23*** (884.81)
Homes	— (—)	−9.30*** (1.21)	— (—)
Homes2	0.99 (2.15)	— (—)	10.25***[a] (2.41)
Homes3	0.63 (1.10)	— (—)	9.90***[a] (1.60)
$\ln(\lambda_h)$	3.46*** (0.67)	−2.94*** (0.90)	— (—)

Parameter τ			Difference in τ (S1–S2)
τ	−2.35*** (0.53)	−1.62** (0.65)	−0.72 (0.86)
Model statistics			
Observations (N)	1023	1164	—
Respondents	341	388	—
χ^2/Significance level	163.52 (16 d.f.)/0.00000001	186.94 (14 d.f.)/0.00000001	—
Pseudo-R^2	0.22	0.22	—

[a]Results show the means and standard deviations of the simulated empirical distributions of implicit price differences. For *homes2* and *homes3*, WTP(S1)–WTP(S2) tests $WTP_{homes2} - WTP_{homes}$ and $WTP_{homes3} - WTP_{homes}$.
***, **, * imply significance at 1 percent, 5 percent and 10 percent level.

Model results also provide insight into preference heterogeneity. For example, results for versions S1 and S2 suggest statistically significant WTP heterogeneity for *neither*, along with systematic variation in WTP associated with sociodemographic variables. Version S1 finds WTP homogeneity for *homes2* and *homes3*, whereas version S2 finds significant WTP heterogeneity for *homes*. Additional discussion of parallels and differences between the two models is provided below.

Overall, the two versions lead to similar general conclusions regarding WTP for adaptation outcomes. These conclusions are consistent with both intuition and results from focus groups. Further, viewed in isolation, both survey versions generate seemingly reasonable WTP estimates (with the exception of results for the ASC in version S1, discussed below). However, despite these general similarities, differences between the two models suggest that the treatment of OU can have non-trivial implications for welfare estimation and choice complexity. The remainder of this paper focuses on these differences, oriented around the three hypotheses described above.

Effects of OU Treatments

Differences between coefficient estimates from versions S1 and S2 are tested using a parametric, empirical convolution of estimate distributions, with significance levels determined through percentiles on the resulting distributions (Poe, Giraud, and Loomis 2005). Implicit price differences are reported as the mean of the resulting simulated distributions (of differences); these hence differ slightly from point estimate differences in WTP means. For the home flooding attributes (where attributes differ across S1 and S2), we test whether implicit prices for either *homes2* or *homes3* in S1 (WTP_{homes2}, WTP_{homes3}) are equal to the implicit price for *homes* in S2 (WTP_{homes}). Other than the expected relationship between WTP for X_{1B} in S1 and X_1 in S2 (discussed above), theory provides ambiguous guidance regarding the expected relationship among these three measures of WTP. Nonetheless, these findings are relevant to whether or not and how OU treatments influence welfare estimates.

Results are shown in the last column of Table 4. Tests for the first hypothesis lead us to reject convergent validity for roughly half of coefficient estimates: *ne × age*, *ne × college*, *hard*, *seawalls*, and *homes3* (in version S1) and *homes* (in version S2) are statistically different between versions S1 and S2. For example, we find relatively large and statistically significant implicit price differences for the binary variable *hard* ($55.50; $p < 0.10$) and the multiplicative interaction terms *ne × age* and *ne × college* (−$6.24 and −$288.37, respectively). That is, when compared across models, alternative treatments of OU for home flooding appear to have an effect on mean implicit prices for multiple adaptation outcomes, including those for the home flooding attributes (*homes*, *homes2* and *homes3*).

Versions S1 and S2 also lead to different conclusions regarding implicit price standard deviations associated with home flooding. We find statistically

insignificant standard deviation estimates for *homes2* and *homes3* in S1. In contrast, the standard deviation estimate on *homes* in S2 is statistically significant. These differences in estimated standard deviations across models are statistically significant at p < 0.01 (i.e., between the standard deviations of *homes2* and *homes*, and between *homes3* and *homes*). Simply put, when home flooding OU is discretized using a single variable (*homes*), we estimate an implicit price with heterogeneity around the mean; this apparent heterogeneity disappears when OU is discretized using two variables (*homes2*, *homes3*).[22]

Considering the second hypothesis, we also reject the null hypothesis that $WTP_{homes2} = WTP_{homes3}$ in version S1 (Wald $\chi^2 = 3.41$; p ≤ 0.0648). As discussed above, there are multiple reasons to expect that these implicit prices might differ,[23] although theory provides ambiguous guidance regarding the sign and magnitude of differences. Regardless, the finding that these implicit prices are not identical suggests that version S1 (the more accurate representation of OU) reveals information about respondents' risk preferences that is not provided by version S2. Results suggest that residents are willing to pay $4.84 more per percentage point reduction in the number of homes expected to flood only in a Category 3 storm than they are for the same reduction in homes expected to flood in Category 2 storm.

This result is particularly interesting because the *ex post* number of homes prevented from flooding is expected to be greater from a one percentage point change in *homes2* than for *homes3* (once the uncertainty is resolved). Despite this, we find that $|WTP_{homes2}| < |WTP_{homes3}|$. This result is consistent with focus group results suggesting that Old Saybrook residents view the protection of homes at very high flood risk (e.g., *homes2*) to be the primary responsibility of the homeowners who purchase this high-risk property. In contrast, residents feel more of an altruistic motive to protect homes in lower-risk locations (*homes3*).[24] Hence, the value per certain home protected is expected to be less for *homes2* than *homes3*, leading to the implicit price patterns of Table 4.[25]

---

[22] One possible explanation of this result is that the apparent welfare heterogeneity in version S2 may be due to different interpretations of the OU representation, rather than true preference heterogeneity. Although this explanation cannot be verified, these results at a minimum indicate that statistical inferences regarding preference heterogeneity are sensitive to the level of discretization of OU.

[23] For example, focus group participants initiated debates about which of the two home groups should be given protection priority, suggesting that residents hold different underlying values for the protection of homes that are at different risks of flooding.

[24] It is also possible that the difference in WTP between the two home groups is due to assumed differences in anticipated damage severity (e.g., the damage caused by 10 cm versus 1 m of water in one's home) that might occur in storms of different intensity. To maintain a manageable level of complexity within the survey, we did not address the extent of flooding to each home; only whether it floods or not.

[25] Focus groups suggested that respondents interpreted the two home attributes as intended, as non-overlapping with *homes3* as the number of additional homes to be flooded in a Category 3 storm. As discussed above, this was clarified using Figure 3, which focus group participants

These findings suggest that the more accurate discretization of OU underlying version S1 may generate additional insights—and different information—than the simpler two-outcome discretization underlying version S2. Results also demonstrate that variations in OU framing for a single attribute (here, home flooding) can lead to non-trivial differences in welfare estimates associated with other model variables, even when other aspects of DCE scenarios are held constant.

The third hypothesis concerns whether or not there are symptoms of additional choice complexity in responses to version S1. Here, there is some evidence that version S1 presented a more complex choice task, leading to greater application of choice heuristics. For example, the very large estimated standard deviation for *neither* in version S1 suggests the potential presence of serial choice behavior (or serial participation / non-participation), in which a subset of respondents universally accepts or rejects most or all status quo options, regardless of other attributes (Kriström 1997, von Haefen, Massey, and Adamowicz 2005). Behavior such as this is often interpreted as a sign of task complexity, in which "the complexity of a decision may preclude the use of all available information" (Mazzotta and Opaluch 1995, p. 500).

The combined magnitudes of the implicit prices associated with *neither*, *ne × age*, *ne × hi_income* and *ne × college* provides an additional sign of complexity associated with version S1. For example, given a mean age of 58 among version S1 respondents, the mean combined point estimate of WTP associated with the status quo option (*neither* = 1 versus 0) for college-educated respondents is −\$983.70, holding all else constant. This improbably large magnitude also suggests that choices may have been influenced by decision heuristics inconsistent with fully compensatory decision-making (e.g., serial anti-status-quo choices).

In contrast, the statistical significance of τ in equation 7 does not suggest additional complexity of version S1. This parameter may be interpreted as the degree to which scale varies across the sample (Fiebig et al. 2010), or the degree of respondent-specific variation in choice randomness or inconsistency within the sample.[26] Although the point estimate of τ is 45 percent greater in version S1 than S2 (−2.35 versus −1.62), the difference between these estimates is not statistically significant. This result suggests

---

appeared to interpret as intended. Discussions with residents and town officials suggested that the majority of residents have experience with the effects of such storm events and how effects vary with storm intensity. This perhaps explains why focus group participants found the distinction between *homes2* and *homes3* relatively easy to understand. Empirical results support the view that respondents interpreted the home groups as intended. Version S1 (the 3-state world) generates a WTP of \$6.75 for *homes2* and \$11.74 for *homes3*. Version S2 (the 2-state world which "pools" the two groups) leads to a WTP of \$7.36. The relative ordering of these three estimates is consistent with respondents (correctly) understanding that the average risk facing *homes* in S2 falls between the risks facing *homes2* and *homes3* in S1.

[26]   In the WTP-space model, this parameter also captures variations in the marginal utility of income (Thiene and Scarpa 2009).

that scale variation across respondents is statistically indistinguishable between the two versions.[27]

Similar findings emerge from multiple alternative WTP- and preference-space model specifications. For illustration, the appendix reports the results of parallel specifications estimated using a mixed logit preference-space model that holds scale constant. Results are generally consistent with those reported above, with variations in marginal utility estimates and statistical significance that are similar to those found in the WTP-space GMNL models (Table 4).[28] Welfare estimates from these models also show similar patterns.

These results highlight potential challenges facing stated preference valuation of outcomes subject to non-trivial OU. Results show that standard two-outcome treatments may suppress some policy-relevant aspects of preferences that are revealed when OU is presented with greater accuracy. That is, the methods used to discretize OU are not welfare-neutral. However, our results also suggest that this additional welfare information may come at a cost of higher task complexity and prevalence of choice heuristics, with symptoms including large implicit prices and implicit price standard deviations associated with the status quo ASC. At the same time, roughly half of implicit price estimates were invariant to the alternative OU treatments. This is an encouraging outcome, suggesting that some welfare inferences from DCEs are robust, notwithstanding other potential tradeoffs associated with alternative approaches to OU framing.

Conclusions

Results of the models shown above suggest the need for additional research on the effect of alternative approaches to OU discretization within DCEs. This is an area that has been subject to relatively little research in the literature, despite the relative frequency with which DCEs address outcomes subject to OU. We also, however, emphasize that the results presented here are specific to our case study of coastal hazard adaptation, and that similar findings may or may not apply to other contexts. For example, case studies addressing very familiar topics and risks might enable respondents to more easily process additional OU resolution. In addition, results might vary in cases where outcome probabilities change across scenarios, as opposed to the present application in which probabilities are purposefully fixed. These and other

---

[27]   This result is consistent with qualitative findings from focus groups. Although version S1 was subjected to fewer pre-tests than version S2, we found no compelling evidence that focus group participants found the more accurate OU scenario depicted in version S1 to be prohibitively complex, or that they misinterpreted or were confused by it.

[28]   Concerns regarding relatively large estimates associated with the status quo option (*neither*) in version S1 prompted us to run dozens of alternatively specified preference- and WTP-space models to check for consistency. Although point estimates of WTP differed slightly across models, the differences between versions S1 and S2 were robust, including WTP associated with the status quo.

caveats aside, the present results suggest that increases in the resolution and accuracy of OU framing within DCEs may require the development of novel methods that enable respondents to understand more accurate probability distributions without concomitant increases in scenario complexity. Pending the development of such methods, researchers may wish to consider the possibility that DCE results may be conditional on the number of intervals used to present OU.

References

Akter, S., J. Bennett, and M.B. Ward. 2012. "Climate Change Scepticism and Public Support for Mitigation: Evidence from an Australian Choice Experiment." *Global Environmental Change* 22(3): 736–745.

Baker, J., W.D. Shaw, D. Bell, S. Brody, M. Riddel, R.T. Woodward, and W. Neilson. 2009. "Explaining Subjective Risks of Hurricanes and the Role of Risks in Intended Moving and Location Choice Models." *Natural Hazards Review* 10(3): 102–112.

Beck, M.W., B. Gilmer, Z. Ferdaña, G.T. Raber, C. Shepard, I. Meliane, J.D. Stone, A.W. Whelchel, and M. Hoover. 2013. "Increasing the Resilience of Human and Natural Communities to Coastal Hazards: Supporting Decisions in New York and Connecticut." In F.G. Renaud, K. Sudmeier-Rieux, and M. Estrella, eds., *The Role of Ecosystems in Disaster Risk Reduction*. Bonn: United Nations University Press, pp. 140–163.

Bliemer, M.C., J.M. Rose, and D.A. Hensher. 2009. "Efficient Stated Choice Experiments for Estimating Nested Logit Models." *Transportation Research Part B: Methodological* 43 (1): 19–35.

Boyle, K.J. and S. Özdemir. 2009. "Convergent Validity of Attribute-based, Choice Questions in Stated-preference Studies." *Environmental and Resource Economics* 42(2): 247–264.

Cameron, T.A., J.R. DeShazo, and E.H. Johnson. 2011. "Scenario Adjustment in Stated Preference Research." *Journal of Choice Modelling* 4(1): 9–43.

Cameron, T.A. and M.D. James. 1987. "Estimating Willingness to Pay from Survey Data: An Alternative Pre-test-market Evaluation Procedure." *Journal of Marketing Research* 24 (4): 389–395.

Carmines, E.G. and R.A. Zeller. 1979. *Reliability and Validity Assessment* (vol. 17). Beverly Hills, California: Sage Publications.

Carson, R.T. and T. Groves. 2007. "Incentive and Informational Properties of Preference Questions." *Environmental and Resource Economics* 37(1): 181–210.

Columbo, S., J. Calatrava-Requena, and N. Hanley. 2007. Testing Choice Experiment for Benefit Transfer with Preference Heterogeneity. *American Journal of Agricultural Economics* 89(1): 135–151.

Corso, P.S., J.K. Hammitt, and J.D. Graham. 2001. "Valuing Mortality-risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation." *Journal of Risk and Uncertainty* 23(2): 165–184.

Daly, A., S. Hess, and K. Train. 2012. "Assuring Finite Moments for Willingness to Pay in Random Coefficient Models." *Transportation* 39(1): 19–31.

Dillman, D.A., G. Phelps, R. Tortora, K. Swift, J. Kohrell, J. Berck, and B.L. Messer. 2009. "Response Rate and Measurement Differences in Mixed-mode Surveys Using Mail, Telephone, Interactive Voice Response (IVR) and the Internet." *Social Science Research* 38(1): 1–18.

Ferrini, S. and R. Scarpa. 2007. "Designs with a Priori Information for Nonmarket Valuation with Choice Experiments: A Monte Carlo Study." *Journal of Environmental Economics and Management* 53(3): 342–363.

Fiebig, D.G., M.P. Keane, J. Louviere, and N. Wasi. 2010. "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity." *Marketing Science* 29(3): 393–421.

Fifer, S., S. Greaves, J. Rose, and R. Ellison. 2011. "A Combined GPS/Stated Choice Experiment to Estimate Values of Crash-risk Reduction." *Journal of Choice Modelling* 4(1): 44–61.

Fischhoff, B., A. Bostrom, and M.J. Quadrei. 1993. "Risk Perception and Communication." *Annual Review of Public Health* 14: 183–203.

Freeman, A.M. III, J.A. Herriges, and C.L. Kling. 2014. *The Measurement of Environmental and Resource Values: Theory and Methods*. New York: Routledge.

Gilboa, G., N. Sochen, and Y.Y. Zeevi. 2002. "Forward-and-backward Diffusion Processes for Adaptive Image Enhancement and Denoising." *IEEE Transactions on Image Processing* 11(7): 689–703.

Gilboa, I., A.W. Postlewaite, and D. Schmeidler. 2008. "Probability and Uncertainty in Economic Modeling." *The Journal of Economic Perspectives* 22(3): 173–188.

Glenk, K. and S. Colombo. 2011. "How Sure Can You Be? A Framework for Considering Delivery Uncertainty in Benefit Assessments Based on Stated Preference Methods." *Journal of Agricultural Economics* 62(1): 25–46.

———. 2013. "Modelling Outcome-Related Risk in Choice Experiments." *Australian Journal of Agricultural and Resource Economics* 57(4): 559–578.

Graham, D.A. 1981. "Cost-benefit Analysis Under Uncertainty." *American Economic Review* 71: 715–725.

Hanley, N., B. Kriström, and J.F. Shogren. 2009. "Coherent Arbitrariness: on Value Uncertainty for Environmental Goods." *Land Economics* 85(1): 41–50.

Hensher, D.A. and W.H. Greene. 2003. "The Mixed Logit Model: The State of Practice." *Transportation* 30(2): 133–176.

———. 2011. "Valuation of Travel Time Savings in WTP and Preference Space in the Presence of Taste and Scale Heterogeneity." *Journal of Transport Economics and Policy* 45(3): 505–525.

Hess, S. and K. Train. 2017. "Correlation and Scale in Mixed Logit Models." *Journal of Choice Modelling* 23: 1–8.

Holland, D.S., J.N. Sanchirico, and R.J. Johnston. 2010. *Economic Analysis for Ecosystem-Based Management: Applications to Marine and Coastal Environments*. New York: Routledge.

Ivanova, G., J. Rolfe, and G. Tucker. 2010. "Using Choice Modelling to Assess the Willingness to Pay of Queensland Households to Reduce Greenhouse Emissions." In 54*th Annual Australian Agricultural and Resource Economics Society Conference, Adelaide*.

Johnston, R.J., K.J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T.A. Cameron, W.M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, and C.A. Vossler. 2017. "Contemporary Guidance for Stated Preference Studies." *Journal of the Association of Environmental and Resource Economists* 4(2): 319–405.

Johnston, R.J. and J.M. Duke. 2010. "Socioeconomic Adjustments and Choice Experiment Benefit Function Transfer: Evaluating the Common Wisdom." *Resource and Energy Economics* 32(3): 421–438.

Johnston, R.J., E.T. Schultz, K. Segerson, E.Y. Besedin, and M. Ramachandran. 2012. "Enhancing the Content Validity of Stated Preference Valuation: The Structure and Function of Ecological Indicators." *Land Economics* 88(1): 102–120.

Johnston, R.J., K. Segerson, E.T. Schultz, E.Y. Besedin, and M. Ramachandran. 2011. "Indices of Biotic Integrity in Stated Preference Valuation of Aquatic Ecosystem Services." *Ecological Economics* 70(11): 1946–1956.

Johnston, R.J. and S.K. Swallow. 1999. "Asymmetries in Ordered Strength of Preference Models: Implications of Focus Shift for Discrete-choice Preference Estimation." *Land Economics* 75(2): 295–310.

Johnston, R.J., T.F. Weaver, L.A. Smith, and S.K. Swallow. 1995. "Contingent Valuation Focus Groups: Insights from Ethnographic Interview Techniques." *Agricultural and Resource Economics Review* 24(1): 56–69.

Just, R.E., D.L. Hueth, and A. Schmitz. 2005. *The Welfare Economics of Public Policy: A Practical Approach to Project and Policy Evaluation*. Northampton, MA: Edward Elgar Publishing.

Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *The American Economic Review* 93(5): 1449–1475.

Kaplowitz, M.D., F. Lupi, and J.P. Hoehn. 2004. "Multiple Methods for Developing and Evaluating a Stated-choice Questionnaire to Value Wetlands." In S. Presser, J.M. Rothget, M.P. Cooper, J.T. Lesser, E. Martin, J. Martin, and E. Singer, eds., *Methods for Testing and Evaluating Survey Questionnaires* (Chapter 24, p. 503). Hoboken, NJ: John Wiley and Sons.

Kriström, B. 1997. "Spike Models in Contingent Valuation." *American Journal of Agricultural Economics* 79(3): 1013–1023.

Lundhede, T., J.B. Jacobsen, N. Hanley, N. Strange, and B.J. Thorsen. 2015. "Incorporating Outcome Uncertainty and Prior Outcome Beliefs in Stated Preferences." *Land Economics* 91(2): 296–316.

Macmillan, D., N. Hanley, and S. Buckland. 1996. "A Contingent Valuation Study of Uncertain Environmental Gains." *Scottish Journal of Political Economy* 43(5): 519–533.

Mazzotta, M.J. and J.J. Opaluch. 1995. "Decision Making When Choices Are Complex: A Test of Heiner's Hypothesis." *Land Economics* 71(4): 500–515.

Miller, A.C. III and T.R. Rice. 1983. "Discrete Approximations of Probability Distributions." *Management Science* 29(3): 352–362.

Morrison, M., J. Bennett, R. Blamey, and J. Louviere. 2002. "Choice Modeling and Tests of Benefit Transfer." *American Journal of Agricultural Economics* 84(1): 161–170.

Patt, A.G. and D.P. Schrag. 2003. "Using Specific Language to Describe Risk and Probability." *Climatic Change* 61(1–2): 17–30.

Peters, E., D. Västfjäll, P. Slovic, C.K. Mertz, K. Mazzocco, and S. Dickert. 2006. "Numeracy and Decision Making." *Psychological Science* 17(5): 407–413.

Phillips, Y. 2011. "August. When the Tide is High: Estimating the Welfare Impact of Coastal Erosion Management." In *New Zealand Agricultural & Resource Economics Society Conference. Waikato Regional Council, Hamilton East, New Zealand*.

Poe, G.L., K.L. Giraud, and J.B. Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87(2): 353–365.

Reynaud, A., C. Aubert, and M.H. Nguyen. 2013. "Living with Floods: Protective Behaviours and Risk Perception of Vietnamese Households." *The Geneva Papers on Risk and Insurance-Issues and Practice* 38(3): 547–579.

Roberts, D.C., T.A. Boyer, and J.L. Lusk. 2008. "Preferences for Environmental Quality Under Uncertainty." *Ecological Economics* 66(4): 584–593.

Rolfe, J. and J. Windle. 2009. "Valuing the Control of Red Imported Fire Ants in Australia Using Choice Modelling." In *Environmental Economics Research Hub Research Reports* (No. 44). Canberra, Australia: Crawford School of Economics and Government, Australian National University.

———. 2010, January. "Valuing Protection of the Great Barrier Reef with Choice Modelling by Management Policy Options." (Paper, BioEcon Conference, Venice, September 27–28, 2010).

———. 2015. "Do Respondents Adjust Their Expected Utility in the Presence of an Outcome Certainty Attribute in a Choice Experiment?" *Environmental and Resource Economics* 60(1): 125–142.

Rose, J.M. and M.C. Bliemer. 2009. "Constructing Efficient Stated Choice Experimental Designs." *Transport Reviews* 29(5): 587–617.

Sandor, Z. and M. Wedel. 2001. "Designing Conjoint Choice Experiments Using Managers' Prior Beliefs." *Journal of Marketing Research* 38(4): 430–444.

Sándor, Z. and M. Wedel. 2002. "Profile Construction in Experimental Choice Designs for Mixed Logit Models." *Marketing Science* 21(4): 455–475.

Scarpa, R. and J.M. Rose. 2008. "Design Efficiency for Non-market Valuation with Choice Modelling: How to Measure It, What to Report and Why." *Australian Journal of Agricultural and Resource Economics* 52(3): 253–282.

Scarpa, R., M. Thiene, and K. Train. 2008. "Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps." *American Journal of Agricultural Economics* 90(4): 994–1010.

Schkade, D.A. and J.W. Payne. 1994. "How People Respond to Contingent Valuation Questions: A Verbal Protocol Analysis of Willingness to Pay for an Environmental Regulation." *Journal of Environmental Economics and Management* 26(1): 88–109.

Shaw, W.D. and J. Baker. 2010. "Models of Location Choice and Willingness to Pay to Avoid Hurricane Risks for Hurricane Katrina Evacuees." *International Journal of Mass Emergencies and Disasters* 28(1): 87–114.

Slovic, P. 1987. Perception of Risk. *Science* 236(4799): 280–285.

Slovic, P., B. Fischhoff, and S. Lichtenstein. 1985. Characterizing Perceived Risk. In R.W. Kates, C. Hohenemser, and J. X. Kasperson, eds., *Perilous Progress: Managing the Hazards of Technology* (Chapter 5, pp. 91–125). Boulder, CO: Westview.

Slovic, P., E. Peters, M.L. Finucane, and D.G. MacGregor. 2005. Affect, Risk, and Decision Making. *Health Psychology* 24(4S): S35.

Starmer, C. 2000. Developments in Non-expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *Journal of Economic Literature*, 38(2): 332–382.

Thiene, M. and R. Scarpa. 2009. Deriving and Testing Efficient Estimates of WTP Distributions in Destination Choice Models. *Environmental and Resource Economics*, 44(3): 379–395.

Torres, C., Faccioli, M. and Font, A.R., 2017. Waiting or acting now? The effect on willingness-to-pay of delivering inherent uncertainty information in choice experiments. *Ecological Economics*, 131(January 2017): 231–240.

Train, K. and M. Weeks. 2005. Discrete Choice Models in Preference Space and Willingness-to-pay Space. In R. Scarpa and A. Alberini, eds., *Applications of Simulation Methods in Environmental and Resource Economics* (Chapter 1, pp. 1–16). New York: Springer.

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

Veronesi, M., Chawla, F., Maurer, M. and Lienert, J., 2014. Climate change and the willingness to pay to reduce ecological and health risks from wastewater flooding in urban centers and the environment. *Ecological Economics*, 98(February 2014): 1–10.

Viscusi, W.K., W.A. Magat, and J. Huber. 1987. An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks. *The RAND Journal of Economics*, 18(4): 465–479.

Von Haefen, R.H., D.M. Massey, and W.L. Adamowicz. 2005. "Serial Nonparticipation in Repeated Discrete Choice Models." *American Journal of Agricultural Economics*, 87(4): 1061–1076.

Wibbenmeyer, M.J., Hand, M.S., Calkin, D.E., Venn, T.J. and Thompson, M.P., 2013. Risk preferences in strategic wildfire decision making: a choice experiment with US wildfire managers. *Risk analysis*, 33(6): 1021–1037.

Wielgus, J., L.R. Gerber, E. Sala, and J. Bennett. 2009. "Including Risk in Stated-preference Economic Valuations: Experiments on Choices for Marine Recreation." *Journal of Environmental Management* 90(11): 3401–3409.

Appendix

Table A1 reports a preference-space, mixed logit counterpart of the WTP-space GMNL models reported in Table 4, estimated to evaluate the robustness of results reported in the main text. Identical specifications were again applied to both DCE versions: parameters on *cost, neither, homes2* and *homes3* (in S1), and *homes* (in S2) were specified random and normally distributed, with *cost* sign-reversed prior to estimation. Estimated coefficients in versions S1 and S2 are again jointly significant at p $\leq$ 0.0001 ($\chi^2 = 153.86$; df $= 15$ and

Table A1. Results of Preference-Space Mixed Logit Models

Choice Attribute	Version S1 (Multiple Outcome Scenario) (Std. Error)	Version S2 (Two-Outcome Scenario) (Std. Error)
Random parameters		
Neither	−1.643 (4.183)	1.026 (2.634)
ne × age	−0.020 (0.061)	−0.038 (0.037)
ne × hi_income	0.646 (1.888)	−0.462 (1.222)
ne × college	−1.680 (1.753)	−1.579 (1.079)
Homes	— (—)	−0.134*** (0.033)
Homes2	−0.064* (0.037)	— (—)
Homes3	−0.091** (0.042)	— (—)
Wetlands	−0.070 (0.049)	−0.115** (0.055)
Beaches	−0.062* (0.029)	−0.099*** (0.031)
Hard	−0.602* (0.352)	−1.073*** (0.340)
Seawalls	−0.026 (0.021)	−0.005 (0.026)
Cost	0.010* (0.005)	0.019*** (0.005)
Standard deviation of random parameters		
Neither	−6.912*** (1.654)	4.417*** (0.975)
Homes	— (—)	0.176*** (0.047)
Homes2	0.047 (0.091)	— (—)
Homes3	−0.091 (0.090)	— (—)
Cost	0.026*** (0.007)	−0.020*** (0.007)
No. of observations (*N*)	1104	1164
Respondents	341	388
χ^2/Significance level	153.86 (15 d.f.)/0.0000	115.78 (13 d.f.)/0.0000
Pseudo-R^2	0.21	0.15
AIC	607.175	697.807
LL function	−288.588	−335.904

***, **, * imply significance at 1 percent, 5 percent, and 10 percent level.

$\chi^2 = 115.78$; df $= 13$) with pseudo-R^2 values of 0.21 and 0.15, respectively. Here, main effect coefficient estimates are interpreted as relative marginal utilities.

Across both models, the signs and statistical significance of coefficient estimates are again identical. The only exception is the significance associated with the main effect coefficient estimates on *wetlands*, which is borderline insignificant in version S1 (Table A1). Most findings of the preference-space models are consistent with results in the main text. There are a few differences in the statistical significance of coefficient estimates between these results and those reported in the main text; these are not surprising given the high level of scale heterogeneity present in both survey versions (preference-space mixed logit models do not account for such variation). Results of these preference-space specifications yield the same general conclusions regarding the effect of OU discretization, with (a) many resulting welfare estimates different across versions S1 and S2, and (b) additional evidence of complexity in version S1.