What to Value and How? Ecological Indicator Choices in Stated Preference Valuation

Minjuan Zhao · Robert J. Johnston · Eric T. Schultz

Abstract The ecological literature accepts that many policy outcomes cannot be observed directly and must be characterized using indicators. Multiple indicators can often be used to communicate similar ecological outcomes. Previous studies using alternative indicators in stated preference surveys suggest that welfare estimates may be indicator-dependent, casting doubt on whether welfare estimates are sufficiently reliable for cost benefit analysis. We suggest that the reason for such indicator dependence may be that indicators used in these prior studies represented different outcomes valued by respondents. This possibility underscores the need for greater attention to selection of indicators and their properties within stated preference survey design. This paper develops a model introducing the concept of outcome equivalent indicators, defined as indicators that provide alternative representations of identical underlying outcomes. To assess empirically whether welfare estimates are indeed robust to indicator choice when alternative indicators are expected to be outcome equivalent, we analyze data from a choice experiment estimating willingness to pay for migratory fish restoration in Rhode Island, USA. Results demonstrate that welfare estimates are robust to the use of alternative ecological indicators within stated preference scenarios.

M. Zhao
College of Economics and Management, Northwest A&F University,
No. 3 Taicheng Road, Yangling 712100, Shaanxi, People's Republic of China
e-mail: minjuan.zhao@nwsuaf.edu.cn

R. J. Johnston (✉)
Department of Economics and George Perkins Marsh Institute,
Clark University, 950 Main St., Worcester, MA 01610, USA
e-mail: rjohnston@clarku.edu

E. T. Schultz
Department of Ecology and Evolutionary Biology, University of Connecticut,
Storrs, CT 06269-3043, USA
e-mail: eric.schultz@uconn.edu

1 Introduction

There is substantial variation in the ecological indicators used to represent similar outcomes within stated preference (SP) surveys (Schultz et al. 2012). Unanticipated sensitivity of welfare estimates to these variations could threaten the perceived reliability of SP welfare estimation for policy analysis. Specifically, if estimated willingness to pay (WTP) for a particular ecological outcome were to vary depending on the specific indicator used to characterize that outcome, the appropriate value to use within cost benefit analysis could be unclear. Moreover, the resulting welfare estimates could be viewed as dependent mostly on researcher discretion (i.e., the particular set of indicators used to characterize a given outcome within a survey instrument).

Despite the relevance of these concerns for welfare analysis, the literature provides little theoretical or empirical guidance regarding whether and how welfare estimates should respond to variations in ecological indicators within SP survey design. This paper models and evaluates the typically unacknowledged implications of these variations for SP welfare estimation, addressing the common situation in which alternative indicators may be chosen by the researcher to communicate change in the same presumptive ecological outcome. We consider two primary questions related to the consequences of such choices: (1) Conceptually, when should SP welfare estimates be robust to changes in ecological indicators used to characterize the same policy outcomes? (2) Empirically, when they are expected to be so, are estimates indeed robust to indicator choice?

1.1 Ecological Indicators and Valued Outcomes in Stated Preference Valuation

The ecological literature has long accepted that many outcomes such as wildlife abundance, biodiversity and water quality cannot be directly observed or adequately quantified using a single metric, and must be communicated via indicators.[1] Multiple indicators can often be used to communicate the same or similar ecological outcomes, or changes in valued ecosystem goods and services. For example, the condition of a wildlife species might be represented using indicators of species abundance (e.g., population size, number of breeding pairs), frequency (e.g., likelihood of encountering the species in a given area, frequency of successful breeding), habitat distribution, population viability, or classified status (e.g., threatened, endangered).[2] Because cognitive limitations constrain the number of attributes that respondents can consider simultaneously within a survey scenario (DeShazo and Fermo 2002), SP survey designers must choose a small set of such indicators to communicate policy outcomes. Even after substantial attention to survey design and testing (e.g., within focus groups and cognitive interviews (Johnston et al. 1995; Kaplowitz et al. 2004; Powe

---

[1] There is a mature literature addressing the use of indicators to communicate unobservable but valued outcomes. Examples include Bortone (2005), Davis and Simon (1995), Jackson et al. (2000), Jordan and Smith (2005), Jørgensen et al. (2004), Naweedi (2005), Niemi and McDonald (2004), Schiller et al. (2001), Suter (1993, 2001).

[2] Population viability reflects the probability that a species will continue to be found in an area after a specified period of time (Boyce 1992; Lee and Rieman 1997). For an application to SP valuation see Johnston et al. (2012).

2007), multiple indicators representing identical or similar outcomes may remain under consideration.[3]

The resulting diversity in survey design possibilities prompts concern over whether alternative indicators communicate the status of *identical* or *related but distinct* ecological outcomes. That is, respondents' assumed relationship between indicators and outcomes is crucial to understanding *what is being valued* within a SP survey, and how this relates to policy outcomes in question and the potential robustness of welfare estimates. In some cases, for example, indicators such as species abundance and frequency might represent alternative means to communicate the same underlying outcome, e.g., species condition. In other cases, such as when values are driven by preferences for recreational encounters, these indicators might represent distinct outcomes. In the latter case, the outcome valued by respondents—and hence WTP—will likely depend on which of these indicators is used to characterize policy outcomes. Such challenges are particularly germane for valuation of policies affecting ecological systems, as these policies often affect dozens of commonly-used ecological indicators, even when the set of primary welfare-relevant outcomes is small.

To clarify subsequent discussion, we define *outcome equivalent* indicators as those that communicate the status of identical underlying outcomes, as perceived by respondents. When two ecological indicators are outcome equivalent, they may in principle be used interchangeably within a SP survey, subject to other indicator properties such as simplicity and ease of communication. That is, one should be able to interchange outcome equivalent indicators within a SP survey without significant effects on welfare estimates. To our knowledge no prior research has demonstrated robustness of welfare estimates to indicator choice in such circumstances.

To the contrary, some prior research suggests that WTP is sensitive to alternative ecological indicators within SP scenarios. Jacobsen et al. (2008), for example, find that WTP estimates in biodiversity valuation differ depending on whether outcomes are characterized in terms of effects on habitat indicators versus indicator species. When interpreting such results, however, it is important to consider whether alternative indicators are outcome equivalent. Does changing ecological indicators in a SP survey influence the perceived outcome being valued by respondents? If so, WTP variability would be expected. That is, when two or more ecological indicators are *not* outcome equivalent, exchanging these indicators within a survey scenario will lead to a change in the underlying outcome being valued, and an attendant change in WTP.

This paper proposes a model and linked approach to SP design that formalizes the concept of outcome equivalence and enables the evaluation of empirical consequences (e.g., for welfare estimates). Illustrative models are estimated and hypothesis tests implemented using choice experiments applied to migratory fish restoration in a Rhode Island (USA) watershed. Results demonstrate that SP welfare estimates can be statistically robust to the use of alternative outcome equivalent ecological indicators within survey scenarios. To our knowledge, this is the first published convergent validity [4] test of independent SP welfare estimates derived using distinct ecological indicators that characterize the same underlying outcomes. Our findings reinforce the reliability of SP models for ecological policy analysis,

---

[3] Van Houtven et al. (2007), Walsh and Wheeler (2011) and Johnston et al. (2005), for example, discuss alternative indicators that may be used to communicate similar changes in water quality. Jacobsen et al. (2008) provide a parallel discussion in the context of wildlife and biodiversity.

[4] Convergent validity reflects whether a measure is correlated with other measures of the same theoretical construct. In valuation research, convergent validity tests typically compare two or more welfare measures that are expected to be similar or identical based on economic theory (Bateman et al. 2002).

and suggest that the sensitivity of welfare estimates to the choice of ecological indicators in prior studies may have been due to changes in the outcomes that were valued by respondents.

2 Theoretical Model of Indicator Selection and Outcome Equivalence

Despite the relevance of ecological indicator use and selection for SP welfare estimation, their implications have been largely overlooked by the valuation literature.[5] Most analyses appear to treat choices among alternative indicators as inconsequential, providing little insight into the process of indicator selection and mapping between indicators and ecological outcomes. This relatively casual treatment of distinctions between indicators and welfare-relevant policy outcomes is somewhat justified, at least in a superficial sense. If indicators within SP surveys are adequately defined and respondents have well-defined preferences, then respondents should be able to make preference-revealing choices; their responses are expected to accurately reflect their interpretation of valuation scenarios (assuming there are no strategic or other biases).

Below the surface of these simple assumptions, however, there may be potential ambiguity about what is being valued by respondents. Scenario interpretation may depart from that which the survey designer intends (Carson 1998; Tonsor 2011), in part because of unclear relationships between indicators and outcomes. For example, researchers presenting a certain set of ecological indicators to respondents might erroneously presume that estimated WTP reflects change to an underlying ecological outcome Q_1, whereas it may reflect assumed changes to other outcomes (e.g., Q_2 or Q_3). Similarly, it may be unclear whether two indicators W_1 and W_2 represent the same outcome Q_1 (in which case only one should likely be used within a SP scenario) or distinct outcomes Q_1 and Q_2 (in which case both might be included to ensure comprehensive welfare estimates). This lack of clarity can engender uncertainty regarding which welfare estimates should be used in cost benefit analysis and whether SP results are sufficiently reliable for policy evaluation.

In this way, the conceptual issue is somewhat related to the concepts of embedding and scope (e.g., Carson and Mitchell 1995; Carson et al. 2001; Jakobsson and Dragun 2001; Powe and Bateman 2004; Veisten et al. 2004). Embedding refers to the issue of whether alternative goods represent different quantities of the same underlying good (i.e. they are nested subsets) or instead represent related but not formally nested goods (Carson and Mitchell 1995), whereas scope refers to the perceived quantity of a single good (Powe and Bateman 2004). Carson and Mitchell (1995) demonstrate that while formal nesting of goods may be straightforward in theory, in practice it may be unclear whether two different measures indeed represent different quantities (or scopes) of the same good. Here the question is similar: whether different indicators represent the same good, albeit perhaps at different scopes, or instead represent similar but distinct goods. As with the matter of embedding, the issue is straightforward in theory but more opaque in practice.

To formalize these issues, we begin with a simple theoretical framework in which utility, $U(\cdot)$, is a function of two ecological public goods or outcomes that cannot be directly observed, Q_1 and Q_2, and a numéraire private good X (e.g., income). Hence, $U(\cdot) = U(Q_1, Q_2, X)$. We assume that the goal of valuation is to establish marginal WTP

[5] The biodiversity valuation literature (e.g., Christie et al. 2006; Czajkowski et al. 2009; Jacobsen et al. 2008; Nunes and van den Bergh 2001; Spash and Hanley 1995) discusses ways in which different aspects of biodiversity may be communicated via indicators. With the exception of Jacobsen et al. (2008) these discussions are primarily qualitative, with no formal empirical analysis of the influence of indicator selection on welfare estimates.

for changes in Q_1. Respondents infer the quantity or quality of the outcome via a relationship between the outcome and an ecological indicator represented as a simple function $Q_1 = f(W_{1A})$, where W_{1A} is an ecological indicator used to communicate the status of Q_1. Alternative indicators W_{1B} or W_{1C} also represent changes in Q_1, where $Q_1 = g(W_{1B})$ and $Q_1 = h(W_{1C})$. That is, W_{1A}, W_{1B} or W_{1C} may all be used to communicate the status of otherwise unobservable outcome Q_1.[6]

While outcome Q_1 may be communicated adequately using only one indicator, we assume that outcome Q_2 requires two distinct ecological indicators, W_{1C} and W_{2A}. That is, $Q_2 = k(W_{1C}, W_{2A})$, where W_{1C} is the same indicator identified in $Q_1 = h(W_{1C})$ above. We emphasize that $f(\cdot)$, $g(\cdot)$, $h(\cdot)$ and $k(\cdot)$ are assumed relationships used by respondents to infer change in outcomes Q_1 and Q_2, based on ecological indicators W_{1A}, W_{1B}, W_{1C}, and W_{2A}. They are not "ecological production functions" in the strict sense (cf. Johnston and Russell 2011), because the outcomes are not produced by the associated indicators.

The use of indicator W_{1A} in SP scenarios to estimate welfare change associated with a change in Q_1 is grounded in the underlying theoretical relationship

$$\frac{\partial U(\cdot)}{\partial W_{1A}} = \left(\frac{\partial U(\cdot)}{\partial Q_1}\right)\left(\frac{\partial Q_1}{\partial W_{1A}}\right) \tag{1}$$

The first term in (1) may be derived from survey responses (e.g., the coefficient estimate on W_{1A} within a choice experiment). The second term reflects the underlying marginal influence of Q_1 on utility. The third term represents the relationship between the indicator (W_{1A}) and the otherwise unobservable outcome (Q_1) assumed by respondents; this may be informed by information provided to respondents within the survey but is also influenced by respondents' prior perceptions and understanding. A parallel structure applies for indicator W_{1B}. That is, W_{1A} and W_{1B} are *outcome equivalent*; they are alternative means to communicate the same underlying outcome and are linked to $\partial U(\cdot)/\partial Q_1$ in the same structural manner. Estimated utility change ($\partial U(\cdot)/\partial Q_1$) from a change in ecological outcome Q_1—and hence WTP for this change—should be equivalent, whether evaluated using either W_{1A} or W_{1B} (one would not use both indicators simultaneously, because they communicate equivalent information). This provides the answer to the conceptual question presented in the previous section (i.e., when should SP welfare estimates be robust to changes in ecological indicators used to characterize policy outcomes?).

Note that this equivalence does not imply that the corresponding WTP associated with the indicators W_{1A} or W_{1B} will be identical, because it will generally be the case that

$$\partial Q_1/\partial W_{1A} \neq \partial Q_1/\partial W_{1B}, \tag{2}$$

so that

$$\partial U(\cdot)/\partial W_{1A} \neq \partial U(\cdot)/\partial W_{1B} \tag{3}$$

That is, the functional relationship between the W_{1A} and the utility-relevant but unobservable outcome Q_1 is not necessarily the same as that between W_{1B} and Q_1, so that the per unit utility gain associated with W_{1A} and W_{1B} is not necessarily identical, even though they are both linked to Q_1. [7]

---

[6] Unbiased welfare estimation requires that these assumed relationships are ecologically valid (Johnston et al. 2012).

[7] For example, both population size (W_{1A}) and the number of breeding pairs (W_{1B}) might be considered outcome equivalent ways communicate the abundance of a bird species (Q_1). Even so, this would not imply that WTP per bird should be equivalent to WTP per breeding pair, because these two indicators are related to species abundance at different scales.

The situation is more complex if W_{1C} is instead used as an indicator of Q_1. In this case,

$$\frac{dU(\cdot)}{dW_{1C}} = \left(\frac{\partial U(\cdot)}{\partial Q_1}\right)\left(\frac{\partial Q_1}{\partial W_{1C}}\right) + \left(\frac{\partial U(\cdot)}{\partial Q_2}\right)\left(\frac{\partial Q_2}{\partial W_{1C}}\right) \qquad (4)$$

Welfare changes that correspond with changes in W_{1C} will capture both the relationship between W_{1C} and Q_1 and between W_{1C} and Q_2. If the latter effect is not anticipated, the result will be a misleading perspective of the value of Q_1. That is, W_{1A} and W_{1C}, while both indicators of Q_1, are not outcome equivalent because W_{1C} is also an indicator of outcome Q_2. Hence, the interchange of W_{1A} or W_{1C} in a survey scenario influences the underlying outcomes that are valued.

The applied valuation literature largely overlooks such concerns, proceeding as if ecological indicators necessarily represent the outcomes to be valued. This assumption is equivalent to a special case in which $f(\cdot)$, $g(\cdot)$, and $h(\cdot)$ are identity functions, such that $W_{1A} = W_{1B} = W_{1C} = Q_1$. In the general case, the outcomes inferred by respondents when viewing a particular ecological indicator may be *ex ante* unclear, and can only be revealed through focus groups, pretests and other best practices in survey development that clarify distinctions between ecological conditions (or processes), indicators and values. For example, in the case of threatened wildlife, do respondents truly and directly value a change in official species status (e.g., endangered to threatened), or does this status convey an underlying but unobservable public good over which respondents have preferences (e.g., the likelihood that a species will not become extinct)? The few past analyses that demonstrate sensitivity of WTP to the communication of ecological outcomes (e.g., Jacobsen et al. 2008) may have inadvertently substituted indicators that were not outcome equivalent, highlighting the importance of more formal treatment of such relationships.

It is also important to emphasize that outcome equivalence as defined here, while related to ecological properties, is a function of respondents' preferences and perceptions. Ecological properties alone cannot determine whether respondents interpret two (or more) indicators as referencing an identical underlying good. Hence, while ecological relationships provide a *necessary* condition for outcome equivalence (i.e., the indicators in question must have some similarity in terms of the outcomes to which they refer, cf. Schiller et al. 2001), they do not provide a *sufficient* condition. The latter requires that respondents use the alternative indicators (e.g., W_{1A} and W_{1B}) to infer change in the same element in their utility functions (i.e., Q_1).

3 Formalizing Indicator Selection in Practice

Lacking attention to the problem, the literature also lacks assurance that there is a workable solution. Is there any evidence that WTP estimates are robust to the alternative use of outcome equivalent ecological indicators in SP scenarios? For example, assume that survey designers have identified outcome equivalent indicators that might be used to characterize a policy outcome, and that design constraints prevent the use of both indicators simultaneously. Are such indicators interchangeable within an SP survey without attendant changes in welfare estimates? Or, will any—even seemingly immaterial—change in ecological indicators influence welfare estimates, leading to a potentially problematic sensitivity of WTP?

We address this issue from two perspectives. First, we present a structured process designed to formalize and clarify the selection of ecological indicators within SP design. This process provides a practical means to address relationships between indicators and underlying goods, as well as distinctions between ecological outcomes or processes, indicators of those out-

comes, and economic values. Second, we present a case study to illustrate the empirical results of such a process and to test outcome equivalence for the resulting set of indicators.

Like all forms of SP content validity (Mitchell and Carson 1989; Bateman et al. 2002), the adequacy ecological information within survey design is context specific. For example, the interpretation of an ecological indicator may vary between two valuation contexts.[8] Because of this, a systematic, context-specific process of indicator selection is required to ensure that indicators reflect utility-relevant outcomes and to evaluate whether alternative indicators are outcome equivalent. Without such a process, there is no way to determine whether any given indicator, regardless of its ecological properties, successfully communicates a particular outcome.

Here, we summarize the process of ecological indicator selection developed for the present case study. While not the only possible means to select indicators within SP design, the proposed methods provide a formal approach that links ecological science to public preferences, and informs linkages between valued outcomes and indicators. The process included four steps:

1. Focus groups and cognitive interviews (Kaplowitz et al. 2004) were first used to identify primary, welfare-relevant outcomes (here resulting from fish restoration in Rhode Island). These initial focus groups and interviews were designed to elicit ways in which respondents valued potential policy outcomes. These outcomes were grouped into categories of welfare-relevant goods and services (e.g., effect on migratory fish, quantity of river habitat restored, etc.), using ethnographic focus group methods outlined by Johnston et al. (1995).

2. Based on focus group and interview results, these goods and services were further disaggregated into final and intermediate outcomes (goods and services). Final goods and services are defined as outcomes that directly enhance respondents' utility, whereas intermediate goods and services are defined as inputs into the biophysical production of final goods and services; they have no direct influence on utility (Johnston and Russell 2011).[9] Following Johnston et al. (2012), only final goods and services were considered for inclusion in survey scenarios. These are akin to *assessment endpoints* in the ecological literature—often unobservable policy goals that influence well-being or affect utility (US EPA 1998).

3. Following Schiller et al. (2001), each assessment endpoint was formally linked to one or more *measurement endpoints*; these are observable ecological indicators used within formal frameworks to communicate, infer, or predict changes in assessment endpoints (US EPA 1998). The result was a conceptual model relating each assessment endpoint (final good or service) to a set of ecologically-linked measurement endpoints (indicators). This step required review of the ecological literature and extensive input from experts in fish restoration ecology. Guidance described in Johnston et al. (2012) was also applied to ensure that candidate indicators had appropriate empirical properties.

4. A sequence of preliminary SP surveys was then developed, grounded in information provided by steps one through three above. Multiple survey variants were tested, each including a different sets of measurement endpoints (indicators) for relevant assessment endpoints. These surveys were evaluated and pretested in a second round of focus groups

---

[8] For example, an identical indicator of species abundance may vary in interpretation depending on whether that species is endemic or invasive within a particular region.

[9] As discussed by Johnston and Russell (2011), an ecological outcome can in principle affect utility both directly as well as indirectly through its contribution to the production of other final outcomes. In this case, the outcome would be considered a final good or service that has an additional intermediate effect on other final goods or services.

and interviews. Within these evaluations, direct questions assessed whether respondents perceived alternative indicators as communicating identical or distinct underlying outcomes (e.g., whether each indicator provided unique information concerning valued goods and services, or whether multiple indicators provided the same information in different ways). When indicators were perceived as communicating identical outcomes, subsequent questions were used to select indicators most easily understood by respondents. Following Schiller et al. (2001) and Johnston et al. (1995), focus groups and interviews were also used to identify the shared, common language best able to communicate indicators.

This four-step process provided justification for the set of indicators included in final survey scenarios, providing evidence (albeit qualitative) of outcome equivalence. Although this process provided unambiguous guidance regarding most indicators, it also revealed a case in which two indicators (i.e., of effects on migratory fish) appeared to be outcome equivalent, and in which respondents did not express a clear preference for either indicator. The final sections of this paper evaluate whether welfare estimates are robust to the alternative use of these two seemingly outcome equivalent indicators within survey scenarios.

4 Case Study of Outcome Equivalence and Welfare Estimation

We address the empirical question of whether SP welfare estimates are robust to choices between putatively commodity-equivalent indicators using a case study of migratory fish restoration in the Pawtuxet Watershed of Rhode Island, USA. At the time of this study the watershed provided no spawning habitat for migratory fish; access to all 4,347 acres of potential habitat was blocked by 22 dams (Erkan 2002).[10] Restoration of fish passage would not only affect fish populations but also other ecosystem outcomes that rely on the presence or abundance of migratory fish. Species that directly benefit from fish passage restoration in this area are alewife (*Alosa pseudoharengus*), blueback herring (*A. aestivalis*), shad (*A. sapidissima*), and American eel (*Anguilla rostrata*). The choice experiment questionnaire (*Rhode Island Rivers: Migratory Fishes and Dams*) estimated the WTP of Rhode Island residents for options that would restore fish passage to between 225 and 900 acres of historical habitat.

Description of the choice experiment is condensed from Johnston et al. (2011); Johnston et al. (2012). The theoretical model was adapted from a standard choice experiment specification in which household h chooses among three policy plans, ($j = A, B, N$), including two multi-attribute restoration options (A, B) and a status quo (N) of no restoration. Choice scenarios and restoration options were informed by data and restoration priorities in the *Strategic Plan for the Restoration of Anadromous Fishes to Rhode Island Coastal Streams* (Erkan 2002). Consistent with the strategic plan, the restoration methods included fish ladders and lifts (Schilt 2007) that neither require dam removal nor would cause appreciable changes in river flows. Information on the ecological roles of the migratory species targeted for restoration (e.g., Loesch 1987) provided the basis for conceptual models linking restoration to valued outcomes identified in focus groups.

The questionnaire was developed and tested over $2\frac{1}{2}$ years in a collaborative process involving economists and ecologists. This included meetings with resource managers, natural scientists, and stakeholder groups, and 12 focus groups. We conducted cognitive interviews (Kaplowitz et al. 2004), verbal protocols (Schkade and Payne 1994) and other pretests, both

---

[10] The first dam at the mouth of the Pawtuxet was removed in August 2011, three years after data collection was completed for this case study.

to collect information prior to survey design and to gain insight into respondents' interpretation of the questionnaire. Survey language and graphics were pretested carefully to ensure respondent comprehension. Particular attention was given to the definition and interpretation of ecological indicators. Prior to presenting choice questions, the survey provided information (1) describing the status of Rhode Island river ecology and migratory fish compared to historical baselines, (2) characterizing affected ecological systems and linkages, (3) describing fish passage restoration, and (4) providing definitions, derivations and interpretations of ecological indicators used in survey scenarios. Information was conveyed via a combination of text, graphics including Geographic Information System (GIS) maps and ecosystem representations, and photographs, all of which were subject to in-depth pretesting. The survey included a number of elements to promote incentive compatibility, including an emphasis on consequentiality, hypothetically binding payments, and instructions to consider each question as an independent choice.[11]

4.1 Ecological Indicators of Fish Passage Outcomes

Based on the process of indicator selection outlined above, choice options were characterized by five ecological indicators, one attribute characterizing public access, and one attribute characterizing annual household cost (Table 1). As described above, ecological indicators serving as attributes in the choice model were selected based on a conceptual model that coordinated ecological science with findings from focus groups. The initial direct ecological effect of restoration is to provide migratory fish with access to additional habitat for spawning and is quantified by the attribute *acres*, based upon restorable Pawtuxet watershed habitat acreage (Erkan 2002). The consequences of greater habitat acreage include increases in both migrating populations and the probability that fish runs will exist in a given area at some future time period. Alternative attributes used to characterize these impacts are the focus of this analysis; these are discussed below. Other indirect impacts include effects on (1) the abundance of non-migratory fish suitable for recreational harvest (*catch*), reflecting abundance measures from statewide sampling; (2) the abundance of fish-dependent wildlife (*wildlife*), reflecting the appearance of identifiable species within restored areas; and (3) overall ecological condition (*IBI*), reflecting the output of a multimetric aquatic ecological condition score [i.e., index of biotic integrity; see Johnston et al. (2011) and Johnston et al. (in press) for additional discussion].

The issue of outcome equivalence arose in survey design when evaluating alternative ecological indicators that could be used to characterize the impacts of restoration on migratory fish. Focus groups revealed that respondents valued, and were willing to pay for enhancements to migratory fish species in the Pawtuxet Watershed, *ceteris paribus*. The specific indicator(s) best suited to quantify this effect, however, were not clear. The ecological literature provides multiple indicators that might be used to quantify effects on migratory fish populations. One indicator is the estimated probability that the restored fish run will exist in 50 years (or an alternative time period), reflecting results calculable through applications of population viability analysis (Lee and Rieman 1997). Another indicator is the number of individual fish passing through fishways (i.e., the channels through which fish pass over or around dams), observed using electronic or visual counts of migrating fish (Erkan 2002). These alternative attributes are entitled *PVA* and *migrants*, respectively (Table 1).

---

[11] As noted by (Vossler et al. 2012, p. 168) the theoretical conditions for incentive compatibility in three-alternative choice experiments have not yet been established, and likely require "require strong restrictions on the utility functions … and beliefs about the preferences of others."

Table 1 Choice experiment attributes and descriptive statistics

Variable	Definition	Pawtuxet *PVA* choice experiment Mean (std. dev.)[a]	Pawtuxet *Migrants* choice experiment Mean (std. dev.)[a]
acres	The number of acres of river habitat accessible to migratory fish, presented as a percentage of the established reference value for the Pawtuxet watershed (Erkan 2002). Range 0–100 %	8.18 (8.16)	8.09 (8.10)
PVA	Population viability analysis (PVA) score. This was described to respondents as the "the probability (in percentage terms) that migratory species will still migrate the river in 50 years, as calculated by scientists." The reference condition is estimated based from surveys of experts in fish restoration, and interpreted following standard mechanisms for PVA models. Range 0–100 %	33.44 (28.12)	–
migrants	Expected number of adult fish that will swim upstream each year, past the first dam on the Pawtuxet River. Presented as a percentage of the reference value for the watershed, calculated using methods in Erkan (2002). This attribute was communicated to respondents via the following text: "expected number of fish that will swim upstream each year. The score (0–100 %) shows the number of fish, as a percent of the maximum possible in the river." Range 0–100 %	–	14.80 (13.18)
catch	The number of catchable-size fish in restored areas, estimated from the number of fish per hour caught by scientific sampling crews. Presented as a percentage of the reference value for the watershed, defined as the highest average level sampled in any Rhode Island river (from Rhode Island Department of Environmental Management sampling data). Range 0–100 %	79.91 (7.58)	79.89 (7.61)

Table 1 continued

wildlife	Number of fish-eating species that are common in restored areas, such as egrets, osprey, otters, eagles, turtles and mink. Presented as a percentage of the reference value for the watershed, quantified from surveys of regional experts in wildlife biology. Range 0–100%	65.01 (10.39)	65.05 (10.43)
IBI	Index of biotic integrity (IBI) score: a linear multimetric index of aquatic ecological condition following Karr (1981), reflecting the similarity of the restored area to the most undisturbed watershed area in Rhode Island. Index components include overall fish abundance, number of mussel species, number of native fish species, number of sensitive fish species, number of feeding types in fishes, percentage of individual fish that are native, percentage of individual fish that are migratory, and percentage of individual fish that are tumor free. Presented as a percentage of the reference condition. Range 0–100%	71.69 (6.08)	71.73 (6.11)
access	Binary (dummy) variable indicating whether the restored area is accessible to the public for walking and fishing; a value of 1 indicates that the public can access the area. Range 0–1	0.33 (0.47)	0.33 (0.47)
cost	Household annual cost, described as the mandatory increase in annual taxes and fees required to implement the restoration plan. Household cost for the status quo is zero. Range 0–25	11.98 (14.10)	12.07 (14.24)
neither	Alternative specific constant (ASC) associated with the status quo, or a choice of neither plan	0.33 (0.47)	0.33 (0.47)
d_mig	Binary (dummy) variable identifying observations from the choice experiment including *migrants* to represent effects on migratory fish. A value of 1 identifies observations from the *migrants* choice experiment; a value of 0 identifies observations from the *PVA* choice experiment	0.00 (0.00)	1.00 (0.00)

[a] Means and standard deviations include status quo option of no restoration

Survey pretests with both individuals and focus groups[12] supported the inclusion of either *PVA* and *migrants* in survey scenarios, but not both. This preference among respondents was due to two factors. First, *PVA* and *migrants* were seen to provide similar information (i.e., effects on migratory fish). Second, choice questions including both of these indicators were more difficult to answer due to the increased number of attributes. While expressing a preference for surveys including either *PVA* or *migrants*, however, respondents did not have a clear preference for one of these indicators over the other. Both were viewed as satisfactory ways to communicate effects on migratory fish, and focus group pretests suggested little effect on the choice frame when these indicators were alternated within the survey design.

Given these pretest results, we implemented two independent versions of the final survey. These were identical in all regards, except for the indicator used to characterize effects on migratory fish and associated explanatory text. One survey version included *PVA*, and the other included *migrants*. That is, each respondent received only one version of the survey, presenting information on either *PVA* or *migrants*, but not both. This enables a direct test of outcome equivalence—whether results are unaffected by this choice of indicators, both of which were deemed satisfactory by survey pretests. To be clear, there is no expectation that implicit prices (marginal WTP for attribute changes, *ceteris paribus*) will be equal for *PVA* or *migrants*, even in the presence of outcome equivalence. This is because the function linking these attributes to the putative underlying outcome (migratory fish) differs [see Eqs. (2) and (3)]. Rather, hypothesis tests consider a more comprehensive set of results from the choice experiment. These include evaluations of (1) potential differences in implicit prices for other model attributes, (2) compensating surplus (CS) measures for restoration programs considered as a whole, and (3) welfare measures calibrated using known ecological relationships between the two indicators.

4.2 Choice Experiment Implementation

Attribute levels within the experimental design (Table 2) were grounded in feasible restoration outcomes identified by ecological models, field studies and expert consultations. Choice scenarios represented each ecological attribute in relative terms with regard to upper and lower reference conditions (i.e., best and worst possible in the Pawtuxet) as defined in survey materials. Relative scores represented percent progress toward the upper reference condition (100 %), starting from the lower reference condition (0 %). Scenarios also presented the cardinal basis for relative scores where applicable. Sample choice questions are illustrated in Fig. 1 for the survey version including *migrants*, and by Fig. 2 for the survey version including *PVA*.

A fractional factorial experimental design was generated using a criterion that minimized D-error for a choice model covariance matrix with both main effects and selected two-way interactions (Kuhfeld 2010; Kuhfeld and Tobias 2005).[13] The final design included 180 profiles blocked into 60 booklets. Each respondent was provided with three choice questions and instructed to consider each as an independent, non-additive choice. Surveys were implemented using a dual wave phone-mail approach during June–August, 2008. An initial random digit dial sample of Rhode Island households was contacted via telephone and asked to participate in a survey addressing Rhode Island "environmental issues and

---

[12] These pretests took place during summer 2007.

[13] The design stacked a 72 profile, 100 % D-efficient design for main effects with a 108 profile design for main effects and selected two-way interactions. Constraints on the latter design prevented calculation of D-efficiency on a relative 100-point scale (Kuhfeld 2010).

Variable	Levels
acres	0 % (0 acres accessible to fish)[a]
	5 % (225 acres accessible to fish)
	10 % (450 acres accessible to fish)
	20 % (900 acres accessible to fish)
PVA	0 % (probability of 50 year fish run survival)[a]
	30 % (probability of 50 year fish run survival)
	50 % (probability of 50 year fish run survival)
	70 % (probability of 50 year fish run survival)
migrants	0 % (0 fish out of 1.2 million possible)[a]
	12 % (150,000 fish out of 1.2 million possible)
	20 % (245,000 fish out of 1.2 million possible)
	33 % (395,000 fish out of 1.2 million possible)
catch	70 % (102 fish/hour sampling abundance)
	80 % (116 fish/hour sampling abundance)[a]
	90 % (130 fish/hour sampling abundance)
wildlife	55 % (20 species common)[a]
	60 % (22 species common)
	70 % (25 species common)
	80 % (28 species common)
IBI	65 % (aquatic ecological condition score)[a]
	70 % (aquatic ecological condition score)
	75 % (aquatic ecological condition score)
	80 % (aquatic ecological condition score)
access	Public cannot walk and fish in area[a]
	Public can walk and fish in area
cost	$0 (cost to household per year)[a]
	$5 (cost to household per year)
	$10 (cost to household per year)
	$15 (cost to household per year)
	$20 (cost to household per year)
	$25 (cost to household per year)

Table 2 Attribute levels in choice experiment design

[a] Status quo value

government programs." Those agreeing to participate were sent the questionnaire via mail, with follow-up mailings to increase response rates (Dillman 2000). Respondents were either sent the survey version including *migrants* or that including *PVA*, but not both (i.e., the samples are independent). For the two survey versions analyzed here, a total of 1,200 questionnaires were sent to Rhode Island residents. These yielded 564 usable returns (47 %), providing provide 1,634 completed responses to choice questions. Response rates were similar across both survey versions.

4.3 Model Specification and Estimation

The random utility models are estimated using simulated likelihood mixed logit (ML) with Halton draws, accounting for correlations in choices from the same respondent. The final model specifications were chosen after the estimation of preliminary models with varying specifications of fixed and random coefficients. Within the final models, coefficients on all non-cost attributes except *catch* are specified as random with a normal distribution.[14] The coefficient on (sign-reversed) *cost* is random with a bounded triangular distribution, ensuring positive marginal utility of income (Hensher and Greene 2003). All variables except *access*

---

[14] In the unrestricted model, we fail to reject the null hypothesis that the standard deviation on this coefficient is zero ($p = 0.52$). Parallel results hold for all estimated models.

Fig. 1 Sample choice experiment question (survey version including *migrants*)

and *cost* reflect represent percent progress towards the upper reference condition (Table 1), so that the coefficient on ecological attribute k, $\hat{\beta}_k$, reflects the relative marginal utility given to a one percentage point change in that attribute.[15]

---

[15] Because choice questions provided the cardinal basis for these percentages, marginal utilities may also presented in cardinal terms, e.g., per acre, per fish, etc.

Question 6. **Projects A** and **B** are possible restoration projects for the Pawtuxet River, and the **Current Situation** is the status quo with no restoration. Given a choice between the three, how would you vote?
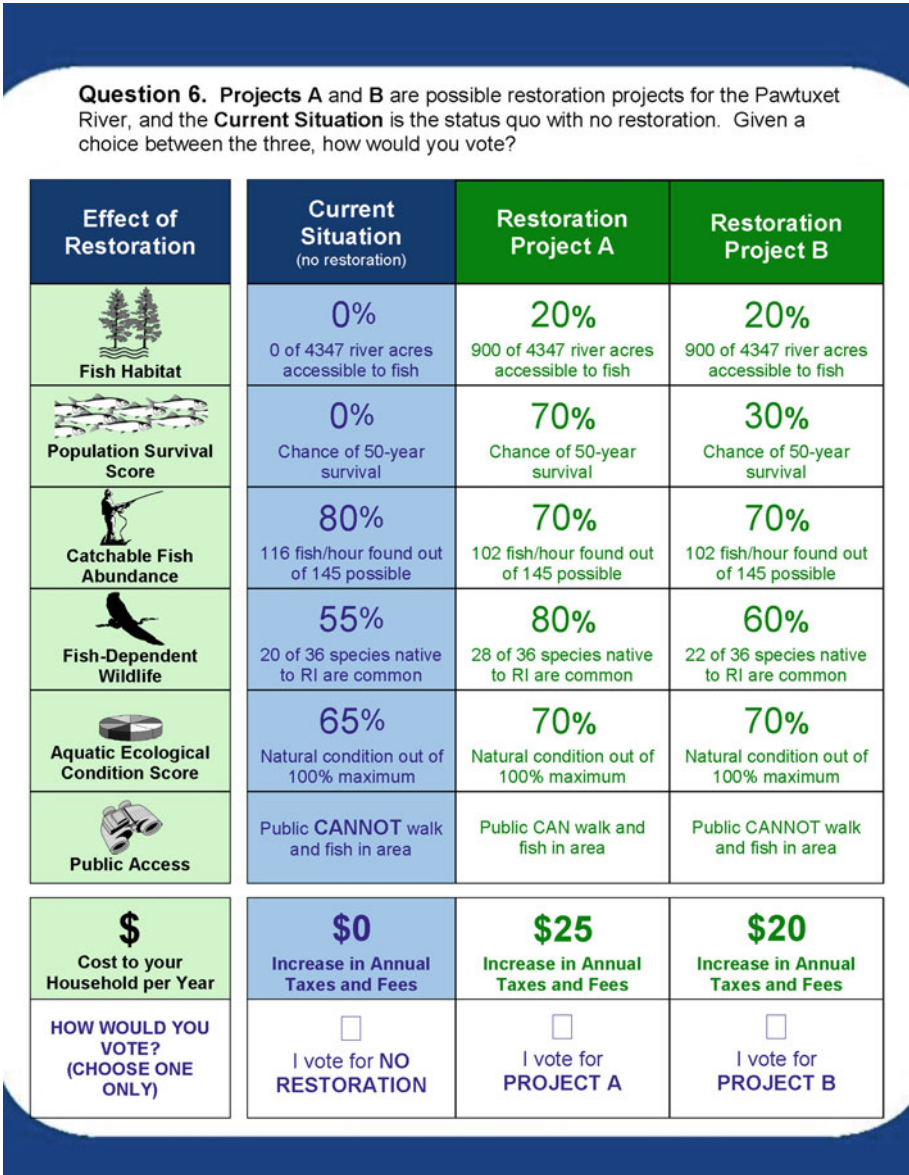
Effect of Restoration	Current Situation (no restoration)	Restoration Project A	Restoration Project B
Fish Habitat	**0%** 0 of 4347 river acres accessible to fish	**20%** 900 of 4347 river acres accessible to fish	**20%** 900 of 4347 river acres accessible to fish
Population Survival Score	**0%** Chance of 50-year survival	**70%** Chance of 50-year survival	**30%** Chance of 50-year survival
Catchable Fish Abundance	**80%** 116 fish/hour found out of 145 possible	**70%** 102 fish/hour found out of 145 possible	**70%** 102 fish/hour found out of 145 possible
Fish-Dependent Wildlife	**55%** 20 of 36 species native to RI are common	**80%** 28 of 36 species native to RI are common	**60%** 22 of 36 species native to RI are common
Aquatic Ecological Condition Score	**65%** Natural condition out of 100% maximum	**70%** Natural condition out of 100% maximum	**70%** Natural condition out of 100% maximum
Public Access	Public **CANNOT** walk and fish in area	Public **CAN** walk and fish in area	Public **CANNOT** walk and fish in area
$ **Cost to your Household per Year**	**$0** Increase in Annual Taxes and Fees	**$25** Increase in Annual Taxes and Fees	**$20** Increase in Annual Taxes and Fees
HOW WOULD YOU VOTE? (CHOOSE ONE ONLY)	☐ I vote for **NO RESTORATION**	☐ I vote for **PROJECT A**	☐ I vote for **PROJECT B**

Fig. 2 Sample choice experiment question (survey version including *PVA*)

Two final, jointly estimated models are illustrated; these are referenced by subscripts *p* (pooled) and *u* (unrestricted).[16] The first model pools observations from the two choice experiments, imposing identical coefficient estimates over all attributes, including the same coefficient on *PVA* and *migrants*. This is accomplished through the creation of a single variable, *fish*, that pools observations on *PVA* and *migrants* across the two choice experiments

---

[16] None of the models estimated in this paper have been published previously. The data from the *PVA* choice experiment have been used as the basis for distinct models published in Johnston et al. (2011, 2012, in press). The data from the *migrants* choice experiment have not been used in any previously published analysis.

(Table 1). The pooled model thereby ignores differences in the relative scale of change across the two fish attributes,[17] imposing an identical coefficient and marginal utility ($\hat{\beta}_{fish,p}$). Other attributes are identical across the two choice experiments and are pooled directly.

The second model is an unrestricted model that allows systematically varying coefficient estimates on all non-cost attributes between the *migrants* and *PVA* choice experiments. This is accomplished through inclusion of multiplicative interactions between each attribute and a dummy variable *d_mig* that identifies observations from the *migrants* choice experiment.[18] Hence, within the unrestricted model, the marginal utility of attribute k in the *PVA* choice experiment is given by $\hat{\beta}_{k,u}$, whereas the marginal utility of the same attribute in the *migrants* choice experiment is given by $(\hat{\beta}_{k,u} + \hat{\beta}_{k \times d\_mig,u})$. The marginal utility of *PVA* itself is given by $\hat{\beta}_{fish,u}$, and the marginal utility of *migrants* is given by $(\hat{\beta}_{fish,u} + \hat{\beta}_{fish \times d\_mig,u})$.

5 Results

Results for the two models are reported in Table 3. Both models are statistically significant at $p < 0.0001$, with pseudo-R^2 statistics in excess of 0.31. Coefficient estimates on all ecological attributes, with the exception of *catch*, are statistically significant at $p < 0.01$ in all models. Signs of statistically significant coefficients match prior expectations. Prior works discuss the general properties and policy relevance of similar model results (Johnston et al. 2011; Johnston et al. 2012); these topics are not repeated here. Rather, we focus on an evaluation of outcome equivalence and implications for policy analysis.

A likelihood ratio test comparing the unrestricted and pooled models fails to show the statistical significance of the imposed restrictions ($\chi^2 = 4.51, df = 7, p = 0.72$). As a result, we cannot reject the pooled model that imposes identical coefficient estimates across the two choice experiments.[19] Similarly, none of the individual coefficient estimates distinguishing the two models ($\hat{\beta}_{k \times d\_mig,u}$) are statistically significant at $p < 0.10$. Hence, at first glance, the choice between *migrants* and *PVA* to represent impacts on migratory fish does not appear to influence model results. Such findings provide initial support for an outcome equivalence hypothesis, and are robust across a wide range of alternative model specifications (including models with alternative specifications and distributions of fixed and random parameters).

As implied by Eqs. (2) and (3) above, however, direct comparison of the coefficient estimates on *migrants* and *PVA* is not necessarily informative; there is no *a priori* reason that these estimates should be identical, even if the two attributes are outcome equivalent. Moreover, the potentially confounding role of the logit scale parameter prevents naïve comparison of individual coefficient estimates across models (Swait and Louviere 1993). Due to these limitations, we continue the analysis through an evaluation of implicit prices and CS estimates

---

[17] Based on guidance from experts in fish restoration and ecology, the maximum feasible level of migrants resulting from fish restoration was assessed to be 33 % of the reference condition for the watershed, whereas the maximum feasible level of PVA was assessed to be 70 % (both out of 100 % maximum, see Tables 1, 2). Hence, while both attributes represent effects on migratory fish, they do so at different scales.

[18] We do not allow the marginal utility of income (i.e., coefficient on cost) to vary between the *PVA* and *migrants* choice experiments. This is because (i) the two choice experiments sample from the same population, (ii) the use of alternative ecological indicators to characterize effects on fish should not affect the marginal utility of income, and (iii) desirable properties of the triangular distribution that apply to $\hat{\beta}_{cost,u}$ would not apply to the sum ($\hat{\beta}_{cost,u} + \hat{\beta}_{cost,u \times d\_mig,u}$), confounding the interpretation of coefficient estimates.

[19] Results are identical if one also allows the coefficient on *cost* to vary between the *PVA* and *migrants* observations.

Table 3 Mixed logit results: Pawtuxet restoration choice experiments

Choice attribute	Pooled model coefficient (std. error)	Unrestricted model coefficient (std. error)
Random parameters		
acres	0.0467*** (0.0085)	0.0463*** (0.0117)
fish (*PVA* and *migrants* pooled)	0.0193*** (0.0039)	0.0169*** (0.0043)
IBI	0.0325*** (0.0120)	0.0497*** (0.0168)
access	1.2809*** (0.1602)	1.1577*** (0.2056)
wildlife	0.0248*** (0.0059)	0.0267*** (0.0083)
neither	−4.3950*** (0.4522)	−4.2235*** (0.4522)
cost (bounded triangular, sign-reversed)	0.0537*** (0.0058)	0.0533*** (0.0058)
Non-random parameters		
catch	−0.0013 (0.0057)	0.0011 (0.0082)
acres × *d_mig*	–	0.0010 (0.0161)
fish × *d_mig*	–	0.0093 (0.0087)
IBI × *d_mig*	–	–0.0345 (0.0229)
access × *d_mig*	–	0.2170 (0.2643)
wildlife × *d_mig*	–	−0.0038 (0.0113)
neither × *d_mig*	–	−0.1865 (0.8233)
catch × *d_mig*	–	−0.0052 (0.0114)
Random parameter distributions		
std. dev. acres	0.0660*** (0.0211)	0.0679*** (0.0216)
std. dev. fish	0.0168* (0.0091)	0.0154 (0.0115)
std. dev. IBI	0.0869*** (0.0277)	0.0816*** (0.0294)
std. dev. access	1.6233*** (0.2478)	1.5873*** (0.2544)
std. dev. wildlife	0.0167 (0.0236)	0.0174 (0.0257)
std. dev. neither	4.8737*** (0.7390)	4.8330*** (0.7627)
spread cost (bounded triangular)	0.0537*** (0.0058)	0.0533*** (0.0058)
−2 Log likelihood χ^2	1,122.75***	1,127.26***
Pseudo-R^2	0.31	0.31
Observations (N)	1,634	1,634

* $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$

(which are more directly relevant for policy and not confounded by scale), combined with welfare comparisons that calibrate for differences in ecological scale between *migrants* and *PVA*.

5.1 Test of Outcome Equivalence: Implicit Prices

Drawing from unrestricted model coefficients, implicit prices are estimated using the welfare simulation described by Johnston and Duke (2007), following the general approach of Hensher and Greene (2003).[20] Presented WTP estimates reflect the mean over the parameter

---

[20] The procedure begins with a parameter simulation following the parametric bootstrap of Krinsky and Robb (1986), with $R = 1,000$ draws taken from the mean parameter vector and covariance matrix. For each draw, the

Table 4 Implicit price differences: *PVA* and *migrants* choice experiments

Choice attribute	Implicit price estimates[a]		Implicit price difference[a]
	PVA choice experiment	*Migrants* choice experiment	
acres	1.2016*** (0.3435)	1.2287*** (0.3176)	0.0271 (0.4283)
PVA	0.4422*** (0.1168)	–	–
migrants	–	0.6929*** (0.2133)	–
access	30.0916*** (6.3454)	36.0333*** (6.4621)	5.9416 (7.0246)
IBI	1.3111*** (0.4437)	0.3952 (0.4427)	−0.9159 (0.5876)
catch	0.0396 (0.2137)	−0.1031 (0.2060)	−0.1427 (0.2904)
wildlife	0.6904*** (0.2204)	0.5931*** (0.2192)	0.0972 (0.2932)

[a] Mean values over empirical distributions; standard errors are in parentheses

* $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$

simulation of mean WTP calculated over the coefficient simulation. Given the model specification, implicit prices are calculated for the *PVA* choice experiment using the general form $\hat{\beta}_{k,u}/\hat{\beta}_{cost,u}$, and for the migrants choice experiment using $(\hat{\beta}_{k,u} + \hat{\beta}_{k \times d\_mig,u})/\hat{\beta}_{cost,u}$.

Results are shown in Table 4. For all attributes except *access*, implicit price results are interpreted as WTP for a marginal, one percentage point increase in the attribute, holding all else constant. For *access*, results indicate WTP for the provision of public access in the restored area, relative to the default of no access. The rightmost column of Table 4 illustrates the difference between implicit prices estimated from the *PVA* and *migrant* models. We calculate statistical significance levels (*p* values for two-tailed tests) for each of the implicit prices and differences. Significance levels are determined through percentiles on the empirical welfare distributions (Poe et al. 2005), with these distributions accounting both for sampling variation reflected in the estimated covariance matrix for model parameters and the estimated distribution of random coefficients (Hensher and Greene 2003).

Results in Table 4 demonstrate that (a) implicit prices vary across distinct ecological effects within each choice experiment,[21] but (b) corresponding implicit prices are statistically indistinguishable across the two choice experiments. For all directly comparable implicit prices across the two choice experiments, we fail to reject the null hypothesis of equivalent implicit prices.

Comparisons involving implicit prices for the attributes *PVA* and *migrants* themselves require additional ecological assumptions. The relationship between the number of migratory fish (*migrants*) in the Pawtuxet watershed and population viability (*PVA*) is conditional on other ecological factors, such as conditions at sea for this migratory species. This caveat aside, attribute levels in the experimental design are conditioned upon an expected 2.37 to 1 ratio, on average, between improvements in population viability and the number of migratory

---

Footnote 20 continued

resulting parameters are used to characterize asymptotically normal empirical densities for fixed and random coefficients. For each of these *R* draws, a coefficient simulation is conducted for each random coefficient, with $S = 1,000$ draws taken from simulated empirical densities (either normal or bounded triangular, depending on the distribution for each coefficient). Welfare measures are calculated for each draw, resulting in a combined empirical distribution of $R \times S$ observations from which summary statistics are derived.

[21] For example, within the model including *PVA*, the implicit price for *acres* differs from the implicit price for *wildlife*. This shows the respondents differentiated across ecological outcomes within each independent choice experiment.

fish.[22] This expectation reflects a monotonically additive relationship (Maunder 2004), based on ecological data for the watershed and consultations with experts in regional migratory fish restoration.

Based on this relationship, and assuming that these indicators are outcome equivalent, one would expect a corresponding but inverse ratio in the implicit prices, with respondents willing to pay approximately 2.37 times more for marginal changes in *migrants* than for parallel changes in *PVA*, in percentage point terms. That is, if these two indicators reflect alternative ways of measuring underlying effects on migratory fish, albeit at different scales, one would expect the WTP ratio for these attributes to be the inverse of the ecological ratio. Empirical results cannot reject this expectation. Based on a comparison of empirical distributions, we fail to reject the null hypothesis ($p = 0.28$) that the implicit price ratio between *migrants* and *PVA* is equal to 2.37.

5.2 Test of Outcome Equivalence: Compensating Surplus

Compensating surplus (CS) measures reflect per household WTP for entire restoration programs that combine outcomes for different attributes. As cost benefit analyses commonly rely on these measures, they arguably provide the most relevant perspective on WTP robustness (Morrison et al. 2002). To assess CS differences, we follow common practice and select a small set of illustrative policy alternatives from thousands of feasible scenarios that can be generated using attribute levels of the experimental design. We choose three illustrative scenarios representing: (1) the smallest positive increase in each ecological outcome; (2) a median increase in each outcome; and, (3) the largest increase in each ecological outcome. CS for each of these three alternatives is calculated both with and without public access, for a total of six policy scenarios. Following Morrison et al. (2002)), all CS estimates incorporate (the negative of) the status quo ASC. When specifying comparable levels of *migrants* and *PVA* across the two models, we assume the same ecological relationship between these attributes that underpins attribute levels in the experimental design. This enables a comparison of CS for equivalent policies across the two models. Empirical distributions are simulated following the a parallel approach to that used for implicit prices above.

Results are shown in Table 5. In addition to the attribute levels defining each policy scenario, Table 5 presents estimated CS for both the *PVA* and *migrants* choice experiments, as well as the difference between these estimates both in cardinal (dollar) and absolute value percentage terms. In all cases, we fail to reject the null hypothesis of zero CS difference. Absolute value percentage differences vary from 1.89 to 7.83 %, with a mean of 4.71 %, indicating nearly identical CS estimates across the two choice experiments. These findings again suggest statistical robustness across the two independent choice experiments, and support the hypothesis of outcome equivalence. Average CS estimates do not vary significantly when alternating between indicators of population size (*migrants*) and viability (*PVA*) within the choice experiment, suggesting that respondents viewed and valued policy scenarios similarly regardless of which indicator was used to characterize direct

---

[22] The exact ratios implied by the experimental design range from 2.12 to 2.50, with the ratio declining at higher levels of both *PVA* and *migrants*. This nonlinearity is approximated by a piecewise linear relationship reflected in the experimental design, and reflects the fact that other exogenous conditions in the Pawtuxet watershed (e.g., water quality) are assumed to prevent the expected 50-year *PVA* score from approaching 100 %, even if the number of migratory fish is relatively high within the next five to ten years. Final attribute levels in the survey were also subject to rounding in response to focus group input.

Table 5 Compensating surplus (CS) estimates and differences: Pawtuxet restoration scenarios

Scenario	Description	Attribute levels defining each scenario								Compensating surplus (CS) estimates		Compensating surplus (CS) differences	
		neither	acres	migrants	PVA	access	IBI	catch	wild	CS migrants choice experiment[a]	CS PVA choice experiment[a]	CS Difference[a]	CS difference, absolute value percent
1	Smallest ecological improvement; no access	−1	5	12	30	0	5	0	5	147.47*** (28.05)	140.09*** (31.45)	7.37 (19.81)	5.26
2	Median ecological improvement; no access	−1	10	20	50	0	10	10	15	174.35*** (29.62)	168.80*** (32.23)	5.55 (20.20)	3.29
3	Largest ecological improvement; no access	−1	20	33	70	0	15	20	25	207.37*** (32.48)	203.52*** (34.21)	3.85 (23.13)	1.89
4	Smallest ecological improvement; access	−1	5	12	30	1	5	0	5	183.50*** (29.96)	170.19*** (33.26)	13.32 (20.58)	7.83
5	Median ecological improvement; access	−1	10	20	50	1	10	10	15	210.38*** (31.80)	198.89*** (34.23)	11.49 (21.17)	5.78
6	Largest ecological improvement; access	−1	20	33	70	1	15	20	25	243.40*** (34.86)	233.61*** (36.35)	9.80 (24.16)	4.20
Mean												8.56	4.71

[a] Mean values over empirical distributions; standard errors are in parentheses * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$

effects on migratory fish. This occurs despite the fact that the two indicators are measured at different scales, with *PVA* effects ranging from 30 to 70 %, and *migrant* effects ranging from 12 to 33 % (approximately 150,000–395,000 fish) from equivalent restoration actions.

6 Conclusion

Compared to prior works in the ecological valuation literature, this analysis promotes a more formal perspective on ecological indicator selection within SP design, specifically invoking the concept of outcome equivalence to characterize cases in which ecological indicators provide alternative means to communicate an identical underlying outcome. The robustness of welfare estimates to ecological indicators is critically important for the use of SP analysis to guide policy: if not robust, welfare estimates could be regarded (or disregarded) as idiosyncratic to the particular set of indicators chosen to communicate a policy outcome. Indeed, when indicators are not outcome equivalent welfare estimates are unlikely to converge. Previous findings that WTP differs when ecological attributes differ are not particularly surprising given that these studies employed indicators that are not *prima facie* outcome equivalent.

We demonstrate that welfare estimates converge when outcome equivalent indicators are used. In contrast to previous studies, the present analysis evaluates welfare implications using independent choice experiments in which effects on an ecological outcome are characterized using two indicators that are expected to be outcome equivalent based on input during survey design. Our findings provide a promising message for SP analysis, and to our knowledge are the first published results demonstrating the convergent validity of SP welfare estimates to alternative uses of outcome equivalent ecological indicators. Welfare estimates are statistically robust whether one characterizes impacts in terms of the size of migrating fish populations (*migrants*) or population viability (*PVA*). We find no significant differences in either implicit prices or CS estimates. Moreover, implicit prices for *migrants* and *PVA* are consistent with the underlying ecological relationship between the two attributes implied by the experimental design. As a result, WTP for a given change in migratory fish—estimated using independent choice experiments—is similar whether the change is measured in terms of population size or viability. This suggests that respondents were able to use either ecological indicator to make equivalent assessments of the welfare gains from migratory fish restoration. If future research establishes similar findings elsewhere, the result should be greater confidence that SP welfare estimates are not overly sensitive to the particular set of ecological indicators used in the questionnaire.

This paper addresses only a few of the many challenges involved in comprehensive coordination of economics and ecology for welfare estimation. Although results here suggest that SP welfare estimates can be robust to the use of alternative ecological indicators within survey scenarios, greater attention is needed to relationships between indicators in survey scenarios and policy outcomes for which welfare estimates are desired. Despite an awareness of challenges involved in communicating ecological changes within SP surveys, the scrutiny given to indicator use and interpretation within the ecological literature has not been matched in the economics literature. The process used to select indicators within survey design is often opaque, and implications for welfare estimates remain obscured. The result can be a lack of transparency in both the ecological outcomes that respondents are being asked to value and ways in which these relate to ecological information presented on the survey page.

References

Bateman IJ, Carson RT, Day B, Hanemann M, Hanley N, Hett T, Jones-Lee M, Loomes G, Mourato S, Ozdemiroglu E, Pearce DW, Sugden R, Swanson J (2002) Economic valuation with stated preference surveys: a manual. Edward Elgar, Northampton, MA

Bortone SA (2005) Estuarine indicators. CRC Press Boca Raton, FL

Boyce MS (1992) Population viability analysis. Annu Rev Ecol Syst 23:481–506

Carson RT (1998) Valuation of tropical rainforests: philosophical and practical issues in the use of contingent valuation. Ecol Econ 24(1):15–29

Carson RT, Flores NE, Meade NF (2001) Contingent valuation: controversies and evidence. Environ Resour Econ 19(2):173–210

Carson RT, Mitchell RC (1995) Sequencing and nesting in contingent valuation surveys. J Environ Econ Manag 28(2):155–173

Christie M, Hanley N, Warren J, Murphy K, Wright R, Hyde T (2006) Valuing the diversity of biodiversity. Ecol Econ 58(2):304–317

Czajkowski M, Buszko-Briggs M, Hanley N (2009) Valuing changes in forest biodiversity. Ecol Econ 68(2):2910–2917

DeShazo JR, Fermo G (2002) Designing choice sets for stated preference methods: the effects of complexity on choice consistency. J Environ Econ Manag 44(1):123–143

Davis WS, Simon TP (eds) (1995) Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton

Dillman DA (2000) Mail and internet surveys: the tailored design method. Wiley, New York, NY

Erkan DE (2002) Strategic plan for the restoration of anadromous fishes to Rhode Island coastal streams. Rhode Island Department of Environmental Management, Division of Fish and Wildlife, Wakefield, RI

Hensher DA, Greene WH (2003) The mixed logit model: the state of practice. Transportation 30(2):133–176

Jackson LE, Kurtz JC, Fisher WS (eds) (2000) Evaluation guidelines for ecological indicators. EPA/620/R-99/005. US EPA, Washington, DC

Jacobsen JB, Boisen JH, Thorsen BJ, Strange N (2008) What's in a name? The use of quantitative measures versus 'iconised' species when valuing biodiversity. Environ Resour Econ 39:247–263

Jakobsson KM, Dragun AK (2001) The worth of a possum: valuing species with the contingent valuation method. Environ Resour Econ 19(3):211–227

Johnston RJ, Besedin EY, IOvanna R, Miller C, Wardwell R, Ranson M (2005) Systematic variation in willingness to pay for aquatic resource improvements and implications for benifit transfer: a meta-analysis. Can J Agric Econ 53(2-3):221–248

Johnston RJ, Duke JM (2007) Willingness to pay for agricultural land preservation and policy process attributes: does the method matter? Am J Agric Econ 89(4):1098–1115

Johnston RJ, Russell M (2011) An operational structure for clarity in ecosystem service values. Ecol Econ 70(12):2243–2249

Johnston RJ, Schultz ET, Segerson K, Besedin EY, Ramachandran M (2012) Enhancing the content validity of stated preference valuation: the structure and function of ecological indicators. Land Econ 88(1):102–120

Johnston RJ, Segerson K, Schultz ET, Besedin EY, Ramachandran M (2011) Indices of biotic integrity in stated preference valuation of aquatic ecosystem services. Ecol Econ 70(11):1946–1956

Johnston RJ, Schultz ET, Segerson K, Besedin EY, Ramachandran M (in press) Stated preferences for intermediate versus final ecosystem services: disentangling willingness to pay for omitted ecological outcomes. Agric Resour Econ Rev 42(1)

Johnston RJ, Weaver TF, Smith LA, Swallow SK (1995) Contingent valuation focus groups: insights from ethnographic interview techniques. Agric Resour Econ Rev 24(1):56–69

Jordan SJ, Smith LM (2005) Indicators of ecosystem integrity for estuaries. In: Bortone SA (ed) Estuarine indicators. CRC Press, Boca Raton, FL

Jørgensen SE, Constanza R, Xu F (2004) Handbook of ecological indicators for assessment of ecosystem health. CRC Press, Boca Raton, FL

Kaplowitz MD, Lupi F, Hoehn JP (2004) Multiple methods for developing and evaluating a stated-choice questionnaire to value wetlands. In: Presser S, Rothget JM, Coupter MP, Lesser JT, Martin E, Martin J, Singer E (eds) Methods for testing and evaluating survey questionnaires, Chapter 24. Wiley, New York

Karr JR (1981) Assessment of biotic integrity using fish communities. Fisheries 6(6):21–27

Krinsky I, Robb AL (1986) On approximating the statistical properties of elasticities. Rev Econ Stat 68(4):715–719

Kuhfeld WF (2010) Marketing research methods in SAS: experimental design, choice, conjoint, and graphical techniques. SAS Institute, Cary, NC

Kuhfeld WF, Tobias RD (2005) Large factorial designs for product engineering and marketing research applications. Technometrics 47:132–141

Lee DC, Rieman BE (1997) Population viability assessment of salmonids by using probabilistic networks. N Am J Fish Manag 17(4):1144–1157

Loesch JG (1987) Overview of life history aspects of anadromous alewife and blueback herring in freshwater habitats. In: Dadswell MJ, Klauda RJ, Moffitt CM, Saunders RL, Rulifson RA, Cooper JE (eds) International symposium on common strategies of anadromous and catadromous fishes. American Fisheries Society, Boston, MA USA, pp 97–103

Maunder MN (2004) Population viability analysis based on combining Bayesian, integrated, and hierarchical analyses. Acta Oecologica. Int J Ecol 26(1):85–94

Mitchell RC, Carson RT (1989) Using surveys to value public goods: the contingent valuation method. Resources for the future, Washington, DC

Morrison M, Bennett J, Blamey R, Louviere J (2002) Choice modeling and tests of benefit transfer. Am J Agric Econ 84(1):161–170

Naweedi MJ (2005) Environmental indicators as performance measures for improving estuarine environmental quality. In: Bortone SA (ed) Estuarine indicators. CRC Press, Boca Raton, FL

Niemi GJ, McDonald ME (2004) Application of ecological indicators. Annul Rev Ecol Syst 35(1):89–111

Nunes PALD, van den Bergh JCJM (2001) Economic valuation of biodiversity: sense or nonsense? Ecol Econ 39(2):203–222

Poe GL, Giraud KL, Loomis JB (2005) Computational methods for measuring the difference in empirical distributions. Am J Agric Econ 87(2):353–365

Powe NA, Bateman IJ (2004) Investigating insensitivity to scope: a split-sample test of perceived scheme realism. Land Econ 80(2):258–271

Powe NA (2007) Redesigning environmental valuation: mixing methods within stated preference techniques. Edward Edgar, Cheltenham, UK

Schiller A, Hunsaker CT, Kane MA, Wolfe AK, Dale VH, Suter GW, Russell CS, Pion G, Jensen MH, Konar VC (2001) Communicating ecological indicators to decision makers and the public. Conserv Ecol 5(1):19. http://www.consecol.org/vol5/iss1/art19/

Schilt CR (2007) Developing fish passage and protection at hydropower dams. Appl Animal Behav Sci 104:295–325

Schkade DA, Payne JW (1994) How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation. J Environ Econ Manag 26(1):88–109

Schultz ET, Johnston RJ, Segerson K, Besedin EY (2012) Integrating ecology and economics for restoration: using ecological indicators in valuation of ecosystem services. Restor Ecol 20(3):304–310

Spash CL, Hanley N (1995) Preferences, information and biodiversity preservation. Ecol Econ 12(3):191–208

Suter GW (1993) Ecological risk assessment. Lewis Publishers, Boca Raton, FL

Suter GW (2001) Applicability of indicator monitoring to ecological risk assessment. Ecol Indic 1:101–112

Swait J, Louviere JJ (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. J Market Res 30(3):305–314

Tonsor GT (2011) Consumer inferences of food safety and quality. Eur Rev Agric Econ 38(2):213–235

US Environmental Protection Agency (EPA) (1998) Guidelines for Ecological Risk Assessments. Risk Assessment Forum, Washington, DC. EPA/630/R-95/002F

Van Houtven G, Powers J, Pattanayak SK (2007) Valuing water quality improvements in the United States using meta-analysis: is the glass half-full or half-empty for national policy analysis? Resour Energy Econ 29(3):206–228

Veisten K, Hoen HF, Navrud S, Strand J (2004) Scope insensitivity in contingent valuation of complex environmental amenities. J Environ Manag 73:317–331

Vossler CA, Doyon M, Rondeau D (2012) Truth in consequentiality: theory and field evidence on discrete choice experiments. Am Econ J Microecon 4(4):145–171

Walsh P, Wheeler W (2011) Water quality index aggregation and cost benefit analysis. US EPA National Center for Environmental Economics working paper, Washington, DC