

# How are Scope and Adding up Relevant for Benefits Transfer?

Catherine L. Kling<sup>1</sup> · Daniel J. Phaneuf<sup>2</sup>

Accepted: 1 December 2017 / Published online: 4 January 2018  
© Springer Science+Business Media B.V., part of Springer Nature 2018

**Abstract** In this paper, we explore ways in which the theoretical constructs of scope and adding up can inform and improve the practice of benefit transfer. Specifically, we examine how the stated preference literature on scope and adding up can inform three critical steps in benefits transfer: study site selection, including studies to select for use in a meta-regression; calibrating benefit functions; and assessing transfer validity.

**Keywords** Nonmarket valuation · Benefits transfer · Benefit-cost analysis

## 1 Introduction

The economic benefits generated by environmental policy are often nonmarket in nature, meaning their measurement for cost-benefit analysis requires application of nonmarket valuation methods. Original studies of this type require a non-trivial commitment of resources and the skills of a specialist. This presents policy analysts with a twofold dilemma: short time-frames and institutional constraints can make execution of an original study infeasible, and many decisions are too limited in scope to justify expenditure on original research. In these cases, an alternative approach is needed to quantitatively measure benefits. Specifically, the

---

Guest Editor: V. Kerry Smith.

---

We appreciate insightful comments and corrections from V. Kerry Smith, Matt Kotchen, Chris Moore, and attendees at the EPA Benefit Transfer workshop held in Washington D.C., December 2016.

---

✉ Catherine L. Kling  
ckling@iastate.edu

✉ Daniel J. Phaneuf  
dphaneuf@wisc.edu

<sup>1</sup> Iowa State University, Ames, IA, USA

<sup>2</sup> University of Wisconsin-Madison, Madison, WI, USA

economic benefits from the policy action need to be measured using previous studies and other existing information. In environmental economics, this process is known as benefits transfer.

The art and science of benefits transfer arose from a set of pragmatic needs that existed prior to economists' efforts to carefully evaluate methods and define best practices. Systematic study of benefits transfer started with a 1992 symposium published in *Water Resources Research*. These papers began the process of defining relevant concepts, outlining methodological issues, suggesting best practice, and identifying validity criteria. The nomenclature that subsequently developed uses the term *policy site* to denote the place or resource for which a nonmarket value is needed, and *study site(s)* to denote places or resources that have been the subject of a past primary study, and that may be relevant for the measurement task at hand. A critical step in any benefits transfer exercise is to select study sites that are of high quality, and similar enough to the policy site to be informative about it.

Quality assessment for nonmarket valuation studies is an enormous area of research that has occupied environmental economists for decades. A number of validity criteria and best practice standards have arisen in conjunction with the commonly applied revealed and stated preference methods. For stated preference, the literature is organized around four validity concepts: criterion, convergent, construct, and content validity. Criterion validity assesses the extent to which an estimate matches a known benchmark, while convergent validity assesses the extent to which two different estimates of the same phenomena are the same. Construct validity gauges the consistency of empirical results with theoretical predictions, and content validity is related to the study's adherence to best practice standards. In the benefits transfer context, each of these is relevant for deciding if an estimate gleaned from a stated preference study should be included among the study sites in a particular transfer exercise.

Two construct validity concepts that grew out of the contingent valuation literature are scope tests and adding up tests. In the context of a stated preference study, a scope test assesses whether or not the willingness to pay for an environmental good is sensitive to its level of provision. For example, intuition suggests people should be willing to pay more to protect a collection of endangered species, than they would pay to protect any one member of that collection. If empirical predictions do not display this, then the study fails the scope test. As usually applied, the scope test is an existence criterion: a statistical test accepts or rejects sensitivity to scope, and the result determines whether the study passes or fails the scope test. In contrast, adding up tests assess whether or not the willingness to pay for the parts of a change, when properly defined and measured, add up to the willingness to pay for the whole of a change. As we discuss below, this is a more demanding standard, which has been less frequently applied than basic scope tests.

In this paper, we explore additional ways in which the theoretical constructs of scope and adding up can inform and improve the practice of benefit transfer. Specifically, we examine how the stated preference literature on scope and adding up can inform three critical steps in benefits transfer: study site selection, including studies to select for use in a meta-regression; calibrating benefit functions; and assessing transfer validity. In doing so, we build on a long line of inquiry by economists who have looked to theory to provide both formal and informal guidance in applied welfare economics.

We begin our discussion in Sect. 2 with a brief overview of how theory and economic intuition have informed the art and practice of applied welfare analysis generally. With the context set, we then define the concepts of scope and adding up as they have evolved in the literature, as well as their historical role in judging stated preference validity. We conclude the section with a description of how the scope and adding up concepts have appeared in the benefits transfer literature. In section three, we describe in detail how scope considerations can explicitly inform and improve the process of benefits transfer through the three routes

noted above: study selection, benefit function construction, and assessment of validity. For this we draw heavily on the concept of scope elasticity, attributable to Whitehead (2016). Discussion and suggestions for moving forward conclude the paper.

## 2 Concepts and Literature

### 2.1 Context

Applied economists working in the policy arena are able to make use of the discipline's theoretical foundations to gauge the plausibility of empirical predictions. Formal theories and informal reasoning based on economic intuition can often be used to set a benchmark for comparison, before an estimate is used to inform a policy decision. To set the context for our study of scope and adding up, we begin this section with a general discussion of the interplay in applied welfare analysis between empirical prediction, plausibility, and theoretical reasoning. We note that theory and economic intuition can provide guidance both at the individual welfare measure level and when aggregating across a population to arrive at a population estimate.

For market commodities, economists are often interested in the welfare impacts of decreased cost, improvements in quality, and innovations leading to new products. Information on prices, quantities, expenditures, and consumer demographics for the affected good and its substitutes provide the inputs for analysis, and tools for assessing plausibility. Consider, for example, processing speed in personal computers among non-business consumers. An inverse demand function regression can reveal the marginal willingness to pay for a faster computer, and multiplying this by the number of computers sold provides one estimate of aggregate value from the change. The plausibility of the estimate can then be assessed by comparing it to annual expenditures in the personal computer market, the number of households that own personal computers, and disposable incomes for the consumer population segment—all based on intuition from consumer choice theory. Alternatively, since a computer is a durable asset, the willingness to pay reflects the present value of the stream of services provided by the additional speed. Annualizing this by assuming a discount rate and computer lifetime allows comparison to rental prices for computers or other technology. Finally, the potential time saving from faster processing could be scaled by the wage rate, as a means of gauging value via saved time. While none of these steps can provide the estimate of interest, they suggest how different forms of economic reasoning can be used to triangulate around the preferred estimate, to avoid first order errors in analysis. Indeed, it seems likely that many a coding error has been rooted out by assessing the extent to which an estimate passes such 'laugh' tests.

For nonmarket commodities, it is usually not possible to assess plausibility by looking directly at prices and expenditures. However, it is possible to assemble information on prices and quantities for related goods. For example, predictions on changes in trip-taking behavior can be benchmarked using information from time use surveys—perhaps via an understanding of elasticities between different types of leisure time activities. Siikamaki (2011) provides an example of how assembling information on time use from multiple sources can be leveraged to understand the role of state parks in nature recreation. The clever use of auxiliary information as done by Siikamaki appears to be an underutilized strategy.

In other instances, theoretical results can provide a direct point of departure for gauging plausibility. Hanemann (1991) derives an expression for the income elasticity of marginal willingness to pay for an environmental good, showing that it depends on the degree of substitutability between market goods and the nonmarket commodity. His expression, in

conjunction with estimates of conventional income elasticities, can be used to gauge the plausibility of differences between willingness to accept and willingness to pay for nonmarket goods. An application of this is Horowitz and McConnell (2003), who examine income effects in the context of meta-data on the WTP/WTA gap. A second example is the logic that Smith (1992) uses to critique the embedding issue in contingent valuation put forth by Kahneman and Knetsch (1992). He uses an expression for the elasticity of marginal willingness to pay for an environmental commodity to argue that the concerns raised by Kahneman and Knetsch can be rationalized by consumer theory, and thus need not invalidate contingent valuation. A further example is the theoretical relationship between marginal and average values. Under the typical assumption of diminishing marginal utility, we know that marginal values should not exceed the average value. Finally, perhaps the consummate example of theory informing plausibility in applied welfare economics is Willig’s well-known bounds (Willig 1976), which provide a plausibility check on the magnitude of errors regarding the differences in Hicksian versus Marshallian welfare measures. In this paper, we suggest that scope can be used in a like manner to assess the adequacy of responsiveness in stated preference studies, and their suitability for use in benefits transfer.

A final context point concerns the role that agency rule making processes may have for the theory and practice of benefit cost analysis and benefits transfer. Many regulations are incremental, in the sense that environmental improvements are realized slowly over time—perhaps due to phased implementation of a series of discrete interventions. In this case there are two possible ways to carry out a benefits study: examine the benefits from the full set of changes at once (with suitable discounting for improvements realized in the future), versus incrementally examining the benefits as they arise with each discrete improvement. The extent to which issues related to adding up may have relevance due to this feature of policy has not been examined in the benefits transfer literature.

### 2.2 Concepts

We now turn specifically to the theoretical constructs of scope and adding up which we focus on for the remaining of the paper. Scope and adding up tests were originally developed in the context of assessing the validity of the contingent valuation method. They are similar concepts, in the sense that both relate to the sensitivity of willingness to pay to changes in the level of provision of the environmental commodity. Likewise, similar terminologies are often used in reference to both concepts—e.g. scope sensitivity, adequacy of scope effects, part-whole bias, and embedding effects. The concepts of both scope and adding up, as well as their relationship to one another, can be seen formally by defining willingness to pay for a change in environmental quality ( $q$ ) from  $q^0$  to  $q^1$  by

$$WTP(q^{0 \rightarrow 1}, u^0) = e(q^0, u^0) - e(q^1, u^0) = y^0 - e(q^1, u^0), \tag{1}$$

where  $e(\cdot)$  is the expenditure function, and  $u^0$  denotes the original level of utility associated with a set of baseline price  $p^0$  and income level  $y^0$ . As defined, this is a compensating variation measure that is positive when  $q^1$  represents an environmental improvement. Similarly, the  $WTP$  for a change from  $q^0$  to  $q^2$  is

$$WTP(q^{0 \rightarrow 2}, u^0) = e(q^0, u^0) - e(q^2, u^0). \tag{2}$$

If  $q^0 < q^1 < q^2$  and  $q$  is a normal good, then the welfare measures are said to show *sensitivity to scope* if  $WTP(q^{0 \rightarrow 1}, u^0) < WTP(q^{0 \rightarrow 2}, u^0)$ .

The relationship between sensitivity to scope and what has come to be referred to as *adding-up* can be seen by further manipulation of (2) to yield<sup>1</sup>

$$\begin{aligned} WTP(q^{0 \rightarrow 2}, u^0) &= e(q^0, u^0) - e(q^2, u^0) \\ &= WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^0). \end{aligned} \quad (3)$$

Note that the second term on the RHS is not a standard welfare construct. For this “adding up” restriction to hold, the *WTP* for a change in environmental quality from  $q^1$  to  $q^2$  (the second term) must be constructed so that it returns the consumer to the original utility level  $u^0$ , rather than the utility level  $u^1$  achieved following provision of the environmental commodity in amount  $q^1$ . This unusual counterfactual makes a strict empirical implementation challenging. Diamond (1996) notes that an approximate adding up test can be undertaken by comparing  $WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^1)$  to  $WTP(q^{0 \rightarrow 2}, u^0)$ , but in fact all that can be said precisely from theory is that the summation of the first two welfare measures should be less than or equal to the third, and strictly less in the presence of income effects (see Hoehn and Randall (1987) and our derivation in the “Appendix”).

Using Diamond’s approximation, we refer to a test of

$$WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^1) \leq WTP(q^{0 \rightarrow 2}, u^0) \quad (4)$$

as a test for the existence of consistency with the adding up property. In contrast, if the second term on the RHS of (3) can plausibly be elicited in a stated preference study, it would be possible to implement a strict adding up test, requiring the equality between the left- and right-hand sides of Eq. (3).

The expressions above can also be used to describe empirical tests of the existence and magnitude of scope effects. A statistical test for the existence of scope is simply to test if  $WTP(q^{0 \rightarrow 1}, u^0) < WTP(q^{0 \rightarrow 2}, u^0)$  in a specific application. A test for the adequacy of scope requires that the analyst have a baseline magnitude of scope against which an application-specific prediction can be compared. Whitehead (2016) has recently suggested the concept of scope elasticity (to which we return below), which he argues is useful for focusing tests of scope on economic, rather than statistical, significance. Given intuition for an elasticity, this suggests there are plausible magnitudes for the size of the scope effect that could be used to benchmark the ‘adequacy’ of scope in any given study. For example, if  $S_p$  is a minimum scope size viewed as economically realistic, based on intuition about an elasticity, then a test of the consistency of the magnitude of a scope effect can be written as

$$WTP(q^{0 \rightarrow 2}, u^0) - WTP(q^{0 \rightarrow 1}, u^0) \geq S_p. \quad (5)$$

Table 1 provides a summary of the potential scope and adding up tests that are possible, divided out by existence and magnitude based criteria.<sup>2</sup>

### 2.3 Scope Literature

Table 1 suggests the literature can be divided into studies that focus on scope, and those that focus on adding up. We review this literature in order to frame our discussion of scope and adding up in the context of benefits transfer. The first suggestion to use a test of scope as the

<sup>1</sup> Derivations are provided in “Appendix 1”.

<sup>2</sup> It is worth noting that we have presented a simple version of the adding up logic, based on a single environmental commodity. Additional complexities based on vectors of environmental commodities were a key aspect of the ‘embedding’ debates that pre-dated our simple expressions. See Kahneman and Knetsch (1992), Smith (1992), and Carson et al. (1998).

**Table 1** Theoretical tests of the consistency of the adding up and scope tests with theory: existence and magnitudes

	Consistency with theoretical magnitude	Consistency with the theoretical existence
Scope	$WTP(q^{0 \rightarrow 2}, u^0) - WTP(q^{0 \rightarrow 1}, u^0) \geq S_p^1$	$WTP(q^{0 \rightarrow 1}, u^0) \leq WTP(q^{0 \rightarrow 2}, u^0)$
Adding up	$WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^0)$ $= WTP(q^{0 \rightarrow 2}, u^0)$	$WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^1)$ $\leq WTP(q^{0 \rightarrow 2}, u^0)$

<sup>1</sup>Here  $S_p$  is a minimum scope size deemed economically realistic

basis for theoretical consistency of stated preference estimates is due to Kahneman (1986); this work was then further developed in Kahneman and Knetsch (1992). The authors used a split sample design, whereby one set of respondents was asked about their willingness to pay to improve environmental services generally. They were then asked to allocate a proportion of their total to a specific component of environmental services (disaster preparedness services, in the application). A second set of respondents was asked only about their willingness to pay for disaster preparedness. Comparing the means and medians among the different samples, the authors conclude that there is little scope sensitivity, in the sense that mean and median sample values for the two treatments are indistinguishable. This analysis helped launch a large research agenda dedicated to refining methodologies for detecting scope effects, and understanding the role that scope sensitivity should play in assessing the validity of stated preference more generally.<sup>3</sup>

Subsequent analyses of scope effects implemented valuation protocols that adhered more closely to current understanding of best stated preference practice. For example, Rollins and Lyke (1998) study Canadian households’ willingness to pay for successive quantities of ecosystem preservation in the country’s remote Northwest Territories. Using a variety of split sample and multiple mode survey approaches, and dichotomous choice elicitation, the authors estimate the value of individual, pairs, and groups of parks to demonstrate sensitivity to scope and diminishing marginal existence value. By recognizing the potential for scope effects to be small at the ‘top’ range of the environmental commodity, when marginal willingness to pay for an additional unit may be small, Rollins and Lyke also contribute to our understanding of when scope effects should and should not be expected.

The emphasis on nuance in understanding the role that scope may play in assessing construct validity is continued in Heberlein et al. (2005). The authors examine four environmental goods in northern Wisconsin, performing split sample and within sample scope tests for each commodity. They show that two goods (water quality and prevention of native spearfishing) pass conventional split sample scope tests, while two others (wolf populations and biodiversity) do not. They then explore a variety of explanations for the scope failures, using within person tests, social psychological theory, and post-study interviews. Based on this broad perspective, they conclude that scope tests are neither necessary nor sufficient to determine the validity or invalidity of a study. Rather, context matters, and the emphasis should be placed on the plausibility of the finding, given what is known about the environmental commodity, population, and survey methodology.

The caution advocated by Heberlein et al. (2005) notwithstanding, meta-analyses suggest that most stated preference studies pass scope tests based on statistical criteria. For example,

<sup>3</sup> As of this writing, there are 1,948 Google Scholar citations and 701 Web of Science citations to Kahneman and Knetsch (1992). A review by Desvousges et al. (2012) identifies 109 studies that conducted scope tests, and met the authors’ criteria for inclusion in their analysis.

Ojea and Loureiro (2011) show that biodiversity valuation studies are on average sensitive to scope. More generally, their Table 1 summarizes meta-analyses for groundwater, coral reefs, wetlands, aquatic resources, endangered species, visibility, and water quality that find evidence of sensitivity to scope.<sup>4</sup> The general notion that the typical stated preference study does not fail scope tests is also supported by Table 1 in Desvousges et al. (2012), which classifies over 100 studies as failing, passing, or having mixed evidence regarding scope. Among these, only 15% are classified as failing to exhibit sensitivity to scope, while the 49% classified as mixed often include specifications where scope effects are absent for expected reasons (see Whitehead 2016, p. 19). Likewise, Carson (2012) argues that well-designed stated preference surveys tend to pass scope tests.

In our review of stated preference validity (Kling et al. 2012), we argued that recent theoretical and empirical developments allow a wider range of empirical outcomes to pass construct validity tests. For example, Amiran and Hagen (2010) show that, under some plausible relationships between market and nonmarket goods, rational behavior does not need to exhibit sensitivity to scope. We conclude from this and the wider body of evidence that contemporary scope tests should not be used as a general litmus test for the overall reliability of stated preference, but should instead be one criterion among several to the gauge the validity of a specific study. For benefits transfer, this means that ‘study sites’ should not be eliminated simply because they do not pass a statistical scope test. Rather, the context should be examined, and additional factors—such as the adequacy of scope when scope effects are expected—should be considered. We return to this theme below.

## 2.4 Adding up Literature

In contrast to the large number of scope tests in the stated preference literature, there is a dearth of genuine adding up tests. Desvousges et al. (2015) identify only four studies whose designs enable tests for public goods based on the incremental welfare measures shown in Eqs. (3) or (4), and fewer still that explicitly execute the test. This is no doubt due to the challenges associated with communicating the unusual counterfactual shown in Eq. (3), whereby respondents must respond conditional on a given level of provision and payment having taken place. For this reason, a small number of studies have investigated adding up using private goods and real payment mechanisms. For example, Bateman et al. (1997) report evidence of adding up failure for components of a restaurant meal in a real payment laboratory experiment. More recently, Elbakidze and Nayga (2016) use a real payment experiment with multiple units of the same private good to investigate adding up, and likewise find that adding up is not satisfied.

Desvousges et al. (2015) is one of the few studies to implement an adding up test in a stated preference context for a public good. They modify a previously-fielded survey that successfully presented respondents with incremental provision levels for the environmental good; the modifications were designed to convey the conditions necessary for the adding up test. Similar to the private good, real payment studies, the authors find that their willingness to pay predictions fail the adding up condition.

It is difficult to use the small adding up literature to draw general conclusions on the validity of stated preference as a valuation method, given the failures seen when the test is applied to private goods with real money payments as well. More pragmatically, evidence on adding up cannot be used to assess the validity of benefits transfer or help guide selection of primary studies for use in a transfer, given the small volume and lack of a criterion that does

<sup>4</sup> Among five other studies listed, three show scope insensitivity and two show negative scope.

exhibit adding up. Indeed, our sense is that adding up will continue to be a useful research area, but test outcomes are unlikely to be informative for stated preference or benefits transfer until adherence to adding up can be found in contexts less challenging, than survey-based public goods valuation.

## 2.5 Scope, Adding up, and Benefits Transfer

The benefits transfer research literature has primarily focused on transfer methodology and transfer validity. Researchers have examined both reduced form and structural approaches to benefit function transfer, defined validity concepts, and carried out validity tests by, for example, comparing transferred estimates to estimates arising from a primary study. The edited volume from Johnston et al. (2015) provides a comprehensive review of this literature. Our sense is that scope concepts as specifically defined in the stated preference literature have not featured prominently in the benefits transfer literature, though concepts related to scope (and adding up) have played a role. In this subsection we comment briefly on this connection.

In performing benefit transfer, study site household values will often need to be adjusted to estimate policy site values to account for differences in the affected areas, size of the quantity changes, and/or baseline levels of quantity. One common approach for accomplishing these adjustments is to compute a unit value (e.g. value per acre of wetland), and apply this unit value to the study site scale of change. Though common, this practice imbeds important assumptions about the quantity of services provided by environmental resources, and how these services relate to observable unit metrics such as acres of wetlands of recreation trips. Mechanically, the unit value approach assumes away diminishing marginal utility, and individual or household level values need to be aggregated to the population level. For the mechanical adjustments, the extent of the market, spatial variation in values, and distance decay must be addressed. More fundamentally, the use of unit value proxies needs to be evaluated in the context of a general model of valuation connecting the units to theoretically consistent measures of value.

These points relate to both aggregation and adjustments for individual preferences, while scope relates only to individual preferences. For example, a primary study reporting the value of services from an acre of wetlands in a specific context could be appropriately sensitive to scope, while a transfer exercise using the unit value in a different context may rely on invalid aggregation assumptions.

The closest direct point of interaction between scope and the benefits transfer literature occurs when sensitivity to scope is used as a criterion for selecting appropriate study sites for inclusion in a transfer exercise. Whitehead et al. (2015) highlight this connection in a review of ways that contingent valuation can be used in benefits transfer. For example, they consider benefits transfer performance using a primary study that is configured to exclude and then include scope sensitivity, and then predict transfer validity using different ways of using findings from the primary study. Specifically, in an application to oyster consumption safety, they report findings from a primary study that predicts the willingness to pay per meal for a program that reduces consumption-related fatality risk. In one specification the data are used to estimate average willingness to pay per meal, while in a second specification, covariates are included that adjust the willingness to pay for changes in risk and other variables, such as frequency of oyster consumption. The two configurations are used to compare unit value transfers using direct predictions from the simple model and fully specified models, and benefit function transfers using the fully specified model. Though the findings are somewhat mixed, the protocol described for assessing validity seems to be one of the few examples whereby sensitivity to scope in the primary studies and transfer methodology were explicitly considered.



A final approach to incorporating scope in benefits transfer is proposed and adopted by Newbold et al. (2016) in this issue. The authors impose scope in their meta-analysis estimating function by beginning with a *marginal willingness to pay function*, and then integrating it to recover a functional form for the *benefit function* that, by construction, is consistent with adding up. They illustrate that there can be significant differences between their utility theoretic approach, and traditional nonstructural approaches.

### 3 Analysis

In this section we present three individual analyses that illustrate ways that scope considerations can influence different aspects of benefits transfer. In the first sub-section we examine the role that scope elasticity could play in study site selection. In the last two sub-sections, we examine how benefit function construction and validity assessment using meta-analysis and preference calibration may be influenced by scope considerations.

#### 3.1 Selecting Study Sites

One of the three ways in which scope and adding up concepts relate to benefits transfer is in the selection of original studies to use in the transfer exercise. The stated preference literature reviewed in the previous section has focused heavily on scope as a construct validity concept, and so one conclusion is that studies should be selected for benefits transfer based on an ‘existence of scope’ criterion (see Table 1). This, however, ignores notions of adequacy of scope, and could exclude studies that are insensitive to scope for plausible reasons. At a more general level, adequate response to scope in stated preference studies has an element of subjectivity, in that we need to judge scope magnitudes without points of comparison. That is, scope effects have not often been examined from criterion or convergent validity perspectives. In this sub-section we explore the possibility of using scope effects in revealed preference studies to better understand scope plausibility in stated preference. For this, the concept of scope elasticity noted above is useful.

Whitehead (2016) defines the elasticity of a welfare measure with respect to the quantity or quality change generating it as

$$\varepsilon_{wtp,q} = \frac{\partial WTP}{\partial(q_1 - q_0)} \frac{(\tilde{q}_1 - \tilde{q}_0)}{WTP}, \quad (6)$$

where  $q$  is the quality or quantity variable of interest,  $WTP$  is the welfare measure for a change in  $q$ , and  $\tilde{q}_1 - \tilde{q}_0$  is the change that generated  $WTP$ , as well as the point of reference for the elasticity. Whitehead refers to this as a scope elasticity, since it provides a unit-free measure of how responsive  $WTP$  is to the size of the change. He suggests that this metric can be computed for functional forms commonly used in the stated preference literature, in order to gain intuition on how to compare scope findings generally. In what follows we apply his logic more broadly to examine scope effects in recreation models, and then for an example stated preference study.

Consider first a continuous demand function model of recreation behavior. The demand for trips  $x$  is given by

$$\ln x = \alpha - \beta p + \phi q, \quad (7)$$

where  $x$  is trips,  $p$  is travel cost, and  $q$  is a measure of site quality. For this zero income effects model, the indirect utility function is

$$V(p, q, y) = y + \frac{\exp(\alpha - \beta p + \phi q)}{\beta}, \tag{8}$$

where  $y$  is income. For this functional form, the willingness to pay for a change in  $q$  can be written as

$$\begin{aligned} WTP(q_1 - q_0) &= \frac{\exp(\alpha - \beta p + \phi q_1) - \exp(\alpha - \beta p + \phi q_0)}{\beta} \\ &= \frac{\exp(\alpha - \beta p)}{\beta} [\exp(\phi q_1) - \exp(\phi q_0)]. \end{aligned} \tag{9}$$

Making the substitution  $q_1 = q_0 + (q_1 - q_0) = q_0 + \Delta q$ , we can write  $WTP$  as

$$WTP(\Delta q) = \frac{\exp(\alpha - \beta p)}{\beta} [\exp(\phi q_0 + \phi \Delta q) - \exp(\phi q_0)]. \tag{10}$$

From Eq. (10), the scope elasticity is<sup>5</sup>

$$\begin{aligned} \frac{\partial WTP(\Delta q)}{\partial(\Delta q)} \frac{\Delta q}{WTP} &= \varepsilon_{wtp,q} = \frac{\frac{\exp(\alpha - \beta p)}{\beta} \exp(\phi q_0 + \phi \Delta q) \phi}{\frac{\exp(\alpha - \beta p)}{\beta} [\exp(\phi q_1) - \exp(\phi q_0)]} \cdot \Delta q \\ &= \frac{\exp(\phi q_1) \phi}{[\exp(\phi q_1) - \exp(\phi q_0)]} \cdot \Delta q. \end{aligned} \tag{11}$$

We use an example from Phaneuf and Requate (2017) to provide an estimate of a scope elasticity for the semi-log recreation model. In their example 17.2 (p. 498), Phaneuf and Requate use data from the Iowa Lakes Project to demonstrate estimation of a system of count data demand equations, where trips to a set of lakes follow a Poisson distribution. Their sample includes a cross section of 2,489 Iowa residents who report visits they made to each of 127 lakes in the state during 2002. Travel costs were imputed for each respondent to each of the destinations using \$0.28 per mile for out of pocket costs, and a fraction of the household wage rate for the opportunity cost of travel time. As an indicator of lake quality, measures of chlorophyll  $a$  (CHL) taken at each of the 127 lakes during the 2002 summer recreation season were included in the specification.<sup>6</sup> Across the set of lakes in 2002, the average CHL concentration was 40.44 micrograms per liter ( $\mu\text{g/l}$ ), with a standard deviation of 38.06  $\mu\text{g/l}$ .

Example estimates from a multiple equation Poisson model are

$$E(x_{i,99}) = \exp(1.34 - 0.024 \cdot p_{i,99} - 0.005 \cdot q_j). \tag{12}$$

For a reference individual with travel cost equal to \$35 and a baseline value of 40  $\mu\text{g/l}$  for CHL, the expected baseline demand is 1.35 trips. A reduction in CHL of 20  $\mu\text{g/l}$  changes the expected demand to 1.50 trips. For this change, the annual willingness to pay is \$6.25. Based on Eq. (11), the scope elasticity for a 20  $\mu\text{g/l}$  change in CHL is 0.95. Thus for this application, and this specification for demand, willingness to pay is nearly unitary elastic with respect to the size of the quality change.

<sup>5</sup> Thanks to Chris Moore for pointing out an error in our original formula for the scope elasticity with this functional form.

<sup>6</sup> Chlorophyll is a useful measure because it reflects the amount of algae and other plant growth responding to nutrient enrichment.

Consider next a simple discrete choice model for recreation site selection. Suppose there are  $J$  recreation sites and that the utility from a visit to site  $j$  is

$$\begin{aligned} V_j &= v_j + \varepsilon_j \\ &= -\beta p_j + \phi q_j + \varepsilon_j. \end{aligned} \tag{13}$$

Applying the familiar log-sum formula, the per trip willingness to pay for an increase in  $q$  at site  $j = 1$  is

$$WTP(q_1^1 - q_1^0) = \frac{1}{\beta} \left\{ \ln \left( \sum_{j=1}^J \exp(v_j^1) \right) - \ln \left( \sum_{j=1}^J \exp(v_j^0) \right) \right\}, \tag{14}$$

where  $q_1^0$  and  $q_1^1$  are the baseline and new quality levels at site 1, respectively, and  $v_1^0$  and  $v_1^1$  are the baseline and new observable utility levels at each site. We show in the appendix that

$$\begin{aligned} \frac{\partial WTP(\Delta q_1)}{\partial \Delta q_1} \frac{\Delta q_1}{WTP} &= \frac{1}{\beta} \frac{\Pr_1^1 \cdot \phi \cdot \Delta q_1}{WTP} \\ &= \frac{\Pr_1^1 \cdot \phi \cdot \Delta q_1}{\ln \left( \sum_{j=1}^J \exp(v_j^1) \right) - \ln \left( \sum_{j=1}^J \exp(v_j^0) \right)}, \end{aligned} \tag{15}$$

where  $\Pr_1^1$  is the probability of a visit to site 1 under the changed conditions.

To understand the magnitude of the scope elasticity in a RUM context, we consider example 17.3 in Phaneuf and Requate (2017, p. 500). Using the same Iowa Lakes Project data as described above, for a trip allocation model (i.e. no repeated choices or participation decision) they present estimates of  $\beta$  and  $\phi$  as follows:

$$U_{ij} = -0.026 \cdot p_{ij} - 0.0091 \cdot q_j + \varepsilon_{ij}. \tag{16}$$

Consider once again a reduction in CHL at a single lake. For Clear Lake in north-central Iowa (one of the lakes in the choice set), the baseline concentration of chlorophyll *a* in 2002 was 30.23  $\mu\text{g/l}$ . We examine a 50% reduction, so that the new level is 15.11  $\mu\text{g/l}$ . Using Eq. (14) for all members of the sample, we calculate an average per trip ex ante willingness to pay for this improvement of \$0.06. The sample average predicted probability of a visit to Clear Lake under improved conditions ( $\bar{P}_1^1$ ) is 0.0114. Substituting this probability, the willingness to pay estimate, and the parameter values into Eq. (15) results in a scope elasticity of per trip willingness to pay for this quality change of 0.502.<sup>7</sup>

These examples illustrate that some intuition on the magnitude of scope effects can be gleaned from commonly used revealed preference models. For completeness, we also predict the scope elasticity for a stated preference study. Houtven et al. (2014) use dichotomous choice contingent valuation to estimate the willingness to pay for water quality improvements in lakes across the state of Virginia. Water quality is described using a five point index, where  $q = 1$  is the highest level of lake water quality in a lake, and  $q = 5$  is the lowest. In one of their specifications (see their model 4 in Table 7), the quality variable is defined as the expected value of the index across all the lakes in Virginia. At baseline conditions the expected index value was reported as  $Q_0 = 3.05$ , and the experimental design presented scenarios that improved the expected index value to  $Q_1 = 2.70, 2.45, 2.30, \text{ or } 2.20$ . Respondents randomly

<sup>7</sup> The demand system and RUM models presented here are only for example. Phaneuf and Requate (2017) present further discussion of these examples, noting in particular that the estimates of the quality impacts are likely to suffer from omitted variable bias.

received one of the four improvement levels. The quality index was log transformed so that the utility difference between the status quo and proposed program was

$$\Delta U = -\beta bid + \phi \ln(Q_0 - Q_1 + 1) + \gamma \cdot Z + \varepsilon, \quad (17)$$

where  $Z$  is a vector of interactions between respondent characteristics and the quality change. The estimates for  $\beta$  and  $\phi$  are  $\beta = 0.00355$  and  $\phi = 1.038$ . For non-college educated respondents who do not take recreation trips to lakes (meaning all elements of  $Z$  are zero), the mean willingness to pay for the program  $\Delta Q = 3.05 - 2.45$  is \$137. This implies a scope elasticity estimate given by

$$\varepsilon_{wtp,Q} = \frac{\phi}{\beta \cdot (\Delta Q + 1)} \frac{\Delta Q}{WTP} = 0.80. \quad (18)$$

How might this be useful for gauging scope effects in stated preference studies, where scope tests are commonly applied? What does this intuition say for how scope magnitudes should be used to select study sites for a benefits transfer? It is interesting to note that the estimates from the two revealed preference models generated scope elasticities that are similar in magnitude to what the stated preference study implies—perhaps providing some intuition on ‘plausible magnitudes’ for sensitivity to scope. This can also be considered an indication of convergent validity.

### 3.2 Constructing Benefit Functions and Assessing BT Validity—Meta Analysis

A recommended approach for benefits transfer is to construct a benefit function, rather than transfer individual  $WTP$  values. A common strategy for benefit function transfer is to use the results of a meta-analysis to construct a relationship between  $WTP$  values from a set of primary studies, and a range of explanatory variables that typically include characteristics of the primary study site and methodological variables. How can scope effects be accommodated in a meta-regression approach to benefit function transfer? How can scope be used to assess the validity of a transfer using a meta-regression based benefit function? In this subsection, we explore the degree to which scope effects can be tested for within a meta-analysis, showing that things depend on the functional form specified for the meta-regression, and the range of welfare estimates available for inclusion in the study.

Bergstrom and Taylor (2006) discuss the development of functional forms based strictly on utility theoretic functions (which they term a ‘strong structural utility theoretic’ approach) and those that are only approximately based on utility function (which they refer to ‘weak structural utility theoretic’). Many applications of meta-analysis using stated or revealed preference studies are best categorized as falling in the latter category. Within this category, Bergstrom and Taylor note that a key decision in performing the meta-analysis is how to incorporate studies that have different scales of change and/or different starting points of the quality or quantity being developed. To address this they note (p. 354) that “At a minimum, reference and target level effects may need to be controlled for on the right hand side of the [meta-analysis-benefit transfer] equation.” The alternative is to make adjustments to the dependent variable, so that the welfare changes that are being aggregated through the meta-regression are of equal scale.

Empirically, the first approach implies a meta-regression specification of the following form:

$$WTP = f(q_1 - q_0, q_0, z), \quad (19)$$

which exhibits scope if the marginal effect of a change in  $(q_1 - q_0)$  is positive. Further, this functional form can be consistent with a declining marginal utility of additional improvements, if the marginal effects of a change in the baseline quality  $q_0$  is negative. This makes the specification in (19) intuitively appealing for analysts interested in explicitly testing for the presence of scope effects in the meta-regression. It also provides a convenient way to test for the expected diminishing marginal value of additional improvements.

However, this functional relationship is not be consistent with the adding up restriction, as it generates ever-increasing *WTP* measures for a given quality change, simply by dividing the size of the change into smaller and smaller increments. Consequently, a functional form with a RHS variable of  $(q_1 - q_0)$  must be understood to be at best an approximation to a benefit function that is utility theoretic.

In addition to considering the size of change as a potential RHS variable, there is also the consideration of how that size of change should be represented: as an absolute quantity, or as a relative change within the original study (i.e. a percentage or proportion). As noted above, Ojea and Loureiro (2011) undertake a meta-analysis of biodiversity values. They consider the implications of these two different representations of the size of change being valued. They interpret the coefficient on the absolute or percentage change as an indication of scope sensitivity. Interestingly, they find evidence of scope when the change is expressed as an *absolute* size change, but not when it is expressed in *relative* terms (as a percentage).

The second approach for making welfare measures associated with different sizes of quality or quantity changes compatible is to convert the *WTP* values into the same units, prior to estimation of the meta-regression. This is particularly common when the meta-analysis summarizes changes in the quantity of a public good, such as the number of wetland acres or visitor day values for recreational trips. In this case, the meta-regression takes one of the following two forms:

$$\frac{WTP}{q_1 - q_0} = f(z), \tag{20}$$

$$\frac{WTP}{q_1 - q_0} = f(q_0, z). \tag{21}$$

Inclusion of the baseline level of the good  $q_0$  on the RHS, as in (21), allows the per unit value to exhibit diminishing marginal value if  $\partial f / \partial q_0 < 0$ . In contrast, the expression in (20) imposes a constant marginal value, regardless of the initial amount of the public good.

In both cases, using an average value on the LHS implies a unitary scope elasticity, as can be seen by rewriting (20) as

$$WTP = f(z)(q_1 - q_0), \tag{22}$$

and constructing the scope elasticity as

$$\frac{\partial WTP}{\partial (q_1 - q_0)} \frac{(q_1 - q_0)}{WTP} = \frac{f(z)(q_1 - q_0)}{WTP} = \frac{f(z)(q_1 - q_0)}{f(z)(q_1 - q_0)} = 1. \tag{23}$$

From this we see that by using average values as the dependent variable in a meta-regression, a unitary scope elasticity is imposed and cannot be inferred from the data.

In the case of the functional forms that are consistent with (19), such as the linear meta-function

$$WTP = \beta_0 + \beta_1(q_1 - q_0) + \beta_2z, \tag{24}$$

the implied scope elasticity can be constructed using the following formula

$$\frac{\partial WTP}{\partial(q_1 - q_0)} \frac{(q_1 - q_0)}{WTP} = \frac{\beta_1(q_1 - q_0)}{WTP}. \tag{25}$$

*Empirical examples* To close this sub-section, we examine three representative meta-analyses to illustrate their implied sensitivity to scope. First, Richardson and Loomis (2009) undertake a meta-analysis of threatened, endangered and rare species, where they estimate a linear meta-regression of the form in (24).<sup>8</sup> Data reported in their Table 1, along with coefficient estimates, allows us to construct a scope elasticity estimate for annual willingness to pay for species of 0.30. Richardson and Loomis also estimate a double log specification, which has an implied scope elasticity of 0.73. These authors do not include the baseline species levels as an explanatory variable, so no direct evidence on diminishing marginal values is available.

Ge et al. (2013) estimate a meta-regression for water quality improvements using both stated and revealed preference data, with a linear functional form. Using pooled observations, their estimates imply a scope elasticity of 0.23. The regressions also demonstrate a clear diminishing marginal value, as the sign on starting water quality levels is negative.

A final example of implied scope elasticity comes from Johnston et al. (2005), who present a meta-analysis of water quality improvements based on stated preference studies. They estimate a specification where the RHS water quality change is crossed with a dummy indicating the type of fishery impacted by the water quality. Their semi-log specification implies a scope elasticity of 0.67. In contrast, their double log elasticity implies a value of 1.68. In both cases, diminishing marginal values are indicated by a negative estimate for the parameter on the baseline level of water quality.

### 3.3 Constructing Benefit Functions and Assessing BT Validity—Structural Estimation

To illustrate the role that scope effects can play in a more structural approach to benefits transfer, we next consider a preference calibration example. As described by Smith et al. (2002), the basic idea behind preference calibration is to select a functional form that represents preferences for the environmental attribute at the policy site, and then calibrate the function using information from the literature. For example, suppose we are interested in assessing the value of water quality changes at a lake that has not been the subject of a primary study. Similar to the example from Sect. 3.1, we assume demand for recreation trips to the lake is given by

$$x(p, q) = \exp(\alpha - \beta p + \phi q). \tag{26}$$

Note that this implies that any income effects are small enough to allow us to treat consumer surplus measures as equivalent to Hicksian measures, which we exploit in the algebra of our calibration, and that the indirect utility function is given by Eq. (8).

Specifically, preference calibration proceeds by first deriving expressions for economic concepts using the assumed functional form. For example, using (26) we can show the following:

$$\begin{aligned} \text{Price elasticity:} & \quad \varepsilon_p = -\beta p \\ \text{WTP for access:} & \quad WTP(x) = \frac{x}{\beta}, \quad WTP(1) = \frac{1}{\beta} \\ \text{Marginal value of quality:} & \quad MWP_q = \frac{\phi x}{\beta}, \quad MWP(1)_q = \frac{\phi}{\beta} \\ \text{WTP for } \Delta q: & \quad WTP(\Delta q) = \frac{x(q_1) - x(q_2)}{\beta} \\ \text{Scope elasticity of WTP:} & \quad \varepsilon_{wtp,q} = \frac{\exp(\phi q_1)\phi}{[\exp(\phi q_1) - \exp(\phi q_0)]} \cdot \Delta q \end{aligned}$$

<sup>8</sup> Their change is expressed in percentages and the formula has been adjusted accordingly.

**Table 2** Example calibration concepts and values

Variable	Functional specification	Value	Condition index
Baseline price	$p_0$	\$60	–
Baseline chlorophyll $a$	$q_0$	50 $\mu\text{g/l}$	–
Baseline demand	$x = \exp(\alpha - \beta p_0 + \phi q_0)$	1.5	1
Price elasticity	$-\beta \cdot p$	– 0.85	2
WTP per trip	$1/$	\$40	3
Marginal WTP quality	$\phi/\beta$	\$0.20	4
WTP( $\Delta q$ ) $\downarrow$ 10 $\mu\text{g/l}$	$\frac{x(q_0)-x(q_1)}{\beta}$	\$8	5
Scope elasticity ( $\Delta q$ ) $\Delta q=20 \mu\text{g/l}$	$\frac{\exp(\phi q_1)\phi \cdot \Delta q}{(\exp(\phi q_1)-\exp(\phi q_0))}$	1.5 or 0.25	6

Preference calibration proceeds by locating estimates of the economic phenomena in the literature (i.e. locating study site estimates for elasticities, WTPs, etc.), and using a minimum distance algorithm to find values for the function parameters that jointly reflect the literature estimates.

In this sub-section we use the preference calibration logic to investigate the importance of including data points in the calibration procedure that reflect scope sensitivities. In particular, we calibrate the parameters in (26) by matching on various sets of the conditions listed above, with and without inclusion of the scope elasticity condition. We then use the calibrated function to predict welfare measures, and compare our predictions under the various calibration strategies. Similar to Sect. 3.1, we assume water quality is measured using the concentration of chlorophyll  $a$ . Table 2 provides several example conditions and values that we calibrate on. The last column provides a reference index for each for the conditions, which we refer to below.

Since there are three parameters and more than three conditions in several of our calibration strategies, we use a minimum distance algorithm, described in general as follows. Suppose there are  $G$  conditions (i.e. relationships between functional expressions and values from the literature) and  $K$  parameters, denoted by the  $K$ -dimensional vector  $\theta$ , to calibrate. In our example,  $\theta = (\alpha, \beta, \phi)$ . Define each condition by the general statement

$$m_g(\theta) = v_g - f_g(\theta), \quad g = 1, \dots, G, \tag{27}$$

where  $f_g(\theta)$  is the expression derived from the structural function (e.g. willingness to pay for a discrete quality change), and  $v_g$  is its value, gleaned from the literature. The objective of the calibration in the over-identified case is to find values for  $\theta$  that make each  $m_g(\theta)$  as close to zero as possible, for all  $g$ . To proceed, denote  $m(\theta)$  as the  $G \times 1$  vector with individual elements  $m_g(\theta)$ . The calibrated parameters are implicitly defined according to

$$\begin{aligned} \theta_{BT} &= \arg \min_{\theta} [m(\theta)' \cdot I_G \cdot m(\theta)] \\ &= \arg \min_{\theta} \left[ \sum_{g=1}^G (v_g - f_g(\theta))^2 \right], \end{aligned} \tag{28}$$

where  $I_G$  is the  $G \times G$  identify matrix.

We implemented the minimization routine in Matlab using six different calibration approaches. The calibration scenarios and our results are shown in Table 3. The first col-

**Table 3** Calibration results

Parameter	Conditions 1, 2, 5	Conditions 1, 2, 5, 6	Conditions 1, 2, 5, 6	Conditions 1, 2, 3, 4, 5	Conditions 1, 2, 3, 4, 5, 6	Conditions 1, 2, 3, 4, 5, 6
$\alpha$	1.6195	1.5805	1.6050	2.5792	2.5846	2.5763
$\beta$	0.0142	0.0136	0.0141	0.0250	0.0250	0.0250
$\phi$	-0.0073	-0.0069	-0.0073	-0.0111	-0.0110	-0.0112
Scope elasticity assumption	NA	0.25	1.50	NA	0.25	1.50
WTP for 20 $\mu\text{g/l}$ $\downarrow$	\$16.55	\$16.55	\$16.53	\$16.79	\$16.79	\$16.82

umn of results is for a simple specification that only uses conditions 1, 2, and 5 in Table 2, without any scope effects. The fourth column of results is a fuller specification that uses all of the conditions in Table 2, less the scope condition. The remaining columns present results that are obtained when scope elasticity conditions set to 0.25 and 1.50 are included in the calibration. For each specification, we present parameter values and the estimated willingness to pay for a 20  $\mu\text{g/l}$  reduction in chlorophyll  $a$ .

The simple and fuller specifications produce somewhat different parameter estimates, but quite similar welfare measures. In addition, there is little difference in parameters or welfare measures when the calibration accounts for scope effects of different sizes. Thus in this case, calibrating on scope elasticity across a wide range of values does not matter for the benefits transfer exercise. One explanation for this is that the structural nature of this type of benefits transfer approach bakes scope effects into the analysis via the assumed functional form and other, more important, calibration conditions.

## 4 Discussion

In this paper, we have considered the question of how scope and/or adding-up tests can be used to inform and improve the practice of benefit transfer. We consider three ways in which these concepts can improve practice: through selection of study sites used directly to transfer values (and/or to include in a meta-analysis), by providing empirical magnitudes to use in calibrating benefit functions via structural benefits transfer, and through ex post assessments of transfer validity. We offer several conclusions.

First, we note that adding-up, while theoretically appealing, is a conceptually difficult test to implement. Further, the few empirical tests undertaken to date using real payments and private goods fail to show consistency with adding-up. This failure, along with the difficulty of implementation in a stated preference framework, suggest to us that appealing to the results of adding-up tests are not likely to be fruitful in informing benefit transfer.

In contrast, scope effects appear to be fertile ground for useful information. Given the lack of clear guidance from theory about the magnitude, and even existence, of scope effects, it would be inappropriate to apply a strict test of either scope existence or magnitude as the only basis on which to select candidate study sites for transfers, or to calibrate a benefit transfer function. However, while a strict test may not be appropriate, both intuition and theory suggest that in many cases, scope effects should be present and can be informative about the validity of a transfer. And, as Whitehead (2016) suggests, the magnitude of those effects can be usefully considered via scope elasticity. In this paper, we suggest that empirical



estimates of scope elasticity from revealed preference studies and meta-analyses could be valuable sources of insight in the benefit transfer process.

To demonstrate the logic, we selected a few examples of recreation demand estimates using both continuous and discrete choice models and computed the implied scope elasticities. While our example estimates are intended only as a proof of concept, a useful next step would be to undertake a more representative assessment of the recreation demand literature to compute implied scope elasticities for both changes in levels of site qualities, and changes in quantities where appropriate. A database of studies, the type of good valued, the implied scope elasticities, and other parameters and study characteristics could provide useful information for benefit transfers accomplished via meta-analysis or benefit function transfer. This same information could be used to inform structural benefit transfer.

While we used examples from recreation demand to demonstrate the method, the same logic and approach can be taken with other revealed preference methods including hedonic models, sorting models, and any other revealed preference approaches to eliciting willingness to pay for quality or quantity changes. A thorough review and assessment of that literature could provide additional indicators of reasonable ranges of scope elasticities for changes in a variety of ecosystem services (e.g., catch rates, water quality measured in Secchi depth, nutrient concentrations, habitat quality, etc.) and public good quantities (number of acres of wetland habitat, user days of recreation sites, etc.).

The logic of employing scope elasticities from revealed preference methods to inform benefits transfer faces a significant limitation in that such methods cannot yield estimates of nonuse values. For insight into plausible scope elasticities for nonuse values, it may be worthwhile to identify stated preference studies that meet a broad set of validity tests (criterion, convergent, construct, and content) for which to compute elasticities.<sup>9</sup>

What does our discussion imply about research needs and opportunities related to benefits transfer methodology? We speculated that the similar scope elasticities in our RP and SP examples implied a type of convergent validity, but did not explore this carefully. Can ideas related to scope elasticity be more fully developed, so as to provide an additional metric for assessing convergent validity—in benefits transfer and valuation generally? Are there systematic difference across types of commodities? When elasticities diverge, are there explanations due to study design that may explain the differences?

Our preference calibration exercise failed to uncover any interesting role for calibrating on scope effects directly. Our model is of course very stylized, and the literature on preference calibration is generally sparse, meaning there is little accumulated wisdom on best practices. Given this, what are the most important pieces of information to calibrate on in richer specifications? Are scope effects adequately represented by calibrating on marginal and discrete choice willingness to pay conditions, making additional information on scope superfluous?

These observations regarding scope are part of a larger theme related to how we use study specific outcomes measured in specific contexts—e.g. consumer surplus per trip, scope elasticity, the marginal willingness to pay, price and other elasticities, etc.—to measure the general contribution to ‘full income’ of environmental services, where full income is defined as both money income and the monetary value of nonmarket services. While the benefits transfer literature contains both reduced form and structural approaches for aggregating study-specific information into policy-site predictions for one-off decision making purposes, a general architecture for measuring the overall contribution of the environment to (full)

<sup>9</sup> There is a risk of circularity in this logic, which analysts need to carefully consider.

income is still missing. This makes it challenging to assess how the many parts constituting knowledge in the nonmarket valuation literature fit into a coherent whole.

As regards the practice of benefits transfer, it will always require good judgment and transparency with regard to assumptions and data on the part of the analyst. The plausibility of the implied scope elasticity could be a valuable indicator of whether a study should be relied upon for use in such an analysis. To support improved benefit transfers, EPA should consider developing and maintaining a database of plausible ranges of scope elasticities from a range of revealed preference (and potentially well done stated preference) studies. These estimates could be used to support benefits transfer through calibration and selection for inclusion in meta-analyses.

### Appendix 1: Definition of the Adding Up Condition

Recall that we define willingness to pay for a change in environmental quality ( $q$ ) from  $q^0$  to  $q^1$  using

$$WTP(q^{0 \rightarrow 1}, u^0) = e(q^0, u^0) - e(q^1, u^0), \tag{A1}$$

Equivalently, the  $WTP$  for a change from  $q^0$  to  $q^2$  can be written

$$\begin{aligned} WTP(q^{0 \rightarrow 2}, u^0) &= e(q^0, u^0) - e(q^2, u^0) \\ &= [e(q^0, u^0) - e(q^1, u^0)] + [e(q^1, u^0) - e(q^2, u^0)] \\ &= WTP(q^{0 \rightarrow 1}, u^0) + [e(q^1, u^0) - e(q^2, u^0)] \\ &= WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^0). \end{aligned} \tag{A2}$$

This demonstrates that the compensation needed to return the consumer to the original utility level after an increase in  $q$  from  $q^0$  to  $q^2$  can be divided into two intermediate components associated with smaller increments in environmental quality (assume  $q^0 < q^1 < q^2$ ). The first term in (A2) is the  $WTP$  for a change from  $q^0$  to  $q^1$ . The second term is similar to a standard compensating variation, but in the move from  $q^1$  to  $q^2$  it requires that consumer's utility be taken back to  $u^0$  instead of  $u^1$ . From an empirical perspective this is important because it is typical to recover welfare estimates that leave consumer as well off as with the original level of environmental quality— $q^0$  in this case. If one were to add the compensation needed to bring the consumer back to  $u^1$ , this would be an underestimate of the full  $WTP$ .

To better understand this last term it is useful to rewrite the  $WTP$  for the change from  $q^0$  to  $q^2$  as the sum of two well understood measures, and an additional term:

$$\begin{aligned} WTP(q^{0 \rightarrow 2}, u^0) &= WTP(q^{0 \rightarrow 1}, u^0) + [e(q^1, u^0) \\ &\quad - e(q^2, u^0) + e(q^1, u^1) - e(q^1, u^1) + e(q^2, u^1) - e(q^2, u^1)] \\ &= WTP(q^{0 \rightarrow 1}, u^0) + [e(q^1, u^1) - e(q^2, u^1)] \\ &\quad + [e(q^1, u^0) - e(q^2, u^0) - e(q^1, u^1) + e(q^2, u^1)] \\ &= WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^1) \\ &\quad - \{[e(q^1, u^1) - e(q^2, u^1)] - [e(q^1, u^0) - e(q^2, u^0)]\} \\ &= WTP(q^{0 \rightarrow 1}, u^0) + WTP(q^{1 \rightarrow 2}, u^1) \\ &\quad - [WTP(q^{1 \rightarrow 2}, u^1) - WTP(q^{1 \rightarrow 2}, u^0)], \end{aligned} \tag{A3}$$

where the term in brackets represents the compensation needed to adjust for the utility difference in moving when using  $u^0$  instead of  $u^1$  as the baseline. These results were originally developed by Hoehn and Randall (1987).

Equation (A3) represents a second version of the adding up condition. In short, the *WTP* for a change in  $q$  from  $q^0$  to  $q^2$  can be decomposed into the sum of the *WTP* for a change from  $q^0$  to  $q^1$ , the *WTP* for a change from  $q^1$  to  $q^2$  (with utility adjusted accordingly) and an income effect.

### Appendix 2: Derivation of Equation (15)

Recall from Eq. (14) that willingness to pay per trip is given by

$$WTP(q_1^1 - q_1^0) = \frac{1}{\beta} \left\{ \ln \left( \sum_{j=1}^J \exp(v_j^1) \right) - \ln \left( \sum_{j=1}^J \exp(v_j^0) \right) \right\}, \tag{A4}$$

which we can rewrite as

$$WTP(\Delta q_1) = \frac{1}{\beta} \left\{ \ln \left( \underbrace{\exp(-\beta p_1 + \phi q_0 + \phi \Delta q_1)}_{\exp(v_1^1)} + \sum_{j=2}^J \exp(v_j^1) \right) - \ln \left( \sum_{j=1}^J \exp(v_j^0) \right) \right\}. \tag{A5}$$

Differentiating (A5) respect to  $\Delta q_1$  results in

$$\begin{aligned} \frac{\partial WTP(\Delta q_1)}{\partial \Delta q_1} &= \frac{1}{\beta} \left\{ \frac{1}{\sum_{j=1}^J \exp(v_j^1)} \cdot \exp(v_1^1) \cdot \phi \right\} \\ &= \frac{1}{\beta} \Pr_1 \cdot \phi, \end{aligned} \tag{A6}$$

so that

$$\frac{\partial WTP(\Delta q_1)}{\partial \Delta q_1} \frac{\Delta q_1}{WTP} = \frac{1}{\beta} \frac{\Pr_1 \cdot \phi \cdot \Delta q_1}{WTP}. \tag{A7}$$

Plugging in the expression for *WTP* leads to Eq. (15).

### References

Amiran EY, Hagen DA (2010) The scope trials: variation in sensitivity to scope and WTP with directionally bounded utility functions. *J Environ Econ Manag* 59:293–301

Bateman Ian, Munro Alistair, Rhodes Bruce, Starmer Chris, Sugden Robert (1997) Does part-whole Bias exist? An experimental investigation. *Econ J* 107:322–332

Bergstrom JC, Taylor LO (2006) Using meta-analysis for benefits transfer: theory and practice. *Ecol Econ* 60:351–360

Carson Richard T (2012) Contingent valuation: a practical alternative when prices aren't available. *J Econ Perspect* 26:27–42

Carson R, Flores N, Hanemann WM (1998) Sequencing and valuing public goods. *J Environ Econ Manag* 36:314–323

Desvousges W, Mathews K, Train Kenneth (2012) Adequate responsiveness to scope in contingent valuation. *Ecol Econ* 84:121–128

- Desvousges W, Mathews K, Train K (2015) An adding-up test on contingent valuation of river and lake quality. *Land Econ* 91:556–571
- Diamond Peter A (1996) Testing the internal consistency of contingent valuation surveys. *J Environ Econ Manag* 30:337–347
- Elbakidze L, Nayga R (2016) Violations of the adding-up test: a non-hypothetical experimental investigation, manuscript
- Ge J, Kling CL, Herriges JA (2013) How much is clean water worth? valuing water quality improvements using a meta analysis, Iowa State University draft manuscript
- Hanemann WM (1991) Willingness to pay and willingness to accept: how much can they differ? *Am Econ Rev* 81:635–647
- Heberlein TA, Wilson MA, Bishop RC, Schaeffer Nora Cate (2005) Rethinking the scope test as a criterion for validity in contingent valuation. *J Environ Econ Manag* 50:1–22
- Hoehn J, Randall A (1987) A satisfactory benefit cost indicator from contingent valuation. *J Environ Econ Manag* 14(3):226–247
- Horowitz JK, McConnell KE (2003) Willingness to accept, willingness to pay, and the income effect. *J Econ Behav Organ.* 51:537–545
- Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (2015) Introduction to benefit transfer methods, chapter 2. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) *Benefit transfer of environmental and resource values: a guide for researchers and practitioners*. Springer, Berlin
- Johnston RJ, Besedin EY, Iovanna R, Miller CJ, Wardwell Ryan F, Ranson Mathew H (2005) Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis. *Can J Agric Econ* 53:221–248
- Kahneman D (1986) Comments by Professor Daniel Kahneman. In: Cummings RG, Brookshire DS, Schulze WD (eds) *Valuing Environmental goods: a state of the art assessment of the contingent valuation method*. Rowman and Allanheld, Upper Saddle River, pp 185–194
- Kahneman D, Knetsch JL (1992) Valuing public goods: the purchase of moral satisfaction. *J Environ Econ Manag* 22:57–70
- Kaul S, Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC (2013) What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. *J Environ Econ Manag* 66:90–104
- Kling CL, Phaneuf DJ, Zhao J (2012) From exxon to BP: has some number become better than no number? *J Econ Perspect* 26:3–26
- Newbold SC, Massey DM, Walsh P, Hewitt J (2016) Estimating benefit transfer functions that add up, draft manuscript
- Ojea E, Loureiro ML (2011) Identifying the scope effect on a meta-analysis of biodiversity valuation studies. *Resour Energy Econ* 33:706–24
- Phaneuf D, Requate T (2017) *A course in environmental economics: theory, policy, and practice*. Cambridge University Press
- Richardson L, Loomis John (2009) The total economic value of threatened, endangered and rare species: an updated meta-analysis. *Ecol Econ* 68:1535–1548
- Rollins K, Lyke A (1998) The case for diminishing marginal existence values. *J Environ Econ Manag* 36(3):324–344
- Siikamaki J (2011) Contributions of the US state park system to nature recreation. *Proc Ntl Acad Sci* 108:14031–14036
- Smith VK (1992) Arbitrary values, good causes, and premature verdicts. *J Environ Econ Manag* 22:71–89
- Smith VK, Van Houtven G, Pattanayak S (2002) Benefit transfer via preference calibration: ‘prudential algebra’ for policy. *Land Econ* 78:132–152
- Van Houtven G, Mansfield C, Phaneuf DJ, von Haefer R, Milstead Bryan, Kenney Melissa A, Reckhow Kenneth H (2014) Combining expert elicitation and stated preference methods to value ecosystem services from improved lake water quality. *Ecol Econ* 99:40–52
- Whitehead JC (2016) Plausible responsiveness to scope in contingent valuation. *Ecol Econ* 128:17–22
- Whitehead JC, Ashton MO, Huth WL (2015) Benefit transfers with the contingent valuation method,” chapter 7. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) *Benefit transfer of environmental and resource values: a guide for researchers and practitioners*. Springer, Berlin
- Willig RD (1976) Consumer surplus without apology. *Am Econ Rev* 66:589–597