Using Meta-Analysis for Large-Scale Ecosystem Service Valuation: Progress, Prospects, and Challenges

Robert J. Johnston ⓘ and Dana Marie Bauer

Abstract

This article discusses prospects and challenges related to the use of meta-regression models (MRMs) for ecosystem service benefit transfer, with an emphasis on validity criteria and post-estimation procedures given sparse attention in the ecosystem services literature. We illustrate these topics using a meta-analysis of willingness to pay for water quality changes that support aquatic ecosystem services and the application of this model to estimate water quality benefits under alternative riparian buffer restoration scenarios in New Hampshire's Great Bay Watershed. These illustrations highlight the advantages of MRM benefit transfers, together with the challenges and data needs encountered when quantifying ecosystem service values.

Key Words: benefit transfer, ecosystem service, meta-analysis, stated preference, valuation, water quality, willingness to pay

JEL Classifications: Q51, Q57

Introduction

The demand for information on ecosystem service values (ESVs), combined with a lack of time and resources required to conduct primary valuation studies, has led to common use of benefit transfer to quantify these values (Plummer 2009; Bateman et al. 2011a, 2011b; Ferrini, Schaafsma, and Bateman 2015; Johnston and Wainger 2015; Richardson et al. 2015; Johnston, Rolfe, and Zawojska 2018). Benefit transfer is defined as the use of research results from preexisting primary studies at one or more sites or policy contexts (called study sites) to predict welfare estimates such as willingness to pay (WTP) or related information for other, typically unstudied sites or policy contexts (called policy sites) (Johnston et al. 2015a). Among the different approaches to benefit transfer in the literature,[1] there is

Robert J. Johnston, George Perkins Marsh Institute and Department of Economics, Clark University, Worcester, MA, USA. Dana Marie Bauer, George Perkins Marsh Institute, Clark University, Worcester, MA, USA. *Correspondence: Robert J. Johnston, George Perkins Marsh Institute Clark University, 950 Main St., Worcester, MA 01610, USA, Email: rjohnston@clarku.edu*

[1] Introductions to these approaches are provided by Johnston et al. (2015a) and Rosenberger and Loomis (2017).

emerging consensus over the advantages of methods that synthesize data from multiple sources, such as meta-analysis (e.g., Rosenberger and Phipps 2007; Boyle et al. 2009; Johnston and Rosenberger 2010; Kaul et al. 2013; Rolfe, Brouwer, and Johnston 2015; Johnston, Rolfe, and Zawojska 2018).

The most common applications of meta-analysis for ESV benefit transfer involve meta-regression models (MRMs). These models generate parametric functions that characterize the systematic influence of economic, ecological (or resource), beneficiary (population), and primary study attributes on comparable measures of value.[2] The resulting functions are used to predict values at policy sites where original valuation studies have not been conducted. Explanatory variables in these MRMs can link directly to biophysical models and enable adjustments in ESV estimates for the attributes of ecosystems, regions, populations, and policy contexts. There are many examples of such MRMs applied to different types of environmental and ecosystem service improvements including water quality (Johnston, Besedin, and Stapler 2017), recreational fishing (Johnston et al. 2006), forest recreation (Zandersen and Tol 2009), and wetlands-provided ecosystem services (Brander et al. 2012), among many others.[3]

Despite this work, further advances in methods and understanding are required if meta-analyses are to be relied on for widespread ESV prediction, particularly within large-scale applications. The capacity of MRMs to support valid and reliable[4] benefit transfers depends on multiple criteria, including the use of appropriate econometrics (Nelson and Kennedy 2009; Boyle and Wooldridge 2018) and specifications able to accommodate expected patterns such as the sensitivity of values to scope, scale, and spatial dimensions (Johnston, Besedin, and Stapler 2017; Johnston, Besedin, and Holland 2018; Johnston, Rolfe, and Zawojska 2018; Kling and Phaneuf 2018; Newbold et al. 2018a, 2018b; Smith 2018).[5] ESV applications are often deficient in these areas (Bateman et al. 2011b; Richardson et al. 2015; Johnston, Rolfe, and Zawojska 2018). Invariance of transferred welfare estimates to dimensions such as these can signal a lack of construct validity (Johnston, Besedin, and Stapler 2017).[6] The estimation of MRMs for ecosystem service valuation also

---

[2] For additional details on MRMs, see Smith and Pattanayak (2002), Bergstrom and Taylor (2006), Moeltner, Boyle, and Paterson (2007), Nelson and Kennedy (2009), Johnston and Rosenberger (2010), Boyle et al. (2013), Boyle, Kaul, and Parmeter (2015), Rolfe, Brouwer, and Johnston (2015), Boyle and Wooldridge (2018), Johnston, Besedin, and Holland (2018), and Johnston, Rolfe, and Zawojska (2018).

[3] For additional examples, see reviews in Nelson and Kennedy (2009), Johnston and Rosenberger (2010), Rolfe, Brouwer, and Johnston (2015), and Johnston, Rolfe, and Zawojska (2018).

[4] For a formal discussion of valuation validity and reliability, see Bishop and Boyle (2019).

[5] Here we define scope as the quantity or quality of an ecosystem service change for which a value is predicted; scale is defined as the geographic area over which the change takes place.

[6] As described by Bishop and Boyle (2019, p. 564), construct validity reflects the extent to which a value measure is consistent with "prior expectations about how the true value ought to be

requires robust primary study data for the type of values under study, with sufficient study reporting for metadata development (Loomis and Rosenberger 2006).

In addition, multiple procedures are required to predict and aggregate values for policy sites using MRM or other transfer functions. These procedures must address challenges such as (1) linking biophysical information on the scope and scale of environmental changes to MRM functions, (2) accounting for the extent of the market and other spatial dimensions of households and environmental changes, (3) addressing practical dissimilarities or gaps between MRM variables and measurable conditions at the policy site, (4) accounting for the systematic influence of methodological factors on value estimates, and (5) ensuring that the resulting predictions have theoretical and empirical properties required to ensure validity for the intended uses.

Concerns such as these are sometimes overlooked in the ecosystem services valuation literature, leading to many MRMs that—despite having superficially acceptable statistical performance—are poorly suited to credible benefit transfer applications. As noted by Johnston, Rolfe, and Zawojska (2018, p. 202), "Many applications [of meta-analysis in the ecosystem services] literature lack minimal properties necessary to promote valid and reliable [benefit transfer] estimates." Even if the underlying MRM is suitable for this purpose, valid transfers also depend on the post-estimation procedures used to predict ESVs from the estimated benefit function. These issues are seldom discussed. The academic literature focuses primarily on estimation and interpretation of the underlying MRMs rather than the evaluation and use of these models for benefit transfer. The practical steps and challenges involved in benefit prediction and aggregation using these MRMs—or properties of the resulting value estimates—are not often considered in detail.

This article discusses prospects and challenges related to the use of MRMs for benefit transfer. We begin with a summary of key issues and challenges associated with the use of MRMs to forecast values for large-scale ESVs, including post-estimation issues given sparse attention in the literature. We illustrate these issues using a meta-analysis of WTP for water quality changes that support aquatic ecosystem services (Johnston, Besedin, and Stapler 2017) and the application of this model to estimate aggregate water quality benefits under alternative riparian buffer restoration scenarios in New Hampshire's Great Bay Watershed. The MRM is reviewed with respect to key validity criteria to highlight both advantages and limitations of the model. Benefit transfers from this MRM are compared for multiple scenarios, scopes, and scales of environmental improvement in the case study area, aggregated over different market extents.

---

related to other variables. Such prior expectations are motivated by theory, intuition, and past empirical evidence."

The goal of this illustration is practical. Unlike publications that focus on MRM estimation and interpretation, here we emphasize often-unreported steps and choices associated with the use of MRM benefit functions for ESV prediction, as well as patterns in the resulting welfare forecasts. Attention is given to topics such as the consistency between variables used in published MRMs (typically reflecting data reported by primary studies in the metadata) and biophysical data of the type available to quantify ecosystem service changes. We also consider the properties and construct validity of the resulting value predictions. These illustrations highlight the advantages of MRM benefit transfers together with challenges and data needs.

Meta-Analysis Properties and Benefit Transfer Validity

The potential advantages of meta-analysis for benefit transfer are established in the literature (Rosenberger and Phipps 2007; Nelson and Kennedy 2009; Rosenberger and Johnston 2009; Johnston and Rosenberger 2010; Kaul et al. 2013; Rolfe, Brouwer, and Johnston 2015; Boyle and Wooldridge 2018; Johnston, Rolfe, and Zawojska 2018). Meta-regression analyses synthesize the results of multiple prior studies into a single set of parametric predictors that can be used within benefit transfer, providing a means to ground transfers in a broad base of prior information.[7] The resulting function allows predicted values to be tailored to the needs of particular policy evaluations.[8]

There is emerging consensus over the advantages of methods that synthesize data from the literature in this way, including potential improvements in benefit transfer reliability—or reductions in generalization error (e.g., Rosenberger and Phipps 2007; Boyle et al. 2009; Johnston and Rosenberger 2010; Johnston and Thomassin 2010; Kaul et al. 2013; Rolfe, Brouwer, and Johnston 2015; Johnston, Rolfe, and Zawojska 2018). The capacity of MRMs to synthesize information across studies that vary across commodity, site, policy, and population factors can obviate the need to find one study that is a close match to the policy site across all dimensions (Stapler and Johnston 2009). Meta-analytic transfers can also reduce the risk of error caused by the use of a study site that differs from the policy site in influential ways or a primary study that suffers from measurement error, and can help diagnose and ameliorate the effect of selection biases in the literature (Stanley 2005; Hoehn 2006; Rosenberger and Johnston 2009; Boyle and Wooldridge 2018).

---

[7]   Results can also help researchers characterize the literature around a topic and allow the analyst to test hypotheses regarding factors that influence study results (Smith and Pattanayak 2002; Boyle and Woodridge 2018).

[8]   Other (non-regression) types of meta-analysis may also be used for benefit transfer, such as the synthesis of prior valuation estimates to generate mean unit values, for example weighted by measures of study precision. For an introduction to different types of meta-analysis applied in economics, see Stanley and Doucouliagos (2012). Here we emphasize MRMs, as these are the most common form of meta-analysis found in the benefit transfer literature.

Despite these advantages, the ability of an MRM to provide valid and reliable predictions depends on multiple factors, including the procedures used for data synthesis, statistical analysis, value prediction, and benefit aggregation (Smith and Pattanayak 2002; Nelson and Kennedy 2009; Boyle et al. 2013; Boyle, Kaul, and Parmeter 2015; Nelson, 2015; Boyle and Wooldridge, 2018; Johnston, Rolfe, and Zawojska 2018). Considerations such as these are frequently overlooked in the ecosystem services literature, leading to transfers that violate conditions for construct validity in welfare estimates (Plummer 2009; Bateman et al. 2011b; Johnston and Wainger 2015; Richardson et al. 2015).

Prior reviews of MRMs in environmental economics find that many violate best practice guidelines for econometric analysis (Nelson and Kennedy 2009). Boyle and Wooldridge (2018) discuss additional issues related to metadata construction and MRM specification, as related to the validity and reliability of predictions. They distinguish between MRMs used to understand a body of empirical literature and those used to predict values. Among their key findings is that "no single estimated meta-equation is suitable for prediction in all benefit-transfer applications" (Boyle and Wooldridge 2018, p. 633). Care must be taken to construct metadata and specify equations to predict values grounded in a consistent theoretical definition and commodity type. Hence, "a one-size-fits-all use of an estimated meta-equation to support benefit transfers is likely not appropriate" (Boyle and Wooldridge 2018, p. 633). Kling and Phaneuf (2018), Newbold et al. (2018b), and Moeltner (2019) consider whether and when MRM specifications should be further restricted to impose theoretical properties such as adding-up on benefit transfers, in some cases reaching contrary conclusions.[9]

Among commonly acknowledged requirements for valid valuation via MRMs is at least a minimal degree of commodity and welfare consistency across metadata observations (Smith and Pattanayak 2002; Bergstrom and Taylor 2006; Loomis and Rosenberger 2006, Nelson and Kennedy 2009; Johnston and Rosenberger 2010; Boyle and Wooldridge 2018; Johnston, Rolfe, and Zawojska 2018). Moeltner and Rosenberger (2014, p. 470) characterize commodity consistency as a situation in which "basic commodities under consideration [within the metadata] must be essentially equivalent." Welfare consistency implies parallel consistency in the theoretical welfare constructs measured across observations. Many ESV MRMs are estimated using primary study data sets that pool value metrics that have no such relationship, such

---

[9] The adding-up property in MRMs relates to sequential use of the model to predict WTP for additive improvements. Assume that WTP_A is predicted WTP for a quality improvement from A to B, WTP_B is predicted WTP for a quality improvement from B to C, and WTP_C is predicted WTP for a combined improvement from A to C. The adding-up property implies that $WTP_A + WTP_B = WTP_C$, when each is predicted separately by the MRM, and utility is held constant at the original level for all calculations. For a more formal definition, see Kling and Phaneuf (2018).

as estimates of consumer value, producer value, and measures that do not reflect welfare-theoretic values of any type. Prior MRMs have also pooled value estimates linked to dissimilar goods and services such as recreation, flood control, fisheries production, carbon sequestration, raw material provision, nutrient cycling, water supply, existence, aesthetics, and others (Johnston, Rolfe, and Zawojska 2018). The lack of metadata consistency within such models raises validity concerns (Bergstrom and Taylor 2006)—neither theory nor economic intuition justifies the inclusion of such dissimilar welfare measures within the estimation of a single meta-analytic value function.[10]

Any remaining differences in commodity, site, or population characteristics that might influence value must be captured using right-hand side variables in the MRM. However, even similar environmental changes may be defined in different ways across the literature, and reconciling these measurements across studies is not always straightforward (Smith and Pattanayak 2002; Johnston et al. 2005; Van Houtven, Powers, and Pattanayak 2007; Rolfe, Brouwer, and Johnston 2015). This can lead to trade-offs between the number of studies (N) and independent variables (K) within an MRM (Moeltner, Boyle, and Paterson 2007, p. 252)—"should the researcher discard explanatory variables that are not common to all studies (thus preserve N at the cost of K) or discard observations that do not include all key regressors (thus preserve K at the cost of N)?" Accuracy can be diminished if the resulting variables are quantified in low-resolution or categorical terms (e.g., high, medium, and low) (Johnston et al. 2012; Johnston, Besedin, and Stapler 2017). Moreover, the usefulness of any valuation exercise is reduced if variables are defined in ways that cannot be directly linked to the quantitative biophysical, health, or engineering outputs used to quantify effects for policy analysis (Loomis and Rosenberger 2006; Johnston et al. 2012; Schultz et al. 2012; Boyd et al. 2016).

The validity of MRM benefit transfers further depends on the extent to which these transfers reflect theoretical patterns expected in welfare estimates (Kling and Phaneuf 2018). For example, microeconomic theory suggests that predictions of per household WTP for most ecosystem service changes should respond to the scope and scale of the change, as well as substitutes and spatial dimensions. For nonmarginal changes, diminishing marginal values are typically expected. Hence, linear "scaling-up" of value predictions

---

[10]   Recent work demonstrates that gains in transfer reliability may sometimes be achieved by relaxing narrow, ex ante commodity and welfare restrictions in favor of those supported by empirical relationships (e.g., Moeltner and Rosenberger 2014; Johnston and Moeltner 2014; Moeltner 2015). For example, in some cases reliability may be improved by pooling metadata on otherwise similar Hicksian and Marshallian welfare measures (Johnston and Moeltner 2014). Pooling of this type within MRMs has been justified based on structural relationships between these two welfare measures, grounded in microeconomic theory (Londoño and Johnston 2012).

is unlikely to produce valid estimates, particularly when scaled to a large degree (Bockstael et al. 2000; Bateman et al. 2011b; Brander et al. 2012; Johnston and Wainger 2015). Theory also suggests that most ESVs should be sensitive to household income, among other population attributes, although naïve income adjustments may not always improve transfers (Johnston and Duke 2010). Expectations such as these influence the type of MRM specifications required for validity.

A final set of questions and concerns—and those perhaps given least attention in the literature—relates to the post-estimation steps required to predict and aggregate values from a metafunction. These procedures can have a profound impact on benefit predictions. For example, as shown by Johnston, Besedin, and Ranson (2006) for the treatment of primary study methodological variables when predicting benefits, the use of ad hoc procedures can cause WTP predictions to vary more than 15-fold. This sensitivity can be reduced with best practices for post-estimation value prediction (Stapler and Johnston 2009; Boyle and Wooldridge 2018).

One often overlooked post-estimation concern relates to variable measurement conventions applied in the MRM (discussed previously), including those used to quantify the biophysical scope and scale of ecosystem service changes. First, one must assess whether the ecosystem service of interest at the policy site is a good match for the service(s) covered by the MRM. Assuming a good match exists, one must then link the relevant biophysical changes predicted at a policy site to ecosystem service measures used within the existing MRM and obtain any required data or biophysical model outputs.

For example, Rolfe, Brouwer, and Johnston (2015) meta-analyze per household WTP per kilometer of river in "good" health, where studies in the metadata define good health using different types of biophysical criteria. The usefulness of such results for benefit transfer requires a defensible and replicable way to predict this outcome (kilometers in good health) using biophysical information available to policy analysts. Even for a given benefit function, different analyst assumptions regarding biophysical measures or indices can lead to different benefit predictions. Walsh and Wheeler (2013, p. 81) demonstrate that different assumptions regarding the mathematical structure of water quality indices "can have a profound effect on benefits." In addition, MRM variables are often defined in ways that are biophysically ambiguous, potentially exacerbating the challenge of linking measured policy effects to MRM benefit estimates.

A second post-estimation concern relates to the role of spatial dimensions in benefit estimation, including the assigned extent of the market for benefit prediction and aggregation. As noted by Bateman et al. (2006, 2011a), Schaafsma (2015), Johnston et al. (2017), and Johnston, Besedin, and Holland (2018), choices such as these can swamp the effects of other procedural choices in benefit transfer. A related challenge is that—depending on the MRM specification—average and aggregate benefits may vary depending on whether average benefits are predicted for one large market or multiple

component submarkets. In such cases, is it sufficient to predict values assuming one constant "average" ecosystem service change and baseline over one large market area? Or, should values be predicted for smaller areas, reflecting the specific changes and baselines in those areas? If one chooses the latter approach, what is the appropriate size of the subareas to be considered? When values depend on spatial dimensions (Schaafsma 2015; Glenk et al. 2019), predicted aggregate values for average environmental changes everywhere will not, in general, be equal to the same values averaged over the population for spatially heterogeneous changes. Similar questions apply to other contextual considerations that vary over space such as income and substitutes.

A third consideration is whether the desired policy site ESV scenarios are suitable for valuation using the MRM in question. For example, would a particular scenario require values to be predicted outside the data range of the primary study metadata? With a few exceptions, the literature provides little guidance for these post-estimation procedures, leading to wide variation in applied practice. Compounding this challenge is a tendency of these procedures and assumptions to remain unreported, leading to lack of methodological transparency.

Illustrative Application: Predicting Water Quality Benefits from Buffer Restoration

This section illustrates the use of an applied MRM for ESV benefit transfer. Among the goals of this illustration are to (a) demonstrate the characteristics of the MRM required for valid benefit transfers, (b) clarify the post-estimation steps and choices required to generate and evaluate predictions using the model, and (c) discuss the advantages and disadvantages of the resulting benefit predictions. The intent is to clarify often-undocumented properties and practices that affect transfer value validity. The empirical case study analysis implements a benefit transfer to quantify per household and aggregate market WTP associated with water quality improvements of the type that could result from the restoration of vegetated buffers within New Hampshire's Great Bay watershed. We use an existing MRM for this purpose to enable focus on model evaluation and post-estimation benefit transfer, rather than on estimation of an original MRM.

Benefits are predicted using an application of the MRM estimated by Johnston, Besedin, and Stapler (2017). This model was designed to support benefit transfer for water quality improvements in U.S. water bodies including rivers, lakes, and estuaries and reflects the type of MRM used by the U.S. Environmental Protection Agency (U.S. EPA) for regulatory analysis (U.S. EPA 2009a, 2010, 2012, 2015; Griffths et al. 2012). The metadata were drawn from primary stated preference studies that estimated total (use and nonuse) per household WTP for water quality changes in water bodies that

support ecosystem services including aquatic life, recreational uses (e.g., fishing, boating, and swimming), and nonuse values.[11] This MRM and metadata have been used as the basis for multiple prior validity tests and analyses (e.g., Newbold et al. 2018b; Moeltner 2019), so that the underlying properties of the meta-analysis are well known. As such, this MRM provides an established foundation from which to demonstrate an applied benefit transfer.

Grounded in this MRM, we present the process used to apply the resulting transfer function, along with assumptions and implications for valid benefit transfer. To illustrate how transferred WTP estimates vary across different types of water bodies, we predict values for improvements to the Great Bay and two of its tributaries, the Squamscott River and Exeter River. Values for improvements to each of these areas are predicted independently, holding conditions at other areas constant. We consider various scenarios of water quality change for these water bodies, based on current baseline quality. We also consider values predicted and aggregated over different market areas. The goal is to consider the responsiveness of WTP estimates to factors such as scope, scale, extent of the market, and baseline condition. We also describe the water quality index (WQI) used to quantify water quality baselines and changes within the MRM, as well as how the index is defined as a function of water quality parameters.

The Meta-Regression Model

As described by Johnston, Besedin, and Stapler (2017), all observations in the original metadata were identified and coded following the guidelines of Stanley et al. (2013). To ensure welfare consistency, observations were restricted to U.S. studies that estimated total (use and nonuse) value, used established stated preference methods, reported comparable Hicksian WTP measures, and provided sufficient detail to verify that valuation methods met minimum quality standards. To ensure commodity consistency, studies were limited to those for which per household WTP estimates could be mapped to water quality changes measured on a standardized 100-point WQI that relates pollutant concentrations to water body suitability for human uses (Walsh and Wheeler 2013). The final metadata included 140 unique observations from 51 stated preference studies published between 1985 and 2013 (Table 1).

The dependent variable in the MRM is the natural log of per household WTP for water quality improvements specified in each original study, with all values adjusted to 2007 USD (these values are subsequently updated to 2016 USD as discussed in Implementing the Benefit Transfer below).

---

[11] The metadata exclude studies focusing on drinking-water quality.

Table 1. Primary Studies in the Metadata (willingness to pay [WTP] is per household per year in 2007 USD)

Reference	Observations	State(s)	Water Body Type(s)	Mean WTP
Aiken (1985)	1	CO	River and lake	193.18
Anderson and Edwards (1986)	1	RI	Salt pond/marsh	180.71
Banzhaf et al. (2006)	2	NY	Lake	57.47
Banzhaf et al. (2011)	1	VA, WV, TN, NC, GA	River/stream	31.30
Bockstael et al. (1988)	1	DC, MD, VA	Estuary	149.03
Bockstael et al. (1989)	2	MD	Estuary	158.30
Borisova et al. (2008)	3	WV, VA	River/stream	44.94
Cameron and Huppert (1989)	1	CA	Estuary	49.53
Carson et al. (1994)	2	CA	Estuary	59.40
Clonts and Malone (1990)	3	AL	River/stream	103.20
Collins and Rosenberger (2007)	1	WV	River/stream	18.19
Collins et al. (2009)	7	WV	River/stream	120.52
Corrigan et al. (2009)	1	IA	Lake	123.30
Croke et al. (1986)	9	IL	River/stream	77.47
De Zoysa (1995)	1	OH	River/stream	70.18
Desvousges et al. (1987)	12	PA	River/stream	59.19
Downstream Strategies (2008)	2	PA	River/stream	12.74
Farber and Griner (2000)	6	PA	River/stream	76.16
Hayes et al. (1992)	2	RI	Estuary	397.44
Herriges and Shogren (1996)	2	IA	Lake	134.55
Hite (2002)	2	MS	River/stream	60.08
Huang et al. (1997)	2	NC	Estuary	258.65
Irvin et al. (2007)	4	OH	All freshwater	21.67
Johnston et al. (1999)	1	RI	River/stream	180.95
Kaoru (1993)	1	MA	Salt pond/marsh	218.61
Lant and Roberts (1990)	3	IA, IL	River/stream	143.93
Lant and Tobin. (1989)	9	IA, IL	River/stream	55.63
Lichtkoppler and Blaine (1999)	1	OH	River and lake	41.93

Table 1. Continued

Reference	Observations	State(s)	Water Body Type(s)	Mean WTP
Lindsey (1994)	8	MD	Estuary	66.80
Lipton (2004)	1	MD	Estuary	63.98
Londoño Cadavid and Ando (2013)	2	IL	River/stream	38.68
Loomis (1996)	1	WA	River/stream	93.07
Lyke (1993)	2	WI	River and lake	78.75
Matthews et al. (1999)	2	MN	River/stream	21.73
Opaluch et al. (1998)	1	NY	Estuary	138.47
Roberts and Leitch (1997)	1	MN, SD	Lake	8.35
Rowe et al. (1985)	1	CO	River/stream	134.59
Sanders et al. (1990)	4	CO	River/stream	160.69
Schulze et al. (1995)	2	MT	River/stream	20.84
Shrestha and Alavalapati (2004)	2	FL	River and lake	156.46
Stumborg et al. (2001)	2	WI	Lake	84.29
Sutherland and Walsh (1985)	1	MT	River and lake	146.03
Takatsuka (2004)	4	TN	River/stream	286.88
Wattage (1993)	3	IA	River/stream	53.89
Welle (1986)	6	MN	Lake	167.28
Welle and Hodgson (2011)	3	MN	Lake	145.10
Wey (1990)	2	RI	Salt pond/marsh	147.26
Whitehead and Groothuis (1992)	3	NC	River/stream	41.01
Whitehead (2006)	3	NC	River/stream	187.18
Whitehead et al. (1995)	2	NC	Estuary	95.44
Whittington et al. (1994)	1	TX	Estuary	194.72

Note: Complete citations to these referenced works are provided by Johnston, Besedin, and Stapler (2017).

Independent variables expected to explain variation in this value measure (and included in the model) characterize (1) the geographic region and affected aquatic resources, (2) affected ecosystem services, (3) populations whose values were measured, (4) baseline resource condition and water quality change, (5) potential substitute resources and complementary land uses, and (6) the primary study methodology and year (Table 2). Emphasis was given to core economic and resource variables directly relevant to

Table 2. Meta-Analysis Variable Descriptions and Mean Metadata Values

Variable	Description	Mean
Ln_BaseQuality	Natural log of the baseline water quality from which improvements would occur, specified on the 100-point water quality index (WQI).	3.589
Ln_QualityChg	Natural log of the change in water quality, specified on the 100-point WQI.	2.907
Ln_Income	Natural log of median household income (in 2007 USD) for the market area based on historical U.S. Census data.	10.745
Non_Users	Binary variable indicating that the survey was implemented over a population of nonusers only (default is a sampled population not restricted to nonusers).	0.086
Swim_Use	Binary variable indicating that changes in swimming uses are specifically noted in the survey.	0.264
Boat_Use	Binary variable indicating that changes in boating uses are specifically noted in the survey.	0.114
Game_Fish	Binary variable indicating that changes in game fishing uses are specifically noted in the survey.	0.057
River	Binary variable indicating that the focal resource is a river or multiple rivers.	0.686
Multi_Body	Binary variable indicating the focal resource includes multiple water body types (e.g., rivers and estuaries combined).	0.078
Ln_PropAgLand	Natural log of the proportion of the land area in all counties that intersect the improved focal resource that is agricultural land based on the National Land Cover Database.	−1.433
Ln_RelativeSize	Natural log of the total shoreline length (in kilometers) of the improved focal resource divided by the size of the market area (in square kilometers). For a river, shoreline length is given by two times the length of the river. For a bay, shoreline length is the perimeter of the bay, not including tributaries.	−1.198
ProportionChg	Proportion of water bodies of the same hydrologic type as the improved focal resource, within affected state(s). For rivers, this is measured as the length of the improved river divided by the length of all rivers of the same or lower order (*PropChgRiver*). For bays and estuaries, this is defined as the shoreline length of the water body as a proportion of all analogous (e.g., coastal) shoreline lengths (*PropChgBay*). *ProportionChg* is defined as the maximum of *PropChgRiver* or *PropChgBay*.	0.188
Northeast_US	Binary variable indicating that the survey included respondents from the U.S. Department of Agriculture (USDA) Northeast region.	0.071

Table 2. Continued

Variable	Description	Mean
Central_US	Binary variable indicating that the survey included respondents from the USDA Midwest or Mountain Plains region.	0.336
Southern_US	Binary variable indicating that the survey included respondents from the USDA Southeast or Southwest.	0.157
MedianWTP	Binary variable indicating that the study's willingness-to-pay (WTP) measure is the median rather than the mean.	0.071
LumpSum	Binary variable indicating that payments were to occur on something other than an annual basis over an extended or indefinite period of time.	0.186
Ln_StudyYear	Natural log of the year in which the primary study was conducted (converted to an index by subtracting 1980, before making the log transformation).	2.212
ChoiceExp	Binary variable with a value of 1 for studies that are choice experiments (default is contingent valuation).	0.107
Thesis	Binary variable with a value of 1 for studies published as thesis projects or dissertations (default is all other publication types, including but not limited to journal articles).	0.144
Voluntary	Binary variable indicating that WTP was estimated using a payment vehicle described as voluntary (default is payment vehicle described as binding or mandatory).	0.086
OutlierBids	Binary variable indicating that outlier bids were excluded when estimating WTP.	0.193
NonParametric	Binary variable indicating that WTP was estimated using nonparametric methods.	0.429
NonReviewed	Binary variable indicating that the study was not published in a peer-reviewed journal.	0.236

benefit transfer. To construct the data set, variables reported directly by primary studies were supplemented with data on study sites, beneficiary populations, and water bodies extracted from readily available national data sets.[12]

Among key variables in the MRM is that characterizing the scope (or size) of the water quality change, measured as a natural log (*Ln_QualityChg*). The model also includes a spatial index variable (*Ln_RelativeSize*) that characterizes the

---

[12] External data were extracted from the National Hydrography Dataset; Hydrologic Unit Code Watershed Boundary Dataset; National Land Cover Database; NOAA Global Self-Consistent, Hierarchical, High-Resolution Geography Database; and U.S. Census.

size of the affected water body measured using shoreline length in kilometers (geospatial scale), relative to the size of the sampled market area measured in square kilometers (market extent). This specification enables the marginal effect of water body size on WTP to decline as size of the sampled market area increases, and vice versa. The inclusion of market area in this composite index also provides a means to proxy for the effect of distance decay on value, as larger sampled market areas generally imply greater distances to affected water bodies, ceteris paribus (Johnston, Besedin, and Stapler 2017).

To characterize the scope of proportional effects on regional (potentially substitute) water bodies, the variable *ProportionChg* measures the proportion of water bodies (of the same hydrologic type) affected by the water quality change, within each state. Potential land use complements to water quality (specifically, the lack of complementary land uses) are characterized using *Ln_PropAgLand*, representing the (natural log of the) proportion the affected area with agricultural land use.[13] The model also includes variables characterizing affected uses (or ecosystem services) mentioned in each primary study such as fishing and swimming (or nonuse only), as well as indicators identifying the region of the United States where changes occurred.

The MRM is estimated using unweighted random effects ordinary least squares with robust standard errors, allowing for cross-sectional correlation among observations from the same study (Nelson and Kennedy 2009).[14] The model is specified as

$$\bar{y}_{js} = \bar{x}_{js}\beta + \varepsilon_{js}, \tag{1}$$

where $\bar{y}_{js}$ is the welfare measure for observation s in study j (the natural log of mean or median WTP per household for the specified water quality change), and $\bar{x}_{js}$ is the vector of independent variables. This vector includes natural logs of household income (*Ln_Income*), water quality change and baseline (*Ln_QualityChg*; *Ln_BaseQuality*), relative geospatial scale (*Ln_RelativeSize*), and agricultural land proportion (*Ln_PropAgLand*). Other variables enter

---

[13] Affected areas are defined as counties that intersect or touch affected water bodies. The rationale for this variable is that nonagricultural land uses (e.g., forests and residential) are often associated with recreational and other human uses that can magnify the per household value of water quality improvements (Johnston, Besedin, and Stapler 2017).

[14] An alternative approach would be to weight metadata observations such that each study (rather than each observation) is given equal weight—both weighted and unweighted approaches are common in the literature (Boyle and Wooldridge 2018). As noted by Johnston et al. (2005, p. 231), "Although weighting methods prevent studies providing multiple observations from unduly influencing model estimation, they also imply that such studies are no more informative, overall, than others (Bateman and Jones 2003)." Hence, the decision to weight observations in an MRM is typically made on a case-by-case basis. Using earlier versions of these metadata, Johnston, Besedin, and Wardwell (2003) and Johnston et al. (2005) find that results are robust to the use of weighted versus unweighted regression.

linearly.[15] The vector β represents a conforming parameter vector. Following standard specifications, $\varepsilon_{js} = e_{js} + u_s$, where u_s represents a systematic, normally distributed, study-level random effect with $E(u_s) = 0$ and $\text{Var}(u_s) = \sigma_u^2$; and e_{js} is a standard iid (independent and identically distributed) estimation level error, distributed with a zero mean and constant variance σ_e^2.

Results are shown in Table 3. Parameter estimates are jointly significant at p<0.01 ($\chi^2 = 729.61$, df = 24), with an R^2 of 0.633. Signs of significant parameter estimates match those suggested by theory. For example, WTP is positively related to the scope of water quality change (*Ln_QualityChg*), household income (*Ln_Income*), improvements to larger water bodies and samples over smaller market extents (*Ln_RelativeSize*), and improvements that affect a larger proportion of surrounding waters (*ProportionChg*). WTP is negatively related to improvements in more agricultural areas (*Ln_PropAgLand*) and samples limited to nonusers (*Non_Users*). The functional form implies diminishing marginal utility with regard to scope and scale. Other properties of the underlying MRM are discussed by Johnston, Besedin, and Stapler (2017).

Implementing the Benefit Transfer

Estimating the MRM is only the first step in benefit transfer. Few articles discuss the procedures required to move from the estimated MRM to benefit transfer predictions, along with the choices and challenges that are involved. The objective of this section is to make these procedures transparent and replicable. The basic steps of an ESV transfer are described by Johnston and Wainger (2015).[16] Two initial steps include establishing the need for benefit transfer and developing the conceptual foundation for the exercise. The latter involves development of an implicit or explicit "conceptual model of relationships between ecosystem processes and human benefits, including the biophysical pathways through which benefits are realized and their connections to different beneficiary groups" (Johnston and Wainger 2015, p. 247). Such a model should clarify the linkages between actions, changes in ecosystem services, and the effect of these changes on utility (Bateman et al. 2011b; Olander et al. 2018). This conceptual model is developed within a valuation context in which management scenarios, ecosystem services, and affected populations can be defined. ESVs at policy sites are predicted conditional on a set of biophysical changes in ecosystem services. These

---

[15] This specification allows the MRM to reflect curvature in the valuation function, a multiplicative rather than additive effect of independent variables, and the implied constraints that WTP approaches zero when water quality change, income, and the geospatial index variable (*Ln_RelativeSize*) approach zero.

[16] Similar steps are described by Bateman et al. (2011b) for ecosystem services valuation in general.

Table 3. Meta-Regression Results—Random Effects Model (Johnston, Besedin, and Stapler et al. 2017)

Variable	Parameter Estimate	Standard Error
Ln_BaseQuality	−0.064	(0.123)
Ln_QualityChg	0.281	(0.106)***
Ln_Income	0.628	(0.375)*
Non_Users	−0.455	(0.121)***
Swim_Use	−0.391	(0.220)*
Boat_Use	−0.314	(0.183)*
Game_Fish	0.303	(0.207)
River	−0.226	(0.128)*
Multi_Body	−0.525	(0.145)***
Ln_PropAgLand	−0.351	(0.095)***
Ln_RelativeSize	0.052	(0.019)***
ProportionChg	0.525	(0.189)***
Northeast_US	0.549	(0.249)**
Central_US	0.601	(0.112)***
Southern_US	1.366	(0.127)***
MedianWTP	−0.264	(0.239)
LumpSum	0.727	(0.136)***
Ln_StudyYear	−0.478	(0.080)***
ChoiceExp	0.487	(0.210)**
Thesis	0.557	(0.195)**
Voluntary	−1.296	(0.209)***
OutlierBids	−0.429	(0.120)***
NonParametric	−0.477	(0.126)***
NonReviewed	−0.679	(0.171)***
Intercept	−2.281	(4.225)
R^2	0.63	
σ_ε	0.541	

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

changes can be quantified in various ways, including biophysical models, field observations, or the use of hypothetical scenarios.

Here, we investigate three distinct water-quality change scenarios for three different water bodies within the Great Bay watershed of New Hampshire, USA (Figure 1): the Great Bay Estuary itself, not including tributaries
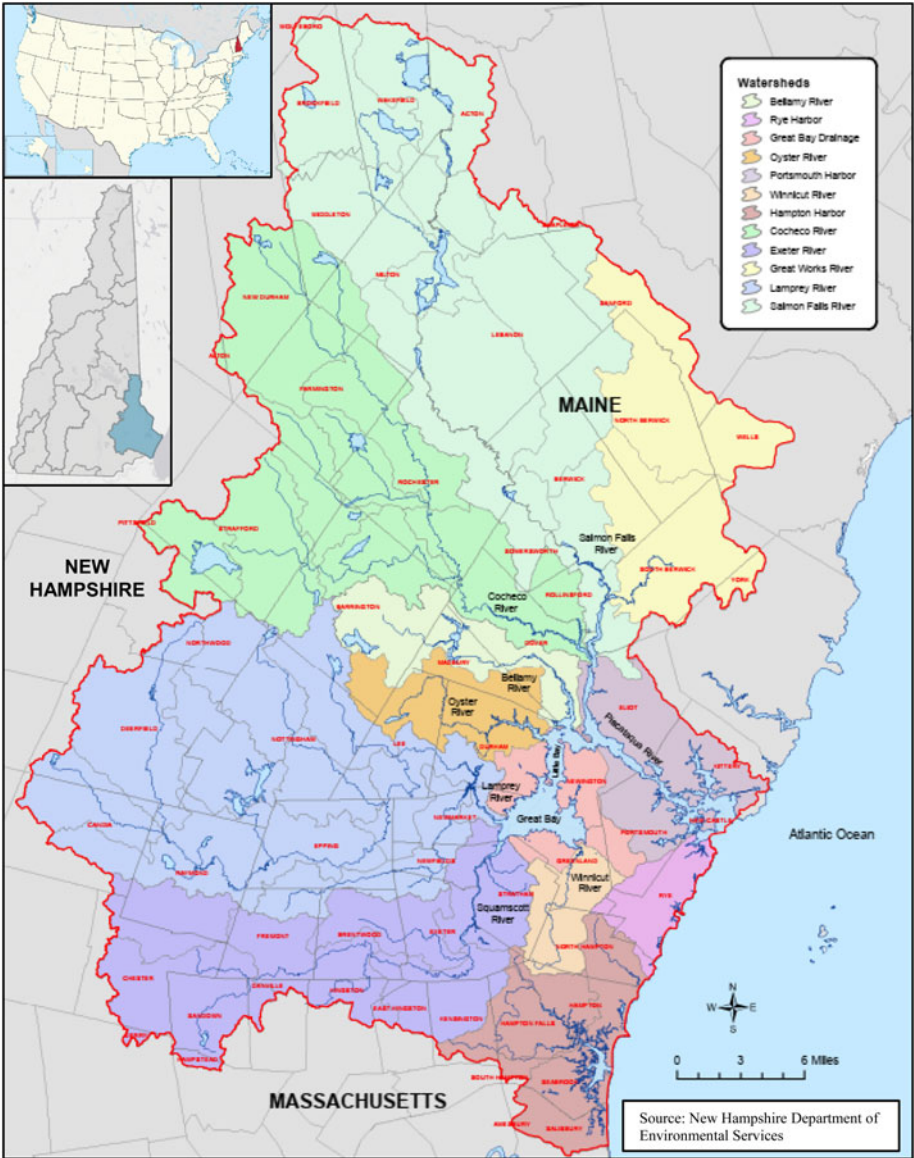
Figure 1. Great Bay Watershed in New Hampshire, USA

([Figure 2](#)), and the freshwater and tidal portions of the Exeter-Squamscott tributary ([Figure 3](#)). These define the geospatial scale of changes to be considered. Market extents must also be defined—the location of households for whom values are estimated. For the Great Bay Estuary, we evaluate WTP over three different market areas: (1) residents in NH towns immediately
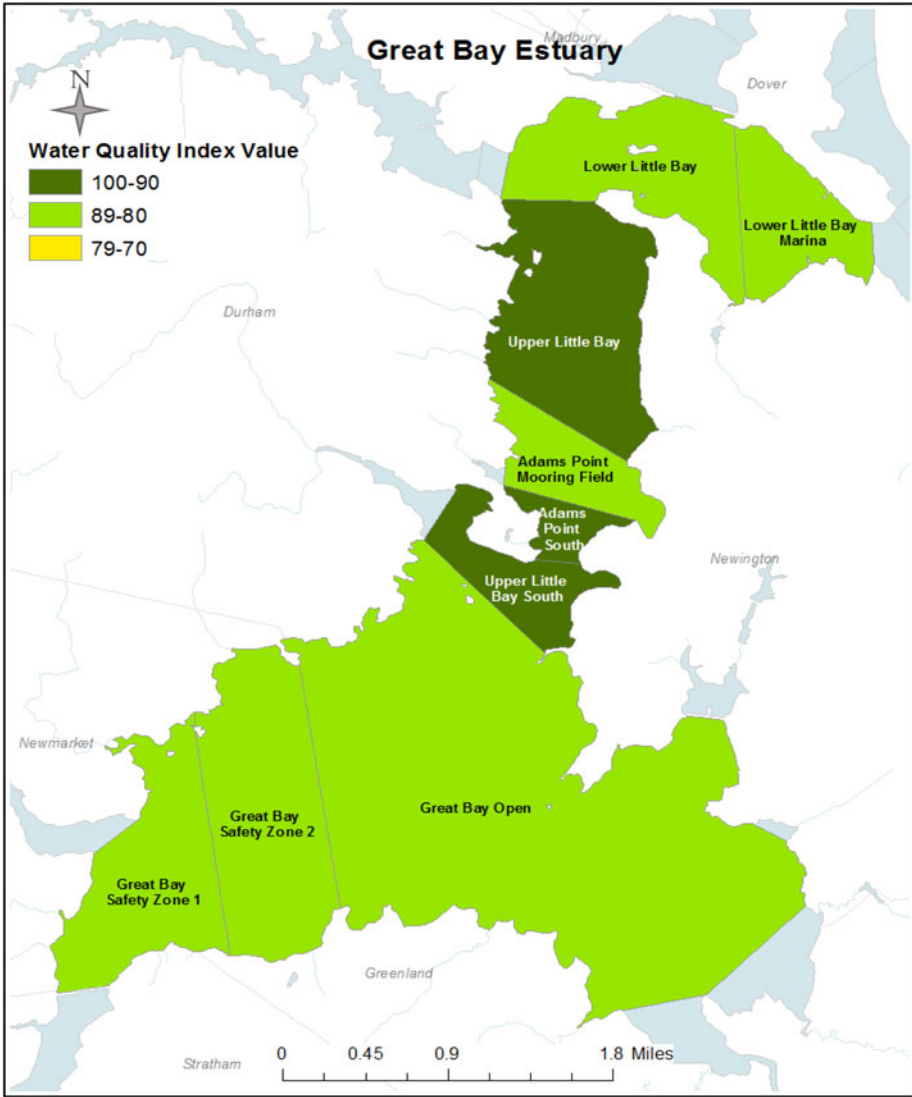
Figure 2. Great Bay Estuary Major Assessment Units and Baseline Water Quality (Table 6)

Note: Water quality index values are calculated using equation (2).

adjacent to the bay (Figure 2), (2) residents of NH towns within the entire Great Bay watershed (Figure 1), and (3) all residents of New Hampshire (Figure 1, inset). For the Exeter-Squamscott River, we evaluate WTP for residents in towns adjacent to the upper or lower portion of the river, respectively (Figure 3). The water quality improvements to be considered include those
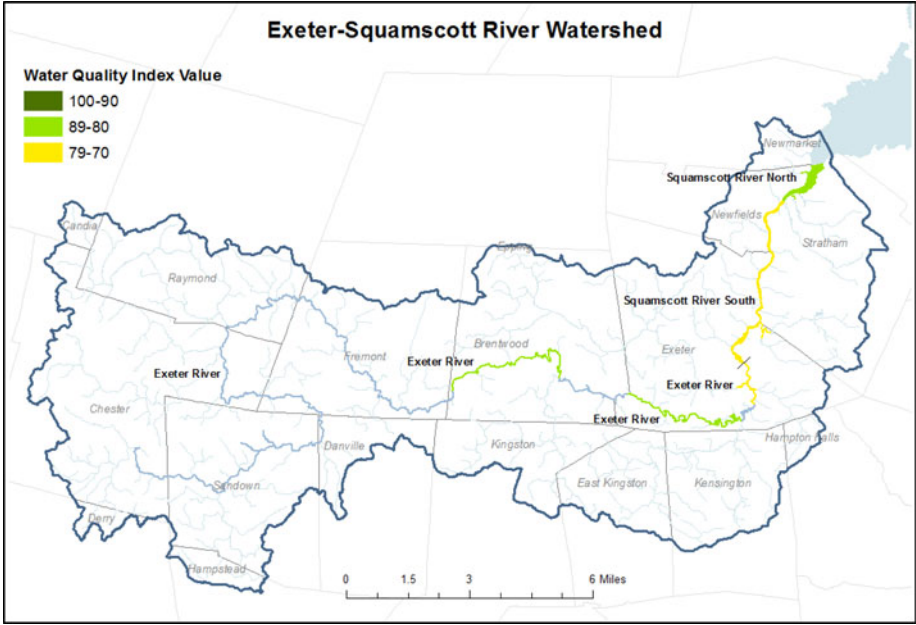
Figure 3. Exeter-Squamscott River Watershed, a Subwatershed in the Southern Portion of the Great Bay Watershed (Hydrologic Unit Code 0106000308)

Notes: The Exeter River is the freshwater portion of the river from the headwaters to the Exeter town center (indicated by hash mark across river), and the Squamscott River is the tidal portion of the river from the Exeter town center to the Great Bay. Baseline water quality is shown for select river segments (Table 6). Water quality index values are calculated using equation (2).

that could potentially arise because of intensive restoration of vegetated riparian buffers around the focal water bodies, of the general type discussed by Johnston, Holland, and Yao (2016) for the nearby Wells National Estuarine Research Reserve. Restoration of this type can improve water quality through various processes, such as nutrient uptake, reduction of storm-water runoff, and shoreline stabilization by riparian buffer vegetation.

Quantifying Water Quality Change

A core component of any ESV benefit transfer is quantification of the change in ecosystem services for which values are to be estimated, in units that are meaningful for value prediction. This requires the definition of biophysical units that link ecosystem service changes at the policy site to variables in the MRM that characterize baselines and changes (i.e., scope). Here, this is done using a standard 100-point WQI to quantify water quality changes of the type that could occur because of riparian buffer restoration (Walsh and Wheeler

2013). As described by Johnston, Besedin, and Holland (2018), WQIs combine information on physical and chemical water quality parameters into a single index linked to the ecosystem services or uses provided by a water body (Abbasi 2012; Walsh and Wheeler 2013; Van Houtven et al. 2014). They are among the most common means to measure water quality for valuation and benefit transfer (Griffiths et al. 2012; Walsh and Wheeler 2013). Here we apply the WQI methodology and classification of U.S. EPA (2009b), adapted from Cude's (2001) Oregon Water Quality Index.

Implementing the WQI to quantify baselines and changes entails three steps: (1) obtaining pollutant data and projections for the water body, (2) transforming these data into subindex values, and (3) combining the subindex values into an aggregate WQI score. The pollutants used by the WQI, along with their required units of measure and associated WQI subindex weights, are shown in Table 4. To establish baselines, pollutant data were obtained from the New Hampshire Department of Environmental Services (NHDES). These data were averaged across all sampling periods and monitoring stations for several NHDES Water Quality Assessment Units in each of our three focal water bodies to produce WQI pollutant parameter values for each pollutant subindex.[17] These pollutant parameter values were transformed into corresponding subindex values using functions and thresholds in Table 5.[18]

Finally, the subindex values weights were used to calculate the WQI for each major water body using the weighted geometric mean, as described by Walsh and Wheeler (2013),

$$WQI = \prod_{i=1}^{6} Q_i^{W_i}, \tag{2}$$

where Q_i is the calculated water quality subindex for parameter i and W_i is the weight of the ith parameter from Table 4. The resulting baseline WQI values for each assessment unit are shown in Table 6 and vary across study areas (Figures 2 and 3). U.S. EPA's (2009b) water quality classification identifies the minimum WQI value on a 100-point scale required for human uses (e.g., drinking, swimming, fishing, and boating), and links to the MRM variables *Ln_QualityChg* and *Ln_BaseQuality*, both of which are measured in WQI units.

---

[17] We investigate water quality in each assessment unit rather than averaging pollutant data across the entire water body in order to produce a range of water quality values that could then be used in sensitivity analyses. Average pollutant concentration values for each assessment unit are listed in Bauer and Johnston (2017).

[18] These functions are drawn from U.S. EPA (2009b, tables 10-1 and 10-3 and appendix F). There are six water quality subindices in each WQI; however, note that the WQI for freshwater rivers and streams includes biochemical oxygen demand, whereas the WQI for estuaries includes chlorophyll-a.

Table 4. Water Quality Index (WQI) Pollutants, Concentration Units, and Index Weights

Pollutant	Unit	Freshwater WQI Weight	Estuarine WQI Weight
Dissolved oxygen (DO)	mg/L	0.24	0.26
Fecal coliform (FC)	colonies/100 mL	0.22	0.25
Total nitrogen (TN)	mg/L	0.14	0.15
Total phosphorous (TP)	mg/L	0.14	0.15
Total suspended solids (TSS)	mg/L	0.11	0.11
Biochemical oxygen demand (BOD)	mg/L	0.15	−–
Chlorophyll-a (ChA)	μg/L	−–	0.08

Table 5. Water Quality Index Parameter-Subindex Equations

Parameter	Value	Subindex
DO	$DO \leq 3.3$	10
	$3.3 < DO < 10.5$	$-80.29 + 31.88*DO - 1.401*DO^2$
	$10.5 \leq DO$	100
FC	$FC \leq 50$	98
	$50 < FC \leq 1{,}600$	$98 * \exp[-0.00099178*(FC - 50)]$
	$1{,}600 < FC$	10
TN	$TN \leq 3$	$100 * \exp(-0.4605*TN)$
	$3 < TN$	10
TP	$TP \leq 0.25$	$100 - 299.5*TP - 0.1384*TP^2$
	$0.25 < TP$	10
TSS	$TSS \leq 28$	100
	$28 < TSS \leq 168$	$158.48 * \exp(-0.0164*TSS)$
	$168 < TSS$	10
BOD	$BOD \leq 8$	$100 * \exp(-0.1993*BOD)$
	$8 < BOD$	10
ChA	$ChA \leq 40$	$100 * \exp(-0.05605*ChA)$
	$40 < ChA$	10

Note: See Table 4 for definition of abbreviations.

Although these calculations estimate water quality baselines for our case study sites, they are not sufficient to predict water quality changes from particular management interventions at these sites. In some cases, these changes can be quantified using modeling exercises that predict changes in

Table 6. Baseline Water Quality Estimates

Water Body	NHDES Assessment Unit (ID)	Type	Baseline WQI
Exeter River[a]	Exeter River–Brentwood (NHRIV600030803-05)	River/Stream	85
Exeter River[a]	Exeter River–Exeter (NHRIV600030805-02)	River/Stream	84
Exeter River[a]	Exeter River–Exeter Dam (NHIMP600030805-04)[b]	Impoundment	77
Squamscott River	Squamscott River South (NHEST600030806-01-01)	Estuary	71
Squamscott River	Squamscott River North (NHEST600030806-01-02)	Estuary	86
Great Bay	Great Bay Safety Zone 1 (NHEST600030904-02)	Estuary	87
Great Bay	Great Bay Safety Zone 2 (NHEST600030904-03)	Estuary	85
Great Bay	Great Bay Open (NHEST600030904-04-05)	Estuary	89
Great Bay	Adams Point South (NHEST600030904-04-06)	Estuary	92
Great Bay	Upper Little Bay South (NHEST600030904-06-12)	Estuary	93
Great Bay	Adams Point Mooring Field (NHEST600030904-06-10)	Estuary	84
Great Bay	Upper Little Bay (NHEST600030904-06-19)	Estuary	91
Great Bay	Lower Little Bay (NHEST600030904-06-18)	Estuary	88
Great Bay	Lower Little Bay Marina (NHEST600030904-06-14)	Estuary	89

[a] Water quality data were limited for much of the Exeter River. The "Brentwood" assessment unit was the farthest upstream unit that contained a relatively complete set of pollutant data.
[b] Beginning in 2016, impoundment area NHIMP600030805-04 behind the Exeter River dam became part of river area NHRIV600030805-32.
Note: NHDES, New Hampshire Department of Environmental Services; WQI, water quality index.

relevant ecosystem service or environmental quality measures that would occur because of specified actions. Spatially explicit forecasts are typically required (Bateman et al. 2011a, 2011b; Ferrini, Schaafsma, and Bateman 2015; Johnston, Besedin, and Holland 2018; Glenk et al. 2019). In many situations, however, models sufficient to predict these changes are unavailable (Olander et al. 2017). In such cases, changes are sometimes quantified using illustrative "what if" scenarios reflecting future conditions that might occur (Johnston and Wainger 2015).

Here, we illustrate benefit transfers using scenarios of potential water quality improvements in affected areas, including 3, 5, 7, and 9-point increases on a 100-point WQI, from minimum baseline conditions in each water body (Table 6). These levels were chosen to fall in the general range of similar changes considered by Johnston, Holland, and Yao (2016) for estuarine systems in Maine, also potentially caused by buffer restoration.[19] Using equation (2) and information in Tables 5 and 6, one can further calculate changes in water quality parameters that would be sufficient to generate these WQI changes. For each level of improvement, we forecast WTP for each geospatial scale (water body) and market extent introduced previously. We predict per household and aggregate population-level WTP for each market area.

Setting MRM Variable Levels

Benefit transfer also requires values (or levels) to be chosen for other independent variables in the MRM (Table 2). These levels are inserted into equation (1) for $\bar{x}_{js}$ to predict WTP. Variable levels are chosen to reflect current conditions at the policy site. Selection of these levels often requires intermediate calculations using external data such as spatial landscape (GIS) metrics and U.S. Census data, whereas other values are chosen based on anticipated policy changes. Variables for study methodology are typically assigned their mean values over the metadata, unless other levels reflect "best practices" associated with reduced measurement errors in primary studies (Johnston, Besedin, and Ranson 2006; Stapler and Johnston 2009; Boyle and Wooldridge 2018).

Within the present application, levels for *Ln_PropAgLand*, *Ln_RelativeSize*, *ProportionChg*, and their underlying geospatial components (e.g., shoreline length, watershed area, town area, county area, and agricultural land area) are calculated using GIS data layers available for the study area, based on our predefined policy scenarios and market areas. These calculations follow variable definitions in Table 2. For example, *Ln_RelativeSize* is calculated by dividing the shoreline length of the focal resource (river or bay) by the size of the market area and then calculating the natural log of this ratio.

Median household income (*Ln_Income*) and number of households for towns, counties, and states were obtained from U.S. Census data for 2015 (2011–2015

---

[19] These water quality changes are larger than one would typically experience due to large-scale regulatory changes but can occur in small areas (e.g., because of extensive restoration). As described by Newbold et al. (2018a, p. 475, citing U.S. EPA 2015), "Most contemporary EPA regulations promulgated under the Clean Water Act are estimated to improve water quality by <1% in the vast majority of water bodies." In contrast, choice experiment attribute levels in Johnston, Holland, and Yao (2016) for intensive riparian restoration at the Wells National Estuarine Research Reserve (restoring up to 500 local riparian acres) allow for up to a maximum of a 20 percentage point increase on a 100-point index of ecological condition, based on "ecological models, field studies, and expert consultations" (p. 743).

American Community Survey 5-year estimates; https://www.census.gov/programs-surveys/acs/). Income for the Great Bay watershed was approximated using a household-weighted average for Rockingham and Strafford counties. Household incomes for groups of communities (e.g., communities adjacent to the Exeter River) were calculated as a household-weighted average across the communities.[20] The resulting variable levels are presented in Table 7.

Values for the remaining (non-methodological) variables were selected based on the scenario. Because none of the scenarios involved multiple geographically distinct water body types, *Multi_Body* = 0. The Squamscott and Exeter scenarios include a river, so *River* = 1. We are interested in forecasting WTP for users and nonusers (*Non_Users* = 0) in New Hampshire (*Northeast_US* = 1) for water bodies supporting three recreational uses (*Swim_Use* = 1, *Boat_Use* = 1, and *Game_Fish* = 1). For methodological variables, we assume an annual, mandatory, mean payment (*LumpSum* = 0, *Voluntary* = 0, and *MedianWTP* = 0), with *StudyYear* = 2017. We also set *OutlierBids* = 1 and *NonReviewed* = 0, under the assumption that these settings reflect best practices.[21] Mean values over the metadata (Table 2) are used for all remaining variables.

Calculating and Aggregating Values

The use of these variable levels for $\bar{x}_{js}$ within equation (1) provides an estimate of $\bar{y}_{js}$, or the natural log of per household WTP, for the particular water quality change scenario, affected area, and assumed market extent. Exponentiating this predicted value according to $WTP = \exp(\hat{y}_{js} + \sigma_\varepsilon^2/2)$, provides an estimate of per household WTP, where σ_ε^2 is the model error variance from Table 3 (Johnston and Besedin 2009). The resulting estimate is denominated in 2007 U.S. dollars and can be converted to current dollars using standard CPI adjustment (here we adjust to 2016 dollars). As this value reflects a mean per household WTP estimate, it can be aggregated across households within the specified market area to estimate aggregate population-level WTP.

This procedure is illustrated in Table 8 for the Squamscott River, 9-point WQI increase scenario, using the minimum baseline water quality from Table 6. Communities in this region include Exeter, Newfields, and Stratham, with a median household income of $86,305.[22] Given these conditions, the benefit

---

[20] Because the meta-regression analysis used 2007 U.S. dollars in its estimation process, we convert U.S. Census income data from 2015 USD to 2007 USD using the consumer price index (CPI) from the U.S. Bureau of Labor Statistics (https://www.bls.gov/cpi/).

[21] This follows Boyle and Wooldridge (2018) in setting levels for selected variables based on assumed "best practices," with others set at mean values. For additional discussion of the treatment of methodological variable levels in MRM predictions, see Johnston, Besedin, and Ranson (2006) and Moeltner, Boyle, and Paterson (2007).

[22] The entry in Table 8, 11.232, reflects the natural log of this value after adjustment from 2015 to 2007 USD.

Table 7. Geospatial and Socioeconomic Data for Benefit Transfer Scenarios

Variable	Units	Exeter River	Squamscott River	Great Bay Towns	Great Bay Counties	Great Bay Statewide
Market Area Towns/ Counties		Brentwood, Chester, Danville, Exeter, Fremont, Raymond, and Sandown	Exeter, Newfields, and Stratham	Dover, Durham, Greenland, Newfields, Newington, Newmarket, and Stratham	Rockingham and Strafford	All New Hampshire Towns
Number of Households	households	18,705	9,637	24,713	165,514	520,251
Household-Weighted Median Income	2015 USD	80,724	86,305	71,668	75,329	66,799
Adjusted Median Income	2007 USD	70,617	75,499	62,695	65,898	58,436
Focal River Length	km	65.3	10.1	N/A	N/A	N/A
Focal Shore Length	km	130.6	20.2	61.3	61.3	61.3
Other River Length	km	1,191	1,191	N/A	N/A	N/A
Other Shore Length	km	2,382	2,382	24.7	24.7	24.7
Market Area	km^2	353	111	302	2873	24,040
County Area	km^2	1,882	1,882	2,873	2,873	2,873
County Ag Land Area	km^2	115	115	181	181	181
Hydrologic Unit Code (HUC) 10 Area	km^2	331	331	1,172	1,172	1,172
HUC 10 Ag Land Area	km^2	38	38	95	95	95

Table 8. Illustrating the Benefit Transfer Process for a 9-Point Increase on the 100-Point Water Quality Index (WQI) in the Squamscott River (baseline WQI = 71)

Variable	(A) Model Coefficients	(B) Selected Values	Data Source	(C) Product (A) * (B)
Ln_BaseQuality	−0.064	4.260	NHDES	−0.273
Ln_QualityChg	0.281	2.197	Scenario	0.617
Ln_Income	0.628	11.232	U.S. Census	7.054
Non_Users	−0.455	0	Scenario	0.000
Swim_Use	−0.391	1	Scenario	−0.391
Boat_Use	−0.314	1	Scenario	−0.314
Game_Fish	0.303	1	Scenario	0.303
River	−0.226	1	Scenario	−0.226
Multi_Body	−0.525	0	Scenario	0.000
Ln_PropAgLand	−0.351	−2.795	GIS calculated	0.981
Ln_RelativeSize	0.052	−1.704	GIS calculated	−0.089
ProportionChg	0.525	0.008	GIS calculated	0.004
Northeast_US	0.549	1	Scenario	0.549
Central_US	0.601	0	Scenario	0.000
Southern_US	1.366	0	Scenario	0.000
MedianWTP	−0.264	0	Scenario	0.000
LumpSum	0.727	0	Scenario	0.000
Ln_StudyYear	−0.478	3.611	Scenario	−1.726
ChoiceExp	0.487	0.107	Metadata	0.052
Thesis	0.557	0.114	Metadata	0.063
Voluntary	−1.296	0	Scenario	0.000
OutlierBids	−0.429	1	Scenario	−0.429
NonParametric	−0.477	0.429	Metadata	−0.205
NonReviewed	−0.679	0	Scenario	0.000
Intercept	−2.281	1		−2.281

Calculation	Data	Result	Value
Sum of column (C)		lnWTP	3.691
$\exp(\text{lnWTP} + \sigma_\varepsilon^2/2)$	$\sigma_\varepsilon = 0.541$	Household WTP$_{2007}$	46.41
$(\text{CPI}_{2016}/\text{CPI}_{2007})*\text{WTP}_{07}$	$\text{CPI}_{2007} = 207.342$ $\text{CPI}_{2016} = 240.007$	**Household WTP$_{2016}$**	53.73
WTP * #Households	#Households = 9,637	**Region-wide WTP**	$517,754

Note: CPI, consumer price index; NHDES, New Hampshire Department of Environmental Services; WTP, willingness to pay.

transfer predicts annual per household WTP = $53.73 (2016 USD) for a WQI change from 71 to 80. When aggregated across all households in the three adjacent communities, the result is a total WTP of $517,754 per year.

Predicted Welfare Patterns and Benefit Transfer Validity

Initial insight into the validity of MRM benefit transfers can be gleaned through consideration of value surface patterns implied by the estimated meta-equation, including parameter estimates and functional form. For example, do parameter estimates indicate intuitive responsiveness of WTP to influences such as scope, scale, and extent of the market? However, other types of insight can only be provided through evaluation of the benefit predictions across different management scenarios. For example, are these predictions consistent with theory, intuition, and results of prior primary studies? Does the MRM generate credible predictions for the particular range of environmental changes under consideration?[23] Results such as these can help determine whether benefit transfer results are defensible for applied use.

Benefit transfers produce a range of WTP forecasts for water quality improvements in the Great Bay watershed, with results varying as expected over different scenarios. Table 9 and Figure 4 illustrate predicted WTP (per household per year) for 3-, 5-, 7-, and 9-point WQI increases for the three focal water bodies, assuming minimum baseline quality from Table 5. Transfers for Great Bay improvements also consider three different market extents when calculating per household WTP (adjacent towns, Rockingham and Strafford counties, and the entire state of New Hampshire). These are shown both numerically (Table 9) and graphically (Figure 4). Table 10 aggregates these per household estimates over all households within the relevant market areas to generate total, population-level predictions. Reported p-values (Tables 9 and 10) are calculated using Wald χ^2 tests (Greene 2012, p. 528), based on the underlying precision of MRM parameter estimates (Table 3).

The results show multiple patterns in WTP that are relevant to validity assessments and prospective policy applications. For example, annual per household WTP increases as the size of the water quality improvement increases (e.g., from a 3- to 9-point increase) for all focal water bodies (Figure 4), but at a decreasing rate. These are intuitive findings consistent

---

[23] For example, "most stated preference water quality valuation studies have examined improvements or decrements in the order of 10-20% of the full range of possible water quality levels represented on the [100-point water quality index] scale" (Newbold et al. (2018a, pp. 474–475). MRMs grounded in this literature typically have few or no data points (in the metadata) for very small water quality changes. WTP predictions for very small quality changes may therefore occur outside the range of available data. This can lead to benefit estimates that lack face validity (e.g., seem unreasonably high) or are sensitive to functional forms imposed by the MRM analyst (Johnston, Rolfe, and Zawojska 2018).

Table 9. Predicted Annual per Household Willingness to Pay to Improve Water Quality Index from Minimum Baselines

Region	3-Point	5-Point	7-Point	9-Point
Exeter River	39.95*	46.12**	50.69**	54.40**
Squamscott River	39.46*	45.55**	50.06**	53.73**
Great Bay Towns	62.69**	72.36**	79.54**	85.36**
Great Bay Counties	57.53**	66.41**	73.00***	78.34***
Great Bay Statewide	47.77**	55.14***	60.61***	65.05***

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$, with p-values calculated using Wald χ^2 tests (Greene 2012, p. 528). All values in 2016 USD.
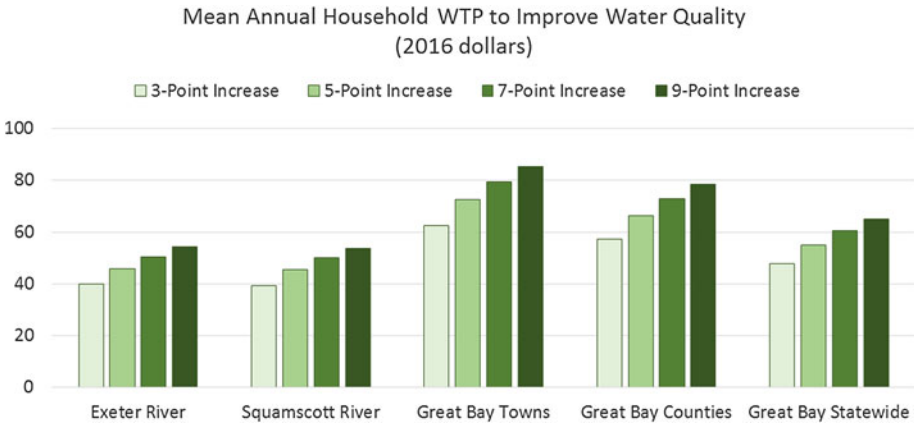
Figure 4. Willingness to Pay (per household per year) for 3-, 5-, 7-, and 9-Point Increases in Water Quality on the 100-Point Water Quality Index (WQI) for Three Water Bodies Using the Minimum Baseline WQI Value for Each Water Body from Table 6

Note: Three market regions (adjacent towns, two counties, and all of New Hampshire) were assessed for the Great Bay.

with positive scope sensitivity and diminishing marginal utility. For the Exeter and Squamscott Rivers, mean WTP transfers range from $39 (3-point) to $54 (9-point) per household per year, for households in adjacent communities. These reflect predictions that are consistent with the range of WTP values for similar improvements found in the prior literature (Table 1) and results of other water quality meta-analyses (e.g., Johnston et al. 2005; Van Houtven, Powers, and Pattanayak 2007).

Table 10. Annual Aggregated Willingness to Pay (WTP; millions of dollars) to Improve Water Quality Index from Minimum Baselines

Region	3-Point	5-Point	7-Point	9-Point
Exeter River Min	0.75*	0.86**	0.95**	1.02**
Squamscott River Min	0.38*	0.44**	0.48**	0.52**
Great Bay Towns Min	1.55**	1.79**	1.97**	2.11**
Great Bay Counties Min	9.52**	10.99**	12.08***	12.97***
Great Bay Statewide Min	24.85***	28.69***	31.53***	33.84***

Notes: *** p < 0.01, ** p < 0.05, * p < 0.10, with p-values calculated using Wald χ^2 tests (Greene 2012, p. 528). All values in 2016 USD. WTP values for Exeter and Squamscott Rivers are aggregated over adjacent towns, as described in the main text.

Per household WTP predictions are also similar across these two rivers, reflecting the offsetting effects of different parameters. Although the baseline water quality is better and the size of the improved water body (i.e., the length of the river) is larger in the Exeter River, median household income is higher in the Squamscott River (Tables 6 and 8). Annual per household WTP is greater ($62–$85) for improvements to Great Bay (despite baseline water quality being higher), primarily because of the size of the water body and the relative lack of substitutes within New Hampshire (Tables 6 and 8). Results such as these demonstrate the combined effects of scope, scale, substitutes, and income on WTP predictions.

Results also demonstrate the effects of market extent. As the market area for the Great Bay benefit transfer increases (from towns to counties to state), annual per household WTP decreases. This reflects intuitive effects of distance decay (Sutherland and Walsh 1985; Bateman et al. 2006, 2011a; Schaafsma 2015; Glenk et al. 2019). That is, larger market areas are associated with larger average distances between individuals and improved resources, ceteris paribus, leading to an expectation of lower mean per household WTP (Johnston, Besedin, and Stapler 2017).

Benefit transfers aggregated over an entire market area (or population) can vary because of differences in per household WTP and in the number of households in the market area. Despite comparable per household WTP measures, regional WTP values aggregated across all households in the adjacent communities for the three-town Squamscott River are lower than values for the larger seven-town Exeter River region, because of the larger number of households in the latter region (Figure 3, Table 10). For example, aggregated WTP for a 3-point WQI increase in the Squamscott River is approximately $0.38 million over the adjacent community population, compared with $0.75 million for the Exeter River (Table 10).

Aggregated values for the seven communities immediately adjacent to the Great Bay exceed those for the Exeter and Squamscott Rivers, because of

larger per household WTP values and the larger number of households in the region. Moreover, as one expands the assumed extent of the market for the Great Bay benefit transfer (from towns to counties to the state), aggregated population-level WTP increases by roughly a factor of 30. Even though mean per household WTP is lower for larger market areas (because of distance decay), the summation of values over (much) larger populations overwhelms this effect when estimating aggregate population-level WTP. Patterns such as these demonstrate the often-dominant effect of assumed market area (the area over which benefits are aggregated) when conducting benefit transfers (Bateman et al. 2006).

Predicted WTP estimates for all scenarios are statistically significant, with significance positively related to the scope of WQI change and the extent of the market (Tables 9 and 10). Hence, variations in the valuation scenario influence not only mean WTP predictions that emerge from an MRM, but also (unsurprisingly) the relative precision of those predictions. Given findings such as these, practitioners may wish to consider not only the predicted WTP point estimate for a transfer, but also whether it is possible to reject the null hypothesis that this estimate is different from zero. We reject this null hypothesis in all cases, in most cases at $p < 0.05$ or better.

Challenges for Large-Scale Ecosystem Service Valuation

Results such as these illustrate many of the patterns desired for large-scale benefit transfers, including responsiveness of welfare estimates to factors that should—according to theory—influence per household and aggregate WTP. Moreover, because the MRM includes an internal means to adjust WTP estimates for variations in scope and scale, there is no need to "scale up" estimates (Brander et al. 2012). However, meta-analytic methods are not a panacea. Although the valuation literature tends to emphasize the advantages of these methods, there are also limitations that should be considered. As discussed previously, the validity of any MRM for benefit transfer depends on multiple factors. However, even with an MRM judged to be sufficient across all of these dimensions, challenges for benefit transfer can still arise. This section discusses some of these challenges, using the previous MRM and case study benefit transfer as an illustration.

Among the primary limitations is that all benefit transfer methods—including MRMs—are constrained in terms of their capacity to account for unique attributes of particular policy sites (Bateman et al. 2011a). All valuation contexts are characterized by unique conditions, including ecosystem services, substitutes and complements, populations, geospatial dimensions, and other factors that influence economic values. Although MRM functions can provide a means to adjust WTP predictions for some of these factors, their ability to do so is limited by the set of variables in the model. This is not an inherent shortcoming in meta-analytic methods—the ability to adjust benefit estimates is usually greater within MRM benefit transfer than within

other benefit transfer approaches (such as unit value transfer, single-study benefit function transfer, and structural preference calibration[24]). However, given limits in the information available from primary valuation studies in the literature, these adjustments are still constrained. For example, the illustrated MRM includes only two variables characterizing affected populations (*Ln_Income* and *Non_Users*) and three variables characterizing affected uses (*Swim_Use*, *Boat_Use*, and *Game_Fish*). Context-specific effects beyond these variables—for example, possible effects on drinking water (which is not valued by this MRM) or on regionally important cultural resources—cannot be accommodated explicitly without adding corresponding control variables and obtaining metadata that provide sufficient variation in these variables.

Among other context-specific features that can only be accommodated in a limited manner is the effect of complements and substitutes. As discussed by Bateman et al. (2011a), Glenk et al. (2019), and Schaafsma (2015), substitutes and complements for specific ecosystem service improvements are often diverse and vary over space. This leads to two linked challenges for benefit transfers: (1) identifying measures of substitutes and complements that are similarly relevant across sites and (2) obtaining data on these measures from primary studies or supplementary sources. Given these challenges, most MRMs omit variables that capture the effect of substitutes and complements on WTP. The illustrated MRM is comparatively superior in this regard but still includes only one variable that captures spatially variable substitute effects (*ProportionChg*) and one that captures potential effects of complements (*Ln_PropAgLand*). As such, the capacity of this (or any) MRM to predict the spatially heterogeneous effects of multiple substitutes and complements on ESVs is limited. If more comprehensive and precise estimates of these effects are required, primary valuation studies should be conducted.[25]

There are also trade-offs implied by the functional forms for MRMs that influence suitability for particular types of benefit transfer. For example, WTP

---

[24] Johnston et al. (2015a) describe these approaches. Preference calibration requires the analyst to specify a structural utility function that describes individuals' choices over market and/or nonmarket goods. One then derives analytical expressions that determine a theoretically consistent relationship between benefit measures from existing primary studies. Given the mathematical complexity of these methods and the need to derive closed-form solutions, approaches of this type almost always incorporate a small number of variables (e.g., Smith, Van Houtven, and Pattanayak 2002; Smith, Pattanayak, and Van Houtven 2006).

[25] The need for high quality, well-documented primary valuation studies to support policy analysis and benefit transfer has been emphasized repeatedly in the literature (Loomis and Rosenberger 2006; McComb et al. 2006; Rosenberger and Johnston 2009; Johnston and Rosenberger 2010; Johnston et al. 2015b). Related issues are summarized by Johnston et al. (2015b, pp. 565–567), who argue that "[those] who support valuation research have thus far been unable to promote development of a large set of empirical valuation studies specifically designed and documented for benefit transfer applications."

patterns implied in Figure 4 may lead to questions regarding how the model can and should be used to estimate benefits for successive, small water quality improvements over time. Even cursory examination of this figure suggests that adding-up likely does not apply—predicted WTP for one 9-unit improvement is much less than WTP for a 3-unit improvement multiplied by 3.[26] Works such as Newbold et al. (2018b) and Moeltner (2019) have responded with alternative functional forms for this MRM that impose the adding-up property. These functions, however, reduce empirical performance —they have poorer fit to the metadata (Johnston, Rolfe, and Zawojska 2018).[27] As noted by Newbold et al. (2018b, p. 544), there is "a tradeoff between improved statistical fit that can be achieved by allowing additional model flexibility […], and consistency with theoretical restrictions that might require reduced flexibility on the other." Kling and Phaneuf (2018, p. 498) further argue that practical failures of adding-up tests in applied market and nonmarket situations suggest that these tests "are not likely to be fruitful in informing benefit transfer."[28]

Trade-offs such as these related to the choice of strong structural versus reduced-form specifications for MRMs are discussed by Bergstrom and Taylor (2006) and Johnston, Rolfe, and Zawojska (2018) and highlight the lack of consensus in this area. As noted by Johnston, Rolfe, and Zawojska (2018, p. 192), "All [MRM model specifications] require assumptions, and the capacity of any assumed specification to approximate all aspects of empirical reality cannot be assured." Disagreements over the appropriateness of strong structural specifications in MRMs are further related to a tension between two competing motivations for economic meta-analysis. Historically, one of the primary motivations for economic meta-analysis was to allow metadata (from many prior studies) to reveal patterns caused by "misspecification biases and specification searching in empirical economic research" (Stanley 2005, p. 205). Viewed from this perspective, the use of strong structural specifications to compel—ex ante—MRMs to produce certain types of empirical results is antithetical to the core purpose of meta-analysis. In contrast, others have argued that structural specifications are required to

---

[26] This is only a crude approximation of adding-up. A formal evaluation would adjust baseline WQI to successively higher levels when calculating WTP for each successive improvement. Adjustments would also be required for income effects (Kling and Phaneuf 2018). However, a more formal evaluation by Moeltner (2019) shows that adding-up is violated when using this MRM for benefit transfer.

[27] This poor fit is not surprising. As noted by Kling and Phaneuf (2018), past evaluations have shown that the adding-up property is routinely violated in both market and nonmarket situations. Hence, these specifications are imposing a property on the MRM that is likely violated by many of the primary studies in the metadata.

[28] Specifications such as that of Johnston, Besedin, and Holland (2018) that define the dependent variable in terms of WTP per unit of change can approximate the adding-up property without sacrificing empirical fit, but this approach does not guarantee that adding-up is satisfied perfectly.

ensure welfare-theoretic consistency within meta-analytic transfers (Smith and Pattanayak 2002; Newbold et al. 2018b).

Potential ramifications of MRM specification are further illustrated using Figure 5, which shows the predicted marginal WTP (or demand) for WQI improvements in the Great Bay, by residents of surrounding towns.[29] By illustrating marginal rather than total WTP, Figure 5 clarifies welfare patterns implied by Figure 4, and particularly WTP patterns for small changes in water quality. As shown by Figure 5, marginal WTP increases sharply for ΔWQI < 3, suggesting high marginal WTP for the first few units of quality change. Because of this pattern, repeated use of the MRM to predict values for successive, small water quality changes will lead to large aggregate WTP estimates for the combined change, ceteris paribus.[30] This is among the motivations used by Newbold et al. (2018b) and Moeltner (2019) to suggest structural adding-up MRM specifications that, among other features, attenuate WTP predictions for small changes.

However, it is important to notice that predictions such as these are outside of the range of data support for the MRM—there are no observations of ΔWQI < 2.5 in the metadata. Moreover, none of the metadata observations report WTP for successive water quality changes—all report WTP for one-time improvements. Structural and nonstructural MRMs estimated from these metadata tend to predict similar WTP estimates for one-time changes in water quality within the range of the data. In contrast, debates over structural MRM functional forms often center on WTP predictions for successive, small water quality changes. Because there are no supporting metadata for predictions of this type, there is little capacity to validate these predictions. More informed debates over WTP predictions for small, successive water quality changes will require primary studies that provide credible estimates of these values.

A related issue concerns the capacity of MRMs to predict ESVs over large scales, such as multistate regions or nationwide. As shown by Table 3, the partial elasticity associated with *Ln_RelativeSize* is small. Recall, this variable is defined as the natural log of the ratio of total shoreline length (kilometers) of the improved resource and the size of the benefit aggregation market area (square kilometers). The inelastic magnitude of this effect implies that per household WTP declines gradually with increases in market area size.[31] Given that only one primary study in the metadata considers more than a two-state market area,[32] it is unclear to what extent this small effect size applies to larger market areas—as these are also (largely) outside the range of the metadata. The implications for large-scale ESV transfers are implied by

---

[29] All other aspects of the valuation scenario mirror those described previously.
[30] Small (negative) WTP adjustments associated with the resulting increases in baseline water quality (*Ln_BaseQuality*) are insufficient to offset this effect.
[31] This pattern implies that distance decay is significant but small in magnitude.
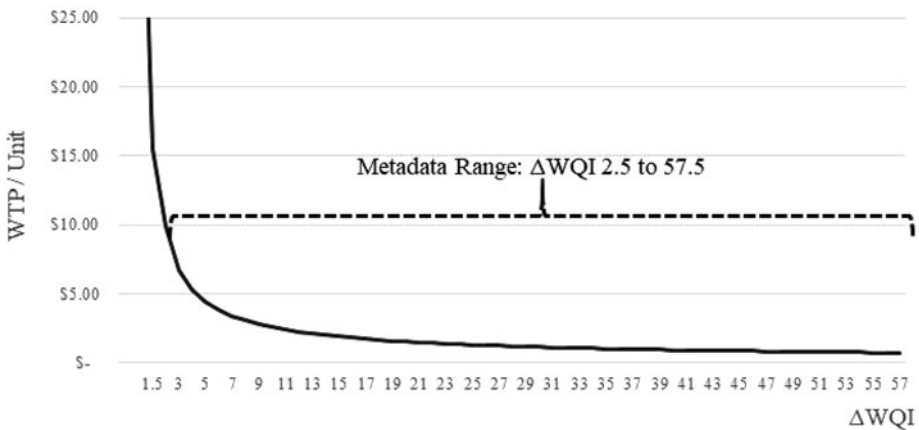[32] Not including the District of Columbia.

Figure 5. Predicted Demand for Water Quality Improvements (ΔWQI), Great Bay by Residents of Adjacent Towns

Notes: Valuation scenario is identical to that in Figure 4 and described in the main text. WQI, water quality index; WTP, willingness to pay.

Table 10—aggregate values increase rapidly with the size of the assumed market area. Hence, the model will potentially predict large aggregate WTP estimates when applied to regions larger than a single state. Given the lack of support in the metadata for such applications, practitioners should proceed with care.

A final note concerns the accuracy of MRM benefit transfer compared with simpler forms of unit value or benefit function transfer, or naïve MRM benefit transfers that overlook the validity concerns discussed previously. Direct comparisons of this type are beyond the scope of this article but have been conducted elsewhere. For example, Johnston, Besedin, and Stapler (2017) and Johnston, Besedin, and Holland (2018) demonstrate the benefit transfer accuracy gains made possible via the inclusion of spatial variables in MRM benefit transfer for different variants of the same water quality metadata used here. Johnston and Thomassin (2010) use an earlier, U.S.-Canada version of the same metadata to demonstrate that MRM benefit transfer reduces transfer errors compared with unit value benefit transfer. Using other sets of metadata, Moelter and Rosenberger (2014) and Johnston and Moeltner (2014) illustrate Bayesian approaches that can be used to evaluate the extent to which commodity and welfare consistency increase benefit transfer efficiency. This and other recent work is summarized by Johnston, Rolfe, and Zawojska's (2018) review of the recent benefit transfer literature.[33]

---

[33] Prior work does not find that more complex benefit function transfers, such as MRMs, always outperform simpler methods (Johnston, Rolfe, and Zawojska 2018). Although benefit function transfers and data synthesis methods have been shown to outperform unit value transfers, on

Conclusion

Methodological advances in meta-regression modeling have improved the capacity of benefit transfers to provide valid and reliable welfare predictions. Variable and functional form specifications for these models can be adapted to meet particular benefit transfer needs and theoretical expectations, and there is increasing recognition of the trade-offs implied by different types of MRM specifications. The future seems bright in terms of continued development of meta-analytic techniques to support ecosystem service valuation.

At the same time, many meta-analyses in the ecosystem services literature omit fundamental elements required to ensure valid value predictions, and the steps required to generate benefit predictions from underlying meta-analytic benefit functions are often obscured. Clarity and transparency on issues such as these are required to ensure credible uses of meta-analytic tools. Few meta-analyses in the literature discuss the extent to which the resulting benefit predictions are credible for applied use or evaluate the impact of post-estimation procedures on benefit transfer results. In some cases, simple unit value transfers may be more accurate than those implemented using meta-analysis, particularly if the latter does not follow best-practice standards. The illustrated case study is designed to demonstrate how evaluations of issues such as these can help clarify the suitability of MRM predictions for benefit transfer.

Recognition of these issues is needed. As noted by Johnston, Rolfe, and Zawojska (2018), there has been a proliferation of benefit transfer tools developed outside of the environmental valuation and benefit transfer literature (Bagstad et al. 2013), and government agencies are applying benefit transfer practices that are considered arbitrary and unjustified by valuation researchers (Boyle, Kotchen, and Smith 2017; Smith 2018). Without "cost-effective, straightforward, transferable, scalable, meaningful, and defensible methods" (Olander et al. 2017, p. 170) for benefit transfer, analysts may apply approaches that are unlikely to provide valid and defensible estimates.

Financial support

---

average (Kaul et al. 2013; Rosenberger 2015), unit value and simpler function transfers have been shown to be more accurate in some cases (e.g., Barton 2002; Lindhjem and Navrud 2008; Johnston and Duke 2010; Bateman et al. 2011a; Klemick et al. 2018; Guignet et al. 2019). The literature has yet to reach consensus on conditions under which this is likely to occur (Johnston, Rolfe, and Zawojska 2018).

Acknowledgments

Conflict of interest

None.

Transparency and openness promotion statement

Data necessary to replicate results are available in this article, through existing data repositories, or by request. Results in Tables 1–3 are available from Johnston, Besedin, and Stapler (2017); data to replicate these results are available by request from the lead author. Geospatial and population data summarized in Tables 6 and 7 are available from the National Hydrography Dataset (http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php); the Hydrologic Unit Code Watershed Boundary Dataset (http://water.usgs.gov/GIS/huc.html); the National Land Cover Database (http://www.mrlc.gov); the NOAA Global Self-Consistent, Hierarchical, High-Resolution Geography Database (http://www.ngdc.noaa.gov/mgg/shorelines/shorelines.html); New Hampshire's Statewide GIS Clearinghouse (http://www.granit.unh.edu/); and the U.S. Census (http://www.census.gov/geo/maps-data/data/tiger.html). Water quality monitoring data are available from the New Hampshire Department of Environmental Services (https://www.des.nh.gov/organization/divisions/water/index.htm).

References

Abbasi, T. 2012. *Water Quality Indices*. Amsterdam: Elsevier.

Bagstad, K.J., D.J. Semmens, S. Waage, and R. Winthrop. 2013. "A Comparative Assessment of Decision-Support Tools for Ecosystem Services Quantification and Valuation." *Ecosystem Services* 5: 27–39.

Barton, D.N. 2002. "The Transferability of Benefit Transfer: Contingent Valuation of Water Quality Improvements in Costa Rica." *Ecological Economics* 42(1–2): 147–164.

Bateman, I.J., R. Brouwer, S. Ferrini, M. Schaafsma, D.N. Barton, A. Dubgaard, B. Hasler, *et al*. 2011a. "Making Benefit Transfers Work: Deriving and Testing Principles for Value Transfers for Similar and Dissimilar Sites Using a Case Study of the Non-market Benefits of Water Quality Improvements across Europe." *Environmental and Resource Economics* 50(3): 365–387.

Bateman, I.J., B.H. Day, S. Georgiou, and I. Lake. 2006. "The Aggregation of Environmental Benefit Values: Welfare Measures, Distance Decay and Total WTP." *Ecological Economics* 60(2): 450–460.

Bateman, I.J., and A.P. Jones. 2003. "Contrasting Conventional with Multi-level Modeling Approaches to Meta-analysis: Expectation Consistency in UK Woodland Recreation Values." *Land Economics* 79(2): 235–258.

Bateman, I.J., G.M. Mace, C. Fezzi, G. Atkinson, and K. Turner. 2011b. "Economic Analysis for Ecosystem Service Assessments." *Environmental and Resource Economics* 48(2): 177–218.

Bauer, D.M., and R.J. Johnston. 2017. *Buffer Options for the Bay: Economic Valuation of Water Quality Ecosystem Services in New Hampshire's Great Bay Watershed*. Great Bay National Estuarine Research Reserve, Greenland, NH.

Bergstrom, J.C., and L.O. Taylor. 2006. "Using Meta-analysis for Benefits Transfer: Theory and Practice." *Ecological Economics* 60(2): 351–360.

Bishop, R.C., and K.J. Boyle. 2019. "Reliability and Validity in Nonmarket Valuation." *Environmental and Resource Economics* 72(2): 559–582.

Bockstael, N.E., A.M. Freeman, R.J. Kopp, P.R. Portney, and V.K. Smith. 2000. "On Measuring Economic Values for Nature." *Environmental Science and Technology* 34(8): 1384–1389.

Boyd, J., P. Ringold, A. Krupnick, R.J. Johnston, M.A. Weber, and K. Hall. 2016. "Ecosystem Services Indicators: Improving the Linkage between Biophysical and Economic Analyses." *International Review of Environmental and Resource Economics* 8(3–4): 359–443.

Boyle, K.J., S. Kaul, and C.F. Parmeter. 2015. "Meta-analysis: Econometric Advances and New Perspectives towards Data Synthesis and Robustness." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 383–418.

Boyle, K.J., M.J. Kotchen, and V.K. Smith. 2017. "Deciphering Dueling Analyses of Clean Water Regulations." *Science* 358(6359): 49–50.

Boyle, K.J., N.V. Kuminoff, C.F. Parmeter, and J.C. Pope. 2009. "Necessary Conditions for Valid Benefit Transfers." *American Journal of Agricultural Economics* 91(5): 1328–1334.

Boyle, K.J., C.F. Parmeter, B.B. Boehlert, and R.W. Paterson. 2013. "Due Diligence in Meta-analyses to Support Benefit Transfers." *Environmental and Resource Economics* 55(3): 357–386.

Boyle, K.J., and J.M. Wooldridge. 2018. "Understanding Error Structures and Exploiting Panel Data in Meta-analytic Benefit Transfers." *Environmental and Resource Economics* 69(3): 609–635.

Brander, L.M., I. Bräuer, H. Gerdes, A. Ghermandi, O. Kuik, A. Markandya, S. Navrud, *et al.* 2012. "Using Meta-analysis and GIS for Value Transfer and Scaling Up: Valuing Climate Change Induced Losses of European Wetlands." *Environmental and Resource Economics* 52(3): 395–413.

Cude, C.G. 2001. "Oregon Water Quality Index: A Tool for Evaluating Water Quality Management Effectiveness." *Journal of the American Water Resources Association* 37(1):125–137.

Ferrini, S., M. Schaafsma, and I.J. Bateman. 2015. "Ecosystem Services Assessment and Benefit Transfer." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 275–305.

Glenk, K., R.J. Johnston, J. Meyerhoff, and J. Sagebiel. 2019. "Spatial Dimensions of Stated Preference Valuation in Environmental and Resource Economics: Methods, Trends and Challenges." *Environmental and Resource Economics*. https://doi.org/10.1007/s10640-018-00311-w

Greene, W.H. 2012. *Econometric Analysis* (7th edition). Upper Saddle River, NJ: Prentice Hall.

Griffiths, C., H. Klemick, M. Massey, C. Moore, S. Newbold, D. Simpson, P. Walsh, and W. Wheeler. 2012. "US Environmental Protection Agency Valuation of Surface Water Quality Improvements." *Review of Environmental Economics and Policy* 6(1): 130–146.

Guignet, D., M.T. Heberling, M. Papenfus, O. Griot, and B. Holland. 2019. "Property Values and Water Quality: A Nationwide Meta-analysis and the Implications for Benefit Transfer." National Center for Environmental Economics Working Paper, 19-05, U.S. Environmental Protection Agency, Washington, DC.

Hoehn, J.P. 2006. "Methods to Address Selection Effects in the Meta Regression and Transfer of Ecosystem Values." *Ecological Economics* 60(2): 389–398.

Johnston, R.J., and E.Y. Besedin. 2009. "Estimating Willingness to Pay for Aquatic Resource Improvements Using Benefits Transfer." In H.W. Thurston, M.T. Heberling, and A. Schrecongost, eds. *Environmental Economics for Watershed Restoration*. Boca Raton, FL: CRC Press, pp. 95–122.

Johnston, R.J., E.Y. Besedin, and B.M. Holland. 2018. "Modeling Distance Decay within Valuation Meta-analysis." *Environmental and Resource Economics* 72(3): 657–690.

Johnston, R.J., E.Y. Besedin, R. Iovanna, C.J. Miller, R.F. Wardwell, and M.H. Ranson. 2005. "Systematic Variation in Willingness to Pay for Aquatic Resource Improvements and Implications for Benefit Transfer: A Meta-analysis." *Canadian Journal of Agricultural Economics* 53(2–3): 221–248.

Johnston, R.J., E.Y. Besedin, and M.H. Ranson. 2006. "Characterizing the Effects of Valuation Methodology in Function-Based Benefits Transfer." *Ecological Economics* 60 (2): 407–419.

Johnston, R.J., E.Y. Besedin, and R. Stapler. 2017. "Enhanced Geospatial Validity for Meta-analysis and Environmental Benefit Transfer: An Application to Water Quality Improvements." *Environmental and Resource Economics* 68(2): 343–375.

Johnston, R.J., E.Y. Besedin, and R.F. Wardwell. 2003. "Modeling Relationships between Use and Nonuse Values for Surface Water Quality: A Meta-analysis." *Water Resources Research* 39(12): 1363–1371.

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

Johnston, R.J., B.M. Holland, and L. Yao. 2016. "Individualized Geocoding in Stated Preference Questionnaires: Implications for Survey Design and Welfare Estimation." *Land Economics* 92(4): 737–759.

Johnston, R.J., and K. Moeltner. 2014. "Meta-modeling and Benefit Transfer: The Empirical Relevance of Source-Consistency in Welfare Measures." *Environmental and Resource Economics* 59(3): 337–361.

Johnston, R.J., M.H. Ranson, E.Y. Besedin, and E.C. Helm. 2006. "What Determines Willingness to Pay per Fish? A Meta-analysis of Recreational Fishing Values." *Marine Resource Economics* 21(1): 1–32.

Johnston, R.J., J. Rolfe, R.S. Rosenberger, and R. Brouwer. 2015a. "Introduction to Benefit Transfer Methods." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 19–59.

Johnston, R.J., J. Rolfe, and E. Zawojska. 2018. "Benefit Transfer of Environmental and Resource Values: Progress, Prospects and Challenges." *International Review of Environmental and Resource Economics* 12(2–3): 177–266.

Johnston, R.J., and R.S. Rosenberger. 2010. "Methods, Trends and Controversies in Contemporary Benefit Transfer." *Journal of Economic Surveys* 24(3): 479–510.

Johnston, R.J., R.S. Rosenberger, J. Rolfe, and R. Brouwer. 2015b. "Benefit Transfer: The Present State and Future Prospects." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 553–574.

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

Johnston, R.J., and P.J. Thomassin. 2010. "Willingness to Pay for Water Quality Improvements in the United States and Canada: Considering Possibilities for International Meta-analysis and Benefit Transfer." *Agricultural and Resource Economics Review* 39(1): 114–131.

Johnston, R.J., and L.A. Wainger. 2015. "Benefit Transfer for Ecosystem Service Valuation: An Introduction to Theory and Methods." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 237–273.

Kaul, S., K.J. Boyle, N.V. Kuminoff, C.F. Parmeter, and J.C. Pope. 2013. "What Can We Learn from Benefit Transfer Errors? Evidence from 20 Years of Research on Convergent Validity." *Journal of Environmental Economics and Management* 66(1): 90–104.

Klemick, H., C. Griffiths, D. Guignet, and P. Walsh. 2018. "Improving Water Quality in an Iconic Estuary: An Internal Meta-analysis of Property Value Impacts around the Chesapeake Bay." *Environmental and Resource Economics* 69(2): 265–292.

Kling, C.L., and D.J. Phaneuf. 2018. "How Are Scope and Adding Up Relevant for Benefits Transfer?" *Environmental and Resource Economics* 69(3): 483–502.

Lindhjem, H., and S. Navrud. 2008. "How Reliable Are Meta-analyses for International Benefit Transfers?" *Ecological Economics* 66(2–3): 425–435.

Londoño, L.M., and R.J. Johnston. 2012. "Enhancing the Reliability of Benefit Transfer over Heterogeneous Sites: A Meta-analysis of International Coral Reef Values." *Ecological Economics* 78: 80–89.

Loomis, J.B., and R.S. Rosenberger. 2006. "Reducing Barriers in Future Benefit Transfers: Needed Improvements in Primary Study Design and Reporting." *Ecological Economics* 60(2): 343–350.

McComb, G., V. Lantz, K. Nash, and R. Rittmaster. 2006. "International Valuation Databases: Overview, Methods and Operational Issues." *Ecological Economics* 60(2): 461–472.

Moeltner, K. 2015. "Optimal Scope and Bayesian Model Search in Benefit Transfer." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 507–523.

Moeltner, K. 2019. "Bayesian Nonlinear Meta Regression for Benefit Transfer." *Journal of Environmental Economics and Management* 93: 44–62.

Moeltner, K., K.J. Boyle, and R.W. Paterson. 2007. "Meta-analysis and Benefit Transfer for Resource Valuation—Addressing Classical Challenges with Bayesian Modeling." *Journal of Environmental Economics and Management* 53(2): 250–269.

Moeltner, K., and R.S. Rosenberger. 2014. "Cross-Context Benefit Transfer: A Bayesian Search for Information Pools." *American Journal of Agricultural Economics* 96(2): 469–488.

Nelson, J.P. 2015. "Meta-analysis: Statistical Methods." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 329–356.

Nelson, J.P., and P.E. Kennedy. 2009. "The Use (and Abuse) of Meta-analysis in Environmental and Resource Economics: An Assessment." *Environmental and Resource Economics* 42(3): 345–377.

Newbold, S., R.D. Simpson, D.M. Massey, M.T. Heberling, W. Wheeler, J. Corona, and J. Hewitt. 2018a. "Benefit Transfer Challenges: Perspectives from U.S. Practitioners." *Environmental and Resource Economics* 69(3): 467–482.

Newbold, S., P.J. Walsh, D.M. Massey, and J. Hewitt. 2018b. "Using Structural Restrictions to Achieve Theoretical Consistency in Benefit Transfers." *Environmental and Resource Economics* 69(3): 529–554.

Olander, L.P., R.J. Johnston, H. Tallis, J. Kagan, L.A. Maguire, S. Polasky, D. Urban, J. Boyd, L. Wainger, and M. Palmer. 2018. "Benefit Relevant Indicators: Ecosystem Services Measures That Link Ecological and Social Outcomes." *Ecological Indicators* 85: 1262–1272.

Olander, L., S. Polasky, J.S. Kagan, R.J. Johnston, L. Wainger, D. Saah, L. Maguire, J. Boyd, and D. Yoskowitz. 2017. "So You Want Your Research to Be Relevant? Building the

Bridge between Ecosystem Services Research and Practice." *Ecosystem Services* 26A: 170–182.

Plummer, M.L. 2009. "Assessing Benefit Transfer for the Valuation of Ecosystem Services." *Frontiers in Ecology and the Environment* 7(1): 38–45.

Richardson, L., J. Loomis, T. Kroeger, and F. Casey. 2015. "The Role of Benefit Transfer in Ecosystem Service Valuation." *Ecological Economics* 115: 51–58.

Rolfe, J., R. Brouwer, and R.J. Johnston. 2015. "Meta-analysis: Rationale, Issues and Applications." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 357–381.

Rosenberger, R.S. 2015. "Benefit Transfer Validity and Reliability." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 307–326.

Rosenberger, R.S., and R.J. Johnston. 2009. "Selection Effects in Meta-analysis and Benefit Transfer: Avoiding Unintended Consequences." *Land Economics* 85(3): 410–428.

Rosenberger, R.S., and J.B. Loomis. 2017. "Benefit Transfer." In P.A. Champ, K.J. Boyle, and T.C. Brown, eds. *A Primer on Nonmarket Valuation* (2nd edition). Dordrecht, the Netherlands: Springer, pp. 431–462.

Rosenberger, R.S., and T.T. Phipps. 2007. "Correspondence and Convergence in Benefit Transfer Accuracy: Meta-analytic Review of the Literature." In S. Navrud and R. Ready, eds. *Environmental Value Transfer: Issues and Methods*. Dordrecht, the Netherlands: Springer, pp. 23–43.

Schaafsma, M. 2015. "Spatial and Geographical Aspects of Benefit Transfer." In R.J. Johnston, J. Rolfe, R.S. Rosenberger, and R. Brouwer, eds. *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Dordrecht, the Netherlands: Springer, pp. 421–439.

Schultz, E.T., R.J. Johnston, K. Segerson, and E.Y. Besedin. 2012. "Integrating Ecology and Economics for Restoration: Using Ecological Indicators in Valuation of Ecosystem Services." *Restoration Ecology* 20(3): 304–310.

Smith, V.K. 2018. "Benefits Transfer: Current Practice and Prospects." *Environmental and Resource Economics* 69(3): 449–466.

Smith, V.K., G. Van Houtven, and S.K. Pattanayak. 2002. "Benefit Transfer via Preference Calibration: 'Prudential Algebra' for Policy." *Land Economics* 78(1): 132–152.

Smith, V.K., and S.K. Pattanayak. 2002. "Is Meta-analysis a Noah's Ark for Non-market Valuation?" *Environmental and Resource Economics* 22(1–2): 271–296.

Smith, V.K., S.K. Pattanayak, and G. Van Houtven. 2006. "Structural Benefit Transfer: An Example Using VSL Estimates." *Ecological Economics* 60(2): 361–371.

Stanley, T.D. 2005. "Beyond Publication Bias." *Journal of Economic Surveys* 19(3): 309–345.

Stanley, T.D., and H. Doucouliagos. 2012. *Meta-regression Analysis in Economics and Business*. New York: Routledge.

Stanley, T.D., H. Doucouliagos, M. Giles, J.H. Heckemeyer, R.J. Johnston, P. Laroche, J.P. Nelson, *et al.* 2013. "Meta-analysis of Economics Research Reporting Guidelines." *Journal of Economic Surveys* 27(2): 390–394.

Stapler, R.W., and R.J. Johnston. 2009. "Meta-analysis, Benefit Transfer, and Methodological Covariates: Implications for Transfer Error." *Environmental and Resource Economics* 42(2): 227–246.

Sutherland, R.J., and R.G. Walsh. 1985. "Effect of Distance on the Preservation Value of Water Quality." *Land Economics* 61(3): 282–290.

U.S. Environmental Protection Agency (U.S. EPA). 2009a. *Environmental Impact and Benefits Assessment for Final Effluent Guidelines and Standards for the Construction and Development Category*. EPA-821-R-09-012, Office of Science and Technology, Office of Water, U.S. EPA, Washington, DC.

U.S. Environmental Protection Agency (U.S. EPA). 2009b. "Nonmarket Benefits from Water Quality Improvements." In *Environmental Impact and Benefits Assessment for Final Effluent Guidelines and Standards for the Construction and Development Category.* EPA-821-R-09-012, U.S. EPA, Washington, DC, pp. 10-1–10-27.

U.S. Environmental Protection Agency (U.S. EPA). 2010. *Economic Analysis of Final Water Quality Standards for Nutrients for Lakes and Flowing Waters in Florida.* Office of Science and Technology, Office of Water, U.S. EPA, Washington, DC.

———. 2012. *Economic Analysis of Proposed Water Quality Standards for the State of Florida's Estuaries, Coastal Waters, and South Florida Inland Flowing Waters.* Office of Science and Technology, Office of Water, U.S. EPA, Washington, DC.

———. 2015. *Benefit and Cost Analysis for the Proposed Effluent Limitations Guidelines and Standards for the Steam Electric Power Generating Point Source Category.* EPA-821-R-15-005, Office of Science and Technology, Office of Water, U.S. EPA, Washington, DC.

Van Houtven, G., C. Mansfield, D.J. Phaneuf, R. von Haefen, B. Milstead, M.A. Kenney, and K.H. Reckhow. 2014. "Combining Expert Elicitation and Stated Preference Methods to Value Ecosystem Services from Improved Lake Water Quality." *Ecological Economics* 99: 40–52

Van Houtven, G., J. Powers, and S.K. Pattanayak. 2007. "Valuing Water Quality Improvements in the United States Using Meta-analysis: Is the Glass Half-Full or Half-Empty for National Policy Analysis?" *Resource and Energy Economics* 29(3): 206–228.

Walsh, J.P., and W.J. Wheeler. 2013. "Water Quality Indices and Benefit-Cost Analysis." *Journal of Benefit-Cost Analysis* 4(1): 81–105.

Zandersen, M., and R.S.J. Tol. 2009. "A Meta-analysis of Forest Recreation Values in Europe." *Journal of Forest Economics* 15(1–2): 109–130.