

Benefit Transfer Challenges: Perspectives from U.S. Practitioners

Accepted: 1 December 2017 / Published online: 28 February 2018 © This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2018

Keywords Benefit transfer · Water quality · Meta-analysis

1 Background

The earliest use of benefit–cost analysis (BCA) within the U.S. Federal government began in the 1930s when large public works projects, such as hydro-electric dams, highways, and harbors, were initiated as part of President Roosevelt's New Deal programs to lift the U.S. economy out of the Great Depression (Hanley and Spash 1993; Hufschmidt 2000; Pearce 2002). While elementary principles of public works valuation had been laid down by Dupuit a century before, new theoretical underpinnings and empirical methods for measuring economic values of goods that were publicly provided were developed in this period (Banzhaf 2010). Later, in the 1960s and 70s, there was an increased emphasis on government efforts

Guest Editor: V. Kerry Smith.

☑ William Wheeler wheeler.william@epa.gov

> Stephen Newbold newbold.steve@epa.gov

The findings, conclusions, and views expressed in this paper are those of the authors and do not necessarily represent those of the U.S. EPA. No Agency endorsement should be inferred. The authors appreciate helpful comments by Kerry Smith. This paper arose out of a workshop co-sponsored by the EPA's Office of Water, Office of Research and Development, and National Center for Environmental Economics. The planning committee for the workshop included Kerry Smith (Arizona State University), Elena Besedin (Abt Associates), and the following EPA personnel: Ashley Allen, Joel Corona, Matt Heberling, Julie Hewitt, Al McGartland, Stephen Newbold, David Simpson, Hale Thurston, Michael Trombley, Patrick Walsh, and William Wheeler.

¹ U.S. EPA, National Center for Environmental Economics, Washington, DC, USA

² U.S. EPA, Office of Research and Development, Cincinnati, OH, USA

³ U.S. EPA, Office of Water, Washington, DC, USA

to improve the environment. The Clean Air Act, Clean Water Act, and other important environmental measures date from this era (Easter and Archibald 1998; Banzhaf 2010; Banzhaf 2016). The year 1970 also marked the consolidation of previously separate Executive Branch regulatory oversight within the new Environmental Protection Agency (EPA). The EPA's regulatory activities and associated analyses since that time have been an important impetus for the development of more sophisticated methods of nonmarket valuation and modern protocols for BCA. These protocols are outlined in various federal agency guidelines, including the EPA's own *Guidelines for Conducting Economic Analyses* (USEPA 2010) and the Office of Management and Budget's (OMB's) Circular A-4 (OMB 2003).

All federal agencies, including the EPA, must follow a formal process when proposing a new rule (Copeland 2011). In order, these steps include: the development, internal review, and approval of the proposed rule by the EPA; external review of the proposed rule by OMB; publication in the U.S. Federal Register of a formal Notice of Proposed Rulemaking with a solicitation for public comments; a response to public comments by the EPA; development and internal review of a final rule by the EPA; OMB review of the final rule; publication of the final rule in the Federal Register; and, finally, implementation of the rule. If the rule is expected to have an impact on the U.S. economy of \$100 million per year or more, then it is deemed "economically significant"¹ and must be accompanied by a formal BCA at both the proposal and final stages (Fraas 1991).

This process can be time-consuming. To give an example, the EPA's regulations concerning discharges from Concentrated Animal Feeding Operations (CAFOs) were formally proposed two years after they were initially conceived and the rule was finalized three years after proposal (USEPA 2009). In another illustrative case, the EPA's Steam Electric effluent guidelines were proposed four years after their conception, and finalized two years after the proposal (USEPA 2015a). Within such timelines an iterative sequence of data collection, analysis, review, and revision must be conducted in compliance with a series of internally and externally imposed intermediate deadlines. The process begins with collecting large amounts of data. (In the case of the Steam Electric rule, for example, a nearly-400-page questionnaire was distributed to each manufacturing facility that might be subject to the new regulation.) The collected data are then used to develop policy options. Environmental engineers then estimate changes in pollution emissions, and water quality scientists produce estimates of changes in ambient water quality levels associated with each option. Economists use these predictions, as well as other information, to estimate the benefits and costs of each option considered for the proposed rule. After the rule is formally proposed, the process pauses for a public comment period—usually lasting between 60 and 120 days—during which interested parties submit comments on the proposal to the EPA. Often a large portion of the public comments are submitted by the regulated industry, and these may include new data and analyses. The EPA then must respond to all submitted public comments and modify the rule options and analyses accordingly. Before a rule can be proposed or finalized, it also must pass through several rounds of internal review, plus external review by other federal agencies and OMB. While the overall time from conception to proposal of a rule, and then from proposal to finalization may stretch into years, the time to conduct a BCA may be more constrained. At each stage of review, EPA staff may be required to analyze new options for the rule on relatively short turn-around times. Furthermore, the policy options as originally configured might be partially or wholly obsolete before a rule-making is completed, and EPA analysts must be prepared to make rapid adjustments to the analysis in response to evolving requests

¹ While the \$100 million threshold is often cited as the distinguishing feature of "significant" rules, Executive Order 12866 also called for analyses of prospective rules that could generate conflicts between agencies, affect entitlements, grants, user fees, or loan programs, or raise novel legal or policy issues.

from managers as the rule-making proceeds.² These factors create a demand for flexible and timely benefit analysis approaches.

In addition to the time pressure benefit–cost analysts may find themselves under, they often also face daunting challenges of scale. EPA often promulgates national regulations, and so must estimate the willingness-to-pay (WTP) for environmental improvements of all households in the U.S. Both the temporal and geographical issues make it unlikely that the Agency will find it possible to design a new nonmarket valuation study tailored to each new proposed regulation, collect and analyze the data, summarize the findings, subject the report to a formal peer review, revise as necessary, and publish the final report before using the results in the BCA for the proposed or even the final regulations.³

Given these facts, it would be useful to have a general purpose integrated framework that combines a comprehensive set of bio-physical models and observations of ambient environmental quality with data on consumer expenditures and preferences that could produce estimates of benefits on a timely basis for new regulations as they are taking shape. Phaneuf et al. (2008) provided a promising proof of concept for such a platform using data on outdoor recreation activities and residential property values in a county in North Carolina. However, sufficient data of comparable quality across the U.S. have not yet been collected or assembled to allow generalizing their approach to the national scale. For the foreseeable future, we expect that the EPA will have to continue to rely heavily on the extrapolation of nonmarket valuation estimates from previous studies to estimate benefits for new policy scenarios in its BCAs.

In the field of environmental economics, the use of benefit estimates reported in existing nonmarket valuation studies to calculate WTP for new policy cases has come to be known as "benefit transfer." To make quantitative statements about the likely effects of public policies, economists must extrapolate findings from previous empirical studies to new policy scenarios. Thus, benefit transfer is just a special case of the general practice of applied policy evaluation using empirical microeconomics. While benefit transfer is sometimes characterized as a method of last resort (OMB 2003), it is impossible to conduct a prospective BCA without the use of at least some form of benefit (and cost) transfers. Even if a new economic study were designed and executed to examine the exact population of households and outcomes that are the target of a proposed regulation shortly before its implementation, the analyst still must assume that the benefits estimated in a study conducted in one year will still be valid in the next year. In most real-world applications, there will be much greater differences between study cases and policy cases than merely a short period of time. Samples are never perfectly representative of their target populations, experimental treatments are never exactly like the policy changes that will be implemented, and control variables are never comprehensively and perfectly measured. For these reasons, there is *always* an element of extrapolation required for prospective policy analyses (Bardach 2004; Steel 2010; Cartwright and Hardie 2012; Howick et al. 2013). Nevertheless, the mere fact that benefit transfer is unavoidable does

 $^{^2}$ There also can be a "hurry up and wait" aspect to regulatory analysis: optimistic projections of the rulemaking timetable lead to requests for expedited analyses, but extraneous factors then intervene and slow down the process. The time that was *actually* available to do an analysis is sometimes longer than what analysts *expected* to be available when they initially developed the plans for their work.

³ The EPA generally requires peer review for "[s]tated preference (e.g., contingent valuation) and revealed preference surveys (e.g., recreational travel cost surveys) developed to assist in the economic analysis of a regulation or program" (USEPA 2015b, p. 50). This requirement could be met by publishing the findings in a peer-reviewed journal or by some other form of external peer review, such as review by the Agency's Science Advisory Board.

nothing to diminish concerns about how it is conducted. Rather, its indispensability heightens the importance of doing it well.

Because the EPA necessarily makes extensive use of benefit transfer, the Agency has a strong interest in developing standards for its application and improving the methods and data available to analysts when conducting BCAs (Iovanna and Griffiths 2006). Improving the methods of benefit transfer has been a focus of research by environmental economists at least since 1992 when a special issue of *Water Resources Research* was devoted to the topic. Since that time, a substantial body of conceptual and empirical work has been conducted to examine the conditions under which benefit transfers can provide sufficiently accurate and precise estimates of total WTP for environmental improvements in new policy cases.

WTP is not directly observable, so assessments of the accuracy of benefit transfer errors are generally based on tests of convergent validity (Rosenberger 2015). This involves comparisons of two or more alternative estimates of the same theoretical construct. That is, benefit estimates based on primary data that pertain specifically to a policy case are compared to alternative benefit estimates based on transfers of unit values or value functions from prior study cases to the policy case, with the accuracy typically expressed as a percentage difference between the two estimates. A recent review of the accuracy of benefit transfers using the convergent validity criterion was conducted by Kaul et al. (2013). The authors examined 31 studies that provide 1071 estimates of benefit transfer errors. They found that the absolute benefit transfer errors ranged from 0% to nearly 7500%, with a mean of 172% and a median of 39%. When the authors excluded the most extreme 14% of the observations, errors ranged between 0 and 172%, the mean was 42%, and the median was 33%. The authors also found that function transfers tend to be more accurate than value transfers; transfers of values for environmental quantity changes tend to be more accurate than those for quality changes; geographic similarity between sites improves the accuracy of transfers, especially for value transfers; combining information from multiple studies improves the accuracy of transfers; and that transfers based on stated preference valuation formats with more options per question, such as choice experiments, have larger transfer errors than methods with fewer choices per question, such as contingent valuation surveys. While considerable progress has been made on understanding benefit transfer errors and refining benefit transfer methods in recent years, more work remains to be done to increase the reliability of benefit transfers and to agree upon best practices for their conduct.⁴

While the EPA relies on benefit transfers in virtually all of its BCAs, some categories of environmental regulations require more complicated transfers than others. Toward the relatively less difficult end of the spectrum, regulations promulgated under the Clean Air Act typically reduce a variety of hazardous air pollutants that are known to increase human mortality risks (USEPA 2011).⁵ Reductions in mortality risks typically account for the lion's share of total benefits in these cases (Cropper et al. 2011). When a single endpoint dominates the benefits, this greatly focuses the analysis. In many cases it also is plausible to assume that the marginal WTP will be roughly constant over the relevant range of mortality risk changes. This means that a simple unit value transfer approach—multiplying an estimate of the average marginal WTP for mortality risk reductions (also known as the "value of statistical life," or VSL) by the change in the expected number deaths each year—is often

⁴ For a recent review and guide for conducting benefit transfers, see Johnston et al. (2015). Another important review of benefit transfer was the special issue of *Ecological Economics* in 2006, which assembled papers from a previous workshop on benefit transfer methods also sponsored by the EPA (Wilson and Hoehn 2006).

⁵ We hasten to emphasize "relatively" in "relatively less difficult;" e.g., see Viscusi (2010) and Cropper et al. (2011).

suitable for estimating the aggregate benefits for most air pollution regulations.⁶ Furthermore, when estimating the VSL, the EPA is able to draw upon a relatively large body of empirical nonmarket valuation research, including hedonic wage studies and stated preference surveys, that estimate marginal WTP for the precise endpoint of interest.

Toward the more difficult end of the spectrum are regulations promulgated under the Clean Water Act. The simple unit value transfer approach commonly used for air quality regulations is often not suitable for surface water quality regulations for several reasons. First, EPA surface water quality regulations typically do not lead to substantial changes in human mortality risks, so the VSL does not play a major role in the benefits assessment; ecological health rather than human health is the focus of the analysis. Second, many ecological endpoints may be affected with no single endpoint dominating the aggregate benefits. Unlike cases where clearly defined human health outcomes are the primary endpoints of interest, many ecological endpoints must be defined by the researchers in the course of a nonmarket valuation study (Boyd and Krupnick 2013). Third, many of the relevant ecological endpoints may be complements or substitutes, and so cannot be examined in isolation. Fourth, WTP for water quality changes may depend on many individual- and neighborhood-level attributes, including the avidity for outdoor recreation activities, the relative scarcity or abundance of water bodies suitable for recreation near the individual's home, and the prevailing level of environmental quality in those water bodies. Fifth, marginal WTP for changes in the affected endpoints may not be constant over the relevant range. All of these complicating factors make it essential to account for changing baseline conditions, relationships among valued ecological endpoints, and the spatial configuration of households with respect to the affected resources, which often requires a high-dimensional benefit function transfer approach. Due to these complications, the EPA has a special interest in improving benefit transfer approaches suitable for valuing surface water quality improvements and related ecological resources (USEPA 2006a).

The EPA also has an interest in advancing the state-of-the-practice for benefit transfers more broadly. An instructive case is a recent study designed to estimate the benefits of water quality improvements under the 2010 Chesapeake Bay Total Maximum Daily Load (TMDL) requirements (Phillips and McGee 2016). The study employed a type of unit-value benefit transfer approach. Estimates of ecosystem service benefits reported in a collection of primary nonmarket valuation studies were normalized by the geographic extent of their associated land use types within their respective study areas, and these normalized values were then transferred to their closest analog land use types across the entire Chesapeake Bay watershed assuming a fixed \$/acre unit value for each land use type. This transfer approach has its origins in the work of Costanza et al. (1997), and may have a superficial plausibility to non-economists because it involves calculations with conformable units: $\alpha \propto \alpha \propto \alpha$ for each land use type summed across all land use types gives a final number in units of \$, which is intended to represent the total value of the ecosystem services in the study area. However, this approach does not account for a variety of factors that are known to have a strong influence on people's WTP and may differ substantially among the study cases and policy cases. These missing factors include the number, proximity, and socio-demographic characteristics of households who might benefit from the environmental improvements, the availability of nearby substitute environmental resources, the baseline environmental quality levels, and the magnitudes of the quality changes being valued. If a benefit-transfer method

⁶ An important implication of this unit value transfer approach is that the baseline distribution of mortality rates among the population of affected individuals has no influence on the estimated benefits, so no complications arise when analyzing multiple regulations simultaneously or sequentially (Newbold et al. 2017).

is to be useful for evaluating regulations under the Clean Water Act, it must be sufficiently flexible to be tailored to local conditions; one size will not fit all for this purpose.

In the remainder of this essay, we describe some of the main benefit transfer challenges that EPA analysts continue to face on a regular basis. Our discussion is loosely structured by reference to the basic steps of an idealized benefit transfer, so we begin with a brief outline of those steps. As we proceed we highlight some of the remaining open questions where we believe further research or refinement of existing benefit transfer methods holds the greatest promise for increasing the credibility of BCAs conducted by the EPA.

2 Steps of an Idealized Benefit Transfer

The benefit transfer process consists of four basic steps (USEPA 2010, pp. 7-45–7-49), which can be summarized as follows.

- *Step 1.* Describe policy case: identify those characteristics of the policy case that are expected to have a measurable influence on total WTP.
- Step 2. Select study cases: develop and apply explicit selection criteria based on indicators of internal and external validity to identify one or more suitable study cases for transfer, accounting for the degree of similarity among the resources being valued, the baseline levels and magnitude of quality or quantity changes, and the affected households.
- Step 3. Transfer values: estimate a unit value or transfer function, based on either a single study or a meta-analysis of multiple studies, and calculate WTP for the policy case.
- *Step 4.* Report results: describe all key judgments and assumptions, and quantify and report the WTP estimates and their uncertainty.

3 EPA Benefit Transfer Challenges

The basic steps outlined above are conceptually straightforward, but many complications invariably arise in practice. First, it is often difficult to confirm the internal validity of value estimates that are reported in existing nonmarket valuation studies and might be used for benefit transfers. That is, before we transfer a benefit estimate reported in an existing study to a new policy case, possibly making adjustments for differences between the two settings, we must first verify that the original study used a reliable nonmarket valuation method and an experimental design that was free of systematic biases. For example, consider the Ecosystem Services Valuation Database (van der Ploeg and de Groot 2010), which was assembled for The Economics of Ecosystems and Biodiversity project (TEEB 2016) and has been used as a source of primary value estimates in a number of subsequent benefit transfer exercises (UK National Ecosystem Assessment 2011; Batker and Schmidt 2015). When analysts consult such a source, it is important to determine the provenance and hence the reliability of the studies cited therein. A review of the TEEB valuation database indicates that more than a third of the recorded values (456 of 1310) are themselves from *other* benefit transfer studies, not from primary valuation studies. Furthermore, many benefit estimates in the database are based on replacement costs, even though a strict set of conditions must be met for replacement costs to be a valid measure of benefits (Bockstael et al. 2000). Other errors, such as confusing marginal and total values, are also common among the value estimates recorded in the TEEB database. In a recent review, Blomqvist and Simpson (2017) found that over half of a sample of 30 estimates chosen at random from the TEEB database were flawed in ways that would have made their use in benefit transfer highly suspect. At a minimum, any estimate of value on which a benefit transfer is to be based should be consistent with received theory; it should not, for example be based on discredited methods such as embodied energy (Johnston et al. 2015, p. 7).

Another potential threat to the validity of primary valuation estimates that might be used for benefit transfers is publication bias, which occurs when the outcome of a study influences the researchers' choice of whether to submit the report for publication or the editors' choice of whether to accept it. Card and Krueger (1995) described three sources of publication bias: reviewer and editors may be more pre-disposed to accept studies that are consistent with the conventional wisdom; researchers may use the conventional wisdom to select models; and everyone may treat statistically significant results more favorably. Publication bias has been a concern for decades, and recent reviews of experimental results in medicine (Ioannides 2005) and psychology (Open Science Collaboration 2015) have raised doubts about the general credibility of empirical claims in these and other fields. Publication bias affects any systematic review of research results and can lead to problems for benefit transfers in particular. As explained by Rosenberger and Stanley (2006), nonmarket valuation studies may be selected for publication based largely on their methodological innovations and not necessarily provide reliable benefit estimates. Rosenberger and Stanley's (2006, p. 376, Table 3) review of previous studies and their own meta-analysis suggest that such selection effects do in fact lead to a bias: published estimates of WTP tend to be smaller on average than unpublished estimates. A number of meta-analysis techniques have been developed to identify and correct for publication bias (Nelson 2015), but these are still not widely used in the development of benefit transfer functions.

Another challenge in the early steps of benefit transfer is determining the suitability of candidate study cases for transfer to the policy cases. Is the environmental resource that was valued in a study case sufficiently similar to the resources that will be affected in the policy case to justify extrapolation? How similar is similar enough? Boyle et al. (2009) explained how structural benefit transfer functions can relax the need for strict site similarity in identifying relevant study cases. However, when study cases do not match the policy case on all relevant dimensions, the primary studies must measure and report the levels of the relevant factors to allow a preference calibration or estimation approach to account for differences among those factors across the study cases and to facilitate the necessary adjustments when transferring the value estimates to the policy cases. When assembling observations of value estimates for use in a meta-regression, EPA analysts have often found it challenging to gather accurate information on all of the important resource and study area attributes to be used as control variables in the regression.

Benefit transfer researchers have raised concerns about incomplete reporting in primary studies for some time (Brookshire and Neill 1992; Boyle and Bergstrom 1992; Loomis and Rosenberger 2006; Johnston et al. 2015). In addition to the details necessary for screening adequate studies, such as reporting on the commodity being valued, the market area and population, and welfare measures (Boyle and Bergstrom 1992; Loomis and Rosenberger 2006), pertinent information is not always reported in the published article. For example, authors may omit descriptions of design features of a stated preference survey instrument that are known to influence respondents' answers, summary statistics describing the distribution of demographic attributes or attitudes in the sample of respondents, secondary estimations that might improve benefit transfer applications (Desvousges et al. 1992), and raw data to allow for additional analyses. Sufficient incentives may not exist for researchers to routinely report information that might be necessary for benefit transfers but is not directly relevant for the

central research questions of the primary study. Furthermore, editors of scholarly journals often favor innovative theoretical or methodological advances over "routine" empirical policy research, so academic researchers have little incentive to design studies, provide data, and report results that are tailored to support government economic analyses (Smith and Pattanayak 2002; Johnston and Rosenberger 2010; Johnston et al. 2015).

A number of proposals to address these reporting issues have been made—such as developing inventories of primary studies, raw data, and questionnaires, or founding a new journal focused on replicating empirical studies—but very few have been implemented (Loomis and Rosenberger 2006). It is becoming more common for peer-reviewed journals to request and for authors to provide supplementary materials to be posted online that may help increase the use of primary studies for benefit transfers. However, authors typically are not required to provide more information than what is needed for replication.

After potentially relevant study cases have been identified and their internal validity verified in step 2 of the benefit transfer process, it is important to compare the baseline and policy levels of environmental quality in both the study and policy cases. A perennial challenge in this step for valuing changes in surface water quality stems from the fact that natural scientists use a wide variety of chemical, physical, and biological measures to assess water quality conditions in different settings. Because there is no single measurement scale for "water quality" that is widely-used among scientists, environmental economists who engage in nonmarket valuation of surface water quality improvements must either adopt one or a small number of the many available indicators of water quality, or they must devise their own water quality scale that best suits the aims of their study. The difficulty in measuring or representing water quality stems partly from the many different human uses of surface waters. For example, swimmers may respond to different physical and chemical characteristics of the water than recreational anglers. The resulting wide range of environmental quality measures that have been used in valuation studies (Abt Associates 2016) can make it difficult to achieve the commodity consistency needed for benefit transfer models (Boutwell and Westra 2013).

To achieve commodity consistency in the face of this diversity of measures, the EPA currently relies on a multi-metric water quality index (WQI) to represent overall water quality conditions for the purposes of regulatory benefits assessments (Walsh and Wheeler 2013). Some primary stated-preference studies in the EPA's meta-dataset used the WQI directly, but for other studies it was necessary for the analysts who assembled the meta-data to use their own best judgments to translate more-or-less distinct measures onto the WQI scale *ex post* (USEPA 2015a). This has allowed the EPA to develop a sufficiently large database of study case observations to support the estimation of a benefit transfer function for surface water quality improvements using a meta-regression approach. While the translation of disparate water quality measures into the WQI allows the inclusion of a larger number of study case observations, it also introduces another source of uncertainty that has not been quantitatively assessed. The uncertainty associated with this step of the benefit transfer process could be reduced by additional empirical research on the influence of various physical, chemical, and biological attributes on people's direct and passive use of surface water bodies.

A further challenge that arises when selecting appropriate study cases for transfer to EPA water quality regulations concerns the magnitude of the environmental quality changes examined. Ideally, both the baseline levels and changes in water quality in the study cases and policy cases would be similar. Most stated preference water quality valuation studies have examined improvements or decrements on the order of 10–20% of the full range of

possible water quality levels represented on the WQI scale.⁷ However, 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 (USEPA 2015a). Similarly, in primary studies that value changes in the expected catch rates of fish by recreational anglers, it is common to value the improvement of expected catch by one or more fish, while most EPA regulations are expected to increase expected catch rates by a small fraction of a fish per trip. The larger are the disparities in the environmental quality levels and changes between the study and policy cases, the more pressure will be put on the assumed form of the benefit transfer function that is estimated or calibrated using the study case values. This is especially important when the benefit transfer function will be extrapolated well outside of the range of environmental quality changes examined in the study cases.

Another challenge EPA analysts often face is that available study cases are thin in many important dimensions. The EPA often promulgates national-level regulations and so must estimate WTP for all households in the U.S. for water quality improvements in many water bodies across the country. Ideally, analysts would be able to draw on a set of primary studies that have valued a wide range of quality improvements in all types of water bodies in many locations across the U.S. However, many gaps remain in the coverage provided by the existing body of empirical water quality valuation studies. The most comprehensive meta-analysis database assembled by the EPA to date was used to estimate a meta-regression benefit transfer function for the Steam Electric rule (USEPA 2015a).⁸ The meta-dataset comprises 140 observations of WTP from 51 stated preference studies. Among the 24 control variables included in the meta-regression estimating equation were dummy variables representing 14 factors with 2 levels and 1 factor with 4 levels. Considering these discrete variables alone, a full factorial experimental design would have $2^{14} \times 4^1$ cells, which of course is far larger than the number of observations in the dataset. Even considering a main-effects only model, there are less than six observations per control variable. Furthermore, as is to be expected of data that were not collected by a controlled experimental protocol, the distribution of observations among control variables is idiosyncratic and many of the design cells are empty. For example, two factors that might be expected to have a large influence on WTP are the geographic location of the affected water bodies and the types of recreational uses that are most prevalent in those water bodies. Table 1 shows the number of meta-data observations by geographic region and recreational use. Many of the primary studies examined water quality improvements confined to a single U.S. state-often focusing on just one or a few specific water bodies—so we can infer that many areas of the country are covered by no or only a very few observations and only a few geographic areas are well covered by studies that examine each of the major recreational use categories. Using a benefit transfer function based on these data to estimate WTP for improvements in the quality of water bodies that are primarily used for boating in the south, for example, would involve extrapolating the average WTP from two observations, probably from only one or two states, to all 13 states in that broad geographic region. Also note that three cells in the table contain no observations, so benefit transfers for each of those uses in each of those regions would rely entirely on the observations for the respective use in other regions and observations for other uses in the respective region.

⁷ The average change on the 100-point water quality scale valued by the studies represented in the EPA's stated preference water quality meta-analysis (USEPA 2015a) is roughly 14 points.

⁸ Rosenberger and colleagues have assembled a large dataset of recreation site access values; see Moeltner and Rosenberger (2014) for a recent application. The dataset is useful for examining site closures or other changes in access but would need to be supplemented with contemporaneous measures of environmental quality to be useful for evaluating environmental policies.

	Swimming	Fishing	Boating	Unspecified activities
Northeast ME, NH, VT, MA, RI, CT, NY	3	0	0	7
Central OH, MI, IN, IL, WI, MN, IA, MO, ND, SD, NE, KS, MT, WY, UT, CO	11	4	11	26
Southeast and Southwest NC, SC, GA, FL, KY, TN, MS, AL, AR, LA, OK, TX, NM	5	0	2	17
Mid-Atlantic and Western NJ, DE, MD, DC, PA, WV, VA, WA, OR, ID, CA, NV, AZ, AK	18	4	3	34

 Table 1
 Number of observations of stated-preference study estimates of WTP for water quality improvements used to estimate the EPA's meta-regression benefit transfer function, by geographic regions and use categories (USEPA 2015a)

None of these observations are meant to diminish the importance of this or other metaanalyses of nonmarket valuation studies. We strongly support making the best possible use of the relevant information that happens to be available, whether it was collected in an experimental or an observational setting. Our main goal here is to highlight the fact that the existing body of stated preference-based water quality valuation studies adds up to sparse and thin coverage of the various geographic regions, water body types, recreational uses, and other factors that might be expected to have a strong influence on WTP for water quality improvements.⁹ Therefore, one straightforward—though not easy or inexpensive—way to improve the quality of benefit transfers would be to increase the quantity of relevant primary valuation studies. One useful preliminary research task would be to identify those regions of the design space where new studies would be expected to provide the most valuable new information for benefit transfers in light of the suite of environmental policy proposals that might be considered in the foreseeable future.

In step 3 of the benefit transfer process, many methodological questions remain about how best to estimate a unit value or function that can be used for benefit transfers, including: What is the proper role for theory in specifying the form of a benefit transfer function (Newbold et al. 2017)? What are the relative advantages and disadvantages of structural, reduced form, and non-parametric estimation approaches (Blow and Blundell 2017), and under what conditions would one of these approaches be recommended over the others? What is the best strategy for maximizing the prediction accuracy of benefit transfer functions, and what is the best way to handle study design control variables when estimating and applying a benefit transfer function (Boyle and Wooldridge 2017)? How should the meta-data observations be weighted to achieve the most efficient transfer model? How should publication bias be diagnosed and accounted for in meta-analyses? Some of these and other related questions were addressed, but not necessarily resolved, in two EPA reports on the use of meta-analysis for estimating the value of statistical life (USEPA 2006b, 2007) and a related article by Nelson and Kennedy (2009).

Another challenge relevant for step 3 involves identifying the geographic extent of the market and the rate of distance decay of households' WTP for environmental quality improvements. The variation of WTP with distance is a crucial element of outdoor recreation demand models, in which costs of access are strongly related to the travel distance between a recreational users' residence and her destination. In hedonic models the distance between a

⁹ This basic conclusion also seems consistent with the broader review of the literature, also including revealed preference studies, conducted by Abt Associates (2016).

property and a river or lake is sometimes included as an explanatory variable, and researchers have generally found that the influence of resource quality on property values declines substantially within 1–2 km away from the resource; e.g., (Walsh et al. 2017). The treatment of distance decay is more speculative with stated preference approaches, however, particularly inasmuch as one common argument for employing stated preference methods is that some of the values they capture may be entirely divorced from use.

Researchers have demonstrated that determining the correct population to whom to attribute benefits can often swamp considerations related to individual value estimates (Smith 1993; Loomis 2000; Bateman et al. 2006). Depending on the size of the jurisdiction, studies based on political boundaries may not capture the full market of beneficiaries, and not allowing for value to vary by distance may overestimate individual WTP from distant households. These issues relate to benefit transfer in at least two ways. First, how do the original studies address extent of market? Second, how and to what population should a transferred unit value or function be applied?

Incorporating distance into original water quality WTP estimates dates back over three decades (Sutherland and Walsh 1985), but many early studies take the approach of using states or other political jurisdictional boundaries to define the market, as well as applying a uniform value to the entire population independent of distance. Subsequent stated preference studies have allowed values to decay more smoothly with distance and examined distance decay variation between users and nonusers (Hanley et al. 2003), iconic environmental goods (Rolfe and Windle 2012; Loomis 1996; Moore et al. 2015), the influence of spatial heterogeneity including the interaction between distance and substitutes (Bateman et al. 2011; Jorgensen et al. 2013; Schaafsma 2015; Schaafsma et al. 2012), and hot spots and related local versus global distance issues (Johnston and Ramachandran 2014).

Standard practice at the EPA for measuring the benefits of surface water quality regulations has evolved from transferring benefits from single studies (USEPA 1982, 1987) to the current approach of transferring benefits from a collection of studies using meta-analysis(USEPA 2015a).¹⁰ The EPA's most recent meta-analysis of stated preference studies developed for the Steam Electric rule includes several study-specific geographic variables related to the extent of the market, but does not incorporate distance decay directly. Instead, the extent of the market for water quality improvements is exogenously imposed by calculating water quality changes within a 100-mile radius around each household (USEPA 2015a). Newer research has incorporated spatially explicit data into benefit transfers to address the extent of market and other geospatial questions (Johnston et al. 2016), but additional work in this area is still needed to develop more comprehensive approaches for discounting water quality improvements over space.

Finally, in step 4 of the benefit transfer process, it remains challenging to characterize uncertainty in benefit transfer estimates. Benefit transfer is inherently an out-of-sample extrapolation problem. The ability of a benefit transfer function to match the WTP estimates for a set of study cases is akin to a within-sample measure of fit, while we are really interested in the accuracy of the function in predicting WTP for new policy cases that were not used to estimate the function in the first place. The study by Kaul et al. (2013) summarized the overall distribution of estimated prediction errors based on benefit transfer convergent validity studies, but more research is needed to develop reliable methods for quantifying the uncertainty in prospective benefit transfers on a case-specific basis. One approach would be to use a form of cross-validation (Hastie et al. 2001) to characterize the out-of-sample

¹⁰ See Griffiths et al. (2012) or Wheeler (2015) for a detailed history of the EPA's benefit estimation approaches for Clean Water Act rules.

prediction accuracy for benefit transfer functions. Stapler and Johnston (2009) used a crossvalidation approach to examine out-of-sample predictions for a meta-regression model of the marginal value of fish from a collection of recreation demand studies; Klemick et al. (2016) examined out-of-sample transfer errors based on hedonic property value estimates of the benefits of increasing water clarity in the Chesapeake Bay; and Newbold et al. (2017) used cross-validation to help guide variable selection for a meta-regression model of WTP estimates from stated preference studies.

4 Conclusions

In this essay we have described why the circumstances of the EPA's BCAs often necessitate reliance on benefit transfers rather than conducting original studies. The fact that benefit transfers are unavoidable expedients does not mean we should ignore or be heedless of their limitations. Results to be transferred from study to policy cases must be reliable, which means that they must be grounded in received theory and derived using valid empirical methods. Moreover, study cases must be selected with care, even if the results of a particular study are internally valid. Publication bias may, for example, mean that analysts would oversample from studies that arrived at significant findings, and no original study, no matter how carefully it was conducted or reported, can shed light on a policy scenario to which it bears no resemblance.

These and other challenges complicate the EPA's task in estimating the benefits (and, *mutatis mutandis*, the costs) of proposed rules. Of course such challenges are not unique to the EPA, nor even just to government regulatory agencies. To the extent that applying the results of *any* empirical study to other circumstances involves some extrapolation of results, any policy choice informed by empirical findings involves some benefit transfer. Guidance concerning best practices in benefit transfer could be of considerable value far beyond the sphere of environmental regulation.

That guidance will necessarily involve two key elements. The first involves better laying out what constitutes not only a *good*, but also a *transferable* study case. As we have noted above, it may not be enough that a study provide evidence concerning values in the area in which it was conducted; there must also be enough information to combine those value estimates with other estimates derived from other places. The second element of guidance will concern how best to synthesize the results of different studies. We might sometimes imagine that eventually we would have original studies conducted in every conceivable location of interest, obviating the need to transfer benefits from where we do have original estimates to where we do not. It seems unreasonable to suppose that this will happen anytime soon—if it ever happens at all. In the interim, we must continue to develop best practices for combining the limited information we do have and putting it to the best possible use for evaluating proposed public policies.

References

Abt Associates (2016) Summary and discussion of literature linking water quality to economic value. Prepared for Office of Research and Development, Office of Policy, and Office of Water, U.S. Environmental Protection Agency under Contract No. EP-C-13-039. Bethesda, MD

Banzhaf HS (2010) Consumer surplus with apology: a historical persepctive on nonmarket valuation and recreation demand. Annu Rev Resour Econ 2:183–207

Banzhaf HS (2016) The environmental turn in natural resource economics. RFF Discussion Paper 16-27. Retrieved from http://www.rff.org/files/document/file/RFF-DP-16-27.pdf

- Bardach E (2004) The extrapolation problem: how can we learn from the experience of others? J Policy Anal Manag 23(2):205–220
- Bateman IJ, Day BH, Georgiou S, Lake I (2006) The aggregation of environmental benefit values: welfare measures, distance decay and total WTP. Ecol Econ 60(2):450–460
- Bateman IJ, Brouwer R, Ferrini S, Schaafsma M, Barton DN, Dugbaard A, Navrud S (2011) 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. Environ Resour Econ 50(3):365–387
- Batker D, Schmidt R (2015) Environmental and Social Benchmarking Analysis of Nautilus Minerals Inc. Solwara 1 Project. Earth Economics, Tacoma, WA. Retrieved from https://drive.google.com/file/d/ 0ByzIUWI76gWVbjM1S045cjUwckk/view
- Blomqvist L, Simpson RD (2017) The value of ecosystem services: what is the evidence. In Kareiva P, Marvier M, Silliman B (eds) Effective conservation science: data not dogma. Oxford University Press, Oxford
- Blow L, Blundell R (2017) A nonparametric revealed preference approach to measuring the value of environmental quality. Environ Resour Econ. https://doi.org/10.1007/s10640-018-0229-9
- Bockstael NE, Freeman AM III, Kopp RJ, Portney PR, Smith VK (2000) On measuring economic values for nature. Eniron Sci Technol 34:1384–1389
- Boutwell JL, Westra JV (2013) Benefit transfer: a review of methodologies and challenges. Resources 2(4):517–527
- Boyd J, Krupnick A (2013) Using ecological production theory to define and select environmental commodities for nonmarket valuation. Agric Resour Econ Rev 42(1):1–32
- Boyle KJ, Bergstrom JC (1992) Benefit transfer studies: myths, pragmatism, and idealism. Water Resour Res 28(3):657–663
- Boyle KJ, Wooldridge JM (2017) Understanding error structures and exploiting panel data in meta-analytic benefit transfers. Environ Resour Econ. https://doi.org/10.1007/s10640-017-0211-y
- Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC (2009) Ncessary conditions for valid benefit transfers. Am J Agric Econ 91(5):1328–1334
- Brookshire DS, Neill HR (1992) Benefit transfers: conceptual and empirical issues. Water Resour Res 28(3):651–655
- Card D, Krueger AB (1995) Time-series minimum-wage studies: a meta-analysis. Am Econ Rev 85:238-243
- Cartwright N, Hardie J (2012) Evidence-based policy: a practical guide to doing it better. Oxford University Press, New York
- Copeland CW (2011) The federal rulemaking process: an overview. CRS Report to Congress, Congressional Research Service
- Costanza R, d'Arge R, de Groot R, Farber S, Grasso M, Hannon B, van den Belt M (1997) The value of the world's ecosystem services and natural capital. Nature 387:253–260
- Cropper M, Hammitt JK, Robinson LA (2011) Valuing mortality risk reductions: progress and challenges. Annu Rev Resour Econ 3:313–336
- Desvousges WH, Naughton MC, Parsons GR (1992) Benefit transfer: conceptual problems in estimating water quality benefits using existing studies. Water Resour Res 28(3):675–683
- Easter KW, Archibald SO (1998) Benefit-cost analysis in U.S. environmental regulatory decisions. In: Proceedings of the sixth joint conference on agriculture, food, and the environment, Minneapolis, MN. Retrieved from http://ageconsearch.umn.edu/record/14475/files/c6east01.pdf
- Fraas A (1991) The role of economic analysis in shaping envrionmental policy. Law Contemp Probl 54(4):113– 125
- Griffiths C, Klemick H, Massey M, Moore C, Newbold S, Simpson D, Wheeler W (2012) EPA valuation of surface water quality improvements. Rev Environ Econ Policy 6:130–146
- Hanley S, Spash C (1993) Cost-benefit analysis and the environment. Edward Elgar, Brookfield
- Hanley N, Schlapfer F, Spurgeon J (2003) Aggregating the benefits of environmental improvements: distance decay functions for use and non-use values. J Environ Manag 68:297–304
- Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning: data mining, inference, and prediction, 2nd edn. Springer, Stanford
- Howick J, Glasziou P, Aronson P (2013) Problems with using mechanisms to solve the prolem of extrapolation. Theor Med Bioeth 34:275–291
- Hufschmidt MM (2000) Benefit-cost analysis: 1933–1985. J Contemp Water Res Educ 116(1):11
- Ioannides JP (2005) Why most published research findings are false. PLoS Med 2(8):696–701
- Iovanna R, Griffiths C (2006) Clean water, ecologial benefits, and benefits transfer: a work in progress at the U.S. EPA. Ecol Econ 60:473–482
- Johnston RJ, Ramachandran M (2014) Modeling spatial patchiness and hot spots in stated preference willingness to pay. Environ Resour Econ 59(3):363–387

- Johnston RJ, Rosenberger RS (2010) Methods, trends and controversies in contemporary benefit transfer. J Econ Surv 24(3):479–510
- Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) (2015) Benefit transfer of environmental and resource values: a guide for researchers and practitioners. Springer, New York
- Johnston RJ, Besedin EY, Stapler R (2016) Enhanced geospatial validity for meta-analysis and environmental benefit transfer: an application to water quality improvements. Environ Resour Econ. https://doi.org/10. 1007/s10640-016-0021-7
- Jorgensen SL, Olsen SB, Ladenburg J, Martinsen L, Svenningsen SR, Hasler B (2013) Spatially induced disparities in users' and non-users' WTP for water quality improvements-testing the effect of multiple substitutes and distance decay. Ecol Econ 92(1):58–66
- Kaul S, Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC (2013) What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. J Environ Econ Manag 66:90–104
- Klemick H, Griffiths C, Guignet D, Walsh P (2016) Improving water quality in an iconic estuary: an internal meta-analysis of property value impacts around the Chesapeake Bay. Environ Resour Econ. https://doi. org/10.1007/s10640-016-0078-3
- Loomis JB (1996) How large is the extent of the market for public goods: evidence from a nationwide contingent valuation survey. Appl Econ 28(7):779–782
- Loomis JB (2000) Vertically summing public good demand curves: an empirical comparison of economic versus political jurisdictions. Land Econ 76:312–321
- Loomis JB, Rosenberger RS (2006) Reducing barriers in future beneit transfers: needed improvements in primary study design and reporting. Ecol Econ 60:343–350
- Moeltner K, Rosenberger RS (2014) Cross-context benefit transfer: a Bayesian search for information pools. Am J Agric Econ 96(2):469–488
- Moore C, Guignet D, Maguire K, Dockins C, Simon N (2015) A stated preference study of the Chesapeake Bay and watershed lakes. NCEE Working Paper #15-06. Accessed at https://www.epa.gov/environmentaleconomics/working-paper-stated-preference-study-chesapeake-bay-and-watershed-lakes
- Nelson JP (2015) Meta-analysis: statistical methods. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) Benefit transfer of environmental and resource values: a guide for researchers and practitionersd, vol 15. Springer, Dordrecht
- Nelson JP, Kennedy PE (2009) The use (and abuse) of meta-analysis in envvironmental and natural resource economics: an assessment. Environ Resour Econ 42:345–377
- Newbold SC, Walsh PJ, Matthew Massey D, Hewitt J (2017) Using structural restrictions to achieve theoretical consistency in benefit transfers. Environ Resour Econ. https://doi.org/10.1007/s10640-017-0209-5
- OMB (2003) Circular A-4: regulatory analysis. Office of the President. Retrieved from http://www.whitehouse. gov/omb/circulars
- Open Science Collaboration (2015) Estimating the reproducibility of psychological science. Science. https:// doi.org/10.1126/science.aac4716
- Pearce D (2002) An intellectual history of environmental economics. Annu Rev Energy Environ 27:57-81
- Phaneuf DJ, Smith VK, Palmquist RB, Pope JC (2008) Integrating property value and local recreation models to value ecosystem services in urban watersheds. Land Econ 84(3):361–381
- Phillips S, McGee B (2016) Ecosystem service benefits of a cleaner Chesapeake Bay. Coast Manag 44(3):241– 258
- Rolfe J, Windle J (2012) Distance decay functions for iconic assets: assessing national values to protect the health of the Great Barrier Reef in Australia. Environ Resour Econ 53:347–365
- Rosenberger RS (2015) Benefit transfer validity and reliability. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) Benefit transfer of environmental and resource values: a guide for researchers and practitioners. Springer, Dordrecht
- Rosenberger RS, Stanley TD (2006) Measurement, generalization, and publication: sources of error in eneit transfers and their management. Ecol Econ 60:372–378
- Schaafsma M (2015) Spatial and geographical aspects of benefit transfer. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) Benefit transfer of environmental and resource values: a guide for researchers and practitioners, vol 18. Springer, Dordrecht
- Schaafsma M, Brouwer R, Rose J (2012) Directional heterogeneity in WTP models for environmental valuation. Ecol Econ 79(1):21–31
- Smith VK (1993) Nonmarket valuation of environmental resources: an interpretive appraisal. Land Econ 69(1):1–26
- Smith VK, Pattanayak SK (2002) Is meta-analysis a Noah's Ark for non-market valuation? Envrion Resour Econ 22:271–296
- Stapler RW, Johnston RJ (2009) Meta-analysis, benefit transfer, and methodological covariates: implications for transfer error. Environ Resour Econ 42:227–246

- Steel D (2010) A new approach to argument by analogy: extrapolation and chain grahs. Philos Sci 77(5):1058– 1069
- Sutherland RJ, Walsh R (1985) Effect of distance on the preservation value of water quality. Land Econ 61:281–291
- TEEB (2016) The economics of ecosystems and biodiversity. Retrieved from http://www.teebweb.org/
- UK National Ecosystem Assessment (2011) UK National Ecosystem Assessment: synthesis of key findings. UNEP-WCMC, Cambridge, UK
- USEPA (1982) Regulatory impact analysis of effluent limitation guidelines regulation for the iron and steel industry. EPA Report W.82.11
- USEPA (1987) Regulatory impact analysis of the effluent guidelines regulation for the organic chemicals, plastics, and synthetic fibers industry. EPA Report W.87.16
- USEPA (2006a) Ecological benefits assessment strategic plan. EPA-240-R-06-001. Retrieved from https:// yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0485-01.pdf/\$File/EE-0485-01.pdf
- USEPA (2006b) Report of the EPA workgroup on VSL meta-analyses. Retrieved from https://yosemite.epa. gov/ee/epa/eerm.nsf/vwRepNumLookup/EE-0494?OpenDocument
- USEPA (2007) SAB advisory on EPA's issues in valuing mortality risk reductions. Retrieved from https:// nepis.epa.gov/Exe/ZyPDF.cgi/P10007U3.PDF?Dockey=P10007U3.PDF
- USEPA (2009) Valuing the protection of ecological systems and services: a report of the EPA science advisory board. Retrieved from https://yosemite.epa.gov/sab%5Csabproduct.nsf/ F3DB1F5C6EF90EE1852575C500589157/\$File/EPA-SAB-09-012-unsigned.pdf
- USEPA (2010) Guidelines for conducting economic analyses. United States Environmental Protection Agency, Office of the Administrator. EPA 240-R-10-001. Retrieved from https://www.epa.gov/environmentaleconomics/guidelines-preparing-economic-analyses
- USEPA (2011) The benefits and costs of the clean air act from 1990 to 2020. U.S. Environmental Protection Agency, Office of Air and Radiation. Retrieved from https://www.epa.gov/sites/production/files/2015-2017/documentts/fullreport_rev_a.pdf
- USEPA (2015a) Benefit and cost analysis for the effluent limitations guidelines and standards for the steam electric power generating point source category. U.S. EPA, Office of Water, Office of Science and Technology
- USEPA (2015b) U.S. environmental protection agency peer review handbook, 4th edn. U.S. EPA, Science and Technology Policy Council, Washington, DC. Retrieved from https://www.epa.gov/sites/production/ files/2016-03/documents/epa_peer_review_handbook_4th_edition.pdf
- van der Ploeg S, de Groot RS (2010) The TEEB valuation database—a searchable database of 1310 estimates of monetary values of ecosystem services. Wageningen, Foundation for Sustainable Development, The Netherlands. Retrieved from https://www.es-partnership.org/services/data-knowledgesharing/ecosystem-service-valuation-database/
- Viscusi WK (2010) Policy challenges of the heterogeneity of the value of statistical life. Found Trends Microecon 6(2):99–172
- Walsh P, Wheeler W (2013) Water quality indices and benefit-cost analysis. J Benefit-Cost Anal 4(1):81–105. https://doi.org/10.1515/jbca-2012-0005
- Walsh P, Griffiths C, Guignet D, Klemick H (2017) Modeling the property price impact of water quality in 14 Chesapeake Bay counties. Ecol Econ 135:103–113
- Wheeler W (2015) Benefit transfer for water quality regulatory rulemaking in the United States. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R (eds) Benefit transfer of environmental and resource values: a guide for researchers and practitioners, vol 6. Springer, Dordrecht
- Wilson MA, Hoehn JP (2006) Valuing environmental goods and services using benefit transfer: the state-of-the art and science. Ecol Econ 60(2):335–342