VALUING IMPROVEMENTS IN THE ECOLOGICAL INTEGRITY OF LOCAL AND REGIONAL WATERS USING THE BIOLOGICAL CONDITION GRADIENT

Authors: Christian A. Vossler[a], Christine L. Dolph[b], Jacques C. Finlay[b], David A. Keiser[c], Catherine L. Kling[d], Daniel J. Phaneuf[e*]

[a]Department of Economics and Howard H. Baker Jr. Center for Public Policy, University of Tennessee, Knoxville, TN 37996
[b]Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108
[c]Department of Resource Economics, University of Massachusetts, Amherst, MA 01003
[d]Charles H. Dyson School of Applied Economics and Management and Cornell Atkinson Center for Sustainability, Cornell University, Ithaca, NY 14853
[e]Department of Agricultural and Applied Economics, University of Wisconsin, Madison, WI 53706

[*]Direct correspondence to Daniel Phaneuf
Email: dphaneuf@wisc.edu

Author Contributions: All authors designed the research, performed the research, and wrote the paper. C.A.V. analyzed the data.
Competing Interest Statement: The authors declare no conflicts of interest.
Classification: Social Sciences, Economic Sciences
Keywords: water pollution, Clean Water Act, stated preferences, willingness to pay

The published version of this study is in the *Proceedings of the National Academy of Sciences*.
The final publication is available at https://doi.org/10.1073/pnas.2120251119.

Abstract (250-word limit): Scientific knowledge related to quantifying the total monetary values for landscape-wide water quality improvements does not meet current regulatory and benefit-cost analysis needs in the US. In this study we address this knowledge gap by incorporating the Biological Condition Gradient (BCG) as a water quality metric into a stated preference survey capable of estimating the total economic value (use and nonuse) for aquatic ecosystem improvements. The BCG is grounded in ecological principles, and generalizable and transferable across space. Moreover, as the BCG translates available data on biological condition into a score on a six-point scale, it provides a simple metric that can be readily communicated to the public. We apply our BCG-based survey instrument to households across the Upper Mississippi, Ohio, and Tennessee River Basins and report values for a range of potential improvements that vary by location, spatial scale, and the scope of the water quality change. We find that people are willing to pay twice as much for an improvement policy that targets their home watershed (defined as a 4-digit hydrologic unit) versus a more distant one. We also find that extending the spatial scale of a local policy beyond the home watershed does not generate additional benefits to the household. Finally, our results suggest that non-use sources of value (e.g., bequest value, intrinsic aesthetic value) are an important component of overall benefits.

Significance Statement (120-word limit): Many lakes, rivers, and streams across the US do not meet water quality goals for ecological health and uses. To help determine where and to what extent water quality improvements should be sought, policymakers must consider the costs of regulations with the monetized values humans place on them. We develop a flexible survey approach for valuing water quality changes that uses a simple quality metric that incorporates both ecological use and ecological health. Our measure that can be broadly applied to different waterbodies and locations, and understood by the public. We surveyed a large number of households across the US Midwest and estimate values for potential policies that vary in their location, spatial scale, and the extent of the water quality improvement. The methods and estimated values have the potential to support various regulatory analyses.

1) Introduction

The US Clean Water Act (CWA) of 1972 is among the most ambitious of federal environmental statutes with a primary objective to "restore and maintain the chemical, physical and biological integrity of the Nation's waters." The CWA called for the elimination of pollution discharges into the nation's waters by 1985, with an interim goal of achieving water quality that is protective of fish, wildlife, and recreation by 1983. Neither goal was met, and many waterbodies remain in poor condition. In its most recently released national assessments, the U.S. Environmental Protection Agency (EPA) reports that 46 percent of river and stream miles are in poor biological condition and 21 percent of the nation's lakes have excessively high levels of nutrients and algae (US EPA 2017). More recent nationwide studies likewise indicate that critical water quality concerns remain (Gilbert 2020; Stets et al. 2020), suggesting the need for additional regulatory and pollution control efforts. Credible quantification of the monetized benefits of water quality will be important for setting reasonable goals and for communicating the rationale for new regulations. As described in Moore et al. (this volume), federal agencies such as the EPA are required to undertake cost-benefit analyses (CBA) to justify the stringency of their rules.

Capturing the economic value of environmental regulations in CBA requires quantifying both the costs and benefits in monetary terms, including those tied to a range of nonmarket services (for which market prices do not adequately capture the benefits). Examples of nonmarket services include recreational swimming and fishing, bird watching, and the desire to preserve ecological integrity and intact natural areas. When there are important sources of nonuse value – value not tied to observable human behavior (e.g., bequest and existence value) – the only established economic method for estimating total economic value (both use and nonuse) involves the application of carefully constructed surveys that ask people to state their preferences for potential policies.

The CBA assessments associated with early CWA rules focused on the benefits related to meeting designated uses of waterbodies (Freeman 1982; Carson and Mitchell 1993). For regulatory purposes a "designated use" serves as a reference point for determining if a water quality goal is met. For example, a use category for primary contact recreation ("swimmable") requires better water quality than protection of a sport fishery ("fishable"), which in turn requires better water quality than secondary contact recreation ("boatable"). If water quality is not sufficient to meet its intended use, a waterbody is deemed "impaired." By design, the early economic studies based on use designation did not measure the value of broader services related to ecological integrity. This reflected the statutory emphasis on recreation in the interim goal as well as limitations in our understanding of the complex relationships between human activities, water pollution, ecological integrity, and the range of ecosystem services provided by aquatic resources. In addition, the methodological innovations and applied experience in nonmarket valuation that enables credible estimation of a wider range of environmental services (e.g., biodiversity and ecological integrity) did not yet exist.

During the last two decades many of these knowledge gaps have been narrowed. Scientists now understand the key role played by pollution in the degradation of aquatic ecosystems, especially eutrophication leading to loss of habitat (e.g. Scheffer et al. 2001), reduced biodiversity (Villeneuve et al. 2015), negative impacts on fishery recruitment of key native species (Jacobson et al. 2017), greenhouse gas production (Gilbert 2020), the proliferation of harmful algal blooms (Michalak et al. 2013), and threats to drinking water quality and treatment costs (Pennino et al. 2017). At the same time, there have been considerable advances in stated preference research methods, including refinements to survey development and implementation, value elicitation, data analysis, and assessments of validity (Kling et al. 2012; Johnston et al. 2017). Theoretical work has improved our understanding of how to design survey instruments that are incentive compatible in the sense that they motivate respondents to truthfully reveal their valuations (Carson and Groves 2007; Vossler et al. 2012; Carson et al. 2014). Finally, researchers have begun to demonstrate feasible strategies for communicating complex ecological concepts in stated preference surveys (e.g., Bateman et al. 2011; Bishop et al. 2017).

In this research we leverage decades of progress in ecology and economics to develop a critical link connecting the aquatic health of waterbodies to economic value (Keeler et al. 2012). We particularly draw on the nonmarket valuation literature related to improving water quality in rivers and streams. Bergstrom and Loomis (2017) provide a thorough review of river restoration and valuation work to date. Our point of departure is to adopt a water quality index new to nonmarket valuation: the Biological Condition Gradient (BCG) (US EPA 2016). The BCG is grounded in ecological principles, generalizable and transferable across space, and consistent with current regulatory decision-making needs. We then demonstrate the use of this index to measure economic values for water quality that include *both* traditional use mechanisms (boatable, fishable, swimmable) and nonuse mechanisms related to ecological integrity and other ecosystem services. We apply our BCG framework by surveying a random panel of 2,000 households located in the Upper Mississippi, Ohio, and Tennessee River Basins (see Figure 1) and report values for a range of quality improvements. With this survey, we provide the first large-scale estimates of the total economic value (use and nonuse) for aquatic ecosystem improvements that are derived from a transferable elicitation method and built on state-of-the-art ecological concepts.

2) Defining the Good

Our valuation methodology uses the BCG to define the commodity for which we measure preferences. The BCG was developed by the EPA (Davies and Jackson 2006) and provides a spatially transferable biological assessment index of how water quality conditions change due to anthropogenic stressors (US EPA 2016). The index depicts departures from a reference "natural" or "undisturbed" condition and is measured by the diversity and relative abundance of freshwater taxa associated with ecosystem integrity for a specific waterbody type. The BCG is attractive for our purpose as it is designed to provide comparable interpretations of biological health across locations and waterbodies (US EPA 2016, p. 30). By using the BCG to define our commodity, we can provide empirical estimates that are broadly comparable across space and linked to a policy-relevant metric of ecological integrity.

The BCG consists of six levels, each associated with a different degree of departure from baseline ecosystem function and integrity. It is analogous to a dose-response curve where the dose represents the degree of anthropogenic stress (including pollution), and biological condition is the response. The degree of biological condition represents the consequences of multiple co-occurring stressors such as nutrient pollution, pesticides, sedimentation, and other physiochemical changes arising from human impacts in watersheds. Biological condition can therefore be improved by actions undertaken to address anthropogenic stressors, such as reducing nitrogen and phosphorus loadings from agricultural land use (e.g., Chambers et al. 2012).

To elicit economic values for changes in BCG levels, we translated the ecological concepts underlying each level into visual and textual representations that are understandable to survey respondents. We first characterized water quality conditions, human uses, and biological diversity supported at each level by identifying physical features of streams and rivers that are (a) visually evident and therefore may affect how people perceive water quality; and (b) known to be important drivers or correlates of ecosystem condition (e.g., Maddock 2001) and therefore likely to be associated with different BCG levels. This step was based on expert judgement supported by available habitat data in our study area. Important visual features included water color/clarity; river channel shape (natural versus channelized); flow conditions (diverse riffles and pools versus homogenous flows); riparian condition (diversity/abundance of streambank vegetation), bank condition (eroded versus vegetated); and in-stream habitat (e.g., accumulated sediments versus gravel beds, submerged plants, and woody debris). We also identified species of fish and aquatic macroinvertebrates that were likely to be associated with each BCG level based on species records from actual stream and river sites in Minnesota where BCG levels were previously assessed (Gerritsen et al. 2017).

Next, we summarized this information in nontechnical language and worked with a graphic artist to develop visual representations of what rivers and riverbanks look like in our study region, corresponding to each of the six levels of the BCG (see Figure 2).[1] For each level, the upper panel provides a stylized visualization. The bottom panel provides a snapshot of biological diversity, with pictures of representative species that could be supported based on the referenced biological condition. Finally, the righthand border of the graphic displays four human use categories consistent with the traditional water quality ladder (Carson and Mitchell 1993), with the addition of a wading category to differentiate full and partial contact uses. We overlaid

---

[1] Photographs of actual locations provided the initial basis for production of these images. The complete survey including all graphics is available in the *SI Appendix*.

use categories onto BCG levels based on our best judgement[2], using information about BCG scores for real stream and river sites in Minnesota and their common real life uses, as well as changes in stream and river condition that would likely correspond to each BCG level and that might subsequently affect use (such as water clarity, the presence of excessive algae, or bacterial contamination; see *SI Appendix*, section S1 for more details). A red circle and slash through the use graphic means that use is not supported. The graphics and associated survey narrative corresponding to the six BCG levels define the water quality commodity.

To develop the spatial dimension of our commodity we then assembled data to accurately assign baseline BCG levels in all watersheds across our study region. BCG levels were based on macroinvertebrate community data collected by twelve state agencies for 19,277 sites across the study region (see *SI Appendix*, Table S1).[3] While both fish and macroinvertebrates are commonly used as indicators of stream biological condition in water quality assessment (US EPA 2013), macroinvertebrate data were more uniformly available across our entire study region. At the time of data acquisition, state agency personnel in four states (IL, IN, MN, OH) had developed BCG scoring criteria for streams and rivers according to the BCG framework outlined by EPA documentation (US EPA 2016). Although the remaining states did not have BCG criteria explicitly developed, they each had biological index scores on which stream condition was evaluated. To develop a high-resolution estimate of biological condition across the entire river basin, we converted the biological index scores used by states without a BCG to

---

[2] As far as we know, no studies have examined the empirical relationship between measures of biological condition and human use of water bodies. This is an area in need of further study.

[3] At the time of data acquisition, no biological condition data was available from the state of Pennsylvania. Data was also not collected from states that intersected only small parts of the study region, including Alabama, Georgia, Mississippi, New York, and South Dakota.

"BCG proxies" based on relationships between biological indices and the BCG previously documented by state agencies, together with narrative criteria used by states to classify streams that could roughly approximate the categories used by the BCG. See section S1 and Table S2 of the *SI Appendix* for additional details on this approach.

BCG scores were averaged across monitoring sites to create a score at the sub-watershed (defined as an 8-digit HUC) level. Finally, we designed color coded maps at different geographical scales to communicate spatial variation in baseline BCG scores across the study region. Figure 3 shows an example for a watershed (defined as a 4-digit HUC) in the eastern part of our study region.[4]

Extensive focus groups and classroom demonstrations were used to develop the final graphics, maps, and the valuation scenarios described in the next section. To prepare respondents for the scenarios, we first provided basic water quality information and incrementally introduced the three elements of the graphics. Maps as in Figure 3 provided the spatial distribution of baseline water quality in a respondent's local watershed, defined as the watershed of residence. The map also provided summary information about the average index score across the area (Figure 3), and respondents were asked to identify the water quality score in their home sub-watershed. During presentation of the graphics and maps, we asked questions to gauge respondents' understanding of the water quality metric and their ability to use the maps to identify water quality levels at points in space.

Our study area (see Figure 1) includes 31 watersheds (HUC4s) that are further divided into 268 sub-watersheds (HUC8s). Current water quality conditions largely consist of BCG

---

[4] In the remainder of this paper, we use the term 'watershed' to correspond to 4-digit hydrologic unit code (HUC4) areas and 'sub-watershed' to correspond to 8-digit HUC (HUC8) areas.

Levels 3 (defined in lay terms as “Some Changes Noticeable”) and 4 (“Many Changes Noticeable”), which constitute 42 percent and 49 percent of the study area, respectively. The remaining areas include 4 percent in Level 2 (“Close to Natural State”) and 5 percent in Level 5 (“Major Degradation”). See the *SI Appendix* (Figure S2) for a map showing the distribution of current BCG levels across the study region.

3) Experiment Design

The BCG is a physical concept that assigns more naturally functioning ecosystems lower numerical scores. The extent to which people prefer more naturally functioning ecosystems is an empirical question which our valuation exercise is designed to estimate. The survey contained six to ten valuation scenarios, details of which varied across respondents, designed to estimate the willingness to pay (WTP) of households for BCG level improvements. To interpret responses to our valuation scenarios as indications of real economic tradeoffs, we followed survey best practices for motivating truthful responses to valuation questions. Informed by the theoretical literature on incentive compatible elicitation in surveys (Carson and Groves 2007; Vossler et al. 2012), we frame the value scenarios as advisory referenda, use a coercive payment mechanism, and ask respondents to treat each referendum independently. We further stress the consequentiality of the survey to participants by informing them that the study is funded by the government, and that the results may be used to inform public policy.

Each scenario is defined by the following attributes: (a) the spatial scale of the policy area; (b) the extent and spatial distribution of the BCG change; (c) whether the policy area included the home watershed; and (d) an increase in household taxes if the policy were implemented. Table 1 shows the range of attribute values we used to define specific valuation scenarios. Household cost was presented as an unavoidable tax increase that would be assessed if the referendum passed. Tax amounts were randomly assigned from the amounts in Table 1 and presented as annual for five years.[5]

We use experimental variation to identify how economic welfare changes with the spatial scale (i.e., size) of the affected area. Our survey presented scenarios in which the water quality improvement was for a single watershed, three contiguous watersheds, and the full study region. To create the middle category, we divided our study area into 10 mutually exclusive, contiguous groupings of three watersheds.[6] To identify the effects of improving water quality, as measured by the BCG scores, we included four different BCG change scenarios in the design (Table 1). For instance, one change scenario is to improve all sub-watersheds within an impacted area to a Level 2. These change scenarios, along with substantial variation in actual (current) conditions, allows the identification of water quality improvements.

Finally, by presenting scenarios that both include, and do not include, the respondent's home watershed, we are able to differentiate economic values for near home versus distant improvements in surface water quality. To facilitate this, we solicited the respondent's zip code at the beginning of the survey, which was then matched to their sub-watershed. Not only did this allow us to create scenarios specific to where the respondent lives, we were able to provide local water quality conditions at the sub-watershed level as an additional "attribute" in the scenario design.

---

[5] Table entries reflect tax amounts in effect for 1875 of the 2000 respondents. We adjusted some of the tax amounts after a soft launch of the survey, which suggested that WTP values were higher relative to the MTurk pilot samples. This is perhaps unsurprising, given that MTurk respondents had lower incomes than the general population. As a result, we included some higher tax amounts, and revised or eliminated some of the lower tax amounts, to better identify the WTP distributions for the various policy scenarios.

[6] In one case, the grouping is four watersheds since our study area consists of 31 watersheds in total.

A valuation scenario consisted of a map showing the policy area and quality improvements (see Figure 4), a table summarizing the area-wide average quality change, the change (if any) to the local sub-watershed, the size of the policy area, household cost, and a vote solicitation framed as a public referendum. Our scenario maps display BCG levels at the sub-watershed level (see Figures 3 and 4) and variation in baseline water quality levels was provided by differences in actual conditions across our study region. Importantly, the BCG changes listed in Table 1 are therefore relative to different baseline conditions. Additional details on how the scenarios were presented are included in the *SI Appendix*, section S2.

We coded the survey using the Qualtrics survey design platform and set it up to be completed by an online panel. The experimental design and survey functionality were tested using an online convenience sample obtained through Amazon's Mechanical Turk (MTurk). Two pilots focused on the state of Illinois. Once we were confident that the mechanics of the survey were working properly, and that the materials and questions were well understood, we piloted the survey a third time with respondents from nine states within the study region to confirm the full survey functionality and to obtain preliminary results for informing the distribution of tax changes to use in the final survey.

4) Data Collection and Results

A sample of 2,000 people residing in our study region, as verified by zip codes, completed the survey experiment between October 15 and November 16, 2021.[7] This sample size was informed by a power analysis using the third MTurk pilot sample, which suggested that

---

[7] Respondents were removed from the sample if they (1) completed the survey in less than 10 minutes and (2) answered more than one of the four questions of understanding incorrectly. The sample size of 2000 is exclusive of these individuals.

2,000 respondents was sufficient to detect a true effect size of $25 with at least 80 percent power when comparing the effect of a one-unit improvement in the BCG score across any two spatial scales (holding location fixed) or testing the difference in WTP between local and non-local policies (holding spatial scale fixed). The surveys were collected by Qualtrics in partnership with NORC at the University of Chicago, using NORC's online probability based AmeriSpeak Panel.[8] Panel members are recruited rather than volunteer or opt-in to the panel, which increases response rates and sample representativeness, and circumvents issues with fraudulent responses (e.g., due to ineligible participants, click farms, and bots). Sample summary statistics are included in the *SI Appendix*, Table S6.

The survey design was informed by economic theory with the goal of providing measures of economic welfare that reflect the true preferences of the target population. We included several questions to help understand whether we were successful. Questions designed to gauge beliefs tied to the sufficiency conditions for incentive compatible elicitation showed that 82 percent voted as if their household would face the stated policy costs, 80 percent voted as if the policies would achieve the stated improvements in water quality, and 76 percent voted as if the data collected will be used to inform policy makers.[9] We also asked how the attributes in our experimental design influenced votes, and the overwhelming majority indicated they were influenced by the size of the area impacted by the policy (75 percent), the improvement in water quality levels (93 percent), and the cost of the policy (88 percent).[10]

---

[8] See https://amerispeak.norc.org/about-amerispeak/Pages/Overview.aspx for additional details on the panel.

[9] The response options were "Disagree", "Neutral", and "Agree", and the percentages reported reflect those who selected the latter option.

[10] The response options were "Little to no effect", "Moderate effect", and "Large effect". The indicated percentages coincide with the latter two options.

The valuation scenario data were analyzed using mixed logit models for repeated choices (Revelt and Train 1998). Model 1 includes the full survey sample and allows WTP to vary according to the BCG score (identified by the variation in the "Change in BCG" attribute along with variation in baseline conditions), spatial scale, and location. All model parameters, except for the parameter associated with the cost of the policy, follow normal distributions. Estimation was carried out using maximum simulated likelihood, using 500 Halton draws. Additional details on the estimation methods, model specification, and parameter estimates are documented in the *SI Appendix*, section S3.

Table 2 presents selected WTP measures for changes in the BCG score and the spatial scale of the water quality improvement. These estimates reflect what the average household is willing to pay per year, over a period of five years, for the improvement. To arrive at these figures, we first calculated WTP for each respondent, considering characteristics of the policy specific to where they live, and then averaged these values over the sample. The delta method is used to compute standard errors.

We find that the WTP for a one BCG level improvement in water quality in the respondent's sub-watershed (HUC8) is \$152.[11] This figure approximately doubles to \$316 and is statistically different (p<0.01) if the affected geographic area includes the respondent's entire local watershed (HUC4). However, further increases in spatial scale to the group of three watersheds and study region levels generate statistically insignificant differences in WTP, relative to the single local watershed level. This provides evidence that the spatial scale of *local*

---

[11] We did not have respondents vote on policies that would only impact their sub-watershed; however, for local policy scenarios, the BCG score of sub-watersheds was included as an attribute in the experimental design. The variation in the policy scenarios along with variation based on where people live allow for identification of WTP for a water quality change at the sub-watershed level.

economic values for water quality improvements does not reach beyond the watershed level in these data.[12] This spatial scale finding may be explained by two non-mutually exclusive factors: a diminishing marginal WTP for an increase in spatial scale and the increasing distance of improved areas from the respondent's home.

A similar pattern regarding the spatial scale of local values emerges for the two "Minimum Level" scenarios. To interpret and compare the point estimates for these two scenarios we note that the change a household experiences is conditional on its local baseline conditions. Only 4 percent of the study region has baseline Level 2 water quality and 91 percent of the study region has baseline Level 3 or 4. The point estimates for the "Minimum Level 2" local scenario therefore mainly reflect household values for one- or two-level changes in the BCG. For this relatively large change in water quality, households are willing to pay on average nearly $500 for a policy in their local watershed. Heterogeneity in local values across the study region shown in Figure 5 is substantial. The map shows the distribution of WTP by local households for improving their local (HUC4) watershed.[13] The range of values is $164 to $810. On the map, the darker colors correspond to higher economic values, and interestingly WTP appears to positively correspond to watersheds in our study region with lower baseline water quality (see *SI Appendix*, Figure S2 for a map of baseline water quality levels).

The point estimates for the "Minimum Level 3" local scenario reflect a more modest change in water quality. Forty-two percent of the study region has baseline Level 4, and only

---

[12] In the *SI Appendix*, section S3 we discuss how our econometric specification accommodates spatial scale and contrast our specification with alternative approaches based on linear distance.

[13] To clarify, for each zip code we calculate values associated with a policy that would only improve water quality throughout the associated local watershed. These estimates therefore ignore values that would accrue to those outside an improvement area. Further, these values are distinct from those associated with a scenario where all sub-watersheds across the entire study region improved to Level 2.

five percent live in areas with even worse water quality. For this scenario less than half the landscape receives an improvement and the sample averages are correspondingly smaller. For example, we find a household average WTP of $217 for a policy that secures a minimum BCG Level 3 for a respondent's local watershed. This is statistically different ($p<0.01$) from the larger estimate of $492 for a policy providing a minimum BCG Level 2.

The righthand columns of Table 2 provide WTP estimates for *non-local* water quality scenarios. The value of a one BCG level improvement in a non-local watershed is $165, suggesting households are willing to pay only half as much ($316 versus $165) for an improvement that does not include their home watershed. Similar patterns emerge for the "Minimum Level" scenarios, and we once again see only modest spatial scale effects when moving from a single to a group of three non-local watersheds.[14]

Table 3 provides additional WTP estimates that allow us to explore in more detail the local/non-local differences. The estimates are derived from Model 2, which utilizes a subset of data for local and non-local voting scenarios in which a single watershed was used as the spatial scale. The specification allows WTP to vary based on the percentage of the policy area located in the respondent's home state. In this sample the average percentage of the policy area that is in-state is 62 percent and 4 percent, respectively, for local and non-local scenarios. Estimates from the model (see *SI Appendix*, Table S5) show that WTP increases by $1.09 per in-state percentage point for a local scenario and by $2.79 per in-state percentage point for a non-local

---

[14] We emphasize that valid local/non-local comparisons in Table 2 require that we compare either the two "watershed" or two "3 watershed" columns directly. Specifically, it is not appropriate to interpret the null difference in WTPs between the local watershed and local 3 watershed scenarios as contradicting the positive WTP for the non-local watershed scenarios. This is because the reference points for payment are different. For the local scenario, people are not willing to pay more for a larger policy area *beyond* what they would already pay for the local watershed area. For the non-local HUC4 scenario no such reference condition applies.

scenario. As shown in the table, these marginal effects yield economically meaningful differences for even modest changes in the percentage of the impacted area located in-state.[15] Figure 6 illustrates this spatial heterogeneity using the example of a single watershed (displayed with a white border) spanning parts of Illinois, Indiana, and Kentucky that improves to a Level 2. The map displays the mean household level WTP for a portion of our study region, derived from zip code level variation in local/non-local impact and the percentage of the impact zone in the home state. Estimates range from $295 for out of state and non-locally affected households to over $600 for largely in-state, locally impacted households.

5) Discussion

We draw four conclusions about the structure of preferences for surface water quality as they relate to BCG levels. First, local improvements – defined here to mean that the respondent's watershed of residence (HUC4) is included in the impacted policy area – are valued approximately twice as much as non-local improvements. Second, the spatial scale of local benefits from an improvement in biological condition does not extend beyond the watershed level. Third, our results suggest an important role for non-use values in respondent preferences. Fourth, the estimated spatial scale of benefits for policy scenarios that do and do not include the home area suggests that values for water quality improvements are locally concentrated.

Each of these has important consequences for understanding the economic benefits of

---

[15] Of households voting on a non-local watershed scenario, 13 percent of the watersheds were partially located in-state. For those voting on non-local policies involving a group of three watersheds (scenarios not used in Model 2), this number nearly doubles to 26 percent. This statistic, along with respondents' willingness to pay more for in-state policies, provides one explanation for why we see a slight increase in WTP in Table 2 when we increase the spatial scope of the non-local policy.

policies aimed at generating water quality improvements in the landscape. Consider, for example, nonuse value. Seventy-one percent of respondents in our sample engage in water-based recreation activities in a typical year. For those that recreate, most respondents (72 percent) report that their *furthest* recreation trip destination is within 150 miles of their home. Of the non-local policy scenarios included in the survey, over 90 percent were further than 150 miles of the respondent's zip code. These figures imply that at least half of the non-local WTP estimates is attributable to nonuse. Similarly, while three quarters of recreation behavior occurs in the local watershed, there is little we can say about the relative magnitude of use and nonuse values in comprising the total value in local watersheds, as our study design was intended to estimate total value only.

Three previous studies provide useful context for our findings. Meyer (2013) and Parthum and Ando (2020) elicit water quality values within our study region, although the smaller spatial scale and types of water quality improvements make their estimates difficult to compare with ours. However, they both find clear evidence of willingness to pay to improve water quality in local rivers and streams. Parthum and Ando estimate an annual WTP of $62 to $85 per household (paid indefinitely) for meeting nutrient reduction goals in the Upper Sangamon River Basin, a small watershed in Central Illinois.[16] An estimate reported by Meyer suggests that the average household is willing to pay $89 per year (over five years) to achieve swimmable conditions throughout the Minnesota River Basin.[17]

---

[16] These numbers are the reported "Mean benefits per household (dollars)" for scenarios (3) and (4) in Table 3 of Parthum and Ando (2020).

[17] Meyer (2013, p. 53) reports an annual WTP of $8.86 for each 1 percent increase in the amount of the river basin that is clean enough to support all recreation activities, including swimming. Full (100 percent) clean up implies a WTP of $8.86×100.

Parthum and Ando (2020) and Meyer (2013) are examples of valuation studies that focus on local-scale resources and achieve internal validity by using of a high degree of local specificity in their experimental design. This is consistent with best practice in SP research, which emphasizes scenario realism and salience. A cost of this local focus, however, is that the results cannot be scaled up to inform policy changes across the wider landscape. For this reason, Carson and Mitchell (1993) conducted a large-scale study to estimate the benefits of national water quality improvements. The authors use a representative sample of US households to estimate the value of maintaining boatable water quality nationwide, as well as improving water quality everywhere to meet "fishable" and then "swimmable" standards. These estimates have served as the cornerstone of many federal regulatory analyses (Griffiths et al. 2012).

Carson and Mitchell's best estimate of achieving swimmable water for all water resources in the nation is $148 per household per year (1990 dollars).[18] Adjusting for inflation and the difference in the number of annual payments yields an estimate of $542 per household per year over five years.[19] This estimate is similar to, but somewhat higher than, our estimate of $463 per year to achieve BCG Level 2 ("swimmable") for our study region (Table 2). Carson and Mitchell find that households are willing to allocate approximately 67 percent of their WTP to within-state improvements and 33 percent to out-of-state improvements in water quality. This is consistent with our finding that households have higher values for in-state improvements.

---

[18] This number is obtained by adding the WTP values to go from "boatable" to "fishable" ($70) and then from "fishable" to "swimmable" ($78) in Carson and Mitchell (1993, Table 3).

[19] Using the CPI inflation calculator (https://www.bls.gov/data/inflation_calculator.htm), $148 in 1990 is equivalent to $304 in 2021. The valuation scenario in Carson and Mitchell proposed a perpetual annual tax whereas our scenario limited payments to a five-year period. We assumed payments end in the Carson and Mitchell scenario after 10 years. Using a 5 percent annual discount rate, $304 paid annually over 10 years is equivalent to paying $542 per year over 5 years.

While the similarities in numbers are interesting, we stress that differences in study design, implementation, and samples imply that our estimates and theirs are not directly comparable. Carson and Mitchell's (1993) survey was fielded in the early 1980s, and much has changed since this time, including baseline water quality levels and survey methodologies. Perhaps most importantly, our results reflect values for residents of the central US whereas their study is nationwide. Still, the comparison allows us to emphasize the utility of large-scale studies such as ours and Carson and Mitchell for analyzing CWA regulations.

Indeed, the magnitude of our estimates suggests that water quality improvements would generate large economic benefits for households in our study region. As a first approximation of this latent value consider the impact of achieving BCG Level 2 ("Close to Natural State") across the Upper Mississippi, Ohio, and Tennessee River Basins. There are approximately 22.6 million households in the counties that lie fully or partially in our study region. Table 2 shows that the average household in our representative sample is willing to pay $463 per year for five years to secure a BCG Level 2 across the full study region. Based on this point estimate, we predict that such a policy would generate over $10.5 billion in economic benefits for our study population annually for five years.

This estimate is derived from an approach that combines the validity advantages of a solicitation technique that displays a high degree of local specificity with a valuation concept that is grounded in ecological principles, transferable across space, and scalable to the national level. In this regard, reducing the tradeoffs between local realism and salience, and relevance for national policy, is the primary innovation in our BCG valuation approach.

We are pursuing additional work that incorporates the BCG. This includes statistical modeling to link policy-based changes in criterion pollutants such as nutrient concentrations to changes in the BCG. Linking criterion pollutants and BCG scores is part of a larger effort to construct an Integrated Assessment Model (IAM) for measuring economic benefits of place-specific policies to reduce nitrogen and phosphorous pollution. In addition, a natural next step is to extend our work to a nationwide valuation study using the BCG framework. In much the same way as the climate community benefits from different models for estimating the social costs of carbon, a national BCG-based study would provide complementary estimates for use in water policy regulatory analyses and other settings. Since we began our project, the BCG has been implemented in a growing number of states including states in the Mid-Atlantic region (Jessup et al. 2019) the Southwest (Jessup and Bradley 2020) and California (Paul et al. 2020), as well as for additional taxonomic groups including diatoms (Charles et al. 2019) and new ecosystems including coral reefs (Santavy et al. 2022). To date, most efforts to develop BCGs have occurred on a state-by-state or regional basis. Ideally, investment by state, tribal and federal agency partners could result in the development of a BCG operable at the continental scale. Such an effort could provide critical information as regulators take further actions at the state, federal, and local levels to achieve the goals of the Clean Water Act and the restoration of the chemical, physical and biological integrity of the Nation's waterbodies.

Data Availability

Anonymized survey data and code for replicating the econometric analysis presented in the paper, and a representative version of the stated preference survey, are available for download at the Harvard Dataverse, https://doi.org/10.7910/DVN/C5XEBF.

References

1. U.S. Environmental Protection Agency, *National Water Quality Inventory: Report to Congress*. EPA-841-R-16-011, U.S. Environmental Protection Agency, Washington, D.C. (2017). https://www.epa.gov/sites/default/files/2017-12/documents/305brtc_finalowow_08302017.pdf
2. P. Gilbert, From hogs to HABs: Impacts of industrial farming in the US on nitrogen and phosphorus and greenhouse gas pollution. *Biogeochemistry* **150**, 139-180 (2020).
3. E. G. Stets et al., Landscape drivers of dynamic change in water quality of U.S. rivers. *Environ. Sci. Technol.* **54**, 4336–4343 (2020). doi:10.1021/acs.est.9b05344
4. C. Moore et al., Measuring the social benefits of water quality improvements to support regulatory objectives: Progress and future directions. *Proc. Natl. Acad. Sci. U.S.A.* (this volume)
5. A. M. Freeman III, *Air and Water Pollution Control: A Benefit-Cost Assessment*. John Wiley & Sons, Inc.: New York (1982).
6. R. T. Carson, R. C. Mitchell, The value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resour. Res.* **29**, 2445-2454.
7. M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems. *Nature* **413**: 591-596 (2001).
8. B. Villeneuve, Y. Souchon, P. Usseglio-Polatera, M. Ferréol, L. Valette. Can we predict biological condition of stream ecosystems? A multi-stressors approach linking three biological indices to physico-chemistry, hydromorphology and land use. *Ecol. Indic.* **48**, 88-98 (2015).
9. P. C. Jacobson, G. J. A. Hansen, B. J. Bethke, T. K. Cross, Disentangling the effects of a century of eutrophication and climate warming on freshwater lake fish assemblages. *PLoS ONE* **12**, e0182667 (2017).
10. A. M. Michalak et al., Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6448-6452 (2013).
11. M. J. Pennino, J. E. Compton, S. G. Leibowitz, S.G., Trends in drinking water nitrate violations across the United States. *Environ. Sci. Technol.* **51**, 13450-13460 (2017).
12. C. L. Kling, D. J. Phaneuf, J. Zhao, From Exxon to BP: Has some number become better than no number?. *J. Econ. Perspect.* **26**, 3-26 (2012).
13. R. J. Johnston et al., Contemporary guidance for stated preference studies. *J. Assoc. Environ. Resour. Econ.* **4**, 319-405 (2017).
14. R. T. Carson, T. Groves, Incentive and informational properties of preference questions. *Environ. Resour. Econ.* **37**, 181-210 (2007).
15. C. A. Vossler, M. Doyon, D. Rondeau, Truth in consequentiality: Theory and field evidence on discrete choice experiments. *Am. Econ. J. Microecon.* **4**, 145-171 (2012).
16. R. T. Carson, T. Groves, J. A. List, Consequentiality: A theoretical and experimental exploration of a single binary choice. *J. Assoc. Environ. Resour. Econ.* **1**, 171-207 (2014).
17. I.J. Bateman *et al*., Making benefit transfers work: deriving and testing principles for value transfers for similar and dissimilar siates using a case study of the non-market benefits of water quality improvements across Europe. *Environ. Resour. Econ.* **50**, 365-387.
18. R. C. Bishop *et al*., Putting a value on injuries to natural assets: The BP oil spill. *Science* **356**, 253-254 (2017).

19. B. L. Keeler *et al*., Linking water quality and well-being for improved assessment and valuation of ecosystem services. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18619-18624 (2012).
20. J. C. Bergstrom, J. B. Loomis, Economic valuation of river restoration: An analysis of the valuation literature and its uses in decision-making. *Water Resour. Econ.* **17**, 9-19 (2017).
21. S. P. Davies, S. K. Jackson, The biological condition gradient: A descriptive model for interpreting change in aquatic ecosystems. *Ecol. Appl.* **16**, 1251-1266 (2006).
22. U.S. Environmental Protection Agency, *A Practitioner's Guide to the Biological Condition Gradient: A Framework to Describe Incremental Change in Aquatic Ecosystems*. EPA-842-R-16-001, U.S. Environmental Protection Agency, Washington, D.C. (2016). phttps://www.epa.gov/sites/default/files/2016-02/documents/bcg-practioners-guide-report.pdf
23. P. A. Chambers *et al*., Development of environmental thresholds for nitrogen and phosphorus in streams. *J. Environ. Qual.* **41**, 7-20 (2012). doi:10.2134/jeq2010.0273
24. I. Maddock, The importance of physical habitat assessment for evaluating river health. *Freshw. Biol.* **41**, 373-391 (2001).
25. J. Gerritsen, W. B. Bouchard, L. Zheng, E. W. Leppo, C. O. Yoder, Calibration of the biological gradient in Minnesota streams: a quantitative expert-based decision system. *Freshw. Sci.* **36**, 427-451 (2017). **Error! Hyperlink reference not valid.**
26. U.S. Environmental Protection Agency, *Biological Assessment Program Review: Assessing Level of Technical Rigor to Support Water Quality Management*. EPA 820-R-13-001, U.S. Environmental Protection Agency, Washington, D.C. (2013). https://www.epa.gov/wqc/biological-assessment-program-review-assessing-level-technical-rigor-support-water-quality
27. D. Revelt, K. Train, Mixed logit with repeated choices: Households' choices of appliance efficiency level. *Rev. Econ. Stat.* **80**, 647-657 (1998).
28. B. Parthum, A. W. Ando, Overlooked benefits of nutrient reductions in the Mississippi River Basin. *Land Econ.* **96**(4), 589-607 (2020).
29. A. Meyer, Intertemporal Valuation of River Restoration. *Environ. Resour. Econ.* **54**, 41-61 (2013).
30. C. Griffiths *et al*., U.S. Environmental Protection Agency valuation of surface water quality improvements. *Rev. Environ. Econ. Policy* **6**, 130–146 (2012).
31. B. Jessup, J. Stamp, M. Paul, E. Leppo, Biological Condition Gradient (BCG) attribute assignments for macroinvertebrates and fish in the Mid-Atlantic Region (Virginia, West Virginia, and Maryland). Virginia Department of Environmental Quality, Richmond, VA, 2019. www.deq.virginia.gov/home/showpublisheddocument/4303/637461491318800000
32. B. Jessup, P. Bradley. Calibration of Biological Condition Gradient models for fish and macroinvertebrates in sandy-bottom rivers in the Southwestern U.S. U.S. Environmental Protection Agency, Washington, D.C., 2020. https://www.env.nm.gov/surface-water-quality/wp-content/uploads/sites/25/2020/12/NMriverBCG_20201208_Final_un-compressed.pdf
33. M. J. Paul, B. Jessup, L.R. Brown, J. J. Carter, M. Cantonati, D. F. Charles, Characterizing benthic macroinvertebrate and algal biological condition gradient models for California wadeable streams, USA. *Ecol. Indic.* **117**, 106618 (2020).
34. D. F. Charles, A.P. Tuccillo, T. J. Belton, Use of diatoms for developing nutrient criteria for rivers and streams: A Biological Condition Gradient approach. *Ecol. Indic.* **96**, 258-269 (2019).

35. D. L. Santavy *et al*., A biological gradient for Caribbean coral reefs: Part II. Numeric rules using sessile benthic organisms. *Ecol. Indic.* **135**, 108576 (2022).

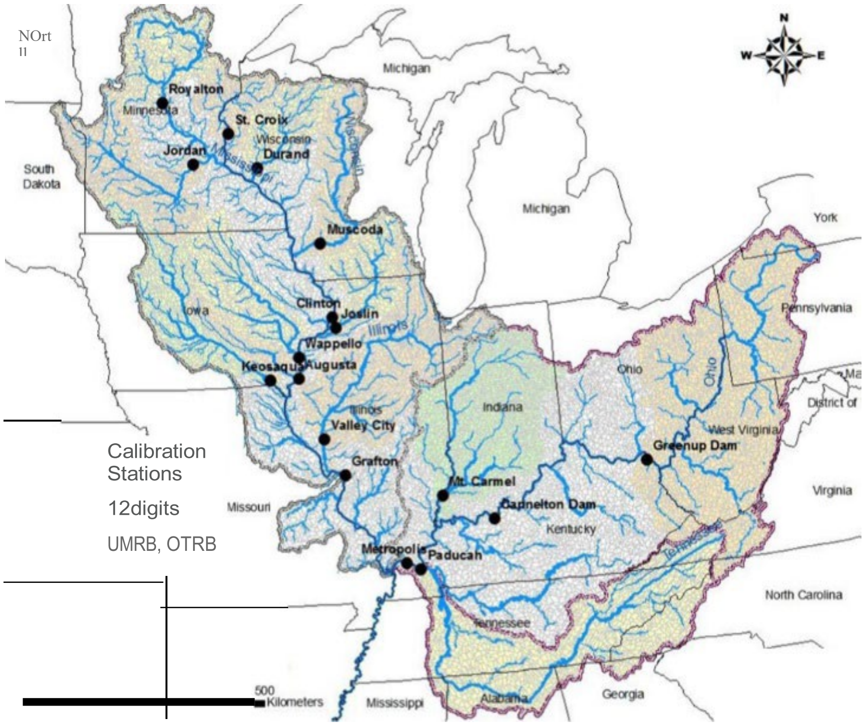
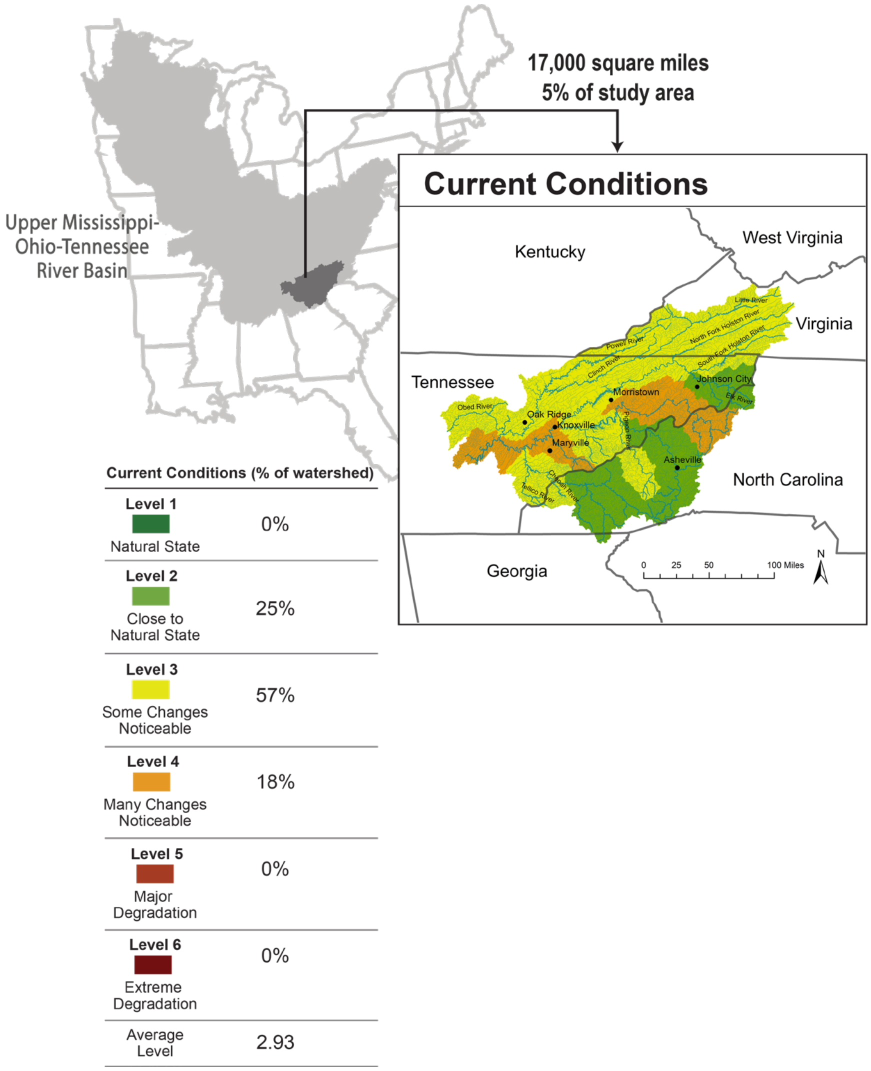


Figure 1: Upper Mississippi, Ohio, and Tennessee River Basins

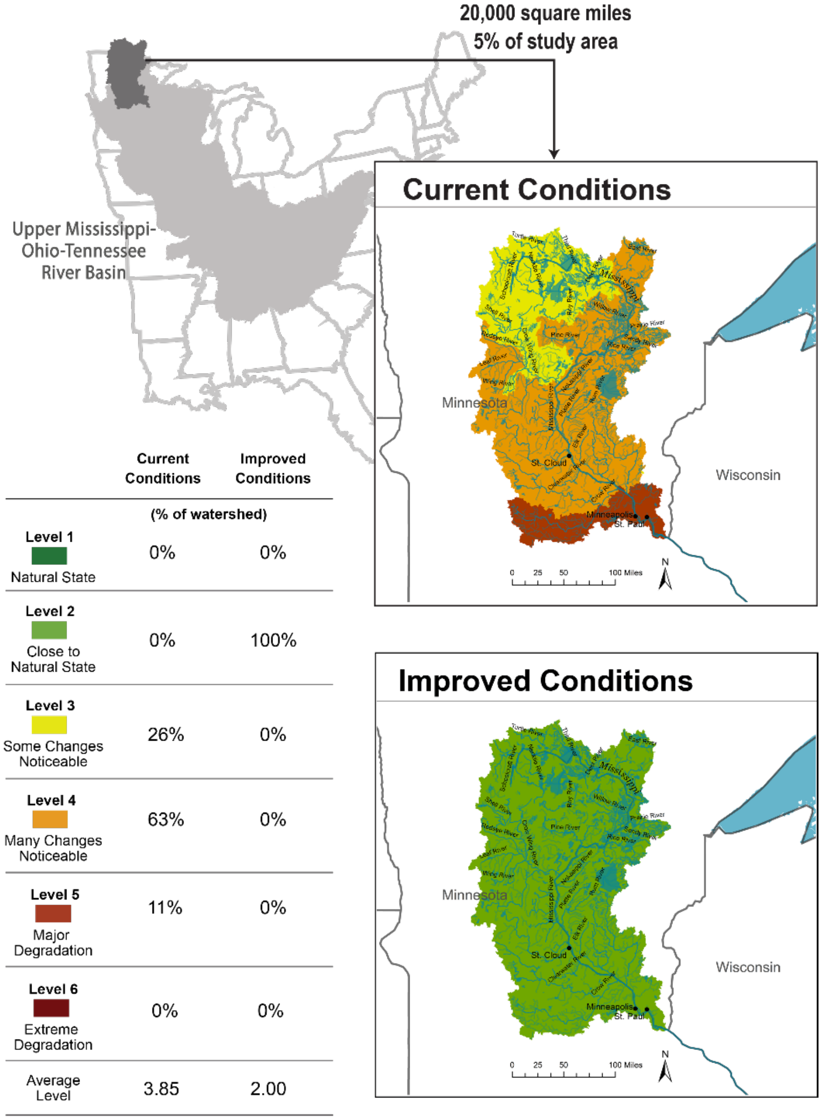
Figure 2. Graphics depicting 6 BCG levels, supported human uses, biodiversity, and visual conditions



Current Conditions (% of watershed)

Level 1 🟩 Natural State	0%
Level 2 🟩 Close to Natural State	25%
Level 3 🟨 Some Changes Noticeable	57%
Level 4 🟧 Many Changes Noticeable	18%
Level 5 🟥 Major Degradation	0%
Level 6 🟥 Extreme Degradation	0%
Average Level	2.93

Figure 3. Visual representation of variation in BCG levels within a 4-digit hydrological unit code (HUC4) watershed



	Current Conditions	Improved Conditions
	(% of watershed)	
Level 1 🟩 Natural State	0%	0%
Level 2 🟩 Close to Natural State	0%	100%
Level 3 🟨 Some Changes Noticeable	26%	0%
Level 4 🟧 Many Changes Noticeable	63%	0%
Level 5 🟥 Major Degradation	11%	0%
Level 6 🟥 Extreme Degradation	0%	0%
Average Level	3.85	2.00

Figure 4. Example water quality change scenario

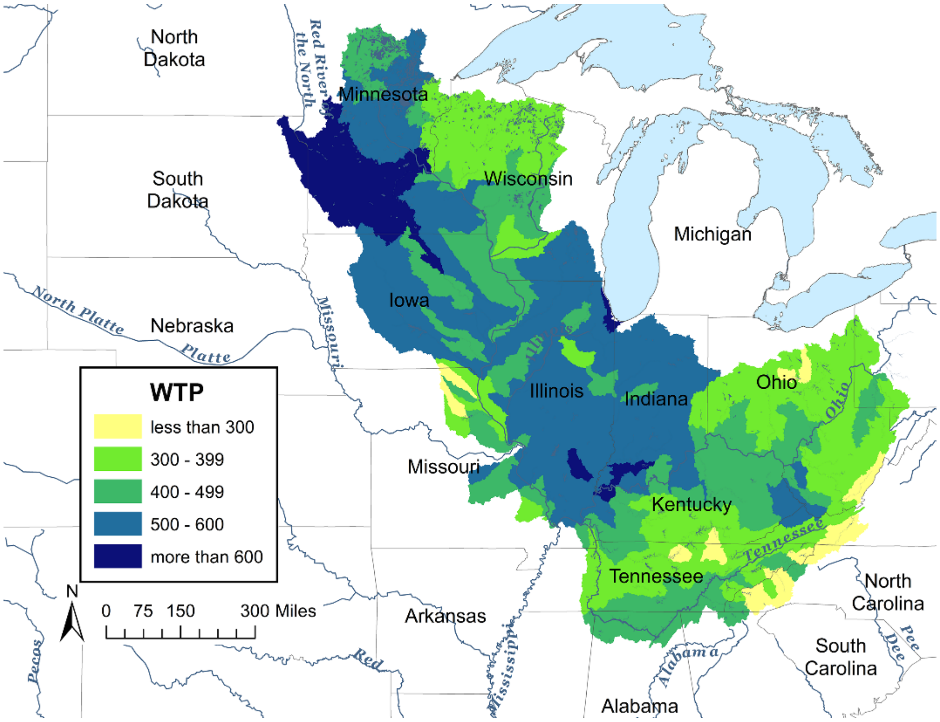


Figure 5. Spatial distribution of local willingness to pay for a minimum BCG Level 2 policy ($ per household in the affected watershed, annual payment for five years)

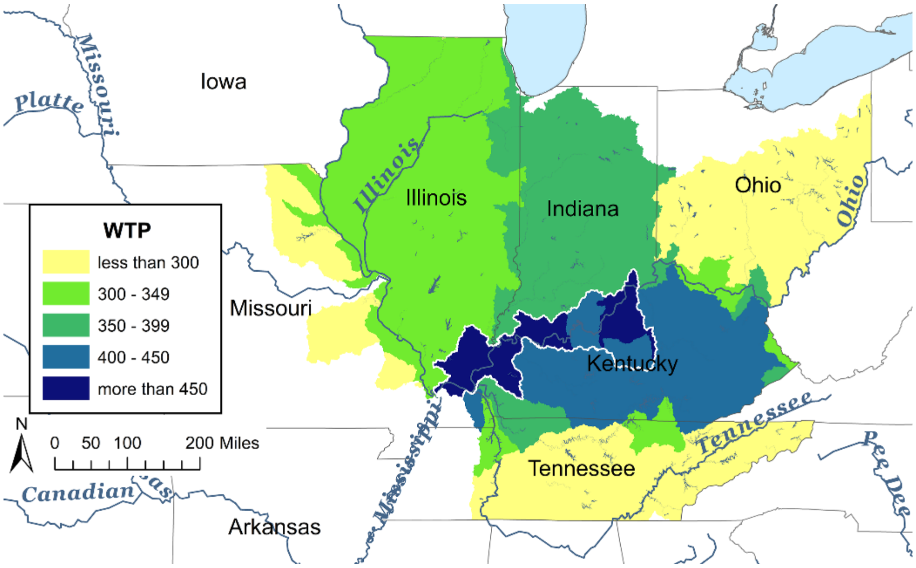


Figure 6. Spatial distribution of willingness to pay for BCG Level 2 in a single watershed (highlighted with a white border) ($ per household, annual payment for five years)

Table 1. Valuation scenario attribute levels

Attributes	Levels
Spatial scale	A single watershed Three contiguous watersheds Full study region
BCG change scenario	One-level BCG improvement in all sub-watersheds Minimum BCG Level 2 Minimum BCG Level 3 Change all BCG Level 3 sub-watersheds to Level 2
Location	Policy area includes home watershed (local) Policy area does not include home watershed (non-local)
Annual tax increase, in effect for five years	\$20, \$50, \$75, \$100, \$150, \$200, \$250, \$350, \$500, \$750

Notes: A watershed corresponds with a 4-digit hydrologic unit code address (HUC4), as defined by the US Geological Survey. The full study region includes the Upper Mississippi, Ohio, and Tennessee River Basins (see Figure 1).

Table 2. Willingness-to-pay for selected water quality improvement scenarios

	Local Changes				Non-Local Changes	
Scenario	Sub-Watershed (HUC8)	Watershed (HUC4)	3 Watersheds (3 HUC4s)	Study Region	Watershed (HUC4)	3 Watersheds (3 HUC4s)
One-level BCG improvement	\$152 (16)	\$316 (13)	\$302 (12)	\$300 (12)	\$165 (11)	\$186 (12)
Minimum BCG Level 2 (“swimmable”)	\$237 (24)	\$492 (21)	\$470 (19)	\$463 (18)	\$225 (15)	\$261 (18)
Minimum BCG Level 3 (“biological”)	\$119 (14)	\$217 (10)	\$209 (9)	\$207 (9)	\$95 (9)	\$112 (9)

Notes: Table entries indicate the mean household willingness-to-pay (in 2021 dollars), per year over a period of five years, for a policy defined by the water quality improvement and the spatial scale. Standard errors in parentheses. A ‘local’ policy is one that improves water quality in the watershed where the household lives and a ‘non-local’ policy does not include the household’s resident watershed. ‘Study Region’ refers to the Upper-Mississippi, Ohio, and Tennessee River Basins. Estimates are derived from Model 1, as described in the *SI Appendix*.

Table 3. Willingness-to-pay for water quality improvement scenarios based on percentage of impacted area located in-state

Scenario	Local policy: impact area 100% in-state	Local policy: impact area 25% in-state	Non-local policy: impact area 25% in-state	Non-local policy: impact area 0% in-state
One-level BCG improvement	\$356 (18)	\$274 (17)	\$228 (19)	\$159 (12)
Minimum BCG Level 2 (“swimmable”)	\$513 (23)	\$432 (27)	\$301 (22)	\$232 (16)
Minimum BCG Level 3 (“biological”)	\$268 (20)	\$187 (13)	\$142 (19)	\$72 (12)

Notes: Table entries indicate the mean household’s willingness-to-pay (in dollars), per year over a period of five years, for a policy defined by the water quality improvement and the spatial scale. Standard errors in parentheses. A “local” policy is one that improves water quality near the household’s residence, whereas a “non-local” policy does not. Estimates are derived from Model 2, as described in the *SI Appendix*.