4-2023

Spatial dimensions of water quality value in New England river networks

Robert J. Johnston

Klaus Moeltner

Seth Peery

Tom Ndebele

Zhenyu Yao

See next page for additional authors

Authors

Robert J. Johnston, Klaus Moeltner, Seth Peery, Tom Ndebele, Zhenyu Yao, Stefano Crema, Wilfred M. Wollheim, and Elena Besedin

PNAS RESEARCH ARTICLE | ECONOMIC SCIENCES SUSTAINABILITY SCIENCE

Spatial dimensions of water quality value in New England river networks

Robert J. Johnston[a,1], Klaus Moeltner[b], Seth Peery[c], Tom Ndebele[a], Zhenyu Yao[b], Stefano Crema[d], Wilfred M. Wollheim[e], and Elena Besedin[f]

Edited by Chris Moore, US Environmental Protection Agency; received January 19, 2022; accepted June 15, 2022 by Editorial Board Member David Zilberman

Households' willingness to pay (WTP) for water quality improvements—representing their economic value—depends on where improvements occur. Households often hold higher values for improvements close to their homes or iconic areas. Are there other areas where improvements might hold high value to individual households, do effects on WTP vary by type of improvement, and can these areas be identified even if they are not anticipated by researchers? To answer these questions, we integrated a water quality model and map-based, interactive choice experiment to estimate households' WTP for water quality improvements throughout a river network covering six New England states. The choice experiment was implemented using a push-to-web survey over a sample of New England households. Voting scenarios used to elicit WTP included interactive geographic information system (GIS) maps that illustrated three water quality measures at various zoom levels across the study domain. We captured data on how respondents maneuvered through these maps prior to answering the value-eliciting questions. Results show that WTP was influenced by regionwide quality improvements and improvements surrounding each respondent's home, as anticipated, but also by improvements in individualized locations identifiable via each respondent's map interactions. These spatial WTP variations only appear for low-quality rivers and are focused around particular areas of New England. The study shows that dynamic map interactions can convey salient information for WTP estimation and that predicting spatial WTP heterogeneity based primarily on home or iconic locations, as typically done, may overlook areas where water quality has high value.

choice experiment | map interaction | spatial | water quality | willingness to pay

Significance

Households' aggregated willingness to pay (WTP) can represent much of the total economic benefit of water quality improvements to society. WTP typically depends on what improvement types occur and where. However, the ways in which spatial dimensions of water quality influence WTP are often modeled primarily with respect to where people live or iconic locations. This study shows that allowing people to interact with water quality change maps during a choice experiment survey can reveal individualized areas wherein some types of water quality improvements have particularly high value, beyond effects related to home locations. Improvements in these areas are associated with significant increases in households' WTP. Overlooking these effects may cause researchers to undervalue some types of water quality change.

Estimates of economic value for water quality improvements are used to inform many policy decisions, for example, through regulatory cost-benefit analyses (1, 2). These values often include benefits realized by households that are quantified using estimates of households' total willingness to pay (WTP), produced using survey-based methods such as stated-preference (SP) choice experiments (3). Choice experiments ask survey respondents to choose among two or more hypothetical but realistic policy scenarios, similar to a public referendum (4). Each scenario describes a set of multiattribute environmental changes that would be obtained and a hypothetically binding monetary cost to the household required to implement the scenario. Data consisting of hypothetically binding choices or votes over many scenarios, by many respondents, allow WTP to be estimated using econometric methods.

These WTP estimates often depend on the spatial dimensions of improvements, e.g., the locations where changes occur relative to the people who value them (5, 6). For example, WTP is often found to decline as a function of the distance between each household's residence and improved bodies of water—called distance decay (7). Values also tend to increase as the size of improved areas increases and in iconic areas (6, 8–10). Empirical models are available to estimate these and other spatial dimensions of WTP (6). However, these models are limited by available data, and choice experiment data are typically produced using a survey architecture that constrains economists' ability to predict relationships between spatial dimensions of quality changes and WTP. Consequently, the ways in which spatial dimensions of water quality influence WTP are often modeled primarily with respect to the distance between water bodies and people's homes (11). Are there other areas where improvements might hold high value to households, and might these effects vary by type of improvement? Can these high-value areas be identified even if they are not anticipated by researchers?

To answer these questions, we integrated a water quality model and map-based, interactive choice experiment to estimate households' WTP for water quality improvements over ~95,800 miles of rivers and streams in Connecticut, Massachusetts, Rhode

Author contributions: R.J.J., K.M., S.P., W.M.W., and E.B. designed research; R.J.J., K.M., S.P., T.N., Z.Y., S.C., and W.M.W. performed research; K.M., S.P., T.N., Z.Y., and S.C. analyzed data; and R.J.J., K.M., S.P., T.N., W.M.W., and E.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. C.M. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: rjohnston@clarku.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2120255119/-/DCSupplemental.

Published April 24, 2023.

Island, Vermont, New Hampshire, and Maine (71,992 square miles of watershed area; map in *SI Appendix*). We estimated WTP for realistic, predicted scenarios of water quality change from a set of possible policy actions throughout this river system. The goal was to provide insight on the extent to which WTP is determined by the spatial distribution of these changes, focusing on household-specific, spatial determinants of value that are unmeasurable using standard approaches. We hypothesized that allowing people to interact with geographic information system (GIS) maps in a valuation survey might provide information that improves the accuracy of WTP prediction by identifying individualized areas where water quality improvements have high value to each respondent.

The online survey was implemented using an address-based, push-to-web sample, drawn randomly from households in the six states. Each choice experiment voting question paired a possible environmental policy scenario with a hypothetically binding household cost required to implement the scenario, compared to a business-as-usual (BAU) status quo with no change in household cost. Each scenario included a spatially explicit prediction of water quality change over the study domain, produced using the Framework for Aquatic Modeling of the Earth System (FrAMES), a process-based water quality model (12, 13). Scenarios illustrated current conditions and prospective changes in three different water quality measures, representing a) safety for human use, b) support for aquatic life (AL), and c) a multimetric indicator of overall water pollution. These measures were communicated in multiple ways, including interactive GIS maps that enabled each measure to be viewed at various zoom levels across the river system. The survey architecture captured data on how respondents maneuvered through each interactive map prior to answering choice experiment questions. Map-interaction data were used to infer where and at what scale water quality might be relevant to each respondent and provided evidence of increased attention to particular areas. The subsequent value-elicitation question was a single, hypothetically binding, binary vote between the BAU and the presented policy scenario, designed for incentive compatibility (14).

Using choice experiment data, a discrete-choice, random-utility model was estimated in WTP space using Bayesian model search (BMS), allowing parameters to be interpreted as dollar-denominated WTP estimates. The model predicts each respondent's vote and corresponding WTP measures as functions of explanatory variables derived from spatially explicit water quality measures in the choice scenario, map interactions, and each respondent's home location. Results show that WTP is influenced by regionwide changes in water quality measures and by the spatial distribution of these changes. With regard to the latter, WTP is influenced by the extent to which changes occur within a) 10 and 25 miles of each respondent's home and b) the geographic area given the longest attention by each respondent during their map interactions, irrespective of home location. These effects vary over different quality measures and are most pronounced for improvements to areas at low baseline quality.

Because the model relies on map-interaction data, it was estimated for the 76% of respondents who completed the full survey and interacted with maps. Supplemental models were estimated to evaluate robustness and test for WTP heterogeneity associated with respondents who interacted with maps (versus those who did not). These models suggest that the presented results are robust (*SI Appendix*).

Results show that the presented architecture for choice experiments can reveal individualized areas wherein water quality changes have high value. The approach allows these areas, and effects on WTP, to be identified even if the reasons why they are valued by each household are not anticipated by researchers. This information can help support more valid and reliable benefit estimation. For example, failure to account for spatial influences on households' WTP can lead to biases in the benefit estimates used to inform policy decisions (6, 9, 11, 15). Methods of this type may be particularly relevant when quality changes are heterogeneous across the landscape and values to each household depend on localized effects that may be unrelated to home locations or areas previously identified by researchers (11, 15). They can also be necessary to characterize benefits to groups of concern (e.g., underserved or environmental justice communities) that may value water quality in different areas or for different reasons than the general population. In cases such as these, provision of accurate benefit estimates requires methods able to identify areas where different types of water quality change are important to particular households and groups, sometimes for reasons that are unforeseen by researchers.

Methods and Results

Results are derived from 1,239 survey responses. The survey elicited choices for WTP estimation using policy scenarios that each reflected a possible set of water quality improvements. Each respondent received a questionnaire presenting one of these scenarios. To develop each scenario, FrAMES was applied to the study domain to characterize baseline conditions and possible changes in three different water quality measures relevant to the public. A water safety (WS) indicator for human use was developed based on fecal coliform guidelines of the US Environmental Protection Agency (EPA) (16). An indicator representing support for AL was developed based on effects of chloride concentrations on aquatic organism survival (17, 18). An indicator for total water pollution (WQ) was developed as a combined metric using all modeled solute concentrations. Each measure was normalized to a 0 (worst) to 100 (best) scale, based on reference conditions for the domain and biophysical thresholds for each indicator. We also binned each indicator into seven color-coded and labeled intervals.

Quality measures were predicted across the domain for a) contemporary conditions; b) a future BAU scenario as of 2025, with no management changes (the status quo); and c) 41 possible policy scenarios reflecting outcomes as of 2025 under a set of alternative policy changes affecting riparian buffers, wastewater treatment plants, stormwater retention, and road salt application. For each scenario, the survey communicated each quality measure in three different ways: a) a normalized, spatial mean value (0 to 100) over the river system; b) a bar chart showing the proportion of total river and stream miles within seven binned quality intervals; and c) high-resolution GIS maps showing water quality predictions over the river system. As an illustration, Fig. 1 shows the survey screen used to explain the WS indicator and characterize contemporary conditions. The *SI Appendix* provides details and shows analogous figures for WQ and AL.

To represent each measure under contemporary conditions, the BAU, and policy scenarios, the survey included nine maps, representing a) WQ contemporary conditions, b) WS contemporary conditions, c) AL contemporary conditions, d) WQ under BAU, e) WQ under the policy scenario, f) WS under BAU, g) WS under the policy scenario, h) AL under BAU, and i) AL under the policy scenario. All respondents were shown all maps in the survey at coarse spatial resolution (one map showing the study domain). All maps also had interactive GIS zoom capability, enabling respondents to click a URL to voluntarily view conditions for any

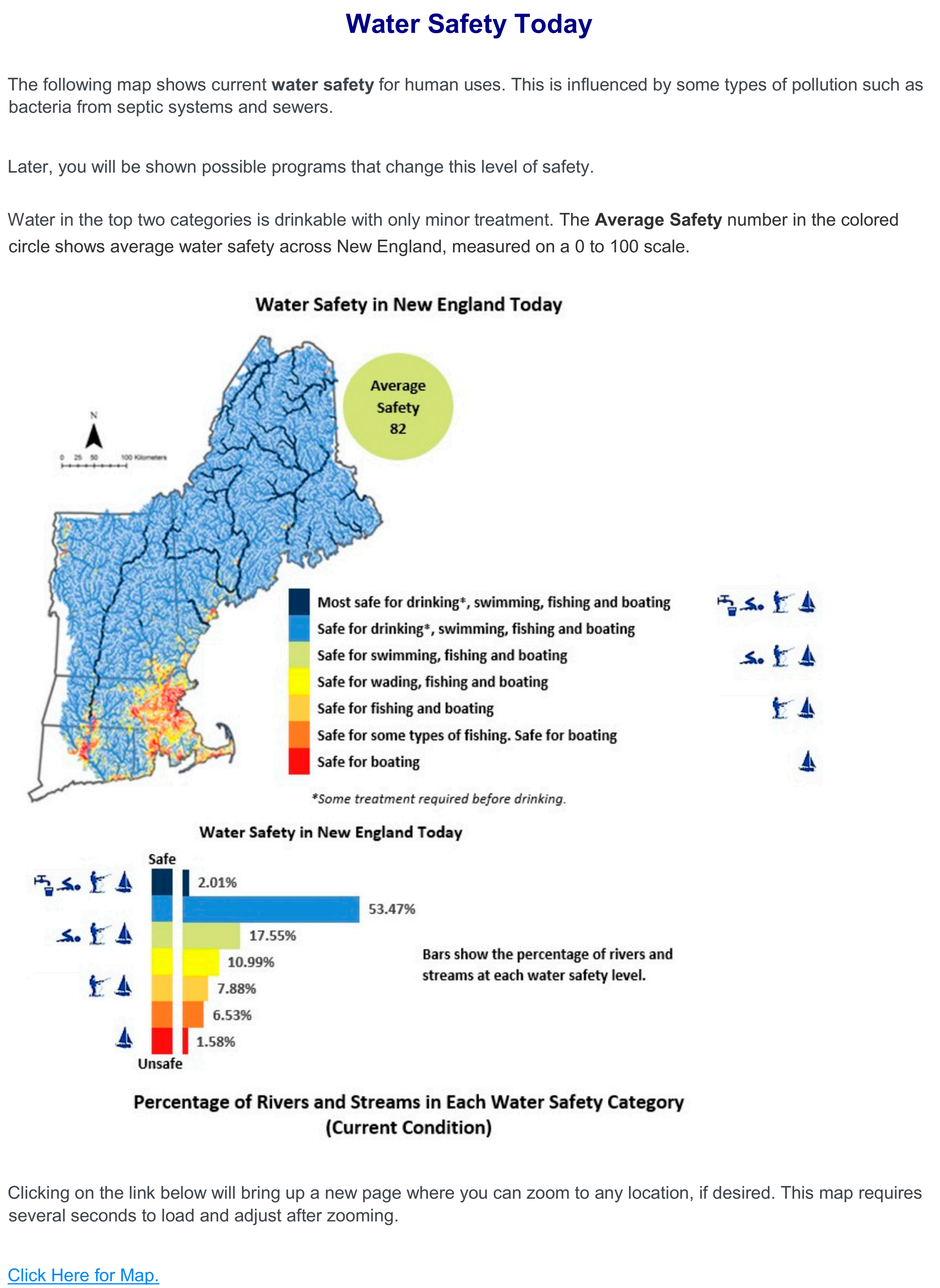

Fig. 1. Survey screen illustrating the WS indicator and showing contemporary conditions for this measure over the study domain. The number in the circle shows the spatial mean value across the domain. The bar chart shows the percentage of river and stream miles at each of seven binned quality levels. Clicking on the lower link on the page allowed the respondent to visit the map in an interactive GIS environment.

desired magnification and area. For example, a respondent could click on the URL for map e to view the WQ policy scenario map in ArcGIS Online, allowing them to pan and zoom. These interactive maps were included based on input from focus groups used to pretest the questionnaire. For example, focus group respondents emphasized that understanding conditions and changes in particular local areas was important when deciding whether to vote for the BAU or the policy scenario (*SI Appendix*).

We used the Esri JavaScript application programming interface (API) to record the zoom level, center point, and extent of the map view each time the respondent panned or zoomed. Captured map-interaction metrics included the time spent viewing each map frame (geospatial extent), with frames defined using geospatial coordinates. The resulting data provided information on all extents viewed by each respondent for each map and the length of time that each was viewed. Following prior studies that used survey engagement metrics to inform and specify choice models [e.g., eye tracking and survey response time (19, 20)], we applied the resulting map-interaction data to infer locations where water quality might be salient to each respondent, based on the time spent viewing each frame. These areas could be anywhere—close to respondents' homes or elsewhere. No prompts were provided to encourage respondents to look at any specific area for any reason, other than to explore water quality prior to answering survey questions.

Where Did Respondents Look? Map-Interaction Results. For each respondent who engaged with at least one of the interactive maps, we identified the three geographical extents (frames) viewed for the longest time in each map with which they interacted. This allowed us to distinguish areas viewed for long periods from those passed through or viewed briefly. This procedure identified at most 27 frames for each respondent (3 frames for each of nine maps). If a respondent did not interact with a map, no frames were identified for that map. We converted the combined set of longest-looked frames to a polygon feature class of overlapping rectangles that reconstructed the regions viewed. To enable visual interpretation, we converted this data to a rasterized representation of the number of polygon intersections per unit area, using a spatial join over a 10-km grid. This data-reduction process identified the combined number of times that each 10-km cell in our domain was viewed by all respondents combined, as part of a longest-looked frame.

When interpreting the data, it is important to realize that views at a large scale have different meaning than views at a small scale. Many views in our dataset represent the default viewing extent at which maps load initially (zoom level 6) or a recentered view at a scale too generalized to convey information beyond that in the static maps included on the main survey screens viewed by all respondents. To orient our analysis around informative scales, we therefore disaggregated analysis of longest-looked frames according to standardized zoom levels (grouped as levels 6 to 10, levels 11 to 13, and levels 14 to 17). Zoom levels are discrete, preset scales at which a map is prerendered on the screen and are used by most modern interactive mapping platforms—analogous to the magnification at which each map was viewed (*SI Appendix*). Levels 6 to 10 represent geographies between the country and city levels (e.g., states and metropolitan areas); levels 11 to 13 represent communities, neighborhoods, and roads; and levels 14 to 17 display smaller streets and structures. As anticipated, analysis of larger-scale views (zoom levels 6 to 10) displayed minimal spatial variation over the region; these views often covered the entire study area and thus did not demonstrate respondents' intentions to view particular areas. Hence, we orient the analysis around views at zoom levels 11 to 13 and 14 to 17 (areas that respondents zoomed in to see at higher resolution).

What areas received most attention? Results are shown in Fig. 2, which aggregates data for all maps. As a precursor to map-interaction results, Fig. 2*A* shows the spatial density of respondents' home addresses, aggregated over the 10-km grid. If respondents primarily viewed areas close to their homes, one would expect to see similar patterns in the map-interaction data. In contrast, evidence of map views in other areas might suggest that respondents viewed areas not immediately surrounding their homes. Fig. 2*B* shows the gridded representation of longest-looked frame density for zoom levels 11 to 13. Fig. 2*C* shows the same representation for levels 14 to 17. These maps may be interpreted as illustrating aggregate respondent interest in the area covered by each 10-km cell, as represented by the count of viewing extents intersecting each cell.

Results for zoom levels 11 to 13 show patterns of intraregional spatial variation. We find respondent interest in areas surrounding Burlington, Vermont; Portland, Maine; a large area covering much of southeastern New Hampshire; the greater Boston metropolitan area in Massachusetts; most of Rhode Island; and a portion of northeastern Connecticut. Although the map exhibits higher respondent interest in cells covering population centers, its pattern of variation is not fully

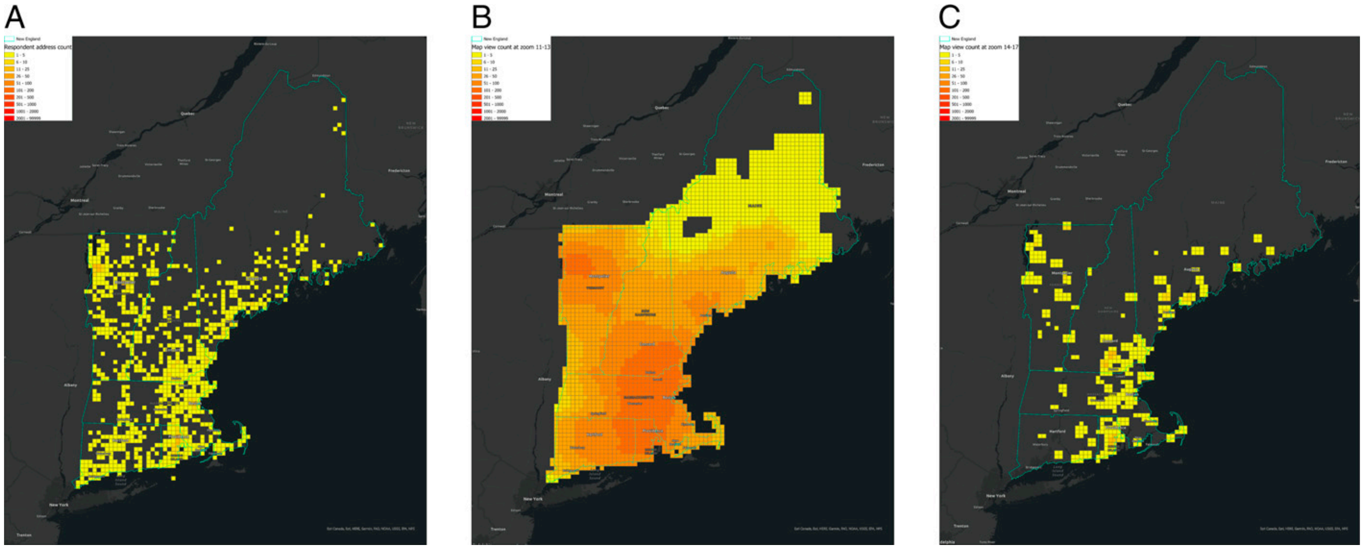

Fig. 2. (*A*) Spatial density of respondent home addresses aggregated to a 10-km grid. (*B*) Number of longest-looked frames aggregated over each 10-km grid cell for map frames viewed at zoom levels 11 to 13 by all respondents who interacted with maps. (*C*) Number of longest-looked frames aggregated over each 10-km grid cell for map frames viewed at zoom levels 14 to 17 by all respondents who interacted with maps. Legends for *B* and *C* indicate the number of longest-looked map views over the grid cell.

explained by the distribution of respondent addresses, as would be the case if respondents only zoomed in to view their own homes or communities. Data for all longest-looked map frames verify that only 56.8% of views at levels 11 to 13 included the respondent's home.

Similar patterns emerge for levels 14 to 17 (Fig. 2*C*). Views at these levels are only captured for respondents who viewed highly specific areas of a map. Viewing frequency at this level also increases for general areas near respondents' homes, but the relative density of map views and home addresses is not identical. Only 35.2% of frames viewed at levels 14 to 17 included the respondent's home. For example, there is a high density of views in southern New Hampshire and some areas of Maine (Fig. 2*C*) that is not matched by a corresponding density of home addresses over identical grids (Fig. 2*A*). Sensitivity analyses conducted using different subsets of the data and measures of map engagement corroborate these patterns (*SI Appendix*).

These data provide a visual representation of the frequency with which respondents viewed particular areas and suggest that respondents might care about water quality in areas that are not immediately centered around their homes. The following section evaluates whether these areas provide information that is relevant to understanding water quality values.

What Did Respondents Value? How WTP Is Affected by Spatial Dimensions. Using combined scenario, survey-response, and map-interaction data, a random-utility model was implemented using Bayesian logit. The model predicts respondents' choices for the policy scenario over the BAU as a function of explanatory variables on quality changes, spatial features of those changes, and household cost. Five categories of variables are included, based on predictions for WQ, WS, and AL. All reflect differences in conditions between the policy scenario presented to each respondent and the BAU.

The first two variable categories are nonspatial. The variables WQ_i, WS_i, and AL_i represent differences in the mean values of each indicator over the domain; these are akin to average regionwide changes, measured in percentage points (±0 to 100). The related nonspatial variables *WQbot3*, *WSbot3*, *ALbot3*, *WQtop3*, *WStop3*, and *ALtop3* quantify changes in the aggregate percentage of river miles (±0 to 100) within the highest three (top3) and lowest three (bot3) binned quality levels for each water quality measure over the domain. The middle category is omitted for identification. These variables provide a means to evaluate whether respondents value changes in the proportion of river length at high or low quality, apart from values for mean regionwide change.

The remaining variables characterize spatial dimensions of quality change. The variables *WQXXbot3*, *WSXXbot3*, *ALXXbot3*, *WQXXtop3*, *WSXXtop3*, and *ALXXtop3* (XX = {10,25}) quantify changes in the aggregate percentage of river miles (±0 to 100) within the top3 and bot3 binned levels for each quality measure within a 10- or 25-mile radius of each respondent's home address. For example, *WS10top3* quantifies the difference in the percentage of river miles within the top3 WS levels (Fig. 1), within a 10-mile radius of the respondent's home. These variables allow evaluation of whether respondents hold higher values for changes that surround their homes within various fixed distances. We measure these differences within circles around each respondent's geocoded home address, following standard approaches (21).

The final six variables (*F1_WQbot3*, *F1_WSbot3*, *F1_ALbot3*, *F1_WQtop3*, *F1_WStop3*, *F1_ALtop3*) are calculated using map-interaction data. To derive them, we first identified the geographical extent of the longest-looked map frame for each respondent. For this area, we calculated the change in the aggregate percentage of river miles (±0 to 100) within the top3 and bot3 quality levels for each measure. For example, the variable *F1_WStop3* quantifies the change in the percentage of river miles in the top3 combined WS levels within the geographical extent of the longest-looked map frame for each respondent. These variables were included as a means to test the hypothesis that quality changes in these potentially salient areas—identifiable via map interactions—exert an additional influence on WTP.

Since theory provides little guidance as to which of these variables have the strongest influence on choices, and therefore WTP predictions, we overlaid our logit estimation with a BMS. The BMS output produces model-averaged parameter estimates and predictions.

Results are shown in Table 1. The first numerical columns present main effect estimates for explanatory variables. For all variables but scale, these are interpreted as marginal WTP estimates (22). The table gives the posterior mean and standard deviation (SD) for these estimates, together with the proportion of posterior draws to the right of 0 ($P > 0$). This shows whether a regressor's effect is predominantly positive ($P > 0 \rightarrow 1$), negative ($P > 0 \rightarrow 0$), or ambiguous ($P > 0 \approx 0.5$). The p(in) column shows how often, in terms of share of total BMS sampler iterations, a variable was included in the model. The closer this value is to 1, the more important the variable is for model fit. We use joint information captured by $P > 0$ and p(in) to assess the strength of a coefficient's signal in terms of the variable's contribution to the model. These signals are represented by asterisks, with three asterisks implying that $\max\left(P > 0,\ 1 - (P > 0)\right)$ exceeds 0.95, two asterisks implying a threshold of 0.90, and one asterisk implying a threshold of 0.80 (in slight abuse of classical tradition). Signals are assigned conditional on inclusion probabilities > 0.5 (beating prior odds).

Results show that spatial dimensions influence WTP. We begin with variables that quantify change surrounding each respondent's home. In contrast to the regionwide (nonspatial) variables, which were clearly presented to all respondents as part of the survey narrative using numbers and bar charts, determining water conditions in one's surrounding area requires additional effort on the part of respondents (identifying the home location on the map and changes surrounding it). Hence, it is ex ante unclear whether these variables might influence WTP. Results suggest that they do. Significant WTP estimates emerge for changes within a 10- and 25-mile radius of respondents' homes, but only for improvements to lower-quality rivers. These effects are additive. For example, respondents are willing to pay $10.66 for each percentage point reduction (improvement) in river length in the bot3 WQ levels within 25 miles of their home (*WQ25bot3*). They are willing to pay an additional $9.85 for analogous improvements within 10 miles (*WQ10bot3*). These results suggest that a large proportion of the total WTP for quality improvements that occurs over the study region is due to changes close to respondents' homes. A corollary general conclusion is that water quality changes in closer proximity to population centers will produce higher aggregate WTP over all households, ceteris paribus.

Results also show that each household's WTP is affected by quality changes within their longest-looked map frame—individualized areas that are only identifiable via each respondent's map interactions. Respondents are willing to pay $7.11 for each percentage point reduction (improvement) in combined river length in the bot3 WQ levels for the area shown in their

Table 1. Bayesian WTP-space logit regression analysis of choice experiment responses

Explanatory variable	Variable definition and units	Parameter mean	Parameter SD	$P>0$	p(in)	Signal
Mean index differences, entire policy domain						
WQ_i	Diff. in index points, policy domain	12.221	10.760	0.741	0.783	
WS_i	Diff. in index points, policy domain	27.823	8.596	0.995	0.996	***
AL_i	Diff. in index points, policy domain	6.310	9.215	0.529	0.625	
Top and bottom level differences, entire policy domain						
WQtop3	Diff. in % of river mi. in top 3 WQ levels	6.282	9.191	0.528	0.624	
WQbot3	Diff. in % of river mi. in bottom 3 WQ levels	−20.774	10.666	0.009	0.936	***
WStop3	Diff. in % of river mi. in top 3 WS levels	16.840	10.265	0.877	0.893	
WSbot3	Diff. in % of river mi. in bottom 3 WS levels	−9.727	10.169	0.057	0.719	**
ALtop3	Diff. in % of river mi. in top 3 AL levels	8.322	9.916	0.607	0.680	
ALbot3	Diff. in % of river mi. in bottom 3 AL levels	−3.407	7.991	0.145	0.546	*
Top- and bottom-level differences, 25-mile radius of respondent's home						
WQ25top3	Diff. in % of river mi. in top 3 WQ levels	1.296	7.052	0.303	0.502	
WQ25bot3	Diff. in % of river mi. in bottom 3 WQ levels	−10.664	9.862	0.043	0.755	***
WS25top3	Diff. in % of river mi. in top 3 WS levels.	3.330	6.904	0.409	0.531	
WS25bot3	Diff. in % of river mi. in bottom 3 WS levels	−4.466	7.523	0.099	0.562	**
AL25top3	Diff. in % of river mi. in top 3 AL levels	5.680	8.778	0.506	0.606	
AL25bot3	Diff. in % of river mi. in bottom 3 AL levels	−3.438	7.903	0.141	0.542	*
Top- and bottom-level differences, 10-mile radius of respondent's home						
WQ10top3	Diff. in % of river mi. in top 3 WQ levels	0.759	6.919	0.280	0.497	
WQ10bot3	Diff. in % of river mi. in bottom 3 WQ levels	−9.852	9.614	0.049	0.739	***
WS10top3	Diff. in % of river mi. in top 3 WS levels	2.662	6.149	0.377	0.501	
WS10bot3	Diff. in % of river mi. in bottom 3 WS levels	−7.083	8.133	0.060	0.668	**
AL10top3	Diff. in % of river mi. in top 3 AL levels	12.440	10.154	0.766	0.800	
AL10bot3	Diff. in % of river mi. in bottom 3 AL levels	−6.860	9.152	0.083	0.642	**
Top- and bottom-level differences, longest-looked map frame by each respondent						
F1_WQtop3	Diff. in % of river mi. in top WQ 3 levels	3.636	7.922	0.415	0.551	
F1_WQbot3	Diff. in % of river mi. in bottom 3 WQ levels	−7.108	8.843	0.071	0.651	**
F1_WStop3	Diff. in % of river mi. in top 3 WS levels	1.882	5.836	0.330	0.476	
F1_WSbot3	Diff. in % of river mi. in bottom 3 WS levels	−8.255	8.496	0.048	0.705	***
F1_ALtop3	Diff. in % of river mi. in top 3 AL levels	2.275	7.105	0.350	0.515	
F1_ALbot3	Diff. in % of river mi. in bottom 3 AL levels	−1.029	6.889	0.207	0.498	
Constant	—	5.416	9.951	0.707	—	
Scale	—	789.591	81.497	1.000	—	—

BMS and averaging results are shown. ***$max\left(P>0,\ 1-(P>0)\right)>0.95$; **threshold > 0.90; *threshold > 0.80. All signals assigned conditional on inclusion probabilities > 0.5 (prior odds). Parameter means on explanatory variables are interpreted as marginal WTP, in US dollars (USD) per year (2021 USD). Diff. in index points, difference in index points (±0–100); Diff. in % of river mi., difference in percentage of river miles (±0–100).

longest-looked map frame (*F1_WQbot3*). Analogous WTP for WS is \$8.26 (*F1_WSbot3*). These results suggest that map interactions yielded information on individualized spatial areas that are relevant to understanding households' values for water quality change. Changes in these areas have systematic effects on WTP, beyond effects related to regionwide changes and changes surrounding respondents' homes at fixed distances.

To illustrate the magnitude of these effects, consider that the mean values for variables *F1_WQbot3* and *F1_WSbot3* within the data are −3.34 and −3.27, respectively (*SI Appendix*). This implies a total contribution of \$7.11 × 3.34 = \$23.75 and \$8.26 × 3.27 = \$26.98, respectively, to each household's total WTP for an average water quality improvement policy shown in the survey, as represented by these variable means. These values are in addition to those linked to overall regional changes and changes within 10 and 25 miles of respondents' homes.

Coordination of these estimates with the map-interaction results in Fig. 2 provides insight into areas where quality changes have particularly high value, ceteris paribus. Fig. 2 *B* and *C* reveal areas where the longest-looked map frames were concentrated. Table 1 shows WTP associated with improvements to these areas. Hence, the maps in Fig. 2 may be interpreted as providing insight into regions of New England where water quality has higher-than-average value to our sample, as revealed by map interactions.

Results also yield conclusions for nonspatial dimensions of WTP. Of the three average-change indicators, WS_i produces the strongest signal with WTP of \$27.82 per 1-unit improvement in WS throughout the area. Regionwide average changes in WQ_i and AL_i, while still producing predominantly positive posterior draws, have less pronounced effects. For regionwide quality-level shares, respondents' choices appear to be motivated primarily by a desire to improve lower-quality river areas. Results suggest that respondents would be willing to pay \$20.77, \$9.72, and \$3.41 for 1–percentage point reductions in regionwide river length in the bot3 quality levels of WQ, WS, and AL, ceteris paribus (*WQbot3*, *WSbot3*, and *ALbot3*).

Discussion

This study evaluates whether allowing people to interact with maps during a choice experiment can reveal areas where water quality has high value. To enable the analysis, we purposefully avoided some of the simplifications that are common in choice experiment scenarios (e.g., limiting scenarios to average changes over large areas). The survey elicited values for a set of predicted water quality changes throughout a regionwide river system—such that changes potentially occur everywhere throughout the domain and were not focused on iconic or recreational areas. This less restricted form of scenario allowed respondents to explore heterogeneous water quality changes using interactive maps and thereby ground their choices in changes and areas that matter to them—rather than having these predefined by the researchers. This structure, paired with map-interaction architecture, allowed us to characterize spatial dimensions of WTP for realistic scenarios of water quality in ways not possible otherwise.

Our results offer evidence that respondents' map interactions convey systematic information that is related to their choices and WTP estimates. Estimated values are influenced by water quality changes close to each respondent's home, as anticipated, but also in locations identifiable via each respondent's map interactions. These spatial effects are pertinent solely for improvements to rivers at low current quality, indicating that spatial WTP heterogeneity depends on whether improvements occur in high- or low-quality waters. Overlooking these effects may cause researchers to undervalue some types of water quality change.

These findings have direct implications for how people value water quality and how values are estimated. As noted above, the dominant spatial paradigm in SP valuation is distance decay around the areas where people live (7). Studies occasionally account for additional influences on WTP, such as spatial scale, effects on iconic or other areas predefined by the researchers, recreational uses, geopolitical boundaries, directionality, substitutes, complements, and conditions surrounding people's homes (6, 8, 21, 23–33). Related work applies spatial econometrics to assess patterns not otherwise explained by observable variables (34–37). Yet although many approaches are available to model spatial dimensions of WTP (6), SP data are almost always generated using survey architectures that provide little or no opportunity for respondents to engage with maps or explore conditions in areas that matter to them. We find that this engagement provides information that explains how and where people value water quality.

This approach integrates map tracking with welfare estimation and thereby suggests multiple avenues for future research. These include whether and how the identified types of WTP patterns might generalize—for example, to other contexts for which water quality values are estimated or different types of nonmarket values. It is also unknown whether analogous results could be obtained using spatial-salience data derived via alternative approaches. These and other areas for future research are discussed in the *SI Appendix*.

Multiple caveats should be considered when interpreting our results. Among these, our results are derived via analysis of survey responses from the realized sample and should be interpreted accordingly. Corresponding to our methodological focus, the results are not intended to provide representative estimates of WTP for statewide populations. We took multiple steps to induce a random address-based sample. However, data screening was required for modeling, the mailing list was restricted to single-family households, the survey could only be taken online, and self-selection can occur during survey response. Data in the *SI Appendix* show that the sample is reasonably representative across some, but not all, demographics compared to US Census averages (e.g., younger and less educated individuals are underrepresented). This is a common property of methodological SP research (3).

Second, following prior choice experiments that used survey metrics based on respondent engagement time [e.g., in Campbell et al. (20)], we treated variables derived from map-interaction data as deterministic. We thus interpret the longest-looked map frames as identifying fixed locations that are potentially salient to each respondent, much as respondents' home locations are interpreted as salient, deterministic locations in distance-decay models. We recognize, however, that one might alternatively interpret map interactions as providing stochastic (or endogenous) indicators of underlying, latent spatial constructs. Corresponding models might seek to explain why individuals chose to interact with certain maps or areas. Our objective was not to explore models of this type, but we acknowledge this as a topic for future exploration.

Third, our scenarios were not created using a typical, mix-and-match combination of environmental attribute levels. To enhance estimation efficiency and reduce cognitive burden, choice experiment scenarios typically show changes using two to six attributes in a simplified matrix (38). For water quality applications, these attributes often represent average conditions over large areas. (In our case, for example, this might have involved scenarios presented solely in terms of mean regionwide WQ_i, WS_i, and AL_i.) Levels for each attribute, within each scenario, are typically assigned using an experimental design that optimizes efficiency for econometric modeling (39), rather than representing predicted future conditions as they might actually occur. To enable the presented analyses, we avoided such ecologically artificial designs and instead produced scenarios using a full factorial of underlying policy actions that are possible in the study domain (*SI Appendix*). These, combined with biophysical processes modeled in FrAMES, produced a set of scenarios with greater realism, correlation among variables, and complexity than are typical in choice experiments. We thus sacrificed estimation efficiency for realism and the capacity to estimate otherwise-obscured spatial effects.

Materials and Methods

Additional methods are summarized below, with details in the *SI Appendix*. The study was approved by the institutional review boards of Clark University and Virginia Tech. Consent was documented using an online consent form that preceded the survey.

Survey Development and Implementation. The questionnaire was designed following best practices as described in the *SI Appendix* (3). Each questionnaire presented one value-elicitation question, along with a) instructions, b) supporting information, c) ancillary and supporting questions, and d) information to enhance incentive compatibility. Each voting question compared predictions for quality measures under the BAU status quo and 1 of the 41 alternative policy scenarios. Generic labels were applied to the choice alternatives, describing the BAU as program A and the alternative policy scenario as program B. For each quality measure, differences between programs A and B were shown using side-by-side maps, bar charts, and spatial mean values. Fig. 3 shows the side-by-side illustration used to communicate differences in the WQ measure for one of the possible policy scenarios. Analogous comparisons were provided for all measures. The survey then presented a binary choice between program A (BAU at \$0 household cost) and program B (the alternative policy scenario at a hypothetically binding, annual cost per household in unavoidable taxes and fees). Possible cost levels were \$30, \$60, \$120, \$240, \$480, \$720, \$960, and \$1,200. The survey was implemented during May to June 2021 using an address-based push-to-web sample. Personalized invitation letters were mailed to 7,167 randomly selected households in each state followed by two reminder mailings. Of 42,979 deliverable invitations, 2,203 total responses were received (5.13% response rate). Of these, 1,698 answered the choice question and had an identifiable home location in or close to the study area.

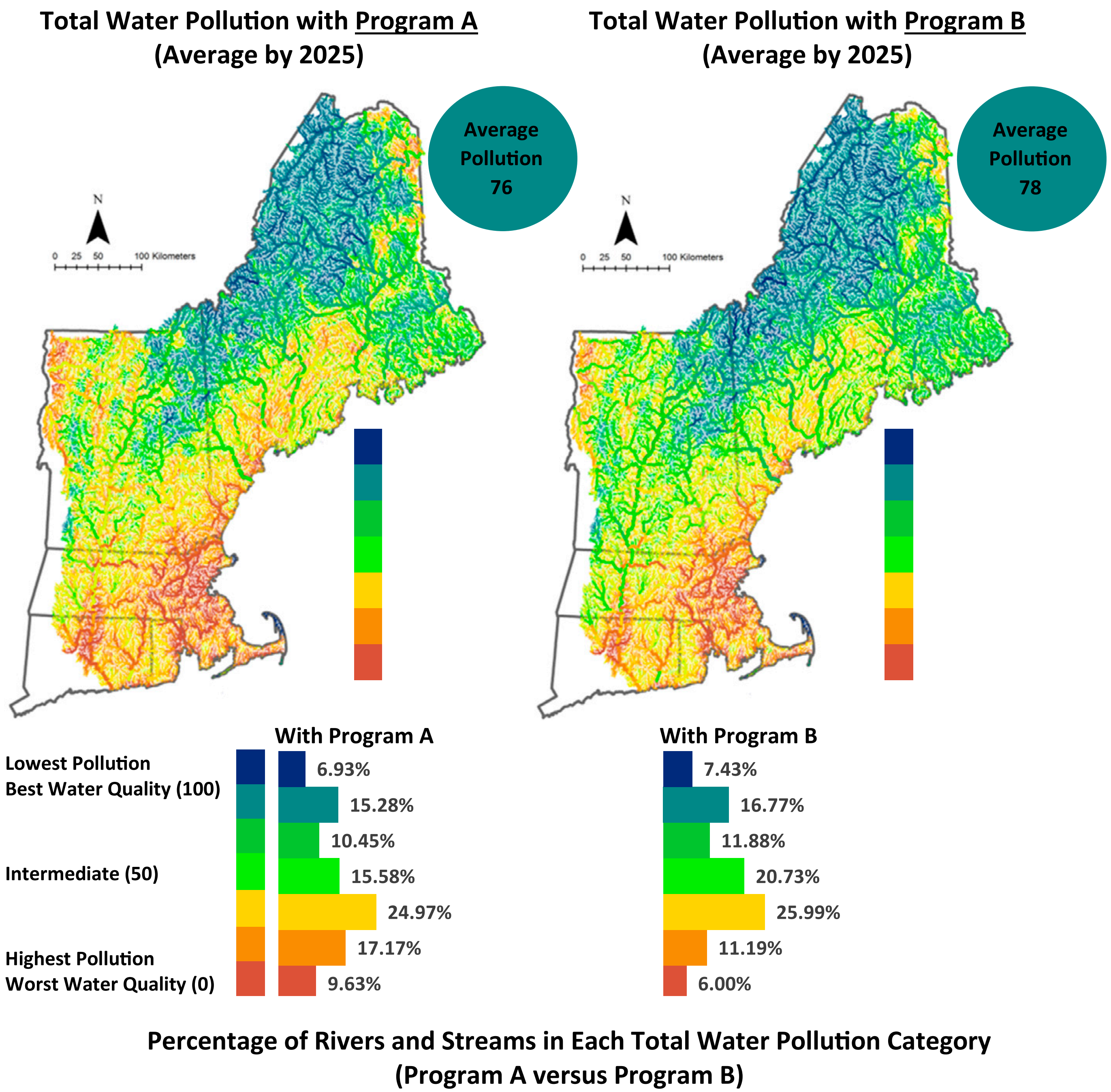

Fig. 3. Example comparison of WQ under program A (the BAU) and program B (the alternative policy scenario), shown as part of the voting question. The figure shows conditions under alternative policy scenario 1 in the experimental design, out of 41 possible alternative scenarios. URL links were also provided on the same survey page, allowing respondents to visit these maps in an interactive GIS environment.

Econometric Methods. The estimation sample was restricted to the $n = 1{,}239$ respondents who interacted with at least one map and lived within 10 miles of the study domain. Each of $i = 1\ldots n$ respondents received one of $s = 1\ldots n$ choice scenarios. Each scenario consisted of a BAU option at zero cost and policy option s at cost P_i. Indirect utilities can be written as follows:

$$\begin{aligned}&\tilde{U}^*_{0i} = \boldsymbol{x}'_{0i}\boldsymbol{\beta}^* + \lambda m_i + \tilde{\varepsilon}^*_{0i}\\&\tilde{U}^*_{si} = \boldsymbol{x}'_{si}\boldsymbol{\beta}^* + \lambda(m_i - P_i) + \tilde{\varepsilon}^*_{si}\\&\tilde{\varepsilon}^*_{ji} \sim EV(0,1), \quad j = 0, s.\end{aligned} \qquad [1]$$

Index 0 is for BAU; regressors $\boldsymbol{x}_{si}$ comprise water quality measures under scenario s, possibly augmented with respondent-specific spatial information; and m_i is income. Error term $\tilde{\varepsilon}^*$ captures unobservables and is assumed to follow a type I extreme value (EV) distribution with zero mean and unit scale. Taking the difference between utility and dividing by the price coefficient λ yields

$$\begin{aligned}&U_i = \boldsymbol{x}'_i\boldsymbol{\beta} - P_i + \varepsilon_i, \quad \text{where}\\&U_i = \frac{(\tilde{U}^*_{si} - \tilde{U}^*_{0i})}{\lambda}, \quad \boldsymbol{x}_i = (\boldsymbol{x}_{si} - \boldsymbol{x}_{0i}), \quad \boldsymbol{\beta} = \frac{\boldsymbol{\beta}^*}{\lambda}, \quad \varepsilon_i = \frac{(\tilde{\varepsilon}^*_{si} - \tilde{\varepsilon}^*_{0i})}{\lambda},\\&\varepsilon_i \sim LOG(0, s), \quad s = \lambda^{-1}.\end{aligned} \qquad [2]$$

Adjusted utility U_i has the interpretation of surplus, defined as the difference between the full WTP to obtain scenario s and the required payment P_i. The estimated coefficients in $\boldsymbol{\beta}$ can therefore be interpreted as marginal WTP. The error scale in the model can, equivalently, be interpreted as the inverse of the price coefficient or marginal utility of income. The differenced and adjusted error follows a logistic (LOG) distribution with a scale equal to the inverted price coefficient [e.g., see Train (40)]. We focus on estimation of latent WTP, $y_i^* = \boldsymbol{x}'_i\boldsymbol{\beta} + \varepsilon_i$, and treat P_i as the decision threshold for the observed yes/no vote for option s over the BAU. We estimate model parameters ($\boldsymbol{\beta}$, s) in a Bayesian framework for reasons of a) nondependence on large-sample (asymptotic) theory, b) obtaining full finite-sample distributions for each parameter, c) ease of deriving full distributions for predictive constructs, and d) ability to perform high-powered model searches to determine which variables likely drive respondents' decisions. Prior examples of Bayesian logit models include Holmes and Held (41), Frühwirth-Schnatter and Frühwirth (42), and Frühwirth-Schnatter and Frühwirth (43). The BMS algorithm is designed to deal with model uncertainty wherein the best combination of explanatory variables is inherently unknown. Model-averaged estimates are shown in Table 1. Details are in the *SI Appendix*.

Data Availability. The data and code that support the findings reported in this study are available publicly as files in the *SI Appendix*. All other study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. This project was developed under Assistance Agreement RD 83616701-1 awarded by the US EPA to Clark University. The views expressed are those of the authors and do not necessarily reflect those of US EPA. Lu Ann Pacenka and Rob Stewart provided assistance.

Author affiliations: [a]George Perkins Marsh Institute, Clark University, Worcester, MA 01610; [b]Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA 24061; [c]Enterprise GIS, Virginia Tech, Blacksburg, VA 24061; [d]Clark Labs, Clark University, Worcester, MA 01610; [e]Department of Natural Resources & the Environment, University of New Hampshire, Durham, NH 03824; and [f]ICF International, Cambridge, MA 02140

1. C. Griffiths *et al.*, US Environmental Protection Agency valuation of surface water quality improvements. *Rev. Environ. Econ. Policy* **6**, 130–146 (2012).
2. D. A. Keiser, C. L. Kling, J. S. Shapiro, The low but uncertain measured benefits of US water quality policy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5262–5269 (2019).
3. R. J. Johnston *et al.*, Contemporary guidance for stated preference studies. *J. Assoc. Environ. Resour. Econ.* **4**, 319–405 (2017).
4. W. Adamowicz, P. Boxall, M. Williams, J. Louviere, Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. *Am. J. Agric. Econ.* **80**, 64–75 (1998).
5. J. De Valck, J. Rolfe, Spatial heterogeneity in stated preference valuation: Status, challenges and road ahead. *Int. Rev. Environ. Resour. Econ.* **11**, 355–422 (2018).
6. K. Glenk, R. J. Johnston, J. Meyerhoff, J. Sagebiel, Spatial dimensions of stated preference valuation in environmental and resource economics: Methods, trends and challenges. *Environ. Resour. Econ.* **75**, 215–242 (2020).
7. R. J. Sutherland, R. G. Walsh, Effect of distance on the preservation value of water quality. *Land Econ.* **61**, 281–291 (1985).
8. J. Rolfe, J. Windle, 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 (2012).
9. R. J. Johnston, E. Y. Besedin, R. Stapler, Enhanced geospatial validity for meta-analysis and environmental benefit transfer: An application to water quality improvements. *Environ. Resour. Econ.* **68**, 343–375 (2017).
10. C. Moore, D. Guignet, K. B. Maguire, C. Dockins, N. B. Simon, Valuing ecological improvements in the Chesapeake Bay and the importance of ancillary benefits. *J. Benefit Cost Anal.* **9**, 1–26 (2018).
11. N. Hanley, F. Schläpfer, J. Spurgeon, Aggregating the benefits of environmental improvements: Distance-decay functions for use and non-use values. *J. Environ. Manage.* **68**, 297–304 (2003).
12. W. M. Wollheim, B. J. Peterson, C. J. Vorosmarty, C. S. Hopkinson, S. A. Thomas, Dynamics of N removal over annual time scales in a suburban river network. *J. Geophys. Res. Biogeosci.* **113**, G03038 (2008).
13. N. Samal *et al.*, Projections of coupled terrestrial and aquatic ecosystem change relevant to ecosystem service valuation at regional scales. *Ecol. Soc.* **22**, 18 (2017).
14. R. T. Carson, T. Groves, Incentive and informational properties of preference questions. *Environ. Resour. Econ.* **37**, 181–210 (2007).
15. I. J. Bateman, B. H. Day, S. Georgiou, I. Lake, The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP. *Ecol. Econ.* **60**, 450–460 (2006).
16. US Environmental Protection Agency, *National Beach Guidance and Required Performance Criteria for Grants* (US Environmental Protection Agency), Washington, DC, 2002).
17. R. P. Morgan, K. M. Kline, M. J. Kline, S. F. Cushman, M. T. Sell, Stream conductivity: Relationships to land use, chloride, and fishes in Maryland streams. *N. Am. J. Fish. Manage.* **32**, 941–952 (2012).
18. S. Zuidema, W. M. Wollheim, M. M. Mineau, M. B. Green, R. J. Stewart, Controls of chloride loading and impairment at the river network scale in New England. *J. Environ. Qual.* **47**, 839–847 (2018).
19. K. Balcombe, I. Fraser, L. Williams, E. McSorley, Examining the relationship between visual attention and stated preferences: A discrete choice experiment using eye-tracking. *J. Econ. Behav. Organ.* **144**, 238–257 (2017).
20. D. Campbell, M. R. Mørkbak, S. B. Olsen, The link between response time and preference, variance and processing heterogeneity in stated choice experiments. *J. Environ. Econ. Manage.* **88**, 18–34 (2018).
21. B. M. Holland, R. J. Johnston, Optimized quantity-within-distance models of spatial welfare heterogeneity. *J. Environ. Econ. Manage.* **85**, 110–129 (2017).
22. K. Train, M. Weeks, "Discrete choice models in preference space and willing-to-pay space" in *Applications of Simulation Methods in Environmental and Resource Economics*, R. Scarpa, A. Alberini, Eds. (Springer, 2005), pp. 1–16.
23. M. Schaafsma, R. Brouwer, J. Rose, Directional heterogeneity in WTP models for environmental valuation. *Ecol. Econ.* **79**, 21–31 (2012).
24. S. L. Jørgensen *et al.*, 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**, 58–66 (2013).
25. S. Lizin, R. Brouwer, I. Liekens, S. Broeckx, Accounting for substitution and spatial heterogeneity in a labelled choice experiment. *J. Environ. Manage.* **181**, 289–297 (2016).
26. M. Czajkowski, W. Budziński, D. Campbell, M. Giergiczny, N. Hanley, Spatial heterogeneity of willingness to pay for forest management. *Environ. Resour. Econ.* **68**, 705–727 (2017).
27. D. S. Choi, R. Ready, Measuring benefits from spatially-explicit surface water quality improvements: The roles of distance, scope, scale, and size. *Resour. Energy Econ.* **63**, 101108 (2021).
28. B. Danley, E. D. Sandorf, D. Campbell, Putting your best fish forward: Investigating distance decay and relative preferences for fish conservation. *J. Environ. Econ. Manage.* **108**, 102475 (2021).
29. R. J. Johnston, J. M. Duke, Willingness to pay for land preservation across states and jurisdictional scale: Implications for benefit transfer. *Land Econ.* **85**, 217–237 (2009).
30. R. Brouwer, J. Martin-Ortega, J. Berbel, Spatial preference heterogeneity: A choice experiment. *Land Econ.* **86**, 552–568 (2010).
31. M. G. Interis, D. R. Petrolia, Location, location, habitat: How the value of ecosystem services varies across location and by habitat. *Land Econ.* **92**, 292–307 (2016).
32. S. B. Olsen, C. U. Jensen, T. E. Panduro, Modelling strategies for discontinuous distance decay in willingness to pay for ecosystem services. *Environ. Resour. Econ.* **75**, 351–386 (2020).
33. T. Badura, S. Ferrini, M. Burton, A. Binner, I. J. Bateman, Using individualised choice maps to capture the spatial dimensions of value within choice experiments. *Environ. Resour. Econ.* **75**, 297–322 (2020).
34. D. Campbell, R. Scarpa, W. G. Hutchinson, Assessing the spatial dependence of welfare estimates obtained from discrete choice experiments. *Lett. Spat. Resour. Sci.* **1**, 117–126 (2008).
35. D. Campbell, W. G. Hutchinson, R. Scarpa, Using choice experiments to explore the spatial distribution of willingness to pay for rural landscape improvements. *Environ. Plann. A* **41**, 97–111 (2009).
36. R. J. Johnston, M. Ramachandran, Modeling spatial patchiness and hot spots in stated preference willingness to pay. *Environ. Resour. Econ.* **59**, 363–387 (2014).
37. R. J. Johnston, D. Jarvis, K. Wallmo, D. Lew, Multi-Scale spatial pattern in nonuse willingness to pay: Applications to threatened and endangered marine species. *Land Econ.* **91**, 739–761 (2015).
38. J. R. DeShazo, G. Fermo, Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *J. Environ. Econ. Manage.* **44**, 123–143 (2002).
39. R. Scarpa, J. Rose, Design efficiency for choice modelling. *Aust. J. Agric. Resour. Econ.* **52**, 253–282 (2008).
40. K. Train, *Discrete Choice Methods with Simulation* (Cambridge University Press, ed. 2, 2009).
41. C. Holmes, L. Held, Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Anal.* **1**, 145–168 (2006).
42. S. Frühwirth-Schnatter, R. Frühwirth, Auxiliary mixture sampling with application to logistic models. *Comput. Stat. Data Anal.* **51**, 3509–3528 (2007).
43. S. Frühwirth-Schnatter, R. Frühwirth, "Data augmentation and MCMC for binary and multinomial logit models" in *Statistical Modelling and Regression Structures*, T. Kneib, G. Tutz, Eds. (Springer-Verlag, 2010), pp. 111–131.