NCEE Working Paper

Hot Spots, Cold Feet, and Warm Glow: Identifying Spatial Heterogeneity in Willingness to Pay

Dennis Guignet, Christopher Moore, and Haoluan Wang

Working Paper 20-01
March, 2020

U.S. Environmental Protection Agency
National Center for Environmental Economics
https://www.epa.gov/environmental-economics

NCEE
NATIONAL CENTER FOR ENVIRONMENTAL ECONOMICS

Hot Spots, Cold Feet, and Warm Glow: Identifying Spatial Heterogeneity in Willingness to Pay

Dennis Guignet*[1], Christopher Moore[2], and Haoluan Wang[3]

Last Revised: March 10, 2020

*Corresponding Author
Department of Economics
Appalachian State University
416 Howard Street
ASU Box 32051
Boone, NC 28608
Phone: +1 828-262-2117
guignetdb@appstate.edu

[1] Department of Economics, Appalachian State University
[2] National Center for Environmental Economics, U.S. EPA
[3] Department of Agricultural and Resource Economics, University of Maryland

The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA. Although the research described in this paper may have been funded entirely or in part by the U.S. EPA, it has not been subjected to the Agency's required peer and policy review. No official Agency endorsement should be inferred. We thank participants at the Northeastern Agricultural and Resource Economics Association's 2018 meeting, the Agricultural and Applied Economics Association's 2019 meeting, the 1st D.C. Area Student/Professor Environmental and Energy Economics Workshop, and faculty of Appalachian State University and Salisbury University for helpful comments. We especially thank Robert Johnston, Adan Martinez-Cruz, and Anna Alberini for providing valuable suggestions on earlier versions of this paper.

Hot Spots, Cold Feet, and Warm Glow: Identifying Spatial Heterogeneity in Willingness to Pay

Abstract:

We propose a novel extension of existing semi-parametric approaches to examine spatial patterns of willingness to pay (WTP) and status quo effects, including tests for global spatial autocorrelation, spatial interpolation techniques, and local hotspot analysis. We are the first to formally account for the fact that observed WTP values are estimates, and to incorporate the statistical precision of those estimates into our spatial analyses. We demonstrate our two-step methodology using data from a stated preference survey that elicited values for improvements in water quality in the Chesapeake Bay and lakes in the surrounding watershed. Our methodology offers a flexible way to identify potential spatial patterns of welfare impacts, with the ultimate goal of facilitating more accurate benefit-cost and distributional analyses, both in terms of defining the appropriate extent of the market and in interpolating values within that market.

JEL Classification: C11 (Bayesian Analysis); C14 (Semiparametric and Nonparametric Methods); Q51 (Valuation of Environmental Effects); Q53 (Air Pollution; Water Pollution; Noise; Hazardous Waste; Solid Waste; Recycling)

Keywords: Bayesian; hotspot analysis; semi-parametric; spatial heterogeneity; stated preference; water quality

1. INTRODUCTION

A better understanding of the spatial distribution of welfare impacts is necessary for conducting accurate benefit-cost and distributional analyses, both in terms of defining the appropriate extent of the market and in interpolating values within that market. There is a rich body of literature on the spatial dimensions of stated preference (SP) studies, focusing on various analytical and statistical methods (see Glenk et al., 2020 for a review). These approaches include the incorporation of space within the survey design (Wang and Swallow, 2016; Badura et al., 2020), the combining of spatial variables with traditional econometric methods (Jørgensen et al., 2013; Schaafsma et al., 2013; Olsen et al., 2020), and the application of spatial econometrics and geo-statistics (Czajkowski et al., 2017; Budziński et al., 2017).

To examine spatial heterogeneity in willingness to pay (WTP), most SP studies have applied the distance decay paradigm, where WTP is hypothesized to diminish with distance from the resource (Bateman et al., 2006). Generally, estimating the effect of distance on WTP depends on the nature of the distance measure (e.g., travel, Euclidean, or geodesic distance) and the econometric model specifications once a particular distance measure is chosen (e.g., linear or non-linear distance decay). While there is a growing number of SP studies using traditional econometric methods to control for "distance decay" or other forms of spatial heterogeneity (e.g., Hanley et al., 2003; Rolfe and Windle, 2012; Olsen et al., 2020), these parametric approaches can sometimes fail to identify existing spatial patterns.

Disciplines in the natural sciences employ more spatially-oriented analytical tools to examine spatial patterns. These tools include tests for global spatial autocorrelation (Getis, 2007), spatial interpolation techniques (Anselin and Gallo, 2006), and local cluster or hotspot analyses (Wang and Qiu, 2017). These tools have been increasingly applied in economics, and in particular, in the nonmarket valuation literature. For example, studies have tested for and generally found positive global spatial autocorrelation of individual-specific WTP values (Campbell et al., 2009; Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015). SP studies have also employed local indicators of spatial association (LISAs) or hotspot analysis to identify local clusters of systematically higher or lower WTP values (i.e., hot and cold spots, respectively) (Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015). In general, these studies found non-continuous, local spatial patterns of WTP.

In contrast to applications of these spatial tools in the natural sciences, the measures under study by economists are often estimates (e.g., WTP), and not observed values. Although previous studies have qualitatively recognized this fact, and its potential importance, no study to date formally accounts for the statistical precision of those estimates when conducting spatial analyses. We are the first to do so by incorporating techniques borrowed from meta-analytic methods into our spatial autocorrelation, interpolation, and hotspot analyses.

We set out to accomplish three main research objectives. First, we develop a two-stage spatial econometric approach to account for the fact that economic analyses typically observe estimated values of WTP and other measures of interest. We use Bayesian modeling techniques to estimate not only household-specific WTP, but also household-level measures of the variances of those

estimates. Doing so can be important because some households may be, for example, less knowledgeable of the environmental commodity, have less defined preferences, or even be less engaged when taking the SP survey. If such households are systematically distributed over space, then not accounting for the statistical precision of the household-specific WTP estimates can confound subsequent spatial analyses. With the household-specific empirical WTP distributions in-hand, we are able to treat the variables of interest not as given data, but as statistically derived estimates. The household-specific variances of the WTP estimates are directly incorporated into the spatial weights used in the second-stage spatial analyses.

Our second objective is to demonstrate our proposed two-stage methodology using data from a SP survey that elicited values for improvements in water quality in the Chesapeake Bay and lakes in the surrounding watershed. Using our proposed variance-adjusted tests for global spatial autocorrelation, spatial interpolation techniques, and hotspot analyses, we examine the spatial distribution and clustering of marginal WTP (MWTP) for improvements in several environmental attributes. We also examine the spatial distribution of status quo (SQ) effects, which are intended to capture potential biases for (e.g., "warm glow") or against (e.g., "cold feet") a policy option that are not explained by the choice attributes defining that policy option. To our knowledge, this is the first study to examine the spatial distribution of respondents exhibiting potential biases associated with SP methods. Such an examination provides insights to improve SP methods, welfare analysis, and future survey designs.

The third objective is to illustrate the potential policy implications of our proposed variance-adjusted spatial analyses. We use our two-stage methodology to estimate total WTP for projected improvements resulting from the Chesapeake Bay Total Maximum Daily Loads (TMDLs). The total benefit estimates are compared to spatial interpolation approaches that do not account for the statistical precision of the WTP estimates, similar to those used in earlier studies (e.g., Campbell et al., 2009; Johnston and Ramachandran, 2014; Johnston et al., 2015), and to conventional models that assume homogeneity of WTP across the population or control for heterogeneity parametrically based on observed household characteristics (Moore et al., 2018).

Our semi-parametric results of the spatial interpolation suggest distinct local patterns in MWTP estimates for all attributes, and evident spatial heterogeneity across the study area. The hotspot analysis confirms statistically significant spatial clusters of high and low MWTP values. Comparison of the conventional spatial analyses to our variance-adjusted results reveals some differences. In general, accounting for the variances of the MWTP estimates diminishes spatial variation, suggesting that not accounting for the statistical precision of the first-stage MWTP estimates could lead analysts to falsely identify patterns of global and local spatial heterogeneity in the second stage. Our analysis of spatial variation of individual-specific SQ effects reveals substantial differences when accounting for the statistical precision of the estimates. In particular, our proposed variance-adjustment leads to an increased identification of clusters of individuals exhibiting "warm glow" or other biases *for* a policy option. Lastly, although differences in local patterns are revealed, our policy simulations suggest that accounting for local spatial heterogeneity (with or without our variance-adjusted extension) may not yield substantial differences in terms of broader welfare implications, at least not in this particular application of water quality and ecosystem improvements in an iconic and well-known estuary.

The remainder of the paper is organized as follows. In the next section we present our two-step empirical methodology. We then discuss the data for the specific application demonstrating that methodology. The empirical results and implications are then presented, followed by some concluding remarks.

2. METHODS

2.1. Random Utility Models (RUM)

Stated choice models are often estimated in a random utility framework, where v_{ij} denotes the deterministic component of utility that respondent *i* receives from alternative *j* in choice occasion *t.* Each respondent is given three choice questions in the application presented, but the choice occasion subscript *t* is omitted for notational ease. The random component of utility is denoted as ε_{ij}. Assuming ε_{ij} is independently and identically distributed following a type I extreme value distribution allows the model to be estimated as a conditional or mixed logit (Maddala, 1983; Greene, 2003; Train, 2003).

The deterministic component of indirect utility v_{ij} is a function of the vector of environmental improvements $\boldsymbol{x}_j$, the cost of living increase incurred by the household $cost_j$, and a binary indicator denoting the status quo option SQ_j. To better capture preference heterogeneity, we interact $\boldsymbol{x}_j$ with a vector of dummy variables denoting whether individual *i* is a user of the corresponding resource $\boldsymbol{user}_i$. We adopt a linear model and log-transform the environmental attributes to capture diminishing marginal utility, while also preserving more degrees of freedom than a model with higher order effects. Marginal utility of income is assumed constant across users and nonusers. Using this specification of $v(\cdot)$, the conditional probability that household *i* would choose alternative *j* is:

$$P_i\left(j \mid \boldsymbol{x}_q, cost_q, SQ_q, user_i\right) = \frac{exp\{ln(\boldsymbol{x}_j)\boldsymbol{\beta}+(ln(\boldsymbol{x}_j)\times user_i)\boldsymbol{\delta}+\gamma cost_j+\varphi SQ_j\}}{\sum_q exp\{ln(\boldsymbol{x}_q)\boldsymbol{\beta}+(ln(\boldsymbol{x}_q)\times user_i)\boldsymbol{\delta}+\gamma cost_q+\varphi SQ_q\}} \quad (1)$$

where q indexes all available alternatives in a given choice occasion. The coefficients to be estimated are $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, γ, and φ. The first three coefficients can then be used to derive the $\boldsymbol{MWTP_i}$ vector for respondent *i*:

$$\boldsymbol{MWTP_i} = \frac{\boldsymbol{\beta}+(user_i)\boldsymbol{\delta}}{-\gamma\tilde{\boldsymbol{x}}} \quad (2)$$

Given the natural log specification in the model, MWTP is nonlinear and varies at different levels of the environmental attribute. In equation (2), $\tilde{\boldsymbol{x}}$ denotes the environmental attribute reference levels from which $\boldsymbol{MWTP_i}$ is calculated. In the empirical analysis, we set $\tilde{\boldsymbol{x}}$ equal to the baseline values shown in the survey.

2.2. Household-level MWTP distributions

To derive the household-level MWTP distributions that inform our spatial analyses, we adopt a random parameters framework to estimate the logit model characterized by equation (1). To simplify notation, we stack the variables *ln(**x**), cost, SQ,* and the user interaction terms into a single *M*-element vector ***z*** and label the corresponding vector of coefficients $\boldsymbol{\lambda}$. The typical exposition of mixed logit models estimates the distribution of the utility parameters over the population $g(\boldsymbol{\lambda}|\upsilon)$, where $\boldsymbol{\upsilon}$ are the parameters of the distribution such as mean and variance. When estimating household-level parameters, the central concept is the distinction between two distributions: the distribution of preferences in the population and the distribution of preferences in the sub-population who make particular choices (Train, 2003, p. 263). To that end, let $\eta(\boldsymbol{\lambda}|\boldsymbol{y},\boldsymbol{z},\boldsymbol{\upsilon})$ be the distribution of $\boldsymbol{\lambda}$ in the sub-population that chose a particular set of responses to the repeated choice experiments contained in the *t*-element vector *y*.

The conditional distribution $\eta(.)$ can be estimated in a classical framework via maximum likelihood, and household-level parameters for that distribution can then be found by substituting in values of ***y*** and ***z*** (Revelt and Train, 2000). Generally, those expressions will be integrals without closed forms and require simulation. We adopt a Bayesian approach to estimation because the household-level distributions are more easily simulated as part of the Metropolis-Hastings (MH) algorithm, from which any moments of those distributions can be found. We provide a derivation of $\eta(.)$ and a detailed description of the estimation algorithm in Appendix A. There is nothing fundamentally different about our approach from the hierarchical Bayes' iterative estimation that has become standard in the literature (Train, 2003). What allows us to characterize household-level parameter distributions is the omission of a step in the standard algorithm that is usually included for computational efficiency.

Typically, when hierarchical Bayes is executed, the draws of the household-level parameters are only used to condition draws of the population-level parameters and then discarded to preserve computational memory and processing speed. In this application, however, we store the household-level parameter estimates and treat them as draws from the empirical distributions of each respondent, $\eta(\boldsymbol{\lambda}|\boldsymbol{y},\boldsymbol{z},\boldsymbol{\upsilon})$. When the simulation is complete, we have a multivariate distribution of $\boldsymbol{\lambda}_i$ for each respondent *i* from which we can calculate draws from the distribution of MWTP via equation (2). Those vectors of MWTP values characterize the distribution for each individual respondent and can be used to calculate the mean and variance of MWTP at the household level. The latter gives us a measure of statistical precision for the MWTP estimates for each respondent.

Except for the coefficient on cost, all parameters are drawn from a multivariate normal distribution. We follow the common practice of treating the cost coefficient as fixed rather than random to ensure MWTP has defined moments (e.g., Revelt and Train, 1998; Layton and Brown, 2000), and we recognize this imposes a fixed marginal utility of income over the population. There are at least two alternatives to modelling the cost parameter that avoid this assumption. One is to choose a distribution with a strictly positive domain for the cost coefficient, such as log-normal. This is problematic in our application because our algorithm requires storing all draws of every coefficient, rather than just the mean value for each iteration. As a result, there can be draws of the cost coefficient small enough to result in very large MWTP values which create computational

difficulties, requiring ad hoc solutions. The second alternative is to estimate the model in WTP space (Scarpa et al., 2008). This method brings its own computational issues and can result in less precise WTP estimates (Hole and Kolstad, 2012), so we opt for the simplest approach of fixing the cost parameter and acknowledge the implication on our results.

2.3. Household-level status quo effect distributions

In addition to household-level MWTP, we closely examine the SQ effects at the household level. Inclusion of the status quo indicator SQ_j allows us to capture a respondent's tendency to vote for or against the SQ option, irrespective of the cost and attribute levels defining the two alternative policy options. Such tendencies estimated by φ may capture respondents' consideration of omitted variables. It could also reflect "warm glow" if negative or "cold feet" if positive. For these reasons we omit φ from the MWTP calculations (see equation (2)). Nonetheless, the spatial distribution of φ could reveal important information about the validity and reliability of SP responses.

To provide an intuitive interpretation of the magnitude of the SQ effects, we express the effects in terms of probability differences. We compare the probability of choosing the SQ option to a constructed alternative that is identical with respect to the attribute levels but omits φ. We start with equation (1), which expresses the multinomial logit probabilities as a function of environmental attributes, the SQ effect, and the cost. We find the probability that each respondent would choose the SQ option (j=*SQ*) in a given choice occasion, $P_i\left(SQ \mid \boldsymbol{x}_q, cost_q, SQ_q\right)$. We then specify a second probability function for a constructed SQ alternative that omits the SQ constant, $\tilde{P}_i\left(SQ|\boldsymbol{x_q}, cost_q\right) = \frac{exp\{ln(x_{SQ})\boldsymbol{\beta}+(ln(x_{SQ})\times user_i)\boldsymbol{\delta}\}}{\sum_q exp\{ln(\boldsymbol{x}_q)\boldsymbol{\beta}+(ln(\boldsymbol{x}_q)\times user_i)\boldsymbol{\delta}+\gamma cost_q\}}$ and estimate the SQ effect as the difference between the two probabilities:

$$SQ\ effect = P_i\left(SQ \mid \boldsymbol{x}_q, cost_q, SQ_q\right) - \tilde{P}_i\left(SQ|\boldsymbol{x_q}, cost_q\right). \quad (3)$$

We perform spatial analyses of the SQ effects and account for household-level variances in the same manner as the MWTP calculations, generating household-level distributions for the probability differences.

2.4. Spatial analyses

The main contribution of this study is the formal incorporation of the underlying statistical precision around each individual household's MWTP estimates into our spatial analyses. To our knowledge spatial clustering and interpolation studies examining spatial variation to date, have treated the variables of interest as observed values, and not as statistically derived estimates (Campbell et al., 2009; Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015; Czajkowski et al., 2017). Although these studies appropriately caveat their findings, none have formally accounted for the underlying statistical precision of the WTP estimates. We do so by borrowing techniques from meta-analytic methods. Conventional meta-analyses synthesize estimates from different primary studies, and in doing so often weight the primary study estimates

according to their inverse variance (Borenstein et al., 2010). We incorporate this same idea into our three sets of spatial analyses – spatial interpolation, global autocorrelation, and hotspot analyses. The following discussion focuses on the MWTP estimates, but the same procedures are applied to the estimated SQ effects.

2.4.1. Variance-adjusted spatial interpolation

Spatial interpolation entails the creation of a raster (or grid) surface that visually depicts the distribution of household MWTP over space. The weighted household-specific MWTP values are used to predict MWTP for all locations in the study area, which in practice are identified as the centroid of each grid cell. The following equation is used to interpolate the MWTP value assigned to each cell *l*:

$$\widehat{MWTP}_l = \sum_{h=1}^{H} (\omega_{lh} \times MWTP_h). \quad (4)$$

$\widehat{MWTP}_l$ is the predicted MWTP estimate at an unsampled location l. $MWTP_h$ is household h's estimated MWTP value, and ω_{lh} is the element from the spatial weights matrix that links locations l and h. We adopt the following functional form for our weighting equation to account for both the spatial relationships and statistical precision of the primary estimates:

$$\omega_{lh} = \begin{cases} \dfrac{\left(\frac{1}{d_{lh}}\right)^{\rho}\left(\frac{1}{v_h}\right)^{1-\rho}}{\sum_{h \in H_l}\left\{\left(\frac{1}{d_{lh}}\right)^{\rho}\left(\frac{1}{v_h}\right)^{1-\rho}\right\}}, & if\ h \in H_l \\ 0, & if\ h \notin H_l \end{cases} \quad (5)$$

where d_{lh} is the distance from the location of the centroid of cell *l* to household *h*, and v_h is the variance of the MWTP estimate derived for household *h*, which comes from the empirical distribution generated through the 10,000 iterations of our Bayesian modelling approach (after burn-in). The summation in the denominator is over the "K-nearest neighbors" to location l (denoted by the set H_l). Households at greater distances than the K-nearest neighbor are given a weight of zero.

Nelson and Boots (2008) have discussed several ways to define spatial weights matrices that include fixed distance, K-nearest neighbor, and shared boundaries. Following Johnston and Ramachandran (2014) and Johnston et al. (2015), we adopt the K-nearest neighbor method (K=8).[1] This spatial weighting scheme is appropriate for several reasons. First, K=8 is the number at which the permutation distribution of the test G_i^* statistic used in the later hotspot analysis approaches normality (Ord and Getis, 1995). Second, this method ensures that very far households across the large study area do not influence the MWTP estimates, so our results are not overly influenced by

[1] Johnston and Ramachandran (2014) conduct sensitivity analyses based on the assumed spatial relationships and found similar results across alternative assumptions.

outliers. Third, the K-nearest neighbor method naturally adapts the neighborhood definition to account for different population densities in urban, suburban, and rural areas across our study area.

Of particular interest in equation (5) is the assumed value for the parameter ρ, which must satisfy $0 \leq \rho \leq 1$. The ρ parameter determines how much influence spatial proximity versus statistical precision of an estimate has on the spatially interpolated MWTP value for cell l. If $\rho = 1$, then equation (5) simplifies to the inverse distance weighting scheme commonly used in past spatial analyses. If $\rho = 0$, then for the K-nearest neighbors, equation (5) is analogous to the common fixed effect size (FES) weighting scheme often utilized in meta-analyses (Borenstein et al., 2010). The choice of ρ is admittedly arbitrary but given our interests in accounting for both statistical precision and spatial patterns, we assume an equal influence of both factors and set $\rho = 0.5$. A sensitivity analysis is then conducted for alternative values of ρ, and most notably for the case where $\rho = 1$, which allows for comparison of our variance-adjusted weights to the conventional spatial weights used in previous studies.

2.4.2. Variance-adjusted global spatial autocorrelation and hotspot analyses

To test for global and local spatial autocorrelation, it is again necessary to define the neighborhood in which relationships across space are evaluated. In contrast to the interpolation exercise above, where the weights matrix denotes the spatial relationships between each interpolated cell centroid and the households in our sample, in the next set of analyses the spatial relationship defined is between each household and the K-nearest neighboring households, including the household itself (i.e., where $d_{ih} = d_{ii} = 0$). This mathematically prevents us from assuming an inverse distance relationship, as done in equation (5). Given our interests in identifying statistically significant high or low clusters of MWTP, a simple uniform weight of 1/K among the K-nearest neighbors (and zero otherwise) is assumed here.

Again, our novel contribution is to account for the statistical precision of the individual household MWTP estimates. More specifically, the weight given to each neighbor is basically re-distributed among the K neighbors, giving more weight to households where the observed MWTP value was estimated with greater statistical precision (i.e., smaller variance). The weight used for household h in explaining the spatial relationship with household i is:

$$\omega_{ih} = \begin{cases} \dfrac{\left(\frac{1}{v_h}\right)^{1-\alpha}}{\sum_{h \in H_i}\left\{\left(\frac{1}{v_h}\right)^{1-\alpha}\right\}}, & if\ h \in H_i \\ 0, & if\ h \notin H_i \end{cases} \tag{6}$$

The summation in the denominator is over the K-nearest neighbors to household i (denoted by the set H_i). The assumed value for the parameter α must satisfy $0 \leq \alpha \leq 1$. Notice that equation (6) is a more general form of the usual K-nearest neighbor weighting scheme. When $\alpha = 1$, ω_{ih} simplifies to 1/K for those K-nearest neighbors. If $\alpha = 0$, then similar to before, for the K-nearest neighbors equation (6) simplifies to the common FES weighting scheme used in meta-analyses (Borenstein et al., 2010).

To test whether the household-specific MWTP values are a result of random spatial process, we apply Moran's I statistic to test for global spatial autocorrelation (Getis, 2010). The underlying expectation of Moran's I test is that the spatial process promoting the observed pattern of the attribute being analyzed is random. A rejection of the null hypothesis suggests that spatial autocorrelation, either spatial clustering or dispersion, exists. Moran's I statistic ranges from −1 to +1, with scores near +1 indicating spatial clustering and scores near –1 indicating spatial dispersion. Moran's I statistic is defined as:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{h=1}^{n}\omega_{ih}} \frac{\sum_{i=1}^{n}\sum_{h=1}^{n}\omega_{ih}(MWTP_i - \overline{MWTP})(MWTP_h - \overline{MWTP})}{\sum_{i=1}^{n}(MWTP_i - \overline{MWTP})^2} \quad (7)$$

where n is the number of households in the data sample, $MWTP_i$ and $MWTP_h$ are the household MWTP values, $\overline{MWTP} = \frac{\sum_{i=1}^{n} MWTP_i}{n}$ is the mean of all households' MWTP values in the sample, and ω_{ih} is the variance-adjusted spatial weight that links households i and h, as defined in equation (6). Notice that equation (7) is essentially the standard Moran's I statistic formula (Getis, 2010), with our adjusted spatial weights illustrated in equation (6).

While Moran's I test for global spatial autocorrelation provides a means to test for spatial patterns across the broader study area, little is revealed about *local* spatial heterogeneity among households. Local indicators of spatial association (LISAs) have been developed to measure local spatial autocorrelation. Commonly referred to as hotspot analysis, the approach provides a statistical test for identifying spatial clusters of high values (hot spots) or low values (cold spots) of a variable of interest beyond what can be explained by random coincidence (Anselin, 1995). Among the most common LISAs is the Getis-Ord G^* statistic proposed by Getis and Ord (1992). Our proposed variance-adjusted Getis-Ord G^* statistic is calculated as follows:

$$G_i^* = \frac{\sum_{h=1}^{n}\{\omega_{ih} MWTP_h\} - \overline{MWTP}}{\sqrt{\frac{\sum_{h=1}^{n}\{MWTP_h{}^2\}}{n} - (\overline{MWTP})^2}\sqrt{\frac{n\sum_{h=1}^{n}\{\omega_{ih}{}^2\} - (\sum_{h=1}^{n}\omega_{ih})^2}{n-1}}}. \quad (8)$$

Equation (8) is simply the standard Getis-Ord G^* statistic (Getis and Ord, 1992), but with our variance-adjusted spatial weights illustrated in equation (6).

3. DATA

We demonstrate our two-step methodology by examining the spatial distribution of household MWTP for improvements in the Chesapeake Bay and freshwater lakes in the broader Chesapeake Bay Watershed. These MWTP values are estimated from data obtained in a stated choice study by Moore et al. (2018), which focused on reductions in nutrient and sediment pollution, and the resulting improvements in conditions for recreation and aquatic wildlife. Three choice questions were included in each survey. Each choice question presented respondents with a SQ alternative, showing current conditions and zero cost, and two policy alternatives with improvements in some or all of the attributes and a positive cost (see Table 1). The cost attribute was expressed as a

permanent cost of living increase shown in annual terms for each household. The attributes defining each policy alternative were improvements in water clarity, striped bass population, blue crab population, and oyster abundance in the Bay. In addition to the Bay itself, freshwater lakes in the watershed benefit from the management practices targeting the Bay. To capture these benefits, an additional attribute reflecting the number of lakes in the watershed with low algae growth was included in the alternatives presented in the survey. Through a series of ten focus groups, 72 one-on-one cognitive interviews, and an extensive pre-test, these attributes were found to be the most salient and important to the population of interest.

The survey was administered via mail to a geographically stratified random sample of households who reside in the District of Columbia or one of the 17 U.S. states that contain at least part of the Chesapeake Bay Watershed or lie within 100 miles of the Eastern coast of the U.S. The survey was sent to 2,829 households, and after adjusting for undeliverable addresses achieved a response rate of 31%. The resulting 671 surveys were screened for protest responses and hypothetical bias, leaving 559 completed surveys with which to estimate our models. Moore et al. (2018) provide more details on the study design and sample characteristics.

4. RESULTS

4.1. Bayesian model results

There are three sets of results from the Bayesian mixed logit that are relevant to our objectives. The first are the summary statistics of the posterior distributions of our estimated model parameters, $g(\lambda|\upsilon)$. From a classical perspective these statistics can be interpreted as the coefficient estimates and standard errors, presented in the first two columns of results in Table 2. Given the inclusion of user-interaction terms in our model, the coefficients on the logged attribute levels are the marginal utilities for nonusers and the coefficients on the interaction terms are the differentials in marginal utilities for users of the corresponding resource. On average, water clarity in the Bay and striped bass population are not significant determinants of choice for non-users but they are for people that recreate in the Chesapeake Bay. Crab populations are significant for non-users, on average, while their contribution to marginal utility is less for users but not significantly so. Water quality in watershed lakes is a significant attribute for the average nonuser and generates a significantly greater marginal utility for lake-users. The MWTP estimates are reported in the last column of Table 2. The MWTP values refer to a one-unit increase relative to the SQ quantities shown in Table 1; for example, a one-inch increase in clarity, a one-million increase in striped bass and crab populations, etc.

The second set of results describe how the household-level parameters are distributed in the population, shown in Table 3. In this case the mean values refer to the average household value and the standard deviation is an indication of how disperse the values in the population are, and *not* an indication of estimation precision. The corresponding MWTP values and the inner 90^{th} percentile for each are shown in the right-hand side of Table 3. The subsequent spatial analyses examine how households' values in these distributions vary over space, after adjusting for statistical precision using the final set of Bayesian results.

The third and final set of Bayesian results relevant to our objectives are the household-level empirical distributions of MWTP that contain information on the central tendency and precision of our estimates at the household level, $\eta(\lambda_i|y_i,z_i,\upsilon)$. Given the number of respondents in the data, it is not practical to report means and standard deviations for each of them here, but those results are available upon request to the authors.

4.2. Spatial interpolation analysis

Figures 1-6 present the MWTP results for *Bass*, *Clarity*, *Crab*, *Lake*, *Oyster*, and the *SQ effects*. In each figure we include two panels covering the results of the spatial interpolation (left) and hotspot analysis (right). Figure 7 shows the distribution of Getis-Ord G_i^* statistics across households in the hotspot analysis. For each set of maps, we present the results from the conventional spatial weights matrix and those from the variance-adjusted spatial weights matrix introduced in this study.

We first demonstrate the spatial interpolation results on "heat" maps, where darker shades identify higher estimates and lighter shades identify lower estimates. In general, visual inspection of the interpolated MWTP surfaces suggest distinct spatial patterns in MWTP estimates for all attributes and spatial heterogeneity across the study area. For instance, we see some of the highest values for striped bass and water clarity among households living nearest to the Chesapeake Bay, or along the Atlantic Coast and just south of the Bay (left panels of Figures 1 and 2, respectively). We also observe relatively higher MWTP values for improvements in freshwater lakes in the Watershed (*Lake*) among households who live within the watershed (left panel, Figure 5). One unexpected spatial pattern, suggested by the left panel in Figure 3, is that households nearest the Chesapeake Bay hold lower values for improvements in the population of blue crabs (*Crab*), an iconic shellfish species in the Bay. Whereas households outside the watershed seem to hold higher values. As suggested by the Bayesian model results in Table 2, this could reflect the relatively large values among nonusers. Visually there is no clear global pattern in values for improved oyster abundance, as seen in the left panel of Figure 4. There are some areas near and just south of the Bay where residents hold relatively high values for increases in oyster populations, but there are other scattered clusters of high MWTP values (e.g., in New England, the gulf-side of Florida, and northwest Pennsylvania).

The interpolated surface of the *SQ effects* suggests an interesting spatial pattern. In West Virginia and in Northern New York near the Great Lakes and Finger Lakes, we see large positive values for the SQ option, suggesting that households in these areas generally hold a preference against any policy options leading to improvements in the Chesapeake Bay. One theory driving this result could be that households in these areas have their own substitute, and possibly equally as iconic, environmental amenities they care about, and thus have a bias against policy options leading to improvements in the Bay. Another possibility is that there is a protest or strategic response against increased regulations that was not completely eliminated using our earlier screening criteria based on responses to debriefing questions. In any case, the potential biases captured by the *SQ effects* are controlled for and are excluded from the MWTP and welfare calculations.

Comparison of the conventional inverse distance interpolated MWTP maps ($\rho = 1$), to our variance-adjusted interpolations ($\rho = 0.5$) in the left panels of Figures 1-5 reveals some differences. In general, we find that the range of the interpolated MWTP estimates across the study area becomes smaller when we account for the variance of individual respondents' estimates. This suggests that accounting for the statistical precision of the first-stage estimates reduces the influence of outlying, often less precisely estimated, values. Although accounting for the variance of the first-stage estimates seems to diminish spatial variation, and reveals some local differences, the general spatial patterns appear similar.

One exception pertains to the interpolated *SQ effect* estimates. As shown in Figure 6, when accounting for the individual variances ($\rho = 0.5$), many of the areas exhibiting biases against a policy option remain, but there is now noticeably more evidence of respondents exhibiting a relatively high preference for a policy option (i.e., $SQ\ effect < 0$), irrespective of the improvements and costs defining the policy options. Such cold spots, for example, are now evident in the area around New York City, as well as in southern Florida and western Maryland. In any case, the interpolation exercise should be interpreted as suggestive at best. Although it visually depicts relevant spatial patterns, whether such patterns are statistically significant remains an open question. To answer that question, we turn to tests for global spatial autocorrelation and the hotspot analysis in sections 4.3 and 4.4, respectively.

4.3. Tests for global spatial autocorrelation

As can be seen in Table 4, the Moran's *I* tests for global spatial autocorrelation reveal broader spatial patterns for some of the MWTP estimates, but not all. MWTP for increases in striped bass and blue crab populations are both spatially correlated over the broader study area. However, such spatial trends are not generally revealed through parametric modelling of the distance gradients (see Appendix B), thus highlighting the importance of considering potentially relevant spatial patterns revealed by non-parametric methods beyond just distance decay. The Moran's *I* test suggests that the SQ effects are also highly correlated over space.

These findings are robust as we vary the value of α. The strongest evidence of global spatial autocorrelation occurs when using the conventional weights ($\alpha = 1$) that do not take into account the statistical precision of the first-stage estimates. These global patterns remain robust but become less significant as we move towards the variance-adjusted weights ($\alpha = 0$). The Moran's *I* tests suggest no significant global spatial patterns in terms of MWTP for improvements in clarity, oyster abundance, and freshwater lakes. We next examine the nature of these global patterns and whether more local spatial patterns may exist that cannot be identified via the Moran's *I* statistic. In specific, we conduct local spatial associations analyses to test for the presence of statistically significant hot or cold spots using the Getis-Ord G_i^* statistic (Getis and Ord, 1992) and our variance-adjusted variant of the G_i^* statistic. We perform this analysis separately for each attribute, where the sampled households are the spatial units.

4.4. Spatial clustering analysis

The hotspot analysis results for the MWTP of each attribute are illustrated by the set of maps in the right panels of Figures 1-5. The maps show the status of each household (i.e., whether it is in a hot spot, cold spot, or demonstrates no statistically significant higher or lower MWTP values relative to other nearby households). These designations are based on the estimated G_i^* statistic for each household. The corresponding distributions of the G_i^* estimates across households are shown in Figure 7 for different values of α. This same information is shown in Table 5, which displays the number of households identified as being in a hot or cold spot.

G_i^* is assumed to be normally distributed under the null hypothesis (Getis and Ord, 1992), and so we fail to reject the null hypothesis when $-1.645 \leq G_i^* \leq 1.645$ (i.e., a statistically insignificant result). Such instances on the maps would suggest that there is no clustering of high or low values around the corresponding household. Hot spots (black points on the figures) represent clusters of atypically high MWTP estimates, indicating a MWTP hot spot significant at the 90%, 95%, or 99% level depending on the size of the dot. Cold spots (white points) represent clusters of atypically low MWTP estimates, those with parallel negative G_i^* statistics indicating a MWTP cold spot at the same levels of significance, again varying by size of the white dot.

The hotspot analysis for *Bass* and *Clarity* reveals clusters of high MWTP values among some households in close proximity to the Bay (Figures 1 and 2, respectively). And for water clarity, we find clusters of systematically lower MWTP values among households near notable substitute waterbodies, like the Great Lakes and Finger Lakes in New York. Comparison to the variance-adjusted hotspot analyses when $\alpha = 0$ reveals similar results, but accounting for statistical precision in the underlying estimates reduces the number of households that belong to a statistically significant local cluster, especially for identified hot spots, as shown in Table 5.

In Figure 3, we find scattered hot spots of MWTP for crabs, mainly outside the watershed. There is also a concentration of cold spots within the watershed, mainly in central Virginia and Maryland. The finding that households in closest proximity to the Bay have the lowest values for improvements in crab populations, and those farthest have the highest values, is again surprising, but is in line with the interpolation exercise. This unexpected spatial pattern could be driven, at least partially, by relatively large nonuse values held by nonusers for this iconic resource. Comparing the conventional hotspot analysis $(\alpha = 1)$ to our variance-adjusted hotspot analysis $(\alpha = 0)$ reveals little difference, but substantially reduces the number of cold spots.

Up until this point we have found little evidence of discernible spatial patterns in oyster populations. The hotspot analysis in Figure 4, however, does suggest statistically significant clusters of high MWTP values for increases in oyster abundance, namely among those living closest to the Bay. There is a noticeable pattern of clustered low MWTP values, particularly around New York City and going North along the Hudson River (near the east-most border of New York state). Again, the variance-adjusted hotspot analysis $(\alpha = 0)$ seems to reduce the number of hotspots, but we find slightly increased evidence of cold spots, as reported in Table 5.

The hotspot analyses for improvements in freshwater lakes in the broader Chesapeake Bay Watershed are displayed in the right panel of Figure 5. As one might expect, the conventional

hotspot analysis reveals evidence of a concentration of statistically higher MWTP estimates among households living in the Watershed for improvements in freshwater lakes within the Watershed. There are also a few scattered cold spots, and most notably a concentration of lower MWTP values just outside the southwest corner of the Watershed; perhaps reflecting that there are several substitute lakes in western Virginia that are outside of the watershed. As suggested by the previous variance-adjusted hotspot analyses, we again see similar patterns in hot and cold spots, but the number of statistically significant local clusters of high MWTP estimates are substantially reduced when accounting for the statistical precision of the first-stage estimates ($\alpha = 0$).

The broader finding that the number of identified clusters are reduced after accounting for the statistical precision of the first-stage estimates is better demonstrated by the distributions of the household-specific G_i^* statistics shown in Figure 7. In general, we see that accounting for the statistical precision of the MWTP estimates makes one less likely to identify statistically significant clusters (i.e., a larger portion of the distribution of G_i^* is located towards zero). This finding is consistent with the hotspot maps in Figures 1-6, and makes intuitive sense. Extreme MWTP estimates are often less precise, and so when these estimates are appropriately discounted due to this lack of precision one is less likely to falsely identify a statistically significant cluster of high or low WTP values. When performing scoping exercises like this to try and identify spatial patterns, this application demonstrates that it may be important to account for the fact that these MWTP values are estimates, and not observed values. Not taking into account the precision of the household MWTP estimates may lead researchers to falsely identify patterns of spatial heterogeneity.

The bottom panel in Figure 7 reveals a finding that is unique to the estimated *SQ effects*. Incorporating household-specific variances into the spatial weights shifts the mass of the G_i^* distributions for the MWTP estimates towards zero. However, for the G_i^* statistics corresponding to the SQ effects, we see the distribution shift more negative. In some cases, as we previously saw, this reduces the number of identified hot spots. For example, the conventional hotspot results ($\alpha = 1$) in Figure 7 suggest that respondents near the Finger Lakes, a notable substitute, are more susceptible to exhibiting potentially biasing behaviors against a policy option that improves water quality in the more distant Chesapeake Bay (e.g., "cold feet). But this identified cluster of significantly high SQ effect estimates disappears once the statistical precision of those underlying estimates is accounted for.

The more unique finding is that accounting for the statistical precision behind the estimated SQ effects identifies *more* statistically significant cold spots. In other words, we identify more clusters of households exhibiting "warm glow" or other potentially biasing behaviors in favor of a policy option. For example, in Table 5 we see a 230% increase in the number of households that belong to a SQ effect cold spot when going from $\alpha = 1$ to $\alpha = 0$. This is also evident in the maps in Figure 6. The rightmost hotspot analysis in Figure 6 where $\alpha = 0$ reveals noticeably more cold spots, especially in western Maryland, and the area around New York City, the Long Island Sound, and Narragansett Bay and Cape Cod. Perhaps respondents near these other iconic estuaries have an implicit bias or strategic response that pushes them towards an option that leads to improvements in the Chesapeake Bay. Alternatively, this may reflect preference heterogeneity in

favor of estuary quality that is not otherwise captured by or correlated with other attributes in the experimental design. In any case, these effects are excluded from the policy illustration we discuss next, but the location and clustering of respondents exhibiting such behaviors are important to keep in mind when designing future stated preference studies and could be important to identify prior to specifying more parametric models to estimate welfare changes. For example, the results of our spatial analyses allude to the potential importance of substitute waterbodies. Insights like this are an example of what can be gained from semi-parametric scoping exercises to examine spatial patterns.

4.5. Policy illustration: Chesapeake Bay Total Maximum Daily Loads

To examine the potential policy implications of accounting for spatial heterogeneity in WTP estimates, and the underlying statistical precision of those estimates, we repeat the benefit calculations carried out by Moore et al. (2018). As reported in Table 9 of Moore et al. (2018), the projected improvements from the Chesapeake Bay TMDLs are an average increase of 4.33 inches in Bay water clarity, 1.03 million striped bass, 41 million blue crab, 541 tons of oysters, and 455 freshwater lakes reaching "low algae" status. Following Holmes and Adamowicz (2003), the annual WTP for each household *i* is calculated as the difference of the deterministic component of the indirect utility, divided by the marginal utility of income:

$$WTP_i = \frac{\boldsymbol{\beta_i} ln(\boldsymbol{x^1}) + \boldsymbol{\delta_i} ln(\boldsymbol{x^1}) user_i - [\boldsymbol{\beta_i} ln(\tilde{\boldsymbol{x}}) + \boldsymbol{\delta_i} ln(\tilde{\boldsymbol{x}}) user_i]}{-\gamma} \quad (9)$$

where $\tilde{\boldsymbol{x}}$ is a vector of the baseline attribute levels (Table 1), and $\boldsymbol{x^1}$ are the projected policy levels (baseline levels plus the improvements). The household-specific WTP estimates are derived for all 559 households in the sample. In order to extrapolate the WTP estimates to the population of 44,353,441 households in the study area, we then create interpolated WTP surfaces using the same procedure described in section 2.4.1. We next take the average of the interpolated cells within each census tract and multiply that by the number of households in that tract according to the 2010 U.S. Census. These total WTP estimates for each tract are then summed over all census tracts in the study area. The resulting total annual WTP estimates are displayed in Table 6. We emphasize that our policy illustration is based solely on the spatial interpolations and does not rely on the analyses of global spatial autocorrelation and local clusters.

The first two columns in Table 6 show the total benefit estimates taken from Moore et al. (2018). Their model 1 estimates assume homogeneity across the entire study area by applying a single average WTP estimate to all households in the population. Their model 2 estimate is based on a similar procedure, and although it does not explicitly account for spatial heterogeneity, it does account for heterogeneity regarding the use of the resource and extrapolates those values based on estimates of the proportion of the population that are users versus nonusers. The next four columns in Table 6 show the results of our spatially-explicit extrapolation exercise, and suggest total benefit estimates for the entire study area ranging from $6.6 to $6.9 billion per year. These total benefit estimates are largely in line with those from Moore et al. (2018). This suggests that accounting for spatial heterogeneity may not yield substantial differences in terms of broader policy implications,

at least not in this specific context and when considering the entire study area as a whole. Such spatially explicit details may be important, however, for more local policies. For example, we do find significant variation in household-level WTPs across the study area, ranging from an annual household WTP of $23 to $312.

Comparison of the total WTP estimates from the conventional spatial interpolation exercise ($\rho = 1$) to our variance-adjusted spatial interpolations ($\rho < 1$) suggests that total WTP estimates decrease as more weight is given to the statistical precision of the first-stage estimates. This is consistent with our broader findings that accounting for statistical precision reduces the influence of less precisely estimated outliers that could otherwise unduly influence empirical analyses. In this particular context, however, the differences in total WTP inferred from the spatial interpolations may not be economically significant. Surprisingly, relatively small differences are also revealed when examining total WTP at more local levels, such as by state or county. In fact, even at the individual tract-level, comparing our variance-adjustment estimates when $\rho = 0.5$ to the benefits inferred from conventional interpolation techniques ($\rho = 1$), suggests that the latter leads to only a 10% difference in total tract-level WTP for the majority (90%) of the 27,117 census tracts in the study area. In short, although we find that accounting for spatial heterogeneity is important, the proposed variance-adjustment may not make much of a practical difference in this particular setting.

5. CONCLUSION

We propose a novel extension of existing semi-parametric techniques to analyze spatial patterns when the variables of interest are estimates and not observed values, as is the case in many applications to nonmarket valuation. When examining spatial welfare patterns, we account for the fact that our first-stage model will estimate some households' values less precisely than others. The methodology in this study estimates household-specific MWTP variances using Bayesian estimation techniques and incorporates that information into the spatial weights matrix used in tests for global spatial autocorrelation, spatial interpolation maps, and hotspot analyses. Similar spatial analyses have been increasingly introduced in the nonmarket valuation literature (e.g., Campbell et al., 2009; Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015; Czajkowski et al., 2017), but our study is the first to formally incorporate the statistical precision of the first-stage estimates into the second-stage spatial analyses.

We demonstrate our two-step methodology using a SP study of water quality improvements in the Chesapeake Bay. Accounting for the statistical precision of the MWTP estimates generally seems to result in less statistically significant evidence of spatial patterns, as reflected by tests for global spatial autocorrelation and the hotspot analysis. This tendency increases as additional weight is given to the statistical precision of the MWTP estimates. A similar finding is found with regards to households exhibiting "cold feet", or a tendency to disproportionally favor the status quo. Overall, the analysis suggests that accounting for the statistical precision of the estimated economic phenomena being analyzed reduces the chances of falsely identifying statistically significant spatial heterogeneity. In contrast, we also find that accounting for the statistical

precision of the underlying estimates can lead to increased identification of areas where respondents disproportionally favor a policy option for reasons not explained by the choice attributes. Identifying locations where households exhibit such "warm glow" or other potentially biasing behaviors can aid in future survey design and help inform econometric model specifications to estimate welfare changes.

We estimate the total benefits projected to result from the Chesapeake Bay TMDLs to examine the importance of our extension of traditional spatial analyses from a practical standpoint. We find that in a broader regional setting, at least with our data, the difference between the benefits inferred from traditional spatial interpolation techniques versus those that accommodate for statistical precision are small. More applications of the methods discussed in this study to SP data valuing other environmental amenities are needed to see whether accounting for the statistical precision of the first-stage WTP estimates reveals similar findings, particularly in cases where the environmental amenities of interest are more local in nature. One might not necessarily expect as much spatial heterogeneity in preferences for a well-known iconic resource, like the Chesapeake Bay. Examination of more localized amenities, perhaps where familiarity with the resource is more varied, may yield different findings in how accounting for statistical precision impacts the identification of spatial patterns.

Although our two-step methodology provides an intuitive path for accounting for the statistical precision of the first-stage estimates when conducting spatial analyses, and presumably allows for more accurate identification of spatial patterns, future simulation studies are needed to formally examine the potential improvements in accuracy. Such studies might entail analysis of simulated data where the researcher knows the true data generating process over space. Nonetheless, given the emphasis of these spatial analytic techniques for purposes of data diagnostics and scoping (Johnston and Ramachandran, 2014; Johnston et al., 2015), we encourage researchers to implement our variance-adjustment methods when attempting to identify potential spatial patterns that may not be immediately apparent through conventional parametric models.

REFERENCES

Anselin, L.,1995. Local indicators of spatial association—LISA. Geographical Analysis 27(2): 93–115.

Anselin, L., Gallo, J.L., 2006. Interpolation of air quality measures in hedonic house price models: Spatial aspects. Spatial Economic Analysis 1(1): 31–52.

Badura, T., Ferrini, S., Burton, M., Binner, A., Bateman, I.J., 2020. Using individualised choice maps to capture the spatial dimensions of value within choice experiments. Environmental and Resource Economics 75: 297–322.

Bateman, I.J., Day, B.H., Georgiou, S., Lake, I., 2006. The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP. Ecological Economics 60(2): 450–460.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. Research Synthesis Methods 1(2): 97–111.

Budziński, W., Campbell, D., Czajkowski, M., Demšar, U., Hanley, N., 2017. Using geographically weighted choice models to account for the spatial heterogeneity of preferences. Journal of Agricultural Economics 69(3): 606–626.

Campbell, D., Hutchison, W.G., Scarpa, R., 2009. Using choice experiments to explore the spatial distribution of willingness to pay for rural landscape improvements. Environment and Planning A 41: 97–111.

Czajkowski, M., Budziński W., Campbell, D., Giergiczny, M., Hanley, N., 2017. Spatial heterogeneity of willingness to pay for forest management. Environmental and Resource Economics 68: 705–727.

Getis, A., 2007. Reflections on spatial autocorrelation. Regional Science and Urban Economics 37(4): 491–496.

Getis, A., 2010. Spatial autocorrelation. In: Fischer, M.M., Getis, A. (Eds.). Handbook of Applied Spatial Analysis. Springer, Heidelberg.

Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24(3): 189–206.

Glenk, K., Johnston, R.J., Meyerhoff, J., Sagebiel, J., 2020. Spatial dimensions of stated preference valuation in environmental and resource economics: Methods, trends and challenges. Environmental and Resource Economics 75: 215–242.

Greene, W.H., 2003. Econometric Analysis. New Jersey: Prentice-Hall.

Hanley, N., Schalpfer, F., Spurgeon J., 2003. Aggregating the benefits of environmental improvements: Distance-decay functions for use and non-use values. Journal of Environmental Management 68: 297–304.

Hole, A.R. Kolstad, J.R., 2012. Mixed logit estimation of willingness to pay distributions: A comparison of models in preference and WTP space using data from a health-related choice experiment. Empirical Economics 42(2): 445–469.

Holmes T.P., Adamowicz, W.L., 2003. Attribute-based methods. In: Champ P.A., Boyle K.J., Brown T.C. (eds) A Primer on Nonmarket Valuation. The Economics of Non-Market Goods and Resources, vol 3. Springer, Dordrecht.

Johnston, R.J., Jarvis, D., Wallmo, K., Lew, D.K., 2015. Multiscale spatial pattern in nonuse willingness to pay: Applications to threatened and endangered marine Species. Land Economics 91(4): 739–761.

Johnston, R.J., Ramachandran, M., 2014. Modeling spatial patchiness and hot spots in stated preference willingness to pay. Environmental and Resource Economics 59: 363–387.

Jørgensen, S.L., Olsen, S.B., Ladenburg, J., et al., 2013. Spatially induced disparities in users' and non-users' WTP for water quality improvements—testing the effect of multiple substitutes and distance decay. Ecological Economics 92: 58–66.

Layton, D.F., Brown, G., 2000. Heterogeneous preferences regarding global climate change. Review of Economics and Statistics 82(4): 616–624.

Maddala, G.S., 1983. Limited-dependent and qualitative variables in economics. New York: Cambridge University Press.

Meyerhoff, J., 2013. Do turbines in the vicinity of respondents' residences influence choices among programmes for future wind power generation? Journal of Choice Modelling 7: 58–71.

Moore, C., Guignet, D., Maguire, K., Simon, N., Dockins, C., 2018. Valuing ecological improvements in the Chesapeake Bay and the importance of ancillary benefits. Journal of Benefit-Cost Analysis 9(1): 1–26.

Nelson, T.A., Boots, B., 2008. Detecting spatially explicit hot spots in landscape-scale ecology. Ecography 31(5): 556–566.

Olsen, S.B., Jensen, C.U., Panduro, T.E., 2020. Modelling strategies for discontinuous distance decay in willingness to pay for ecosystem services. Environmental and Resource Economics 75: 351–386.

Ord, J.K., Getis, A., 1995. Local spatial autocorrelation statistics: Distributional issues and an application. Geographical Analysis 27(4): 286–306.

Pate, J., Loomis J.B., 1997. The effect of distance on willingness to pay values: A case study of wetlands and salmon in California. Ecological Economics 20: 199–207.

Revelt, D., Train, K., 2000. Customer-specific taste parameters and mixed logit: Households' choice of electricity supplier. Working paper No. E00-274, Department of Economics University of California Berkeley.

Rolfe, J., Windle, J., 2012. Distance decay functions for iconic assets: Assessing national values to protect the health of the great barrier reef in Australia. Environmental and Resource Economics 53: 347–365.

Scarpa, R., Thiene, M., Train, K., 2008. Utility in willingness to pay space: A tool to address confounding random scale effects in destination choice to the Alps. American Journal of Agricultural Economics 90(4): 994–1010.

Schaafsma, M., Brouwer, R., Gilbert, A., van den Bergh, J., Wagtendonk, A., 2013. Estimation of distance-decay functions to account for substitution and spatial heterogeneity in stated preference research. Land Economics 89: 514–537.

Train, K.E., 2003. Discrete choice analysis with simulation. Cambridge, Cambridge, UK.

Wang, H, Qiu, F., 2017. Investigation of the dynamics of agricultural land at the urban fringe: A comparison of two peri-urban areas in Canada. The Canadian Geographer 61(3): 457–470.

Wang, H., Swallow, B.M., 2016. Optimizing expenditures for agricultural land conservation: Spatially-explicit estimation of benefits, budgets, costs and targets. Land Use Policy 59: 272–283.

TABLES AND FIGURES

Table 1. Attribute Descriptions and Levels

Attribute	Description	Status Quo Level	Post-Policy Levels
Bay Water Clarity	Number of feet below the water surface you can see	3 feet	3; 3.5; 4.5
Striped Bass Population	Number of adult striped bass in the Chesapeake Bay (millions)	24 million	24; 30; 36
Blue Crab Population	Number of adult blue crab in the Chesapeake Bay (millions)	250 million	250; 285; 328
Oyster Abundance	Tons of oysters living in the Chesapeake Bay	3,300 tons	3,300; 5,500; 10,000
Low Algae Lakes	Out of 4,200 freshwater lakes in the Chesapeake Bay Watershed, number with low algae levels	2,900 lakes	2,900; 3,300; 3,850
Annual Cost to Household	Permanent increase in the annual cost of living starting the following calendar year	$0 per year	$20; $40; $60; $180; $250; $500

Table 2. Posterior Distributions of Coefficient Estimates

	Mean	SD	Non-User MWTP
ln(clarity)[a]	0.505	0.548	1.47
ln(bass)	0.874	0.584	3.83*
ln(crab)	2.070***	0.638	0.87***
ln(oyster)	0.198	0.206	0.01
ln(lake)	3.769***	0.824	0.14***
			User MWTP
user × ln(clarity)	1.101***	0.720	4.69***
user × ln(bass)	0.615**	0.863	6.52***
user × ln(crab)	-0.366	0.933	0.72
user × ln(oyster)	0.355	0.320	0.02
user × ln(lake)	1.305*	1.069	0.18**
SQC	-1.938***	0.352	
cost	-0.009***	0.001	

*** p<0.01, ** p<0.05, * p<0.1. All coefficient estimates modelled as random, except the coefficient on cost is treated as fixed to ensure MWTP distributions are finite. Marginal willingness to pay (MWTP) estimates expressed in 2014$. (a) Note that clarity is expressed as inches in the empirical models and subsequent MWTP estimates. All other environmental attributes are expressed in the same units originally specified in the survey, and as reported in Table 1 (i.e., millions of bass, millions of crabs, tons of oysters, and the number of low algae lakes).

Table 3. Distribution of Coefficients and MWTP in the Population

	Mean	SD	Mean Non-User MWTP	Inner 90th Percentile MWTP	
ln(clarity)	0.4327	3.0252	1.38	-5.69	9.52
ln(bass)	0.5519	2.2868	3.84	1.63	6.92
ln(crab)	2.058	0.6847	0.87	0.74	1.03
ln(oyster)	0.2532	1.3006	0.01	-0.01	0.03
ln(lake)	3.6197	1.3837	0.14	0.08	0.24
			Mean User MWTP		
user × ln(clarity)	1.5235	1.6738	4.91	-4.55	14.34
user × ln(bass)	1.0359	1.2955	6.55	0.37	13.35
user × ln(crab)	-0.9166	1.4795	0.72	0.44	0.99
user × ln(oyster)	0.0865	0.6691	0.02	-0.08	0.26
user × ln(lake)	1.4167	2.1951	0.17	0.07	0.27

Marginal willingness to pay (MWTP) estimates expressed in 2014$. MWTP for an increase in Bay water clarity is expressed in inches. MWTP for all other attributes are expressed in the same units originally specified in the survey, and as reported in Table 1 (i.e., millions of bass, millions of crabs, tons of oysters, and the number of low algae lakes).

Table 4. Moran's *I* Tests for Global Spatial Autocorrelation

	α = 1.0		α = 0.5		α = 0.0	
	Moran's *I*	z-score	Moran's *I*	z-score	Moran's *I*	z-score
Clarity	-0.0047	-0.1408	-0.0056	-0.1824	-0.0058	-0.1947
Striped Bass	0.0459	2.3157**	0.0419	2.1029**	0.0355	1.7525*
Blue Crab	0.0577	2.8897***	0.0533	2.6599***	0.0476	2.3411**
Oysters	0.0305	1.5682	0.0300	1.5368	0.0294	1.4872
Lakes	0.0194	1.0276	0.0163	0.8737	0.0131	0.7116
SQ Effect	0.0726	4.0285***	0.0705	3.692***	0.0708	3.2323***

*** p<0.01, ** p<0.05, * p<0.1

Table 5. Number of Households Identified as Being in a Hot or Cold Spot

	Numbers of Hot Spots			Numbers of Cold Spots		
	α = 1.0	**α = 0.5**	**α = 0.0**	**α = 1.0**	**α = 0.5**	**α = 0.0**
Clarity	24	21	16	20	21	20
Striped Bass	62	42	33	1	2	1
Blue Crab	34	41	46	46	35	20
Oysters	49	38	27	23	24	31
Lakes	47	35	16	21	19	23
SQ Effect	57	37	24	52	132	172

The displayed counts show the number of households (out of the sample of 559) that are identified as being part of a spatial cluster of statistically higher (hot spot) or lower (cold spot) values. Hot spots are those with $G_i^* \geq 1.645$ and cold spots are those with $G_i^* \leq -1.645$.

Table 6. Total Annual Willingness to Pay for Improvements under Total Maximum Daily Loads (2014$, billions)

Moore et al. (2018)		Spatial Interpolation in this Study			
Model 1	Model 2	ρ = 1.00	ρ = 0.75	ρ = 0.50	ρ = 0.25
$6.813	$6.488	$6.870	$6.790	$6.711	$6.635

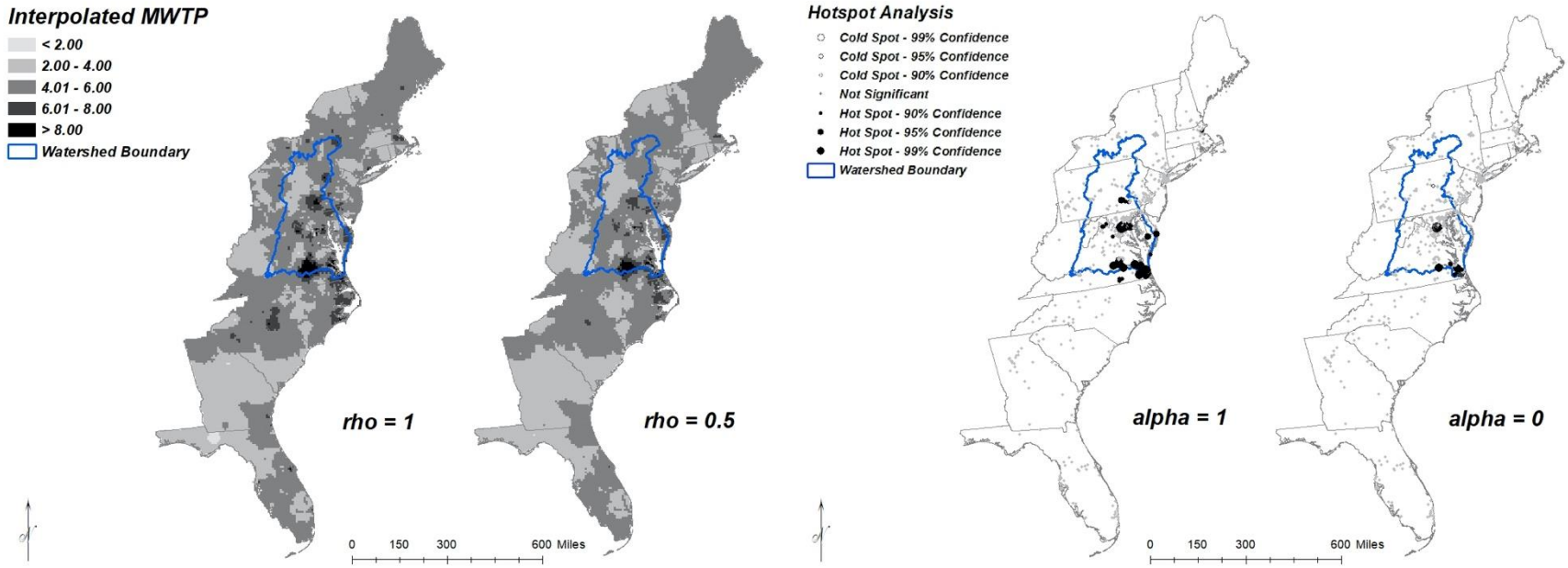

Figure 1. MWTP for Bass

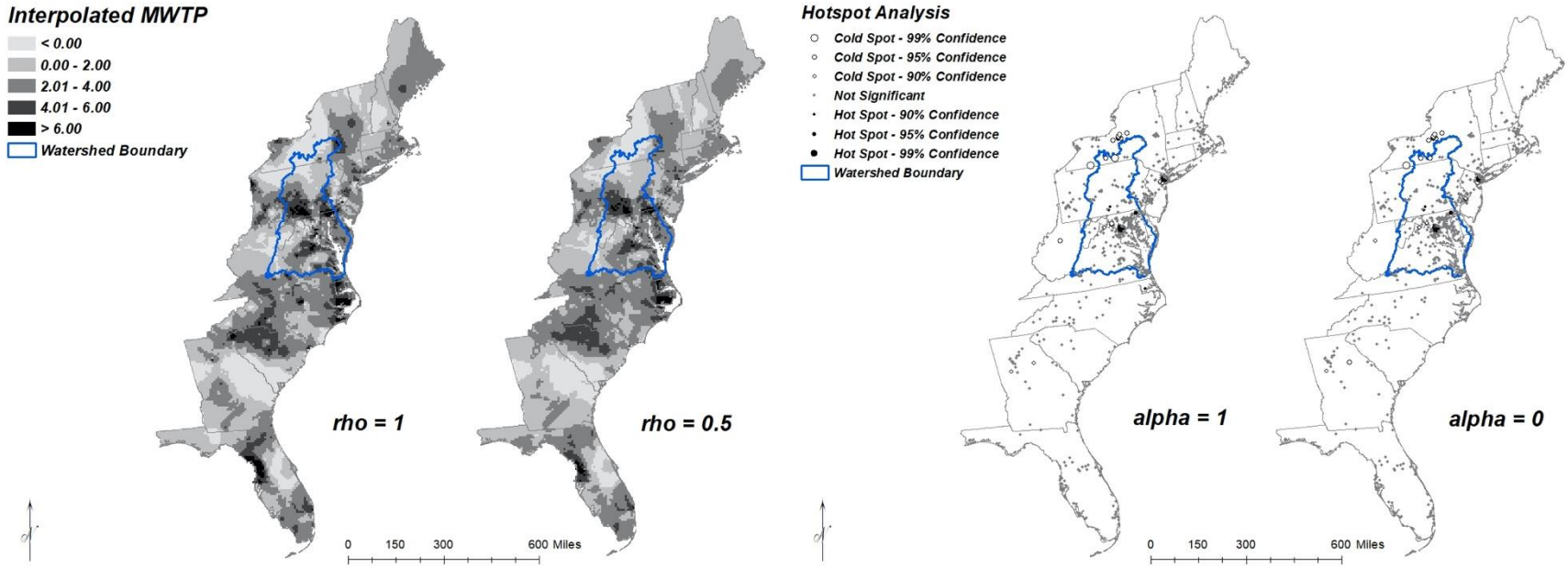

Figure 2. MWTP for Clarity

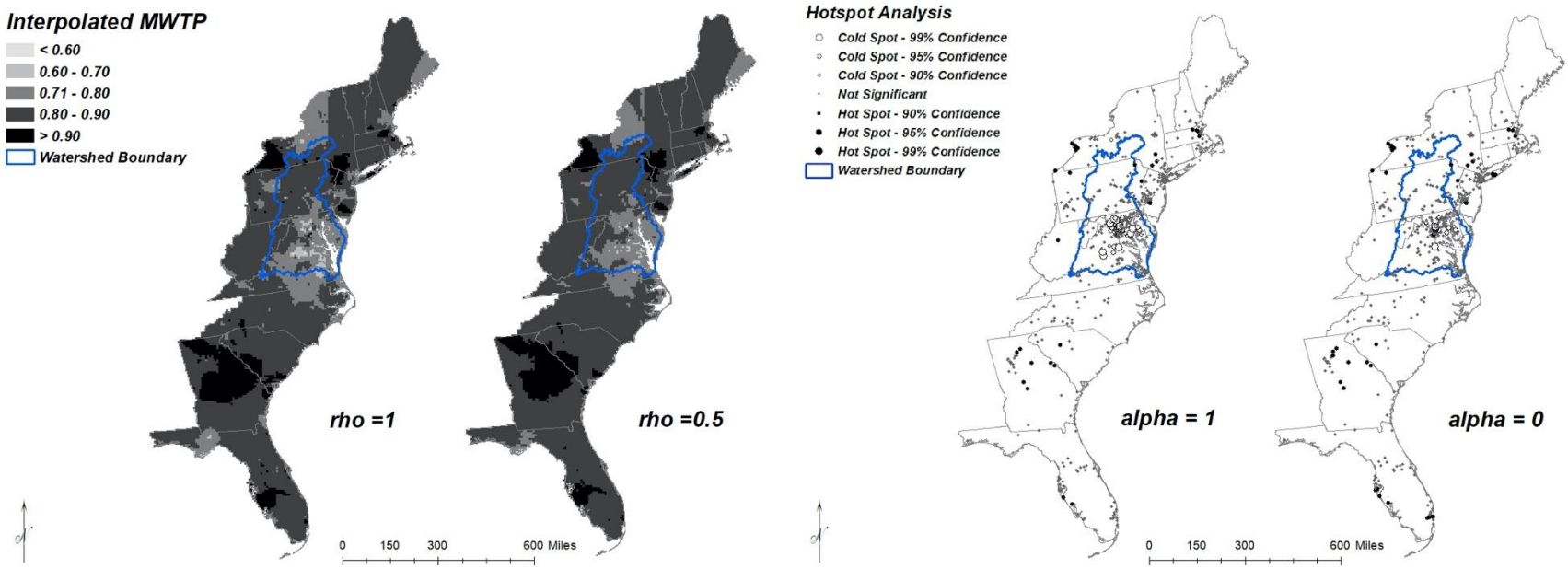

Figure 3. MWTP for Crabs

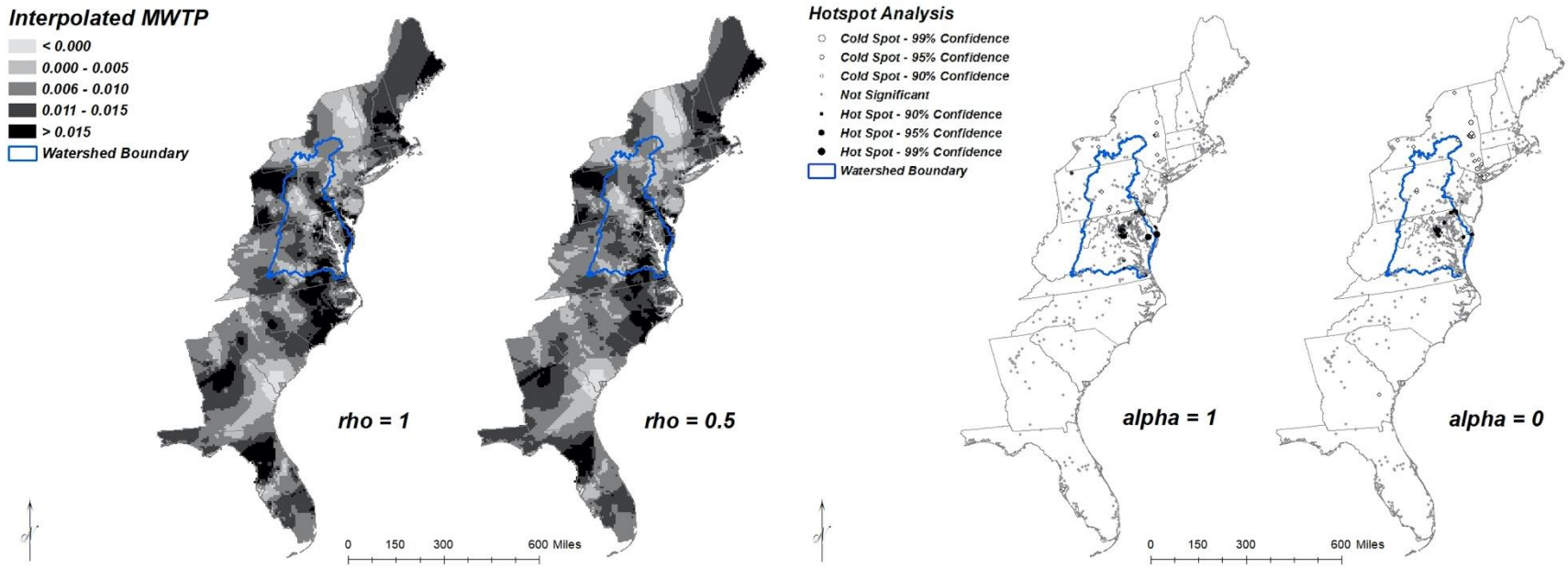

Figure 4. MWTP for Oysters

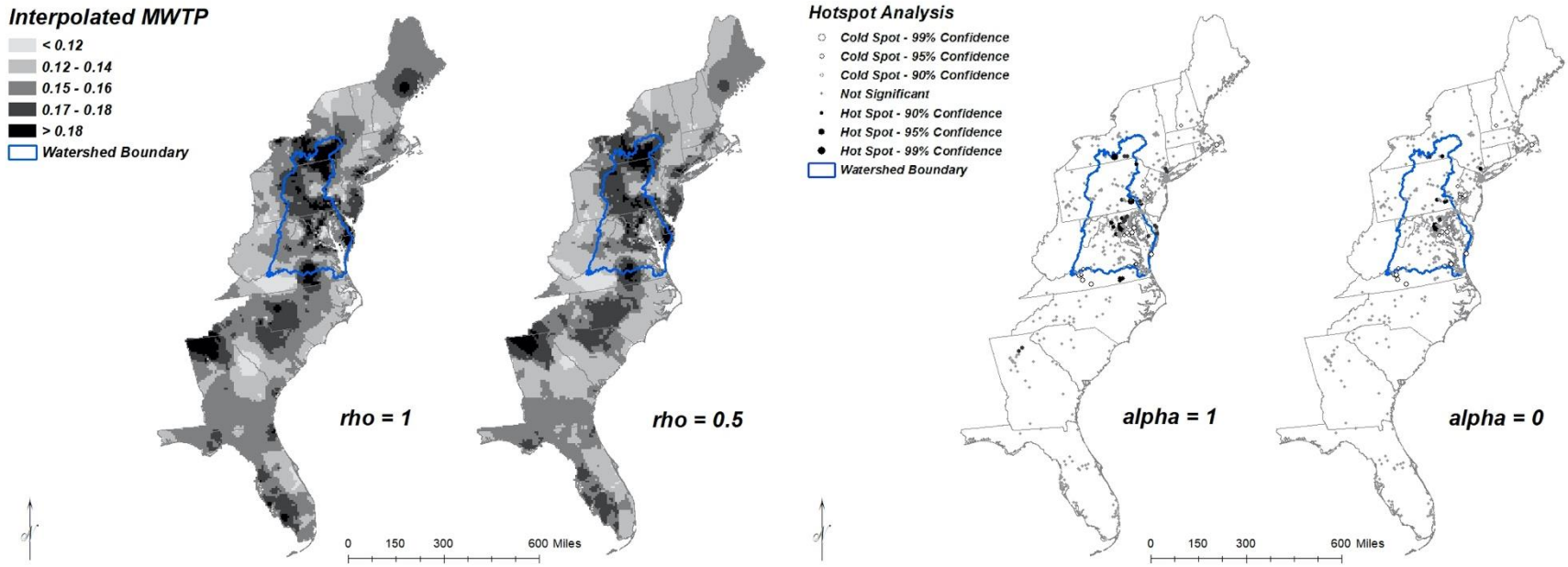

Figure 5. MWTP for Lakes

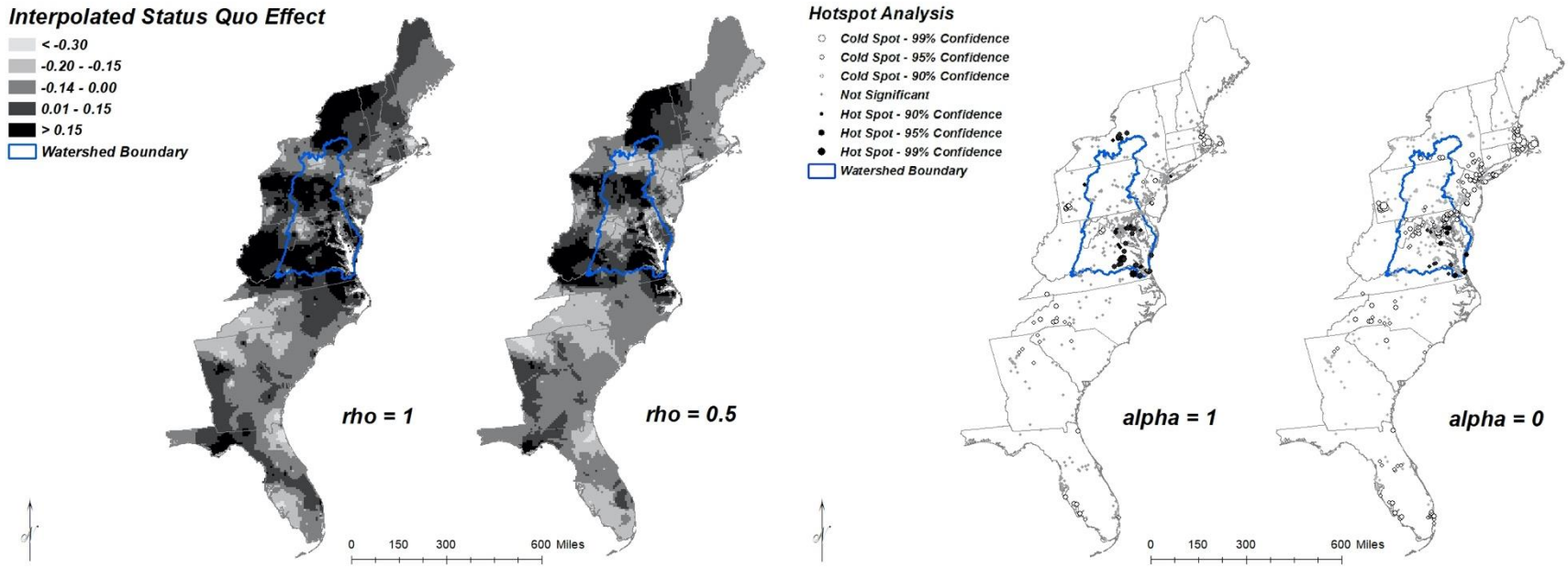

Figure 6. Status Quo Effects

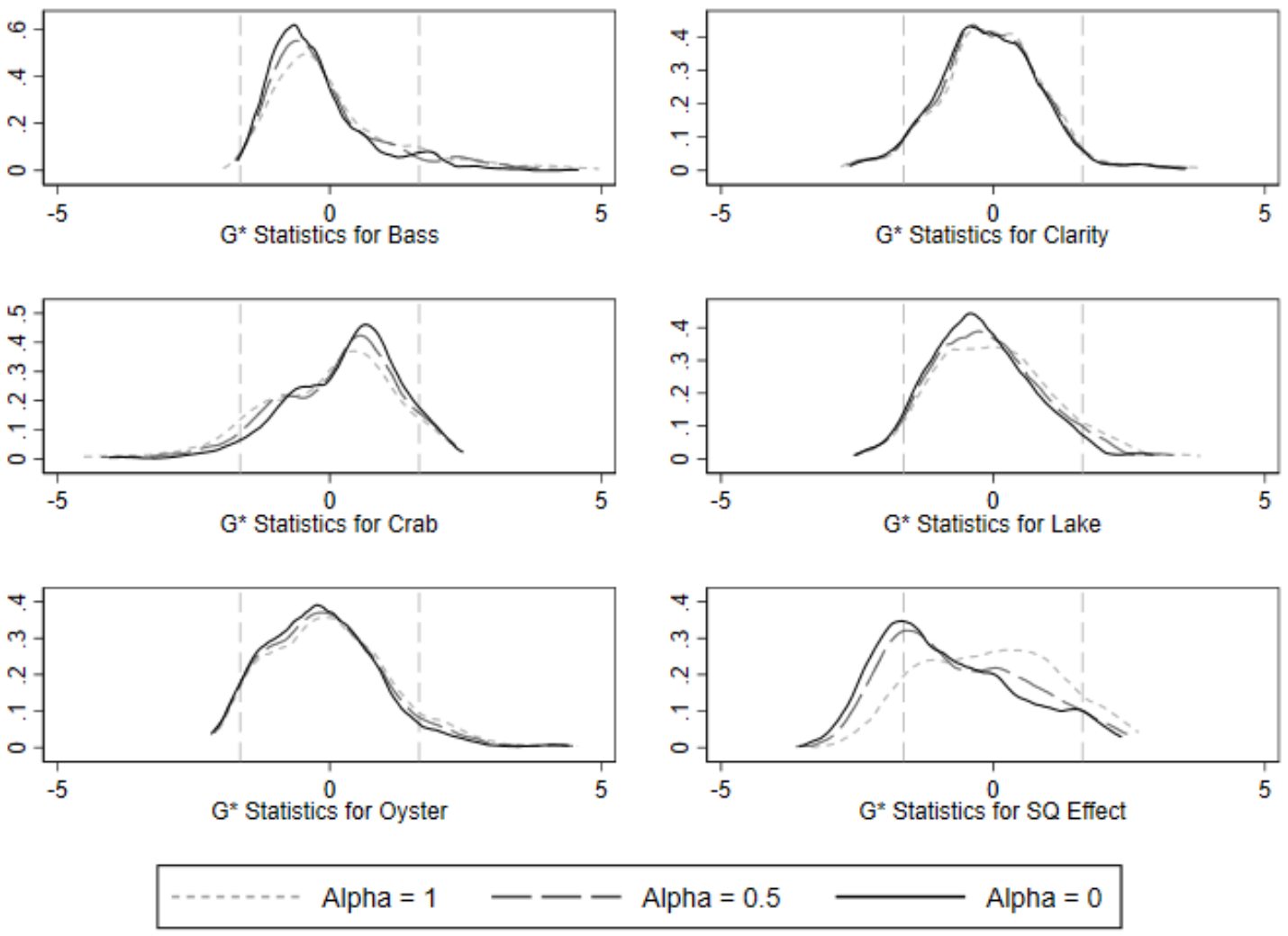

Figure 7. Distribution of Hotspot Analysis Getis-Ord Statistics

Note: Vertical long-dashed grey lines denote 90% confidence interval.

APPENDIX

Appendix A: Description of Bayesian Estimation Routine

We can appeal to Bayes' Rule to define the relationship between the conditional distribution $\eta(\boldsymbol{\lambda}|y,z,\upsilon)$ and the distribution over the population $g(\boldsymbol{\lambda}|\upsilon)$. First consider the probability for the mixed logit:

$$P\left(y_i \mid z_i, \upsilon\right) = \int P\left(y_i \mid z_i, \lambda\right) g\left(\lambda \mid \upsilon\right) d\lambda \qquad \text{(A1)}$$

which gives the probability of household *i*'s set of responses, given the data and the parameters of the population level parameter distribution. Using Bayes Rule, we express $\eta(\boldsymbol{\lambda}|y,z,\upsilon)$ as $\frac{P\left(y_i \mid z_i, \lambda\right) g\left(\lambda \mid \upsilon\right)}{P\left(y_i \mid z_i, \upsilon\right)}$ and since the denominator is constant with respect to $\boldsymbol{\lambda}$, $\eta(\boldsymbol{\lambda}|y,z,\upsilon)$ is proportional to the numerator which provides a useful interpretation of $\eta(.)$. The density of $\boldsymbol{\lambda}$ in the subpopulation that chose y_i when faced with z_i is proportional to the density of $\boldsymbol{\lambda}$ in the entire population, given by $g(.)$, multiplied by the probability that someone would choose y_i given the data and the set of parameters $\boldsymbol{\lambda}$.

Estimating the mixed logit model via hierarchical Bayes' requires iteratively drawing from the distributions for the mean and variance of $\boldsymbol{\lambda}$ and the household-level parameters while always conditioning on the most recent draw of the other parameters. We refer the reader to Train (2003, pages 302-308) for a complete description of the algorithm but for our purposes, consider a simulation that begins with starting values for the mean vector of $\boldsymbol{\lambda}$, a covariance matrix W, and a vector of household-level parameters for each respondent $\boldsymbol{\lambda}_i$. The first step of the algorithm draws a realization of the mean vector $\bar{\lambda}$ conditional on $\boldsymbol{\lambda}_i$ and W which is distributed $N\left(\sum_i \frac{\lambda_i}{N}, W\right)$. The second step draws the covariance matrix W from an inverse Wishart distribution with M+N degrees of freedom and scale matrix $\frac{MI+N\bar{S}}{M+N}$, where I is a M-dimensional identity matrix and $\bar{S} = \frac{\sum_i \left(\lambda_i - \bar{\lambda}\right)\left(\lambda_i - \bar{\lambda}\right)'}{N}$. The third and final step of the algorithm draws household-level parameter vectors from a density proportional to $\prod_t \frac{e^{\lambda_i' z_{iy_{it}t}}}{\sum_i e^{\lambda_i' z_{it}}} \phi\left(\lambda_i \mid \bar{\lambda}, W\right)$ which requires a MH algorithm.

After a burn-in period, the draws will converge to the joint posterior distribution of the model parameters.

Appendix B: Parametric Estimation of Distance Decay

As a preliminary step to examine whether preferences for environmental improvements vary with distance to the resource, researchers often include a measure of the distance of respondent *i* to the resource (e.g., Pate and Loomis, 1997; Hanley et al., 2003; Bateman et al., 2006). We denote this distance measure as $f(d_i)$, and examine various functional forms for such a global distance gradient. The model in equation (1) of section 2.1 in the main text is thus augmented as follows:

$$P_i(j \mid \boldsymbol{x}_q, cost_q, SQ_q, d_q) = \frac{exp\{ln(\boldsymbol{x}_j)\boldsymbol{\beta}+[ln(\boldsymbol{x}_j)\times f(d_i)]\theta+\gamma cost_j+\varphi SQ_j\}}{\sum_q exp\{ln(\boldsymbol{x}_q)\boldsymbol{\beta}+[ln(\boldsymbol{x}_q)\times f(d_i)]\theta+\gamma cost_q+\varphi SQ_q\}} \quad \text{(B1)}$$

In these preliminary regression models, equation (B1) is estimated as a conditional logit model. The added parameter vector θ will reflect any spatial heterogeneity in respondent's preferences for improvements in the environmental commodity over space.

If estimates of θ are statistically significant, then it would suggest that preferences, and hence subsequent MWTP calculations, vary with distance. If we are unable to find a specification of $f(d_i)$ that yields statistically significant estimates of θ, then it is reasonable to conclude that MWTP does not change in a smooth, parametric way over space. Such a finding motivates our more flexible semi-parametric examination of spatial heterogeneity in MWTP.

The results of a series of conditional logit models following equation (B1) are estimated using different functional form assumptions for $f(d_i)$, including linear distance, the natural log of distance, inverse distance, a quadratic specification, and various stepwise functions (e.g., within versus outside a 50 kilometer buffer, within versus outside of the watershed, and based on three geographic strata in the experimental design of the survey application). The results are presented in Table B1 below. The individual coefficient results are difficult to interpret given the numerous interaction terms, and in any case, are not of primary interest here.

The results of interest pertain to the statistical significance of the coefficients, or lack thereof, in the lower panel of Table B1. These estimates correspond to θ in equation (B1). As can be seen, the estimates are often statistically insignificant, both individually and jointly. This demonstrates that, at least under the lens of this conventional parametric approach, there is little evidence of spatial heterogeneity in MWTP. Researchers and practitioners may often stop here and assume spatial homogeneity in the household benefits going forward. Such a conclusion, however, may be premature. Statistically significant spatial heterogeneity can still be identified using the two-step semi-parametric approaches we discuss in the main paper.

Table B1. Conditional Logit Models with Parametric Distance Gradient

VARIABLES	Linear distance (1)	ln(distance) (2)	Inverse distance (3)	Quadratic distance (4)		Stepwise: 50km (5)	Stepwise: In watershed (6)	Stepwise: Geographic Strata (7)	
ln(clarity)	2.336e-01	1.973e-01	2.589e-01	1.746e-01		3.183e-01	3.367e-01	7.411e-03	
	(0.353)	(0.934)	(0.259)	(0.466)		(0.287)	(0.336)	(0.430)	
ln(bass)	-5.943e-01	-1.388e+00	-2.395e-01	-8.126e-01*		-2.248e-02	-1.869e-02	-6.930e-01	
	(0.371)	(0.982)	(0.276)	(0.486)		(0.310)	(0.363)	(0.447)	
ln(crab)	-1.859e-01	-5.876e-01	3.363e-01	-2.027e-01		4.288e-01	7.597e-01	-5.320e-01	
	(0.519)	(1.386)	(0.387)	(0.691)		(0.428)	(0.480)	(0.636)	
ln(oyster)	-1.131e-01	-7.380e-02	-5.586e-02	-7.431e-02		-6.705e-02	-1.439e-03	-1.870e-02	
	(0.127)	(0.335)	(0.095)	(0.166)		(0.107)	(0.123)	(0.150)	
ln(lake)	1.160e+00**	4.660e-01	1.573e+00***	1.359e+00**		1.883e+00***	1.788e+00***	8.586e-01	
	(0.492)	(1.423)	(0.358)	(0.662)		(0.386)	(0.451)	(0.613)	
SQC	-8.971e-01***	-1.293e+00***	-7.081e-01***	-1.051e+00***		-5.598e-01***	-5.090e-01***	-1.063e+00***	
	(0.178)	(0.432)	(0.137)	(0.227)		(0.151)	(0.172)	(0.215)	
cost	-5.008e-03***	-5.009e-03***	-5.045e-03***	-5.005e-03***		-5.035e-03***	-5.004e-03***	-5.064e-03***	
	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	
				distance	**distance^2**			**watershed states**	**other east coast states**
ln(clarity) × f(dist)	1.591e-05	1.022e-02	-1.376e-02	4.576e-04	-4.662e-07	-3.273e-01	-2.313e-01	6.531e-01	1.747e-02
	(0.001)	(0.178)	(0.474)	(0.003)	(0.000)	(0.668)	(0.529)	(0.628)	(0.639)
ln(bass)× f(dist)	1.257e-03	2.352e-01	3.654e-01	3.060e-03	-1.689e-06	-1.010e+00	-5.013e-01	3.860e-01	1.242e+00*
	(0.001)	(0.191)	(0.577)	(0.003)	(0.000)	(0.662)	(0.553)	(0.626)	(0.727)
ln(crab) × f(dist)	1.862e-03	1.920e-01	7.844e-01	1.956e-03	3.062e-08	-2.114e-01	-9.659e-01	6.992e-01	2.407e+00**
	(0.001)	(0.269)	(0.589)	(0.004)	(0.000)	(0.950)	(0.794)	(0.875)	(1.005)
ln(oyster) × f(dist)	2.262e-04	4.815e-03	2.075e-01	-1.094e-04	3.101e-07	9.369e-02	-1.228e-01	-2.197e-01	2.024e-01
	(0.000)	(0.065)	(0.161)	(0.001)	(0.000)	(0.224)	(0.190)	(0.223)	(0.230)
ln(lake) × f(dist)	1.551e-03	2.283e-01	3.689e-01	-1.544e-04	1.592e-06	-1.455e+00	-5.109e-01	6.058e-01	2.075e+00**
	(0.001)	(0.273)	(1.147)	(0.004)	(0.000)	(0.977)	(0.739)	(0.841)	(0.938)
Joint significance	$\chi2(4) = 4.64$	$\chi2(4) = 1.94$	$\chi2(4) = 6.30$	$\chi2(8) = 5.51$		$\chi2(4) = 2.56$	$\chi2(4) = 2.78$	$\chi2(8) = 12.40$	
	p = 0.3267	p = 0.7464	p = 0.1779	p = 0.7014		p = 0.6334	p = 0.5958	p = 0.1341	

SQC × f(dist)	6.697e-04	1.189e-01	6.904e-02	1.896e-03	-1.130e-06	-7.387e-01**	-4.840e-01*	5.166e-01*	6.248e-01*
	(0.000)	(0.084)	(0.100)	(0.001)	(0.000)	(0.304)	(0.260)	(0.296)	(0.330)
Observations	4,719	4,719	4,719	4,719		4,719	4,719	4,719	
LL	-1567.4058	-1569.1278	-1568.1169	-1564.6207		-1566.0111	-1568.7936	-1555.4703	

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.