Department of Economics Working Paper

Number 22-06 | August 2022

Water Quality and Hedonic Models: A Meta-Analysis of Commodity, Market, and Methodological Characteristics

Matthew T. Heberling
US Environmental Protection Agency, Office of Research and Development

Dennis Guignet
Appalachian State University

Michael Papenfus
US Environmental Protection Agency, Office of Research and Development

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Water Quality and Hedonic Models: A Meta-Analysis of Commodity, Market, and Methodological Characteristics

Matthew T. Heberling[a,*]; Dennis Guignet[b]; and Michael Papenfus[c]

[a]US Environmental Protection Agency, Office of Research and Development, Cincinnati, OH, 45268, USA.

[b]Appalachian State University, Department of Economics, Boone, NC, 28608-2051, USA.

[c]US Environmental Protection Agency, Office of Research and Development, Corvallis, OR, 97330, USA.

*Corresponding author: M. T. Heberling, US EPA/ORD/CEMM (MS 587), 26 W MLK Dr., Cincinnati, OH 45268, USA. Telephone: +1.513.569.7917; Email: Heberling.Matt@epa.gov

ABSTRACT: This study quantitatively reviews the hedonic literature examining surface water quality to assess how attributes of the commodity, housing market, and methodological choices affect the significance and expected sign of the estimated property value effects. Using meta-analysis, we provide evidence that many of the definitions and decisions, including type of waterbody, water quality categories, and the region of the United States, made in primary studies do affect the estimated relationship between water quality and home prices. Methodological choices appear to have a critical role in determining the estimated relationships. Our findings can inform future hedonic study designs, help identify potential concerns with data and modeling choices, and guide decision-makers when considering what studies to use to inform management and policy decisions.

Keywords: Hedonic, Water quality, Meta-analysis, Property value, Water pollution

1. Introduction

The number of hedonic property value studies continues to grow quickly as data on sales and house attributes, including surrounding environmental conditions, become more readily accessible (Bishop et al. 2020; Petrolia et al. 2021; Guignet and Lee 2021). This growth in the literature provides an opportunity to analyze how environmental commodity definition, market characteristics, and methodological choices affect the results of the hedonic property value model.

Our synthesis focuses on hedonic models that examine the capitalization of water quality in surrounding housing values to consider whether the literature generally supports a significant and theoretically consistent relationship. Although several literature reviews of the hedonic property value literature exist (e.g., Boyle and Kiel 2001; Crompton 2001; Kiel 2006; Wilde et al. 2012), only Nicholls and Crompton (2018) review the hedonic literature focused specifically on water quality. Our study goes beyond a narrative review by using meta-analysis to quantitatively assess the literature based on the observations gathered from 29 unique hedonic studies. Guignet et al. (2022) was the first meta-analysis of the hedonic literature examining surface water quality, but their objective was in improving benefit transfer, and their meta-regression analysis focused exclusively on a subset of 18 hedonic studies specifically on water clarity (i.e., Secchi disk depth). We build on Guignet et al.'s work by examining hedonic studies that used any objective measure of water quality, not just clarity, and by focusing on how the study design and methodological choices affect the likelihood of finding a significant and theoretically consistent relationship.

Meta-analysis uses a variety of statistical approaches to analyze previously reported scientific results and draw broader conclusions (Stanley 2001; Nelson and Kennedy 2009; Stanley and Doucouliagos 2012). The conclusions depend on the purpose of the meta-analysis, for which there are generally two possible objectives. A meta-analysis can be used for purposes of benefit transfer, or to draw inferences from the collective body of literature (Boyle and Wooldridge, 2018). Our focus for this study is on the latter.

One may generally expect water quality to be capitalized into waterfront home values. However, these estimated capitalization effects may vary not only due to variations in the commodity itself, but also variation in study design and methodological choices. A meta-regression model allows one to systematically examine how such variation impacts the primary study results. Often the dependent variable in a meta-regression model is the estimated effect size or summary statistic of interest from the primary study. The inclusion of various independent variables representing different study characteristics such as methodology, data, and functional form, allow one to identify the effect of these choices on the results (Stanley 2001). The dependent variable can also be a binary variable that represents whether the effect size is significant and has a theoretically consistent or expected sign (Smith and Huang 1993; Kiel and Williams 2007). The results of meta-analyses with such binary outcomes can inform future hedonic studies, help identify potential concerns with data or modeling choices, support expectations based on economic theory, and inform decision-makers to the quality of the results (Smith and Huang 1993; Kiel and Williams 2007; Nelson and Kennedy 2009; Stanley and Doucouliagos 2012).

Insight for this paper was provided by Smith and Huang (1993), who estimated meta-regressions examining the effects of air pollution on property values, Kuminoff et al. (2010) who examined how omitted variables affect hedonic models, and three recent "best practices" articles (Taylor 2017; Bishop et al. 2020; Guignet and Lee 2021). Smith and Huang (1993), who had the same objectives as this paper, evaluated how study characteristics related to finding a negative and theoretically consistent relationship between air pollution and property sales. They found that the use of actual sales price (as opposed to census data or appraisal values), estimating a linear model, and using more than one air pollution measure all decreased the likelihood of finding a significant and theoretically expected result. Market conditions, as measured by home vacancy rates, was also included in their meta-analysis. Kuminoff et al. (2010: p. 157) ran an internal meta-analysis with their set of simulations to assess how omitted variables affect marginal implicit price estimates. They suggest controlling for functional form, sample size, whether the primary study controls for omitted variables, and whether the study estimates time-constant implicit prices, are important when estimating a meta-regression.

The three "best practices" articles play an important role in ensuring we include recommended modeling choices in our meta-analysis and allow us to compare the differences in results to those recommended approaches. The purpose of our meta-analysis is not to promote hedonic modeling choices that lead to significant and expected results; rather, we intend to provide practitioners and decision-makers with information to aid in the interpretation and application of hedonic model estimates.

2. Meta-dataset

Our synthesis of surface water quality and housing values is based on a meta-dataset developed by Guignet et al. (2022). The complete meta-dataset contains a comprehensive set of hedonic property value studies that examined surface water quality in the US and were published or released between 1979-2017. Details on the development of the meta-dataset, including the search protocol and summary statistics, can be found in Guignet et al. (2022) and their online supplementary material. The meta-data are publicly available at US EPA's Environmental Dataset Gateway (https://doi.org/10.23719/1518489).

We started with the 36 primary studies in the meta-dataset that examined surface water quality in the US using objective water quality measures, but we focus on only 29 primary studies and the 290 unique house price elasticity estimates that correspond to waterfront homes.[1] A variety of water quality measures have been examined in the hedonic literature; 17 different measures are observed in our meta-dataset. For each measure, we have an estimate of the house price elasticity with respect to water quality, as well as an estimated standard error. Some of the elasticities represent improvement in water quality, while the majority of elasticities relate to water quality measures where a higher value denotes a degradation (e.g., higher fecal coliform indicates lesser water quality). The elasticities for Secchi disk depth, percent water visibility, and dissolved oxygen (DO) are all generally expected to be positive because increases in these three measures are often considered improvements. Table S1 lists the water quality measure, expected sign, study citation, and the number of observations in our final meta-dataset.

[1] We drop observations from Walsh et al. (2017) and Guignet et al. (2017) that were based on Secchi disk depth, because such observations are redundant with those pertaining to light attenuation, which are maintained for our meta-analysis. We also drop an additional six observations where average home price is missing and ten observations from studies that analyzed the effect of pH on home prices. The estimated house price elasticities corresponding to pH range from -0.82 to 12.80. We drop pH because it is difficult to identify what constitutes a degradation or improvement in water quality and those observations were often extreme outliers.

Using the standard errors derived and included in the meta-dataset, we calculate whether the estimated elasticity for each meta-observation is significantly different from zero. With the expected sign of the elasticity based on the water quality measure and the estimated standard errors, we characterize whether an elasticity is significant and theoretically consistent (e.g., Smith and Huang 1993). The conversion of water quality elasticities into a binary variable, Y, does lose information about the size of the effect. For example, studies with small and large positive effects are treated equally.[2] Nonetheless, we focus on this coarser binary dependent variable for two reasons. First, the size of the effect has already been analyzed by Guignet et al. (2022), at least with respect to water clarity. Second, it is unclear whether pooling and analyzing price elasticities with respect to very different water quality measures is appropriate. A one-percent change in Secchi disk depth is very different than a one-percent change in the concentration of nitrogen or fecal coliform count, for example. Pooling the meta-data across water quality measures and examining whether the primary study elasticity estimates were statistically significant and of the theoretically expected sign does not present such concerns.

The dependent variable in our main model uses $p < 0.05$ from a standard two-tailed t-test to determine if the estimated elasticity is statistically significant.[3] We have 161 observations (56%) that are statistically significant elasticity estimates (i.e., with a p-value less than 0.05), regardless of the expected sign. About 52% of the estimated elasticities are considered significant *and* have the expected sign (151 observations). There are 129 observations that are statistically equal to

[2] As the literature continues to grow, it may be important to focus on whether a theoretically consistent and significant relationship exists, or if the translation of the elasticities to a binary dependent variable artificially creates the relationship, as suggested by Stanley and Doucouliagos (2012: p. 16), through the publication selection process.
[3] We later examine the robustness of our results to alternative cutoffs for determining statistical significance.

zero (83 of those insignificant observations have the expected sign and 46 do not). The primary objective of our quantitative review is to examine under what conditions the hedonic model tends to yield the expected results, focusing specifically on the commodity attributes, market characteristics, and methodological variables discussed next.

3. Commodity, Market, and Methodological Characteristics

While the definition of the water quality commodity and the market definition are often chosen by the researcher to match the study objectives and are sometimes constrained by the availability of both water quality and property sales data, the methodological and estimation choices made by researchers can have an important influence on a given study's results. Including methodological characteristics in the meta-analysis is motivated in part to help understand the importance of these decisions.

Descriptive statistics can be found Table 1. Several of the variables are binary indicators or dummy variables, and in such cases the mean describes the percentage of observations where that variable equals one. Figures presenting significance and theoretical consistency for some of the variables can be found in the online supporting material. The 29 studies included in the meta-data were published between 1985-2017 and provided elasticity estimates ranging from -2.64 to 8.32. Generally, more water quality hedonic studies were published later in this time period, but in 2000 and 2007, we see a larger number of observations. The number of significant elasticities with the expected sign certainly varies from year to year (see Figure S1 in the supplementary material).

3.1. Environmental commodity

To synthesize the water quality hedonic literature, we begin with the environmental commodity (i.e., the type of waterbody and water quality measure examined in the primary studies). The hedonic property value studies focusing on surface water quality tend to examine lakes or reservoirs (57%) as a group compared to estuaries or rivers. Hedonic studies related to lakes/reservoirs also have a higher frequency of elasticities that are significant and meet expectations, as compared to estuaries and rivers (see Figure S2 in the supplementary material).

For tractability, we organized the 17 different objective water quality measures into five broader categories – Clarity, Nutrients, Sediment, Bacteria, and Biochemical (see Table 2). Some categories are more likely to be directly observed by homebuyers, and others could be considered proxies for perceived water quality. Our categorization is intended to group measures that reflect similar water quality issues and processes, and that also may be perceived similarly by homebuyers and sellers. Water clarity is the most common measure in this meta-dataset (62%). Bacteria and nutrients combined make up a much smaller proportion of the meta-dataset, while biochemical water quality measures and sediment have the fewest observations.

3.2. Study Area and Housing Market Characteristics

Most elasticities were estimated for housing markets and waterbodies in the south (39%), followed by the northeast region, midwest and west. The south region has the largest number of elasticities that are significant and have the expected sign, while the west region has very few. (Figure S3 in the supplementary material).

The characteristics defining the assumed housing market in a primary study may play an important role in whether a study yields the expected result. Bishop et al. (2020), Taylor et al. (2017), and others emphasize the need to define a market, both geographically and temporally, by the "law of one price" – meaning that within the assumed market identical housing bundles will sell for the same price. In other words, a single hedonic equilibrium price surface should apply throughout the entire housing market, and that equilibrium is not changing over the assumed time period and spatial definition of that market. The housing market characteristics considered in our meta-analysis are the average house price, the spatial and temporal definitions of the market (i.e., whether multiple counties were pooled together and the number of years in the study period, respectively), and whether the sample years include the 2006-2009 housing market bubble burst. The market bubble issue has received significant attention by researchers (e.g., Boyle et al., 2012).

Using the consumer price index (CPI), we update the average house price to 2018$ in each study based on the reported year or, if not reported, the last year of the sample. We used the last year of the sample for 98 of the 290 observations.

To assess the effects of spatial definitions on elasticities, we initially identify whether observations defined a market as a subcounty area, multiple subcounties, or multiple counties. No study in this literature was based on all property sales from a single county. Hedonic property value studies examining water quality almost always focus on homes within some distance of the waterbody, and do not utilize the entire set of transactions in a county. We characterize such market definitions as subcounty. Subcounty samples are constrained to a county, but do not include all properties from that county (e.g., only sales in a county that are within a certain distance to a waterbody). Most observations were at the subcounty level (73%). We combined multiple counties and multiple subcounties as the other market definition (Table 1).

In order to examine the temporal definition of the market, we use the number of years in the sample. Our interest lies in whether hedonic results vary when estimated from sales data over longer study periods, and hence where the "law of one price" assumption is less plausible. The variable *Sample Years* has noticeable variation, ranging from 1 to 24 years. As suggested by Figure S4 in the supplementary material, the number of elasticities that are significant and of the expected signs tends to be relatively higher when estimated from studies based on between five and ten years of transaction data. Starting around 12 years, however, we see that trend change. Studies covering periods longer than 12 years tend to produce a relatively lower number of results that are significant and of the expected sign.

Economists have discussed the implications of the 2006-2009 housing market bubble and burst on both hedonic methods and the interpretation of results (e.g., Boyle et al. 2012; Taylor 2017; Bishop et al. 2020). The structural shifts that occurred during the housing market bubble and subsequent burst clearly affected market equilibriums. Hedonic models that are estimated with samples that included transactions both pre- and post- the market bubble burst, and that do not properly allow the entire hedonic surface to shift with that new equilibrium in their models, violate the "law of one price" and are theoretically invalid (Bishop et al., 2020). Defining the burst from 2006 to 2009, which matches Taylor's (2017) definition, we create a dummy variable

(*Bubble*) that is equal to 1 if any of the sample years include the burst (about 37% of the observations). Elasticities more often tend to be significant and of the expected sign when primary studies do not include any years during the 2006-2009 bubble burst (see Figure S5).

3.3. Methodological decisions

Independent variables that characterize methodological decisions can also lead to variation in the significance and the expected sign of elasticity estimates. The methodological variables examined include choices of functional form, methods to account for spatial dependence, whether actual transaction prices are used, and decisions about the water quality data and variables included in the hedonic models.

The choice of functional form leads to different interpretations of the coefficients, but there is still relatively little guidance on the appropriate functional form assumptions for hedonic price models. Cropper et al.'s (1988) seminal study suggested that simpler functional forms (e.g., log-linear) outperform more complex models in the presence of omitted variables, but more recently Kuminoff et al. (2010) found that flexible functional forms may perform better when combined with spatial and temporal fixed effects and quasi-experimental methods. Otherwise, the only firm guidance is that a linear specification is generally not theoretically appropriate (Bockstael and McConnell 2007; Bishop et al. 2020; Taylor 2017).

Several functional forms are used by the primary studies in the meta-data, including linear and non-linear forms. Most observations come from a double-log specification, but linear-log and log-linear are not far behind. The fewest observations come from linear models and log-quadratic specifications. Both linear models and linear-log models have a relatively greater number of significant and expected signed elasticities (see Figure S6 in supplementary material).

Spatial dependence and the potential for (spatially correlated) omitted variable bias is a long-standing concern in the hedonic property value literature (Kuminoff et al. 2010; Guignet and Lee 2021). A variety of approaches have emerged to address the issue, such as spatial econometric specifications (Anselin and Lozano-Gracia 2009), quasi-experimental methods (Parmeter and Pope 2013), and spatial fixed effects (Guignet and Lee 2021; Taylor 2017). With the exception of Olden and Tamayo (2014), who use an instrumental variable approach to address endogeneity concerns, no studies in our meta-data utilized quasi-experimental methods (e.g., difference-in-difference, regression discontinuity). About 41% of the observations did not use any approach to explicitly address spatial dependence (see Table 1), while 170 observations (59%) used some combination of spatial fixed effects, spatial lag models, and/or spatial autocorrelation approaches. Eighty-four observations were derived from models that included spatial fixed effects (e.g., neighborhood and watershed, town, city, or lake level). While no observation was based only on a spatial autoregressive (SAR) model, which includes a spatial lag of price (LeSage and Pace 2009), 95 observations were derived from these models in combination with the other approaches. In addition, 95 observations accounted for spatial autocorrelation either through a formal spatial error model (LeSage and Pace 2009) or allowing for clustered errors within a spatially defined group.[4] Studies not accounting for spatial dependence have many

[4] Observations from two studies in the original meta-dataset used by Guignet et al. (2022) were coded as using all three spatial dependence modeling approaches. For this methodological meta-analysis, after reviewing the papers, we recoded Netusil et al. (2014) to not using spatial autocorrelation approaches, and Liu et al. (2017) was recoded as not using spatial lag models (depending on the specific observation). Olden and Tamayo (2014) was originally coded as not using spatial methods. This study, however, uses two-stage least squares and instrumental variables to

more elasticities appearing significant with the expected sign compared to elasticities appearing insignificant and/or an unexpected sign (Figure S7).

Sales price data have become more accessible, either through private companies (e.g., CoreLogic, Zillow) or directly from county and state property assessor offices. However, data of assessed or predicted property values are sometimes easier to acquire and are available for a larger sample of homes. Although more comprehensive by not just reflecting homes that are sold, assessed or predicted values do not directly reflect market transactions, and hence revealed preferences. As identified by Bishop et al. (2020), the assessed or predicted values may have measurement error that in turn could affect the results from subsequent hedonic models. At the same time, Smith and Huang (1993) found that the use of actual sales prices reduced the likelihood of finding a significant and theoretically consistent elasticity. They suggested that this may occur because of higher variability or noise in actual sales data.

In the current meta-dataset, most observations are based on actual sales prices. When studies used actual sales data, the number of elasticities that are significant and theoretically consistent are about equal to those that are insignificant and/or have an unexpected sign (see Figure S8). Results from models using assessed or predicted values tend to have significant results consistent with the expected sign more often and are thus consistent with Smith and Huang's (1993) meta-analysis of air pollution.

---

deal with endogeneity. Although not a spatial econometric approach per se, it does set out to address the same spatially correlated omitted variable issue, and so we recoded this study to using spatial methods, but leave spatial fixed effects, spatial lag, and spatial autocorrelation variables in the meta-regressions equal to 0.

The hedonic literature uses a variety of approaches for acquiring measures of water quality, including: *in situ* measurements, spatial interpolation, model prediction, and satellite imagery (e.g., remote sensing). Most of the observations in the meta-data are from studies that used *in situ* measurement. Water quality measures based on spatial interpolated data were the second most common, followed by predicted measurements from water quality models. A relatively new approach for hedonic models that will certainly become increasingly common (e.g., Wolf and Kemp 2021; Zhang et al. 2022) is the use of satellite-based measures. During the period of our meta-data, however, Horsch and Lewis (2009) are the only ones to use water quality measures based on remote sensing data. For *in situ*, satellite, and predicted measurements, we see that elasticities are more likely to be significant and have the expected sign, compared to those studies that used spatial interpolation (Figure S9 in the supplementary material).

Another methodological decision is how many water quality variables to include in the primary hedonic model. For our study, most observations included just one water quality variable. However, the remaining 24% used up to seven water quality variables in a single model, with most using either two or five (e.g., Walsh and Milon 2016; Netusil et al. 2014; Bin and Czajkowski 2013). In their meta-analysis for air pollution, Smith and Huang (1993) found that as the number of air pollutant variables included in the hedonic model increased, the results were less likely to yield a statistically significant relationship with house prices. We see a similar trend in our meta-data (Figure S10 in the supplementary material).

4. Meta-analysis

In describing the meta-data above, we highlighted some patterns observed in the data, but primary study decisions like the commodity and housing market to analyze, and methodological and data choices, are not made independently by practitioners. Meta-regression analysis allows us to investigate potential relationships more formally between these dimensions and see how they might influence the statistical significance and theoretical consistency of the hedonic results.

Because our interest is in whether an elasticity is significant and theoretically consistent, we require a model that can handle a binary outcome variable. We start with the framework for Generalized Linear Models (GLM, McCullagh and Nelder 1989; Wilson and Lorenz 2015) where the dependent variable does not have to follow a normal distribution. Many common models fit under the GLM framework, and it provides an approach for creating a linear relationship even if a dependent variable has a nonlinear relationship with its independent variables. An important assumption for the GLM is that all observations are independent. Because our meta-data sometimes contains observations from the same study or data set, we use an extension of the GLM called the Generalized Estimating Equation (GEE; see Liang and Zeger 1986; Zeger and Liang 1986; Cameron and Miller 2011; Wilson and Lorenz 2015).

Following Guignet et al. (2022), clusters are defined as unique study and housing market combinations, leading to a total of J=98 clusters in the meta-data. Each cluster has a total number of observations defined as N_j. Y_{ij} is the binary outcome variable denoting whether the corresponding primary study elasticity estimate is significant and of the expected sign for observation i in cluster j. We define p_{ij} as the probability that Y_{ij} is equal to 1 and define a function that connects p_{ij} to the linear predictor variables ($\boldsymbol{x_{ij}}$). The inverse of this function defines the link function. For this study, the function $\Phi(\cdot)$ is the standard normal cumulative distribution function for the population-averaged probit (Cameron and Miller 2011).

$$Prob\left(Y = 1 \middle| \boldsymbol{x_{ij}}\right) = E\left(Y_{ij} \middle| \boldsymbol{x_{ij}}\right) = p_{ij} \tag{1}$$

$$= \Phi\left(\boldsymbol{x_{ij}}'\boldsymbol{\beta}\right)$$

Estimating a standard probit model that ignores *j* provides a simple approach for estimating **β**, provided that cluster-robust standard errors are also estimated (Cameron and Miller 2011). However, other approaches exist to address the cluster nature of our data (Cameron and Miller 2011).

One approach is to estimate a cluster-specific model that uses the standard probit but adds a cluster-specific variable, α_j, such that

$$Prob\left(Y = 1 \middle| \boldsymbol{x_{ij}}, \alpha_j\right) = p_{ij} = \Phi(\boldsymbol{x}'_{\boldsymbol{ij}}\boldsymbol{\beta} + \alpha_j) \tag{2}$$

where α_j can be estimated as a random or fixed effect (Cameron and Miller 2011). However, we are less interested in results for specific clusters and more interested in the average, or population, effects of hedonic study choices on significance and theoretical consistency.

The GEE, using a quasi-likelihood methodology which requires few assumptions about the distribution of *Y*, provides greater flexibility in identifying the correlation structure within clusters, and provides population-average results (Liang and Zeger 1986; Zeger and Liang 1986; Cameron and Miller 2011). The family of GEE models have rarely been used in environmental economics, but there are a few examples (e.g., Johnston et al. 2002; King and Anderson 2004).

With GEE models, a cluster-specific variable is not specified (as in Eq. 2), so we estimate (Eq. 1), but do not ignore j. Instead, expectations are defined for the jth cluster

$$E(\boldsymbol{Y}_j|\boldsymbol{x}_j) = \boldsymbol{p}_j(\beta). \qquad (3)$$

where $\mathbf{p}_j=[p_{i,1}, \ldots, p_{i,Nj}]$ is the marginal expectation of $\mathbf{Y}_j$ (Pendergast et al. 1996; Cameron and Miller 2011). Using the quasi-likelihood method, the set of GEE parameters, $\boldsymbol{\beta}$, solves (Pendergast et al. 1996; Cameron and Miller 2011):

$$S(\boldsymbol{\beta}) = \sum_{j=1}^{J} \frac{\partial \boldsymbol{p}_j{'}}{\partial \boldsymbol{\beta}} V_j^{-1}\left(\boldsymbol{Y}_j - \boldsymbol{p}_j(\boldsymbol{\beta})\right) = 0 \qquad (4)$$

If this were the GLM with independent observations within a cluster, we would have the variance matrix $V_j = A_j$, where A_j is a diagonal matrix of variances of p_{ij} as the jth diagonal element (Liang and Zeger 1986; Pendergast et al. 1996).

Unlike the GLM, Liang and Zeger (1986) broadened the choices of correlation possibilities within clusters with $R_j(\boldsymbol{\alpha})$, defined as the working correlation matrix. A_j is again the diagonal matrix of variances and ϕ is a scale parameter.

$$V_j = A_j^{1/2} R_j(\boldsymbol{\alpha}) A_j^{1/2} / \phi \qquad (5)$$

The matrix V_j for cluster j is what differentiates the GEE model from the GLM (Pendergast et al. 1996).

Although no specific approach exists for identifying the correct correlation structure, a number of choices exist (Zorn 2001). The working correlation matrices typically used include: independence, exchangeable, unstructured, and user-defined matrices (see Wilson and Lorenz 2015). The independence correlation matrix assumes observations within the same cluster are not correlated. The exchangeable matrix assumes observations within the same cluster have the same correlation. If an unstructured matrix is chosen, each pairwise correlation is estimated but having too many observations in a cluster or having unbalanced clusters can cause problems with the model (e.g., Shults et al. 2009). The more representative the working correlation structure is of the data, the more efficient the estimators. An incorrect choice of the working correlation structure does not affect the asymptotic consistency of the estimators as $J \rightarrow \infty$, but it can affect the consistency of the variance estimate (Zorn 2001). Therefore, a cluster-robust estimate of the variance-covariance matrix is almost always recommended because it is consistent as long as $J \rightarrow \infty$ is met (Liang and Zeger 1986; Zorn 2001; Cameron and Miller 2011).

Because the GEE uses quasi-likelihood, and not maximum likelihood like the GLM, an alternative approach for evaluating relative model performance is needed. Pan (2001) developed the quasi-likelihood for independence criterion (QIC) which is similar to the Akaike Information Criterion (AIC; see also Hardin and Hilbe 2003). The smallest QIC can help identify the appropriate working correlation structure and best model fit. A simplification of the QIC, the QICu – which substitutes in a penalty for the number of parameters – can also be used, but only to identify the appropriate set of variables (Pan 2001). We use a probit link function and test independence and exchangeable working correlation matrices using the QIC. Once the appropriate correlation structure is identified we choose the preferred set of variables using the QIC and QICu.

As described earlier, p_{ij} is the probability that Y_{ij} is equal to one. The linear predictor variables described above can be divided into three vectors, $\mathbf{q}_{ij}$, $\mathbf{m}_{ij}$, and $\mathbf{z}_{ij}$, along with the corresponding coefficients, $\boldsymbol{\beta}$, so that we have the function:

$$p_{ij} = \Phi(\boldsymbol{q}_{ij}\boldsymbol{\beta}_q + \boldsymbol{m}_{ij}\boldsymbol{\beta}_m + \boldsymbol{z}_{ij}\boldsymbol{\beta}_z) \qquad (6)$$

The vector $\mathbf{q_{ij}}$ denotes variables describing the commodity (i.e., type of waterbody and water quality measure). The vector $\mathbf{m_{ij}}$ represents market characteristics such as the average home price, study region and spatial and temporal definitions of the market. The vector $\mathbf{z_{ij}}$ represents methodological choices (e.g., use of assessed housing price, functional form, spatial dependence, approach for acquiring measures of water quality). To complete the set of independent variables, we include a study year trend based on the year of publication (range: 1985-2017), where 1985=0, 1986=1, and continuing to 2017=32. Such a trend variable may partially capture changes in methods, data like water quality monitoring, preferences, and perceptions of water quality over time that may affect significance and theoretical consistency of the primary study estimates. Table 1 displays the variables for each category.

5. Results

5.1 Probit Generalized Estimating Equation Results

We estimate separate probit GEE meta-regressions for the independence and exchangeable correlation structures using SAS 9.4 (SAS Institute 2013). Our first GEE model tests independence correlation structure which assumes there is no correlation within each cluster but allows for cluster-robust standard errors (SAS Institute 2013). We use the 98 unique study-housing market combinations to define the clusters across the 290 observations. The results are presented in Table 3. Model (1) includes the vector $\mathbf{q_{ij}}$ that represents the variables describing the type of water body and water quality measures examined. For this model, none of the variables are significant.

Model (2) adds variables denoting the study region of the US. For studies that estimated hedonic models for estuaries, the coefficient is now negative and significant. This suggests that compared to lakes/reservoirs, hedonic studies of estuaries are less likely to yield statistically significant results with the expected sign. Studies of waterbodies and housing markets in the south tend to be more likely to yield significant and theoretically expected results, relative to studies in the northeast region (the omitted category). None of the variables representing the categories of water quality are statistically significant, suggesting that, at least in this model, hedonic studies of the various water quality categories are equally as likely to yield the expected result (or not), holding all else constant.

Model (3) in Table 3 adds the remaining market definition variables from vector $\mathbf{m}_{ij}$. The previous findings regarding estuaries and the south region remain robust, as is the finding that the types of water quality variable are statistically insignificant. The market characteristic variables – *Mean House Price* and *Multiple Counties/Subcounties* – are insignificant. However, *Sample Years* is significant and negative, suggesting that as the study period (and hence the duration of the assumed hedonic equilibrium) increases in length, the likelihood of the elasticity being insignificant or theoretically inconsistent increases. However, this result is not robust in subsequent meta-regression models that control for methodological features of the primary study.

Model (4) in Table 3 adds the final vector representing methodological choices. The previous findings regarding estuaries and studies in the south are robust, but otherwise we see large variability in the results. For example, the coefficient corresponding to *River* is now positive and significant, suggesting that hedonic studies examining rivers are more likely to yield the expected results, all else constant. The variable *Nutrients* is now significant, suggesting that studies examining the impact of nutrients on house prices have a higher likelihood of yielding an elasticity estimate that is significant and of the expected sign, compared to studies of water clarity (the omitted category). In Model (4), the evidence suggests that hedonic studies of waterbodies and housing markets in the west are less likely to yield the expected results (compared to studies of the northeast). The remaining market characteristics are insignificant.

Six of the seven methodological variables in Model (4) are significant, emphasizing the importance of primary researchers' data decisions and modeling assumptions. For example, not using *in situ* water quality data (*Not In Situ*) and not using actual sales prices (*Use of Assessed Housing Price*) lead to a higher probability of a study yielding statistically significant results that are of the expected sign. These methodological choices may reduce the variability in the data, possibly facilitating more precise estimates in the primary hedonic studies (Smith and Huang 1993). We emphasize, however, that finding the expected result does not necessarily imply the correct result. Including more than one water quality variable in the model leads to a higher likelihood that the elasticity will be insignificant and/or have an unexpected sign. This result is also in line with Smith and Huang's (1993) meta-analysis of hedonic studies on air quality. Our initial modeling approach for examining the role of controlling for spatial dependence on primary study results was to include a dummy variable that was equal to 1 if any spatial modeling approach was used. Because the coefficient was negative and significant, we were interested to see how the different approaches for modeling spatial dependence (i.e., *Spatial Fixed Effects*, *Spatial Lag*, and *Spatial Autocorrelation*) affect significance and expected sign of the elasticities. Somewhat surprisingly, all the corresponding coefficients for all variables representing spatial dependence modeling approaches are significant and negative.

Comparing Model (4) to Model (5), we see very little change in the results. The added publication time trend in Model (5) is insignificant, suggesting that, all else constant, the likelihood of hedonic studies yielding significant results of the expected sign has not changed over time.

The same five models are re-estimated using the probit link function and exchangeable correlation structure (Table 4). There are some differences using this correlation structure. Model (1) has a negative and significant coefficient for *Estuary* while two categories of water quality are significant. Home price elasticity estimates with respect to nutrients are again more likely to be significant and of the expected sign, compared to water clarity, while price elasticities with respect to biochemical water quality measures are less likely to be of the expected sign and significant (although this result is not significant in subsequent models).

Model (2) in Table 4 also has a negative and significant coefficient for *Estuary*, as well as a positive and significant coefficient for southern observations. In addition, elasticity estimates from study areas in the west demonstrate a lower tendency to be statistically significant and of the expected sign. The previous results pertaining to *Nutrients* and *Biochem* are not robust to the inclusion of these regional variables, but the *Nutrients* variable does become significant again in later models where we control for methodological choices.

When we add the remaining market characteristics for Model (3), the coefficient corresponding to the variable *Multiple Counties/Subcounties* suggests a higher likelihood that the elasticity will be significant and consistent with theory. This finding is stronger in subsequent models. It seems that primary studies covering broader study areas are more likely to yield significant results that match expectations of how the corresponding water quality measure should impact home prices.

For Models (4) and (5) in Table 4, *River*, *Estuary*, and *Nutrients* are similar to the independence correlation structure models, as are the regional variables, *West* and *South*. The variable *Multiple Counties/Subcounties* is positive and significant, unlike the corresponding independence model (Model (4) in Table 3). The methodological variables also have similar effects on elasticity estimates as in Table 3, including significant and negative coefficients corresponding to all of the approaches for modeling spatial dependence. Similar to the independence correlation structure models, the time trend variable added in Model (5) is statistically insignificant.

Although not reported, we also ran Models (3), (4), and (5) with the dummy variable equal to 1 if any of the sample years included the housing market bubble burst (2006-2009). The results were very similar to those already presented in Tables 3 and 4, but the market bubble burst coefficient was statistically insignificant. This is surprising given the patterns seen in Figure S5, which suggested elasticities were more likely to be significant and of the expected sign when studies did not include sample years during the market bubble burst (2006-2009).

When comparing results that use different correlation structures, the model yielding the smallest QIC is the one that best fits the meta-data (Pan 2001). Across the board, the independence correlation structure models in Table 3 yield the lowest QIC. When identifying the most appropriate model within a correlation structure, the smallest QIC_u can also be used. For those models using the independence structure, Model (4) (in Table 3) appears to best fit the meta-data.

5.2 Robustness Check

Robustness checks on our results for an elasticity that is significant at the 5% level can be found in the online supplementary material. The robustness models use alternative dependent variable definitions. The first is a more stringent definition, where $Y_{ij}=1$ only if the elasticity was significant with a p-value less than 0.01 and had a sign consistent with economic theory, and the second is less stringent, where $Y_{ij}=1$ if the primary study elasticity estimate is of the expected sign and statistically significant based on a p-value less than 0.10. For the more stringent dependent variable we have 114 observations where $Y_{ij}=1$, and for the less stringent dependent variable we have 173 observations where $Y_{ij}=1$. These can be compared to our main dependent variable definition, where 151 observations had $Y_{ij}=1$ (see Table S2 for summary statistics).

Because of the change in observations with a 1 or a 0 for the dependent variable, some independent variables can no longer be included in the model because of little (to no) variation in the dependent variable for a given value of the independent variable. We drop the variables *Nutrients*, *Sediment*, *Linear* and *Use of Assessed Housing Price* for all robustness checks because they are nearly perfect predictors under our alternative definitions for Y_{ij}. We identify the results from the independence correlation structure as most appropriate (see Table S3 in online supplementary material). When comparing the robustness results for the independence correlation structure, the estimates are fairly robust. *Estuary*, *South*, and including more than one water quality variable remain significant and of the same sign across all models. We point out that the variable *Multiple Counties/Subcounties* differs with $Y_{0.01}$ and $Y_{0.10}$ compared to Y (Table S3). None of the spatial methods are significant for $Y_{0.01}$, but *Spatial Autocorrelation* is negative and significant for Y and $Y_{0.10}$.

6. Discussion

The primary objective of this study was to answer the following question:

> What does the hedonic literature examining surface water quality generally reveal about how the type of commodity, market characteristics, and methodological decisions affect the significance and theoretical consistency of the estimated property value impacts?

The results from the GEE meta-regression models provide evidence that several of the definitions and decisions made in primary studies do affect the estimated relationship between water quality and property value impacts. The type of waterbody, a focus on nutrients, the region of the country, and many of the methodological choices play a role in the estimated impacts of water quality on waterfront home values. Our meta-regression results also provide evidence that, all else constant, the significance and theoretical consistency of the estimated house price elasticities with respect to water quality are not increasing or decreasing through time.

For many of the independent variables, we were unsure *a priori* as to what the estimated effect on the probability of the expected price effect would be. For those variables where we were able to hypothesize the effect, we did not always find the expected result. Smith and Huang (1993) hypothesized and confirmed that both linear specifications and hedonic models that included more than one air pollution measure would lead to insignificant or inconsistent results. In contrast, in the context of water quality we find a null effect with respect to assuming a linear model. On the other hand, when more than one water quality parameter was included in a hedonic model, we find a negative and significant result, which agrees with Smith and Huang's (1993) meta-analysis.

When considering the commodity definition (i.e., waterbody type and water quality category), hedonic analyses of estuaries appear to be less likely to yield a significant and theoretically expected result, compared to studies of lakes and reservoirs. Given the salience of water clarity as a measure of water quality, it is surprising that a focus on nutrients tends to yield a higher likelihood of the expected result compared to clarity.

One market factor that we thought might be important was the 2006-2009 housing market bubble burst. There has been plenty written about the topic, and in particular, whether implicit prices of interest could vary during such shocks (Boyle et al. 2012; Taylor 2017; Bishop et al. 2020). Smith and Huang (1993) included the vacancy rate as a proxy for market conditions. They found a higher likelihood of a significant result in terms of the implicit price of air pollution as vacancy rates increased. In our meta-analysis, however, controlling for whether a primary study sample included transactions during the 2006-2009 bubble burst did not yield a significant result. Although we find no effect, a better understanding of housing market expansions and contractions on implicit price estimates of interest should be examined more closely in future research.

Methodological choices appear to have a very important role in determining the estimated relationship between water quality and housing prices. The use of assessed housing prices and predicted or modeled water quality data lead to a similar, higher likelihood of finding a significant estimated price impact that is of the expected sign. Actual housing prices and *in situ* measurements may have more random variation in the data, which could obscure the expected results. Assessed housing prices and modeled water quality may reduce this variability and should be recognized by decision-makers interested in hedonic property value results. On the other hand, assessed values do not directly reflect market behavior and modeled water quality values can introduce prediction error, so there is a tradeoff that researchers must consider when designing a new hedonic study, and for practitioners to consider when evaluating a study to inform policy. Finding the expected result in in a hedonic analysis, does not necessarily mean it is the "correct" result, and in general best practices should be followed (Bishop et al. 2020).

Nonetheless, with increased data accessibility and computing power, new studies are being published at broader, even national, scales (e.g., Moore et al. 2020; Zhang et al. 2022). Finding consistent water quality measures across the country is difficult, meaning that studies going forward will likely rely heavily on modeled water quality or data generated from algorithms and satellite imagery.

Given the importance of trying to minimize spatially correlated omitted variable bias in hedonic property value models, we paid particular attention to methodological choices meant to account for spatial dependence. We find that controlling for spatial dependence actually decreases the likelihood of a primary study yielding a significant result of the expected sign. Although speculative, one possible explanation is that the true water quality price effect is relatively small and controlling for spatially correlated confounders better identifies that near zero effect. On the other hand, if the role of spatial dependence in the true data generating process is minimal, then spatial fixed effects and other spatial modeling approaches may be over-parameterizing the models, making it less likely a study would identify a significant effect if there is one. It is also possible that some approaches to address spatial dependence, such as spatial fixed effects, may be absorbing much of the price variation of interest (Abbot and Klaiber 2011). Variation in housing prices due to water quality may often be more due to spatial, rather than temporal, variation in water quality (Kung et al. 2022). In such cases, it is more difficult to isolate variation due to water quality from the spatially correlated omitted variables; making it less likely that one would find the expected result because the resulting omitted variable bias would still be present.

For the 120 observations that were derived from models where no spatial methods were implemented, 78 (65%) were significant and matched expectations (Figure S7). It is possible that no spatial dependence was found in some of these cases; in such instances finding the expected result is reasonable. Two studies (out of the 15 that do not use spatial methods in some or all of their models) test for spatial autocorrelation and do not find it in their data (Feather et al. 1992; Liao et al. 2016). Ten studies do not mention spatial dependence in their papers, suggesting that the elasticity estimates could potentially be biased, or at least inefficient depending on the nature of the spatial dependence (e.g., Nelson 2008; Chi and Zhu 2020). The last three studies use spatial methods in some of their observations as a comparison to observations that do not. The negative sign corresponding to estimates from models that did account for spatial autocorrelation suggests that standard errors could be underestimated when no spatial approach is used (Nelson 2008). Testing for spatial dependence and using spatial methods, when appropriate, is generally considered "best practice" (Taylor 2017; Bishop et al. 2020), so decision-makers should consider how practitioners addressed these issues before applying or extrapolating hedonic results.

7.0 Conclusion

Hedonic property value methods represent a large and growing branch of the nonmarket valuation literature. As we move forward and continue to apply and advance the methodology, it is important to look back and take stock on what has been done and the empirical implications of past analyses and modeling decisions. Our meta-analysis attempts to do just that, by systematically and quantitatively reviewing the hedonic property value literature on the price effects of water quality.

With the intention of providing information to assess hedonic models and the estimated price effects, we highlight three key points. First, our meta-regression results are limited by the existing literature, so we encourage researchers to fill in areas where the literature is scarce. In particular, in the context of the US, more hedonic studies examining water quality in the west and midwest regions, and for rivers and estuaries, are needed. Second, we demonstrate that study design choices and modeling assumptions have a large influence in determining the estimated price effects. When assessing what studies to use to inform water quality management and policy decisions, practitioners, and decision-makers must consider the implications of our meta-analysis, along with how closely a primary study follows best practices. Third, we unequivocally recommend that researchers continue to follow contemporary guidance in hedonic modeling (Bishop et al. 2020). Our meta-regression results suggest that practices currently considered to be subpar in many applications, like using assessed housing values and not accounting for spatial dependence, may increase the tendency for a hedonic analysis to yield the hypothesized result. Therefore, we caution that the expected result is not necessarily the correct result, and researchers should continue to assess the robustness of their findings against their study design and methodological choices.

ACKNOWLEDGMENTS

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the US EPA. The research described in this article has been funded wholly or in part by the US EPA under contract EP-C-13-039 to Abt Associates. Any mention of trade names, products, or services does not imply an endorsement by the US Government or the US EPA. The authors declare they have no actual or potential competing financial or other conflicts of interests.

REFERENCES

Abbot, J.K. and Klaiber, H.A. 2011. An embarrassment of riches: Confronting the omitted variable bias and multi-scale capitalization in hedonic price models. Review of Economics and Statistics, 93(4), pp. 1331-1342.

Anselin, L. and Lozano-Gracia, N. 2009. Spatial Hedonic Models. In T.C. Mills and K. Patterson (eds), Palgrave Handbook of Econometrics (pp 1213-1250). London, United Kingdom: Palgrave Macmillan.

Bin, O. and Czajkowski, J., 2013. The impact of technical and non-technical measures of water quality on coastal waterfront property values in South Florida. Marine Resource Economics, 28(1), pp.43-63.

Bishop, K.C., Kuminoff, N.V., Banzhaf, S., Boyle, K.J., von Gravenitz, K., Pope, J.C., Smith, V.K., Timmins, C.D., 2020. Best practices in using hedonic property value models for welfare measurement. Review of Environmental Economics and Policy, 14(2), pp. 260-281.

Bockstael, N.E., and McConnell, K. E., 2007. Environmental and Resource Valuation with Revealed Preferences: A Theoretical Guide to Empirical Methods. Springer, Dordrecht, The Netherlands.

Boyle, M. and Kiel, K., 2001. A survey of house price hedonic studies of the impact of environmental externalities. Journal of real estate literature, 9(2), pp.117-144.

Boyle, K., Lewis, L., Pope, J. and Zabel, J., 2012. Valuation in a bubble: Hedonic modeling pre- and post-housing market collapse. AERE Newsletter, 32(2), pp.24-31.

Boyle, K.J. and Wooldridge, J.M., 2018. Understanding error structures and exploiting panel data in meta-analytic benefit transfers. Environmental and resource economics, 69(3), pp.609-635.

Cameron, A.C. and Miller, D.L., 2011. "Robust Inference with Clustered Data." In Handbook of Empirical Economics and Finance, eds A. Ullah, D.E. Giles. Chapman and Hall/CRC. Boca Raton, FL., pp. 1-28

Chi, G., and Zhu, J., 2020. Spatial regression models for the social sciences. SAGE Publications, Inc. https://dx.doi.org/10.4135/9781544302096

Crompton, J.L., 2001. The impact of parks on property values: A review of the empirical evidence. Journal of leisure research, 33(1), pp.1-31.

Cropper, M.L., Deck, L.B. and McConnell, K.E., 1988. On the choice of funtional form for hedonic price functions. The review of economics and statistics, pp.668-675.

Feather, T.D., Pettit, E.M. and Ventikos, P., 1992. Valuation of lake resources through hedonic pricing (No. IRW-92-R-8). Army Engineer Inst for Water Resources. Alexandria, VA.

Guignet, D., Griffiths, C., Klemick, H. and Walsh, P.J., 2017. The implicit price of aquatic grasses. Marine Resource Economics, 32(1), pp.21-41.

Guignet, D., Heberling, M., Papenfus, M. and Griot, O., 2022. Property values, water quality, and benefit transfer: A nationwide meta-analysis. Land Economics 98(2), May 2022

Guignet, D. and Lee, J., 2021. State of the Art of Hedonic Pricing. In Oxford Research Encyclopedia of Environmental Science. DOI: 10.1093/acrefore/9780199389414.013.555

Hardin, J.W. and Hilbe, J.M., 2003. Generalized estimating equations. Chapman and Hall/CRC. Boca Raton, FL.

Horsch, E.J. and Lewis, D.J., 2009. The effects of aquatic invasive species on property values: evidence from a quasi-experiment. Land Economics, 85(3), pp.391-409.

Johnston, R.J., Swallow, S.K. and Bauer, D.M., 2002. Spatial factors and stated preference values for public goods: considerations for rural land use. Land economics, 78(4), pp.481-500.

Kiel, K., 2006. Environmental Contamination and House Values. Environmental Valuation: Interregional and Intraregional Perspectives, J. I. Carruthers and B. Mundy, B. (eds.), Ashgate: Burlington, VT, pp. 121-146.

Kiel, K.A. and Williams, M., 2007. The impact of Superfund sites on local property values: Are all sites the same?. Journal of Urban Economics, 61(1), pp.170-192.

King, J.R. and Anderson, C.M., 2004. Marginal property tax effects of conservation easements: a Vermont case study. American Journal of Agricultural Economics, 86(4), pp.919-932.

Kuminoff, N.V., Parmeter, C.F. and Pope, J.C., 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?. Journal of environmental economics and management, 60(3), pp.145-160.

Kung, M., Guignet, D. and Walsh, P., 2022. Comparing pollution where you live and play: A hedonic analysis of enterococcus in the Long Island Sound. Marine Resource Economics, 37(1), pp. 65-89.

LeSage, J. and Pace, R.K., 2009. Introduction to spatial econometrics. Chapman and Hall/CRC.

Liang, K.Y. and Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika, 73, pp.13-22.

Liao, F.H., Wilhelm, F.M., and Solomon, M., 2016. The Effects of Ambient Water Quality and Eurasian Watermilfoil on Lakefront Property Values in the Coeur d'Alene Area of Northern Idaho, USA. Sustainability, 8(1), p.44.

Liu, T., Opaluch, J.J., and Uchida, E., 2017. The impact of water quality in Narragansett Bay on housing prices. Water Resources Research, 53(8), pp.6454-6471.

McCullagh, P. and Nelder, J.A., 1989. Generalized linear models. 2nd Edition. Chapman and Hall. New York.

Moore, M.R., Doubek, J.P., Xu, H. and Cardinale, B.J., 2020. Hedonic price estimates of lake water quality: valued attribute, instrumental variables, and ecological-economic benefits. Ecological Economics, 176, p.106692.

Nelson, J.P., 2008. Hedonic Property Value Studies of Transportation Noise: Aircraft and Road Traffic. In: Baranzini A., Ramirez J., Schaerer C., Thalmann P. (eds) Hedonic Methods in Housing Markets. Springer, New York, NY. (pp.57-82). https://doi.org/10.1007/978-0-387-76815-1_4

Nelson, J.P., and Kennedy, P., 2009. The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. Environmental and Resource Economics, 42(3), pp.345-377.

Netusil, N.R., Kincaid, M., and Chang, H., 2014. Valuing water quality in urban watersheds: A comparative analysis of Johnson Creek, Oregon, and Burnt Bridge Creek, Washington. Water Resources Research, 50(5), pp.4254-4268.

Nicholls, S., and Crompton, J., 2018. A Comprehensive Review of the Evidence of the Impact of Surface Water Quality on Property Values. Sustainability, 10(500), pp.1-30.

Olden, J.D. and Tamayo, M., 2014. Incentivizing the public to support invasive species management: Eurasian milfoil reduces lakefront property values. PloS one, 9(10), p.e110458.

Pan, W., 2001. Akaike's information criterion in generalized estimating equations. Biometrics, 57(1), pp.120-125.

Parmeter, C. F. and Pope, J. C., 2013. Quasi-experiments and Hedonic Property Value Methods. In J. A. List and M. K. Price (eds), Handbook on Experimental Economics and the Environment (pp 3-66). Cheltenham, United Kingdom: Edward Elgar Publishing.

Pendergast, J.F., Gange, S.J., Newton, M.A., Lindstrom, M.J., Palta, M. and Fisher, M.R., 1996. A survey of methods for analyzing clustered binary response data. International Statistical Review/Revue Internationale de Statistique, pp.89-118.

Petrolia, D.R., Guignet, D., Whitehead, J., Kent, C., Caulder, C. and Amon, K., 2021. Nonmarket Valuation in the Environmental Protection Agency's Regulatory Process. Applied Economic Perspectives and Policy, 43(3), pp.952-969. DOI: 10.1002/aepp.13106.

SAS Institute Inc., 2013. SAS® 9.4 Statements: Reference. Cary, NC: SAS Institute Inc.

Shults, J., Sun, W., Tu, X., Kim, H., Amsterdam, J., Hilbe, J.M. and Ten-Have, T., 2009. A comparison of several approaches for choosing between working correlation structures in generalized estimating equation analysis of longitudinal binary data. Statistics in medicine, 28(18), pp.2338-2355.

Smith, V.K. and Huang, J.C., 1993. Hedonic models and air pollution: twenty-five years and counting. Environmental and Resource Economics, 3(4), pp.381-394.

Stanley, T.D., 2001. Wheat from chaff: Meta-analysis as quantitative literature review. Journal of economic perspectives, 15(3), pp.131-150.

Stanley, T.D. and Doucouliagos, H., 2012. Meta-regression analysis in economics and business. Routledge. New York, NY.

Taylor, L.O., 2017. Hedonics. In Champ P., Boyle K., Brown T. (eds), A primer on nonmarket valuation. The Economics of Non-Market Goods and Resources (pp. 235-292). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7104-8_7

Walsh, P., Griffiths, C., Guignet, D., and Klemick, H., 2017. Modeling the property price impact of water quality in 14 Chesapeake Bay Counties. Ecological Economics, 135, pp.103-113.

Walsh, P.J. and Milon, J.W., 2016. Nutrient standards, water quality indicators, and economic benefits from water quality regulations. Environmental and Resource Economics, 64(4), pp.643-661.

Wilde, L., Loos, C. and Williamson, J., 2012. Pipelines and property values: An eclectic review of the literature. Journal of Real Estate Literature, 20(2), pp.245-259.

Wilson, J.R. and Lorenz, K.A., 2015. Modeling binary correlated responses using SAS, SPSS and R (Vol. 9). Springer. https://doi.org/10.1007/978-3-319-23805-0

Wolf, D. and Kemp, T., 2021. Convergent Validity of Satellite and Secchi Disk Measures of Water Clarity in Hedonic Models. Land Economics, 97(1): 39-58. https://doi.org/10.3368/wple.97.1.050819-0062R1

Zeger, S.L. and Liang, K.Y., 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics, pp.121-130.

Zhang, J., Phaneuf, D.J. and Schaeffer, B.A., 2022. Property values and cyanobacterial algal blooms: Evidence from satellite monitoring of Inland Lakes. Ecological Economics, 199, p.107481. https://doi.org/10.1016/j.ecolecon.2022.107481

Zorn, C.J., 2001. Generalized estimating equation models for correlated data: A review with applications. American Journal of Political Science, pp.470-490.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Elasticity	0.0876	0.7002	-2.6376	8.3202
Dependent variable (Significant with a p-value < 0.05 and theoretically consistent=1**):**				
Y	0.5207	0.5004	0	1
Environmental commodity variables:				
Lake or Reservoir[b]	0.5724	0.4956	0	1
Estuary	0.3379	0.4738	0	1
River	0.0897	0.2862	0	1
Clarity[b]	0.6172	0.4869	0	1
Nutrients	0.1276	0.3342	0	1
Sediment	0.0517	0.2219	0	1
Bacteria	0.1414	0.3490	0	1
Biochem	0.0621	0.2417	0	1
Study area and market characteristics:				
Northeast[b]	0.3034	0.4605	0	1
Midwest	0.1793	0.3843	0	1
West	0.1241	0.3303	0	1
South	0.3931	0.4893	0	1
Mean House Price (thousands, 2018$)[a]	317.9502	220.6630	8.0196	1245.9600
Single Subcounty[b]	0.7276	0.4460	0	1
Multiple Counties/Subcounties	0.2724	0.4460	0	1
Sample Years[a]	8.8586	5.0435	1	24
Bubble (=1 if sample includes 2006-2009)	0.3655	0.4824	0	1
Methodological variables:				
Double-log[b]	0.3172	0.4662	0	1
Log-linear[b]	0.2966	0.4575	0	1
Linear-log[b]	0.2759	0.4477	0	1
Log-quadratic[b]	0.0448	0.2073	0	1
Linear	0.0655	0.2479	0	1
No Spatial Method[b]	0.4138	0.4934	0	1
Spatial Fixed Effects	0.2897	0.4544	0	1
Spatial Lag	0.3276	0.4701	0	1
Spatial Autocorrelation	0.3276	0.4701	0	1

Use of Assessed Housing Price	0.0828	0.2760	0	1
Not In Situ	0.3483	0.4772	0	1
More than One WQ Variable	0.2448	0.4307	0	1
Time trend:				
Time Trend (year published)[a]	24.1379	7.5817	0	32

Unweighted descriptive statistics presented for n=290 unique elasticity estimates in meta-dataset. (a) Denotes independent variables that are continuous. (b) Denotes reference category.

Table 2: Categories of Water Quality Measures (for all 17 measures of water quality)

Water Quality Variable	Water Quality Measure in Hedonic Model
Clarity	Secchi disk depth Light attenuation Percent water visibility
Nutrients	Nitrogen Phosphorus Lake trophic state index Chlorophyll a Trophic state index
Sediment	Sediment Total suspended solids Sedimentation rate Turbidity
Bacteria	Fecal coliform *E. coli*
Biochem	Dissolved oxygen Temperature Salinity

Table 3: GEE Probit Meta-Regression Results (Independence Correlation Structure)

VARIABLES	(1)	(2)	(3)	(4)	(5)
Intercept	0.16	0.18	0.44	-0.09	0.02
	(0.204)	(0.253)	(0.365)	(0.402)	(0.596)
River	-0.25	0.06	0.21	3.53***	3.60***
	(0.497)	(0.651)	(0.630)	(1.274)	(1.273)
Estuary	-0.39	-1.62***	-1.05**	-2.01***	-2.05***
	(0.340)	(0.359)	(0.501)	(0.641)	(0.638)
Nutrients	0.46	0.44	0.90	1.79**	1.78**
	(0.472)	(0.369)	(0.651)	(0.758)	(0.746)
Sediment	-0.49	-0.26	-0.78	-0.11	-0.15
	(0.577)	(0.613)	(0.544)	(0.671)	(0.650)
Bacteria	0.33	0.43	0.09	0.33	0.29
	(0.641)	(0.515)	(0.394)	(0.312)	(0.350)
Biochem	-0.64	-0.50	-0.96	-0.11	-0.13
	(0.605)	(0.629)	(0.806)	(0.716)	(0.711)
Midwest		-0.35	0.11	-0.50	-0.54
		(0.405)	(0.447)	(0.612)	(0.647)
West		-0.59	-0.21	-1.71***	-1.64**
		(0.595)	(0.590)	(0.643)	(0.653)
South		1.35***	1.51***	1.76***	1.79***
		(0.355)	(0.524)	(0.597)	(0.595)
Mean House Price (thousands, 2018$)			-1.00E-04	1.70E-03	1.80E-03
			(0.001)	(0.001)	(0.001)
Sample Years			-0.08**	-0.03	-0.02
			(0.040)	(0.051)	(0.050)
Multiple Counties/Subcounties			0.48	0.53	0.52
			(0.405)	(0.454)	(0.452)
Linear				0.06	0.04
				(0.338)	(0.332)
Spatial Fixed Effects				-0.88**	-0.88**
				(0.378)	(0.373)
Spatial Lag				-0.64**	-0.62**
				(0.303)	(0.314)
Spatial Autocorrelation				-0.88**	-0.85**
				(0.346)	(0.346)
Use of Assessed Housing Price				1.96***	1.99***
				(0.687)	(0.705)
Not In Situ				1.50***	1.50***
				(0.435)	(0.428)
More than One WQ Variable				-1.29**	-1.31**
				(0.576)	(0.564)

Time Trend (year published)					-0.01
					(0.028)
Observations	290	290	290	290	290
QIC	436.12	398.19	391.38	351.96	354.08
QICu	395.99	369.06	361.42	332.42	334.24

Dependent variable: Y based on p<0.05. *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses; clustered according to the J=98 study-housing market combinations.

Table 4: GEE Probit Meta-Regression Results (Exchangeable Correlation Structure)

VARIABLES	(1)	(2)	(3)	(4)	(5)
Intercept	0.42**	0.37*	0.26	-0.06	-0.01
	(0.166)	(0.195)	(0.289)	(0.372)	(0.574)
River	-0.26	0.48	0.73	3.57***	3.61***
	(0.494)	(0.646)	(0.762)	(1.088)	(1.102)
Estuary	-0.66**	-1.70***	-1.59***	-2.49***	-2.50***
	(0.306)	(0.363)	(0.447)	(0.557)	(0.555)
Nutrients	1.03***	0.48	0.55	1.64**	1.64**
	(0.387)	(0.334)	(0.428)	(0.693)	(0.689)
Sediment	-0.49	-0.52	-0.75	-0.16	-0.18
	(0.544)	(0.547)	(0.524)	(0.588)	(0.575)
Bacteria	-0.36	-0.24	-0.48	0.15	0.13
	(0.537)	(0.609)	(0.609)	(0.406)	(0.455)
Biochem	-1.14*	-0.84	-1.22	-0.35	-0.36
	(0.607)	(0.577)	(0.872)	(0.747)	(0.747)
Midwest		0.17	0.14	-0.37	-0.38
		(0.362)	(0.403)	(0.464)	(0.476)
West		-1.04*	-0.87	-1.64**	-1.62**
		(0.589)	(0.679)	(0.662)	(0.672)
South		1.16***	1.35***	1.80***	1.81***
		(0.354)	(0.500)	(0.595)	(0.596)
Mean House Price (thousands, 2018$)			4.00E-04	2.00E-03*	2.00E-03*
			(0.001)	(0.001)	(0.001)
Sample Years			-0.03	-0.04	-0.04
			(0.032)	(0.049)	(0.050)
Multiple Counties/Subcounties			0.57*	0.75**	0.75**
			(0.308)	(0.340)	(0.342)
Linear				0.01	0.00
				(0.222)	(0.223)
Spatial Fixed Effects				-1.01***	-1.01***
				(0.369)	(0.372)
Spatial Lag				-0.64***	-0.64***
				(0.233)	(0.237)
Spatial Autocorrelation				-0.61**	-0.61**
				(0.298)	(0.296)
Use of Assessed Housing Price				1.35***	1.36***
				(0.456)	(0.465)
Not In Situ				1.66***	1.66***
				(0.396)	(0.395)
More than One WQ Variable				-1.31***	-1.31***
				(0.479)	(0.475)

Time Trend (year published)					-3.80E-03
					(0.026)
Observations	290	290	290	290	290
QIC	446.93	416.81	415.99	359.39	362.10
QICu	421.29	391.47	384.14	341.62	343.41
Working Correlation	0.890	0.624	0.585	0.267	0.268

Dependent variable: Y based on p<0.05. *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses; clustered according to the J=98 study-housing market combinations.

Supporting Information for:

Water Quality and Hedonic Models: A Meta-Analysis of Commodity, Market, and Methodological Characteristics

Matthew T. Heberling[a,*]; Dennis Guignet[b]; and Michael Papenfus[c]

[a]US Environmental Protection Agency, Office of Research and Development, Cincinnati, OH 45268, USA.

[b]Appalachian State University, Department of Economics, Boone, NC, 28608-2051, USA.

[c]US Environmental Protection Agency, Office of Research and Development, Corvallis, OR, 97330, USA.

*Corresponding author: M. T. Heberling, US EPA/ORD/CEMM (MS 587), 26 W MLK Dr., Cincinnati, OH 45268, USA. Telephone: +1.513.569.7917; Email: Heberling.Matt@epa.gov

Contents of this file

Introduction

This supporting information provides figures and tables for comparing variables and meta-regression models. We present a set of variables that represent environmental commodity definition, market characteristics, and methodological choices and the distribution of elasticities that are insignificant and/or have unexpected sign vs. elasticities that are significant and have expected sign. In addition, we present the robustness check dependent variable statistics and meta-regression results to compare with the main results.

Supplementary Figures

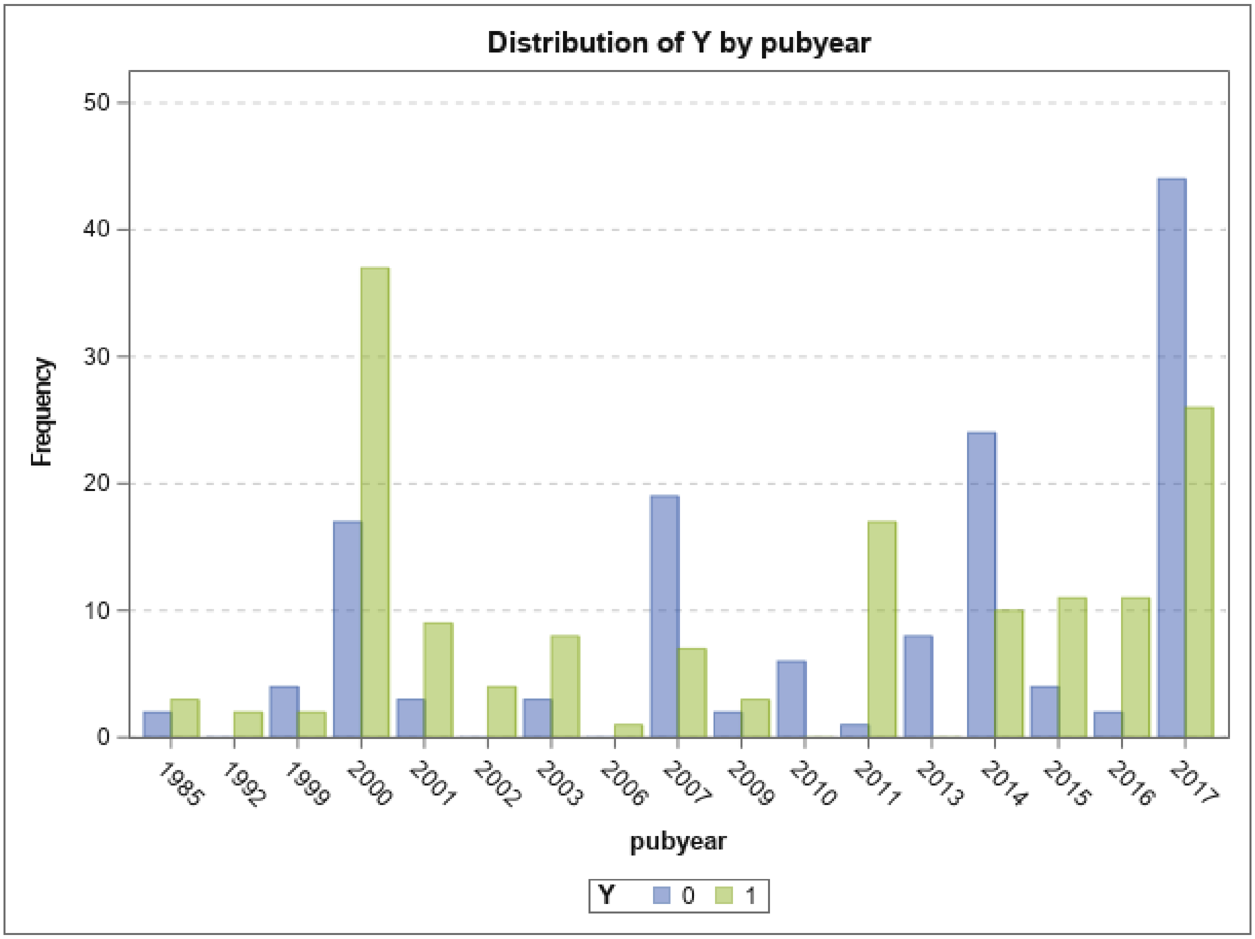

Figure S1: Publication year (pubyear) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

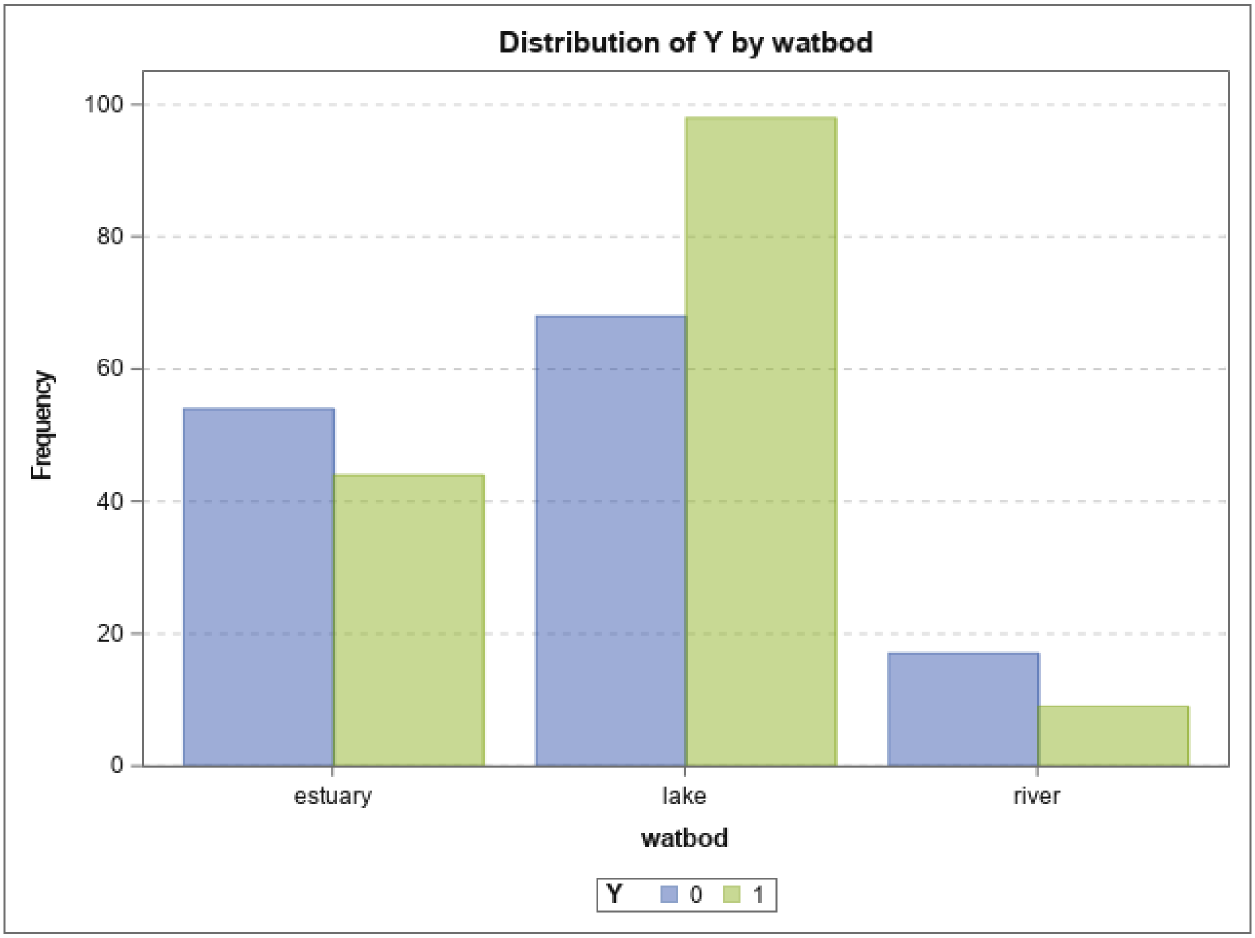

Figure S2: Waterbody type (watbod) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

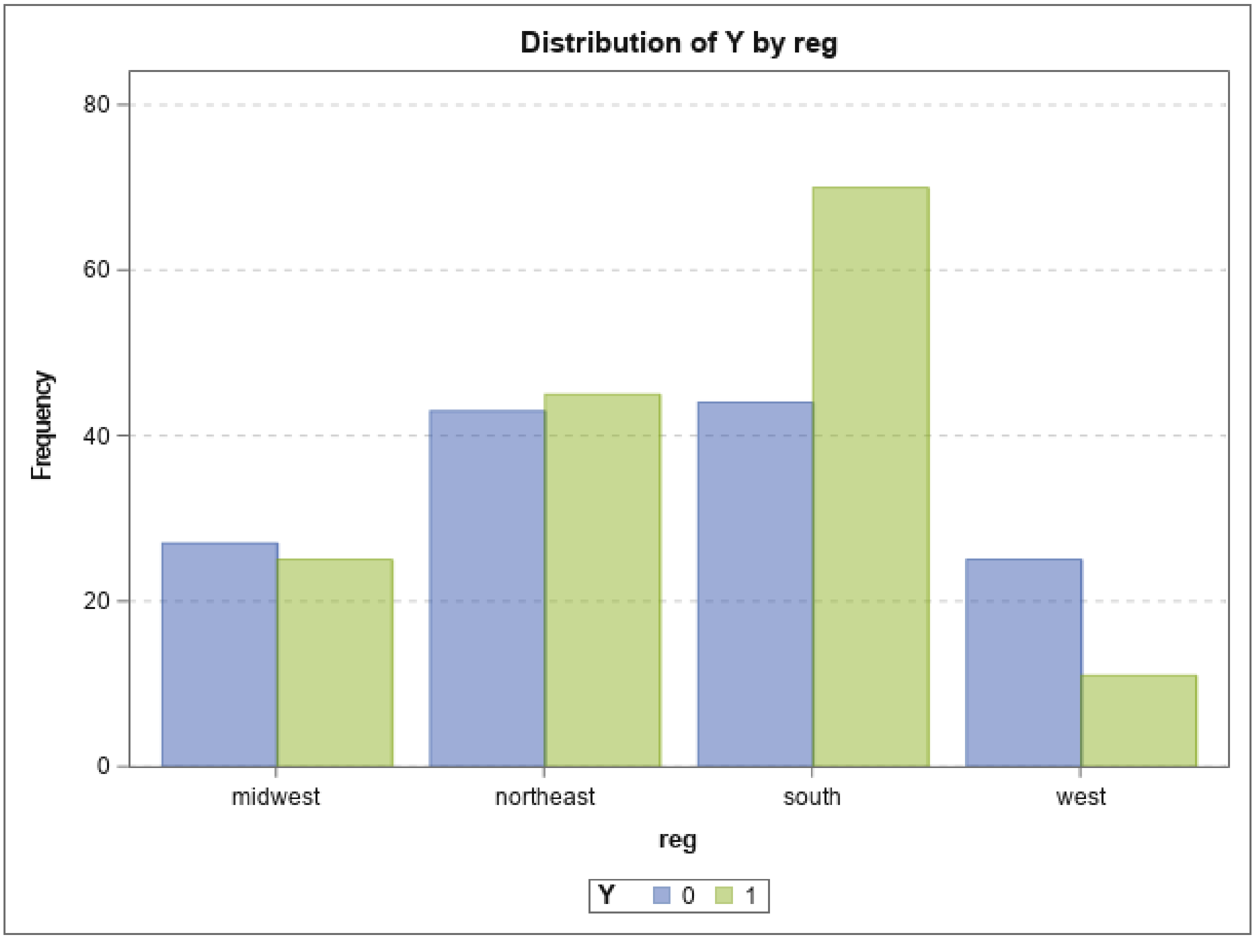

Figure S3: Regions (reg) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

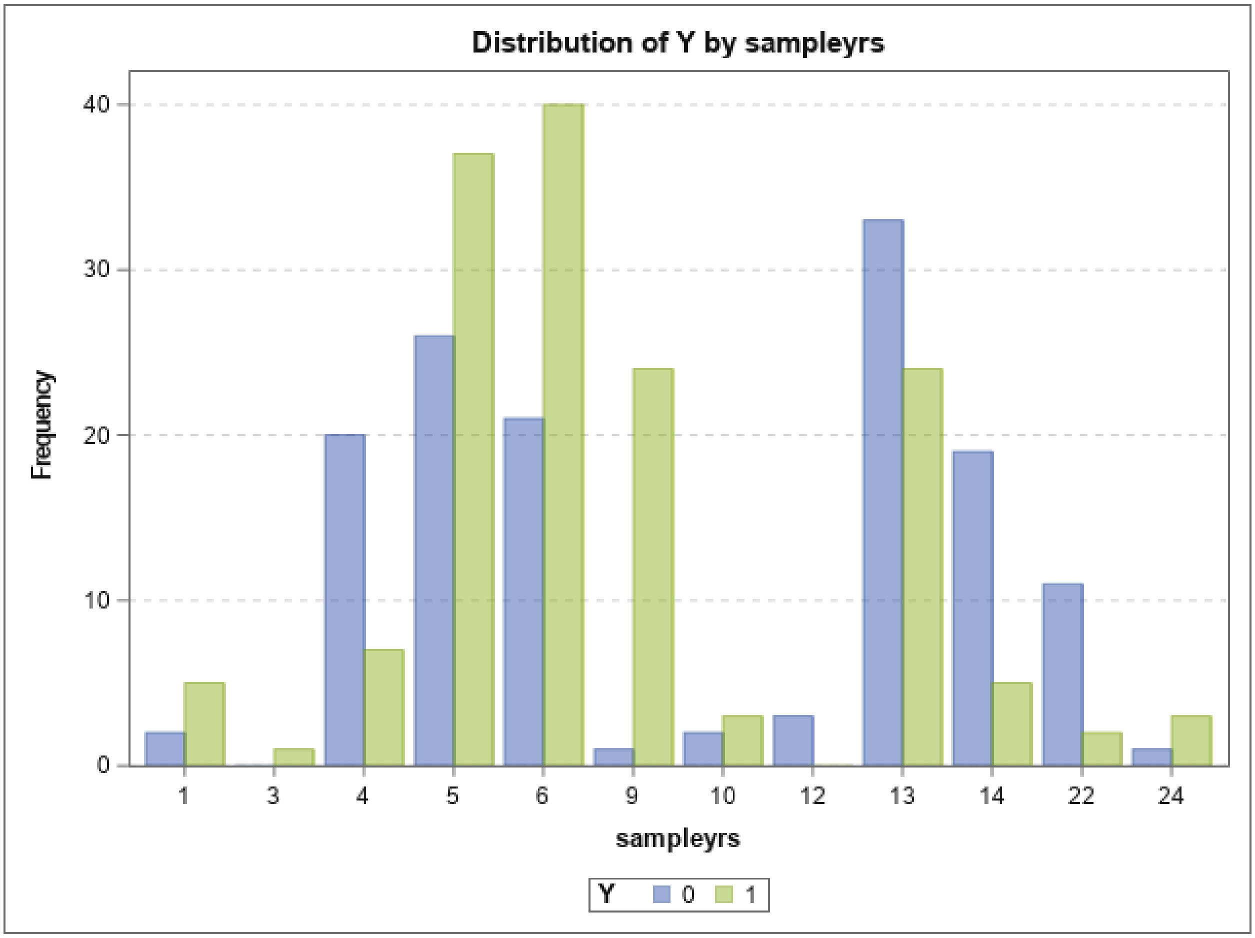

Figure S4: The number of sample years (sampleyrs) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

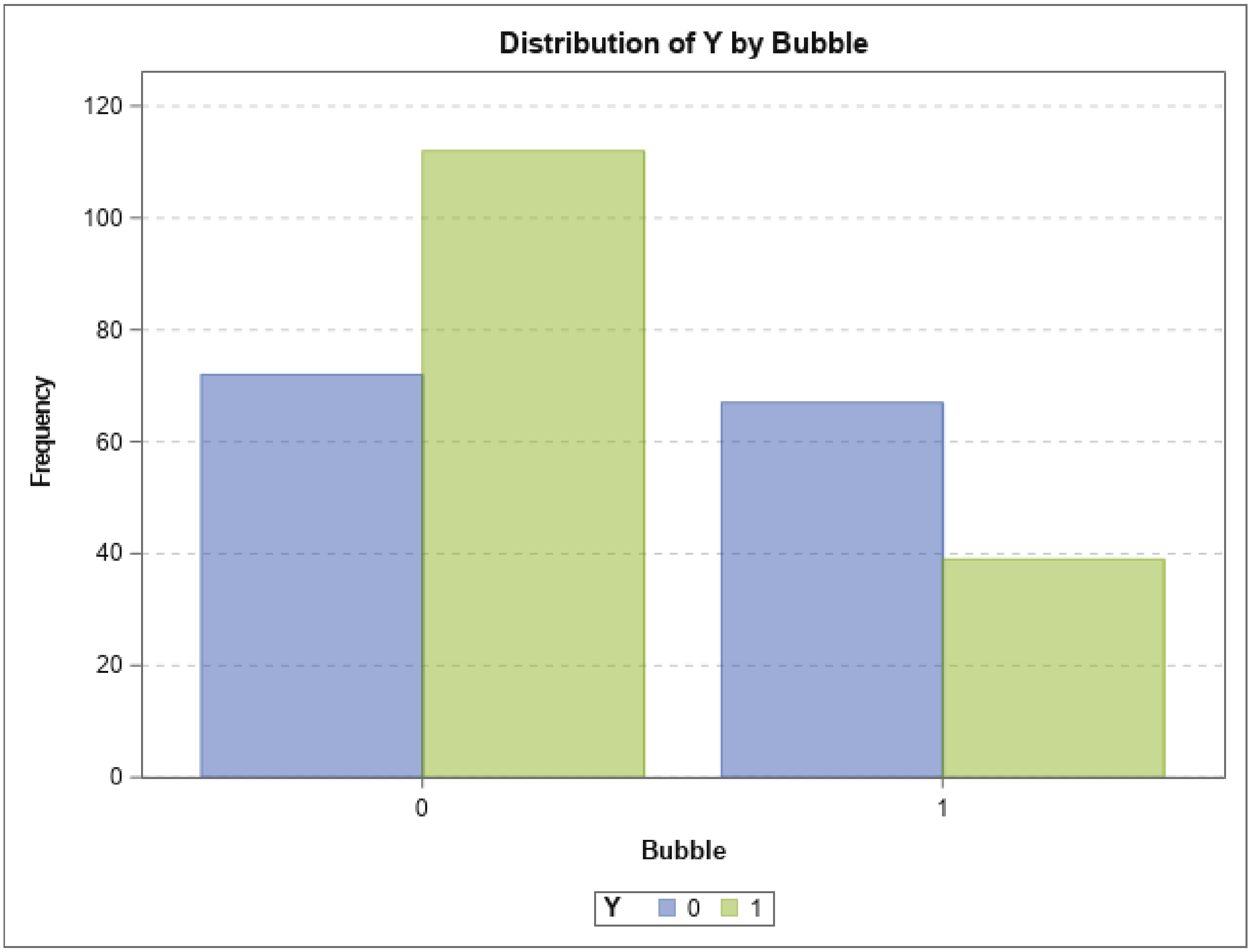

Figure S5: Whether the study included housing bubble years (Bubble=1 if sample includes 2006-2009) in sample and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

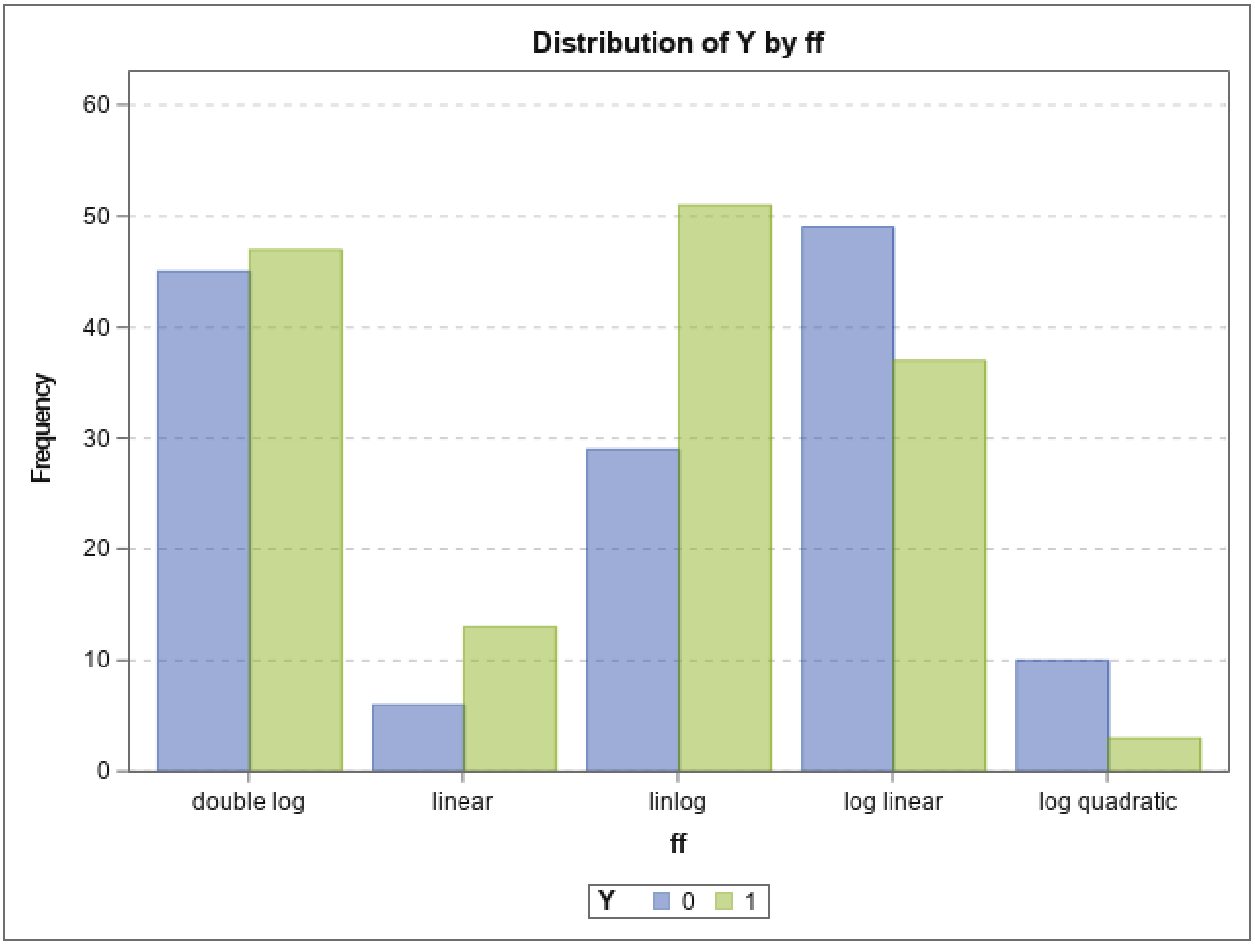

Figure S6: Categories of functional form (ff) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

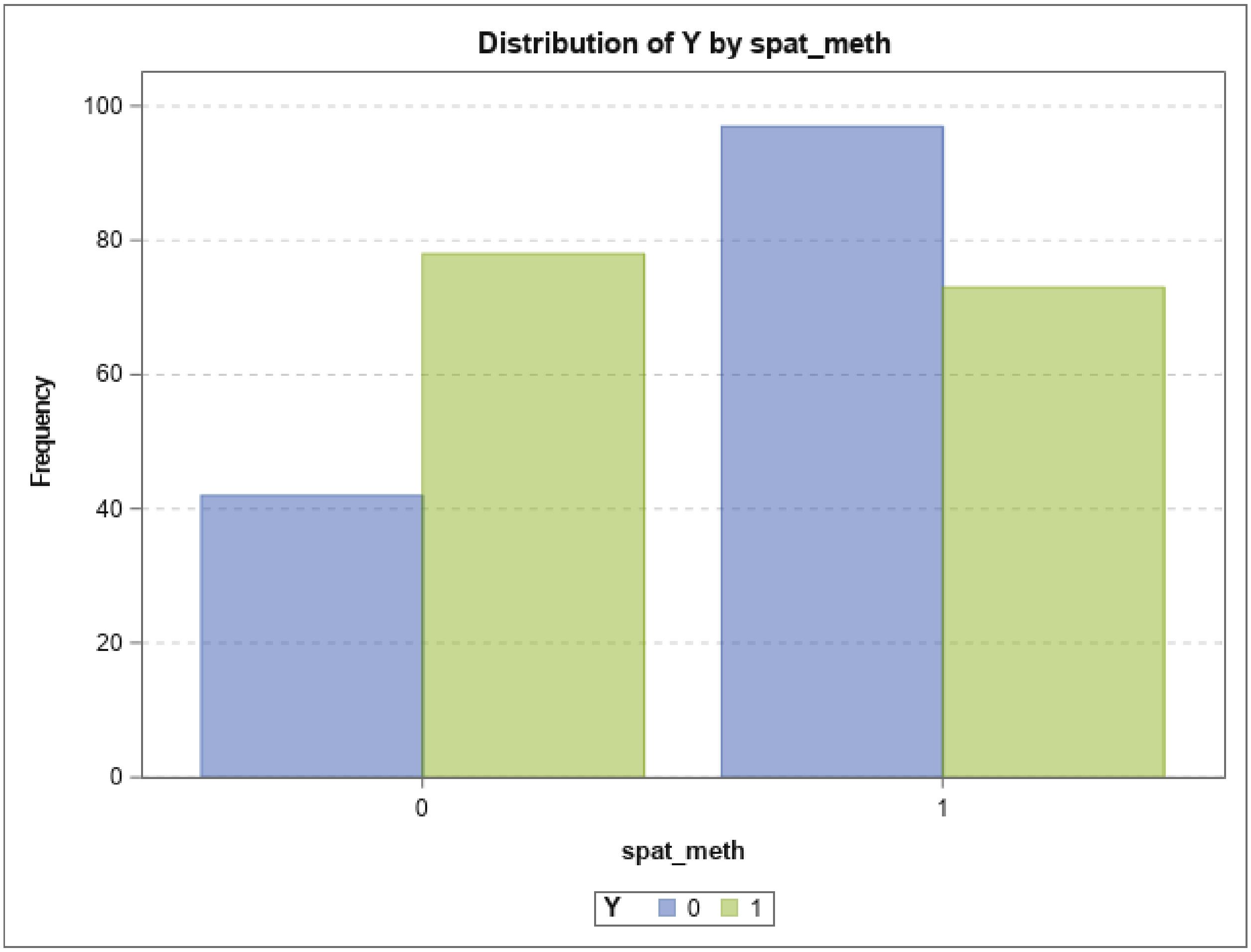

Figure S7: Whether the study addressed spatial dependence (spat_meth=1) or did not (spat_meth=0) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

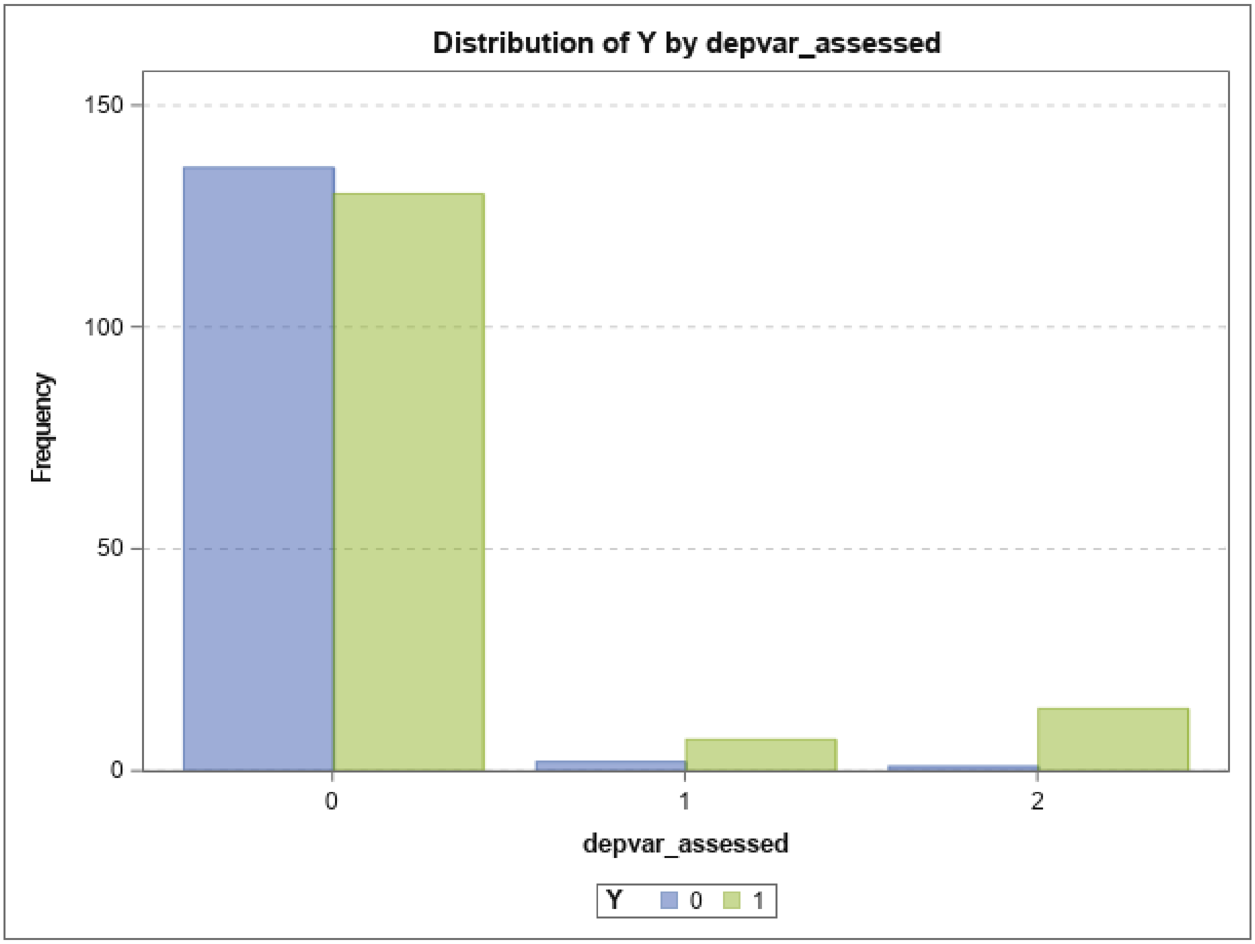

Figure S8: Housing price type (depvar_assessed: 0 sales, 1 assessed, 2 other) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

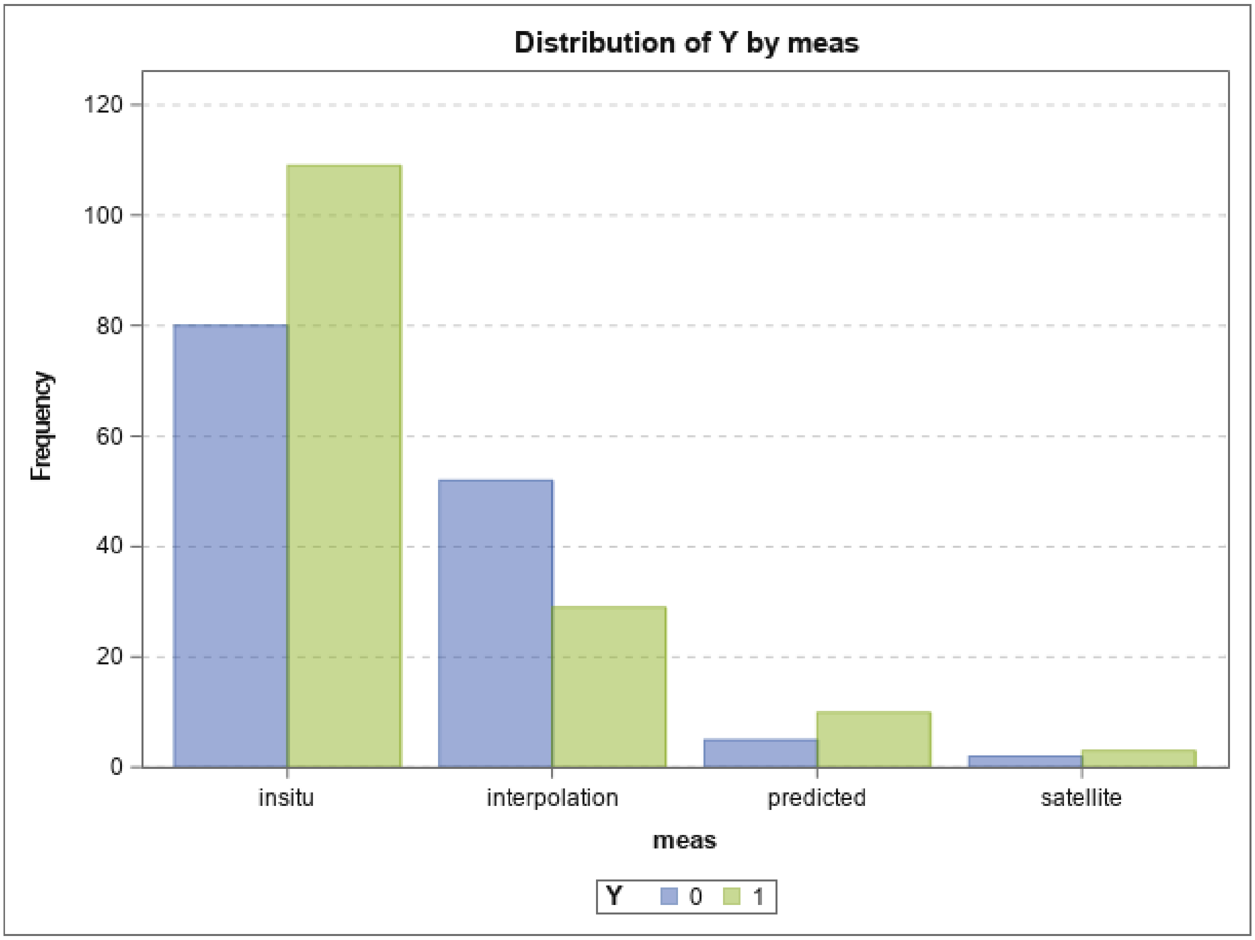

Figure S9: Categories of water quality measurements (meas) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

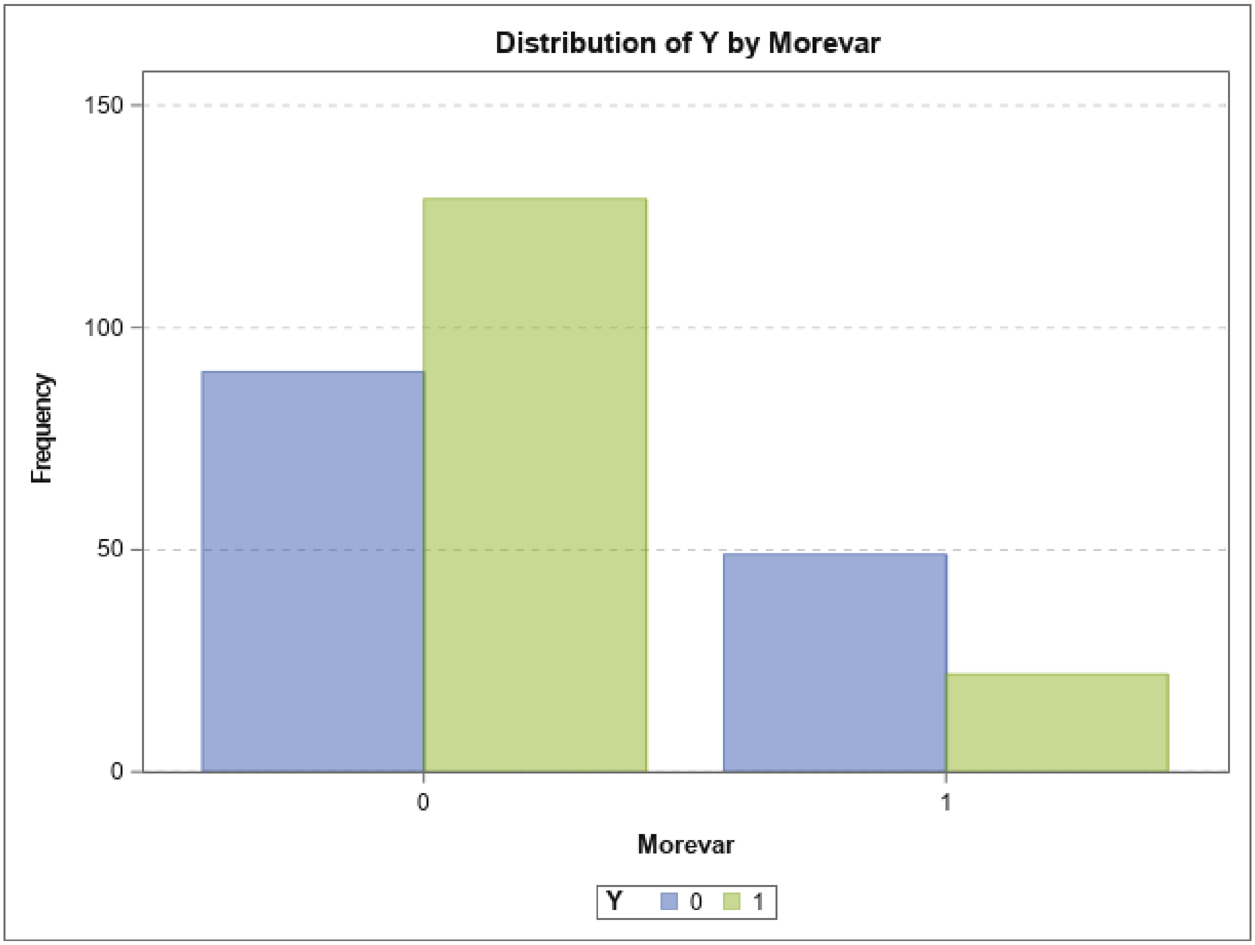

Figure S10: Whether the hedonic model included one water quality variable (morevar=0) or more than one (morevar=1) and distribution of elasticities that are insignificant and/or have unexpected sign (0) vs. elasticities that are significant and have expected sign (1).

Supplementary Tables

Table S1: Water Quality Measures, Expected Sign, Study Citation, and the Number of Observations in Meta-dataset

WQ Measure in Hedonic Model	Expected Sign	Study Citations	Observations
Chlorophyll a	-	Liu et al. 2017; Walsh & Milon 2016; Walsh et al 2011b	18
Dissolved oxygen (DO)	+	Bin & Czajkowski 2013; Netusil et al. 2014	10
E. coli	-	Netusil et al. 2014	5
Fecal coliform	-	Ara 2007; Brashares 1985; Leggett & Bockstael 2000; Netusil et al. 2014	36
Lake trophic state index	-	Feather et al. 1992	2
Light attenuation	-	Guignet et al. 2017; Walsh et al. 2017	57
Nitrogen	-	Liu et al. 2014; Poor et al. 2007; Walsh & Milon 2016; Walsh et al. 2011b	7
Percent water visibility	+	Bin & Czajkowski 2013	2
Phosphorus	-	Liu et al. 2014; Walsh & Milon 2016; Walsh et al. 2011b	6
Salinity	-	Bin & Czajkowski 2013	2
Sediment	-	Liu et al. 2014; Yoo et al. 2014	4
Sedimentation rate	-	Bejranonda et al. 1999	2
Temperature	-	Netusil et al. 2014	6
Total suspended solids	-	Netusil et al. 2014; Poor et al. 2007	7
Turbidity	-	Brashares 1985	2
Trophic state index	-	Walsh & Milon 2016; Walsh et al. 2011b	4
Water clarity (Secchi disk depth)	+	Ara 2007; Boyle et al. 1999; Boyle and Taylor 2001; Gibbs et al. 2002; Horsch & Lewis 2009; Hsu 2000; Kashian et al. 2006; Krysel et al. 2003; Liao et al. 2016; Liu et al. 2014; Michael et al. 2000; Olden & Tamayo 2014; Poor et al. 2001; Walsh et al. 2011a; Zhang & Boyle 2010; Zhang et al. 2015	120

Table S2: Descriptive Statistics for Different Dependent Variables

Dependent variable		Mean	Std Dev	Minimum	Maximum
Y	Same as defined in Table 1	0.521	0.500	0	1
$Y_{0.01}$	Significant with a p-value < 0.01 and theoretically consistent=1	0.393	0.489	0	1
$Y_{0.10}$	Significant with a p-value < 0.10 and theoretically consistent=1	0.597	0.491	0	1

n=290.

Table S3: Robustness check 10% vs. 5% vs. 1% (Model 4: Independence Correlation Structure)

VARIABLES[a]	Dependent (p<0.05) Full Model	Dependent (p<0.05) Four variables dropped	Dependent (p<0.10) Four variables dropped	Dependent (p<0.01) Four variables dropped
Intercept	-0.09	-0.18	0.01	-0.74**
	(0.402)	(0.366)	(0.380)	(0.348)
River	3.53***	1.53	0.60	-0.32
	(1.274)	(1.230)	(1.491)	(0.891)
Estuary	-2.01***	-1.80***	-1.64**	-1.72***
	(0.641)	(0.588)	(0.672)	(0.464)
Nutrients	1.79**			
	(0.758)			
Sediment	-0.11			
	(0.671)			
Bacteria	0.33	0.25	0.44	0.32
	(0.312)	(0.330)	(0.354)	(0.302)
Biochem	-0.11	-0.21	0.01	0.30
	(0.716)	(0.793)	(0.859)	(0.702)
Midwest	-0.50	0.17	-0.03	0.11
	(0.612)	(0.470)	(0.327)	(0.442)
West	-1.71***	-0.25	0.36	0.40
	(0.643)	(0.890)	(1.108)	(0.817)
South	1.76***	2.48***	2.97***	1.78**
	(0.597)	(0.531)	(0.626)	(0.705)
Mean House Price (thousands, 2018$)	1.70E-03	6.00E-04	-0.00	4.00E-04
	(0.001)	(0.001)	(0.001)	(0.001)
Sample Years	-0.03	0.02	0.02	0.02
	(0.051)	(0.040)	(0.045)	(0.035)
Multiple Counties/Subcounties	0.53	0.52	0.98***	0.69*
	(0.454)	(0.424)	(0.293)	(0.362)
Linear	0.06			
	(0.338)			
Spatial Fixed Effects	-0.88**	-0.28	-0.05	-0.25
	(0.378)	(0.383)	(0.436)	(0.317)
Spatial Lag	-0.64**	-0.60*	-0.61	0.10
	(0.303)	(0.310)	(0.393)	(0.533)
Spatial Autocorrelation	-0.88**	-0.82**	-1.04***	-0.26
	(0.346)	(0.327)	(0.324)	(0.243)
Use of Assessed Housing Price	1.96***			
	(0.687)			
Not In Situ	1.50***	0.25	0.07	0.05
	(0.435)	(0.364)	(0.361)	(0.371)

More than One WQ Variable	-1.29**	-1.13**	-1.04*	-0.71*
	(0.576)	(0.471)	(0.597)	(0.395)
Observations	290	290	290	290
QIC	351.96	375.98	362.69	383.88
QICu	332.42	349.12	332.03	361.13

Dependent variable: Y based on either $p<0.05$, $p<0.10$, or $p<0.01$. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Cluster-robust standard errors in parentheses; clustered according to the J=98 study-housing market combinations. For robustness models, we drop *Nutrients*, *Sediment*, *Linear* or *Use of Assessed Housing Price* due to near perfect correlation with dependent variable.

REFERENCES

Ara, S., 2007. The influence of water quality on the demand for residential development around Lake Erie. Doctoral Dissertation, The Ohio State University.

Bejranonda, S., Hitzhusen, F.J., and Hite, D., 1999. Agricultural sedimentation impacts on lakeside property values. Agricultural and Resource Economics Review, 28(2), pp.208-218.

Bin, O. and Czajkowski, J., 2013. The impact of technical and non-technical measures of water quality on coastal waterfront property values in South Florida. Marine Resource Economics, 28(1), pp.43-63.

Boyle, K.J. and Taylor, L.O., 2001. Does the measurement of property and structural characteristics affect estimated implicit prices for environmental amenities in a hedonic model? The Journal of Real Estate Finance and Economics, 22(2-3), pp.303-318.

Boyle, K.J., Poor, P.J. and Taylor, L.O., 1999. Estimating the demand for protecting freshwater lakes from eutrophication. American Journal of Agricultural Economics, 81(5), pp.1118-1122.

Brashares, E., 1985. Estimating the instream value of lake water quality in southeast Michigan. Doctoral Dissertation, University of Michigan.

Feather, T.D., Pettit, E.M. and Ventikos, P., 1992. Valuation of lake resources through hedonic pricing (No. IRW-92-R-8). Army Engineer Inst for Water Resources. Alexandria, VA.

Gibbs, J.P., Halstead, J.M., Boyle, K.J., and Huang, J.C., 2002. A hedonic analysis of the effects of lake water clarity on New Hampshire lakefront properties. Agricultural and Resource Economics Review, 31(1), pp.39-46.

Guignet, D., Griffiths, C., Klemick, H. and Walsh, P.J., 2017. The implicit price of aquatic grasses. Marine Resource Economics, 32(1), pp.21-41.

Horsch, E.J. and Lewis, D.J., 2009. The effects of aquatic invasive species on property values: evidence from a quasi-experiment. Land Economics, 85(3), pp.391-409.

Hsu, T.I., 2000. A hedonic study of the effects of lake-water clarity and aquatic plants on lakefront property prices in Vermont. Doctoral Dissertation, University of Maine.

Kashian, R., Eiswerth, M.E., and Skidmore, M., 2006. Lake rehabilitation and the value of shoreline real estate: Evidence from Delavan, Wisconsin. The Review of Regional Studies, 36(2), p.221.

Krysel, C., Boyer, E.M., Parson, C., and Welle, P., 2003. Lakeshore property values and water quality: Evidence from property sales in the Mississippi Headwaters Region. Submitted to the Legislative Commission on Minnesota Resources by the Mississippi Headwaters Board and Bemidji State University.

Leggett, C.G. and Bockstael, N.E., 2000. Evidence of the effects of water quality on residential land prices. Journal of Environmental Economics and Management, 39(2), pp.121-144.

Liao, F.H., Wilhelm, F.M., and Solomon, M., 2016. The Effects of Ambient Water Quality and Eurasian Watermilfoil on Lakefront Property Values in the Coeur d'Alene Area of Northern Idaho, USA. Sustainability, 8(1), p.44.

Liu, H., Gopalakrishnan, S., Browning, D., Herak, P., and Sivandran, G., 2014. Estimating the impact of water quality on surrounding property values in Upper Big Walnut Creek Watershed in Ohio for dynamic optimal control. In 2014 AAEA Annual Meeting, July 27-29, Minneapolis, MN.

Liu, T., Opaluch, J.J., and Uchida, E., 2017. The impact of water quality in Narragansett Bay on housing prices. Water Resources Research, 53(8), pp.6454-6471.

Michael, H.J., Boyle, K.J., and Bouchard, R., 2000. Does the measurement of environmental quality affect implicit prices estimated from hedonic models? Land Economics, 72(2), pp.283-298.

Netusil, N.R., Kincaid, M., and Chang, H., 2014. Valuing water quality in urban watersheds: A comparative analysis of Johnson Creek, Oregon, and Burnt Bridge Creek, Washington. Water Resources Research, 50(5), pp.4254-4268.

Olden, J.D. and Tamayo, M., 2014. Incentivizing the public to support invasive species management: Eurasian milfoil reduces lakefront property values. PloS one, 9(10), p.e110458.

Poor, P.J., Boyle, K.J., Taylor, L.O., and Bouchard, R., 2001. Objective versus subjective measures of water clarity in hedonic property value models. Land Economics, 77(4), pp.482-493.

Poor, P.J., Pessagno, K.L., and Paul, R.W., 2007. Exploring the hedonic value of ambient water quality: A local watershed-based study. Ecological Economics, 60(4), pp.797-806.

Walsh, P.J. and Milon, J.W., 2016. Nutrient standards, water quality indicators, and economic benefits from water quality regulations. Environmental and Resource Economics, 64(4), pp.643-661.

Walsh, P.J., Milon, J.W., and Scrogin, D.O., 2011a. The spatial extent of water quality benefits in urban housing markets. Land Economics, 87(4), pp.628-644.

Walsh, P.J., Milon, J.W., and Scrogin, D.O., 2011b. The property-price effects of abating nutrient pollutants in urban housing markets. In Economic Incentives for Stormwater Control. Ed. Thurston, H.W. CRC Press - Taylor & Francis Group, LLC, Boca Raton, FL.

Walsh, P., Griffiths, C., Guignet, D., and Klemick, H., 2017. Modeling the property price impact of water quality in 14 Chesapeake Bay Counties. Ecological Economics, 135, pp.103-113.

Yoo, J., Simonit, S., Connors, J.P., Kinzig, A.P., and Perrings, C., 2014. The valuation of off-site ecosystem service flows: Deforestation, erosion and the amenity value of lakes in Prescott, Arizona. Ecological Economics, 97, pp.74-83.

Zhang, C. and Boyle, K.J., 2010. The effect of an aquatic invasive species (Eurasian watermilfoil) on lakefront property values. Ecological Economics, 70(2), pp.394-404.

Zhang, C., Boyle, K.J., and Kuminoff, N.V., 2015. Partial identification of amenity demand functions. Journal of Environmental Economics and Management, 71, pp.180-197.