Using Hedonic Property Models to Value Public Water Bodies: A Note Regarding Specification Issues

by

Nicholas Z. Muller

October 2007

MIDDLEBURY COLLEGE ECONOMICS DISCUSSION PAPER NO. 07-21

DEPARTMENT OF ECONOMICS
MIDDLEBURY COLLEGE
MIDDLEBURY, VERMONT 05753

http://www.middlebury.edu/~econ

Using Hedonic Property Models to Value Public Water Bodies: A Note Regarding Specification Issues.

Nicholas Z. Muller
Department of Economics
Environmental Studies Program
Middlebury College
Middlebury, VT 05753
nmuller@middlebury.edu
802-443-5918
July 6^{th}, 2007

Abstract

The hedonic literature has established that public water bodies provide external benefits that are reflected in the value of nearby residential real estate. The literature has employed several approaches to quantify these nonmarket services. With a residential hedonic model, this paper tests whether model specification affects resource valuation using an actively managed reservoir in Indiana and a passively managed lake in Connecticut. The results indicate that valuation is quite sensitive to model specification, and that omitting either the waterview or waterfront variables from the hedonic function likely results in a misspecificed model. The findings from this study are important for researchers and public agencies charged with managing water resources to bear in mind as the external benefits from existing or proposed man-made lakes and reservoirs are estimated. Therefore, while it requires considerably more effort to determine which properties are in waterfront loactions and which properties have a view, the potential misspecification of "distance-only" models likely justifies these extra research costs. Further, the findings in this analysis call into question results from "distance-only" models in the literature.

AGU Index Terms: 6304 Benefit-cost analysis
Keywords: Hedonic Model, Public Water Bodies, Valuation

Acknowledgements: The author wishes to thank the Glaser Progress Foundation for funding this project. I would also like to thank Robert Mendelsohn, Nathaniel Keohane, and participants in the Yale Environmental Economics Seminar for their helpful comments.

1 Introduction

Economists using hedonic property models have established that public water bodies provide external benefits that are reflected in the value of nearby residential real estate. The literature has employed a number of approaches to quantify these nonmarket services. Several studies have used distance to the water to measure the nonmarket services generated by public water bodies [Mahan, Polasky, Adams, 2000]. Other authors have controlled for both distance and view services [Brown, Pollakowski, 1977]. Another strategy measures three distinct services: access (measured by distance to the water), a view of the water, and adjacency [Lansford, Jones, 1995; Loomis, Feldman, 2003]. Since hedonic models are often used to value water bodies, specification is especially critical. Bias due to the omission of relevant variables affects parameter estimates in the regression function which, in turn, impacts valuation. This paper tests how amenity measurement influences the benefit estimates for public water bodies.

The importance of specification is tested using the three approaches to amenity measurement outlined above. Each model is used to value two public lakes. Along with a standard assortment of structural, economic and neighborhood variables, the first model employs only distance to value the lakes. The second model includes both distance and water view, and the third model uses distance, water view and waterfront regressors to measure the services produced by the lakes. This experimental design facilitates an assessment of how specification, through amenity measurement, affects valuation. A Hausman test [Hausman, 1978] is used to test whether distance to the water is an endogenous variable when the water view and the waterfront variables are omitted from the hedonic model. Further, the Hausman test is used to test whether water view is endogenous when the waterfront control is omitted.

The data consist of sales from housing markets surrounding two lakes; Lake Monroe, an actively managed Army Corps of Engineers reservoir in Indiana, and Candlewood Lake, a passively managed lake in Connecticut. Lake Monroe is managed for flood control, and as a source of irrigation and drinking water. Management strategies dictate dramatic seasonal fluctuations in water levels. In contrast, Candlewood Lake is predominantly a recreation resource which does not experience drastic fluctuations in water levels. The hedonic models are applied to each sample separately.

This empirical setting also addresses an interesting management issue raised by Loomis and Feldman [2003] and Lansford and Jones [1995]. These authors indicate that anthropogenic reductions in lake levels constitute a significant disamenity, negatively affecting housing prices. This finding has important implications for cost benefit analyses of lakes managed for irrigation and hydroelectricity production. Although this analysis does not directly estimate the implicit price of water levels, I qualitatively test this hypothesis in the following manner. Fluctuations in water levels likely have the most tangible market implications for waterfront properties. Therefore, if water levels do affect the price of such homes, one would expect to observe a difference in the waterfront premium affixed to properties on Candlewood Lake, which has a stable water level, and Lake Monroe which experiences substantial seasonal lake level changes. In the current setting, a significantly smaller premium for waterfront houses on Lake Monroe would provide qualitative support for the hypothesis that active lake management affects the amenity value of lakes.

The results indicate that valuation is quite sensitive to model specification. Using the Lake Monroe sample, the distance-only model produces an annual amenity value that is 40% greater than that produced when controlling for view and adjacency. Using the Candlewood Lake sample, the total value generated by the distance-only model is 25% greater than the total amenity value produced when controlling for view and adjacency. The difference in total lake values is driven by the parameter estimate for distance to the lake: in both samples this coefficient declines precipitously when view and adjacency are added to the distance-only model. The Hausman test provides strong evidence that lake distance is endogenous when the view and waterfront controls are omitted from the hedonic price function. Taken together, it is likely that the lake distance coefficient suffers from omitted variable bias in distance-only models. The strong effect that this bias bears upon the total value of the lakes suggests that the application of distance-only models to estimate external benefits generated by public water bodies is quite problematic. Finally, waterfront and access services are generated by the passively managed, Candlewood Lake. In contrast, only view and access services are provided by Lake Monroe, the reservoir. This resource does not generate waterfront services. This may suggest that fluctuations in water levels due to management practices produces a significant disamenity for waterfront homeowners.

2 Methods

2.1 Model Specification

The econometric framework is built upon a log-linear hedonic model. This specification was chosen for a number of reasons. First, imposing a linear form on the hedonic function implies, undesirably, that perfect repackaging of property characteristics is possible [Freeman, 2003]. Second, the hedonic models employ proxy variables and Cropper, Deck and McConnell [1988] suggest that linear and log-linear functional forms outperform other forms when proxy measurements are employed. Third, since the estimated regression functions are used to generate amenity values, it is preferable to maintain a relatively simple form. Employing more complicated functional forms impedes both interpretation of the parameter estimates and their use in valuation. Finally, preliminary regression results indicated that the log-linear form exhibited an excellent fit to the data. The full list of variables included in the regression models is shown in table 1. Descriptive statistics are shown in tables 2 and 3.

Using the log-linear form, the empirics begin with the most parsimonious specification found in the literature. The benchmark strategy, model (1), employs the distance to the water (D) as the solitary amenity measure in the hedonic price function. Mahan, Polasky, and Adams [2000] employed the distance-only model to value urban wetlands. Model (2) controls for view services (V) and distance. Brown and Pollakowski [1977] used this approach to value shoreline. Model (3) incorporates adjacency (F), view, and distance. This specification stems from the work of Lansford and Jones [1995] and Loomis and Feldman [2003]. In a departure from the literature, this study uses a more flexible functional form for proximity; models (1,2,3) are estimated employing cubic splines to capture the relationship between proximity and price. The $\mathrm{m}_\lambda(\mathrm{D})$ notation represents the spline function.

This experimental design tests the implications of specification, through environmental amenity measurement, for valuation. Since the structural, economic and neighborhood controls are the same in each model, the differences in values produced each model are directly attributable to the different amenity measurement strategies.

$$log(P_i) = \alpha_0 + \alpha_1\mathbf{X}_i + \alpha_2\mathbf{N}_i + m_\lambda(D_i) + \varepsilon_i \tag{1}$$

where: $\alpha_0, ..., \alpha_n, \lambda$ Statistically Estimated Parameters
$\mathbf{X}_i$ Physical/structural characteristics property (i)
$\mathbf{N}$ Neighborhood characteristics
D Straight-Line Distance to Nearest Shore Point
P sales price ($2000)
ε error term

$$log(P_i) = \alpha_0 + \alpha_1\mathbf{X}_i + \alpha_2\mathbf{N}_i + m_\lambda(D_i) + \alpha_4 V_i + \varepsilon_i \tag{2}$$

where: V Water view Binary

$$log(P_i) = \alpha_0 + \alpha_1\mathbf{X}_i + \alpha_2\mathbf{N}_i + m_\lambda(D_i) + \alpha_4 V_i + \alpha_5 F_i + \varepsilon_i \tag{3}$$

where: F Water front Binary

Although none of the above models address the necessarily ad hoc nature of the hedonic property model, prior to estimation there is reason to believe that model (3) embodies the preferable approach. First, the goal in the hedonic function is to use an amenity measurement that mimics consumers' perceptions of the amount of the amenity embodied in each property [Freeman, 2003]. Of the models found in the literature, model (3) is the most effective at capturing the criteria that consumers use when ranking and evaluating residential properties close to a public water body. These criteria are evident in how realtors advertise such properties: waterfront homes, homes with a water view, and homes a number of blocks from the water. Omitting any of these attributes reduces the correspondence between amenity measurement in the hedonic function and the criteria used by consumers.

Model (3) is also preferable because of the clear potential for measurement error in models (1) and (2). In model (1), distance to the water presumably captures the bundle of water-related services embodied in each property. In model (2) view likely captures the price effects of both view and adjacency services since the view and (omitted) adjacency variables are likely to be highly correlated. In this setting, the role of distance changes. This variable now captures the effect which the degree of access to the water bears upon housing prices. Model (3) eliminates this imprecision by including explicit controls for each of these three services.

The non-parametric approach to modeling the relationship between distance to the lakes and housing prices using cubic splines requires determining the appropriate smoothing parameter (λ), which controls the number of knots (inflection points) in the spline function. This analysis uses a data-driven mechanism in order to select an appropriate value for (λ). Specifically, a variety of degrees of smoothness for the spline function are tested and the fit of each is evaluated using the leave-one-out-method (Hardle, 1990; Stone, 1974). The mean square error (MSE) is estimated for each sample.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(P_i - \hat{P}_i \right)^2 \qquad (4)$$

where: P_i observed sale price of property (i).

$\hat{P}_i$ model prediction of sale price for property (i).

The value of (λ) that minimizes the MSE is selected as the preferred smoothing parameter for the econometric and valuation experiments. The optimal (error-minimizing) smoothing parameter is permitted to vary between the two samples.

For the remaining portions of this paper, values stemming from the distance to the water variable are referred to as access services.

2.2 Valuation Methodology

The hedonic literature has explored using estimated hedonic property models to determine the welfare effects of changes in environmental quality [Freeman, 2003; Palmquist, 1992]. Using the log-linear functional form, the marginal willingness-to-pay (MWTP) for the environmental amenity is equivalent to the regression coefficient corresponding to the amenity (α) times the price (P), [Loomis, Feldman, 2003].

$$MWTP = P \times \alpha \tag{5}$$

However, Freeman [2003] and others have discussed the difficulties involved with using the implicit prices estimated in hedonic property models to compute welfare effects for non-marginal changes in an environmental amenity. In this study, there are two types of environmental changes that affect the value of residential real estate. Variation in the distance between properties and the water body is a continuous function, so small relocations along the distance gradient may be viewed as marginal changes. In contrast, the waterview and waterfront categories are discrete categories; either a property is adjacent to the water or it is not. Movement of a property into and out of this category is clearly a non-marginal change. Therefore, in light of Freeman [2003], attempting to assess the welfare implications of the waterfront and water view amenities using implicit prices is problematic.

However, Palmquist [1992] points out that if the number of properties affected by an amenity is small relative to a larger real estate market, then the hedonic price function does not shift for non-marginal changes in the amenity. In this case, the change in price associated with changes to the amenity is correctly interpreted as a welfare measurement. This approach also requires the assumption of low moving costs. In each sample, this study focuses on roughly 300 properties within larger real estate markets. The assumptions embodied in the work of Palmquist [1992] are invoked in order to use the methods outlined above to compute the aggregate welfare effect of the lakes in this study.

In order to convert the market prices, which represent the present value of a stream of rental income, to the more conventional welfare measure of annualized benefits, Freeman [2003] points out that the following calculations are necessary. Using (r) the social rate of return on investment, and (t) the local property tax rate, and the housing price (P_i), annualized rental value of property (i), denoted (R_i) is equivalent to equation (6).

The Lake Monroe sample is situated within the Bloomington, Indiana market. This city contains 26,000 households. Similarly, the Candlewood Lake sample is nested within both the New Milford and New Fairfield markets; these towns contain 15,000 households. Clearly the number of properties affected by these two amenities is quite small relative to each residential real estate market.

$$R_i = P_i(r + t) \quad (6)$$

With estimates of (t) and (r) being readily available, using this approach all values reported in this study are annualized benefits; interpreted as constant year-2000 dollars per year.

The approach to estimating the value attributed to each service involves first using the statistically estimated models to derive predicted real estate prices. Next, the amount of each service (access, view, adjacency) is systematically manipulated to determine the corresponding change in price. Holding constant all other variables in the hedonic price function isolates the change in price due to varying the levels of each service. Employing model (3), the log-linear form dictates that the price of a parcel of property in a waterfront location (P_f) is equivalent to the following expression. (Note that the predicted value produced by a log-linear model requires the addition of one-half the empirical variance of the error term: υ=1/2(s^2), [Heien, 1968].)

$$(P_f) = \exp^{(\alpha_0+\alpha_1\mathbf{X}+\alpha_2\mathbf{N}+m_\lambda(D)+\alpha_4 V+\alpha_5 1+\upsilon)} \quad (7)$$

The price of an identical home not in a waterfront location is equal to:

$$(P_0) = \exp^{(\alpha_0+\alpha_1\mathbf{X}+\alpha_2\mathbf{N}+m_\lambda(D)+\alpha_4 V+\alpha_5 0+\upsilon)} \quad (8)$$

For each sample (s), the parameter estimates in the hedonic function are used to estimate (P_{fs}) and (P_{0s}) using the mean independent variable values for waterfront homes. By holding constant all covariates other than the waterfront variable, we extract the price change strictly attributable to adjacency services. The adjacency premium for sample (s), denoted (Φ_s) in equation (9), is expressed as a proportion of the waterfront price.

$$\Phi_s = \frac{(P_{fs}) - (P_{0s})}{(P_{fs})} \quad (9)$$

Converting the premium to monetary units simply involves multiplying (Φ_s) times the price of an average waterfront parcel. In order to generate an aggregate adjacency service value, the following calculations are necessary. First, the housing density in each sample area, denoted (D_s), is determined using U.S. Census data (U.S. Census, 2006). Next, the total number of homes in waterfront locations (H_{fs}) is determined by multiplying the sample housing density (D_s) by the land area in water front zones (A_{fs}), (see appendix A.2).

$$H_{fs} = (D_s) \times (A_{fs}) \tag{10}$$

The aggregate value of waterfront services for each sample (WF_s) is found by multiplying the observed average price of a waterfront home $\left(\bar{P}_{fs}\right)$ by the number of homes in waterfront locations (H_{fs}) and by the estimated waterfront premium (Φ_s).

$$WF_s = (H_{fs}) \times (\bar{P}_{fs}) \times (\Phi_s) \tag{11}$$

The valuation procedure for view services proceeds in an analogous manner. The estimation of aggregate access values relies on a slightly different approach than that used for adjacency and view values. Valuing access services requires first determining the distance at which the water bodies' influence on housing prices diminishes. In order to determine the boundary of each lake's effect on housing prices, the spline model fitted to the lake distance variable, $m_\lambda(D)$, is examined for a local minimum point. Using the Lake Monroe sample, the $m_\lambda(D)$ function minimizes at 1.8 miles from the lake: this point is used as the boundary of the lake's influence on property values. Using the Connecticut sample, the fitted spline model, $m_\lambda(D)$, minimizes at 1.7 miles from the lake. This point is used as the boundary of the lake's influence on property values for the non-parametric models applied to the Connecticut sample.

For each sample, the number of homes in the zone from the lakeshore to the sample-specific boundary points (1.7 and 1.8 miles) is estimated. This necessitates determining the total land area in this lake-influenced zone and multiplying the area times the observed housing density (U.S. Census, 2006). The number of homes in the land area corresponding to sample (s) is denoted (H_s).

Then the estimated regression models are used to predict the mean price (P_{ms}) for properties in each sample, excluding properties beyond the boundary points in each sample. This entails inserting the mean values for each covariate, from the spatially restricted samples, into the estimated model. Next, the boundary price is generated by inserting the boundary distance into the estimated regression model, while all other independent variables are held at their mean levels. This isolates the price effect of relocating a property from the mean distance to the boundary distance. The access premia (Ψ_s) are calculated by subtracting the predicted price of a home with mean attributes at the boundary point (P_{bs}) from (P_{ms}) and then dividing by (P_{ms}).

$$\Psi_s = \frac{(P_{ms}) - (P_{bs})}{(P_{ms})} \qquad (12)$$

To compute the total access value generated by eack lake, the number of homes is multiplied times the observed mean house price for homes within the boundary distance ($\bar{\mathrm{P}}_s$) times the estimated access premium. The aggregate access value for each sample is denoted (AC_s).

$$AC_s = (H_s) \times (\bar{P}_s) \times (\Psi_s) \qquad (13)$$

2.3 Data

Lake Monroe was constructed in 1965 by the U.S. Army Corps of Engineers as a flood control reservoir, an irrigation source and as the primary source of drinking water for Bloomington, Indiana. The 10,750 acre lake spans two counties in south central Indiana. Candlewood Lake was constructed in 1929 as a source of hydroelectric power. It is managed by the Northeast Generating Company. The 5240 acre lake spans two counties in northwestern Connecticut.

Definitions of the variables are provided in table 1, while descriptive statistics are shown in tables 2 and 3. The Lake Monroe sample consists of approximately 330 sales disclosures gathered at the Monroe County (Indiana) Property Tax Assessor's Office. This data set includes sales of residential properties in three townships proximate to the lake from 1999 to 2001. The Candlewood Lake sample consists of approximately 320 sales disclosures gathered using a real estate records service. The sample includes sales from 1999 to 2003 in two towns nearby to the lake. Sales disclosures were deleted from the samples if the buyer and seller had the same last name, if the sales were tax exempt or if certain critical data were missing. The tax assessors also provided property tax rates for each town in the samples. The mill rate (t) is 0.03 in the Connecticut sample and (t) is 0.01 in the Indiana sample. Although a number of values for the rate of return on investment (r) are possible, (r) is set to 0.04 for both samples.

The environmental variables include the straight-line distance from each property to the nearest shore point on the lakes (measured in miles), whether the property has a view of the lake, and whether the property is adjacent to the lakeshore. All properties in each sample were visited in order to determine if the property is adjacent to the lake or if it has a view of the lake. The properties were visited during the summer months, so it is likely that certain properties that do not have summertime views have a view in the winter when the predominantly deciduous trees in each study area lose their leaves.

The dependent variable is the natural log of the sales price adjusted to \$2000. In models (1), (2), and (3) the physical and structural characteristics include; acreage of the lot, living area in the structure, and number of stories in the structure. This approach relies on the living area of each structure as a proxy for more detailed structural characteristics such as the number of bedrooms, bathrooms, and fireplaces. Brookshire et al. [1982] employ this approach. The Lake Monroe sample includes condominiums, mobile homes and single-family houses. Therefore, the models applied to that sample include dummy variables for condominiums and mobile homes. Single-family homes are the default case. In the Candlewood Lake sample all observations are single-family houses. The inclusion of these particular house and property attributes is rooted in previous residential hedonic models [Brown, Pollakowski, 1977; Earnhart, 2001; Mahan, Polasky, Adams, 2000].

The vector of neighborhood characteristics includes distance to the nearest town in both linear and quadratic forms. Both samples include measurements of the year in which the sale occurred. The year of sale variables are included as a proxy measurement for temporal changes in each local housing market. In particular, these variables are intended to capture annual variation associated with property tax assessments and property tax rates; both of which have clear implications for housing prices. Since these are binary variables it necessary to omit one of the year of sale variables from the models. The 2001 variable is omitted from the models. Further, the annual average prime rate is included as a proxy variable for mortgage interest rates and annual GDP is included to proxy for aggregate income. Structure age data was only available for the Candlewood Lake sample. In the Lake Monroe sample, properties were drawn from three townships and two elementary school districts. The Candlewood Lake sample contains properties in two towns. In both cases, these geographic distinctions function as proxy measurements for neighborhood and community characteristics.

3 Results

3.1 Econometrics

Table 4 displays the regression results for both samples. The column headings indicate the sample and the model specification (shown in parentheses). All regression analyses were conducted using ordinary least squares with heteroskedasticity-robust standard errors, [White, 1990]. While included in the models, the covariates that were consistently not significant at $\alpha = 0.10$ are omitted from Table 2. In each of the models, the covariates measuring the structural aspects of the properties behave as intuition would predict. The positive impact which lot size and the size of the structure have on price are notable examples of this conformity. This lends a degree of credibility to the models that is desirable given the use of the regression coefficients to value the non-market services produced by the lakes.

The most notable econometric result is the difference in the mix of services produced by each lake. Using model (3), adjacency and access services are produced by the passively managed Candlewood Lake. View services are marginally significant. View and access, but not adjacency, are produced by Lake Monroe, the reservoir. The coefficient associated with adjacency in the Candlewood Lake sample is approximately 0.61. Thus, the price of a home adjacent to the lake is approximately 60% greater than an equivalent home not in a waterfront location. Although marginally significant, the view coefficient is 0.07: this implies that a view of the lake increases a property's value by 7% above an equal property without a view. There is no premium affixed to waterfront properties in the Lake Monroe sample. Another important econometric result evident in table 4 is that the first segment of the spline function, $\mathrm{m}_\lambda(\mathrm{D}_1)$, measuring distance to the water is negative and statistically significant in each model applied to both samples. Intuitively, this implies a premium for properties located close to the lakes. However, the magnitude of the coefficient corresponding to the first spline segment is quite sensitive to model specification. In both samples, the coefficient corresponding to the first spline segment is largest in model (1). Comparing model (1) to model (2) applied to the Lake Monroe sample, this coefficient declines from 0.26 to 0.15: a 40% reduction. Employing the Candlewood Lake sample, again comparing model (1) to model (2), the coefficient on the first spline segment drops from 0.18 to 0.08, a 55% decline. The decrease in the lake-distance coefficient when the view control is added to the model suggests that the measurement of distance to the water is capturing view services when the waterview variable is omitted in model (1).

Using the Lake Monroe sample, adding the adjacency control causes no change to the coefficients affixed to water view or lake distance in model (2). In contrast, employing the Candlewood Lake sample, adding the adjacency regressor reduces the view coefficient estimated using model (2) from 0.30, significant at $\alpha = 0.01$, to 0.07 approaching significance at $\alpha = 0.10$. The reduction in the view coefficient when the waterfront variable is added indicates that the view variable is capturing waterfront services in model (2). The marked sensitivity of the coefficients for distance to the water in both samples and water view in the Connecticut sample evident in table 4 are both strongly suggestive that model (1) and (2) are misspecified. A Hausman test is used to test for evidence of endogeneity in these regressors in models (1) and (2).

The results of the Hausman test are shown in table 5. Significant t-statistics indicate evidence of engodeneity for the variable shown in the right-hand column of table 5. Hence, it is clear that when used solitarily, in model (1), the lake distance variable is endogenous in both samples. The Hausman test reveals that using distance to the lake as the only amenity measurement in the hedonic function likely results in a misspecified model. This result meshes with the changes in the magnitude of the lake distance coefficient between models (1) and (2) observed in table 4. The significant decrease in the size of the distance coefficient suggests that in model (1) lake distance captures some portion of view and adjacency services which are omitted from model (1).

When model (2) is employed, the Hausman test reveals evidence that the view covariate is endogenous only using the Candlewood Lake sample. So, the Hausman test indicates that model (2) is likely to be misspecified when using this sample. This result meshes with the decrease in the magnitude and the significance of the view coefficient upon adding the control for adjacency services. This significant reduction suggests that in model (2) view captures some portion of adjacency services which are omitted from model (2). This holds only for the Candlewood Lake sample. The following general finding emerges from the Hausman test; the coefficients corresponding to the variables which the Hausman test indicates are likely endogenous are biased upwards. The remaining question is to explore how this bias affects valuation.

3.2 Valuation

Table 6 displays the valuation methodology as it is applied to model (3). Recall that the mill rate (t) is 0.03 in the Connecticut sample and (t) is 0.01 in the Indiana sample. The rate of return on investment (r) is set to 0.04 for both samples. To compute the annual value derived from access services from Lake Monroe, the method involves multiplying the annualized value of a property with average characteristics (all independent variables except distance to the lake are set to the sample mean values), which in this case is $130,000 times (r+t = 0.05), times the access premium of 0.19. This yields an annual value of access services for an average property of $1235. Then, since there are 1710 homes within 1.8 miles of the lake, the per unit premium is mutliplied times 1710 yielding an estimate of total annual access services of $2.0 million. This procedure is repeated for each service in each model and for both samples. Table 7 displays the valuation results derived from the application of these methods to all of the estimated models.

Employing model (1), the total lake values are $15.6 million and $4.6 million attributed to Candlewood and Monroe lakes, respectively. On a per (property) unit basis, the total lake values are $5270 and $2700 (again for Candlewood and Monroe lakes, respectively). When the view covariate is added to model (1), the total amenity value of both lakes declines significantly: employing model (2) the values for Candlewood and Monroe are $9.8 million and $2.8 million and the per unit values are $3310 and $1640, respectively. The total amenity value attributed to both lakes drops by roughly 40%. In both samples, the decrease in total amenity values is driven by the change in the magnitude of the effect that access services have on sales prices. Recall that adding the view control to model (1) reduced the magnitude of the effect that proximity to the lakes has on the price of property. This is clearly reflected in the difference in access services estimated using model (1) compared to model (2): access services decline by 46% in the Indiana sample, and by 50% in the Connecticut sample. Thus, the change in the parameter estimate for lake-distance has a dramatic effect on the total lake values.

Since the waterfront coefficient is insignificant in model (3) applied to Lake Monroe, there is no change to the total lake value when comparing the values predicted by model (2) and model (3): it remains $2.8 million. In contrast, the total amenity value derived from model (3) applied to the Candlewood Lake sample is $11.7 million: this is 15% greater than the value derived from model (2) . Note that $11.7 million is 25% less than the value generated by model (1). In model (2), view services generate $2 million in value, annually. Yet, when the waterfront variable is added in model (3), view services are effectively $0 while adjacency services are worth $2.5 million, annually. Recall from table 4 that adding the waterfront control to model (2) dramatically reduced the effect that a view of the lake has on the price of property: the view coefficient is 75% smaller and marginally significant in model (3). Table 7 shows that the sensitivity of the water view coefficient to model specification has a tangible impact on the total amenity values generated by the two models.

The valuation results reflect the findings discussed in section 3.1. Comparing model (1) to models (2) and (3), table 4 shows that the distance coefficient declines markedly when the other amenity variables are added to the model. This suggests that the lake-distance coefficient is biased upwards as this variable captures the price effects of view and adjacency which are omitted from model (1). Supporting the notion that the distance parameter estimate is biased in model (1), is the Hausman test which provides evidence that in model (1) the lake distance variable is endogenous in both samples. This influence of the sensitivity of the distance coefficient to model specification is evident in table 7; using both samples, the access value is substantially lower in model (2) and (3), after adding view and adjacency, than in model (1). A similar issue Similarly, when comparing model (2) and (3) applied to the Connecticut sample, table 4 shows that in model (2) the view coefficient declines in magnitude and significance when adjacency is added to the model. It is likely that the view coefficient is biased upward because the view variable captures the price effect of omitted adjacency services. As above, the Hausman test provides further evidence that the view covariate is endogenous (in the Candlewood Lake sample). The effect of the lake coefficient on the total amenity value is evident in table 7. First, the total values for Lake Monroe produced using models (2) and (3) are the same, supporting the Hausman test finding of no evidence of endogeneity when omitting adjacency from model

(2). However, comparing the service values for Candlewood Lake generated by models (2) and (3) reveals evidence of this bias. That is, the value of view services is roughly \$2 million in model (2) yet it is only \$0 in model (3) when the waterfront variable is added to the hedonic function. Equally problematic is the fact that the biased price for view services in model (2) underestimates the value of waterfront services estimated using model (3). In sum, the valuation experiments reveal that the biased coefficients, lake-distance in model (1) and view in model (2), have tangible implications for valuation.

4 Conclusions

This paper finds that how one measures the services provided by public water bodies in an hedonic price function has a substantial impact on the total estimated value of the amenity. Evidence uncovered in this study suggests that omitting controls for either view or adjacency services may result in misspecifation of the hedonic price function. The implicit prices derived from hedonic models without such measurements may be biased. Total amenity values estimated using such models will likely embody this bias. This analysis indicates that such specification errors may dramatically affect valuation. This is important for researchers and public agencies charged with managing water resources to bear in mind as the external benefits from existing or proposed man-made lakes and reservoirs are estimated. Therefore, while it requires considerably more effort to determine which properties are in waterfront loactions and which properties have a view, the potential misspecification of "distance-only" models likely justifies these extra research costs. Further, the findings in this analysis call into question results from "distance-only" models in the literature.

The findings in this paper suggest that not all public water bodies provide adjacency services. Recall that Candlewood Lake generates adjacency, view (which is marginally significant), and access services while Lake Monroe, the reservoir, provides only view and access services. Comparing the management strategies employed on the lakes provides a plausible explanation for these results. Lake Monroe is a flood control reservoir, and a source of irrigation and drinking water which experiences management-induced seasonal fluctuations in water levels. This likely has tangible implications for waterfront homeowners. For instance, drastic reductions in water levels may present waterfront homeowners with considerable land area normally underwater. In addition to impeding access, this area would likely possess undesirable qualities; no vegetation, sunken debris, and offensive odors. If consumers are aware of such conditions, this situation would likely affect the price of waterfront homes. In contrast, the management strategies applied to Candlewood Lake do not involve drastic fluctuations in water levels. Instead, a rather constant water level is maintained. In this more passive management setting, there is a substantial premium affixed to waterfront locations. Thus, intensive management necessitating unnatural fluctuations in water levels may impinge upon the flow of services to waterfront homeowners.

Prior studies provide quantitative evidence of this causal mechanism [Loomis, Feldman, 2003; Lansford, Jones, 1995]. These authors estimate the effect which reduced lake levels have on housing prices. In both studies, the effect is significant and negative, indicating that reductions in lake levels constitute a tangible disamenity. Further, in both articles, the observed fluctuations in lake levels are due to management strategies much like those employed on Lake Monroe. Therefore, similarities between management techniques employed at Lake Monroe and the lakes used in these prior studies points to anthropogenic fluctuations as a likely contributing factor to the absence of a waterfront premium on Lake Monroe.

Table 1: Variable Descriptions

Variable	Description
Price ($)	Sales Price ($ 2000)
Water Front	1 = Waterfront Property
Water View	1 = Waterview Property
Distance to Lake (mi)	Straight-Line Distance from Property to Nearest Lake Shore
Distance to Town (mi)	Straight-Line Distance from Property to Nearest Town Center
Lot Size (Acres)	Natural log of the acreage of the property
Living Area (ft^2)	Natural log of the square-footage of finished living area
Number of Stories	Natural log of the total number of stories in structure
Age of Structure	Natural log of the age at time of sale
New Milford Twnshp	1 = in New Milford Township
Condominium	1 = Unit is a Condominium
Mobile Home	1 = Unit is a Mobile Home
Clear Creek Township	1 = Unit is in Clear Creek Township
Polk Township	1 = Unit is in Polk Township
Unit Sold in 1999	1 = Unit was sold in 1999
Unit Sold in 2000	1 = Unit was sold in 2000
Unit Sold in 2002	1 = Unit was sold in 2002
Unit Sold in 2003	1 = Unit was sold in 2003
Elementary District 1	1 = Unit is in Elementary School District 1

Table 2: Descriptive Statistics Indiana Sample (n =329)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Price ($)	125,648	80,731.6	25,000	453,000
Water Front	0.06	0.25	0	1
Water View	0.24	0.43	0	1
Distance to Lake (mi)	0.84	1.05	0.02	5.4
Distance to Town (mi)	8.29	1.63	3.45	17
Lot Size (Acres)	1.88	5.78	0.011	73.7
Living Area (ft^2)	1,852.9	1,052.1	380	5,664
Number of Stories	1.691	0.76	1	4
Condominium	0.57	0.5	0	1
Mobile Home	0.02	0.13	0	1
Clear Creek Township	0.86	0.35	0	1
Polk Township	0.02	0.13	0	1
Unit Sold in 1999	0.50	0.5	0	1
Unit Sold in 2000	0.47	0.5	0	1
Elementary District 1	0.86	0.35	0	1

Table 3: Descriptive Statistics Connecticut Sample (n = 326)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Price ($)	294,007	225,337	35,000	1,775,000
Water Front	0.17	0.38	0	1
Water View	0.34	0.48	0	1
Distance to Lake (mi)	0.68	0.98	0	5.2
Distance to Town (mi)	3.01	1.40	0	5.9
Lot Size (Acres)	0.92	1.29	0.1	12.8
Living Area (ft^2)	1,684.9	838.9	410	5,728
Number of Stories	1.4	0.38	1	2
Age of Structure	37.33	29.88	0	201
Unit Sold in 1999	0.14	0.35	0	1
Unit Sold in 2000	0.25	0.44	0	1
Unit Sold in 2002	0.03	0.18	0	1
Unit Sold in 2003	0.02	0.13	0	1
New Milford Twnshp	0.6	0.49	0	1

Table 4: Regression Results, Dependent Variable: log(Price) *** 0.01, ** 0.05, * 0.10 level of significance

Independent Variable (t-statistics)	Lake Monroe Model (1)	Lake Monroe Model (2)	Lake Monroe Model (3)	Cand. Lake Model (1)	Cand. Lake Model (2)	Cand. Lake Model (3)
Constant	6.14** (16.61)	6.33** (17.23)	6.35** (16.76)	8.94 (0.35)	15.0 (0.70)	20.4* (1.81)
Water Front			−0.03 (−0.30)			0.61*** (9.30)
Water View		0.34** (5.87)	0.34** (5.21)		0.30*** (5.35)	0.07 (1.29)
Distance to Water: $m_\lambda(D_1)$	-0.26** (-3.73)	−0.15* (−2.20)	−0.15* (−2.22)	-0.18*** (-4.37)	-0.08** (-2.35)	-0.10*** (-3.05)
Distance to Water: $m_\lambda(D_2)$	-0.02 (-0.25)	−0.07 (−0.86)	−0.07 (−0.86)	0.19** (2.51)	0.11 (1.53)	0.10 (1.44)
Distance to Water: $m_\lambda(D_3)$	0.09 (0.83)	0.17 (1.54)	0.17 (1.52)	-0.07 (-0.62)	-0.04 (-0.33)	-0.01 (-0.11)
Living Area (ft^2)	0.81** (22.38)	0.77** (20.59)	0.77** (20.21)	0.95*** (16.19)	0.85*** (14.19)	0.68*** (11.27)
Lot Size	0.06** (3.65)	0.06** (3.68)	0.06** (3.67)	-0.01 (-0.19)	0.04* (1.39)	0.10*** (3.51)
Age at Sale				-0.02 (-1.05)	-0.03 (-1.39)	-0.05*** (-2.91)
Mobile Home	−0.41* (−2.12)	−0.44* (−2.38)	−0.45* (−2.38)			
School District	−0.21 (−1.54)	−0.34** (−3.32)	−0.34** (−3.22)			
1999	0.35* (4.01)	0.30** (3.49)	0.30** (3.48)	0.41 (0.88)	0.54 (1.30)	0.74** (2.45)
2000	0.14 (1.40)	0.12 (1.26)	0.12 (1.28)	0.77 (0.79)	1.10 (1.26)	1.59*** (2.93)
2002				-0.88 (-0.66)	-1.37 (-1.17)	-1.99*** (-2.70)
2003				-1.40 (-0.81)	-2.02 (-1.30)	-2.88*** (-2.85)
Prime Rate				-2.81 (-0.82)	-3.98 (-1.31)	-5.55*** (-2.95)
GDP				0.24 (0.11)	-0.11 (-0.06)	-0.22 (-0.22)
Clear Creek Township	0.16 (1.23)	0.31** (3.25)	0.30** (3.15)			
Polk Township	−0.63** (−2.80)	−0.40 (−1.84)	−0.41** (−1.88)			
New Milford Town				-0.22*** (-4.41)	-0.14*** (-2.74)	-0.11*** (-2.97)
Distance to Town	−0.14* (−2.39)	−0.10 (−1.86)	−0.11 (−1.89)	0.08 (1.59)	0.06 (1.05)	0.08* (1.80)
$(\text{Distance to Town})^2$	0.01** (3.09)	0.01 (1.96)	0.01 (1.96)	-0.01 (-1.54)	-0.01 (-0.99)	-0.01* (-1.74)
R^2	0.66	0.69	0.69	0.63	0.67	0.76
F-Stat.	66.5	75.2	72.1	30.2	32.1	69.0

Table 5: The Hausman Test. * 0.01, ** 0.05 level of significance

Sample (Model)	**t-stat.**	**Endogenous Var.**
Lake Monroe (1)	5.83**	Distance to Water
Candlewood Lake (1)	2.89**	Distance to Water
Lake Monroe (2)	0.30	Water View
Candlewood Lake (2)	9.56**	Water View

Table 6: Valuation Methodology and Results Using Model (3)

Service (Sample)	# Houses	Price ($1000)	Service Premium	Annual Service Value ($1000/year)
(Lake Monroe)				
Access	1710	130	0.19	$(1,710) \times (0.19) \times (130,000 \times (r+t)) = 2000$
View	120	180	0.29	$(118) \times (0.29) \times (180,000 \times (r+t)) = 300$
Adjacency	60	180	0.00	$(60) \times (0.00) \times (180,000 \times (r+t)) = 0$
				Total Amenity Value = 2300
(Candlewood Lake)				
Access	2960	300	0.13	$(1,115) \times (0.13) \times (300,000 \times (r+t)) = 9200$
View	230	411	0.07	$(230) \times (0.07) \times (411,000 \times (r+t)) = 500$
Adjacency	120	562	0.46	$(116) \times (0.46) \times (562,000 \times (r+t)) = 2500$
				Total Amenity Value = 12200

Table 7: Total Lake Valuation ($1,000/year)

Sample (Model)	**Access**	**View**	**Adjacency**	**Total**
Lake Monroe (1)	4600	*	*	4600
Lake Monroe (2)	2500	300	*	2800
Lake Monroe (3)	2500	300	0	2800
Candlewood Lake (1)	15,600	*	*	15,600
Candlewood Lake (2)	7800	2000	*	9800
Candlewood Lake (3)	9200	500	2500	12,200

A.1 Hausman Test Procedures

Let X denote the sub-set of exogenous variables included in model (1) and D denote lake-distance (the variable suspected of correlation with the error term) in the following regression model:

$$log(P) = \beta\mathbf{X} + \theta D + \varepsilon \quad (14)$$

Then D is projected onto the full set of exogenous variables (denoted (Z)) which includes the water front and water view variables.

$$D = \delta\mathbf{Z} + \upsilon \quad (15)$$

The next step is to capture the residuals $(\hat{\upsilon})$ from this projection and then estimate the following model:

$$log(P) = \beta\mathbf{X} + \theta D + \rho\hat{\upsilon} + r \quad (16)$$

The t-test statistic for ρ using ordinary least squares is an appropriate test of

$$H_0 : \rho = 0 \quad (17)$$

$$H_0 : \rho \neq 0 \quad (18)$$

If we can reject the null hypothesis (H_0) then D shows evidence of correlation with the structural error term.

A.2 Valuation Methods

A square schematic diagram of each lake is used to determine the land area in each of the distance categories, the water front and the water view zones. Brown and Pollakowski [1977] used a similar approach. The schematics are squares with areas equal to the square mileage of each lake. Water front zones are within 500 feet of the lake shore. Water view zones are within 1000 feet. The distance categories are concentric bands 1 mile wide around each square schematic. It is likely that using square schematics to represent the lakes imparts some degree of error in the land area calculations. Since each sample area spans multiple census tracts, the average housing density (units per square mile) across tracts within each sample is used.

References

[1] Brookshire, D.S., M.A. Thayer, W.D. Schulze, R.C. d'Arge. March 1982. "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches." American Economic Review. 72: 165-177.

[2] Brown, G.M. Jr., H.O. Pollakowski. August 1977. "Economic Value of Shoreline" Review of Economics and Statistics. 59: 272-278

[3] Cropper, M.L., L.B. Deck, K.E. McConnell, 1988. "On the Choice of Functional Form for Hedonic Price Functions." The Review of Economics and Statistics, 70(4): 668-675.

[4] Earnhart, D., 2001. "Combining Revealed and Stated Preference Methods to Value Environmental Amenities at Residential Locations." Land Economics. 77(1): 12-29.

[5] Freeman, A.M. "The Measurement of Environmental and Resource Values." 2nd Ed. Resources for the Future. Washington D.C. c2003.

[6] Hardle, W. "Applied Non-Parametric Regression", Econometric Society Monograph, No. 19, Cambridge University Press, Cambridge, U.K.,1990

[7] Heien, D.M. 1968. "A Note on Log-Linear Regression." Journal of the American Statistical Association. 63(323): 1034-1038.

[8] Hausman, J.A. 1978. "Specification Tests in Econometrics." Econometrica, 46(6):1251-1271.

[9] Lansford, N.H., L.L. Jones. 1995. "Recreational and Aesthetic Value of Water Using Hedonic Price Analysis." Journal of Agricultural and Resource Economics. 20(2): 341-355.

[10] Loomis, J., M. Feldman. 2003. "Estimating the Benefits of Maintaining Adequate Lake Levels to Homeowners Using the Hedonic Property Method." Water Resour. Res., 39(9).

[11] Mahan, B.L., S. Polasky, R.M. Adams. 2000. "Valuing Urban Wetlands: A Property Price Approach." Land Economics. 76(1): 100-113.

[12] Palmquist, R.B., "Valuing Localized Externalities." Journal of Urban Economics. 31(1):59-68.

[13] Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," Journal of Political Economy. 82: 34-55.

[14] Stone, M. 1974. "Cross-Validatory Choice and Assessment of Statistical Predictions." Journal of the Royal Statistical Society. Series B (methodological). 36 (2): 111-147

[15] U.S. Census Bureau. Block Maps. http://ftp2.census.gov/plmap/pl_blk/st18_Indiana/c18105_Monroe/PB18105_000.pdf.

[16] U.S. Census Bureau. Block Maps. http://ftp2.census.gov/plmap/pl_blk/st09_Connecticut/c09001_Fairfield/PB09001_000.pdf.

[17] U.S. Census Bureau. Block Maps. http://ftp2.census.gov/plmap/pl_blk/st09_Connecticut/c09005_Litchfield/PB09005_000.pdf.

[18] U.S. Census Bureau. http://www.census.gov/census2000/states/ct.html.

[19] White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity" Econometrica, 48 (4): 817-838.