2009

Hedonic Property Value Modeling Of Water Quality, Lake Proximity, And Spatial Dependence In Central Florida

Patrick Walsh
University of Central Florida

HEDONIC PROPERTY VALUE MODELING OF WATER QUALITY, LAKE PROXIMITY, AND SPATIAL DEPENDENCE IN CENTRAL FLORIDA

by:

PATRICK WALSH
B.A. Salisbury University, 2004
B.A. Salisbury University, 2004
M.S. University of Central Florida, 2007

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Economics
in the College of Business
at the University of Central Florida
Orlando, Florida

Summer Term
2009

Major Professor: J. Walter Milon

ABSTRACT

Hedonic property value analysis is one of the leading methods of environmental valuation. This non-market technique uses variation in home sales to infer the values of amenities or disamenities. While there have been numerous studies about air quality and hazardous waste, the number of papers focusing on water quality is much smaller. Consequently, there are still many unanswered questions about the proper handling of water quality through hedonic methods. Furthermore, estimates from hedonic property price analyses are rarely used in government cost benefit analyses. This dissertation investigates several important hedonic issues in a large analysis of water quality in central Florida.

The first chapter of this paper explores the extent of water quality benefits. Almost all past studies have focused exclusively on waterfront homes. The present paper includes non-waterfront homes and investigates three hypotheses about the marginal impact of water quality. The first hypothesis is that non-waterfront homes are positively affected by water quality, but by a smaller amount than waterfront homes. The second hypothesis is about the effect of lake distance on the relationship between water quality and property prices: this relationship should be negative. The third hypothesis states that properties near larger lakes have a higher implicit price for water quality than homes around smaller lakes, all else constant. These three hypotheses are investigated in each chapter of the dissertation, and provide a unifying theme to the paper. Results from Chapter 1 support all three hypotheses. Most importantly, the empirical estimates indicate that water quality benefits extend beyond the waterfront in a declining gradient. Excluding non-lakefront homes from the analysis can therefore substantially underestimate the total benefits of a water quality improvement. Estimates of the total property price benefits from

a one foot increase in water quality were found to double with the addition of non-waterfront homes.

The second chapter examines the sensitivity of results to several spatial specifications. Spatial issues can be a problem in analyses of real estate data because of spatially correlated variables, unobservable neighborhood codes and covenants, identical or similar builders, and property appraisal valuation techniques. The focus of the chapter is on the spatial weights matrix (SWM). Six different SWM's are constructed, which are based on popular specifications encountered in the current spatial hedonic literature. An out-of-sample forecasting exercise is used to compare multiple spatial specifications. Results indicate that certain spatial models may be sensitive to the specification of the weights matrix. Furthermore, many popular models currently used in the literature could be improved by allowing more non-zero elements in the SWM.

The third chapter investigates the definition of "water quality" and uses several additional quality indicators. Choosing the proper pollution indicator is an issue that has plagued many areas of the valuation literature. While clarity indicators have become popular in hedonic property price analysis, they are not used for the purposes of regulation by many state environmental departments. This chapter uses several indicators that are used by the state of Florida to classify lakes and implement policy. Implicit prices are computed for all of the indicators and issues of benefit extent and total benefits are explored. Instead of finding an optimal indicator for all situations, results indicate that the use of at least two types of indicators may capture a larger range of the true total benefits.

The final chapter uses a repeat sales model to address potential problems with omitted variable bias. Due to the size of the data set in this paper, there are a substantial number of

homes that have sold more than once. The repeat sales model analyzes differences in property sales prices for the same home over time. The three hypotheses of the first chapter are explored in this alternative model. The implicit price obtained from the repeat sales model is much larger than the regular hedonic model. However, there are some concerns with the smaller population of repeat sales.

I would like to dedicate this dissertation to my parents and brothers for their constant support.

ACKNOWLEDGMENTS

Special thanks go to J. Walter Milon, David Scrogin, Shelby Gerking, Martin Wanielista, and Mark Schneider for assistance with this paper. Without valuable support from these people, this paper would not exist. Dr. Milon provided continuous guidance, expert advice, and constant help with data. Without constant supervision and professional guidance from Dr. Scrogin, my Stata programming abilities would be severely limited. Dr. Gerking provided a firm foundation of econometric advice and was always available to steer me on the right path. I would also like to thank others from the Department of Economics who helped me along the way, including Michael Caputo, Ujjayant Chakravorty, Mireya Cortes, Mark Dickie, Glenn Harrison, Ossama Mikhail, Judy Minei, John Paul, Fran Percival, and Elisabet Rutstrom. Also, Chen Ling, Randy Moffett, Dmitriy Popov, David Rivenbark, Shabori Sen, and the other graduate students in the Department of Economics were always a source of inspiration. Finally, thanks to Nicole for putting up with the stress, late nights, and other externalities of a doctoral program.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Hedonic property price analysis has become a preferred method of environmental valuation. Instead of relying on subjective surveys and techniques, as in the stated preference category of environmental valuation, hedonic property price analysis relies on actual market behavior. It is a revealed preference method with a rich empirical background, and is used in a variety of areas, ranging from the construction of the consumer price index to the prediction of mortgage defaults.

In the field of environmental policy, hedonic property price analysis has had varying impacts. While it is not frequently used as a primary estimate of benefits in a cost-benefit analysis, Smith and Palmquist (2002) state that it is frequently used as a double check, or "second opinion" on primary estimates. Also, they point out that hedonic analysis has seen significant success in litigation, where demonstrated real effects on property price as a result of pollution or other environmental damages can be particularly convincing to judges and juries. The most common applications of hedonic analysis in the environmental literature are air quality, hazardous waste, and water quality. While air quality and hazardous waste have received significant attention recently, water quality has lagged behind.

The last two decades of hedonic literature have seen a gradual increase in the use of spatial econometrics. Hedonic property price analysis uses real estate data, which are spatially distributed across a landscape. This spatial configuration of real estate data makes it especially prone to spatial autocorrelation. While it is typically easy to quantify a property's structural characteristics, locational attributes are much harder to observe. Neighborhood characteristics and location, which can heavily influence a home's price, are typically included in a hedonic analysis only through proxy variables. These proxy variables—such as census tracts—are almost always estimated with error (Dubin 1998), and rarely capture true neighborhood boundaries or characteristics. Furthermore, similar building techniques, neighborhood attributes, and other omitted variables are not accounted for in a traditional ordinary least squares regression (OLS). These issues result in spatial dependence in the data, which cause OLS estimates to be biased.

A variety of methods have been developed to deal with spatial dependence. Spatial econometrics has emerged as a wide field, with its roots in geography. Applications of spatial econometrics are appearing in a wide variety of areas, from transportation economics to the political economy of voting districts. Spatial techniques have appeared in limited quantities in hedonic studies from all three of the main applications, with the most papers dedicated to analyzing air quality. However, much of the literature has been slow to adopt these techniques, as there are no concrete theoretical recommendations for the "correct" spatial functional forms and approaches to utilize in a hedonic analysis. The approach to spatial modeling can appear rather arbitrary at times, though data size and the particular context can provide a rough guide to the proper approach.

To cope with the theoretical shortcomings of spatial econometrics, this paper will compare a variety of the main techniques in order to find one, or several, preferred spatial functional forms. The dissertation enhances its use of spatial econometrics with tools and methods from the field of property appraisal. Since hedonic modeling and property appraisal can have similar goals, and because hedonic regression is sometimes employed in property appraisal, these two fields yield complementary procedures. The methods use here should simultaneously allow a thorough hedonic analysis of lake water quality and a critical exploration of spatial techniques.

The dissertation is composed of four chapters. The first chapter is a spatial hedonic analysis of water quality. The focus of this chapter is the spatial extent of benefits from an improvement in water quality. The prevailing wisdom in the literature is that these benefits are restricted to the population of waterfront homeowners. Chapter 1 includes non-waterfront homes in the analysis and tests three hypotheses about the marginal effect of water quality on property prices. The second chapter explores some of the spatial econometric issues encountered in the first chapter. A central component of a spatial analysis is the construction of the spatial weights matrix, yet there is no theoretical guidance for the proper formulation of this matrix. This chapter uses six weights matrices and two spatial models to determine the spatial sensitivity of the results from Chapter 1.

The third chapter explores the definition of water quality. Clarity indicators are typically employed in hedonic property price analyses. On the other hand, Florida regulators use “water chemistry” indicators to manage lakes. The clarity indicators may be measuring more aesthetic benefits than the water chemistry indicators. Chapter 3 explores several additional water quality indicators and their impact on the hedonic property model results.

The final chapter uses an alternative property price model to investigate the impact of water quality changes on property prices. A hybrid repeat sales/hedonic model has been shown to control for omitted variable bias. There are a large number of repeat sales in the data for this dissertation. The hybrid model is used to explore the three hypotheses from the first chapter about the marginal effect of water quality.

CHAPTER 1: HEDONIC PROPERTY VALUE ANALYSIS OF WATER QUALITY

1.1. Introduction

The Clean Water Act (CWA) is the primary federal legislation regulating US surface water. Originally passed in 1972 and amended in 1977, the CWA has directed billions of dollars into water programs since its inception. To assess the efficiency of expenditures, federal regulations require the Environmental Protection Agency to measure the benefits and costs of CWA programs.[1] This task is complicated because many of the benefits of improved water quality are non-market in nature, so must be estimated. Hedonic analysis, a revealed preference valuation method, has significant potential in this area but remains under-utilized in benefit cost analyses of water quality (Palmquist and Smith 2001). The EPA does not currently use hedonic estimates to calculate the benefits of water policy; the agency favors recreation demand and stated preference (such as contingent valuation) estimates of benefits (Morgan and Owens 2001; Iovanna and Griffiths 2006). In contrast, hedonic methods are regularly used by federal agencies to value hazardous waste impacts (Jackson 2001; Case et al 2006), air quality benefits (Boyle and Kiel 2001), and mortality and morbidity risk (Viscusi and Aldy 2003).

Several issues and assumptions need to be addressed before hedonic estimates can be broadly applied in cost benefit analyses of water quality. A prevalent concern in the environmental valuation literature is the extent of benefits or damages (Smith 1993). For example, travel cost studies critically examine the spatial boundaries of recreation benefits (Smith and Kopp 1980), and the full population affected by an environmental change is a current issue in contingent valuation surveys (Vajjhala, John, and Evans 2008). The extent of benefits

[1] Executive Orders 12291 and 12866 require that all new US regulations above a certain monetary threshold be subjected to a benefit-cost analysis.

has not yet been analyzed in hedonic studies of water quality, where the conventional wisdom is that these benefits only extend a short distance from a water body (Palmquist and Smith 2001). This convention is illustrated by the focus of past studies on waterfront homes; all past hedonic studies of water quality (e.g. Michael et al 1996; Leggett and Bockstael 2000; Poor et al 2001; Gibbs et al 2002) except one (Poor, Pessagno, and Paul 2007) restrict the sample of property sales to this boundary. This may be realistic in certain rural areas where there are not many non-lakefront homes. However, in other areas with more dense population the benefits may extend beyond waterfront homes, and this assumption may result in a considerable underestimation of total water quality benefits. The current paper investigates the diffusion of water quality benefits by including both waterfront and non-waterfront homes in a spatial hedonic analysis of water quality.

Concern about the wider spatial impact of water quality changes can be found in the hedonic literature on open space, which has consistently shown that proximity to lakes has a positive effect on property prices several hundred meters from the lake (Palmquist and Fulcher 2006). Also, hedonic property studies of environmental disamenities, such as hazardous waste, frequently find that negative effects on property prices penetrate deep into the surrounding community (Jackson 2001; Deaton and Hoehn 2004).

A complete evaluation of the interaction between location and the value of water quality to date remains largely unexplored. The current paper seeks to bridge the gap between the hedonic lake proximity and water quality literatures by including non-lakefront homes in the sample and controlling for distance and location. Three main hypotheses relating to the value of water quality are tested. First, non-lakefront homes are affected by changes in the water quality of local lakes, but by a smaller amount than lakefront homes. Second, the value of water quality depends on lake proximity, diminishing as lake distance increases. The third hypothesis asserts that the value of water quality increases in larger lakes.

These hypotheses are investigated in a database of more than 54,000 home sales and 146 lakes in Orange County, Florida with water quality data covering a nine year time period. To analyze this large sample of spatially distributed real estate data, two issues must be confronted. The first is the selection of the functional form of the analysis and the second is the treatment of spatial dependence. A further complication is that there may be interaction and tradeoffs between the two issues. These econometric concerns are tackled in a spatial hedonic analysis using a spatial lag model in conjunction with several joint tests of functional form and spatial dependence.

Results indicate that water quality benefits extend beyond the waterfront in a declining gradient, supporting all three hypotheses and yielding several policy implications. First, excluding the non-lakefront home benefits can substantially underestimate the total benefits from a water quality improvement. Second, the total amount of home price appreciation and increase in property tax revenues due to improvements in water quality is larger than previously thought. Third, more efficient funding mechanisms to promote water quality improvement can be identified using these estimates.

1.2. Data and Setting

The setting of the study is Orange County, Florida, which has an abundance of lakes and provides an ideal location for a hedonic water quality analysis. The property information was obtained from the Orange County Property Appraiser (OCPA).[2] These data contain sales prices

[2] Residential sales that were not qualified (according to OCPA criteria) or did not meet a minimum or maximum threshold value for the deflated sales price were eliminated from the database. These thresholds were instituted to

and other property characteristics such as the number of bedrooms, the number of bathrooms, parcel size, home age, etc. Single family property sales within 1,000 meters of a lake in Orange County, Florida during the period 1996-2004 are used for estimation. The 1,000 meter boundary was implemented to exclude properties that are located far from lakes. Several papers in the open space literature find that home prices are unaffected by water bodies beyond this distance (Lansford and Jones 1995).

The OCPA was also the source of several layers of GIS maps, which were used to define each property's spatial location. Using the property's latitude and longitude coordinates, ARC GIS was employed to calculate the spatial variables, such as the distance to the nearest lake and the central business district (CBD, downtown Orlando). Distances were calculated from the centroid of the parcel, correcting for potential problems with lakeshore definitions on particular lakes.

Additional data were integrated by geo-referencing the parcel location with the 2000 US Census. These data contain demographic and economic information on income, racial composition, and the senior population at the census tract level. Summary statistics for the dependent and independent variables appear in Table 2, separated by lakefront status. As expected, clear differences can be seen between these two groups. Waterfront homes are more expensive, larger, and newer, on average. The mean price for a waterfront home is more than double the mean price of a non-waterfront property. Nonetheless, the average distance from downtown Orlando is very similar between the two groups, as are the average latitude and longitude coordinates. Also, looking at the distribution of sales over time, the two groups sold at comparable rates over the time period.

---

eliminate relative bequests, gifts, and other sub-market property transactions and outliers, a common practice in past literature (Krysel, et al., 2003)

Obtaining reliable water quality data for the complete set of lakes over a nine year period involved four different sources and an extensive search. The Orange County Department of Environmental Protection provided the largest portion of the lake quality data. For the named lakes not appearing within these data, three individual municipalities provided records of lake monitoring activities, which were then recorded and geo-referenced into the database.

A wide variety of water quality indicators have been used in past analyses. In this chapter, a clarity indicator is employed, Secchi Disk Measurement (SDM).[3] Several studies have recommended this indicator because it is easily seen by consumers and can be a physical manifestation of eutrophication (Michael, Boyle, and Bouchard 1996; Michael, Boyle, and Bouchard 2000; Poor et al. 2001; Gibbs, Halstead, and Boyle 2002). Eutrophication results in excessive algal growth and significantly decreases the appearance and recreational benefits of the lake, and should affect consumers' purchasing decisions (Michael, Boyle, and Bouchard 1996). This use of additional indicators of water quality is pursued further in Chapter 3.

The data for this analysis contain 146 lakes, with wide variation in both lake size and water quality. Past studies have debated the use of maximum, minimum, or mean values for each individual lake, as in Boyle and Taylor (2001). The present study uses the annual mean value of SDM for each lake, as it is a more conservative estimate than the maximum or minimum value (Krysel et al. 2003).

There are no "point sources" of pollution in this area, such as factories or industrial plants, so eutrophication is mostly the result of non-point sources such as residential or commercial development and sewer and stormwater runoff within each lake's watershed (FDEP 2006). Table 1 summarizes SDM over the sample period. Note that the number of lakes in each

---

[3] Obtaining an SDM reading involves lowering a black and white checkered disk into the water and recording the depth at which it disappears from sight.

year can vary as there may not be a home sale near each lake in each year.[4] Out of the full set of lakes, 100 appear in every year.[5] SDM levels fluctuate from a minimum of 0.66 feet in 2002 to a maximum of 18.05 feet in 1997, exhibiting substantial variation in water quality over lakes. Each home sale is matched to the average annual SDM level in the nearest lake.

1.3. Literature Review

The valuation of environmental amenities cannot be done through typical market valuation. There is no direct market where one can purchase water quality, so there are no price signals available. To overcome this problem, economists have developed several non-market valuation techniques, which include hedonic property price analysis. The most prevalent types of non-market valuation studies are stated preference analyses, where values are directly elicited from subjects, typically using surveys or experiments. Stated preference methods include contingent valuation, which has been used extensively in environmental valuation. The alternative form of valuation available is revealed preference, which extracts values from consumer purchasing behavior. Hedonic analysis is a form of revealed preference valuation.

The term "hedonic price" was first introduced by A.T. Court in 1939, in a paper about the price of automobiles. The practice of conducting a hedonic analysis has evolved substantially since the earlier studies. Technological advances in computing power in the latter half of the 20th century saw a large increase in the number of hedonic analyses. The principal idea of hedonic price analysis is to use variation in the price of heterogeneous goods resulting from variation in the characteristics of those goods to estimate marginal values for the individual characteristics.

[4] None of the lakes in these data appear on the 303(d) list of impaired waters for Florida, so fish kills and other extreme conditions were not observed in our sample. Also, fecal coliform is not a problem in any of the lakes.
[5] While this core set of lakes could have been used, it was decided that the additional variation in SDM resulting from the additional lakes should be included. Also, the additional observations did not change the qualitative results.

A simple example can show the main ideas of hedonic property analysis. Suppose two homes are for sale and are identical except for the size of the parcels: one home is on a larger plot of land than the other. The difference in the prices of the two homes should be due to the difference in land sizes, and should reflect the marginal value of land. If the discrepancy between the two home prices is less than the marginal value of land, bidding by homes buyers will cause the price to rise (Haab and McConnell 2002). This basic idea can be extended to additional characteristics of non-identical homes with a large enough sample and several assumptions: perfect competition in the market, full available information, full mobility of consumers, and the maximization of well-behaved preferences (Chay and Greenstone 2005).

1.3.1. Theoretical Background

The earliest application of the hedonic model to property prices was in the 1960s, and focused on the effect of air pollution on home prices (Haab and McConnell 2002). The hedonic model was theoretically established in Rosen (1974), who formulated an equilibrium model of buyers and sellers based in utility maximization.[6] Rosen described a situation involving a differentiated commodity that could be described by a "basket" of characteristics, $\boldsymbol{z} = (z_1,\ldots, z_n)$. Homes are generally thought to fit this description well, as they are composed of a wide variety of attributes that are pursued by consumers and supplied by sellers and producers. Consumers shop around in a market composed of heterogeneous homes, which are described and advertised by the characteristic vector $\boldsymbol{z}$.

There are several main assumptions that Rosen institutes. First, both buyers and sellers are assumed to have full information about the prices of other homes and about the characteristics of homes. Additionally, it is assumed that the market has a large enough amount

[6] This model is summarized in Haab and McConnell (2002), from which the current paper draws.

of homes for sale at any given time so that the buyer faces a continuously varying set of attributes. The size of the market is large enough that it can reach equilibrium, which may be an unrealistic assumption in certain areas, but in a metropolitan area such as Orlando the market should be sufficiently large. It is generally assumed that each buyer is only purchasing one home at a time. Finally, it is assumed that the size of the market results in competitive behavior from both buyers and sellers, and that action from individuals on either side of the market does not influence the market price.

The model describes the interaction of utility maximizing consumers and profit maximizing producers. Consumers are expected to choose a particular set of characteristics of the good so that the marginal price of each consumer is set equal to the marginal rate of substitution of a compound good (which represents other consumption). In the context of home sales, we assume households have a utility function that represents their preferences, in the form $u(\boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\beta})$. The vector of physical, spatial, and environmental housing attributes is represented by $\boldsymbol{z}$, $\boldsymbol{x}$ is a composite bundle representing other commodities, and $\boldsymbol{\beta}$ is a vector of parameters of the household preference function. In most applied settings, only a portion of the total characteristics of a home that influence price and utility will be observable. The household's budget constraint is given by $y = h(\boldsymbol{z}) + \boldsymbol{x}$, where the price of $\boldsymbol{x}$ is normalized to one, it is assumed that the household is only purchasing one home, and y represents household income. The function $h(\boldsymbol{z})$ represents the hedonic price function. In equilibrium, the household maximizes their utility subject to the budget constraint. The slope of the hedonic function with respect to a particular characteristic represents the individual's marginal willingness to pay for that characteristic.

To determine the first order conditions of this maximization procedure, first substitute for the composite bundle within the utility function as $\boldsymbol{x} = y - h(\boldsymbol{z})$ to obtain:

$$U_z(y-h(z),z,\beta)-U_x(y-h(z),z,\beta)h_z(z)=0 \quad (1.1)$$

This first order condition can be rearranged to demonstrate an important characteristic of the slope of the hedonic function:

$$h_z(z)=\frac{U_z(y-h(z),z,\beta)}{U_x(y-h(z),z,\beta)} \quad (1.2)$$

Where it is clear that the slope of the hedonic function with respect to a particular attribute in ***z*** is equal to the marginal rate of substitution between that attribute and the composite bundle, all else constant (Parmeter and Pope 2009).

A main part of Rosen's theoretical model involves the interaction of consumer bid functions and producer offer functions. The consumer's (buyer's) bid for a property can be represented by the function $\theta(\boldsymbol{z};u,y)$, which holds constant utility and income. This function describes the maximum bid a consumer will place for a given vector of home attributes at given constant levels of utility and income. Using this function, it is possible to obtain a family of indifference curves illustrating tradeoffs between the composite bundle and home attributes, or characteristics, within ***z***. The bid function can be used to obtain a new form of the first order condition (1.1) by substituting it for $h(\boldsymbol{z})$.

$$U_z(y-\theta(z;u,y),z,\beta)-U_x(y-\theta(z;u,y),z,\beta)\theta_z(z;u,y)=0 \quad (1.3)$$

With the second order condition appearing as:

$$\theta_{zz}=(U_x^2U_{zz}-2U_xU_zU_{xz}+U_z^2U_{xx})/U_x^3<0 \quad (1.4)$$

The equivalence of equation (1.3) to equation (1.1) illustrates an important characteristic of Rosen's consumer model. The individual consumer bid functions are tangent to the market hedonic price function in equilibrium. The bid functions represent the maximum bid the individual would be willing to pay at a particular level of attributes, ***z***. The hedonic function represents the minimum market price that they would have to pay for a home with those characteristics. Therefore the hedonic price function forms an upper envelope of the consumer bid functions. Figure 1 contains Rosen's example of this part of the consumer equilibrium, where he uses p(**z**) as notation for the hedonic price function h(***z***). This figure shows an example of two consumers and their corresponding tangencies, where consumer 2 has a bundle with more z_1 than consumer 1.

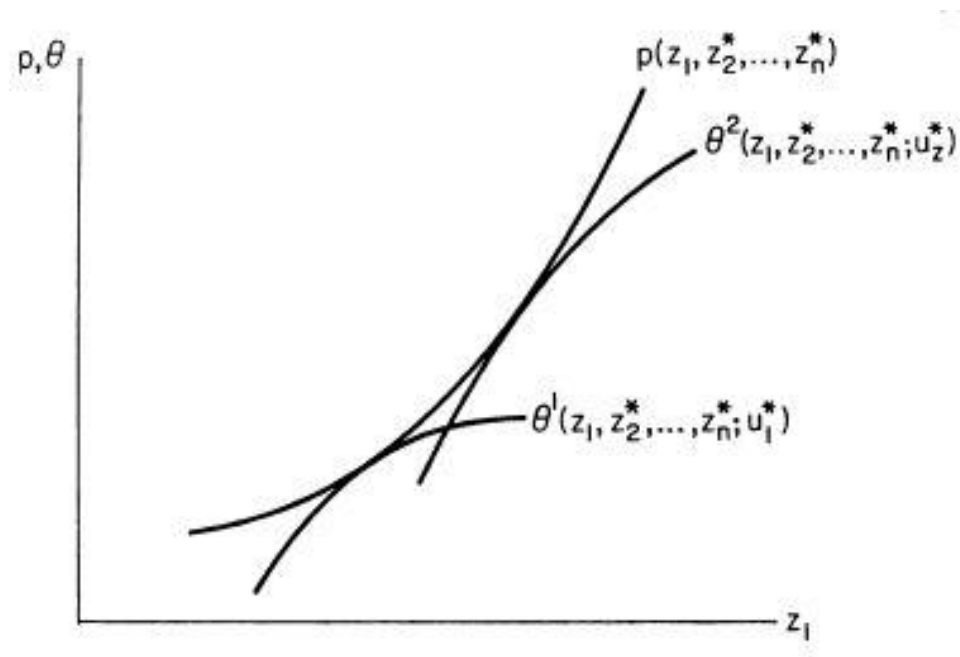

Figure 1: Consumer Bid Functions

Producers are profit maximizers, and sell various bundles of goods with characteristics $(z_1, \ldots, z_n)$. Each producer sells $M(\boldsymbol{z})$ units offering attribute bundle $\boldsymbol{z}$. Each producer faces costs determined by the industry, represented by $C(M, \boldsymbol{z}; \boldsymbol{\gamma})$, where $\boldsymbol{\gamma}$ represent individual cost and production differences between the various producers. This cost function is assumed to have traditional assumptions with respect to its curvature properties: C_m and C_x are greater than zero and C is convex.

The profit of each firm is determined by

$$\pi = Mh(z) - C(M, z) \tag{1.5}$$

which yields first order conditions (1.6) and (1.7).

$$h_i(z) = C_{z_i}(M, z) / M, i = 1, ..., n \tag{1.6}$$

$$h(z) = C_M(M, z) \tag{1.7}$$

where equation (1.6) illustrates that at the optimum, the slope of the hedonic price function is equal to the marginal costs of production; marginal revenue is set equal to marginal costs. From equation (1.7), the hedonic price function (equal to marginal revenue for one unit) is set equal to the marginal cost of selling one additional unit.

While consumers had bid functions, the symmetrical function for producers is the offer function. This function, $\phi(\boldsymbol{z}; \pi, \boldsymbol{\gamma})$ defines the unit prices that the firm is willing to accept for specific attribute bundles and a constant level of profit. From this we obtain a family of production indifference curves, or iso-profit functions. Similar to the approach in the consumer side, we substitute the seller's offer price function for the producers offer function and then obtain first order conditions.

$$\phi_{z_i} = C_{z_i} / M > 0 \tag{1.8}$$

$$\phi_{\pi} = 1 / M > 0 \tag{1.9}$$

From these new conditions it can be seen that when profit is held constant, the marginal offer price will be equal to the marginal cost of production. Also, when attribute levels are held constant, the marginal offer price is equal to the constant $1/M$.

To interpret the analysis of producers, we can proceed in a similar manner to the consumer approach. By definition, the offer price is the amount that the particular producer will accept for a particular bundle of attributes $\boldsymbol{z}$ and constant level of profit π. On the other hand, $h(\boldsymbol{z})$ is the maximum price that they can get for a model with the particular attribute bundle in the market. At the equilibrium both the producer offer functions and the market hedonic price function will satisfy essentially the same first order conditions just described. Consequently, the producer equilibrium will be characterized by a series of tangencies between individual offer functions and the market hedonic price function. Figure 2 illustrates Rosen's depiction of this situation (p. 43).

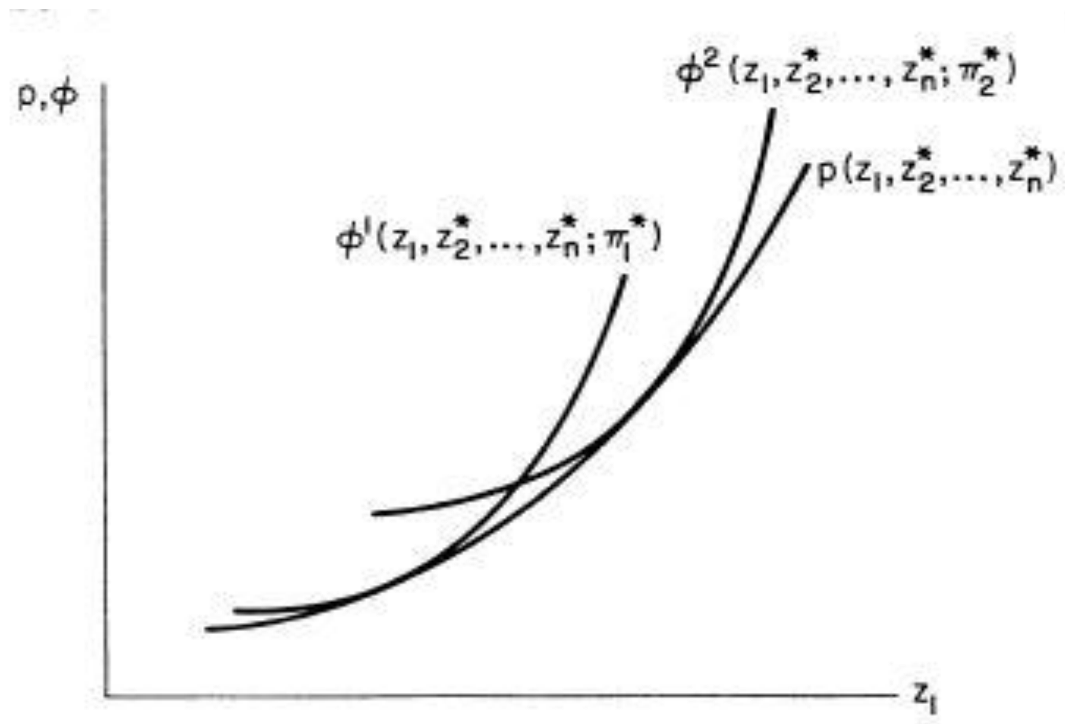

Figure 2: Producer Offer Functions

In equilibrium, the interaction of consumers and producers will result in the formation of the hedonic price function, which is an upper envelope to the consumers' offer functions and a lower envelope of the producers' bid functions. This equilibrium price function represents a locus of equilibrium tangencies between consumer bid functions and producer offer functions (Parmeter and Pope 2009). This situation is pictured in Figure 3. Since all points on this locus represent tangencies between the bid functions and offer functions, the following equation holds.

$$\frac{\partial h(\mathbf{z})}{\partial z} = \frac{\partial \theta(\mathbf{z})}{\partial z} = \frac{\partial \phi(\mathbf{z})}{\partial z} \tag{1.10}$$

Equation (1.10) indicates that the marginal price obtained from the hedonic price function is equal to a person's marginal willingness to pay for an attribute *z* (Rosen 1974; Parmeter and Pope 2009).

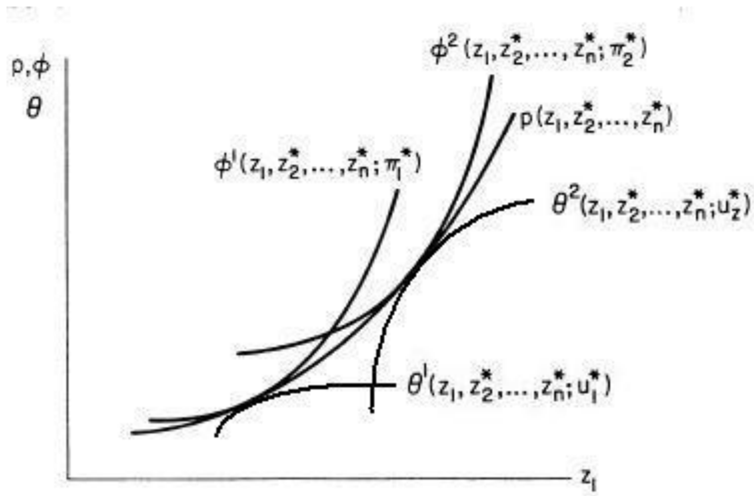

Figure 3: Hedonic Equilibrium

The estimation of the marginal values in equation (1.10) is the focus of the present paper, and is referred to a "first stage" analysis. These represent the marginal rate of substitution between the home value and the individual attributes. In Rosen's paper, he continued the theoretical derivation of this model to a "second stage," where demand functions for the attributes are estimated. Since the publication of Rosen's (1974), several studies have shown that the second stage requires strict assumptions to be identified. This is because the second stage requires variation in the implicit prices obtained from the first stage. One way to confront this problem is through the use of multiple markets. If separate hedonic price functions are estimated for each market the implicit price will change over areas and the second stage demand functions can be identified. This approach is attractive since it does not require the assumption of additive separability of property characteristics used in closed form solutions. However the definition of an individual market is never clear, which can complicate the situation (Boyle, Poor, and Taylor 1999). The second stage will not be addressed in the current paper.

1.3.2. Environmental Applications: Air and Hazardous Waste

The hedonic model has been widely used in the environmental literature because it uses actual market transactions to estimate the value of non-market externalities. The market usually analyzed is the residential property market where the good is a collection of structural, location, and environmental attributes. Since water quality in local lakes is an attribute “purchased” along with homes, variation in this attribute should be reflected by variation in home prices. If the property market is competitive and buyers and sellers are fully informed price takers, the hedonic price function can be used to extract the implicit price of improved water quality.

A variety of environmental issues have been studied using hedonic property price analysis. Compared to other areas of the hedonic literature, the number of papers analyzing water quality is somewhat small (Leggett and Bockstael 2000). The first study to analyze air quality in a hedonic framework was Ridker and Henning (1967). This early paper employed a measure of sulfation levels to analyze air quality, and find that this form of air pollution is significantly related to property prices.[7] In a study of air quality in Boston, Massachusetts, Harrison and Rubinfeld (1978) devote special attention to the process of estimating a willingness to pay function for improvements in air quality. Harrison and Rubinfeld found problems with past studies that use values obtained from marginal changes to infer benefits at other levels of pollution. They caution against this practice, as it assumes that the marginal damage function for air pollution is linear. In reality, marginal damages vary with the level of pollution. To deal with

[7] For “sulfation levels,” a variable indicating the existence of SO_2, SO_3, H_2S, and H_2SO_4 was used.

these issues, Harrison and Rubinfeld provide a four step procedure to evaluate the benefits of a pollution control strategy. The first two steps involve the calculation of each household's marginal willingness to pay for improvements in clean air. The latter two steps use Rosen's second stage procedure, where demand functions for clean air are constructed.

Graves et al. (1988) investigate several prominent issues encountered in hedonic property value studies, including variable selection, measurement error, functional form, and the distribution of the errors. The particular application is air pollution in California and its effect on property prices. Several variations of each of the four prominent issues are explored. Graves et al. find that the results are sensitive to these four issues, yielding a wide range of marginal benefit estimates. Graves et al. recommend careful robustness checks for future studies, as the effects of collinearity and measurement error may have significant consequences for benefit estimates. Smith and Huang (1996) also examine the sensitivity of hedonic property price estimates to specification issues in a meta-analysis of air pollution models between 1967 and 1988. They find that the MWTP estimated are influenced by several important factors within each city, such as the starting level of air quality and the average income of the residents. Furthermore, Smith and Huang state that there is evidence of "publication bias," as they find significantly smaller MWTP estimates in published studies.

Zabel and Kiel (2000) use a data set of properties in four cities over the period 1974-1991 to estimate demand equations for air quality. They compare the three main specifications of the hedonic equation: the linear, log-linear, and log-log models. The linear is found to be the worst, but the other two forms yield relatively similar results. Also, this study draws attention to a critical point about individuals' perceptions of pollution. Zabel and Kiel use four different indicators of air pollution, and question which one most accurately represents individuals' perceived pollution levels. This is a definite issue in the water quality literature, and is explored in Poor et al. (2001) and later addressed in Chapter 3 of this dissertation. There are also several more recent analyses of air pollution, but discussion of these papers is delayed until Chapter 2, as they have a more direct focus on spatial econometrics.

Hedonic analysis has also been used extensively to analyze the effect of proximity to environmental hazards on property prices. A review of the literature up to the year 2001 is provided in Jackson (2001). In one of the earliest papers in this category, Blomquist (1974) investigates the effect of distance to an electrical power plant on property prices. Blomquist finds that even smaller, cleaner power plants located in communities can cause significant damage to home prices at distances as far as two miles away. The paper uses census blocks as the unit of observation, with average property price serving as the dependent variable.

The incident at the nuclear power plant on Three Mile Island in March, 1972 received broad coverage in the media. Nelson (1981), however, finds that this event was not capitalized into the prices of homes located within four miles of the plant. Nelson studied home prices several months after the accident using hedonic analysis. There were no significant differences between the home prices in the area near the power plant and two control areas located much farther away.

Ketkar (1992) conducts a hedonic property price analysis of hazardous waste in New Jersey. Results indicate that the benefits of fast cleanup of sites exceed the costs. The cleanup of one site in a municipality is found to increase the median property value in that municipality by approximately $1,300 to $2,000. Ketkar also looks at implications of the Coase theorem to the cleanup of hazardous waste. He finds that it would be beneficial for property owners to contribute to local hazardous waste cleanup, above the minimum amount done by the government.

While many papers identify one particular hazardous waste variable, Deaton and Hoehn (2004) look at potential clustering of several perceived hazards in industrial zones. They show that traditional hedonic analysis of hazards encounters problems when zoning and other laws group several industrial disamenities together. Deaton and Hoehn argue that the failure of past studies to differentiate between these bundled disamenities may be responsible for some implausible results. They mitigate this dilemma by including a measure of industrial activity in their analysis of superfund sites, and find that failure to include this new variable may bias estimates of the negative effect of hazardous waste upwards.

Messer et al. (2006) look at the psychological effects of superfund sites on property values. They find that psychological stigma may significantly influence home values, and may take several years to wear off. Messer et al. say that hazardous waste cleanup should be done as quickly as possible in a way that causes the least media attention. The media attention may stigmatize the area, which can cause long-lasting damage to local homes.

Kiel and Williams (2007) examine several superfund sites across the United States, and look at their effects on property prices. They are one of the first studies to enact an extensive analysis across multiple areas. The tools of meta-analysis are used to synthesize the results from the different areas. One of the main results is that the size of the site is related to its effect, with larger sites having a larger negative impact on local home prices.

1.3.3 Environmental Applications: Water Quality

The issue of water quality has not been as heavily investigated as the above environmental issues. Due to variation in data and setting, comparisons between past hedonic water quality studies are problematic. Water quality can be measured by both subjective indicators (surveys), and objective indicators collected by sampling the waterbody or using scientific instruments to measure certain characteristics (such as temperature or pH). There is no agreement as to the “best” indicator, and choices are frequently made on availability (Poor et al 2001). Many studies rely upon water quality data collected by state agencies which are required to monitor water quality under the CWA. Yet in order to properly account for local conditions, states are given some freedom in the quality indicators they monitor (FDEP 2006). Consequently, a wide range of both objective and subjective water quality indicators have been used in past hedonic literature.

In an early hedonic paper, David (1968) used Wisconsin Department of Conservation ratings of water quality that included “poor”, “moderate” and “good.” A better rating was found to significantly increase property prices. Brashares (1985) investigated a variety of definitions of water quality and found that indicators that are directly visible to consumers, such as clarity measures, are more likely to be reflected in the property price than physical measures like dissolved oxygen or pH. Epp and Al-Ani (1979) performed a hedonic analysis on streams in Pennsylvania. They obtained data on the pH of the water, along with dissolved oxygen, biochemical oxygen demand, acid from minerals, acid from carbon dioxide, and nitrate and phosphate concentrations. Each of these variables was used as a water quality indicator in separate hedonic equations. Their conclusions indicated that the water quality of local streams affected property prices, with the strength of this relationship depending on the indicator. The pH of the water was found to have an effect on property prices near clean streams but not near polluted ones. The authors postulated that this was due to the increase in recreational activities available on clean streams. Brashares (1985) also examined the use of multiple water quality indicators in the hedonic equation. Brashares recommended the use of indicators which are visible to property owners, as these indicators are more likely to be absorbed into property prices.

In a study of St. Albans Bay in Vermont, Young (1984) analyzed the effect of decreased water quality conditions on local property prices. The decreased conditions in the Bay, which is a part of the larger Lake Champaign, were a result of a malfunctioning treatment plant. Two methods were used to estimate the impact of the decreased water quality on property prices. The first was a simple dummy variable that indicated whether a home was located on the bay or on Lake Champaign. The second was a subjective indicator, based on a scale of one to ten, solicited from local "water quality experts", realtors, and officials. Two regressions were performed, one for each water quality indicator. In the first regression, the dummy variable had a coefficient of -4,690.2, indicating that homes located on the bay had average property prices of approximately $4,700 less than those on the lake. Young attributes this difference to the decrease in water quality. In the latter regression, the subjective indicator had a coefficient of 1,417.1. Since the average water quality rating in the lake was 6.4, while the average rating in the bay was only 3.4, the difference in property values is stated to be approximately $4,200. Since the average home price in the sample was $22,400, these values represent approximately a 20% impact on home values.

To control for some of the population sorting issues encountered when using water quality in a hedonic property price analysis, Steinnes (1992) uses the value of a property's land as the dependent variable. Normally the value of the entire parcel, including the land and structure, is used as the dependent variable. The land values were individually estimated using appraisal techniques. Furthermore, land values are aggregated by the lake level and expressed in foot frontage. Steinnes found hedonic techniques can be used to estimate the effect of water quality on land values, and found that higher SDM readings were correlated with higher values.

There have been several recent hedonic property price analyses of water quality in Maine lakes. Michael, Boyle, and Bouchard (1996) is a report done for the Maine Department of Environmental Protection that finds a positive relationship between water quality and property prices. Implicit prices indicate a one meter improvement in lake water clarity results in an increase in the average property price of $11 to $200, expressed in benefit per foot frontage on the lake and varying over setting. Boyle, Poor, and Taylor (1999) investigate the demand for lake preservation in freshwater lakes. They use several real-estate markets to identify demand shifters, and find significant evidence that water quality is positively related to property prices. Unlike the present paper, Boyle et al. focus on the second stage of hedonic analysis, which focuses on estimating demand. The present stage instead concentrates on the first stage and implicit prices.

Michael, Boyle, and Bouchard (2000) explore the differences in estimated values caused by using different environmental quality indicators in a hedonic regression. They were concerned that the subjective choice of indicator by the researcher can affect the results and conclusions of the analysis. Several combinations of SDM, which include historical values and future expectations, are examined. They find that the various indicators, which represent differing points of view about water quality perception, result in significantly different estimates of value.

Boyle and Taylor (2001) analyze error in the measurement of structural and neighborhood characteristics, and its effects on a hedonic regression of water quality. Boyle and Taylor were worried that there may be significant measurement error in data from tax collectors and other municipal sources. They compare regression coefficients resulting from several sources of data, and find positive coefficients on the water quality variables from each source.

The paper supports the use of tax assessor data, and finds that there is less error than previously assumed in these data.

In the latest Maine study, Poor et al. (2001) follow the work of Brashares (1985) and Michael, Boyle, and Bouchard (2000) and compare objective and subjective indicators of water quality. Their results support the use of objective measures, such as SDM, as opposed to subjective survey measures. The debate over subjective versus objective indicators of water quality is still an active research focus. In a recent meta-analysis of the broader water quality valuation literature, Van Houtven, Powers, and Pattanayak (2007) support a movement towards the "designated use" classifications of the CWA. These categories sort waterbodies into what use society has designated for them (such as drinking water, water-based recreation, fishing, and agricultural water supply) and Van Houtven et al say provide a common framework for standardization. Nonetheless, within the more specific hedonic literature, there appears to be a general trend towards the use of SDM. This indicator is easily seen by home buyers, collected on a wide scale, and is a proxy for eutrophication.

Leggett and Bockstael (2000) use fecal coliform as the quality indicator in an analysis of waterfront homes on the Chesapeake Bay. It was expected that levels of fecal coliform bacteria would be incorporated into property prices because they were regularly announced in the media, and were the cause of frequent beach closings. In their study area sanitation plants are the source of fecal coliform. However, these plants are also associated with obnoxious noise and smells and can cause omitted variable bias if plant proximity is not incorporated in the hedonic model. This is the only hedonic analysis of water quality to control for spatial dependence.

Other papers include an applied hedonic investigation by Gibbs, Halstead, and Boyle (2002), who estimate the implicit price of water quality using a variety of lakefront homes in several markets in New Hampshire. Implicit prices are estimated for all markets, the majority of which are positive. The magnitude of the implicit price of water quality is found to be dependent on the study area. Results also indicate that policies that improve water quality can result in increased property taxes.

In a state government report in Minnesota, Krysel et al. (2003) estimated a hedonic analysis of water quality on several lake areas. Their main model was based on the previous Maine papers, to which they added slight variations to accommodate the setting. They found water quality to be a significant contributor to lakefront property prices in all lake areas studied. The potential benefit from water quality improvement programs was found to be in a range of tens of thousands to millions of dollars.

The aforementioned hedonic studies have investigated a variety of issues in several different markets; however they share a common characteristic: they all focus on waterfront properties. In rural areas of Maine (as in Michael et al (1996), (2000) and Poor et al (2001)) where there may not be many properties past the lakefront, this focus may be justified. However, in more densely populated areas it is natural to question if benefits extend past the lakefront. If non-waterfront homes do in fact have a positive implicit price of water quality, a cost benefit analysis that ignores these values may substantially understate the true benefits. Both Gibbs et al (2002) and Leggett and Bockstael (2000) acknowledge that non-lakefront homes may be affected by water quality changes, but neither study confronts the issue.

The only hedonic study to include non-lakefront homes is an analysis of “ambient” water quality by Poor, Pessagno, and Paul (2007). Homes are grouped into broad geographic areas based on 22 water quality monitoring stations; each home within the area is assigned the same level of water quality. Properties are therefore not associated with a particular waterbody, so issues of water proximity are not addressed and waterfront and non-waterfront homes are indistinguishable in their model.[8] Total suspended solids and total inorganic nitrogen are used as the water quality indicators. Results indicate that the implicit price of a water quality improvement is positive for the average home; since 98% of the properties are non-waterfront they conclude that non-waterfront homes are positively affected by water quality.

Evidence for the influence of location and the spatial extent of benefits can be found in other areas of the hedonic literature. For instance, hedonic studies of open space proximity commonly find that distance to a water body is positively valued (Irwin 2002; McConnell and Walls 2005; Anderson and West 2006; and Cho et al 2009). Brown and Polasky (1977), Lansford and Jones (1995) and Palmquist and Fulcher (2006) focus specifically on distance to water bodies and results indicate proximity to lakes is positively related to property prices, with the effect decreasing over distance. It is therefore reasonable to hypothesize that the value of lake proximity is affected by changes in water quality (Palmquist and Smith 2001).

Distance effects have also received careful treatment in hedonic studies of environmental disamenities (Kolhase 1991; Deaton and Hoehn 2004; Cameron 2006; Kiel and Williams 2007; Greenstone and Gallagher 2008). Cameron (2006) provides several models for incorporating both distance and direction in a hedonic model. A common result in hazardous waste papers is that the negative impact of the disamenity extends far into the surrounding community (Jackson 2001). Additionally, benefit extent is a common problem in travel cost papers (Smith and Kopp 1980; Smith 1993), illustrated by choices between day trips and overnight trips and travel

[8] They state that ambient water quality is appropriate because the area has a long history of fishing and water-based recreation, and is an integral part of the culture. Water quality deterioration in the area is a regular topic in local politics.

boundaries. Finally, market and benefit extent issues have recently been explored in contingent valuation surveys (Vajjhala, John, and Evans 2008).

A cursory analysis of several other fields of the hedonic literature indicates that in many cases the effect of an amenity (or disamenity) can extend sizable distances. The extent of benefits may not be of paramount importance in rural areas with low population density around lakes. However, in more urban areas lakes represent a large source of recreation and aesthetic benefits to a broad group of recipients, so it is reasonable to propose that their impact on the local community is not confined to lakefront homes. A complete hedonic analysis of lake water quality should draw from these other fields to expand the focus beyond the lakefront.

1.4. Methods

The data for this analysis contain more than 54,000 property sales near 146 lakes, over a nine year time period. These data provide ample variation in homes and water quality, yielding a unique setting to analyze the extent of benefits. The econometric model is a first-stage hedonic analysis that estimates the implicit prices of attributes and does not go into the second stage outlined by Rosen (1974). This model is motivated by three main hypotheses:

H^1: *The Edge Effect*: The prices of non-lakefront homes are positively affected by increases in water quality, though by a smaller amount than lakefront properties. This issue has not been directly dealt with in the past (Poor, Pessagno, and Paul 2007) and is the first step in extending the boundary of water quality benefits.

H^2: *The Proximity Effect*: Given that water quality benefits extend past waterfront homes, this effect declines with distance. Motivated by several other areas of the valuation literature, this hypothesis introduces distance into the relationship between property prices and water quality.

<u>H</u>[3]: *The Area Effect:* The value of water quality depends on the size of the lake. Larger lakes permit recreational opportunities such as boating, swimming, and water skiing, which are enhanced by higher quality water. Several papers have found that the implicit price of water quality is higher in larger lakes, including Boyle et al (1999), Poor et al (2001), and Gibbs et al (2002). Also, in studies of open space, Anderson and West (2006) and Cho et al (2009) find that the implicit price of proximity to lakes also increases with lake size.

All three hypotheses are used to construct the econometric model. The main model appears in equation (1), which highlights the water quality and location interactions:

$$\begin{aligned} P &= \beta_0 + \beta_1 WF + \beta_2 SDM + \beta_3 WF * SDM + \beta_4 Distance + \beta_5 Distance * SDM \\ &+ \beta_6 Area * SDM + \gamma'\mathbf{x} + \varphi'\mathbf{y} + \psi'\mathbf{l} + \delta'\mathbf{t} + \varepsilon \end{aligned} \tag{1.11}$$

The dependent variable *P* is the deflated sale price, the waterfront dummy variable is *WF*, the water clarity variable is *SDM*, *Distance* is the variable that measures the distance from each home to the nearest lake, and *Area* represents the size of the nearest lake. The bold characters represent vectors. The additional property characteristics are contained in the **x** vector, the **y** vector contains location-based variables such as latitude and longitude coordinates and distance to the central business district, lake effects appear in the **l** vector, time effects for each year appear in the **t** vector, and ε is the error term. The lake effects control for lake specific attributes, such as docks, accessibility, and boat ramps.

Equation (1.11) also contains interaction variables between SDM and three other variables: *WF*, *Distance*, and *Area*. Restrictions on these variables allow tests of the three main hypotheses. The Edge Effect is tested through the restriction β_3=0; if this cannot be rejected, the implicit price of water quality does not differ between waterfront and non-waterfront homes.[9]

[9] In all models, waterfront and non-waterfront homes are pooled. Without pooling, nearby non-lakefront homes are omitted from the neighbor set defined in the Spatial Weights Matrix (SWM) for lakefront homes. Properties in both

The Proximity Effect can be tested with the null hypothesis $\beta_5 = 0$. If the Proximity Effect exists, this coefficient is expected to be negative as the value of SDM decreases with distance. The final hypothesis, the Area Effect, can be tested through the coefficient β_6, which is expected to be positive and significant. Three main models are estimated to highlight these restrictions. The first model omits the interaction terms *SDM*Distance* and *SDM*Area*. The second model includes *SDM*Distance*, and the third model contains the full form of equation (1.11).

To econometrically estimate the model in (1.11) two important specification issues must be confronted: functional form and spatial dependence. There is no theoretical foundation for the functional form of a hedonic price model (Leggett and Bockstael 2000), so decisions about this choice are typically based on either assumptions about the data, a Box-Cox test, or the "fit" of the various models. On the other hand, spatial dependence can be a problem in analyses of real estate data because of spatially correlated omitted variables, unobservable neighborhood codes and covenants, identical or similar builders, and property appraisal valuation techniques (Anselin 1999; Kim et al 2003; Palmquist and Fulcher 2006). These spatial influences violate traditional regression assumptions about ε in equation (1.11), which will result in biased coefficients (Anselin 1999). There are currently no well defined procedures for the joint determination of the functional form and type of spatial dependence, although recent research indicates that they are directly related (McMillen 2003).

The most commonly used functional forms are the linear, semi-log and double-log models. Water clarity variables are typically used in their natural log form because at higher levels of clarity it is more difficult to perceive a marginal change (Michael et al 1996; Poor et al

---

these groups share many unobservable neighborhood characteristics, and if the lakefront homes are separated, the SWM for that analysis will be extremely sparse and the value added by the spatial regression will be severely limited. Also, past studies that split the two groups have found it difficult to separate the effect of water quality on sales prices from other benefits associated with living on the water, such as view, recreation, and open space (Haab and McConnell 2002).

2001; Boyle and Taylor 2001). When modeling water quality variables, ecologists normally use this transformation to "accommodate heterogeneity of variance" (Brown et al. 2000).

The two most common forms of spatial dependence are the spatial lag and spatial error models. The spatial lag model uses a spatially weighted average of nearby home prices as a dependent variable to incorporate spatial influences. The spatial error model assumes that there are omitted variables that are spatially related to other variables in the analysis, and uses a non-spherical error covariance to correct for this (Anselin 1988, 1999).[10] These two forms of spatial dependence have been widely applied in the literature and are fully discussed in the second chapter of this dissertation. The spatial lag model has been used in several recent hedonic studies of air quality (Kim et al 2003; Anselin and LeGallo 2006; Anselin and Lozano-Gracia 2008) and the spatial error model has been used in the hedonic water quality literature (Leggett and Bockstael 2000). Both of these models employ an *n* x *n* spatial weights matrix, *W*, which specifies the spatial configuration of the data by defining a neighbor set for each observation. In the current paper, the spatial weights matrix is based on time and inverse distance. W_{ij} is nonzero if the homes *i* and *j* are within 200 meters of each other and sold either 6 months before or 3 months after each other.[11] Several recent studies, such as Pace et al. (1998), Gelfand et al. (2004), and Case et al. (2004) have stressed the incorporation of both space and time into the spatial weights matrix.

These two econometric issues–functional form and spatial dependence–have traditionally been treated independently. For instance, Patton and McErlean (2003) test for functional form

---

[10] The spatial lag model is represented by $P = \rho WP + X\beta + \varepsilon$, where ρ is a spatial autocorrelation parameter similar to those used in time series analysis. *W* is a spatial weights matrix that specifies the spatial configuration of the data by defining a neighbor set for each observation. The spatial error model is expressed as $P = X\beta + \varepsilon$, where $\varepsilon = W\varepsilon + u$, and $u \sim N(0, \sigma^2 I)$ and the parameter λ is a coefficient on the spatially correlated errors. A non-zero value of λ indicates the influence of spatial dependence in the errors which can cause inefficient but unbiased regression estimates (LeSage 1999).

[11] This is fully explained in the next chapter.

using the Box-Cox model and then employ Lagrange Multiplier (LM) tests to confirm the existence of spatial dependence. Kim et al (2003) choose the semi-log functional form based on measures of collinearity, and then use LM tests to select the appropriate spatial model. However, recent evidence indicates that these two issues need to be considered jointly. McMillen (2003) shows that spatial tests may improperly detect spatial dependence if the functional form is incorrectly specified.

Baltagi and Li (2001, 2004) developed LM tests that jointly test functional form and either spatial lag or spatial error dependence. These tests allow an exploration of spatial dependence under alternative functional forms and are based on a spatial Box-Cox transformation. Boxall et al (2005) use these tests to determine functional form in a hedonic analysis of oil and natural gas facilities. Le and Li (2008) and Le (2009) recently developed Double-Length Regression (DLR) tests that yield the same results as the Baltagi and Li LM tests but do not require the Hessian or second derivatives of the log-likelihood function. Due to the size of the dataset in the present dissertation, the DLR tests present less computational difficulties than the LM tests. The DLR tests are used in the present paper to examine issues of both spatial dependence and functional form.

1.5. Results

1.5.1. Estimation

Since there is no theoretical basis for the functional form of the hedonic property price model (Le and Li 2009), several tests are performed. The traditional method of testing functional forms involves the use of a Box-Cox test. This test, which is explained in Chapter 2, uses a flexible functional form that can be employed to evaluate hypotheses about the functional form of the model. Results of this test supported the double log regression over the linear or semi-log specifications.[12] Cropper et al. (1988) show that the double-log model performs particularly well in hedonic models in the face of misspecification and Kang and Reichert (1996) find that the double log model performed the best at forecasting property prices.

To further investigate functional form specification and spatial dependence, several DLR tests are performed.[13] First, DLR tests of spatial lag and spatial error dependence are carried out under both linear and double log specifications. In each case the null hypothesis of no spatial dependence is rejected. Next, joint tests of spatial dependence and the linear and double-log functional forms are implemented. The results of these joint tests, which are presented in Chapter 2, support the existence of spatial dependence and the use of the double-log functional form.

The DLR tests of spatial dependence do not consistently reject either spatial lag or spatial error dependence at a higher level than the other, which is the typical method of selection between the two forms (Mueller and Loomis 2008). To choose between the spatial lag and spatial error models, the second chapter of this dissertation examines both models using out-of-sample prediction errors with the same data as this chapter. Based on seven categories of prediction error used in Case et al. (2004), the spatial lag model decisively outperforms the spatial error model. Therefore the spatial lag model is used.[14]

For purposes of comparison, both spatial and non-spatial results are presented. The spatial and non-spatial models were both estimated using maximum likelihood regression, and selected estimation results are reported in Table 3. Breusch-Pagan tests indicated the presence of

[12] The procedure is fully explained in Chapter 2, where estimation results are also presented.
[13] Since the eigenvalues of the SWM need to be calculated, the size of the data impose some computational constraints with these tests. To compensate for this, the data is split into the individual municipalities for DLR testing. This testing procedure is fully explained in Chapter 2.
[14] The non-spatial models are estimated on Stata. All of the spatial work, including SWM construction and spatial lag estimation, is done on Matlab. The code for the spatial regressions can be found on www.spatial-econometrics.com.

heteroskedasticity, so White's corrected standard errors are used. The complete set of coefficients, including time and lake fixed effects, appear in Table 6. Models 1, 2, and 3 are the non-spatial models, while Models 1S, 2S and 3S are the spatial lag models.

The three different model variations are used to test the hypotheses of Section III. Hypothesis 1 can be tested through Model 1, where the coefficient on *Ln(SDM)*Ln(Distance)* is 0.116 and significant at the 0.01 level, indicating there is an edge effect present in the data. Water quality has a larger effect on waterfront homes than non-waterfront homes, illustrated by the positive and significant coefficient on this variable in all models of Table 3. Also, consistent with Poor et al (2007), the positive and significant coefficient on *Ln(SDM)* in Model 1 indicates that water quality has a positive effect on non-waterfront homes.

To investigate the character and extent of benefits beyond the lakefront, Model 2 contains the interaction term *Ln(SDM)*Ln(Distance)*. The coefficient on this variable is -0.017 and significant, illustrating a negative relationship between the value of water quality and lake distance. The proximity effect of Hypothesis 2 is therefore supported by this result and the effect of a water quality change is dependent on the location of the home.

The Area Effect can be examined in Model 3, where the addition of *Ln(SDM)*Ln(Area)* yields the full form of Equation (1). The positive and significant coefficient on this interaction term indicates that the effect of water quality on property prices is positively related to the size of the lake. With all three interaction terms in the model the two previous hypotheses still hold, but the magnitude of the *Ln(SDM)*Lakefront* coefficient decreases by approximately 30%. Consequently, the benefit to waterfront homes may be overstated if area and distance are not controlled for. Alternatively, the effect of lake proximity on water quality is relatively strong, the coefficient on *Ln(SDM)*Ln(Distance)* remains constant over all specifications. A likelihood ratio test of the joint significance of the three water quality interaction terms confirms their inclusion at the 99% level and Model 3 is the preferred specification.

For the coefficients not appearing in Table 3, all structural home characteristics (such as the number of bathrooms, home size, and parcel size) were positive and significant, as presented in Table 6. The location and neighborhood variables all had expected signs except the variable indicating proximity to an airport, which was positive, and the "percent black" and "percent over 65 variables," which were insignificant. All of the annual dummy variables in both models were significant at the 99% level and negative (but increasing slightly each year). The variable left out of the equation was the 2004 indicator, so these negative and increasing coefficients capture the effect of escalating home prices in Orange County during this period. An F-test of the hypothesis that the year dummies are equal to zero is rejected at the 99% level ($F_{8;\,54,539} = 7,875.9$).

Although the spatial parameter ρ is significant in each model, its size is small and the differences between the spatial and non-spatial coefficients are not substantial. None of the three hypotheses are affected when spatial dependence is controlled for. The largest disparity between models is in the latitude and longitude variables, which partially represented spatial influences in the non-spatial lag model. Leggett and Bockstael (2000) and Mueller and Loomis (2008) also find that controlling for detected spatial dependence does not substantially change the magnitude of the coefficients in their hedonic models.

1.5.2. Marginal Analysis

To illustrate the policy relevance of these results, the effect of a one foot change in SDM on the average home price (expressed in dollars) appears in Table 4. For the mean level of SDM, this represents a 17% change in water quality. These implicit prices are estimated at mean lake distance values for the first two rows; in the next five rows the distance to the lake is varied. The final row in the table contains the average implicit price of lake proximity. Kim et al (2003) illustrate the derivation of implicit prices for a spatial lag model; for the present paper the implicit price of water quality is:

$$\frac{\partial P}{\partial SDM}=\left(\frac{1}{1-\rho}\right)\left(\frac{P}{SDM}\right)\left(\beta_{SDM}+\beta_{WF*SDM}*WF+\beta_{Dist*SDM}*\ln(Dist)+\beta_{Area*SDM}*\ln(Area)\right) \tag{1.12}$$

The mean implicit price of water quality for non-lakefront homes is positive in all models, and varies from a minimum of $435 in Model 2 to a maximum of $891 in Model 3S. The implicit price of water quality for a waterfront home is substantially larger than the non-waterfront estimate in all models. The Model 3S waterfront implicit price of $11,784.28 represents a 2.6% improvement in the average lakefront home price for a one foot change in SDM.

In the next five rows of Table 4 the implicit price is evaluated at several distance intervals from the lake, which allows a more complete picture of the relationship between water quality, lake proximity, and non-waterfront home prices. Models 1 and 1S do not allow distance to effect the implicit price of water quality so only average values for lakefront and non-lakefront can be evaluated. In the other models, the Proximity Effect can be clearly seen as implicit prices diminish with distance. For example, the implicit price of non-lakefront homes in Model 3S decreases from $1,786.31 at 100 meters from the lake to $509.71 at a distance of 900 meters away. The importance of including lake size can be seen in the differences between the Model 2S and 3S implicit prices, especially with non-lakefront homes. All of the 3S values are substantially higher, with the average non-lakefront implicit price twice as high as the 2S value. Overall, the dollar values in Table 4 indicate that sizable welfare benefits would be excluded if the analysis was restricted to waterfront homes.

The results of this paper clearly indicate that non-lakefront homes have a positive implicit price for water quality in local lakes. Furthermore, this value differs between waterfront and non-waterfront homes, and varies with location and lake area. These are critical issues in the calculation of water quality benefits, and could result in large differences in policy recommendations. To illustrate these differences, three representative lakes were chosen for benefit estimates; a small lake (70 acres), a medium sized lake (271 acres), and a large lake (just over 1,000 acres). The assessed market values of all lakefront and non-lakefront homes within 1,000 meters of the lake were used to calculate the total benefits of an improvement of one foot of water clarity using the estimates of the full spatial lag model. To accurately measure total benefits, this calculation contains all home values in the area, not just the property sales (as in the data used to calculate the hedonic model).[15]

Table 5 contains the results of the benefit calculations, in all three of the lakes the total non-lakefront benefits were larger than the total lakefront benefits. In the small lake, lakefront benefits were $971,452 and non-lakefront benefits were $1,079,068, representing a doubling of total benefits. Including the non-lakefront homes increases the total benefits by a factor of 4.6 in the medium sized lake and in the large lake total non-lakefront benefits are $1,992,136 larger than lakefront benefits. The last column of Table 5 contains the total non-lakefront benefits within 500 meters of the lake. Even when this smaller sample is used, the inclusion of non-lakefront benefits approximately doubles the total estimate of benefits in each case, illustrating the importance of these values.

[15] It is recognized that the assessed value of these homes represent an imperfect estimate of their potential selling price. However, the focus of this exercise is a comparison of the total non-waterfront benefits to the total waterfront benefits. If all homes are undervalued by this assessment procedure by a similar amount, the ratio of waterfront to non-waterfront benefits should change only marginally.

1.6. Conclusion

Water quality is an important policy issue in the United States, illustrated by the extensive requirements of the original CWA and the recently proposed "Clean Water Restoration Act," which seeks to expand the jurisdiction of the CWA from "navigable waters" (as defined in the original CWA) to all waters in the US.[16] The benefits of improved water quality accrue in several areas, including multiple forms of recreation, aesthetic appeal, ecosystem support, and property price appreciation. In past literature, it was assumed that the property price benefits only represented a small fraction of the total benefits and were confined to lakefront properties. The present paper explores the extent of property price benefits by including both lakefront and non-lakefront homes in a large spatial hedonic analysis in Central Florida.

Three hypotheses about the effect of water quality are tested and supported. First, a waterfront edge effect is found, indicating that waterfront and non-waterfront properties have a significantly different implicit price of water quality. Second, the implicit prices are directly affected by lake proximity, diminishing as homes are located farther from the lake. Third, the area of a lake also affects the implicit price of water quality, with larger lakes having higher value implicit prices. Overall, the results of this paper yield considerable support for the inclusion of non-lakefront homes in hedonic property analysis. Water quality benefits extend far beyond the lakefront. These results yield several policy recommendations. One of the main objectives of conducting environmental valuation is to inform cost benefit analysis. In each of the three lakes analyzed in this paper, including the gains to non-lakefront properties would more than double the total estimate of water quality benefits. This represents a considerable increase in

[16] H.R. 2421, introduced to the US House of Representatives by James L. Oberstar (5/22/2007) and S.1870, introduced to the US Senate by Russell D. Feingold (04/09/2008)

benefits and shows that restricting the sample to waterfront homes could cause a severe downward bias in a cost benefit analysis.

Furthermore, the shape and magnitude of the implicit price gradient could be used to design more effective taxing regimes around lakes. Past taxing schemes for lake improvement activities in Orange County assign an MSTU (municipal service taxing unit) fee to the property tax of each home in a particular neighborhood surrounding a lake. These MSTU fees are similar to recycling or trash collection fees and are instituted through a majority vote by the neighborhood.[17] When approved, every home in the neighborhood is assigned the same tax rate increase. The results of the present study indicate that these uniform fees do not accurately represent the distribution of benefits from lake improvements, especially with arbitrarily defined neighborhood boundaries. If the neighborhood is small, there is the problem of free riding by those outside the defined neighborhood. If the neighborhood is large, those at the periphery may be unfairly taxed. It would be more efficient to tax homes differentially based on lake proximity and waterfront status.

The magnitude of the implicit prices in this study also indicate that widespread lake improvement programs could cause a considerable increase in required property taxes through home price appreciation. Given the size of the total estimated property price benefits for the three lake examples in this study ($700,000 - $5,000,000), lake water quality improvement projects could yield a sizeable amount of tax revenue. This tax revenue could be used to fund additional lake programs. Given that 45% of our nation's lakes are still classified as "impaired" by the EPA (US EPA 2003), these results represent an opportunity for significant improvement.

[17] There have only been a handful of these MSTU programs implemented. Only 5 out of the 146 lakes in the present study contain MSTU programs, which are only active during some years. Attempts to relate MSTU funding to water quality found no correlation between the two.

Several open questions remain after this exercise, some of which are addressed in later chapters. First, there are multiple potential approaches to the spatial modeling of the hedonic equation. Chapter 2 explores several different methods of incorporating spatial dependence. Second, SDM has been used as the only indicator of water quality. However, this is only one of several available indicators; alternative measurements may imply different implicit prices (Michael et al. 2000). Chapter 3 examines several alternative indicators of water quality that are used by state regulators to assess the health of lakes. Furthermore, additional research should investigate the functional form of the implicit price gradients. In the current paper these gradients are functions of the lake distance and lake area. The specific form may impact the identified benefit extent.

Tables and Figures: Chapter 1

Table 1: Summary Statistics on Water Quality Over Time

	Secchi Disk (feet)				
Year	Mean	Std Dev	Minimum	Maximum	Lakes (N)
1990	5.53	3.25	0.66	15.49	101
1991	5.30	3.18	1.31	15.39	104
1992	5.51	3.35	1.31	15.06	103
1993	5.75	3.52	1.25	16.67	112
1994	5.86	3.37	1.51	16.40	137
1995	5.76	3.33	1.21	16.47	136
1996	5.86	3.45	1.54	15.49	143
1997	6.18	3.73	1.64	18.05	152
1998	5.23	2.81	1.54	17.26	148
1999	5.25	3.04	1.48	15.35	152
2000	5.23	3.20	0.89	13.88	137
2001	5.27	2.84	0.98	13.12	134
2002	4.95	2.90	0.66	14.76	126
2003	5.23	2.69	1.31	14.53	129
2004	5.33	2.69	1.15	12.99	125

Table 2: Summary Statistics of Property Sales: Lakefront and Non-Lakefront Homes

		Lakefront (N =1,496)		Non-Lakefront (N =53,216)	
Variable	Units	Mean	Std Dev	Mean	Std Dev
Property Characteristics					
Sales Price	2002 Dollars	452,646.2	370,667.8	199,982.4	201,110.9
Heated Area	Square Feet	2,769.7	1,295.9	1,962	949.5
Area of Parcel	Square Feet	30,002.3	30,487.7	11,580.9	11,737.0
Number of Bedrooms	--	3.6	1.0	3.3	0.8
Number of Bathrooms	--	2.8	1.1	2.2	0.9
Home Age	Years	24.1	13.5	18.6	15.1
% With Pool	--	20.1	---	20.7	---
Spatial Characteristics					
Distance to Nearest Lake	Meters	42.37	23.70	467.13	267.7
Area of Nearest Lake	Acres	519.3	635.2	278.1	445.2
Distance to CBD	Meters	8,824.4	5,123.9	9,265.8	5,150.2
Latitude Coordinate	Degrees	654,756.9	7,415.4	653,223.8	7,940.8
Longitude Coordinate	Degrees	505,615.8	5,123.9	505,664.6	6,272.2
% In airport noise zone	--	8.9	--	15.9	--
Census Block Characteristics					
% of Population White	--	88.6	--	77.9	--
% of Population Black	--	4.9	--	12.5	--
% of Population > 65	--	15.2	--	11.1	--
Median Household Income	2002 Dollars	67,355.2	30,343.2	58,876.8	24,863.1
Distribution of Sales by Year					
% of sales in 1996	--	10.4	--	9.2	--
% of sales in 1997	--	11.1	--	10.9	--
% of sales in 1998	--	12.9	--	12.1	--
% of sales in 1999	--	14.0	--	13.4	--
% of sales in 2000	--	11.0	--	11.4	--
% of sales in 2001	--	8.9	--	9.9	--
% of sales in 2002	--	11.3	--	9.9	--
% of sales in 2003	--	11.0	--	11.0	--
% of sales in 2004	--	9.4	--	12.2	--

Table 3: Selected Hedonic Estimation Results

	Non-Spatial						Spatial					
	Model 1		Model 2		Model 3		Model 1S		Model 2S		Model 3S	
Variable	β	SE	β	SE	β	SE	β	SE	β	SE	β	SE
Lakefront	0.087*	0.026	0.143*	0.028	0.146*	0.028	0.087*	0.019	0.143*	0.020	0.146*	0.020
Ln(Distance)	-0.062*	0.002	-0.035*	0.004	-0.035*	0.004	-0.062*	0.001	-0.035*	0.004	-0.035*	0.004
Ln(SDM)	0.017*	0.003	0.117*	0.017	-0.028	0.032	0.017*	0.004	0.118*	0.014	-0.030	0.031
Ln(SDM)*Lakefront	0.116*	0.015	0.080*	0.016	0.078*	0.016	0.116*	0.010	0.081*	0.011	0.079*	0.011
Ln(SDM)*ln(dist)	---	---	-0.017*	0.003	-0.017*	0.003	---	---	-0.017*	0.002	-0.017*	0.002
Ln(SDM)*ln(area)	---	---	---	---	0.011*	0.002	---	---	---	---	0.012*	0.002
Ln(X_coord)	-6.273*	1.056	-6.290*	1.055	-6.442*	1.055	-6.335*	0.106	-6.362*	0.034	-6.509*	0.053
Ln(Y_coord)	3.689*	0.695	3.509*	0.695	3.389*	0.694	3.763*	0.656	3.583*	0.662	3.462*	0.666
Time variables	*Jointly Significant*		*Jointly Significant*		*Jointly Significant*		*Jointly Significant*		*Jointly Significant*		*Jointly Significant*	
Lake variables	*Jointly Significant*		*Jointly Significant*		*Jointly Significant*		*Jointly Significant*		*Jointly Significant*		*Jointly Significant*	
ρ	---		---		---		0.001*	0.000	0.001*	0.000	0.001*	0.000
Pseudo adj-R^2	0.893		0.894		0.894		0.893		0.893		0.893	

Note: * denotes significance at the 1% level. The full set of estimation results is reported in the Appendix. Also note that White's robust standard errors are used to control for heteroskedasticity in Models 1, 2, and 3.

Table 4: Implicit Price of a One Foot Increase in Water Clarity

Scenario	Model 1	Model 2	Model 3	Model 1S	Model 2S	Model 3S
Edge Effect						
Waterfront	$10,187.37	$10,295.22	$11,663.86	$10,270.89	$10,382.14	$11,784.28
	(1,162.90)	(1,161.89)	(1,190.984)	(816.66)	(816.39)	(857.10)
Non-Waterfront	$568.12	$434.84	$877.92	$572.16	$437.46	$890.71
	(118.41)	(115.35)	(151.06)	(125.07)	(126.25)	(151.95)
Proximity Effect						
Non-Waterfront (300m)	---	$690.91	$1,132.23	---	$696.47	$1,148.01
		(120.74)	(156.34)		(126.06)	(151.62)
Non-Waterfront (700m)	---	$200.98	$645.65	---	$200.91	$655.73
		(122.99)	(155.91)		(134.12)	(158.71)
Area Effect						
Non-Waterfront (100 acres)	---	---	$485.42	---	---	$489.26
			(115.57)			(126.59)
Non-Waterfront (1000 acres)	---	---	$1,368.89	---	---	$1,392.89
			(226.29)			(218.57)

Implicit prices are evaluated at the sample means. Standard errors computed using the delta method.

Table 5: Total Water Quality Benefits from a 1 Foot Increase in SDM

Lake Size	Lakefront Benefits	Non-Lakefront Benefits (Homes within 1,000m)	Non-Lakefront Benefits (500m)
70 Acres	$971,452	$1,079,068	$728,822
271 Acres	$698,144	$3,219,550	$1,991,001
1000 acres	$5,441,246	$7,433,382	$5,057,007

Table 6: Full Regression Results

	Model 1	Model 2	Model 3	Model 1S	Model 2S	Model 3S
Waterfront	0.087*	0.143*	0.146*	0.087*	0.143*	0.146*
(1 if WF, 0 if non-WF)	(0.026)	(0.028)	(0.028)	(0.019)	(0.020)	(0.020)
Ln(lake dist)	-0.062*	-0.035*	-0.035*	-0.062*	-0.035*	-0.035*
(Distance to lake in m)	(0.002)	(0.004)	(0.004)	(0.001)	(0.004)	(0.004)
Ln(SDM)	0.017*	0.117*	-0.028	0.017*	0.118*	-0.030
(Feet of clarity)	(0.003)	(0.017)	(0.032)	(0.004)	(0.014)	(0.031)
ln(SDM)*Waterfront	0.116*	0.080*	0.078*	0.116*	0.081*	0.079*
	(0.015)	(0.016)	(0.016)	(0.01)	(0.011)	(0.011)
Ln(SDM)*ln(dist)	---	-0.017*	-0.017*	---	-0.017*	-0.017*
		(0.003)	(0.003)		(0.002)	(0.002)
Ln(SDM)*ln(lake area)	---	---	0.011*	---	---	0.012*
(Lake area in acres)			(0.002)			(0.002)
ln(x_coord)	-6.273*	-6.299*	-6.442*	-6.335*	-6.362*	-6.509*
	(1.057)	(1.055)	(1.056)	(0.106)	(0.034)	(0.053)
ln(y_coord)	3.689*	3.509*	3.389*	3.763*	3.583*	3.462*
	(0.696)	(0.695)	(0.694)	(0.656)	(0.662)	(0.666)
Canalfront	0.112*	0.104*	0.104*	0.113*	0.105*	0.105*
(1 if next to canal, 0 otherwise)	(0.016)	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)
Golffront	0.331*	0.332*	0.332*	0.332*	0.333*	0.334*
(1 if next to golf course, 0 otherwise)	(0.038)	(0.038)	(0.038)	(0.033)	(0.033)	(0.033)
Ln(bath)	0.119*	0.120*	0.120*	0.119*	0.120*	0.120*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Pool	0.014*	0.014*	0.014*	0.014*	0.014*	0.014*
(1 if home has poo1, 0 otherwise)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(age)	-0.070*	-0.071*	-0.071*	-0.070*	-0.070*	-0.070*
(Age of home in years)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(area_heated)	0.673*	0.673*	0.673*	0.673*	0.672*	0.672*
(Area in sq. feet)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Ln(area parcel)	0.165*	0.166*	0.165*	0.166*	0.167*	0.167*
(Area in sq. feet)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Ln(dist CBD)	-0.187*	-0.191*	-0.189*	-0.186*	-0.190*	-0.189*
(Distance in m)	(0.011)	(0.011)	(0.011)	0.001)	(0.010)	(0.010)
Near airport	0.027*	0.029*	0.029*	0.027*	0.028*	0.028*
(1 if near airport, 0 otherwise)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln(Med income)	0.106*	0.108*	0.107*	0.105*	0.107*	0.106*
(Income in 2002 dollars)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Percent white00	0.249*	0.247*	0.246*	0.248*	0.245*	0.244*
(In percentage of census block)	(0.031)	(0.031)	(0.031)	(0.032)	(0.032)	(0.032)
Percent black00	0.052	0.053	0.052	0.052	0.054	0.053

(In percentage of census block)	(0.035)	(0.035)	(0.035)	(0.034)	(0.034)	(0.034)
Percent over6500	0.015	0.023	0.022	0.016	0.024	0.023
(In percentage of census block)	(0.022)	(0.022)	(0.022)	0.020)	(0.020)	(0.020)
year_1996†	-0.820*	-0.820*	-0.821*	-0.818*	-0.818*	-0.819*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_1997	-0.762*	-0.762*	-0.762*	-0.760*	-0.760*	-0.760*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_1998	-0.693*	-0.693*	-0.692*	-0.692*	-0.692*	-0.690*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_1999	-0.607*	-0.606*	-0.605*	-0.606*	-0.605*	-0.603*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2000	-0.498*	-0.498*	-0.498*	-0.497*	-0.497*	-0.497*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2001	-0.380*	-0.380*	-0.379*	-0.379*	-0.379*	-0.378*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2002	-0.287*	-0.287*	-0.285*	-0.286*	-0.286*	-0.284*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2003	-0.156*	-0.157*	-0.156*	-0.156*	-0.156*	-0.155*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
BASS_LAKE‡	-0.127*	-0.130*	-0.080*	-0.125*	-0.128*	-0.077*
	(0.025)	(0.025)	(0.027)	(0.016)	(0.016)	(0.018)
BAY_LAKE_B	-0.306*	-0.297*	-0.260*	-0.302*	-0.293*	-0.255*
	(0.033)	(0.033)	(0.034)	(0.032)	(0.032)	(0.033)
BEARHEAD_L~E	-0.089*	-0.090*	-0.041	-0.086*	-0.087*	-0.037
	(0.025)	(0.025)	(0.027)	(0.026)	(0.025)	(0.026)
BIG_SAND_L~E	0.203*	0.198*	0.162*	0.203*	0.198*	0.162*
	(0.014)	(0.014)	(0.015)	(0.014)	(0.014)	(0.016)
CLEAR_LAKE	-0.380*	-0.375*	-0.361*	-0.378*	-0.373*	-0.359*
	(0.020)	(0.020)	(0.021)	(0.018)	(0.018)	(0.018)
DEEP_LAKE	-0.082***	-0.078*	0.003	-0.080*	-0.077**	0.006
	(0.044)	(0.044)	(0.046)	(0.030)	(0.031)	(0.036)
KASEY_LAKE	-0.259*	-0.256*	-0.197*	-0.261*	-0.258*	-0.198*
	(0.019)	(0.019)	(0.022)	(0.020)	(0.021)	(0.024)
KELLY_LAKE	-0.325*	-0.322*	-0.222*	-0.326*	-0.324*	-0.222*
	(0.019)	(0.019)	(0.027)	(0.018)	(0.019)	(0.027)
KRISTY_LAKE	-0.334*	-0.333*	-0.239*	-0.335*	-0.334*	-0.238*
	(0.018)	(0.018)	(0.026)	(0.019)	(0.019)	(0.026)
LAKE_ADAIR	0.092*	0.091*	0.134*	0.092*	0.091*	0.135*
	(0.024)	(0.024)	(0.026)	(0.021)	(0.021)	(0.023)
LAKE_ANDER~N	-0.091*	-0.093*	-0.038	-0.088*	-0.090*	-0.034**
	(0.024)	(0.024)	(0.027)	(0.016)	(0.016)	(0.018)

LAKE_ANGEL	-0.492*	-0.499*	-0.435*	-0.490*	-0.496*	-0.431*
	(0.029)	(0.029)	(0.032)	(0.023)	(0.023)	(0.025)
LAKE_ARNOLD	-0.132*	-0.134*	-0.085*	-0.130*	-0.133*	-0.083*
	(0.025)	(0.025)	(0.027)	(0.018)	(0.018)	(0.020)
LAKE_BALDWIN	0.010	0.012	0.026	0.013	0.016	0.030
	(0.030)	(0.030)	(0.030)	(0.018)	(0.019)	(0.020)
LAKE_BARTON	-0.214*	-0.215*	-0.193*	-0.211*	-0.212*	-0.190*
	(0.029)	(0.029)	(0.029)	(0.016)	(0.016)	(0.017)
LAKE_BEARD~L	-0.551*	-0.558*	-0.471*	-0.544*	-0.551*	-0.462*
	(0.061)	(0.061)	(0.064)	(0.048)	(0.048)	(0.051)
LAKE_BEAUTY	-0.150*	-0.149*	-0.032	-0.150*	-0.149*	-0.029
	(0.043)	(0.043)	(0.049)	(0.039)	(0.039)	(0.045)
LAKE_BELL	-0.206*	-0.203*	-0.160*	-0.206*	-0.203*	-0.159*
	(0.029)	(0.029)	(0.030)	(0.024)	(0.025)	(0.027)
LAKE_BERRY	0.151*	0.151*	0.183*	0.151*	0.151*	0.184*
	(0.031)	(0.031)	(0.031)	(0.018)	(0.019)	(0.021)
LAKE_BESSIE	0.542*	0.527*	0.532*	0.544*	0.528*	0.533*
	(0.048)	(0.048)	(0.048)	(0.023)	(0.023)	(0.023)
LAKE_BLANCHE	0.059*	0.057*	0.070*	0.059*	0.057*	0.071*
	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.014)
LAKE_BUCHA~N	-0.151*	-0.152*	-0.113*	-0.149*	-0.150*	-0.111*
	(0.025)	(0.025)	(0.026)	(0.029)	(0.029)	(0.030)
LAKE_BUCK	0.513*	0.512*	0.532*	0.520*	0.520*	0.540*
	(0.059)	(0.059)	(0.059)	(0.025)	(0.024)	(0.022)
LAKE_BUMBY	-0.134*	-0.134*	-0.065**	-0.131*	-0.132*	-0.061**
	(0.023)	(0.023)	(0.026)	(0.026)	(0.026)	(0.029)
LAKE_BURKETT	0.008	0.013	0.050	0.011	0.016	0.054**
	(0.042)	(0.042)	(0.043)	(0.024)	(0.026)	(0.028)
LAKE_BUTLER	-0.012	-0.015	-0.066*	-0.01	-0.013	-0.066*
	(0.019)	(0.019)	(0.022)	(0.012)	(0.013)	(0.016)
LAKE_C	-0.170*	-0.172*	-0.105*	-0.168*	-0.170*	-0.102*
	(0.028)	(0.028)	(0.031)	(0.017)	(0.017)	(0.021)
LAKE_CANE_A	-0.045*	-0.047*	-0.022**	-0.045*	-0.047*	-0.021
	(0.010)	(0.010)	(0.011)	(0.013)	(0.013)	(0.014)
LAKE_CATHE~B	0.139***	0.149**	0.201*	0.140*	0.150*	0.203*
	(0.072)	(0.072)	(0.073)	(0.046)	(0.046)	(0.047)
LAKE_CAY_DEE	-0.113*	-0.116*	-0.034	-0.112*	-0.115*	-0.032
	(0.030)	(0.030)	(0.034)	(0.021)	(0.021)	(0.026)
LAKE_CHARITY	0.129*	0.132*	0.181*	0.128*	0.131*	0.180*
	(0.046)	(0.046)	(0.047)	(0.035)	(0.036)	(0.037)
LAKE_CHASE	0.487*	0.478*	0.490*	0.489*	0.480*	0.492*

	(0.066)	(0.065)	(0.065)	(0.031)	(0.031)	(0.031)
LAKE_CHERO~E	0.055***	0.053***	0.109*	0.056**	0.054**	0.111*
	(0.032)	(0.032)	(0.034)	(0.027)	(0.027)	(0.029)
LAKE_CHRIS~E	-0.158*	-0.159*	-0.103*	-0.156*	-0.156*	-0.099*
	(0.025)	(0.025)	(0.028)	(0.033)	(0.033)	(0.035)
LAKE_COMO	-0.009	-0.015	0.088*	-0.008	-0.013	0.091*
	(0.026)	(0.026)	(0.033)	(0.019)	(0.019)	(0.027)
LAKE_CONCORD	-0.044	-0.044	-0.010	-0.043	-0.043	-0.009
	(0.032)	(0.032)	(0.033)	(0.028)	(0.028)	(0.029)
LAKE_CONWAY	-0.043***	-0.047**	-0.096*	-0.040*	-0.044*	-0.094*
	(0.023)	(0.023)	(0.025)	(0.012)	(0.011)	(0.015)
LAKE_COPEL~D	-0.018	-0.019	0.032	-0.016	-0.017	0.036
	(0.038)	(0.038)	(0.040)	(0.030)	(0.030)	(0.031)
LAKE_DANIEL	-0.135*	-0.131*	-0.064**	-0.134*	-0.131*	-0.062***
	(0.028)	(0.028)	(0.031)	(0.030)	(0.030)	(0.033)
LAKE_DESTINY	0.223*	0.220*	0.281*	0.221*	0.219*	0.281*
	(0.030)	(0.030)	(0.032)	0.032)	(0.032)	(0.035)
LAKE_DOT	-0.485*	-0.492*	-0.413*	-0.482*	-0.490*	-0.409*
	(0.045)	(0.045)	(0.047)	(0.035)	(0.035)	(0.038)
LAKE_DOVER	-0.094*	-0.096*	-0.009	-0.092**	-0.094*	-0.005
	(0.030)	(0.030)	(0.035)	(0.036)	(0.036)	(0.040)
LAKE_DOWN	0.031*	0.034*	-0.005	0.032*	0.035*	-0.006
	(0.011)	(0.011)	(0.013)	(0.011)	(0.011)	(0.013)
LAKE_DOWNEY	-0.005	0.004	0.074***	-0.003	0.007	0.078*
	(0.041)	(0.041)	(0.043)	(0.014)	(0.016)	(0.022)
LAKE_DRUID	-0.005	-0.010	0.049***	-0.004	-0.009	0.051**
	(0.025)	(0.025)	(0.028)	(0.019)	(0.019)	(0.022)
LAKE_EMERALD	0.077	0.077	0.141*	0.081*	0.081*	0.146*
	(0.048)	(0.048)	(0.050)	(0.023)	(0.024)	(0.028)
LAKE_EOLA	-0.105**	-0.112**	-0.062	-0.102*	-0.109*	-0.059
	(0.045)	(0.045)	(0.046)	(0.035)	(0.035)	(0.036)
LAKE_ESTELLE	0.031	0.031	0.071**	0.031	0.031	0.071**
	(0.032)	(0.032)	(0.033)	(0.028)	(0.029)	(0.030)
LAKE_FAIRH~E	-0.072***	-0.066***	-0.006	-0.069**	-0.062***	-0.001
	(0.040)	(0.040)	(0.041)	(0.033)	(0.033)	(0.035)
LAKE_FAIRV~W	-0.259*	-0.258*	-0.258*	-0.258*	-0.257*	-0.257*
	(0.023)	(0.022)	(0.022)	(0.018)	(0.018)	(0.019)
LAKE_FAITH	-0.005	-0.004	0.056	-0.003	-0.002	0.060
	(0.049)	(0.048)	(0.050)	(0.035)	(0.035)	(0.038)
LAKE_FARRAR	-0.090*	-0.101*	-0.018	-0.088*	-0.100*	-0.015
	(0.026)	(0.026)	(0.031)	(0.019)	(0.019)	(0.024)

LAKE_FORMOSA	-0.036	-0.036	0.007	-0.035	-0.035	0.009
	(0.030)	(0.029)	(0.031)	(0.024)	(0.024)	(0.026)
LAKE_FREDR~A	0.146*	0.152*	0.188*	0.148*	0.155*	0.191*
	(0.030)	(0.030)	(0.031)	(0.017)	(0.017)	(0.018)
LAKE_GATLIN	0.032	0.040	0.071**	0.035	0.044***	0.076*
	(0.032)	(0.032)	(0.032)	(0.025)	(0.024)	(0.025)
LAKE_GEAR	-0.049	-0.054***	0.032	-0.046***	-0.052***	0.037
	(0.033)	(0.033)	(0.037)	(0.027)	(0.027)	(0.032)
LAKE_GEM_A	-0.102*	-0.098*	-0.024	-0.101*	-0.098*	-0.022
	(0.032)	(0.032)	(0.035)	(0.023)	(0.024)	(0.029)
LAKE_GEM_M~Y	-0.048***	-0.051***	0.007	-0.047**	-0.050**	0.010
	(0.027)	(0.027)	(0.029)	(0.021)	(0.021)	(0.023)
LAKE_GEORGE	-0.024	-0.023	0.010	-0.021	-0.021	0.013
	(0.025)	(0.025)	(0.026)	(0.014)	(0.014)	(0.015)
LAKE_GEORGIA	0.064	0.073***	0.106**	0.065*	0.074*	0.108*
	(0.044)	(0.044)	(0.044)	(0.017)	(0.019)	(0.022)
LAKE_GILES	-0.162*	-0.162*	-0.118*	-0.160*	-0.160*	-0.115*
	(0.026)	(0.026)	(0.028)	(0.019)	(0.019)	(0.021)
LAKE_GLORIA	-0.064*	-0.067*	-0.029	-0.062*	-0.065*	-0.025
	(0.020)	(0.020)	(0.022)	(0.018)	(0.018)	(0.019)
LAKE_GREEN~D	-0.068***	-0.076**	0.019	-0.066**	-0.074**	0.023
	(0.037)	(0.037)	(0.041)	(0.030)	(0.030)	(0.035)
LAKE_HART	0.248*	0.240*	0.249*	0.254*	0.245*	0.255*
	(0.046)	(0.046)	(0.046)	(0.025)	(0.023)	(0.021)
LAKE_HIAWA~E	-0.132*	-0.132*	-0.121*	-0.133*	-0.133*	-0.121*
	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)
LAKE_HIGHL~D	0.004	-0.001	0.042	0.005	0.000	0.044***
	(0.028)	(0.028)	(0.029)	(0.022)	(0.022)	(0.024)
LAKE_HOLDEN	-0.347*	-0.344*	-0.325*	-0.345*	-0.342*	-0.323*
	(0.019)	(0.019)	(0.020)	(0.015)	(0.015)	(0.015)
LAKE_HOPE	-0.160*	-0.156*	-0.102***	-0.154*	-0.150*	-0.094***
	(0.061)	(0.060)	(0.061)	(0.051)	(0.052)	(0.053)
LAKE_HUNGE~D	-0.371*	-0.368*	-0.298*	-0.369*	-0.366*	-0.294*
	(0.037)	(0.037)	(0.039)	(0.030)	(0.030)	(0.033)
LAKE_IRMA	0.007	0.012	0.040	0.008	0.013	0.041**
	(0.039)	(0.038)	(0.039)	(0.015)	(0.017)	(0.019)
LAKE_ISLEW~H	0.329*	0.322*	0.358*	0.331*	0.324*	0.362*
	(0.083)	(0.082)	(0.082)	(0.039)	(0.039)	(0.040)
LAKE_IVANHOE	0.068*	0.066*	0.090*	0.068*	0.066*	0.091*
	(0.023)	(0.023)	(0.023)	(0.018)	(0.018)	(0.019)
LAKE_JACKSON	0.090**	0.094**	0.144*	0.091*	0.095*	0.147*

	(0.045)	(0.045)	(0.046)	(0.030)	(0.030)	(0.032)
LK_JEN_JEWEL	-0.151*	-0.150*	-0.120*	-0.149*	-0.148*	-0.117*
	(0.024)	(0.024)	(0.025)	(0.021)	(0.020)	(0.021)
LAKE_JESSA~E	-0.245*	-0.248*	-0.242*	-0.242*	-0.245*	-0.239*
	(0.018)	(0.018)	(0.018)	(0.013)	(0.013)	(0.013)
LAKE_KILLA~Y	-0.111*	-0.109*	-0.096*	-0.110*	-0.108*	-0.095*
	(0.026)	(0.026)	(0.026)	(0.018)	(0.018)	(0.019)
LAKE_KOZART	-0.303*	-0.315*	-0.277*	-0.302*	-0.314*	-0.275*
	(0.017)	(0.017)	(0.019)	(0.017)	(0.017)	(0.018)
LAKE_LANCA~R	-0.003	-0.003	0.033	-0.002	-0.002	0.036***
	(0.025)	(0.025)	(0.026)	(0.018)	(0.018)	(0.019)
LAKE_LAWSONA	0.115*	0.109*	0.172*	0.116*	0.110*	0.175*
	(0.028)	(0.028)	(0.031)	(0.023)	(0.023)	(0.026)
LAKE_LOUIS~B	0.404*	0.399*	0.414*	0.405*	0.401*	0.416*
	(0.052)	(0.052)	(0.052)	(0.023)	(0.023)	(0.024)
LAKE_LOVE	0.235*	0.229*	0.339*	0.233*	0.226*	0.339*
	(0.034)	(0.034)	(0.040)	(0.041)	(0.042)	(0.047)
LAKE_LOVELY	-0.288*	-0.289*	-0.248*	-0.288*	-0.289*	-0.247*
	(0.025)	(0.025)	(0.026)	(0.021)	(0.022)	(0.024)
LAKE_LURNA	0.001	-0.000	0.065**	0.002	0.001	0.068*
	(0.029)	(0.029)	(0.032)	(0.022)	(0.022)	(0.025)
LAKE_MABEL	0.163*	0.157*	0.158*	0.165*	0.158*	0.160*
	(0.027)	(0.026)	(0.026)	(0.023)	(0.024)	(0.024)
LAKE_MAITL~D	0.327*	0.329*	0.336*	0.327*	0.39*	0.36*
	(0.030)	(0.030)	(0.030)	(0.017)	(0.018)	(0.020)
LAKE_MANN	-0.406*	-0.408*	-0.390*	-0.403*	-0.405*	-0.387*
	(0.023)	(0.023)	(0.023)	(0.020)	(0.020)	(0.021)
LAKE_MARSHA	-0.006	-0.008	0.015	-0.006	-0.008	0.016
	(0.011)	(0.011)	(0.011)	(0.014)	(0.014)	(0.015)
LAKE_MINNE~A	0.125*	0.130*	0.159*	0.125*	0.130*	0.160*
	(0.032)	(0.032)	(0.032)	(0.021)	(0.022)	(0.024)
LAKE_MIZELL	0.320*	0.322*	0.359*	0.321*	0.323*	0.361*
	(0.041)	(0.041)	(0.042)	(0.027)	(0.027)	(0.029)
LAKE_NAN	-0.017	-0.012	0.042	-0.015	-0.010	0.046
	(0.040)	(0.040)	(0.041)	(0.027)	(0.028)	(0.031)
LAKE_NONA	0.433*	0.440*	0.419*	0.438*	0.446*	0.424*
	(0.048)	(0.047)	(0.047)	(0.027)	(0.026)	(0.025)
LAKE_OFWOODS	-0.691*	-0.710*	-0.653*	-0.687*	-0.706*	-0.648*
	(0.055)	(0.055)	(0.056)	(0.051)	(0.051)	(0.052)
LAKE_OLIVE	-0.126*	-0.126*	-0.054	-0.12*	-0.123*	-0.049
	(0.044)	(0.044)	(0.046)	(0.035)	(0.034)	(0.037)

LAKE_OLYMPIA	-0.267*	-0.264*	-0.242*	-0.268*	-0.265*	-0.242*
	(0.014)	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)
LAKE_ORLANDO	-0.243*	-0.241*	-0.219*	-0.244*	-0.241*	-0.219*
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)
LAKE_OSCEOLA	0.360*	0.362*	0.384*	0.360*	0.362*	0.384*
	(0.033)	(0.033)	(0.033)	(0.017)	(0.018)	(0.020)
LAKE_PAMELA	-0.359*	-0.359*	-0.297*	-0.358*	-0.358*	-0.295*
	(0.034)	(0.034)	(0.036)	(0.036)	(0.036)	(0.038)
LAKE_PEARL_B	-0.177*	-0.184*	-0.145*	-0.177*	-0.184*	-0.144*
	(0.012)	(0.012)	(0.015)	(0.014)	(0.013)	(0.015)
LAKE_PICKETT	0.257*	0.266*	0.256*	0.269*	0.277*	0.268*
	(0.078)	(0.078)	(0.078)	(0.050)	(0.051)	(0.052)
LAKE_PINEL~H	-0.109*	-0.112*	-0.077*	-0.107*	-0.110*	-0.074*
	(0.023)	(0.023)	(0.024)	(0.016)	(0.016)	(0.017)
LAKE_PORTER	-0.080*	-0.086*	-0.037	-0.078*	-0.085*	-0.035**
	(0.026)	(0.026)	(0.027)	(0.014)	(0.014)	(0.017)
LAKE_RABAMA	-0.120*	-0.123*	-0.038	-0.118*	-0.121*	-0.034
	(0.026)	(0.026)	(0.031)	(0.017)	(0.017)	(0.023)
LAKE_RICHM~D	-0.324*	-0.322*	-0.292*	-0.321*	-0.318*	-0.287*
	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.021)
LAKE_ROBERTS	0.027	0.022	0.042***	0.028*	0.022	0.043**
	(0.023)	(0.023)	(0.023)	(0.017)	(0.017)	(0.017)
LAKE_ROSE_B	-0.132*	-0.131*	-0.108*	-0.132*	-0.132*	-0.108*
	(0.010)	(0.010)	(0.011)	(0.012)	(0.012)	(0.013)
LAKE_ROWENA	-0.009	-0.012	0.024	-0.008	-0.011	0.026
	(0.027)	(0.027)	(0.028)	(0.020)	(0.021)	(0.022)
LAKE_SANTI~O	-0.117*	-0.120*	-0.054***	-0.115*	-0.117*	-0.050**
	(0.028)	(0.028)	(0.031)	(0.018)	(0.018)	(0.022)
LAKE_SARAH	-0.123*	-0.119*	-0.059	-0.123*	-0.19*	-0.058
	(0.036)	(0.036)	(0.038)	(0.041)	(0.041)	(0.043)
LAKE_SHADOW	-0.151*	-0.147*	-0.115*	-0.151*	-0.148*	-0.115*
	(0.025)	(0.025)	(0.025)	(0.022)	(0.023)	(0.024)
LAKE_SHANNON	-0.098*	-0.100*	-0.011	-0.097*	-0.099*	-0.008
	(0.028)	(0.028)	(0.032)	(0.017)	(0.018)	(0.025)
LAKE_SHEEN	0.223*	0.220*	0.201*	0.223*	0.221*	0.202*
	(0.019)	(0.019)	(0.019)	(0.016)	(0.016)	(0.017)
LAKE_SHERW~D	-0.185*	-0.189*	-0.168*	-0.186*	-0.190*	-0.169*
	(0.012)	(0.012)	(0.012)	(0.014)	(0.014)	(0.014)
LAKE_SILVER	0.018	0.018	0.051**	0.018	0.018	0.052*
	(0.021)	(0.021)	(0.022)	(0.017)	(0.017)	(0.018)
LAKE_STARKE	-0.302*	-0.298*	-0.287*	-0.302*	-0.298*	-0.287*

	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)
LAKE_SUE	0.112*	0.112*	0.135*	0.113*	0.113*	0.136*
	(0.027)	(0.027)	(0.028)	(0.017)	(0.017)	(0.019)
LAKE_SUNSET	-0.486*	-0.493*	-0.443*	-0.481*	-0.488*	-0.437*
	(0.039)	(0.039)	(0.041)	(0.033)	(0.033)	(0.034)
LAKE_SUSAN~H	-0.090*	-0.092*	-0.057***	-0.087*	-0.089*	-0.054*
	(0.029)	(0.029)	(0.030)	(0.016)	(0.017)	(0.019)
LAKE_SYBELIA	-0.024	-0.024	0.010	-0.024	-0.024	0.011
	(0.036)	(0.036)	(0.036)	(0.026)	(0.027)	(0.029)
LAKE_SYLVAN	0.264*	0.266*	0.328*	0.264*	0.266*	0.329*
	(0.034)	(0.034)	(0.036)	(0.021)	(0.022)	(0.026)
LAKE_TENNE~E	-0.177*	-0.181*	-0.111*	-0.174*	-0.178*	-0.106*
	(0.027)	(0.027)	(0.030)	(0.018)	(0.018)	(0.022)
LAKE_TERRACE	-0.187*	-0.190*	-0.104*	-0.185*	-0.188*	-0.100*
	(0.024)	(0.024)	(0.030)	(0.016)	(0.016)	(0.023)
LAKE_THERESA	-0.064**	-0.066**	0.057	-0.063*	-0.065*	0.060***
	(0.030)	(0.030)	(0.038)	(0.024)	(0.024)	(0.034)
LAKE_TIBET	0.297*	0.295*	0.253*	0.298*	0.296*	0.254*
	(0.014)	(0.014)	(0.016)	(0.013)	(0.013)	(0.015)
LAKE_UNDER~L	-0.129*	-0.131*	-0.108*	-0.126*	-0.129*	-0.105*
	(0.027)	(0.027)	(0.028)	(0.020)	(0.020)	(0.020)
LAKE_VIRGI~A	0.286*	0.286*	0.303*	0.286*	0.286*	0.303*
	(0.029)	(0.029)	(0.029)	(0.017)	(0.018)	(0.019)
LAKE_WADE	-0.101*	-0.105*	-0.040	-0.099*	-0.103*	-0.037
	(0.026)	(0.026)	(0.029)	(0.019)	(0.019)	(0.023)
LAKE_WALKER	-0.385*	-0.384*	-0.335*	-0.384*	-0.382*	-0.332*
	(0.036)	(0.036)	(0.037)	(0.029)	(0.029)	(0.030)
LAKE_WARREN	-0.070*	-0.069*	-0.029	-0.068*	-0.066*	-0.025
	(0.025)	(0.025)	(0.027)	(0.021)	(0.020)	(0.021)
LAKE_WAUNA~A	0.060	0.063***	0.099**	0.061*	0.063*	0.101*
	(0.038)	(0.038)	(0.039)	(0.020)	(0.021)	(0.024)
LAKE_WELDONA	-0.046	-0.049	0.000	-0.044***	-0.047**	0.003
	(0.031)	(0.031)	(0.032)	(0.024)	(0.024)	(0.026)
LAKE_WESTON	-0.343*	-0.340*	-0.298*	-0.344*	-0.340*	-0.297*
	(0.023)	(0.023)	(0.025)	(0.022)	(0.022)	(0.024)
LAKE_WHIPP~R	0.131*	0.123*	0.134*	0.137*	0.129*	0.141*
	(0.045)	(0.045)	(0.045)	(0.026)	(0.025)	(0.023)
LAKE_WINYAH	0.103*	0.100*	0.151*	0.103*	0.100*	0.151*
	(0.024)	(0.024)	(0.026)	(0.017)	(0.018)	(0.021)
LAWNE_LAKE	-0.290*	-0.296*	-0.273*	-0.291*	-0.297*	-0.274*
	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)	(0.017)

LITTLE_FISH	0.480*	0.476*	0.516*	0.480*	0.477*	0.518*
	(0.017)	(0.017)	(0.019)	(0.020)	(0.020)	(0.021)
LIT_LK_FAIR	-0.136*	-0.136*	-0.105*	-0.136*	-0.136*	-0.105*
	(0.025)	(0.025)	(0.026)	(0.020)	(0.021)	(0.022)
LITTLE_SAND	0.231*	0.235*	0.247*	0.231*	0.236*	0.248*
	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)
LONG_LAKE	-0.348*	-0.346*	-0.322*	-0.350*	-0.348*	-0.323*
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)
MUD_LAKE_C	0.094**	0.106**	0.171*	0.096**	0.107*	0.174*
	(0.047)	(0.046)	(0.048)	(0.039)	(0.040)	(0.042)
PALM_LAKE	0.035*	0.038*	0.103*	0.035*	0.038*	0.105*
	(0.010)	(0.010)	(0.016)	(0.012)	(0.012)	(0.017)
PARK_LAKE_B	0.030	0.024	0.090**	0.032	0.026	0.093*
	(0.034)	(0.034)	(0.036)	(0.029)	(0.029)	(0.032)
POCKET_LAKE	0.282*	0.280*	0.292*	0.283*	0.280*	0.293*
	(0.019)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)
ROCK_LAKE	-0.480*	-0.486*	-0.444*	-0.478*	-0.484*	-0.441*
	(0.026)	(0.025)	(0.027)	(0.025)	(0.025)	(0.026)
SPRING_LK_B	-0.013	-0.018	0.022	-0.01	-0.015	0.025
	(0.024)	(0.024)	(0.025)	(0.021)	(0.021)	(0.023)
SPRING_LK_C	0.180*	0.181*	0.197*	0.183*	0.183*	0.200*
	(0.012)	(0.012)	(0.013)	(0.014)	(0.014)	(0.015)
Constant	42.555**	45.105**	48.569*	42.389*	44.962*	48.499*
	(17.705)	(17.662)	(17.669)	(7.215)	(8.276)	(9.458)
Rho	---	---	---	0.001*	0.001*	0.001*
				(0.000)	(0.000)	(0.000)

† All time variables are dummies = 1 if home sold in indicated year, 0 otherwise
‡ All lake variables are dummies = 1 if home is closest to indicated lake, 0 otherwise.
Note: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Standard errors appear in parentheses.

CHAPTER 2: EXAMINATION OF SPATIAL DEPENDENCE IN THE HEDONIC MODELING OF WATER QUALITY

2.1. Introduction

The treatment of spatially correlated influences has become a popular research topic in recent years. Traditional methods of controlling for location in hedonic property price analysis—such as the use of dummy variables to indicate neighborhood, census tract, or town—have been shown to insufficiently capture the underlying spatial structure (Dubin 1998; Anselin 1999; Case et al. 2004; Anselin and Le Gallo 2006; Anselin and Lozano-Gracia 2008). The broad field of spatial econometrics has been developed to address these issues and provides tools to more effectively capture spatial structure and correlations. This remains, however, a relatively new field that lacks guidance in several areas.

The purpose of this chapter is to both explore some unanswered spatial questions and to further investigate the spatial sensitivity of the regressions in the first chapter. A specific focus of this chapter is differences caused by the spatial weights matrix (SWM). Most approaches to spatial analysis use a SWM to represent the spatial configuration of the data. This matrix plays a key role in the modeling process, yet there is limited theoretical guidance about its proper form (Mueller and Loomis 2008). This chapter investigates the impact of varying the SWM on the two core spatial models and compares the resulting differences. Another area with limited theoretical guidance is the functional form of the hedonic model. The most popular method of functional form testing involves the Box-Cox model; however the use of spatial econometric techniques complicates the Box-Cox test. Recent evidence indicates that the functional form of the regression can affect tests of spatial dependence and spatial modeling (McMillen 2003; Boxall, Chan, and McMillan 2005). This chapter uses a new double-length regression (DLR) test to explore a spatial Box-Cox test, which is used to jointly test issues of spatial dependence and functional form.

Two central spatial specifications are investigated: the spatial error and spatial lag models. Six total variations of the two most popular types of weights matrices are used in conjunction with each of these models. The first type of spatial weights matrix is based on specifying a fixed number of neighbors for each parcel, while the second type defines the neighbor set using spatial and temporal distances between the parcels. These two types of SWM's and their variations are the most popular configurations encountered in the literature. However, the choice of SWM in most spatial papers is not explained. This chapter hypothesizes that the distance-based SWM's are superior to the nearest neighbor specifications. They introduce more information into the SWM and mitigate the influence of outliers.

A forecasting exercise is used to compare the various spatial models. This procedure splits the data and uses out-of-sample forecasting to gauge the fit of the models. Several forms of forecasting error are used to evaluate each model. Case et al. (2006) use a similar procedure to compare several more exotic spatial models.

2.2. Data

The data in this chapter come from the same dataset used in Chapter 1. The analysis begins with 54,712 property sales from Orange County, Florida. As in Chapter 1, these data have been matched with US Census variables, water quality variables, and GIS maps. As the purpose of this chapter is to estimate and then forecast out-of-sample, the data are randomly split into two groups. The first group contains 44,712 property sales and the second group contains the remaining 10,000 property sales. The first group is used to estimate the various spatial models while the second group is used for forecasting.

The proportion of observations in the prediction group and the forecasting group falls between two recent studies that employ a similar approach. Case et al. (2004) use approximately a 90/10 percent split in their sample and Bourassa, Cantoni, and Hoesli (2007) use an 80/20 percent split. The split of 84/26% in the current paper falls in between the two previous papers.

Summary statistics for this new smaller dataset appear in Table 7. The mean waterfront property price in the sample of 45,712 observations is $451,086.90 and the mean non-waterfront property price is $200,383.40, which are very close to the full sample values of $452,646.20 and $199,982.40 (appearing in Table 2). Also, the average heated area for both waterfront and non-waterfront homes only change by approximately one foot and the average home ages are almost identical. Overall, Table 7 shows only minor changes in the average variables from the full dataset.

2.3. Literature Review

The influence of spatial econometrics has spread to a variety of economic fields in recent years. For instance, in 2008 the Nobel prize was won by Paul Krugman, whose contributions to the "New Economic Geography" arose out of issues in regional science such as imperfect competition and increasing returns to scale (Florax and Van der Vlist 2003). Arbia (2006) discusses recent attempts to update the Solow Model and estimate cross-country convergence in growth levels using spatial econometric techniques. Spatial correlation was found to be an important influence on regional growth levels. In the sphere of public policy, Lacombe (2004) analyzes the use of spatial econometric techniques in analyses of public policy, with a specific application to Food Stamp programs.

Anselin (1988) roughly defines spatial dependence as the presence of a functional relationship between occurrences at one particular point in space and similar occurrences at other points. Early work on spatial issues arose from several statisticians in the 1940s and 1950s, with notable contributions by Moran (who later developed a test for spatial dependence), Geary, and Whittle (Florax and Van der Vlist 2003). Serious progress in the area began with the publication of *Spatial Autocorrelation* in 1973 (Cliff and Ord 1973), which covers many of the early approaches to handling spatial statistics and was the first source to suggest the handling of spatial data with maximum likelihood techniques. The first comprehensive approach to the econometric modeling of spatial dependence was Anselin (1988). This book treats several of the main econometric approaches to handling spatial data, with a focus on the use of maximum likelihood techniques to analyze spatial data.

In any analysis where location is an integral component, spatial autocorrelation is likely a problem (Dubin 1998). Real estate data are particularly susceptible to spatial dependence, as location is an essential determinant of home prices. In addition to location, there are several area-specific determinants which can also affect the property prices of nearby homes and can be correlated in a regression. The neighborhood effects are rarely measured properly since proxies are typically used—such as census tracts. Consequently, when conducting a hedonic analysis of property prices these spatial error issues must be addressed. Failing to correct for spatial autocorrelation can result in inefficient OLS estimators and biased variance estimates (Anselin 1999).

2.3.1. Past Use of Spatial Models in Hedonic Property Value Analysis

The influence of location and spatial dependencies on property prices has given spatial econometrics a prominent role in modern hedonic analysis. These spatial dependencies arise from several sources, in particular the practice of property appraisal which uses several nearby "comparable sales" to create an estimate of the property price. This section discusses some of the primary hedonic analyses that have used spatial econometric techniques.

In hedonic analyses of property prices, the first applications of spatial econometrics appeared in Dubin (1988, 1992) and Can (1990, 1992). Dubin (1988) draws on earlier spatial studies in the fields of geography and geology to use kriging to estimate the variance structure of the data, based on the distance between homes. Maximum likelihood techniques are used to estimate a hedonic regression model, and spatial techniques perform significantly better than traditional OLS. Can (1990) introduces a spatial expansion model that allows the structural coefficients to vary with neighborhood quality. Can (1992) examines the influence of both spatial dependence and spatial heterogeneity and investigates four different SWM's, as well as linear and semi-log functional forms. Several specification tests are explored, and models are compared by log-likelihood values. Dubin (1992) examines the incorporation of neighborhood quality into hedonic analyses. Dubin recommends a spatial error model to treat neighborhood quality issues directly in the error term, instead of addressing it with dummy variables.

Using the data from the influential Harrison and Rubinfeld (1978) hedonic analysis of air quality, Pace and Gilley (1997) re-estimate the analysis while correcting for spatial autocorrelation through a simultaneous autoregression. This model uses a spatially weighted average of the errors on nearby properties that reduced estimated errors by 44% lower than the previous OLS regression. Dubin, Pace, and Thibodeau (1999) investigate several spatial error modeling techniques in the context of real estate data. They explain the three types of functions that have been used for this purpose—correlograms, covariograms, and semivariograms—and illustrate their usefulness in a hedonic example. Gawande and Jenkins-Smith (2001) use a spatial hedonic model to estimate the impact of highly publicized shipments of nuclear waste on home prices along the shipment route.

Kim, Phipps, and Anselin (2003) use a spatial model to estimate the marginal value of air quality in Seoul, South Korea. Both spatial lag and spatial error models are estimated and the spatial dependence is an important characteristic of the data. An important contribution of this paper is its explanation of the marginal effects in a spatial lag model. These marginal effects will be different due to the appearance of the spatially lagged dependent variable appearing on the right hand side of the equation. This will be discussed further in section 5 of this chapter. Using the spatial lag model, they find that a small positive change in air pollution of approximately 4% results in a property price increase of $2,333, or 1.43% of the average home price.

Brasington and Hite (2005) analyze the demand for environmental quality in Ohio. Their measure of quality is the distance to nearby environmental hazards and point sources of pollution. Significant spatial effects are detected and a spatial Durbin model is used for estimation. This is one of the few papers in the environmental hazards literature to use spatial econometrics and one of the only papers in the broader hedonic literature to use a spatial Durbin model.

In another analysis of air quality, Anselin and Le Gallo (2006) examine the interpolation of air pollution values to home locations in a hedonic model. A spatial lag model is used to control for spatial autocorrelation. Strong spatial dependence is detected in the data even after controlling for house and neighborhood characteristics, as done in previous hedonic analyses of the same market. This paper devotes significant attention to the assignment of air pollution variables to homes using a kriging process. This approach is used because air pollution readings are typically only taken at specific monitoring sites. Kriging is a minimum mean square error method of spatial prediction that is popular in the geosciences literature. It uses best linear unbiased prediction (BLUP) to assign weights to observations based on distances (Dubin, Pace, and Thibodeau 1999). It is commonly used to interpolate the values of a pollutant over a continuous surface when the only data available are a fixed number of monitoring stations (Anselin and Le Gallo 2006).

One of the most recent spatial hedonic analyses of air quality is Anselin and Lozano-Gracia (2008). Using the data from Anselin and Le Gallo (2006), Anselin and Lozano-Gracia (2008) takes an "errors in variables" approach to deal with the endogeneity of the air pollution variable. There are two potential sources of endogeneity. The first source involves residential sorting, where preferences for pollution affect the selection of home location. The second source of endogeneity arises from errors in the assignment of the pollution value, which may be correlated with other spatial variables and omitted variables. This paper directly deals with the latter type of endogeneity but they stress that, from an empirical perspective, the source of the endogeneity does not matter for the treatment. A spatial two stage least squares approach is used to confront the endogeneity, where the latitude, longitude, and the product of the two are used as instruments for the endogenous air pollution indicators (two are used).

One of the few hedonic papers that evaluates differences in spatial specifications and SWM's is Mueller and Loomis (2008). In a hedonic analysis of the impact of wildfires on home prices, they find evidence that in many studies spatial corrections do not cause substantive differences from the OLS estimates. They further investigate this phenomenon and use three different SWM's: four nearest neighbors, eight nearest neighbors, and inverse distance. The three

SWM's cause minor changes in the magnitudes of the coefficients. There are no qualitative differences in the implicit prices. Nonetheless, they stress that the lack of theoretical guidance for the SWM is a void in the literature that needs to be further analyzed.

While a broad array of spatial techniques have been used in the hedonic literature, the incorporation of spatial techniques into hedonic analyses of water quality and lake proximity has been rather limited. Leggett and Bockstael (2000) utilize a spatial error model to estimate a hedonic model of water quality in the Chesapeake Bay in Maryland. While a series of tests indicated the presence of spatial autocorrelation, they found that correcting for spatial autocorrelation did not seem to produce any consistent direction in the bias in the standard errors in the variables of importance to the study.

Palmquist and Fulcher (2006) use the spatial error model in an analysis of lake proximity. The purpose of the article is to replicate the study by Brown and Polakowski (1977), with corrections for spatial dependence. Palmquist and Fulcher repeat the qualitative results of Brown and Pollakowski and find lake proximity to be positively related with home prices. Also, the implicit price of lake proximity increases with the spatial error model. The purpose of the present analysis is to expand the application of spatial econometrics to a hedonic analysis of water quality beyond a basic spatial error model.

2.3.2. Functional Form and Spatial Dependence Tests

As discussed in the first chapter, two econometric concerns that must be jointly confronted are functional form and spatial dependence. The functional form of the regression is a contentious issue in hedonic regression and is further aggravated by the introduction of spatial dependence. When the hedonic model was theoretically developed in Rosen (1974), there was no prescription for the proper functional form of the regression. The two econometric issues of functional form and spatial dependence are interrelated and must be confronted jointly. McMillen (2003) uses a Monte Carlo experiment to show that spatial tests may improperly detect spatial autocorrelation if the functional form is incorrect.

A popular method of testing for functional form is the Box-Cox model (Haab and McConnell 2002). This model uses a transformation of the dependent and independent variables that can be used to test between functional forms. The most common form of the Box-Cox model appears below:

$$y^{(r)} = X^{(r)}\beta + Z\gamma + \varepsilon \tag{2.1}$$

where

$$y^{(r)} = \begin{cases} \dfrac{y^r - 1}{r}, & \text{if } r \neq 0 \\ \log(y), & \text{if } r = 0 \end{cases} \tag{2.2}$$

In equation (2.1) X is a matrix of continuous independent variables and Z is a matrix of non-continuous independent variables that are not subject to the Box-Cox transformation. When r = 1 the equation becomes linear, when r = 0 the transformation becomes the double log specification. This flexible functional form has not commonly been used in spatial models since it significantly complicates the implementation of the model and the interpretation of the results (Kim et al. 2003). However, it is sometimes used as a preliminary check of the functional form. For instance, Patton and McErlean (2003) use a Box-Cox test before spatial analysis to explore the most appropriate functional form.

A test commonly used for spatial dependence is the Lagrange Multiplier (LM) test (Mueller and Loomis 2008), but the normal version of this test does not account for alternative functional forms. Baltagi and Li (2001) and Baltagi and Li (2004) developed LM tests to jointly test functional form restrictions and both spatial error and spatial lag dependence. These tests allow an exploration of spatial dependence under alternative functional forms and are based on the Box-Cox model. Boxall, Chan, and McMillan (2005) use these LM tests to support the results of a Box-Cox test in a hedonic regression of the effects of oil and natural gas facilities on home prices. The main problem with the LM test, however, is that it is extremely difficult to estimate with large datasets. Le and Li (2008) and Le (2009) develop Double-Length Regression (DLR) tests that mimic the Baltagi and Li LM tests, but unlike the latter they do not require the computation of the Hessian or the second derivatives of the log likelihood function. The DLR tests can therefore be used on larger datasets and are employed in the present paper to jointly test functional form and spatial dependence.

The DLR tests first appeared in Davidson and MacKinnon (1985) for non-spatial tests of Box-Cox functional forms. The application of the DLR test to spatial Box-Cox models is the subject of two recent papers. The first paper is Le and Li (2008), which develops the DLR test for a spatial error model. The test begins with the Box-Cox model in equations (2.1) and (2.2), but the error is assumed to be of the form:

$$\varepsilon = \lambda W \varepsilon + u \qquad (2.3)$$

where λ is a spatial autoregressive coefficient on the errors, *W* is a spatial weights matrix and *u* is an *nx1* vector with $u \sim N(0,\sigma^2 I)$—the basic spatial error model. The foundation of the DLR model is composed of two equations from the log-likelihood function:

$$\log L = -\frac{n}{2}\log 2\pi + \sum_{i=1}^{n} k_i(y,\theta) + \frac{1}{2}\sum_{i=1}^{n} f_i^2(y,\theta) \qquad (2.4)$$

where

$$f(y,\theta) \equiv \frac{1}{\sigma}(I-\lambda W)(y^{(r)} - X^{(r)}\beta - Z\gamma) \qquad (2.5)$$

and

$$k_i(y,\theta) \equiv -\log\sigma + \log(1-\lambda\omega_i) + (r-1)\log y_i \tag{2.6}$$

Which is an element of the column vector $\boldsymbol{k}(y,\theta)$, and in a similar fashion an individual element of $\boldsymbol{f}(y,\theta)$ is $f_i(y\ \theta)$. To define (2.6), Le and Li use a result from Anselin (1988):

$$\log|I-\lambda W| = \sum_{i=1}^{n}\log(1-\lambda\omega_i) \tag{2.7}$$

where ω_i are the eigenvalues of the SWM W. The parameters of the equation are contained in the vector $\theta = (\sigma, \beta', \gamma', \lambda', r)$.

Now, define:

$$F_{ij} \equiv \frac{\partial f_i(y,\theta)}{\partial\theta_j} \tag{2.8}$$

$$K_{ij}(y,\theta) \equiv \frac{\partial k_i(y,\theta)}{\partial\theta_j} \tag{2.9}$$

Given these definitions, the DLR regression is performed by estimating the following model:

$$\begin{bmatrix} f(y,\theta) \\ \iota_n \end{bmatrix} = \begin{bmatrix} -F(y,\theta) \\ K(y,\theta) \end{bmatrix}\delta + e \tag{2.10}$$

where $\boldsymbol{F}(y,\theta)$ and $\boldsymbol{K}(y,\theta)$ are n x (K+L+3) (y is n x 1, X is n x K, and z is n x L), δ is (K+L+3) x 1, ι_n is an n x 1 vector of ones, and e is the residual. Following Davidson and MacKinnon (1985), to compute the DLR test for a null hypothesis, the restricted maximum likelihood estimates of θ, F, and K are computed and the regression in (2.10) is performed. The test statistic is the explained sum of squares in this regression, which is distributed as χ^2_R, where the degrees of freedom are the number of restrictions under the null hypothesis.

The spatial lag version of the DLR test defined in Le (2009) is a slight variation of the above procedure. The starting point is a spatial lag Box-Cox equation:

$$y^{(r)} = \rho W y^{(r)} + X^{(r)}\beta + Z\gamma + \varepsilon \tag{2.11}$$

where ρ is the familiar spatial lag coefficient and X and Z are defined as in equation (2.1). This equation can be rewritten as;

$$\frac{1}{\sigma}(I-\rho W)y^{(r)}=\frac{1}{\sigma}(X^{(r)}\beta+Z\gamma+\varepsilon) \tag{2.12}$$

Where both sides of the equation have been divided by the standard deviation so that the error terms are normally distributed, or (ε/σ)~N(0,I).

The f and k functions are now defined as:

$$f(y,\theta)=\frac{1}{\sigma}[(I-\rho W)y^{(r)}-X^{(r)}\beta-Z\gamma] \tag{2.13}$$

$$k(y,\theta)=-T\log(\sigma)+\sum_{i=1}^{n}\log(1-\rho w_i)+(r-1)\sum_{i=1}^{n}\log(y_i) \tag{2.14}$$

Given these new equations, $\boldsymbol{F}(y,\theta)$ and $\boldsymbol{K}(y,\theta)$ are defined as in (2.8) and (2.9). The DLR test statistic is computed the same way as with the spatial error model, using the regression given by (2.10) with the new ***F*** and ***K*** functions.

2.3.3. Spatial Weights Matrix

Anselin (1999) shows that spatial autocorrelation can be formally expressed by the moment condition: $\mathrm{Cov}[y_i,y_j]=\mathrm{E}[y_iy_j]-\mathrm{E}[y_i]\cdot\mathrm{E}[y_j]\neq 0$, for $i\neq j$, where i, j refer to individual observations (locations). In most settings, there is inadequate information present to estimate this covariance matrix directly from the data. As the number of observations (N) grows, the covariance matrix grows by N^2. The typical approach for dealing with this information shortfall is to impose a structure on the functional relationships between observations in space. This is done through the construction of an N x N SWM. The imposition of a spatial structure on the underlying spatial system allows empirical estimation and testing of other relationships in the data.

The original methods espoused for constructing a SWM were advanced by Moran (1948) and Geary (1954) and were based on the notion of contiguity. SWM's were constructed by entering a 1 in the w_{ij} position of the SWM if observation i and observation j are spatially contiguous. This approach to the classification of space is useful when analyzing geographical units such as counties, states, or voting districts. However, in applications where direct contiguity is not common and individual observations are much smaller, such as real estate sales data where direct neighbors will not frequently sell their homes within similar time periods, this type of SWM fails to adequately capture space. More appropriate definitions of the SWM are based on defining a set of neighbors that influence each observation. Anselin (1988) gives several possibilities for this, and formally defines the concept as follows. Assume there exists a system S of N spatial observations and a variable x that is observed for each of these observations. The set of neighbors for observation i is J such that $P[x_i \mid x] = P[x_i \mid x_J]$, where P[.] indicates probability. A less strict definition is $(j \mid P[x_i] \neq P[x_i \mid x_J])$. A more strict definition that explicitly includes concerns with space and distance, and is recommended by Anselin, is $(j \mid P[x_i] \neq P[x_i \mid x_J]$ and $d_{ij} < \epsilon_{ij})$, where ϵ_{ij} is a cutoff point for each spatial unit (which almost always is the same for each unit) and d_{ij} is the distance between observations i and j.

A primary objective of the construction of the SWM is to define the spatial configuration of the data. As discussed in Chapter 1, there are many unobserved spatially-correlated characteristics that may affect the prices of nearby homes. Individual neighborhoods may have unobservable codes or covenants. Also, certain neighborhoods may have local parks, proximity to shopping areas, playgrounds, and other amenities or disamenities that may affect local housing prices in a similar fashion. Additionally, the size of an individual home in relation to other homes in a neighborhood may affect the home price. Traditional ways to control for these neighborhood effects involve using indicator variables or latitude and longitude coordinates of homes. However, unlike time variables, where organization is usually intuitive (such as annual or monthly dummies), spatial configurations are less clear. The optimal spatial configuration of neighborhood indicators can also vary significantly over a landscape. Dubin (1998) finds that the use of neighborhood indicators is typically unable to adequately capture the spatial configuration of homes. Pace et al. (1998) compare a hedonic model with 28 variables that uses spatial indicators to a simpler 12 variable spatial model that accounts for space and time in a SWM. The spatial model significantly outperforms the OLS model with indicator variables.

The most common forms of SWM's are based on either distance or *n* nearest neighbors. Both of these forms fit into the above neighbor definition. The former includes in the neighbor set all homes within a given distance and uses one of several options for the individual elements in the SWM. In one variation, if observation j is within the distance boundary of observation i, $w_{ij} = 1$, otherwise $w_{ij} = 0$. This specification, however, more closely resembles the nearest neighbor specification described below. Alternatively, one could use a function of the actual distance in w_{ij}, such as the inverse distance between i and j. For instance, if inverse distance is used, $w_{ij} = (1/d_{ij})$ if $d_{ij} > B$, where B is a maximum distance boundary, and $w_{ij} = 0$ otherwise. Inverse distance squared has also been used, where $w_{ij} = (1/d_{ij}^2)$ (Can 1992). Functions of the inverse distance are popular because homes that are farther away from the subject property receive less weight in the SWM.

In some of the original versions of weights matrices, Cliff and Ord (1973) recommend the use of distance in the SWM. They propose several versions of the SWM that use a combination of distance and the size of the regional unit (since the applications were mostly in regional science, this was appropriate). Weights matrices based on distance have become popular in the literature, as they allow a relatively flexible definition of the neighborhood. Also, they admit more information into the SWM with respect to the relationship between individual observations.

While distance has been incorporated into spatial analyses, the influence of time has only recently been included in the construction of the SWM. The incorporation of time was briefly discussed in Chapter 1, where it was emphasized that it is prudent to recognize the dynamic variation in spatial attributes if the data span several years. During the time period analyzed in this paper, the housing market was experiencing a significant expansion. The construction of new homes, apartments, and condos accelerated in Orlando and the neighboring areas. Consequently, the composition of many neighborhoods within the area changed, so that omitted neighborhood variables in one year may not correspond to those in the next year.

Several papers have attempted to control for these issues in recent years. Pace et al. (1998) propose a spatiotemporal autoregressive (STAR) model, and use a time weights matrix as well as a spatial weights matrix. A function of these two matrices is used as the overall weights matrix in the spatial regressions. Their weights matrix is fundamentally different than most used in the literature because it is constructed to be a lower diagonal matrix organized by sales dates from oldest to newest. This aids in later filtering and computation using OLS instead of the typical maximum-likelihood approach. Case et al. (2004), in a study previously mentioned in this paper, use multiple techniques in an attempt to control for both spatial and temporal influences. Pace and Gilley (1998) also use both a SWM (S) and a time weights matrix (T), and multiply (I – S) by (I – T) to obtain the full weights matrix for the analysis. All of these studies agree that time and distance should both be recognized in the analysis. This can be done through the particular configuration of the weights matrix.

2.3.4. Forecasting

There are several potential methods available to examine the strengths of the various spatial approaches. This chapter uses a forecasting method which had seen significant use lately in the hedonic and real estate literature (Dubin 1988; LeSage 1999; Pace et al. 2000; Case et al. 2004; Bourassa, Cantoni, and Hoesli 2007). The forecasting exercise allows a comparison of the various SWM's by determining which model can best predict property prices. To do this, the data is split into two parts. The larger of these sections is used for estimation. The coefficients from the larger data set are then used to predict the property prices from the smaller dataset. The forecasted values can then be compared to the actual values. Based on these results, several comparisons can be made.

Pace et al. (2000) use out-of-sample forecasting to evaluate the contribution of temporal restrictions on the SWM. They also examine the effect of varying the number of neighbors in the SWM and find that this has little effect on the model's goodness of fit. However, they only examined the range of 160-200 nearest neighbors. A wider range of neighbors may have resulted in a larger impact of the SWM definition on the results. Also, temporal restrictions were found to have a positive impact on the goodness of fit.

Much of the inspiration for the current chapter arises from Case et al. (2004), who compare several spatial approaches with the same dataset. An open call was made to spatial researchers to participate in the study and a large database of home sales and characteristics from Fairfax, Virginia was made available to participants. There was also an out-of-sample group of observations withheld from participants, and stayed in the possession of one individual responsible for calculating out-of-sample statistics. Once the out-of-sample forecasting was performed, and errors calculated, several statistics were calculated to compare the various models. The statistics used for comparison were mean and median prediction error, standard deviation of the prediction error, root mean and median squared prediction error, and mean and median absolute value percent error. Another interesting characteristic of the data was that participants were provided with a significant deal of information on nearest neighbors. The 15 nearest neighbors (for both in and out-of-sample observations) were identified for each observation, as well as the corresponding distances between observations.

There were three participants who answered the call for participation: Bradford Case, Robin Dubin, and John Clapp. The models compared were three unorthodox spatial econometric models that have been espoused by each author but have not seen wide application in the literature. Case used a variation of his process of dividing a housing market into its various districts. Clapp used a kriging version of his local regression model. Dubin used two versions of her kriging model. An OLS regression was also estimated for purposes of comparison. Of the three models, Clapp’s local regression model performed poorest. The superior performance of the other two models was attributed to their incorporation of nearest neighbors. A second round of forecasting was allowed, where authors could attempt to improve their models. The largest improvement was a result of incorporating more information about nearest neighbors, similar to expanding the number of neighbors in a SWM. The maximum number of nearest neighbors given was 15, and results indicated that moving beyond this threshold would likely further improve results.

Bourassa, Cantoni, and Hoesli (2007) use forecasting to evaluate geostatistical and lattice models in a hedonic analysis of New Zealand homes. The focus of the paper is on housing submarkets and issues more closely related to spatial heterogeneity than spatial dependence. They find that the inclusion of accurate submarket variables in an OLS regression result in larger predictive gains than the use of geostatistical or lattice methods.

This dissertation chapter employs the basic approach of Case et al. (2004), although with more of a focus on the SWM than on more exotic models of spatial dependence. To evaluate the various models, this chapter employs the error statistics from Case et al. (2004), including mean and median prediction error, standard deviation of the prediction error, root mean and median squared prediction error, and mean and median absolute value percent error.

2.4. Methods

A variety of models are examined in this chapter, using a large dataset of approximately 55,000 observations. Two key econometric issues are encountered when using large real estate databases. The first is the functional form of the regression equation. The second issue is the handling of spatial dependence. These issues must addressed individually and in tandem, as recent evidence suggests that the functional form of the regression can alter the results of spatial tests (McMillen 2003). Two main spatial models and six SWM's are evaluated in this chapter. To assess all of these models an out-of-sample forecasting exercise is performed. The present section explains the construction of the SWM's, the joint tests of functional form and spatial dependence, and the spatial error and spatial lag models.

In Chapter 1 the full version of the hedonic model—with three interaction terms between SDM and the waterfront, distance, and area variables—was chosen as the preferred model. The interaction terms allow the implicit price of water quality to vary over the landscape, instead of restricting it to be the same over lakes, types of property, and location. In light of these observations, this chapter will use the full version of the hedonic model from Chapter 1, again appearing below:

$$\begin{aligned} P = {} & \beta_0 + \beta_1 WF + \beta_2 SDM + \beta_3 WF * SDM + \beta_4 Distance + \beta_5 Distance * SDM \\ & + \beta_6 Area * SDM + \gamma'\mathbf{x} + \varphi'\mathbf{y} + \psi'\mathbf{l} + \delta'\mathbf{t} + \varepsilon \end{aligned} \tag{2.15}$$

The variables not directly depicted in this equation are the same as in Chapter 1, and again include structural and census tract variables (**x**), location variables (**y**), and lake (**l**) and time (**t**) fixed effects.

2.4.1. Functional Form and Spatial Testing

Several restrictions can be used in both of the DLR tests to test for spatial dependence. These tests cannot, however, be used to select between the spatial error and spatial lag model. Forecasting is used later in the chapter to select between the two models. The first hypotheses to consider are joint hypotheses of spatial dependence and functional form. These tests are useful for comparing the linear and double-log models. The two joint hypotheses to be tested are H_{00}: $r = 0$, $\rho = 0$ (or $\lambda = 0$ for the spatial error model), or a double-log model with no spatial dependence, and H_{10}: $r = 1$, $\rho = 1$ ($\lambda = 1$), or a linear model with no spatial dependence. The alternative hypothesis for both these tests would be a general Box-Cox model with parameter r and the presence of spatial dependence.

Since a rejection of the joint hypotheses could reflect either improper functional form of the existence of spatial dependence, "one-direction" DLR tests are used to confirm the existence of spatial dependence. The first of these tests is H_0: $\rho = 0$ ($\lambda = 0$) assuming $r = 0$, or a test for spatial dependence assuming the proper model is double log. The second one-direction test is H_1: $\rho = 0$ ($\lambda = 0$) assuming $r = 1$, or a test for spatial dependence assuming the proper model is linear. Spatial lag and spatial error versions of both joint and one-direction tests will be performed.

There is one additional constraint on the implementation of these tests. As shown in equation (2.7), the n eigenvalues of the n x n SWM need to be computed. The size of the data is a barrier to this calculation, so the data are split into smaller subsets. The data are first split by individual municipalities within Orange County, and for the two largest municipalities the data are randomly split into smaller subsets. Municipalities are used because much of the spatial dependence should be contained within these areas. Within a given municipality residents pay the same taxes and have access to the same municipal facilities.

2.4.2. Spatial Weights Matrix

The distance and time-based SWM employed in the first chapter uses a distance boundary of 200 meters and a time boundary of six months previous back and three months forward. With a 200 meter boundary, all homes within the circle centered on each observation with a diameter of 400 meters, or a quarter mile, are potential neighbors. The distance boundary was believed to adequately capture neighborhood effects in the dense urban area under study. The time boundary was chosen with the practices of property appraisers in mind, who typically use homes sold within six months as "comparables." For a discussion of these issues in the context of spatial analysis, see Pace et al. (1998), Pace and Gilley (1998), and Pace et al. (2000). The three month forward boundary was chosen due to lags between the actual close of a sale and its withdrawal from the market and subsequent reporting. Can and Megbolugbe (1997) use a similar SWM and discuss some of the positive aspects of this formulation.

This chapter conducts a more thorough examination of the SWM by varying the boundaries of the neighborhood definition. Six different SWM's are used to explore the effect of the SWM definition on the estimated coefficients and implicit prices. These SWM's are defined and described in Table 8, where the names appear in the first column. The names are simply based on the type of boundaries. For instance, the first SWM is a basic inverse distance matrix with a distance boundary of 200 meters so its name is "200". The second SWM, which is the default model from the first chapter, has time boundaries of 6 months back and 2 months forward, so its name is "20063". The distance and temporal boundaries of each SWM are defined in columns 2-4 of Table 8. Columns 5-7 give summary statistics about the number of neighbors for each observation. Since SWM 200 has no time boundaries, its average number of neighbors is approximately 40 more than the 20063 SWM. These columns illustrate the differences in the SWM caused by the different SWM definitions. Note that the NN15 and NN20 SWM's have the same number of neighbors over all observations. Finally, the last column of Table 8 contains the full number of non-zero elements in the SWM. This column shows that the 200 SWM clearly has the most total non-zero elements.

The other SWM's used in this chapter represent further permutations of the 20063 SWM. The third SWM (200123) extends the time boundary of the default model back from 6 months to 12 months. Increasing the time boundary to 12 months expands the average number of neighbors in the SWM to 9 and the maximum to 119. The fourth model keeps the 6 month time boundary but extends the distance boundary to 400 meters. Doubling the distance boundary to 400 meters more than doubles the mean number of neighbors to 14, with a maximum neighbor set equal to 137. The last two SWM's are based on the nearest neighbor definition; one uses 15 nearest neighbors and the other uses 20 nearest neighbors.

In order to estimate the distance and time based SWM, an iterative routine was programmed for use in Matlab. The code draws on a previous program written by Shawn Bucholtz at the US Department of Agriculture. Spatial distances are measured between each observation using the latitude and longitude coordinates. The sale dates of the homes are used to determine temporal distances between property sales. The iterative procedure computes both the spatial and temporal distances between each observation and imposes the boundary conditions to construct the SWM.

The other form of SWM—the *n* nearest neighbor approach—contains a 1 in the SWM at entry w_{ij} if *j* is one of the *n* closest neighbors to *i*. The particular density of homes can vary substantially over a particular setting in both space and time. Pace et al. (1998) point out that a particular distance boundary near a city center may represent significantly different population density than the same boundary applied farther away. They also draw attention to the fact that real estate appraisers, who specialize in predicting short-term fluctuations in home prices, use a fixed number of neighbors in space, although the distance between neighbors may vary with local market conditions.

The nearest neighbor based SWM is the most common form of SWM used in the literature, yet there is significant variation in the construction and composition of these types of matrices. Kim, Phipps, and Anselin (2003) have data on 78 individual districts within Seoul, South Korea that properties reside in. All homes within the same district are considered neighbors, as well as all homes within neighboring districts that have a centroid less than 4 km away. Both Pace et al. (1998) and Pace and Gilley (1998) use weights matrices based on 15 nearest neighbors. Anselin and Lozano-Gracia (2008) use two nearest neighbor weights matrices, based on 6 and 12 nearest neighbors.

The creation of the nearest neighbor weights matrix in the current paper is accomplished through a program in the spatial toolbox found at www.spatial-econometrics.com (LeSage 1999), and based on a Delaunay triangularization. This procedure uses triangles created by forming line segments between points in a plane. The key idea is to get the smallest number of triangles made by these connecting segments, without crossing segments or points within triangles. Once these triangles have been created, the program uses them to define neighbors based on the triangles and the neighbors of neighbors. For this paper, the highest order of four is selected, so that neighbors of neighbors of neighbors of neighbors are analyzed to determine the set of n nearest neighbors. When the program is used with the default second order, only slight differences emerge. Brasington (2006) uses a nearest neighbor SWM based on continuity through the same Delaunay triangularization used here. The nearest neighbor matrix for the current paper uses 15 nearest neighbors, as this has been determined to adequately capture the spatial configuration of data in several papers, including Pace et al. (1998), Pace, Sirmans, and Slawson (2002), and Case et al. (2004).

Finally, two additional characteristics of each SWM include row standardization and zero diagonal entries. Row standardization requires that each row sum to one, formally: $\sum_{j=1}^{n} w_{ij} = 1$. This transformation eases interpretation and aids in estimation, and is recommended by most sources in the literature (Anselin 1988; Can 1992; Anselin 1999). The diagonal elements of the SWM are constrained to equal zero, or $w_{ii} = 0$, so that individual observations are not their own neighbors, or that individual observations do not predict themselves (Anselin and Lozano-Gracia 2008). This chapter will use the six types of SWM's for each model analyzed. There are several pros and cons for each type of weights matrix, and a comparison of them over several models should provide significant insight into their respective merits.

2.4.3. Spatial Lag Model

This section explains the econometric estimation of the spatial lag model through maximum likelihood techniques. This model uses a spatially lagged dependent variable on the right hand side of the regression model, *Wy*. This model is also called the mixed regressive-spatial autoregressive model, due to its lagged dependent variable. The formal expression of the spatial lag model is:

$$y = \rho Wy + X\beta + \varepsilon \quad (2.16)$$

where y is the dependent variable, W is the SWM, X is a matrix of independent variables, β is a vector of estimated coefficients, ε is the error term, and ρ is a spatial autoregressive coefficient, similar to an autoregressive coefficient used in time series analysis. An alternative expression of the spatial lag model is as follows.

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1}\varepsilon \tag{2.17}$$

The Wy term captures the influence of the sales prices of other homes in the area. Due to the use of "comparable sales" in the property appraisal practice, which use nearby home sales to construct an estimate of a home price, property prices may be spatially correlated. Also, as discussed earlier, homes in the same neighborhood will have similar attributes and be located near the same local amenities or disamenities (many of which, such as parks, trails, "rough areas," local shopping centers, and good landscaping, are difficult to quantitatively represent). If the SWM used is based on distance or inverse distance and is row standardized, Wy will represent a spatially-weighted (by distance) average of nearby home prices. If the SWM is based on the nearest neighbors criterion and is row standardized, the Wy term represents an average of neighboring home prices. While it would seem that the estimation would parallel the approach of time series—as the above equation resembles the first order autocorrelation equation from time series—this is not the case. In time series the correlation is linear in nature, either forward or backward in time. Spatial correlation operates in two dimensions and requires different tools for estimation.

Through the reduced form of the spatial lag model, a final interpretation, or representation, is explained in Anselin and Lozano-Gracia (2008). Under standard regularity conditions, it is possible to represent the inverse $(I-\rho W)^{-1}$ using a power expansion as follows:

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \ldots \tag{2.18}$$

when this expression is inserted into the reduced form of the spatial lag model, it is easily seen that the property price is a function of both its own characteristics (through X) and the characteristics of neighboring properties (through, for instance, WX and W^2X).

Anselin (1988) explains some of the differences between spatial autocorrelation and temporal autocorrelation. A well known result in econometrics is that the OLS model will remain consistent in the face of a lagged dependent variable, as in the case of first order autocorrelation, as long as serial correlation is not present (Wooldridge 2003). Even though the estimator will be biased it can still be used for inference because it is consistent. The same situation does not hold with a spatially lagged dependent variable. For example, with a simple first order autoregressive model, $y=\rho Wy+\varepsilon$, there are two conditions for consistency. The first is: $plim\ N^{-1}(Wy'Wy) = Q$, a finite and non-singular matrix, and the second is $plim\ N^{-1}(Wy'\varepsilon) = 0$. The first of these consistency conditions is feasibly satisfied with several simple assumptions. The second condition will not hold, which is easily seen by solving for *y* and inserting the solution. This yields $plimN^{-1}(Wy'\varepsilon)=plimN^{-1}\varepsilon'W(I-\rho W)^{-1}\varepsilon$, which will not equal zero. The results of this simple demonstration translate directly into the spatial lag model which includes *X* variables alongside the *Wy* term on the right hand side of the regression equation. Consequently, if OLS is used to estimate this model, the parameters will be both biased and inconsistent.

To confront these OLS estimation issues, maximum likelihood techniques are used to estimate regression parameters in the face of spatial dependence. These techniques are sufficiently outlined in Anselin (1988) and were first derived in Ord and Cliff (1973). The log-likelihood for the spatial lag model is:

$$\ln L = -(N/2)\ln(2\pi) - (N/2)\ln\sigma^2 + \ln|I-\rho W| - (1/2\sigma^2)(y-\rho Wy - X\beta)'(y-\rho Wy - X\beta) \tag{2.19}$$

In a traditional regression model, the joint likelihood simply equals the sum of the individual log likelihoods. This is not the case in a spatial lag model estimated by maximum likelihood. In this instance, the spatial dependence introduces a Jacobian term that is the determinant of a full *N* x *N* matrix (Anselin 1999).

The first order condition of this model, from Anselin (1999), is:

$$\hat{\beta}_{ML} = (X'X)^{-1}X'(y - \rho Wy) \tag{2.20}$$

Anselin points out that this first order condition can be used to obtain the maximum likelihood estimates using a five step procedure. This procedure is used in the econometric programming of the spatial lag model used in this paper, from the econometrics toolkit (LeSage 1999). The first order condition above is equivalent to:

$$\hat{\beta}_{ML} = (X'X)^{-1}X'y - \rho(X'X)^{-1}XWy = \hat{\beta}_O - \rho\hat{\beta}_L \tag{2.21}$$

With this equation in mind, the first step in the process is to use an OLS regression of *y* on *X* to obtain:

$$\hat{\beta}_O = (X'X)^{-1}X'y \tag{2.22}$$

The second step is to perform another OLS regression of *Wy* on *X* to obtain the following equation:

$$\hat{\beta}_L = (X'X)^{-1}X'Wy \tag{2.23}$$

The third step in the process is to calculate residuals. Equations (2.22) and (2.23) will yield residual vectors e_O and e_L equal to:

$$e_O = y - X\hat{\beta}_O \text{ and} \tag{2.24}$$

$$e_L = Wy - X\hat{\beta}_L. \tag{2.25}$$

From additional first order conditions appearing in chapter 6 of Anselin (1988), these two residual vectors will later be used to compute an estimate for the error variance:

$$\hat{\sigma}^2 = (1/N)(e_O - \rho e_L)'(e_O - \rho e_L) \quad (2.26)$$

The fourth step in the procedure is where the actual maximum likelihood process takes place. A concentrated log-likelihood of the following form is employed:

$$L_C = C - (N/2)\ln[(1/N)(e_O - \rho e_L)'(e_O - \rho e_L) + \ln|I - \rho W| \quad (2.27)$$

where the C term appearing in the equation contains the familiar constant parts of the log-likelihood equation. Given e_O and e_L from the previous step, the maximization of this equation involves the maximization of a function of only one parameter, ρ, which is the fourth step in the process.

The fifth and final step in the procedure is to use the estimate of ρ to compute the following two estimates:

$$\hat{\beta}_{ML} = \hat{\beta}_O - \rho\hat{\beta}_L \quad (2.28)$$

$$\hat{\sigma}^2 = (1/N)(e_O - \rho e_L)'(e_O - \rho e_L) \quad (2.29)$$

In order to obtain estimates of variance, the maximum likelihood estimates of ρ and σ^2 are used to compute the numerical hessian matrix. An internal Matlab function is used to compute the hessian, and the negative of the inverse of this matrix is used to obtain asymptotic estimates of the variance-covariance matrix.

2.4.4. Spatial Error Model

The spatial error model is the second of the two most common spatial models. Unlike the spatial lag model, if the OLS model is estimated in the face of spatial dependence in the error term, the parameters will be unbiased but inefficient. The inefficiency will result from the spatially correlated configuration of the variance-covariance matrix. The spatial error model is a variation of a regression that uses a spatially weighted error term, and the basic structure appears as follows:

$$y = X\beta + \varepsilon \tag{2.30}$$

where $\varepsilon = \lambda W\varepsilon + u$ and $u \sim N(0, \sigma^2 I)$.

This functional form has been particularly popular in the hedonic literature. In empirical studies of spatial autocorrelation, it is the most common model employed (Arbia 2006). In fact, it is the only specification that has been used thus far in the water quality literature (Leggett and Bockstael 2000). Unlike the spatial lag model, the calculation of the implicit prices in a spatial error model remains the same as in an OLS model.

The full log likelihood for this model is:

$$\ln L = -(N/2)\ln\pi - (N/2)\ln\sigma^2 + \ln|I - \lambda W| - (1/2\sigma^2)(y - X\beta)'(I - \lambda W)'(I - \lambda W)(y - X\beta) \tag{2.31}$$

In order to compute estimates of the spatial error model, an iterative approach, similar to that of the spatial lag model, is used. This iterative approach originally appeared in chapter 12 of Anselin (1988) and is used for the programming of the spatial error code in the present analysis (described in LeSage (1999)). This procedure again takes advantage of some of the first order conditions of the spatial maximum likelihood model (presented in Chapter 6 of Anselin (1988)) to use a concentrated form of the log likelihood:

$$L_c = C - (N/2)\ln[(1/N)(y - X\beta)'(I - \lambda W)'(I - \lambda W)(y - X\beta)] + \ln|I - \lambda W| \tag{2.32}$$

The approach for the spatial error model is slightly more complicated than the approach for the spatial lag model. While the maximization of the concentrated log likelihood for the spatial lag model involved one parameter, ρ, with the spatial error model the residual vector is also indirectly a function of λ. Consequently, a simple optimization of L_c with respect to λ is not sufficient.

The full procedure, from Anselin (1988, p. 183), is as follows:

1. Compute OLS estimates of a regression of y on X to obtain $\hat{\beta}_{OLS}$
2. From the OLS estimates, obtain the residuals e_{OLS}.
3. Given the residual vector, find an initial value of λ that maximizes L_c.
4. Using this initial value of λ, carry out estimated generalized least squares (EGLS) to obtain $\hat{\beta}_{EGLS}$ .The EGLS estimate of β is obtained through:

   $$\hat{\beta}_{EGLS} = [X'(I-\lambda W)'(I-\lambda W)X]^{-1} X'(I-\lambda W)'(I-\lambda W)y \tag{2.33}$$

   where $\sigma^2 = (1/N)(y - X\hat{\beta}_{EGLS})'(I-\lambda W)'(I-\lambda W)(y - X\hat{\beta}_{EGLS})$ (2.34)

5. Compute the residuals from the EGLS procedure:

   $$e_{EGLS} = (y - X\hat{\beta}_{EGLS}) \tag{2.35}$$

6. If a previously defined convergence criteria is satisfied, continue on to step 7. Otherwise, return to step 3.
7. Given the estimates of e_{EGLS} and λ, compute the final estimate of:

   $$\sigma^2 = (1/N)(y - X\hat{\beta}_{EGLS})'(I-\lambda W)'(I-\lambda W)(y - X\hat{\beta}_{EGLS}) \tag{2.36}$$

This process will produce estimates for the spatial error model and can be used with a relatively large number of observations.

2.5. Results

2.5.1. Functional Form and Spatial Dependence

The first step in the determination of the functional form is a Box-Cox test. When this transformation is applied to the property sales in the current study, a statistically significant estimate of *r* of -0.145 is obtained (results appear in Appendix). This estimate is much closer to the double-log than the linear specification. Results of the regression appear in the appendix to the dissertation. It is important to interpret the coefficients of the regression with caution, however. Since the focus of the Box-Cox model is on the fit of the data, the coefficients are not typically the best linear unbiased estimators (Zabel and Kiel 2000).

In order to fully test for spatial dependence, issues of functional form and spatial dependence must be considered jointly. Preliminary tests of spatial dependence use the DLR one directional test of spatial dependence. This is a test for either spatial lag or spatial error dependence while assuming a particular functional form. Four of these tests are performed: two tests of spatial lag dependence (one for the linear functional form and one for the double-log functional form) and two tests of spatial error dependence for both functional forms. Two different SWM's are used for these tests, one based on distance and time (with a distance boundary of 200 meters and a time window of 6 months back, 3 months forward) and a 15 nearest neighbor specification. The results of the 20063 one directional tests appear in Table 9 and the results of the NN15 tests appear in Table 10. The results of all tests are unanimous: the null hypothesis of no spatial dependence (in either spatial lag or spatial error form) is rejected at the 99% significance level.

The one direction tests therefore all agree on the existence of spatial dependence. An interesting characteristic of the two tables is that varying the weights matrix does not affect the results of these tests. There are only small changes in the magnitudes of the test results between Table 9 and Table 10. Similar results are discussed in Anselin (1988). Additionally, the municipalities with the largest number of properties (the three Orlando samples and four Unincorporated samples) reject spatial dependence at the highest levels.

To further analyze the existence of spatial dependence while controlling for functional form, several joint DLR tests are performed. These joint tests use the null hypotheses H_{00}: $r = 0$, $\rho = 0$ ($\lambda = 0$ for spatial error dependence), and H_{10}: $r = 1$, $\rho = 0$ ($\lambda = 0$). Both the spatial lag (Le 2009) and spatial error (Le and Li 2008) versions of these tests are performed, and two types of SWM are again used. The results of these tests appear in Table 11 and Table 12. The chi-squared critical value for two degrees of freedom (equal to the number of restrictions) at the 0.001 level is 13.82. All of the DLR tests appearing in these tables are again rejected, indicating that there is an improper specification of functional form or spatial dependence (or both) for each scenario in the table. Given that the one-directional tests found the presence of spatial dependence, the important part of these tables is that for both the spatial lag and spatial error versions the H_{10} test is rejected at a much higher level than the H_{00}. In fact, the difference is approximately 10 times greater in almost all cases. This indicates that the linear model is always rejected at a much higher rate, holding spatial dependence constant.

2.5.2. Spatial Lag and Spatial Error Models

The various spatial models are estimated in Matlab with the full 54,712 observations.[18] Selected results of the spatial lag (spatial autocorrelation) models appear in Table 13 (full results appear in Table 18). All coefficients appear in this table except the time and lake effects. The

[18] Due to the size of the weights matrices and the number of computations involved, it took an average of two and a half hours to construct each SWM. All of the computations were done on a laptop with a 2.2 Ghz processor with 3 gigabytes of RAM. It is likely that advances in computing power will decrease the time required to construct these matrices in the near future.

labels at the top of the table indicate which SWM is used in each model. "Sar" denotes the spatial autoregressive, or spatial lag, model. The autoregressive parameter ρ appears in the first row of Table 13. Although it is significant in all models except for the 20 nearest neighbor specification, the absolute magnitude of this parameter is small. Also, ρ is negative in the nearest neighbor models and positive in the distance and time models. The insignificance of ρ in the 20NN model may arise from the inclusion of distant properties into nearest neighbor sets, a consequence of the SWM construction. If a home does not have many nearby property sales, the specification will still include 20 neighbors, some of which may be separated by large distances. These remote homes may have no correlation with the price of the subject property, and if enough of these homes are included in the SWM it may dull the contribution of the spatial analysis.

The other coefficients in the model are relatively stable between specifications. The coefficient on *Ln(SDM)* varies from a minimum of -0.052 in the Sar200 model to a maximum of -0.011 in the SarNN15 model. With the distance and time SWM's, the coefficients only vary from -0.029 to -0.032, a fairly small range. The coefficients on the interaction term *Waterfront*Ln(SDM)* are even closer in range, all are within 0.001 of each other and significant at the 0.01 level except the Sar200 model. The water quality interaction terms exhibit similar stability. More importantly, the Edge Effect, Proximity Effect, and Area effect are all present in each of the spatial lag models. The variance in SWM's has not had a major impact on these main hypotheses. The majority of the coefficients in Table 13 exhibit stability across specifications, with the largest differences appearing in the latitude and longitude coefficients. As these coefficients also deal with spatial influences and location, it is no surprise that they are sensitive to the definition of the SWM. Finally, the two variables that are consistently insignificant are

%Black and *%Over 65*. The stability of the coefficients between SWM's is similar to the results of Mueller and Loomis (2008), who also found that their model's results were not particularly sensitive to specification. They point out that from a statistical perspective, the spatial models are superior. However, from a practical point of view, the differences in the coefficients may not be significant enough to alter policy.

The results of the spatial error models appear in Table 14, where the "Sem" in the column labels denotes spatial error and the numbers refer to the particular SWM (full coefficient results appear in Table 19). The first row contains the spatial error parameter λ, which is positive and significant in all models but displays much more variation than the spatial lag parameter ρ. Also, the SDM coefficients in this table display more variation between specifications than the spatial lag estimates. For instance, the coefficient on *ln(SDM)* varies from -0.07 to 0.06 and the *ln(Distance)*ln(SDM)* coefficient is insignificant in the Sem200 model but is highly significant in all other models. The Proximity Effect is therefore not present in the Sem200 model, further indicating that the spatial error model is more affected by changes in the weights matrix than the spatial lag model. The magnitudes of most of the other coefficients in Table 14 do not vary substantially over specifications. The interaction terms *Waterfront*Ln(SDM)* and *Ln(Area)*Ln(SDM)* are relatively stable across the six models, as are most of the home attribute variables. The Edge Effect and the Area Effect are therefore present in all specifications.

To analyze the differences in the various models in terms of dollar estimates, the marginal effects, or implicit prices, are computed for the spatial lag and spatial error models. The formula for the spatial lag implicit price is given in Chapter 1. The formula for the spatial error implicit prices appears in equation (2.37):

$$\frac{\partial P}{\partial SDM} = \left(\frac{P}{SDM}\right)\left(\beta_{SDM} + \beta_{WF*SDM} * WF + \beta_{Dist*SDM} * \ln(Dist) + \beta_{Area*SDM} * \ln(Area)\right) \tag{2.37}$$

The calculation of this implicit price is the same as the calculation of an OLS, or non-spatial, implicit price since there is no spatially weighted lag term.

Table 15 contains the spatial lag implicit prices for the six different SWM's. For the waterfront prices, the 15 nearest neighbor SWM model has the lowest value at $11,360.75. The maximum implicit price appears in the 200 regression, at $12,274.90. The values in the 200 specification are higher than the other estimates for all implicit prices. The 200123 and 40063 models illustrate that the effect of increasing the neighbor boundaries is to increase the waterfront implicit prices. This is not the case with non-waterfront implicit prices, as these have slightly different rates of decay with lake distance due to small differences in the *ln(Distance)*ln(SDM)* coefficient and differences in the value of ρ. Based on these observations, there is no clear relationship between the number of neighbors and the implicit prices. Nonetheless, changes in SWM only result in small percentage changes in the implicit prices in the spatial lag model.

The spatial error estimates appear in Table 16. These values are both considerably less than the spatial lag values and more sensitive to the SWM. The mean non-lakefront spatial error implicit prices are approximately half the value of the non-lakefront spatial lag prices. Also, whereas the 200 waterfront estimates were the largest in

Table 15, they are the smallest here. The 200 model has the most deviant values in Table 16 compared to the other, starting at a lower value but leveling off as distance is increased. Strangely, the implicit price of water quality for non-lakefront homes only changes by $46.55 from 500 to 900 meters. This anomalous behavior is contrary to all other spatial error models and may indicate that this particular SWM is ill-specified.

The effect of increasing the neighborhood boundary in Table 16 differs depending on the type of boundary relaxed. In the nearest neighbor specifications, including more neighbors depresses the implicit prices. With the distance/time based SWM's, increasing the time boundary decreases the implicit prices, while expanding the distance boundary increases the lakefront implicit price but slightly decreases the non-lakefront implicit price. There appears to be a complex relationship between the specification of the SWM and the implicit prices, which is more erratic with the spatial error model than the spatial lag model.

Comparing the implicit prices of the error and lag specifications, some general observations can be drawn. The spatial error regressions are much more sensitive to the particular specification of the SWM. For instance, the use of the 200 SWM results in a loss of the Proximity Effect. The spatial lag implicit prices are both larger than the spatial error models and closer to the original non-spatial implicit prices from Chapter 1. Overall, the coefficients and implicit prices are much more sensitive to the model specification (error vs lag) than to changes in the SWM. These differences between models will be further examined in the upcoming section on forecasting and errors.

2.5.3. Forecasting Errors

In order to examine the fit of the models that have been presented and discussed in the last two sections, a forecasting exercise is performed. This procedure follows four main steps. In the first step, the data are split into two portions: one containing 45,712 observations (referred to as the in-sample) and another containing 10,000 observations (the out-sample). In the second step, all of the models are estimated using the in-sample. In the third step, the estimates from the second step are applied to the out-sample to obtain predicted estimates. In the fourth step, the seven forms of prediction error from Case et al. (2004) are calculated and analyzed. These error variables include mean and median prediction error, the standard deviation of the errors, root mean and median square error, and mean and median absolute value percent error.

The prediction error variables for the spatial lag and spatial error models appear in Table 17. The first column in this table contains the mean prediction error, where all of the spatial lag models have lower (absolute magnitude) errors than the corresponding spatial error models. For example, the SARNN15 lag model has a mean prediction error of 0.00271 and the SEMNN15 model has a mean prediction error of -0.01277. The second column contains the median prediction errors, where all the spatial error models but one have lower median error than the corresponding spatial lag estimates. In the third column, all of the spatial lag models have a lower standard deviation of the prediction error than the spatial error models. The same also holds true for the root mean square error, root median square error, mean absolute value % error, and median absolute value % error. Given the superior performance of the spatial lag model in all columns but one, the spatial lag model fits the data much better than the spatial error model.

The results of Table 17 can also be used to compare the various weights matrices. In both the spatial lag and spatial error results, all columns indicate that the distance and time based SWM's have lower prediction error than the nearest neighbor specifications. For instance, the more accurate nearest neighbor model, NN20, has a mean prediction error that is 440% larger than the 40063 model. The out-of-sample predictive ability of the distance-based SWM's is clearly superior to that of the nearest neighbor specifications.

Comparisons can also be made within the two main types of SWM. With the distance and time based models, increasing the number of neighbors improves forecasting. In the three distance and time-based SWM SAR models, all columns of Table 17 show a slight improvement in prediction error when the distance or time boundary is slightly relaxed. In 5 of the 6 columns of Table 17 the SAR200123 and SAR40063 models have lower error than the more tightly defined SAR20063 model. The 200 model represents a complete relaxation of the time boundary, and although this model has approximately twice the mean prediction error as the SAR distance and time-based models, it has a superior performance in all other columns. In fact, the 200 model has the best error performance over all columns, between both SEM and SAR models. The increase in additional information in the SWM from the relaxation of this boundary appears to considerably improve the fit of the model. Also, in both SEM and SAR models and in all columns of Table 17, the NN20 model performed better than the NN15. Again, it appears that introducing more information into the SWM results in a better fit of the model. Case et al. (2004) used models based on a maximum of 15 nearest neighbors and suggest that results indicated that the inclusion of more neighbors could have improved the models. This general trend suggests that future analyses may want to increase the average number of items in each row of the SWM.

Future research should further examine the proper specification of the SWM. Numerous configurations are currently used without much theoretical or empirical motivation. Many hedonic analyses also use SWM's based on nearest neighbors and contiguity that have a mean number of neighbors under 10 (for instance, Anselin and Le Gallo (2006), Anselin and Lozano-Gracia (2008)). The present study and Case et al. (2004) both find evidence that it may be better to increase the number of neighbors in the SWM beyond these low levels. Until tests or theoretical developments emerge for the specification of the SWM, forecasting represents a reliable way to compare alternatives.

2.6. Conclusions

Spatial econometric techniques have seen a rapid increase in use over the last twenty years. Many fields that had traditionally been estimated with ordinary least squares or other common techniques are now being re-estimated to account for spatial dependence. Studies of regional GDP, international growth, and property prices, for example, have been shown to be critically affected by spatial influences. To account for spatial dependence, a wide variety of techniques, using maximum likelihood, GMM, and Bayesian methods are available (Arbia 2006; Anselin and Lozano-Gracia 2008). A common aspect of most of these models is that the spatial correlation between data points must be exogenously specified through an SWM. The construction of the SWM has not yet received critical attention, as computing advances have only recently allowed a variety of options. This chapter examines the behavior of several different SWM's while using three different spatial models.

After investigating the out-of-sample predictive ability of the various weights matrices and spatial models, some general conclusions arise. First, in this particular application, the spatial lag model fit the data much better than the spatial error model. The spatial error model actually performed worse than the original non-spatial model in terms of prediction error.

There are also some general trends observed from varying the SWM in the spatial error and spatial lag models. First, the distance and time based weights matrices performed much better than the nearest neighbor matrices. Also, when the distance based matrices are used, increasing the total number of neighbors improves the fit as more information is embedded in the SWM. These are critical results, as many recent papers use nearest neighbor matrices, such as Kim, Phipps, and Anselin (2003), Anselin and Le Gallo (2006), and Anselin and Lozano-Gracia (2008). In addition, it is clear that different SWM's can result in quite different results. In the spatial error model, the six SWM specifications result in waterfront implicit prices that vary from \$5,072 to \$10,251, a 100% change in magnitude. Given this wide range of values, the only current method of determining which value is right is through a forecasting application, which may not always be feasible. This clearly represents a gap in the literature on spatial dependence.

The one area of optimism with respect to the SWM definition is spatial testing. Varying the SWM had no effect on any of the qualitative results of the DLR tests. A similar result was found by Anselin (1988), which is replicated in the current paper. Nonetheless, further exploration of simultaneous tests of spatial dependence and functional form should be explored.

Tables and Figures: Chapter 2

Table 7: Summary Statistics for "In-Sample"

		Lakefront (N =1,218)		Non-Lakefront (N =43,494)	
Variable	Units	Mean	Std Dev	Mean	Std Dev
Property Characteristics					
Sales Price	2002 Dollars	451,086.90	363,414.50	200,383.40	201,669.20
Heated Area	Square Feet	2,761.88	1,290.42	1,963.40	946.19
Area of Parcel	Square Meters	2,798.07	2,757.67	1,079.35	1,140.27
Number of Bathrooms	--	2.75	1.11	2.22	0.88
Home Age	Years	24.18	13.35	18.61	15.05
% With Pool	--	20.11	--	20.65	--
Spatial Characteristics					
Distance to Nearest Lake	Meters	42.32	23.69	467.00	267.95
Area of Nearest Lake	Acres	2,044,963.00	2,543,847.0	1,132,663.00	1,809,132.00
Distance to CBD	Meters	8,844.95	5,149.64	9,269.03	5,151.67
Latitude Coordinate	Degrees	654,732.60	7,463.92	653,235.50	7,949.14
Longitude Coordinate	Degrees	505,761.60	6,757.01	505,668.70	6,273.74
% In airport noise zone	--	9.03	--	15.73	--
Census Block Characteristics					
% of Population White	--	88.6	--	0.22	--
% of Population Black	--	4.9	--	0.20	--
% of Population > 65	--	15.2	--	0.07	--
Median Household Income	2002 Dollars	64,053.01	28,880.8	56,400.49	23,795.07
Distribution of Sales by Year					
% of sales in 1996	--	0.10	--	0.09	--
% of sales in 1997	--	0.11	--	0.11	--
% of sales in 1998	--	0.14	--	0.12	--
% of sales in 1999	--	0.14	--	0.13	--
% of sales in 2000	--	0.11	--	0.11	--
% of sales in 2001	--	0.08	--	0.10	--
% of sales in 2002	--	0.12	--	0.10	--
% of sales in 2003	--	0.12	--	0.11	--
% of sales in 2004	--	0.09	--	0.12	--

Table 8: Spatial Weights Matrix Descriptions and Summary Statistics

SWM Name	Distance/Time Boundaries: Distance	Time backward	Time Forward	# of Neighbors: Mean	Min	Max	# of Non-Zero Elements in SWM
200	200	--	--	46.724	0	214	2,556,364
20063	200	6	3	5.6191	0	111	307,432
200123	200	12	3	8.6256	0	119	471,924
40063	400	6	3	14.4779	0	137	792,114
NN15	--	--	--	15	15	15	627,945
NN20	--	--	--	20	20	20	837,260

Table 9: One Direction DLR Tests of Spatial Dependence, DT20063 SWM

	Spatial Lag		Spatial Error	
	ρ=0 | r=0	ρ=0 |r=1	λ=0 | r=0	λ=0 | r=1
Belle Isle	30.92	30.69	29.87	48.73
Edgewood	32.72	120.45	39.41	99.31
Maitland/Eatonville	46.72	65.19	54.31	46.72
Ocoee/Winter Garden	27.63	27.37	27.36	29.61
Windermere	29.96	36.14	31.46	34.59
Winter Park	41.14	50.65	42.78	75.37
Orlando 1	106.11	107.81	106.43	114.76
Orlando 2	106.02	155.90	105.99	240.51
Orlando 3	106.41	114.15	109.99	112.45
Unincorporated 1	121.91	231.72	121.80	312.66
Unincorporated 2	124.32	145.92	124.33	133.29
Unincorporated 3	122.84	123.26	121.07	139.33
Unincorporated 4	125.20	142.24	136.18	147.48

Table 10: One Direction DLR Tests of Spatial Dependence, NN15 SWM

	Spatial Lag		Spatial Error	
	ρ=0 | r=0	ρ=0 |r=1	λ=0 | r=0	λ=0 | r=1
Belle Isle	41.12	55.75	34.29	62.42
Edgewood	31.49	33.79	38.13	38.13
Maitland/Eatonville	58.87	47.76	61.07	48.22
Ocoee/Winter Garden	28.69	27.29	30.21	27.66
Windermere	29.32	30.62	29.26	27.50
Winter Park	42.11	369.74	57.26	141.33
Orlando 1	106.02	108.08	106.67	115.44
Orlando 2	105.95	148.61	107.45	168.06
Orlando 3	106.17	115.01	106.87	116.34
Unincorporated 1	122.95	155.95	124.16	145.19
Unincorporated 2	124.24	161.62	126.86	173.78
Unincorporated 3	121.09	129.20	121.98	122.05
Unincorporated 4	126.97	257.82	135.48	243.18

Table 11: Joint DLR Tests of Functional Form and Spatial Dependence, DT20063 SWM

	Spatial Lag		Spatial Error	
	r=0, ρ=0	r=1, ρ=0	r=0, λ=0	r=1, λ=0
Belle Isle	55.67	1,235.20	54.44	1,236.55
Edgewood	32.81	407.56	39.44	404.67
Maitland/Eatonville	47.28	1,857.65	54.75	1,858.27
Ocoee/Winter Garder	40.25	448.05	39.82	448.04
Windermere	51.31	457.56	51.96	456.11
Winter Park	51.67	4,331.14	53.58	4,337.97
Orlando 1	120.78	6,013.77	121.05	6,015.75
Orlando 2	122.48	6,027.86	122.48	6,027.96
Orlando 3	153.23	6,501.46	156.87	6,500.64
Unincorporated 1	481.12	8,681.12	483.72	8,680.89
Unincorporated 2	446.85	8,979.65	447.11	8,980.24
Unincorporated 3	275.95	8,607.37	274.38	8,606.78
Unincorporated 4	406.07	9,367.29	417.15	9,368.49

Table 12: Joint DLR Tests of Functional Form and Spatial Dependence, NN15 SWM

	Spatial Lag		Spatial Error	
	r=0, ρ=0	r=1, ρ=0	r=0, λ=0	r=1, λ=0
Belle Isle	62.96	1,256.32	57.66	1,236.64
Edgewood	31.55	402.99	38.16	403.42
Maitland/Eatonville	59.57	1,856.24	61.65	1,856.36
Ocoee/Winter Garder	41.49	448.89	41.61	455.21
Windermere	50.95	453.88	51.40	454.03
Winter Park	52.16	4,332.80	65.09	4,349.56
Orlando 1	120.69	6,013.86	121.29	6,015.00
Orlando 2	122.42	6,026.44	124.05	6,027.80
Orlando 3	153.17	6,497.36	153.94	6,498.89
Unincorporated 1	481.03	8,684.35	482.16	8,683.27
Unincorporated 2	446.98	8,980.48	448.91	8,980.41
Unincorporated 3	274.35	8,606.86	275.69	8,608.44
Unincorporated 4	406.65	9,366.79	419.70	9,369.99

Table 13: Spatial Lag Coefficients

	SarNN15	SarNN20	Sar200	Sar 20063	Sar 200123	Sar 40063
Rho	-0.0027*	-0.0001	0.1090*	0.0014*	0.0026*	0.0083*
	(0.0002)	(0.0002)	(0.0077)	(0.0003)	(0.0004)	(0.0007)
Ln(SDM)	-0.011	-0.028	-0.052	-0.030	-0.029	-0.032
	(0.031)	(0.031)	(0.030)	(0.031)	(0.031)	(0.031)
Waterfront*Ln(SDM)	0.078*	0.078*	0.068*	0.079*	0.079*	0.079*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Ln(Distance)*Ln(SDM)	-0.017*	-0.017*	-0.015*	-0.017*	-0.017*	-0.017*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(Area)*Ln(SDM)	0.010*	0.011*	0.013*	0.012*	0.011*	0.012*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Waterfront	0.147*	0.146*	0.152*	0.146*	0.146*	0.147*
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Ln(Distance Lake)	-0.035*	-0.035*	-0.029*	-0.035*	-0.035*	-0.034*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Canalfront	0.106*	0.105*	0.099*	0.105*	0.104*	0.104*
	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)	(0.018)
Golffront	0.331*	0.332*	0.308*	0.334*	0.335*	0.341*
	(0.033)	(0.033)	(0.032)	(0.033)	(0.033)	(0.033)
Ln(Bath)	0.119*	0.120*	0.111*	0.120*	0.120*	0.120*
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Pool	0.013*	0.014*	0.013*	0.014*	0.014*	0.014*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Home Age)	-0.072*	-0.071*	-0.064*	-0.070*	-0.070*	-0.070*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Home Area)	0.675*	0.673*	0.621*	0.672*	0.672*	0.670*
	(0.005)	(0.005)	(0.003)	(0.005)	(0.005)	(0.005)
Ln(Parcel Area)	0.166*	0.165*	0.162*	0.167*	0.167*	0.167*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Distance CBD)	-0.189*	-0.189*	-0.169*	-0.189*	-0.189*	-0.189*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Near Airport	0.028*	0.029*	0.028*	0.028*	0.028*	0.028*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln(Lattitude)	-6.633*	-6.445*	-5.680*	-6.509*	-6.555*	-6.727*
	(0.014)	(0.070)	(0.199)	(0.053)	(0.074)	(0.204)
Ln(Longitude)	3.620*	3.395*	2.761*	3.462*	3.440*	3.257*
	(0.664)	(0.666)	(0.640)	(0.666)	(0.665)	(0.657)
Ln(Median Income)	0.107*	0.107*	0.087*	0.106*	0.106*	0.103*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
% White (2000)	0.249*	0.246*	0.179*	0.244*	0.244*	0.238*
	(0.032)	(0.032)	(0.031)	(0.032)	(0.032)	(0.032)
% Black (2000)	0.056***	0.052	0.005	0.053	0.053	0.047
	(0.034)	(0.034)	(0.033)	(0.034)	(0.034)	(0.034)
% Over 65 (2000)	0.021	0.022	0.025	0.023	0.022	0.022
	(0.020)	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)
R^2	0.8940	0.8935	0.8907	0.8935	0.8934	0.8933

Note: *, **, and *** denotes significance at the 1%, 5% and 10 % levels. Standard errors appear in parentheses.

Table 14: Spatial Error Coefficients

	semNN15	semNN20	sem200	sem 20063	sem 200123	sem 40063
Lambda	0.7560*	0.7820*	0.5940*	0.3500*	0.4480*	0.5070*
	(0.0057)	(0.0045	(0.0013)	(0.0025)	(0.0020)	(0.0017)
Ln(SDM)	-0.005	-0.011	-0.078*	0.024	0.064***	0.006
	(0.030)	(0.030)	(0.028)	(0.037)	(0.038)	(0.042)
Waterfront*Ln(SDM)	0.061*	0.061*	0.034*	0.058*	0.060*	0.067*
	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Ln(Distance)*Ln(SDM)	-0.011*	-0.010*	-0.002	-0.020*	-0.021*	-0.019*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Area)*Ln(SDM)	0.006*	0.006*	0.008*	0.008*	0.005***	0.009*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Waterfront	0.157*	0.154*	0.197*	0.171*	0.165*	0.157*
	(0.019)	(0.019)	(0.021)	(0.019)	(0.019)	(0.019)
Ln(Distance Lake)	-0.066*	-0.068*	-0.078*	-0.042*	-0.046*	-0.044*
	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Canalfront	0.070*	0.067*	0.066*	0.076*	0.069*	0.088*
	(0.018)	(0.018)	(0.021)	(0.017)	(0.017)	(0.017)
Golffront	0.231*	0.221*	0.236*	0.280*	0.217*	0.245*
	(0.031)	(0.031)	(0.038)	(0.030)	(0.030)	(0.030)
Ln(Bath)	0.089*	0.092*	0.098*	0.104*	0.098*	0.101*
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Pool	0.008*	0.007*	0.007*	0.009*	0.007*	0.009*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(Home Age)	-0.085*	-0.086*	-0.086*	-0.079*	-0.081*	-0.081*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Home Area)	0.557*	0.563*	0.539*	0.603*	0.575*	0.584*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Ln(Parcel Area)	0.172*	0.170*	0.179*	0.173*	0.174*	0.172*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Distance CBD)	-0.203*	-0.166*	-0.192*	-0.188*	-0.190*	-0.187*
	(0.015)	(0.015)	(0.018)	(0.013)	(0.014)	(0.014)
Near Airport	0.030*	0.026*	0.010	0.020*	0.015***	0.014***
	(0.008)	(0.008)	(0.009)	(0.007)	(0.008)	(0.007)
Ln(Lattitude)	-4.378*	-3.485*	-4.096	-5.358*	-5.066*	-2.774*
	(0.018)	(0.770)	(0.413)	(0.673)	(0.440)	(0.497)
Ln(Longitude)	2.530*	3.038*	2.785**	2.247*	2.410*	2.134**
	(1.319)	(0.994)	(1.162)	(0.765)	(0.905)	(0.912)
Ln(Median Income)	0.121*	0.117*	0.127	0.121*	0.125*	0.118*
	(0.007)	(0.007)	(0.009)	(0.006)	(0.007)	(0.007)
% White (2000)	0.184*	0.189*	0.237*	0.266*	0.284*	0.224*
	(0.049)	(0.049)	(0.057)	(0.040)	(0.044)	(0.044)
% Black (2000)	-0.058	-0.054	0.036*	0.040	0.042	0.012
	(0.053)	(0.053)	(0.060)	(0.042)	(0.047)	(0.047)
% Over 65 (2000)	0.091*	0.101*	0.043*	0.050**	0.039	0.045***
	(0.027)	(0.027)	(0.033)	(0.024)	(0.026)	(0.025)
R^2	0.9171	0.9157	0.9259	0.9093	0.9145	0.9142

Note: *, **, and *** denotes significance at the 1%, 5% and 10 % levels. Standard errors appear in parentheses.

Table 15: Spatial Lag Implicit Prices

	NN15	NN20	D200	DT20063	DT200123	DT40063
SDM_WF	11,360.75	11,668.39	12,274.90	11,784.28	11,842.51	11,978.16
	(852.18)	(855.77)	(934.37)	(857.10)	(858.02)	(862.12)
SDM_NWF	803.55	879.64	1,118.96	890.72	887.94	918.62
	(151.16)	(151.77)	(165.71)	(151.95)	(152.10)	(152.87)
SDM_100	1,678.36	1,764.62	2,017.99	1,786.31	1,784.34	1,821.63
	(179.69)	(180.42)	(197.34)	(180.68)	(180.84)	(181.76)
SDM_300	1,054.88	1,133.88	1,377.24	1,148.01	1,145.46	1,178.05
	(150.82)	(151.42)	(165.44)	(151.62)	(151.76)	(152.54)
SDM_500	764.97	840.61	1,079.31	851.22	848.40	878.80
	(151.87)	(152.49)	(166.48)	(152.67)	(152.82)	(153.60)
SDM_700	574.02	647.43	883.07	655.73	652.74	681.69
	(157.87)	(158.52)	(172.99)	(158.71)	(158.87)	(159.66)
SDM_900	431.39	503.15	736.50	509.71	506.59	534.47
	(164.82)	(165.50)	(180.57)	(165.70)	(165.87)	(166.69)
WF	128,783.97	128,563.36	138,756.89	129,408.11	129,788.42	131,127.18
	(3,103.21)	(3,116.71)	(3,385.91)	(3,123.71)	(3,126.72)	(3,141.40)
Distance (M)	-27.75	-27.79	-27.08	-27.97	-27.94	-27.92
	(0.64)	(0.64)	(0.64)	(0.64)	(0.64)	(0.64)

Standard errors appear in parentheses.

Table 16: Spatial Error Implicit Prices

	NN15	NN20	D200	DT20063	DT200123	DT40063
SDM_WF	7,784.93 (816.23)	7,399.15 (819.40)	5,072.04 (856.18)	9,691.46 (885.07)	9,244.18 (889.51)	10,251.20 (935.17)
SDM_NWF	336.41 (142.11)	252.86 (141.86)	705.20 (125.30)	489.42 (189.37)	254.90 (189.19)	487.60 (222.71)
SDM_100	906.55 (181.98)	770.38 (182.46)	827.27 (171.20)	1,542.16 (219.31)	1,336.42 (223.59)	1,490.51 (253.19)
SDM_300	500.21 (140.85)	401.53 (140.85)	740.27 (124.10)	791.86 (188.55)	565.61 (188.55)	775.73 (222.14)
SDM_500	311.26 (143.32)	230.03 (143.04)	699.82 (126.68)	442.99 (190.22)	207.20 (190.10)	443.37 (223.50)
SDM_700	186.81 (152.93)	117.07 (152.49)	673.17 (137.62)	213.19 (197.08)	---	224.45 (230.09)
SDM_900	93.86 (163.55)	32.69 (163.02)	653.27 (149.59)	41.56 (204.91)	---	60.94 (237.74)
WF	120,116.49 (3,043.52)	118,934.86 (3,043.96)	116,723.81 (3,566.52)	124,039.05 (3,029.50)	122,664.99 (3,039.89)	125,039.50 (2,979.51)
Dist	-36.35 (0.90)	-36.64 (0.89)	-35.02 (1.05)	-33.22 (0.76)	-35.39 (0.83)	-33.42 (0.82)

Standard errors appear in parentheses.

Table 17: Forecasting Error, Spatial Lag and Spatial Error Models

Model	Mean Error	Median Error	St. Dev. Error	Root mean sq error	Root median sq error	Mean abs value % error	Median abs value % error
Spatial Lag							
SAR NN15	0.00271	0.01070	0.22151	0.16213	0.12395	0.01352	0.01037
SAR NN20	0.00250	0.01057	0.22150	0.16210	0.12394	0.01352	0.01037
SAR200	0.00100	0.00709	0.21752	0.15638	0.11877	0.01305	0.00992
SAR20063	-0.00053	0.00737	0.22132	0.16182	0.12381	0.01350	0.01035
SAR200123	-0.00047	0.00745	0.22132	0.16176	0.12354	0.01349	0.01034
SAR40063	-0.00046	0.00762	0.22117	0.16161	0.12323	0.01348	0.01034
Spatial Error							
SEM NN15	-0.01277	-0.00385	0.22747	0.16830	0.12917	0.01403	0.01077
SEM NN20	-0.01254	-0.00378	0.22742	0.16828	0.12863	0.01403	0.01072
SEM200	0.00258	0.00943	0.22624	0.16684	0.12795	0.01389	0.01077
SEM20063	-0.00183	0.00620	0.22282	0.16401	0.12587	0.01367	0.01054
SEM200123	-0.00400	0.00485	0.22400	0.16511	0.12734	0.01376	0.01057
SEM40063	-0.00473	0.00324	0.22404	0.16501	0.12755	0.01375	0.01066

Table 18: Full Spatial Lag Results

	SarNN15	SarNN20	Sar200	Sar20063	Sar200123	Sar40063
Constant	48.103 (8.868)	48.532 (9.686)	45.760 (11.084)	48.499 (9.458)	49.395 (9.729)	54.059 (11.371)
Ln(SDM)	-0.011 (0.031)	-0.028 (0.031)	-0.052 (0.030)	-0.030 (0.031)	-0.029 (0.031)	-0.032 (0.031)
ln(SDM)*Waterfront	0.078 (0.011)	0.078 (0.011)	0.068 (0.011)	0.079 (0.011)	0.079 (0.011)	0.079 (0.011)
Ln(SDM)*ln(dist)	-0.017 (0.002)	-0.017 (0.002)	-0.015 (0.002)	-0.017 (0.002)	-0.017 (0.002)	-0.017 (0.002)
Ln(SDM)*ln(lake area)	0.010 (0.002)	0.011 (0.002)	0.013 (0.002)	0.012 (0.002)	0.011 (0.002)	0.012 (0.002)
Waterfront	0.147 (0.020)	0.146 (0.020)	0.152 (0.020)	0.146 (0.020)	0.146 (0.020)	0.147 (0.020)
Ln(lake dist)	-0.035 (0.004)	-0.035 (0.004)	-0.029 (0.004)	-0.035 (0.004)	-0.035 (0.004)	-0.034 (0.004)
Canalfront	0.106 (0.018)	0.105 (0.018)	0.099 (0.017)	0.105 (0.018)	0.104 (0.018)	0.104 (0.018)
Golffront	0.331 (0.033)	0.332 (0.033)	0.308 (0.032)	0.334 (0.033)	0.335 (0.033)	0.341 (0.033)
Ln(bath)	0.119 (0.005)	0.120 (0.005)	0.111 (0.004)	0.120 (0.005)	0.120 (0.005)	0.120 (0.005)
Pool	0.013 (0.003)	0.014 (0.003)	0.013 (0.003)	0.014 (0.003)	0.014 (0.003)	0.014 (0.003)
Ln(age)	-0.072 (0.001)	-0.071 (0.001)	-0.064 (0.001)	-0.070 (0.001)	-0.070 (0.001)	-0.070 (0.001)
Ln(area_heated)	0.675 (0.005)	0.673 (0.005)	0.621 (0.003)	0.672 (0.005)	0.672 (0.005)	0.670 (0.005)
Ln(area parcel)	0.166 (0.003)	0.165 (0.003)	0.162 (0.003)	0.167 (0.003)	0.167 (0.003)	0.167 (0.003)
Ln(dist CBD)	-0.189 (0.010)	-0.189 (0.010)	-0.169 (0.010)	-0.189 (0.010)	-0.189 (0.010)	-0.189 (0.010)
Near airport	0.028 (0.006)	0.029 (0.006)	0.028 (0.006)	0.028 (0.006)	0.028 (0.006)	0.028 (0.006)
ln(x_coord)	-6.633 (0.014)	-6.445 (0.070)	-5.680 (0.199)	-6.509 (0.053)	-6.555 (0.074)	-6.727 (0.204)
ln(y_coord)	3.620 (0.664)	3.395 (0.666)	2.761 (0.640)	3.462 (0.666)	3.440 (0.665)	3.257 (0.657)
Ln(Med income)	0.107 (0.005)	0.107 (0.005)	0.087 (0.005)	0.106 (0.005)	0.106 (0.005)	0.103 (0.005)
Percent white00	0.249 (0.032)	0.246 (0.032)	0.179 (0.031)	0.244 (0.032)	0.244 (0.032)	0.238 (0.032)
Percent black00	0.056 (0.034)	0.052 (0.034)	0.005 (0.033)	0.053 (0.034)	0.053 (0.034)	0.047 (0.034)
Percent over6500	0.021 (0.020)	0.022 (0.020)	0.025 (0.019)	0.023 (0.020)	0.022 (0.020)	0.022 (0.020)
year_1996	-0.814 (0.004)	-0.821 (0.004)	-0.818 (0.004)	-0.819 (0.004)	-0.816 (0.004)	-0.812 (0.004)
year_1997	-0.754	-0.762	-0.759	-0.760	-0.759	-0.755

	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_1998	-0.687	-0.692	-0.690	-0.690	-0.690	-0.685
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_1999	-0.598	-0.605	-0.602	-0.603	-0.603	-0.599
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2000	-0.493	-0.498	-0.496	-0.497	-0.497	-0.494
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2001	-0.376	-0.379	-0.376	-0.378	-0.377	-0.375
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2002	-0.284	-0.285	-0.282	-0.284	-0.284	-0.282
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
year_2003	-0.156	-0.156	-0.155	-0.155	-0.155	-0.154
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
BASS_LAKE	-0.079	-0.080	-0.035	-0.077	-0.076	-0.070
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)
BAY_LAKE_B	-0.265	-0.260	-0.223	-0.255	-0.253	-0.234
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)
BEARHEAD_LAKE	-0.039	-0.041	-0.013	-0.037	-0.036	-0.033
	(0.026)	(0.026)	(0.025)	(0.026)	(0.026)	(0.026)
BIG_SAND_LAKE	0.168	0.162	0.134	0.162	0.162	0.157
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
CLEAR_LAKE	-0.359	-0.361	-0.295	-0.359	-0.358	-0.353
	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)	(0.018)
DEEP_LAKE	-0.004	0.003	0.043	0.006	0.009	0.022
	(0.035)	(0.036)	(0.037)	(0.036)	(0.036)	(0.038)
KASEY_LAKE	-0.208	-0.197	-0.140	-0.198	-0.197	-0.190
	(0.024)	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)
KELLY_LAKE	-0.238	-0.222	-0.152	-0.222	-0.221	-0.212
	(0.027)	(0.027)	(0.026)	(0.027)	(0.027)	(0.027)
KRISTY_LAKE	-0.252	-0.238	-0.168	-0.238	-0.237	-0.229
	(0.026)	(0.026)	(0.025)	(0.026)	(0.026)	(0.027)
LAKE_ADAIR	0.130	0.135	0.155	0.135	0.136	0.141
	(0.023)	(0.023)	(0.022)	(0.023)	(0.023)	(0.023)
LAKE_ANDERSON	-0.037	-0.037	-0.007	-0.034	-0.033	-0.027
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)
LAKE_ANGEL	-0.438	-0.435	-0.354	-0.431	-0.431	-0.425
	(0.025)	(0.025)	(0.024)	(0.025)	(0.025)	(0.026)
LAKE_ARNOLD	-0.086	-0.085	-0.050	-0.083	-0.082	-0.075
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
LAKE_BALDWIN	0.031	0.026	0.027	0.030	0.032	0.038
	(0.019)	(0.020)	(0.021)	(0.020)	(0.020)	(0.022)
LAKE_BARTON	-0.189	-0.193	-0.160	-0.190	-0.188	-0.181
	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.019)
LAKE_BEARDALL	-0.479	-0.471	-0.345	-0.462	-0.459	-0.437
	(0.051)	(0.051)	(0.049)	(0.051)	(0.051)	(0.051)
LAKE_BEAUTY	-0.046	-0.032	0.005	-0.029	-0.029	-0.024
	(0.045)	(0.045)	(0.044)	(0.045)	(0.045)	(0.045)
LAKE_BELL	-0.166	-0.160	-0.115	-0.159	-0.159	-0.150
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)

LAKE_BERRY	0.183	0.184	0.204	0.184	0.186	0.193
	(0.021)	(0.021)	(0.022)	(0.021)	(0.021)	(0.023)
LAKE_BESSIE	0.530	0.532	0.468	0.533	0.533	0.531
	(0.023)	(0.023)	(0.022)	(0.023)	(0.023)	(0.023)
LAKE_BLANCHE	0.067	0.070	0.073	0.071	0.070	0.068
	(0.013)	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)
LAKE_BUCHANNAN	-0.113	-0.113	-0.029	-0.111	-0.109	-0.105
	(0.030)	(0.030)	(0.028)	(0.030)	(0.030)	(0.030)
LAKE_BUCK	0.548	0.532	0.450	0.540	0.542	0.552
	(0.023)	(0.022)	(0.020)	(0.022)	(0.022)	(0.021)
LAKE_BUMBY	-0.067	-0.064	-0.010	-0.061	-0.060	-0.056
	(0.029)	(0.029)	(0.028)	(0.029)	(0.029)	(0.029)
LAKE_BURKETT	0.048	0.051	0.082	0.054	0.057	0.071
	(0.027)	(0.028)	(0.030)	(0.028)	(0.029)	(0.031)
LAKE_BUTLER	-0.060	-0.067	-0.081	-0.066	-0.066	-0.067
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
LAKE_C	-0.108	-0.105	-0.065	-0.102	-0.101	-0.091
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)
LAKE_CANE_A	-0.025	-0.022	-0.005	-0.021	-0.021	-0.021
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
LAKE_CATHERINE_B	0.194	0.201	0.191	0.203	0.206	0.218
	(0.047)	(0.048)	(0.047)	(0.047)	(0.048)	(0.048)
LAKE_CAY_DEE	-0.043	-0.034	0.010	-0.032	-0.031	-0.022
	(0.026)	(0.027)	(0.026)	(0.026)	(0.027)	(0.027)
LAKE_CHARITY	0.174	0.181	0.155	0.180	0.181	0.187
	(0.037)	(0.038)	(0.038)	(0.037)	(0.038)	(0.039)
LAKE_CHASE	0.490	0.490	0.429	0.492	0.491	0.486
	(0.031)	(0.031)	(0.030)	(0.031)	(0.031)	(0.031)
LAKE_CHEROKEE	0.104	0.109	0.123	0.111	0.111	0.115
	(0.029)	(0.029)	(0.028)	(0.029)	(0.029)	(0.029)
LAKE_CHRISTIE	-0.100	-0.102	-0.058	-0.099	-0.099	-0.090
	(0.035)	(0.035)	(0.033)	(0.035)	(0.035)	(0.034)
LAKE_COMO	0.078	0.088	0.124	0.091	0.092	0.100
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)
LAKE_CONCORD	-0.015	-0.010	0.025	-0.009	-0.008	-0.002
	(0.029)	(0.029)	(0.028)	(0.029)	(0.029)	(0.029)
LAKE_CONWAY	-0.082	-0.096	-0.091	-0.094	-0.094	-0.092
	(0.015)	(0.014)	(0.014)	(0.015)	(0.014)	(0.014)
LAKE_COPELAND	0.029	0.033	0.065	0.036	0.036	0.041
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
LAKE_DANIEL	-0.072	-0.064	-0.018	-0.062	-0.063	-0.056
	(0.032)	(0.033)	(0.032)	(0.033)	(0.033)	(0.033)
LAKE_DESTINY	0.267	0.281	0.291	0.281	0.283	0.292
	(0.034)	(0.035)	(0.035)	(0.035)	(0.035)	(0.036)
LAKE_DOT	-0.420	-0.413	-0.296	-0.409	-0.407	-0.399
	(0.038)	(0.038)	(0.036)	(0.038)	(0.038)	(0.038)
LAKE_DOVER	-0.010	-0.008	0.048	-0.005	-0.005	0.004
	(0.040)	(0.040)	(0.039)	(0.040)	(0.040)	(0.040)
LAKE_DOWN	0.000	-0.006	-0.029	-0.006	-0.006	-0.009

	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)
LAKE_DOWNEY	0.072	0.075	0.112	0.078	0.081	0.093
	(0.021)	(0.022)	(0.024)	(0.022)	(0.022)	(0.025)
LAKE_DRUID	0.042	0.049	0.080	0.051	0.052	0.059
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
LAKE_EMERALD	0.141	0.141	0.171	0.146	0.150	0.173
	(0.027)	(0.028)	(0.030)	(0.028)	(0.028)	(0.031)
LAKE_EOLA	-0.068	-0.062	-0.026	-0.059	-0.058	-0.054
	(0.036)	(0.036)	(0.035)	(0.036)	(0.036)	(0.036)
LAKE_ESTELLE	0.066	0.071	0.084	0.071	0.072	0.078
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)
LAKE_FAIRHOPE	-0.010	-0.006	0.044	-0.001	-0.001	0.005
	(0.035)	(0.035)	(0.034)	(0.035)	(0.035)	(0.036)
LAKE_FAIRVIEW	-0.259	-0.258	-0.206	-0.257	-0.257	-0.249
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)
LAKE_FAITH	0.050	0.057	0.078	0.060	0.061	0.072
	(0.037)	(0.038)	(0.038)	(0.038)	(0.038)	(0.039)
LAKE_FARRAR	-0.022	-0.018	0.012	-0.015	-0.014	-0.007
	(0.024)	(0.025)	(0.024)	(0.024)	(0.025)	(0.025)
LAKE_FORMOSA	0.001	0.007	0.039	0.009	0.010	0.016
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)
LAKE_FREDRICA	0.190	0.188	0.186	0.191	0.192	0.198
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)
LAKE_GATLIN	0.074	0.071	0.065	0.076	0.075	0.076
	(0.025)	(0.025)	(0.024)	(0.025)	(0.025)	(0.025)
LAKE_GEAR	0.032	0.033	0.041	0.037	0.038	0.048
	(0.031)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
LAKE_GEM_A	-0.034	-0.024	0.024	-0.022	-0.020	-0.011
	(0.028)	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)
LAKE_GEM_MARY	0.005	0.008	0.029	0.010	0.010	0.015
	(0.023)	(0.023)	(0.022)	(0.023)	(0.023)	(0.023)
LAKE_GEORGE	0.013	0.010	0.019	0.013	0.014	0.017
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)
LAKE_GEORGIA	0.106	0.106	0.135	0.108	0.110	0.121
	(0.021)	(0.022)	(0.025)	(0.022)	(0.023)	(0.026)
LAKE_GILES	-0.116	-0.118	-0.084	-0.115	-0.115	-0.108
	(0.021)	(0.021)	(0.020)	(0.021)	(0.021)	(0.021)
LAKE_GLORIA	-0.025	-0.028	0.018	-0.025	-0.025	-0.023
	(0.019)	(0.019)	(0.017)	(0.019)	(0.019)	(0.018)
LAKE_GREENWOOD	0.011	0.020	0.053	0.023	0.023	0.029
	(0.035)	(0.035)	(0.034)	(0.035)	(0.035)	(0.035)
LAKE_HART	0.267	0.249	0.197	0.255	0.257	0.265
	(0.022)	(0.021)	(0.018)	(0.021)	(0.021)	(0.018)
LAKE_HIAWASSEE	-0.124	-0.121	-0.091	-0.121	-0.121	-0.118
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
LAKE_HIGHLAND	0.037	0.042	0.072	0.044	0.045	0.051
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
LAKE_HOLDEN	-0.322	-0.325	-0.266	-0.323	-0.322	-0.316
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)

LAKE_HOPE	-0.104 (0.053)	-0.101 (0.053)	-0.071 (0.053)	-0.094 (0.053)	-0.088 (0.053)	-0.062 (0.054)
LAKE_HUNGERFORD	-0.306 (0.033)	-0.298 (0.033)	-0.231 (0.033)	-0.294 (0.033)	-0.291 (0.034)	-0.279 (0.034)
LAKE_IRMA	0.043 (0.018)	0.040 (0.020)	0.069 (0.022)	0.041 (0.019)	0.044 (0.020)	0.053 (0.023)
LAKE_ISLEWORTH	0.358 (0.040)	0.359 (0.040)	0.297 (0.039)	0.362 (0.040)	0.363 (0.040)	0.367 (0.040)
LAKE_IVANHOE	0.086 (0.019)	0.090 (0.019)	0.114 (0.019)	0.091 (0.019)	0.092 (0.019)	0.097 (0.020)
LAKE_JACKSON	0.134 (0.032)	0.145 (0.032)	0.163 (0.033)	0.147 (0.032)	0.149 (0.032)	0.158 (0.033)
LK_JEN_JEWEL	-0.117 (0.021)	-0.119 (0.021)	-0.092 (0.020)	-0.117 (0.021)	-0.117 (0.021)	-0.113 (0.021)
LAKE_JESSAMINE	-0.236 (0.013)	-0.242 (0.012)	-0.207 (0.011)	-0.239 (0.013)	-0.238 (0.012)	-0.235 (0.012)
LAKE_KILLARNEY	-0.096 (0.019)	-0.096 (0.020)	-0.053 (0.020)	-0.095 (0.019)	-0.094 (0.020)	-0.086 (0.021)
LAKE_KOZART	-0.279 (0.018)	-0.277 (0.018)	-0.219 (0.018)	-0.275 (0.018)	-0.275 (0.018)	-0.269 (0.019)
LAKE_LANCASTER	0.032 (0.019)	0.033 (0.019)	0.063 (0.019)	0.036 (0.019)	0.036 (0.019)	0.041 (0.019)
LAKE_LAWSONA	0.165 (0.026)	0.172 (0.026)	0.194 (0.025)	0.175 (0.026)	0.175 (0.026)	0.181 (0.026)
LAKE_LOUISE_B	0.417 (0.024)	0.414 (0.024)	0.352 (0.023)	0.416 (0.024)	0.417 (0.024)	0.418 (0.024)
LAKE_LOVE	0.321 (0.047)	0.339 (0.047)	0.325 (0.047)	0.339 (0.047)	0.339 (0.047)	0.346 (0.048)
LAKE_LOVELY	-0.256 (0.024)	-0.247 (0.024)	-0.172 (0.024)	-0.247 (0.024)	-0.245 (0.024)	-0.235 (0.025)
LAKE_LURNA	0.058 (0.025)	0.066 (0.025)	0.096 (0.024)	0.068 (0.025)	0.069 (0.025)	0.073 (0.025)
LAKE_MABEL	0.164 (0.024)	0.158 (0.024)	0.116 (0.024)	0.160 (0.024)	0.160 (0.024)	0.158 (0.024)
LAKE_MAITLAND	0.334 (0.019)	0.336 (0.020)	0.318 (0.022)	0.336 (0.020)	0.337 (0.020)	0.344 (0.022)
LAKE_MANN	-0.390 (0.020)	-0.390 (0.021)	-0.315 (0.019)	-0.387 (0.021)	-0.385 (0.021)	-0.378 (0.021)
LAKE_MARSHA	0.016 (0.014)	0.015 (0.015)	0.051 (0.014)	0.016 (0.015)	0.015 (0.014)	0.015 (0.014)
LAKE_MINNEHAHA	0.155 (0.023)	0.159 (0.024)	0.167 (0.026)	0.160 (0.024)	0.161 (0.024)	0.171 (0.026)
LAKE_MIZELL	0.358 (0.028)	0.359 (0.029)	0.347 (0.030)	0.361 (0.029)	0.362 (0.029)	0.369 (0.030)
LAKE_NAN	0.039 (0.030)	0.042 (0.031)	0.083 (0.033)	0.046 (0.031)	0.047 (0.031)	0.061 (0.034)
LAKE_NONA	0.438 (0.026)	0.419 (0.025)	0.348 (0.023)	0.424 (0.025)	0.425 (0.025)	0.429 (0.024)
LAKE_OFWOODS	-0.652	-0.653	-0.569	-0.648	-0.648	-0.630

	(0.052)	(0.052)	(0.050)	(0.052)	(0.052)	(0.052)
LAKE_OLIVE	-0.059	-0.054	-0.009	-0.049	-0.049	-0.044
	(0.037)	(0.037)	(0.036)	(0.037)	(0.037)	(0.037)
LAKE_OLYMPIA	-0.247	-0.242	-0.175	-0.242	-0.242	-0.239
	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)
LAKE_ORLANDO	-0.224	-0.219	-0.174	-0.219	-0.219	-0.211
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)
LAKE_OSCEOLA	0.380	0.384	0.363	0.384	0.385	0.393
	(0.019)	(0.020)	(0.022)	(0.020)	(0.020)	(0.022)
LAKE_PAMELA	-0.306	-0.297	-0.231	-0.295	-0.295	-0.290
	(0.038)	(0.038)	(0.037)	(0.038)	(0.038)	(0.038)
LAKE_PEARL_B	-0.154	-0.144	-0.084	-0.144	-0.144	-0.141
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)
LAKE_PICKETT	0.265	0.256	0.581	0.268	0.280	0.326
	(0.051)	(0.053)	(0.048)	(0.052)	(0.053)	(0.056)
LAKE_PINELOCH	-0.076	-0.077	-0.046	-0.074	-0.074	-0.069
	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)
LAKE_PORTER	-0.037	-0.037	-0.007	-0.035	-0.033	-0.027
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
LAKE_RABAMA	-0.041	-0.037	-0.007	-0.034	-0.033	-0.026
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)
LAKE_RICHMOND	-0.292	-0.291	-0.226	-0.287	-0.288	-0.283
	(0.021)	(0.021)	(0.020)	(0.021)	(0.021)	(0.021)
LAKE_ROBERTS	0.037	0.042	0.059	0.043	0.042	0.044
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
LAKE_ROSE_B	-0.112	-0.108	-0.064	-0.108	-0.107	-0.105
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)
LAKE_ROWENA	0.021	0.024	0.050	0.026	0.027	0.033
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
LAKE_SANTIAGO	-0.055	-0.053	-0.024	-0.050	-0.049	-0.042
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
LAKE_SARAH	-0.067	-0.059	-0.038	-0.058	-0.057	-0.054
	(0.043)	(0.043)	(0.042)	(0.043)	(0.043)	(0.043)
LAKE_SHADOW	-0.126	-0.115	-0.070	-0.115	-0.114	-0.105
	(0.024)	(0.025)	(0.025)	(0.024)	(0.025)	(0.026)
LAKE_SHANNON	-0.015	-0.010	0.016	-0.008	-0.006	0.002
	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)	(0.026)
LAKE_SHEEN	0.208	0.201	0.157	0.202	0.201	0.196
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
LAKE_SHERWOOD	-0.172	-0.168	-0.104	-0.169	-0.169	-0.166
	(0.015)	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)
LAKE_SILVER	0.046	0.051	0.082	0.052	0.053	0.060
	(0.018)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)
LAKE_STARKE	-0.292	-0.287	-0.204	-0.287	-0.287	-0.281
	(0.017)	(0.017)	(0.015)	(0.017)	(0.017)	(0.016)
LAKE_SUE	0.134	0.135	0.143	0.136	0.137	0.143
	(0.018)	(0.019)	(0.020)	(0.019)	(0.019)	(0.020)
LAKE_SUNSET	-0.445	-0.443	-0.357	-0.437	-0.434	-0.419
	(0.034)	(0.034)	(0.033)	(0.034)	(0.034)	(0.034)

LAKE_SUSANNAH	-0.052 (0.018)	-0.057 (0.019)	-0.046 (0.020)	-0.054 (0.019)	-0.052 (0.019)	-0.044 (0.020)
LAKE_SYBELIA	0.008 (0.028)	0.010 (0.029)	0.024 (0.029)	0.011 (0.029)	0.012 (0.029)	0.023 (0.030)
LAKE_SYLVAN	0.328 (0.025)	0.328 (0.026)	0.308 (0.027)	0.329 (0.026)	0.331 (0.026)	0.340 (0.028)
LAKE_TENNESSEE	-0.114 (0.022)	-0.110 (0.022)	-0.066 (0.022)	-0.106 (0.022)	-0.106 (0.022)	-0.098 (0.023)
LAKE_TERRACE	-0.111 (0.023)	-0.104 (0.023)	-0.062 (0.022)	-0.100 (0.023)	-0.099 (0.023)	-0.092 (0.023)
LAKE_THERESA	0.046 (0.033)	0.057 (0.034)	0.090 (0.033)	0.060 (0.034)	0.062 (0.034)	0.072 (0.035)
LAKE_TIBET	0.260 (0.015)	0.253 (0.015)	0.202 (0.015)	0.254 (0.015)	0.252 (0.015)	0.246 (0.016)
LAKE_UNDERHILL	-0.106 (0.020)	-0.108 (0.020)	-0.070 (0.020)	-0.105 (0.020)	-0.104 (0.020)	-0.099 (0.021)
LAKE_VIRGINIA	0.301 (0.019)	0.303 (0.019)	0.292 (0.020)	0.303 (0.019)	0.304 (0.019)	0.308 (0.021)
LAKE_WADE	-0.046 (0.023)	-0.040 (0.023)	0.002 (0.022)	-0.037 (0.023)	-0.036 (0.023)	-0.030 (0.023)
LAKE_WALKER	-0.338 (0.030)	-0.335 (0.030)	-0.264 (0.029)	-0.332 (0.030)	-0.332 (0.030)	-0.321 (0.030)
LAKE_WARREN	-0.021 (0.021)	-0.028 (0.021)	-0.027 (0.020)	-0.025 (0.021)	-0.024 (0.021)	-0.019 (0.020)
LAKE_WAUNATTA	0.099 (0.023)	0.100 (0.024)	0.126 (0.026)	0.101 (0.024)	0.102 (0.024)	0.111 (0.027)
LAKE_WELDONA	-0.002 (0.026)	0.001 (0.026)	0.033 (0.025)	0.003 (0.026)	0.004 (0.026)	0.010 (0.026)
LAKE_WESTON	-0.302 (0.024)	-0.297 (0.024)	-0.217 (0.024)	-0.297 (0.024)	-0.296 (0.024)	-0.285 (0.025)
LAKE_WHIPPOOR	0.152 (0.024)	0.135 (0.023)	0.103 (0.020)	0.141 (0.023)	0.144 (0.023)	0.154 (0.021)
LAKE_WINYAH	0.144 (0.020)	0.151 (0.021)	0.168 (0.021)	0.151 (0.021)	0.152 (0.021)	0.159 (0.022)
LAWNE_LAKE	-0.276 (0.016)	-0.273 (0.017)	-0.212 (0.016)	-0.274 (0.017)	-0.273 (0.017)	-0.267 (0.017)
LITTLE_FISH	0.511 (0.021)	0.517 (0.021)	0.464 (0.021)	0.518 (0.021)	0.517 (0.021)	0.511 (0.022)
LIT_LK_FAIR	-0.107 (0.022)	-0.105 (0.022)	-0.035 (0.022)	-0.105 (0.022)	-0.103 (0.022)	-0.097 (0.023)
LITTLE_SAND	0.251 (0.021)	0.247 (0.021)	0.261 (0.021)	0.248 (0.021)	0.248 (0.021)	0.246 (0.021)
LONG_LAKE	-0.328 (0.021)	-0.322 (0.021)	-0.240 (0.019)	-0.323 (0.021)	-0.322 (0.021)	-0.315 (0.021)
MUD_LAKE_C	0.164 (0.042)	0.171 (0.042)	0.164 (0.042)	0.174 (0.042)	0.175 (0.042)	0.181 (0.043)
PALM_LAKE	0.096 (0.017)	0.104 (0.017)	0.106 (0.017)	0.105 (0.017)	0.104 (0.017)	0.103 (0.017)
PARK_LAKE_B	0.084	0.090	0.118	0.093	0.093	0.098

	(0.032)	(0.032)	(0.031)	(0.032)	(0.032)	(0.032)
POCKET_LAKE	0.290	0.292	0.253	0.293	0.292	0.286
	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.019)
ROCK_LAKE	-0.446	-0.443	-0.373	-0.441	-0.439	-0.433
	(0.026)	(0.026)	(0.025)	(0.026)	(0.026)	(0.027)
SPRING_LK_B	0.019	0.022	0.048	0.025	0.026	0.031
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
SPRING_LK_C	0.199	0.197	0.196	0.200	0.199	0.198
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)
rho	-0.003	0.000	0.109	0.001	0.003	0.008
	(0.000)	(0.000)	(0.008)	(0.000)	(0.000)	(0.001)

Standard errors appear in parentheses.

Table 19: Full Spatial Error Results

	SemNN15	SemNN20	Sem200	Sem20063	Sem200123	Sem40063
Constant	33.395	14.478	26.185	49.449	43.610	16.605
	(17.489)	(2.782)	(9.764)	(19.075)	(17.782)	(5.352)
Ln(SDM)	-0.005	-0.011	-0.078	0.024	0.064	0.006
	(0.030)	(0.030)	(0.028)	(0.037)	(0.038)	(0.042)
ln(SDM)*Waterfront	0.061	0.061	0.034	0.058	0.060	0.067
	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Ln(SDM)*ln(dist)	-0.011	-0.010	-0.002	-0.020	-0.021	-0.019
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(SDM)*ln(lake area)	0.006	0.006	0.008	0.008	0.005	0.009
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Waterfront	0.157	0.154	0.197	0.171	0.165	0.157
	(0.019)	(0.019)	(0.021)	(0.019)	(0.019)	(0.019)
Ln(lake dist)	-0.066	-0.068	-0.078	-0.042	-0.046	-0.044
	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Canalfront	0.070	0.067	0.066	0.076	0.069	0.088
	(0.018)	(0.018)	(0.021)	(0.017)	(0.017)	(0.017)
Golffront	0.231	0.221	0.236	0.280	0.217	0.245
	(0.031)	(0.031)	(0.038)	(0.030)	(0.030)	(0.030)
Ln(bath)	0.089	0.092	0.098	0.104	0.098	0.101
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Pool	0.008	0.007	0.007	0.009	0.007	0.009
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ln(age)	-0.085	-0.086	-0.086	-0.079	-0.081	-0.081
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(area_heated)	0.557	0.563	0.539	0.603	0.575	0.584
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Ln(area parcel)	0.172	0.170	0.179	0.173	0.174	0.172
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(dist CBD)	-0.203	-0.166	-0.192	-0.188	-0.190	-0.187
	(0.015)	(0.015)	(0.018)	(0.013)	(0.014)	(0.014)
Near airport	0.030	0.026	0.010	0.020	0.015	0.014
	(0.008)	(0.008)	(0.009)	(0.007)	(0.008)	(0.007)
ln(x_coord)	-4.378	-3.485	-4.096	-5.358	-5.066	-2.774
	(0.018)	(0.770)	(0.413)	(0.673)	(0.440)	(0.497)
ln(y_coord)	2.530	3.038	2.785	2.247	2.410	2.134
	(1.319)	(0.994)	(1.162)	(0.765)	(0.905)	(0.912)
Ln(Med income)	0.121	0.117	0.127	0.121	0.125	0.118
	(0.007)	(0.007)	(0.009)	(0.006)	(0.007)	(0.007)
Percent white00	0.184	0.189	0.237	0.266	0.284	0.224
	(0.049)	(0.049)	(0.057)	(0.040)	(0.044)	(0.044)
Percent black00	-0.058	-0.054	0.036	0.040	0.042	0.012
	(0.053)	(0.053)	(0.060)	(0.042)	(0.047)	(0.047)
Percent over6500	0.091	0.101	0.043	0.050	0.039	0.045
	(0.027)	(0.027)	(0.033)	(0.024)	(0.026)	(0.025)
year_1996	-0.838	-0.836	-0.866	-0.816	-0.812	-0.804
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)
year_1997	-0.780	-0.780	-0.803	-0.759	-0.756	-0.754

	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)
year_1998	-0.712	-0.712	-0.730	-0.687	-0.688	-0.682
	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)
year_1999	-0.621	-0.621	-0.638	-0.601	-0.600	-0.596
	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)
year_2000	-0.513	-0.511	-0.526	-0.494	-0.493	-0.489
	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)
year_2001	-0.397	-0.393	-0.401	-0.380	-0.378	-0.379
	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)
year_2002	-0.295	-0.286	-0.293	-0.279	-0.276	-0.271
	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.006)
year_2003	-0.162	-0.156	-0.167	-0.153	-0.155	-0.148
	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)	(0.006)
BASS_LAKE	-0.206	-0.194	-0.166	-0.143	-0.168	-0.190
	(0.027)	(0.031)	(0.030)	(0.026)	(0.027)	(0.029)
BAY_LAKE_B	-0.347	-0.300	-0.265	-0.252	-0.254	-0.262
	(0.055)	(0.051)	(0.063)	(0.040)	(0.045)	(0.044)
BEARHEAD_LAKE	-0.176	-0.170	-0.123	-0.124	-0.145	-0.147
	(0.038)	(0.043)	(0.046)	(0.031)	(0.033)	(0.039)
BIG_SAND_LAKE	0.184	0.156	0.185	0.157	0.177	0.170
	(0.026)	(0.022)	(0.026)	(0.020)	(0.023)	(0.023)
CLEAR_LAKE	-0.452	-0.443	-0.427	-0.380	-0.386	-0.415
	(0.027)	(0.029)	(0.032)	(0.023)	(0.025)	(0.027)
DEEP_LAKE	-0.086	-0.132	-0.052	-0.029	-0.047	-0.069
	(0.052)	(0.040)	(0.048)	(0.055)	(0.054)	(0.041)
KASEY_LAKE	-0.263	-0.268	-0.242	-0.200	-0.227	-0.229
	(0.038)	(0.031)	(0.036)	(0.031)	(0.034)	(0.033)
KELLY_LAKE	-0.292	-0.328	-0.268	-0.237	-0.269	-0.245
	(0.038)	(0.032)	(0.036)	(0.034)	(0.037)	(0.038)
KRISTY_LAKE	-0.320	-0.330	-0.284	-0.250	-0.287	-0.250
	(0.038)	(0.033)	(0.036)	(0.034)	(0.036)	(0.037)
LAKE_ADAIR	0.051	0.055	0.076	0.109	0.084	0.077
	(0.032)	(0.030)	(0.037)	(0.031)	(0.033)	(0.031)
LAKE_ANDERSON	-0.131	-0.156	-0.113	-0.083	-0.109	-0.124
	(0.025)	(0.031)	(0.031)	(0.026)	(0.026)	(0.029)
LAKE_ANGEL	-0.557	-0.564	-0.493	-0.465	-0.490	-0.504
	(0.038)	(0.040)	(0.043)	(0.032)	(0.035)	(0.039)
LAKE_ARNOLD	-0.185	-0.206	-0.176	-0.135	-0.156	-0.167
	(0.027)	(0.030)	(0.033)	(0.029)	(0.029)	(0.029)
LAKE_BALDWIN	0.000	-0.055	-0.019	0.010	0.005	-0.048
	(0.033)	(0.028)	(0.032)	(0.033)	(0.033)	(0.025)
LAKE_BARTON	-0.293	-0.342	-0.268	-0.225	-0.239	-0.266
	(0.027)	(0.026)	(0.029)	(0.030)	(0.029)	(0.024)
LAKE_BEARDALL	-0.519	-0.593	-0.555	-0.472	-0.501	-0.515
	(0.079)	(0.077)	(0.090)	(0.059)	(0.064)	(0.069)
LAKE_BEAUTY	-0.215	-0.110	-0.080	-0.085	-0.109	-0.084
	(0.053)	(0.055)	(0.072)	(0.058)	(0.062)	(0.065)
LAKE_BELL	-0.195	-0.262	-0.235	-0.188	-0.214	-0.205
	(0.043)	(0.034)	(0.043)	(0.038)	(0.041)	(0.035)

LAKE_BERRY	0.129	0.065	0.129	0.164	0.146	0.134
	(0.035)	(0.028)	(0.033)	(0.036)	(0.036)	(0.028)
LAKE_BESSIE	0.471	0.443	0.492	0.569	0.545	0.508
	(0.029)	(0.029)	(0.040)	(0.027)	(0.029)	(0.029)
LAKE_BLANCHE	0.093	0.060	0.105	0.090	0.089	0.104
	(0.019)	(0.019)	(0.025)	(0.018)	(0.020)	(0.021)
LAKE_BUCHANNAN	-0.254	-0.272	-0.172	-0.155	-0.173	-0.218
	(0.048)	(0.050)	(0.046)	(0.036)	(0.039)	(0.041)
LAKE_BUCK	0.671	0.635	0.444	0.613	0.637	0.565
	(0.047)	(0.062)	(0.052)	(0.028)	(0.028)	(0.045)
LAKE_BUMBY	-0.287	-0.270	-0.169	-0.154	-0.194	-0.233
	(0.047)	(0.049)	(0.045)	(0.037)	(0.039)	(0.042)
LAKE_BURKETT	-0.032	-0.074	-0.004	0.041	0.022	-0.025
	(0.043)	(0.030)	(0.038)	(0.047)	(0.045)	(0.031)
LAKE_BUTLER	0.090	0.101	0.022	0.007	0.052	0.062
	(0.021)	(0.022)	(0.025)	(0.022)	(0.023)	(0.024)
LAKE_C	-0.242	-0.279	-0.189	-0.152	-0.179	-0.197
	(0.029)	(0.031)	(0.033)	(0.033)	(0.033)	(0.032)
LAKE_CANE_A	-0.039	-0.045	-0.029	-0.039	-0.042	-0.032
	(0.020)	(0.020)	(0.024)	(0.018)	(0.020)	(0.021)
LAKE_CATHERINE_B	0.187	0.077	0.151	0.152	0.144	0.107
	(0.066)	(0.058)	(0.071)	(0.055)	(0.058)	(0.054)
LAKE_CAY_DEE	-0.071	-0.161	-0.082	-0.064	-0.076	-0.048
	(0.034)	(0.032)	(0.037)	(0.037)	(0.038)	(0.034)
LAKE_CHARITY	0.202	0.177	0.173	0.157	0.149	0.122
	(0.057)	(0.047)	(0.059)	(0.049)	(0.053)	(0.047)
LAKE_CHASE	0.501	0.469	0.417	0.523	0.506	0.446
	(0.038)	(0.037)	(0.047)	(0.034)	(0.036)	(0.034)
LAKE_CHEROKEE	-0.021	0.027	0.058	0.071	0.046	0.035
	(0.039)	(0.041)	(0.049)	(0.037)	(0.040)	(0.041)
LAKE_CHRISTIE	-0.256	-0.237	-0.212	-0.177	-0.211	-0.224
	(0.078)	(0.083)	(0.068)	(0.043)	(0.050)	(0.056)
LAKE_COMO	-0.057	-0.068	-0.038	0.021	-0.018	-0.017
	(0.033)	(0.035)	(0.038)	(0.037)	(0.037)	(0.039)
LAKE_CONCORD	-0.093	-0.061	-0.047	-0.008	-0.032	-0.042
	(0.038)	(0.038)	(0.048)	(0.037)	(0.040)	(0.038)
LAKE_CONWAY	-0.130	-0.154	-0.119	-0.113	-0.098	-0.169
	(0.020)	(0.028)	(0.026)	(0.019)	(0.019)	(0.026)
LAKE_COPELAND	-0.008	-0.045	-0.064	0.004	-0.023	-0.028
	(0.042)	(0.043)	(0.050)	(0.038)	(0.041)	(0.040)
LAKE_DANIEL	-0.139	-0.184	-0.143	-0.101	-0.123	-0.092
	(0.045)	(0.041)	(0.049)	(0.040)	(0.043)	(0.040)
LAKE_DESTINY	0.248	0.178	0.261	0.290	0.266	0.265
	(0.055)	(0.042)	(0.058)	(0.049)	(0.052)	(0.048)
LAKE_DOT	-0.455	-0.450	-0.383	-0.430	-0.443	-0.422
	(0.051)	(0.051)	(0.060)	(0.047)	(0.050)	(0.050)
LAKE_DOVER	-0.128	-0.144	-0.064	-0.069	-0.099	-0.139
	(0.054)	(0.054)	(0.050)	(0.051)	(0.053)	(0.053)
LAKE_DOWN	0.122	0.105	0.069	0.049	0.075	0.078

	(0.017)	(0.018)	(0.021)	(0.017)	(0.018)	(0.021)
LAKE_DOWNEY	-0.057	-0.141	-0.057	0.000	-0.037	-0.086
	(0.034)	(0.029)	(0.027)	(0.043)	(0.039)	(0.028)
LAKE_DRUID	-0.042	-0.090	-0.058	-0.004	-0.029	-0.040
	(0.030)	(0.029)	(0.034)	(0.032)	(0.033)	(0.031)
LAKE_EMERALD	0.005	-0.084	0.009	0.063	0.021	-0.055
	(0.049)	(0.047)	(0.043)	(0.050)	(0.046)	(0.036)
LAKE_EOLA	-0.132	-0.163	-0.087	-0.078	-0.093	-0.113
	(0.047)	(0.048)	(0.059)	(0.044)	(0.047)	(0.047)
LAKE_ESTELLE	-0.022	-0.066	-0.011	0.035	0.011	-0.022
	(0.039)	(0.036)	(0.044)	(0.039)	(0.040)	(0.035)
LAKE_FAIRHOPE	-0.158	-0.155	-0.070	-0.038	-0.070	-0.086
	(0.050)	(0.045)	(0.051)	(0.044)	(0.046)	(0.045)
LAKE_FAIRVIEW	-0.302	-0.334	-0.291	-0.267	-0.271	-0.260
	(0.032)	(0.024)	(0.032)	(0.028)	(0.030)	(0.025)
LAKE_FAITH	0.059	0.003	-0.007	0.015	-0.021	0.005
	(0.061)	(0.050)	(0.061)	(0.050)	(0.053)	(0.047)
LAKE_FARRAR	-0.156	-0.162	-0.073	-0.069	-0.096	-0.109
	(0.031)	(0.035)	(0.036)	(0.034)	(0.034)	(0.036)
LAKE_FORMOSA	-0.056	-0.127	-0.066	-0.018	-0.038	-0.037
	(0.036)	(0.034)	(0.044)	(0.036)	(0.038)	(0.036)
LAKE_FREDRICA	0.103	0.055	0.129	0.160	0.154	0.092
	(0.026)	(0.032)	(0.033)	(0.030)	(0.029)	(0.031)
LAKE_GATLIN	-0.020	-0.019	0.027	0.025	0.012	-0.035
	(0.035)	(0.040)	(0.043)	(0.030)	(0.033)	(0.035)
LAKE_GEAR	-0.156	-0.215	-0.054	-0.009	-0.031	-0.064
	(0.047)	(0.046)	(0.046)	(0.044)	(0.045)	(0.042)
LAKE_GEM_A	-0.031	-0.105	-0.072	-0.035	-0.054	-0.026
	(0.043)	(0.032)	(0.042)	(0.041)	(0.043)	(0.036)
LAKE_GEM_MARY	-0.091	-0.115	-0.079	-0.038	-0.064	-0.087
	(0.030)	(0.035)	(0.038)	(0.030)	(0.031)	(0.034)
LAKE_GEORGE	-0.064	-0.066	-0.027	-0.013	-0.017	-0.070
	(0.021)	(0.029)	(0.028)	(0.023)	(0.022)	(0.026)
LAKE_GEORGIA	0.015	-0.064	0.022	0.067	0.040	-0.029
	(0.038)	(0.022)	(0.027)	(0.045)	(0.042)	(0.022)
LAKE_GILES	-0.212	-0.231	-0.189	-0.155	-0.177	-0.197
	(0.030)	(0.033)	(0.035)	(0.030)	(0.031)	(0.031)
LAKE_GLORIA	-0.140	-0.185	-0.099	-0.098	-0.115	-0.148
	(0.036)	(0.041)	(0.036)	(0.024)	(0.026)	(0.033)
LAKE_GREENWOOD	-0.143	-0.142	-0.085	-0.042	-0.083	-0.062
	(0.043)	(0.045)	(0.054)	(0.044)	(0.047)	(0.048)
LAKE_HART	0.216	0.137	0.121	0.166	0.155	0.085
	(0.044)	(0.063)	(0.052)	(0.025)	(0.025)	(0.048)
LAKE_HIAWASSEE	-0.142	-0.150	-0.138	-0.128	-0.135	-0.130
	(0.016)	(0.016)	(0.020)	(0.014)	(0.016)	(0.017)
LAKE_HIGHLAND	-0.027	-0.053	-0.024	0.015	0.003	-0.009
	(0.033)	(0.033)	(0.040)	(0.033)	(0.035)	(0.033)
LAKE_HOLDEN	-0.413	-0.439	-0.402	-0.367	-0.380	-0.389
	(0.024)	(0.027)	(0.028)	(0.021)	(0.022)	(0.025)

LAKE_HOPE	0.078 (0.085)	0.074 (0.074)	-0.027 (0.081)	-0.101 (0.060)	-0.105 (0.061)	-0.073 (0.057)
LAKE_HUNGERFORD	-0.316 (0.051)	-0.401 (0.042)	-0.288 (0.051)	-0.325 (0.043)	-0.324 (0.045)	-0.299 (0.040)
LAKE_IRMA	-0.044 (0.036)	-0.082 (0.023)	-0.035 (0.025)	0.011 (0.040)	-0.013 (0.037)	-0.068 (0.021)
LAKE_ISLEWORTH	0.568 (0.053)	0.521 (0.052)	0.357 (0.062)	0.450 (0.047)	0.464 (0.047)	0.431 (0.047)
LAKE_IVANHOE	0.022 (0.029)	0.003 (0.026)	0.020 (0.032)	0.056 (0.028)	0.036 (0.029)	0.020 (0.026)
LAKE_JACKSON	0.062 (0.050)	0.001 (0.039)	0.134 (0.054)	0.092 (0.044)	0.041 (0.046)	0.025 (0.042)
LK_JEN_JEWEL	-0.158 (0.031)	-0.179 (0.035)	-0.191 (0.037)	-0.147 (0.027)	-0.162 (0.028)	-0.205 (0.031)
LAKE_JESSAMINE	-0.311 (0.021)	-0.317 (0.026)	-0.287 (0.026)	-0.271 (0.016)	-0.270 (0.018)	-0.295 (0.023)
LAKE_KILLARNEY	-0.144 (0.033)	-0.207 (0.023)	-0.156 (0.031)	-0.129 (0.030)	-0.130 (0.032)	-0.139 (0.024)
LAKE_KOZART	-0.362 (0.028)	-0.390 (0.028)	-0.359 (0.031)	-0.315 (0.024)	-0.333 (0.027)	-0.345 (0.028)
LAKE_LANCASTER	-0.053 (0.027)	-0.080 (0.030)	-0.065 (0.033)	-0.016 (0.027)	-0.030 (0.028)	-0.051 (0.029)
LAKE_LAWSONA	0.105 (0.034)	0.026 (0.036)	0.086 (0.042)	0.125 (0.035)	0.096 (0.037)	0.085 (0.037)
LAKE_LOUISE_B	0.431 (0.035)	0.428 (0.034)	0.488 (0.041)	0.499 (0.030)	0.525 (0.032)	0.537 (0.032)
LAKE_LOVE	0.281 (0.065)	0.270 (0.055)	0.330 (0.072)	0.315 (0.060)	0.281 (0.064)	0.227 (0.060)
LAKE_LOVELY	-0.262 (0.042)	-0.303 (0.030)	-0.281 (0.037)	-0.267 (0.033)	-0.286 (0.036)	-0.289 (0.031)
LAKE_LURNA	-0.049 (0.032)	-0.073 (0.035)	-0.042 (0.041)	-0.004 (0.033)	-0.034 (0.035)	-0.056 (0.036)
LAKE_MABEL	0.212 (0.043)	0.159 (0.040)	0.193 (0.042)	0.136 (0.033)	0.141 (0.035)	0.158 (0.034)
LAKE_MAITLAND	0.341 (0.035)	0.244 (0.018)	0.328 (0.028)	0.338 (0.035)	0.340 (0.034)	0.296 (0.021)
LAKE_MANN	-0.464 (0.032)	-0.501 (0.032)	-0.457 (0.036)	-0.409 (0.026)	-0.419 (0.028)	-0.441 (0.029)
LAKE_MARSHA	0.017 (0.025)	0.011 (0.025)	0.048 (0.026)	0.003 (0.019)	0.005 (0.021)	0.029 (0.023)
LAKE_MINNEHAHA	0.132 (0.042)	0.077 (0.025)	0.151 (0.036)	0.161 (0.039)	0.151 (0.040)	0.144 (0.028)
LAKE_MIZELL	0.257 (0.042)	0.248 (0.033)	0.308 (0.041)	0.337 (0.041)	0.315 (0.042)	0.279 (0.033)
LAKE_NAN	-0.049 (0.051)	-0.097 (0.040)	-0.041 (0.047)	0.020 (0.049)	-0.008 (0.050)	-0.045 (0.037)
LAKE_NONA	0.567 (0.048)	0.517 (0.061)	0.360 (0.056)	0.426 (0.033)	0.436 (0.035)	0.376 (0.050)
LAKE_OFWOODS	-0.671	-0.771	-0.667	-0.669	-0.702	-0.701

	(0.078)	(0.077)	(0.089)	(0.063)	(0.072)	(0.069)
LAKE_OLIVE	-0.099	-0.135	-0.104	-0.082	-0.107	-0.131
	(0.048)	(0.049)	(0.058)	(0.045)	(0.048)	(0.048)
LAKE_OLYMPIA	-0.246	-0.260	-0.230	-0.237	-0.245	-0.224
	(0.032)	(0.032)	(0.031)	(0.019)	(0.022)	(0.027)
LAKE_ORLANDO	-0.250	-0.262	-0.227	-0.206	-0.217	-0.229
	(0.034)	(0.027)	(0.034)	(0.027)	(0.030)	(0.028)
LAKE_OSCEOLA	0.327	0.268	0.377	0.394	0.387	0.349
	(0.034)	(0.019)	(0.029)	(0.035)	(0.034)	(0.022)
LAKE_PAMELA	-0.337	-0.452	-0.315	-0.332	-0.331	-0.383
	(0.053)	(0.052)	(0.070)	(0.049)	(0.056)	(0.054)
LAKE_PEARL_B	-0.153	-0.185	-0.133	-0.162	-0.179	-0.137
	(0.024)	(0.026)	(0.027)	(0.019)	(0.021)	(0.026)
LAKE_PICKETT	0.089	-0.003	0.211	0.208	0.202	0.022
	(0.116)	(0.117)	(0.067)	(0.073)	(0.068)	(0.054)
LAKE_PINELOCH	-0.151	-0.219	-0.134	-0.126	-0.143	-0.159
	(0.024)	(0.029)	(0.030)	(0.023)	(0.024)	(0.027)
LAKE_PORTER	-0.161	-0.169	-0.102	-0.078	-0.099	-0.132
	(0.023)	(0.028)	(0.028)	(0.027)	(0.026)	(0.027)
LAKE_RABAMA	-0.162	-0.200	-0.121	-0.077	-0.108	-0.137
	(0.029)	(0.033)	(0.034)	(0.033)	(0.033)	(0.035)
LAKE_RICHMOND	-0.452	-0.395	-0.357	-0.320	-0.321	-0.353
	(0.034)	(0.034)	(0.038)	(0.026)	(0.030)	(0.032)
LAKE_ROBERTS	0.123	0.120	0.121	0.069	0.072	0.092
	(0.032)	(0.035)	(0.033)	(0.023)	(0.025)	(0.029)
LAKE_ROSE_B	-0.120	-0.132	-0.098	-0.116	-0.125	-0.101
	(0.022)	(0.022)	(0.024)	(0.016)	(0.018)	(0.021)
LAKE_ROWENA	-0.045	-0.089	-0.025	0.014	-0.004	-0.012
	(0.031)	(0.029)	(0.035)	(0.032)	(0.033)	(0.029)
LAKE_SANTIAGO	-0.181	-0.227	-0.130	-0.112	-0.134	-0.167
	(0.030)	(0.033)	(0.035)	(0.033)	(0.033)	(0.033)
LAKE_SARAH	-0.134	-0.201	-0.126	-0.092	-0.135	-0.144
	(0.056)	(0.052)	(0.064)	(0.050)	(0.052)	(0.048)
LAKE_SHADOW	-0.169	-0.226	-0.216	-0.152	-0.181	-0.201
	(0.041)	(0.029)	(0.038)	(0.035)	(0.037)	(0.031)
LAKE_SHANNON	-0.132	-0.181	-0.099	-0.058	-0.085	-0.083
	(0.033)	(0.030)	(0.033)	(0.037)	(0.037)	(0.033)
LAKE_SHEEN	0.307	0.266	0.252	0.225	0.249	0.254
	(0.027)	(0.023)	(0.029)	(0.023)	(0.025)	(0.024)
LAKE_SHERWOOD	-0.187	-0.191	-0.165	-0.172	-0.183	-0.176
	(0.026)	(0.025)	(0.026)	(0.018)	(0.021)	(0.023)
LAKE_SILVER	-0.033	-0.041	-0.013	0.027	0.010	-0.001
	(0.028)	(0.024)	(0.030)	(0.027)	(0.028)	(0.025)
LAKE_STARKE	-0.279	-0.312	-0.289	-0.304	-0.319	-0.271
	(0.032)	(0.033)	(0.033)	(0.018)	(0.022)	(0.028)
LAKE_SUE	0.083	0.037	0.089	0.123	0.111	0.074
	(0.029)	(0.024)	(0.030)	(0.030)	(0.030)	(0.024)
LAKE_SUNSET	-0.570	-0.558	-0.488	-0.461	-0.483	-0.471
	(0.052)	(0.051)	(0.058)	(0.040)	(0.043)	(0.045)

LAKE_SUSANNAH	-0.207 (0.031)	-0.248 (0.028)	-0.111 (0.029)	-0.094 (0.031)	-0.098 (0.031)	-0.141 (0.025)
LAKE_SYBELIA	0.028 (0.049)	-0.053 (0.038)	-0.013 (0.045)	0.011 (0.040)	-0.012 (0.043)	-0.011 (0.036)
LAKE_SYLVAN	0.226 (0.040)	0.201 (0.030)	0.278 (0.036)	0.294 (0.040)	0.273 (0.041)	0.230 (0.032)
LAKE_TENNESSEE	-0.219 (0.028)	-0.240 (0.033)	-0.197 (0.035)	-0.174 (0.031)	-0.196 (0.031)	-0.204 (0.033)
LAKE_TERRACE	-0.221 (0.028)	-0.228 (0.032)	-0.186 (0.033)	-0.165 (0.032)	-0.194 (0.032)	-0.193 (0.034)
LAKE_THERESA	-0.111 (0.041)	-0.162 (0.040)	-0.085 (0.043)	-0.021 (0.046)	-0.075 (0.047)	-0.053 (0.045)
LAKE_TIBET	0.345 (0.022)	0.300 (0.020)	0.325 (0.024)	0.284 (0.021)	0.321 (0.022)	0.308 (0.022)
LAKE_UNDERHILL	-0.212 (0.029)	-0.231 (0.031)	-0.207 (0.035)	-0.152 (0.029)	-0.171 (0.030)	-0.197 (0.029)
LAKE_VIRGINIA	0.252 (0.030)	0.193 (0.023)	0.237 (0.031)	0.280 (0.032)	0.259 (0.032)	0.207 (0.025)
LAKE_WADE	-0.114 (0.029)	-0.171 (0.033)	-0.146 (0.037)	-0.092 (0.030)	-0.119 (0.032)	-0.109 (0.034)
LAKE_WALKER	-0.443 (0.043)	-0.466 (0.043)	-0.480 (0.050)	-0.386 (0.037)	-0.404 (0.040)	-0.421 (0.041)
LAKE_WARREN	-0.145 (0.033)	-0.178 (0.040)	-0.117 (0.039)	-0.081 (0.028)	-0.096 (0.030)	-0.117 (0.036)
LAKE_WAUNATTA	0.040 (0.043)	-0.018 (0.031)	0.044 (0.033)	0.076 (0.043)	0.054 (0.042)	0.003 (0.028)
LAKE_WELDONA	-0.099 (0.034)	-0.150 (0.036)	-0.074 (0.041)	-0.060 (0.033)	-0.076 (0.035)	-0.091 (0.035)
LAKE_WESTON	-0.325 (0.043)	-0.354 (0.034)	-0.319 (0.038)	-0.309 (0.034)	-0.328 (0.036)	-0.325 (0.033)
LAKE_WHIPPOOR	0.092 (0.048)	-0.020 (0.064)	0.014 (0.055)	0.056 (0.028)	0.046 (0.029)	-0.023 (0.049)
LAKE_WINYAH	0.066 (0.030)	0.022 (0.024)	0.060 (0.031)	0.104 (0.031)	0.080 (0.032)	0.063 (0.028)
LAWNE_LAKE	-0.326 (0.026)	-0.363 (0.023)	-0.305 (0.028)	-0.279 (0.022)	-0.291 (0.025)	-0.307 (0.024)
LITTLE_FISH	0.495 (0.033)	0.485 (0.028)	0.510 (0.036)	0.482 (0.028)	0.471 (0.031)	0.470 (0.031)
LIT_LK_FAIR	-0.165 (0.038)	-0.218 (0.032)	-0.127 (0.034)	-0.127 (0.032)	-0.144 (0.034)	-0.135 (0.030)
LITTLE_SAND	0.226 (0.037)	0.203 (0.036)	0.229 (0.035)	0.218 (0.028)	0.211 (0.031)	0.208 (0.032)
LONG_LAKE	-0.362 (0.041)	-0.410 (0.034)	-0.339 (0.036)	-0.318 (0.026)	-0.334 (0.030)	-0.326 (0.030)
MUD_LAKE_C	0.154 (0.059)	0.088 (0.050)	0.177 (0.067)	0.182 (0.053)	0.151 (0.057)	0.181 (0.048)
PALM_LAKE	0.105 (0.022)	0.109 (0.021)	0.114 (0.025)	0.092 (0.022)	0.080 (0.024)	0.109 (0.026)
PARK_LAKE_B	0.019	-0.008	0.016	0.057	0.032	0.032

	(0.043)	(0.044)	(0.051)	(0.040)	(0.043)	(0.043)
POCKET_LAKE	0.314	0.291	0.337	0.297	0.314	0.312
	(0.030)	(0.025)	(0.032)	(0.024)	(0.027)	(0.026)
ROCK_LAKE	-0.482	-0.459	-0.487	-0.449	-0.458	-0.479
	(0.040)	(0.041)	(0.047)	(0.034)	(0.037)	(0.039)
SPRING_LK_B	-0.058	-0.057	-0.017	0.006	-0.023	-0.047
	(0.033)	(0.032)	(0.038)	(0.030)	(0.033)	(0.032)
SPRING_LK_C	0.168	0.184	0.192	0.174	0.179	0.181
	(0.023)	(0.022)	(0.027)	(0.019)	(0.022)	(0.023)
Lambda	0.756	0.782	0.594	0.350	0.448	0.507
	(0.006)	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)

CHAPTER 3: WATER QUALITY INDICATORS IN THE HEDONIC PROPERTY PRICE MODEL

3.1. Introduction

Although significant progress in cleaning up US waterbodies has been made since the passage of the CWA, non-point source pollution still remains a threat to lakes. The EPA estimates that almost half of the nation's lakes are classified as impaired, with nutrients the leading source of impairment (US EPA 2003). Numerous state and local programs have been enacted to combat lake pollution. These programs typically focus on reductions in nutrients loadings and improvements in other physical measures. Chapters 1 and 2 used hedonic property analysis to show that considerable benefits result from lake programs that increase water quality. However, an issue that has received scant attention is whether variables used to represent water quality in hedonic analysis imply the same concept of "quality" used to set regulatory goals. Since Secchi Disk Measurement (SDM) is not directly used to regulate lakes, there may be differences in the goals of regulators and the values of property owners that could bias the results of a cost-benefit analysis. If property owners are not fully informed about the ecological and ecosystem benefits, the values obtained from property prices may understate these benefits and favor the value of aesthetic benefits.

A chronic issue in the valuation of water quality is the choice of water quality indicator. In papers that compare multiple water quality indicators, the typical goal is to find the one "best" measure. However, evidence from the fields of ecology and biology suggests that a focus on a single indicator may be too narrow (Hoyer et al. 2002). "Water quality" itself is a multi-dimensional concept that depends on the designated use of the waterbody and varies over the population. Recreational anglers may value greener water as a good fish habitat, whereas swimmers desire higher transparency (Hoyer, Brown, and Canfield Jr. 2004). These differences are recognized in the Clean Water Act, which classifies lakes and their water quality targets by designated uses.

How individuals perceive water pollution is a major concern in this area. The hedonic model assumes that consumers know the level of water quality and translate this knowledge into their purchasing decision. Zabel and Kiel (2000) discuss a similar issue in the context of air quality and note that very little research into individual’s perceptions of pollution has been conducted. Delucchi, Murphy, and McCubbin (2002) look at differences in consumer’s values for health and visibility benefits from improved air quality. They point out that is hard to determine which of these categories consumers have in mind when purchasing a house. Imperfect information exists, since individuals likely obtain their perceptions from visibility and the media.

The current water quality regulation in Florida was constructed from a biological and ecological point of view, where visibility only plays a small role. Lake regulation is done with the health of the ecosystem in mind, and policy variables are chosen to reflect this. If clarity indicators like SDM are not directly correlated with physical indicators, hedonic analyses may not be measuring the desired benefits. Furthermore, it would be easier for regulators to use hedonic estimates if they were formulated in terms of variables they have control over. Cropper (2000) notes that to be effective, estimates of value associated water quality changes must be relevant to environmental management.

Several past studies have questioned the validity of SDM as an indicator of ecosystem health (Leggett and Bockstael 2000). Transparency can represent the process of eutrophication in certain settings, but this relationship has been shown to be very loose in Florida (FDEP 1996).

Tannins released by plants and trees can cause less transparent water with no quality change. Furthermore, Leggett and Bockstael (2000) point out that lakes in the northeast US that have been made very clear by acid rain but are unable to support an ecosystem. However, while there may not always be an ecological value linked with clear water, it is important to recognize that there are considerable economic benefits associated with it. Chapter 1 and several studies mentioned therein illustrate sizeable property price benefits resulting from increased clarity. In order to create the most efficient policy that properly considers social benefits and costs, regulators should not ignore these economic benefits.

This chapter investigates several water quality indicators in a spatial hedonic property price analysis in order to compare the resulting implicit prices and benefit estimates. The choice of indicators was driven by official Florida state water quality policy. The four additional indicators are total nitrogen, total phosphorous, chlorophyll, and TSI. The first three are regularly used nationwide to classify lakes and are deemed the most important variables (alongside SDM) for nutrient criteria by the EPA (EPA 2000). The fourth indicator, TSI, plays a primary role in Florida lake classification and regulation. The goal of this chapter is to compare the alternate indicators to the previous SDM estimates and to determine if the previous three hypotheses (the Edge Effect, Proximity Effect, and Area Effect) hold when other indicators are used. Results indicate that consumer’s valuations of water quality differ over indicators. Future policy and research can gain from accounting for the value of improvements from more than one indicator.

3.2. Literature Review

In the early years of the CWA, the majority of effort and legislation focused on the regulation of point sources. Significant progress was made in this area, which led regulators to realize the large impact of non-point sources of pollution. In 1987, the US Congress amended the CWA to include non-point sources. These alternative forms of pollution were addressed under section 319 of the CWA, which calls on states to develop programs to manage and monitor non-point sources. These state reports now show that non-point sources of pollution are the main cause of impairment of US surface waters (US EPA 2003). According to the EPA, the main source of non-point source pollution is agriculture.

In Orange County, the majority of water pollution is a result of non-point sources. Whereas agriculture is a large contributor to water pollution in rural areas, urban areas are much more affected by pollution from stormwater runoff, precipitation, atmospheric deposition, and seepage. (US EPA 2003). The increased amount of impervious surfaces in urban areas creates a different situation for water pollution than rural areas with increased farming activities.

Heterogeneity in pollution definitions is not unique to water quality. Studies of air quality (Smith and Huang 1995; Zabel and Kiel 2000), open space (Anderson and West 2006; Cho et al. 2009), and hazardous waste (Jackson 2001) have all debated about proper representations of the relevant amenity or disamenity. To fully understand concerns with the definition of water quality, lessons from these other areas can be incorporated.

This literature review begins with two sections about consumer perceptions, which draws on papers from several areas of valuation. Next, indicators from past studies of water quality are discussed. In the short history of this field, a wide variety of water quality indicators have been used. The final section of this literature review summarizes water quality legislation and the indicators that are used in the State of Florida for water regulation.

3.2.1. Perceptions

There are many influential reasons to use SDM to measure water quality. Water clarity is easily perceived by consumers and past hedonic papers have found that indicators that are visible to the public are more likely to affect property prices (Brashares 1985). People's perceptions are likely to be the impetus of their behavior (Egan et al. 2009), so it makes sense to use an indicator that reflects this fact. Also, SDM is cheap to obtain and is widely collected by many organizations.

On the other hand, people's perceptions of water quality may not match up with the physical measures of water quality (also called "water chemistry" indicators) used by states to regulate waterbodies. Pendleton, Martin, and Webster (2001) conducted a survey in California to determine the accuracy of people's perceptions about beach water quality and find that the public's perceptions often conflict with the actual situation. Respondents generally overstated pollution levels, could not properly sort beaches by water quality, and were typically unable to identify the true causes of water degradation (storm water run-off and sewage overflow).

Jeon et al. (2005) examine differences in water quality perceptions and several physical measures in an analysis of recreation demand. A survey was sent to a random sample of Iowans which collected recreation demand and perception information about 131 Iowa lakes. The surveys were designed to elicit answers that were easily compared to physical indicators of water quality. In their recreation demand model they find that the use of either the survey perceptions or physical measures alone paints an incomplete picture of behavior. To fully explain respondents' behavior, they recommend the use of both classes of variables. Additionally, they find that individuals have different perceptions of water quality than the EPA or scientists. Another paper to compare individual perceptions to objective measures is Adamowicz et al. (1997), who investigate discrete choice models of the demand for moose hunting. They find that the indicators based on perceptions slightly outperform the objective measures in explaining behavior and that there may be some differences between the two categories. However, similar to Jeon et al. (2005), a joint model that combines both forms of information outperforms the two separate models.

Hoyer, Brown, and Canfield Jr. (2004) compare individual's perceptions to physical measures in Central Florida. They survey volunteers of the Florida Lakewatch program on the same day that water quality samples are taken to determine if there is a relation between user perceptions and water chemistry and visibility data. A direct relationship between perceptions and water chemistry measures is found, although there is significant noise in the relationship and perceptions varied significantly between individuals. Furthermore, they note that their results are hard to translate to the general population because they survey active participants in lake monitoring.

In most regions of the country, SDM is seen as a physical, objective measure of water quality. In Florida, however, there are particular biological concerns that make clarity measures more subjective. In particular, "dark water" can occur because local plants and trees release tannins into the water, which typically have no effect on the quality of the water but can substantially affect clarity. This dark water can potentially affect consumers' personal perceptions of water quality. In a Florida Department of Environmental Protection report, it is stated that "Attempts in previous reports to include secchi depth have been unsuccessful in dark-water lakes and estuaries, where dark waters rather than algae diminish transparency" (FDEP, 2006).

Another issue relating to consumer perception is consumer's awareness of non-transparency indicators. Variation in these variables may not be easily identified and have a smaller correlation with property prices than more visible indicators (Brashares 1985). These issues of perceptions have important implications for hedonic property price analysis since past research has shown that the choice of water quality indicator can affect the significance and magnitude of the estimated implicit prices (Michael, Boyle, and Bouchard 2000). Since these implicit prices are used to construct estimates of the total benefits of water quality improvements, the results of a cost benefit analysis can depend on the particular indicator used.

A hedonic study that represents a preliminary investigation of this problem is Poor et al. (2001), who investigate differences between objective and subjective water quality indicators. Their subjective indicator is based on a survey of waterfront homes and their objective indicator is SDM; they find that the objective indicator is preferred to the survey based indicator in its ability to explain variation in home prices. The current paper takes Poor et al.'s analysis one step further by questioning the reliability of SDM as an objective indicator. Water clarity indicators may only measure one dimension of water quality related to perceptions and ignore other dimensions related to state regulatory targets.

Issues of perception have seen some interest in recent air quality studies. Zabel and Kiel (2000) call for more research into how individuals form their perceptions of quality and how these impact values obtained from property prices. Delucchi, Murphy, and McCubbin (2002) distinguish health benefits from aesthetic benefits in estimates of the value of air quality. They assert that individuals form their perceptions from visibility and the media, which is likely to cause incomplete information. When individuals purchase homes and bid more for those with better quality air, their definitions of quality may be more influenced by visibility than health effects. In a contingent valuation paper, Brookshire, Thayer, and Schulze (1982) find that approximately 34% of the value of air quality benefits is from visibility and 66% is from health impacts. Since visibility is easily perceived, it is likely that value estimates from hedonic property analysis capture the majority of this aesthetic benefit. However, due to imperfect information, the value of health benefits is likely biased downwards (Cropper 2000).

These issues of perception in the air quality literature parallel those in the current analysis of water quality. Home prices may be conveying two distinct values: aesthetic benefits and ecosystem or environmental benefits. When consumers purchase homes, it is likely that they care about both of these issues. However, imperfect information about the determinants of ecosystem health may cause hedonic estimates to undervalue these issues.

3.2.2. Consumer Information and Indicator Choice

. The use of property prices to infer the value of water quality may introduce some restrictions on the choice of water quality indicator, since consumers (property buyers and sellers) must have access to information on that indicator. This information requirement can introduce a tradeoff between indicators that better represent the health of a lake and indicators that are most easily perceived by home buyers. While it is relatively easy for consumers to see clarity as represented by SDM, finding information on TSI and nutrients may be slightly more involved. If consumers do not have access to information about the water chemistry indicators, they are not likely to influence home prices. For instance Leggett and Bockstael (2000) stressed that consumers regularly encountered information about their physical indicator (fecal coliform) since its levels were regularly broadcast on the radio, published in local newspapers, and were the cause of frequent beach closings.

In order to use nitrogen, phosphorous, chlorophyll a, and TSI as water quality indicators in a hedonic property regression in central Florida, it is important that individuals have access to information about them. Several pieces of evidence indicate that this is the case. First, Central

Floridians can download a wealth of data from the STORET database (an online database started by the EPA). This is a national database that is updated by individual states. However, the database requires a degree of technical sophistication and may only be used by a small fraction of homeowners.

In a more user friendly format, the city of Orlando has a comprehensive listing of local lakes on its website (http://www.cityoforlando.net/public_works/stormwater/lakes/index.htm), which contains a summary of recent TSI, nitrogen, phosphorous, and chlorophyll a levels for each lake. This water quality information is on the same website that contains official information about garbage and recycling collection, job openings, community events, local parks, police and fire departments, and other local information that prospective homebuyers are bound to come across. The other municipalities in Orange County also disseminate lake information. For instance, the City of Winter Park distributes a "Lake Management Report and Lake Users Guide" to its residents. This pamphlet contains information on SDM, water temperature, pH, dissolved oxygen, conductivity, oxidation reduction potential, total nitrogen, total phosphorous, chlorophyll a TSI, and turbidity, for all lakes in the municipality. Historical trends with lines of best fit are presented for the majority of the indicators, and interpretations are provided about lake conditions.

Finally, the Florida Lakewatch program is an active network of volunteers that sample and record lake readings on a monthly basis. Volunteers are required to complete a training course, and are given equipment for sampling and secchi depth measurements. Water samples and secchi readings are deposited at several collection centers for recording and analysis. The result is an extremely active network, which produces and distributes quarterly newsletters throughout the community that detail local water quality information and data. Furthermore, data is available from their website, and all indicators used in this chapter are freely available. These data regularly appear in local news stories about lake issues.

Water quality information is therefore distributed on the state, municipality, and local level, and home buyers have access to the information on the four additional indicators. Furthermore, the importance of lake recreation in the area results in water quality issues being regularly raised in local news and politics. Due to all of the above sources, it is clear that residents can easily access information about nutrients, TSI, and chlorophyll a.

3.2.3. Water Quality Indicators in Environmental Valuation

A wide variety of water indicators have been used in the hedonic water quality literature, several of which were discussed in the first chapter. In early hedonic property analyses, survey methods were used to measure the effect of water quality on property prices. David (1968) used this method in Wisconsin by having representatives from the Wisconsin Department of Conservation rate a lake as poor, moderate, or good in terms of water quality. She concluded that the water quality of nearby lakes definitely affects property prices. Other studies that use surveys include Epp and Al-Ani, (1979), Young (1984), Michael, Boyle, and Bouchard (2000), and Poor, et al., (2001). There is widespread skepticism with surveys, however, since the implicit price of water quality computed could be measuring a perceived, rather than actual, measure of water quality (Steinnes, 1992; Poor et al. 2001).

Another class of water quality indicators used in hedonic property analysis is based on the chemical composition of the water body, frequently referred to as “physical measures” or “water chemistry.” For instance, Epp and Al-Ani (1979) use the pH of the water along with dissolved oxygen, biochemical oxygen demand, acid from minerals, acid from carbon dioxide, and nitrate and phosphate concentrations. They used each of these in a separate hedonic equation, though they were looking at streams instead of lakes. Their conclusions indicated that the water quality of local streams affected property prices. Feenberg and Mills (1980) used 13 different physical quality measures and found that indicators that are visible to people, such as oil content and turbidity, were most correlated with property prices. Other indicators of this type that have been used in hedonic pricing equations include fecal coliform, (Leggett and Bockstael, 2000) and dissolved inorganic nitrogen and total suspended solids (Poor, Pessagno, and Paul 2007).

As discussed in the first chapter, the most common indicator of water quality used in hedonic price analyses in recent studies is SDM. This indicator is relatively simple to obtain and is recorded by most state-level departments of environmental protection for large water bodies. The ease of measurement and widespread use are frequently cited as reasons for its prevalent use in water quality research dealing with property prices (Poor et al. 2001).

Michael, Boyle, and Bouchard (2000) focused on the selection of the water quality indicator in their study of water pollution in Maine lakes. They stated (p.283) “It appears that most hedonic studies are conducted using whatever empirical measure(s) of environmental quality is (are) available as environmental variable(s)…” They also observed “Consequently, an issue that appears to have been overlooked in the literature is the appropriate selection of an environmental variable to include in the hedonic price equation.” The focus of the paper was on temporal variation in water quality, and whether people consider long term trends or short term changes. A survey about respondent’s perceptions of water quality was distributed to local property owners. Based on the results of the survey, nine variations of SDM were used in a hedonic property price analysis around several Maine lakes. The nine variations were created to represent differing perceptions about past, current, and future impressions of water quality. The results do not provide clear evidence for a preferred indicator and the confidence intervals for all of the estimated implicit prices overlap. Nonetheless, the magnitudes of the point estimates are very different between indicators, with some half that of others. They conclude that the choice of indicator can influence policy recommendations resulting from a cost benefit analysis.

Studies that use SDM as a water quality indicator typically claim that the clarity of the water can be interpreted as the quality of the water (Michael, Boyle, and Bouchard 2000). This may be true in certain pristine environments like the Maine lakes encountered in Michael, Boyle, and Bouchard (2000), however many others are questioning whether this can be extended to other settings. For example, Steinnes, (1992) suggests that SDM may be capturing a perceived, rather than actual, measure of water quality. He cites a litigation case in Minnesota dealing with acid rain. Because acid rain can kill certain algae, it can actually increase water transparency. Therefore, paradoxically, acid rain and water pollution may appear to be positively correlated with property values when using SDM as an indicator of water quality. Leggett and Bockstael (2000) also talk about this point, stating "benefits of improvements in water clarity have ambiguous ecological merit: mountain lakes plagued by acid rain can be crystal clear, yet completely sterile" (Leggett and Bockstael, 2000, p.122).

A variety of indicators have also been used in other areas of the environmental valuation literature. In a random utility model about recreation demand for sport fishing, Karou, Smith, and Liu (1995) connect regulations on effluent loadings to measures of environmental quality that influence people's behavior. They focused on non-point source pollution, and linked changes in nutrients (in the form of nitrogen) and biochemical oxygen demand with predictions of changes in fish catch. These water quality variables are found to significantly influence the demand for sport fishing. von Haefen (2003) also examines recreation demand in random utility models, but is concerned with welfare measures and several specific functional forms. To represent water quality at each of 219 potential sites for water based recreation, von Haefen uses the Trophic State Index (TSI, discussed below). The Pennsylvania TSI uses a combination of phosphorous and SDM to calculate the trophic state of a waterbody.

Atasoy, Palmquist, and Phaneuf (2006) use a spatial econometric model to analyze the effect of residential development and impervious surfaces on water quality. Total suspended solids, total nitrogen, and total phosphorous are used to represent water quality. Both residential development and changes in development contribute to increased TN and TP levels and degrade water quality. Only new construction is found to increase the level of total suspended solids.

In a meta-analysis of water quality valuation studies, Van Houtven, Powers, and Pattanayak (2007) summarize much of the previous water quality valuation literature. As discussed in Chapter 1, they had trouble combining numerous studies because of differences in water quality indicators. In order to include the most past comparable papers, the meta-analysis is restricted to previous stated preference papers that express water quality in a 10-point scale. They find that WTP values for water quality changes vary in expected ways with socioeconomic factors in the population. In order to standardize future analyses of water quality in stated preference research, Van Houtven et al support the use of indicators based on the "designated use" classifications of the CWA.

Egan et al. (2009) directly examine the choice of water quality indicator in a recreation demand model. They focus on indicators used by the EPA to classify lakes as "impaired," including SDM, total nitrogen, total phosphorous, chlorophyll, bacteria (cyanobacteria and total phytoplankton), and suspended solids (organic and inorganic). They split the data and perform out-of-sample prediction to determine goodness of fit, similar to the second chapter in this dissertation. Results indicate that SDM is "clearly" the best indicator to use due to its pervasive statistical significance and the low cost of obtaining measurements. The second best indicator is found to be total phosphorous, while chlorophyll and suspended solids fare the worst. They also find the surprising result that increased chlorophyll levels do not negatively impact recreation decisions, indicating that the average respondent may not mind some "greenish" water due to higher chlorophyll levels. The tolerance for greenish water highlights the point that water quality is a much larger concept than simply clarity. There are many ecosystem services provided by lakes that may not always be directly correlated with clarity (Hoyer, Brown, and Canfield Jr. 2004). Overall, this study provides significant evidence that recreation decisions are influenced by water quality, including characteristics represented by physical measures such as total phosphorous and total nitrogen.

3.2.4. Water Quality Regulation

The Clean Water Act (CWA) provides an overall framework under which all states must cooperate. Legislation and regulation of water quality in most states is directly influenced by the mandates of the CWA. In the official Florida regulation dealing with water quality, the CWA is frequently mentioned as a motivation (FDEP 2008). The present section discusses the CWA and the relevant Florida laws that pertain to the maintenance of water quality.

Under the CWA each state has to submit a 305(b) report and a 303(d) list every two years. Arising from section 305(b) of the CWA, the first report requires individual states to conduct water quality monitoring activities and describe current conditions and activities. The second report, the 303(d) list, contains a list of water bodies that do not meet specific standards.

In the 305(b) report, each waterbody in the state is sorted into one of three categories: fully meets use designations, partially meets use designations, or does not meet use designations

(DeBusk 2002). Under the CWA, waterbodies are assigned “designated uses” such as class I: potable water supply, class II: shellfish propagation or harvesting, and class III: recreation, propagation and maintenance of a healthy, well balanced population of fish and wildlife. There are three main sources of data that the state of Florida uses to assess whether the designated uses are met. The first source involves water chemistry data from the Florida STORET database, which is the source of most of the water quality data in the present paper. The second source is biological data from the FDEP Biological Database, and the third source is the Department of Health’s fish advisory data (DeBusk 2002). For lakes a TSI (discussed in the next section) is created from these data and used to determine if the relevant designated use is met. In Orange County, the only water body that does not meet its designated use is Lake Apopka, which is not included in the data for this dissertation.

Recognizing the differences in pollution sources between states, the CWA allows some flexibility in individual state water quality monitoring. The official rules on monitoring and water quality regulation were drafted by the Florida Department of Environmental Protection and approved by the legislature; these various forms of legislation appear in the Florida Administrative Code (found at www.flrules.org/) and the Florida Statutes (found at www.flsenate.gov). Florida Statute 403 deals with environmental control and pollution control and contains some of the main laws about the regulation of water quality. Section 403.067 defines “the establishment and implementation of total maximum daily loads” and outlines some of the main state procedures that are required to comply with the CWA. This law essentially states that any waterbody that does not meet specific criteria will be listed and a TMDL will be established.

The specific water quality criteria are defined in the Florida Administrative Code, in Chapter 62-303. Section 62-303.310 of the FAC states that a waterbody shall be placed on a "planning list" if one of four thresholds is exceeded. Waterbodies included on the planning list will be analyzed for inclusion on the 303(d) list. The first threshold is an aquatic life-based criteria outlined in 62-303.320. The second threshold deals with biological assessments which are defined in 62-303.330. The third threshold concerns toxic components as defined in 62-303.340 and the final threshold is based on nutrients as defined in 62-303.350. To illustrate the particular indicators used for regulation, each of these thresholds will be briefly discussed.

The aquatic life-based criteria in 62-303.320 include several chemicals and substances that have been found to be toxic to aquatic life, including aluminum, arsenic, chlorine, cyanide, DDT, silver, and several others. The actual thresholds are based on the proportion of samples taken that contain a certain level of the chemical or substance.

The biological thresholds of section 62-303.330 only apply to lakes with a color classified as less than 20 platinum cobalt units (extremely clear). This section declares that biological data must meet the requirements of 62-303.320, and refers to an expanded list of chemicals and substances that are defined in section 62-302.530. Section 62-303.340 states that any waterbodies with two samples that indicate either acute or chronic toxicity will immediately be placed on the planning list.

The fourth threshold contained in section 62-302.350 is based on nutrients, and states that the TSI and annual mean chlorophyll a values are to be used as the primary indicators of nutrient impairment. Since the majority of Florida lakes are impaired by nutrients (FDEP 2006), these thresholds are the most commonly employed. The specific nutrient thresholds for each class of water defined by the CWA are discussed in the 305(b) reports.

The FDEP is also currently developing a set of numeric nutrient criteria to protect against the further nutrient enrichment of state waters (FDEP 2009). These are being designed to support the narrative nutrient criteria of the state, which declares "in no case shall nutrient concentrations of a body of water be altered so as to cause an imbalance in natural populations of flora or fauna." These new nutrient criteria are for phosphorous, nitrogen, chlorophyll a and potentially transparency.

These above FAC thresholds provide specific guidance about water quality criteria used by the state of Florida to classify lakes. These were passed by the Florida state government and guided by the CWA. None of the aforementioned documents specifically mention SDM as a means of classifying or regulating lakes. Consequently, when Florida regulators set goals for water quality and conduct management projects, SDM may not be directly used. This could be a problem for hedonic property analysis that uses SDM, as the goals of regulators may differ from what is being used as the water quality indicator. This begs an important policy question: does the actual health of the lake get incorporated into home prices, or is it only lake visibility that affects home prices? If visibility is the most important characteristic, this has welfare implications for policy makers. Since current policy targets physical indicators over clarity indicators, changes to these physical measures may have unintended impacts on property prices and the resulting welfare calculations (Jeon et al. 2005).

3.3. Data

This chapter will use the same housing data as the first chapter, augmented with four additional water quality variables. These new variables are total nitrogen, total phosphorous, chlorophyll a, and TSI. As discussed in the past section, these indicators are used by the state of

Florida to classify lakes and institute policy. These additional indicators have been shown to accurately gauge the level of water pollution resulting from both point sources and non-point sources, though they are particularly useful for non-point sources (FDEP 2008).

Nutrients exist naturally in low levels in most lakes. However, when their levels are elevated, the process of eutrophication occurs. Eutrophication is harmful to many types of organisms and can significantly limit the ability of the lake to achieve its designated use. When the additional nutrients are introduced, certain aquatic plant life and algal bloom productivity substantially increases. The increase in aquatic plants can decrease the amount of light and subsequent photosynthesis taking place in the deeper areas of lakes. Also, when this excessive aquatic plant life dies, bacteria that decompose this organic matter appear in abundance. These bacteria tend to deplete the oxygen supply for other organisms and can produce a bad odor, as well as significantly decreasing visibility (US EPA 2003). Chlorophyll a has been found to be a useful indicator for the prevalence of algal blooms. When other data are not freely available, Florida reporting agencies track levels of Chlorophyll a.

Another reason to use an indicator like chlorophyll is that "chlorophyll is the dominant factor determining Secchi Depth" in Florida waters, and in Florida lakes chlorophyll accounts for as much as 72% of the variance in Secchi Depth (Hoyer et al. 2002). Focusing on chlorophyll may better capture the many elements of water quality, and net out many of the problems with perceptions. Nitrogen and phosphorous are also useful in this respect, as they more accurately gauge the level of eutrophication in a lake. Furthermore, total phosphorous is the principal limiting nutrient in Florida Lakes, and accounts for approximately 79% of the variation in chlorophyll (Hoyer et al. 2002).

There are two main classifications of the trophic status of lakes: oligotrophic ("little food" in Greek) and eutrophic ("well fed"). A third class is generally called mesotrophic and refers to waterbodies in transition between oligotrophic and eutrophic (Ryding and Rast 1989). These trophic states are generally used to describe the "health" of the lake and its nutrient contents. Oligotrophic lakes generally have low nutrient concentrations, diverse plant and animal life, low primary productivity and biomass, and generally good water quality for the majority of uses (Ryding and Rast 1989). On the other hand, eutrophic lakes are characterized by high productivity and frequent algal blooms, anoxic bottom waters during thermal stratification, less diverse plant and animal communities, higher growth of littoral zone aquatic plants, and generally poor water quality for the majority of uses (Ryding and Rast 1989). Nonetheless eutrophic waters can support increased fish productivity, though the diversity of fish is typically much less.

A measure has been developed to measure the trophic state of lakes: TSI. This measure was developed in Carlson (1977) as a way to address the multidimensional concept that is a lake's trophic state. At the time it was believed that a lakes trophic status could not be summarized by one measure because it is due to many different influences such as nutrient loading and concentration, algal productivity, and local plant activity. Carlson was concerned with the usefulness of multi-parameter indices and designed an indicator based on chlorophyll a, SDM, and total phosphorous. The TSI has since evolved from its original form to be adapted to specific areas.

The Florida TSI was developed by the FDEP and is a combination of total nitrogen, total phosphorous, and chlorophyll a. The dark water issue is the reason that the FDEP defines its TSI differently than most other states. In the face of dark water, which affects the clarity but not quality of the water, limnologists found that TSI measures based on SDM could be misleading (FDEP 1996). The TSI is defined in the FDEP 1996 305(d) report to the EPA, and is based on statistical analyses of nutrient levels in Florida lakes. Figure 4 is from that report, and gives the precise equations that define the TSI. The index uses the three inputs to produce a number between 1 and 100. A TSI value between 0-59 is seen as good, from 60-69 is fair, and 70 to 100 is classified as poor.

Trophic State Index (TSI)
for lakes and estuaries

For lakes: 0-59 is good, 60-69 is fair, 70-100 is poor.
For estuaries: 0-49 is good, 50-59 is fair, 60-100 is poor.

Trophic State Index	Chlorophyll CHLA/ micrograms per liter (µg/1)	Total Phosphorus TP/ milligrams of phosphorus per liter (mgP/1)	Total Nitrogen TN/ milligrams of nitrogen per liter (mgN/1)
0	0.3	0.003	0.06
10	0.6	0.005	0.10
20	1.3	0.009	0.16
30	2.5	0.01	0.27
40	5.0	0.02	0.45
50	10.0	0.04	0.70
60	20.0	0.07	1.2
70	40	0.12	2.0
80	80	0.20	3.4
90	160	0.34	5.6
100	320	0.58	9.3

Trophic State Index equations that generate the above criteria (LN = Natural Log):

$CHLA_{TSI} = 16.8 + [14.4 \times LN (CHLA)]$

$TN_{TSI} = 56 + [19.8 \times LN (TN)]$

$TN2_{TSI} = 10 \times [5.96 + 2.15 \times LN(TN + .0001)]$

$TP_{TSI} = [18.6 \times LN (TP \times 1000)] - 18.4$

$TP2_{TSI} = 10 \times [2.36 \times LN(TP \times 1000) - 2.38]$

** Limiting nutrient considerations for calculating $NUTR_{TSI}$*

If $TN/TP > 30$ then $NUTR_{TSI} = TP2_{TSI}$

If $TN/TP < 10$ then $NUTR_{TSI} = TN2_{TSI}$

If $10 < TN/TP < 30$ then $NUTR_{TSI} = (TP_{TSI} + TN_{TSI}) / 2$

$TSI = (CHLA_{TSI} + NUTR_{TSI}{}^{*}) / 2$

Figure 4: Trophic State Index Definition

From the figure it is clear that TSI is not a continuous function. This is because the TSI is based on the limiting nutrient concept. While the level of both nitrogen and phosphorous in a lake is important to gauge its health, the ratio of nitrogen to phosphorous can also play a significant role. For instance, a lake is declared phosphorous limited if the ratio is greater than 30 and nitrogen limited if the ratio is less than 10. These situations can imply different levels of eutrophication and water quality (FDEP 1996). In the bottom of Figure 4, these unbalanced ratios result in different calculations of TSI which take into account the limiting factor.

Table 20 summarizes the TSI rating of lakes used in this study. The first row presents the average of all individual lake samples in the data. Approximately 88% of observations are labeled good, 12% are fair, and 1% of the observations are classified as poor. The third row of Table 20 contains numerical counts of lake ratings, from which the first row was obtained. Only 10 annual TSI readings in the data are classified as poor.

Table 20: TSI Ranking Exploration

	Good	Fair	Poor
Average TSI across all lake samples	0.876	0.124	0.008
Total number of samples in each category	1035	146	10

The TSI, nutrient, and chlorophyll a data for the present chapter was obtained from the Florida STORET database, with some missing observations filled in by data from three municipalities (the same sources as the SDM data in Chapter 1). Table 21 shows mean TSI, total nitrogen, total phosphorous, and chlorophyll a values over all lakes in each year. From this table it can be seen that the highest years of lake water quality were between 1998 and 2000. This is clearly reflected in both TSI and chlorophyll a values, which both see a slight decrease in those years. This table also illustrates the similarity between these four variables, as they all move together. In particular, as nutrients increase there is an increase in algal mass, illustrated by the subsequent increase in chlorophyll a.

Table 21: Water Quality Indicators

	TSI	TN	TP	Chlor a
Mean Indicator Level by Year				
1996	46.828	0.872	0.039	12.753
1997	45.983	0.819	0.036	11.809
1998	42.782	0.706	0.034	9.947
1999	42.460	0.849	0.032	9.744
2000	41.549	0.859	0.034	9.117
2001	43.407	0.830	0.036	10.873
2002	45.746	0.850	0.037	10.712
2003	47.712	0.856	0.038	11.567
2004	46.111	0.828	0.037	11.956
Summary Statistics Over all Years				
Mean	44.677	0.829	0.036	10.925
St. Dev	13.256	0.334	0.034	11.142
Min	11.185	0.154	0.003	0.35
Max	87.49	2.565	0.35	72.431

Table 21 also shows the overall summary statistics for all lakes in all years for each of the sampling-based water quality indicators. Similar to the result in Table 20, the overall TSI mean in Table 21 is less than 59, indicating a rating of “good”. In the max level row a wide range of TSI values are represented: there was a lake that had a very poor rating of 87.49, and the best lake rating is 11.185.

3.4. Methods

With the introduction of the alternative water quality indicators, the hedonic model from Chapter 1 can be re-estimated. In order to facilitate comparison between estimates, the basic structure of the final model from Chapter 1 will remain the same except SDM is replaced with one of the other indicators. Therefore four new models will be estimated, one for each alternative water quality indicator. These new models appear below.

TSI:

$$P = \beta_0 + \beta_1 WF + \beta_2 TSI + \beta_3 WF * TSI + \beta_4 Distance + \beta_5 Distance * TSI + \beta_6 Area * TSI + \gamma' \mathbf{x} + \varphi' \mathbf{y} + \psi' \mathbf{l} + \delta' \mathbf{t} + \varepsilon \quad (3.1)$$

Total Nitrogen (TN):

$$P = \beta_0 + \beta_1 WF + \beta_2 TN + \beta_3 WF * TN + \beta_4 Distance + \beta_5 Distance * TN + \beta_6 Area * TN + \gamma' \mathbf{x} + \varphi' \mathbf{y} + \psi' \mathbf{l} + \delta' \mathbf{t} + \varepsilon \quad (3.2)$$

Total Phosphorous (TP):

$$P = \beta_0 + \beta_1 WF + \beta_2 TP + \beta_3 WF * TP + \beta_4 Distance + \beta_5 Distance * TP + \beta_6 Area * TP + \gamma' \mathbf{x} + \varphi' \mathbf{y} + \psi' \mathbf{l} + \delta' \mathbf{t} + \varepsilon \quad (3.3)$$

Chlorophyll (Ch):

$$P = \beta_0 + \beta_1 WF + \beta_2 Ch + \beta_3 WF * Ch + \beta_4 Distance + \beta_5 Distance * Ch + \beta_6 Area * Ch + \gamma' \mathbf{x} + \varphi' \mathbf{y} + \psi' \mathbf{l} + \delta' \mathbf{t} + \varepsilon \quad (3.4)$$

In each equation, the variables are defined as in Chapter 1. For all four of these models, three regressions are estimated: a non-spatial regression and two spatial regressions—one with 15 nearest neighbors and another with distance and time boundaries of 200 meters and 6 months back, three months forward. In Chapter 2, there were not large qualitative differences between the various SWM's in the spatial lag model. Nonetheless, two of the main types of SWM are employed to ensure these results are not affected by the specification of the weights matrix.

After the regressions are estimated, implicit prices can be constructed and compared. The spatial extent of these new implicit prices is of particular interest. If the magnitude and spatial extent of the physical indicators of water quality are different than the corresponding SDM values, regulators may be faced with some policy tradeoffs. Following Michael, Boyle, and Bouchard (2000) and Poor et al. (2001), the a priori expectation is that these implicit prices will be different than the SDM values. Finally, the total benefits for three representative lakes of the first chapter will be calculated. These total benefits should better illustrate differences between water quality indicators.

3.5. Results

The non-spatial regressions are estimated using maximum likelihood in Stata while the spatial regressions were estimated in Matlab. In order to conserve space the 146 lake dummies are not presented in the following tables. The coefficients on these dummies are not of any particular interest for interpretation. An F-test of the significance of the dummies was estimated for each regression. The hypothesis that the coefficients were all equal to zero was strongly rejected in each case.

The results of the TSI and Chlorophyll a regressions appear in Table 22. Starting with the first column of non-spatial coefficients, the variables unrelated to lakes all display the expected signs. With the lake variables, the expected signs are the opposite of the SDM regressions since higher values of TSI correspond to lower water quality. Therefore a negative coefficient on the TSI variable indicates that increases in TSI and worse water quality are negatively related to home prices, as expected. In fact, all three of the hypotheses from the first chapter about the effect of water quality on property prices are confirmed in this regression. The negative and significant coefficient on the TSI and waterfront variable indicates that waterfront homes are more negatively affected by increases in TSI than non-waterfront homes, illustrating the Edge Effect. The coefficient on the distance and TSI interaction term is positive and significant, indicating that the negative impact of higher TSI on property prices diminishes with distance, in line with the Proximity Effect. Also, the Area Effect is confirmed by the negative and significant coefficient on *WQ*Ln(Area)*, although this coefficient is now only significant at the 5% level.

Overall, these results exhibit the same qualitative implications as the model based on SDM measures of water quality.

As in the previous chapters, there is not much variation between the spatial and non-spatial coefficients, which are quite stable across specifications. The standard errors in the spatial regressions are lower for every variable except *Canalfront*, where there is a difference of 0.001. Consequently, the main impact of controlling for spatial influences and omitted variables is a narrowing of confidence intervals. The spatial parameter ρ is highly significant in each spatial model, but relatively small in magnitude, at 0.003 for the nearest neighbor SWM and 0.001 for the distance/time SWM.

The results of the three chlorophyll a regressions appear on the right side of Table 22. The Edge Effect, Proximity Effect, and Area Effect are all supported in these regressions and all of the water quality interaction terms are significant at the 99% level. The non-lake related coefficients all conform to the expected signs. Furthermore, their magnitudes are very similar to those in the TSI regressions, illustrating stability over specifications. For instance, the coefficients on the *Ln(Bath)* variable in all TSI and chlorophyll a regressions are within 0.001 of 0.120. The chlorophyll a coefficients are also stable across spatial and non-spatial specifications, with only minor differences in magnitude and almost all standard errors are smaller.

Table 23 contains the results of the TN and TP regressions. The TN regressions exhibit the same general characteristics as the TSI and chlorophyll a regressions. The Edge Effect, Proximity Effect, and Area Effect all hold and the nutrient variable and its interactions are all significant at the 99% level. Also, ρ has the same significance and magnitude as the TSI and Chlorophyll a regressions, and the main difference between the spatial and non-spatial models is a decrease in standard error. In the TP regressions, however, there are some notable differences with the other models. First, the Area Effect is not supported in any of the three TP models as the area and TP interaction term is not significant. The effect of TP on property prices does not change with the size of the lake, contrary to expectations.

In Table 24 and Table 25 the implicit prices for the nutrient regressions appear. Due to differences in units, these variables are calculated differently than SDM. The implicit price of SDM was expressed in one unit, or one foot, of SDM. A one unit change in each of the alternative indicators implies a much different change; since TSI is based on a scale of 1-100 a unit change is miniscule, while a one unit change in TP is enormous (the mean TP level is 0.036). It is important that the changes in lake quality not be too large; non-marginal changes may cause the demand function for housing and environmental attributes to shift (Haab and McConnell 2002). While recognizing that the difference between the various water quality indicators will cause any choice of units to be rather imperfect, the implicit prices of the alternative water quality measures are instead expressed in percentage changes. As a one unit change in SDM corresponds to approximately 17% of the mean level of SDM, the same percentage change is used for the other indicators.

The TSI implicit prices appear in the first three columns of Table 24, where the most visible changes from the SDM estimates are the missing values in the table. The blanks in the table correspond to implicit prices that have dropped to zero, where a change in water quality is no longer valued. For all of the TSI models the non-waterfront implicit price evaluated at the average distance from the lake is zero. In fact, the implicit price for TSI drops to zero at approximately 441 meters from the lake; considerably shorter than the extent of the implicit price of SDM, which was positive beyond the 1000 meter boundary. On the other hand, the implicit price of TSI is relatively high for lakefront homes and non-lakefront homes within 100 feet of the lake; they are both almost twice as large as the SDM implicit prices of $11,784.28 for waterfront homes and $1,786.31 for non-waterfront homes 100 meters from the lake. These differences in implicit price magnitudes and extent may be due to the desired use of the lake. Waterfront property owners have a greater interest in maintaining the overall health of the lake for multiple uses. Conversely, those located farther from the lake may have more topical interests in the lake, based more on view and aesthetic benefits. SDM is therefore more “valuable” to properties located beyond 500 meters from the lake than TSI.

The chlorophyll a implicit prices appear in the third through sixth columns of Table 24. While the implicit price of chlorophyll a is still positive for the mean distance non-lakefront home, these implicit prices are quite small. The extent of improved chlorophyll a benefits completely diminishes at about 500 meters from the lake. All of the chlorophyll a implicit prices are less than the SDM implicit prices.

Table 25 contains the implicit prices of the total nitrogen and total phosphorous regressions. The benefits of a 17% improvement of TN extend farther than the TSI and chlorophyll a benefits. Although the mean waterfront implicit price of TN is greater than the waterfront SDM implicit price, the mean non-waterfront implicit price for the spatial model with a distance/time SWM is $361, almost half the value of the non-waterfront implicit price of SDM. Also, the extent of TN benefits is less than the extent of SDM benefits. The phosphorous implicit prices appear in the last three columns of Table 25. The extent of TP benefits is even less than those of TSI. The implicit price of TP has a steep gradient: the waterfront implicit price is larger than that of SDM, but the implicit price diminishes to zero closer to the lake.

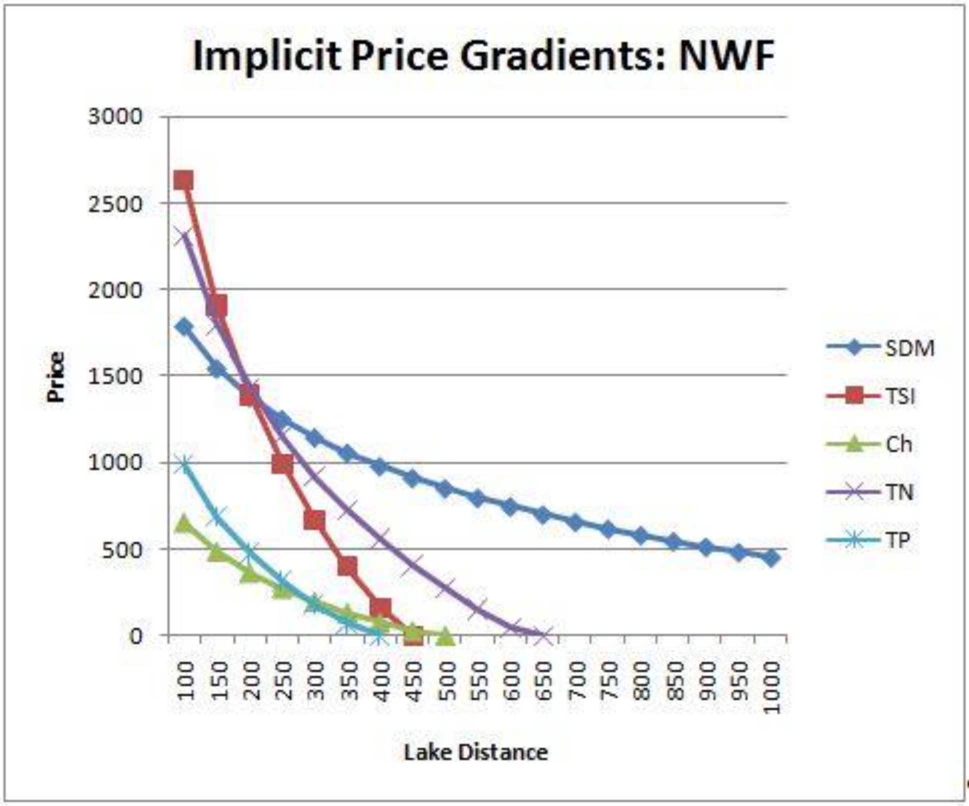

Figure 5: Implicit Price Gradients: Non-Waterfront

Figure 5 contains a graph of the non-waterfront implicit prices and lake distance, where the differences in the slopes of the gradients are clearly illustrated. SDM begins relatively high and gradually decreases. On the other hand, TSI has the highest y-intercept but the steepest slope, rapidly decreasing to zero at approximately 441 meters from the lake. The implicit price gradient of chlorophyll a has the most similar slope to SDM. This is not too surprising since Hoyer et al. (2002) show that chlorophyll a is the largest contributor to lake clarity. It is, however, surprising that the chlorophyll a curve has such a low y-intercept and extent. One explanation for this may be the popularity of fishing in central Florida lakes. Florida is regularly referred to as the fishing capital of the world, and central Florida attracts numerous anglers each year. Hoyer, Brown, and (Canfield Jr. 2004) state that citizens who are interested in fishing may view green water (a symptom of high chlorophyll a levels) as a supportive habitat for healthy fish. Also, in their survey of Florida lake users they find that people who live in lake areas characterized by green waters are more likely to be tolerant of it, dulling the effect of chlorophyll a on property prices.

The total benefits implied by these implicit prices appear in Table 26. Similar to Chapter 1, three lakes of different sizes are used as examples. The first row contains the benefits for a one foot increase in SDM. As this represents a much larger percentage change in Lake Mann than the other two lakes, the second row contains the total benefits of a 17% increase in SDM. The next four rows contain the benefits of a 17% improvement in the other indicators, with lake descriptive statistics appearing in the lower rows. Benefits of the increase in water quality are separated into three categories: waterfront benefits, non-waterfront benefits and total benefits. Due to the broad extent of its benefits, SDM has larger total non-lakefront benefits than the other indicators. Alternatively, since TSI had the highest implicit price for homes nearest the lake, improvements in TSI yield the largest lakefront benefits in all lakes. In each lake, either SDM or TSI yields the highest level of benefits. In all lakes chlorophyll a yields the lowest benefits for a 17% improvement.

The above results show that different water quality indicators can result in substantially different benefit estimates, implicit price magnitudes and benefit extent. Although all of the indicators exhibit the Proximity Effect, the extent of benefits is dependent on the indicator. The Area Effect varies over indicator choice; this hypothesis was not present in the TP regressions. Furthermore, benefits from the physical indicators are concentrated closer to the lake, while SDM benefits extend past 1,000 meters from the lake. These large differences in results indicate that the sole use of SDM in a hedonic analysis of water quality may not fully capture the full range of benefits from a water quality improvement. The shapes of the gradients, along with evidence from past papers, indicate that changes in SDM are valued for aesthetic and visual benefits. Conversely, improvements in physical measures condensed in the TSI are valued for deeper ecosystem and lake recreation benefits. A proper benefit cost analysis of water quality improvements should account for all of these benefits. If SDM is the sole measure of water quality in a hedonic analysis, the resulting estimate of benefits will improperly discount the values of homes directly near the lake.

3.6. Conclusion and Policy Discussion

Current Florida water quality policy weighs improvements in nutrient indicators much more heavily than improvements in clarity (FDEP 2009). While these changes are not mutually exclusive, results of the present study show that they may imply considerably different economic impacts. While the ecological benefits of transparency are suspect (Leggett and Bockstael 2000), the economic benefits of improvements in lake clarity are hard to ignore. Consequently, lake management activities that improve nutrient levels may have unintended welfare impacts on the local community. Since lake management programs are funded by local taxpayers, the economic impacts of transparency changes should be given more consideration. In contrast, the current hedonic water quality literature may be placing too much weight on transparency indicators. In certain areas these indicators may not fully reflect important ecosystem and recreation services.

This section uses several different indicators of water quality in a spatial hedonic analysis. Implicit prices are computed for all of the indicators and issues of benefit extent and total benefits are explored. Instead of finding an optimal indicator for all situations, results indicate that the use of at least two types of indicators may capture a larger range of the true total benefits.

All indicators examined were found to be significantly related to property prices. Results show that the two indicators that yield the greatest total benefits are SDM and TSI. TSI has a large effect on waterfront homes and homes close to the lake but has a short extent, diminishing just before 500 meters away. The implicit price gradient of SDM is less steeply sloped and the implicit prices of waterfront homeowners are not as high as those from TSI, but the extent of SDM benefits is much longer, extending past 1,000 meters. One explanation of these differences in the magnitude and extent of benefits is related to lake usage. Property owners close to the lake value the eutrophic status of the lake, which supports a variety of uses and amenities. Conversely, property owners further from a lake have less of an interest in the eutrophic state of the lake and more of an interest in aesthetic and purely recreation-based activities. Alternatively, property owners closer to the lake may have a greater incentive to obtain information about physical indicators, while owners located farther away base their purchasing decision on visibility.

TSI has only been used previously in studies of recreation demand (Jeon et al. 2005). The present section represents the first time this indicator has been used in a hedonic analysis of water quality. This is an important step since this is the main indicator used by the state of Florida (and several other states) to classify and regulate lakes. When lake managers aim to improve the water quality in a particular lake, one of their main targets is TSI, illustrating the importance of this indicator. TSI holds considerable potential for future hedonic analysis. It is measured on a 100-point scale that can be adapted to different states and settings and was created to condense several other indicators of eutrophication into one convenient scale (Carlson 1977).

Furthermore, since it is based on several measures of water chemistry and based on the work of limnologists, it is an objective indicator.

Alternatively, SDM has received the most attention in hedonic analyses of water quality. Although there are concerns with dark water Florida, variation in SDM is significantly related to property prices and yields higher total benefits than each of the individual components of TSI. Given the solid performance of both indicators, future studies should consider both in hedonic property price analysis. Instead of focusing on a single indicator to estimate benefits it may be better to incorporate both. Since the concept of water quality may vary over different portions of the population (Hoyer, Brown, and Canfield Jr. 2004), the use of two indicators may capture a broader estimate of benefits.

Concerns with water quality definitions point to some of the fundamental problems in valuing water quality. Quality is a different concept to different portions of the population, and is dependent on the designated use of the lake (Hoyer, Brown, and Canfield Jr. 2004). These differences have been incorporated into the foundation of the CWA through the assignment of differing designated uses to different lakes and should be integrated into cost benefit analyses of water quality changes. Future hedonic analyses of water quality should use both visibility and physical measures to reflect differences in designated uses and perceptions. Similar to air quality studies that differentiate between health benefits and visibility benefits (Delucchi, Murphy, and McCubbin 2002), water quality hedonic studies should recognize the differences between aesthetic visibility values and ecosystem or environmental values.

The recognition of these different categories of values is also important for lake management officials. An optimal policy should weigh the full social costs and benefits. If regulators focus solely on the ecosystem benefits of a policy and ignore the economic benefits of improved visibility there may be potential gains left on the table. For instance, TSI improvements can be made through changes in its three individual components. By evaluating which changes will yield the desired TSI improvement while simultaneously increasing visibility, lake managers can maximize the positive impact of lake programs.

Future research should include stated preference surveys to elicit more accurate data on consumers' water quality preferences and perceptions. Similar to the stated preference work in the air quality literature that differentiated between health benefits and visibility benefits, water quality surveys should differentiate between aesthetic benefits and ecosystem or environmental benefits. These surveys could guide decisions about the proper set of indicators to use in a hedonic property price regression.

Tables and Figures: Chapter 3

Table 22: Regression Coefficients: TSI, Chlorophyll a

	TSI			Chlorophyll a		
	Non-Spatial	NN15	Dist/Time	Non-Spatial	NN15	Dist/Time
WQ Variable	-0.205*	-0.216*	-0.209*	-0.025***	-0.026***	-0.026***
	(0.050)	(0.048)	(0.048)	(0.015)	(0.014)	(0.014)
WQ*Waterfront	-0.161*	-0.160*	-0.161*	-0.063*	-0.062*	-0.063*
	(0.026)	(0.016)	(0.016)	(0.008)	(0.005)	(0.005)
WQ*Ln(Distance Lake)	0.052*	0.052*	0.052*	0.012*	0.012*	0.012*
	(0.004)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
WQ*Ln(Area)	-0.008**	-0.007**	-0.008**	-0.004*	-0.004*	-0.004*
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Waterfront	0.854*	0.851*	0.858*	0.365*	0.366*	0.366*
	(0.094)	(0.059)	(0.059)	(0.016)	(0.010)	(0.010)
Ln(Distance Lake)	-0.251*	-0.251*	-0.253*	-0.081*	-0.081*	-0.081*
	(0.017)	(0.013)	(0.013)	(0.003)	(0.002)	(0.002)
Canalfront	0.096*	0.097*	0.096*	0.101*	0.103*	0.101*
	(0.017)	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)
Golffront	0.336*	0.335*	0.338*	0.335*	0.333*	0.336*
	(0.038)	(0.033)	(0.033)	(0.038)	(0.033)	(0.033)
Ln(Bath)	0.121*	0.119*	0.121*	0.120*	0.119*	0.120*
	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)
Pool	0.014*	0.013*	0.014*	0.014*	0.013*	0.014*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Home Age)	-0.071*	-0.072*	-0.071*	-0.071*	-0.072*	-0.070*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Home Area)	0.673*	0.675*	0.673*	0.675*	0.677*	0.674*
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Ln(Parcel Area)	0.164*	0.165*	0.165*	0.163*	0.164*	0.165*
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Ln(Distance CBD)	-0.196*	-0.195*	-0.195*	-0.194*	-0.193*	-0.193*
	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
Near Airport	0.029*	0.029*	0.029*	0.029*	0.028*	0.028*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln(Lattitude)	-6.360*	-6.569*	-6.426*	-6.046*	-6.265*	-6.107*
	(1.055)	(0.036)	(0.014)	(1.056)	(0.172)	(0.157)
Ln(Longitude)	3.528*	3.746*	3.609*	3.711*	3.930*	3.788*
	(0.691)	(0.659)	(0.661)	(0.692)	(0.644)	(0.647)
Ln(Median Income)	0.112*	0.112*	0.111*	0.111*	0.111*	0.110*
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
% White (2000)	0.223*	0.225	0.221*	0.224*	0.227*	0.222*
	(0.031)	(0.032)	(0.032)	(0.031)	(0.032)	(0.032)
% Black (2000)	0.039	0.043	0.039	0.041	0.044	0.041
	(0.035)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)
% Over 65 (2000)	0.040***	0.038**	0.041**	0.036	0.035***	0.037***
	(0.022)	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)

ρ	---	-0.003*	0.001*	---	-0.003*	0.001*
		(0.000)	(0.000)		(0.000)	(0.000)
R^2	0.8942	0.8947	0.8942	0.8940	0.8945	0.8939

*, **, and *** indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors appear in parentheses.

Table 23: Regression Coefficients: TN, TP

	TN			TP		
	Non-Spatial	NN15	Dist/Time	Non-Spatial	NN15	Dist/Time
WQ Variable	-0.120*	-0.130*	-0.118*	-0.108*	-0.121*	-0.110*
	(0.042)	(0.041)	(0.041)	(0.031)	(0.029)	(0.029)
WQ*Waterfront	-0.118*	-0.117*	-0.119*	-0.122*	-0.121*	-0.122*
	(0.025)	(0.019)	(0.019)	(0.012)	(0.009)	(0.009)
WQ*Ln(Distance Lake)	0.037*	0.037*	0.037*	0.021*	0.021*	0.022*
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
WQ*Ln(Area)	-0.008*	-0.008*	-0.009*	-0.001	0.000	-0.001
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Waterfront	0.238*	0.240*	0.239*	-0.223*	-0.220*	-0.223*
	(0.012)	(0.009)	(0.009)	(0.050)	(0.037)	(0.037)
Ln(Distance Lake)	-0.052*	-0.052*	-0.052*	0.020*	0.020*	0.021*
	(0.002)	(0.002)	(0.002)	(0.008)	(0.006)	(0.006)
Canalfront	0.098*	0.100*	0.099*	0.099*	0.100*	0.099*
	(0.017)	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)
Golffront	0.330*	0.329*	0.331*	0.334*	0.333*	0.336*
	(0.038)	(0.033)	(0.033)	(0.038)	(0.033)	(0.033)
Ln(Bath)	0.120*	0.119*	0.120*	0.120*	0.118*	0.120*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
Pool	0.014*	0.014*	0.014*	0.014*	0.013*	0.014*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Ln(Home Age)	-0.071*	-0.072*	-0.070*	-0.071*	-0.072*	-0.071*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Home Area)	0.672*	0.674*	0.671*	0.672*	0.675*	0.672*
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Ln(Parcel Area)	0.166*	0.166*	0.167*	0.164*	0.165*	0.165*
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Ln(Distance CBD)	-0.189*	-0.188*	-0.189*	-0.197*	-0.196*	-0.197*
	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
Near Airport	0.028*	0.027*	0.028*	0.033*	0.033*	0.033*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Ln(Lattitude)	-5.675*	-5.894*	-5.735*	-6.221*	-6.429*	-6.287*
	(1.059)	(0.193)	(0.174)	(1.054)	(0.119)	(0.096)
Ln(Longitude)	3.540*	3.766*	3.613*	3.616*	3.840*	3.696*
	(0.696)	(0.642)	(0.646)	(0.692)	(0.651)	(0.655)
Ln(Median Income)	0.108*	0.108*	0.107*	0.111*	0.111*	0.110*
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
% White (2000)	0.258*	0.261*	0.256*	0.225*	0.228*	0.223*
	(0.031)	(0.032)	(0.032)	(0.031)	(0.032)	(0.032)
% Black (2000)	0.067***	0.070**	0.067**	0.045	0.049	0.045

	(0.035)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)
% Over 65 (2000)	0.014 0.022)	0.013 (0.020)	0.015 (0.020)	0.035*** (0.021)	0.034*** (0.020)	0.036*** (0.020)
ρ	---	-0.003* (0.000)	0.001* (0.000)	---	0.001* (0.000)	-0.003* (0.000)
R^2	0.8936	0.8941	0.8935	0.8942	0.8947	0.8942

*, **, and *** indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors appear in parentheses.

Table 24: Implicit Prices: TSI, Chlorophyll a

	TSI 17% Improvement in Water Quality			Chlorophyll a 17% Improvement in Water Quality		
	Non-Spatial	NN15	Dist/Time	Non-Spatial	NN15	Dist/Time
WQ: Mean WF	22,108.74 (1,823.18)	21,877.70 (1,157.72)	22,235.27 (1,164.58)	7,269.29 (575.96)	7,252.65 (397.76)	7,294.69 (400.11)
WQ: Mean NWF	---	---	---	17.87 (80.86)	29.24 (66.31)	12.10 (66.70)
WQ 100m	2,612.54 (304.68)	2,585.90 (251.00)	2,626.72 (252.48)	651.19 (104.37)	659.38 (80.80)	652.82 (81.27)
WQ 300m	677.91 (215.88)	656.80 (202.74)	672.84 (203.93)	199.81 (82.06)	210.27 (65.84)	196.17 (66.23)
WQ 500m	---	---	---	---	1.45 (66.74)	---
WQ 700m	---	---	---	---	---	---
WQ 900m	---	---	---	---	---	---
Waterfront	115,612.40 (4,461.77)	115,897.04 (3,196.53)	116,439.02 (3,217.49)	102,945.40 (4,857.12)	103,286.80 (3,656.63)	103,726.77 (3,680.11)
Lake Proximity	24.95 (0.68)	24.94 (0.62)	25.11 (0.63)	23.16 (0.76)	23.16 (0.69)	23.28 (0.69)

Standard errors appear in parentheses, obtained using the delta method.

Table 25: Implicit Prices: TN, TP

	TN 17% Improvement in Water Quality			TP 17% Improvement in Water Quality		
	Non-Spatial	NN15	Dist/Time	Non-Spatial	NN15	Dist/Time
WQ WF	17,129.74 (1,818.46)	16,854.86 (1,377.73)	17,264.24 (1,385.82)	13,062.97 (954.72)	12,808.45 (715.39)	13,136.85 (719.49)
WQ NWF	354.54 (203.74)	298.71 (200.78)	361.17 (201.91)	---	---	---
WQ 100m	2,296.92 (283.14)	2,241.79 (248.99)	2,309.08 (250.41)	978.85 (175.85)	922.45 (158.35)	987.09 (159.24)
WQ 300m	912.57 (208.08)	856.94 (198.71)	920.78 (199.83)	180.52 (143.41)	128.83 (140.87)	179.80 (141.65)
WQ 500m	268.88 (204.70)	213.02 (202.34)	275.26 (203.48)	---	---	---
WQ 700m	---	---	---	---	---	---
WQ 900m	---	---	---	---	---	---
Waterfront	118,634.20 (4,621.78)	118,911.82 (3,242.98)	119,377.74 (3,263.89)	91,734.62 (5,010.16)	92,194.68 (3,806.71)	92,454.68 (3,830.39)
Lake Proximity	25.55 (0.71)	25.53 (0.62)	25.70 (0.63)	23.23 (0.70)	23.23 (0.66)	23.37 (0.67)

Standard errors appear in parentheses, obtained using the delta method.

Table 26: Total Benefits of an Increase in Water Quality

	Lake Silver (70 Acres)			Lake Mann (271 acres)			Lake Conway (1000 acres)		
WQ Indicator	WF Benefits	NWF Benefits	Total Benefits	WF Benefits	NWF Benefits	Total Benefits	WF Benefits	NWF Benefits	Total Benefits
SDM*	971,452	1,079,068	2,050,520	698,144	3,219,550	3,917,694	5,441,246	7,433,382	12,874,628
SDM (17%)	950,586	1,055,890	2,006,476	255,053	1,176,198	1,431,251	9,461,946	12,926,130	22,388,076
TSI*	1,925,818	650,091	2,575,909	478,823	466,424	945,247	15,421,028	6,905,877	22,326,905
Total Nitrogen	1,456,301	635,259	2,091,560	368,364	562,615	930,979	12,262,230	8,174,365	20,436,595
Total Phosphorous	1,174,163	263,384	1,437,547	285,331	155,257	440,588	9,208,511	2,055,284	11,263,795
Chlorophyll	613,778	130,846	744,624	155,311	125,104	280,415	5,290,549	2,309,189	7,599,738
Number Homes	62	1,928	1,990	63	2,635	2,698	669	6,300	6,969
Average SDM		5.756			2.149			10.229	
Average TSI		46.235			57.308			27.580	
Average TN		0.718			1.439			0.529	
Average TP		0.022			0.030			0.009	
Average Chlor		7.663			17.178			1.417	

*For the first SDM row, a one foot increase is used. For the other rows a 17% increase is used. Estimates from the distance/time SWM spatial lag model used for calculation.

CHAPTER 4: WATER QUALITY IN A HYBRID REPEAT SALES MODEL

4.1. Introduction

The focus of the first three chapters of the dissertation is hedonic modeling. Hedonic property models represent a popular revealed preference method of environmental valuation. However, there have been some critiques of the approach which deserve some attention. This chapter uses an alternative method to estimate the value of water quality, which also uses property prices: a repeat sales model. A pervasive concern in hedonic property analysis is omitted variable bias, which arises when important variables are excluded from the hedonic equation. For instance, over the population there may be significant heterogeneity in people's tastes for environmental quality which may cause them to self select into some areas instead of others (Chay and Greenstone 2005). Some problems with omitted variables may be muted by the use of spatial models that control for spatially correlated omitted variables (Brasington and Hite 2005), but there is no guarantee that it will be a full solution. Variables may be excluded for a variety of reasons, including lack of data, measurement problems or empirical constraints. No matter the reason, these omitted variables can result in biased coefficients and implicit prices (Palmquist 1982). The repeat sales model controls for omitted variable bias and can be used to estimate the value of water quality.

Due to the size of the data set in this paper, there are a substantial number of homes that have sold more than once. By analyzing differences in home price for the same home over repeat sales, it is possible to control for the effect of omitted variables (Case and Quigley 1991). This type of analysis has only been used to analyze a small number of environmental problems; in each case the issue was a discrete event such as the recognition of hazardous waste contamination. The present section represents an attempt to investigate a continuously changing environmental amenity using a hybrid repeat sales/hedonic analysis. The goal of this section is to see if water quality has an impact on non-waterfront homes after controlling for omitted variable bias, and to investigate further the Edge Effect, Proximity Effect, and Area Effect.

4.2. Literature Review

There are two main areas where repeat property sales have been used to analyze the effect of environmental amenities (or disamenities). The first area deals with the construction of real estate price indices, and has evolved to incorporate environmental effects. The second area uses repeat sales in "quasi-experimental" analyses. The former has gradually evolved over the last 40 years and has recently combined aspects of regular hedonic analysis in order to include home attributes that change over time. The latter is a newer approach that uses difference-in-difference estimators to assess the impact of a quasi-random event, such as the landfall of a hurricane. Due to its ability to accommodate changing attributes, this paper uses the price index approach to estimate the impact of water quality on repeat home sales.

The earliest application of repeat sales was in Bailey, Muth, and Nourse (1963), who were interested in controlling for quality differences between homes to obtain a more accurate construction of price indices. They proposed a method that uses sales of the same property at different time in a simple regression framework. Ordinary least squares regression is used to estimate the following equation:

$$r_{itt'} = \sum_{j=1}^{T} b_j x_j + u_{itt'} \qquad (4.1)$$

Where $r_{itt'} = \ln(p_{it})\text{-}\ln(p_{it'})$, or the difference between two sales of the same home taking place in years t and t'. The x_j variables are non-binary dummy variables for each year, equal to -1 in the *t*-th year (the initial year) and 1 in the *t'*-th year. This method is attractive because the use of repeat sales controls for many unobserved attributes. A large literature has been spawned from this early paper, and most official home price indices now use a variation of this basic approach (Shiller 1991).

Palmquist (1982) is the first paper to combine hedonic analysis with repeat sales in a hybrid model. Palmquist recognized the potential efficiency in using repeat sales due to the ability to control for omitted variables. He proposed a model that combined the price index construction of Bailey, Muth, and Nourse (1963) with environmental variables that change over time. The environmental issue analyzed in Palmquist (1982) is traffic noise. In an area north of Seattle, Washington a large highway was constructed, rapidly increasing noise levels. Home prices that occurred before and after the highway was built were incorporated into the model, and corrections were made for error correlation resulting from homes that sold more than twice (discussed in the methods section).

Case and Quigley (1991) were concerned with the drop in sample size resulting from the use of repeat sales. They propose an estimation method that allows the use of both repeat sales and single sales, in a hybrid variation of Palmquist (1982). This method is shown to produce tighter confidence intervals as compared to the hedonic or repeat sales estimates alone. However, their sample size is rather small with 418 total property sales and only 47 repeat sales. It is therefore no surprise that the repeat sales estimate of the price index has a relatively wide confidence interval.

Shiller (1993) also estimates a hybrid repeat sales/hedonic model that incorporates changes in attributes between sales. Shiller was concerned with quality differences between observations which may not be recognized in traditional repeat sales measures. An example in the derivatives market is used to illustrate the construction of price indices. Shiller's model focuses on including additional attributes that vary over time into the repeat sales model. This was a critical step in bridging the gap between hedonic and repeat sales models, and set the foundation for an analysis of environmental issues that change over time.

Case et al. (2006) use a hybrid repeat sales/hedonic model to analyze environmental externalities. Their unique dataset on groundwater contamination in Scottsdale Arizona was the result of the court case *Baker vs. Motorola*, in which many homeowners sued a manufacturer after the discovery of long term contamination. They use condominium sales, which limits the changes in structural attributes. Their basic model appears in equation (4.2):

$$\ln(P_i) - \ln(\tilde{P}_i) = \ln(\gamma) - \ln(\tilde{\gamma}) + \tau_1(T_{i1} - \tilde{T}_{i1}) + \tau_2(T_{i2} - \tilde{T}_{i2}) + ... + \tau_n(T_{in} - \tilde{T}_{in}) + \beta_0 CNTM(\phi_{82,i} - \tilde{\phi}_{82,i}) \tag{4.2}$$

In this equation, the *T* variables represent the year dummy variables from Bailey, Muth, and Nourse (1963), where a tilde above the variable indicates the initial sale. The CNTM variable represents contamination level, and varied between several models corresponding to different risk perception interpretations. These contamination variables distinguish the hybrid model from traditional repeat sales methods. After estimating multiple models, results indicate that the contaminated groundwater had a negative and significant effect on property prices, varying from two percent of home value to 15 percent of home value. Additionally, the effect of the groundwater contamination on property prices did not fully mature until several years after discovery.

The other area of the hedonic literature that uses repeat sales is the group of “quasi-experimental” analyses of environmental amenities. The main goal of these papers is to use a “random” event as a natural experiment and analyze the effect on property prices. For example, Chay and Greenstone (2005) estimate the marginal willingness to pay for air quality in the

United States, using a national sample. Under the amendments to the CAA, individual counties were designated as either “attainment” or “non-attainment” based on their levels of air quality. Using counties that are just near this cutoff, they use the designation of this status as the “random event.” The main regression in this paper is a semi-log model that uses first-differenced data between the years 1970 and 1980 since the event occurred between these years. Property prices are aggregated by county instead of using individual home sales. Several forms of the regression are explored that confront endogeneity resulting from both omitted variables and sorting, and their results indicate that omitted variable bias is a much bigger issue to control for than selectivity bias.

In a paper about health risk and housing values, Davis (2004) uses a difference in difference estimator to examine the effect of an increase in cancer risk. The random event in the natural experiment of this paper was a sudden public awareness of a large increase in cases of pediatric leukemia. The cause of the leukemia is currently unknown. The difference in difference estimator was employed to mitigate concerns with preference-based sorting and omitted variable bias. The increased cancer risk was found to decrease home values in the affected area by as much as 12-16 percent.

Hallstrom and Smith (2005) use the landfall of Hurricane Andrew to define a quasi-random experiment in Lee County Florida. This particular county was a “near-miss,” so was not directly damaged by the hurricane. A difference in difference model is used to estimate the effect of a perceived increase in hurricane risk on property prices. This model uses repeat sales with an initial sale before the event and a final sale after the hurricane passed. Homes both in and out of FEMA flooding hazards are used to control for differences in severity, and a fixed effects regression is used. They find that Andrew lead to an average 19 percent decrease in the sale prices of homes contained within flood hazard areas.

Pope (2008) uses a difference in difference analysis with repeat sales to investigate airport noise in North Carolina. A regulation was passed in Wake County, North Carolina that required home sellers to provide buyers a disclosure about airport noise. The timing of this regulation was treated as the random event in the quasi-experiment. In the area classified by high noise, Pope finds that the disclosure decreased home values by 2.9 percent.

Parmeter and Pope (2009) summarize the difference in difference/quasi-experimental approach to hedonic analysis. This paper critiques several of the above papers and provides recommendations for how to conduct a similar study. Ten steps are provided to ensure the consistency and reliability of further quasi-experimental studies. This paper illustrates the value of repeat sales analysis in the context of a quasi random event, and also discusses some of the important limitations.

The use of repeat sales can potentially have advantages over a traditional hedonic analysis due to its ability to control for omitted factors. However, repeat sales data has several problems. Gatzlaff and Haurin (1997) examine issues of sample selection bias in repeat sales index construction. They stress that while repeat sales analysis holds constant home quality, the sales may have taken place in different economic conditions. They find that bias in repeat sales indices is directly related to variance in economic conditions between sales. Another recognized problem is that the use of repeat sales restricts the sample to be much smaller than the full group of home sales. This smaller sample may be non-random and overrepresented by “lemons” (Case and Quigley 1991). Furthermore, the repeat sales model assumes that the implicit prices of variables remain constant over time; otherwise the attributes would not cancel out between sales.

The use of repeat sales in environmental valuation holds significant potential. The ability to mitigate omitted variable bias and preference-based sorting is a promising strength of this approach. As both the quasi-experimental approach and the hybrid approach are still relatively new, they represent two future paths of the hedonic literature. Both types of models have performed well in the face of discrete events; the current paper will analyze the model's ability to analyze a non-discrete change in an environmental amenity.

4.3. Data

The Orange County, Florida housing dataset from the past chapters serves as the starting point for the data used in this chapter. First, all homes that sold only once are eliminated from the database leaving 23,381 properties. Since the time units of the water quality variable is years, homes that sold twice within the same year are not used in the analysis. This is a common procedure in the repeat sales literature, and may help reduce the number of outliers (Case et al. 2006). After removing these observations (along with the other home sales that are now no longer repeat sales) there are 21,732 observations. The new dataset is obtained by differencing between sales, leaving 11,785 observations. After these elimination steps, the proportion of waterfront homes in the data drops considerably, leaving only slightly over 1% of the data composed of unique waterfront home sales.

Summary statistics for the repeat sales data appear in Table 27. Due to the small number of waterfront homes, summary statistics for both groups are combined.[19] The average sale price in these data is over $55,000 more than the average non-lakefront home price in the full dataset. [20] Conversely, the average home is a similar size and age as those in the full data. The other

[19] Deflated sales prices (2001) are used in these data, as in the previous chapters.
[20] Excluding the waterfront homes only decreases this average by $7,000, still much larger than before.

attributes of the average home in these data are comparable to the attributes of the average non-lakefront home in the full data. Nonetheless, the difference in average home prices and the proportion of waterfront homes is a concern.

4.4. Methods

4.4.1. Models

Of the two types of repeat sales analysis presented in the literature review, the price index-based hybrid approach represents the only feasible option for the present analysis. The quasi-experimental/difference in difference method requires the occurrence of a discrete random event, such as a hurricane, the passage of a law, or a sudden increase in media coverage of a health risk. In the present setting the environmental amenity continuously changes over the time period, excluding a difference-in-difference analysis from the set of econometric approaches.

The original repeat sales model appears in Bailey, Muth, and Nourse (1963), where the objective was to create accurate price indices. Their analysis starts with the basic hedonic property model appearing in equation (4.3):

$$P_{it} = \sum_{j=1}^{J} \beta_j x_{jit} + \sum_{t=1}^{T} c_t D_t + e_{it} \tag{4.3}$$

In this model, P_{it} represents the natural log of the sales price of home i during the time period t, β_j is a vector of estimated coefficients corresponding to the property attributes, and c_t is a vector of coefficients for each of the binary time dummy variables. When repeat sales are introduced, the natural log of the initial sale is subtracted from the natural log of the second sale, yielding equation (4.4):

$$P_i^2 - P_i^1 = \left(\sum_{j=1}^{J} \beta_j x_{ji}^2 + \sum_{t=1}^{T} c_t D_t^2 \right) - \left(\sum_{j=1}^{J} \beta_j x_{ji}^1 + \sum_{t=1}^{T} c_t D_t^1 \right) + e^{21} \tag{4.4}$$

The basic repeat sales model of Bailey, Muth, and Nourse (1963) assumes that property attributes do not change over time, so the *x* variables cancel out, leaving the common repeat sales equation (4.5):

$$P_i^2 - P_i^1 = \sum_{t=1}^{T} c_t (D_i^2 - D_i^1) + e_i^{21} \tag{4.5}$$

To empirically estimate this equation, the dependent variable is the difference in the natural logs of the two property prices. For the independent variables, the time dummy variables are set to -1 for the initial sale date, +1 at the final sale date, and zero otherwise. The simplicity of (4.5) is one of its primary advantages (Shiller 1993). Several papers have since updated this model to include an intercept (Case et al. 2006), as in:

$$\ln(P_i) - \ln(\tilde{P}_i) = \ln\left(\frac{\gamma}{\tilde{\gamma}}\right) + \tau_1(T_{i1} - \tilde{T}_{i1}) + \cdots + \tau_n(T_{in} - \tilde{T}_{in}) \tag{4.6}$$

where (4.6) uses the notation of Case et al. (2006) illustrated in the literature review. Note that this equation contains the intercept term, $\ln\left(\gamma/\hat{\gamma}\right)$, since these intercepts may not cancel out. Case et al. 2006 state that this term should capture nontemporal aspects of home price appreciation, which may result from changes in the physical attributes of the property. Equation (4.6) will serve as the baseline model for the present section, and is estimated to give an illustration of the price changes that occurred during the sample period. As in equation (4.2), a tilde above a variable denotes the initial sale.

To include water quality in the analysis there are two main options. One basic way is to use the difference in SDM between sales as an explanatory variable. This hybrid repeat sales/hedonic regression is the second model estimated, which introduces the variables that change over time:

$$\ln(P_i) - \ln(\tilde{P}_i) = \ln\left(\frac{\gamma}{\tilde{\gamma}}\right) + \tau_1(T_{i1} - \tilde{T}_{i1}) + \cdots + \tau_n(T_{in} - \tilde{T}_{in}) + \beta_1(S_{it} - \tilde{S}_{it}) + \beta_2 I \qquad (4.7)$$

Two additional variables appear in this regression: the difference in SDM, $(S_{it} - \tilde{S}_{it})$, and the value of home improvements between sales, *I*. The difference in SDM is calculated as the difference in annual mean SDM between sales. The home improvement variable is provided by the Orange County Property Appraisers Office for each home sale.

There are some important limitations of this second model that must be acknowledged. Since SDM enters as a differenced variable, it omits information about the initial level of SDM. As discussed in Chapter 1, a one foot change in SDM may be much less visible in a relatively clear lake than a lake with limited visibility. While the use of repeat sales data confronts problems of omitted variable bias, there may be other costs related to this loss of information. These costs are present in the investigation of the Edge, Proximity, and Area Effects. Initial attempts to include these three effects as interaction variables resulted in counter-intuitive or insignificant signs.

To further investigate the effect of distance and area on implicit prices, several dummy interaction terms are used which allow the implicit price to vary over different groups. For example, to examine the Proximity Effect, four dummy interaction variables are used:

$$\ln(P_i) - \ln(\tilde{P}_i) = \ln\left(\frac{\gamma}{\tilde{\gamma}}\right) + \tau_1(T_{i1} - \tilde{T}_{i1}) + \cdots + \tau_n(T_{in} - \tilde{T}_{in}) + \beta_1 WFs_i + \beta_2 300s_i + \beta_3 700s_i + \beta_4 1000s_i + \beta_5 I \qquad (4.8)$$

Where $300s_i$ equals 1 if the home is within 300 meters of a lake and SDM improved between the two property sales.

To investigate the Area Effect, homes are split into four groups, based on an equal split of the data. These are small lakes (under 30 acres), medium size lakes (between 30 and 100 acres) large lakes (100-250 acres) and extra large lakes (greater than 250 acres). Similar to (4.8), these are interacted with a dummy variable that indicates if the level of SDM improved between the two sales. This formulation appears in equation (4.9):

$$\ln(P_i) - \ln(\tilde{P}_i) = \ln\left(\frac{\gamma}{\tilde{\gamma}}\right) + \tau_1(T_{i1} - \tilde{T}_{i1}) + \cdots + \tau_n(T_{in} - \tilde{T}_{in}) + \beta_1 SMs_i + \beta_2 MEDs_i + \beta_3 LGs_i + \beta_4 ELs_i + \beta_5 I \tag{4.9}$$

The use of these interaction terms should allow an analysis of the Edge, Proximity, and Area Effects and an isolation of potential nonlinearities in the implicit price gradient.

4.4.2. Error Corrections

When using repeat sales data, it is important to recognize correlations between error terms. If there are more than two repeat sales, the errors of these multiple sales will be correlated. A simple OLS regression on the differences in home sales will ignore the fact that these multiple sales come from the same home and failure to correct for this correlation will result in biased estimation (Case et al. 2006). Bailey, Muth, and Nourse (1963) proposed a weighted sales regression (generalized least squares) to control for these correlations and showed that the weighted method performed better at home price prediction than a regular regression. Under a set of assumptions about the correlation in the errors of the multiple sales, the variance-covariance matrix M of the residuals is known up to a scalar. With this estimator, Bailey, Muth, and Nourse (1963) show that the minimum variance linear unbiased estimator is

$$\hat{b} = (x'M^{-1}x)^{-1}(x'M^{-1}P) \tag{4.10}$$

This correction method is explained in both Palmquist (1982) and Case et al. (2006), and is based on Aitken (1935). This weighted regression method is used in the present chapter.

The original model of Bailey, Muth, and Nourse (1963) assumed that the variance of the error term is constant across homes, and under this assumption their log price index is the best linear unbiased estimate of the log of the price level. Case and Shiller (1987) refuted the validity of this assumption, and presented evidence that this variance is related to the interval of time between sales. Case and Shiller propose the use of a weighted, or generalized least squares (GLS), regression procedure, outlined in both Case and Shiller (1987) and Case and Shiller (1989). This weighted procedure, which corrects for heteroskedasticity due to the time period between sales, is composed of three steps. In the first step, the Bailey, Muth, and Nourse (1963) regression is performed, and a vector of regression residuals is computed. In the second step a regression of the residuals is performed on a constant and the time interval between sales. The third step consists of two parts. First, each observation is divided by the fitted value from the second step. Next, a GLS regression of the same form as the first step is performed with these transformed observations to obtain the new heteroskedasticity-corrected estimates.

The impacts of the GLS corrections are explored for each repeat sales model, resulting in three variations. The first is a basic OLS model, the second regression corrects for heteroskedasticity due to differences in time between sales, and the third uses both the heteroskedasticity correction and the Aitken (1935) error adjustment to correct for homes that sold more than twice. These variations should illustrate the effect of water quality on home prices while controlling for a variety of omitted influences.

4.5. Results

Results of the OLS regressions appear in Table 28.[21] The first column contains the basic repeat sales model, which only uses the non-binary dummy variables. The adjusted R^2 is 31.7%, and the hypothesis that all annual price changes are zero is rejected ($F_{8,\ 11776} = 683.55$). The coefficients on these variables exhibit similar qualitative characteristics as the binary time dummy variables in the other chapters; they are negative and significant in each period and increase monotonically over time. Again, the 2004 variable was excluded from the equation, so the coefficients are expected to be negative to reflect increasing home prices during this time period. To interpret the individual coefficients, the particular functional form must be accounted for. As the dependent variable is the difference in the natural log of home price, the 1996 coefficient means that the average home sold in 1996 for 47% less than the average home sold in 2004 ($1 - e^{-0.637} = 47.11\%$) (Case et al. 2006).[22] Likewise, the average home sold for 32.9% ($1 - e^{-0.399} = 32.90\%$) less in the year 2000 than the average home in 2004. Although this seems like a very large difference, the property price bubble that affected the United States during this time period was pervasive in the central Florida area.

The second column of Table 28 contains the estimates of the hybrid hedonic/repeat sales model, where the difference in SDM between sales is used as a variable. The adjusted R^2 of this model increases slightly to 32.7%, and an F test again indicates that the hypothesis that all

[21] All repeat sales models, including the GLS error correction models, are estimated in Stata.

[22] $-0.637 = \ln\left(\frac{P}{\tilde{P}}\right) - \ln\left(\frac{P'}{\tilde{P}}\right)$

$-0.637 = \ln\left(\frac{P}{P'}\right)$

$e^{-0.637} = \frac{P}{P'}$ Then $1 - e^{-0.637} = \frac{P' - P}{P'}$ since $P' > P$. Conversely, if 0.637 was the coefficient and $P > P'$ we have $e^{0.637} - 1 = \frac{P - P'}{P'}$

coefficients are zero is rejected ($F_{10,\,11774}$ = 574.25). The coefficients on the year variables remain negative, significant, and increasing in each year, and are similar to the basic repeat sales time coefficients.

The coefficient on the difference in SDM in column two of Table 28 is positive and significant. After controlling for omitted variable bias, improvements in water quality still have a positive effect on the average home. Furthermore, since 99% of the homes in the sample are non-waterfront, this model gives further evidence that those homes have a positive implicit price of water quality. The coefficient indicates that a one unit increase in SDM between two home sales results in an increase in the average home price of 2.33% ($e^{0.023}$-1 = 0.02326). Since the mean home sale for the 21,732 repeat sales is $217,836.90, this amounts to an increase in home value of $5,068.41. Since the average lake distance is 445 meters and the average lake area is 283 acres, this represents a significant increase from the implicit prices presented in the first chapter.[23]

The full coefficient on the home improvement variable in the second column of this table is -0.00000236, indicating that a one dollar improvement between home sales will cause a $e^{-0.00000236} - 1 = 0.0002359\%$ decrease in the home price. This negative result at first seems counter-intuitive, that improving homes decreases their value. However, it may be the case that the homes that needed improvements were in worse shape. An improvement in the form of an additional deck or room is probably much less common than fixing up various depreciated parts of an older home.

The third column of Table 28 contains the water quality interaction dummies that separate water quality by lake distance. The adjusted R^2 of this regression is 0.326, 0.001 smaller

[23] Significantly different based on a Student's t-test.

than the previous regression, and an *F* test rejects the hypothesis that all coefficients are zero ($F_{12,\,11772}$ = 474.77). The four water quality and distance variables display a positive but decreasing trend as lake distance increases, starting at 0.093 for waterfront homes and decreasing to 0.001 for homes located 700 to 1,000 meters from a lake. The first three interaction coefficients are significant at the 99% level, but the 1000 meter variable is not significant. These coefficients indicate that water quality has a larger effect on homes closer to the lake, which potentially tapers off beyond 700 meters from the lake. Evidence for the Proximity and Edge Effects are therefore found in the repeat sales data, although at a weaker level than in the full hedonic model (since they only appeared when the dummy interaction variables were used).

The Area Effect is not found in the repeat sales model. The fourth column of Table 28 contains the repeat sales model with the area and water quality improvement interactions. From this model, homes near smaller lakes have higher values than homes on larger lakes, contrary to the Area Effect. However, there are some clear non-linearities in the impact of lake size. The coefficient on *WQImproveMd* is smaller than both the *WQImproveSm* coefficient and the *WQImproveLg* coefficient. This strange behavior is still present when the distance interactions are also present (not contained in the table, as it did not add anything to the analysis), and may represent the impact of a few outliers. Meese and Wallace (1997) find that the repeat sales model is much more susceptible to the influence of outliers than the regular hedonic model. Nonetheless, this strange result opposes the previous hedonic results from this study and several others in the hedonic literature.

The GLS error corrections based on heteroskedasticity due to age appear in Table 29. The qualitative results of the basic repeat sales model are maintained in this table, although the coefficients on the time variables change slightly. Conversely, the other variables in the hybrid regression do not change very much. The main difference is a drop in the R^2 for each model. Finally, the models with both GLS error corrections appear in Table 30. The main effect of this additional error correction is to increase the water quality-related coefficients in all models. Also, the R^2's rise back to the levels of Table 28. The Area Effect is still not present, although the Edge and Proximity Effects remain.

After investigating several variations of the hybrid hedonic/repeat sales model, it is clear that several issues still need to be confronted before a variable such as water quality can be reliably analyzed. First, the average home that sold multiple times had a much higher selling price than the full set of home sales. This subset of data may not be a representative sample of homes in the central Florida area, particularly heading into the property bubble. Second, the proper representation of water quality in the model is not clear. Due to the temporal distance between sales, the water quality in the nearest lake may have been characterized by behavior not fully captured by the difference in SDM between sales. Future research may benefit from the use of water quality trends instead of point observations. Furthermore, while the hybrid repeat sales/hedonic model mitigates omitted variable bias, it washes away initial conditions that may be an important part of the analysis. For instance, a one foot change in water clarity in a poor quality lake is much more noticeable than a one foot change in a lake with an average of 12 feet of clarity. Tackling these issues should help with the fit of the model, which is currently quite low. Adjusted R^2 values in all of the repeat sales models were much lower than the full hedonic model. Similar results were found in Zabel and Kiel (2000), who attempted to use the hybrid repeat sales/hedonic model in an analysis of air quality but ended up not reporting the results of the regression due to its poor performance compared to the full hedonic model.

4.6. Conclusion

In order to address issues of omitted variable bias, this section pursues an alternative to the hedonic price model. This is done by analyzing differences in property prices between home sales. The subsample of homes that sold twice or more are used in a hybrid repeat sales/hedonic model. The results of these regressions are somewhat mixed. On a positive note, the model that uses the difference in SDM values between sales indicates that this variable is positively related to home prices, yielding an implicit price that is higher than the full hedonic models. Since 99% of the homes in the sample are non-lakefront, this supports one of the main assertions of this dissertation: non-lakefront homes should be included in an analysis of water quality.

In contrast, when the Edge, Proximity, and Area Effects are investigated, the results are not as clear. Attempts to include interactions with the difference in SDM and area and distance yielded insignificant and counter-intuitive results. Only the weaker interactions involving an indicator of a water quality improvement between sales illustrated the Proximity and Edge Effects. The Area Effect was not supported in any of the model variations. Furthermore, the repeat sales models were all characterized by low R^2 values.

There are several potential explanations for the poor performance of these models, which suggest that the hybrid repeat sales/hedonic model is still in need of substantial development. First, the sample of homes that sold more than once were characterized by much higher sales prices than the full dataset. Also, lakefront homes represented a much smaller proportion of the data. The causes for these data differences are unclear, but may have resulted in an over-representation of “lemons” or home “flipping.” Furthermore, additional techniques may be needed to use a continuous variable in a repeat sales model. The two main variations here may not accurately capture differences in water quality from the consumer's perspective. These problems may explain the dearth of studies that use this approach for environmental valuation.

Future research should use a difference in difference model to estimate the value of water quality. A quasi-experimental model would likely have greater success in identifying this value. A discrete change in water quality would be needed, with enough of a lasting impact to influence the local property market. Since there are billions of dollars spent on water quality improvement projects each year, this represents a promising avenue for further research. Additionally, more work needs to be done on the use of continuous environmental variables in a hybrid hedonic/repeat sales model. Since there has traditionally been more data available on air quality, a venture in that field may yield more robust results.

Tables and Figures: Section 3.3

Table 27: Repeat Sales Summary Statistics

Variable	Units	N= 11,785 Mean	Std. Dev.
Property Characteristics			
Sales Price	2002 Dollars	255,628.80	229,761.20
Heated Area	Square Feet	1,979.87	961.41
Area of Parcel	Square Feet	11,889.37	15,548.78
Number of Bedrooms	--	3.27	0.86
Number of Bathrooms	--	2.25	0.92
Home Age	Years	19.67	13.84
% With Pool	--	20.28	40.21
Spatial Characteristics			
Distance to Nearest Lake	Meters	444.40	263.54
Area of Nearest Lake	Acres	282.82	445.10
Distance to CBD	Meters	8,709.04	5,244.47
Latitude Coordinate	Degrees	653,192.90	7,341.98
Longitude Coordinate	Degrees	505,450.50	6,216.74
% In airport noise zone	--	18.02	38.44
Census Block Characteristics			
% of Population White	--	81.01	19.21
% of Population Black	--	9.84	17.36
% of Population > 65	--	11.64	7.67
Median Household Income	2002 Dollars	58,081.30	24,454.32

Table 28: Basic Repeat Sales Models

	Basic RS	Hybrid 1	Distance	Area
	β/(SE)	β/(SE)	β/(SE)	β/(SE)
$T_{96}^{2}-T_{96}^{1}$	-0.637*	-0.649*	-0.644*	-0.643*
	(0.010)	(0.010)	(0.010)	(0.010)
$T_{97}^{2}-T_{97}^{1}$	-0.582*	-0.595*	-0.591*	-0.591*
	(0.009)	(0.009)	(0.009)	(0.009)
$T_{98}^{2}-T_{98}^{1}$	-0.547*	-0.551*	-0.551*	-0.553*
	(0.008)	(0.008)	(0.008)	(0.008)
$T_{99}^{2}-T_{99}^{1}$	-0.486*	-0.488*	-0.489*	-0.491*
	(0.007)	(0.007)	(0.007)	(0.007)
$T_{00}^{2}-T_{00}^{1}$	-0.399*	-0.402*	-0.402*	-0.402*
	(0.007)	(0.007)	(0.007)	(0.007)
$T_{01}^{2}-T_{01}^{1}$	-0.310*	-0.309*	-0.312*	-0.313*
	(0.006)	(0.006)	(0.006)	(0.006)
$T_{02}^{2}-T_{02}^{1}$	-0.222*	-0.220*	-0.224*	-0.224*
	(0.006)	(0.006)	(0.006)	(0.006)
$T_{03}^{2}-T_{03}^{1}$	-0.135*	-0.137*	-0.139*	-0.139*
	(0.006)	(0.006)	(0.006)	(0.006)
Diff SDM		0.023*		
		(0.003)		
Market Change		-0.000*	-0.000*	-0.000*
		(0.000)	(0.000)	(0.000)
WQImprove WF			0.093*	
			(0.020)	
WQImprove300			0.030*	
			(0.006)	
WQImprove700			0.023*	
			(0.005)	
WQImprove1000			0.001	
			(0.007)	
WQImprove Sm				0.033*
				(0.006)
WQImprove Md				0.024*
				(0.007)
WQ Improve Lg				0.033*
				(0.007)
WQ Improve El				0.001
				(0.007)
Intercept	0.106*	0.117*	0.108*	0.106*
	(0.004)	(0.004)	(0.005)	(0.005)

Adjusted R^2	0.317	0.328	0.327	0.326
N	11,785	11,785	11,785	11785.000

Note: * denotes significance at the 1% level.

Table 29: Repeat Sales with Age-Heteroskedasticity Error Correction

	Basic RS Weighted β/(SE)	Hybrid Weighted β/(SE)	Distance Weighted β/(SE)	Area Weighted β/(SE)
$T_{96}^{2}-T_{96}^{1}$	-0.615*	-0.625*	-0.621*	-0.621*
	(0.011)	(0.011)	(0.011)	(0.011)
$T_{97}^{2}-T_{97}^{1}$	-0.563*	-0.575*	-0.571*	-0.572*
	(0.010)	(0.010)	(0.010)	(0.010)
$T_{98}^{2}-T_{98}^{1}$	-0.531*	-0.534*	-0.535*	-0.537*
	(0.009)	(0.009)	(0.009)	(0.009)
$T_{99}^{2}-T_{99}^{1}$	-0.472*	-0.474*	-0.476*	-0.478*
	(0.008)	(0.008)	(0.008)	(0.008)
$T_{00}^{2}-T_{00}^{1}$	-0.388*	-0.391*	-0.391*	-0.391*
	(0.007)	(0.007)	(0.007)	(0.007)
$T_{01}^{2}-T_{01}^{1}$	-0.302*	-0.300*	-0.304*	-0.304*
	(0.006)	(0.006)	(0.006)	(0.006)
$T_{02}^{2}-T_{02}^{1}$	-0.216*	-0.213*	-0.218*	-0.218*
	(0.006)	(0.006)	(0.006)	(0.006)
$T_{03}^{2}-T_{03}^{1}$	-0.132*	-0.134*	-0.136*	-0.136*
	(0.006)	(0.006)	(0.006)	(0.006)
Diff SDM		0.023*		
		(0.003)		
Market Change		-0.000*	-0.000*	-0.000*
		(0.000)	(0.000)	(0.000)
WQ Improve WF			0.091*	
			(0.020)	
WQ Improve300			0.033*	
			(0.006)	
WQ Improve700			0.022*	
			(0.005)	
WQ Improve1000			0.001	
			(0.007)	
WQ Improve Sm				0.033*

				(0.006)
WQ Improve Md				0.023*
				(0.007)
WQ Improve Lg				0.032*
				(0.007)
WQ Improve El				0.000
				(0.007)
Intercept	0.527*	0.582*	0.536*	0.529*
	(0.020)	(0.020)	(0.023)	(0.023)
Adjusted R^2	0.277	0.288	0.286	0.287
N	11,785	11,785	11,785	11785.000

Note: * denotes significance at the 1% level.

Table 30: Repeat Sales with Heteroskedasticity and Aitken Corrections

	Basic RS Weighted Aitken β/(SE)	Hybrid Weighted Aitken β/(SE)	Distance Weighted Aitken β/(SE)	Area Weighted Aitken β/(SE)
$T_{96}^2 - T_{96}^1$	-0.651*	-0.659*	-0.653*	-0.653*
	(0.010)	(0.010)	(0.010)	-0.01
$T_{97}^2 - T_{97}^1$	-0.594*	-0.603*	-0.599*	-0.600*
	(0.010)	(0.010)	(0.010)	-0.01
$T_{98}^2 - T_{98}^1$	-0.558*	-0.556*	-0.556*	-0.558*
	(0.009)	(0.009)	(0.009)	-0.009
$T_{99}^2 - T_{99}^1$	-0.498*	-0.495*	-0.496*	-0.498*
	(0.008)	(0.008)	(0.008)	-0.008
$T_{00}^2 - T_{00}^1$	-0.408*	-0.406*	-0.405*	-0.406*
	(0.007)	(0.007)	(0.007)	-0.007
$T_{01}^2 - T_{01}^1$	-0.314*	-0.308*	-0.311*	-0.312*
	(0.007)	(0.007)	(0.007)	-0.007
$T_{02}^2 - T_{02}^1$	-0.223*	-0.217*	-0.220*	-0.221*
	(0.006)	(0.006)	(0.006)	-0.006
$T_{03}^2 - T_{03}^1$	-0.136*	-0.137*	-0.138*	-0.139*
	(0.006)	(0.006)	(0.006)	-0.006
Diff SDM		0.025*		
		(0.003)		
Market Change		-0.000*	-0.000*	-0.000*

		(0.000)	(0.000)	(0.000)
WQ Improve WF			0.094*	
			(0.019)	
WQ Improve300			0.043*	
			(0.006)	
WQ Improve700			0.032*	
			(0.005)	
WQ Improve1000			0.010	
			(0.007)	
WQ Improve Sm				0.044*
				-0.006
WQ Improve Md				0.031*
				-0.007
WQ Improve Lg				0.043*
				-0.007
WQ Improve El				0.01
				-0.007
Intercept	0.455*	0.521*	0.456*	0.447*
	(0.019)	(0.020)	(0.022)	-0.022
Adjusted R^2	0.302	0.314	0.314	0.314
N	11,785	11,785	11,785	11,785

Note: * denotes significance at the 1% level.

APPENDIX A: BOX-COX MODEL

The following table presents the results of the flexible Box-Cox model. This model is estimated in Stata (version 9), using the “lambda” option of the “boxcox” command. The variables are split into two groups: transformed and no-transform. The no-transform group is not subjected to the Box-Cox transformation because they do not contain strictly positive elements. Dummy variables and variables that contain negative values are therefore contained in the no-transform group. The first row of Table 31 contains the results for the estimated Box-Cox parameter “r”. The rest of the table contains the coefficients for the independent variables in the Box-Cox model.

Table 31: Box-Cox Model from Chapters 1 and 2

		Coefficient	Std. Err.	P>|z|
	r	-0.145	0.004	0.000
Trans				
	sdm	-0.023		
	Dst*SDM	0.057		
	Lake Area*SDM	-0.142		
	Lake Dist	-0.066		
	bath	0.028		
	Home Age	-0.010		
	Heated Area	0.367		
	Parcel Area	0.041		
	Dist to CBD	-0.111		
	x_coord	-15.401		
	y_coord	7.616		
	Median HH income	0.073		
	/sigma	0.057		
Notrans				
	waterfront	0.035		
	WF*sdmFT	0.001		
	canalfront	0.016		
	golffront	0.052		
	pool	0.003		
	nearair	0.001		
	Percent white	0.099		
	Percent black	0.060		
	Percent over 65	-0.026		
	BASS_LAKE	-0.088		
	BAY_LAKE_B	-0.098		
	BEARHEAD_L~E	-0.066		
	BIG_SAND_L~E	0.051		
	CLEAR_LAKE	-0.084		

DEEP_LAKE	-0.068
KASEY_LAKE	-0.188
KELLY_LAKE	-0.182
KRISTY_LAKE	-0.184
LAKE_ADAIR	-0.060
LAKE_ANDER~N	-0.084
LAKE_ANGEL	-0.203
LAKE_ARNOLD	-0.068
LAKE_BALDWIN	0.020
LAKE_BARTON	-0.033
LAKE_BEARD~L	-0.200
LAKE_BEAUTY	-0.117
LAKE_BELL	-0.088
LAKE_BERRY	-0.029
LAKE_BESSIE	0.055
LAKE_BLANCHE	-0.012
LAKE_BUCHA~N	-0.104
LAKE_BUCK	0.145
LAKE_BUMBY	-0.108
LAKE_BURKETT	-0.031
LAKE_BUTLER	0.021
LAKE_C	-0.120
LAKE_CANE_A	-0.036
LAKE_CATHE~B	-0.051
LAKE_CAY_DEE	-0.081
LAKE_CHARITY	-0.007
LAKE_CHASE	0.052
LAKE_CHERO~E	-0.069
LAKE_CHRIS~E	-0.112
LAKE_COMO	-0.111
LAKE_CONCORD	-0.058
LAKE_CONWAY	0.030
LAKE_COPEL~D	-0.087
LAKE_DANIEL	-0.110
LAKE_DESTINY	-0.015
LAKE_DOT	-0.182
LAKE_DOVER	-0.192
LAKE_DOWN	0.025
LAKE_DOWNEY	-0.048
LAKE_DRUID	-0.055
LAKE_EMERALD	-0.064

LAKE_EOLA	-0.082
LAKE_ESTELLE	-0.043
LAKE_FAIRH~E	-0.197
LAKE_FAIRV~W	-0.052
LAKE_FAITH	-0.042
LAKE_FARRAR	-0.067
LAKE_FORMOSA	-0.055
LAKE_FREDR~A	0.032
LAKE_GATLIN	-0.021
LAKE_GEAR	-0.038
LAKE_GEM_A	-0.100
LAKE_GEM_M~Y	-0.072
LAKE_GEORGE	-0.003
LAKE_GEORGIA	-0.017
LAKE_GILES	-0.075
LAKE_GLORIA	-0.085
LAKE_GREEN~D	-0.100
LAKE_HART	0.136
LAKE_HIAWA~E	-0.044
LAKE_HIGHL~D	-0.054
LAKE_HOLDEN	-0.082
LAKE_HOPE	-0.074
LAKE_HUNGE~D	-0.121
LAKE_IRMA	-0.027
LAKE_ISLEW~H	0.048
LAKE_IVANHOE	-0.024
LAKE_JACKSON	-0.066
LK_JEN_JEWEL	-0.063
LAKE_JESSA~E	-0.032
LAKE_KILLA~Y	-0.046
LAKE_KOZART	-0.162
LAKE_LANCA~R	-0.055
LAKE_LAWSONA	-0.080
LAKE_LOUIS~B	0.065
LAKE_LOVE	-0.044
LAKE_LOVELY	-0.158
LAKE_LURNA	-0.090
LAKE_MABEL	0.055
LAKE_MAITL~D	0.040
LAKE_MANN	-0.091
LAKE_MARSHA	-0.063

LAKE_MINNE~A	-0.015
LAKE_MIZELL	0.021
LAKE_NAN	-0.051
LAKE_NONA	0.158
LAKE_OFWOODS	-0.205
LAKE_OLIVE	-0.140
LAKE_OLYMPIA	-0.120
LAKE_ORLANDO	-0.084
LAKE_OSCEOLA	0.036
LAKE_PAMELA	-0.135
LAKE_PEARL_B	-0.155
LAKE_PICKETT	0.070
LAKE_PINEL~H	-0.048
LAKE_PORTER	-0.038
LAKE_RABAMA	-0.081
LAKE_RICHM~D	-0.133
LAKE_ROBERTS	-0.075
LAKE_ROSE_B	-0.079
LAKE_ROWENA	-0.042
LAKE_SANTI~O	-0.111
LAKE_SARAH	-0.106
LAKE_SHADOW	-0.066
LAKE_SHANNON	-0.046
LAKE_SHEEN	0.067
LAKE_SHERW~D	-0.118
LAKE_SILVER	-0.034
LAKE_STARKE	-0.138
LAKE_SUE	-0.009
LAKE_SUNSET	-0.140
LAKE_SUSAN~H	0.003
LAKE_SYBELIA	-0.030
LAKE_SYLVAN	0.009
LAKE_TENNE~E	-0.085
LAKE_TERRACE	-0.117
LAKE_THERESA	-0.126
LAKE_TIBET	0.068
LAKE_UNDER~L	-0.041
LAKE_VIRGI~A	0.026
LAKE_WADE	-0.141
LAKE_WALKER	-0.192
LAKE_WARREN	-0.006

LAKE_WAUNA~A	-0.035
LAKE_WELDONA	-0.104
LAKE_WESTON	-0.158
LAKE_WHIPP~R	0.100
LAKE_WINYAH	-0.048
LAWNE_LAKE	-0.105
LITTLE_FISH	0.025
LIT_LK_FAIR	-0.115
LITTLE_SAND	-0.023
LONG_LAKE	-0.138
MUD_LAKE_C	-0.058
PALM_LAKE	-0.038
PARK_LAKE_B	-0.080
POCKET_LAKE	0.029
ROCK_LAKE	-0.128
SPRING_LK_B	-0.060
SPRING_LK_C	0.009
_cons	51.042

Note: Stata does not report standard errors for the Box-Cox model with untransformed variables, except on the r variable.

APPENDIX B: MATLAB PROGRAMS

B.1. SWM Programs

Note: the spatial lag and spatial error programs used in this paper are from the spatial toolbox, available at www.spatial-econometrics.com, and written by James P. LeSage and R. Kelley Pace. The following programs are used to create the various spatial weights matrices (SWM) employed in the present dissertation.

B.1.1. Distance/Time SWM

The following programs are used to create an inverse distance-based SWM, which includes distance and time boundaries. Three programs are required to create each SWM. The first program creates an *N* x *N* matrix that contains the distance between each home and its neighbors (as defined by the distance and time boundaries). This code was adapted from a program by Shawn Bucholtz at the USDA, the original program did contain time boundaries. The second program was written for this dissertation, and simply inverts each element of the *N* x *N* matrix. The third program row standardizes the SWM, as explained in Chapter 2. An example of how these three programs would be computed appears below:

```
pdT200_6_3=pdweightT(xc,yc,sale_date,0,200,6,3,0);
%now, invert each element.
IpdT200_6_3II=invdist1(pdT200_6_3);
%To finish SWM, use final file 'normbig'
ISpdT200_6_3II=normbig(IpdT200_6_3II);
```

B.1.1.1. SWM Program 1: pdweightT

```
function g=pdweightT(xcoord,ycoord,date,lower,upper,t1,t2,RowStdOpt);
% PURPOSE: uses x,y coordinates and dates to produce distance and time-based
spatial weight matrices
%          The user is asked to input x and y coordinates, as well as a lower
and upper
%          bound cutoff for the neighborhood.
% ------------------------------------------------------
% USAGE: W = pdweightT(xcoord,ycoord,lower,upper,t1,t2,rowstdopt)
% where:     xcoord = x-direction coordinate
%            ycoord = y-direction coordinate
%            date = the sale date transformed into a conutinuous variable,
%            for example, instead of 199001:199112 denoting year and month,
%            we have date = 1:24
%            lower  = lower bound distance cut-off
%            uppper = upper bound distance cut-off
%            t1 = the length of time back to consider
%            t2 = the length of time forward to consider
%            rowstdopt = 1 for row-standardization
%                      = 0 for no standardization
% ------------------------------------------------------
% RETURNS:
%          W = a sparse weight matrix based on distance and time cut-offs
%              set by lower, upper, t1 and t2 input options
% ------------------------------------------------------

% written by: Patrick Walsh
% PatJwalsh@gmail.com
% University of Central Florida
% Department of Economics
%
% adapting most of the code
% written by: Shawn Bucholtz
% SBUCHOLTZ@ers.usda.gov
% USDA-ERS-ISD-ADB

%This function builds a sparse spatial weight matrix.
%The user is asked to input x and y coordinates, as well as a lower and upper
%bound cutoff for neighborhood, and time limits for neighborhood.
%The user is given the option of indicating if they would like
%to Row-standarize the weight matrix (1=Row Standardized, 0=Not Row-
Standardized)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Declare inital values;
i=1;
p2=[1 1 -8888];
%Begin loop for each observation;

for i=1:(length(xcoord));
      j=1;
```

```
      p1=[1 1 -9999];%Make an inital value;
        %Begin loop to compute distance from observation i to all other
observations;

            for j=1:(length(xcoord));

                if xcoord(i)==xcoord(j) & ycoord(i)==ycoord(j);
                    D=1;%This ensures that two places with the same x,y
location will
                        %be defaulted to 1 distance unit away from eachother.
                        %This can happen in the case of two condos who are
assocaited with one building location.
                else;
                    Xdist=abs(xcoord(i)-xcoord(j));
                    Ydist=abs(ycoord(i)-ycoord(j));
                    D = sqrt(Xdist^2 + Ydist^2);

                end;
                T = date(i)-date(j);
                p = [i j D];

                %A loop to generate a list of neighbors j of current obs i
                    if D > lower & D <= upper & T < t1 & T > -t2;  %Check to
see if j meets the neighborhood cutoff criteria

                        if p1(3) == -9999;  %If this is the first j in the
list of neighbors, then
                                            %make it the inital data set;
                            p1=p;
                        else;
                            p1 = [p1;p];%If this is not the first j in the
list of neighbors, then append this
                                        %to previous list of all neighbors j
for observation i.
                        end;

                    else;
                        p1=p1;%If distance between i and j did not meet
neighborhood cutoff criteria, then
                              %do not append this to previous list of
neighbors for obs. i;
                    end;
                j=j+1; %Step to next neighbor j of observation i

            end;

            if p1(3) == 0;%A Check to see if observation i had any neighbors
within cutoff
                i=i+1;%If it did not, then pass loop to next i
            else;


                    %If observation i had at least 1 neighbor within cutoff,
then append that data to
                    %previous data set for i-1
```

```
                    if RowStdOpt == 1;%Row standardize the weight matrix
                        p1(:,3) = p1(:,3) ./ sum(p1(:,3));
                    end;
                    if p2(3) == -8888;%If the current observation i is the
first observationervation to have any neighbors
                                        %then make it the first data set;
                        p2=p1;
                    else;
                        p2=[p2;p1];%If the current observation i is not the
first observation to have any neighbors
                                    %then append current observation i's
neighbors to all other previous i's neighbors
                        clear p1;
                    end;

                 i=i+1;%pass loop to next i
            end;
  end;

  %Generate the sparse matrix
  g=sparse(p2(:,1),p2(:,2),p2(:,3));
```

B.1.1.2. SWM Program 2: invdist1

```
function S = invdist1(SW);
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Purpose: Transforms a distance weights matrix into an inverse distance
%          weights matrix
% Inputs:
%     WD = An n x n spatial weights matrix based on distance
%     This matrix should NOT be row-standardized.
%     To row standardize, use the file 'normbig'
%
% Returns: Sparse nxn matrix
%
% Author: Patrick Walsh
%         Department of Economics
%         University of Central Florida
%         PatJwalsh@gmail.com
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


[n j] = size(SW);
for i=1:n
    for j=1:n
       if i==j;
            SW(i,j)=0; %make sure the i,i th element =0
        end;
    end;
end;
for i=1:n;
    for j=1:n;
        if (SW(i,j)>0);
            SW(i,j)=1/SW(i,j); %Turn the distances to inverse distances
        else
            SW(i,j)=0;
        end;
    end;
end;
S = sparse(SW);
```

B.1.1.3. SWM Program 3: normbig

```
function SN=normbig(SW);
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Purpose: Transforms an unstandardized weights matrix into a standardized
%          weights matrix, where all rows sum to one. Is useful with large
%          matrices.
% Inputs:
%     W = An n x n spatial weights matrix
%     This matrix should NOT be row-standardized.
%
%
% Returns: Sparse nxn matrix
% Author: Patrick Walsh
%         Department of Economics
%         University of Central Florida
%         PatJwalsh@gmail.com
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[n j] = size(SW);
L=sum(SW');
T=L';
Q=logical(T); %make sure we are not dividing by zero if no neighbors
ones1=ones(n,1);
Q2=Q-ones1;
Q3=Q2*-1;
T2=T+Q3;
for i=1:n;
    SW(:,i)=SW(:,i)./T2;
end;
SN=sparse(SW);
```

B.1.2. Nearest Neighbor SWM

The nearest neighbor SWM is created using the program “make_nnw”, from the spatial toolbox.

No additions to this program are made. An example of how it is used appears below:

```
W15 = make_nnw(xc,yc,15,4);
```

B.2. Implicit Price Program

The following program is the Matlab program used to compute the implicit prices of a spatial lag model. This program is flexible: it is designed to be used with any of the water quality indicators.

```
function
b=sarimp3(sar,meanwfp,meannwfp,meanWQ,meanwfdist,meannwfdist,lnmeanwfarea,lnm
eannwfarea);
% PURPOSE: Uses the results from a spatial autoregressive model, from a
%          specific application with water quality, and calculates implicit
%          prices
%          Note that the 3 in the title indicates that this file is to be
%          used with specification 3, where all 3 interaction terms are
%          used.
%          The ordering of the variables in the regression is critical.
%          It must follow the previous order, or will not work.
% ------------------------------------------------------
% USAGE: W =sarimp3(sar,meanwfp,meannwfp,meanWQ,meanwfdist,        ...
%                  meannwfdist,lnmeanwfarea,lnmeannwfarea);
% where: sar = the name of the sar results
%        meanwfp = The mean waterfront price
%        meannwfp = The mean nonwaterfront price
%        meanWQ = The mean of the WQ variable
%        meanwfdist = The mean waterfront distance to lake
%        meannwfdist = The mean nonwaterfront distance to lake
%        lnmeanwfarea = The mean nonwaterfront lake area
%        lnmeannwfarea = The mean waterfront lake area
% ------------------------------------------------------
% RETURNS: A set of implicit prices at both mean values and at several
%          distances from the lake
%
%
% ------------------------------------------------------
% written by: Patrick Walsh
% PatJwalsh@gmail.com
% University of Central Florida
% Department of Economics


lnmeanwfdist=log(meanwfdist);
lnmeannwfdist=log(meannwfdist);
lnmeanWQ=log(meanWQ);

%implicit price of WQ
MP3sarwqwf=1/(1-
sar.rho)*(meanwfp/meanWQ)*(sar.beta(2,1)+sar.beta(3,1)+sar.beta(4,1)*lnmeanwf
dist+sar.beta(5,1)*lnmeanwfarea);
MP3sarwqnwf=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*lnmeannwfdist+sar.bet
a(5,1)*lnmeannwfarea);

% Implicit price of waterfront
MP3sarwqLF=1/(1-sar.rho)*meanwfp*(sar.beta(6,1)+sar.beta(3,1)*lnmeanWQ);
```

```
% Implicit price of distance
MP3sarwqdist=1/(1-
sar.rho)*(meannwfp/meannwfdist)*(sar.beta(7,1)+sar.beta(4,1)*lnmeanWQ);

% Implicit prices as distance changes
MP3sarWQnwf100=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*log(100)+sar.beta(5,1
)*lnmeannwfarea);
MP3sarWQnwf300=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*log(300)+sar.beta(5,1
)*lnmeannwfarea);
MP3sarWQnwf500=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*log(500)+sar.beta(5,1
)*lnmeannwfarea);
MP3sarWQnwf700=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*log(700)+sar.beta(5,1
)*lnmeannwfarea);
MP3sarWQnwf900=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*log(900)+sar.beta(5,1
)*lnmeannwfarea);

% Implicit Prices as Area Changes
% 100 acres is  404685.643 sq m, and the ln of this is 12.91086585
% 1000 acres is 4046856.43 sq m, and ln of this is 15.21345095
ln100ac=12.91086585;
ln1000ac=15.21345095;
% Waterfront
MP3sarwqwfA100=1/(1-
sar.rho)*(meanwfp/meanWQ)*(sar.beta(2,1)+sar.beta(3,1)+sar.beta(4,1)*lnmeanwf
dist+sar.beta(5,1)*ln100ac);
MP3sarwqwfA1000=1/(1-
sar.rho)*(meanwfp/meanWQ)*(sar.beta(2,1)+sar.beta(3,1)+sar.beta(4,1)*lnmeanwf
dist+sar.beta(5,1)*ln1000ac);
% Non-Waterfront
MP3sarwqnwfA100=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*lnmeannwfdist+sar.bet
a(5,1)*ln100ac);
MP3sarwqnwfA1000=1/(1-
sar.rho)*(meannwfp/meanWQ)*(sar.beta(2,1)+sar.beta(4,1)*lnmeannwfdist+sar.bet
a(5,1)*ln1000ac);

b=[MP3sarwqwf,MP3sarwqnwf,MP3sarwqLF,MP3sarwqdist,MP3sarWQnwf100,MP3sarWQnwf3
00,MP3sarWQnwf500,MP3sarWQnwf700,MP3sarWQnwf900,MP3sarwqwfA100,MP3sarwqwfA100
0,MP3sarwqnwfA100,MP3sarwqnwfA1000 ];
```

B.3. Standard Error Program

The following is an example of one of the programs used to calculate the standard errors of the implicit prices using the delta method.

```
function
b=sarSE3(p,dhessn,parm,meanwfp,meannwfp,meanWQ,meanwfdist,meannwfdist,lnmeanw
farea,lnmeannwfarea);
% PURPOSE: Calculates standard errors for the implicit prices from a water
%          quality hedonic regression. Must be used with a spatial lag
%          regression.
%          THE FOLLOWING CODE MUST BE RUN BEFORE THE PROGRAM CAN BE USED:
%          dbstop at 191 in sar;
%          sarTSI'SWM'=sar(y,x3TSI,'SWM');
%          Note that the 3 in the title indicates that this file is to be
%          used with specification 3, where all 3 interaction terms are
%          used. Also, it is critical that the x vector is the exact one
%          used in the WQ_1996IIb.m file, a version of the "xF3b". If the x
variables are in a
%          different configuration, the results will be crap.
% -------------------------------------------------------
% USAGE: W =sarSE3(SWM,y,xF3b,meanwfp,meannwfp,meanWQ,meanwfdist,        ...
%                  meannwfdist,lnmeanwfarea,lnmeannwfarea);
% where: SWM = The spatial weights matrix to be used in the sar
%        y = the indep variable, ln_price
%        xF3b = the x vector, discussed above
%        meanwfp = The mean waterfront price
%        meannwfp = The mean nonwaterfront price
%        meanWQ = The mean of the WQ variable
%        meanwfdist = The mean waterfront distance to lake
%        meannwfdist = The mean nonwaterfront distance to lake
%        lnmeanwfarea = The mean nonwaterfront lake area
%        lnmeannwfarea = The mean waterfront lake area
% -------------------------------------------------------
% RETURNS: A set of standard errors for the implicit prices in a water
%        quality regression, written for my dissertation
% -------------------------------------------------------
% written by: Patrick Walsh
% PatJwalsh@gmail.com
% University of Central Florida
% Department of Economics


% dbstop at 191 in sar;
% sar3=sar(y,xF3b,SWM);

rho=p;
xpxi1=invpd(-dhessn);

lnmeanwfdist=log(meanwfdist);
lnmeannwfdist=log(meannwfdist);
lnmeanWQ=log(meanWQ);
%Waterfront WQ
```

```
WFmult=(1/(1-rho))*(meanwfp/meanWQ);
aWFWQ1=zeros(1,177);
aWFWQ1(1,2)=1;
aWFWQ1(1,3)=1;
aWFWQ1(1,4)=lnmeanwfdist;
aWFWQ1(1,5)=lnmeanwfarea;
aWFWQ=aWFWQ1*WFmult;
seWFWQ=sqrt(aWFWQ*xpxi1*aWFWQ');
%check:
WFWQ=aWFWQ*parm;
% Now for non-WF WQ
NWFmult=(1/(1-rho))*(meannwfp/meanWQ);
aNWFWQ1 = zeros(1,177);
aNWFWQ1(1,2)=1;
aNWFWQ1(1,4)=lnmeannwfdist;
aNWFWQ1(1,5)=lnmeannwfarea;
aNWFWQ=aNWFWQ1*NWFmult;
seNWFWQ = sqrt(aNWFWQ*xpxi1*aNWFWQ');
NWFWQ = aNWFWQ*parm;
%now, WQ NWF as distance changes
%NWF at 100 feet:
aNWFWQ1001 = aNWFWQ1;
aNWFWQ1001(1,4)=log(100);
aNWFWQ100=aNWFWQ1001*NWFmult;
seNWFWQ100= sqrt(aNWFWQ100*xpxi1*aNWFWQ100');
NWFWQ100 = aNWFWQ100*parm;
NWFWQ100cilow = NWFWQ100-seNWFWQ100*1.96;
NWFWQ100cihi = NWFWQ100+seNWFWQ100*1.96;
% now 300 feet
aNWFWQ3001 = aNWFWQ1;
aNWFWQ3001(1,4)=log(300);
aNWFWQ300=aNWFWQ3001*NWFmult;
seNWFWQ300= sqrt(aNWFWQ300*xpxi1*aNWFWQ300');
NWFWQ300 = aNWFWQ300*parm;
NWFWQ300cilow = NWFWQ300-seNWFWQ300*1.96;
NWFWQ300cihi = NWFWQ300+seNWFWQ300*1.96;
% 500 feet
aNWFWQ5001 = aNWFWQ1;
aNWFWQ5001(1,4)=log(500);
aNWFWQ500=aNWFWQ5001*NWFmult;
seNWFWQ500= sqrt(aNWFWQ500*xpxi1*aNWFWQ500');
NWFWQ500 = aNWFWQ500*parm;
NWFWQ500cilow = NWFWQ500-seNWFWQ500*1.96;
NWFWQ500cihi = NWFWQ500+seNWFWQ500*1.96;
%700 feet
aNWFWQ7001 = aNWFWQ1;
aNWFWQ7001(1,4)=log(700);
aNWFWQ700=aNWFWQ7001*NWFmult;
seNWFWQ700= sqrt(aNWFWQ700*xpxi1*aNWFWQ700');
NWFWQ700 = aNWFWQ700*parm;
NWFWQ700cilow = NWFWQ700-seNWFWQ700*1.96;
NWFWQ700cihi = NWFWQ700+seNWFWQ700*1.96;
%900 feet
aNWFWQ9001 = aNWFWQ1;
aNWFWQ9001(1,4)=log(900);
aNWFWQ900=aNWFWQ9001*NWFmult;
seNWFWQ900= sqrt(aNWFWQ900*xpxi1*aNWFWQ900');
```

```
NWFWQ900 = aNWFWQ900*parm;
NWFWQ900cilow = NWFWQ900-seNWFWQ900*1.96;
NWFWQ900cihi = NWFWQ900+seNWFWQ900*1.96;
% Now, WQ as Area changes
% 100 acres is  404685.643 sq m, and the ln of this is 12.91086585
% 1000 acres is 4046856.43 sq m, and ln of this is 15.21345095
ln100ac=12.91086585;
ln1000ac=15.21345095;
% Waterfront 100
aWFWQAr1001=aWFWQ1;
aWFWQAr1001(1,5)=ln100ac;
aWFWQAr100=aWFWQAr1001*WFmult;
WFWQAr100=aWFWQAr100*parm;
seWFWQAr100=sqrt(aWFWQAr100*xpxi1*aWFWQAr100');
% Waterfront 1000
aWFWQAr10001=aWFWQ1;
aWFWQAr10001(1,5)=ln1000ac;
aWFWQAr1000=aWFWQAr10001*WFmult;
WFWQAr1000=aWFWQAr1000*parm;
seWFWQAr1000=sqrt(aWFWQAr1000*xpxi1*aWFWQAr1000');
% Non-WF 100
aNWFWQAr1001 = aNWFWQ1;
aNWFWQAr1001(1,5)=ln100ac;
aNWFWQAr100=aNWFWQAr1001*NWFmult;
NWFWQAr100=aNWFWQAr100*parm;
seNWFWQAr100=sqrt(aNWFWQAr100*xpxi1*aNWFWQAr100');
% Non-WF 1000
aNWFWQAr10001 = aNWFWQ1;
aNWFWQAr10001(1,5)=ln1000ac;
aNWFWQAr1000=aNWFWQAr10001*NWFmult;
NWFWQAr1000=aNWFWQAr1000*parm;
seNWFWQAr1000=sqrt(aNWFWQAr1000*xpxi1*aNWFWQAr1000');
% Now, distance:
aDst1=zeros(1,177);
aDst1(1,4)=lnmeanWQ;
aDst1(1,7)=1;
aDst=aDst1*(1/(1-rho))*(meannwfp/meannwfdist);
seDst = sqrt(aDst*xpxi1*aDst');
Dst = aDst*parm;
%finally, waterfront premium
aW1 = zeros(1,177);
aW1(1,3)=lnmeanWQ;
aW1(1,6)=1;
aW=aW1*(1/(1-rho))*meanwfp;
seW = sqrt(aW*xpxi1*aW');
WFprem = aW*parm;

b=[seWFWQ, WFWQ, seNWFWQ, NWFWQ;
    seNWFWQ100, NWFWQ100, NWFWQ100cilow, NWFWQ100cihi;
    seNWFWQ300, NWFWQ300, NWFWQ300cilow, NWFWQ300cihi;
    seNWFWQ500, NWFWQ500, NWFWQ500cilow, NWFWQ500cihi;
    seNWFWQ700, NWFWQ700, NWFWQ700cilow, NWFWQ700cihi;
    seNWFWQ900, NWFWQ900, NWFWQ900cilow, NWFWQ900cihi;
    seDst, Dst, seW, WFprem;
    seWFWQAr100, WFWQAr100, seWFWQAr1000, WFWQAr1000;
    seNWFWQAr100, NWFWQAr100, seNWFWQAr1000, NWFWQAr1000];
```

LIST OF REFERENCES

Adamowicz, Wiktor, Joffre Swait, Peter Boxall, Jordan Louviere, and Michael Williams. 1997. Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. *Journal of Environmental Economics and Management* 32:65-84.

Aitken, A.C. 1935. On Least Squares and Linear Combinations of Observations. *Proceedings of the Royal Society of Edinburgh* 55:42-48.

Anderson, S.T. , and S.E. West. 2006. Open Space, Residential Property Values, and Spatial Context. *Regional Science and Urban Economics* 36:773-789.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Norwell, MA: Kluwer Academic Publishers.

———. 1999. Spatial Econometrics. In *Working Paper*. Dallas, Tx: University of Texas at Dallas.

Anselin, Luc, and Julie Le Gallo. 2006. Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects. *Spatial Economic Analysis* 1 (1):31-51.

Anselin, Luc, and Nancy Lozano-Gracia. 2008. Errors in Variables and Spatial Effects in Hedonic House Price Models of Ambient Air Quality. *Empirical Economics* 34 (5):5-34.

Arbia, Giuseppe. 2006. *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Edited by M. M. Fischer, G. J. D. Hewings, P. Nijkamp and F. Snickars, *Advances in Spatial Science*. Berlin: Springer-Verlag.

Atasoy, Mary, Raymond B. Palmquist, and Daniel J. Phaneuf. 2006. Estimating the Effects of Urban Residential Development on Water Quality Using Microdata. *Journal of Environmental Management* 79:399-408.

Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse. 1963. A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association* 58 (304):933-942.

Baltagi, Badi H., and Dong Li. 2001. LM Tests for Functional Form and Spatial Error Correlation. *International Regional Science Review* 24 (2):194-225.

———. 2004. Testing for Linear and Log-Linear Models Against Box-Cox Alternatives with Spatial Lag Dependence. In *Advances in Econometrics, Volume 18*, edited by J. P. LeSage. Bingley, UK: Emerald Group Publishing Limited.

Blomquist, Glenn. 1974. The Effect of Electric Utility Power Plant Location on Area Property Value. *Land Economics* 50 (1):97-100.

Bourassa, Steven C., Eva Cantoni, and Martin Hoesli. 2007. Spatial Dependence, Housing Submarkets, and House Price Prediction. *Journal of Real Estate Finance and Economics* 35:143-160.

Boxall, Peter C., Wing H. Chan, and Melville L. McMillan. 2005. The Impact of Oil and Natural Gas Facilities on Rural Residential Property Values: a Spatial Hedonic Analysis. *Resource and Energy Economics* 27:248-269.

Boyle, Kevin J., P. Joan Poor, and Laura O. Taylor. 1999. Estimating the Demand for Protecting Freshwater Lakes from Eutrophication. *American Journal of Agricultural Economics* 81 (5):1118-1122.

Boyle, Kevin J., and Laura O. Taylor. 2001. Does the Measurement of Property and Structural Characteristics Affect Estimated Implicit Prices for Environmental Amenities in a Hedonic Model. *Journal of Real Estate Finance and Economics* 22 (2/3):303-318.

Boyle, Melissa A., and Katherine A. Kiel. 2001. A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities. *Journal of Real Estate Literature* 9 (2):117-144.

Brashares, Edith. 1985. Estimating the instream Value of Lake Water Quality in Southeast Michigan, University of Michigan.

Brasington, David M. 2006. School Choice and the Flight to Private Schools: To What Extent are Public and Private Schools Substitutes? In *LSU Department of Economics Working Paper Series*. Baton Rouge: Louisiana State University.

Brasington, David M., and Diane Hite. 2005. Demand for Environmental Quality: a Spatial Hedonic Analysis. *Regional Science and Urban Economics* 35:57-82.

Brookshire, D.S., Mark A. Thayer, and William D. Schulze. 1982. Valuing Public Goods: A Comparison of Survey and Hedonic Approaches. *American Economic Review* 72:165-177.

Brown, Claude D., Mark V. Hoyer, Roger W. Bachmann, and Daniel E. Jr.   Canfield. 2000. Nutrient-Chlorophyll Relationships: an Evaluation of Empirical Nutrient-Chlorophyll Models Using Florida and North-Temperate Lake Data. *Canadian Journal of Fisheries and Aquatic Science* 57:1574-1583.

Brown, Gardner M., and Henry O. Polakowski. 1977. Economic Valuation of Shoreline. *The Review of Economics and Statistics* 59 (3):272-278.

Cameron, Trudy A. 2006. Directional Heterogeneity in Distance Profiles in Hedonic Property Value Models. *Journal of Environmental Economics and Management* 51:26-45.

Can, Ayse. 1990. The Measurement of Neighborhood Dynamics in Urban House Prices. *Economic Geography* 66 (3):254-272.

———. 1992. Specification and Estimation of Hedonic Housing Price Models. *Regional Science and Urban Economics* 22:453-474.

Can, Ayse, and Isaac Megbolugbe. 1997. Spatial Dependence and House Price Index Construction. *Journal of Real Estate Finance and Economics* 14:203-222.

Carlson, Robert E. 1977. A Trophic State Index for Lakes. *Limnology and Oceanography* 22 (2):361-369.

Case, Bradford, John Clapp, Robin Dubin, and Mauricio Rodriguez. 2004. Modeling Spatial and Temporal House Price Patterns: A Comparison of Four Models. *Journal of Real Estate Finance and Economics* 29 (2):167-191.

Case, Bradford, Peter F. Colwell, Chris Leishman, and Craig Watkins. 2006. The Impact of Environmental Contamination on Condo Prices: A Hybrid Repeat-Sale/Hedonic Approach. *Real Estate Economics* 34 (1):77-107.

Case, Bradford, and John M. Quigley. 1991. The Dynamics of Real Estate Prices. *The Review of Economics and Statistics* 73 (1):50-58.

Case, Karl E. , and Robert J. Shiller. 1989. The Efficiency of the Market for Single-Family Homes. *The American Economic Review* 79 (1):125-137.

Case, Karl E., and Robert J. Shiller. 1987. Prices of Single Family Homes Since 1970: New Indexes for Four Cities. In *Cowles Foundation Discussion Papers*. New Haven, Connecticut: Yale University.

Chay, Kenneth Y., and Michael Greenstone. 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy* 115 (2):376-424.

Cho, Seong-Hoon, Christopher D. Clark, William M. Park, and Seung Gyu Kim. 2009. Spatial and Temporal Variation in the Housing Market Values of Lot Size and Open Space. *Land Economics* 85 (1):51-73.
Cliff, A., and J. Ord. 1973. *Spatial Autocorrelation*. London: Pion.
Cropper, Maureen L. 2000. Has Economic Research Answered the Needs of Environmental Policy? *Journal of Environmental Economics and Management* 39:328-350.
Davidson, R., and James G. MacKinnon. 1985. Testing Linear and Loglinear Regressions against Box-Cox Alternatives. *The Canadian Journal of Economics* 18 (3):499-517.
Davis, Lucas W. 2004. The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster. *American Economic Review* 94:1693-1704.
Deaton, B. James, and John P. Hoehn. 2004. Hedonic Analysis of Hazardous Waste Sites in the Presence of Other Urban Disamenities. *Environmental Science & Policy* 7:499-508.
DeBusk, William. 2002. Surface Water Quality Assessment in Florida: The 305(b) Report and 303(d) List. In *Working Paper Series*: Soil and Water Sciences Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, http://edis.ifas.ufl.edu.
Delucchi, Mark A., James J. Murphy, and Donald R. McCubbin. 2002. The Health and Visibility Cost of Air Pollution: A Comparison of Estimation Methods. *Journal of Environmental Management* 64:139-152.
Dubin, Robin A. 1992. Spatial Autocorrelation and Neighborhood Quality. *Regional Science and Urban Economics* 22:433-452.
Dubin, Robin A. . 1988. Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms. *The Review of Economics and Statistics* 70 (3):466-474.
———. 1998. Spatial Autocorrelation: A Primer. *Journal of Housing Economics* 7:304-327.
Dubin, Robin A., R. Kelley Pace, and Thomas G. Thibodeau. 1999. Spatial Autoregression Techniques for Real Estate Data. *Journal of Real Estate Literature* 7:79-95.
Egan, Kevin J., Joseph A. Herriges, Catherine L. Kling, and John A. Downing. 2009. Valuing Water Quality as a Function of Water Quality Measures. *American Journal of Agricultural Economics* 91 (1):106-123.
EPA, US. 2000. Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs. Washington, DC: US EPA Office of Water.
FDEP. 1996. 1996 Water-Quality Assessment for the State of Florida: Section 305 b Main Report, edited by FDEP. Tallahassee: Bureau of Water Resources Protection.
———. 2006. Integrated Water Quality Assessment for Florida: 2006 305 (b) Report and 303 (d) List Update, edited by Florida Department of Environmental Protection.
———. 2008. Integrated Water Quality Assessment for Florida: 2008 305 b Report and 303 d List Update, edited by Florida Department of Environmental Protection. Tallahassee.
———. 2009. State of Florida Numeric Nutrient Criteria Development Plan, edited by Bureau of Assessment and Restoration Support; Division of Environmental Assessment and Restoration; FDEP. Tallahassee, Fl.
Feenberg, Daniel, and Edwin Mills. 1980. *Measuring the Benefits of Water Pollution Abatement*. New York: Academic Press.
Florax, Raymond J., and Arno J. Van der Vlist. 2003. Spatial Econometric Data Analysis: Moving Beyond Traditional Models. *International Regional Science Review* 26 (3):223-243.

Gatzlaff, Dean H., and Donald R. Haurin. 1997. Sample Selection Bias and Repeat-Sales Index Estimates. *Journal of Real Estate Finance and Economics* 14:33-50.

Gawande, Kishore, and Hank Jenkins-Smith. 2001. Nuclear Waste Transport and Residential Property Values: Estimating the Effects of Perceived Risks. *Journal of Environmental Economics and Management* 42:207-233.

Gelfand, Alan E., Mark D. Ecker, John R. Knight, and C.F. Sirmans. 2004. The Dynamics of Location in Home Price. *Journal of Real Estate Finance and Economics* 29 (2):149-166.

Gibbs, Julie P., John M. Halstead, and Kevin J. Boyle. 2002. An Hedonic Analysis of the Effects of Lake Water Clarity on New Hampshire Lakefront Properties. *Agricultural and Resource Economics Review* 31 (1):39-46.

Graves, Phil, James C. Murdoch, Mark A. Thayer, and Don Waldman. 1988. The Robustness of Hedonic Price Estimation: Urban Air Quality. *Land Economics* 64 (3):220-233.

Greenstone, Michael, and Justin Gallagher. 2008. Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. In *MIT Department of Economics Working Paper no.05-27, Available at SSRN: http://ssrn.com/abstract=840207*: Massachusets Institute of Technology.

Haab, Timothy C., and Kenneth E. McConnell. 2002. *Valuing Environmental and Natural Resources--The Econometrics of Non-Market Valuation*. Northampton, MA: Edward Elgar Publishers.

Hallstrom, Daniel G., and V. Kerry Smith. 2005. Market Responses to Hurricanes. *Journal of Environmental Economics and Management* 50 (3):541-561.

Harrison, D., Jr., and D.L. Rubinfeld. 1978. Hedonic Housing Prices and the Demand for Clean Air. *Journal of Environmental Economics and Management* 5:81-102.

Hoyer, Mark V., Claude D. Brown, and Daniel E. Canfield Jr. 2004. Relations Between Water Chemistry and Water Quality as Defined by Lake Users in Florida. *Lake and Reservoir Management* 20 (3):240-248.

Hoyer, Mark V., Thomas K. Frazer, Sky K. Notestein, and Daniel E. Canfield Jr. 2002. Nutrient, Chlorophyll, and Water Clarity Relationships in Florida's Nearshore Coastal Waters with Comparisons to Freshwater Lakes. *Canadian Journal of Fisheries and Aquatic Science* 59:1024-1031.

Iovanna, Richard, and Charles Griffiths. 2006. Clean Water, Ecological Benefits, and Benefits Transfer: a Work in Progress at the U.S. EPA. *Ecological Economics* 60 (3):473-482.

Jackson, Thomas O. 2001. The Effects of Environmental Contamination of Real Estate: A Literature Review. *Journal of Real Estate Literature* 9 (2):93-116.

Jeon, Yongsik, Joseph A. Herriges, Catherine L. Kling, and John A. Downing. 2005. The Role of Water Quality Perceptions in Modeling Lake Recreation Demand. In *Iowa State University Department of Economics Working Paper Series*. Ames, Iowa: Iowa State University.

Karou, Yoshiaki, V. Kerry Smith, and Jin Long Liu. 1995. Using Random Utility Models to Estimate the Recreational Value of Estuarine Resources. *American Journal of Agricultural Economics* 77 (1):141-151.

Ketkar, Kusum. 1992. Hazardous Waste Sites and Property Values in the State of New Jersey. *Applied Economics* 24 (6):647-659.

Kiel, Katherine A., and Micheal Williams. 2007. The Impact of Superfund Sites on Local Property Values: Are all Sites the Same? *Journal of Urban Economics* 61:170-192.

Kim, Chong Won, Tim T. Phipps, and Luc Anselin. 2003. Measuring the Benefits of air Quality Improvement: a Spatial Hedonic Approach. *Journal of Environmental Economics and Management* 45:24-39.
Kolhase, Janet E. 1991. The Impact of Toxic Waste Sites on Housing Values. *Journal of Urban Economics* 30 (1):1-26.
Krysel, Charles, Elizabeth Marsh Boyer, Charles Parson, and Patrick Welle. 2003. Lakeshore Property Values and Water Quality: Evidence from Property Sales in the Mississippi Headwaters Region, edited by L. C. o. M. Resources. Bemidji: Mississippi Headwaters Board and Bemidji State University.
Krysel, Charles, Elizabeth Marsh Boyer, Charles Parson, and Patrick Welle. 2003. Lakeshore Property Values and Water Quality: Evidence from Property Sales in the Mississippi Headwaters Region. Bemidji: Mississippi Headwaters Board and Bemidji State University.
Lacombe, Donald J. 2004. Does Econometric Methodology Matter? An Analysis of Public Policy Using Spatial Econometric Techniques. *Geographic Analysis* 36 (2):105-118.
Le, Canh Quang. 2009. Three Essays in Spatial Economietrics and Labor Economics, Department of Economics, Kansas State University, Manhattan, Kansas.
Le, Canh Quang, and Dong Li. 2008. Double-Length Regression Tests for Testing Functional Forms and Spatial Error Dependence. *Economics Letters* 101:253-257.
Leggett, C. G., and N. E. Bockstael. 2000. Evidence of the Effects of Water Quality on Residential Land Prices. *Journal of Environmental Economics and Management* 39:121-144.
LeSage, James P. 1999. The Theory and Practice of Spatial Econometrics. In *Documentation for Spatial Econometrics Toolbox for MATLAB*. www.spatial-econometrics.com: University of Toledo.
McMillen, Daniel P. 2003. Spatial Autocorrelation or Model Misspecification? *International Regional Science Review* 26 (2):208-217.
Meese, Richard A., and Nancy E. Wallace. 1997. The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic Regression, and Hybrid Approaches. *Journal of Real Estate Finance and Economics* 14:51-73.
Messer, Kent D., William D. Schulze, Katherine F. Hackett, Trudy A. Cameron, and Gary H. McClelland. 2006. Can Stigma Explain Large Property Value Losses? The Psychology and Economics of Superfund. *Environmental and Resource Economics* 33:299-324.
Michael, Holly J., Kevin J. Boyle, and Roy Bouchard. 1996. Water Quality Affects Property Prices: a case Study of Selected Maine Lakes. In *Miscellaneous Report 398*: Maine Agricultural and Forest Experiment Station, University of Maine.
———. 2000. Does the Measurement of Environmental Quality Affect Implicit Prices Estimated from Hedonic Models? *Land Economics* 76 (2):283-298.
Morgan, Cynthia, and Nicole Owens. 2001. Benefits of Water Quality Policies: The Chesapeake Bay. *Ecological Economics* 39 (2):271-284.
Mueller, Julie M. , and John B. Loomis. 2008. Spatial Dependence in Hedonic Property Models: Do Different Corrections for Spatial Dependence Result in Economically Significant Differences in Estimated Implicit Prices? *Journal of Agricultural and Resource Economics* 33 (2):212-231.
Nelson, John P. 1981. Three Mile Island and Residential Property Values: Empirical Analysis and Policy Implications. *Land Economics* 57 (3):363-372.

Ord, J. , and A. Cliff. 1973. *Spatial Autocorrelation*. London: Pion.
Pace, R. Kelley, Ronald Barry, John M. Clapp, and Mauricio Rodriguez. 1998. Spatiotemporal Autoregressive Models of Neighborhood Effects. *Journal of Real Estate Finance and Economics* 17 (1):15-33.
Pace, R. Kelley, Ronald Barry, Otis W. Gilley, and C.F. Sirmans. 2000. A Method for Spatial-Temporal Forecasting with an Application to Real Estate Prices. *International Journal of Forecasting* 16:229-246.
Pace, R. Kelley, and Otis W. Gilley. 1997. Using the Spatial Configuration of the Data to Improve Estimation. *Journal of Real Estate Finance and Economics* 14 (3):333-340.
———. 1998. Generalizing the OLS and Grid Estimators. *Real Estate Economics* 26 (2):331-347.
Pace, R. Kelley, C.F. Sirmans, and V. Carlos Jr. Slawson. 2002. Are Appraisers Statisticians? In *Real Estate Valuation theory*, edited by K. Wang and M. L. Wolverton. Norwell, MA: Kluwer Academic Publishers.
Palmquist, Raymond B. 1982. Measuring Environmental Effects on Property Values without Hedonic Regressions. *Journal of Urban Economics* 11:333-347.
Palmquist, Raymond B., and Charles M. Fulcher. 2006. The Economic Valuation of Shoreline: 30 Years Later. In *Explorations in Environmental and Natural Resource Economics: Essays in Honor of Gardner M. Brown, Jr*, edited by R. Halvorsen and D. F. Layton. Northampton, Ma: Edward Elgar Publishing, Inc.
Palmquist, Raymond B., and V. Kerry Smith. 2001. The Use of Hedonic Property Value Techniques for Policy and Litigation. In *International Yearbook of Environmental and Resource Economics*, edited by T. Tietenberg and H. Folmer. Northampton, MA: Edward Elgar Publishing, Inc.
Parmeter, Christopher F., and Jaren C. Pope. 2009. Quasi-Experiments and Hedonic Property Value Methods. In *The Handbook on Experimental Economics and the Environment* edited by J. A. List and M. K. Price. Forthcoming 2009: Edward Elgar Publishers.
Patton, Myles, and Seamus McErlean. 2003. Spatial Effects within the Agricultural Land Market in Northern Ireland. *Journal of Agricultural Economics* 54 (1):35-54.
Pendleton, Linwood, Nicole Martin, and D. G. Webster. 2001. Public Perceptions of Environmental Quality: A Survey Study of Beach Use and Perceptions in Los Angeles County. *Marine Pollution Bulletin* 42 (11):1155-1160.
Poor, P. Joan, Kevin J. Boyle, Laura O. Taylor, and Roy Bouchard. 2001. Objective versus Subjective Measures of Water Clarity in Hedonic Property Value Models. *Land Economics* 77 (4):482-493.
Poor, P. Joan, Keri L. Pessagno, and Robert W. Paul. 2007. Exploring the hedonic value of ambient water quality: A local watershed-based study. *Ecological Economics* 60 (4):797-806.
Pope, Jaren C. 2008. Buyer Information and the Hedonic: The Impact of a Seller Disclosure on the Implicit Price for Airport Noise. *Journal of Urban Economics* 63:498-516.
Ridker, R. G., and J.A. Henning. 1967. The Determinants of Residential Property Values with Special Reference to Air Pollution. *Review of Economics and Statistics* 49 (2):246-257.
Rosen, Sherwin. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy* 82 (1):34-55.

Ryding, Sven-Olof, and Walter Rast. 1989. *The Control of Eutrophication of Lakes and Reservoirs*. Edited by J. N. R. Jeffers. Vol. 1, *Man and the Biosphere Series*. Paris, France: UNESCO and The Parthenon Publishing Group.

Shiller, Robert J. 1991. Arithmetic Repeat Sales Price Estimators. *Journal of Housing Economics* (1):110-126.

Shiller, Robert J. . 1993. Measuring Asset Values for Cash Settlement in Derivative Markets: Hedonic Repeated Measures Indices and Perpetual Futures. *The Journal of Finance* 48 (3):911-931.

Smith, V. Kerry. 1993. Welfare Effects, Omitted Variables, and the Extent of Market. *Land Economics* 69:121-131.

Smith, V. Kerry, and Ju-Chin Huang. 1995. Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. *The Journal of Political Economy* 103 (1):209-227.

Smith, V. Kerry, and R.T. Kopp. 1980. The Spatial Limits of the Travel Cost Recreation Demand Model. *Land Economics* 59:64-72.

Smith, V. Kerry, and Raymond B. Palmquist. 2002. The Use of Hedonic Property Value Techniques for Policy and Litigation. In *International Yearbook of Environmental and Resource Economics*, edited by T. Tietenberg and H. Folmer. Northampton, UK: Edward Elgar Publishing.

Steinnes, Donald N. 1992. Measuring the Economic Value of Water Quality: The Case of Lakeshore Land. *The Annals of Regional Science* 26 (2):171-76.

US EPA. 2003. National Management Measures for the Control of Nonpoint Pollution from Agriculture, edited by Office of Water. Washington DC: US Environmental Protection Agency,.

Vajjhala, Shalini P., Anna Mische John, and David A. Evans. 2008. Determining the Extent of Market and Extent of Resource for Stated Preference Survey Design Using Mapping Methods. In *RFF Discussion Papers*. Washington, DC: Resources for the Future.

Van Houtven, George, John Powers, and Subhrendu K. Pattanayak. 2007. Valuing Water Quality Improvements in the United States Using Meta-Analysis: Is the Glass Half-Full of Half-Empty for National Policy Analysis? *Resource and Energy Economics* 29:206-228.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. The Value of Statistical Life: A Critical Review of Market Estimates Throughout the World. *Journal of Risk and Uncertainty* 27 (1):5-76.

von Haefen, Roger H. 2003. Incorporating Observed Choice into the Construction of Welfare Measures from Random Utilty Models. *Journal of Environmental Economics and Management* 45:145-165.

Wooldridge, Jeffrey M. 2003. *Introductory Econometrics: A Modern Approach*. 2d ed. ed. New York: South-Western College Publishing.

Young, C. Edwin. 1984. Perceived Water Quality and the Value of Seasonal Homes. *Water Resources Bulletin* 20 (2):163-166.

Zabel, J.E., and Katherine A. Kiel. 2000. Estimating the Demand for Air Quality in Four U.S. Cities. *Land Economics* 76 (2):174-194.