College of Saint Benedict and Saint John's University
DigitalCommons@CSB/SJU

CSBSJU Distinguished Thesis | Undergraduate Research

2021

Effect of the Mississippi River on Property Values in Anoka County: A Hedonic Price Analysis

Vince Parisi
vparisi001@csbsju.edu

Follow this and additional works at: https://digitalcommons.csbsju.edu/ur_thesis

Recommended Citation

Parisi, Vince, "Effect of the Mississippi River on Property Values in Anoka County: A Hedonic Price Analysis" (2021). *CSBSJU Distinguished Thesis*. 17.
https://digitalcommons.csbsju.edu/ur_thesis/17

This Thesis is brought to you for free and open access by DigitalCommons@CSB/SJU. It has been accepted for inclusion in CSBSJU Distinguished Thesis by an authorized administrator of DigitalCommons@CSB/SJU. For more information, please contact digitalcommons@csbsju.edu.

Effect of the Mississippi River on Property Values in Anoka County: A Hedonic Price Analysis

By Vince Parisi

Approved by:

Samrat Kunwar, Ph.D

Thesis Advisor
Assistant Professor of Economics

Sucharita Sinha Mukherjee, Ph.D

Faculty Reader
Department Chair of Economics

Yasin Arafat, Ph.D

Faculty Reader
Assistant Visiting Professor of Economics

Lindsey Gunnerson Gutsch, M.S.

Program Director
CSB/SJU Distinguished Thesis

ABSTRACT: Minnesota is known as a land of plentiful lakes, most of which provide a high economic value to the communities that surround them. This led us to question whether rivers play the same role to surrounding real estate. We obtained special data and housing and neighborhood characteristics from the Minnesota Geospatial Commons for our sample of 4125 single family homes. We collected crime rate data from a neighborhood evaluation website, educational data from the Minnesota Department of Education and water quality data from the MPCA. Distance to the river was calculated with ArcMAP GIS software. A buffer was also used to group properties based on water quality monitoring stations. The data was summarized and then three separate regressions were run. The first regression, the basic model, looked at the relationship between house price and the two environmental variables. The secondary model included those variables along with more house characteristics and the full regression included all our experimental variables. Distance to river was significant through all three models and was shown to have a positive effect on house price when comparing the first buffer distance to the second. This study shows support for the economic value of riverside development, results showed that property adjacent to the river is valued 32.2% higher on the market. As there is much undeveloped riverside property in Anoka County we advocate for sustainable development and redevelopment

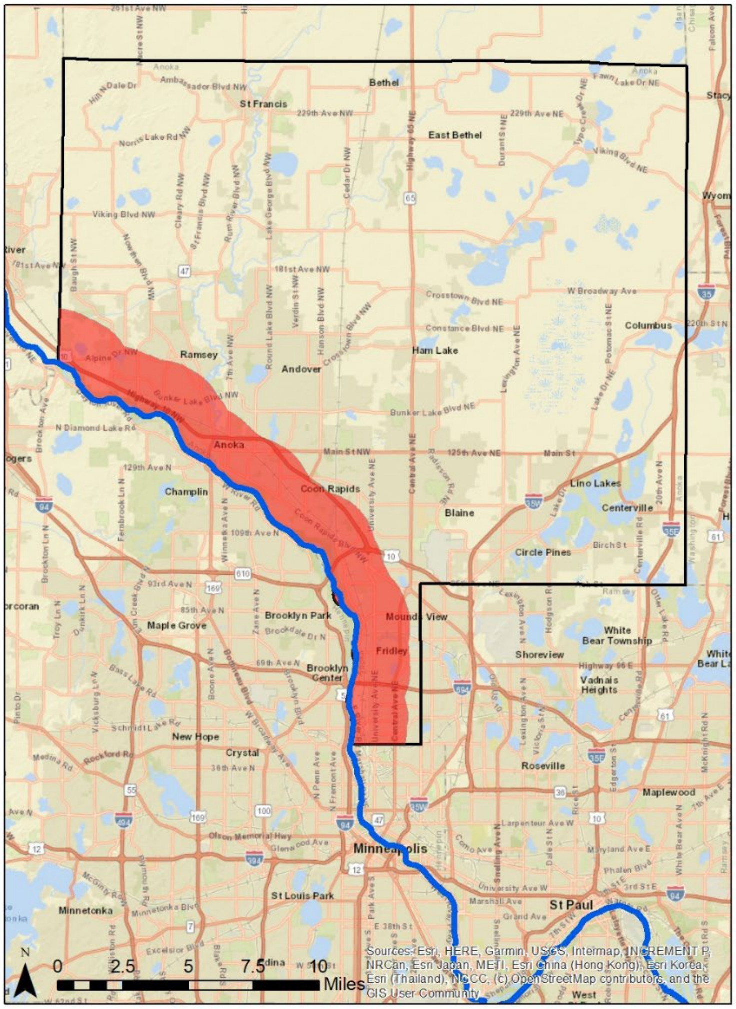
1. Introduction

Water is a fundamental part of Minnesota. The state is known to be a land of 10,000 lakes but not well recognized as the home to the headwaters of the Mississippi river. The Mississippi River "has long been one of the defining natural feature of the United States" (Weller & Russell, 2016) and a valued resource long before the British came across the Atlantic. Water from Minnesota flows north to the Hudson Bay, east to the Atlantic, and south through the Mississippi to the Gulf of Mexico. While there are over 69,200 miles of streams in the state, total wetland acreage has decreased by 8 acres since 1850 ("Lake, River and Wetland Facts", 2013). There are 680 miles of the Mississippi in Minnesota out of the total 2,350 miles, roughly thirty percent.

The Mississippi River is a fundamental part of Minnesota and a foundational natural feature of the United States. As the second longest river in North America, it runs for approximately 2,350 miles from Lake Itasca in upper Minnesota all the way to the Gulf of Mexico (National Park Service, 2021). All or parts of 32 states and 2 Canadian Provinces drain into the Mississippi, making the Greater Mississippi Watershed accountable for about 40% of the drainage in the United States (Weller & Russell, 2016). Native Americans have lived along and used the river for thousands of years before the colonization of the US. The name "Mississippi" is closely related to the what the Ojibwe Indians of Northern Minnesota called the river "Messipi" or "Big River" (Weller & Russell, 2016). The river is home to at least 260 species of fish, 50 mammal species, 145 species of amphibians and reptiles, and 38 species of mussels (Weller & Russell, 2016). The river also provides drinking water to 18 million Americans, 1 million of which are in Minnesota (Weller & Russell, 2016). There are 29 dams on the river that allow water-bound navigation from Baton Rouge, La to Minneapolis, MN (National Park Service, 2021). Additionally, the river connects the state the rest of the world, allowing goods to be imported and exported to and from other states and other countries. The main commodities that Minnesota sends downstream on the river are grain (87%), and asphalt (6%). Commodities imported by river to Minnesota are sand and gravel (49%), fertilizer (24%), salt (9%), and cement (10%) (National Park Service, 1999).

Rivers, by nature, have a more complex relationship to their surrounding environment than lakes do. Benefitting economically from a river does not require the same conditions that aesthetic and recreational benefit require (Parker and Oates, 2016). For example, mills mainly depend on the flow of a river but for a river to be a recreational resource for a city it must be safe to swim and fish in. Rivers can also be a disservice when flooding occurs, or pollution impacts the health of native ecosystems. Polluting is harmful in many ways. It can cause physical harm and bring disease to those who ingest it and prevent the use of a natural resource for recreation. Pollution also causes harm to the ecosystem by putting wildlife in danger, and can shorten the life of a resource, potentially barring its use from future generations. It is important to assign monetary values to environmental resources for a multitude of reasons, but this is a prominent one. In the case of a damage to the environmental resource it is important to have assigned a monetary value to the resource to fight the polluter in court (Tietenberg & Lewis, 2018). For example, in the Deepwater Horizon BP oil spill, the settlement paid by BP for damages was 8.8 billion while cleanup efforts cost 30 billion. Since the resource cannot represent itself in a court of law, a fair estimate of its worth is as close as a river, forest, endangered species, etc. can get to physical representation. Damage estimates are also used in policy decisions like determining acceptable levels of pollution in a river (Tietenberg & Lewis, 2018).

Figure 1



Anoka county location relative to the Mississippi river and Twin City area

A healthy and clean river can benefit society in a multitude of ways. Parker and Oates summarize some of these benefits in their 2016 paper titled, How do Healthy Rivers Benefit Society. To summarize, healthy rivers can provide a space for fisheries and allow for recreational fishing. Currently, much of the river is impaired and has limits on consumption from fish harvested from the river due to unsafe levels of mercury. Rivers are a necessary source of freshwater for the communities that surround them, and if polluted can cause greatly increase the cost of water. Rivers provide a great aesthetic value to those who live close to them and have been "closely associated with the development of human culture". An unhealthy and polluted river can negate all potential benefits. It can reverse the role of a river to its surrounding community from a generator of property value and industry potential to a major health risk and financial burden.

Valuation methods can be broken down into two categories: stated preference and revealed preference methods. Hedonic valuation is a revealed preference model as it is observable and involves actual market behavior (Tietenberg & Lewis, 2018). The goal of this Hedonic valuation is to provide empirical support for the relationship between property value and distance from the river. If living near the river increases property value holding all other factors constant, policymakers may have an increased incentive to keep the river clean and allow it to support more recreational use. In our study we use the dependent variable property value and its relationship to distance from the river along with many other explanatory variables to infer the effect the Mississippi river has on surrounding property. Since property value is influenced by a multitude of factors, we aim to control for as many of these as we can. Variables that measure the housing characteristics like the presence of a garage were included along with neighborhood characteristics like crime rate and percentage of graduates from the corresponding school district level were included.

Distance_to_river was shown to have a significant positive relationship with *lnsale_value* when comparing the first buffer distance to the second. This relationship held at all levels of our regression analysis. In the full model *RIVERADJ* was significant at the 1% level and suggested that adjacent to river location of property lead to a 032 percent increase in sale value. Our water quality variable *do* was significant and positive in our basic model but not through the secondary and full models. We suspect that this variable was calculated inaccurately, and, in a follow up study, will hope to have more measurements of water quality. As our study provides support for increasing the development of riverside property we advocate for it sustainably, with an emphasis on redevelopment. We advocate for development to include permeable pavement and implementation of urban bioswales to reduce runoff as much as possible.

The paper proceeds as follows: the introduction above is followed by section II, the literature review, followed by the hypothesis section. This is followed by the theoretical and empirical model, which contains the data source and variable section as well as graphical and regression analyses. This is then followed by the results section and then the conclusion, which contains policy extensions of the research.

2. Literature Review

Our study was most closely influenced by Bonetti, Corsi, Orsi, and De Noni (2016) who investigated the relationship between proximity to urban rivers and streams and the water quality of the waterbodies in Northern Italy. The main difference from our research is that this study aims to evaluate the distance between natural streams and man-made canals. The survey area was the province of Milan in Northern Italy, the largest real estate market in the country. The data collection methods and theoretical model of this paper heavily influenced our method. The over 10,000 data points came from a web scraping software and was analyzed using GIS, while ours came directly from a GIS database. The hedonic price model was used, and the composition of variables was again like our variable set. The researchers calculated proximities "as the crow flies" using GIS software. We grouped our properties into 16 buffer zones instead of calculating each parcels individual distance to the river. The study also included interaction variables between water quality and proximity to river. We would have liked to include this in our model but could not due to both being categorical variables. The model was semi-logarithmic, with the dependent variable being the natural log of sale price. The researchers found a negative relationship between proximity to the river and property value (sale price) while proximity to artificial canals related positively to the dependent variable. Water quality was found to have an independent effect from proximity was always positively related to the dependent variable.

Krysel, Boyer, Parson, and Welle (2003) wanted to determine whether water quality in the Mississippi headwaters region affects lakeshore property values and hypothesized that it does. The study focused on the water quality of lakes because they are an important part of Minnesota's cultural, recreational, and economic makeup. As custom in hedonic studies the researchers controlled for other property characteristics to isolate the effect of water quality. Water quality was a concern for the researchers because in the past 50 years lakefront development has become increasingly prominent. The researchers state that development negatively impacts water quality by increasing levels sediment and pollution. This occurs because instead of sinking into the ground water travels across paved surfaces to drain and brings pollutants with it into the water. Krysel et al. obtained water quality data from the Minnesota Pollution Control Agency. The data obtained from each property was purchase price, feet of frontage on the lake, lot size, tax rate, square footage, stories, fireplace, central heating, and garage. The water quality measure used was the mean Secchi disk reading for the lake in the year the property sold. Lakes were grouped by similar water quality values and in the hedonic model used, purchase price is a function of property characteristics, structural characteristics, land characteristics, the natural log of water quality, multiplied by the size of the lake. Krysel et al. found that increased runoff resulted in a decreased aesthetic quality for the lake, decreased recreational benefits, lowered the overall property value, and discussed that water quality was a significant explanatory variable in all the lake groups. They believe their study is important for policy decisions because precautionary treatment is more economical than ecosystem restoration.

While not a polluted lake in Minnesota, the Buffalo River, New York Area of Concern is a degraded area in New York state designated for priority remedial action. The area of concern (AOC) lasts for 6.2 miles and is known to contain toxic chemicals in its sediment. This is largely due to the decline of an industrial complex on the banks of the river. AOC's not only danger local ecosystems but can lower the values of property that surrounds them. The Great Lakes Regional Collaboration determined that remediation of all US AOC's can cost up to 4.5 billion. It was important for Braden et al. (2008) to determine whether the degraded area does negatively impact property value since the remediation prices are so high. The authors conducted both stated preference and revealed preference methods. The researchers used data from 3,474 single family homes within five miles of the AOC. The researchers used GIS software to measure geographical variable distances as we did in our study. Braden et al. (2008) did a hedonic analysis based on distance from the AOC, which is very similar to the design of our experiment. The hedonic model stated that property value is a function of structural characteristics like stories or fireplaces, lot characteristics like acreage, and neighborhood characteristics like school district and proximity to the city center. The hedonic models revealed a significant impact of proximity to the AOC in the southern portion of the study area but not the north. The specific results of the regression reveal that for a 1 percent increase in distance from the AOC the property value increases by $67.69. The impact to the north was small and not very significant. This is likely since very few residential properties exist to the north of the river until the central business district is reached. The property value loss to the south is significant, within the five miles to the south of the river property value losses exceed 118 million. Braden et al. (2008) were able to come up with a formula to determine this loss and it was essentially the predicted value of the property at a certain boundary distance from the AOC minus its actual value.

Lewis, Bohlen, and Wilson examined the impacts of dam removal on property value removal in a study conducted in 2008. Removal of dams have shown to increase the reach of fish populations and increase river recreation. Since the removal of the Edwards Dam in 1999 many other dams on Maine rivers have been removed, and the researchers aim to see if this is a valuable endeavor for the local real estate. Since dam use is expensive and controversial it is important to understand all the effects, more than just the effect on real estate. Opposition to dam removal is usually over the issue that it will lower the water level and destroy certain recreational and aesthetic aspects. Opposition the removal of the Edwards dam was strong, but since its removal fishing and boating have improved in the affected area. Real estate data from 18 towns bordering the Kennebec River was obtained from 1997 to 2005 in the study. There was a total of 7874 data points in the initial set. The data included sales price and other structural characteristics. Properties were geocoded with GIS software to control for the components of its surrounding environment. The researchers chose to focus on lots with 1 acre or less of property. After filtering the data, the final data set consisted of 1200 properties. Distances of the properties from the river and the dam sites were calculated like how we calculate driving distance to certain environmental and neighborhood sites in our study. Land use variables for each property were obtained from a Maine GIS set. Open space was calculated by pixelating a map of the area and looking for the ratio of open to used. The natural log of the housing price is used as the dependent variable. According to the study it appeared that homes lose value with age, and that housing prices rise by about 6% in Augusta and 5% in Waterville. The results also show that, before the removal of the dam, homeowners were willing to pay $2000 to be a half mile away from the dam and after the removal that price shrinks to $134. This may be due to improved fishing and recreation above the removed dam. It is interesting to note that living close to or near the rivers in the study has a negative relationship with property value as this is contrary to our hypothesis.

Lewis and Landry (2017) conducted a hedonic analysis of the value of living close to three rivers and four markets in Maine. Like Lewis et al. the study takes place in Maine where rivers are known to have a negative relationship to property value due to damming and heavy industrial use. The removal of the Edwards Dam in 1999 set a precedent for better river treatment and the negative relationship is beginning to change. This study evaluates existing hedonic models to see how benefits compare across different rivers and different stages of restoration. Benefit transfer applies existing benefits of removing a resource (a dam) to another resource (river) and according to the researchers, few benefit transfer analyses of hedonic valuation have been conducted. The hedonic model of the study follows the standard format where the dependent variable, ln of house price, is influenced by environmental, geographical, and structural characteristics. The data set includes information on 12000 retail properties from the years 1997 to 2006. GIS software was used to calculate geographical data and other data came from the US Census. Most of the structural characteristics were statistically significant and willingness to pay for property increased as distance from the river increased. Until recently, Maines's rivers were badly polluted and smelled unpleasant in the summer. Additionally, much of the river shoreline is filled with industrial buildings like mills and factories. This in turn caused consumers to negatively value living close to the river. This provides an economic incentive for river remediation, and this is something we will explore in our study. Our paper is one of the few papers to focus on the hedonic value of the Mississippi river. Like how Krysel et al. state that there is already significant pollution in Minnesota lakes, many parts of the Mississippi River are impaired as well. The segment of the river that we focus on in our study was last assessed in 2019 and was found to be impaired in the three assessment categories of Aquatic Consumption, Aquatic Life, and Aquatic Recreation (Minnesota Pollution Control Agency). In our study as we analyze dissolved oxygen concentration, pH and distance from the river in different reaches of the Anoka County segment and regress it with the property value. We follow a similar approach to Bonetti et al., in our analysis of distance from the river and its relation to property value but selected a sample of approximately 4000 single family homes within 2 miles of the river area and did not expect proximity to the river to decrease property value. Lewis and Landry and Braden et. al (2008) use Euclidian distance to the river but in our study, we grouped the properties into 1/8-mile buffer increments.

3. Theoretical Model

Theory:

As explained above, the basic hedonic model uses existing market transactions to infer the value of a resource, in our case the Mississippi River. By observing the consumer purchases in our selected sample, we can recover marginal willingness to pay for certain housing attributes. The typical hedonic model is as follows:

$$1 \rightarrow P_h = f(S, N, Z)$$

$$2 \rightarrow P_h = f(dissolved\ oxygen, RIVERADJ, acres\ poly, year\ built, Garage, BASEMENT, HEATING, home\ style, percent\ graduate, CRIME\ RATE)$$

Where P_h is the real price of the housing unit and it is a function of S, structural characteristics like square footage, N, neighborhood characteristics like crime rate, and Z, which are locational and environmental characteristics like driving distance to the river and phosphorus concentration. We decided on the variables used in each category from a in depth analysis of existing hedonic studies of rivers and lakes. Our variables for the regression analysis are broken down into three categories: housing, neighborhood, and environmental.

4. Hypothesis

While living close to lakes or oceans is seen as desirable, we believe that living close to rivers does not cause the same property value increase. Rivers can pose a tremendous flood risk which has happened multiple times in the twin cities. While residential flooding has been contained in recent years hydrologists suspect that flood walls may drive floods higher by not giving the river a natural floodplain (Woltman, 2019). Rivers are also subject to pollution, as there is more industry that depend and use the river than lakes or oceans. Due to these reasons, we hypothesize that as house distance from the river decreases property value will also decrease, however we believe it will decrease more in less transparent and polluted segments of the Mississippi.

5. Empirical Model

Basic regression: Isolating the effect of the environmental variables on ln*sale_value*

$$3 \rightarrow lnsale\_value = \beta_0 + \beta_1 DISTANCE\_TO\_RIVER + \beta_2 ph + \beta_3 dissolved\_ox + \beta_3 RIVERADJ + e_n$$

Secondary regression: ln*sale_value* vs environment variables and housing characteristics

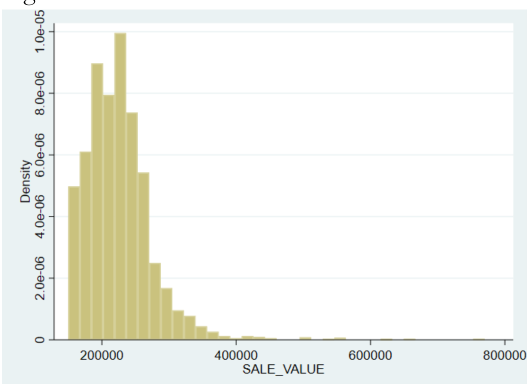
$$4 \rightarrow lnsale\_value = \beta_0 + \beta_1 DISTANCE\_TO\_RIVER + \beta_2 ph + \beta_3 dissolved\_ox + \beta_4 RIVERADJ + \beta_5 acres\_poly + \beta_6 year\_built + \beta_7 Garage + \beta_8 BASEMENT + \beta_9 HEATING + \beta_{10} home\_style + e_n$$

Full regression model: ln*sale_value* vs environment variables, housing characteristics and neighborhood characteristics

$$5 \rightarrow lnsale\_value = \beta_0 + \beta_1 DISTANCE\_TO\_RIVER + \beta_2 ph + \beta_3 dissolved\_ox + \beta_4 RIVERADJ + \beta_5 fireplace + \beta_6 acres\_poly + \beta_7 year\_built + \beta_8 Garage + \beta_{10} BASEMENT + \beta_{10} HEATING + \beta_{10} home\_style + \beta_{10} percent\_graduate + \beta_{10} CRIME\_RATE + e_n$$

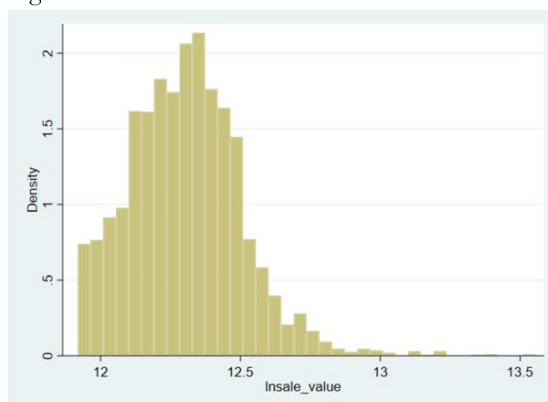
We used ln*sale_value* instead of *sale_value* because the natural log of the variable is more normally distributed (Figure 4), and we wanted our dependent variable to be consistent with the assumptions of the OLS method and the normal distribution of the error term.

Figure 2



Skewed distribution of SALE_VALUE

Figure 3



Distribution of *lnsale_value*

6. Results

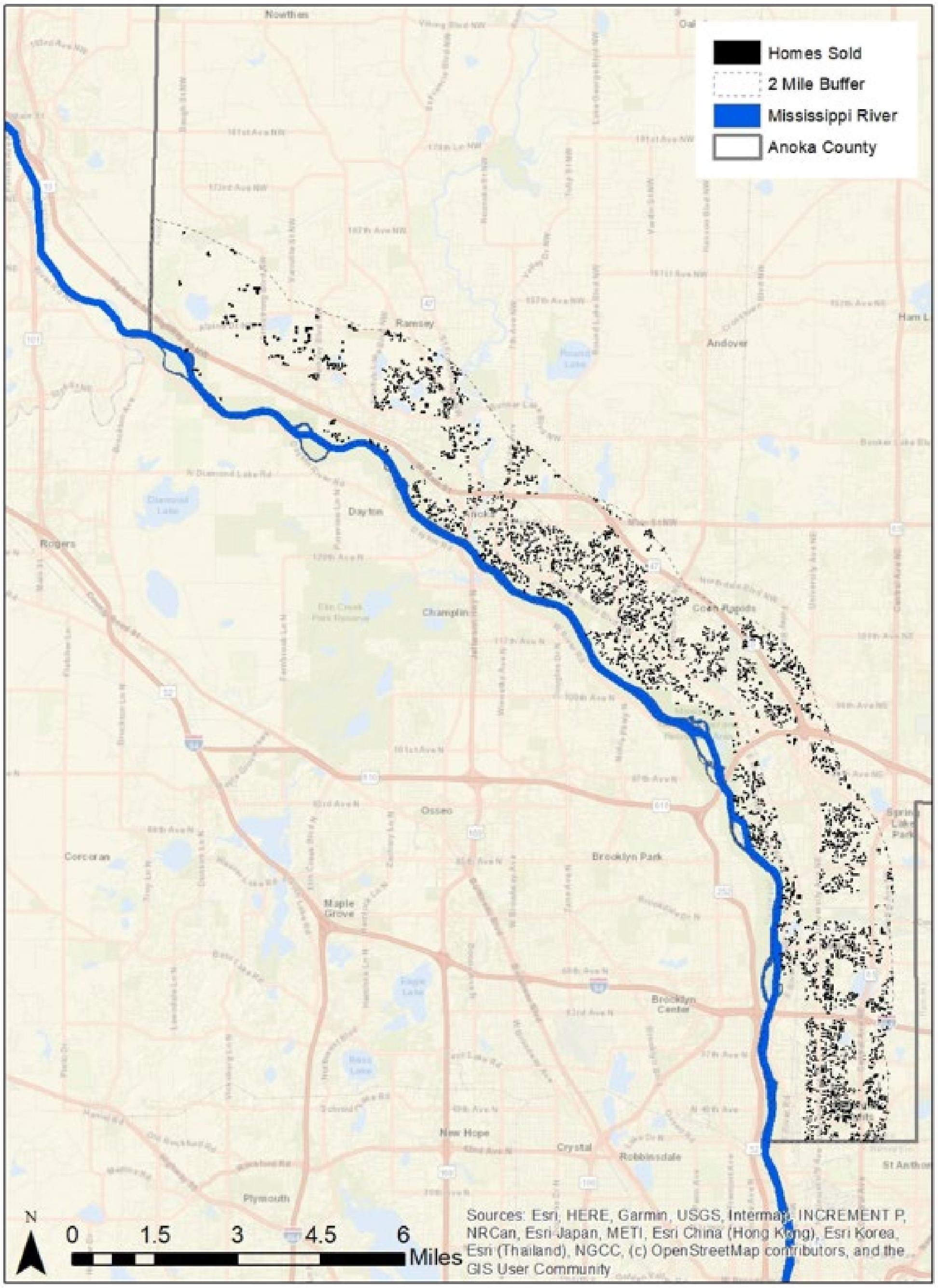
6.1 Data Source and Variables

The final dataset is a sample of the Metro Regional Parcel Dataset from the Minnesota Geospatial Commons. It includes 4125 single family homes in Anoka County that are contained within two miles of the Mississippi river. We used ArcMap GIS software to calculate distance from the river, sort data, and extract housing characteristics like, finished square feet, home style, acreage, year built, type of heating and presence of a garage and a basement. Data from the original dataset was filtered from over 100,000 observations to approximately 10,000 by eliminating parcels that didn't have sale value data and resided on plots of land greater than 1 acre. This dataset was further filtered by removing properties sold before 2016 and after 2019. We chose to do this to retain recent value after inflation but remove the effect of the pandemic on home prices. We chose not to include parcels with a sale value of under 150,000 or over 1,000,000 for consistency and an aimed isolation of environmental variables. We also filtered out new construction after 2016 as new construction will naturally sell for a higher value. A series of other variables from alternative data sources were then collected. First, we collected the crime per 100k people in the zip code where each property was listed. The data came from areavibes.com[1]. Then, we collected the percentage of graduation rate in each homes corresponding school district. This data came from the Minnesota Department of Education Report Card. Next, we collected water quality data in the form of dissolved oxygen concentration. The data for this variable comes from the Minnesota Pollution Control Agency. Dissolved Oxygen (DO) is a measure of how much usable oxygen exists for river organisms. Low levels of DO can result in unsightly and undesired agal blooms and DO levels are usually measured with a water quality probe meter. Dissolved oxygen is also a categorical variable with the first level having the highest DO measurement and the third level being the lowest value. We grouped our 4125 observations into three zones based on the locations of the stations on the river. This was done due to sparce water quality measurements of the Mississippi River in Anoka County. In extensions of this research, we aim to see the effect of more accurate and representative water quality measurements. Additionally, we would like to see this relationship across a greater reach of the river, for

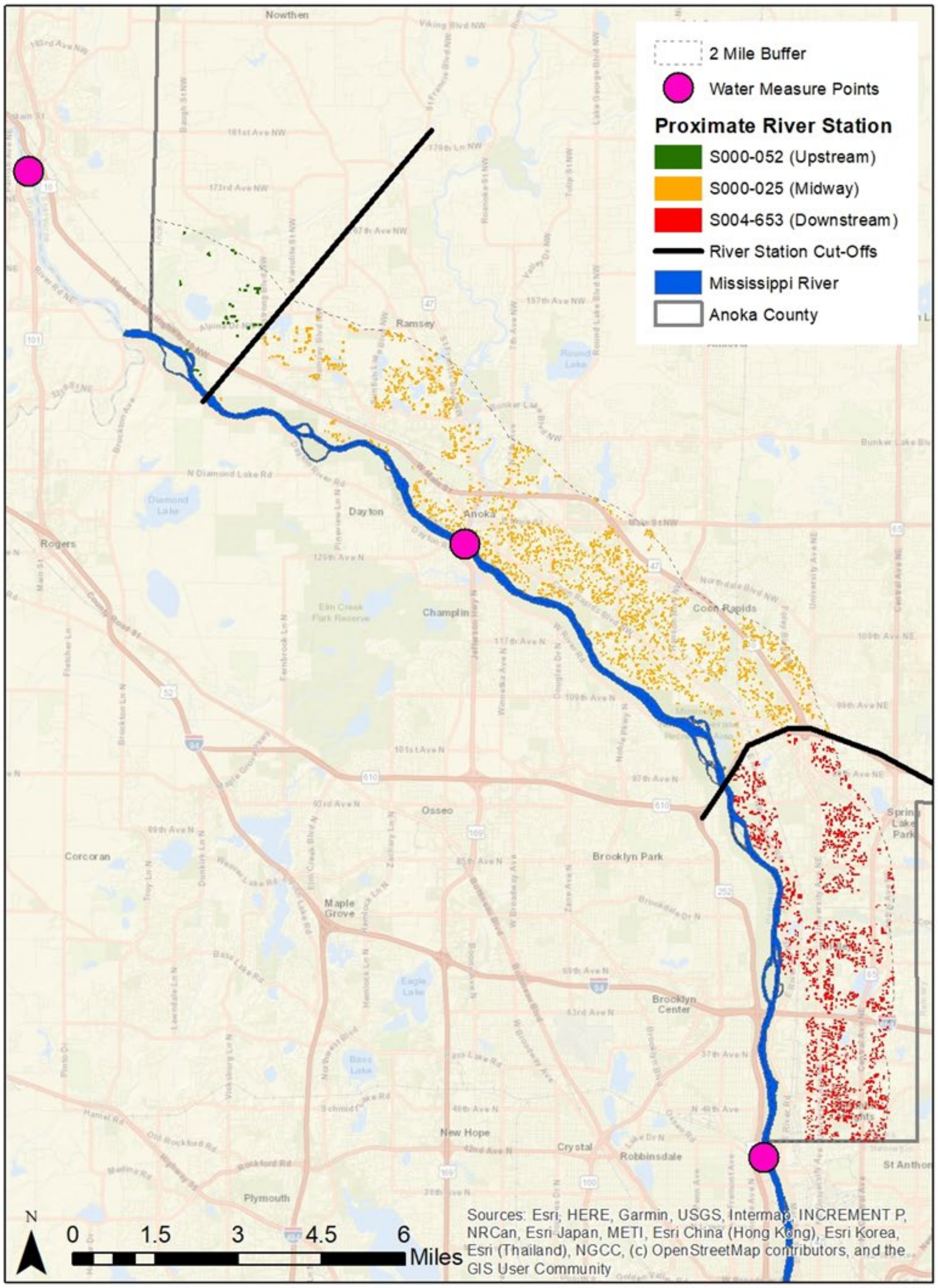
[1] https://www.areavibes.com/

example throughout the entire seven-county metro area. Our distance variable is divided into 5 buffers, with each buffer encompassing 3300 feet. This was done instead of calculating each individual parcels distance due to limitations in GIS ability and time constraints in the research. In a future study we would like to calculate the distance of each individual property to the river and have a continuous linear variable instead of a five-level categorical variable.

Figure 4



All parcels included in final dataset
Figure 5



Zoning of parcels in relation to water quality monitoring stations

Table 1

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Dependent						
sale_value	Price of home during year of sale	4124	25615.21	50630.23	150000	770000
Housing						
acres_poly	Lot size of home in Acres	4124	.25	.15	0	1
fin_sq_ft	Finished square feet in home	4124	1217.34	351.89	540	5056
year_built	Year home was built	4124	1970.01	22.12	1876	2016
Garage	Dichotomous: presence of a garage	4124	1.89	.309	1	2
home_style	Categorical variable: levels of home	4124	12.76	4.86	1	22
BASEMENT	Dichotomous: presence of a basement	4124	2.86	.52	1	3
HEATING	Categorical: type of heating	4124	3.96	.37	1	6
Neighborhood						
percent_graduate	Graduation rate of school district	4124	86.07	2.89	77.8	89.7
crime_rate	Crime rate in ZIP	4124	2723.77	1395.34	851	4930
Environmental						
do	Concentration of dissolved oxygen	4124	9.46	.05	9.41	9.75
RIVERADJ	Dichotomous: Home has river view	4124	1.01	.086	1	2
distance_to_river	Categorical (buffer): distance from river	4124	8.62	4.36	1	16

The mean sale value for the homes in our dataset is $225,615.21 with a standard deviation of 50630.24, implying a relatively big spread of house price values from the mean. The average number of acres was .245 with a low standard deviation of 0.148. The average number of finished square feet was 1217.338 with a standard deviation of 351.892. The average year that single family home in our dataset were built in was 1970.01 with a standard deviation of 22.116. Garage was a dichotomous variable with two levels, presence of a garage and absence of a garage. Home style was also a categorical variable with three levels that correspond to floors of a home. Basement was another dichotomous variable with two levels identical to Garage. Heating is a categorical variable with 5 levels. The mean graduation rate of the sample was 86.067 percent with a standard deviation of 2.885. The mean crime rate of the sample was 2723.77 with a high standard deviation of 1395.34. The mean dissolved oxygen concentration was 9.46 with a standard deviation of .05. RIVERADJ was a dichotomous variable. Distance to the river was also a categorical variable with five levels.

6.2 Graphical Analysis

We used Stata-64 to carry out our regression and correlations, to obtain our graphical figures, and we also manipulated two variables using the software. The *sale_price* had a skewed distribution, so we used the *lnsale_value* instead of sale value as it was more normally distributed. Figure 6 is a scatter of our distance variable, *d*, versus our dependent variable, *lnsale_value* this scatter shows an initial negative relationship and then a positive relationship. Figure 7 is a scatter of *dissolved_oxygen* versus *lnsale_value* and shows a slightly positive relationship. Figure 8 is again a scatter, of *RiverAdj* versus *lnsale_value* and shows that river adjacent properties generally have a greater sale value.

Figure 6

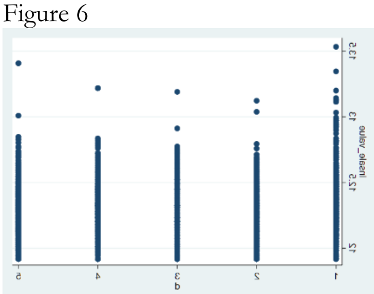


Figure 7

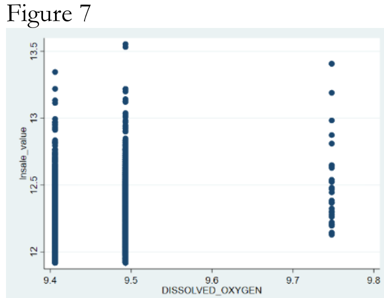
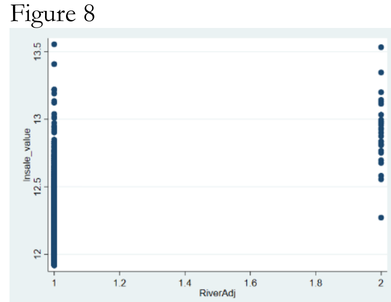


Figure 8



Scatter of lnsale_value vs d

Scatter of dissolved_oxygen vs lnsale_value

Scatter of RiverAdj vs lnsale_value

6.3 Regression Analysis

VARIABLES	(1) Basic Model	(2) Secondary Model	(3) Full Model
2.do	0.0449***	-0.0184***	0.0120
	(0.00616)	(0.00533)	(0.00877)
3.do	0.164***	-0.0999***	-0.0603*
	(0.0350)	(0.0296)	(0.0311)
2.d	-0.103***	-0.0442***	-0.0448***
	(0.0108)	(0.00862)	(0.00860)
3.d	-0.0553***	-0.0155**	-0.0164**
	(0.00995)	(0.00789)	(0.00787)
4.d	-0.0365***	-0.00925	-0.00842
	(0.0101)	(0.00798)	(0.00797)
5.d	-0.0549***	-0.0177**	-0.0152*
	(0.0109)	(0.00856)	(0.00856)
2.RIVERADJ	0.548***	0.321***	0.323***
	(0.0350)	(0.0282)	(0.0281)
acres_poly		0.256***	0.260***
		(0.0199)	(0.0201)
fin_sq_ft		0.000244***	0.000246***
		(7.76e-06)	(7.77e-06)
year_built		0.00241***	0.00244***
		(0.000125)	(0.000126)
Garage			0.158***
			(0.00861)
BASEMENT		0.0838***	0.0829***
		(0.00500)	(0.00500)
HEATING		0.00764	0.00686
		(0.00646)	(0.00644)
2.home_style		0.00386	0.00680
		(0.00559)	(0.00563)
3.home_style		0.0605***	0.0598***
		(0.00793)	(0.00791)
crime_rate			1.42e-05***
			(3.15e-06)
percent_graduate			0.00323***
			(0.000949)
2.Garage		0.157***	
		(0.00862)	
Constant	12.32***	6.805***	6.254***
	(0.00892)	(0.247)	(0.275)
Observations	3,877	3,833	3,833
R-squared	0.113	0.471	0.475

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

Our basic regression model isolated the effects of 3,877 observations of the environmental variables *d* (distance from river), *RIVERADJ*, *dissolved_oxygen* and *ph* on our dependent variable, *lnsale_value*. The variables *do*, *d*, and *RIVERADJ* are all categorical variables meaning that the coefficients are in relation to the reference category. In relation to the reference category d1 d2 was significant with a p-value of 0.000. From buffer distance d1 to d2 the natural log of sale_value decreases by 10.2 percent. D3, d4 and d4 were additionally all significant implying a positive relationship between distance to river and property value. It is important to note that the relationship between buffer levels is not linear in its positive relationship. The value lost as distance increases goes from a 10.3% decrease in value in *d2* to -5.53% in *d3* to 3.65% in *d4* and then to 5.48% in *d5*. This could result from a multitude of reasons, but we hypothesize that distance from the river does not have as strong of an effect if the property is outside the second buffer. Our basic regression model isolated the effects of 3,877 observations of the environmental variables *d* (distance from river), *RIVERADJ*, *dissolved_oxygen* and *ph* on our dependent variable, *lnsale_value*. The variables *do*, *d*, and *RIVERADJ* are all categorical variables meaning that the coefficients are in relation to the reference category. In relation to the reference category *d1, d2* was significant with a p-value of 0.000. From buffer distance *d1* to *d2* the natural log of *sale_value* decreases by 10.2 percent. *D3, d4* and *d4* were additionally all significant implying a positive relationship between distance to river and property value. It is important to note that the relationship between buffer levels is not linear in its positive relationship. The value lost as distance increases goes from a negative 10.3% decrease in value in *d2* to -5.53% in *d3* to 3.65% in *d4* and then to 5.48% in *d5*. This could result from a multitude of reasons, but we hypothesize that distance from the river does not have as strong of an effect if the property is outside the second buffer. Past buffer distance 2 we expect that other factors that determine value become more important than distance from the river which is why there is no linear relationship with levels 3 through 5. In the basic regression the second and third levels of *do* have positive and significant relationships. *D2* has a greater coefficient than *d3* which makes sense because *d1* has the highest *do* value, *d2* the second highest and *d3* the lowest.

In the secondary regression we see *do* flip signs from a positive coefficient to negative. Initially I thought this would be corrected once *ph* was removed from the model. Once this happened it was not the case. There are likely reasons resulting from how the variable is set up. The first level of the variable has only 32 observations and that is likely causing this issue. Additional significant variables are *2.do* (1%), *3.do* (1%), *2.d* (1%), *3.d* (5%), *5.d* (5%), *RIVERADJ* (1%), *acres_poly* (1%), *fin_sq_ft* (1%), *year_built* (1%), *Garage* (1%), *BASEMENT* (1%) and *3.home_style* (1%). We expected *RIVERADJ* to have a positive coefficient but did not realize it would have such a large effect on property value. Additionally, we expected the relationships between our structural variables and *lnsale_value*. These findings supported our hypothesis.

For the final regression the significant variables were *3.do* (10%), *2.d* (1%), *3.d* (5%), *5.d* (10%), *RIVERADJ* (1%), *acres_poly* (1%), *fin_sq_ft* (1%), *year_built* (1%), *Garage* (1%), *BASEMENT* (1%), *3.home_style* (1%), *crime_rate* (1%) and *percent_graduate* (1%). The relationship between *RIVERADJ* and *lnsale_value* held throughout all three models. In the full model, a location adjacent to the river is responsible for a 32% increase in value. This finding implies powerful policy extensions for undeveloped land adjacent to the river. The same relationship between *d1* and *d2* held throughout all three levels of the regression analysis. In the full model, an increase in distance from buffer distance 1 to distance 2 is responsible for a 4.5% decrease in property value. Crime_rate and percent_graduate, the two neighborhood variables in our study have significant relationships to lnsale_value, with coefficients of 1.42e-05 and 0.00323 respectively. This was a smaller relationship than expected.

After originally omitting *ph* due to a near perfectly correlation with *do* with value of .96, we tested all levels of the model for multicollinearity. The mean VIF of Basic Regression was 1.55, 1.47 of the Secondary Model and 1.72 of the Full Model. From these values, we concluded that there is not a problem of multicollinearity within these models. Next, we tested the models' levels for heteroskedasticity. The Breusch-Pagan Test value for Basic model was 0.0000, 0.0002 for the Secondary Model and 0.0003 for the Full Model. After the three values rejecedt the null hypothesis of no variation in the variance, we determined there is heteroskedasticity at all levels of the model. As a result, we ran a robust regression of the model but did not see a change in any relationships.

Table 3

VARIABLES	(1) Basic (Robust)	(2) Secondary (Robust)	(3) Full (Robust)
2.do	0.0449***	-0.0184***	0.0120
	(0.00605)	(0.00544)	(0.00873)
3.do	0.164***	-0.0999**	-0.0603
	(0.0612)	(0.0442)	(0.0456)
2.d	-0.103***	-0.0442***	-0.0448***
	(0.0114)	(0.00890)	(0.00887)
3.d	-0.0553***	-0.0155**	-0.0164**
	(0.0104)	(0.00789)	(0.00788)
4.d	-0.0365***	-0.00925	-0.00842
	(0.0109)	(0.00855)	(0.00855)
5.d	-0.0549***	-0.0177**	-0.0152*
	(0.0116)	(0.00886)	(0.00887)
2.RIVERADJ	0.548***	0.321***	0.323***
	(0.0439)	(0.0332)	(0.0328)
acres_poly		0.256***	0.260***

		(0.0244)	(0.0246)
fin_sq_ft		0.000244***	0.000246***
		(9.62e-06)	(9.58e-06)
year_built		0.00241***	0.00244***
		(0.000141)	(0.000144)
Garage			0.158***
			(0.0108)
BASEMENT		0.0838***	0.0829***
		(0.00473)	(0.00471)
HEATING		0.00764	0.00686
		(0.00719)	(0.00728)
2.home_style		0.00386	0.00680
		(0.00578)	(0.00587)
3.home_style		0.0605***	0.0598***
		(0.00858)	(0.00857)
crime_rate			1.42e-05***
			(3.10e-06)
percent_graduate			0.00323***
			(0.000937)
2.Garage		0.157***	
		(0.0109)	
Constant	12.32***	6.805***	6.254***
	(0.00967)	(0.278)	(0.309)
Observations	3,877	3,833	3,833
R-squared	0.113	0.471	0.475

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

The robust regression did not differ from the Full Model.

Due to a perceived limitation with the *do* variable in our analysis, we conducted the likelihood ratio test aiming to see if the fit of regression variables would be improved after adding *do* to the model. We regressed *lnsale_value* on our full set of variables excluding do and then ran a full model regression including do to see if the addition of this variable would improve the fit of the model.

Table 4

Likelihood-ratio test	LR chi2(2) = 15.35
	Prob > chi2 = 0.0005

From this probability we can conclude that *do* still improves the fit of the model, however its increase and effect on the Adjusted R^2 value is so minimal (.4648 with *do* included in the model vs .4630 with *do* excluded) that it may make sense for these results to omit do from all levels of analysis.

7 Conclusion

In our study we aimed to examine the effect of driving distance to river and transparency on house price. We collected data from 4125 single family homes on whether the homes had a garage or basement, the lot size of each property, the finished square feet inside the property, the type of home style, distance from the river, year the home was built, whether it was located adjacent to the river, the dissolved oxygen concentration, the crime rate of the corresponding zip, and percent graduation rate of the school district the property is located in. We obtained this data from the Minnesota Geospatial Commons, the Minnesota Pollution Control Agency, the Minnesota Department of Education and Areavibes.com. We divided our data into a 5-category buffer of distances from the river, each buffer measuring 2112 feet. We divided our data sample into 3 zones based on geographical proximity to the river water quality monitoring stations where the measurements were collected. We added one variable to our regression analysis, *lnsale_value*, to improve the quality of our analysis. We ran regressions at three levels and our distance variable *distance_to_river* remained significant along with *RIVERADJ* at all three levels while *do* remained significant only in the first model.

This leads us to believe that there are problems with the variable *do* as it may not have been collected properly. We transpose only three water quality measurements to over three thousand observations which is not standard in hedonic studies. We were limited by collection methods and data availability and as a result do not strongly rely on our interpretation of the effect of dissolved oxygen concentration on property value.

Rivers pose a more complex relationship to the surrounding property than lakes do. Urban rivers are often not seen as an economic asset, as they are subject to pollution and flooding. Yaacovi et al (2021) argue that rivers are a vital urban resource and provide high natural economic value due to their connectedness with nature and benefit to social well-being. Especially in places like Minneapolis and St. Paul, where approximately 46% of Minnesota's 1.3 million anglers reside (U.S. Department of the Interior, Fish and Wildlife Service, 2002), a clean and healthy river can provide high recreational value for the cities it runs through. Our research shows a clear drop off in value from within 2112 feet of the river to outside of 2112 feet from the river. With property values decreasing by 4.5% in the full model once it is located outside the first buffer zone. Additionally, we see that location of a property adjacent to the river increases its sale value by 32%. This provides significant support for riverside development and redevelopment of vacant land adjacent to the river. It provides a strong policy incentive for riverside development and development of property within 2112 feet of the Mississippi in Anoka County.

While we were not able to provide significant support for the valuation of increased dissolved oxygen levels, it should be a concern of local governments to maintain and increase the clarity of the Mississippi in Anoka County. Possible policy implication and urban runoff solutions like increased vegetation on the rivers banks to absorb runoff and permeated urban pavement so runoff can absorb into the ground should still be considered by local leaders. Additional avenues of research may be to use the Contingent Valuation Method to gauge the recreational value of the river in the twin cities. It is also of concern to us to see if the relationship between sale value and distance holds outside the study area and to see the impact distance to the river has on surrounding real estate.

References

National Park Service. (n.d.). *"Working Mississippi": Moving Goods on the River.* https://www.nps.gov/miss/learn/education/upload/WorkingMiss_30x40.pdf.

U.S. Department of the Interior. (2021, April 14). *Mississippi River Facts.* National Parks Service. https://www.nps.gov/miss/riverfacts.htm.

Lakes, rivers, and wetlands facts. Minnesota Department of Natural Resources. (2013). https://www.dnr.state.mn.us/faq/mnfacts/water.html.

Weller, L & Russell, T. A. (2016). *State of the River Report.* Friends of the Mississippi River. https://www.nps.gov/miss/learn/education/upload/WorkingMiss_30x40.pdf.

Tietenberg, T. H., & Lewis, L. (2018). Valuing the Environment: Methods. In *Environmental and natural resource economics* (pp. 89–91). essay, Routledge.

Parker, H., & Oates, N. (2016). How do healthy rivers benefit society. *A review of the evidence. Overseas Development Institute (ODI), London, United Kingdom. Report to WWF-UK, Woking, United Kingdom*, 73.

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

Braden, J. B., Taylor, L. O., Won, D., Mays, N., Cangelosi, A., & Patunru, A. A. (2008). Economic benefits of remediating the Buffalo River, New York area of concern. *Journal of Great Lakes Research, 34(4), 631-648.*

Lewis, L. Y., Bohlen, C., & Wilson, S. (2008). Dams, dam removal, and river restoration: A hedonic property value analysis. *Contemporary Economic Policy*, *26*(2), 175-186.

Lewis, L. Y., & Landry, C. E. (2017). River restoration and hedonic property value analyses: Guidance for effective benefit transfer. *Water resources and economics*, *17*, 20-31.

Woltman, N. (2019, March 27). Flooding along the Mississippi River has menaced St. Paul from its earliest days. Twin Cities. https://www.twincities.com/2019/03/27/3317687/.

Hanley, N., Colombo, S., Tinch, D., Black, A., & Aftab, A. (2006). Estimating the benefits of water quality improvements under the Water Framework Directive: are benefits transferable? *European Review of Agricultural Economics*, *33*(3), 391-413.

U.S. Department of the Interior, Fish and Wildlife Service, and U.S. Department of Commerce, Bureau of the Census. (2002). *2001 National survey of fishing, fishing, and wildlife-associated recreation. Washington, DC*: U.S. Government Printing Office.

Yaacobi, Y., Gasith, A., & Becker, N. (2021). How much is an urban stream worth? Using land senses and economic assessment of an urban stream restoration. International Journal of Sustainable Development & World Ecology, 1-10.

Bonetti F, Corsi S, Orsi L, De Noni I. Canals vs. Streams: To What Extent Do Water Quality and Proximity Affect Real Estate Values? A Hedonic Approach Analysis. *Water*. 2016; 8(12):577. https://doi.org/10.3390/w8120577