Residential Demand for Sediment Remediation to Restore Water Quality: Evidence from Milwaukee

Richard T. Melstrom*
School of Environmental Sustainability
Loyola University Chicago

Selected paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA, July 31 – August 2.

*Address: 424 BVM Hall, School of Environmental Sustainability, 1032 W. Sheridan Road, Chicago, IL 60660. Email: rmelstrom@luc.com; Phone: 773-508-2948

Residential Demand for Sediment Remediation to Restore Water Quality: Evidence from Milwaukee

Abstract

This paper examines the effect of removing pollutants, including polychlorinated biphenyls (PCBs), on property values in Great Lakes Areas of Concern (AOCs). AOCs are heavily polluted locations identified as priorities for restoration under the Great Lakes Water Quality Agreement (GLWQA) between the United States and Canada. Since the signing of the GLWQA, AOCs have undergone cleanup actions that could encourage local redevelopment and raise nearby property values. This paper focuses on the Milwaukee Estuary AOC and estimates property owner willingness to pay using a sorting model and home sales data before and after a major sediment remediation project. Results indicate that owner willingness to pay for cleanup is significant. This paper also examines potential differences in willingness to pay across income and race groups. Results fail to indicate any significant group-level differences in willingness to pay, although they do indicate systematic inequalities as a result of pre-existing sorting patterns.

Keywords: Property values; pollution; polychlorinated biphenyls; PCBs; environmental policy

JEL codes: Q51, Q53, R21, R23

1 Introduction

Water pollution is one of the world's greatest environmental concerns. When asked about environmental problems, ninety-three percent of people globally report being either very or somewhat concerned about water pollution, more than the percentages for air pollution and climate change (GlobeScan, 2009). In the United States, residents are more concerned about water pollution than nearly any other topic, including medical bills and terrorism, and in Canada two-thirds of residents report being very concerned about water quality (Chapman University, 2018; Gallup, 2019). Surveys like these reveal that public concerns about water quality remain elevated despite decades of strict pollutant discharge regulations and cleanup efforts. Lawmakers, public officials and environmental organizations therefore face continuous pressure to address water pollution problems.

Economists have responded to these concerns and the growth in pollution control policies by generating information about the economic benefits and costs of cleaning up water. Since the 1970s, the United States has devoted 0.8 percent of GDP to control water pollution (Keiser and Shapiro, 2019). A large literature examines the economic benefits of these efforts, in terms of residential activities (e.g. Poor et al. (2001, 2007); Tuttle and Heintzelman (2015)), health (e.g. Dwight et al. (2005); McConnell and Rosado (2000)), and recreation (e.g. Bockstael et al. (1987); Carson and Mitchell (1993)). Nevertheless, there remains a pressing need for research on the value of clean water, particularly research examining cleanup actions, as recent evaluations have thrown the welfare consequences of cleanup into question (Keiser et al., 2019).

In this paper, I examine whether recent actions to restore water quality through the Great Lakes Water Quality Agreement (GLWQA) have yielded benefits in one part of the Great Lakes. The GLWQA is a commitment between the United States and Canada to rid pollution from the Great Lakes—a water-rich area that, for many people, epitomizes the potential of pollution control and remediation programs to rehabilitate water resources. To achieve this, the GLWQA has focused water pollution control and remediation actions on heavily polluted locations known as Areas of Concern (AOCs). Canada and the United States have collectively spent $23 billion cleaning up these areas (Hartig et al., 2020). I estimate the benefit of cleaning up part of the Milwaukee Estuary AOC using home sales data before and after a major cleanup event. My identification strategy exploits variation in where remediation occurred to attribute household location decisions and willingness to pay to improvements in the AOC.

This paper makes two novel contributions to research on the economics of water quality. First, it estimates the benefit of an AOC remediation action ex post. For two decades, the standard approach has relied on property value hedonics and sales data collected in the ex ante period. To estimate the benefits of remediation before actual cleanup, economists have relied on proximity to the AOC as a disamenity in the hedonic price function (Braden et al., 2004, 2008; McMillen, 2017; Patunru et al., 2007; Stoll et al., 2002).[1] A weakness of this approach is that unobserved attributes could be correlated with proximity to the AOC. To get around this endogeneity problem, several papers have used choice experiments or combined actual and hypothetical property sales data (Chattopadhyay et al., 2005; Phaneuf et al., 2013). However, the accuracy of these estimates—whether based on distance and actual home sales, or on hypothetical cleanup—remains unclear. This paper helps fill this information gap by measuring the effect of cleanup using pre and post home sales and spatial variation in remediation. This pre-post, control-treated comparison reduces the channels of bias that could have affected prior hedonic estimates. My approach is also unique in that it estimates the benefits of cleanup using a residential sorting model. Sorting models are well-suited to measuring the effects of water quality on residential locations and willingness to pay

[1]Isely et al. (2018) uses a hedonic property value model and data in the ex post period to value AOC shoreline improvements, but does not use a quasi-experimental design to measure the effect of shoreline quality on prices.

because they combine information about housing characteristics with the structure of the choice process, allowing them to account for heterogeneous preferences while predicting the outcomes of large shocks (Kuminoff et al., 2013). Yet little prior work has used sorting models to value water quality improvements.

Second, this paper attempts to shine new light on underlying inequities by examining group-level heterogeneity in residential sorting patterns. There is growing concern among community advocates and government officials that cleaning up pollution tends to benefit some groups more than others due to differences in income, political power, and discrimination. Cleanup actions that ignore these structural factors can fail to correct inequitable exposure and contribute to environmental gentrification (Banzhaf et al., 2019; Melstrom and Mohammdi, 2021). Environmental gentrification occurs when environmentally motivated migration patterns alter the demographic mix in a neighborhood, for example when rich replace poor households due to differences in budget constraints, affordability, and willingness to pay (Banzhaf and Walsh, 2008). This kind of sorting means that disparities in pollution exposure (e.g. between rich and poor households) will persist even if cleanup occurs in disadvantaged neighborhoods, which can fuel environmental justice concerns. By looking for evidence of income and race group-based sorting on cleanup, this paper contributes to an expanding pool of research on the distributional consequences of environmental policies (Bakkensen and Ma, 2020; Banzhaf, 2012; Banzhaf et al., 2019; Depro et al., 2015).

The remainder of the paper is organized as follows. Section 2 briefly describes the history of the GLWQA and AOCs and the geography of the study area. Section 3 describes the residential sorting model. Section 4 summarizes the data. Section 5 presents and discusses the results of the sorting model and willingness to pay estimates. Section 6 concludes.

2 Background on the study area

In 1972, Canada and the United States signed the Great Lakes Water Quality Agreement (GLWQA) to address widespread, deteriorating water quality conditions in the Great Lakes. The GLWQA committed the countries to the adoption of common water quality standards and legislation controlling municipal, industrial and agricultural pollution flowing into the lakes. Water pollution is a prevalent and harmful outcome of economic activity in the Great Lakes because the lakes border many of the largest metropolitan areas in the United States and Canada, including Chicago, Detroit and Toronto, while serving as a source of drinking water and recreation for over 30 million people. The GLWQA was revised significantly in 1987 to focus restoration efforts on 43 Areas of Concern (AOCs) suffering from persistent industrial contamination. As defined in the GLWQA, AOCs are areas that as a result of toxic substances fail to support one or more 14 beneficial uses, which range from animal deformities to beach closings. Often associated with the locations of Superfund sites in the United States, AOCs require extensive remediation before they can be considered free of pollution-related health risks.

Cleaning up AOCs has proven challenging and expensive (International Joint Commission, 2021). Between 1985 and 2019, the United States and Canada spent $22.78 billion on cleanup actions, or nearly half a billion dollars per AOC (Hartig et al., 2020). Wastewater treatment plant upgrades account for the largest share of expenditures (72%), followed by sediment remediation and brownfield cleanup (23%) (Hartig et al., 2020). Seven AOCs have been "delisted" and two more are currently "in recovery" while officials verify the restoration of beneficial uses (Collaboration for Great Lakes Water Quality, 2021). Restoration and rehabilitation actions have occurred at all of the remaining 34 AOCs, which suggests the potential for at least partial recovery at each site.

The Milwaukee Estuary AOC was designated in 1987. The AOC includes nearly all of Milwaukee's Lake Michigan shoreline, as well as its harbor area and connecting tributaries, including the Milwaukee River. Polychlorinated biphenyls (PCBs) have been the focus of cleanup actions since $3 billion in improvements to the regional wastewater treatment system in the 1990s and 2000s. PCBs are linked to cancer and compromised immune, reproductive and endocrine systems. In Milwaukee, PCBs are concentrated in sediments near impoundments and river bends, and are responsible for most of the beneficial use impairments (BUIs) in the AOC. Between 2011 and 2015, the U.S. Environmental Protection Agency led action on one of the most significant PCB deposits in the upper estuary, in the Milwaukee River near Lincoln Park (Dow, 2020). The source of these PCBs is unknown but likely from one or more defunct industrial operations in northwest Milwaukee. This area of the AOC was responsible for increasing PCB transport from about 5 kg/year upstream to about 15 kg/year downstream (Wisconsin DNR, 2005). Cleanup removed contaminated sediments at a cost of $43 million, the largest expenditure on sediment removal in the AOC to date.[2] The next largest suspected source of PCBs is located in the lower estuary adjacent to the city center, approximately five miles downstream from Lincoln Park (Dow, 2020).

Cleanup of the upper Milwaukee River occurred in two phases. Phase 1 removed 5,000 pounds of PCBs as well as 4,000 pounds of polynuclear aromatic hydrocarbons (PAHs) in 2011-2012. Phase 2 removed 2,300 pounds of PCBs and 12,700 pounds of PAHs in 2014-2015. In addition to PCB and PAH contaminated sediment removal, the remediation project restored wetland habitat. Both phases were advertised by local environmental groups and media as a success story in cleaning up the river.[3] The public

[2]This remediation project targeted the following BUIs in the Milwaukee River: restrictions on fish and wildlife consumption, degradation of fish and wildlife habitat, fish tumors or other deformities, degradation of benthos, restrictions on dredging activities, and loss of fish and wildlife habitat.

[3]For example, see the reports by Fox 6 Milwaukee (2012), Healing Our Waters Great Lakes Coalition (2012), and Urban Milwaukee (2015). In 2014, the *Milwaukee Journal Sentinel* ran a series of "River Reborn" stories that ascribed growing housing demand near the Milwaukee River to local water quality improvements (Milwaukee Journal Sentinel, 2014).

could have known about PCBs in this section of the river since 1981, when the Wisconsin Department of Natural Resources issued the first fish advisory (Urban Milwaukee, 2015). However, awareness of contamination likely increased after 2008, when officials updated the AOC boundary to include portions of the upper Milwaukee River and connecting tributaries. Thus, public perceptions of contamination in the Milwaukee River could have begun to shift in 2011, once it became known that water quality in the river was improving.

3 Empirical strategy

3.1 Sorting model

I measure the effect of cleanup on Milwaukee residents using a sorting model. The sorting model describes a household's residential location as a function of housing cost, neighborhood attributes, and household characteristics. The choice set consists of census tracts, which I will refer to as residential locations. Each household i chooses to live in residential location $j \in A_t$ at time t with utility U_{ijt}. The choice set A_t is time-specific because access to residential locations can shift from year to year depending on sale locations. Utility depends on δ_{jt} common to all homebuyers as well an individual-specific component λ^i_{jt}, both of which are a function of observable location attributes. Utility also depends on unobservable location attributes ξ_{jt} and idiosyncratic tastes ϵ^i_{jt}:

$$V^i_{jt} = \delta_{jt} + \xi_{jt} + \lambda^i_{jt} + \epsilon^i_{jt}. \tag{1}$$

The utility function in equation (1) is based on the models developed by Bayer et al. (2007) and Bakkensen and Ma (2020). Let P_{jt} denote the price of living in j and X_{1jt} other location attributes that households have homogeneous preferences over. Then δ_{jt} can be written as

$$\delta_{jt} = \alpha_{X_1} X_{1jt} - \alpha_P P_{jt} \tag{2}$$

where α_{X1} and α_P are preference parameters.

I model the effect of cleanup in the individual-specific component λ^i_{jt}. Research on AOCs suggests that the damages of pollution fall rapidly with distance, with the largest effects within a few blocks of polluted shoreline (Braden et al., 2008; McMillen, 2017). The effect of cleanup on utility is thus likely to be concentrated on households that live closest to the water quality improvement. Let d_j measure the distance from location j to the nearest point on the AOC. Furthermore, let $TREAT_j$ indicate the locations whose nearest point experienced a water quality improvement, and $POST_t$ indicate the time after the improvement. Then I model the effect of cleanup in Milwaukee as

$$\lambda^i_{jt} = \alpha^i_{AOC}\frac{1}{d_{jt}} + \alpha^i_T \frac{TREAT_j}{d_{jt}} + \alpha^i_{TP}\frac{TREAT_j}{d_{jt}} \times POST_t + \alpha^i_{X_2} X_{2jt} \tag{3}$$

where X_{2jt} includes any other location attributes that households may have individual-specific tastes for. The term $1/d_{jt}$ is a gravity index that measures the proximity of a location to the AOC. Distance-based measures of pollution exposure that place more weight on near than far locations are widely used in hedonic property value research (e.g. Banzhaf et al. (2019); Cameron (2006); Isely et al. (2018); Kiel and Williams (2007)). The gravity index used here assumes that the proximity effect increases at an increasing rate as one moves toward the AOC. The coefficient α^i_{AOC} measures the desirability of living near the AOC. Whether the coefficient is positive or negative will depend on the precise mix of impaired and unimpaired beneficial uses in the AOC, although prior research implies that the sign is negative—i.e. moving away from the AOC increases utility. The coefficient α^i_T measures any difference in desirability between living near the treated area and the rest of the AOC, before cleanup. The coefficient of interest, α^i_{TP}, measures the effect of cleanup on the desirability of living in the treated area. If households prefer locations near better water quality, then $\alpha^i_{TP} > 0$.

I define $TREAT_j$ as the upper estuary of the Milwaukee River running downstream from the remediated sediment area. The upper estuary includes reaches 1-3 in the AOC's Remedial Action Plan, between Lincoln Park and the Humboldt Avenue Bridge, which covers a distance of five miles (Dow, 2020). I focus on the effect in the upper estuary, although restoration actions could have improved water quality in the lower estuary, for two reasons: first, there are sharp differences in aesthetics and water-related activities between the sections and, second, water quality improvements may have attenuated downstream. Between Lincoln Park and the Humboldt Avenue Bridge, the river is publicly accessible via a linear park and greenway, and there are no waterfront properties. Below the bridge, the riverfront is highly developed with private ownership, but the water is more likely to be impacted by urban runoff and sewer overflow. Hedonic research on AOCs has found the effect of pollution can vary between parts of an AOC (Braden et al., 2008). I explore alternative definitions of the treated area, including portions of the lower estuary, in an appendix.

I define $POST_t$ as 2017, a full year after completion of the last cleanup action in the upper Milwaukee River. To provide a clean pre-post comparison in estimation, I restrict t to the two years that bookend the cleanup, rather than including every year between 2011 and 2017. The extent that household behaviors began to shift in 2011 or earlier, in anticipation of cleanup, is difficult to evaluate because of limited sales data before this time, though Braden et al. (2008) finds that households continued to discount proximity to the Sheboygan River AOC even after remediation plans were announced. An increase in demand for homes in anticipation of cleanup would attenuate the coefficient of interest, providing an interpretation that potentially understates the importance of water quality in location decisions. Nevertheless, if the housing market takes several years to adjust to a new equilibrium following "disturbance" by the water quality improvement, then the empirical strategy used here should provide at least a partial estimate of the total effect.

Initially, I assume that household preferences are homogenous by fixing $\alpha_\ell^i = \alpha_\ell$

for $\ell = AOC, T, TP, X_2$. This simplification is important for two reasons. First, as described below, I can estimate equations (2) and (3) assuming $\alpha_\ell^i = \alpha_\ell$ using arms-length sales records alone, without relying on a match to buyer characteristics; this match is imperfect and thus a potential source of measurement error. Second, it makes for a cleaner comparison between the results in this paper and research that uses the hedonic property value model and assumes preference homogeneity.

In addition to the baseline specification above, I estimate versions of the model that include constants for Public Use Microdata Areas (PUMAs), which are Census geographies that nest tracts, as well as PUMA-year effects (i.e. PUMA-by-POST effects) to control for unobservable location attributes. I also perform a placebo test of the identification strategy by replacing $TREAT_j$ with a placebo indicator for locations whose closest point to the AOC is in the upper Menominee River. The Menominee River passes through Milwaukee and is part of the AOC but did not experience any cleanup actions between 2011 and 2017. There is no reason for household preferences for this area to change during the study period, so a placebo effect different from zero would suggest bias in the research design.[4] I also estimate the specifications described above using census block groups rather than tracts as residential locations. Finding the same or similar coefficients when location demand is modeled at a finer level of spatial precision should ease concerns that the estimates are biased due to the ecological fallacy (Banzhaf et al., 2019).

Next, I estimate the model with location choices based on Home Mortgage Disclosure Act (HMDA) records rather than arms-length sales. HMDA records provide a source of buyer characteristics that I can use to account for heterogeneous preferences. Like arms-length sales, mortgage-backed purchases make up a large share of real estate transactions, so one would expect the results to be similar if I based residential locations on either

[4] To be clear, in a residential sorting context, a placebo effect like this could be real for a sufficiently large shift in the actual effect, although in the opposite direction. That is, if households re-sorted toward the Milwaukee River because they valued cleanup, then demand for homes would be lower in other locations like near the Menominee River, and the placebo effect would appear to be negative.

arms-length sales or mortgage-backed purchases.

3.2 Preference heterogeneity

Finally, I estimate the model allowing households to have heterogeneous preferences based on their race and income. HMDA records include buyer demographics, so I can use the mortgage-backed purchase data to estimate equation (3) with the coefficients a function of race and income characteristics z_k^i:

$$\alpha_\ell^i = \alpha_{0,\ell} + \sum \alpha_{k,\ell} z_k^i. \tag{4}$$

where $\ell = AOC, T, TP, X_2$. The parameter $\alpha_{0,\ell}$ measures the effect of a location attribute on utility in the base group and $\alpha_{k,\ell} z_k^i$ measures the additional effect on utility associated with household characteristic z_k^i. The base group includes white, Asian and Native American persons of average income. Characteristics z_k^i include indicators for whether the buyer identifies as Black or Hispanic and a variable measuring demeaned household income. If households benefit from cleanup regardless of their characteristics, then $\alpha_{0,TP}^i + \alpha_{k,TP}^i z_k^i > 0$ for any k. Willingness to pay, however, may not be positive for households who face structural constraints that make buying a home near the water challenging or impossible. This could occur because of, for example, constraints on ability to pay or discriminatory practices that steer away certain demographic groups (Christensen et al., 2020).

3.3 Estimation

I estimate the model using conditional logit regression assuming households live in location j where $V_{jt}^i > V_{ht}^i$ for all $j \neq h$ and that ϵ_{jt}^i is distributed extreme value. I measure d_{jt} as the distance in kilometers from the centroid of geocoded home sales in a location to the nearest point on the AOC, and P_{jt} as a quality-adjusted price index based on sales prices.[5] The set of attributes X_{1jt} includes tract-level median income, population density, percent high school degree, crime density and three land use variables measuring the share of developed land, owner-occupied housing and tree cover within a 1 kilometer radius. The attributes X_{2jt} include the percent of Black and percent of Hispanic residents.

I estimate the model using maximum likelihood. In contrast to other sorting models in the literature, I do not use contraction mapping or a two stage procedure to estimate and then decompose the mean utilities (i.e. δ_{jt}) to estimate the utility function coefficients. Rather, I estimate the coefficients in one step. As with other sorting models in the literature, though, I need to employ an instrumental variables strategy to control for unobserved attributes and price endogeneity.

I correct for price endogneneity using a control function, which is an extra term in the utility function that conditions out the unobserved variables causing endogeneity (Petrin and Train, 2010). The method works as follows: First, I construct a set of instruments correlated with price that do not enter utility directly following the logic of Bayer and Timmins (2007) and Bakkensen and Ma (2020). The logic is that price depends on observable and unobservable attributes in j but also the attributes of other residential locations $-j$, because demand must be affected by the availability and quality of substitutes. The attributes of distant substitute locations are a source of instruments because they will correlate with price without directly affecting the utility of living in j. I construct instruments by calculating the shares of developed land, owner-occupied housing and tree cover in Milwaukee within 1, 3, and 3+ kilometer radii around the centroid of each residential location. I allow the attributes of the nearest locations (i.e. 1 kilometer) to directly enter the utility function and use the attributes of the most distant locations (i.e. 3+ kilometers) as instruments. Second, I calculate the residuals

[5] I construct the index by regressing individual transactions on housing characteristics, including indicators for duplexes and condominiums and polynomials for age and square footage, as well tract fixed effects. I then use the tract fixed effects estimates as tract-level price indices.

μ_{jt} from the regression of price on the location attributes that directly enter the utility function and the instruments. Third, I insert the residuals into the utility function with the following flexible form $\lambda_1\mu_{jt} + \lambda_2\mu_{jt}/d_{jt}$, where λ_1 and λ_2 are parameters to be estimated. With valid instruments, the control function becomes a proxy for local unobservables, rendering the remaining variation in P_{jt} exogenous. Control functions are well suited to correcting endogeneity in discrete choice models, although standard errors need to be corrected for two-stage estimation, which I do by bootstrapping (Liu et al., 2010; Malone and Lusk, 2017; Wrenn et al., 2017).

4 Data

The primary data set comes from the Milwaukee city assessor's office and includes the address, price, property type, year built, square footage, number of stories, number of rooms, and building style for 34,523 arms-length transactions between 2002 and 2018.[6] I narrow sales to detached single family homes, duplexes and condominiums sold in 2011 and 2017 with a price per square between \$25 and \$500. To link sales to tracts, I use the U.S. Census Bureau's batch geocode tool, and I assign any sale unmatched by the tool to the nearest tract based on the coordinates of the home and the centroid of each tract. These refinements leave 5,807 sales to estimate the model.

I then calculate the distance to the AOC using the coordinates of each home and the coordinates of the Milwaukee, Menominee and Kinnickinnick Rivers as well as the Lake Michigan shoreline. The average distance is 3.47 kilometers, with a range of 0.07 to 12.98 kilometers. One kilometer is about six residential blocks in Milwaukee. There are 623 sales within 1 kilometer of the upper Milwaukee River. More properties were sold in this area after cleanup, from 11.5% of all sales in 2011 to 12.5% of all sales in 2017. After calculating the gravity index and the treatment variables, I match each sale to

[6]This data can be accessed at `https://data.milwaukee.gov/dataset/property-sales-data`. For less than 1% of sales I imputed missing age and square footage using the average of the observed values.

location attribute data based on tracts, and then construct the choice set by collapsing the sales data into a database of tract and year averages.

The tract attributes include median income, population density, percent Black residents, percent Hispanic residents, percent high school graduates and percent owner occupied housing from the American Community Survey (ACS). I use the ACS five-year samples in the years preceding the transactions made in 2011 and 2017, respectively. The attributes also include crime data from the Wisconsin Incident Based Reporting System (WIBRS), which is Wisconsin's version of the FBI's National Incident-Based Reporting System. Crime incidents in Milwaukee's WIBRS include arson, assault, burglary, criminal damage, homicide, robbery, sexual offense and theft. To measure crime, I calculate the average number of annual incidents in each tract in 2006-2010 and 2012-2016, divided by tract area, which I assign to transactions made in 2011 and 2017, respectively. The assumption here is that sales are based on the demographics and crimes that households' could have known about in the years leading up to the move. I measure the amount of developed land and tree cover in each tract using data from the EPA's EnviroAtlas Land Cover Summary (LCSum) file.

Table 1 presents statistics from the sale microdata matched to the tract attributes, with Panel A showing homes within 1 kilometer of the AOC whose closest point lies on the upper Milwaukee River (i.e. where water quality improved), and Panel B showing all other homes. Variables denominated in dollars area adjusted to 2017 dollars using the Consumer Price Index for urban consumers in the Milwaukee-Racine metropolitan area. Panel A shows that the average price was \$205,000 in 2011 and \$209,000 in 2017 for homes nearest the upper Milwaukee River, after adjusting for inflation. Panel B shows that the average price was \$155,000 in both 2011 and 2017 for homes in the rest of Milwaukee. So between 2011 and 2017, average homes prices near where water quality improved increased modestly while prices in the rest of the city changed very little.

The third data set comes from Home Mortgage Disclosure Act (HMDA) records published by the Consumer Financial Protection Bureau. This mortgage-backed buyer data includes buyers' income, race, loan amount and county, and the census tract of their home. In the heterogeneous sorting model, I use buyer data from mortgages for first lien, owner-occupied, 1-4 family homes purchased in 2012 and 2017. The first year is 2012 rather than 2011 because prior to 2012 mortgage companies reported home locations using Census 2000 tracts rather than Census 2010 tracts. I therefore use sales from 2012 to describe the tracts buyers in 2012 could have bought their home in. This could create a downward bias in the cleanup effect because buyers could have begun pricing in benefits by 2012, coinciding with the end of Phase 1 cleanup. However, using the arms-length sales data I show in an appendix that willingness to pay estimates are not systematically lower when the base year is changed from 2011 to 2012.

The mortgage-backed sales provide a sample of 6,172 buyers. Average income is $69,790, with a demographic breakdown of 16% Black buyers (considerably less than the percent of black residents in Milwaukee as whole) and 14% Hispanic buyers (just below the percent of Hispanic residents in Milwaukee). Because white buyers make up most of the remaining racial and ethnic breakdown, with a small percentage of Asian and Native American buyers, I refer to the group of non-Black and non-Hispanic buyers as "predominantly white."[7] Home locations in the mortgage-backed sales data do not perfectly match the locations in the arms-length sales data because the former includes sales between related parties, while the latter includes cash sales. Nevertheless, tract-level sales counts in the two databases show a high level of correspondence, with a correlation coefficient of 0.91.

[7] Approximately 6% of buyers are Asian and less than 1% are Native American.

5 Results and discussion

5.1 Baseline model and sensitivity analyses

Table 2 presents the key coefficients in the baseline model, estimated using location choices based on arms-length sales in 2011 and 2017. The positive and significant coefficients in the control function confirm the role of omitted location attributes correlated with price.[8] Next, note the positive and significant coefficient on the gravity index (90% CI 0.079 to 0.155). This indicates that households actually prefer locations near the AOC. In contrast to prior research on AOCs, in Milwaukee the amenities associated with living near the water appear valuable enough that households place a premium on rather than discounting proximity to the water, despite the large number of BUIs. Furthermore, the positive and significant coefficient on treated area ($TREAT \times 1/d$; 90% CI 0.026-0.155) shows that households place a large premium on living near the upper estuary portion of the Milwaukee River. The water quality improvement that is the focus here therefore appears to have occurred near locations that were already relatively desirable. Turning to the cleanup effect, the coefficient is positive with a confidence interval that just overlaps zero (90% CI -0.006 to 0.124). We can convert this coefficient into a measure of average willingness to pay for cleanup by multiplying it by the average gravity index in the treated area, which is 0.67, and then dividing by the price coefficient; this indicates an average willingness to pay of $4,863 for cleanup.

Now consider the regressions in columns (2) and (3), which employ different varieties of fixed effects to control for unobservable measures of quality in different parts of the city. Column (2) adds five constants for Milwaukee's PUMAs, which closely align with the city's center, north, south, east and west sides, and column (3) adds ten PUMA-

[8]The instruments provide significant explanatory power in the first stage regression, with a joint test F-statistic of 9.40. Without the control function, the price coefficient is essentially zero (0.0009), so the control function adjusts the price coefficient in the expected direction. Results without the control function available upon request.

year effects. Overall, the coefficients in columns (1) and (2) are similar, suggesting little influence of correlated unobservables at the PUMA level. The coefficient of interest is just significantly different from zero (90% CI 0.0003 to 0.133) in column (2), with an implied willingness to pay for cleanup of \$6,549. However, the coefficient increases in magnitude and significance (90% CI 0.123-0.282) when the regression includes PUMA-year effects, with an implied willingness to pay of \$19,031. We can clearly reject the null of no time-varying unobservables at the PUMA level, so this regression should increase confidence that the water quality improvement has a real effect on household surplus.[9]

The regression in column (4) replaces the treated area indicator with a placebo indicator for locations nearest to the upper Menominee River. The coefficient on the placebo area gravity index is significantly negative (90% CI -0.250 to -0.045), which shows that living near this section of the AOC is less desirable than other sections. However, the coefficient measuring the placebo cleanup effect itself is close to zero (90% CI -0.121 to 0.121), which suggests very little change in the desirability of this part of the city during the study period. For an alternative placebo test area, see the appendix.

Overall, the three specifications in Table 2 provide evidence that households value the water quality improvement in the Milwaukee River. The coefficient of interest is not significantly different from zero in the most basic regression but increases in size and passes conventional significance levels when the analysis controls for richer sources of unobservables. Average willingness to pay for the water quality improvement ranges from \$4,863 to \$19,031, depending on the regression, or approximately 4% to 15% of the average sale price. These estimates lie firmly in the range reported in prior research, which run from Phaneuf et al. (2013), who estimate that households living within one kilometer have willingness to pay of \$4,339-\$10,831 or 4-11% of mean sale price to clean up the Buffalo River AOC, to Patunru et al. (2007), who estimate willingness to pay of \$54,807 or 24% of average home value to clean up the Waukegan Harbor AOC. These

[9]The χ^2 statistic in a likelihood ratio test is 263, compared with critical value of about 18 at the 0.05 significance level when there are 10 degrees of freedom.

estimates therefore affirm the benefit estimates in prior research on AOCs. An important caveat, however, is that the cleanup action I focus on produced a partial water quality improvement and that further remediation needs to be done in the Milwaukee Estuary AOC.

Now consider the results in Table 3, which presents the coefficients when the model uses block groups as residential locations. In the first specification, the coefficient of interest is positive and significant (90% CI 0.091-0.191).[10] The coefficient is also significantly positive in the other two specifications, and insignificant in the placebo test, which provides strong evidence that households value the water quality improvement in the Milwaukee River. Converting the coefficients into estimates of willingness to pay, the first three specifications in Table 3 indicate that cleanup is worth \$9,290, \$11,902 and \$19,577, respectively, to households in the treated area.

Table 4 presents the coefficients when I base location choices on mortgage-backed purchases in 2012 and 2017 (for results based on arms-length sales in 2012 and 2017, see the appendix). In columns (1) and (2), the coefficient of interest is not significantly different from zero but yields an average willingness to pay of \$4,467 and \$9,187, respectively, which are in-line with the estimates in Table 2. In column (3), the coefficient is significantly positive (90% CI 0.078 to 0.273), with an implied willingness to pay of \$30,882. Overall, results are similar if location choices are based on arms-length sales or mortgage-backed purchases.

Between two different definitions of the location alternatives and two sources of choice data, I find evidence Milwaukee households value water quality improvements. I subjected the identification strategy to a placebo test, which different versions of the model passed. Estimates indicate that the property market does not fully discount proximity to the AOC; in fact, evidence indicates that households prefer living near the water. Nevertheless, demand for homes increased further in the areas where water

[10]Switching from tracts to block groups, the instruments are stronger in the price regression used to generate the control function, with a joint test F-statistic of 18.74.

quality improved.

5.2 Preference heterogeneity

I now turn to the question of who benefits from cleanup. Table 5 presents the key coefficients in the heterogeneous sorting model, with location choices based on mortgage-backed purchases. This model interacts the gravity index and treated area terms with three buyer characteristics: income and Black and Hispanic identity. Although not shown in the table, the model also includes interactions between the shares of Black and Hispanic residents in a location and the buyer characteristics, to control for any group-specific preferences for amenities that could be correlated with local demographics. The main coefficient of interest (i.e. $\alpha_{0,TP}$ in equation 4) measures the effect of cleanup on average income, white household utility, while the coefficients on the interactions measure the additional effects of cleanup on the utility of Black, Hispanic and higher-income households. In all three specifications, the coefficients on the interactions are statistically insignificant. Put simply, the results provide no evidence of systematic differences in the cleanup effect between race and income groups. However, the interaction effects associated with the treated area itself imply that the water quality improvement occurred in an area preferred mainly by high income, predominantly white households. The coefficient on the treated area gravity index (i.e. $\alpha_{0,T}$) is significantly negative, implying that predominantly white households tend to locate near other parts of the AOC rather than the upper Milwaukee River, other things being equal. However, the coefficients on the Black, Hispanic and income interactions are significantly negative, negative and positive, respectively, which implies that Black, Hispanic and low-income households are even less likely to locate near the upper Milwaukee River. This means the benefits of cleanup, which occurred in the upper Milwaukee River, disproportionately flowed toward predominantly white, high-income households.

The potential role of heterogeneous sorting in concentrating pollution among poor and minority households can make cleanup efforts controversial. Inequalities in affordability and discrimination can lead to income and race-based differences in who moves in and who moves out as competition for housing near the cleanup area heats up. Cleanup can become controversial because of perceptions that high income, white households are the main beneficiaries. The results in this paper, however, do not provide significant evidence of income and race-based sorting after cleanup. If there was, then we would see income and racial differences in household preferences for the treated area after water quality improved. Preference parameters associated with the improvement are too imprecisely measured to conclude that cleanup contributed to re-sorting along income and race. This could be interpreted as good news for community and environmental justice advocates concerned about the effect of cleanup on displacement.

The lack of significant evidence of heterogeneous sorting on cleanup in this paper comes with several caveats, however. First, there is substantial uncertainty associated with the coefficients measuring heterogeneity in the model, so it is not necessarily the case that cleanup affects income and race groups similarly. For example, the hypothesis that predominantly white willingness to pay exceeds Black willingness to pay for cleanup by $10,000 cannot be rejected at conventional levels based on the specification with PUMA-year effects. Second, the results show that high income and predominantly white households were more likely to locate in the treated area before water quality improved than low income, Black and Hispanic households. Remediation may have had little effect on sorting patterns because the groups most likely to move toward the water were already living there. Third, the lack of significant evidence contrasts with prior research that finds heterogeneous sorting contributes to disproportionate pollution exposure among poor and minority households (Christensen et al., 2020; Depro et al., 2015; Melstrom and Mohammdi, 2021).

5.3 Additional discussion

The results provide evidence that households value water quality improvements through sediment remediation. The baseline model shows that residential demand increased where water quality improved. Measuring proximity using a gravity index shows that the housing market puts a premium on living near the AOC, likely due to the amenity benefits of water, and that this premium increased the most in locations closest to the improvement. Spatial attenuation in benefits is consistent with prior research on the extent of water quality impacts in the housing market around AOCs. McMillen (2017) estimates that cleanup of a portion of the Grand Calumet River AOC increased the sale price of properties adjacent to the river by 27%, and for properties within three blocks by 18%. Phaneuf et al. (2013) estimates damages in the 6-15% range for homes within 0.5 kilometers and 3-8% for homes within 1.5 kilometers of the Buffalo River AOC. Braden et al. (2008) estimates losses of 12-20% for adjacent properties and losses of 3-5% for properties three kilometers away from the Sheboygan River AOC. Using the baseline model with PUMA-year effects, I find that owner willingness to pay is about 29% of home values at 0.5 km and 12% of home values at 1.5 km from the Milwaukee River.

Back of the envelop calculations suggest that cleanup passes the benefit-cost test. Remediation of the Milwaukee River stretch studied here cost $43 million, not including earlier improvements in wastewater treatment and sewer overflow. About 46,000 households live nearest the part of the AOC where water quality improved. If average household willingness to pay for cleanup is $19,031, then the aggregate benefit would be $875 million. This is in line with other estimates, in particular, Braden et al. (2010) used benefit transfer to estimate the benefits of fully remediating the Milwaukee AOC to be worth $1.6 billion, adjusted for inflation.

However, it should be noted that the results are somewhat sensitive to model specification. Estimates of willingness to pay were lower in model specifications that did not include the richest variety of fixed effects. These fixed effects narrow the potential channels of omitted variables bias, but they could also amplify attenuation bias in the price coefficient, potentially inflating willingness to pay estimates. Nevertheless, even if willingness to pay for the water quality improvement was as low as, say, $4,863, based on the baseline model without fixed effects, the aggregate benefit would still exceed the cost ($224 million vs $43 million). Another important caveat is that the benefit estimates are based on single-family home sales, though nearly 60% of homes in Milwaukee are renter-occupied. Actual benefits could be lower than the estimates reported here, due to lower willingness to pay among renters as a result of the short-term nature of rental contracts, differences in income between owners and renters, and discriminatory practices in the rental market.

6 Conclusions

With water quality a major public concern and growing doubts about the welfare gains from national cleanup policies, there is a pressing need for information about the value of further cleanup efforts (Keiser et al., 2019). In this paper, I measured willingness to pay in the housing market for water quality improvements in the Milwaukee Estuary Area of Concern. Based on variation in the timing and location of remediation actions, I found that willingness to pay increased substantially to live in neighborhoods near where water quality improved.

The results were partially inconclusive on the question of who benefits from cleaning up water. Environmental justice advocates have raised concerns that cleanup actions often fail to benefit low income and minority households as a result of environmental gentrification, move-in and displacement by high income and white households. With no significant differences in preferences for cleanup across income and race groups in the model, though, I cannot say that neighborhoods became increasingly high income or white as a result of cleanup. However, the results do suggest that high income and white households benefited disproportionately because water quality improved near already affluent neighborhoods. This begs the question, would environmental gentrification have occured if fewer high income and white households lived near the water prior to remediation? Future research could answer this question by studying water cleanup actions closer to low income and minority neighborhoods.

Equity implications notwithstanding, the results can be used to inform discussions about cleaning up water pollution at highly contaminated urban sites. The benefit amounts reported in this paper are similar to previously published estimates on the value of AOC remediation, which is good news because most prior research did not use natural experiments to identify the effects of pollution and cleanup. Water quality improvements in the Milwaukee Estuary AOC appear to have delivered hundreds of millions in benefits to the housing market. As benefit-cost analysis often informs the decisions that policy makers and environmental agencies make, this outcome provides important economic justification for restoring water quality in urban settings.

References

Bakkensen, L. A. and L. Ma (2020). Sorting over flood risk and implications for policy reform. *Journal of Environmental Economics and Management 104*, 102362.

Banzhaf, H. S. and R. P. Walsh (2008). Do people vote with their feet? an empirical test of tiebout. *American Economic Review 98*(3), 843–63.

Banzhaf, S. (2012). *The Political Economy of Environmental Justice*. Stanford University Press.

Banzhaf, S., L. Ma, and C. Timmins (2019). Environmental justice: The economics of race, place, and pollution. *Journal of Economic Perspectives 33*(1), 185–208.

Bayer, P., F. Ferreira, and R. McMillan (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bayer, P. and C. Timmins (2007). Estimating equilibrium models of sorting across locations. *The Economic Journal 117*(518), 353–374.

Bockstael, N. E., W. M. Hanemann, and C. L. Kling (1987). Estimating the value of water quality improvements in a recreational demand framework. *Water Resources Research 23*(5), 951–960.

Braden, J. B., X. Feng, L. Freitas, and D. Won (2010). Meta-functional transfer of hedonic property values: Application to great lakes areas of concern. *Agricultural and Resource Economics Review 39*(1), 101–113.

Braden, J. B., A. A. Patunru, S. Chattopadhyay, and N. Mays (2004). Contaminant cleanup in the Waukegan Harbor Area of Concern: Homeowner attitudes and economic benefits. *Journal of Great Lakes Research 30*(4), 474–491.

Braden, J. B., D. Won, L. O. Taylor, N. Mays, A. Cangelosi, and A. A. Patunru (2008). Economic benefits of remediating the Sheboygan River, Wisconsin Area of Concern. *Journal of Great Lakes Research 34*(4), 649–660.

Cameron, T. A. (2006). Directional heterogeneity in distance profiles in hedonic property value models. *Journal of Environmental Economics and Management 51*(1), 26–45.

Carson, R. T. and R. C. Mitchell (1993). The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research 29*(7), 2445–2454.

Chapman University (2018). America's Top Fears 2018 Chapman University Survey of American Fears. Web. Accessed 3-11-2021.

Chattopadhyay, S., J. B. Braden, and A. Patunru (2005). Benefits of hazardous waste cleanup: New evidence from survey-and market-based property value approaches. *Contemporary Economic Policy 23*(3), 357–375.

Christensen, P., I. Sarmiento-Barbieri, and C. Timmins (2020). Housing discrimination and the toxics exposure gap in the United States: Evidence from the rental market. *Review of Economics and Statistics*, 1–37.

Collaboration for Great Lakes Water Quality (2021). About the Great Lakes Water Quality Agreement. Web. Accessed 3-10-2021.

Depro, B., C. Timmins, and M. ONeil (2015). White flight and coming to the nuisance: can residential mobility explain environmental injustice? *Journal of the Association of Environmental and Resource Economists 2*(3), 439–468.

Dow, B. (2020). Remedial Action Plan Update for the Milwaukee Estuary Area of Concern. Wisconsin Department of Natural Resources.

Dwight, R. H., L. M. Fernandez, D. B. Baker, J. C. Semenza, and B. H. Olson (2005). Estimating the economic burden from illnesses associated with recreational coastal water pollution–a case study in Orange County, California. *Journal of Environmental Management 76*(2), 95–103.

Fox 6 Milwaukee (2012). Cleanup of Milwaukee River, Lincoln Creek nearly complete. Web. Accessed 3-10-2021.

Gallup (2019). Environment. Web. Accessed 3-11-2021.

GlobeScan (2009). Water Issues Research. Report for Circle of Blue. Web. Accessed 3-11-2021.

Hartig, J. H., G. Krantzberg, and P. Alsip (2020). Thirty-five years of restoring Great Lakes Areas of Concern: Gradual progress, hopeful future. *Journal of Great Lakes Research*.

Healing Our Waters Great Lakes Coalition (2012). Huge source of PCBs removed from Milwaukee River. Web. Accessed 3-10-2021.

International Joint Commission (2021). Great Lakes Areas of Concern. Web. Accessed 3-10-2021.

Isely, P., E. S. Isely, C. Hause, and A. D. Steinman (2018). A socioeconomic analysis of habitat restoration in the Muskegon Lake Area of Concern. *Journal of Great Lakes Research 44*(2), 330–339.

Keiser, D. A., C. L. Kling, and J. S. Shapiro (2019). The low but uncertain measured benefits of us water quality policy. *Proceedings of the National Academy of Sciences 116*(12), 5262–5269.

Keiser, D. A. and J. S. Shapiro (2019). US water pollution regulation over the past half century: Burning waters to crystal springs? *Journal of Economic Perspectives 33*(4), 51–75.

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

Kuminoff, N. V., V. K. Smith, and C. Timmins (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature 51*(4), 1007–62.

Liu, X., M. E. Lovely, and J. Ondrich (2010). The location decisions of foreign investors in China: Untangling the effect of wages using a control function approach. *Review of Economics and Statistics 92*(1), 160–166.

Malone, T. and J. L. Lusk (2017). Taste trumps health and safety: Incorporating consumer perceptions into a discrete choice experiment for meat. *Journal of Agricultural and Applied Economics 49*(1), 139–157.

McConnell, K. E. and M. A. Rosado (2000). Valuing discrete improvements in drinking water quality through revealed preferences. *Water Resources Research 36*(6), 1575–1582.

McMillen, D. P. (2017). The benefits of environmental improvements in a low-income area: The Grand Calumet River dredging plan in Gary, Indiana. *Environmental Valuation: Interregional and Intraregional Perspectives*, 147.

Melstrom, R. T. and R. Mohammdi (2021). Residential mobility, brownfield remediation and environmental gentrification in Chicago. *Land Economics 97*(3).

Milwaukee Journal Sentinel (2014). Rivers Reborn — Special Report. Web. Accessed 3-29-2022.

Patunru, A. A., J. B. Braden, and S. Chattopadhyay (2007). Who cares about environmental stigmas and does it matter? a latent segmentation analysis of stated preferences for real estate. *American Journal of Agricultural Economics 89*(3), 712–726.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Phaneuf, D. J., L. O. Taylor, and J. B. Braden (2013). Combining revealed and stated preference data to estimate preferences for residential amenities: A GMM approach. *Land Economics 89*(1), 30–52.

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

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

Stoll, J. R., R. C. Bishop, and J. P. Keillor (2002). Estimating economic benefits of cleaning up contaminated sediments in Great Lakes Areas of Concern.

Tuttle, C. M. and M. D. Heintzelman (2015). A loon on every lake: A hedonic analysis of lake water quality in the Adirondacks. *Resource and Energy Economics 39*, 1–15.

Urban Milwaukee (2015). Estabrook Dams Golden Age for Swimming. Web. Accessed 3-10-2021.

Wisconsin DNR (2005). Estabrook Impoundment Sediment Remediation Pre-Design StudyProject Completion Report to USEPA. Wisconsin Department of Natural Resources PUBL-WT 826.

Wrenn, D. H., H. A. Klaiber, and D. A. Newburn (2017). Confronting price endogeneity in a duration model of residential subdivision development. *Journal of Applied Econometrics 32*(3), 661–682.

Table 1: Summary statistics of the location attributes before and after cleanup.

A. Treated locations 1km from AOC	Attributes in 2010		Attributes in 2016	
Variable	Mean	St. Dev.	Mean	St. Dev.
Price (in 1,000s of $)	204.812	45.966	208.538	44.873
Gravity index	2.389	2.066	2.618	1.962
Median income (in 1,000s of $)	41.035	12.350	46.438	11.967
Population density (1,000/sqmi)	15.035	6.302	14.824	5.527
Percent Black	13.863	16.227	15.145	15.781
Percent Hispanic	6.824	4.554	8.330	5.449
Percent high school degree	90.447	6.119	92.798	4.982
Crime density (1,000/sqmi)	7.087	2.815	4.726	1.843
Sales	156		477	

B. All other locations	Attributes in 2010		Attributes in 2016	
Variable	Mean	St. Dev.	Mean	St. Dev.
Price (in 1,000s of $)	155.372	68.519	155.358	64.585
Gravity index	0.966	1.826	0.743	1.342
Median income (in 1,000s of $)	53.099	14.991	50.871	15.731
Population density (1,000/sqmi)	7.920	5.138	7.895	4.912
Percent Black	21.785	27.241	23.729	28.546
Percent Hispanic	12.745	16.865	15.014	17.604
Percent high school degree	87.166	10.008	87.778	9.094
Crime density (1,000/sqmi)	2.886	2.530	2.153	1.975
Sales	1,358		3,816	

Table 2: Utility function coefficients from the sorting model

Variable	(1)	(2)	(3)	(4)
Price (in 1,000s of \$)	-0.0079^{**}	-0.0066^{**}	-0.0070^{**}	-0.0089^{**}
	(0.0011)	(0.0013)	(0.0013)	(0.0013)
$1/d$	0.1190^{**}	0.1392^{**}	0.1387^{**}	0.2253^{**}
	(0.0230)	(0.0234)	(0.0214)	(0.0214)
$TREAT \times 1/d$	0.0936^{**}	0.1158^{**}	0.0149	-0.1274^{**}
	(0.0394)	(0.0409)	(0.0652)	(0.0652)
$TREAT \times 1/d \times POST$	0.0576	0.0647^{*}	0.2002^{**}	-0.0154
	(0.0393)	(0.0399)	(0.0763)	(0.0763)
Residual for $Price$	0.0062^{**}	0.0047^{**}	0.0046^{**}	0.0068^{**}
	(0.0012)	(0.0013)	(0.0014)	(0.0014)
Residual for $Price \times 1/d$	0.0023^{**}	0.0021^{**}	0.0023^{**}	0.0014^{**}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Tract attributes	✓	✓	✓	✓
PUMA fixed effects		✓		
PUMA-year effects			✓	✓
Log-likelihood	-28896.049	-28784.217	-28764.497	-28792.818

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 1000 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

Table 3: Utility function coefficients when the model uses block groups as locations

Variable	(1)	(2)	(3)	(4)
Price (in 1,000s of $)	−0.0096**	−0.0069**	−0.0071**	−0.0080**
	(0.0012)	(0.0016)	(0.0016)	(0.0016)
$1/d$	0.2665**	0.2653**	0.2631**	0.2631**
	(0.0222)	(0.0224)	(0.0226)	(0.0194)
$TREAT{\times}1/d$	−0.0752**	−0.0863**	−0.1383**	−0.1192**
	(0.0271)	(0.0253)	(0.0309)	(0.0879)
$TREAT{\times}1/d \times POST$	0.1369**	0.1270**	0.2124**	0.0269
	(0.0299)	(0.0299)	(0.0393)	(0.0890)
Residual for $Price$	0.0085**	0.0056**	0.0054**	0.0067**
	(0.0013)	(0.0017)	(0.0017)	(0.0018)
Residual for $Price \times 1/d$	0.0021**	0.0020**	0.0022**	0.0018**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Tract attributes	✓	✓	✓	✓
PUMA fixed effects		✓		
PUMA-year effects			✓	✓
Log-likelihood	−28908.491	−34103.879	−34088.926	−34104.450

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 1000 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

Table 4: Utility function coefficients when the model uses location choices based on mortgage-backed purchases

Variable	(1)	(2)	(3)	(4)
Price (in 1,000s of $)	−0.0042**	−0.0033**	−0.0038**	−0.0041**
	(0.0011)	(0.0012)	(0.0012)	(0.0013)
$1/d$	0.0702**	0.0892**	0.0929**	0.0863**
	(0.0221)	(0.0230)	(0.0227)	(0.0250)
$TREAT \times 1/d$	−0.0783*	−0.0804*	−0.1674**	0.0094
	(0.0437)	(0.0452)	(0.0488)	(0.0600)
$TREAT \times 1/d \times POST$	0.0280	0.0449	0.1739**	−0.0024
	(0.0526)	(0.0535)	(0.0606)	(0.0668)
Residual for $Price$	0.0015	0.0000	0.0001	0.0006
	(0.0012)	(0.0013)	(0.0013)	(0.0013)
Residual for $Price \times 1/d$	0.0024**	0.0022**	0.0022**	0.0023**
	(0.0004)	(0.0003)	(0.0003)	(0.0004)
Tract attributes	✓	✓	✓	✓
PUMA fixed effects		✓		
PUMA-year effects			✓	✓
Log-likelihood	−30491.326	−30364.414	−30350.379	−30356.227

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 1000 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

Table 5: Heterogeneous sorting model with location choices based on mortgage-backed purchases

Variable	(1)	(2)	(3)	(4)
Price (in 1,000s of \$)	-0.0045^{**}	-0.0025^{**}	-0.0030^{**}	-0.0026^{**}
	(0.0012)	(0.0013)	(0.0013)	(0.0013)
$1/d$	0.0618^{**}	0.0720^{**}	0.0769^{**}	0.0257
	(0.0235)	(0.0244)	(0.0240)	(0.0276)
$TREAT \times 1/d$	−0.0438	−0.0733	-0.1648^{**}	−0.0268
	(0.0481)	(0.0499)	(0.0525)	(0.1094)
$\times Black$	-0.6283^{**}	-0.5696^{**}	-0.5433^{**}	−0.0974
	(0.3167)	(0.3261)	(0.3074)	(0.1293)
$\times Hispanic$	-0.9460^{**}	-0.9446^{**}	-0.9415^{**}	−0.0831
	(0.9120)	(0.9711)	(0.8851)	(0.1319)
$\times$*Income (in 1,000s of \$)*	0.0015^{**}	0.0013^{**}	0.0015^{**}	-0.0077^{**}
	(0.0004)	(0.0004)	(0.0004)	(0.0025)
$TREAT \times 1/d \times POST$	0.0535	0.0685	0.2044^{**}	0.1291
	(0.0562)	(0.0567)	(0.0632)	(0.1178)
$\times Black$	−0.0882	−0.0730	−0.1244	−0.0428
	(0.3753)	(0.3853)	(0.3778)	(0.1612)
$\times Hispanic$	0.4477	0.4599	0.4550	0.1565
	(0.9298)	(0.9877)	(0.9045)	(0.1553)
$\times$*Income (in 1,000s of \$)*	−0.0004	−0.0004	−0.0006	0.0042
	(0.0006)	(0.0006)	(0.0006)	(0.0027)
Residual for *Price*	0.0018	−0.0009	−0.0007	−0.0014
	(0.0013)	(0.0014)	(0.0014)	(0.0014)
Residual for $Price \times 1/d$	0.0023^{**}	0.0022^{**}	0.0023^{**}	0.0029^{**}
	(0.0004)	(0.0004)	(0.0003)	(0.0004)
Tract attributes	✓	✓	✓	✓
PUMA fixed effects		✓		
PUMA-year effects			✓	✓
Log-likelihood	−28995.093	−28805.633	−28787.754	−28808.472

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 1000 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

Appendix

A1. Complete set of coefficient estimates

Table A1 presents all of the coefficients in the baseline specification and the specification with PUMA-year effects. Most of the signs are consistent with expectations. Complete estimates of the other models are available upon request.

Table A1: Coefficients in the sorting model

	Baseline specification		Including PUMA-year effects	
Variable	Coefficient	St. Error	Coefficient	St. Error
Price (in 1,000s of \$)	-0.0079^{**}	0.0011	-0.0070^{**}	0.0013
$1/d$	0.1190^{**}	0.0230	0.1387^{**}	0.0214
$TREAT \times 1/d$	0.0936^{**}	0.0394	0.0149	0.0652
$TREAT \times 1/d \times POST$	0.0576	0.0393	0.2002^{**}	0.0763
Median household income	0.0055^{**}	0.0017	0.0037^{**}	0.0017
Population density	0.0464^{**}	0.0064	0.0385^{**}	0.0069
Percent Black	-0.0152^{**}	0.0012	-0.0184^{**}	0.0013
Percent Hispanic	-0.0094^{**}	0.0019	-0.0049^{**}	0.0022
Percent high school degree	0.0369^{**}	0.0037	0.0312^{**}	0.0038
Crime density	-0.0976^{**}	0.0161	-0.1037^{**}	0.0180
Percent owner-occupied housing	0.0170^{**}	0.0018	0.0116^{**}	0.0017
Percent tree cover	0.0397^{**}	0.0093	0.0166	0.0103
Percent developed	0.0297^{**}	0.0033	0.0290^{**}	0.0033
Residual for *Price*	0.0062^{**}	0.0012	0.0046^{**}	0.0014
Residual for $Price \times 1/d$	0.0023^{**}	0.0003	0.0023^{**}	0.0003
PUMA-year effects				✓

Standard errors calculated using 1000 bootstraps. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

A2. Alternative placebo test

This section presents a version of the placebo test in which I calculate a separate gravity index based on distance from downtown Milwaukee. This test is intended to address concerns that, because the Milaukee River runs through downtown Milwaukee, the treated effect is driven by downtown development rather than water quality improvements per se. Recall that the placebo test in the main text redefines the treated area to be the portion of the Menominee River above downtown. Like the Milwaukee River, the Menominee River is surrounded by a greenway, which makes it a good placebo. However, additional placebo tests can increase confidence in the results. The table below presents a placebo test in which the placebo is a gravity index based on distance from downtown (43.0375, -87.9190). The columns show the estimates using different varieties of fixed effects. The placebo effect is negative and significant in the first specification, indicating that living near downtown became less desirable between 2011 and 2017; however, the effect moves toward zero and loses significance in the specification with PUMA-YEAR effects.

Table A2: Model estimates with alternative placebo test

Variable	(1)	(2)	(3)
Price (in 1,000s of $)	-0.0081^{**}	-0.0070^{**}	-0.0077^{**}
	(0.0011)	(0.0013)	(0.0013)
$1/d_{aoc}$	0.1806^{**}	0.1928^{**}	0.1986^{**}
	(0.0203)	(0.0208)	(0.0207)
$TREAT \times 1/d_{placebo}$	-0.0101	0.1755^{*}	0.0257
	(0.0831)	(0.0899)	(0.1114)
$TREAT \times 1/d_{placebo} \times POST$	-0.1894^{*}	-0.2144^{**}	0.0268
	(0.1001)	(0.0991)	(0.1170)
Residual for $Price$	0.0068^{**}	0.0052^{**}	0.0056^{**}
	(0.0012)	(0.0013)	(0.0013)
Residual for $Price \times 1/d_{aoc}$	0.0019^{**}	0.0015^{**}	0.0016^{**}
	(0.0003)	(0.0002)	(0.0002)
Tract attributes	✓	✓	✓
PUMA fixed effects		✓	
PUMA-year effects			✓
Log-likelihood	-28912.509	-28807.787	-28798.756

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 100 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

A3. The effect of changing the base year from 2011 to 2012

Cleanup took place between 2011 and 2015, with phased completions in 2012 and 2015. Ideally we could observe home buying activity before and after these dates, but the city's arms-length sales data appears to be incomplete prior to 2011 and HMDA's mortgage-backed sales data link to Census 2010 tracts only for mortgages filed in 2012 and after. One concern is that using 2012 rather than 2011 (or 2011 rather than 2010, for that matter) as the first year could downward bias willingness to pay for cleanup, if buyers began pricing in benefits prior to 2012. To gauge the amount of this bias, I estimate the sorting model using arms-length sales in different years: Table 2 presents the results using 2011 and 2017, while Table A3 below presents the results using 2012 and 2017. Note that column (4) in Table A3 presents the results of the placebo test. Willingness to pay estimates are not lower using the 2012 and 2017 data. Across columns (1) through (3) in Table 2, the coefficients indicate willingness to pay is \$4,863, \$6,549, \$19,031, respectively, given a mean gravity index in the treated area of 0.67. In contrast, the same columns in Table A3 imply willingness to pay is \$7,761, \$12,992, and \$35,140, respectively, given a mean gravity index in the treated area of 0.55. Based on bootstrapped samples, the first two estimates are not significantly different from eachother, while the third is different at the 10% significance level. Thus, there is weak evidence that home buying may have in fact shifted away from the treated area between 2011 and 2012, which does not support the hypothesis that buyers began price in benefits prior to 2012. Estimates of willingness to pay for cleanup are higher—not lower—using 2012 rather than 2011 as the base year, although the difference is not highly significant.

This comparison indicates that differences between the arms-length sales and mortgage-backed purchase data are likely driven by the change in base year rather than underlying systematic differences in populations the samples are drawn from, i.e. lower average wealth among bank-financed versus cash buyers, who will appear in the arms-length sales data but not in mortgage records.

Table A3: Model estimates with location choices based on arms-length sales in 2012 and 2017

Variable	(1)	(2)	(3)	(4)
Price (in 1,000s of \$)	−0.0035**	−0.0029**	−0.0034**	−0.0050**
	(0.0012)	(0.0013)	(0.0013)	(0.0013)
$1/d$	0.1527**	0.1699**	0.1727**	0.1925**
	(0.0196)	(0.0203)	(0.0203)	(0.0228)
$TREAT \times 1/d$	−0.0820**	−0.0660	−0.1751**	−0.0487
	(0.0420)	(0.0428)	(0.0489)	(0.0577)
$TREAT \times 1/d \times POST$	0.0500	0.0677**	0.2199**	−0.0757
	(0.0467)	(0.0477)	(0.0558)	(0.0714)
Residual for $Price$	0.0014	0.0004	0.0005	0.0020
	(0.0013)	(0.0013)	(0.0013)	(0.0014)
Residual for $Price \times 1/d$	0.0019**	0.0017**	0.0018**	0.0017**
	(0.0003)	(0.0003)	(0.0002)	(0.0003)
Tract attributes	✓	✓	✓	✓
PUMA fixed effects		✓		
PUMA-year effects			✓	✓
Log-likelihood	−29975.385	−29864.015	−29846.985	−29851.684

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 100 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.

A4. Sensitivity analysis of the water quality improvement area

This section examines the sensitivity of the results in the main text to alternative definitions of the treated area, i.e. the area with water quality improvements. To generate these results I defined the treated area as the Milwaukee River running from the remediation site to point that the upper and lower estuary meet, which is a distance of five miles. The table below presents estimates when I extend or shorten this distance based by one or two miles, using the specification with PUMA-year effects. Extending the distance brings more locations into the treated area, and shortening the distance reduces the treated area. The top of each column reports the distance. 7 miles essentially includes the upper and lower estuary of the Milwaukee River, which adds a large number of homes with direct water access. In contrast, 3 miles contains very little housing near the river. Consequently, the mean gravity index in the treated area falls from 0.85 to 0.33 moving from 7 miles to 3 miles. The coefficients in Table A4 indicate that willingness to pay is \$12,478 when the treated area runs a distance of seven miles, \$16,971 at six miles, \$5,073 at four miles, and \$2,410 at 3 miles. The estimates at distances that include the lower estuary are lower but comparable to the benchmark. Shortening the distance results in significantly lower and highly imprecise estimates likely due to a combination of limited area, fewer homes located near the river (there are only three tracts within one kilometer of the treated area at 3 miles, so much of the identification using a smaller treated area will come from homes much farther away), and confounding of treated and comparison groups.

Table A4: Model estimates with alternative definitions of the treated area

Variable	(1)	(2)	(3)	(4)
	7 miles	6 miles	4 miles	3 miles
Price (in 1,000s of $)	−0.0029**	−0.0067**	−0.0063**	−0.0066**
	(0.0014)	(0.0013)	(0.0013)	(0.0013)
$1/d$	−0.1821**	0.1226**	0.1916**	0.1897**
	(0.0397)	(0.0229)	(0.0189)	(0.0190)
$TREAT{\times}1/d$	0.5184**	0.1335**	−0.1314**	−0.1052
	(0.0503)	(0.0359)	(0.0636)	(0.1296)
$TREAT{\times}1/d \times POST$	0.0429	0.1479**	0.0658**	0.0492
	(0.0361)	(0.0392)	(0.0714)	(0.1361)
Residual for $Price$	0.0041**	0.0039**	0.0045**	0.0046**
	(0.0014)	(0.0013)	(0.0013)	(0.0013)
Residual for $Price \times 1/d$	0.0005**	0.0027**	0.0015**	0.0015**
	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Tract attributes	✓	✓	✓	✓
PUMA-year effects	✓	✓	✓	✓
Log-likelihood	−28692.008	−28761.571	−28804.976	−28804.977

Tract attributes not reported include median household income, population density, share Black, share Hispanic, percent high school degree, crime rate, percent owner-occupied housing, percent tree cover and percent developed. Standard errors calculated using 100 bootstraps in parentheses below coefficients. $**$ and $*$ indicate significance at the 0.05 and 0.10 levels, respectively, based on the share of bootstrapped estimates that cross zero.