Bridging the racial disparity in wealth creation in Milwaukee

Kishor, N. Kundan and Konkel, Rebecca and Yoon, Jangsu and Zhao, Tian

University of Wisconsin-Milwaukee

June 2023

Bridging the Racial Disparity in Wealth Creation in Milwaukee

N. Kundan Kishor, Professor, Department of Economics

Rebecca H. Konkel, Assistant Professor, Department of Criminal Justice & Criminology

Jangsu Yoon, Assistant Professor, Department of Economics

Tian Zhao, Associate Professor, Department of Computer Science

Executive Summary

Milwaukee's segregation and racial disparities in homeownership and mortgage application denial rates have been chronicled in numerous research papers and news articles through the years. Although the impediments to fair housing have been documented[1] and many initiatives through the years have attempted to address these disparities, few strategies have moved the needle in a positive and sustained direction.

Notably, homeownership rates for Black and Hispanic households in Milwaukee continue to lag the homeownership rate among White households. In the City of Milwaukee, Black and Hispanic homeownership rates (27% and 37%, respectively) are dramatically lower than the White/Not Hispanic homeownership rate (56%).[2] The lack of homeownership in communities of color is a significant contributor to the large wealth gap for Black and Hispanic households relative to White households.[3] As such, it is no surprise that these persistent disparities have given rise to a new collective action to address racial equity in homeownership in Milwaukee. This effort, which is codified in the City of Milwaukee Collective Affordable Housing Plan (housingplan.org), represents a bright, new opportunity involving many key stakeholders in Milwaukee, who span the public, private, non-profit, policy, and academic sectors. Rooted in analyses, which were based on quantitative and qualitative data collected from approximately 80 stakeholders across a one-year period, the Plan represents the most promising effort to date to address disparities in Milwaukee homeownership rates.

This executive summary has three goals. First, the findings presented herein are targeted to be a useful tool to inform public policy, as well as social and commercial investment solutions that align with and advance the City of Milwaukee's Collective Affordable Housing Plan. Second, the quantitative methodologies and data science techniques deployed in this research effort aim to provide a foundational approach that can further be built upon in Milwaukee. Here, there is the potential for the inclusion of additional data sets to train models used in order to identify other factors that may contribute to disparities. Further, the methodological approaches and models constructed may be salient to other researchers who would like to adopt similar methodological techniques in other cities, perform similar evaluation, or for benchmarking. Third, the models used for this evaluation, along with the findings, provide a durable metric that investors can leverage to establish social impact goals, and furthermore, to access their achievement of such goals.

A Wealth Creation Index

The evaluation stemming from this project resulted in the development of an econometric machine learning model that produces output we refer to as the Wealth Creation Index (hereafter, the "Index"). As with any index, it is the sum of its parts, or components. Each potential and measurable contributary to wealth disparity is a component of the Index, which when aggregated to a defined geographic area (e.g., a collection of Census tracts), enables us to compare the wealth creativeness of one geography relative to another. The geography was chosen to allow for cohort comparisons, which resulted in the Index focusing on areas where place-based strategies may be most salient to needs.

The risk in this approach is that it may validate that less wealth is created in geographies predominated by households of color and, in doing so, reinforce long-held prejudices which have manifested in redlining practices.[4][5] The benefits, however, far outweigh this risk. While the Index may highlight racial/ethnic disparities, it also quantifies them and, in doing so, provides an ability to apply a consistent measurement over time. This will be critical to policymakers who will want to benchmark and measure the impact of their policies. For investors focused on the triple-bottom line[6], it provides a way to quantify the social impact of their investments. Not only will it be important to measure Index values over time, but just as important to use Index values to evaluate if disparities are diminishing or, perhaps, getting worse.

Based on the models used in this evaluation, it was found that, indeed, racial/ethnic disparities in wealth creation are pronounced within Milwaukee, net other key household and property factors linked with homeownership. Further, it was identified that additional environmental factors, such as crime and proximity to quality high schools, are capable of amplifying or mitigating these disparities in wealth creation. Moreover, results indicated that there is a correlation between wealth creation and a neighborhood's homeownership rate, with a higher rate of ownership being associated with higher wealth accretion. Beyond these findings, and perhaps most notable, is the differential impact of foreclosures on home values, in which foreclosures have a greater negative impact on the values of homes owned by individuals of color, compared to the values of homes owned by White individuals.

Estimating Housing Returns and Disparities

In this evaluation, housing returns were employed as a proxy measure for wealth creation through homeownership versus renting a home. Housing return is defined as the gain/loss in a property's value from the point of purchase to disposition. Disposition could occur for several reasons. For instance, a homeowner could sell a property, or a homeowner could lose their home as the result of a foreclosure or an alternative transfer of title (e.g., a deed-in-lieu of foreclosure). In this context, housing returns are based on the difference in cost/value of a property between the start date (date of purchase) and an end date (date of sale, foreclosure, or transfer of title), in which these values are available through public records, such as a mortgage or deed of trust.

In the evaluation report, the quantitative methodologies used for data collection and analyses are described in detail and include information on each dataset used, matching techniques deployed, and how these datasets were appended to provide a more complete picture of the environment, properties, households, mortgage activities, and several outcomes within the 20-year study period. In terms of the methods used for model building, this project estimates an econometric machine learning model, which focused on exploring disparities in housing returns and factors that may aid in bridging the gap in housing returns across different populations. The estimation strategy used involves three steps. First, using machine learning algorithms, data from multiple sources that contain socioeconomic characteristics and individual housing transactions from the Milwaukee metropolitan area were merged with one another to create one, complete dataset. Second, using this merged dataset, we estimate the factors that explain wealth creation among households and the extent of racial/ethnic gap in housing returns. Third, in a more nuanced approach, we identify the key factors associated with disparities in housing returns.

Key Findings

Based on the methods and analyses described above, several novel and important findings on disparities in wealth creation through homeownership were identified. Although each of these findings are discussed in detail in the complete report, the most noteworthy findings are highlighted below.

- In Milwaukee, homeownership is associated with wealth creation. Specifically, wealth accumulation is 13% higher among those who own their homes, compared to among those who rent their place of residence.
- Home values appreciate at a lower rate for minority homeowners, in which, on average, homes owned by Black and Hispanic homeowners have lower appreciation values compared to homes owned by White homeowners (-6.8% and -3.0%), respectively, conditional on other characteristics.
- On average, female homeowners experience a 1% lower appreciation in home value, compared to male homeowners, conditional on other characteristics.
- On average, foreclosed homes yield lower housing returns. Moreover, these returns are even lower when the home was owned by a minority or female homeowner. This may be especially detrimental to individuals from these populations, as Black and Hispanic homeowners are more likely to experience a foreclosure than White homeowners, and similarly, female homeowners are more likely to experience a foreclosure than male homeowners. Here, we note compounding of effects, as individuals from these populations are subjected to receiving lower home value appreciation in general, are more likely to have a home foreclosed on, and in cases of foreclosure, receive an even lower valuation.
- Foreclosure has a non-linear effect. When the percentage of foreclosed houses within a given area is below the 95th percentile, foreclosure reduces home value by 1.2%. On the other hand, if a house has been foreclosed on in a neighborhood with foreclosure rates at or above the 95th percentile, foreclosure reduces the home value by 9.6%. Stated differently, in neighborhoods with extremely high foreclosure rates (≥95%), on average, the impact of devaluation on a given home is 8 times greater than those with lower foreclosure rates (≤95%). These have the ability to lead to the lowest housing returns in neighborhoods that experience the highest levels of foreclosures.
- The ownership ratio in a neighborhood affects home values. Here, the cut-off threshold for the percentage of owner-occupied housing units is 30%. Specifically, when the share of owner-occupied housing units within a given neighborhood is equal to or exceeds 30%, there is a positive impact on home values. Conversely, when less than 30% of housing units are owner-occupied (i.e., renter-occupied), home values are lower.
- Household income is a key determinant of home value, in which each 1% increase in income is associated with a 0.1% increase in home values. Stated differently, a home's value is approximately 1% higher for a homeowner who earns 10% more than a comparable homeowner.
- Education plays a significant role in determining home values. There is a clear dichotomy in home values based on educational attainment of neighborhood residents. Here, home values are lower in neighborhoods comprised of more residents who have lower levels of educational attainment (i.e., not more than a high school diploma), whereas home values are higher among houses in neighborhoods with greater percentages of

residents with higher levels of educational attainment (i.e., a bachelor's degree or higher).
- Proximity to a higher-quality high school is a key determinant of home values. The distance to a high school exceeding expectations in terms of quality of education is important. Home value decreases as the distance between a given house and a high-performing high school increased.
- Crime negatively affects home values. Based on a measure that captured several types of reported criminal offenses, it was found that houses located in neighborhoods with more crime are valued lower than those with less crime.
- Lot size is also an important determinant of home values. A 1% increase in lot size correlates with around a 1.7% increase in home values.

In the subsequent sections of this report, we provide an overview of the extant literature on homeownership and home value disparities. Next, we detail the methodologies employed to obtain, clean, and model the data used for analyses. The following section describes each of the key findings that emerged based on analyses. Lastly, rooted in the results from analyses, we conclude with a discussion of potential implications for policy development and growth.

1. Introduction

Housing wealth constitutes the largest portion of households' balance sheets worldwide. Appreciation in housing values serves as a valuable hedge against inflation and is a primary source of wealth accumulation for many families. This is especially true for low- to middle-income households, for whom housing accounts for a larger fraction of their income, leaving few opportunities for alternative investments. Despite significant progress in homeownership opportunities for historically disadvantaged minorities, racial wealth disparities in the U.S. remain large and persistent. According to Kuhn et al. (2020), the median wealth of Black households has rarely exceeded one-tenth of the median wealth of White households since 1949.

The challenges faced by racial minorities in wealth accumulation are particularly magnified in Milwaukee, which is the most racially and economically segregated metropolitan area in the United States (Massey & Tannen, 2015). Therefore, this study aims to examine the factors that contribute to wealth creation among households, with a focus on housing returns. To accomplish this, we utilized data from multiple sources, including individual housing transactions within the Milwaukee metropolitan area's housing markets. Detailed housing transaction data for Milwaukee County was obtained from Black Knight; however, this dataset does not include socio-economic features of homeowners. To address this, we obtained data from the Home Mortgage Disclosure Act (HMDA) to include these variables in our models. Additionally, we extracted assessed home price data from Milwaukee County's Master Property File (MPROP) to track the evolution of house prices over time. To further support our analysis, we collected detailed data on school quality, crime, and amenities. We employed a machine learning approach to merge data from all these sources. The merged dataset enables us to track wealth creation among households over time.

In theory, observed wealth creation can be influenced by various socioeconomic factors. One of the key advantages of our approach is that it enables us to estimate racial disparities in housing returns and identify the critical factors that contribute to these disparities. We utilize an econometric/machine learning approach for variable selection to quantify the roles of various factors in the evolution of wealth over time. By analyzing household-level data, we also investigate factors that may impact racial disparities in wealth creation in the Milwaukee metropolitan area.

We find that household income and lot size are positively associated with home values. Using a comprehensive measure of crime, our study suggests that houses located in neighborhoods with higher crime rates tend to experience slower appreciation in value on average. This effect is particularly pronounced in ZIP Codes with majority-minority populations that also have significantly higher crime rates. This impact is present in our model that also controls for other socioeconomic variables, including income and education.

In addition, education and school quality are significant factors in home price appreciation. Homes in neighborhoods with a higher percentage of residents with lower levels of education (e.g., high school degree or lower) tend to appreciate at a slower rate. Conversely, in neighborhoods with a higher percentage of residents with higher levels of education (e.g., bachelor's degree or higher), home value appreciation is greater. Furthermore, distance to a higher-quality high school is a critical factor associated with home values. The proximity of a high-performing school is important, as home values decrease as the distance between a given house and a high-performing school increases. Overall, our findings suggest that home values are influenced by a range of socioeconomic factors, including income, lot size, crime rates, education levels, and school quality.

We also find a significant role of homeownership in house price changes in our analysis. Here, houses have higher values when they are located within neighborhoods where 30% or more of the housing units are owner-occupied. On the other hand, houses have lower values when they are situated in neighborhoods where less than 30% of housing units are owner-occupied (i.e., renter-occupied). Consistent with findings in the extant literature, we find that foreclosure has a significant and negative impact on house wealth creation. Specifically, foreclosed properties on average appreciate at a lower percentage than corresponding non-foreclosed properties. In addition, foreclosure has a non-linear effect on property values. When the percentage of foreclosed houses is below the 95th percentile in a ZIP Code, foreclosure has a negative impact on home value by 1.2%. On the other hand, neighborhoods with foreclosure rates above the 95th percentile experience a decline of 9.57% in home value if a house has been foreclosed on.

A significant part of our study examines housing return disparities in Milwaukee County across different racial and gender groups. We found evidence of a racial and gender gap in realized housing returns in Milwaukee County. Specifically, we found that when the percentage of foreclosed houses within a neighborhood is below the 95th percentile, the value of an individual home that has been foreclosed on declines by 1.2%. However, if a house has been foreclosed on in a neighborhood with foreclosure rates above the 95th percentile, these houses experience a more significant decline in home value of 9.6%.

Next, we find that, on average, Black homeowners in Milwaukee County experience a 6.8% lower price appreciation on their homes, compared to those of the baseline White homeowners. Controlling amenities like school quality and crime reduces this disparity to 4.6 percentage points. We also find a disproportionate impact of foreclosure on Black homeowners. This is particularly pronounced in areas with high foreclosure rates. In neighborhoods with foreclosure rates above the 95th percentile, we find that Black homeowners on average observe home value returns that are 16% lower than returns among comparable White homeowners.

Literature Review

Although housing is the single largest asset class held by middle-class households (Campbell, 2006), with returns to housing often exceeding those of alternative investments (Jordà et al., 2019), the wealth held by middle-class racial minorities has remained alarmingly low. Homeownership is a key contributor to racial wealth disparities, by both the proportion of each racial group who owns and the values of homes among each group (Charles & Hurst, 2002; Squires & Kubrin, 2006). The literature has highlighted multiple reasons why this is the case. One strand of literature has focused on the role of distressed sales. The underlying argument is that there is a significant disparity in the number of houses sold via distressed sales, which therefore may lead to a racial disparity in house price appreciation.[1] Kermani and Wong (2021) contribute to this literature by showing that neighborhood-level differences in house price appreciation (and thus ignoring distressed sales) greatly underestimates the differences in realized housing returns by race. They show that Black and Hispanic homeowners are not only more likely to experience a distressed sale, but moreover, are also more likely to live in neighborhoods where distressed sales erase greater proportions of house values. Importantly, they highlight that in absence of financial distress, houses owned by minorities do not appreciate at slower rates than houses owned by non-minorities.

The literature on financial distress is also related to findings on the racial disparity in income stability (Hardy et al., 2018; Wrigley-Field & Seltzer, 2020). This literature underscores the importance of income volatility for mortgage default and wealth accumulation, which aligns

[1]See Anacker (2010), Faber and Ellen (2016), Flippen (2004), and Kahn (2021), among others.

with recent evidence that has demonstrated that a large share of mortgage defaults can be attributed to income shocks and other adverse life events (Ganong & Noel, 2020). In addition to income volatility, there is a significant disparity in liquid wealth holdings, which can play a key role in financial distress. According to a recent estimate, the median liquid wealth for Hispanic homeowners is more than twice of that of Black homeowners ($5,400 and $2,400, respectively) (Kermani & Wong, 2021). Kermani and Wong (2021) focus on income instability and illiquidity as key factors that drive racial differences in mortgage default. This relates to prior research examining racial differences in income shocks (Ganong & Noel, 2020; Ritter & Taylor, 2011; Wrigley-Field & Seltzer, 2020). In this line of inquiry, findings show that these factors are associated with disparities in housing wealth accumulation, which connects with the larger literature on the racial wealth gap (Altonji & Doraszelski, 2005; Barsky et al., 2002; Blau & Graham, 1990; Gittleman & Wolff, 2004; Hamilton & Darity Jr., 2010; Kuhn & Ploj, 2020).

Another strand of literature has examined racial disparity in housing markets. Discriminatory practices, as outlined in Akbar et al. (2019), have shifted from overt discrimination (e.g., denial of home loan applications, explicit racism) to more covert means of discrimination (e.g., discriminatory financing). For example, Bayer et al. (2017), Ihlanfeldt and Mayock (2009), and Myers (2004) found that minority homeowners pay more for identical housing than White homeowners. Other studies have found that minorities pay higher housing costs through unfavorable tax assessments (Avenancio-Leon & Howard, 2019), interest rates and fees (Ambrose et al., 2020; Bartlett et al., 2022; Bhutta & Hizmo, 2021; Fuster et al., 2020), and refinancing-related practices (Gerardi et al., 2020).

The literature also suggests that gentrification plays a role in disparities in housing returns, in which minority homeowners may buy at higher prices and sell for lower prices. Using U.S.

Census data from 1930, Akbar et al. (2019) show that Black families entering a previously White neighborhood paid a house price premium of 28%, only to see the value of their homes fall as White homeowners left the neighborhood. Further, Bayer et al. (2017) show that Black and Hispanic homebuyers pay more for houses than their White counterparts, and moreover, Perry et al. (2018) document substantial undervaluation of properties located in neighborhoods predominantly comprised of Black residents.

Below, we turn to a more nuanced review of the literature regarding the segregation of housing across varying neighborhood contexts.

Neighborhood Effects, Residential Segregation, and Milwaukee's Housing Market

Milwaukee is one of the most racially segregated metropolitan areas within the U.S. (Massey & Tannen, 2015); however, this stratification does not exist in a racial vacuum, but rather, spans several socio-structural planes. For example, racial segregation in Milwaukee is compounded by vast economic segregation. Here, the areas with the highest levels of poverty largely mirror the neighborhoods called home by higher proportions of Black and Hispanic individuals (see Figure 1). The tangling of economic and racial segregation is not unique to Milwaukee. Following the widening of income inequalities and the economic decline that took place in the early 2000s, research has indicated increases in racial segregation of Black and Hispanic residents within high-poverty neighborhoods (Logan, 2011; Reardon & Bischoff, 2011). Although findings suggest a greater level of racial heterogeneity within middle-class neighborhoods, Black middle-class individuals report being more likely and preferring to reside in neighborhoods with a greater proportion (at least 50%) of Black residents (Adelman, 2004), whereas White middle-class individuals are most likely to prefer to reside within homogenous White neighborhoods (Charles, 2006; Lewis et al., 2011; Swaroop & Krysan, 2011).

Along this same vein, White residential segregation is found to decrease as Black and Hispanic socioeconomic status (SES) increases, whereas White segregation increases as White SES increases (Sharp & Iceland, 2013). Further, research suggests that such effects are not consistent across all non-White races. For example, Sharp and Iceland (2013) surmount that the interaction between SES and race is inapplicable to Asian residents, and therefore, they assert that Black and Hispanic individuals are viewed as disadvantaged, whereas Asian and White individuals are deemed to be more advantaged. Others have identified that racial segregation patterns are influenced by additional socio-structural factors (i.e., education, family composition). When considering other indices related to socioeconomic disadvantage, residential integration of Hispanic and Asian individuals has been more successful, whereas racial neighborhood boundaries have not been crossed as often by Black individuals (Fischer, 2003; Iceland & Wilkes, 2006; Logan et al., 2004). Although findings on the level of assimilation of Hispanic residents into predominantly White neighborhoods have varied, findings indicating lower levels of integration among Black residents and higher levels of integration among Asian residents have remained consistent.

Race and economic standing not only impact where different racial groups reside, but also are associated with the likelihood of owning a house. Congruent with the literature highlighting that White middle- to upper-class individuals are more likely to own homes than are those of other races (Ruggles et al., 2010), across Milwaukee, Black and Hispanic individuals are more likely to reside in areas with lower percentages of owner-occupied homes (see Figure 2). When assessing the level of desirability for homeownership, potential buyers take many elements of the neighborhood into account. Proximal access to private/public institutions (e.g., schools), institutions (e.g., organizations, services), and other amenities (e.g., parks) are deemed to be a critical factor accounted for by potential buyers (Charles & Hurst, 2002; Friedman & Rosenbaum, 2004; Shapiro, 2004; Swaroop & Krysan, 2011). Perhaps amplifying the intersection of SES and residential segregation by race, access to such institutions and amenities are often unevenly dispersed across neighborhoods (Alba & Logan, 1993; Sampson & Wilson, 1995; Wilson, 2012). In this light, Wilson (2012) highlights that diminished institutional bases (e.g., services, institutions, amenities) are common across poor and minority neighborhoods, which has the potential to lead to social isolation. Further, this may decrease the attractiveness of minority and lower-income neighborhoods among potential homebuyers.

Another component often considered in homebuying is the safety of the larger neighborhood, in which neighborhoods with lower crime rates are seen as more desirable (Pynoos & Nishita, 2006). The literature identifies White homebuyers often stereotype neighborhoods comprised predominantly of Black and Hispanic residents as having higher crime rates (Charles, 2006; Ellen, 2001). Countless studies have identified that the majority of crimes (at least 80%) occur within few locations (less than 20%) (Weisburd, 2015) (i.e., “The Pareto Principle”), and further, that crime typically concentrates within residentially unstable (measured as percent renter-occupied units and percent residents moving into the neighborhood within the past five years) and disadvantaged neighborhoods (for examples, see Bellair, 1997; Hipp, 2007; Sampson & Groves, 1989; Shaw & McKay, 1942; Warner & Pierce, 1993).

Beyond influencing where individuals reside and whether they rent or own a home, home value is impacted by the racial composition of neighborhoods (Bianchi et al., 1982). Within Milwaukee, median home value is inversely associated with non-White populations, in which areas with higher percentages of Black or Hispanic residents have lower median home values when compared to predominantly White neighborhoods (see Figure 3). The literature identifies neighborhood typecasting may be one factor that leads to disparate housing values, in which White homebuyers view houses in predominantly Black and Hispanic neighborhoods as having declining values (Charles, 2006; Ellen, 2001). Further, houses within neighborhoods with greater residential instability, disadvantaged youth, and crime/disorder typically have lower values (Hipp et al., 2009; Schwartz et al., 2003; Tita et al., 2006).

2. Empirical Methodology

2.1 Methodology

We suggest a Wealth Creation Index using predictors of wealth components. For a residence α,

$$WCI_{\alpha} = \frac{1}{N_{\alpha}} \sum_{i \in R_{\alpha}} [1\{Home_i = 1\}\widehat{HR_\iota}(X) + 1\{Home_i = 0\}\widehat{NR_\iota}(X)]$$

where N_{α} is the number of residents in area α, R_{α} is the group of residents in area α, $1\{Home_i = 1\}$ is the indicator of housing occupancy, $\widehat{HR_\iota}(X)$ is the predictor of housing value for houseowners, and $\widehat{NR_\iota}(\mathrm{X})$ is a predictor of wealth for non-houseowners. X denotes a set of covariates (e.g., income, race, credit history, time trend, housing price shifters). This index captures the aggregate wealth creation in Milwaukee County as measured by changes in house prices.

The first step to compute the Wealth Creation Index is to generate a training dataset for wealth predictors. We start by estimating the racial gap in housing returns of Milwaukee. Kermani and Wong (2021) suggested two measures of housing returns. The first (and the most intuitive) measure is r_i^u, the annual unlevered rate of return for owner i:

$$1 + r_i^u = \left(\frac{P_{i1}}{P_{i0}}\right)^{\frac{1}{T_{i1}-T_{i0}}}$$

where P_{i0} and P_{i1} are the property purchase and sale prices, and $T_{i1} - T_{i0}$ is the length of the ownership in years. Since we have yearly assessed property value for all houses in Milwaukee, we approximate year t's return by:

$$1 + \mathrm{r}_{\mathrm{i(t+1)}} = \left(\frac{\mathrm{P}_{\mathrm{i(t+1)}}}{\mathrm{P}_{\mathrm{it}}}\right)$$

even if the property was not actually sold. We observe the property purchase price and date from the City of Milwaukee Property Assessment Data (https://assessments.milwaukee.gov/) and the currently assessed value of the property from Milwaukee's Master Property File (MPROP). The MPROP (https://data.milwaukee.gov/dataset/mprop) also provides any rise or decline of the value from the previous year. We merge the two datasets using TAXKEY, a unique ten-digit number assigned by the City of Milwaukee to each property. Milwaukee Open Data Portal (https://data.milwaukee.gov/) also offers additional spatial or mapped data.

Based on the listed data sources, we use the following variables, including the year built, lot size, number of rooms, tax rate, parking type and spaces, neighborhood (crime rate, distance to groceries, park, bus stops, schools), and previous and current property values. We compute the current annual unlevered rate of return r_i^u and the predicted rate of return using the variables.

The second housing return measure is the levered rate of return r_i^l, satisfying

$$DownPay_{i0} = \sum_{t=1}^{T_i-1} \frac{rent_{it} - pymt_{it}}{\left(1 + r_i^l\right)^t} + \frac{max\{0.01, rent_{iT} - pymt_{iT} + 0.95P_{iT} - UPB_{iT}\}}{\left(1 + r_i^l\right)^{T_i}}$$

where $DownPay_{i0}$ is the down payment, $rent_{it}$ is monthly rent, $pymt_{it}$ is monthly housing costs, P_{iT} is the sale price, and UPB_{iT} is the outstanding principal balance at the time of sale.

We could impute these variables using the Home Mortgage Disclosure Act (HMDA) and Black Knight datasets; however, due to a high proportion of missing observations on $rent_{it}$ and UPB_{iT}, the analysis based on the levered rate of return was unfavorable in our analysis. The reported results in the following sections are all based on the assessed yearly return r_{it}.

The computation of the return r_{it} requires the merged HMDA and Black Knight dataset in Milwaukee. We requested detailed mortgage characteristics (income, loan amount, FICO credit score, interest rate) and loan performance information (missed payments, delinquency rate, foreclosure rate) from Black Knight so that we successfully conduct a matching with the existing HMDA information.

Using the described housing return measures, we run a basic linear regression:

$$r_{it} = \alpha_0 1\{Black_{it}\} + \alpha_1 1\{Hispanic_{it}\} + \alpha_2 1\{Sex_{it}\} + \mu_{c(i),t} + \varepsilon_i$$

where r_{it} is the housing return of house i at year t, $c(i)$ is house i's ZIP Code or Census tract, and t is the property assessment year. The indicator function $1\{Black_{it}\}$ is a binary variable whose value is 1 if the homeowner is Black and 0 otherwise. Similarly, other indicator functions show if the homeowner (or the primary mortgage applicant) is Hispanic or female. Thus, the parameter α_0 captures the average housing return gap between Black and other racial groups.

In the case that one finds a more detailed dataset with fewer missing observations, we can apply the same identification strategy using the approximated levered house return. The HMDA directly covers institutions to report every mortgage application they receive, along with the action taken (e.g., originated, denied), the amount and purpose of the loan, the Census tract of the property, the applicant's income, and race as well as several other variables. HMDA covers the vast majority of home purchase lending in the U.S., excluding only some exceedingly small lenders.

Black Knight data contain information on both mortgage characteristics measured at origination (e.g., loan-to-value ratio, property value, and borrower credit scores) and loan performance information, including monthly loan balance, payment amount, delinquency, and foreclosure. To get borrower income and tract matched to this data, Black Knight and HMDA records can be matched based on loan amount, purpose, type, origination date, occupancy type and a ZIP Code to Census tract crosswalk.

The HMDA-Black Knight merged data identify the components of the levered rate of return r_i^l. The down payment $DownPay_{i0}$ is the sum of initial equity and closing costs. Equity is defined by the difference between purchase price and loan values. The purchase price is from the City of Milwaukee Property Assessment Data and loan values are from HMDA. The closing costs are measured by the difference of total loan costs and lender credits, which are available in HMDA. The monthly rent $rent_{it}$ follows the median annual county-level rental costs. The monthly housing costs $pymt_{it}$ consist of principal and interest payments and insurance payments (escrow). The monthly mortgage payments are computed by the loan size and interest rate observed in the Black Knight dataset. The escrow payments are observed in the Black Knight dataset for the sample of first-lien mortgages. We can predict the escrow payments for all homeowners using the machine learning-based predictions. Lastly, we compute the principal balance UPB_{iT} using the standard amortization formulas assuming a 30-year term. The loan size and monthly loan balance are available in the Black Knight dataset.

2.2 Data Description (Applicant and Parcel-Level Data)

Home Mortgage Disclosure Act (HMDA): HMDA directs covered institutions to report every mortgage application they receive, along with the action taken (e.g., originated or denied), the *amount and purpose of the loan*, the Census tract of the property, the race and income of the applicant, and other variables. HMDA covers the vast majority of home purchase lending in the U.S., excluding only exceedingly small lenders.

Black Knight Public Property Data: The Black Knight dataset contains records of individual mortgage loans serviced by large servicers in the United States. It contains detailed information on origination and monthly loan performance. The loan and borrower information at origination includes the loan amount, interest rate type (fixed rate versus adjustable rate), amortization type and term, interest-only (IO) term, option adjustable-rate mortgage (ARM) indicator, loan-to-value ratio (LTV), maturity term, documentation type (full, low, no or reduced), borrower's credit score (FICO), debt-to-income ratio (DTI), loan purpose (e.g., home purchase versus refinance), occupancy status (e.g. owner-occupied versus investment); property type (e.g., single-family house versus condominium), property location (e.g., state, county, MSA, and ZIP Code), and so forth. This dataset also has information on the loan's status in each month following loan origination, so we can identify delinquency and default. Black Knight and HMDA records can be matched based on the loan amount, purpose, type, origination date, occupancy type, and ZIP Code to Census tract crosswalk. Further, the Black Knight-HMDA merged data provide additional borrower demographic information (e.g., race, ethnicity), which are important demand-side variables for mortgage loans.

Milwaukee's Master Property File (MPROP): MPROP data were obtained to examine the current value of each housing unit and *how the value has changed over time*. Specifically, these data include (1) the current assessed land value, (2) the current assessed value of all improvements on the property, (3) the current sum of land and improvement value, (4) the previous year's assessed land value, (5) the previous year's assessed value of all improvements on the property, and (6) the previous year's sum of land and improvement value. Also, historical MPROP data were requested to allow for the examination of longitudinal trends and how these trends vary with changes in neighborhood context.

American Community Survey (ACS): Based on the literature outlined above, we posit that although individual-level effects of homeowners are important, neighborhood-level effects are also likely to exist. For this reason, data on several neighborhood-level (i.e., Census tract) socio-structural measures were collected from the Census Bureau's ACS (2019, 5-year estimates). Because data on individual-level measures span from 2000 to 2020, data from the ACS reflect the 5-year estimates on data reflecting Census measures in 2010, 2015, and 2020.

To examine the racial distribution across Milwaukee tracts, data on the races/ethnicities of tract residents were collected from the ACS based on each of the three survey years listed above (5-year estimates). Specifically, these data were used to calculate measures reflecting the percentage of tract residents who identified as each of the following races/ethnicities: Asian, Black, Hispanic, non-Hispanic White, or "other" (i.e., those who identified as a race not previously listed, including American Indian/Alaska Native, Native Hawaiian or Pacific Islander, or as bi-/multi-racial).

Additionally, several other tract-level variables were collected for the ACS and were used in analysis. To reflect educational attainment, a series of three variables were collected (and are based on the timeframes listed above), including the percentage of tract residents who reported having a level of education that was less than obtaining their high school degree, the percentage of tract residents who reported earning at least their bachelor's degree, and the percentage of tract residents who reported educational attainment that was greater than their bachelor's degree (e.g., graduate school, law school). Next, a measure reflecting unemployment was created based on the percentage of tract residents who reported being unemployed. Further, three measures reflecting the level of economic disadvantage within a given tract were created, and based on the percentage of tract residents who reported household earnings that were less than the federally defined poverty line for the respective year, the percentage of tract residents who reported receiving Supplemental Nutrition Assistance Program ("SNAP") vouchers (formerly known as "food stamps") for a given year, and the percentage of tract residents who reported receiving public assistance for a given year.

To examine family/household dynamics, three additional measures were created. Here, measures were created that represent the percentage of households that had children under the age of 18 years, which were headed by a single male, a single female, or simply a single householder (i.e., the sum of single female- and single male-headed households). In addition, measures reflecting the percentage of tract residents who reported moving into their current dwelling/unit within the past five years were used to capture the tenure of householders.

Additionally, two additional variables were collected that represent housing attributes, and included the reported median value of houses located within a given tract and the reported median rent of rental properties located within a given tract.

Using these data, a series of principle component, regression-based factor analyses with a varimax rotation were conducted. Here, data were entered for each of the timeframes above, measures entered into the factor analysis included the percentage of residents identifying as White, the percentage of residents who reported receiving SNAP vouchers, the median home value, the percentage of female-headed households with minor children, the percentage of residents who reported being unemployed, the percentage of residents who reported not having earned their high school diploma/GED, the percentage of residents who reported renting their home, and the percentage of residents who reported moving into their current home within the past five years.

Based on the factor analysis, seven variables loaded onto the first factor and two variables loaded onto the second factor. The variables that loaded onto the first factor are those that are traditionally identified to be reflective of "neighborhood socioeconomic disadvantage," and included the percentage of residents identifying as White, the percentage of residents who reported receiving SNAP vouchers, the median home value, the percentage of female-headed households with minor children, the percentage of residents who reported being unemployed, and the percentage of residents who reported not having earned their high school diploma/GED. Two variables loaded onto the second factor included the percentage of owner-occupied units and the percentage of renter-occupied units. These measures are traditionally deemed to be reflective of "residential mobility."

To disaggregate neighborhoods based on racial composition, three measures are created to reflect the percentage of tract residents who identify as (1) Black, (2) Hispanic, and (3) non-Hispanic White. This allowed for the identification of housing and socio-structural contextual patterns across neighborhoods with varying racial compositions. In addition, the median net worth of tract residents and the percentage of residents who report being married are obtained from the ACS. Furthermore, the median age of tract residents was collected. Housing attribute data were collected from the ACS and measured at the tract level. These data captured the (1) median age of the home, (2) the average number of residents per square foot, (3) the percentage of households that are single-family homes (as compared to multi-family homes), (4) home appreciation price, and (5) median home value. Next, the locations of violent and property crimes are obtained from the Milwaukee Open Data Portal.

Milwaukee Open Data Portal: The Milwaukee Open Data Portal was used to obtain data on the locations of crimes reported to Milwaukee Police Department, which were identified as Wisconsin Incident Based Reports (WIBR) Group A Offenses. Address data on all WIBR Group A Offenses were collected from January 1, 2005, through December 31, 2020. WIBR Group A Offenses taking place within this timeframe included arson (N=4,740), assault (N= 159,912), burglary (N= 91,740), criminal damage (N= 98,601), homicide (N=1,696), robbery (N=49,037), theft (N= 158,298), and vehicle theft (N= 90,841). These data were geocoded using ArcGISPro and then joined to the tract to create a measure reflecting the numbers of violent offenses of each type of crime and of property crimes that took place within each tract during a given year. Since we are interested in the general effect of crime on home values, we construct a single multi-dimensional measure of crime. To do so, individual crime types were entered into principal components factor analysis to create a singular measure of neighborhood crime.

Other Geographical/Location Data: Data on institutions, resources, and other amenities were also collected. Data on the locations of county and state parks and bus routes are obtained from the Milwaukee Open Data Portal. Data on public and private services and amenities, including the locations of medical services (e.g., hospitals, urgent care facilities, general care providers), pharmacies, dentists and optometrists, counseling services, youth programs, social service agencies/programs, homeless shelters, and various types of grocery stores/food retailors (convenience stores versus grocery stores versus storefront grocery stores) were collected via Data Axel. All addresses were geocoded and then joined to the tract to create a count-based measure of each type of institution/service listed above.

Lastly, data on the locations of schools was collected from Data Axle. Schools were disaggregated into elementary, middle, and high schools. These address-based data were geocoded in ArcGISPro. Based on these data, several measures were created. First, measures of the number of each type of school per tract were created. Next, using point data, measures were created that reflected the distance (in miles) between any given parcel within the sample, and the nearest school. Three separate distance analyses were used to measure the distance from a given parcel to the nearest elementary, middle, and high schools. Next, because not all schools have equal performance, schools were further disaggregated to create groups of schools that "meet expectations" and that "exceed expectations." Schools deemed to be at least "meeting expectations" include schools categorized as meeting, exceeding, or significantly exceeding performance expectations during the 2020-2021 academic year, as identified by the Wisconsin Department of Public Instruction. Schools deemed to be "exceeding expectations" include schools categorized as exceeding or significantly exceeding performance expectations during the 2020-2021 academic year, as identified by the Wisconsin Department of Public Instruction (Wisconsin Office of Educational Accountability, 2021). Using the same distance analysis methods discussed above, measures on the distance between a given parcel and schools that met expectations and schools that exceeded expectations were created for each level of school (i.e., elementary, middle, and high school).

Survey of Income and Program Participation (SIPP) and Integrated Public Use Microdata Series (IPUMS): We observe economic well-being and family income details from the Survey of Income and Program Participation (SIPP) and Integrated Public Use Microdata Series (IPUMS) datasets. Both datasets have similar variables. The data has annual household-level income dynamics. The data particularly provides the (1) residence, (2) race, (3) home occupancy status (homeowner or renter), and (4) income details (salary income, investment, or property income). We identify the heterogeneous wealth creation processes between homeowners and renters across racial groups using the dataset.

Merging Black Knight, HMDA and MPROP: A Machine Learning Approach

To create our primary dataset, we combine variables from three sources: Black Knight, HMDA, and MPROP. Because both the Black Knight and MPROP datasets included a variable containing each property's exact address, it was possible to employ a simple data merger, which allowed for the retention of all relevant variables across the two datasets. The primary issue in merging the HMDA data with the other datasets was the result of a limited number of identifiers made available in the HMDA data. Specifically, HMDA does not include a home address identifier within the dataset. Below we detail the methods used to match and merge the data across these three datasets.

First, within the HMDA data, Milwaukee County contains 609,869 records in the dataset (2007-2020). Of those, 311,066 records led to the origination of a loan. Key variables that were used for property identification in the HMDA data were Census tract, loan type, loan amount, and loan origination year. Note, the complete loan origination date is not made available in public use HMDA data.

Second, the Black Knight data contained 442,983 loans that were originated in Milwaukee County (2005-2021). Several key variables were used for property identification, including Census tract, ZIP Code, full street address, loan type, loan amount, and loan origination date. These variables aided in matching these data with the HMDA data; however, this did not allow for a 100% successful match rate. Therefore, we developed additional criteria and restrictions to ensure the properties from Black Knight matched with the correct property owners from HMDA. To this end, we limited the sample from HMDA to include only those loan applications that led to the origination of a loan, were listed as "secured by the first lien," had a loan purpose listed as a "home purchase," and had an occupancy status as "owner-occupied as primary dwelling." These decisions were made in line with recent literature on methods used to remedy issues related to merging HMDA data.

After making these restrictions, we created further restrictions that mandated that HMDA entries could only be matched with Black Knight entries if they had the same Census tract, loan origination year, and loan type, as well as if they had loan amounts that were within 2% of one another. This filter led to an 89.8% match rate (*N*= 87,918). This matched database became our joint Black Knight-HMDA dataset, which was then merged with MPROP through matching property addresses in the way previously outlined, resulting in a complete Black Knight-HMDA-MPROP dataset.

2.3 Empirical Testing and Validation

We provide a structural approach to identify the source of the housing return gap between different racial groups. The dependent variable is the housing return measure r_{it}. The individual i follows different wealth accumulation paths depending on the homeownership and potential financial difficulties. In particular, the financial distress during the mortgage repayment period may result in default. In such a case, the property will be in short-sale or under foreclosure and the homeowner's realized housing return will be significantly reduced.

As the wealth accumulation process is summarized in housing returns, we construct a model:

$$\begin{aligned} E[Wealth_i \mid X] = {} & P(Mortgage_i = 0|X)\, E[Wealth_i \mid X, Mortgage_i = 0] \\ & + P(Mortgage_i = 1|X)\, P(Default_i = 1 \mid X, Mortgage_i = 1)\, E[Wealth_i \mid X, Mortgage_i = 1, Default_i = 1] \\ & + P(Mortgage_i = 1|X)\, P(Default_i = 0 \mid X, Mortgage_i = 1)\, E[Wealth_i \mid X, Mortgage_i = 1, Default_i = 0]. \end{aligned}$$

The group without applying for the mortgage or denied from the mortgage application ($Mortgage_i = 0$) accumulates wealth without relying on housing returns. The probability

$P(Mortgage_i = 1|X)$ is directly identified by the share of renters relative to homeowners from the ACS, SIPP, and IPUMS datasets. We predicted the approval probability using other explanatory variables in ACS, SIPP, and IPUMS, including the racial, regional, and economic variables (Wang et al., 2019) and marital status (Ishii-Kuntz et al., 2004). The ACS and IPUMS datasets also provide the median net worth and household income for renters (Choi et al., 2019). We assume that the wealth accumulation process for the no-housing group depends on their income only. We can directly observe the housing returns for groups who got approval from the mortgage application. The conditional probability of default $P(Default_i = 1 \,|X, Mortgage_i = 1)$ is identified by the Black Knight dataset. We also predicted the default probability using the Black Knight variables, including the credit score, income, racial indicators, and educational attainment.

The last stage estimation involves the prediction of two conditional expectations on wealth, $E[Wealth_i \mid X, Mortgage_i = 1, Default_i = 1], E[Wealth_i \mid X, Mortgage_i = 1, Default_i = 0]$. The machine learning-based models work to predict the housing returns based on the realized returns. The predictors derive counterfactual mortgage approval probability, default probability, and housing returns with respect to the change of policy variables.

To select the most suitable predictors, a number of machine learning algorithms were evaluated by dividing the preprocessed datasets into training, validation, and test datasets. We selected from several supervised machine learning models that include decision trees (Shalev-Shwartz & Ben-David, 2014), factorization machines (Rendle, 2010), gradient-boosted trees (Friedman, 1999), generalized linear regression (Madsen & Thyregod, 2011), and random decision forests (Ho, 1998). Most of these algorithms are implemented in frameworks like Spark, which supports distributed and parallel computation. We also evaluated the performance of deep neural network-based regression models by starting with a simple feedforward network (i.e., a multilayer perceptron) (Hastie et al., 2009) that consisted of several dense hidden layers followed by nonlinear activation functions.

3. Empirical Results

3.1 The Determinants of Home Value

3.1.1 Baseline regression

In this section, we investigate the factors that drive home value over time. We focus on the group of homeowners who got mortgage approval in this section and discuss the comparison of renters and homeowners in Section 5. For the identification of $E[Wealth_i \mid X, Mortgage_i = 1]$, we estimate the following equation using the Ordinary Least Squares with Zip Code and year fixed effects.

$$hval_{it} = \beta_0 + \beta_1 ethnicity_{it} + \beta_2 race_{it} + \beta_3 sex_{it} + \ln(income_{it}) + \ln(loan_{it}) + \epsilon_{it}$$

where the dependent variable $hval_{it}$ is the natural logarithm of home values. The baseline set of regressors includes ethnicity, race, sex, income, and loan size.

Ethnicity is a dummy variable indicating whether the homeowner is Hispanic (coded as 1) or non-Hispanic (coded as 0). Race is a categorical variable indicating the primary mortgage applicant's race and taking the following values: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. Sex is a dummy variable, in which the value of 1 represents that the primary mortgage applicant is female.

Due to the fixed effects controlling the ZIP Code-level and year-level unobserved heterogeneity, the estimation requires a package for the linear regression with high-dimensional fixed effects. The standard error of the regression is cluster-robust at the ZIP Code and year levels, as the houses in the same ZIP Code or assessed in the same year may share the same price shocks unobservable to the researchers.

Table 1 presents the results of the baseline regression. We begin the discussion with the coefficient for ethnicity (Model 1). The reported coefficient for Hispanic homeowner on home value is negative and highly significant (b=-0.0364; p<.001). Specifically, homeowners who identify as Hispanic had home values that were 3.64% lower than those who identified as non-Hispanic owners, while holding all other variables constant. Similarly, homeowners who identified as Black had home values that were significantly lower than those who identified as White (b=-0.0688; p<.001), in which Black owners had homes that were valued 6.88% lower than homes owned by White homeowners. Additionally, houses owned by female homeowners had values that were 1.42% lower than those owned by male homeowners, while holding all other variables constant (b=-0.0142; p<.01). Note that we reported only Black homeowner effects relative to White homeowners even though we controlled all race categorical values.

Other factors that drive up home value are the tenant's income, loan amount and the size of the lot. Each 1% increase in income is associated with an increase in home value by 0.1149% (b=-0.1149; p<.001). Further, for each 1% increase in loan amount, homes experienced an increased value by 0.2120% (b=.2120; p<.001), and larger lot size was associated with a 0.0174% increase in home value (b=.0174; p<.01).

Next, we turn to neighborhood-level indices that may impact the value of a home. First, we found that neighborhood crime rates play a critical role in the value of homes within a given neighborhood (Model 2). Specifically, for each one-unit increase in the crime factor score, a 4.9% decrease in home value was observed (b=-.0490; p<=.001). Also, at the neighborhood level, the value of a given house was found to be associated with the measures of educational attainment of residents of the neighborhood in which the home was situated (Model 3). Here, we employ four discrete measures of educational attainment, which include the percentage of tract residents who reported having earned less than a high school diploma (i.e., the percentage of residents who did not graduate from high school), the percentage of tract residents who had earned their high school degree, the percentage of tract residents who reported earning their bachelor's degree, and the percentage of residents who indicated that they had attained a level of education past their bachelor's degree (i.e., master's degree, doctoral degree/Ph.D., Juris Doctor/JD degree). The results show a clear dichotomy in the effects of education on home values. Neighborhoods that have higher percentages of residents with *lower* levels of educational attainment (i.e., higher percentages of residents who reported not earning more than their high school degrees; those who did not graduate from college) experienced lower-appraised home values. Conversely, homes located in neighborhoods with higher percentages of residents with *higher* levels of educational attainment (i.e., higher percentages of residents who reported earning at least their bachelor's degree) were found to have higher home values.

Table 1. Baseline

	(1)	(2)	(3)	(4)
Hispanic	-0.0364***	-0.0364***	-0.0121	-0.0125
	(-4.10)	(-4.13)	(-1.51)	(-1.55)
Black	-0.0688***	-0.0664***	-0.0471***	-0.0466***
	(-6.46)	(-6.43)	(-5.37)	(-5.32)
Female	-0.0142**	-0.0135**	-0.0139**	-0.0136**
	(-3.04)	(-2.91)	(-3.22)	(-3.18)
Log(income)	0.1149***	0.1140***	0.1015***	0.1014***
	(16.95)	(16.91)	(16.47)	(16.47)
Log(lot size)	0.0174**	0.0166*	0.0181**	0.0178**
	(2.75)	(2.59)	(2.93)	(2.87)
Log(loan)	0.2120***	0.2109***	0.1903***	0.1903***
	(20.11)	(20.03)	(19.35)	(19.32)
Crime		-0.0490***		-0.0175
		(-3.86)		(-1.58)
No HS percentage			-0.0077***	-0.0075***
			(-8.24)	(-8.06)
HS graduate percentage			-0.0030**	-0.0029*
			(-2.68)	(-2.59)
Bachelor percentage			0.0051***	0.0050***
			(3.58)	(3.44)
Above bachelor percentage			0.0068***	0.0067***
			(4.03)	(3.98)
Zip code dummies	Yes	Yes	Yes	Yes
Assessment year dummies	Yes	Yes	Yes	Yes
Observations	227631	227631	227631	227631
R^2	0.413	0.417	0.469	0.470

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the results of a regression where home value is the dependent variable. Columns (1) – (4) present the different model specifications. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

In Model 4, we reintroduced the crime variable to the model. In this full model, the crime variable no longer reaches a level of statistical significance. Here, we note that this may be due to the other variables within the model absorbing the majority of the variation within home values. Furthermore, there may be a correlation between educational attainment and crime and/or selection biases, as more educated people may choose to live in areas with lower crime rates. Additionally, we note that the association between Hispanic homeowners and home value decreased by approximately two-thirds. Here, we propose that educational attainment has a more pronounced effect on home value and, in turn, may mitigate the impact of ethnicity, specifically Hispanic ethnicity, on home value.

In summary, the findings from Table 1 indicate that homeowner sociodemographic indices (e.g., homeowner race, ethnicity, sex), as well as homeowner income, loan amount, and lot size play a significant role in determining home values. Moreover, although houses located in high crime areas have lower home values, neighborhood educational attainment was able to suppress the effects of crime on housing value.

3.1.2 The role of amenities and distance to school

Prospective homebuyers consider two important aspects: neighborhood amenities and distance to school. Amenities enhance the attractive features of a neighborhood, including infrastructure and community programs. For this study, we focus on several amenities, such as grocery stores, schools, youth programs, social services, pharmacies, dental/vision care, state parks, county parks, medical care, and homeless shelters. To simplify our analysis, we combined these variables by taking their simple average. We call the obtained variable "amenities."

Table 2 presents findings from the baseline model that has been extended to include measures related to neighborhood amenities. Model 1 includes the composite measure of amenities, which is discussed above. Here, although the coefficient for the impact of amenities on home value is significant, the coefficient is relatively small and does not meet the threshold for statistical significance. We point out that the inclusion of the amenities variable does not increase the explanatory power of our regression. Therefore, we surmise that it is likely that the other variables within the model, such as income, lot size, and demographic characteristics, hold greater importance in determining home value than do local amenities.

A more nuanced approach to examining the importance of schools as a specific type of amenity revealed that the distance between a given home and a school is a key factor associated with housing values. To further analyze schools, we created measures reflective of the distance between a given parcel and three types of schools: "any school" (i.e., no restriction on quality or grade level of school), a school "meeting expectations," and a school "exceeding expectations" (see methods section for definitions of these categories). The results indicate that the quality and distance between a home and a school do indeed impact the value of the home. Notably, as the distance between a house and a school that exceeds expectations increases, the value of the home decreases ($b=.0768$; $p<.05$) (Model 2).

Next, we further disaggregate schools by education level to create three measures that capture the distance between a given house and the locations of an elementary school, a middle school, and a high school that exceed expectations (Model 3). Within this model, the distance between a parcel and both elementary and middle schools are statistically significant; however, the coefficient measuring the distance between a given parcel and a high-performing elementary school on home value is not in the anticipated direction. Here, we observe a positive coefficient, which indicates that housing values are higher for homes that are located a farther distance from an elementary school that exceeds expectations ($b=.0459$; $p<.05$). In terms of the distance between a given house and a middle school that exceeds expectations, we observe the hypothesized effect, in which houses that are farther away from a high-performing middle school have lower assessed values ($b=-.0523$; $p<.05$). In the final model, we extend the baseline to include only high schools exceeding expectations (Model 4). Here, the impact of distance between a given house and a high-performing high school on housing value is negative, which indicates that as the distance increases, the value of the home decreases ($b=-.0460$; $p<.05$). Again, we note that this effect is only significant when the elementary and middle school measures are omitted from the model, which indicates that the impact of proximal high schools on the value of a home is suppressed by the impact of proximal elementary and middle schools.

Table 2. Amenities

	(1)	(2)	(3)	(4)
Hispanic	-0.0391***	-0.0384***	-0.0370***	-0.0398***
	(-4.39)	(-4.32)	(-4.24)	(-4.46)
Black	-0.0698***	-0.0670***	-0.0645***	-0.0674***
	(-6.37)	(-6.48)	(-6.23)	(-6.41)
Female	-0.0120*	-0.0130**	-0.0134**	-0.0130**
	(-2.59)	(-2.76)	(-2.86)	(-2.78)
Log(income)	0.1074***	0.1069***	0.1065***	0.1072***
	(16.38)	(16.44)	(16.02)	(16.20)
Log(lot size)	0.0171**	0.0164*	0.0180**	0.0161*
	(2.77)	(2.53)	(2.90)	(2.59)
Log(loan)	0.2045***	0.2042***	0.2039***	0.2045***
	(19.96)	(19.90)	(19.51)	(19.68)
Amenities	0.0030	0.0029	0.0034	0.0038
	(1.23)	(1.23)	(1.48)	(1.62)
Distance to:				
Any school		0.0682		
		(1.82)		
School meeting expectations		-0.0039		
		(-0.07)		
School exceeding expectations		-0.0768*		
		(-2.12)		
Elementary school exceeding expectations			0.0459*	
			(2.00)	
Middle school exceeding expectations			-0.0523*	
			(-2.20)	
High school exceeding expectations			-0.0343	-0.0460*
			(-1.74)	(-2.38)
Zip code FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	216482	216482	216482	216482
R^2	0.425	0.428	0.430	0.427

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the results of a regression where home value is the dependent variable. Columns (1) – (4) present the different model specifications. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

To conclude, while general amenities are likely considered by prospective homebuyers, our findings suggest that they do not significantly affect home value. However, distance to schools, particularly high-quality schools, was found to be a factor associated with house value. The effects of school quality and distance varied by grade level, with greater distances to middle or high schools that exceeded expectations being associated with lower home values. Conversely, as the distance to an elementary school that exceeded expectations increased, housing value also increased.

3.1.3 Homeownership threshold

Next, we examine whether the percentage of owner-occupied housing units is associated with house value. More specifically, we aim to identify the threshold for the percentage of owner-occupied houses on housing values (i.e., at what percentage does the percentage of owner-occupied households exist versus disperse). This line of inquiry is motivated by the literature on urban economics that examines the positive externality generated by homeownership, which provides a rationale for government subsidy to homeownership (for example, see Coulson et al., 2002).

Findings from the analysis examining this relationship are presented in Table 3. Here, the results demonstrate that when the percentage of owner-occupied housing units within a given Census tract is less than 30% (i.e., 70% of the housing units are either renter-occupied or vacant), there is a negative effect on house value. In neighborhoods where less than 30% of housing units are owner-occupied, each additional 1% decrease in the percentage of owner-occupied housing units is associated with a 0.6% decrease in housing value. Stated differently, a home that is in a tract where only 24% of houses are owner-occupied has, on average, a house value that is 6% lower than a comparable home located within a tract where 30% of the housing units are owner-occupied.

Table 3. Ownership threshold

	(1) Below 30%	(2) Above 30%
Hispanic	-0.0376	-0.0333***
	(-1.01)	(-3.87)
Black	-0.1872***	-0.0414***
	(-4.65)	(-5.25)
Female	-0.0222	-0.0115**
	(-1.06)	(-2.84)
Log(income)	0.1150***	0.1074***
	(6.32)	(16.08)
Log(lot size)	0.0015	0.0195**
	(0.10)	(2.68)
Log(loan)	0.3312***	0.1857***
	(8.21)	(20.58)
Owner percentage	-0.0059*	0.0032***
	(-2.42)	(5.64)
Zip Code FE	Yes	Yes
Year FE	Yes	Yes
Observations	23117	204514
R^2	0.517	0.415

Note: This table presents the results of a regression where home value is the dependent variable. Columns (1) – (2) present two different samples based on ownership rates. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

3.1.4 Heterogeneity across ZIP Codes

In this section, we explore heterogeneity across different ZIP Codes in our sample. As explained in the introduction, Milwaukee is one of the most racially segregated cities in the United States. To examine whether housing markets in majority-minority ZIP Codes behave differently, we focus on two ZIP Codes: 53206 and 53209. Within 53206, 85.12% (3,495 out of 4,106 residents) identified as Black, while in 53209, 50.90% (5,718 out of 11,232 residents) identified as Black. The first column of Table 4 presents the results of analysis from the baseline models discussed in the previous sections. The second and third models are based on analysis using data drawn exclusively from the 53206 and 53209 ZIP Codes, respectively.

From Table 4, three factors that affect home value stand out when comparing the two ZIP Codes, 53206 and 53209, to the baseline ZIP Code: crime, sex, and education. First, we find that crime is associated with a decrease in home value in certain ZIP Codes. In 53206, each one-unit increase in crime is associated with a 45.50% decrease in home value ($b=-0.4550$, $p>.05$). In comparison, within the baseline model, a one-unit increase in crime is associated with a 0.43% decrease in home value. Second, regarding the association between homeowner sex and home value, being a female homeowner in 53206 was associated with an 18.21% decrease in home value, while being a female homeowner in 53209 was associated with a 7.46% decrease in home value compared to the baseline model. Third, the level of neighborhood education had a greater impact on homeowners in 53206 than in the baseline and 53209 models. Here, each one-unit increase in the percentage of tract residents without at least a high school degree or higher was associated with a 2.10% decrease in home value. On the other hand, each one-unit increase in the percentage of tract residents who reported having earned at least a bachelor's degree was associated with a 0.82% increase in home value.

Table 4. Heterogeneity across ZIP Codes

	Baseline	53206	53209
Hispanic	0.0117	0.1264	0.0458
	(0.73)	(1.50)	(1.10)
Female	-0.0575***	-0.1821*	-0.0746**
	(-5.60)	(-2.37)	(-2.92)
Log(income)	0.0994***	0.1007	0.1330***
	(7.70)	(1.24)	(4.31)
Log(lot size)	0.0553***	0.4984***	0.3756***
	(6.23)	(3.81)	(4.21)
Log(loan)	0.1161***	0.0380	0.0337
	(7.86)	(0.41)	(1.29)
Crime	-0.0043	-0.4550	0.0150
	(-0.42)	(-1.84)	(0.26)
No HS percentage	-0.0076***	-0.0201**	-0.0087
	(-7.01)	(-3.25)	(-1.80)
HS graduate percentage	-0.0017	-0.0020	-0.0016
	(-1.37)	(-0.89)	(-0.42)
Bachelor percentage	0.0047**	0.0104	0.0020
	(2.89)	(1.04)	(0.46)
Above bachelor percentage	0.0057**	0.0082*	-0.0012
	(3.26)	(2.09)	(-0.48)
Zip code dummies	Yes	No	No
Assessment year dummies	Yes	Yes	Yes
Observations	32243	1023	2575
R^2	0.514	0.462	0.530

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents the results of a regression where home value is the dependent variable. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

3.2 The Effect of Foreclosure

3.2.1 The determinants of foreclosed home values

In this section, we restrict our baseline regression to the sample of houses that have experienced foreclosure. The goal of this section is to identify the conditional probability of default for homeowners, $P(Default_i = 1 \mid X, Mortgage_i = 1)$ (see Section 3.2.4), and the expected wealth of those who have ever experienced foreclosure or other financial difficulties due to mortgage repayment failure, $E[Wealth_i \mid X, Mortgage_i = 1, Default_i = 1]$ (see Sections 3.2.1 to 3.2.3) in our model structure. The literature on housing market returns has shown the adverse effect of foreclosure on housing prices, and as such, we examine the role of foreclosure on housing wealth accumulation in Milwaukee County. Below, we discuss how the findings within the foreclosed home models differ from those in the baseline model (Table 55).

Table 5. Determinants of foreclosed home values

	(1)	(2)	(3)	(4)
Hispanic	-0.0120	-0.0134	0.0014	0.0011
	(-0.70)	(-0.78)	(0.09)	(0.07)
Black	-0.0689***	-0.0677***	-0.0537**	-0.0536**
	(-3.85)	(-3.81)	(-3.23)	(-3.22)
Female	-0.0590***	-0.0586***	-0.0527***	-0.0526***
	(-5.32)	(-5.31)	(-4.90)	(-4.90)
Log(income)	0.1062***	0.1062***	0.0959***	0.0960***
	(7.37)	(7.37)	(7.54)	(7.53)
Log(lot size)	0.0549***	0.0547***	0.0550***	0.0550***
	(6.37)	(6.35)	(6.26)	(6.26)
Log(loan)	0.1336***	0.1334***	0.1167***	0.1167***
	(8.78)	(8.76)	(7.95)	(7.95)
Crime		-0.0236*		-0.0037
		(-2.39)		(-0.37)
No HS percentage			-0.0076***	-0.0076***
			(-6.84)	(-6.94)
HS graduate percentage			-0.0016	-0.0016
			(-1.35)	(-1.34)
Bachelor percentage			0.0047**	0.0047**
			(2.94)	(2.91)
Above bachelor percentage			0.0056**	0.0056**
			(3.22)	(3.22)
Zip code dummies	Yes	Yes	Yes	Yes
Assessment year dummies	Yes	Yes	Yes	Yes
Observations	32243	32243	32243	32243
R^2	0.466	0.467	0.516	0.516

Note: This table presents the results of a regression where home value is the dependent variable. The sample is restricted to homes which experienced foreclosure t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

When examining the impact of being a Hispanic homeowner on home value, we found that the association between Hispanic homeownership and home value was weaker when considering only foreclosed homes, compared to the full baseline model that included both foreclosed and non-foreclosed homes. However, these effects remained statistically non-significant. In contrast, the coefficient between being a Black homeowner and home value remained consistent and significant across the baseline models and those examining only foreclosed homes. Specifically, Black homeowners experienced home values that were 5.4% to 6.8% lower than non-Black homeowners, depending on other variables included in the model. Lastly, the effect of being female on home value was approximately four times greater in homes that had experienced foreclosure compared to the baseline model that included all homes. In the baseline model, female homeowners had home values that were approximately 1.4% lower than male homeowners when controlling for other variables. However, when examining only homes that had been foreclosed, female homeowners had home values that were 5.3% to 5.9% lower than those of male homeowners. These findings suggest that the values of foreclosed homes are more sensitive to homeowner sex and less responsive to ethnicity.

3.2.2 Effect of foreclosure on Black and female homeowners

In the previous section, we show that foreclosure is more sensitive to certain variables. To investigate this further, we extend our baseline regression model by interacting race and sex to a foreclosure dummy variable. Results are reported in Table 6.

Table 6. Foreclosure interacted with race and sex

	(1)	(2)	(3)
Hispanic	-0.0300***	-0.0295***	-0.0332***
	(-3.52)	(-3.45)	(-3.86)
Black	-0.0611***		-0.0596***
	(-6.06)		(-5.92)
Female	-0.0141**	-0.0171***	
	(-3.13)	(-3.84)	
Log(income)	0.1101***	0.1153***	0.1128***
	(16.58)	(16.92)	(16.93)
Log(lot size)	0.0169**	0.0172**	0.0172**
	(2.69)	(2.70)	(2.72)
Log(loan)	0.2173***	0.2135***	0.2150***
	(20.24)	(19.99)	(20.21)
Foreclosed	-0.1061***		
	(-13.45)		
Black × Foreclosed		-0.1100***	
		(-7.71)	
Female × Foreclosed			-0.1222***
			(-12.45)
Constant	8.2088***	8.1858***	8.1969***
	(40.95)	(40.16)	(40.93)
Zip code FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	227631	227631	227631
R^2	0.417	0.412	0.415

Note: This table presents the results of a regression where home value is the dependent variable. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

We extended our baseline model to include a foreclosure dummy variable that takes a value of 1 if a house has experienced foreclosure in the past and 0 if not (see Table 6, Model 1). The results showed that foreclosure is associated with a decrease in home value of 10.61%. Next, we interacted the foreclosure dummy variable with measures reflective of Black and female homeowners (Models 2 and 3, respectively). These models revealed that the negative effect of foreclosure on home value is amplified when the homeowner is either Black or female, with the devaluation of foreclosed homes owned by a Black individual being approximately 11%, and that of foreclosed homes owned by a female being 12.22%. In summary, our findings suggest that past foreclosure results in lower home values, and this effect is even more pronounced when the homeowner is either Black or female.

3.2.3 Foreclosure threshold

We have highlighted that foreclosure has a negative impact on home value. A question that needs to be addressed is what percentage of houses in a certain neighborhood need to be foreclosed on to negatively affect the values of the surrounding houses. Table 7 reports the results.

Table 7. Foreclosure threshold

	(1) Full Sample	(2) Below 95th	(3) Above 95th
Hispanic	-0.1715***	-0.1692***	0.0502
	(-8.91)	(-8.59)	(1.17)
Black	-0.0520***	-0.0373*	-0.1568***
	(-3.43)	(-2.41)	(-5.08)
Female	-0.0038	-0.0051	-0.0725**
	(-0.75)	(-0.99)	(-3.15)
Log(income)	0.1375***	0.1478***	0.0981***
	(15.54)	(16.45)	(4.44)
Log(lot size)	0.0206**	0.0204**	-0.0415
	(3.07)	(3.02)	(-1.49)
Log(loan)	0.2431***	0.2552***	0.1929***
	(20.39)	(20.28)	(5.35)
Foreclosure	-0.0134***	-0.0120***	-0.0957**
	(-5.74)	(-4.86)	(-2.73)
Year FE	Yes	Yes	Yes
Observations	216482	216222	9687
R^2	0.340	0.324	0.428

Note: This table presents the results of a regression where home value is the dependent variable. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

Model 1 of Table 7 reports our extended baseline regression with the foreclosure variable. On average, foreclosure lowers home value by 1.34%. To further explore this, we continued analysis using a restricted sample, as well as a foreclosure variable varying with time and ZIP Code. Next, we ranked foreclosure from highest to lowest. Further, we created a 95th percentile threshold to employ in these analyses.

Our findings suggest that the effect of foreclosure is non-linear. When the percentage of foreclosed houses in a ZIP Code is below the 95th percentile, foreclosure reduces home value by 1.2%, which is less than the average reported in Model 1. However, when foreclosed homes are in neighborhoods with foreclosure rates above the 95th percentile, they experience a 9.57% decline in home value. These results indicate that the impact of foreclosure is disproportionate in areas with higher foreclosure rates. Specifically, the effect of unit-specific and neighborhood foreclosures on home values is non-linear. Homes in ZIP Codes with exceptionally high foreclosure rates (i.e., 95th percentile or higher) have significantly lower appreciation rates.

3.2.4 The determinants of foreclosure

In this section, we examine additional drivers of home foreclosures. Our focus in this subsection is to identify $P(Default_i = 1 \mid X, Mortgage_i = 1)$. We run the following linear model where the dependent variable is a binary variable.[2] This model is known as the linear probability model with ZIP Code and year fixed effects. We choose this approach to allow for a more meaningful interpretation of the coefficients.

$$fc_{it} = \beta_0 + \beta_1 ethnicity_{it} + \beta_2 race_{it} + \beta_3 sex_{it} + \ln(income_{it}) + \ln(loan_{it}) + \epsilon_{it}$$

where the dependent variable fc_{it} is the foreclosure dummy variable taking the value of 1 if a house has been foreclosed on and 0 otherwise. The baseline set of regressors includes key variables, including ethnicity, race, sex, income, and loan amount. The definitions of the regressors are discussed in the previous section.

[2] Other non-linear estimation strategies include the probit and logit models, but these methods are computationally burdensome to deal with a large-size dataset and high-dimensional fixed effects.

Ethnicity is a dummy variable indicating whether the primary loan applicant is Hispanic or not. Sex is a dummy variable taking the value of 1 if the primary loan applicant is female. Race indicates the primary mortgage applicant's race, which is a categorical variable taking the following values: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. Results are reported in Table 8. Since White is the baseline race in our specification, the coefficients on ethnicity represent the effect relative to White Americans.

Model 1 in Table 8 reports the result for the baseline regression model. First, in terms of sociodemographic factors, it was found that if the homeowner is Hispanic, Black, or female, they are more likely to experience foreclosure. Specifically, the likelihood of foreclosure is 4.02% higher for Hispanic homeowners and 3.81% for Black homeowners, when compared to the referent categories ($p<.001$). Although female homeowners are slightly more likely to own a home that is foreclosed on (0.92%) than are male homeowners, this association failed to meet a level of statistical significance.

Table 8. Determinants of foreclosure

	(1)	(2)	(3)	(4)
Hispanic	0.0402***	0.0402***	0.0362***	0.0363***
	(4.24)	(4.24)	(3.81)	(3.83)
Black	0.0281***	0.0275***	0.0244**	0.0242**
	(3.54)	(3.49)	(3.14)	(3.13)
Female	0.0092	0.0090	0.0091	0.0090
	(1.93)	(1.90)	(1.92)	(1.90)
Log(income)	-0.0462***	-0.0460***	-0.0440***	-0.0440***
	(-10.20)	(-10.18)	(-9.68)	(-9.68)
Log(lot size)	-0.0047**	-0.0045*	-0.0048*	-0.0047*
	(-2.62)	(-2.49)	(-2.52)	(-2.45)
Log(loan)	0.0489***	0.0492***	0.0525***	0.0525***
	(11.08)	(11.16)	(12.05)	(12.05)
Crime		0.0115**		0.0064
		(2.97)		(1.86)
No HS percentage			0.0012***	0.0011***
			(4.13)	(3.94)
HS graduate percentage			0.0009***	0.0009***
			(4.00)	(3.89)
Bachelor percentage			-0.0007*	-0.0007
			(-1.97)	(-1.87)
Above bachelor percentage			-0.0010**	-0.0010*
			(-2.63)	(-2.56)
Zip code dummies	Yes	Yes	Yes	Yes
Assessment year dummies	Yes	Yes	Yes	Yes
Observations	227631	227631	227631	227631
R^2	0.034	0.034	0.038	0.038

Note: This table presents the results of a regression where foreclosure is the dependent variable. Foreclosure is a dummy variable taking the value of 1 if a house has been foreclosed on and 0 otherwise. t-statistics are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

In addition, those who have higher incomes and those who own homes with larger lot sizes are less likely to own a home that has been foreclosed on. On the other hand, those with larger loan amounts are more likely to own a home that has been foreclosed on.

4. Counterfactual Exercises

In this section, we explore the changes in house value across years and subsamples. We perform various counterfactual exercises. For example, we are interested in examining what would have been the price appreciation of minority homeowners in the absence of foreclosure. Our dataset covers house prices from 2000 to 2020 for the different Milwaukee ZIP Codes. We can then look at the average 21-year house price appreciation. We extend this analysis to account for different race subsamples and houses that previously experienced foreclosure.

Our methodology is as follows.

- First, we calculate the average house prices in 2000 for each ZIP Code. We use these initial house prices as the baseline value of the index. As a result, the price index is 100 in 2000 under the column “all houses” in Table 9. We calculate the price index as follows:

$$Price\ Index_{i,t} = \frac{Average\ house\ price_{i,t}}{Average\ house\ price_{i=all\ house\ ,\ \ t=200}}$$

  where *i* represents the race subsample and *t* the time.

- Now we can begin to look at the average house price for different subsamples in the year 2000. For example, we calculate the average house prices for White individuals for each ZIP Code. We then divide the obtained number to the average house prices including all houses in the corresponding ZIP Code to obtain the price index.
- We extend our calculations in 2000 for a subsample of Black individuals, foreclosed houses, foreclosed houses belonging to White individuals, and foreclosed houses belonging to Black individuals.

- We continue our analysis and now look at the house prices in 2020. We calculate the price index for the same subsample before.

Table 9. Counterfactual exercises

2000

Zip Code	All Houses	White	Black	Foreclosed	Foreclosed White	Foreclosed Black
53202	100	103	69	77	103	-
53204	100	99	76	79	99	62
53206	100	99	100	103	99	107
53207	100	100	94	80	100	92
53208	100	104	76	77	104	51
53209	100	114	90	87	114	84
53210	100	106	85	81	106	74
53212	100	101	79	83	101	59
53213	100	100	103	93	100	89
53214	100	101	110	100	101	86
53215	100	99	102	92	99	80
53216	100	101	98	96	101	95
53218	100	98	101	99	98	97
53219	100	100	95	97	100	80
53220	100	101	100	93	101	98
53221	100	100	96	82	100	87
53222	100	101	96	92	101	89
53223	100	99	87	84	99	77
53224	100	103	98	90	103	89
53225	100	103	95	91	103	77

2020

Zip Code	All Houses	White	Black	Foreclosed	Foreclosed White	Foreclosed Black
53202	167	171	139	127	148	-
53204	215	211	218	135	133	117
53206	138	167	136	148	161	144
53207	212	211	191	167	169	177
53208	190	200	140	153	162	90
53209	151	167	139	131	138	126
53210	160	174	128	119	128	101
53212	306	314	213	253	284	149
53213	171	168	189	148	148	151
53214	150	152	142	156	162	134
53215	166	166	156	148	150	127
53216	146	147	135	142	133	144
53218	133	130	186	133	140	127
53219	151	151	149	146	147	122
53220	149	150	169	137	137	149
53221	154	154	125	125	124	165
53222	165	167	126	151	154	147
53223	150	152	262	120	122	113
53224	166	171	147	141	144	148
53225	155	159	100	134	134	134

House Value Appreciation
> 100%
< 50%

Table 9 presents the average growth in house prices in the Milwaukee area from 2000 to 2020, which is 68%. However, there is some variation in the growth rates across ZIP Codes. Specifically, the three ZIP Codes highlighted in red experienced a growth of over 100%, while the two ZIP Codes highlighted in blue had growth rates of less than 50%.

Moreover, we analyzed the values of houses based on the owner's race. In 2000, houses owned by White individuals had higher values compared to houses owned by individuals of other races, and this trend remained consistent in 2020. We then investigated whether foreclosure affects the value of a house over time and if this effect differs based on the owner's race. We found that foreclosure leads to a decrease in the house's value. As shown in Table 9, the value of foreclosed houses is less than the average value of all houses in the sample.

Furthermore, we examined the foreclosed houses based on the race of their owners. We found that houses owned by White individuals had higher values. However, we were unable to determine whether they retain their value more after foreclosure compared to houses owned by individuals of other races.

5. Disparity Between Homeowners and Renters

To examine the roles of homeownership and renting in wealth creation, we employed the Integrated Public Use Microdata Series (IPUMS) USA dataset. It's worth noting that the dataset used in this study is less comprehensive than the one used in previous analyses. Nonetheless, it enabled us to collect cross-sectional data on both renters and owners in Milwaukee at the county level between 2000 and 2020. We gathered pertinent information on detailed income data, mortgage status, monthly rent, annual property tax, and other housing maintenance items.

Our analysis of the 2020 IPUMS cross-sectional data indicated that there are racial disparities in homeownership. The dataset contained 608,709 observations for Wisconsin and 54,650 for Milwaukee County, of which 65% were White, 23% were Black, and 4% were Asian. Further examination revealed that 69% of White individuals were homeowners in the sample, compared to 56% of Asian individuals and 35% of Black individuals. The objective of this section was to provide a broad overview of the divergence in wealth creation between owners and renters.

5.1 The Determinants of Homeownership

The summary statistics above reveal racial disparity in homeownership. Most White households are homeowners, whereas most Black households are renters. In light of these preliminary findings, we set out to study the determinants of ownership. We run the following probit regression, assuming that the probability of becoming a homeowner is a nonlinear function of regressors.

$$\Pr(Owner_{it} = 1|X) = \Phi(\beta_0 + X_{it}'\beta)$$

where the dependent variable *Owner* is a dummy variable taking the value of 1 if the person occupying the house owns it and 0 otherwise. X_{it} is a set of regressors that includes information on age, education, employment status, sex, income family size, marital status, and race. The regression informs about $P(Mortgage_i = 0|X)$, which is part of a non-homeowner's wealth accumulation process.

Specifically, variables are coded as follows: the education variable college is a dummy variable taking the value of 1 if the person holds a bachelor's degree and 0 otherwise. Sex is coded as a dummy variable, in which female is coded as 1, and male is coded as 0. The employment status variable, "employed," takes the value of 1 if the person is currently employed and 0 if the individual is not currently employed. The race variables Asian, Black, and White are also coded as dummy variables. Age, income, and family size are continuous variables.

Table 10 presents the results of our estimation. In Model 1, we find that if the person went to college, the likelihood of being a homeowner increases by 9%. An employed person is 11% more likely to own a house, compared to someone who is not currently employed. Females are 4% less likely to own a house, compared to males. This magnitude decreases as we control for more variables, as seen in Models 2 and 3. In Model 2, we introduce family size and marital status to our baseline regression. Larger families are more likely to own their house, compared to families with fewer people. Additionally, being married increases the likelihood of being an owner by 23%.

Table 10. Likelihood of ownership

Probit Regression	Ownership		
	Model 1	Model 2	Model 3
Age	0.01 ***	0.01 ***	0.01 ***
	(0.00)	(0.00)	(0.00)
College	0.09 ***	0.09 ***	0.05 ***
	(0.01)	(0.00)	(0.00)
Employed	0.11 ***	0.07 ***	0.05 ***
	(0.01)	(0.01)	(0.01)
Female	-0.04 ***	-0.02 ***	-0.01 **
	(0.00)	(0.00)	(0.00)
Income (in logs)	0.00 ***	0.00 ***	0.00 ***
	(0.00)	(0.00)	(0.00)
Family size		0.04 ***	0.05 ***
		(0.00)	(0.00)
Marital status		0.23 ***	0.18 ***
		(0.01)	(0.01)
Asian			0.05 ***
			(0.01)
Black			-0.08 ***
			(0.01)
White			0.19 ***
			(0.01)
Observations	46208	46208	46208

Note: This table presents the results of a probit regression where owner is the dependent variable. Owner is a dummy variable taking the value of 1 if the occupant of the house owns it and 0 otherwise. Robust standard errors are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

The specification in Model 3 sheds light on the disparity of house ownership among races. Specifically, if a person is White, they are 19% more likely to become homeowners compared to those who identify as another race. Individuals who identify as Asian are 5% more likely to own a house, whereas individuals who identify as Black are 8% less likely to own a house.

5.2 The Determinants of Wealth Among Owners and Renters

Using the IPUMS data, we create a wealth index for renters and homeowners. For the renters, we calculate their wealth according to the following. The wealth of renters is the sum of their total wage income plus investment/dividend income minus rent and taxes. The wealth of homeowners instead is the sum of their total wage income plus investment plus home value increase minus property tax and mortgage payments. The difference is that the house is an asset for the homeowners.

Next, our focus is to figure out the expected wealth $E[Wealth_i \mid X, Mortgage_i = 0]$ for non-homeowners. We run the following linear regression using the ordinary least squares (OLS) method.

$$Wealth_{it} = \beta_0 + X_{it}'\beta + \xi_{it}$$

where wealth is the dependent variable constructed as described above. X_{it} is a set of regressors including age, sex, income, education, and race. These variables are coded using the same methodology discussed above. Table 11 reports the results of our estimation. First, we find that wealth is significantly determined by and positively correlated with age, income, education, and race. Although being female elicited a negative effect, this coefficient did not meet a level of statistical significance.

Second, we take a closer look at the racial disparity of homeownership. Among White individuals, being an owner was associated with wealth growth that was approximately twice as great compared to those who were renters (25% and 12%, respectively). Surprisingly, among Black individuals, a decrease in wealth regardless of ownership versus rentership, was found, in which a 32% decrease in wealth was observed. Also, contrary to expectations, among Asian individuals, those who rented experienced wealth growth that was five times greater than those who owned a home (41% and 8%, respectively).

Table 11. Determinants of wealth among owners and renters

	Dependent variable:	
	Wealth (in logs	
	Owner	Renter
Age	0.0005***	0.002***
	(0.0002)	(0.0003)
Age squared	-0.0000	-0.0000***
	(0.0000)	(0.0000)
Female	-0.001	-0.003
	(0.002)	(0.003)
Income (in logs)	0.002***	0.01***
	(0.0005)	(0.001)
College	0.01***	0.01***
	(0.002)	(0.003)
Asian	0.08***	0.41***
	(0.005)	(0.01)
Black	-0.32***	-0.33***
	(0.004)	(0.01)
White	0.25***	0.12***
	(0.003)	(0.01)
Constant	12.12***	9.73***
	(0.02)	(0.02)
Observations	27,574	18,634
R^2	0.70	0.59

Note: This table presents the results of a linear regression where wealth is the dependent variable. We divide the sample to distinguish between homeowners and renters. Robust standard errors are presented in parentheses. ***, **, and * denote statistical significance at 99, 95, and 90 percent confidence levels.

In summary, racial disparities persist not only among homeownership, but in the effect of homeownership on building one's wealth. While White individuals experience benefits in gains related to wealth, this benefit is smaller among Asian individuals. Moreover, homeownership was found to be a disadvantage in terms of wealth growth among Black individuals.

6. Policy Implications

The results obtained in our analysis have significant public policy implications, particularly regarding the role of foreclosure, amenities, and homeownership rates in wealth creation in Milwaukee. It cannot be overstated that foreclosed homes appreciate at a lower rate than non-foreclosed homes, and the degree of disparity that foreclosures create for Black homeowners in ZIP Codes with high foreclosure rates is surprising. The negative externality generated by high foreclosure rates warrants a public policy response with significant implications for wealth creation among minority households. High foreclosure rates may affect the local housing markets by creating a vicious cycle of oversupply and lower prices, and a disamenity effect where properties may not be maintained and may remain vacant for long periods. Consideration should be given to a public policy response that provides short-term aid to individuals to avert some foreclosures.

Several strategies may help mitigate the odds and frequencies of home foreclosures, such as providing financial assistance to homeowners who are at risk of losing their homes. This may include programs that offer low-interest loans to help homeowners pay their mortgages or financial incentives for lenders to modify the terms of existing mortgages to make them more affordable. Policies that aim to stabilize the housing market and prevent widespread declines in property values may also help to reduce the number of foreclosures, such as government policies that encourage banks to lend to qualified homebuyers or provide tax incentives for individuals and families to purchase homes.

The finding that there is a threshold effect in the impact of homeownership on home values in a neighborhood provides a rationale for government support of homeownership. Developing affordable housing should be considered an important goal for policymakers, including measures such as implementing zoning laws that encourage the construction of new homes or providing tax breaks or other incentives to developers who build affordable housing units. Regulators and policymakers can play an important role in stabilizing the housing market by ensuring responsible lending and providing support to homeowners who are at risk of foreclosure.

Our findings shed light on the disparity in housing market returns for minority homeowners in Milwaukee County. Black homeowners witness a 6% to 7% lower appreciation in their home values, partly explained by amenities, school quality, and crime rate. After controlling for these covariates, the disparity in returns declines to 4.6%, showing that investment in amenities in minority neighborhoods can be used to reduce housing wealth disparity in Milwaukee. It is conceivable that income shocks, adverse life events, and a lack of legacy savings to fall back on may play a role for minority homeowners. The disproportionate impact of foreclosures on Black homeowners results in a higher probability of going into foreclosure and more negative returns conditional on getting foreclosed on. Neighborhoods with high foreclosure rates witness a much larger negative impact on house prices, reinforcing the vital role foreclosure plays in housing return. A well-targeted policy that targets areas with high foreclosure rates can stabilize the housing market and lead to a significant reduction in disparity in housing returns for minority homeowners.

Acknowledgement

We would like to acknowledge the invaluable contributions of our research assistants, Zo Andriantomanga and Matthew Waller. Mortgage Guaranty Insurance Corporation (MGIC) has provided the funding for this project.

References:

Adelman, R. M. (2004). Neighborhood opportunities, race, and class: The Black middle class and residential segregation. *City & Community*, *3*(1), 43-63.

Akbar, P. A., Li, S., Shertzer, A., & Walsh, R. P. (2019). *Racial segregation in housing markets and the erosion of black wealth* (No. w25805). National Bureau of Economic Research.

Alba, R., & Logan, J. (1993). Minority proximity to Whites in the suburbs: An individual-level analysis of segregation. *The American Journal of Sociology, 98*, 1388–1427.

Altonji, J. G., & Doraszelski, U. (2005). The role of permanent income and demographics in black/white differences in wealth. *Journal of Human Resources*, *40*(1), 1-30.

Ambrose, B. W., Conklin, J. N., & Lopez, L. A. (2020). Does borrower and broker race affect the cost of mortgage credit?. *The Review of Financial Studies*, *34*(2), 790-826.

Anacker, K. B. (2010). Still paying the race tax? Analyzing property values in homogeneous and mixed-race suburbs. *Journal of Urban Affairs*, *32*(1), 55-77.

Avenancio-Leon, C., & Howard, T. (2019). The assessment gap: Racial inequalities in property taxation. Available at: *SSRN 3465010*.

Barsky, R., Bound, J., Charles, K. K., & Lupton, J. P. (2002). Accounting for the black–white wealth gap: A nonparametric approach. *Journal of the American Statistical Association*, *97*(459), 663-673.

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*.

Bayer, P., Casey, M., Ferreira, F., & McMillan, R. (2017). Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, *102*, 91-105.

Bellair, P. E. (1997). Social interaction and community crime: Examining the importance of neighbor networks. *Criminology*, *35*(4), 677-704.

Bhutta, N., & Hizmo, A. (2021). Do minorities pay more for mortgages?. *The Review of Financial Studies*, *34*(2), 763-789.

Bianchi, S. M., Farley, R., & Spain, D. (1982). Racial inequalities in housing: An examination of recent trends. *Demography*, *19*(1), 37-51.

Blau, F. D., & Graham, J. W. (1990). Black-white differences in wealth and asset composition. *The Quarterly Journal of Economics*, *105*(2), 321-339.

Campbell, J. Y. (2006). Household finance. *The Journal of Finance*, *61*(4), 1553-1604.

Charles, C. Z. (2006). *Won't you be my neighbor: Race, class, and residence in Los Angeles*. Russell Sage Foundation.

Charles, K. K., & Hurst, E. (2002). The transition to home ownership and the black-white wealth gap. *Review of Economics and Statistics*, *84*(2), 281-297.

Choi, J. H., McCargo, A., Neal, M., Goodman, L., & Young, C. (2019). Explaining the Black-White homeownership gap. *Washington, DC: Urban Institute. Retrieved March*, *25*, 2021.

Coulson, N. E., Hwang, S. J., & Imai, S. (2002). The value of owner occupation in neighborhoods. *Journal of Housing Research*, *13*(2), 153-174.

Dupond, S. (2019). A thorough review on the current advance of neural network structures. *Annual Reviews in Control*, *14*, 200-230.

Ellen, I. G. (2001). *Sharing America's neighborhoods*. Harvard University Press.

Faber, J. W., & Ellen, I. G. (2016). Race and the housing cycle: Differences in home equity trends among long-term homeowners. *Housing Policy Debate*, *26*(3), 456-473.

Fischer, M. J. (2003). The relative importance of income and race in determining residential outcomes in US urban areas, 1970-2000. *Urban Affairs Review*, *38*(5), 669-696.

Flippen, C. (2004). Unequal returns to housing investments? A study of real housing appreciation among black, white, and Hispanic households. *Social Forces*, *82*(4), 1523-1551.

Friedman, J. H. (1999). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189-1232.

Friedman, S., & Rosenbaum, E. (2004). Nativity status and racial/ethnic differences in access to quality housing: Does homeownership bring greater parity?. *Housing Policy Debate*, *15*(4), 865-901.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2020). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance*, *77*, 5-47.

Ganong, P., & Noel, P. (2020). Liquidity versus wealth in household debt obligations: Evidence from housing policy in the Great Recession. *American Economic Review*, *110*(10), 3100-3138.

Gerardi, K., Willen, P., & Zhang, D. H. (2020). Mortgage prepayment, race, and monetary policy. *Federal Reserve Bank of Boston Research Department Working Papers No. 20-7*.

Gittleman, M., & Wolff, E. N. (2004). Racial differences in patterns of wealth accumulation. *Journal of Human Resources*, *39*(1), 193-227.

Hamilton, D., & Darity Jr., W. (2010). Can 'baby bonds' eliminate the racial wealth gap in putative post-racial America?. *The Review of Black Political Economy*, *37*(3-4), 207-216.

Hardy, B., Morduch, J., Darity Jr., W., & Hamilton, D. (2018). Wealth inequality, income volatility, and race. *Working Paper*.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

Hipp, J. R. (2007). Block, tract, and levels of aggregation: Neighborhood structure and crime and disorder as a case in point. *American Sociological Review*, *72*(5), 659-680.

Hipp, J. R., Tita, G. E., & Greenbaum, R. T. (2009). Drive-bys and trade-ups: Examining the directionality of the crime and residential instability relationship. *Social Forces*, *87*(4), 1778-1812.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8), 832-844.

Iceland, J., & Wilkes, R. (2006). Does socioeconomic status matter? Race, class, and residential segregation. *Social Problems*, *53*(2), 248-273.

Ihlanfeldt, K., & Mayock, T. (2009). Price discrimination in the housing market. *Journal of Urban Economics*, *66*(2), 125-140.

Ishii-Kuntz, M., Makino, K., Kato, K., & Tsuchiya, M. (2004). Japanese fathers of preschoolers and their involvement in child care. *Journal of Marriage and Family*, *66*(3), 779-791.

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M., & Taylor, A. M. (2019). The rate of return on everything, 1870–2015. *The Quarterly Journal of Economics*, *134*(3), 1225-1298.

Kahn, M. E. (2021). *Racial and Ethnic Differences in the Financial Returns to Home Purchases from 2007 to 2020* (No. w28759). National Bureau of Economic Research.

Kermani, A., & Wong, F. (2021). *Racial Disparities in Housing Returns* (No. w29306). National Bureau of Economic Research.

Kuhn, M., & Ploj, G. (2020). Job stability, earnings dynamics, and life-cycle savings. *IZA Discussion Paper No. 13887*.

Kuhn, M., Schularick, M., & Steins, U. I. (2020). Income and wealth inequality in America, 1949–2016. *Journal of Political Economy*, *128*(9), 3469-3519.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324.

Lewis, V. A., Emerson, M. O., & Klineberg, S. L. (2011). Who we'll live with: Neighborhood racial composition preferences of whites, blacks and Latinos. *Social Forces*, *89*(4), 1385-1407.

Logan, J. R. (2011). Separate and unequal: The neighborhood gap for blacks, Hispanics and Asians in Metropolitan America. *Project US2010 Report*, 1-22.

Logan, J. R., Stults, B. J., & Farley, R. (2004). Segregation of minorities in the metropolis: Two decades of change. *Demography*, *41*(1), 1-22.

Madsen, H., & Thyregod, P. (2011). *Introduction to general and generalized linear models*. Chapman & Hall/CRCC.

Massey, D. S., & Tannen, J. (2015). A research note on trends in black hypersegregation. *Demography*, *52*(3), 1025-1034.

Myers, C. K. (2004). Discrimination and neighborhood effects: Understanding racial differentials in US housing prices. *Journal of Urban Economics*, *56*(2), 279-302.

Perry, A., Rothwell, J., & Harshbarger, D. (2018). *The devaluation of assets in black neighborhoods: The case of residential property*. Metropolitan Policy Program at Brookings.

Pynoos, J., & Nishita, C. M. (2006). The elderly and a right to housing. In R. G. Bratt, M. E. Stone, & C. Hartman (Eds.), *A right to housing: Foundation for a new social agenda* (pp. 279-295). Temple University Press.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*(4), 1092-1153.

Rendle, S. (2010). Factorization Machines. *2010 IEEE International Conference on Data Mining*, pp. 995-1000. DOI: 10.1109/ICDM.2010.127

Ritter, J. A., & Taylor, L. J. (2011). Racial disparity in unemployment. *The Review of Economics and Statistics*, *93*(1), 30-42.

Ruggles, S. J., Alexander, T., Genadek, K., Goeken, R., Schroeder, M. B., & Sobek, M. (2010). *Integrated public use microdata series: Version 5.0*. Minneapolis: University of Minnesota.

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *The American Journal of Sociology, 94*(4), 774–802.

Sampson, R. J., & Wilson, W. J. (1995). Toward a theory of race, crime, and urban inequality. In M. Vogel (Ed.), *Crime, inequality and the state* (pp. 312-325). Routledge.

Schwartz, A. E., Susin, S., & Voicu, I. (2003). Has falling crime driven New York City's real estate boom?. *Journal of Housing Research*, *14*(1), 101-135.

Shalev-Shwartz, S., & Ben-David, S. (2014). Decision trees. *Understanding machine learning*. Cambridge University Press.

Shapiro, T. M. (2004). *The hidden cost of being African American: How wealth perpetuates inequalities*. Oxford University Press.

Sharp, G., & Iceland, J. (2013). The residential segregation patterns of whites by socioeconomic status, 2000–2011. *Social Science Research*, *42*(4), 1046-1060.

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas.* University of Chicago Press.

Squires, G. D., & Kubrin, C. E. (2006). *Privileged places: Race, residence, and the structure of opportunity*. Lynne Rienner Publishers.

Swaroop, S., & Krysan, M. (2011). The determinants of neighborhood satisfaction: Racial proxy revisited. *Demography*, *48*(3), 1203-1229.

Tita, G. E., Petras, T. L., & Greenbaum, R. T. (2006). Crime and residential choice: A neighborhood level analysis of the impact of crime on housing prices. *Journal of Quantitative Criminology*, *22*(4), 299-317.

Wang, K., & Immergluck, D. (2019). Housing vacancy and urban growth: Explaining changes in long-term vacancy after the US foreclosure crisis. *Journal of Housing and the Built Environment*, *34*(2), 511-532.

Warner, B. D., & Pierce, G. L. (1993). Reexamining social disorganization theory using calls to the police as a measure of crime. *Criminology*, *31*(4), 493-517.

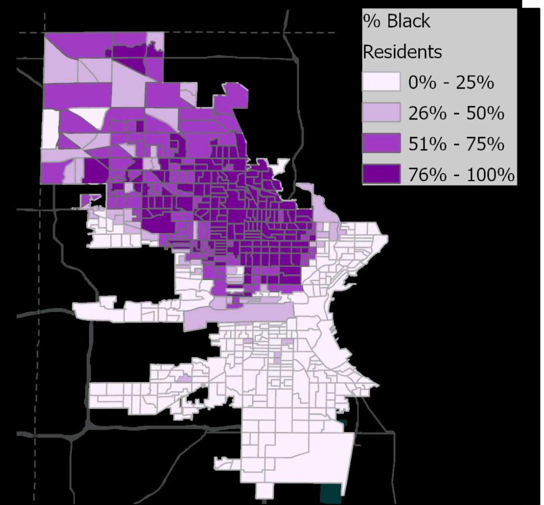
Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology*, *53*(2), 133-157.

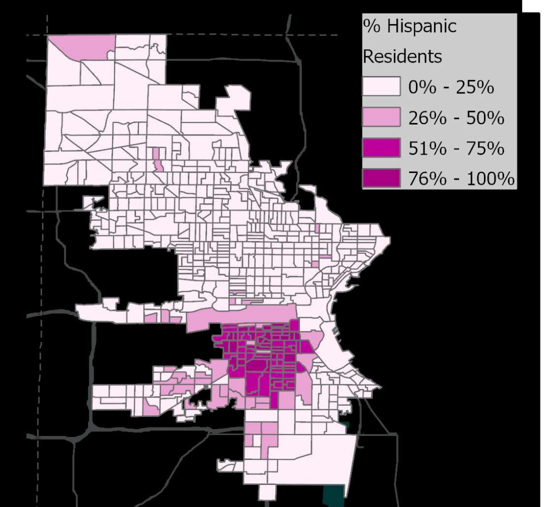
Wilson, W. J. (2012). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

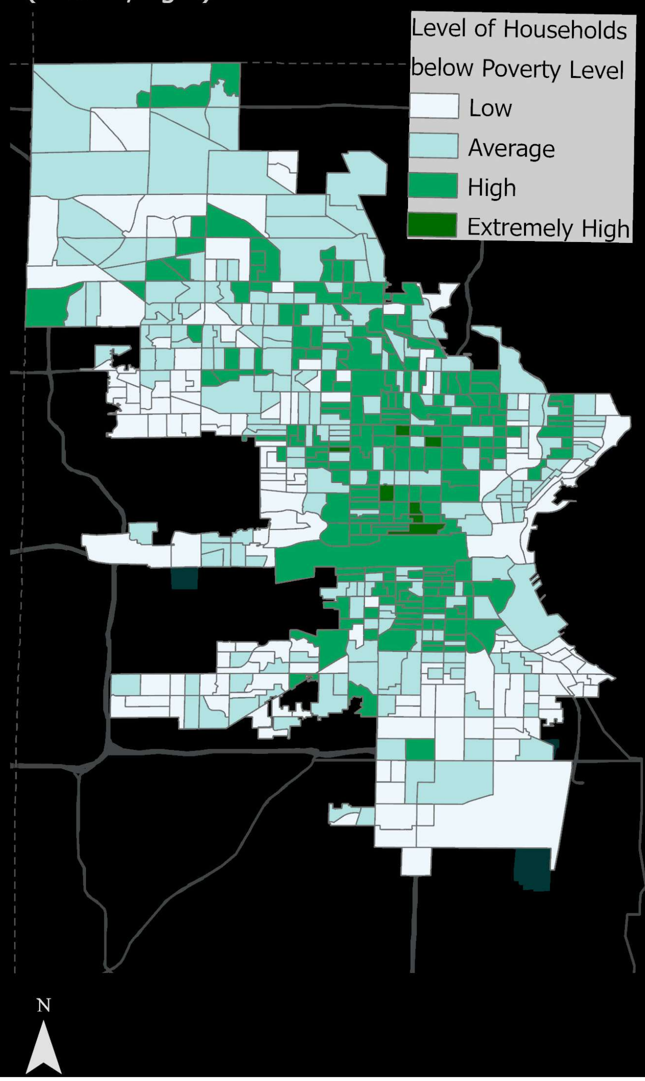
Wrigley-Field, E., & Seltzer, N. (2020). Unequally insecure: Rising black/white disparities in job displacement, 1981–2017. *Washington Center for Equitable Growth Working Paper Series*.

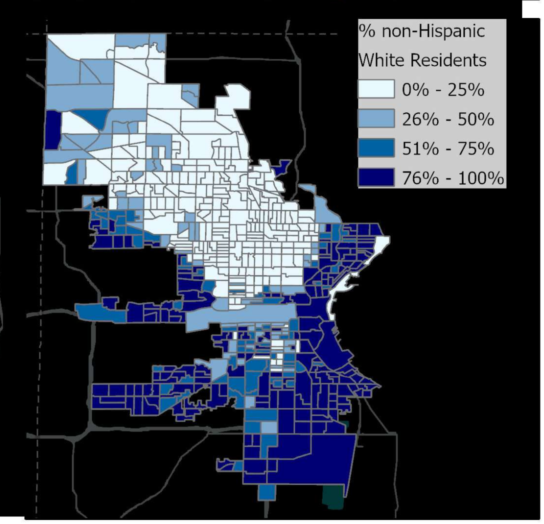
Milwaukee, Wisconsin

Based on standard deviation breaks, this product shows the distribution of the percentage of residents with annual earning below the poverty line across block groups (bottom, left). Additionally, this product shows the percent of block group residents who identify as Black top, right), the percent of block group residents who identify as Hispanic (center, right), and the percent of block group residents who identify as non-Hispanic White (bottom, right).



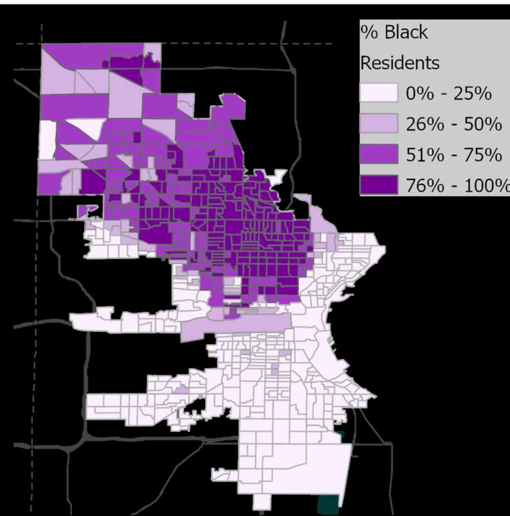


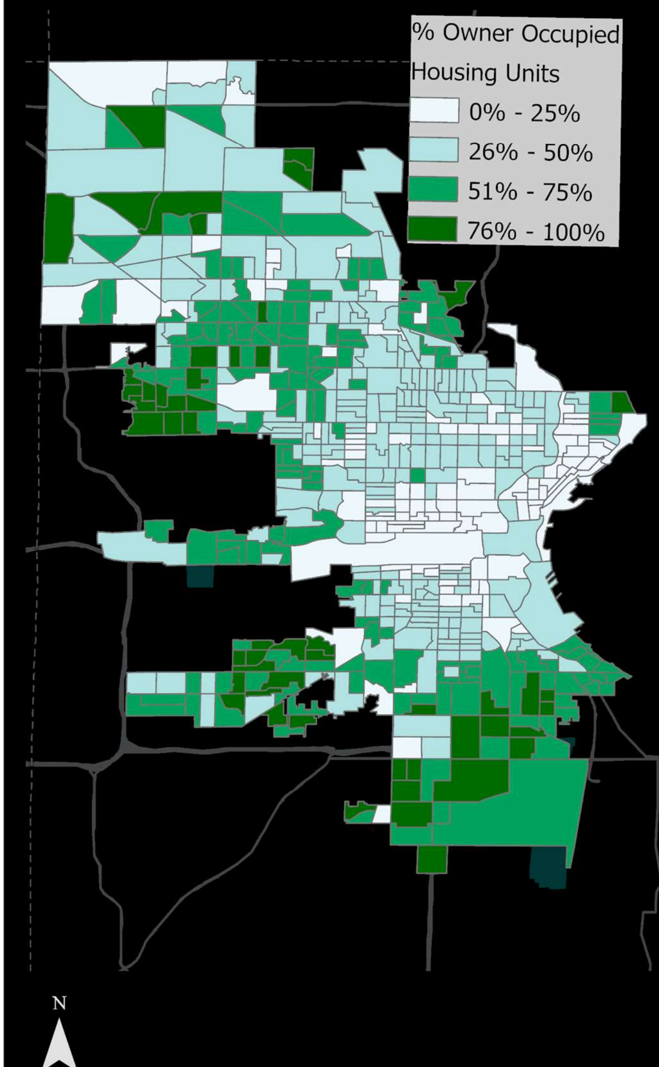


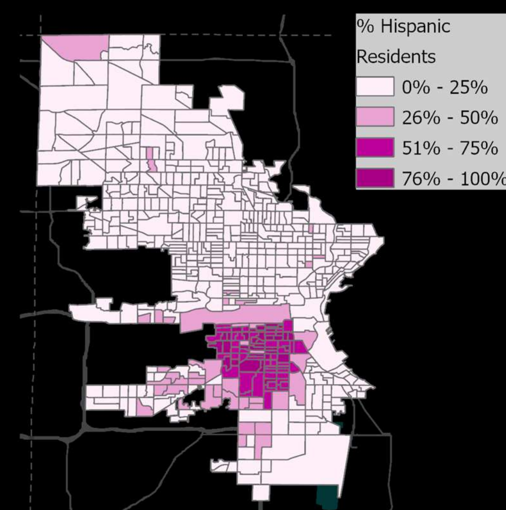


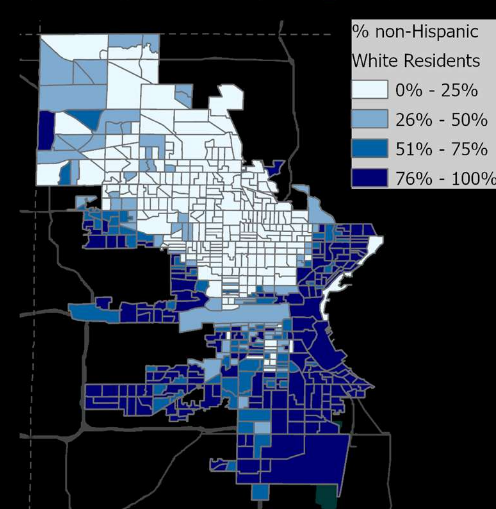
Milwaukee, Wisconsin

Based on equal interval breaks, this product shows the distribution of the percentage of owner occupied housing units across block groups (bottom, left). Additionally, this product shows the percent of block group residents who identify as Black top, right), the percent of block group residents who identify as Hispanic (center, right), and the percent of block group residents who identify as non-Hispanic White (bottom, right).









Milwaukee, Wisconsin

Based on standard deviation breaks, this product shows the distribution of block group median home value (bottom, left). Additionally, this product shows the percent of block group residents who identify as Black top, right), the percent of block group residents who identify as Hispanic (center, right), and the percent of block group residents who identify as non-Hispanic White (bottom, right).