

NBER WORKING PAPER SERIES

RACIAL SEGREGATION IN HOUSING MARKETS AND THE EROSION OF BLACK WEALTH

Prottoy A. Akbar
Sijie Li
Allison Shertzer
Randall P. Walsh

Working Paper 25805
<http://www.nber.org/papers/w25805>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2019

Support for this research was provided by the National Science Foundation (SES-1459847). We thank Werner Troesken, Arie Beresteneau, Dan Bogart, Ed Coulson, William Even, Thomas Davidoff, Leah Boustan, Hoyt Bleakley, Bill Collins, Ross Mattheis, Christian Dippel and seminar participants at the NBER (DAE), Arizona, UC-Irvine, Miami of Ohio, Pitt, Harvard, the Philadelphia Fed, Tennessee, Arizona State, Colorado, and Illinois for helpful comments. We thank John Logan for assistance with enumeration district mapping and Carlos Villarreal and the Union Army Project (www.uadata.org) for the 1930 street files. Antonio Diaz-Guy, Phil Wetzel, Jeremy Brown, Andrew O'Rourke, Aly Caito, Loleta Lee, Zach Gozlan, and Sai Konduru provided outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Prottoy A. Akbar, Sijie Li, Allison Shertzer, and Randall P. Walsh. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Racial Segregation in Housing Markets and the Erosion of Black Wealth
Prottoy A. Akbar, Sijie Li, Allison Shertzer, and Randall P. Walsh
NBER Working Paper No. 25805
May 2019, Revised January 2020
JEL No. J15,N12,R31

ABSTRACT

Housing is the most important asset for the vast majority of American households and a key driver of racial disparities in wealth. This paper studies how residential segregation by race eroded black wealth in prewar urban areas. Using a novel sample of matched addresses from prewar American cities, we find that over a single decade rental prices soared by roughly 50 percent on city blocks that transitioned from all white to majority black. Meanwhile, pioneering black families paid a 28 percent premium to buy a home on a majority white block. These homes then lost 10 percent of their original value as the block became majority black. These findings strongly suggest that segregated housing markets cost black families much of the gains associated with migrating to the North.

Prottoy A. Akbar
University of Pittsburgh
Dietrich School of Arts and Sciences
Department of Economics
4901 Wesley W. Posvar Hall
230 South Bouquet Street
Pittsburgh, PA 15260
prottoyamanakbar@pitt.edu

Allison Shertzer
Department of Economics
University of Pittsburgh
4901 WW Posvar Hall
230 South Bouquet Street
Pittsburgh, PA 15260
and NBER
shertzer@pitt.edu

Sijie Li
Department of Economics
University of Pittsburgh
4519 Wesley W. Posvar Hall
Pittsburgh, PA 15260
sijie.li@pitt.edu

Randall P. Walsh
Department of Economics
University of Pittsburgh
4901 WW Posvar Hall
230 S. Bouquet St.
Pittsburgh, PA 15260
and NBER
walshr@pitt.edu

“Daisy and Bill Myers, the first black family to move into Levittown, Pennsylvania, were greeted with protests and a burning cross. A neighbor who opposed the family said that Bill Myers was ‘probably a nice guy, but every time I look at him I see \$2,000 drop off the value of my house.’ ”

- Ta-Nehisi Coates, *We Were Eight Years in Power: An American Tragedy* (2017)

“During the early nineteen twenties it is estimated that more than 200,000 Negroes migrated to Harlem... It was a typical slum and tenement area little different from many others in New York except for the fact that in Harlem rents were higher... Before Negroes inhabited them, they could be let for virtually a song. Afterwards, however, they brought handsome incomes.”

- Frank Boyd, American Life Histories Manuscripts (WPA Federal Writers' Project, 1938)

I. Introduction

The Great Migration – which saw millions of African Americans depart the Jim Crow South for northern cities – is a key channel through which black families sought to improve their economic standing in the middle decades of the twentieth century. Moving to the North was associated with increased wages and improved occupational status (Myrdal 1944, Collins and Wanamaker 2014, Boustan 2016). Yet these earnings gains failed to close the racial wealth gap, which persisted and, in some cases, worsened over the ensuing decades (Blau and Graham 1990; Bound and Freeman 1992). To explain the persistence of these disparities, economists have long highlighted the role of pervasive discrimination in the labor market and in the educational system in preventing black families from accumulating wealth at the same rate as whites (Smith and Welch 1989, Collins and Wanamaker 2017, Bayer and Charles 2018).

Recent scholarship has also highlighted the role of discriminatory government policies in supporting residential segregation by race and disadvantaging black wealth accumulation through home ownership (Rothstein 2017). While there is a clear consensus that real estate markets and housing policy were integral in fostering the disadvantage faced by black families, we know surprisingly little about the specific channels through which segregated housing

markets served to erode black wealth. Closing this gap is particularly relevant to our understanding of how residential segregation worsened and solidified the racial wealth disparities in American cities during the Great Migration.

The extant lack of empirical work stems largely from the fact that urban segregation during this time period was a block-level phenomenon and researchers have historically lacked data linking housing prices and demographic characteristics at the fine level of spatial detail needed to explore these dynamics. In this paper we introduce two major data innovations to overcome these limitations. First, using the full-count censuses of 1930 and 1940, we create detailed demographic data, including housing values and rents, at the city-block level for ten major northern cities. These two censuses were the first to ask about home values and rents, and they are also free of confidentiality restrictions, enabling us to observe the address for the universe of individual census records in each of our ten cities.² As a result, we can geocode the vast majority of these blocks. Our second data innovation is to match addresses across these two censuses. This matching allows us to track the evolution of an individual home's price and its exposure to city-block-level changes in racial composition during the Great Depression, a decade when black neighborhoods saw major expansions.

Our analysis begins by providing the first description of residential segregation by race at the city-block level in prewar American cities. We find high degrees of concentration, with the percent of black families living on blocks that were virtually all black (> 90 percent) increasing from 52 percent in 1930 to 63 percent at the end of the decade. Racial transition was also central to the black experience in this decade. While only 4 percent of black families lived on blocks that were predominantly white (> 75 percent) in 1930, fully 13 percent lived on blocks that had been

² There is no systematic, address-level housing price data available for 1920 or earlier years. The 1950 census is not yet digitized and publicly available.

predominantly white in 1930. The price dynamics on transitioning blocks represent an important component of black housing during the decade. Importantly, these blocks can be used in conjunction with our new linked sample to recover estimates of the causal impact of racial segregation and white flight on the housing prices and rents faced by black households more broadly.

Using repeated observations of addresses that were occupied by white homeowners at the start of the decade, we show that black families that were renters paid a premium of roughly 50 percent relative to white families renting in equivalent housing on blocks that did not transition. Occupancy soared on these blocks as well, increasing by roughly 47 percent. While price dynamics differed, home ownership provided black families no escape. To this point, we find that the first black arrivals on a newly transitioning block were much more likely to buy their home than to rent. To induce incumbent white owners to sell to a black family, these pioneers paid a premium of roughly 28 percent relative to the prices that white homeowners were paying on the same block. Moreover, once these early pioneers had locked in their ownership at an inflated price, we find that home values declined significantly throughout the transition process. By the time a block had transitioned to majority black, homes had on average lost nearly 10 percent of their original (no premium) value.

To better understand the divergence between prices and rents, we propose a simple no-arbitrage condition to fix the relationship between rents and home prices (Kearl 1979; Poterba 1992). Through this lens, investors demanded higher rents to compensate for declines over time in the price of their rental properties that were anticipated as a result of racial transition. While one potential driver of the observed rent increases is the increase in occupancy rates that occurred during the racial transition process, we demonstrate that crowding cannot explain the

majority of the key price changes. We also find that landlords provided steep discounts to white renters who stayed on blocks during the process of transition. We argue that the presence of these rental discounts for remaining white families is direct evidence of landlord discrimination. These rental markets are particularly important given that 86 percent of blacks on fully transitioned blocks were renters in our sample.

While ours is the first work of which we are aware to document that segregated housing markets led to both *elevated* rents and *declining* home values for black households,³ these findings are consistent with the historical record of the period. Real estate historians have argued that the urban color line moved because black families who demanded better quality housing outbid whites for the purchase of homes in neighborhoods just outside of the established ghetto (Mehlhorn 1998, Troesken and Walsh 2019). In response to these new black arrivals, and at least in part compelled by concerns about falling home values and the quality of public services, white households subsequently fled these transitioning areas (Boustan 2010, Derenoncourt 2018, Shertzer and Walsh 2019). These transitions were then associated with an increase in absentee landlords, as many former residents either rented out their home or sold it to a (white) investor.⁴

In total, our investigation documents that segregation and the process of black neighborhood expansion left African Americans both living in declining neighborhoods and

³ The consensus in the literature is that segregation that arises from constraints on black housing supply will result in black families paying higher prices for similar housing relative to whites. Indeed, most papers that examine racial housing price disparities between 1940 and 1970 have argued that blacks paid such a premium (King and Mieszkowski 1973; Yinger 1978; Schafter 1979). The passage of the Fair Housing Act in 1968 reduced the tools available to white families to maintain the color line, and most papers working with data from after 1970 argue that segregation was maintained by whites paying a premium to avoid black neighbors (Follain and Malpezzi 1981; Chambers 1992). Using the interaction between black household and measured racial segregation in a particular city, Cutler Glaeser and Vigdor conclude that blacks paid a premium in the 1940s and whites a premium by the 1990s.

⁴ See for instance United States Congress House Committee on the District of Columbia, 1935, *Rent Commission: Hearings before the subcommittee on Fiscal Affairs on H.R. 3809*, p. 7. The investors are described as follows: “It is a certain class of individuals in a great many cases that buys up these properties and gets as much out of them as they possibly can until the properties are condemned or fall down or are converted to some other use... In a great many other cases [the houses] have been in the family for years, and the family does not know how to get rid of it, so they just keep renting the house.”

doubly poorer. Applying our estimates to the modal black renter in our sample of northern cities, we find that the price shifts associated with racial transition erased roughly forty percent of the annual income gain associated with moving from the South to the North. This calculation suggests that a significant share of cumulative gains in occupational standing and earnings achieved by black families who migrated to the North were canceled out by the market dynamics associated with segregated housing markets. Segregation thus limited the degree to which black Americans could move to opportunity over the course of the Great Migration.

II. Historical Background

The Great Migration saw millions of African Americans leave the poverty and oppression of the Jim Crow South for better lives in northern cities. However, they soon discovered that the North maintained its own system of racial segregation, particularly in housing markets. Black families found themselves largely restricted to homes in existing black neighborhoods through a mixture of threats, actual violence, and discriminatory real estate practices. The narrative history emphasizes collective action taken by whites to maintain the color line, which shifted over time from angry mobs in the early days of the Great Migration to the later establishment of genteel neighborhood “improvement” associations (Massey and Denton, 1993). Such associations were created in part to lower the costs of adopting restrictive covenants, which were deed provisions prohibiting the sale of a house to a black family. Such covenants had effect until 1948 when the Supreme Court struck down their enforcement in *Shelley v. Kraemer*.

Still, the color line was not inviolate. The 1920s and 1930s saw significant expansions of existing black neighborhoods in most northern cities. Urban historians underscore the desperation of black families for better housing and their tendency to outbid whites for homes

near the ghetto. At the same time, real estate professionals and academics were united in their belief that black entry would harm home values.⁵ Such expectations made banks reluctant to underwrite a mortgage for a “pioneer” black family entering a white neighborhood where the lending institution already held loans. One urban historian summarized the dichotomy thusly:

“One of the most interesting points made in the [real estate] broker comments is the recurring theme that while sellers may not get their price from whites (who are reluctant to consider an area undergoing racial transition), they probably can from nonwhites. This is quite different from the unqualified prediction that all prices in an ‘invaded’ area fall” (Laurenti 1960, p. 20).

The fact that black neighborhoods expanded even though black families on average had fewer assets to use for a down payment suggests that some banks did in fact underwrite mortgages for them. While banks were typically reluctant to initiate racial transition on a block, they appear willing to have made loans in neighborhoods likely to transition. Surveys of real estate brokers from the period suggest that the first family to enter a white neighborhood often sought a mortgage from a distant bank that did not have exposure to the area in question (Schietinger 1953, p. 172). The narrative history on the issue of mortgage terms is mixed, with some surveys finding blacks and whites received similar terms (Rapkin and Grigsby 1960, p. 77) and other scholars arguing that African American borrowers were steered towards installment contracts where they could lose possession of their home if they were late on a single payment (Satter 2009, p. 4).

Government policies also influenced black families’ ability to finance home purchases. Beginning in 1934, at the height of the Depression, the Federal Housing Authority initiated underwriting mortgages and imposed policies that would disadvantage the low income and transitioning central-city neighborhoods where black families were likely to buy. However, FHA

⁵ Some social scientists had a more nuanced view of the process. For instance, Gunnar Myrdal argued in *An American Dilemma* that white racism was the primary cause of drops in home values as a block began transitioning and that prices should recover once the neighborhood was majority black (p. 623).

underwriting was still a nascent process during our sample period, particularly so in the neighborhoods that we study. As of the end of 1940, the FHA had underwritten only 60,339 mortgages on existing homes across the entire metropolitan areas of the cities we study in this paper (our analysis is limited to existing homes).⁶ Further, federal urban renewal policies did not begin until the 1949 Housing Act (Collins and Shester, 2013; LaVoice, 2019). It is thus unlikely that federal housing or lending policies can explain our findings.

Of course, not all black families bought their home; in fact, the majority were renters. As we discuss below, we find that the proportion of renters increased throughout the transition process. The question of who owned properties rented to black families is thus important for interpreting our results. The census does not allow us to observe the identity of property owners in the case where the occupants are renters. We thus turn to the narrative history, which suggests white investors purchased properties in the black ghetto with the perhaps self-fulfilling expectation that their investment would sharply depreciate over time. Real estate brokers believed that houses that were converted to multi-family rentals would lose value over time and were generally unwilling to make loans for the purchase of such properties (McEntire 1960, ch. xiii). It would thus be necessary to buy these properties with cash. It is also likely the case that some landlords were former homeowners who decided to convert the house into a rental property instead of selling. Both considerations underscore the fact that in our setting the owners of rental properties were most likely white.

⁶ This figure comes from the FHA's Annual report for 1940 (FHA, 1941). We have been unable to identify exactly how large the metropolitan areas were for this reported data. However, as an example, the FHA reported more homes insured in the New York City Metropolitan area than it reported for the entire state of New York, suggesting that they used broad metropolitan area definitions. Thus, this number should likely be viewed as a very conservative upper bound and FHA penetration into our city neighborhoods would still have been quite limited as of 1940 (likely representing well fewer than 2 percent of the homes in our sample).

III. Data

For this paper we constructed a novel dataset composed of the universe of addresses in ten major cities matched across the 1930 and 1940 censuses. The sample cities are Baltimore, Boston, the Brooklyn and Manhattan boroughs of New York, Chicago, Cincinnati, Cleveland, Detroit, Philadelphia, Pittsburgh, and St. Louis. To create the set of addresses matched over time, we developed an algorithm designed to ensure that only true address matches are included in our sample and to prevent the inclusion of buildings for which we can't be certain about their occupants. Our basic approach is as follows:

1. We first assign every individual living in one of our sample cities in either 1930 or 1940 to an address that is consistent across all household members. If an address is missing, we impute it using another member of the household (households with inconsistent addresses are dropped).
2. We standardize street names to deal with variations of directional prefixes and typical suffixes ("First" vs. "1st", "st" vs. "Street"). We cross-reference street names using a digitized street file for each city: if there is no corresponding street name in the neighborhood in the digitized data, we drop everyone with an address on that street from the census data.
3. We conduct a series of consistency checks to identify the types of errors and omissions that are common in the address field, including making sure neighbors on the same street have street numbers that change monotonically as we move down a manuscript page.
4. We retain only observations on streets that pass our quality checks and have no address inconsistencies.
5. We merge across the 1930 and 1940 census using standardized street names and house numbers, yielding a sample of both single-family homes and apartment buildings.

Our algorithm is conservative in that we discard everyone associated with a particular address *and* everyone associated with an adjacent address on the manuscript when there is a potential problem with the census data, minimizing the risk of missing true occupants of a particular address in our final dataset. Because we wish to examine both occupancy rates and

prices in our matched sample, developing an accurate count of household members is essential. Further details of the address data construction can be found in the Appendices I and II. Our final sample contains 591,780 unique addresses that could be located in both 1930 and 1940 from about 100,000 city-blocks across the sample cities (see Appendix Tables I and II).⁷ We have on average 10 to 15 addresses per city-block in the unmatched data, depending on the city. We compare addresses that could be cleaned and matched to the universe of addresses in Appendix Table III. There is some evidence that addresses with fewer occupants were more likely to be matched although the differences are economically small.⁸

We aggregate households in addresses with multiple units to obtain aggregate rents and occupancy. Addresses that report both an owner and a renter are dropped from the sample. One potential concern with using self-reported valuations as a measure of home prices is accuracy. To verify the accuracy of our price data, for a sample of homes in Pittsburgh that sold in 1930 or 1940, we identified the corresponding address in the county Recorder of Deeds office and obtained the actual sales price. We plot the differential between the census valuation and the actual sales price in Appendix Figure I. The figure suggests that there is no systematic bias.

In previous work, we constructed fine-grained, spatially-identified demographic data for neighborhoods in ten of the largest northern cities for 1900, 1910, 1920, and 1930 (Shertzer, Walsh, and Logan 2016). We are thus able to measure a relatively broad set of baseline neighborhood characteristics at a small unit of geography, specifically at the level of the 1930 census enumeration district (typically around four city blocks in urban areas). Using our address

⁷ To obtain the final address-level dataset, we trim outliers that are likely transcription errors or records associated with institutionalized individuals. In particular, we drop any households with more than 10 members, any household with more than three heads, any addresses with monthly rent greater than \$100, and any addresses with a value greater than \$20,000.

⁸ For instance, there were 7.51 individuals per address in the universe of addresses compared with 6.81 individuals on average in our matched addresses. Because of the large sample size, nearly every difference in Appendix Table III is statistically significant.

data, we are further able to measure racial composition and other key variables at the city-block level. Blocks are delineated using postal service convention with street number intervals in the hundreds.

Figure 1 summarizes the block-level racial patterns in this newly constructed data. Panel A illustrates the distribution of percent black on the blocks where black families lived in 1930. Panel B presents the same data for 1940. These distributions document extremely high, and increasing, levels of segregation at the city-block level.⁹ By 1940, 63 percent of all black families lived on a block that was more than 90 percent black and more than four out of every five black families lived on a block that was at least 75 percent black. Conversely, only 4 percent of black families lived on blocks that were greater than 75 percent white. Concurrent with the increase in segregation, the number of black households in our sample increased by 16 percent. Much of this growth in households was facilitated by the racial transition of previously all white blocks. As is illustrated in Panel C, nearly 8 percent of black households in 1940 lived on a city block that was less than 10 percent white in 1930.

These newly transitioned blocks are at the core of our broader identification strategy which focuses on a sample of single-family owner-occupied homes located on blocks that were all white in 1930. We present summary statistics for this sample in Table 1, subdividing the sample by whether the block had begun undergoing racial transition or not in 1940 (defined as having any black population in 1940). We first note the enormous drop in nominal home prices that accompanied the Great Depression, with homes on all blocks losing about 40 percent of their value between 1930 and 1940. Blocks that transitioned were located in neighborhoods closer to

⁹ Between 1930 and 1940 city-block level isolation (dissimilarity) indices increased from 80% (95%) to 85% (96%), calculated at the household level based on the matched address data set.

existing black neighborhoods in 1930, but other differences in neighborhood characteristics are relatively small.

Homes on blocks that transitioned were actually slightly more expensive on average in 1930, a finding we explore in more detail below. Average rents on these blocks for homes that switched to rentals in 1940 were higher relative to homes on blocks that remained white (\$39.30 versus \$35.10, respectively). Finally, homes on blocks that transitioned gained more occupants relative to homes on blocks that remained white (1.18 versus 1.06, respectively).

IV. Semi-Parametric Analysis

We begin with a discussion of the overall patterns in our data, Figures 2 and 3 present the semiparametric relationship between racial transition and rents and home prices estimated using the Robinson's double residual method (Robinson, 1988). The figures are based on our baseline matched sample of homes that were single family, owner occupied and located on a block that was all white in 1930. They visualize the non-parametric relationship between the level of racial transition as of 1940 (horizontal axis) and rent or price in 1940 (vertical axis), controlling parametrically for a full set of controls including the home's value in 1930.¹⁰ We begin in Panel A of Figure 2 by showing the relationship between 1940 black share and the log rent for all houses that had switched to being rentals by 1940.

Rental prices are relatively flat at low levels of racial transition process. However, prices increase rapidly after a block attains majority black status. In total, Figure 2 suggests that rents increased by 25 percent over the course of racial transition. In Panel B, we decompose rents

¹⁰ That is, we estimate $\ln Price_{i40} = X_i' \beta + f(Black\ share_{i40}) + \epsilon_i$ where X_i includes controls for the log of 1930 price, occupancy at the address level, share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level.

based on the race of the occupant. The figure shows that while blacks in general always paid a premium relative to whites to rent on the same block, this premium grew significantly along with the level of transition, accelerating its growth rate at approximately 40 percent black. Thus, above this point white renters demanded, and landlords were willing to provide, a significant discount to remain on transitioned blocks. The most direct explanation for these differentials is racial discrimination.

We now turn from rents to valuations of home prices. If racial market dynamics were driven solely by supply restrictions in the market for black housing related to the enforcement of segregated neighborhoods, we would expect the value of owner-occupied homes in black neighborhoods to experience similar increases in valuations upon racial transition. Yet, as is shown in Panel A of Figure 3, overall home values in fully transitioned neighborhoods declined by about 35 log points (about 40 percent). In Panel B we decompose this relationship by the race of the owner. This figure suggests that the overall drop in home values was, in part, driven by the “pioneer” premium paid by black families buying a home on a mostly white block. We estimate that this premium is about 30 log points (35 percent). Further, homes purchased on transitioning blocks lost value throughout the transition process. By the time the block was mostly black, such homes had lost about 11 percent of their *original* value.

To better understand the impact of these price and rent dynamics, Figure 4 documents black homeownership rates over the range of transition. Panel A documents high home ownership rates (low rental rates) among the pioneering black families. Black ownership rates then decrease with transition level, plateauing when blocks become majority black. This pattern indicates that black families were most likely to buy homes when the purchase premium required from them was highest and most likely to rent when the rent premium was highest. Taken

together, these results suggest that there was no escape from the disadvantageous housing market faced by black families. In contrast, there is no trend in ownership for white families (Panel B).

We now turn to a more parametrized analysis. Here, we have two primary goals. First, we wish to better understand the divergence between rents and owner-occupied housing values. Second, we seek to better identify the relationships we document in the semi-parametric analysis and demonstrate that they are causal.

V. Capitalization Framework and Parametric Analysis

Our parametric framework models the relationship between rents, property values, and the racial composition of neighborhoods from the perspective of an arbitraging real estate investor. To fix ideas, we denote the price (rent or own) of an individual building as follows:

$$P_i = \begin{cases} \text{annual rent}_i & \text{if tenure} = r \\ \text{sales value}_i & \text{if tenure} = o \end{cases}$$

For a given owner occupied house, its price in year t is given by:

$$P_{it} = c_t * \rho_t * Q(Z_i) \tag{1}$$

where ρ_t is the city-specific price level at time t, Z_i is a vector of housing and neighborhood characteristics that are particular to the given house, $Q(.)$ is a quantity function that maps these characteristics into a unidimensional measure of service flow, and c_t is a capitalization rate that captures the equilibrium relationship between sales price and annual rent.

Thus, we follow Poterba (1992) in conceptualizing the capitalization rate as follows:

$$\text{rent}_{it} = c_t * \text{sale}_{it} \tag{2}$$

The capitalization rate, c_t can be decomposed as follows:

$$c_t = i + \tau_p + \text{risk} + \text{maintenance} + \text{depreciation} - \text{appreciation}$$

where i is the risk-free interest. Intuitively, the real estate investor must receive a return on her investment equal to the risk-free interest rate available in the broader market place. This risk-free rate of return is adjusted for additional costs and benefits associated with owning the property. In particular: tax benefits or costs associated with owning a home (τ_p), a risk premium associated with housing price uncertainty (*risk*), costs for maintaining the property (*maintenance*), physical depreciation (*depreciation*), and appreciation of the home's value net of the overall inflation rate (*appreciation*) - with all of these terms expressed as percentages of the property's values. We also note that the results presented in Panel B of Figure 2 suggest augmenting the basic model to include an additional term to account for landlord preferences over the race of potential tenants (taste-based discrimination).

This no-arbitrage relationship is central to understanding Figures 2 and 3 that show racial transition being associated with lower sales values and higher rents. While perhaps surprising at first, this dichotomy can be rationalized by investors having exceedingly pessimistic expectations regarding the impact of racial transition on housing price appreciation (expectations of rapidly declining values), physical depreciation or maintenance costs.

To operationalize this relationship, we begin by combining equations (1) and (2) to derive a unified expression for P_{it} :

$$P_{it} = \rho_t * c_t^{I_{rent}} * Q(Z_i) \quad (3)$$

where I_{rent} is an indicator variable which equals 1 if the house is rented. Taking the log of both sides yields the following:

$$\ln P_{it} = \ln \rho_t + \ln c_t * I_{rent} + q(Z_i) \quad (4)$$

where, $q(Z_i) = \ln Q(Z_i)$. In our application, we don't directly observe characteristics Z_i , but we do observe prices in both 1940 and 1930. As we detail below, we can use this information to effectively control for these unobserved characteristics.

Solving the 1930 iteration of equation (4) for $q(Z_i)$ gives: $q(Z_i) = \ln P_{it} - \ln \rho_t - \ln c_t * I_{rent}$. Assuming that Z_i is time invariant, limiting our sample to houses that were owner occupied in 1930 (we relax both of these restrictions later), and substituting this expression into the 1940 version of equation (4) yields the following expression for 1940 prices:

$$\ln P_{i40} = \ln \rho_{40} - \ln \rho_{30} + \ln P_{i30} + \ln c_{t40} * I_{rent40} . \quad (5)$$

Thus, ignoring for the moment neighborhood racial transition, we have the following model:

$$\ln P_{i40} = \alpha + \beta * I_{rent40} + \gamma \ln P_{i30} + \epsilon_i . \quad (6)$$

We can interpret the key coefficients in equation (6) as follows: α is the difference in the (logged) price levels between 1940 and 1930 and β is the logged capitalization rate in 1940. Further, inclusion of the 1930 house price effectively controls for all time-invariant house and neighborhood characteristics.¹¹

To build on this basic empirical specification, we begin by limiting our sample to houses located on city blocks that were all white 1930. We then generate an indicator variable for racial transition (I_{trans_i}). Finally, we add the transition variable and its interaction with the rent indicator to equation (6) yielding our basic specification:

$$\ln P_{i40} = \alpha + \beta_{trans} * I_{trans_i} + \beta_{rent} * I_{rent_i} + \beta_{trans \times rent} * I_{trans \times rent_i} + \gamma \ln P_{i30} + \epsilon_i \quad (7)$$

In this specification, $\exp(\hat{\beta}_{trans})$ provides an estimate of the percent difference in sales prices between blocks that transitioned and those that did not. Further, $\exp(\hat{\beta}_{trans} + \hat{\beta}_{trans \times rent})$

¹¹ One could restrict the coefficient γ to be equal to 1. However, not doing so allows for the possibility that price deflation between 1930 and 1940 varied across the distribution of housing quantities.

provides an estimate of the percent difference in rental prices across transitioning and non-transitioning blocks.¹²

One potential concern is that certain characteristics of houses (or their neighborhoods) might change in systematic ways between 1930 and 1940. We control for this possibility in two separate ways. First, we directly include controls for a number of 1930 characteristics at the address, block, and neighborhood level that may be predictive of these systematic changes. Specifically, we control for the occupancy at the address level, share renters and total number of addresses at the city-block level, and at the neighborhood level we control for share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score. Second, we drop the neighborhood-level controls (keeping the house and city-block-level controls) and instead include ED-level fixed effects. These fixed effects will absorb any time changing characteristics that are shared at the ED-level (recall that EDs in our sample are typically approximately four city blocks).

A. Baseline Results

We begin our parametric analysis by relating changes in block-level racial composition to changes in housing prices over the 1930s. For our baseline specification, we consider the impact of city block-level racial change as measured by a variable that equals 1 if a formerly white block became majority black by 1940 and 0 otherwise. Column (1) of Panel A in Table 2 reports the empirical estimate of equation (7), restricting the sample to single-family, owner-occupied homes and controlling only for price and occupancy in 1930. The second column adds neighborhood-level controls and the third incorporates both neighborhood fixed effects as well as

¹² We also note that $\hat{\beta}_{rent}$ and $\hat{\beta}_{transXrent}$ allow for the recovery of effective capitalization rates in transitioned and un-transitioned neighborhoods.

block-level controls for share renters and number of households in 1930. While results are qualitatively consistent across specifications, the model presented in column (3) is the most robust in terms of controls. We therefore view it as our preferred specification.¹³

The coefficient on the rent indicator (-2.214) reflects the log of the capitalization rate for blocks that did not experience racial transition. It implies a baseline capitalization rate of 10.9 percent. Thus, in white neighborhoods the annual rent that a real estate investor should have expected to receive on a given property was roughly 11 percent of its sales value.¹⁴ The coefficient on the racial transition variable (-.096) implies that houses on blocks that saw an influx of blacks lost 9.1 percent of their value relative to blocks that remained white. Meanwhile, in conjunction with the coefficient on the interaction between rented and transition (.503), this estimate implies that rents on these blocks increased by 50.2 percent relative to non-transitioning blocks. Finally, the estimated capitalization rate for transitioned neighborhoods is 18.1 percent (computed as the exponent of the sum of the rented coefficient and the interaction of transition and rented).

Although we prefer to restrict our attention to single-family, owner-occupied homes for the purpose of identification of the transition effect, we also present results for a larger sample of addresses in column 4. Specifically, we also include buildings that were rented in 1930.¹⁵ Our estimates are quantitatively similar. In column (5) we report results focusing solely on the sample of addresses that were rented in 1930, again finding similar results but with a smaller

¹³ Appendix Figures 4 and 5 demonstrate the general robustness of our analysis to different choices of baseline blocks (i.e. choosing as a baseline all city-blocks that were less than x% black in 1930) and different thresholds for racial transition (i.e. defining as transitioned all city-blocks that were > x% black in 1940).

¹⁴ The Great Depression was associated with substantial housing price deflation which outpaced the concomitant declines in rents, and thus we should expect capitalization rates that are in general larger than those from the current day, which tend to center around 6 percent (see for instance Davis et al. 2008).

¹⁵ This specification requires additional controls for tenure status in 1930. We do not include mixed-tenure or multiple owner addresses in this analysis because it is unclear how to aggregate a mix of valuations or valuations and rents into an address-level price.

rental premium for homes on blocks that underwent racial transition and remained rented in 1940 as well. As we show below in Section V.E. and in Appendix Figure 5, these results are robust to alternative definitions of racial transition.

In all specifications we find that racial transition was accompanied by falling home values, sharply increasing rents, and a substantially higher capitalization rate. As discussed above, the finding that rents and valuations diverged on transitioned blocks, can be rationalized by investors having exceedingly pessimistic expectations regarding housing prices, physical depreciation, or maintenance costs. We note that the channel through which racial preferences impact prices and rents is less direct here, where we are comparing average prices and rents across blocks at *different stages of racial transition*, than it was in panel B of Figure 1 where we compared differences in rent paid by black and white households on blocks at *identical stages of transition*. In the earlier comparison across the race of individual renters, differentials are likely driven by white landlords preferring to rent to white tenants and thus charging blacks higher rents than whites for identical properties. Here we are instead focused on how rents change with block-level transitions, independent of the race of a home's renter or owner. The primary channel through which racial preferences drive cross-block market dynamics on this dimension are more likely white flight and related expectations about future declines in price.

Independent of racial preferences or price expectations, contemporaneous narratives suggest that one channel through which capitalization rates (and thus rents) could have been higher for buildings on blocks undergoing racial transition is through higher occupancy rates, either due to subdividing single-family housing into multiple rental units or as a result of black families taking on boarders to help cover the steep rents that they faced. Managing contracts with

multiple households could have imposed direct costs, while increased occupancy itself could have led to more rapid physical depreciation.

We explore the impact of transition on occupancy rates in panel B of Table 2, which replicates panel A using log of aggregate occupancy as the outcome variable. Occupancy results are generally similar across all models. Houses on un-transitioned blocks that switched from being owned to rented saw increases in their average aggregate occupancy rate of approximately 20 percent. These occupancy rate increases were even larger when the move to rental status was associated with racial transition. The estimates from column (3) indicate that rental occupancy soared by 47 percent in such homes relative to homes that remained owned on blocks that did not transition.¹⁶ For owner occupied housing, the main effect of racial transition is very small or negative in all specifications (-0.032 and statistically insignificant in our preferred specification). This finding is consistent with the narrative evidence that higher-socioeconomic-status black families, who would not need to bring on boarders or live in subdivided units to afford housing, were the first to arrive on a transitioning block and bought their homes rather than renting them (e.g. Massey and Denton, 1993).

Our occupancy results raise the possibility that the observed increase in capitalization rates on transitioned blocks, and the associated rent spikes, could simply be the direct result of increases in maintenance or physical depreciation costs arising from higher-density habitation. To examine this issue directly, in Table 3 we consider how capitalization rates varied with both occupancy and racial transition, augmenting our baseline log-price specification to consider multiple levels of racial transition and splitting the sample between houses that experienced increased occupancy rates and those that experienced decreased or unchanged occupancy rates.

¹⁶ That is, $\exp(.184-.032+.233)=1.47$.

The results suggest that while occupancy rates had a small impact on capitalization rates on blocks that remained white, the magnitudes are too small to explain the bulk of the rent hikes experienced in transitioning blocks. In Table 4, we go further and focus on subsamples comprised only of addresses that gained between one and four members located on blocks that remained white between 1930 and 1940. Even for addresses that gained at least four members, a very large occupancy increase, the coefficient estimates imply capitalization rates that never exceed 11.5 percent. We thus reject the notion that our results are driven solely by occupation rates.

B. Discriminatory Premiums in the Housing Market

We next seek to estimate the premiums required of black families in owned and rented housing markets suggested by the semiparametric results from Section IV. Recall that black families appeared to pay a premium to buy a house on a white block at the earliest stages of transition (Figure 3.B). To estimate this premium, we modify our baseline specification (7) by adding an indicator for a black household and interacting this indicator with rental status. We also drop the indicator for racial transition and instead restrict analysis to blocks that experienced low levels of racial transition (black share in 1940). Column 1 (column 2) of Table 5 presents results for the sample of blocks that were all white in 1930 and less than five percent (ten percent) black in 1940. The highly significant point estimate on black household of 0.32 (0.25) translates to pioneering black families paying a 38 percent (28 percent) discriminatory premium. It is unlikely that these premiums reflect simple differences in perceptions since the black homebuyers would have just recently purchased their homes in white neighborhoods.

Turning to the rental market, Figure 2.B suggests that white families received a large discount if they remained in their rental units during the racial transition process, particularly

once the block experienced marked racial mixing. To estimate these premia, we again rerun our modified specification but for blocks where racial transition was well underway. Column 3 (column 4) shows the results for blocks that were at least 40 percent (60 percent) black in 1940. The coefficient estimates suggest that on such blocks black renters paid 34 percent (44 percent) more than white renters for identical housing. Again, the most direct explanation for these premiums is racial discrimination on the part of landlords. We discuss the implications of the rental and homeownership premiums faced by black households in Section VI.

C. Selection

One potential concern with our empirical results is the possibility that neighborhoods that were already destined to experience declining values (or higher rents) were differentially targeted for racial expansion, even after controlling for price in 1930. Perhaps most concerning is the role played by proximity to existing black neighborhoods. Our data clearly document that proximity to an existing black neighborhood was a strong predictor of racial transition. If these proximate neighborhoods were also destined to see systematic departures from price trends, for instance because of reduced city services or other forms of disinvestment, our results could be biased. The inclusion of enumeration district (ED) fixed effects in our preferred specification is largely a response to this concern as they will control for all factors affecting prices that are constant over very small neighborhood definitions. However, it is still possible that even differences in black neighborhood proximity across a few city blocks could lead to selection problems.

As a first test of our fixed effects strategy we evaluate the effectiveness of using enumeration district fixed effects to absorb control for the correlation between 1930 demographic measures and racial transition over the following decade. Table 6 presents the results of a block-level estimation of the determinants of racial transition for blocks that had at

least one owner-occupied single-family home and were all white in 1930. Columns (1) and (4) present regressions on the entire sample that include ED fixed effects in addition to controls for household head age, share laborer, foreign-born share, average rent per person and homeownership share. For ease of interpretation, all explanatory variables are expressed in terms of their standard deviations. Thus, the coefficient estimate of $-.001$ on average age of household head implies that, once one controls for enumeration district fixed effects, a one standard deviation decrease in average household head age in 1930 is on average associated with a one-tenth of a one percentage point increase in percent black on the block in 1940. While a number of demographic variables are statistically significant in columns 1 and 4, the coefficient estimates are all quite small, and from an economic perspective, they are precisely estimated zeros.

Omitted from these first two regressions are controls for proximity to existing black neighborhoods, perhaps the most important potential confound (see Table 1). To address the possibility that distance to an existing black neighborhood is biasing our baseline results, we geocoded our sample of city blocks.¹⁷ This geocoded subsample allows us to directly test the efficacy of our ED fixed effects in controlling for existing black neighborhood proximity. Appendix Figure II presents a visualization of our geocoded blocks for Detroit, which is typical of all of our sample cities. A limitation of our geocoding is that we were only able to geocode approximately 87 percent of our sample. One concern is that this subsample will vary systematically from our main sample as addresses that were targeted for urban renewal and demolition in the 1960s and 1970s may be overwhelmingly represented in the set of addresses that could not be geocoded.

¹⁷ See the Appendix III for a description of this process.

Columns (2) and (5) of Table 6 replicate the regressions of columns 1 and 4 on the geocoded subsample; while columns (3) and (6) add distance to the nearest black neighborhood (ED > 15 percent black in 1930) to the regressions (normalized by standard deviation). Consistent with expectations, such proximity is a stronger predictor of transition than were our demographic variables. Nonetheless, ED fixed effects are still effective in absorbing proximity's impact. Controlling for ED, on average, a one standard deviation decrease in proximity in 1930 is associated with only a one-half of a one percentage point increase in percent black on a block in 1940. These results suggest that our price and occupancy results are driven by racial transition and not by other factors.

As a final test on this dimension, we evaluate the impact of incorporating distance to existing black neighborhoods directly into our main specification. Columns (2) and (5) of Table 7 replicate our baseline results for the geocoded sample; for comparison, baseline results for the entire sample are repeated in Columns (1) and (4). While qualitatively similar to the full-sample estimates, the interaction between rented and majority black is smaller in the geocoded sample versus the full sample (.345 versus .503, respectively). Thus, it is important to focus within the geocoded subsample when assessing the impact of controls for distance to the nearest black neighborhood on our coefficients of interest. Columns (3) and (6) add a control for distance to the nearest black neighborhood (defined as miles to an enumeration district that was at least 15 percent black) to the model. Comparing these results to those in columns (2) and (5) demonstrates that while distance to the nearest black neighborhood is negatively associated with price, all other coefficient estimates are virtually unchanged by its inclusion in the regression, suggesting that enumeration district fixed effects provide sufficient controls for this source of

selection bias. Additionally, distance to a black neighborhood is not associated with occupancy (column 6).

D. Further robustness and concordance with non-parametric analysis

Our parametric results are based on a relatively granular characterization of the racial transition of city blocks: moving from all white in 1930 to majority black in 1940. To develop a richer understanding of the underlying process, we explore the impact of racial transition on prices and occupancy over the range of 1940 black share. This approach echoes our semiparametric analysis and provides insight into price dynamics on blocks that were at different stages of racial transition. Specifically, we partition our sample of blocks that were white in 1930 into four groups: those that remained white, those that had strictly between 0 and 10 percent black population in 1940, those that had between 10 and 50 percent black population in 1940, and those that had over 50 percent black population in 1940. The coefficient estimates from this analysis are presented in Appendix Table IV. For ease of interpretation, we also summarize the effects on prices, rents and capitalization rates for this specification in the top panel of Table 3.

Recall that our preferred specification includes neighborhood (ED) fixed effects along with block-level controls. Thus, identification comes from variation in block-level racial composition from within a very small neighborhood and beyond that which can be predicted by residential density and rental share. In terms of average prices and rents, our findings are robust to this disaggregation. The finding that transition impacts are muted below 50 percent black in our most robust specification (with ED fixed effects) is consistent with the non-parametric analysis of Tables 2 and 3 and the sensitivity analysis presented in Appendix Figure V. Appendix Figure V also demonstrates our results' robustness to the choice of transition threshold.

Additionally, we note here that our results are robust to alternate definitions of a white block in 1930 (see Appendix Figure IV).

VI. Discussion

The housing market dynamics underlying our analysis arose under, and were shaped by, the Great Migration. During this period, black migrants were both pushed and pulled to the North by a myriad of factors; the opportunity for economic advancement playing a central role. In 1940, black men between the age of 18 and 60 working for wages in the states of the former Confederacy earned \$475 per year. In contrast, their counterparts living in one of our ten sample cities earned on average \$994 per year. Upon migrating North, these individuals fared slightly less well than longer-term black residents of the North, earning on average approximately 15 percent less than the extant work force (\$839 per year), but they still experienced a greater than 75 percent increase in average wages.¹⁸

In contrast, our analysis demonstrates that this promise of higher wages was offset, at least partially, by forces at work in the North's segregated housing markets. At this time, the large majority of blacks living in our sample cities were renters, paying an average rent of \$36.88 per month. At this level of rent, our estimated 50.2 percent rent premium translated to an annual cost of \$147.97, or just over 40 percent of the income gain associated with northern migration. One way that black renters attempted to ameliorate these costs was to live more densely relative to white renters (on the order of 22 percent according to the estimates presented in Table 2 column 3). But of course this increase in density was not without costs, and our estimates assume that the transition from white block to black block was associated with no other changes in

¹⁸ Averages computed from the 100 percent sample of the 1940 US census accessed through IPUMS-USA. We restrict the sample to wage workers who reported working at least 50 weeks in the prior year.

neighborhood quality, such as those that might be expected if the racial transition was associated with a reduction in the quality of city services provided to its residents (Derenoncourt, 2018).

Some migrants would have avoided these steep rent premiums by instead purchasing their home. As is shown in Panel A of Figure 4, by far the highest black ownership rates occurred on predominantly white blocks where black ownership rates were on the order of 70 percent (as opposed to black ownership rates of closer to 20 percent on predominantly black blocks). Blacks purchasing on these blocks were also disadvantaged by the segregated housing markets (see Panel B of Figure 3 and Table 5). Black families who purchased homes on city blocks that were less than 10 percent white paid on average \$4,166.74 for their home and bought at a premium of almost 28 percent relative to what white families would have paid on the same block. Further, the arrival of a few pioneering black families led to a significant racial transition, switching the block from white to black. Then, this transition would have on average eroded the home's value to roughly 10 percent its initial value (See Panel A of Figure 3 and Table 2). The net effect of the initial race premium and ensuing price erosion on such a house represents a loss of \$1,218.72, or nearly 3.5 years' worth of the migration-driven gain in wages for a typical black migrant.

Largely unanswered in this discussion is an identification of the specific role that racism played in underpinning the price dynamics of this period's segregated housing markets and the costs that these markets extracted from black families. The potential channels are numerous and interwoven. Clearly white flight from the arrival of pioneering black families is an important channel.¹⁹ While such flight was likely at least partially driven by white attitudes about black families, flight could also be partially explained by expectations about how a block's racial

¹⁹ Shertzer and Walsh (2019), Boustan (2011).

transition could lead to a reduction in the quality of city services, with such reductions themselves being the result of racist city policies. Similarly, the price premium demanded of pioneering black families can be viewed as evidence of preference-based discrimination by sellers or of concerns about sanctions that would be faced by either the selling agent or the selling family at the hands of remaining owners. Of course these sanctions reflected underlying racist attitudes or a fear of future price declines, which themselves were likely rooted in racist attitudes and city policies. To some degree, likely all of these channels were at work and served to reinforce themselves in an unfortunate circle of causality.

VII. Conclusion

In this paper we constructed a novel dataset of rents, home values, and the racial composition of city blocks in interwar American cities to systematically investigate the housing market dynamics associated with black entry into white neighborhoods. We find that racial transition was associated with both increases in aggregate rental prices and decreases in property values. To our knowledge this is the first paper to demonstrate that black entry into a neighborhood caused the price of owned and rented housing to diverge, a finding that is consistent with much of the narrative history.

Impacts of racial transition were large. We find that rental prices soared by 50 percent in blocks that transitioned from all white to majority black. In contrast, home values fell by 10 percent relative to blocks that remained all white. The impact of these market dynamics for racial wealth inequality were further exacerbated by our finding that pioneering black families paid a significant premium for homes on majority white blocks at the early stages of transition. Similarly, rent discounts to white families that remained on transitioning blocks later into the

process also further eroded black wealth relative to that of whites. Our conservative calculation suggests that the rental premium required of black families was roughly 40 percent of the wage gain for the average unskilled laborer of moving North relative to remaining in the South. Our findings strongly indicate that segregated housing markets eroded a large fraction of the potential return to migrating to higher-paying labor markets for African Americans.

The dramatic decline in property values had important implications for city budgets and real estate investors alike. Rental property owners, faced with the costs of creating and maintaining rental units that were going to depreciate in value and with a ready supply of black households desperate for housing outside of the already underserved ghetto, were able to charge high enough rental prices to make their investment worthwhile. These processes overlapped and reinforced each other, during which entire sections of cities transitioned from being all white to majority black over a relatively short period, with devastating results for black household wealth. Our results highlight the importance of private market dynamics that occurred at the block level prior to the heyday of the FHA and suggest that racial disparities in wealth accumulation would likely have emerged absent discriminatory federal policies. Government at all levels missed the opportunity to change the trajectory of private housing markets.

References

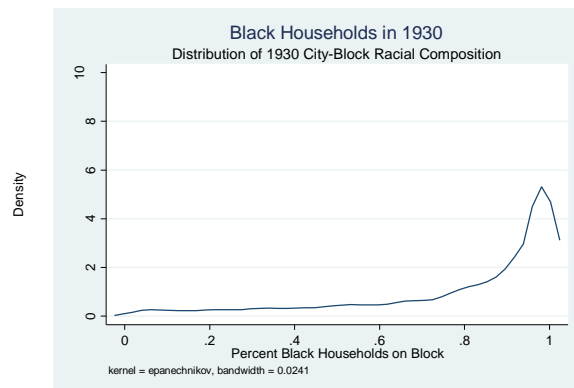
- Albouy, David, and Mike Zabek. *Housing inequality*. No. w21916. National Bureau of Economic Research, 2016.
- Ananat, Elizabeth Oltmans. "The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality." *American Economic Journal: Applied Economics* 3.2 (2011): 34-66.
- Bayer, Patrick, and Kerwin Kofi Charles. "Divergent paths: A new perspective on earnings differences between black and white men since 1940." *The Quarterly Journal of Economics* 133.3 (2018): 1459-1501.
- Blau, Francine D., and John W. Graham. "Black-white differences in wealth and asset composition." *The Quarterly Journal of Economics* 105.2 (1990): 321-339.
- Bound, John, and Richard B. Freeman. "What went wrong? The erosion of relative earnings and employment among young black men in the 1980s." *The Quarterly Journal of Economics* 107.1 (1992): 201-232.
- Boustan, Leah Platt. "Was postwar suburbanization "white flight"? Evidence from the black migration." *The Quarterly Journal of Economics* 125.1 (2010): 417-443.
- Boustan, Leah P., and Robert A. Margo. "A silver lining to white flight? White suburbanization and African-American homeownership, 1940-1980." *Journal of Urban Economics* 78 (2013): 71-80.
- Boustan, Leah Platt. *Competition in the promised land: Black migrants in northern cities and labor markets*. Princeton University Press, 2016.
- Card, David, Alexandre Mas, and Jesse Rothstein. "Tipping and the Dynamics of Segregation." *The Quarterly Journal of Economics* 123.1 (2008): 177-218.
- Chambers, Daniel N. "The racial housing price differential and racially transitional neighborhoods." *Journal of Urban Economics* 32.2 (1992): 214-232.
- Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics* 129.4 (2014): 1553-1623.
- Collins, William J., and Katharine L. Shester. "Slum clearance and urban renewal in the United States." *American Economic Journal: Applied Economics* 5.1 (2013): 239-73.
- Collins, William J., and Marianne H. Wanamaker. "Selection and economic gains in the great migration of African Americans: new evidence from linked census data." *American Economic Journal: Applied Economics* 6.1 (2014): 220-52.
- Collins, William J., and Marianne H. Wanamaker. *Up from slavery? African American intergenerational economic mobility since 1880*. No. w23395. National Bureau of Economic Research, 2017.

- Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. "The rise and decline of the American ghetto." *Journal of political economy* 107.3 (1999): 455-506.
- Davis, Morris A., Andreas Lehnert, and Robert F. Martin. "The Rent-price ratio for the aggregate stock of owner-occupied housing." *Review of Income and Wealth* 54.2 (2008): 279-284.
- Derenoncourt, Ellora. *Can you move to opportunity? Evidence from the Great Migration*. Harvard University, mimeo, 2018.
- Follain Jr, James R., and Stephen Malpezzi. "Another look at racial differences in housing prices." *Urban Studies* 18.2 (1981): 195-203.
- Kearl, James R. "Inflation, mortgage, and housing." *Journal of Political Economy* 87.5, Part 1 (1979): 1115-1138.
- King, A. Thomas, and Peter Mieszkowski. "Racial discrimination, segregation, and the price of housing." *Journal of Political Economy* 81.3 (1973): 590-606.
- Laurenti, Luigi. *Property Values and Race*. University of Berkeley Press, 1960.
- Lavoice, Jessica. "The Long-Run Implications of Slum Clearance: A Neighborhood Analysis." Manuscript, 2019.
- John R. Logan and Weiwei Zhang. 2017. "Developing GIS Maps for U.S. Cities in 1930 and 1940" In Don Lafreniere, Ian Gregory, and Don Debats (editors), *The Routledge Handbook of Spatial History*. Routledge: UK.
- Massey, Douglas S., and Nancy A. Denton. *American apartheid: Segregation and the making of the underclass*. Harvard University Press, 1993.
- McEntire, Davis. *Residence and race*. Univ of California Press, 1960.
- Mehlhorn, Dmitri. "A requiem for blockbusting: Law, economics, and race-based real estate speculation." *Fordham L. Rev.* 67 (1998): 1145.
- Myrdal, Gunnar. *An American dilemma; the Negro problem and modern democracy*. Harper and Bros., 1944.
- Poterba, James M. "Taxation and Housing: Old Questions, New Answers." *The American Economic Review* 82.2 (1992): 237-242.
- Rapkin, Chester, and William G. Grigsby. *The Demand for Housing in Racially Mixed Areas: A Study of the Nature of Neighborhood Change*. Univ of California Press, 1960.
- Robinson, Peter M. "Root-N-consistent semiparametric regression." *Econometrica: Journal of the Econometric Society* (1988): 931-954.
- Rothstein, Richard. *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing, 2017.

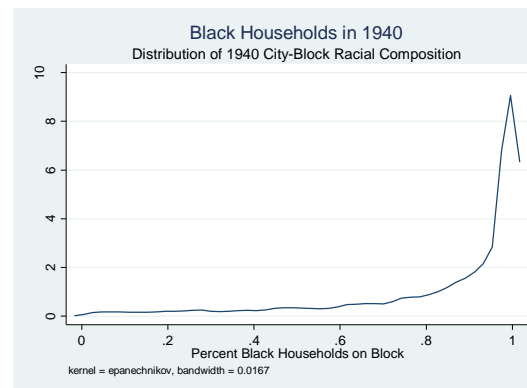
- Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2019.
<https://doi.org/10.18128/D010.V9.0>
- Satter, Beryl. *Family properties: Race, real estate, and the exploitation of black urban America*. Macmillan, 2009.
- Schafer, Robert. "Racial discrimination in the Boston housing market." *Journal of Urban Economics* 6.2 (1979): 176-196.
- Schelling, Thomas C. "Dynamic models of segregation." *Journal of mathematical sociology* 1.2 (1971): 143-186.
- Schietinger, Egbert Frederick. "Racial succession and changing property values in residential Chicago." PhD dissertation, University of Chicago, Department of Sociology, 1953.
- Shertzer, Allison, and Randall P. Walsh. "Racial sorting and the emergence of segregation in American cities." *Review of Economics and Statistics* 101.3 (2019): 415-427.
- Shertzer, Allison, Randall P. Walsh, and John R. Logan. "Segregation and neighborhood change in northern cities: New historical GIS data from 1900–1930." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 49.4 (2016): 187-197.
- Smith, James P., and Finis R. Welch. "Black economic progress after Myrdal." *Journal of economic literature* 27.2 (1989): 519-564.
- Troesken, Werner, and Randall Walsh. "Collective Action, White Flight, and the Origins of Racial Zoning Laws." *The Journal of Law, Economics, and Organization* 35.2 (2019): 289-318.
- Wilson, William Julius. *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press, 2012.
- Wolff, Edward N. *Household Wealth Trends in the United States, 1962-2013: What Happened over the Great Recession?*. No. w20733. National Bureau of Economic Research, 2014.
- Yinger, John. "The black-white price differential in housing: some further evidence." *Land Economics* 54.2 (1978): 187-206

Figure 1. Average City-Block-Level Percent Black Experienced by Black Families

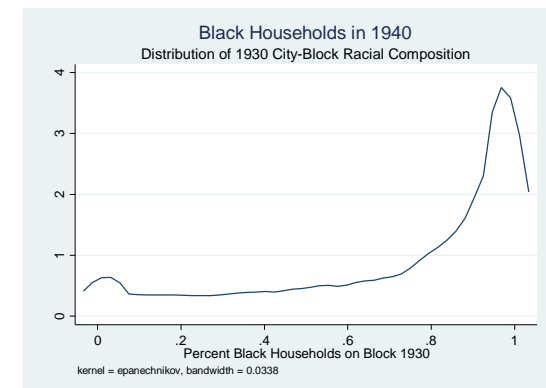
Panel A. 1930



Panel B. 1940



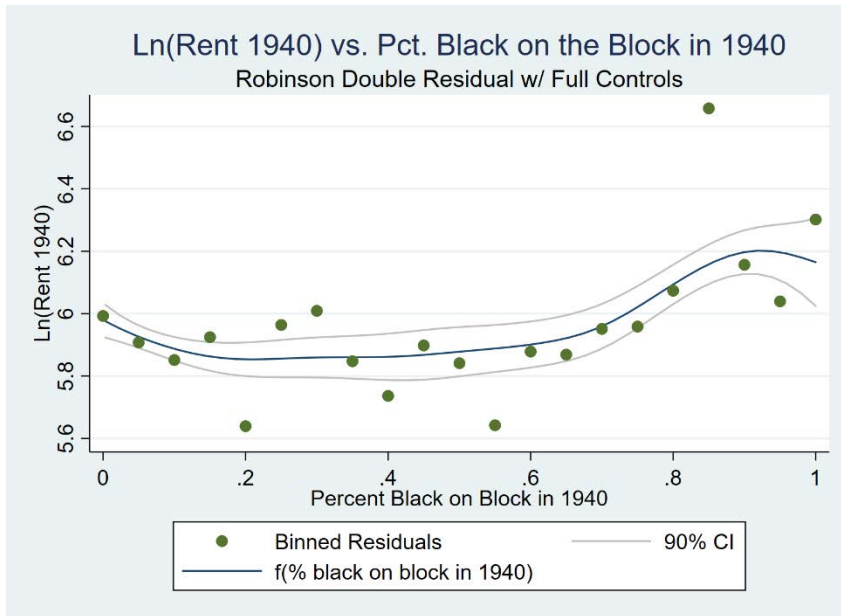
Panel C. 1940 (level in 1930)



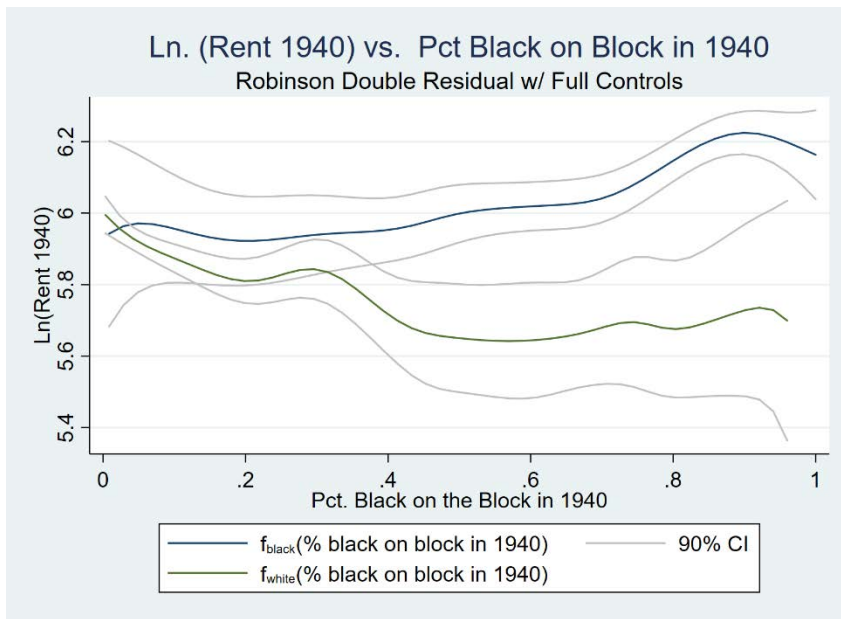
Notes: These figures show the distribution of percent black experienced by black households in our matched sample. Panels A and B report contemporaneous distributions (i.e. 1930 percent black on blocks where black families live in 1930). Panel C reports the distribution of percent black in 1930 for the blocks were black families were living in 1940. The basic unit of observation underlying these distributions is a black household head as identified in the 1930 or 1940 census. Note the change in scale between Panel C and Panel's A and B.

Figure 2. Semiparametric relationship between Percent Black and Rents

Panel A. Baseline Sample



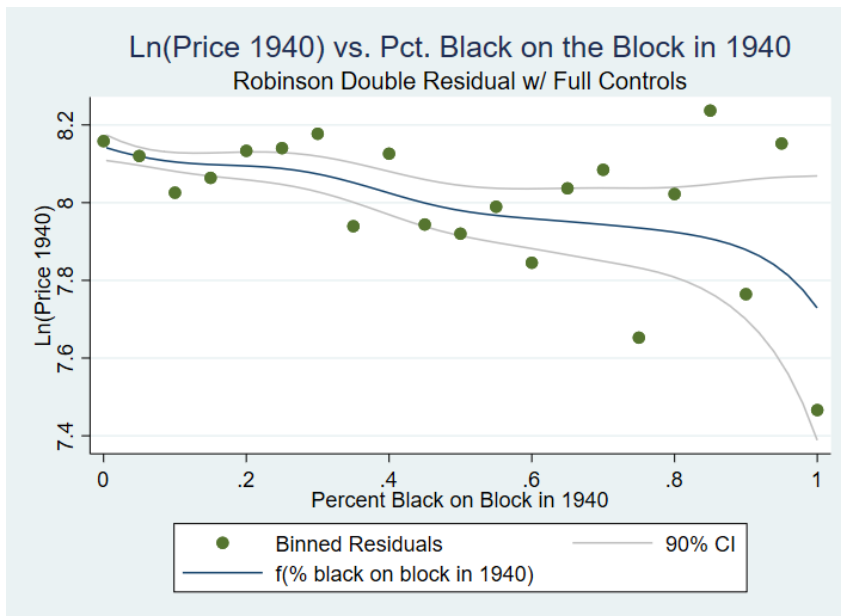
Panel B. Black and White Households Separately



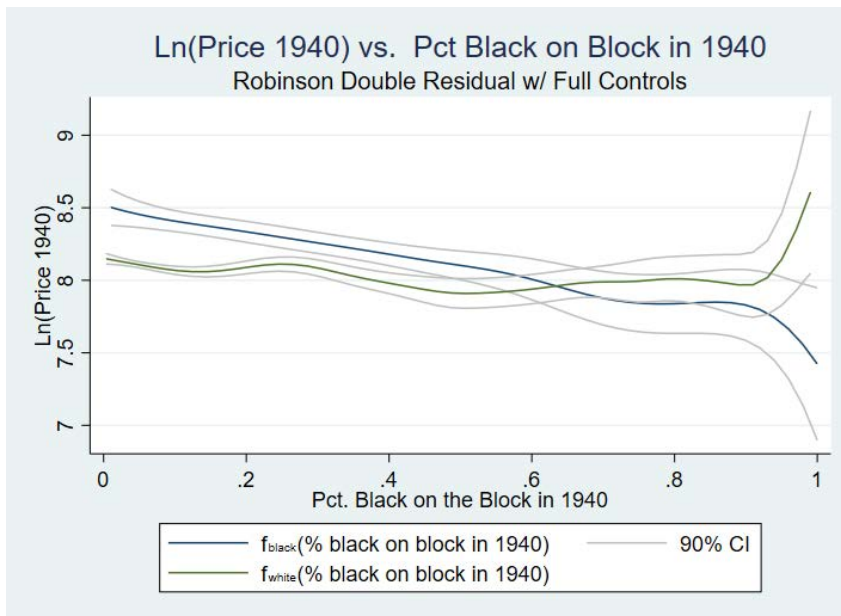
Notes: These figures show the semiparametric relationship between percent black on the block in 1940 (independent variable) and log rent in 1940 (dependent variable) on our baseline sample of homes that were single family, owner occupied, and located on a block that had no blacks in 1930. Controls are included for 1930 price and occupancy at the address level, share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level. The estimation method is Robinson’s double residual method (1988). We also include binned residuals from the regression on each chart.

Figure 3. Semiparametric relationship between Percent Black and Home Values

Panel A. Baseline Sample



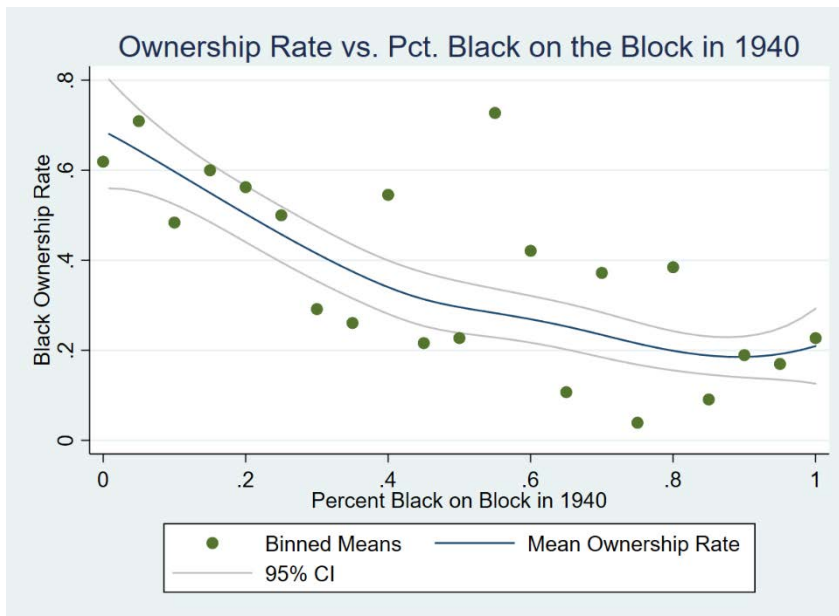
Panel B. Black and White Households Separately



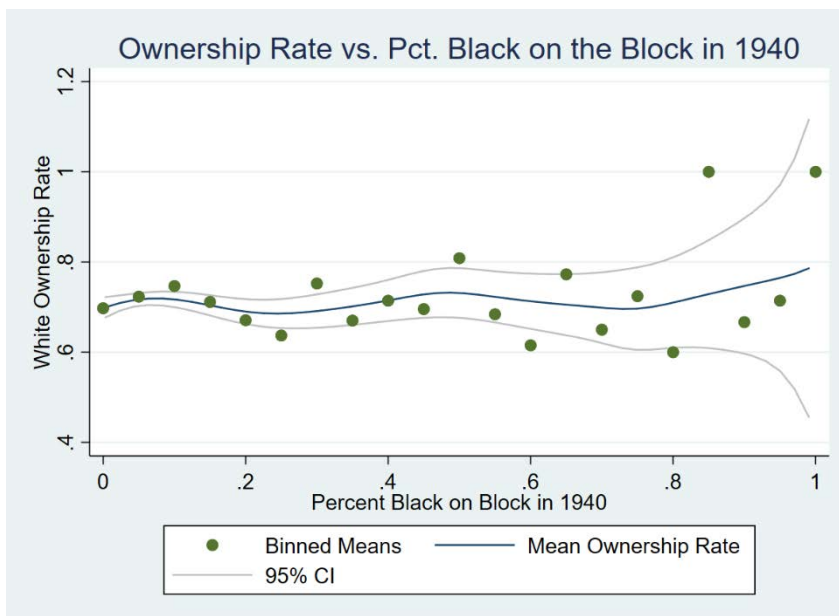
Notes: These figures show the semiparametric relationship between percent black on the block in 1940 (independent variable) and log home price in 1940 (dependent variable) on our baseline sample of homes that were single family, owner occupied, and located on a block that had no blacks in 1930. Controls are included for 1930 price and occupancy at the address level, share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level. The estimation method is Robinson’s double residual method (1988). We also include binned residuals from the regression on each chart.

Figure 4. The Relationship between Percent Black and Ownership Rates

Panel A. Black Ownership Rate



Panel B. White Ownership Rate



Notes: These figures show the semiparametric relationship between percent black on the block in 1940 (independent variable) and log rent in 1940 (dependent variable) on our baseline sample of homes that were single family, owner occupied, and located on a block that had no blacks in 1930. Controls are included for 1930 price and occupancy at the address level, share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level. The estimation method is Robinson’s double residual method (1988). We also include binned residuals from the regression on each chart

Table 1. Summary Statistics for Single-Family Home Addresses

	Blocks stayed white in 1940		Blocks > 0% black in 1940	
	1930	1940	1930	1940
<i>Address Level:</i>				
Nominal home value	6280.37 (3466.17)	3897.07 (2322.96)	6654.52 (4312.74)	4095.07 (2739.39)
Aggregate monthly rent	-	35.10 (23.64)	-	39.30 (29.02)
Aggregate occupancy	4.31 (1.84)	4.15 (2.01)	4.59 (2.03)	4.72 (2.62)
Aggregate households	1.00 (0.05)	1.06 (0.36)	1.00 (0.07)	1.18 (0.70)
<i>Block Level:</i>				
Black Share	-	-	-	0.17 (0.24)
<i>Enumeration District Level:</i>				
Laborer Share	0.03 (0.03)	0.03 (0.03)	0.04 (0.04)	0.05 (0.04)
Foreign-Born Share	0.18 (0.10)	0.15 (0.08)	0.21 (0.12)	0.16 (0.10)
Mean Age of Head	29.45 (3.22)	32.80 (2.60)	30.11 (3.51)	32.64 2.91
Share of Homes Owned	1	0.76 (0.43)	1	0.66 (0.47)
Miles to nearest black ED in 1930	1.29 (1.46)	-	0.64 (1.11)	-
Number of owner-occupied houses	247191	187962	5590	3687
Number of rented houses	-	59229	-	1903

Notes: This table reports statistics on our baseline sample of homes that were single family, owner occupied, located on a block with no black residents in 1930, and could be matched across the 1930 and 1940 censuses.

Table 2. Main Results: Price and Occupancy

Panel A: Log price	No	Controls	ED FE	All Obs.	Rental FE
	Controls			FE	
	(1)	(2)	(3)	(4)	(5)
Rented	-2.243*** (0.003)	-2.235*** (0.003)	-2.214*** (0.003)	-2.157*** (0.002)	-2.052*** (0.004)
Racial Transition	-0.266*** (0.039)	-0.166*** (0.038)	-0.096** (0.043)	-0.135*** (0.036)	-0.105 (0.074)
Rented x Transition	0.460*** (0.050)	0.496*** (0.048)	0.503*** (0.049)	0.371*** (0.038)	0.272*** (0.075)
Observations	242,441	241,793	242,441	399,964	151,501
R-squared	0.738	0.757	0.802	0.819	0.731

Panel B: Log occupancy	No	Controls	ED FE	All Obs.	Only
	Controls			FE	Rentals
	(1)	(2)	(3)	(4)	(5)
Rented	0.167*** (0.002)	0.182*** (0.002)	0.184*** (0.002)	0.185*** (0.002)	0.166*** (0.004)
Racial Transition	-0.020 (0.030)	-0.049* (0.027)	-0.032 (0.033)	-0.033 (0.029)	0.068 (0.063)
Rented x Transition	0.278*** (0.038)	0.279*** (0.035)	0.233*** (0.038)	0.163*** (0.030)	0.023 (0.063)
Observations	252,781	246,234	247,169	407,954	154,626
R-squared	0.024	0.211	0.250	0.361	0.445

Notes: The first three columns report the OLS estimation of equation (7) on our baseline sample of homes that were single family, owner occupied, and located on a block with no black residents in 1930. The first column controls only for price and occupancy of the address in 1930. The second column adds controls share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level. The third column drops the neighborhood controls and includes ED fixed effects. The fourth column adds addresses that were rented in 1930 to the sample and uses the specification from column (3) with an additional control for tenure status in 1930. The fifth column uses only rented homes in 1930. The transition indicator is equal to one if the block became more than 50 percent black by 1940.

Table 3. Price Shifts and Capitalization Rates by Occupancy Change

All Addresses			
Percent black on block in 1940	Sales	Rent	Cap. Rate
0	100.00%	100.00%	10.90%
0-10%	98.22%	107.36%	11.92%
10-50%	98.51%	110.08%	12.18%
50-100%	91.48%	151.74%	18.09%
Observations			242,441
Occupancy Increased			
Percent black on block in 1940	Sales	Rent	Cap. Rate
0	100.00%	100.00%	11.24%
0-10%	93.43%	111.18%	13.37%
10-50%	97.92%	116.53%	12.68%
50-100%	76.64%	129.30%	18.96%
Observations			71,943
Occupancy Decreased or Same			
Percent black on block in 1940	Sales	Rent	Cap. Rate
0	100.00%	100.00%	10.32%
0-10%	101.01%	101.82%	10.40%
10-50%	100.40%	100.20%	10.30%
50-100%	99.80%	151.29%	15.65%
Observations			170,498

Notes: The table reports the implied capitalization rates from an OLS estimation of equation (7) on our baseline sample of homes that were single family, owner occupied, and located on a block with no black residents in 1930. Regressions include controls for price and occupancy of the address in 1930, share renters and total number of addresses at the block level, and ED fixed effects. See text for details on how to compute the capitalization rate from regression coefficients.

Table 4. Coefficient on Rental Indicator for Addresses on Blocks that Remained White

	No Controls (1)	Controls (2)	ED FE (3)	All Obs FE (4)
All addresses	-2.244*** (0.003)	-2.237*** (0.003)	-2.216*** (0.003)	-2.160*** (0.002)
Addresses that gained at least 2 members	-2.186*** (0.009)	-2.171*** (0.009)	-2.177*** (0.010)	-2.114*** (0.008)
Addresses that gained at least 3 members	-2.164*** (0.013)	-2.153*** (0.013)	-2.179*** (0.016)	-2.106*** (0.011)
Addresses that gained at least 4 members	-2.139*** (0.019)	-2.118*** (0.018)	-2.161*** (0.025)	-2.080*** (0.017)

Notes: The table reports the rental indicator from OLS estimation of equation (7) on our baseline sample of homes that were single family, owner occupied, and located on a block that had no blacks in both 1930 and 1940. The first column controls only for price and occupancy of the address in 1930. The second column adds controls share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level. The third column drops the neighborhood controls and includes ED fixed effects. The last column adds addresses that were rented in 1930 to the sample and uses the specification from column (3) with an additional control for tenure status in 1930. The table reports the coefficient on the “rented” variable, which is an indicator for whether the house switched to being a rental in 1940.

Table 5. Discrimination in Owned and Rented Housing

	Black < 5%	Black < 10%	Black > 40%	Black > 60%
	(1)	(2)	(3)	(4)
Rented	-2.216*** (0.003)	-2.215*** (0.003)	-2.093*** (0.080)	-2.033*** (0.104)
Black Household	0.322*** (0.106)	0.250*** (0.074)	0.097 (0.079)	-0.023 (0.095)
Rented x Black Household	0.213 (0.182)	0.133 (0.128)	0.295*** (0.111)	0.368*** (0.133)
Black Pioneer Premium	0.38	0.28		
Black Rental Premium			0.34	0.44
Observations	3,044	1,991	837	533
R-squared	0.858	0.846	0.822	0.840

Notes: The table reports the OLS estimation of the modified version of equation (7) that includes an indicator for black household interacted with the rental indicator. The racial transition indicator is dropped and the modified specification run on different parts of the black share distribution. All specifications include enumeration district fixed effects and are run on our baseline sample of homes that were single family, owner occupied, and located on a block with no black residents in 1930.

Table 6. Predicting Racial Transition in Baseline Sample

Block characteristics in 1930:	Percent Black in 1940			Percent Black in 1940 > 50%		
	(1)	(2)	(3)	(4)	(5)	(6)
Average age of heads of HH	-0.001*** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001*** (0.000)	-0.001* (0.000)	-0.001* (0.000)
Share laborer heads of HH	0.001*** (0.000)	0.000 (0.000)	0.000 (0.000)	0.001*** (0.000)	0.001** (0.000)	0.001** (0.000)
Share foreign born heads of HH	0.001*** (0.000)	0.000 (0.000)	0.000 (0.000)	0.001* (0.000)	-0.000 (0.000)	-0.000 (0.000)
Average rent per person	-0.001** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.001 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Ownership share	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Distance to nearest blk nbhd			-0.005* (0.003)			-0.002 (0.003)
Sample	All	Geo.	Geo.	All	Geo.	Geo.
Observations	51,859	43,819	43,819	53,478	43,819	43,819
R-squared	0.585	0.654	0.654	0.563	0.609	0.609

Notes: This table reports OLS estimations of selection into racial transition using our baseline sample of blocks that had at least one owner-occupied, single-family home and no black residents in 1930.

Table 7. Results for Racial Transition and Proximity to Nearest Black Neighborhood

	Dependent variable = log price			Dependent variable = log occupancy		
	All blocks (1)	Geocoded (2)	Geocoded (3)	All blocks (4)	Geocoded (5)	Geocoded (6)
Rented	-2.214*** (0.003)	-2.222*** (0.003)	-2.222*** (0.003)	0.186*** (0.002)	0.184*** (0.002)	0.184*** (0.002)
Transition	-0.096** (0.043)	-0.202*** (0.058)	-0.201*** (0.058)	-0.039 (0.033)	-0.042 (0.044)	-0.042 (0.044)
Rented x Transition	0.503*** (0.049)	0.345*** (0.066)	0.345*** (0.066)	0.240*** (0.037)	0.120** (0.050)	0.120** (0.050)
Distance to Blk Nbhd			0.102*** (0.012)			-0.001 (0.009)
Observations	242,441	211,964	211,964	252,781	221,054	221,054
R-squared	0.802	0.805	0.805	0.250	0.252	0.252

Notes: The table reports the OLS estimation of equation (7) on our baseline sample of homes that were single family, owner occupied, and located on a block with no black residents in 1930. Regressions include controls for price and occupancy of the address in 1930, share renters and total number of addresses at the block level, and ED fixed effects. The second and third columns include only addresses that could be geocoded. Black neighborhoods are defined as EDs with at least 15 percent black population.

Data Appendix

I. Constructing the Matched Address Sample

Each record in the census data represents an individual in a household. Each household has a head and related members who share the same address. An address is the combination of a house number and a street name. In an ideal world, we would know the number of individuals and households residing at a given address. However, either the house number or the street name entry for an individual could have been mis-recorded by the census enumerators or mis-digitized by the contemporary census digitization workers. Therefore, some households have incorrect or incomplete addresses, possibly leading to inaccurate counts of households in any building. This appendix describes the algorithm we used to construct a representative set of households for our sample cities in 1930 and 1940, focusing in particular on the challenge of assigning all individuals to the correct address.

We first need to make sure that no household is either missing an address or assigned more than one. We assume that the enumeration districts (EDs) and tracts reported in the census data were transcribed correctly. A tiny fraction of EDs and tracts from the census do not coincide with the list of EDs that we use to define our cities. We drop those EDs or tracts, as they are likely to be institutions that were given a separate ED number.

We have digitized 1930 enumeration district boundaries (Shertzer et al. 2016) and obtained census tract boundary files from the National Historical Geographic Information System (NHGIS). We cross-check census address data by “fuzzy” matching each census street name to a list of street names from the corresponding ED/tract obtained from the spatial datasets. We exclude addresses on streets that have either no reasonable match or too many potential matches among the digitized streets.

Census enumerators were instructed to survey households as they moved along a street, and thus we do not expect to see house numbers within a street jump around. Thus, the order in which households appear on the manuscripts should generally reflect their location within the ED relative to neighboring households.²⁰ To ensure that we have all the households living in each address in our sample, we also drop any address that shares a street-block (or the entire street-ED when the block cannot be identified) with an address that is potentially out of order on the manuscript. We provide further details of the process below.

II. Details on Matching Methodology

We make sure that every household has exactly one address composed of a street name and house number. To begin, we assign the address information from the household head to everyone in his/her household. When the household head has partial (e.g. only a street name or only the house number) or no information on address, we fill in information from the household's non-head member. We perform a series of quality checks on these imputed addresses that are described below. If the household head is missing an address and household members disagree on either street name or house number, we impute the missing address information from those of households listed just before this one in the census manuscripts and flag these households.

In the case of multiple households sharing the same dwelling unit, we will have more than one household head. When these household heads disagree on the address, we compare each component of the addresses (the street names and house numbers) to those of adjacent households and keep the one(s) that matches that of the most number of neighbors. We flag all

²⁰ Our indicators of manuscript page and line numbers are not very reliable, so we use the household IDs assigned by IPUMS as proxy for the order in which households appear in the original census manuscripts.

addresses imputed from adjacent households. A very small number of dwellings from the 1940 Census seem to have members belonging to different EDs/tracts. As with street names and house numbers, we assume the household head's ED/tract is the correct one. In the case of multi-family households, we compare each candidate ED/tract with those of households appearing immediately before and after on the census manuscripts, and only retain the EDs/tracts with the highest number of matches. We have a few households located at the intersection of EDs/tracts, and we flag these as well.

Then we standardize street names in the census, which are noisy and frequently riddled with typos. We first standardize all the directional prefix and street suffix, convert ordinal street numbers to their cardinal text forms, and remove any redundant information from street name (such as "Block A"). We then match these formatted street names to our digitized 1930 city streets to standardize further the names. We create a crosswalk of digitized street names, 1930 EDs, and 1940 tracts and fuzzy match them with the set of unique census street names by ED/tract (allowing some margin of error in the string match). We use STATA's `relink2` command for this task. If a census street matches to more than one digitized street (a "one-to-many" match) within an ED/tract, then we flag all the digitized streets that were a match. Eventually we drop all Census records where the street does not match a digitized street or matches one that is flagged as part of a one-to-many match. Note that the process is sensitive to the margin of error that we allow in our string match. A wide error margin means we will have more one-to-many matches and fewer non-matches, whereas with a narrow error margin, we will have more non-matches and fewer one-to-many matches. The former introduces false one-to-one matches that might otherwise stay unmatched, whereas the latter introduces false one-to-one matches that might otherwise be matched to many. Thus, a conservative approach is to allow a

wide margin of error, but narrow enough that we are still left with a reasonably sized sample after dropping one-to-many and non-matches.

House numbers, like street names, are also prone to errors and typos. The next step is to standardize house numbers as best as we can across ED/tracts and census years. When the house number variable is just one clear number, we leave it as it is. When it is not (e.g., “945/6”, “4531 667” or “1??2”), we try to identify a minimum and a maximum possible house number. For instance, when the reported house number is “4531 667”, we treat it as ranging from 667 to 4531 and flag all addresses on the same street block and ED with house numbers in that range.²¹ We assume a “?” can range from 0 to 9, so that house number “1??2” ranges from 1002 to 1992. We treat separators like “/”, “-”, “&”, “+”, “~” and “,” as spaces when identifying the range, while we ignore alphabets (treating “5a” as “5”) and other non-alphanumeric characters (e.g. parentheses and brackets). All problematic addresses are flagged.

We do not have digitized historical house numbers as with street names to validate our cleaning process. Instead, we perform a number of quality checks based on the ordering of households in the census manuscripts and flag households that fail to satisfy these checks. Failing one or more of these reality checks implies that the re-formatted and standardized addresses are unlikely to be correct. These flagged households include cases where:

- (1) the address differs from that of adjacent households on the manuscript when adjacent households share an address,

²¹ There are alternative ways of interpreting a reported house number of “4531 467”. The second number might be an apartment number within the building, or the building might span house numbers 4531 to 4667. However, given that we eventually drop all street blocks intersecting this range of numbers, we believe our range assignment is the most conservative in dealing with such ambiguity. The street block is defined by the street name and the hundreds of the house number.

- (2) only the house number matches that of one adjacent household, and only the street name matches that of the other adjacent household,
- (3) the house number differs from adjacent house numbers by more than 10 along the same street,
- (4) the house number changes non-monotonically (and differs from adjacent house numbers by at least 4) along the same street, and
- (5) the address is (either partially or completely) imputed from that of the preceding household when adjacent street names differ.

We drop households in all addresses that were flagged in any of the previous steps. If a household's address is flagged, the correct address is likely to be that of adjacent households on the manuscript, given our assumption on the path of the enumerators. To avoid undercounting the individuals in these adjacent addresses, we also drop all addresses adjacent to flagged addresses on the manuscript. Thus, we generate a sample of addresses that are correct with a reasonable degree of accuracy that is our baseline.²²

Finally, from each sample, we retain only the addresses that appear in both the 1930 and the 1940 Census. Since we have digitized 1930 ED boundaries and 1940 tract boundaries, we further make sure that the reported EDs (in 1930) and tracts (in 1940) corresponding to each address overlap spatially.

III. Geocoding Addresses

We geocode all formatted address strings on Google Maps' Directions API. We include all 1930 and 1940 addresses with non-missing street names and house numbers, including those we

²² If the street block of a flagged address cannot be identified credibly (e.g. when the house number is completely non-numeric or the range of house numbers is unrealistically large), we drop all addresses on the same street and ED.

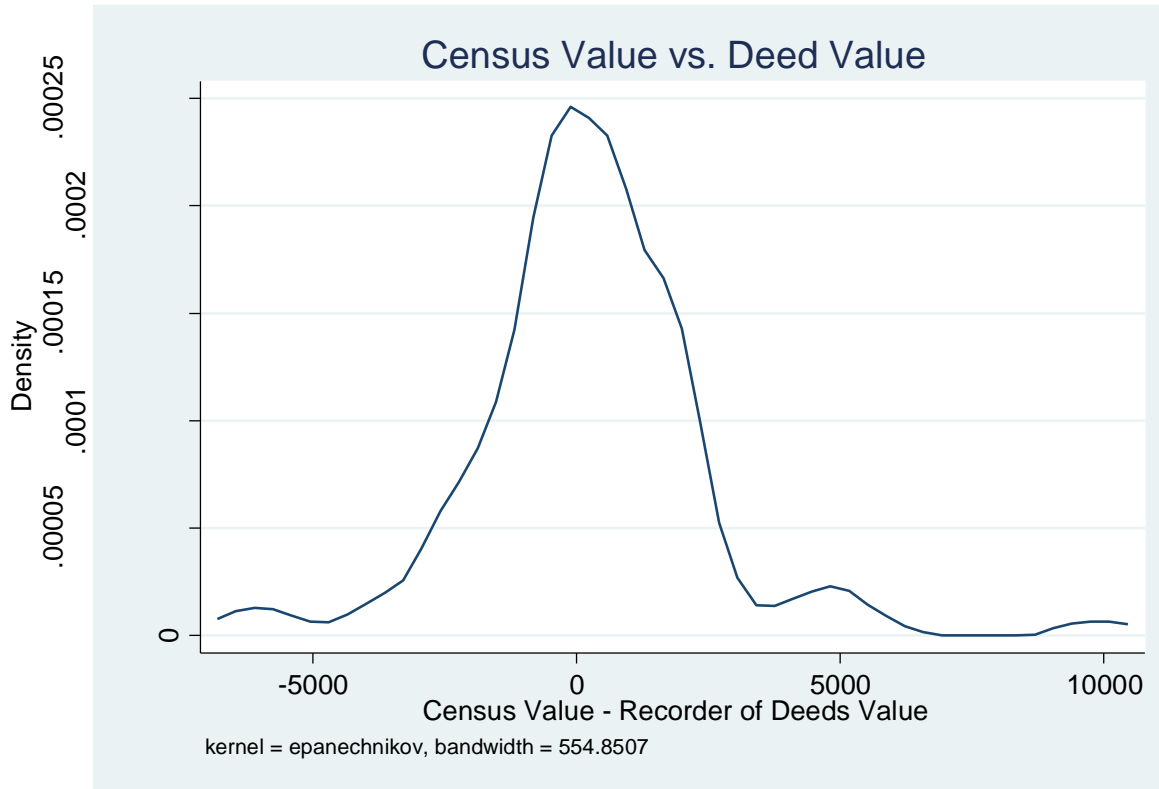
have flagged as potentially erroneous. The Directions API does a fuzzy name match of our input strings with addresses on Google’s database and returns none, one or multiple location matches. For each location match, the API returns the geographic coordinates, the level of precision of the geocoding (e.g. “street address”, “route”, “intersection”, “ward”, etc.), and any administrative/political areas that the geocoded location falls within (e.g. the county, city, state, postal code or other well-defined “neighborhoods”). We drop any matches where the precision of the geocoding is an administrative area (e.g. a ward, a neighborhood, a city, etc.) or if the state differs from that of our city.

We then map each geocoded location to our 1930 ED and 1940 tract boundaries, and drop any location matches that do not coincide with either the ED or the tract associated with the input address. From each remaining geocoded location matches, we compute straight-line distances to the nearest black neighborhood. If there are still multiple location matches for an address, we keep the location match whose distance to the nearest black neighborhood is closest to the average distance from all location matches of addresses in the same block. Finally, in a small number of cases, when location matches for an address are tied in their deviation from the average distance to the nearest black neighborhood in the block, we pick the location match that appears first in the Directions API’s sorting of results.²³

To compute block-level distances to the nearest black neighborhood, we take the average of distances from each address in the block. As long as a block includes at least one address that is not flagged as problematic, we exclude distances from flagged addresses.

²³ The sorting reflects the “prominence” of the location, which is Google’s measure of how likely the location is to be the result of a search.

Appendix Figure I. Self-Reported Value vs. Deed Value from County Records



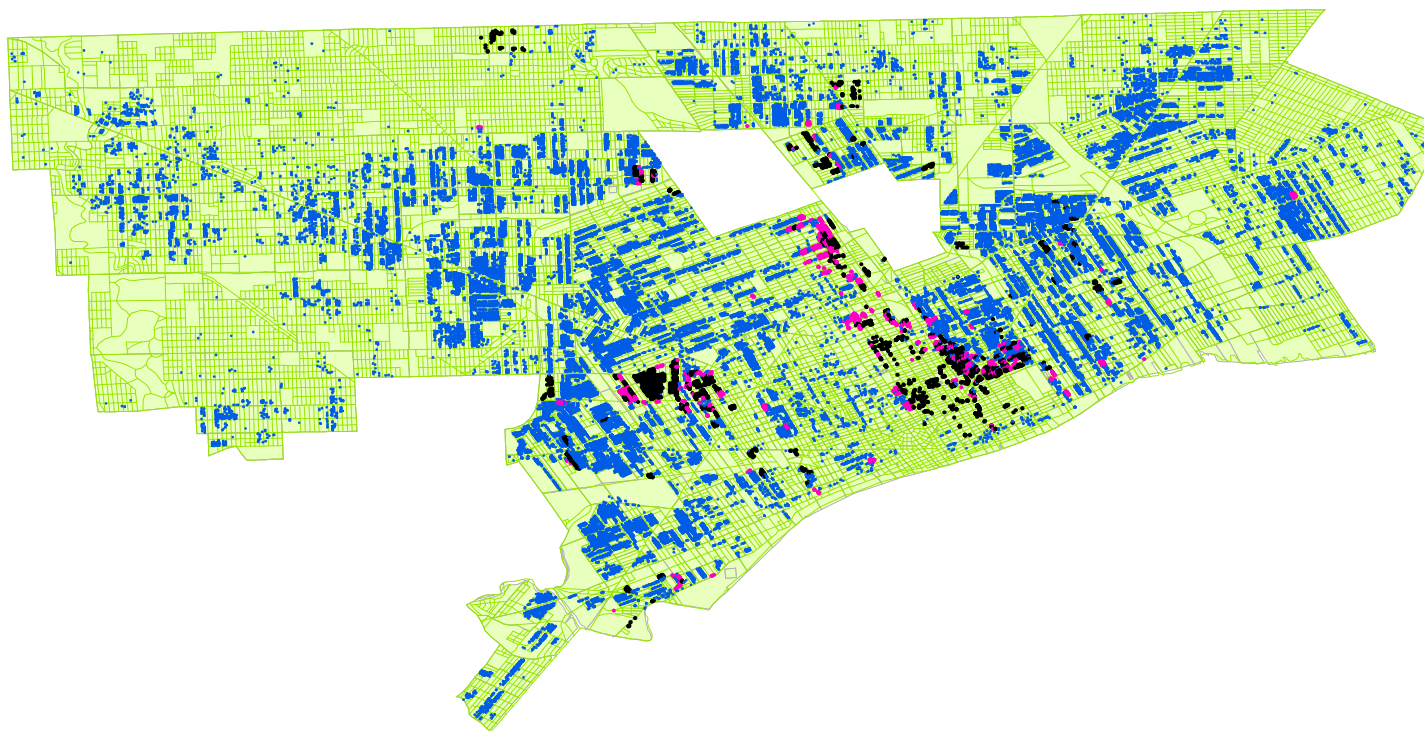
Notes: the figure presents a Kernel Density Estimate of the PDF of differences between self-reported home valuations as recorded in the decennial census and sales amounts as recorded by the Allegheny County Recorder of Deeds for a sample of 404 owner-occupied homes in the city of Pittsburgh. The data were constructed by identifying homes in the recorder of deed's records that were sold in either 1930 or 1940 and then hand matching them to the appropriate individual census record based on the home's address.

Appendix Figure II. Geocoded Detroit Addresses



Notes: the figure shows the addresses in our sample for the city of Detroit that could be geocoded against a map of 1940 enumeration districts produced by Logan and Zhang (2017).

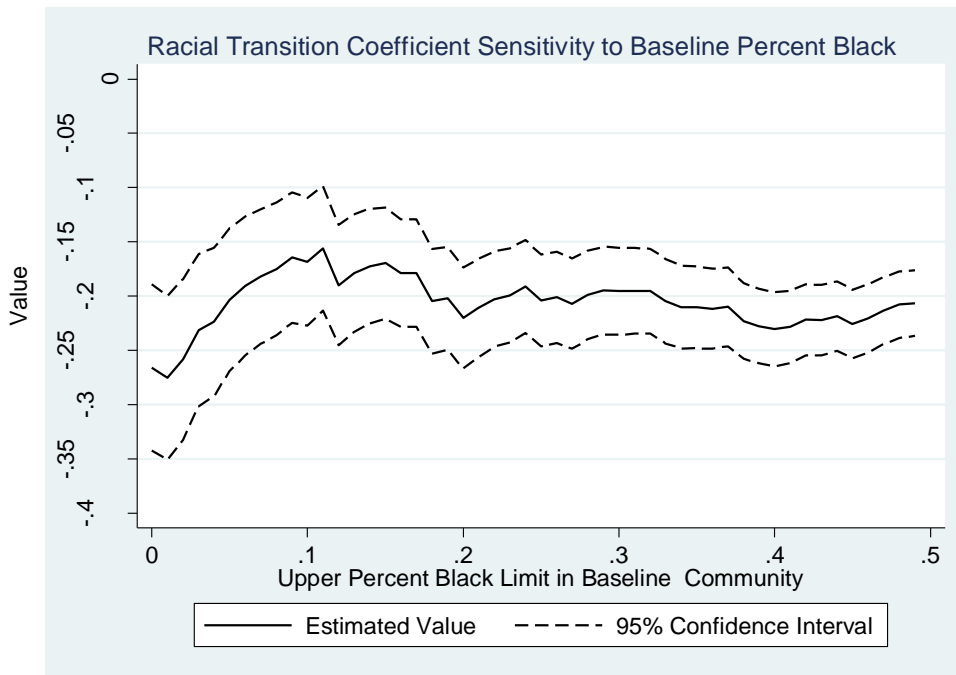
Appendix Figure III. Racial Transition in Geocoded Blocks in Detroit



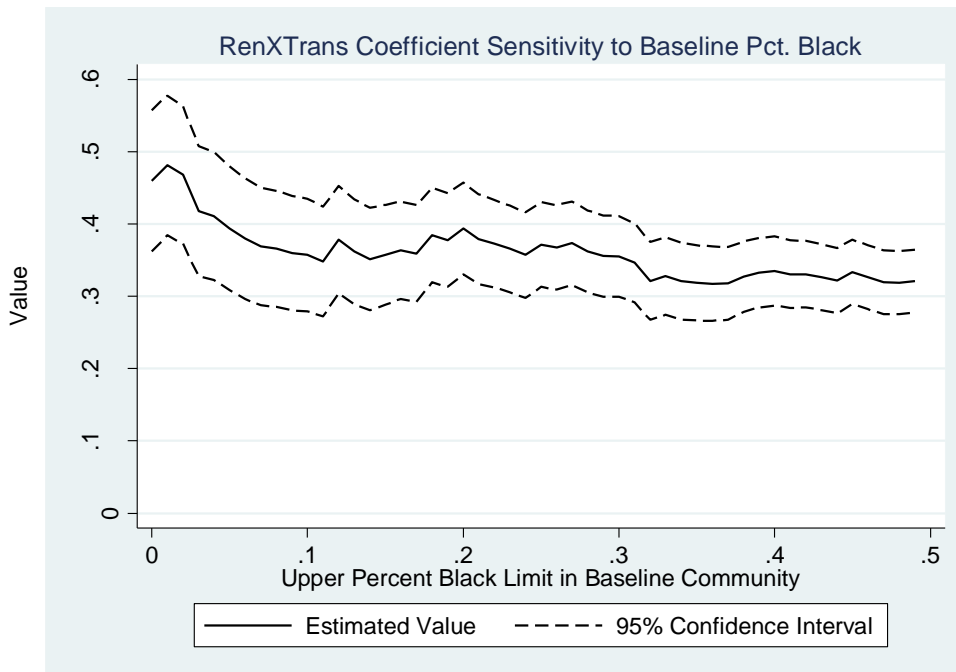
Notes: the figure shows the addresses in our sample for the city of Detroit that could be geocoded against a map of 1940 enumeration districts produced by Logan and Zhang (2017). Blocks are color-coded as follows: blue blocks were less than 5 percent black in both 1930 and 1940, pink blocks were less than 5 percent black in 1930 and more than 5 percent black in 1940, and black blocks were over 5 percent black in both 1930 and 1940.

Appendix Figure IV. Robustness to Definition of a White Block in 1930

Panel A. Racial Transition Coefficient



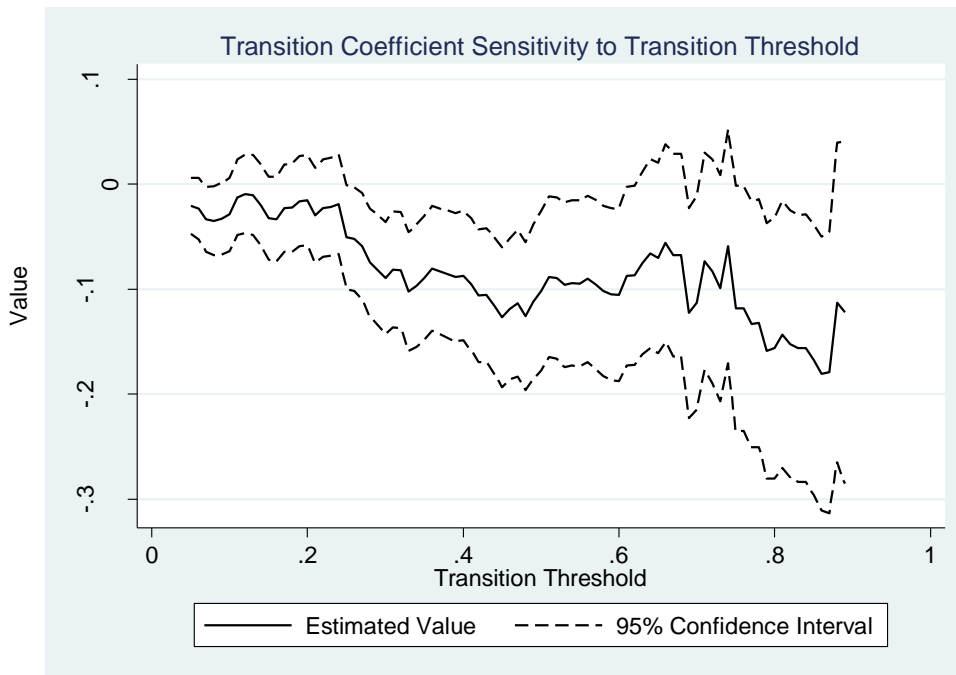
Panel B. Racial Transition Rent Interaction Coefficient



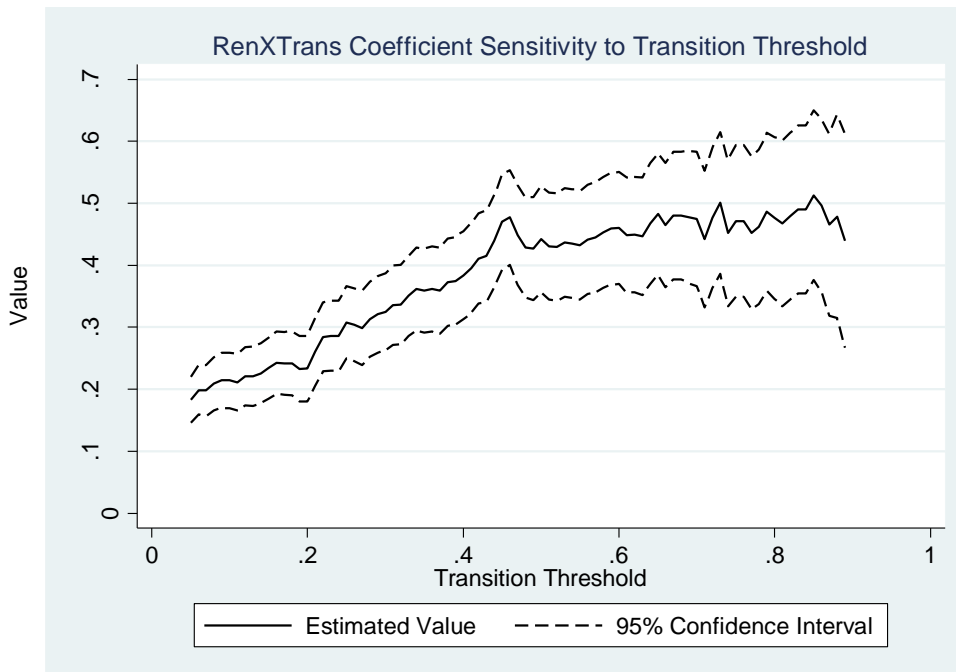
Notes: These figures present parameter estimates consistent with those presented in Column 3 of Table 2, Panel A under different definitions of a baseline block (in Table 2, only blocks with pct. black = 0 in 1930 were included in the analysis).

Appendix Figure V. Robustness to Definition of Racial Transition

Panel A. Racial Transition Coefficient



Panel B. Racial Transition Rent Interaction Coefficient



Notes: These figures present parameter estimates consistent with those presented in Column 3 of Table 2, Panel A under different definitions of a racial transition (in Table 2, transition was defined as all city-blocks that were > 50% black in 1940).

Appendix Table I. Address Statistics for Block Sample

	All Households				Addresses		Blocks		Addresses per Block	
	Total Households		Quality Address		Unique Addresses		Unique Blocks		per Block	
	1930	1940	1930	1940	1930	1940	1930	1940	1930	1940
Baltimore	193,979	245,862	147,962	132,680	118,741	97,264	8,249	7,831	14.4	12.4
Boston	182,090	211,731	132,944	135,944	62,913	61,052	4,090	4,051	15.4	15.1
Brooklyn	614,082	752,606	390,826	358,432	157,005	125,803	8,935	7,450	17.6	16.9
Chicago	845,436	1,025,731	545,383	437,973	278,694	198,297	20,530	17,766	13.6	11.2
Cincinnati	124,321	143,864	87,188	65,169	51,436	38,384	4,898	4,009	10.5	9.6
Cleveland	222,856	247,713	129,774	99,907	86,588	65,744	10,991	8,745	7.9	7.5
Detroit	370,556	451,198	225,457	219,961	168,955	163,406	18,380	18,169	9.2	9.0
Manhattan	470,552	614,786	188,258	191,471	25,178	20,876	1,854	1,856	13.6	11.2
Philadelphia	459,749	515,472	338,928	254,737	291,919	211,705	15,054	12,033	19.4	17.6
Pittsburgh	153,628	185,039	107,276	102,587	78,809	66,712	7,878	7,134	10.0	9.4
St. Louis	216,133	225,794	116,945	111,305	77,551	72,166	7,117	6,560	10.9	11.0
Total/Average	3,853,382	4,619,796	2,410,941	2,110,166	1,397,789	1,121,409	107,976	95,604	12.9	11.9

Notes: The first two columns report the number of households reported in the census in each city. “Quality addresses” are the households for which we were able to assign an address that passed all quality checks described in the Data Appendix. “Unique addresses” are addresses that both pass the quality checks and are unique with a street name, street number, and 1930 enumeration district. We use postal service convention and assign house numbers to blocks using hundreds within a given street name. “Unique blocks” are the number of unique blocks represented by our sample of unique addresses. The last column of the table reports the number of unique addresses per unique block. This is the sample of addresses we used to construct our block sample.

Appendix Table II. Address Sample Statistics

	Households with address found in both census years						Addresses		Households per Address	
	Total		Trimmed Sample		Quality Address		Unique Addresses		1930	1940
	1930	1940	1930	1940	1930	1940	1930	1940		
Baltimore	110,312	125,598	98,780	111,757	67,925	75,888	57,287	57,287	1.2	1.3
Boston	122,353	136,230	100,785	109,334	67,008	71,343	33,267	33,267	2.0	2.1
Brooklyn	365,589	413,796	254,723	286,483	144,116	159,787	62,108	62,108	2.3	2.6
Chicago	443,948	497,700	355,109	389,081	190,802	207,627	104,553	104,553	1.8	2.0
Cincinnati	78,245	85,719	67,293	72,601	33,132	35,112	20,967	20,967	1.6	1.7
Cleveland	124,151	135,182	111,170	118,948	48,676	51,126	34,843	34,843	1.4	1.5
Detroit	212,211	228,290	184,660	194,112	95,309	98,560	76,845	76,845	1.2	1.3
Manhattan	235,841	299,774	95,304	119,594	29,369	36,001	3,913	3,913	7.5	9.2
Philadelphia	227,479	244,202	206,716	218,856	145,313	152,650	131,469	131,469	1.1	1.2
Pittsburgh	84,028	94,428	73,731	81,806	43,172	47,767	32,289	32,289	1.3	1.5
St. Louis	141,183	148,756	124,771	130,522	48,185	50,361	34,239	34,239	1.4	1.5
Total/Average	2,145,340	2,409,675	1,673,042	1,833,094	913,007	986,222	591,780	591,780	2.1	2.3

Notes: The “Total” columns report the number of households with addresses we were able to locate in both the 1930 and 1940 censuses. We trimmed this sample to eliminate transcription errors and institutions (we drop any households with more than 10 members, any household with more than three heads, any addresses with monthly rent greater than \$100, and finally any addresses with a value greater than \$20,000). The “Trimmed Sample” columns report the number of households without problematic census values in both 1930 and 1940. The “Quality Address” columns report the number of households without problematic census values that passed the address quality checks described in the Data Appendix. The “Unique Addresses” columns report the number of addresses represented by this sample of households. This is the sample of addresses we used in our address-level analysis.

Appendix Table III. Selection into Sample

	Year	All	Quality Address	Matched Address
Individuals	1930	15,591,308	9,894,466	4,495,743
	1940	15,729,224	7,560,898	4,345,911
Households	1930	3,845,617	2,406,975	1,082,691
	1940	4,610,562	2,106,438	1,180,009
Addresses	1930	2,077,442	1,407,878	659,688
	1940	2,217,640	1,125,845	659,688
Households per address	1930	1.85	1.71	1.64
	1940	2.08	1.87	1.79
Individuals per address	1930	7.51	7.03	6.81
	1940	7.09	6.72	6.59
Average household size	1930	4.40	4.39	4.39
	1940	3.87	3.96	3.99
Distance to CBD (tract centroid)	1930	4.43	4.32	4.30
	1940	4.61	4.47	4.30
Population density (tract)	1930	0.013	0.013	0.013
	1940	0.012	0.012	0.013
Percent black (tract)	1930	0.076	0.073	0.066
	1940	0.084	0.078	0.078

Notes: The “All” column reports statistics for the full sample of census records across all ten cities. The “Quality Address” column reports statistics for census records that had an address that passed our quality checks as described in the Data Appendix. The “Matched Address” column reports statistics for the sample of quality addresses that could be matched across the 1930 and 1940 census. The distance to CBD is defined as the distance from the central business district to the centroid of the 1940 tract. All tract variables refer to the 1940 census tract.

Appendix Table IV. Decomposing Transition

	Log Price			Log Occupancy		
	No Controls (1)	Controls (2)	ED FE (3)	No Controls (4)	Controls (5)	ED FE (6)
Rented	-2.244*** (0.003)	-2.238*** (0.003)	-2.216*** (0.003)	0.184*** (0.002)	0.181*** (0.002)	0.183*** (0.002)
Transition < 10%	0.005 (0.012)	-0.007 (0.012)	-0.018 (0.013)	0.022*** (0.009)	0.017* (0.009)	-0.008 (0.010)
Rented < 10%	0.071*** (0.023)	0.087*** (0.022)	0.089*** (0.021)	0.066*** (0.016)	0.065*** (0.016)	0.053*** (0.017)
Transition 10-50%	-0.142*** (0.020)	-0.080*** (0.020)	-0.015 (0.022)	-0.000 (0.014)	-0.012 (0.014)	-0.025 (0.017)
Rented x 10-50%	0.105*** (0.034)	0.141*** (0.033)	0.111*** (0.032)	0.152*** (0.024)	0.153*** (0.024)	0.137*** (0.025)
Transition 50-100%	-0.266*** (0.039)	-0.166*** (0.038)	-0.089** (0.043)	-0.042 (0.027)	-0.046* (0.027)	-0.028 (0.033)
Rented x 50-100%	0.461*** (0.050)	0.499*** (0.048)	0.506*** (0.049)	0.290*** (0.035)	0.286*** (0.035)	0.238*** (0.038)
Observations	242,441	241,793	242,441	247,169	246,234	247,169
R-squared	0.738	0.757	0.802	0.202	0.212	0.251

Notes: The first three columns report the OLS estimation of equation (7) on our baseline sample of homes that were single family, owner occupied, and located on a block with no black residents in 1930 with racial transition disaggregated to black share strictly between zero and ten percent, black share between ten and fifty percent, and black share strictly above 50 percent. The first and fourth columns controls only for price and occupancy of the address in 1930. The second and fifth columns add controls share renters and total number of addresses at the block level, and share black, share immigrant, share laborer, mean age, median home value, median rent, and median occupational score at the neighborhood level. The third and sixth columns drop the neighborhood controls and includes ED fixed effects.