NBER WORKING PAPER SERIES

DISTINGUISHING CAUSES OF NEIGHBORHOOD RACIAL CHANGE: A NEAREST NEIGHBOR DESIGN

Patrick Bayer
Marcus D. Casey
W. Ben McCartney
John Orellana-Li
Calvin S. Zhang

Working Paper 30487
http://www.nber.org/papers/w30487

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2022, Revised December 2024

Distinguishing Causes of Neighborhood Racial Change: A Nearest Neighbor Design
Patrick Bayer, Marcus D. Casey, W. Ben McCartney, John Orellana-Li, and Calvin S. Zhang
NBER Working Paper No. 30487
September 2022, Revised December 2024
JEL No. J15, R23, R31

ABSTRACT

We study neighborhood choice using a novel research design that contrasts the move rate of homeowners who receive a new different-race neighbor immediately next-door versus slightly further away on the same block. This approach isolates a component of household preferences directly attributable to their neighbors' identities. Both Black and White homeowners are more likely to move after receiving a new different-race neighbor. These findings are robust to additional controls (e.g., income) and alternative research designs. We find evidence of heterogeneity in responses, especially associated with housing density, which has implications for understanding contemporary neighborhood racial change and the prospects for maintaining stable, integrated neighborhoods.

Patrick Bayer
Department of Economics
Duke University
213 Social Sciences
Durham, NC 27708
and NBER
patrick.bayer@duke.edu

Marcus D. Casey
Department of Economics
University of Illinois at Chicago
601 South Morgan UH725 M/C144
Chicago, IL 60607
mcasey@uic.edu

W. Ben McCartney
University of Virginia
Rouss Hall and Robertson Hall
140 Hospital Dr
Charlottesville, VA 22903
mccartney.ben@gmail.com

John Orellana-Li
The Graduate Center
City University of New York
365 5th Avenue
New York, NY 10016
jorellana@gradcenter.cuny.edu

Calvin S. Zhang
University of Oregon
calvin.zhang@phil.frb.org

1 Introduction

Two decades into the 21st century, US neighborhoods remain racially stratified. This phenomenon is well-documented, yet there remains little consensus on how the identities of neighbors themselves drive racial sorting and persistent stratification in modern housing markets. At issue is whether the attributes of individual neighbors directly enter household residential choices. Conventional tipping theories since Schelling (1971) emphasize preferences over the racial composition of the neighborhood as a principal catalyst for neighborhood turnover. This widely accepted understanding of the relationship between race and neighborhood transition has been complicated by recent work noting that neighborhood composition changes are often strongly bundled with gentrification and related processes where individual racial preferences may be unimportant.[1] These observations coupled with an ostensible expansion in progressive public–facing racial attitudes have led some observers to argue that in contemporary housing markets, a neighbor's race may be of secondary or almost no importance in individual household neighborhood choices despite the persistence of racial stratification. Krysan and Crowder (2017), for example, discusses how seemingly race-neutral neighborhood search processes can result in maintaining stratified neighborhoods while Gould Ellen (2000) notes the absence of traditional White flight behavior in response to Black entry in contemporary housing markets.[2] Understanding the source of the dynamics maintaining stratification is no idle concern. A broad literature establishes that neighborhoods, and who lives in them, matter for a wide range of economic, social, and health outcomes (Chetty et al., 2016, 2022a,b; Chyn and Katz, 2021).

Credibly distinguishing whether households respond directly to the attributes of their neighbors or factors coincidental with new neighborhood entrants has proven difficult. Two fundamental challenges confound identification of preferences over the identities of new neighbors. First, neighborhood demographic change is typically accompanied by shifts among a wide range of other neighborhood amenities, including public goods (e.g. schooling and safety) and private goods and services (e.g. shops and restaurants). Many important amenities are both likely to be unobserved and to respond endogenously to a change in neighborhood racial composition. Second, household location decisions are naturally affected by both current conditions and future expectations. Some current residents may perceive new entrants of a different race negatively because of concerns about future

[1] Couture et al. (2019) and Baum-Snow and Hartley (2020) document, for example, the sharp rise in demand for center-city neighborhoods by high income households since 1990, while Almagro and Dominguez-Iino (2020) highlights how both neighborhood demographic composition *and* nearby amenities change endogenously in response to the rapid expansion of AirBnB in major cities in recent years.

[2] Instead Gould Ellen (2000) attributes the persistent stratification primarily to avoidance of Black neighborhoods on the part of White households.

entry by others of that group (see e.g., Gould Ellen (2000) and Casey (2020)). Together these factors make it difficult to determine empirically whether households are reacting to the *identities* of their neighbors directly or instead using information about neighborhood composition to form expectations about the future evolution of the neighborhood.

Evidence from surveys that attempt to directly measure preferences by providing "neighborhood cards" describing different configurations of racial integration introduced in Farley et al. (1978) and subsequent followup work (e.g., Farley and Frey (1994); Logan et al. (2004)) for the Detroit area notes that although White tolerance for Black neighbors has increased, surveyed households voiced discomfort with the prospect of a majority Black neighborhood.[3] While useful, this hypothetical evidence provides limited insight into how such stated preferences play out in reality, as surveys necessarily have to abstract from other features of the neighborhood and respondents may find it difficult to "hold all else equal." In recognition of the need of better estimates of preferences, several recent papers propose research designs aimed at distinguishing preferences for neighborhood composition from other neighborhood attributes/amenities including Almagro et al. (2021), Caetano and Maheshri (2021), Davis et al. (2021), and Li (2021).

In this paper, we study how incumbent residents respond to the receipt of new different-race neighbors to better understand the nature of these responses and how they may shape the character of neighborhood racial change in contemporary housing markets. To do so, we propose a novel approach that exploits highly localized variation in exposure to new neighbors on a residential block. Our primary research design contrasts the propensity to move in response to receiving a new neighbor of a different race immediately next door versus slightly further away on the same side of street on the same block.[4]

We motivate this approach with a dynamic model of neighborhood choice presented in Section 2. We use the model to highlight several fundamental challenges to identifying preferences for neighbors' race. We formally introduce our "nearest neighbor" research design in Section 3 and show how,

[3] By contrast, Black respondents were fine with a broader array of neighborhoods, though they did express distaste for nearly all-White contexts. Charles (2000), extending this approach to Los Angeles to measure preferences for integration in a multiracial context, finds that all groups express a preference for living in neighborhoods with some degree of integration. However, Whites and non-Black groups admitted discomfort at the prospect of living in neighborhoods with higher shares of Black neighbors.

[4] McCartney et al. (2024) use a closely-related research design based on contrasting the move decisions of households within the same census block group who receive a new Democratic or Republican neighbor to study political polarization. We explore a similar approach. Bayer et al. (2008), Bayer et al. (2021), and McCartney and Shah (2022) use a related research design based on contrasts between households on the same block versus a block or two away to study the role of neighborhood social interactions on job referrals, household finance, and investment activity. In these papers, the thinness of the owner-occupied housing market provides the primary basis for the argument that the assignment of neighbors at such fine geographic scales is essentially random.

under reasonable identifying assumptions, it overcomes these identification challenges to isolate a component of move propensities attributable to preferences for the neighbors' race. In the empirical analysis that follows, we provide supporting evidence that these identifying assumptions do, in fact, hold in our sample.

We implement our strategy using a national set of neighborhoods experiencing racial change using data that combine detailed housing transactions with demographic information on households available in home mortgage loan application registry (LAR) files collected as required by the Home Mortgage Disclosure Act (HMDA).[5] The housing transactions data provide information about the housing unit as well as the exact timing of home sales. The HMDA files provide important demographic details on buyers including race of the applicant and/or co-applicant, income, and key loan-related information. We match these data to construct a house-level panel dataset that allows us to observe racial transition dynamics at the block level. We focus specifically on the well-documented, historically salient, and stubborn Black-White neighborhood stratification.[6] Moreover, the places we study using our research design include those with the highest potential to foster stable, integrated neighborhoods.

Our baseline results reveal that both Black and White homeowners have a higher propensity to move in response to receiving a new different-race neighbor immediately next door versus just two to three doors away. The magnitude and statistical significance of these results are essentially unaffected by the inclusion of a broad set of building, homeowner, and mortgage characteristics. For Black households, the implied effect size corresponds to a roughly six percent higher move propensity, while the corresponding estimate represents a three percent increase for White households. Importantly, these estimates isolate only the incremental response of incumbent homeowners residing right next to the new neighbor relative to other homeowners a few doors further away. The latter homeowners may, of course, also react to the new different-race neighbor and, as a result, our estimate naturally provides a lower bound on the magnitude of the full response to the receipt of a new different-race next-door neighbor.

Next, we estimate an alternative set of models that expand the set of homes considered beyond the same side of the immediate block. We examine responses to receiving a new nearest neighbor

[5] To combat redlining and mortgage discrimination, the Home Mortgage Disclosure Act, enacted originally in 1975 and expanded several times since, requires financial institutions to collect and disclose information on mortgage applicants and borrowers.

[6] The continued salience of Black-White stratification and its significance in governing the dynamics of modern housing markets remains a topic of active research. See, e.g., Lewis et al. (2011), Hwang and Sampson (2014), or Christensen and Timmins (2022).

of a different race among the 40 nearest parcels by distance within the same census block group. These parcels potentially include homes down the block, across the street, and behind the home of interest. Similar to our baseline strategy, we find evidence consistent with a hyper localized response. In particular, both Black and White incumbent households receiving a new different race neighbor within the 1 or 2 closest parcels are statistically more likely to leave than if the new different-race neighbor arrived among the 3rd, 4th, or 5th closest parcels. The magnitude of these results are similar to the baseline within-block results.

There are several potential challenges to interpreting these reactions as a response to the race of a new neighbor. We address four major concerns in our analysis. First, there might be non-random selection in where the new neighbor arrives on a block, even at the fine geographic scale that we use for these empirical contrasts. Using a set of observable homeowner and housing characteristics, we show that the arrival of a new neighbor one versus two to three doors away appears to be as good as randomly assigned. A second potential concern is that the move responses we identify might somehow be driven by within-block changes in amenities that vary even at the fine geographic scale we study. To test for this possibility, we examine how the receipt of a different race neighbor one versus two to three doors down affects subsequent resale housing prices. This analysis results in a precisely-estimated null effect for both Black and White current residents, providing empirical support for the claim that our baseline results are not capturing changes in highly localized variation in unobserved aspects of housing or neighborhood quality.

A third potential concern with the interpretation of our baseline estimates is the possibility that incumbent residents may be responding to the general disruption arising from the receipt of any new next-door neighbor, regardless of the neighbor's race. To study this possibility, we present results from an alternative estimation approach that compares current residents who received a different-race neighbor with those who received a same-race neighbor in the same Census block group and quarter. For both Black and White incumbent households, we again find evidence of differential move responses among those receiving a new different race neighbor immediately next door versus a few doors away. These estimates are statistically significant with magnitudes similar to our baseline estimates. In contrast, the estimated responses of both Black and White incumbent homeowners to receiving a new same-race neighbor immediately next door are less than half the magnitude.

A final potential concern about our baseline results is that the estimated response to receiving a new neighbor of a different race may not be capturing a racial response per sé, but instead reflect household preferences for other attributes of their neighbors, such as income, which are correlated

with race. To address this concern, we are able to leverage the income measures provided in the HMDA data. Adding income as a control to the specifications described in the previous paragraph reveals a strong income response – incumbent move propensities are inversely related to the income of their new next-door neighbor – but, remarkably, has no effect on the estimated response to the race of the new neighbor. That controlling for income does not affect the estimated racial response is directly attributable to the extremely fine geographic contrast at the heart of our research design. That is, despite the strong correlation of income and race at the population level, income is essentially randomly assigned among new different-race neighbors one versus two to three doors away in the same Census block group. The robustness of our result to income greatly strengthens the likelihood that the differential responses we estimate to the receipt of new different-race neighbors do, in fact, represent responses to neighbors' race rather than some other correlated attribute.

We close the paper by studying heterogeneity in the racial responses estimated in our baseline results. The main goal of this final section is to provide a sense of the places in American society where neighborhood race-based preferences remain strong drivers of stratification versus those spaces where such preferences are weaker and, therefore, may be more amenable to stable racial integration. The most compelling and robust heterogeneity that we estimate is related to housing density. In particular, we find that the estimated move response to a new different race neighbor is especially strong when homes are in very close proximity, especially for White incumbents receiving a new Black neighbor. This response decays sharply with distance, falling to zero when homes are beyond approximately 25 meters apart. In this way, racial responses are especially strong at urban densities and much weaker at suburban densities. This finding is consistent with recent descriptive evidence that shows declining racial segregation and increases in racial tipping points in suburban areas of many American cities (Bartik and Mast, 2021).

For income heterogeneity, we explore how racial responses vary with the income of the (i) incumbent resident, (ii) new neighbor, and (iii) neighborhood. These results are somewhat noisier but suggest that racial responses are greater when the incumbent resident is relatively high-income and when the new different-race neighbor is relatively low-income. Finally, we estimate stronger racial responses by White homeowners in Northern versus Southern states. This result is in line with the especially high levels of racial segregation, historically and currently, in Northern cities (Boustan, 2010; Derenoncourt, 2022), as well as increased racial integration in the suburbs of Sunbelt cities.

Our paper makes several broad contributions to the existing literature. First, the empirical results provide new evidence that race remains an independent force shaping household sorting

decisions in contemporary housing markets. This result has important implications for racial inequality in current and future generations, since (i) racial sorting drives enormous differences in neighborhood quality for Black and White households with identical levels of household income and wealth (Aliprantis et al., 2022; Bayer and McMillan, 2005) and (ii) differences in neighborhood quality have substantial causal effects on many social and economic outcomes, especially for children (Bayer et al., 2008; Chetty and Hendren, 2018; Chetty et al., 2016; Chyn, 2018; Chyn and Katz, 2021). Putting these two effects together, Chetty et al. (2020) shows that neighborhood differences have an independent effect, over and above parental differences, on the Black-White gap in absolute intergenerational mobility. Likewise, Chetty et al. (2022a) and Chetty et al. (2022b) demonstrate the importance of social connectedness within neighborhoods for upward mobility outcomes.

Our results also imply that immediate responses to the racial identity of neighbors – which may be compounded, in turn, by accompanying changes in amenities, prices, and expectations – contribute to the dynamic patterns of racial tipping, "White flight", and neighborhood racial transition (documented in, for example, Blair (2017); Boustan (2010); Card et al. (2011); Casey (2020); Derenoncourt (2022); Gould Ellen (2000)), making it difficult to sustain racially integrated neighborhoods. An extensive literature in economics has documented the causal benefits of racial integration (Billings et al., 2022; Johnson, 2011, 2019; Reber, 2011; Tuttle, 2019) and costs of residential segregation (Ananat, 2011; Andrews et al., 2017; Chyn et al., 2022; Cox et al., 2022; Lutz, 2011), for both Black and White children. In addition, recognition of the potential benefits of racial integration has spurred the passage of a number of public policies aimed at broadening access to better neighborhoods and fostering stable neighborhood integration and these results suggest nuance in interpreting the efficacy of such policies in modern housing markets.[7] We discuss the implications of our findings for the sustainability of racially integrated neighborhoods further in the concluding section of the paper.

A final empirical implication of our paper is that social interactions that occur at the level of one's immediate neighbors remain important in a modern context. Interactions at this hyper-local level have motivated identification strategies for estimating social interactions and local spillovers in Bayer et al. (2008), Anenberg and Kung (2014), Bayer et al. (2021), and McCartney and Shah (2022) and the neighbor-based segregation index developed by Logan and Parman (2017).

Our paper also makes a methodological contribution to the literature that has attempted to distinguish the role of racial preferences in neighborhood sorting. In Section 3, we discuss in detail how

[7] These interventions broadly include anti-discrimination legislation such as the Fair Housing Act of 1968, housing vouchers, scatter site housing, and related policies. See Galster (1992) and Massey and Denton (1993) for additional discussion.

our research design works to solve the problem of separately identifying direct preferences for the attributes or identities of one's neighbors from associated changes in neighborhood amenities and expectations about the future evolution of the neighborhood. This issue is well known in the literature on neighborhood sorting and many papers – e.g., Bayer et al. (2007) – simply acknowledge the inability to distinguish these components as a limitation of the analysis.

Another issue that naturally arises in the neighborhood sorting literature is how to distinguish the role of preferences in household location choices from various forms of housing discrimination.[8] Bayer et al. (2007) and Krysan and Crowder (2017) discuss the conceptual difficulty of differentiating whether observed neighborhood choices are driven by preferences of households for neighborhood composition versus discriminatory constraints that effectively restrict the options available to certain households.[9] By focusing on the exit decision of existing homeowners in response to changes in the demographic characteristics of their neighbors, our methodological approach avoids the difficulty of trying to identify whether neighborhood entry choices are driven by preferences or constraints.

2 A Conceptual Framework

We begin by introducing a stylized theoretical model of an existing homeowner's decision to remain in their neighborhood or move, as their neighborhood evolves. We assume this decision is inherently dynamic insofar as the homeowner bases their decision on both current neighborhood conditions and their expectations about how the relevant factors will evolve going forward. We use the homeowner's dynamic decision problem to highlight two fundamental identification issues facing researchers seeking to study whether the neighbors' identities directly affect residential location decisions. We formally introduce our nearest neighbor research design in the next section of the paper, explaining there how it helps to overcome these important identification issues.

We characterize the dynamic problem of a homeowner deciding whether to stay in their existing residence or move in each period. Households are forward-looking with preferences defined over the characteristics of their neighborhood, including the composition of their neighbors. Households also understand that neighborhoods are constantly evolving and form expectations about how their

[8] An extensive literature has documented housing discrimination through many channels including differential willingness to rent/sell to Black renters/buyers and neighborhood steering by real estate agents (Bayer et al., 2017; Christensen and Timmins, 2021b, 2022; Hanson and Hawley, 2011; Ondrich et al., 2003; Page, 1995; Yinger, 1986).

[9] Christensen and Timmins (2021a) uses a novel combination of an audit study and choice date to estimate a model that simultaneously captures both racial preferences and housing discrimination. Li (2021) uses detailed historical data to separately identify the role of preferences versus constraints in driving segregation in Northern cities during the Great Migration.

neighborhood is likely to change going forward given its current state.

We model the decision of an existing homeowner i with observable attributes Z_i to stay or leave their current residence in neighborhood j as a dynamic binary choice model in discrete time. We characterize the per-period utility, U that i receives from their current neighborhood at time t as:

$$U_{i,j,t} = f(Z_i, p_{i,j,t}, X_{j,t}, \xi_{j,t}, \alpha) + \sum_k g(Z_i, Z_{k,t}, D_{i,k}, \beta) + \epsilon_{i,j,t} \tag{1}$$

where:

- $f(\cdot)$ captures utility from neighborhood amenities, both observed $X_{j,t}$ and unobserved $\xi_{j,t}$, as well as the value of household i's home, $p_{i,j,t}$,
- $g(\cdot)$ captures utility associated with the attributes of each neighbor k, located a distance $D_{i,k}$ away,
- $\epsilon_{i,j,t}$ captures the idiosyncratic taste of household i for living in their current residence.

This general formulation of the per-period utility function allows homeowners to have preferences over both neighborhood amenities and the identities/attributes of their neighbors, and to potentially care more about their immediate neighbors than those a bit further away. The flexible form of $f(\cdot)$ also permits homeowners to care about the value of their home both as a measure of the cost of living in this location and because they benefit from any appreciation that occurs over time.

The Unobserved Amenity Problem. Inspection of the components of the per-period utility function illustrate a principal reason why it is challenging to separately identify and isolate independent causes of household move decisions. Since many neighborhood amenities are likely to be unobserved in any data set, distinguishing preferences for the identities/attributes of one's neighbors – captured in $g(\cdot)$ – from tastes for unobserved neighborhood amenities in ξ is difficult. Are households responding directly to the changing identity of their neighbors or to other aspects of the neighborhood – e.g., schools, shops, restaurants, churches – that may be changing at the same time? Answering this question is made all the more difficult by the fact that many of these amenities may evolve endogenously in response to changes in the neighborhood demographic and socioeconomic composition. Such endogenous amenities include public goods like school quality and public safety as well as any private goods and services that respond to local demand such as local restaurants, shops, and churches that may be highly correlated with neighborhood demographics.

The Neighborhood Expectations Problem. Given the characterization of static utility above, we can recursively define the household's present discounted value of remaining in their current

residence $V_{i,j,t}$ using the Bellman equation:

$$V_{i,j,t} = f(Z_i, p_{i,j,t}, X_{j,t}, \xi_{j,t}, \alpha) + \sum_k g(Z_i, Z_{k,t}, D_{i,k}, \beta) + \delta EMax(V_{i,j,t+1}, 0) + \epsilon_{i,j,t} \quad (2)$$

where δ is the discount rate and for simplicity, and without loss of generality, we normalize the value of moving away to zero.

The right hand side of the Bellman equation highlights a second challenging identification problem related to the dynamic nature of the decision problem. In particular, the continuation value term $EMax(V_{i,j,t+1}, 0)$ in equation 2 captures expectations about the future evolution of the neighborhood along a number of dimensions, including household i's house price p. The inclusion of this term in the homeowner's decision problem makes it difficult to ascertain whether households respond to changes in neighborhood amenities and demographic composition because of the direct effect on their own utility or because they provide new information about the future evolution of the neighborhood. In the context of racial tipping, for example, this forward-looking behavior makes it difficult to distinguish whether households care directly about the race of their neighbors or are motivated instead by what neighborhood demographic change might signal about the future evolution of the neighborhood. In many historical contexts, in particular, fears about future price depreciation have been put forth as a primary explanation/justification for "White flight" in reaction to the initial entry of Black residents on a block.

These fundamental identification problems are the primary reason the question of what fundamentally drives neighborhood racial change remains an open academic and policy question more than half a century after realtors openly practiced blockbusting in many American cities and Schelling (1971) formalized a dynamic model of neighborhood tipping. No existing paper has been able to fully separate the independent contribution of the identity of one's neighbors from the associated simultaneous (and potentially endogenous) changes in local amenities and what those identities might signal about the future evolution of the neighborhood.

3 Nearest Neighbor Research Design

The research design we use in this study is based on an empirical contrast between the behavior of existing homeowners who reside on the same residential block. Specifically, we compare how households of the same race react to receiving a new neighbor of a different race depending on whether the new different-race neighbor moves in immediately next door versus slightly further away on the

same side of the street on the same block. In practice, our main empirical results focus on comparing the reaction of current residents to receiving new next-door neighbors to those receiving a new neighbor just two to three doors away.[10] Intuitively, this fine geographic contrast directly addresses the fundamental identification challenges discussed above. First, because these homeowners live so close to one another, any differences in the neighborhood amenities they experience are likely to be very small and idiosyncratic. And, second, the arrival of the new neighbor should provide nearby existing homeowners with the same new information about how the neighborhood is likely to evolve over the coming years, regardless of whether they move in one versus two to three doors away. As a result, any systematic differences in move propensities in response to the arrival of the new neighbor can plausibly be attributed to differences in preferences related to proximity to the new neighbor.

3.1 A Potential Outcomes Framework

We introduce the nearest neighbor research design using a potential outcomes framework. Our main goal is to characterize the identified causal effect and the key identifying assumptions underpinning the design in a concise way. We discuss the implications of slightly weaker assumptions for the interpretation of the identified effect and foreshadow some empirical tests of these identifying assumptions that we implement later in the paper.

Consider existing homeowners i of race r who reside on block j with neighboring homes d doors away. We model each incumbent household's move propensity, Y, in response to the arrival of a new neighbor of a different race r' arriving d doors away as:

$$Y_i(r',d) = P(r',d) + \beta_i A(r',d) + \nu_i + \omega_j \tag{3}$$

where the four factors on the right hand side of equation 3 capture the impact on move propensities due to:

- preferences for living near a different-race r' neighbor d doors away: $P(r',d)$,
- *within-block* differences in future amenities or housing prices related to the arrival of the different-race r' neighbor d doors away: $\beta_i A(r',d)$,
- idiosyncratic factors affecting household i's move propensity: ν_i, and
- factors affecting block j as a whole, ω_j, including any expected future changes to amenities, composition, and housing prices at the block level related to the arrival of the new neighbor.

[10] In some specifications, we expand this contrast to within 2 parcels away relative to 3, 4, or 5 parcels away.

The nearest neighbor research design seeks to estimate the component of move propensities attributable to preferences for living near a different race neighbor, $P(r',d)$, by contrasting the move responses of incumbent households immediately next door, $d = 1$, versus those a few doors away, to the arrival of a new neighbor of race r'. We are specifically interested in identifying $P(r',1)$, which requires three assumptions:

Assumption 1: Quasi-Random Arrival Location within the Block ($d \perp\!\!\!\perp v$). The first key assumption underlying the nearest neighbor design is that $d \perp\!\!\!\perp v$ – i.e., the location of the newly-arrived neighbor of race r' is quasi-randomly assigned to existing owners residing a few houses from the new neighbor on the same block.[11] Our empirical focus on owner-occupied single family homes is motivated by this requirement.[12] Two aspects of the market for single family homes help make this a reasonable assumption empirically. First, while the size, age, and quality of single family homes vary greatly across a city, homes are much more homogeneous on a given residential block. Second, the market for single family homes is typically quite thin, with only a small fraction listed for sale at a given moment in time. As a result, while a household might have a preference for a particular type of home in a particular neighborhood, the exact home they wind up buying within a block is largely a function of homes listed for sale at the time of their search.

This assumption is testable on observable household and housing attributes, although obviously not on unobserved factors. As we show in the empirical analysis, there are no systematic differences in the observable attributes of incumbent households one versus two to three doors away from a new neighbor of a different race on the same block in our large sample.

Assumption 2: Limited Geographic Scope ($P(r',d) = 0 \ \forall \ d > 1$). This assumption requires that only close next-door neighbors are directly affected by the racial identity of the newly-arrived neighbor, over and above any impact on the future amenities, composition, and price of the block j as a whole. In particular, this assumption ensures that the component of move propensities due directly to preferences over the racial identity of the new neighbor can be identified by comparing the responses of immediate next-door neighbors, $d = 1$, and those just a few houses away – e.g., $d \in (2,3)$. Admittedly, this requirement is quite strong and unlikely to be satisfied in all but the most extreme circumstances.

[11] The independence assumption also implicitly requires all incumbent households to observe the arrival of the new neighbor, allowing them to update beliefs about the expected future evolution of the neighborhood accordingly.

[12] Our estimation sample also includes townhomes and a small number of multiplexes, if the units are arranged such that we can confidently identify next-door neighbors.

A weaker, and more reasonable, version of this assumption is that the preferences of neighbors two to three doors down are affected by the new neighbor in a qualitatively identical but quantitatively weaker way, $P(r',1) > P(r',d)\ \forall\ d > 1$.[13] In this case, our estimated effect provides a *lower bound* on the true strength of the move response to a new next-door neighbor due directly to preferences over their racial identity.

Assumption 3: No Effect of Arrival Location on Expectations of Future Amenities or House Prices ($E[A(r',d)] = 0\ \forall\ d$). This final key identifying assumption ensures that all households on block j within a few homes of the newly-arrived neighbor of race r' experience the same (block-level) impact on expectations regarding future amenities and house prices captured in ω. We use the term amenities broadly here to include anything that affects the value households receive from their residential location besides that due directly to the identities/composition of their neighbors. For most local public and private goods this is likely to be a reasonable assumption empirically, given the fine geographic scale of the analysis – i.e., differences in distances to shops, churches, transportation, and employment opportunities will generally be quite small between same-block neighbors living just a few doors apart from one another. All homes on the same side of a residential block are also almost always assigned to the same local schools and the likelihood of crime victimization is unlikely to vary much within a residential block.[14]

While we do not observe measures of neighborhood amenities that vary within blocks, we are able to examine whether proximity to the newly arriving neighbor, d, is correlated with future home sales price in the empirical analysis below. Housing price is an especially attractive variable because it serves as a meaningful summary statistic for the combination of all neighborhood amenities and housing attributes. We find negligible effects in our analysis, suggesting that there is essentially no correlation between localized neighborhood amenities and d. This result also suggests that the racial identity of the newly-arrived neighbor does not have a meaningful impact on future housing prices of immediate next-door neighbors versus those a few doors away, which helps rule out any concern that incumbents are potentially responding to any (positive or negative) changes to the neighbors' home or property rather than their racial identity.

[13] This statement of the assumption implicitly assumes that $P(r',d) > 0\ \forall\ d$.

[14] A somewhat subtler point about this assumption is that arrival distance must not affect *expectations* about future amenities and prices. Even if there is no differential effect on future amenity levels, this assumption might fail if distance affects the salience of the information provided by the arrival of the new different race neighbor. An attractive feature of the hyper-local contrast between neighbors arriving one versus 2-3 doors away that we use in the empirical analysis below is that the salience of the arrival of the new neighbor is likely to be similar over such small distances.

3.2 Implementing the Nearest Neighbor Design

The most straightforward implementation of our nearest neighbor research design would be to observe J blocks, each of which contain exactly two incumbent homeowners, a and b, of race r, living $d = 1$ and $d = 2$ doors away, respectively, from a newly-arriving neighbor of race $r' \neq r$. We refer to the next-door neighbor a as being "treated" and the slightly more distant neighbor b as a "control". Differencing the move propensities of the treated and control households yields:

$$\big(Y_a(1) - Y_b(2)\big) = \big(P(r',1) - P(r',2)\big) + \big(\beta_a A(r',1) - \beta_b A(r',2))\big) + \big(\nu_a - \nu_b\big) + \big(\omega_j - \omega_j\big) \tag{4}$$

Assumption 2 reduces the first term on the right hand of this equation to $P(r',1)$. Assumption 3 eliminates the second term. The fourth term drops out, leaving:

$$Y_a(1) - Y_b(2) = P(r',1) + \big(\nu_a - \nu_b\big) \tag{5}$$

Assumption 1 implies that $\big(\nu_a - \nu_b\big) \perp\!\!\!\perp P(r',1)$. Thus, averaging over the J blocks provides the estimated treatment effect, $\overline{Y(1) - Y(2)}$, which converges to $P(r',1)$ as the number of blocks goes to infinity.

Our empirical approach approximates this direct implementation. Specifically, we estimate the following equation:

$$M_{i,j,t} = \pi_1 I(r', d = 1)_{ijt} + \pi_2 I(r', d = 2/3)_{ijt} + \pi_3 I(r', d = 4 \text{ or more})_{ijt} + \lambda Z_{i,j,t} + \omega_{j,t} + \nu_{i,j,t} \tag{6}$$

where:

- $M_{i,j,t}$ is an indicator ($\times 100$) for whether household i moves within a given time period following the receipt of a new different race r' neighbor.
- $I(r', d = k)_{ijt}$ indicate whether a new different-race r' neighbor moves in k={1,2/3,4+} doors from household i, on the same side of the street within block j.
- $Z_{i,j,t}$ are any other time-varying observable factors at the household or block level that might affect the likelihood of moving.
- $\omega_{j,t}$ are block-by-quarter fixed effects.

Our parameter of interest is $\pi_1 - \pi_2$, which characterizes the additional propensity to move in response to receiving a new different-race neighbor immediately next door versus two to three doors away. The inclusion of block-by-quarter fixed effects ensures that this parameter is identified only by

comparisons of households residing on the same block during the same time period. All of the tests of the identifying assumptions mentioned above and shown below are conducted using the same structure for the right hand side of this equation.[15]

As equation 6 makes clear, our main specification includes three potential treatments at distances 1, 2/3, and 4+ doors away, respectively, from the incumbent's house. While the hyper-local contrast, $\pi_1 - \pi_2$, is our main parameter of interest, the inclusion of the third treatment is helpful for adding statistical power – i.e., for increasing the number of experiments and more precisely estimating block-by-quarter fixed effects. As Borusyak and Hull (2023) show, a potential issue with including more than two treatments is that Angrist (1998)'s result ensuring a weighted-average interpretation for regression estimands may no longer hold. To address this potential concern, we report results below for alternative specifications that reduce the number of treatments to two by either dropping the third treatment or combining the second and third ones into a single control group.

Race Versus Other Household Covariates. A final issue worth noting in the context of laying out our research design is the interpretation of the estimated response to a neighbor's identity as a reaction to the neighbor's race, per sé. It is important to recognize that we observe only a small set of observable attributes for the homeowners in our sample. Thus, it will be impossible, in general, to rule out that any reactions we detect are instead related to other unobserved neighbor characteristics correlated with race. That said, we do observe a measure of household income at the time of entry into the neighborhood. And, strikingly, controlling separately for both current resident and relative difference between the new neighbor's and the current resident's income in the analysis has little effect on estimated response to receiving a new neighbor of a different race. This ostensibly surprising result occurs because, despite the strong correlation of income and race in the population, there is no systematic relationship between income and proximity to the newly arriving different-race neighbor, d, at the finely differentiated distances we study. To the extent this holds for other household attributes, our analysis will, in fact, isolate a response to new different-race neighbors that is indeed attributable to race itself.

[15] Note that including control variables Z is not required for identification but can be used and to improve the precision of the estimates, as well as an implicit test of the random arrival location assumption. It is also worth noting that when treatment effects are heterogeneous the regression coefficient recovers a convex average of treatment effects, with weights that depend on the conditioning variables, including the fixed effects (see, for example, Angrist (1998) and Borusyak and Hull (2023)).

3.3 The Strengths and Limitations of the Nearest Neighbor Design

In our view, the great advantage of the nearest neighbor approach is that it can be used to isolate a component of the move response to the receipt of a new different-race neighbor that can be attributed to preferences for the neighbors' racial identity under a set of assumptions that are both reasonable and can be tested/refuted in the data. Identification of such responses in a manner that holds constant changes in neighborhood amenities and prices (both current and expected future) has proven extraordinarily difficult in the literature to date.

That said, there are several limitations of the nearest neighbors approach that are important to emphasize. First, as mentioned above, under the weaker, and more reasonable, version of the Limited Geographic Scope assumption, the estimated causal effect provides a lower bound on the strength of preference-related response to the receipt of new next-door neighbor of a different race. Second, the mapping between value functions in Section 2 and the move probabilities in equation 3 will generally be a function of the variance of ϵ and all of the other components of the value function. Without estimating a fully-specified dynamic model, it is impossible to translate differences in move probabilities back to structural preference parameters, or to convert them to dollar values or other meaningful units for assessing differences in welfare.[16]

Finally, in the presence of heterogeneous treatment effects, the estimated causal effect will be a weighted average of the treatment effects for incumbent households living on blocks where experiments occur. As we show below, and not surprisingly, more racially integrated neighborhoods are over represented in the sample of experiments, which raises the obvious possibility that the set of treated households are more racially "tolerant" and may have weaker responses to the receipt of a new different-race next-door neighbor than, for example, households residing in very segregated neighborhoods.[17]

4 Data

The data we use to study household moves and neighborhood dynamics are drawn primarily from two sources. The first consists of detailed housing transactions collected and made available by CoreLogic

[16] Researchers and policy makers may also be interested in other measures of the response to a new different-race neighbor – e.g., a total response rather than the component due only to preferences over the neighbors' racial identity. Our focus lies in trying to isolate the latter response, while other papers may seek to identify a different "estimand."

[17] Other aspects of the experimental design may push in the other direction, however, if, for example, the blocks where experiments occur have more churn and, as a result, incumbent households are generally closer to the margin of moving relative to the typical household.

Real Estate Solutions. These data feature the near-universe of housing transactions for counties with digitized records, which by the mid-2000s includes almost all counties in the country. These data include substantial information on homes including: actual transaction prices, transaction closing date, structural characteristics such as square footage and year built as well as each home's precise location. Importantly, each home is assigned a unique identifier that allows observation of a home and any change in ownership over time.

We match these data to information available in publicly-available loan application registry (LAR) archives collected as required by the Home Mortgage Disclosure Act of 1975 (HMDA). The HMDA files focus particularly on mortgage applicant data relevant to monitoring potential redlining and mortgage discrimination behavior by lenders including race, ethnicity, gender, and household income of all applicants and co-applicants. Additional house and financial variables such as the transaction date, the purpose of the loan (purchase or refinance), the census tract of the home, and characteristics of the loan originated are reported as well. These variables help facilitate the match with the transactions data.

The CoreLogic and HMDA match uses a multi-step algorithm that exploits several key variables including the census tract of the home, the loan amount, the purpose of the loan, the year of the transaction, and name of the lender who originated the home.[18] The procedure initially matches the LAR files from HMDA with the lender information in the transactions data, then the two datasets are joined using the matching keys.[19] Overall, the match procedure is generally successful with nearly 60 percent of all mortgages in our CoreLogic sample matched to a mortgage application in the HMDA data. Some homeowners remain unmatched due to either having no unique match in the HMDA data or no transaction to match to, as happens in the case of cash purchases. These omissions reduce our sample size but do not threaten our strategy for identifying differential move responses since our empirical analysis focuses only on circumstances in which both incumbent homeowners and newly arriving neighbors have successful HMDA matches, which are necessary for characterizing the homeowner's race and household income.

We focus on transactions that occurred between 2005 and 2023 to build a panel of parcels and their owners across time. To build the panel, we use the full sample of mortgages when matching with HMDA. Crucially, not all mortgages correspond to changes in ownership and not all ownership changes involve mortgages. For our purposes, a transaction represents a change in ownership if it

[18] A number of papers in the literature use matched HMDA-transactions data. See, e.g., Bayer et al. (2017, 2018, 2016).
[19] We provide additional details on the match in Online Appendix A.

does not include a mortgage tagged as a refinance and does not represent a transfer of a property into a trust, a deed correction, or a quitclaim. New neighbors are therefore defined as those who arrive as a consequence of transactions involving an ownership change. Likewise, we define a move as a transaction involving a current resident involved in an ownership change that is not tagged as a foreclosure.

4.1 Identifying Nearest Neighbors

Our empirical strategy depends on correctly defining nearby neighbors. To properly assign neighbors in our estimation sample, our primary approach uses an algorithm that exploits the availability of exact addresses for each home. The algorithm first uses sequential numbering within a block to assign homes on the same side of the street as 1 door away, 2 doors away, 3 doors away, and beyond. To deal with cases of imperfect sequential numbering among the street addresses and other non-standard configurations, the algorithm imposes additional restrictions using Euclidean distances between properties.[20] We also use an alternative approach of assigning nearby neighbors that explicitly uses distance, "as the crow flies," relative to each home in the sample. Specifically, we characterize the 40 nearest parcels in the same block group as the home.

Figure 1 illustrates these two approaches to assigning nearby neighbors. The figure features a common configuration of a set of blocks within a neighborhood in many areas within the US and is typical of many of the neighborhoods where our experiments occur. The unnumbered black house is our house of interest. Our primary approach defines the "nearest" neighbor as homes next door and on the same side of the street. This definition corresponds to the homes shaded in dark gray in the figure. The homes shaded in light gray are defined as the homes two or three doors down and thus would serve as control comparisons in the context of our house of interest receiving the treatment of a different-race neighbor next door.

[FIGURE 1 HERE]

The alternative approach we use can be illustrated by considering instead the numbered homes in the figure. The number corresponds to the distance away from the house shaded black, i.e., the house numbered 1 is the closest house and the house numbered 7 is the 7th closest house. In this case, and in most cases in the data, the two closest homes are also the two homes next-door. However, this definition allows us to consider other nearby homes that are not on the same side of the street;

[20] We provide additional details on the algorithm that identifies nearest neighbors in Online Appendix B.

in this case, the across the street houses numbered 5, 7, and 10 as well as homes 6, 8, and 9 which are located behind the home of interest.

4.2 Estimation Sample

Our primary estimation sample is built by imposing a set of restrictions on the transactions matched with the HMDA data for virtually all MSAs in the country. All current residents (and new neighbors) must appear in the the matched data. We impose the following restrictions and refinements on the included homes and neighborhoods to arrive at our estimation sample. Specifically, we drop all homes initially built before 1900 or with building square feet over 6,000 to avoid comparisons between homes vastly different in size or age. For similar reasons, we also drop households with incomes under $10,000 or greater than $2,000,000. To ensure that the experiments studied are likely to satisfy the identifying assumptions, we further require that: (i) the census block be made up entirely of single family residences, duplexes, triplexes, or quadplexes, (ii) the interquartile range of age of the block's homes be less than 30 years, (iii) the number of housing units on the block be between 20 and 500 as per the 2010 census, and (iv) the population density be between 500 and 10,000 persons per square mile as per the 2010 census.

Next, we restrict our analysis to Black and White households. This choice is driven by several considerations. First, there is a longstanding and well-studied historical aversion and hostility to Black entry in predominantly White neighborhoods. In addition, this focus avoids some of the difficulties that arise in classifying Hispanic homeowners racially, as many are phenotypically similar to, and may be perceived as, White homeowners, especially in wealthier neighborhoods. As such, to mitigate potential confounding in interpretation, we focus on block groups in which the Hispanic population share is less than 10 percent as per the 2010 census. We also drop Asian households from the analysis, due primarily to the relatively small number of experiments that would be present in the sample. To focus on incumbents who are plausibly at risk of leaving, we limit the sample to current residents who have owned their residence for at least four quarters and that their next transaction, if they have one, not be a foreclosure.[21] Finally, because we are interested in studying salient local social interactions among Blacks and Whites in the context of neighborhood change, we restrict our experiments to those in which the new next-door neighbor arrives in a home that is within 25 meters of the current resident.[22] Together, these requirements allow us to appropriately

[21] Their first quarter at the residence is either the quarter they purchase the home or the quarter that CoreLogic data first covers their county.

[22] We present evidence on the reasonableness of this assumption below, finding evidence that there is virtually no evi-

classify nearest neighbors and to focus on more homogeneous blocks. We further explore potential heterogeneity driven by these choices in Table C6.

Figure 2 presents a national map showing the counties where a valid experiment occurs. Darker shade indicates that the incidence of experiments, relative to the population, is higher. Experiments are more common in the eastern half of the US, primarily in the Northeast, Midwest, and Sunbelt regions.

[FIGURE 2 HERE]

The high-frequency nature of these data across both space and time allows us to characterize local neighborhood dynamics at the level of a residential block and year-quarter. Our principal outcome measures whether an existing homeowner moves within 2 years of receiving a new different-race neighbor nearby on their block. We assign treatment and control status on the basis of race and the timing of when the home transacts. A household is classified as being "treated" if it receives a new next-door neighbor of a different race in a particular year-quarter and does not move in the same quarter that the new neighbor arrived or the subsequent quarter. Importantly, in the cases where households do leave in the immediate quarter subsequent to the arrival of a new neighbor of a different race, they are not classified as a control in the year-quarter of the arrival either.[23] Corresponding "control" households for a treated household are those of the same race on the same block. We further delineate control households by the distance to the new neighbor, focusing most of our analysis on a control group of incumbent homeowners located two to three houses away from the new different-race neighbor on the same side of the street on the same block.

For a visual representation of where our experiments of interest are located within a typical metropolitan county, the second panel in Figure 2 presents a block-level map of Charlotte and the surrounding areas of Mecklenburg County. The dark gray areas map the census blocks where at least one experiment occurs over the sample observation period. The spatial distribution of these locations highlights the fact that although these experiments take place all over the county, the majority are in the periphery of the county. This is consistent with recent descriptive research that demonstrates a substantial increase in Black suburbanization in many metropolitan areas (Bartik and Mast, 2021).

dence that incumbent residents have a response beyond 25 meters.

[23] This restriction on the timing of moves for treatment assignment helps to ensure we do not capture moves that happen to be coincidental with the arrival of new different-race neighbors, as home transactions are often formally recorded several weeks/months after a deal is reached.

4.3 Summary Statistics: Where Different-Race Neighbor Experiments Occur

Table 1 presents summary statistics, in separate panels for Black and White homeowners in the sample. The first column in each panel presents statistics for the full set of CoreLogic-HMDA matched homeowners that meet the sample restrictions as defined in Section 4.2. An observation is available for any year-quarter in which a homeowner is observed and, thus, at risk of receiving a new next-door neighbor. The second column shows analogous statistics for those homeowner-by-quarter observations where the homeowner is treated by the receipt of a new next-door neighbor of a different race. As we discuss in detail below, a comparison of these columns highlights the kinds of homes, homeowners, blocks, and neighborhoods where experiments are more likely to occur within the study area.

[TABLE 1 HERE]

Comparing the first two columns of Table 1 highlights the differences between Black homeowners that are treated relative to the full set of Black homeowners in the sample. A first item to note is how rare it is for an existing Black homeowner to receive a new White next-door neighbor in our sample. Out of approximately 18.5 million quarters in which an existing Black homeowner is at risk of receiving a new next-door neighbor, there are just under 72,000 treated Black household-quarters. The first row of the table highlights a second interesting statistic: treated Black households are much more likely to move over the next two years than homeowners in the full Black sample, 6.28 vs. 4.56 percent. In addition to any direct response to the receipt of a new White neighbor, this raw difference could, of course, also be due to selection, changing neighborhood amenities, and many other confounding factors. Dealing with these potential confounders is exactly what motivates our nearest neighbor research design.

Comparing the first two columns for the remaining rows of Table 1 reveals that treated Black households tend to live in larger and more recently built homes, with an average year built of 1986 and square footage of nearly 2,000 square feet, compared to 1976 and 1,766 square feet for Black-owned homes as a whole. Interestingly, treated Black households also tend to have higher household incomes, $77,000 vs. $71,000, than the full sample of Black homeowners. The neighborhoods where treated Black homeowners reside also differ from the sample as a whole. Reflecting increased Black suburbanization in recent decades and greater racial integration with White households in the suburbs, the neighborhoods where treated Black households reside tend to be much less dense and have lower Black and higher White population shares than those for Black homeowners as a whole. In

addition, treated Black households tend to live in neighborhoods with higher median household incomes, 75,000 dollars, as compared to the overall average of about 65,500 dollars.

The contrast between treated White households and the overall sample of White homeowners also exhibits several striking differences. Out of the roughly 134 million observations in the data, only 104,402 White household-quarters are treated, an arrival rate for new different-race next door neighbors that is even lower than for Black homeowners. Like treated Black households, treated White households are more likely to move, live in newer houses, and on blocks that are less dense compared to the overall White homeowner population. In contrast to their Black counterparts, however, treated White households have lower incomes, 80,375 versus 89,281 dollars, and tend to live in neighborhoods that have lower average incomes compared to the full sample of White homeowners.

Perhaps the most noteworthy feature of Table 1 is that *treated* Black and White homeowners are quite similar to one another along almost every dimension: household income, the age and size of their homes, neighborhood racial composition and income, and population density. This suggests that the places where experiments occur for Black and White homeowners are quite similar and that these happen most often in middle-income, newer, suburban-density neighborhoods with a meaningful amount of racial integration.

4.4 Balance Tests

In reporting the results for our main analysis below, we focus on the main parameter of interest $\pi_1 - \pi_2$ from equation 6. This parameter captures the incremental difference in move propensities for treated households located immediately next door to a new different-race neighbor relative to control households of the same race located two to three houses away, on the same side of the street on the same block in the same quarter. Before showing the main results, we first formally test "balance" between the treatment and control groups associated with this fine geographic contrast for observable characteristics, reporting these results in Table 2. These specifications use an estimating equation that is identical to one used in our main analysis but replace the left hand side variable with a housing or household attribute:

$$X_{i,j,t} = \phi_1 I(r', d = 1)_{ijt} + \phi_2 I(r', d = 2/3)_{ijt} + \phi_3 I(r', d = 4+)_{ijt} + \lambda Z_{i,j,t} + \omega_{j,t} + \nu_{i,j,t} \quad (7)$$

where X is an observable attribute. The first eight columns of Table 2 present the results of separate regressions for square footage, year built, household income, presence of a mortgage co-applicant,

the length of time (in quarters) we have observed the homeowner in the home, the age (in quarters) of their mortgage, the purpose of the mortgage, and whether the loan is a conventional mortgage.

The final column reports results for an index of these eight housing and household variables. The weights for this index are the estimated coefficients of a regression model that predicts exit over the next two years as a function of these observable household and housing variables. This index not only collapses the information in the observable variables into a single variable, but the resulting magnitude of the coefficient shown in column (9) can be interpreted as the incremental likelihood that treated households would move in the next two years compared to control households two to three doors away, as predicted by the full set of observable characteristics.

[TABLE 2 HERE]

Turning first to Panel A, which reports results for Black incumbent homeowners, we find negligible differences between the treatment and control households for all eight variables. In fact, the coefficients amount to precise null estimates in all cases. The estimate reported in column (1), for example, implies a difference of about two square feet in homes that receive a new different-race neighbor immediately next door versus two to three doors away. Corresponding differences in year built and income amount to only 10 *days* and 16 dollars. All of the implied t-stats for these tests are small, implying that, in addition to little economic significance, the estimates are also statistically indistinguishable from zero. The final column of Panel A reports the differences between treatment and control for the index of observable attributes. The estimate of 0.001 means that treated households are 0.001 percentage points more likely to move over the next two years than the control households, as predicted by these observable characteristics. This number is less than 1/4000th the mean, implying no meaningful difference in expected exit behavior of treated and control households on the basis of observed attributes.

Panel B presents analogous results for White current residents. The estimates again imply negligible differences in covariates for treated and control White incumbents. We find a difference of about one square foot in the size of their homes, 7 days in the age of the home, and roughly 170 dollars in income, on average, across the treatment and control groups. Examining the joint predictive power of these attributes using the index, we again find tiny differences between treatment and control groups: 0.02 percentage points on a mean of 10.90 percentage points. Although the difference is statistically significant, the predicted power of observable characteristics is again on the order of 1/500th the mean exit rate for White homeowners.

Taken as a whole, the estimates reported in Table 2 strongly suggest that the arrival of a new different-race neighbor one versus two to three doors down on the same side of the street on the same block is as good as randomly assigned. For both Black and White current residents, observable differences in predicted move rates based on income, home size and age, and mortgage information between treatment and control households represent a tiny fraction of the mean.

5 Main Results

Table 3 reports the estimated effect of receiving a new different-race next-door neighbor on the decision to move within two years. The sample includes only treatment and control homeowners on the same residential block. The table reports our main parameter of interest, $\pi_1 - \pi_2$ from equation 6, which again can be interpreted as the incremental move response of incumbent homeowners who receive a new different-race neighbor immediately next door relative to same race homeowners two or three doors away, on the same side of the street on the same block within the same quarter. Each specification includes block-by-quarter fixed effects. Column (1) includes no additional control variables. Column (2) adds property attributes including controls for square footage and year built. Column (3) adds available household attributes including indicators for income bins and presence of a co-applicant. Column (4) adds mortgage attributes, including type of loan, age, and purpose.

[TABLE 3 HERE]

Looking first at the results for Black current residents in Panel A, column (1) presents the estimate from a baseline model excluding house and homeowner controls. The estimate indicates a positive and statistically significant increase in move propensity of 0.377. Adding controls for building, resident, and loan attributes as we move across columns (2) through (4) has little impact on the estimates. In the specification that includes the full set of controls presented in column (4), the coefficient is only slightly lower at 0.375. Notice that the difference between the estimate reported in column (4) and the one shown in column (1) is equal to the estimate for the predicted index in Table 2. As mentioned above, this difference is less than $1/4000th$ of the mean move propensity or less than $1/100th$ of the main effect size shown in Table 3.

Panel B presents estimates from analogous models for White current residents. Column (1) again presents the estimate from our baseline model with no additional controls, yielding a statistically significant increase in the relative move propensity of 0.313. Similar to the results for Black current

residents, including additional controls in columns (2), (3), and (4) has virtually no effect on the estimates. In column (4), conditional on the full set of controls, the estimated coefficient is only marginally smaller than the model with no controls at 0.292.

Overall, the incremental responses of both Black and White homeowners to receiving a new different-race neighbor immediately next door are similar in magnitude. Both are also statistically significant at conventional levels. Another way to view the size of these coefficients is to compare them to the mean of the dependent variable, reported in a lower row of each panel. The average move rates over the next two years are 5.87 percent and 10.90 percent for Black and White households, respectively, which means these estimates translate into roughly 6 and 3 percent increases in exit rates relative to households two or three doors away.

The robust responses estimated here are consistent with a model in which households have preferences directly over the race of their neighbors. Such preferences might be driven by a desire to live near neighbors of the same race and/or to explicitly avoid neighbors of another race.[24] Notice also that a desire to move away from new different-race neighbors could result from an incumbent household's own racism or as a reaction to the racism of the newly arriving neighbor. An especially racist homeowner, for example, might flee their existing neighborhood quickly upon receiving a new different-race neighbor and create a hostile relationship upon their arrival as a new neighbor. In this way, the symmetry of the responses of Black and White incumbent households that we estimate here may, in part, reflect the fact that the basis for the response *in both cases* is a new neighbor cross-race relationship subject to the potential racism of one party.

The main advantage of the nearest neighbor design applied at the fine geographic scale we use here is that it credibly attributes this incremental component of neighborhood choice directly to household preferences for the identities of their neighbors. As discussed more fully in Section 3, a limitation of our approach, however, is that these estimated effects capture only the incremental response of homeowners living immediately next door to the newly-arrived neighbor over and above the response of those just a two to three doors further away. As a result, these estimates naturally serve as lower bounds on the full direct response to the receipt of a new different-race neighbor, to the extent that the receipt of a new different-race neighbor may also directly enter the preferences of households two to three doors away.

The standard errors reported in Table 3 and throughout the analyses presented in the paper are

[24] These are isomorphic in observational data. As such, the results are consistent with recent work suggesting homophily as one of several key drivers of sorting responses (Aliprantis et al., 2022; Bruch and Mare, 2006).

adjusted for clustering at the tract-year level. Our choice of this level of clustering is motivated by the fact that a given block-by-quarter is included in our sample analysis only if it has moves of the appropriate type, and our identifying assumption is that the location of the move is then as good as random. This corresponds to a version of clustered sampling, plus random assignment within cluster, which is one motivation for clustered standard errors. This motivation suggests that clustering at the tract level may be a bit conservative. Appendix Table C2 reports results for a wide variety of alternative choices for the level clustering. The estimated standard errors indeed imply that our approach tends to be conservative. The statistical significance of the main results at conventional levels are robust to this choice.[25]

5.1 Alternative Measurement of Nearest Neighbors

To examine the robustness of our main results to alternative ways of constructing fine geographic contrasts among close neighbors, we expand the set of comparisons beyond the immediate block by studying move responses to new different-race neighbors within the 40 nearest parcels based on distance. In contrast to our baseline research design, this set of homes can include houses on surrounding census blocks (but always in the same census block group), such as those across the street. We focus particularly on moves as a response to receiving a new different-race neighbor among the two nearest parcels in comparison to responses to a new-different race neighbor within the 3, 4, or 5 closest parcels. All models include block group-by-quarter fixed effects.[26]

The first and third columns of Table 4 reports balance results for the index of all eight observable attributes for Black and White current residents. For Black incumbent homeowners, this coefficient is -0.0002, again suggesting no systematic selection on the basis of observables in this alternate design. For White current residents, the index is slightly larger and statistically significant, but at

[25] To study the inference problem more directly, we implemented a placebo test that randomly re-assigns treatment status within a block-by-quarter. Specifically, for each "experiment", we randomly re-assign the treatment dummies (d=1, d=2/3, d=4+) that occurred in the data to the incumbent households observed on that block and included in our original analysis. As a result, the new placebo sample has exactly the same number of (i) observations, (ii) each treatment assignment type, and (iii) fixed effect cells as our main sample. We repeated this exercise 1,000 times and report the empirical distribution of the estimates in Appendix Figure C3. The mean and standard deviations for these estimates are 0.011 and 0.133 for Black incumbents and 0.001 and 0.126 for white incumbents, respectively. That the means are near zero provides a clean placebo test for our main estimates. That the estimated standard deviations are almost identical to the estimated standard errors for our main specifications provides additional assurance that the clustering of the standard errors is working properly. We thank an anonymous referee for suggesting this test.

[26] Balance tests for this design, analogous to those presented in Table 2 above, are shown in Appendix Table C3. We again find small differences between the two groups – e.g., for Black incumbent homeowners, we find a difference of less than 1 square foot in the size of the homes, no difference in year built, no difference in the likelihood of having a co-applicant, and less than 200 dollars in annual income. For White incumbent homeowners, we again find generally small differences for these observables: less than 5 square feet, about 0.1 of a year, virtually no difference in share of co-applicants on mortgage, and 90 dollars of income.

0.023 remains small in economic terms. Overall, although the balance tests were slightly better for our preferred research design, we interpret these balance tests as supportive of the validity of this alternative research design.

[TABLE 4 HERE]

In the second and fourth column of Table 4, we present the results for this alternative research design for Black and White current residents, respectively. In each case, the estimates suggest the presence of a highly localized differential move response. The estimated coefficients for receiving a different race in the two closest parcels are positive and statistically significant for Black and White current residents, with similar magnitudes of 0.274 and 0.366 percentage points, respectively. In general, the main findings are robust to the exact set of neighbors used as the closest controls.

5.2 Additional Robustness Checks

In this subsection, we provide a number of additional checks designed to study the robustness of our results to alternative specifications and sample restrictions. First, as mentioned above, Appendix Table C5 reports results for two additional specifications that reduce the number of treatments to two, so that the specification fits within the scope of Angrist (1998)'s result. In the first, we drop the treatment 4+ doors away, leaving just our main treatment and control groups. In the second, we combine 2-3 doors away and 4+ doors away on the same block into a single control group, 2+ doors away. In both cases, the relevant point estimates are a bit larger for both Black and White incumbents and not statistically significantly different in these alternative specifications compared to our baseline specification.

Appendix Table C6 reports results analogous to the final column of Table 3 for a number of alternative sample restrictions. Column (1) of this table reproduces the estimates shown in column (4) of Table 3. Columns (2), (3), (4), (5), and (6) each relax one of the sample restrictions we apply when making our estimation sample and Column (7) adds a sample restriction. Specifically, Column (2) includes homes built before 1900 and with building square feet over 6,000. Column (3) includes blocks with a housing unit count under 20 and population density under 500 as per the 2010 census. Column (4) includes block groups with Hispanic share over 10%. Column (5) includes blocks where the interquartile range of age homes on the block is greater than 30 years. Column (6) includes all current residents, regardless of tenure. Finally, column (7) restricts our main sample to just those counties in metropolitan statistical areas with at least 200,000 residents and in which the Black

share is at least as large as the Black share in the US population, 12 percent, during the study period. The effect of receiving a new next-door neighbor of a different race is remarkably stable across all specifications and is in all cases both statistically and economically significant.

6 Testing Identifying Assumptions and Alternative Explanations

In this section, we provide results from additional empirical analyses designed to test features of the identifying assumptions described in Section 3 and to examine alternative explanations for the main results presented in Table 3 above.

The estimates reported for the balance tests in Table 2 provide evidence in support of Assumption 1 – i.e., quasi-random arrival location within the block. In particular, the negligible effects reported there imply the proximity to a new different-race neighbor at the very fine geographic scale that we study (next door vs. two to three houses away) appears to be as good as randomly assigned, to the extent this can be tested on key observable attributes such as income, home size, and age.

Assumption 3 requires that proximity to the newly-arrived neighbor has no impact on future amenities or house prices within the block. While we do not observe neighborhood amenities at this fine of a geographic scale, we can observe the price of any home that sells during our sample period. Examining future sales prices allows us to (i) directly assess whether the receipt of a new different-race neighbor immediately next door versus two or three doors further away has any direct effect on housing prices and (ii) to test for the presence of any highly localized amenities that would be significant enough to affect housing prices.

Table 5 reports results for a series of specifications that compare house prices for homes that sell in the same quarter. Comparisons are restricted to the same block group and quarter in columns (1) and (3), and, more finely, the same block and quarter in columns (2) and (4). The estimated parameter returns the incremental effect on housing prices for houses that recently received a new different-race neighbor immediately next door compared to others two to three doors away from a new different-race neighbor in the same geographic area and quarter. The sample for each column includes all treated and control observations two to three doors away for which a home sells following the receipt of a new different-race neighbor plus all other homes that sold within the same geographic area and quarter. All columns include all eight controls as before, Z, and the estimating equation can be written:

$$lnp_{i,j,t} = \psi_1 I(r', d = 1)_{ijt} + \psi_2 I(r', d = 2/3)_{ijt} + \psi_3 I(r', d = 4+)_{ijt} + \lambda Z_{i,j,t} + \omega_{j,t} + \nu_{i,j,t} \quad (8)$$

where our parameter of interest is $\psi_1 - \psi_2$ and j indicates the level of geography for the given specifications: block group in columns (1) and (3), and block in (2) and (4).

[TABLE 5 HERE]

We turn first to the results for Black current residents shown in column (1), which conditions on group-by-quarter fixed effects. The difference between treated and control homes is 0.02. Restricting geography further to the block in column (2) moves the point estimate closer to zero, 0.008. This implies that the receipt of a new White neighbor next door one versus two to three doors away has a statistically insignificant 0.8 percent impact on subsequent sales price. For White households, the block-by-quarter fixed effects specification shown in column (4) has a point estimate of -0.007 and a standard error of 0.006, again implying a statistically null effect on future resale prices. Overall, there is no evidence of any systematic relationship between the receipt of a new different-race neighbor immediately next door versus two to three doors away on future housing prices.

Taken together, the results presented in Table 5 suggest that sales prices are not affected much by the receipt of a new different-race neighbor immediately next door versus two to three doors away within the same neighborhood. These results help diminish concerns about (i) highly localized variation in current or future amenities that might be correlated with the receipt of a new different-race next door neighbor and (ii) that the presence of the new neighbor might affect future sales prices at this fine geographic scale.

6.1 An Alternative Empirical Approach

In Section 1, we discussed two additional potential concerns that naturally arise in interpreting the effects from our main analysis as a response to the race of the newly arriving neighbor. First, households might respond more generally to the disruption resulting from the receipt of a new neighbor, regardless of the new neighbor's race. And, second, households might respond to other attributes of their neighbors, such as income, which are highly correlated with race at the population level.

To address these concerns directly, we consider an alternative empirical approach that complements our main analysis. In this case, we compare the responses of households who receive new different-race neighbors to those who receive new same-race neighbors, immediately next door versus two to three doors away, within the same Census block group in the same year-quarter. To do so,

we first restrict the sample to those homeowners who received new different- or same-race neighbors within three doors on the same block on the same side of the street at some time during the sample. We then include block group-by-quarter fixed effects in the analysis to ensure that all comparisons of move propensities are made among households who received either different- or same-race neighbors within three houses in the same block group and quarter. Our baseline specification for this research design can be written:

$$M_{i,j,t} = \alpha_1 I(r',1)_{ijt} + \alpha_2 I(r',2/3)_{ijt} + \alpha_3 I(r,1)_{ijt} + \lambda z_{i,j,t} + \theta_{j,t} + \nu_{i,j,t} \tag{9}$$

where $\theta_{j,t}$ are block group-by-quarter fixed effects, $I(r',1)_{ijt}$ and $I(r',2/3)_{ijt}$ are indicators for whether a new different-race r' neighbor moves in next door and two to three doors away, respectively, and $I(r,1)_{ijt}$ is an indicator for whether a new same-race r neighbor moves in next door. All three measures apply to houses on the same side of the same street and the omitted category includes homeowners who receive a new same-race neighbor two to three doors away.

The parameters reported in Table 6 are $\alpha_1 - \alpha_2$, which measures the incremental effect of receiving a new different-race neighbor immediately next door versus two to three houses away, and α_3, which measures the incremental effect of receiving a new same-race neighbor immediately next door versus two to three houses away. All specifications include the full set of property and homeowner control variables used in our main analysis.

6.1.1 Different- Vs. Same-Race Neighbors

Column (1) in each panel of Table 6 highlights the difference in responses to different- versus same-race next door neighbors. Turning first to the results for incumbent Black homeowners shown in Panel A, the estimate of the impact of receiving a new White neighbor immediately next door are positive, similar in magnitude to those reported in our main analysis (0.514 here versus 0.375 in Table 3), and statistically significant at conventional levels. In contrast, the estimates for Black current residents receiving new Black neighbors next door versus two or three doors away, while positive, are less than half of magnitude than the estimate for receiving a new White neighbor and statistically insignificant. .

[TABLE 6 HERE]

The results for White incumbent homeowners paint a similar picture. Looking at the estimates reported in column (1) of Panel B, the response of White current residents to the receipt of a new

Black neighbor immediately next door versus two to three doors away is 0.434, which is again slightly larger than that reported in our baseline specification. Like for Black homeowners, the estimated response of White homeowners to the receipt of a new same-race neighbor next door versus two to three doors away is less than half the magnitude. In this case, the estimate is 0.200. Thus, for both Black and White residents the response to receiving a new different-race neighbor appears to be substantially greater than any general response to the receipt of a next-door neighbor.

6.1.2 Adding Controls for Income

A key advantage of this alternative research design, compared to our main approach, is that it permits the simultaneous consideration of both the race and income of the new neighbor. This is attractive for several reasons. First, it allows us to explore whether the estimated responses to the receipt of a new different-race neighbor might not be a response to the neighbor's race, per sé, but instead reflect a response to other attributes that are correlated with race in the population, of which income serves as a prime example. Second, including income as a control in the analysis provides a complementary estimate of how homeowners respond directly to the income of their neighbors, which is of independent interest in the residential sorting literature. And, finally, we can broaden the analysis to consider various measures of how a new neighbor's income compares to either the neighborhood or the incumbent homeowner to gain a richer understanding of what might be motivating incumbent homeowners' reactions.

Returning to Table 6, columns (2) through (4) report estimates of specifications that are analogous to the specification reported in column (1) but also include control for the incomes of new neighbors in several ways. The specification reported in column (2) includes five bins for income of the new neighbors. The estimates for the direct role of income are shown in Appendix Table C7. These point estimates imply that both Black and White incumbent homeowners have a strong inverse response to the income of their new neighbors. Compared to the receipt of a new neighbor with less than $\$50k$ in income, for example, Black homeowners have 0.570 and 0.252 percentage point decreased rates of moving over the next two years in response to a new neighbor with $\$100k - \$150k$ and over $\$150k$ in income, respectively. The comparable numbers are similar for White incumbent homeowners at a 0.384 and 0.469 percentage point decreased rate of moving. In this way, both Black and White homeowners systematically prefer to have close neighbors with higher incomes, and this relationship appears to be roughly monotonic in income.

Remarkably, given the substantial preferences estimated for the new neighbor's income, the in-

clusion of these income controls has almost no impact on the estimated responses to the race of new neighbors. That is, there is no change in the estimated coefficients related to race in moving from column (1) to column (2). At first glance, this result might seem surprising given the correlation between race and income in the general population. But it is actually just another implication of the very fine geographic scale of the comparisons made in applying our nearest neighbor research design. In particular, that controlling for neighbor income has no effect on the race coefficients (despite having a strong direct effect) implies that there is essentially no correlation between income and proximity of new neighbors at this geographic scale – i.e., a new different-race neighbor's income is not correlated with whether they arrive one versus two to three doors away.

In this way, the estimates reported in column (2) confirm that the estimated responses to the neighbor's race in our analysis is not an income effect. The orthogonality of an important attribute like income to race at these fine geographic scales is suggestive that other household attributes are also likely to have little correlation with race within the context of our preferred research design. As such, this finding lends greater credence to the interpretation of our main results as a response to race itself, rather than potential correlates of race.

Instead of direct controls for the income of the new neighbor, the remaining columns of Table 6 include different income measures. The specification shown in column (3), for example, includes bins for the difference between the new neighbor's income and the block group median household income. Column (4) instead includes bins capturing the difference between the current resident and new neighbor's income. As in Column (2), the inclusion of these alternative measures again has no impact on the estimated coefficients related to race.[27]

7 Heterogeneity in the Racial Response

We conclude the presentation of our main results by examining heterogeneity in the estimated response to a new neighbor's race. Rather than just an interesting elaboration, exploring heterogeneity in these responses has important real-world consequences in that it can help distinguish contexts where race-based preferences remain strong drivers of neighborhood stratification from those where such preferences likely have a smaller role. The latter may help distinguish places in American society that have become more amenable to stable racial integration in recent decades.

[27] The expanded results presented in Appendix Table C7 also point to some interesting heterogeneity, suggesting that high-income incumbent residents (both Black and White) may also have the strongest reactions to the income of their new neighbors.

7.1 Heterogeneity in Distance to Neighbors

Table 7 reports a set of results that examine heterogeneity in the racial response as a function of distance between neighboring houses. Intuitively, when this distance is small, neighbors are forced to live closer to one another and may interact more often or more intensely. With greater distance between homes in suburban and exurban settings, neighbors may be able to come and go without interacting much at all.

[TABLE 7 HERE]

Aspects of this intuition are present in Table 7 where we report results for two specifications that allow the distance between the homes of the current resident and the new different race neighbor to vary more continuously. The first and third columns augment our main specification with an interaction of our treatment indicator with distance measured in decameters. The sample is restricted to experiments that occur for homes that are within 25 meters of one another, as in our main specification. The second and fourth columns expand the sample to include experiments for homes up to 100 meters of one another, reporting estimates from an alternative model that includes a set of indicators for the the distance of the new arrival: between 0 and 12.5 meters, 12.5 - 25 meters, 25 - 100 meters.

Overall, the results support the notion that localized social interactions among nearest neighbors are mediated by distance. The point estimates indicate that both Black and White current residents respond especially strongly to new different-race neighbors whose homes are close and this effect seems to decay monotonically with distance from the home of the new arrival. As seen in columns (2) and (4), beyond 25 meters the move response is essentially negligible. Interestingly, for White incumbent homeowners the racial response is especially strong when neighboring homes are very close together and declines sharply with distance: the point estimate falls from 1.248 between 0 and 12.5 meters to 0.331 between 12.5 and 25 meters, and to roughly 0 thereafter.

7.2 Regional Heterogeneity in Responses to Race

Given the distinct history surrounding cross-racial interaction across the country, we next explore regional heterogeneity in estimated racial responses. To that end, Table 8 reports results from our preferred specification where the sample split by region focuses specifically on a comparison of Southern versus Northern states.[28] The number of observations shown in each column of the table make

[28] We follow the region definitions from the US Census characterizing the Southern states as those in the South region and the Northern states as those in the Northeast and Midwest.

clear that the majority of our sample for both Black and White homeowners comes from the Southern states.

[TABLE 8 HERE]

For Black homeowners in both regions, the point estimates suggest a similarly-sized racial response to the main results presented in Table 3, although the estimated response in the Northern region is not statistically significant. In contrast, White homeowners in the North have significantly stronger responses to receiving new different-race neighbors than their White counterparts in the South as well as Black homeowners in either region. An especially strong racial response by White homeowners in Northern cities is in line with the history of White flight in response to the in-migration of Black residents during the Great Migration (Boustan (2010), Boustan (2017), Derenoncourt (2022)) and the accompanying legal and extra-legal efforts to confine Black households to highly segregated neighborhoods in these cities. That White racial responses remain stronger in these Northern cities into the 21st Century is also consistent with the observation that residential racial segregation remains significantly higher in these cities than elsewhere in the country.

7.3 Income Heterogeneity

We close this section by examining heterogeneity in responses to new different-race neighbors along three dimensions of income: income of the (i) neighborhood, (ii) incumbent resident, and (iii) new neighbor. The columns of Table 9 report the results for expanded specifications of our main analysis that interact the treatment indicator with an indicator for whether the corresponding income measure is above or below the median..[29] The first three columns report results for incumbent Black homeowners who receive a White neighbor next door versus two to three houses away, while the final three columns show the responses of incumbent White households.

[TABLE 9 HERE]

Overall, while a bit noisy in places, the broad pattern of income heterogeneity is quite similar for Black and White incumbent homeowners. For both Black and White households, there does not seem to be any systematic heterogeneity on the basis of neighborhood income, columns (1) and (4). In contrast, the relative incomes of both incumbent and new residents appear to make more of a

[29]The third column includes slightly fewer observations because we restrict the sample to block-quarters that receive exactly one new different-race neighbor so that income can be cleanly categorized.

difference. The results shown in column (2) imply that relatively high-income Black incumbent residents have stronger move responses, with an estimated move response of 0.558 for those with above median income for the neighborhood versus 0.194 for those below median income. This difference is statistically significant at conventional levels. While somewhat less precise, the same pattern holds for White incumbents – see column (5). The point estimates in Columns (3) and (6) suggest that incumbent residents of both races are more likely to move in response to a new different-race neighbor with relatively low incomes. For both Black and White incumbents, although the response to new neighbors with incomes below the median is statistically significant for both Black and White incumbents, the difference in the responses to higher versus lower income is not significant.

8 Conclusion

Racial stratification remains a defining feature of every major city in the United States. Even as individual neighborhoods evolve, the overall segregated structure of American cities has been continuously renewed and reinforced by the ways that households create neighborhood change by both pushing into new neighborhoods or by moving away. Social scientists have documented the kinds of dynamic patterns – neighborhood tipping, white flight, gentrification, and aversive sorting – that help maintain and reinforce racial segregation (Boustan, 2010; Caetano and Maheshri, 2017; Card et al., 2008; Casey, 2020; Gould Ellen, 2000; Guerrieri et al., 2013; Krysan and Crowder, 2017). But a long-standing question in this literature has remained open for decades: to what extent are households responding to the identities of their new neighbors versus the kinds of endogenous (current and future) changes to the neighborhood that accompany them? In the days of blockbusting, for example, realtors would stoke white fears about what Black in-migration would mean for their future home values (Boustan, 2017). And it is often the rapid changes in local businesses (e.g., a new Starbucks or Whole Foods), reacting to increased local (high-income) demand, that serve as the most obvious markers of modern gentrifying neighborhoods (Couture and Handbury, 2020; Glaeser et al., 2018).

The main contribution of this paper is the development and application of a new nearest neighbor research design that seeks to separately identify a component of neighborhood racial change attributable directly to neighbors' identities rather than any associated neighborhood amenity changes, current or future. Our approach bases these estimates on an empirical contrast between the out-migration decisions of two single-family homeowners of the same race on the same residential block – one immediately nearby, one a bit further away – in reaction to receiving a new neighbor of a dif-

ferent race. The core identifying assumption is that where the new neighbor arrives on the block is as good as randomly assigned and has the same implications for current and future neighborhood changes, including for house prices.

Our findings indicate strong, statistically significant responses of both Black and White homeowners to receiving a new neighbor of a different race.[30] For both Black and White homeowners, the estimates are driven by experiments in the data that are especially likely to occur in more newly built, middle-income, suburban-density neighborhoods in Southern states. We demonstrate robustness of these findings across a number of dimensions. The magnitude and qualitative implications of these results remain when we consider different definitions of nearest neighbors, use alternative research designs, and control directly for the income of new different-race neighbors.

These findings have implications for the potential sustainability of racially integrated neighborhoods over long periods of time. If neighborhood racial change were fully attributable to broader changes in neighborhood amenities and house prices, policy responses, such as anti-discrimination enforcement, affordable housing initiatives, and efforts to maintain access to existing public and private goods and services, might be enough to effectively foster and maintain racially integrated neighborhoods. If, as the evidence indicates, direct responses to neighbors' identities have a significant role in the decision to move away then, if integrated neighborhoods are to be sustained, it suggests that many of the routine, daily social interactions among neighbors require the attention of activists and policy makers. Given the responses estimated in this paper, public policies that aim to foster positive social interactions, especially among new neighbors, are likely critical to maintaining racially integrated neighborhoods in many settings.

Finally, the estimated racial responses exhibit heterogeneity by density and region that has implications for contexts in which racial preferences continue to harden historical segregation patterns as well as places in modern American society with greater potential for racial integration. In particular, the estimated racial responses for White homeowners reveal stronger responses to new Black neighbors in the Northern states and in particularly dense areas. In contrast, White racial responses are much more muted in suburban-density spaces and in the Southern states. Taken as a whole, these results suggest that emergent neighborhoods with substantial racial integration in suburban regions, especially of Southern cities, are likely much more sustainable than in previous generations.

[30] The examination of neighborhood racial change in both directions distinguishes our paper from most of the literature, which, for historical reasons, has generally focused on White responses to Black in-migration. For more discussion in the context of tipping, see, e.g., Card et al. (2011).

References

ALIPRANTIS, D., CARROLL, D., AND YOUNG, E. 2022. What explains neighborhood sorting by income and race? *Federal Reserve Bank of Cleveland Working Paper* .

ALMAGRO, M., CHYN, E., AND STUART, B. 2021. The effects of urban renewal programs on gentrification and inequality: Evidence from chicago public housing demolitions. *U of Chicago Working Paper* .

ALMAGRO, M. AND DOMINGUEZ-IINO 2020. Location sorting and endogenous amenities: Evidence from amsterdam. *U of Chicago Working Paper* .

ANANAT, E. 2011. The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics* 3:34–66.

ANDREWS, R., CASEY, M., HARDY, B. L., AND LOGAN, T. D. 2017. Location matters: Historical racial segregation and intergenerational mobility. *Economics Letters* 158:67–72.

ANENBERG, E. AND KUNG, E. 2014. Estimates of the size and source of price declines due to nearby foreclosures. *American Economic Review* 104:2527–2551.

ANGRIST, J. D. 1998. Estimating the labor market impact of voluntary military service using social security data on military applicants. *Econometrica* 66:249 –288.

BARTIK, A. AND MAST, E. 2021. Black suburbanization and neighborhood inequality. *University of Illinois Working Paper* .

BAUM-SNOW, N. AND HARTLEY, D. 2020. Accounting for central neighborhood change. *Journal of Urban Economics* 117:103 – 228.

BAYER, P., CASEY, M., FERREIRA, F., AND MCMILLAN, R. 2017. Racial and ethnic price differentials in the housing market. *Journal of Urban Economics* 102:91–105.

BAYER, P., FERREIRA, F., AND MCMILLAN, R. 2007. A unified framework for measuring preferences for schools and neighborhoods. *Journal of political economy* 115:588–638.

BAYER, P., FERREIRA, F., AND ROSS, S. L. 2018. What drives racial and ethnic differences in high-cost mortgages? the role of high-risk lenders. *The Review of Financial Studies* 31:175–205.

BAYER, P., MANGUM, K., AND ROBERTS, J. 2021. Speculative fever: Investor contagion in the housing bubble. *American Economic Review* 111:609–651.

BAYER, P. AND MCMILLAN, R. 2005. Racial sorting and neighborhood quality. *National Bureau of Economic Research Working Paper 11813* .

BAYER, P., MCMILLAN, R., MURPHY, A., AND TIMMINS, C. 2016. A dynamic model of demand for houses and neighborhoods. *Econometrica* 84:893–942.

BAYER, P., ROSS, S., AND TOPA, G. 2008. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy* 116:1150–1196.

BILLINGS, S., CHYN, E., AND HAGGAG, K. 2022. The long-run effects of school racial diversity on political identity. *American Economic Review: Insights* 3:267–284.

BLAIR, P. 2017. Outside options (now) more important than race in explaining tipping points in US neighborhoods. *Harvard University Working Paper* .

BORUSYAK, K. AND HULL, P. 2023. Negative weights are no concern in design-based specifications. *AEA Papers and Proceedings* forthcoming.

BOUSTAN, L. P. 2010. Was postwar suburbanization "white flight"? evidence from the black migration. *The Quarterly Journal of Economics* 125:417–443.

BOUSTAN, L. P. 2017. Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets. National Bureau of Economic Research Publications.

BRUCH, E. E. AND MARE, R. D. 2006. Neighborhood choice and neighborhood change. *American Journal of Sociology* 112:667 – 709.

CAETANO, G. AND MAHESHRI, V. 2017. School segregation and the identification of tipping behavior. *Journal of Public Economics* 148:115–135.

CAETANO, G. AND MAHESHRI, V. 2021. A unified empirical framework to study segregation. *University of Georgia working paper* .

CARD, D., MAS, A., AND ROTHSTEIN, J. 2008. Tipping and the dynamics of segregation. *The Quarterly Journal of Economics* 123:177–218.

CARD, D., MAS, A., AND ROTHSTEIN, J. 2011. Are Mixed Neighborhoods Always Ustable. *In* H. Newburger, E. Birch, and S. Wachter (eds.), Neighborhood and Life Chances: How Place Matters in Modern America. Philadelphia: University of Pennsylvania Press.

CASEY, M. 2020. The Microdynamics of Neighborhood Transition. *U. of Illinois-Chicago Working Paper* .

CHARLES, C. Z. 2000. Neighborhood racial-composition preferences: Evidence from a multiethnic metropolis. *Social problems* 47:379–407.

CHETTY, R. AND HENDREN, N. 2018. The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics* 133:1107–1162.

CHETTY, R., HENDREN, N., JONES, M. R., AND PORTER, S. R. 2020. Race and economic opportunity in the united states: An intergenerational perspective. *The Quarterly Journal of Economics* 135:711–783.

CHETTY, R., HENDREN, N., AND KATZ, L. F. 2016. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review* 106:855–902.

CHETTY, R., JACKSON, M. O., KUCHLER, T., AND ET AL. 2022a. Social capital i: measurement and associations with economic mobility. *Nature* 608:108–121.

CHETTY, R., JACKSON, M. O., KUCHLER, T., AND ET AL. 2022b. Social capital ii: determinants of economic connectedness. *Nature* 608:122 – 134.

CHRISTENSEN, P. AND TIMMINS, C. 2021a. The damages and distortions from discrimination in the rental housing market. Technical report, National Bureau of Economic Research.

CHRISTENSEN, P. AND TIMMINS, C. 2021b. Racial discrimination and housing outcomes in the united states rental market. *NBER Working Paper 29516* .

CHRISTENSEN, P. AND TIMMINS, C. 2022. Sorting or steering: Experimental evidence on the economic effects of housing discrimination. *Journal of Political Economy* .

CHYN, E. 2018. Moved to opportunity: The long-run effect of public housing demolition on labor market outcomes of children. *American Economic Review* .

CHYN, E., HAGGAG, K., AND STUART, B. 2022. The effects of racial segregation on intergenerational mobility: Evidence from historical railroad placement. *Dartmouth University Working Paper* .

CHYN, E. AND KATZ, L. 2021. Neighborhoods matter: Assessing the evidence for place effects. *Journal of Economic Perspectives* 35:197–222.

COUTURE, V., GAUBERT, C., HANDBURY, J., AND HURST, E. 2019. Income growth and the distributional effects of urban spatial sorting. *NBER Working Paper* .

COUTURE, V. AND HANDBURY, J. 2020. Urban revival in america. *Journal of Urban Economics* 119.

COX, R., CUNNINGHAM, J., ORTEGA, A., AND WHALEY, K. 2022. Black lives: The high cost of segregation. *University of Southern California Working Paper* .

DAVIS, M., GREGORY, J., AND HARTLEY, D. 2021. The long-run effects of low-income housing on neighborhood composition. *Rutgers University working paper* .

DERENONCOURT, E. 2022. Can you move to opportunity? evidence from the great migration. *American Economic Review* pp. 369–408.

FARLEY, R. AND FREY, W. H. 1994. Changes in the segregation of whites from blacks during the 1980s: Small steps toward a more integrated society. *American sociological review* pp. 23–45.

FARLEY, R., SCHUMAN, H., BIANCHI, S., COLASANTO, D., AND HATCHETT, S. 1978. "chocolate city, vanilla suburbs:" will the trend toward racially separate communities continue? *Social Science Research* 7:319–344.

GALSTER, G. C. 1992. The case for racial integration. *In* G. C. Galster and E. W. Hill (eds.), The Metropolis in Black and White: Place, Power, and Polarization. Routledge.

GLAESER, E. L., KIM, H., AND LUCA, M. 2018. Measuring gentrification: Using yelp data to quantify neighborhood change. *NBER Working Paper 24952* .

GOULD ELLEN, I. 2000. Sharing America's Neighborhoods: The Prospects for Stable Racial Integration. Havard University Press.

GUERRIERI, V., HARTLEY, D., AND HURST, E. 2013. Endogenous gentrification and housing price dynamics. *Journal of Public Economics* 100:45–60.

HANSON, A. AND HAWLEY, Z. 2011. Do landlords discriminate in the rental housing market? evidence from an internet field experiment in us cities. *Journal of Urban Economics* 70:99–114.

HWANG, J. AND SAMPSON, R. J. 2014. Divergent pathways of gentrification: Racial inequality and the social order of renewal in chicago neighborhoods. *American Sociological Review* 79:726 – 751.

JOHNSON, R. 2011. Long-run impacts of school desegregation school quality on adult attainments. *NBER Working Paper 16664* .

JOHNSON, R. 2019. Children of the dream: Why school integration works. Basic Books.

KRYSAN, M. AND CROWDER, K. 2017. Cycle of Segregation: Social Processes and Residential Stratification. Russell Sage Foundation.

LEWIS, V. A., EMERSON, M. O., AND KLINEBERG, S. L. 2011. Who we'll live with: Neighborhood composition choices of whites, blacks and latinos. *Social Forces* 89:1385–1407.

LI, N. 2021. Housing market channels of segregation. *Consumer Financial Protection Bureau working paper* .

LOGAN, J. R., STULTS, B. J., AND FARLEY, R. 2004. Segregation of minorities in the metropolis: Two decades of change. *Demography* 41:1–22.

LOGAN, T. D. AND PARMAN, J. M. 2017. The national rise in residential segregation. *The Journal of Economic History* 77:127–170.

LUTZ, B. 2011. The end of court-ordered desegregation. *American Economic Journal: Economic Policy* 3:130–168.

MASSEY, D. AND DENTON, N. 1993. American Apartheid: Segregation and the Making of the Underclass. Harvard University Press.

MCCARTNEY, W. B., ORELLANA-LI, J., AND ZHANG, C. 2024. Political polarization affects households' financial decisions: Evidence from home sales. *Journal of Finance* 79:795–841.

MCCARTNEY, W. B. AND SHAH, A. M. 2022. Household mortgage refinancing decisions are neighbor influenced, especially along racial lines. *Journal of Urban Economics* 128:103409.

ONDRICH, J., ROSS, S., AND YINGER, J. 2003. Now you see it, now you don't: why do real estate agents withhold available houses from black customers? *Review of Economics and Statistics* 85:854–873.

PAGE, M. 1995. Racial and ethnic discrimination in urban housing markets: Evidence from a recent audit study. *Journal of Urban Economics* 38:183–206.

REBER, S. 2011. School desegregation and educational attainment for blacks. *Journal of Human Resources* 45:893–914.

SCHELLING, T. C. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1:143–186.

TUTTLE, C. 2019. The long-run economic effects of school desegregation. *University of Maryland Working Paper* .

YINGER, J. 1986. Measuring racial discrimination with fair housing audits: Caught in the act. *The American Economic Review* pp. 881—-893.

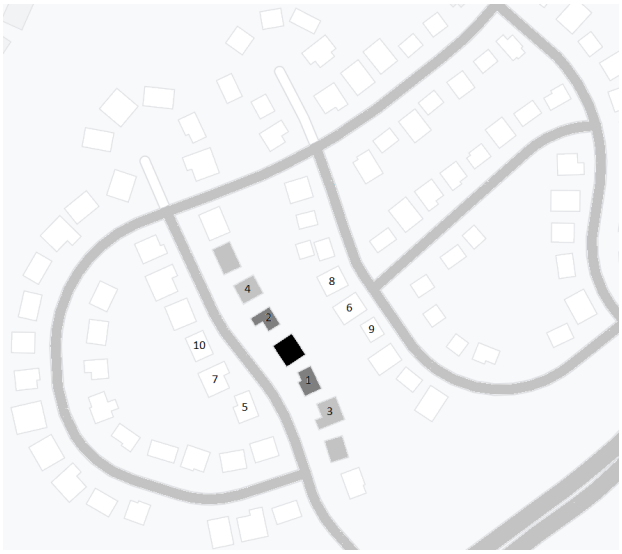


Figure 1: Defining Nearest Neighbors

Notes: This figure visualizes how we define a household's nearest neighbors. Homes and roads are outlined in light gray. Census block borders are thick lines shaded light-gray and often overlapping with streets. We consider a sample household shaded black. Our first definition of nearby neighbors uses a street address algorithm to define adjacent parcels. This household has two next-door neighbors (shaded in dark gray) and four two- or three-doors down neighbors (shaded in light gray). Our second methodology for defining a household's nearest neighbors uses an as-the-crow-flies distance measure. Specifically, we measure the geodistance between each parcel and all other parcels and then identify the one that is closest, the one that is second closest, and so on. The ten homes closest to the given, black-shaded home are marked with a number corresponding to how close they are.

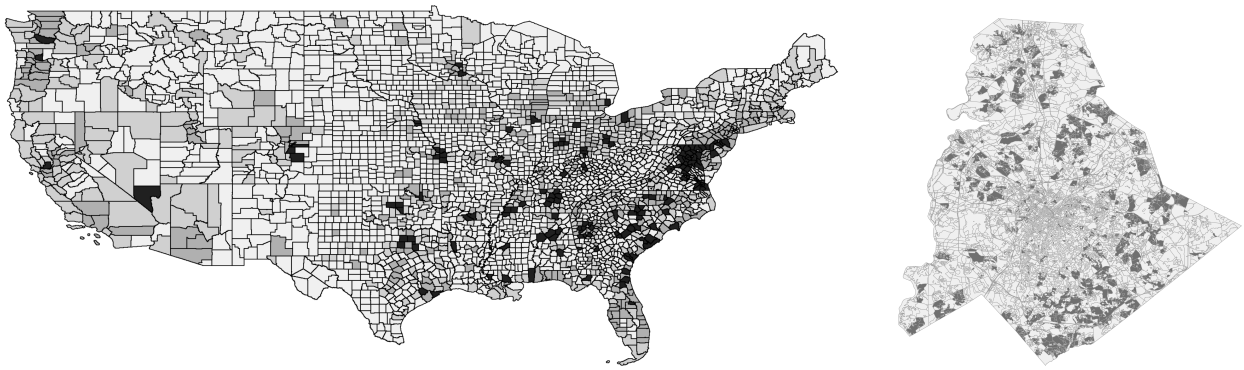


Figure 2: Geographic Coverage of the Sample

Notes: The left panel of this figure shades in the US counties where our "experiments," as defined in the text and described in Table 1, occur. The higher the incidence of experiments, weighted by population, the more darkly the county is shaded. In this way, those places that contribute disproportionately to our sample are the most darkly shaded. The right panel shades in dark gray the census blocks in Mecklenburg County (Charlotte) where at least one "experiment" occurs at some point. Higher resolution figures can be found in the Appendix, see Figure C1 and Figure C2.

Table 1: Summary Stats

	Black Current Residents		*White Current Residents*	
	All	*New 1-Door White Nbr*	*All*	*New 1-Door Black Nbr*
Dependent Variable				
Sell within 2 Years (=100)	4.56	6.28	8.30	10.54
New Nbrs				
New Nbr Different Race 1 Door (=1)	0.01	1.00	0.00	1.00
Property Characteristics				
Year Built	1976	1986	1975	1982
Building Sq Ft	1,766	1,921	1,829	1,848
Resident Characteristics				
Has Co-Borrower (=1)	0.44	0.48	0.61	0.57
Owner Income	$70,658	$77,066	$89,281	$80,375
Tenure (quarters)	29.16	28.45	28.52	29.21
Loan Characteristics				
Loan Age (quarters)	15.83	14.66	13.05	14.26
Refinance (=1)	0.70	0.67	0.72	0.67
Conventional Loan (=1)	0.70	0.65	0.83	0.76
Neighborhood Characteristics				
Census Block Group Median Income	$65,528	$74,907	$80,078	$75,609
Census Block Population Density	5,303	4,851	5,215	5,010
Census Block Group Black Share	0.45	0.26	0.07	0.17
Census Block Group White Share	0.42	0.59	0.80	0.69
Observations	18,509,592	71,724	133,967,854	104,402

Notes: This table describes the sample of current resident-by-quarter observations where the current resident and, if applicable, new neighbor exists in the merged CoreLogic Solutions Real Estate and Home Mortgage Disclosure Act (HMDA) data set. Columns (1) and (3) describe the sample of all "at-risk" Black and White homeowner-by-quarter observations, respectively, as defined in the text. In column (2), we describe the subsample of Black household-by-quarter observations where the Black resident was involved in a valid experiment and received a new White neighbor right next-door. Column (4) does likewise for White current resident-by-quarter observations. Sell within two years is a dummy equal to 100 if the current resident has sold their home within 2 years of the given quarter. New neighbor different race is a dummy equal to 1 if the current resident-by-quarter gets a new different-race neighbor. Property age, property size, whether the loan has a co-borrower, resident tenure, and mortgage characteristics come from CoreLogic Solutions. Owner income at time of mortgage application and race come from HMDA. Median block group income come from the Census Bureau's 2015 American Community Survey. Population density is defined as the population per square mile; population measures comes from the 2010 census. Black share is the share that is either Hispanic or non-Hispanic Black. White share is the share that is non-Hispanic White.

Table 2: Balance Tests

Panel A. Black Current Residents

Dependent Variable:	Bldg Sqft	Year Built	Owner Income	Co-Borrower	Tenure	Loan Age	Refinance	Conventional	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
New 1-Door White Nbr	-2.449	0.0282	-15.70	-0.000644	0.0263	-0.0621	-0.000360	-0.00197	0.00149
	(2.189)	(0.0326)	(218.1)	(0.00247)	(0.0518)	(0.0522)	(0.00185)	(0.00232)	(0.00471)
Controls	X	X	X	X	X	X	X	X	
Fixed Effects									
Block × Quarter	X	X	X	X	X	X	X	X	X
Counts									
N	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785
Fixed Effect Cells	203,931	203,931	203,931	203,931	203,931	203,931	203,931	203,931	203,931
Sample Means									
Dependent Variable	1978	1991	$76,627	0.46	28.41	14.90	0.66	0.63	5.87

Panel B. White Current Residents

Dependent Variable:	Bldg Sqft	Year Built	Owner Income	Co-Borrower	Tenure	Loan Age	Refinance	Conventional	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
New 1-Door Black Nbr	1.348	0.0187	170.2	-0.00239	-0.0392	-0.0204	-0.00327**	0.00115	0.0210***
	(1.590)	(0.0256)	(188.3)	(0.00187)	(0.0414)	(0.0387)	(0.00141)	(0.00162)	(0.00775)
Controls	X	X	X	X	X	X	X	X	
Fixed Effects									
Block × Quarter	X	X	X	X	X	X	X	X	X
Counts									
N	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864
Fixed Effect Cells	254,859	254,859	254,859	254,859	254,859	254,859	254,859	254,859	254,859
Sample Means									
Dependent Variable	1,955	1990	$84,682	0.59	28.87	13.96	0.66	0.77	10.90

Notes: This table estimates the "effect" of a new different-race neighbor on current residents' properties, personal characteristics, and mortgage attributes. The sample includes all current residents who received a new different-race neighbor either next-door or two-/three-doors down. We further include all other current residents on the same block and say that they received a new different-race neighbor elsewhere on the block. Current residents in this sample can receive a new different-race neighbor in up to three non-mutually exclusive distance bins: next-door, one-to-three-doors down (repressed for brevity), and elsewhere on the same block (the omitted group). Control variables include four square feet bins, four year built bins, five income bins, a dummy indicating a co-borrower, four resident tenure bins, five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). We omit the attribute as a control variable when that attribute is the outcome variable. To calculate the index, we first regress sell-next-two-years on these eight attributes and block-by-quarter fixed effects. We then regress the predicted values from that regression on our treatment arms and block-by-quarter fixed effects. The number of fixed effect cells is the number of unique block-by-quarter cells in the estimation sample. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 3: New Different-Race Nearby Neighbors Cause Current Resident Move-Outs

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door White Nbr	0.377***	0.375***	0.374***	0.375***
	(0.128)	(0.128)	(0.128)	(0.128)
Controls				
Building		X	X	X
Resident			X	X
Loan				X
Fixed Effects				
Block × Quarter	X	X	X	X
Counts				
N	1,400,785	1,400,785	1,400,785	1,400,785
Fixed Effect Cells	203,931	203,931	203,931	203,931
Sample Means				
Dependent Variable	5.87	5.87	5.87	5.87

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door Black Nbr	0.313**	0.313**	0.298**	0.292**
	(0.127)	(0.127)	(0.127)	(0.126)
Controls				
Building		X	X	X
Resident			X	X
Loan				X
Fixed Effects				
Block × Quarter	X	X	X	X
Counts				
N	3,699,864	3,699,864	3,699,864	3,699,864
Fixed Effect Cells	254,859	254,859	254,859	254,859
Sample Means				
Dependent Variable	10.90	10.90	10.90	10.90

Notes: This table estimates the effect of a new different-race neighbor on current residents' likelihood of selling their homes within the next two years. The sample includes all current residents who received a new different-race neighbor either next-door or two-/three-doors down. We further include all other current residents on the same block and say that they received a new different-race neighbor elsewhere on the block. Current residents in this sample can receive a new different-race neighbor in up to three non-mutually exclusive distance bins: next-door, one-to-three-doors down (repressed for brevity), and elsewhere on the same block (the omitted group). Building control variables include four square feet bins and four year built bins. Resident control variables include five income bins, a dummy indicating a co-borrower, and four resident tenure bins. Loan control variables include five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). The number of fixed effect cells is the number of unique block-by-quarter cells in the estimation sample. The full set of coefficient estimates is reported in Online Appendix Table C1. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 4: Parcels 1-40, Results

Sample:	Black Current Residents		White Current Residents	
Dependent Variable:	Index	Sold within 2 Years	Index	Sold within 2 Years
	(1)	(2)	(3)	(4)
New 1 to 2 Closest Parcels Diff-Race Nbr	-0.000220	0.274**	0.0229***	0.366***
	(0.00488)	(0.118)	(0.00771)	(0.123)
Controls		X		X
Fixed Effects				
Group × Quarter	X	X	X	X
Counts				
N	1,154,893	1,154,893	1,932,917	1,932,917
Fixed Effect Cells	231,206	231,206	247,301	247,301
Sample Means				
Dependent Variable	5.96	5.96	10.27	10.27

Notes: This table estimates the effect of a new different-race neighbor on current residents' likelihood of selling their homes within the next two years. The sample is current residents who received at least one new different-race neighbor in one of the 40 parcels nearest to them. Current residents can receive a new different-race neighbor in one of the two closest parcels, one of the five closest parcels, one of the parcels between six and ten parcels away, one of the parcels between 11 and 20 parcels away, one of the parcels between 21 and 30 parcels away, and/or one of the parcels between 31 and 40 parcels away, all subject to those parcels being on the same census block group as the current resident. Control variables include four square feet bins, four year built bins, five income bins, a dummy indicating a co-borrower, four resident tenure bins, five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). The full set of coefficient estimates for columns (2) and (4) is reported in Online Appendix Table C4. To calculate the index, we first regress sell-next-two-years on the eight control variables and block group-by-quarter fixed effects. We then regress the predicted values from that regression on our treatment arms and block group-by-quarter fixed effects. The number of fixed effect cells is the number of unique block group-by-quarter cells in the estimation sample. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 5: Price Effects

Dependent Variable:	Log Sale Price			
Sample:	Black Current Residents		White Current Residents	
Sample:	Who Sold Their Homes		Who Sold Their Homes	
	(1)	(2)	(3)	(4)
New 1-Door Diff-Race Nbr	0.0240	0.00829	0.000431	-0.00729
	(0.0165)	(0.0237)	(0.00446)	(0.00593)
Controls	X	X	X	X
Fixed Effects				
Group × Quarter	X		X	
Block × Quarter		X		X
Counts				
N	23,261	7,670	909,908	319,634
Fixed Effect Cells	9,967	3,564	314,100	141,406
Sample Means				
Dependent Variable	12.15	12.22	12.42	12.45

Notes: To create this table, we start with the sample of house sales. We then classify sales as one of three non-mutually exclusive types: those that received a new next-door neighbor of a different race within the last two years, those that received a new 1- to 3-doors down neighbors of a different race within the last two years, and all other sales that occurred in the same quarter and within the same block group (columns (1) and (3)) or block (columns (2) and (4)). Sales might be in this other category because the current residents received no new nearby neighbors recently, because they did, but the new neighbor was of the same race, or because they did, but the new neighbor's race is missing. Control variables include four square feet bins, four year built bins, five income bins, a dummy indicating a co-borrower, four resident tenure bins, five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6: Restrictive Control Group

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door White Nbr	0.514***	0.514***	0.500***	0.512***
	(0.179)	(0.179)	(0.180)	(0.179)
New 1-Door Black Nbr	0.157	0.157	0.170	0.157
	(0.139)	(0.139)	(0.140)	(0.139)
Controls				
Building	X	X	X	X
Resident	X	X	X	X
Loan	X	X	X	X
Nbr Income Bin		X		
Block Group Income Diff Bin			X	
Self Income Diff Bin				X
Fixed Effects				
Group × Quarter	X	X	X	X
Counts				
N	245,198	245,198	243,733	245,198
Fixed Effect Cells	82,870	82,870	82,288	82,870
Sample Means				
Dependent Variable	5.50	5.50	5.50	5.50

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door Black Nbr	0.434***	0.434***	0.427***	0.434***
	(0.152)	(0.152)	(0.153)	(0.152)
New 1-Door White Nbr	0.200***	0.200***	0.201***	0.200***
	(0.0371)	(0.0371)	(0.0372)	(0.0371)
Controls				
Building	X	X	X	X
Resident	X	X	X	X
Loan	X	X	X	X
Nbr Income Bin		X		
Block Group Income Diff Bin			X	
Self Income Diff Bin				X
Fixed Effects				
Group × Quarter	X	X	X	X
Counts				
N	3,396,416	3,396,416	3,382,676	3,396,416
Fixed Effect Cells	838,887	838,887	834,377	838,887
Sample Means				
Dependent Variable	9.78	9.78	9.78	9.78

Notes: To create this table, we use the sample of current residents who received a new neighbor, either same-race or different-race, within three doors. We drop current residents who received both a different-race and a same-race neighbor within three doors in the same quarter. To estimate the effect of receiving new neighbors of a different race we proceed as follows. We include as independent variables in our regression specification two mutually exclusive effects: a dummy for receiving a new same-race neighbor within three doors (omitted group) and a dummy for receiving a new different-race neighbor within three doors (repressed for readability). We simultaneously estimate (and report) the out-sized effect of the new different-race or same-race neighbor being immediately next-door. Control variables include four square feet bins, four year built bins, five income bins, a dummy indicating a co-borrower, four resident tenure bins, five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). Column (2) further controls for five new neighbor income bins. Column (3) controls for six bins for the difference between the new neighbor's income and the block group median income (as measured in the 2015 ACS). In column (4), we include six bins for the difference between the new neighbor's income and the current resident's income. In the event that a current resident received multiple new neighbors, we use the income of the nearest one when calculating new neighbor income. A more complete set of coefficient estimates is reported in Online Appendix Table C7. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 7: Heterogeneity as a Function of Distance Between Neighbors

Dependent Variable:	Sold within 2 Years (=100)			
Sample:	Black Current Residents		White Current Residents	
Subsample:	Nbr <25m	Nbr <100m	Nbr <25m	Nbr<100m
	(1)	(2)	(3)	(4)
New 1-Door Diff-Race Nbr	0.632***		1.132***	
	(0.227)		(0.234)	
New 1-Door Diff-Race Nbr × Distance	-0.130		-0.418***	
	(0.0928)		(0.0954)	
New 0-12.5m, 1-Door Diff-Race Nbr		0.588**		1.248***
		(0.290)		(0.316)
New 12.5-25m, 1-Door Diff-Race Nbr		0.400***		0.331***
		(0.125)		(0.124)
New 25m-100m, 1-Door Diff-Race Nbr		0.0819		-0.0254
		(0.121)		(0.119)
Controls	X	X	X	X
Fixed Effects				
Block × Quarter	X	X	X	X
Counts				
N	1,400,785	2,453,248	3,699,864	7,187,730
Fixed Effect Cells	203,931	306,627	254,859	338,112
Sample Means				
Dependent Variable	5.87	5.42	10.90	9.95
New 1-Door Nbr Distance (Decameters)	1.98	2.62	2.01	2.70

Notes: This table estimates the effect of a new different-race neighbor on current residents' likelihood of selling their homes within the next two years as in column (4) of Table 3. In columns (1) and (3), we keep our sample restriction that drops residents who have no neighbors within 25 meters and then allow the effect of receiving a new next-door neighbor of a different race to vary as a function of the distance (measured in decameters) between the current resident and their new next-door neighbor. In columns (2) and (4), we relax the sample restriction, dropping just residents whose nearest neighbor is farther than 100 meters away. As before, current residents in this sample can receive a new different-race neighbor one-to-three-doors down (repressed for brevity) and elsewhere on the same block (the omitted group). But here, instead of a single dummy variable for receipt of a new different-race neighbor immediately next-door, we split this treatment category into three mutually exclusive categories depending on how nearby the next-door house is. Control variables include four square feet bins, four year built bins, five income bins, a dummy indicating a co-borrower, four resident tenure bins, five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). The number of fixed effect cells is the number of unique block-by-quarter cells in the estimation sample. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 8: Regional Heterogeneity

Dependent Variable:	Sold within 2 Years (=100)			
Sample:	Black Current Residents		White Current Residents	
Subsample:	South	Northern	South	Northern
	(1)	(2)	(3)	(4)
New 1-Door Diff-Race Nbr	0.375**	0.297	0.239	0.510**
	(0.158)	(0.252)	(0.185)	(0.201)
Controls	X	X	X	X
Fixed Effects				
Block × Quarter	X	X	X	X
Counts				
N	1,064,518	243,290	2,167,527	1,028,132
Fixed Effect Cells	132,252	50,127	135,420	88,972
Sample Means				
Dependent Variable	5.95	5.32	11.67	9.50

Notes: This table estimates the effect of a new different-race neighbor on current residents' likelihood of selling their homes within the next two years as in column (4) of Table 3. The sample is split into groups based on the state of the residents. We follow the region definitions from the US Census characterizing the Southern states as those in the South region and the Northern states as those in the Northeast and Midwest, see https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 9: Interacting Treatment Effects with Neighbor Income and Current Resident Income

Dependent Variable:	Current Resident Sold within 2 Years (=100)					
Sample:	Black Current Residents			White Current Residents		
Income Measure:	Block Group	Resident's	New Neighbor's	Block Group	Resident's	New Neighbor's
	(1)	(2)	(3)	(4)	(5)	(6)
New 1-Door Diff-Race Nbr × Above Median Income	0.326*	0.558***	0.241	0.208	0.433**	0.159
	(0.189)	(0.186)	(0.200)	(0.194)	(0.187)	(0.195)
New 1-Door Diff-Race Nbr × Below Median Income	0.389**	0.194	0.385**	0.358**	0.175	0.433**
	(0.175)	(0.177)	(0.189)	(0.167)	(0.173)	(0.179)
Controls	X	X	X	X	X	X
Fixed Effects						
Block × Quarter	X	X	X	X	X	X
Counts						
N	1,395,971	1,400,785	897,303	3,688,974	3,699,864	2,470,259
Fixed Effect Cells	202,900	203,931	134,810	253,620	254,859	160,028
Sample Means						
Dependent Variable	5.88	5.87	5.84	10.91	10.90	10.91

Notes: In this table, we investigate how the effect of receiving a new different-race neighbor varies with three different income measures. Our first income measure is the median income of all residents of the block group as per the 2015 American Community Survey. Our second measure is the current resident's income. And our third is the income of the new neighbor. To create this table, we first begin with the same sample used to create column (4) of Table 3. We then create two dummy variables, above and below median, for each of the three measures and interact these dummy variables with a dummy for receiving a new different-race neighbor immediately next-door. We also include a dummy for receiving a new different-race neighbor within 3 doors, and a dummy for receiving a new different-race neighbor elsewhere on the block. Columns (1) and (4) therefore compare the treatment effect in high income vs low income neighborhoods. Columns (2) and (5) compare the treatment effect between high income and low income current residents. And columns (3) and (6) compare the treatment effect between current residents who receive a high income versus low income new neighbor. Note that to unambiguously define the income of the new neighbor, we further restrict the sample that created column (4) of Table 3 to just those block-quarters that received exactly one new different-race neighbor. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Distinguishing Causes of Neighborhood Racial Change: A Nearest Neighbor Design

Pat Bayer, Marcus Casey, Ben McCartney, John Orellana-Li, and Calvin Zhang

Online Appendix

A Describing the CoreLogic - HMDA Match

To build our crosswalk between CoreLogic and the publicly available HMDA data, we first clean and standardize the following variables in both data sets: census tract (being careful to use either the 2000 on 2010 tract definitions as appropriate), year of mortgage application, loan purpose (purchase or refinance), loan type (conventional, FHA, VA, or FSA/RHS), presence of a coapplicant/coborrower, mortgage amount, and lender name. All together, there are seven variables along which a CoreLogic mortgage and HMDA mortgage can match.

Second, we join the CoreLogic and HMDA datasets in seven rounds. In our first round, we require matches on all seven variables. Census tract, year of application, purpose, type, and co-borrower need to match exactly. We round the mortgage amount in CoreLogic using the same rounding rules as HMDA and this rounded mortgage amount must also match exactly. To compare lender names in CoreLogic and HMDA we calculate the string difference in the cleaned lender names using the Stata command strdist and then divide that distance by the length of the lender's name in CoreLogic. We say that the lender names match if this lender comparison variable is less than 0.5. If loans match one-to-one we remove those loans from the CoreLogic and HMDA lists of loans and then attempt to rematch. If there are multiple matches, we keep the one with the closest lender name and then remove both the CoreLogic and HMDA lists.

In the following rounds, we relax the matching requirements. In our second round, we join the unmatched loans again and require a match on all seven variables except presence of a co-applicant. Our third requires a match on all seven variables except loan type. Our fourth round requires a match on all seven variables except presence of a co-applicant and loan type. In our fifth round, we match on all six of the numeric variables and require that the lender name sound is the same (using the Stata command soundex). Round six is the same as round five but without the requirement of same presence of co-applicant and loan type. Our seventh and final attempt and finding a match uses all of the seven matching variables except loan amount and then declares a match successful if there is a unique loan with a loan amount less than $2,000 different for loans between 2005 and 2017 or less than $5,000 different for loans between 2018 and 2022.

Relative to some CoreLogic - HMDA crosswalks, we take a very conservative approach to ensure the fidelity of our measure of race for households in our sample. Our overall match rate, mortgages in

CoreLogic for which we find a unique match in HMDA, is 90,338,856 out of 152,727,515 mortgages, or 59.2%. Of these 90,338,856 matches, 76% are matched in phase 1 (that requires exact matching on all seven matching variables) and a further 14% are matched in phase 2 (that relaxes the requirement of a matching presence of co-applicant).

B Describing the Neighbor Identifying Algorithm

To identify each property's nearest neighbors, we first build a cross-section of all unique properties, defined using CoreLogic's unique parcel identifier, in CoreLogic as follows. We use five historical assessor files corresponding, roughly, to the years 2018 through 2022, though the dates that each assessment took place depend on when the county last assessed that particular property. We define each property based on its most recent assessment up to 2020. To be included in our sample we require each property to have non-missing geolocations, be either a single family residence, duplex, triplex, or quadplex and have a residential land use code. We further require that street addresses be unique. Note, these parcels are the universe we will define neighbors for and require only coverage in CoreLogic's assessor files. We do not require that these parcels have ever changed ownership or have had any mortgages issued against them.

Our neighbor matching algorithm uses this cross-section and creates two new variables that assign each property a side-of-street and block-of-street. We say that odd numbered homes and even numbered homes are on opposite sides of the street. To define the block of the street each parcel is on, we rely on the convention that homes on the same street, but on different street blocks, have house numbers with different 100s digits (e.g. 1203 is on a different block than 1153). We thus denote street blocks through a variable that takes the floor of the house number divided by 100. We then order homes within block-of-street and side-of-street and use house numbers to define neighbors. We define the neighboring property with the smallest (absolute) difference in street numbers as "next-door", second smallest difference as two-door-down, and third smallest difference as three-doors-down. Thus, every property is assigned up to six same street side neighbors, three on the left, and three on the right.

To ensure that this algorithm works as intended, we drop the following edge cases. First, we drop homes with neighbors that are "misordered" in terms of Euclidean distance. That is, we require (i)

that a neighbor we define as three-doors down is not closer in distance than either of the neighbors one- or two-doors down in the same direction and (ii) that a neighbor we say is two-doors-down is not closer than the next-door neighbor in the same direction. Second, we drop all census blocks where two or more distinct parcels have identical geolocations, which drops blocks with imprecise geolocations. Third, we drop census blocks with an instance of five or more distinct parcels having the same street address. Up to four represents a quadplex (which are included subject to our other requirements), but more than that is likely a mislabeled apartment building. Fourth, we drop census blocks that are not 100% single or multi-family (up to quadplex). There are some cases where residences are separated by something other than another residence or empty space, and we do not want to define them as neighbors. Fifth, and finally, we require that next-door neighbors be not farther apart than 0.1 kilometers.

C Supplemental Figures and Tables

Figure C1: Counties in the Sample

This figure is a larger, more detailed version of the left panel of Figure 2.

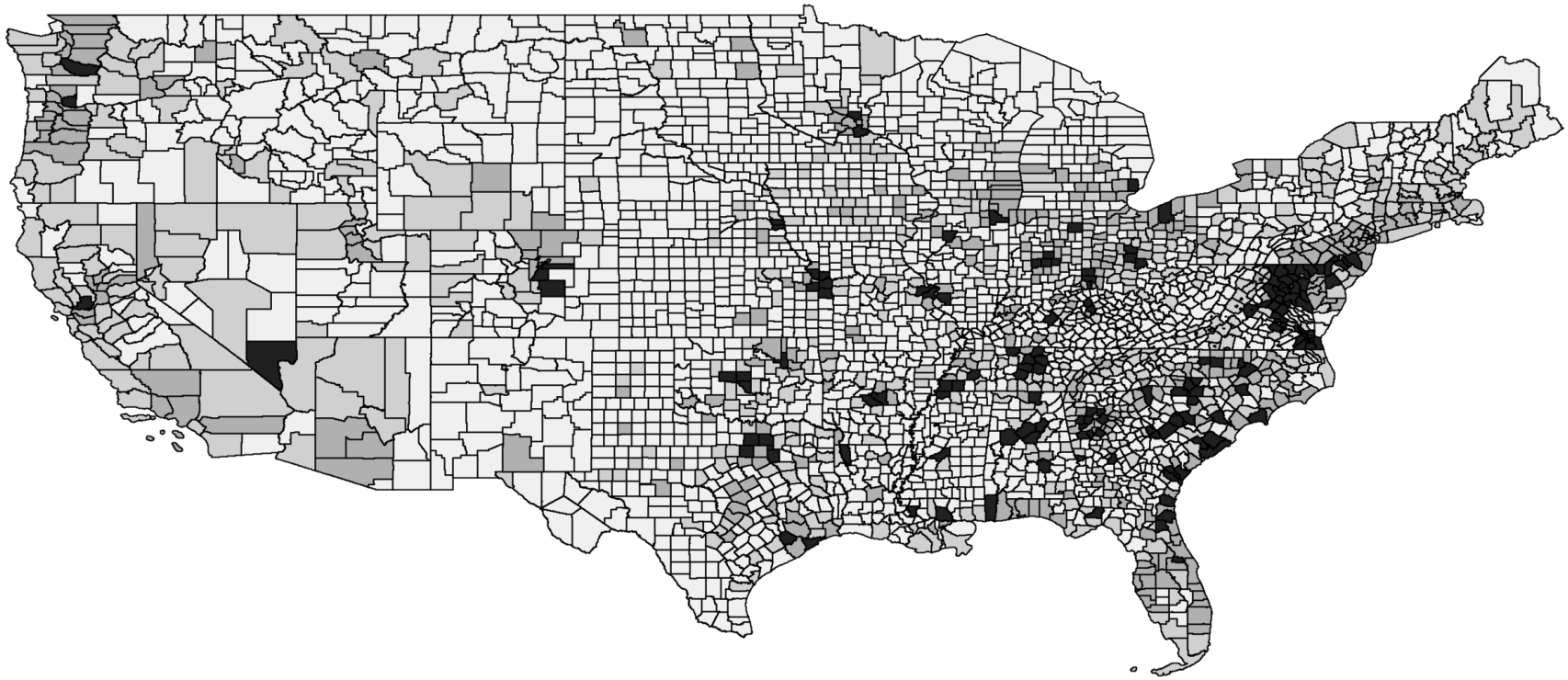
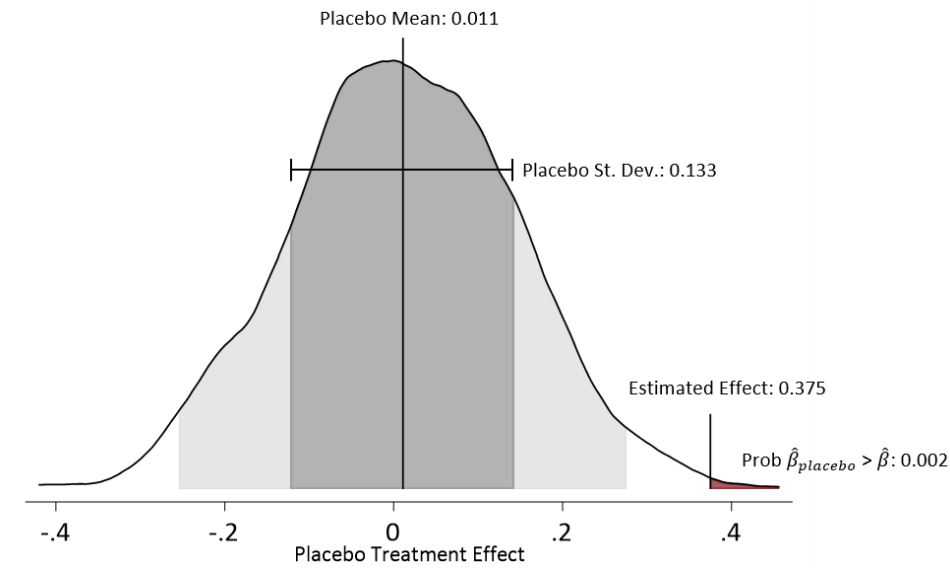


Figure C2: Census Blocks in the Sample (Charlotte, NC)

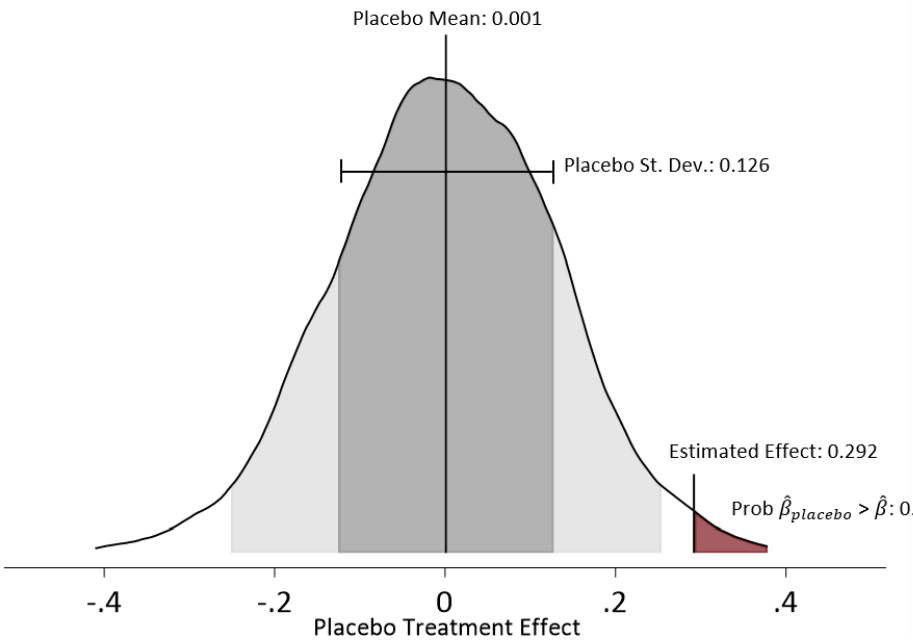
This figure is a larger, more detailed version of the right panel of Figure 2.

Figure C3: Random Re-Assignment of Treatment Placebo Test

Panel A. Black Current Residents



Panel B. White Current Residents



Notes: To create this figure we conduct 1,000 placebo tests for each of the Black and White samples. Specifically, we use the same sample as column (4) of Table 3, but randomly assign each current resident to a treatment status, subject to each block-by-quarter fixed effect cell containing the same count of each type of treatment as the real sample. We then estimate the treatment effect using each of the 1,000 placebo samples. This figure plots the probability density function of these treatment effects.

Table C1: Complete Set of Estimated Coefficients for Table 3

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door White Nbr	0.377***	0.375***	0.374***	0.375***
	(0.128)	(0.128)	(0.128)	(0.128)
New 1-/3-Door White Nbr	0.0607	0.0550	0.0549	0.0540
	(0.0823)	(0.0824)	(0.0823)	(0.0823)
Building Controls				
Bldg Sq Ft: Less than 1,250 (omitted)				
Bldg Sq Ft: 1,250 to 1,999		0.00248	-0.0648	-0.0667
		(0.104)	(0.105)	(0.104)
Bldg Sq Ft: 2,000 to 2,999		-0.448***	-0.624***	-0.615***
		(0.125)	(0.126)	(0.126)
Bldg Sq Ft: 3,000 or more		-1.043***	-1.346***	-1.336***
		(0.171)	(0.174)	(0.174)
Year Built: Before 1960 (omitted)				
Year Built: 1960 to 1979		-0.324	-0.334	-0.338
		(0.209)	(0.209)	(0.209)
Year Built: 1980 to 1999		-0.122	-0.198	-0.223
		(0.243)	(0.242)	(0.242)
Year Built: 2000 or later		0.732***	0.551**	0.497*
		(0.259)	(0.258)	(0.258)
Resident Controls				
Resident Income: $50,000 or less (omitted)				
Resident Income: $50,001 to $75,000			0.386***	0.403***
			(0.0737)	(0.0738)
Resident Income: $75,001 to $100,000			0.590***	0.638***
			(0.0898)	(0.0902)
Resident Income: $100,001 to $150,000			0.856***	0.938***
			(0.103)	(0.104)
Resident Income: $150,001 or more			1.296***	1.410***
			(0.154)	(0.154)
Co-Borrower (=1)			-0.120**	-0.125**
			(0.0605)	(0.0606)
Tenure: 5 to 8 quarters (omitted)				
Tenure: 9 to 16 quarters			1.003***	0.875***
			(0.106)	(0.114)
Tenure: 17 to 24 quarters			1.214***	1.262***
			(0.120)	(0.136)
Tenure: 25 quarters or more			-0.215**	0.162
			(0.104)	(0.138)

table continued on next page...

...table continued from previous page

Outstanding Loan Controls				
Loan Age: 4 quarters or fewer (omitted)				
Loan Age: 5 to 8 quarters				0.545***
				(0.0762)
Loan Age: 9 to 16 quarters				0.885***
				(0.0829)
Loan Age: 17 to 24 quarters				0.884***
				(0.101)
Loan Age: 25 quarters or more				0.926***
				(0.102)
Loan Purpose: Refinance (=1)				-0.605***
				(0.0819)
Loan Type: Conventional (omitted)				
Loan Type: FHA Loan				0.497***
				(0.0720)
Loan Type: VA Loan				1.377***
				(0.113)
Fixed Effects				
Block × Quarter	X	X	X	X
Counts				
N	1,400,785	1,400,785	1,400,785	1,400,785
Fixed Effect Cells	203,931	203,931	203,931	203,931
Sample Means				
Dependent Variable	5.87	5.87	5.87	5.87

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door Black Nbr	0.313**	0.313**	0.298**	0.292**
	(0.127)	(0.127)	(0.127)	(0.126)
New 1/3-Door Black Nbr	0.155*	0.140*	0.139*	0.138*
	(0.0821)	(0.0821)	(0.0820)	(0.0820)
Building Controls				
Bldg Sq Ft: Less than 1,250 (omitted)				
Bldg Sq Ft: 1,250 to 1,999		-0.258***	-0.287***	-0.241***
		(0.0746)	(0.0745)	(0.0744)
Bldg Sq Ft: 2,000 to 2,999		-0.838***	-0.919***	-0.801***
		(0.0899)	(0.0908)	(0.0907)
Bldg Sq Ft: 3,000 or more		-1.242***	-1.459***	-1.272***
		(0.125)	(0.127)	(0.126)
Year Built: Before 1960 (omitted)				
Year Built: 1960 to 1979		-0.105	-0.0729	-0.0589
		(0.138)	(0.137)	(0.137)
Year Built: 1980 to 1999		0.473***	0.431***	0.443***
		(0.157)	(0.156)	(0.156)
Year Built: 2000 or later		1.974***	1.643***	1.643***
		(0.173)	(0.172)	(0.172)
Resident Controls				
Resident Income: $50,000 or less (omitted)				
Resident Income: $50,001 to $75,000			0.401***	0.423***
			(0.0571)	(0.0570)
Resident Income: $75,001 to $100,000			0.560***	0.652***
			(0.0658)	(0.0656)
Resident Income: $100,001 to $150,000			0.691***	0.864***
			(0.0700)	(0.0700)
Resident Income: $150,001 or more			1.248***	1.465***
			(0.0926)	(0.0926)
Co-Borrower (=1)			-0.373***	-0.346***
			(0.0454)	(0.0454)
Tenure: 5 to 8 quarters (omitted)				
Tenure: 9 to 16 quarters			1.716***	1.911***
			(0.0804)	(0.0863)
Tenure: 17 to 24 quarters			1.249***	2.037***
			(0.0889)	(0.100)
Tenure: 25 quarters or more			-2.119***	-0.294***
			(0.0751)	(0.1000)

table continued on next page...

...table continued from previous page

Outstanding Loan Controls				
Loan Age: 4 quarters or fewer (omitted)				
Loan Age: 5 to 8 quarters				0.705***
				(0.0563)
Loan Age: 9 to 16 quarters				0.862***
				(0.0591)
Loan Age: 17 to 24 quarters				0.628***
				(0.0718)
Loan Age: 25 quarters or more				0.361***
				(0.0735)
Loan Purpose: Refinance (=1)				-1.829***
				(0.0602)
Loan Type: Conventional (omitted)				
Loan Type: FHA Loan				1.455***
				(0.0620)
Loan Type: VA Loan				2.641***
				(0.0980)
Fixed Effects				
Block × Quarter	X	X	X	X
Counts				
N	3,699,864	3,699,864	3,699,864	3,699,864
Fixed Effect Cells	254,859	254,859	254,859	254,859
Sample Means				
Dependent Variable	10.90	10.90	10.90	10.90

Notes: This table shows the full estimates of Table 3.

Table C2: Alternative Choices of Clustering Level(s)

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New 1-Door White Nbr	0.375***	0.375***	0.375***	0.375***	0.375***	0.375***	0.375***	0.375***
	(0.128)	(0.129)	(0.128)	(0.128)	(0.120)	(0.139)	(0.121)	(0.138)
Controls	X	X	X	X	X	X	X	X
Fixed Effects:								
Block × Quarter	X	X	X	X	X	X	X	X
Cluster Level	Single	Single	Single	Single	Single	Single	Double	Double
Cluster Detail								
Tract × Year	X							
Block × Quarter		X						
Tract			X				X	
Block				X				X
Year					X		X	
Quarter						X		X
Counts								
N	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785	1,400,785
Fixed Effect Cells	203,931	203,931	203,931	203,931	203,931	203,931	203,931	203,931
Sample Means								
Dependent Variable	5.87	5.87	5.87	5.87	5.87	5.87	5.87	5.87

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New 1-Door Black Nbr	0.292**	0.292**	0.292**	0.292**	0.292**	0.292**	0.292**	0.292**
	(0.126)	(0.127)	(0.128)	(0.127)	(0.127)	(0.112)	(0.128)	(0.112)
Controls	X	X	X	X	X	X	X	X
Fixed Effects:								
Block × Quarter	X	X	X	X	X	X	X	X
Cluster Level	Single	Single	Single	Single	Single	Single	Double	Double
Cluster Detail								
Tract × Year	X							
Block × Quarter		X						
Tract			X				X	
Block				X				X
Year					X		X	
Quarter						X		X
Counts								
N	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864	3,699,864
Fixed Effect Cells	254,859	254,859	254,859	254,859	254,859	254,859	254,859	254,859
Sample Means								
Dependent Variable	10.90	10.90	10.90	10.90	10.90	10.90	10.90	10.90

Notes: This table presents the results of estimating column (4) of Table 3 under different clustering regimes. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table C3: Parcels 1-40, Balance

Panel A. Black Current Residents

	Bldg Sqft	Year Built	Owner Income	Co-Borrower	Tenure	Loan Age	Refinance	Conventional	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
New 1 or 2 Closest Parcels White Nbr	0.774	-0.000913	180.7	-0.000691	-0.0295	-0.0157	-0.000172	-0.00222	-0.000220
	(2.141)	(0.0336)	(196.6)	(0.00226)	(0.0496)	(0.0492)	(0.00174)	(0.00212)	(0.00488)
Controls	X	X	X	X	X	X	X	X	
Fixed Effects									
Group × Quarter	X	X	X	X	X	X	X	X	X
Counts									
N	1,154,893	1,154,893	1,154,893	1,154,893	1,154,893	1,154,893	1,154,893	1,154,893	1,154,893
Fixed Effect Cells	231,206	231,206	231,206	231,206	231,206	231,206	231,206	231,206	231,206
Sample Means									
Dependent Variable Mean	1,933	1986	$76,707	0.48	27.78	15.04	0.67	0.65	5.96

Panel B. White Current Residents

	Bldg Sqft	Year Built	Owner Income	Co-Borrower	Tenure	Loan Age	Refinance	Conventional	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
New 1 or 2 Closest Parcels Black Nbr	4.368***	0.102***	-92.83	-0.00486***	-0.0731*	-0.0770*	-0.00271*	0.00118	0.0229***
	(1.565)	(0.0260)	(185.1)	(0.00184)	(0.0411)	(0.0394)	(0.00141)	(0.00159)	(0.00771)
Controls	X	X	X	X	X	X	X	X	
Fixed Effects									
Group × Quarter	X	X	X	X	X	X	X	X	X
Counts									
N	1,932,917	1,932,917	1,932,917	1,932,917	1,932,917	1,932,917	1,932,917	1,932,917	1,932,917
Fixed Effect Cells	247,301	247,301	247,301	247,301	247,301	247,301	247,301	247,301	247,301
Sample Means									
Dependent Variable Mean	1,844	1982	$80,812	0.57	28.67	14.54	0.68	0.77	10.27

Notes: This table estimates the "effect" of receiving a new different-race neighbor on current residents' properties, personal characteristics, and mortgage attributes. The sample includes all Current residents can receive a new different-race neighbor in one of the two closest parcels, one of the five closest parcels, one of the parcels between six and ten parcels away, one of the parcels between 11 and 20 parcels away, one of the parcels between 21 and 30 parcels away, and/or one of the parcels between 31 and 40 parcels away, all subject to those parcels being on the same census block group as the current resident. Control variables include four square feet bins, four year built bins, five income bins, a dummy indicating a co-borrower, four resident tenure bins, five loan age bins, a dummy for the loan's purpose (refinance or purchase) and a categorical variable for loan type (Conventional, FHA, and VA). We omit the attribute as a control variable when that attribute is the outcome variable.To calculate the index, we first regress sell-next-two-years on these eight control variables and block group-by-quarter fixed effects. We then regress the predicted values from that regression on our treatment arms and block group-by-quarter fixed effects. The number of fixed effect cells is the number of unique block group-by-quarter cells in the estimation sample. Standard errors, adjusted for clustering at the tract-year level, are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table C4: Complete Set of Estimated Coefficients for Table 4

Dependent Variable:	Current Resident Sold within 2 Years (=100)	
Sample:	Black Current Residents	White Current Residents
	(1)	(2)
New 1 or 2 Closest Parcels Diff-Race Nbr	0.274**	0.366***
	(0.118)	(0.123)
New 1 to 5 Closest Parcels Diff-Race Nbr	0.159*	0.171*
	(0.0486)	(0.0908)
New 6 to 10 Closest Parcels Diff-Race Nbr	0.119	0.183**
	(0.0784)	(0.0797)
New 11 to 20 Closest Parcels Diff-Race Nbr	0.152**	0.0940
	(0.0650)	(0.0655)
New 21 to 30 Closest Parcels Diff-Race Nbr	0.0694	0.0890
	(0.0641)	(0.0647)
New 31 to 40 Closest Parcels Diff-Race Nbr (omitted)		
Building Controls		
Bldg Sq Ft: Less than 1,250 (omitted)		
Bldg Sq Ft: 1,250 to 1,999	-0.113	-0.153*
	(0.0998)	(0.0837)
Bldg Sq Ft: 2,000 to 2,999	-0.641***	-0.656***
	(0.122)	(0.107)
Bldg Sq Ft: 3,000 or more	-1.462***	-0.929***
	(0.176)	(0.164)
Year Built: Before 1960 (omitted)		
Year Built: 1960 to 1979	-0.220	-0.474***
	(0.176)	(0.152)
Year Built: 1980 to 1999	0.128	0.171
	(0.209)	(0.191)
Year Built: 2000 or later	1.228***	1.335***
	(0.224)	(0.216)
Resident Controls		
Resident Income: $50,000 or less (omitted)		
Resident Income: $50,001 to $75,000	0.509***	0.468***
	(0.0759)	(0.0686)
Resident Income: $75,001 to $100,000	0.874***	0.613***
	(0.0926)	(0.0801)
Resident Income: $100,001 to $150,000	1.001***	0.954***
	(0.106)	(0.0885)
Resident Income: $150,001 or more	1.577***	1.562***
	(0.154)	(0.119)
Co-Borrower (=1)	-0.291***	-0.320***
	(0.0624)	(0.0556)
Tenure: 5 to 8 quarters (omitted)		
Tenure: 9 to 16 quarters	0.874***	1.795***
	(0.119)	(0.112)
Tenure: 17 to 24 quarters	1.151***	2.152***
	(0.138)	(0.127)
Tenure: 25 quarters or more	-0.00866	-0.103
	(0.136)	(0.125)

table continued on next page...

...table continued from previous page

Outstanding Loan Controls		
Loan Age: 4 quarters or fewer (omitted)		
Loan Age: 5 to 8 quarters	0.573***	0.724***
	(0.0836)	(0.0755)
Loan Age: 9 to 16 quarters	0.767***	0.899***
	(0.0868)	(0.0763)
Loan Age: 17 to 24 quarters	0.770***	0.677***
	(0.104)	(0.0894)
Loan Age: 25 quarters or more	0.778***	0.269***
	(0.102)	(0.0892)
Loan Purpose: Refinance (=1)	-0.834***	-2.069***
	(0.0822)	(0.0737)
Loan Type: Conventional (omitted)		
Loan Type: FHA Loan	0.571***	1.513***
	(0.0750)	(0.0753)
Loan Type: VA Loan	1.302***	2.845***
	(0.115)	(0.127)
Fixed Effects		
Group × Quarter	X	X
Counts		
N	1,154,893	1,932,917
Fixed Effect Cells	231,206	247,301
Sample Means		
Dependent Variable Mean	5.96	10.27

Notes: This table shows the full estimates of Table 4.

Table C5: Robustness to Binary Treatment Specifications

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)		
Research Design			
Treatment Category 1:	New 1-Door Diff Race Nbr	New 1-Door Diff Race Nbr	New 1-Door Diff Race Nbr
Treatment Category 2:	New 1/3-Door Diff Race Nbr		
Control Group:	New 4+ Diff Race Nbr	New 2/3-Door Diff Race Nbr	New 2+ Diff Race Nbr
	(1)	(2)	(3)
New 1-Door Diff-Race Nbr	0.375***	0.491**	0.425***
	(0.128)	(0.227)	(0.103)
Controls	X	X	X
Fixed Effects			
Block × Quarter	X	X	X
Counts			
N	1,400,785	67,795	1,400,785
Fixed Effect Cells	203,931	29,778	203,931
Sample Means			
Dependent Variable	5.87	5.86	5.87

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)		
Research Design			
Treatment Category 1:	New 1-Door Diff Race Nbr	New 1-Door Diff Race Nbr	New 1-Door Diff Race Nbr
Treatment Category 2:	New 1/3-Door Diff Race Nbr		
Control Group:	New 4+ Diff Race Nbr	New 2/3-Door Diff Race Nbr	New 2+ Diff Race Nbr
	(1)	(2)	(3)
New 1-Door Diff-Race Nbr	0.292**	0.443**	0.420***
	(0.126)	(0.181)	(0.102)
Controls	X	X	X
Fixed Effects			
Block × Quarter	X	X	X
Counts			
N	3,699,864	166,620	3,699,864
Fixed Effect Cells	254,859	67,955	254,859
Sample Means			
Dependent Variable	10.90	10.77	10.90

Notes: Column (1) of this table reproduces the estimates show in (4) of Table 3. Columns (2) and (3) each make a simplification that makes the main independent variable a dummy. Column (2) limits the sample to just those current residents who received a new different-race neighbor within 3 doors and then compares those whose new new neighbor was right next-door to those whose new neighbor was two- or three-doors down. Column (3) uses the same sample as the main sample but compares those whose new new neighbor was right next-door to those whose new neighbor was somewhere else on the block. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table C6: Robustness to Alternative Samples

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)						
Sample Change:	None	No Prop Char Reqs	No Min Density	No Hisp Share Max	No IQR Max	No Tenure Min	High Black Share MSAs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New 1-Door White Nbr	0.375***	0.369***	0.249**	0.382***	0.323***	0.354***	0.301**
	(0.128)	(0.128)	(0.121)	(0.110)	(0.125)	(0.118)	(0.142)
Controls	X	X	X	X	X	X	X
Fixed Effects							
Block × Quarter	X	X	X	X	X	X	X
Counts							
N	1,400,785	1,401,903	1,573,109	1,910,553	1,458,557	1,581,581	1,191,027
Fixed Effect Cells	203,931	204,188	234,367	279,857	213,581	221,793	154,836
Sample Means							
Dependent Variable	5.87	5.87	6.16	6.07	5.84	5.67	5.75

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)						
Sample Change:	None	No Prop Char Reqs	No Min Density	No Hisp Share Max	No IQR Max	No Tenure Min	High Black Share MSAs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New 1-Door Black Nbr	0.292**	0.292**	0.250**	0.240**	0.292**	0.239**	0.412***
	(0.126)	(0.126)	(0.116)	(0.109)	(0.123)	(0.118)	(0.151)
Controls	X	X	X	X	X	X	X
Fixed Effects							
Block × Quarter	X	X	X	X	X	X	X
Counts							
N	3,699,864	3,703,501	4,196,242	4,961,672	3,844,582	4,174,187	2,670,678
Fixed Effect Cells	254,859	255,058	308,656	344,299	269,030	271,049	178,116
Sample Means							
Dependent Variable	10.90	10.90	11.33	10.96	10.84	10.54	10.86

Notes: Column (1) of this table reproduces the estimates shown in (4) of Table 3. Columns (2), (3), (4), (5), and (6) each relax one of the sample restrictions we apply when making our estimation sample and Column (7) adds a sample restriction. Specifically, Column (2) includes homes built before 1900 and with building square feet over 6,000. Column (3) includes blocks with a housing unit count under 20 and population density under 500 as per the 2010 census. Column (4) includes block groups with Hispanic share over 10%. Column (5) includes blocks where the interquartile range of age homes on the block is greater than 30 years. Column (6) includes all current residents, regardless of tenure. Finally, column (7) restricts our main sample to just those counties in metropolitan statistical areas with (1) at least 200,000 residents and (2) in which the Black share is at least as large as the Black share in the US population, 12 percent, during the study period. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table C7: Complete Set of Estimated Coefficients for Table 6

Panel A. Black Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door White Nbr	0.514***	0.514***	0.500***	0.512***
	(0.179)	(0.179)	(0.180)	(0.179)
New 1-Door Black Nbr	0.157	0.157	0.170	0.157
	(0.139)	(0.139)	(0.140)	(0.139)
New 1/3-Door White Nbr	-0.0429	-0.0232	-0.0115	-0.0201
	(0.205)	(0.205)	(0.205)	(0.205)
New 1/3-Door Black Nbr (omitted)				
New Neighbor Income				
New Nbr Income: $50,000 or less (omitted)				
New Nbr Income: $50,001 to $75,000		0.144		
		(0.199)		
New Nbr Income: $75,001 to $100,000		-0.0352		
		(0.242)		
New Nbr Income: $100,001 to $150,000		-0.570**		
		(0.279)		
New Nbr Income: $150,001 or more		-0.252		
		(0.413)		
New Neighbor's Income - Block Group Median Income				
Difference: -$50,001 or less (omitted)				
Difference: -$50,000 to -$20,001			-0.406	
			(0.410)	
Difference: -$20,000 to -$1			-0.633	
			(0.430)	
Difference: $0 to $19,999			-0.750*	
			(0.443)	
Difference: $20,000 to $49,999			-0.869*	
			(0.464)	
Difference: $50,000 or more			-0.829*	
			(0.497)	

table continued on next page...

...table continued from previous page

New Neighbor's Income - Current Resident's Income				
Difference: -$50,001 or less (omitted)				
Difference: -$50,000 to -$20,001				-0.505*
				(0.277)
Difference: -$20,000 to -$1				-0.506*
				(0.298)
Difference: $0 to $19,999				-0.534*
				(0.318)
Difference: $20,000 to $49,999				-0.510
				(0.344)
Difference: $50,000 or more				-1.102***
				(0.392)
Controls				
Building	X	X	X	X
Resident	X	X	X	X
Loan	X	X	X	X
Fixed Effects				
Group × Quarter	X	X	X	X
Counts				
N	245,198	245,198	243,733	245,198
Fixed Effect Cells	82,870	82,870	82,288	82,870
Sample Means				
Dependent Variable Mean	5.50	5.50	5.50	5.50

Panel B. White Current Residents

Dependent Variable:	Current Resident Sold within 2 Years (=100)			
	(1)	(2)	(3)	(4)
New 1-Door Black Nbr	0.434***	0.434***	0.427***	0.434***
	(0.152)	(0.152)	(0.153)	(0.152)
New 1-Door White Nbr	0.200***	0.200***	0.201***	0.200***
	(0.0371)	(0.0371)	(0.0372)	(0.0371)
New 1/3-Door Black Nbr	0.0814	0.0736	0.0804	0.0759
	(0.123)	(0.123)	(0.123)	(0.123)
New 1/3-Door White Nbr (omitted)				
New Neighbor Income				
New Nbr Income: $50,000 or less (omitted)				
New Nbr Income: $50,001 to $75,000		0.0595		
		(0.0707)		
New Nbr Income: $75,001 to $100,000		-0.0918		
		(0.0803)		
New Nbr Income: $100,001 to $150,000		-0.384***		
		(0.0846)		
New Nbr Income: $150,001 or more		-0.469***		
		(0.105)		
New Neighbor's Income - Block Group Median Income				
Difference: -$50,001 or less (omitted) (omitted)				
Difference: -$50,000 to -$20,001			-0.227**	
			(0.116)	
Difference: -$20,000 to -$1			-0.314***	
			(0.120)	
Difference: $0 to $19,999			-0.422***	
			(0.124)	
Difference: $20,000 to $49,999			-0.549***	
			(0.127)	
Difference: $50,000 or more			-0.726***	
			(0.129)	

table continued on next page...

...table continued from previous page

New Neighbor's Income - Current Resident's Income				
Difference: -$50,001 or less (omitted)				
Difference: -$50,000 to -$20,001				-0.332***
				(0.0789)
Difference: -$20,000 to -$1				-0.410***
				(0.0844)
Difference: $0 to $19,999				-0.527***
				(0.0886)
Difference: $20,000 to $49,999				-0.683***
				(0.0937)
Difference: $50,000 or more				-0.929***
				(0.102)
Controls				
Building	X	X	X	X
Resident	X	X	X	X
Loan	X	X	X	X
Fixed Effects				
Group × Quarter	X	X	X	X
Counts				
N	3,396,416	3,396,416	3,382,676	3,396,416
Fixed Effect Cells	838,887	838,887	834,377	838,887
Sample Means				
Dependent Variable Mean	9.78	9.78	9.78	9.78

Notes: This table shows the full estimates of Table 6.