Grantham Research Institute on Climate Change and the Environment

Disparities in pollution capitalization rates

Joshua Graff Zivin and Gregor Singer

June 2026

Grantham Research Institute on Climate Change and the Environment Working Paper No. 392

ISSN 2515-5717 (Online)

Grantham Foundation for the Protection of the Environment

LSE Global School of Sustainability

The Grantham Research Institute on Climate Change and the Environment was established in 2008 at the London School of Economics and Political Science. The Institute brings together international expertise on economics, as well as finance, geography, the environment, international development and political economy to establish a world-leading centre for policy-relevant research, teaching and training in climate change and the environment. It is hosted by the Global School of Sustainability at LSE and is funded by the Grantham Foundation for the Protection of the Environment.

www.lse.ac.uk/granthaminstitute

This paper was first published in January 2023 by the Grantham Research Institute on Climate Change and the Environment at the London School of Economics and Political Science. An interim version was substantially revised in December 2024 and updated further in July 2025 and November 2025. The current version was updated in June 2026.

Suggested citation:

Graff Zivin J and Singer G (2026) Disparities in pollution capitalization rates. Grantham Research Institute on Climate Change and the Environment Working Paper 392. London: London School of Economics and Political Science

Disparities in Pollution Capitalization Rates

Joshua Graff Zivin* Gregor Singer*

UC San Diego & NBER LSE

June 17th, 2026

LINK TO LATEST VERSION

Abstract

We examine how exogenous changes in air pollution exposure over the past two decades have altered disparities in home values between Black and White homeowners. We find that air quality capitalization rates are significantly lower for Black homeowners, so much lower that, despite secular reductions in the Black-White pollution exposure gap, housing value disparities increased. An exploration of mechanisms suggests 60% of this difference is attributable to seller race and 40% to racial neighborhood composition, robust to a range of seller, property and neighborhood characteristics. We discuss our findings through the lens of direct and systemic discrimination in housing markets.

Keywords: house prices, environmental justice, air pollution, race, discrimination

JEL codes: Q51, R30, J15

*Joshua Graff Zivin: University of California, San Diego, CA 92093, E-mail: jgraffzivin@ucsd.edu. Gregor Singer: London School of Economics, London WC2A 2AE, UK, E-mail: g.a.singer@lse.ac.uk. We thank Milena Almagro, Stephie Fried, Akshaya Jha, Matthew Kahn, Andrea La Nauze, Craig McIntosh, Jeffrey Shrader and Anant Sudarshan, and participants at the LSE GRI workshop, UC Riverside, UC Irvine, Toulouse School of Economics, the Gran Sasso Science Institute, the 9th NTHU-UNSW Symposium on Sustainable Finance and Economics, the Mapping the Effects of Environmental Policies workshop at U Hamburg, the Luxembourg Institute of Socio-Economic Research, University of Tokyo, EPIC-China, the Inequality and Climate Change workshop at the University of Amsterdam, University of Sydney, University of Oregon, the AAEA Health Economics seminar and the CESifo workshop for helpful discussion. We thank Zillow for providing data through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author and do not reflect the position of the Zillow Group. G.S. acknowledges support from the Grantham Research Institute on Climate Change and the Environment at the London School of Economics.

I. Introduction

While racial segregation in the United States formally ended more than half a century ago, the existence of predominantly Black and White neighborhoods persists, and they continue to differ on a wide range of dimensions.[1] One important dimension is pollution, where Black communities are disproportionately exposed to poor air quality relative to their White counterparts (Jbaily et al. 2022). Since the harms from air pollution, which include health as well as other human capital impairments (Graff Zivin & Neidell 2012), have been shown to capitalize into housing values[2] (Chay & Greenstone 2005, Bayer et al. 2009, Grainger 2012, Bajari et al. 2012, Currie et al. 2015, Bento et al. 2015, Bayer et al. 2016, Sager & Singer 2025) it may also contribute to the well-documented racial disparities in housing values across and within neighborhoods (Myers 2004, Faber & Ellen 2016, Bayer et al. 2017, Perry et al. 2018, Kermani & Wong 2021, Kahn 2021, Higgins 2023, Diamond & Diamond 2024). In this paper, we examine this relationship directly by examining how changes in exposure to air pollution over the past two decades have altered the disparities in home values between Black and White homeowners, the single biggest factor in household wealth.

We begin by noting that there are good reasons to be optimistic. The Clean Air Act Amendments and other secular trends have led to significant air quality improvements (Colmer et al. 2020), and those improvements were larger in Black communities thereby reducing the Black-White exposure gap (Currie et al. 2023, Sager & Singer 2025). Whether this also reduced disparities in housing values, however, depends not only on relative exposure, but also on whether this amenity capitalizes similarly across homeowners of different race. To analyze this relationship, we combine three types of administrative data, including address-level housing characteristics and transaction information from Zillow, homeowner characteristics including race from HMDA (2022), and neighborhood characteristics from the Census, with fine particulate matter pollution exposure (PM2.5) in the US at a fine level of spatial resolution. Our main sample includes around 9 million transactions where we observe seller race in the contiguous US to estimate the relationship between air quality and house prices by seller race.

Our estimation strategy relies on a hedonic design that controls for observed house and neighborhood characteristics, as well as unobserved time-invariant property characteristics or amenities that differ across communities and race at the most granular administrative level of Census blocks (around 53 individuals on average).[3] We allow for unobserved flexible trends at the state or county

[1] Although the Fair Housing Act of 1968 outlawed explicit discrimination in housing, substantial racial segregation and wealth gaps persist today.

[2] Capitalization could be driven by the buyer or seller side. Note that even if air quality were not a salient feature on the buyer side, capitalization could be driven by the seller side, where air quality may be more easily inferred, e.g. some residents detecting dark smoke plumes or experiencing breathing difficulties and moving out as in Tiebout sorting (Banzhaf & Walsh 2008). It has been shown that air pollution capitalizes in house prices at least since Ridker & Henning (1967), with mounting evidence over the past two decades.

[3] In this context, a related approach to the widely used hedonic design (e.g. Chay & Greenstone 2005, Bajari et al. 2012) are structural equilibrium sorting models (e.g. Bayer et al. 2009, Depro et al. 2015, Bayer et al. 2016). Cassidy et al. (2024)

level and by degree of urbanicity. A main challenge is that changes in pollution are likely to be correlated with changes in other amenities that affect house prices, such as economic activity. We overcome this challenge using a well-established instrumental variable strategy that exploits Clean Air Act rules that led to plausibly exogenous differential changes in air quality across counties (Chay & Greenstone 2005). We follow Sager & Singer (2025) to account for the bias in the first stage arising from confounding trends between treated and control units due to differences in pre-sample pollution, and allow for heterogeneous effects within nonattainment areas based on pre-sample pollution levels (Auffhammer et al. 2009, Bishop et al. 2023).[4]

Because neighborhood characteristics may themselves respond to air quality improvements, we fix neighborhood characteristics, including racial composition, at the beginning of our sample. We show that changes in air quality do not affect sorting into neighborhoods based on race,[5] and show robustness to using a sub-sample that focuses on areas with little change in racial composition during our study period. Transaction prices remain free to adjust and thus incorporate any equilibrium sorting responses. It is important to note that since we estimate capitalization rate outcomes over time in a two-sided market with intermediation, they are not necessarily equivalent to the buyers' marginal willingness-to-pay as defined in standard frictionless cross-sectional hedonic models (Rosen 1974).[6] While preferences of buyers affect sorting, our focus is on sellers and their primary objective is to maximize sales price irrespective of buyer preferences. As in Greenstone & Gallagher (2008), improvements in amenities such as air quality accrue to incumbent property owners (e.g. sellers), irrespective of the subsequent sorting induced by the marginal willingness-to-pay for air quality of buyers.

Our first set of results show that a one-unit decrease in $PM_{2.5}$ increases house prices by 6.3%, a figure consistent with previous estimates (Bento et al. 2015, Sager & Singer 2025), but this average figure masks considerable heterogeneity across racial groups. While the Non-Hispanic White (NHW) pollution capitalization rate is 7.5%, the Black capitalization rate is only 5.3%, a difference of 42% in relative terms, and over 100% in absolute terms since price levels are higher for NHW homeowners.[7] Importantly, the results are almost unchanged when including property fixed effects, when conditioning on property size or seller income fully interacted with air quality, or when interacting air quality with a leave-one-out measure of property quality constructed from repeat-

find no differential sorting along socio-economic dimensions as a response to waste cleanup across the US and similar estimates as with a hedonic approach.

[4]We address potential nonrandom exposure to the policy shock via re-centering instruments following Borusyak & Hull (2023, 2025).

[5]We acknowledge that the ideal test for sorting in response to changes in air quality requires not only data on changes in population shares, but also knowledge of source-destination movement matrices (Depro et al. 2015).

[6]We discuss this further in Section VI..

[7]This implies heterogeneous private capitalization of public improvements (Tsivanidis Forthcoming). Capitalization rates can vary because property characteristics, including amenities and the identity of the seller, are valued as a bundle rather than independently (Rosen 1974), that is, complementarities usually exist among characteristics within the bundle.

sales premia or discounts relative to highly local price indices.[8] This implies that the disparity by race persists irrespective of the stance one takes on whether these characteristics should be viewed as the result of some form of discrimination or that discrimination should be defined after conditioning on them.[9] Despite the larger decrease in $PM_{2.5}$ for Black homeowners (6.6 units) relative to NHW homeowners (4.9 units) from 2000-2019, the much lower Black capitalization rate per unit of cleaner air means that the Black-White housing-value gap actually *increased* as a result of those pollution reductions. To be clear, both groups still experience gains from cleaner air, but at differential rates. Indeed, if Black homeowners had the same capitalization rate as their NHW counterparts, their home values would have been 16% higher by the end of 2019.

Since our results with observed seller race are based on a sample that is restricted to properties or transactions that involved a mortgage or loan, we also show results using seller race predicted by a neural network algorithm based on first and surnames. Using this sample that encompasses over four times the observations, we find that our results are externally valid and, based on an analysis of measurement error due to prediction, conclude that the difference in capitalization rates may be even larger. We also show that our results by seller race are not driven by homophily in transactions. Simultaneously including observed or predicted buyer race does not explain the difference in capitalization rates by seller, and if anything, Black buyers pay a premium on air quality improvements, consistent with the literature focused on buyer race (Bayer et al. 2017, Higgins 2023). This asymmetry between the discount received by Black sellers and the premium paid by Black buyers for air quality improvements reveals a potential double jeopardy in terms of the welfare of Black homeowners.

We then ask whether the capitalization gap is associated primarily with the race of the seller or with the racial composition of the neighborhood in which the home is located. This distinction matters because seller race and neighborhood racial composition are correlated, but they correspond to different potential channels. The seller-race component captures differences in capitalization associated with the identity of the seller but is not a physical attribute of the house or neighborhood. Racial composition of the neighborhood is a characteristic of the place in which the asset is located, and may proxy for amenities, investment histories, information, or perceptions. We decompose the effect into these two components by simultaneously interacting air quality with baseline neighborhood racial composition and seller race. We find that approximately 60% of the difference in capitalization rates is associated with seller race, with the remaining 40% associated with racial neighborhood composition. We show that this split is robust to a wide range of neighborhood

[8] We show that this is in part driven by our identification strategy that exploits shocks to air quality. As one might expect, if we conduct an unconditional analysis of house price disparities, controlling for income or property size affects the measured disparity significantly. Focusing on exposure itself, Colmer et al. (2024) find that equalizing the Black-White income gap would only reduce the pollution exposure gap by 10%.

[9] Focusing on repeat sales, we also find that NHW sellers obtain a relative premium over prevailing local market conditions compared to Black sellers selling the exact same property, a result which holds across the distribution of property quality.

characteristics interacted with air quality, including neighborhood income, house prices and rents, measures of greenness, crime, or supply elasticities, as well as restricting our samples to areas that had little change in racial composition over time.

To help contextualize these findings, we discuss the extent to which we can interpret these components through the lens of direct and systemic discrimination (Bohren, Hull & Imas 2025), while emphasizing that the estimates are conditional capitalization rates rather than structural measures of marginal willingness to pay. Under this interpretation, the seller-race component may capture forms of direct discrimination after holding neighborhood racial composition or other house or seller characteristics fixed. The neighborhood-composition component, conditional on other amenities, may capture forms of systemic discrimination including historical constraints on residential sorting, investment, and perceptions that often persist today (Aaronson et al. 2021, Sood & Ehrman-Solberg 2026).[10] Both components could operate through buyers, real estate agents, appraisers, lenders, bargaining, or search and steering frictions in an intermediated housing market (Christensen & Timmins 2022, 2023).[11] We explore the role of direct and systemic discrimination in driving our results, limitations in this attribution, as well as alternative explanations for the capitalization rate gap. Two additional pieces of evidence are consistent with discrimination interpretations. First, both the seller-race and neighborhood-composition components are driven by areas with high measures of housing-market discrimination based on experimental data from paired testers. Second, if our results are indeed driven by discrimination, we would expect areas with more racial residential segregation to show larger disparities in capitalization rates. Consistent with this, the disparities are much larger in more segregated areas. Taken together, the evidence suggests that the lower capitalization of clean air for Black homeowners reflects both seller-race transaction channels and racialized place-based channels.

This paper contributes to the growing literature on environmental justice (Banzhaf et al. 2019), and connects the literature on housing prices and racial groups (Aaronson et al. 2021, Kahn 2021, Kermani & Wong 2021, Akbar et al. 2022, Higgins 2023, Diamond & Diamond 2024) with that on housing prices and pollution (Chay & Greenstone 2005, Bayer et al. 2009, Bajari et al. 2012, Grainger 2012, Currie et al. 2015, Bento et al. 2015, Bayer et al. 2016, Sager & Singer 2025). Our findings that the pollution capitalization rate differs by race provides novel insights into how the marginal effects of pollution exposure differ across the population, which is critical for understanding the distributional effects of air quality policies (Hsiang et al. 2019). Furthermore, our analysis of mechanisms is, to our knowledge, the first to unpack the relative roles of seller-race and neighborhood-racial-

[10] The neighborhood sorting literature often distinguishes preferences about racial composition from discriminatory constraints on residential choice. In our setting, however, preferences over racial composition are not merely a sorting primitive: if they affect the prices received by homeowners conditional on other amenities, they constitute a racialized place-based channel through which capitalization differs across communities.

[11] As an illustrative recent example, a Black couple filed a lawsuit after they received substantially higher valuations from an appraiser when a White colleague posed as the homeowner (US District Court 2022).

composition components. Our findings are consistent with recent evidence on discriminatory pathways, such as racial steering in the housing market (Christensen & Timmins 2022, 2023, Christensen et al. 2022), racial disparities in mortgage lending and refinancing practices (Munnell et al. 1996, Charles & Hurst 2002, Ambrose et al. 2021, Bhutta et al. 2022), and lower offers for Black sellers in other marketplaces (List 2004, Doleac & Stein 2013, Barnes & Stein 2024).

The paper proceeds with a simple conceptual framework in Section II. We describe our data and provide some motivating empirical facts in Section III. We discuss our empirical strategy in Section IV and show our main results in Section V. In Section VI, we discuss and interpret our results before we conclude in Section VII.

II. A Simple Conceptual Framework

Before turning to the data and empirical specifications, we outline the conceptual framework that links air-quality improvements to racial disparities in housing wealth appreciation. The key object is a clean-air capitalization slope: the change in the sale price received by an incumbent owner when local air quality improves. This slope can differ across seller groups because Black and NHW sellers may sell different houses, live in neighborhoods where clean air capitalizes differently, or face different transaction processes conditional on the house and neighborhood. We use the framework to define the seller-group capitalization gap and then decompose that gap into a seller-race component and a neighborhood racial composition component.

Define $q_{bt} \equiv -\mathrm{PM}_{bt}$ as clean air in Census block b and year t, so that a one-unit increase in q_{bt} corresponds to a one-unit decrease in $\mathrm{PM}_{2.5}$. For a transaction i, the clean-air capitalization slope is

$$\kappa_i \equiv \frac{\partial \log(P_i)}{\partial q_{bt}}. \tag{1}$$

This slope is the percentage change in the transaction price received by the seller for a one-unit improvement in local air quality. For racial group g, the housing wealth incidence of air quality improvements can be approximated by:

$$\Delta \log(P_g) \approx \kappa_g \Delta q_g. \tag{2}$$

This expression highlights the two margins that determine whether air quality improvements reduce or increase racial gaps in housing wealth: which group experiences larger pollution reductions (Δq_g), and how strongly those reductions are capitalized into sale prices (κ_g).

It is important to note that κ_i is not a structural marginal willingness-to-pay parameter. A transaction price is an equilibrium outcome in a two-sided housing market with search, bargaining, financing, and intermediation. Our empirical estimates recover the combined pass-through of demand and supply into prices and are therefore capitalization rates. A simple market-clearing expression illustrates why house and neighborhood characteristics can affect this slope. Consider a local market m for a given house-neighborhood bundle, with demand and supply satisfying

$$D_m(P, q; X, A, N, L) = S_m(P, q; X, A, N, L), \tag{3}$$

where X denotes property attributes and other transaction-level characteristics, including seller characteristics that can affect reservation values, A nonracial neighborhood amenities, N baseline neighborhood racial composition, and L broader market conditions. Differentiating this condition shows that the log-price capitalization slope depends on how clean air shifts demand and supply, scaled by the price responsiveness of the two sides of the market:[12]

$$\frac{d\log(P_m)}{dq} = \frac{1}{P_m}\frac{D_{mq} - S_{mq}}{S_{mP} - D_{mP}}. \tag{4}$$

This is not a structural model we estimate, but the expression is useful for interpretation. It shows why capitalization slopes may differ across houses and places. On the demand side, clean air may shift demand more in markets where marginal buyers have greater willingness or ability to pay for lower exposure. These differences may reflect income and credit constraints, information about pollution, health valuation, expected exposure, expected tenure, perceived neighborhood trajectory, or complementary amenities, i.e. the extent to which cleaner air is more valuable when bundled with other amenities such as green space, safety, schools, or economic opportunity. Property characteristics can also affect capitalization if clean air is more valuable for larger or higher-quality houses, or if those characteristics proxy for different buyer pools. On the supply side, the same demand shift may translate into different price responses where housing supply is more or less elastic, or where seller reservation values, proxied for example by income, differ. Thus, heterogeneous capitalization by seller group can arise from property, seller, or neighborhood differences that are correlated with seller race.

Seller race is conceptually different from asset and place characteristics. In a frictionless world, the race of the seller should not affect the price response of the same house in the same neighborhood to the same clean-air improvement. We write this distinction as:

$$\log(P_i) = h(q_{bt}, X_i, A_b, N_b, L_{ct}) + \omega(q_{bt}, R_i \mid X_i, A_b, N_b, L_{ct}) + \varepsilon_i, \tag{5}$$

where $h(\cdot)$ is the value of the asset and its place for the reference seller group, and $\omega(\cdot)$ is a transaction wedge that may vary with seller race R_i, conditional on those same characteristics. We decompose the seller-group capitalization gap into seller race and neighborhood racial composition,

[12] Totally differentiating the market-clearing condition gives $D_{mP}\frac{dP_m}{dq} + D_{mq} = S_{mP}\frac{dP_m}{dq} + S_{mq}$,, which implies $\frac{dP_m}{dq} = \frac{D_{mq}-S_{mq}}{S_{mP}-D_{mP}}$. Dividing by P_m gives the corresponding log-price expression: $\frac{d\log(P_m)}{dq} = \frac{1}{P_m}\frac{D_{mq}-S_{mq}}{S_{mP}-D_{mP}}$.

conditional on controls and fixed effects:

$$\kappa_i = \kappa_0 + \boldsymbol{\kappa}_R' R_i + \boldsymbol{\kappa}_N' N_b + \boldsymbol{\zeta}' Z_i. \tag{6}$$

This is an accounting device for the capitalization slope, not a structural decomposition of the primitives in Equation 4. Both components $\boldsymbol{\kappa}_R$ and $\boldsymbol{\kappa}_N$ can reflect search frictions, steering, appraisal practices, listing exposure, bargaining, or beliefs about the seller or neighborhood. We allow controls collected in Z_i to interact with clean air, so they can also affect capitalization rates. The distinction between both components matters. Block composition absorbs all other unobserved place-level differences that may bias estimation of $\boldsymbol{\kappa}_R$ if racial composition of sellers mirrors local block composition.[13] Furthermore, an overall seller-race gap in capitalization can arise even if $\boldsymbol{\kappa}_R = 0$. If Black and NHW sellers tend to sell in neighborhoods with different baseline racial composition, and capitalization rates vary with N_b, then the average capitalization rate will differ across seller groups even without a seller-race-specific wedge. Our decomposition below separates this neighborhood-composition channel from the component associated with seller race itself.

An ideal experiment would vary air quality across otherwise identical transactions, holding fixed the property, nonracial seller characteristics, nonracial neighborhood amenities, and market conditions, while varying only seller race or neighborhood racial composition. Since such experimental variation is not feasible, our empirical approach uses plausibly exogenous air-quality improvements to estimate capitalization slopes and then asks how the seller-group gap in those slopes changes as we sequentially hold fixed property, seller, and neighborhood characteristics, allowing each to affect the slope with respect to clean air.

Empirically, we proceed in four steps. Before we introduce neighborhood racial composition or amenities, we first estimate overall seller-race specific capitalization rates (only including $\boldsymbol{\kappa}_R' R_i$), noting that the seller-race coefficient may capture neighborhood racial composition or other correlated factors, even in our repeat sales designs. We then ask whether this gap changes when we add factors tied to the seller or property itself, including seller income, square footage, or a measure of unobserved property quality, all interacted with air quality improvements. We next introduce baseline neighborhood racial composition that acts as a summary place characteristic that captures the fact that Black and NHW sellers tend to transact in different neighborhoods. This step separates the broad seller-group gap into a seller-race component and a place-composition component. Finally, we interact clean air with additional neighborhood and market characteristics, including income, house values, rents, greenness, crime, opportunity, and housing supply elasticity, to ask whether

[13] If, within each block, the racial composition of sellers mirrors the racial composition of residents, the residual seller-race variation after conditioning on baseline block composition is orthogonal to omitted block-level amenities. Let R_i denote a Black seller, N_b the baseline Black population share, and A_b any block-level amenity omitted from Z_i. If $E[R_i \mid b] = N_b$, then $R_i = N_b + \eta_i$, where $E[\eta_i \mid b] = 0$. Since A_b is fixed within block, residual seller variation is uncorrelated with unobserved amenities: $E[\eta_i A_b] = E[A_b E[\eta_i \mid b]] = 0$.

these variables account for either component that allow us to interpret the decomposition and link it to the housing discrimination literature.[14]

We focus on homeowners who sell during the sample period, because transaction prices measure the capitalization of cleaner air into the value received by incumbent owners. The incidence for renters is conceptually different: lower capitalization into rents may benefit renters even when it lowers the asset value received by owners. We therefore treat rental capitalization separately in Appendix Table A.12. The framework can in principle be generalized beyond $PM_{2.5}$ to any local amenity shock that can affect racial housing wealth gaps through both exposure and capitalization. We focus on air quality because the Clean Air Act provides a large, observed, and plausibly exogenous source of variation in a local amenity. We show, however, some non-parametric within-property evidence on seller-race price differences in Figure 3 that does not rely on any specific amenity shock as the source of variation, and which can speak to whether this pattern may be generalizable.

III. Data and Descriptives

A. Transaction-Level House Price Data

We use two databases from Zillow (2020) that allow us to obtain house prices and basic hedonic characteristics at the transaction level from 2000-2019 for the contiguous US. The first database are transactions (ZTransaction) sourced from county recorder's offices with information including transaction price (deflated to 2012 US$), type of deed and date of sale. The second database contains hedonic information (ZAssessment), sourced from county assessor's offices, including square footage (SQFT) and geolocation.[15] For counties that report details such as transaction price, this should capture the universe of transactions, but not all counties report prices, e.g. few do in Texas (see spatial coverage below).[16] Importantly, we only use arm's length transactions and residential properties, dropping transactions such as refinancing or foreclosures, and use historic assessment data to reduce missing values of hedonic information (for details on data cleaning see Appendix A.4). Since we use log prices and state-by-year or county-by-year fixed effects for our analysis, our data are effectively deflated with state or county deflators. We map the geolocation of each transacted property to US Census blocks using the 2010 US Census boundaries.

[14] A fully saturated specification with all seller, property, neighborhood, and market interactions would have limited power because each additional $PM_{2.5}$ interaction requires corresponding interactions of the nonattainment instruments. We therefore use a sequence of fully instrumented specifications to preserve first-stage strength and show how each set of interactions changes the decomposition.

[15] We only use property location and size, as other hedonic information is often missing.

[16] Much of our analysis relies on within county variation.

B. Observed and Predicted Race and Income at Transaction Level

The Zillow data contains no information on buyer or seller race. We use two separate approaches to obtain observed and predicted race respectively for each transaction.

First, for observed race we use administrative data from the Home Mortgage Disclosure Act (HMDA 2022) that contains information on the universe of mortgage applications from all banks, savings associations, and credit unions with assets above a threshold ($39 million in 2010).[17] Importantly, HMDA data contains information on race/ethnicity and income of mortgage applicants. We use three racial groups, Black and Non-Hispanic White ("NHW") Americans (as in Currie et al. (2023)), and a third group for Other Americans ("Other"). We match Zillow with HMDA data based on common variables including Census tract, year of transaction, loan amount rounded to the nearest thousand and name of lending institution.[18] For joint sellers or buyers, we assign the race as Black or NHW if at least one of the joint sellers/buyers is Black or NHW, and the other is not of the other race (i.e. we drop mixed race Black-NHW joint applicants, but one applicant is allowed to be of a different ethnicity).

We are primarily interested in seller race. To obtain seller race for a particular transaction we actually don't require a HMDA entry for the corresponding transaction. Instead, we need to match the seller of this specific transaction to a previous transaction on the same property where that seller was a buyer or borrower and thus recorded in HMDA. We accomplish this using the first and surnames of sellers and previous buyers/borrowers of the same property and fuzzy string matching where at least two of three state-of-the-art algorithms agree.[19] We dramatically improve our success rate of finding a previous transaction from a specific seller by also including all previous non-sales transactions for this step, such as refinancing, home improvements, or HELOCs, as these transactions are contained in both Zillow and HMDA data.[20] With the identified previous transaction of a specific seller, we can match these previous transactions to HMDA to obtain seller race. For obtaining buyer race, we can match the transaction to HMDA and use recorded buyer race directly, provided that the transaction involved a qualified mortgage.

We further restrict our sample by requiring a minimum match quality on lender name between Zillow and HMDA. Our baseline version uses a cutoff of 60% bi-gram match resulting in 9.2 million

[17] In 2017, 92% of originated mortgage loans nationwide were covered in HMDA (Consumer Financial Protection Bureau 2018).

[18] In the rare occasion that there are multiple matches between Zillow and HMDA, we use the lender name with the best fuzzy string match within each Census tract, year, loan amount correspondence.

[19] We use the Jaro-Winkler, the Jaccard and the Damerau-Levenshtein distances from van der Loo (2014) for fuzzy string matching. For joint buyers or sellers we allow for swaps in who is listed as primary or secondary buyer/seller.

[20] This helps as some buyers may have originally bought a home before our sample begins. We use all transactions from 1992 to identify sellers that are previous buyers/borrowers. In case we have multiple matched previous transactions per seller, we use an iterative procedure that prioritizes those matches for which we have non-missing reported race and income on previous transactions and are closer to the date of the eventual sales transaction. We deflate recorded seller income to 2012.

observations with observed seller race, and we show robustness to stricter cutoffs at 75% and 90%, which reduces observations to 8.1 and 5.2 million, respectively.[21] The number of observations where we observe both seller and buyer race is 2.6 million.[22]

Second, as an alternative to observed race, we use predicted buyer and seller race. One shortcoming of using observed race from HMDA data is that we can only include transactions on properties that involved a mortgage from a reporting lender (for sellers: previously involved financing on same property). Therefore we also construct predicted race of buyers and sellers for each transaction based on full names of buyers and sellers. In particular, we use a neural network algorithm trained by Xie (2022) who uses Florida voter registration data with a focus on minority groups. The algorithm calculates probabilities for belonging to different racial groups for each first and surname pair of each buyer and seller. We classify a buyer or seller as belonging to a racial group when the probability of belonging to a single racial group is at least 70% (and show robustness to different thresholds), allowing us to identify predicted seller race for 40.1 million transactions, and both predicted seller and buyer race for 23.7 million transactions. Overall the accuracy of the prediction (share of correct predictions in observations) is 78%, with 94%, 79%, 84% accuracy for Black, NHW, and Other respectively, measured in the sample where we have both predicted and observed race. The accuracy conditional on observed race (true positives divided by all positives) is 75%, 74%, 96% respectively. To provide additional reassurance about the quality of our prediction, we test how well the prediction performs by comparing our main estimates based on predicted race for the observations where we also observe race, before showing estimates using the larger sample with predicted race only.

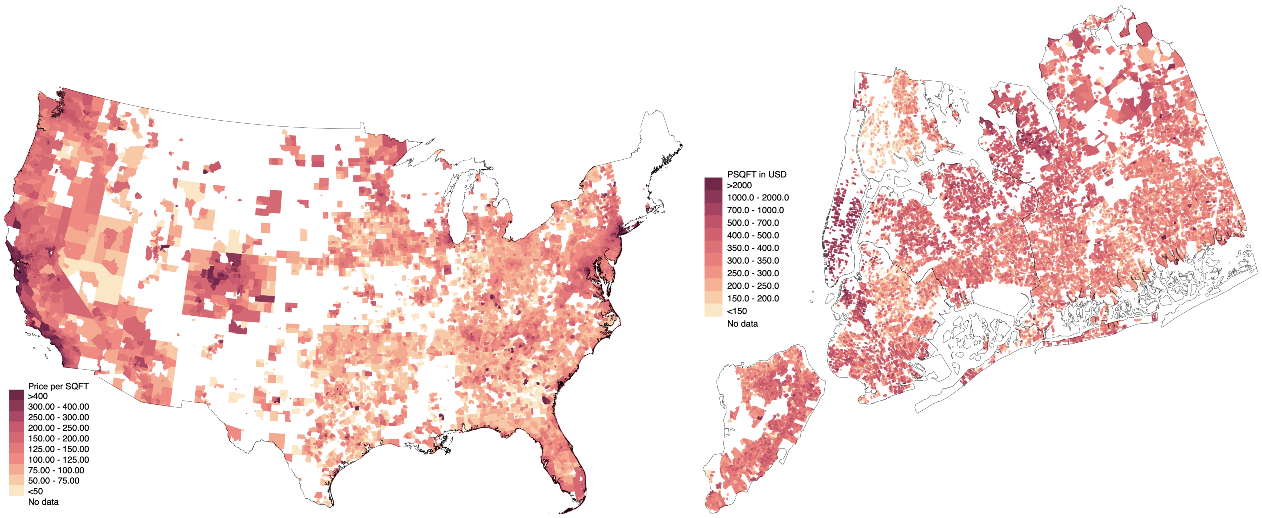
The advantage of using observed seller race is that we have no measurement error from prediction, and the advantage of using predicted race is that we can include transactions which are not linked to a mortgage resulting in a much larger sample while addressing concerns of possible selection into the observed race sample.

C. Pollution Data and Clean Air Act Nonattainment Areas

We use annual data on fine particulate matter concentrations ($PM_{2.5}$) at the 1km-by-1km resolution from van Donkelaar et al. (2021), which is constructed by combining ground-based measurements, satellite images and chemical transport models. We map the $PM_{2.5}$ data into Census blocks using

[21] Note that observations reported in regression tables may be slightly lower as fully partialed out observations through fixed effects are not counted.

[22] The number of observations where we have observed buyer race irrespective of observing seller race is 12.3 million. As a point of comparison: Of the 64.7 million transactions in the Zillow data in the contiguous US with non-missing transaction price and square footage, 48% involve a mortgage with information on loan amount. Of these transactions we match 49% to the HMDA data based on Census tract, year and loan amount to obtain buyer race (i.e. for 24% of all observations). This is a similar ratio as for San Francisco (Bayer et al. 2016) or Florida alone (Graff Zivin, Liao & Panassie 2023). Applying our match quality thresholds regarding lender name and thresholds for defining race including no mixed race applicants reduces this to 19% of total observations for buyer race.



(a) Price/SQFT aggregated up to Census tracts: USA (b) Price/SQFT aggregated up to Census blocks: NYC

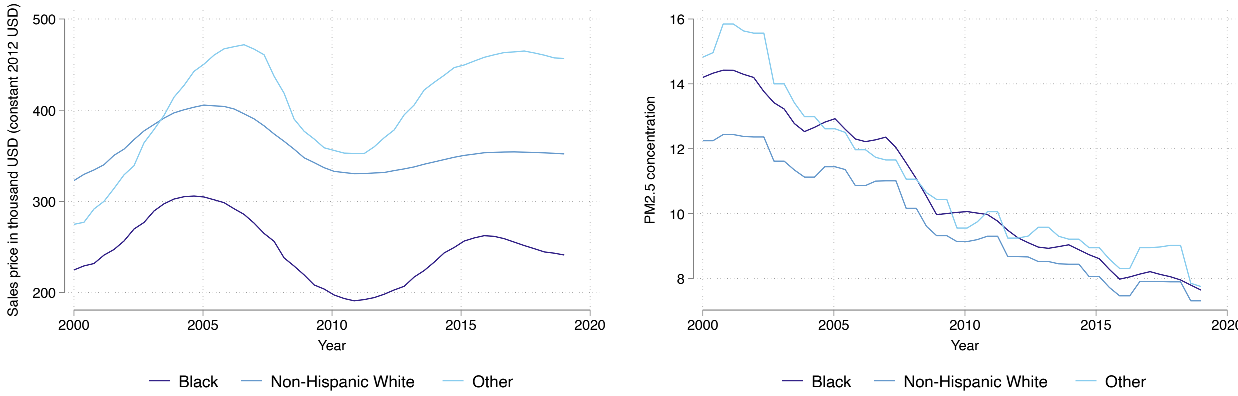
Figure 1: Spatial distribution of price per SQFT

Notes: Panel (a) and (b) show the spatial distribution of price per SQFT in our sample pooled across time. Panel (a) aggregates up to Census tract averages. Panel (b) shows the variation around New York City aggregated to the more granular Census block level. All monetary values are deflated to 2012 US$.

the closest pollution grid point to the Census block centroid. To identify the effect of pollution on house prices, we make use of the 2005 Clean Air Act rules for $PM_{2.5}$, following Sager & Singer (2025) who provide a detailed account of this regulation. We use the 208 counties from the EPA (2005) that became regulated in 2005 as the treated group, because they did not meet the necessary threshold of 15 $\mu g/m^3$ for the three-year average of annual mean $PM_{2.5}$ concentrations. These counties were assigned into nonattainment, and were subject to stricter action to reach air pollution standards from the Environmental Protection Agency. It is worth noting that $PM_{2.5}$ is measured with error. Our results are robust to using alternative $PM_{2.5}$ datasets. Sager & Singer (2025) show that using information on uncertainty in measurement in this data by excluding areas with larger uncertainty, or alternatively using ground-monitor data only, yields similar estimated effects of the 2005 Clean Air Act designations.

D. Neighborhood Racial Composition and Additional Place-Based Data

We combine our transaction information with data from the 2000 Census on population counts by race (Manson et al. 2022). We calculate the share of Black, NHW and Other at the Census block level in 2000 and the share of that block that is urban/rural. Census blocks are the most granular administrative unit, and in 2000 there were 5.3 million non-empty Census blocks in the contiguous US with an average population of 53 individuals and 77% of housing units owner-occupied. We use the first year of the sample for neighborhood characteristics throughout to exclude variation coming from spatial sorting during our sample.



(a) Sales price by racial group

(b) PM2.5 concentration by racial group

Figure 2: Evolution of prices and PM2.5 by seller racial groups

Notes: Panel (a) shows the evolution of average sales price (real 2012 US$) by racial group. Panel (b) shows the evolution of average PM2.5 concentrations of home sellers by racial group. All monetary values are deflated to 2012 US$.

We use additional information on characteristics of neighborhoods. At the Census block level this includes the Normalized Difference Vegetation Index (NDVI) in 2000 derived from satellite imagery at the 1km-by-1km resolution based on Didan (2021), which provide a proxy measure for local green spaces like parks. At the Census block group level (1315 individuals per block group on average) this includes the share of the population in poverty, median household income, median personal income, median rent and median house values from the 2000 Census, all deflated to 2012 US$. At the Census tract level (4319 individuals per tract on average), we use the proportion of land that is developed with high intensity ("BuiltHigh") from NHGIS (Manson et al. 2022), that is with at least 80% share of impervious surfaces (e.g. asphalt and concrete), derived from the National Land Cover Database in 2001. Based on Chetty et al. (2018), we construct economic opportunities at the tract level using the average percentile in the 2014-2015 income distribution for children born between 1978-1983. We use housing supply elasticities at the tract level from Baum-Snow & Han (2024). We calculate an index of racial residential segregation within tracts based on the 2000 Census that measures how uniformly residents of different races are mixing across Census blocks within Census tracts (e.g. high index if racial groups are living in separate blocks), following Reardon & Firebaugh (2002). At the county level (89,927 individuals per county), we measure arrest rates from the FBI (2006) in 2000 (results are similar with crime rates).

E. *Descriptive Statistics on Racial Housing and Pollution Disparities*

Figure 1a provides an overview of the spatial coverage of our data where we observe seller race, showing price-per-sqft (PSQFT) aggregated to the Census tract averages for better visualization across the entire US. This masks a large degree of spatial granularity within Census tracts. Figure

Table 1: Summary statistics in sample: averages of selected variables by racial groups

Group	PM2.5	Price (th.)	SQFT	HH income	Med income	Poverty	Spop B	Spop NHW	Spop OTH	Spop URB
Black	10.10	251.2	1798	98.6	48.06	0.12	0.35	0.47	0.18	0.92
NHW	9.07	357.0	1918	122.5	56.16	0.07	0.05	0.81	0.14	0.85
Other	10.95	418.4	1781	116.9	56.12	0.10	0.07	0.53	0.40	0.94
Total	9.28	357.9	1900	121.0	55.81	0.08	0.06	0.77	0.17	0.86

Notes: The table shows averages of the indicated variables by racial group based on the sample used in estimation. PM2.5 is in $\mu g/m^3$, "Price" is in thousand US$, "HH income" stands for seller income at the transaction level in thousand US$, "Med income" stands for median income of the Census block group in 2000 in thousand US$, and all monetary values deflated to 2012 US$. "Poverty" stands for poverty rate at the Census block group in 2000, and Spop B, Spop NHW, Spop OTH, and Spop URB stand for share of population in Census block in 2000 that is Black, NHW, Other, and urban. Appendix Table A.1 provides further summary statistics.

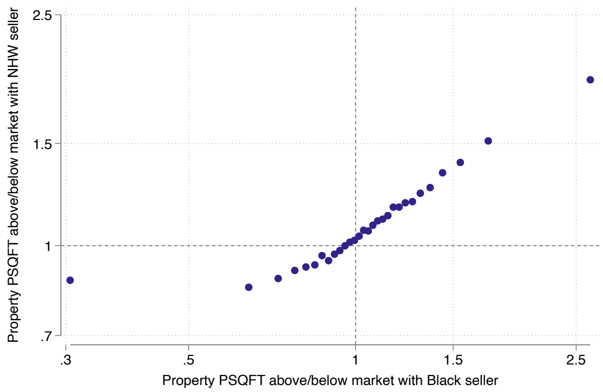
1b show the variation in a few counties around New York City, aggregated to the Census block level instead. This illustrates that, even across a few city blocks, house prices can vary substantially due to differences in amenities among other things.

Figure 2a shows the house price gaps and how real house prices evolved differently by observed seller race. The housing crisis hit Black and Other sellers particularly hard, consistent with Faber & Ellen (2016). While prices recovered for Other relative to NHW homeowners, the recovery for Black homeowners was more modest, resulting in a persistent gap of house prices between Black and NHW homeowners.

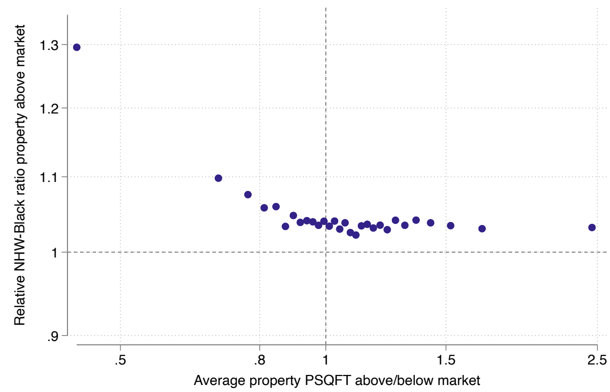
Figure 2b shows falling PM2.5 concentrations by observed seller race. Black homeowners faced higher pollution levels on average, but the gap between Black and NHW homeowners narrowed over time. This is consistent with some recent empirical work focused on environmental justice (Jbaily et al. 2022, Currie et al. 2023), and shows that this relationship also holds for the subset of the general population who own a home that is sold during our study period.

Table 1 shows summary statistics by observed seller race. Black homeowners sell slightly smaller houses and have a lower income. They also live in neighborhoods (block groups) that are lower income, have a higher poverty rate, and are characterized by a significantly larger share of Black population relative to NHW. Appendix Table A.1 shows more detailed summary statistics for all variables used in the analysis.

As a final descriptive exercise, Figure 3 compares sales prices by Black and NHW sellers of the exact same property at different points in time. We first generate a highly local price index for PSQFT using locally weighted regressions as in Ahlfeldt et al. (2023). We then calculate how much a property is above or below current local market PSQFT (the "delta"), which nets out local house price cycles. Using repeat sales of the same properties where we have both a Black and NHW seller at some point, we calculate the average delta and compare Black and NHW sellers only within properties. Figure 3a shows that for the same property, the "above market" deltas for Black and NHW sellers are highly correlated. Above market properties tend to be above market in other points in time as well driven by inherent unobserved quality of the property. However, as Figure 3b shows on the vertical axis, the "above market" deltas are consistently higher for NHW sellers than Black



(a) Premium for NHW and Black for same property



(b) NHW-Black premium ratio by property premium

Figure 3: Within-property racial difference in PSQFT premia above local market average

Notes: The figure plots binscatters based on repeat-sale properties that were each sold at least once by a Black seller and once by a NHW seller. We compare transaction prices to highly local price indices for PSQFT using locally weighted regressions as in Ahlfeldt et al. (2023). Panel (a) plots price-per-square-foot (PSQFT) premium over the local market when a NHW owner vs when a Black owner sells the same property. Panel (b) converts those two premiums into a within-property ratio (NHW-Black) and plots it against the property's mean PSQFT premium, which proxies for overall property quality.

sellers within the exact same property, netting out local price cycles that adjust for selling at different times. NHW sellers receive an around 5% higher premium relative to local market conditions compared to Black sellers on the same property, and the figure shows that this is consistent along the entire distribution of property quality on the horizontal axis. Table A.6 shows an average NHW seller premium of 5% in a regression framework.[23] These striking patterns motivate the main analysis below, where we introduce a plausibly exogenous amenity shock via air-quality improvements to formally analyze capitalization differences by seller race.

IV. Empirical Strategy

To formally explore how pollution reductions affect home sales prices P_i across racial groups, we run regressions at the transaction level i in year t, Census block b, county c, state s, and racial group j:

$$\begin{aligned}\log(\mathrm{P}_i) =& \alpha \mathrm{PM}_{bt} + \sum_j \left(\beta_j S_i^j \mathrm{PM}_{bt}\right) + \sum_j \left(\gamma_j Spop_b^j \mathrm{PM}_{bt}\right) \\ &+ \boldsymbol{\delta}_1 \boldsymbol{X}_i + \boldsymbol{\delta}_2 \boldsymbol{X}_i \mathrm{PM}_{bt} + \delta_3 \boldsymbol{W}_b \mathrm{PM}_{bt} + \boldsymbol{\xi}_{jb} + \boldsymbol{\lambda}_{(s \text{ or } c)t} + \sum_t \left(\tau_t Urb_b t\right) + \varepsilon_i\end{aligned} \quad (7)$$

[23] These results are consistent with Drukker & Ma (2025) who also find lower housing returns for Black sellers, after ruling out other factors that could drive racial differences, including home renovations, seller liquidity pressure, and agent quality.

where PM_{bt} denotes pollution concentrations of PM2.5 in $\mu g/m^3$, S_i^j race of seller in a transaction, and $Spop_b^j$ the racial composition of the block (shares). We later also enrich this specification to include buyer race. $\boldsymbol{X}_i$ is a vector of seller or property characteristics (including property fixed effects for part of the analysis) and $\boldsymbol{W}_b$ a vector of neighborhood characteristics, interacted with PM_{bt}, which we discuss later.[24]

We use seller-race-by-block fixed effects $\boldsymbol{\xi}_{jb}$ that net out time-invariant differences in average house prices by seller race. We allow these absorbed differences in levels to differ by block, as different neighborhoods could have differential average house quality across racial groups, which may otherwise introduce spurious correlation with air quality averages. That is, we only rely on variation in the housing *returns* of air quality in our analysis, using our shock to air quality to identify differences in pollution capitalization rates. These fixed effects also capture all other time-invariant amenities and other block characteristics.

State-by-year or county-by-year fixed effects $\boldsymbol{\lambda}_{(s \text{ or } c)t}$ allow for flexible confounding trends differentiated by state or county, including unobserved changes in local labor markets. Bayer et al. (2009) argue that the larger the unobserved moving costs, the more attenuated the estimated hedonic willingness-to-pay for air quality. Our county-by-year fixed effects also help to reduce the impacts of moving cost differences as moving costs within county are lower than outside of the county (or state). Note that differences in moving costs across race that could affect our coefficients of interest are likely negligible compared to the level of moving costs that would primarily affect α.[25] Finally, heterogeneous slopes of the urban share of block by year $\sum_t (\tau_t Urb_b t)$ allow for confounding trends based on the urbanicity of neighborhoods.

We fix all of our neighborhood characteristics, including racial composition of the block, at the start of our sample period in 2000 to isolate the source of temporal variation that comes from pollution. Since neighborhood characteristics can themselves respond to policy-induced changes in air quality, conditioning on time-varying neighborhood covariates could introduce post-treatment bias. Transaction prices are free to adjust and thus incorporate any equilibrium sorting responses. As in Greenstone & Gallagher (2008), benefits from improvements in amenities affect incumbent property owners, irrespective of subsequent sorting by marginal willingness-to-pay for air quality.[26] Note that sorting in response to air quality changes is less problematic due to our instrument for air quality described below. Importantly, Table A.2 shows that changes in racial block composition are essentially uncorrelated with changes in pollution over time (relative or absolute), with a ten

[24] Note that uninteracted $\boldsymbol{W}_b$ would be absorbed by the $\boldsymbol{\xi}_{jb}$ fixed effect.

[25] If moving costs are higher for Black sellers, this would reduce measured capitalization rates for them. It is plausible, however, that seller income differences can proxy for differential moving costs, as moving costs are characterized by lumpy up-front liquidity or financing needs that can introduce frictions that bind more tightly with lower income. When we include seller income fully interacted with our racial capitalization rate the gap remains robust indicating that moving cost differences by race are likely not a major concern.

[26] Individuals with a higher marginal willingness to pay for air quality may want to move to areas with relatively improved air quality, but also drive up house prices as a result.

percent increase in pollution increasing the share of the Black population by 0.001. This echoes findings in the literature that finds racial composition of neighborhoods to be quite persistent (Bayer et al. 2016). Finally, to test the robustness of our approach, we use the 2020 Census and calculate the change in population shares for each racial group at the block (or tract) level, focusing only on those blocks (or tracts) where racial composition changed little.

A. Instrumenting Air Pollution with Regulatory Nonattainment

Despite the use of fixed effects, changes in air pollution are likely to be correlated with changes in unobserved amenities that also impact house prices. For example, changes in economic activity or infrastructure are likely to drive both, pollution and house prices. We therefore use the 2005 $PM_{2.5}$ Clean Air Act regulation that induced changes in pollution in nonattainment counties as an instrument. The identifying assumption is that, conditional on our fixed effects and controls, the regulation only shifted pollution, and no other unobservables correlated with house prices. We examine potential threats to this assumption in the next section. We follow Sager & Singer (2025) to address bias from underlying trends that differ by baseline pollution and attainment status by including pre-period pollution levels from 1998-99 (PMpre_b) interacted with year dummies as part of our controls $\boldsymbol{X}_i$.[27] This also addresses concerns that areas that are initially more polluted may be on different house price trends or that more polluted blocks have differential trends in block-level economic activity.

As Auffhammer et al. (2009) show, nonattainment effects are often stronger in those parts of nonattainment areas that are closer to monitors that have higher initial pollution readings. To allow for such heterogeneous effects, we additionally interact nonattainment status with pre-period pollution concentrations PMpre_b (see also Bishop et al. (2023)).[28] We include instruments for each term that contains PM_{bt} in Equation 7. The first stage, for example, for PM_{bt} itself, is:

$$\begin{aligned}\text{PM}_{bt} =& \theta_0 \text{NA}_{bt} + \theta_1 \text{NA}_{bt}\text{PMpre}_b \\ &+ \sum_j \left(\eta_j S_i^j \text{NA}_{bt} + \rho_j S_i^j \text{NA}_{bt}\text{PMpre}_b \right) + \sum_j \left(\phi_j Spop_b^j \text{NA}_{bt} + \kappa_j Spop_b^j \text{NA}_{bt}\text{PMpre}_b \right) \\ &+ \boldsymbol{\omega}_1 \boldsymbol{X}_i + \boldsymbol{\omega}_2 \boldsymbol{X}_i \text{NA}_{bt} + \boldsymbol{\omega}_3 \boldsymbol{X}_i \text{NA}_{bt}\text{PMpre}_b + \boldsymbol{\omega}_4 \boldsymbol{W}_b \text{NA}_{bt} + \boldsymbol{\omega}_5 \boldsymbol{W}_b \text{NA}_{bt}\text{PMpre}_b \\ &+ \boldsymbol{\psi}_{jb} + \varsigma_{(s \text{ or } c)t} + \sum_t (\sigma_t Urb_b t) + \mu_i\end{aligned} \tag{8}$$

Our set of instruments vary at the block level, but we allow for spatial correlation by clustering

[27] Note that this also helps to address concerns over shifting hedonic price functions that can conflate willingness-to-pay with such shifts in price functions (Kuminoff & Pope 2014), as this effectively controls for the changing value of baseline air quality, similar to Banzhaf (2021).

[28] Note that Auffhammer et al. (2009) find that air quality increases more in those areas closer to the worst offending monitors within treated areas. Our strategy generalizes this logic, i.e. higher induced clean up rates in initially more polluted areas within treated areas. This is also in line with findings in Sager & Singer (2025).

standard errors at the tract level, which contains an average of 151 blocks.[29] Appendix Table A.4 shows the first stages for the three endogenous variables in our initial specification using only seller race without neighborhood composition. Reassuringly, the exogenous interactions between nonattainment and racial groups affect the corresponding endogenous interactions between change in $\text{PM}_{2.5}$ and racial groups. The Kleibergen-Paap F statistic is high (≥ 100) for the vast majority of our results and reported in our main tables.

Since our instruments also contain interactions of nonattainment with $PMpre_b$ and seller race or block composition, they can be interpreted as formula based IV.[30] Following Borusyak & Hull (2023) and Borusyak & Hull (2025), we additionally re-center instruments to purge any correlation from non-random exposure to the policy shock in robustness checks. Specifically, we assign counties into 10 bins based on the tract with the highest average $\text{PM}_{2.5}$ in the years 2001-2003 within each county, following the logic of EPA assignment rules based on readings of the worst offending monitors during these years. For each of the 10 bins we calculate the average probability of a county being assigned into nonattainment. We then subtract this average probability from the actual nonattainment assignment for all years after 2005 to re-center the IVs. For example, we use $(\text{NA}_{bt} - \overline{\text{NA}_{bt}})S_i^j\text{PMpre}_b$ as the re-centered IV, essentially only capturing "surprise" policy shocks.

V. Results

A. Disparities in Pollution Capitalization Rates

We begin by showing results from estimating versions of Equation 7 that omit neighborhood racial composition in Table 2. Column 1 also omits seller race interactions, and shows that air quality capitalizes into house prices. The omitted variable bias is sizable and positive in the OLS estimates in Panel (a) when comparing with the results in Panel (b) that use our regulatory instruments. This is consistent with the notion that economic activity is accompanied by beneficial amenities that push up house prices, while simultaneously increasing pollution. A one-unit decrease in $\text{PM}_{2.5}$ increases house prices by 6.3%.[31] This corresponds to an overall elasticity of -0.58, broadly in line with Sager & Singer (2025) in 2001-13 and Bento et al. (2015) in the 1990s.

Column 2 adds interactions with seller race. Panel (b) shows that the capitalization rate for Black homeowners is lower by one third, which is robust throughout our analysis (we turn to visualizations in Figure 4 later). This implies that racial house price disparities not only exist in levels, but also in house price changes resulting from plausibly exogenous changes in amenities, here

[29]This is more conservative than clustering at the block group level in Bishop et al. (2023), who use a similar instrument by interacting nonattainment with historic pollution levels to examine the impacts of pollution on dementia.

[30]Note that the Sager & Singer (2025) control already purges confounding trends based on nonrandom assignment of the nonattainment instrument by itself, so the re-centering helps most with interactions of the nonattainment instrument.

[31]Since the outcome is in logs, the coefficient translates into an exact percent change of $\exp(0.061) - 1 = 6.3\%$, and the elasticity is calculated as $(\exp(0.061) - 1) * 9.23$ using the overall endline period mean of $\text{PM}_{2.5}$.

Table 2: Capitalization rates with observed race of seller

	Transaction Price (log)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel (a) OLS results:									
PM2.5	-0.018***	-0.018***	-0.014***	-0.014***	-0.011***	0.026***	0.002*	-0.014***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.001)	(0.001)	(0.002)
PM2.5 * Black seller		0.036***	0.038***	0.038***	0.042***	0.038***	0.034***	0.037***	0.036***
		(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * Other seller		0.000	0.007***	0.008***	0.006***	0.007***	-0.000	0.003***	0.003***
		(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
log(SQFT)	0.519***	0.519***	0.520***	0.498***		0.569***	0.496***		0.519***
	(0.005)	(0.005)	(0.005)	(0.005)		(0.012)	(0.005)		(0.005)
log(HH income)				0.060***			0.099***		
				(0.001)			(0.001)		
PM2.5 * log(SQFT)						-0.005***			
						(0.001)			
PM2.5 * log(HH income)							-0.004***		
							(0.000)		
PM2.5 * log(Quality)								-0.001	
								(0.000)	
PM2.5 * PM2.5									-0.001***
									(0.000)
Panel (b) IV results:									
PM2.5	-0.061***	-0.061***	-0.070***	-0.071***	-0.060***	-0.239***	-0.093***	-0.051***	-0.019***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.022)	(0.004)	(0.006)	(0.007)
PM2.5 * Black seller		0.021***	0.024***	0.025***	0.021***	0.025***	0.025***	0.020***	0.023***
		(0.002)	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.003)	(0.002)
PM2.5 * Other seller		0.001**	0.004***	0.004***	0.003***	0.005***	0.004***	0.001*	0.003***
		(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(SQFT)	0.520***	0.520***	0.520***	0.498***		0.329***	0.497***		0.520***
	(0.005)	(0.005)	(0.005)	(0.005)		(0.028)	(0.005)		(0.005)
log(HH income)				0.060***			-0.001		
				(0.001)			(0.006)		
PM2.5 * log(SQFT)						0.021***			
						(0.003)			
PM2.5 * log(HH income)							0.007***		
							(0.001)		
PM2.5 * log(Quality)								0.016***	
								(0.002)	
PM2.5 * PM2.5									-0.001***
									(0.000)
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Property FE					Yes			Yes	
State by year FE	Yes	Yes					Yes		Yes
County by year FE			Yes	Yes	Yes	Yes		Yes	
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,745,962	7,745,962	7,743,104	7,376,157	2,297,704	7,743,104	7,379,025	2,127,829	7,745,962
First-stage F (KP)	680.555	279.610	249.262	248.774	200.619	276.349	19.529	130.125	10.218

Notes: The table shows regression estimates using OLS in Panel (a) and IV in Panel (b) with log transaction price as the dependent variable. The columns show results with varying controls and fixed effects as indicated. "HH income" means seller income at the transaction level, SQFT stands for square footage of the property, and "Quality" is our metric for property quality. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

air quality. The capitalization rate for Other homeowners is similar to that of NHW homeowners throughout, so we will focus on the Black-NHW gap in the remainder.

In Column 3 we include county-by-year instead of state-by-year fixed effects. One concern may be that our instrument also affects local industry or labor markets (Greenstone 2002, Walker 2013), potentially violating the exclusion restriction. As labor markets are usually defined at the county or MSA level, using county-by-year fixed effects provides reassurance on the validity of our instruments, as we are only exploiting the heterogeneous policy effects within nonattainment counties, rather than the policy effect across attainment and nonattainment counties. While it is possible that our identifying variation is correlated with employment effects even within counties, such localized impacts would require labor market frictions that are quite atypical in the US. The fact that our key estimate does not change much going from state-by-year to county-by-year fixed effects is also reassuring in this regard. Another concern is that unobserved moving costs attenuate the willingness-to-pay for air quality (Bayer et al. 2009). Both concerns affect primarily our estimate of the uninteracted $PM_{2.5}$ and would bias it upward. Indeed, when controlling for county-by-year fixed effects, this estimate becomes slightly more negative, but our estimated disparity in pollution capitalization rates remains nearly unchanged at one third. For the remainder, we continue to use county-by-year fixed effects.

Finally, our results are robust to re-centering our interaction instruments as in Borusyak & Hull (2023, 2025) as Column 8 of Table A.7 shows.

B. *Conditioning on Further Seller or Property Characteristics and Nonlinearities*

So far, the only property control included is log(SQFT). While homeowners are generally drawn from higher SES groups than non-homeowners, we also see income differences among homeowners along racial lines. In particular, Table 1 shows that Black sellers have a lower income on average than NHW sellers, and this could also affect house prices. In Column 4 of Table 2 we include seller income from HMDA data, which hardly affects our estimates. There may, however, still be several confounding property characteristics that are unobserved, such as the configuration of rooms or quality of the house. To address this concern, we include property fixed effects in Column 5 that accounts for average price and characteristics of each home, but requires repeat sales in the data. This reduces our sample size by two thirds, but the results are highly robust. This corroborates our non-parametric descriptive patterns with repeat sales in Figure 3b. It is also worth noting that since we estimate capitalization rate differences in relative terms with log prices as the dependent variable, we should, at least partially, address differences in level effects from baseline home values. Note that the Census block by seller race fixed effects control for the correlation of sales price with seller race, with the relationship being allowed to differ across space.

While this may control for differential property characteristics across racial groups, a remaining concern is that these property characteristics may themselves be driving different capitalization rates from air quality and confound our main estimate, i.e. we know that Black homeowners have smaller homes, on average, so it could be that SQFT is driving the difference in capitalization rates. In Columns 6 and 7 we interact log(SQFT) and log seller income respectively with $PM_{2.5}$. While these characteristics indeed affect capitalization rates themselves (to a small degree), they do not confound the estimated capitalization rate differences between Black and NHW sellers. While there may be other seller characteristics that we cannot observe, income is likely one of the most relevant, if not the most relevant, characteristic one may want to condition on to disentangle whether it is race itself or correlated characteristics driving the difference in capitalization rates.[32] Similarly, size of the property is likely the most important property characteristic. Since we obtain almost identical results on the racial disparity, irrespective of whether we condition on the most salient characteristics (income or property size) interacted with $PM_{2.5}$, these results point towards some form of discrimination, which we will unpack later. While the fact that controlling for income or property size does not affect our estimates may seem surprising, it is worth noting that our identification strategy relies on shocks to home values based on instrumented air quality improvements conditional on fixed effects, which helps to address potential bias in the racial disparity arising from correlated characteristics. This is consistent with Bayer et al. (2025), who note that, despite the strong aggregate correlation between race and income, controlling for neighbor income leaves their estimated racialized response to new different-race neighbor essentially unchanged. In Appendix Table A.5, we show that in a simple analysis of the racial disparity in housing prices per se, the estimated disparity due to seller race drops significantly once we condition on seller income.

To address remaining concerns that the capitalization-rate gap merely reflects unobserved differences in property characteristics, we build a leave-one-out property quality metric using repeat sales. For each property, we calculate the average percentage premium (or discount) of its PSQFT relative to the local market average PSQFT across all of the property's transactions, excluding the focal sale.[33] This is conceptually related to Diamond & Diamond (2024) who use past resale prices as a proxy for unobserved quality to impute rents. We interact this measure of property quality with air quality, fully instrumented, and show in Column 8 that our estimate of the capitalization gap remains consistent when controlling for capitalization due to property quality differences.[34]

Finally, since Black homeowners experience higher pollution reductions (see Figure 2b), there could be confounding nonlinearities in $PM_{2.5}$, so we include a quadratic $PM_{2.5}$ term in Column 9. Our estimates remain robust which rules out that our racial differences in capitalization rates are driven by non-linear effects of air quality improvements across racial groups. In Appendix Table

[32]Including income interacted with $PM_{2.5}$ also hardly affects our estimates when included together with neighborhood racial composition.

[33]We construct local market PSQFT index using local spatially weighted regressions centered on Census block groups following Ahlfeldt et al. (2023).

[34]To construct our metric of property quality, we have to rely on the repeat sales sample and also include property fixed effects.

A.7, Columns 6-7, we show similar robustness to including the interaction of $PM_{2.5}$ with baseline pollution, as a related concern for nonlinearity may be that Black homeowners live in more polluted areas at baseline and experience different effects for the same unit reductions as a result. Appendix Table A.7 Columns 1-3 also shows the robustness of Columns 6, 7 and 9 to including property fixed effects.[35] Table A.8 Columns 1-3 shows robustness to different thresholds for the match quality of lender names between Zillow and HMDA data (60% vs 75% vs 90% threshold).

C. Observed vs. Predicted Seller Race

We now turn to using predicted seller race instead of observed seller race, which greatly expands our sample size with the advantage of exploring external validity beyond transactions connected to a mortgage. In our case, predictions of race based on names may actually be an advantage when in practice only names are known and discrimination is based on beliefs derived from names. Column 1 in Table 3 replicates our baseline Column 3 of Table 2 for OLS and IV for convenience.

Column 2 shows a similar estimate for the difference in capitalization rates across race when using predicted race, with a sample that is more than four times as large. Column 3 shows robustness to adding property fixed effects in a repeat sales analysis. Since we know that there is measurement error in *predicted* seller race by construction, the true difference in capitalization rates is likely larger if the measurement error is classical. Indeed, in Table A.8, we vary the prediction threshold from 70% required probability of race classification in Column 4 to 60% in Column 5 where attenuation is larger yielding a smaller estimate, and 80% in Column 6, where attenuation is lower yielding a larger estimate. In Table A.9 Column 2, we use the same sample as Column 1 of Table 3, but use predicted race on the subset of the observed-race sample for which predicted race is also available, resulting in an attenuated coefficient compared to using observed seller race. Taken together, these results demonstrate the external validity of our estimated disparity in capitalization rates beyond transactions connected to mortgages, addressing concerns of potential selection into the observed race sample that could be correlated with pollution and house prices. If anything, these results imply that the true disparities may be even larger outside of our sample with observed race.

D. Seller vs. Buyer Race

While the disparity in capitalization rates across races is concerning, it is a priori unclear whether it is driven by seller race or buyer race. If Black homeowners only sell to Black buyers, then these buyers would effectively be purchasing air quality at a discount, limiting the welfare implications

[35]Note that for Columns 7 and 9 in Table 2, we only include state by year fixed effects. Appendix Table A.7 Columns 4-5 shows these results with county by year fixed effects instead, with almost identical coefficients, but a lower first stage F-stat. For some specifications, the F-stat is lower as we need instruments for each interaction with $PM_{2.5}$, and in some first stages the interaction of specific control variables with nonattainment has relatively lower t-stats, resulting in lower F-stats.

Table 3: Capitalization rates: Observed vs. predicted race and seller vs. buyer race

	Transaction Price (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a) OLS results:						
PM2.5	-0.014***	-0.021***	-0.023***	-0.015***	-0.022***	-0.033***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
PM2.5 * Black seller	0.038***	0.025***	0.029***	0.022***	0.026***	0.029***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * Other seller	0.007***	0.011***	0.012***	0.005***	0.010***	0.023***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
PM2.5 * Black buyer				0.003***	0.009***	0.010***
				(0.000)	(0.000)	(0.001)
PM2.5 * Other buyer				-0.000*	0.003***	0.003***
				(0.000)	(0.000)	(0.001)
log(SQFT)	0.520***	0.513***		0.518***	0.535***	
	(0.005)	(0.004)		(0.006)	(0.003)	
Panel (b) IV results:						
PM2.5	-0.070***	-0.118***	-0.142***	-0.060***	-0.136***	-0.177***
	(0.005)	(0.008)	(0.010)	(0.006)	(0.009)	(0.011)
PM2.5 * Black seller	0.024***	0.027***	0.031***	0.007***	0.034***	0.041***
	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.004)
PM2.5 * Other seller	0.004***	0.016***	0.018***	0.003***	0.016***	0.023***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
PM2.5 * Black buyer				-0.007***	-0.020***	-0.026***
				(0.001)	(0.003)	(0.005)
PM2.5 * Other buyer				-0.004***	-0.004	0.015***
				(0.001)	(0.003)	(0.005)
log(SQFT)	0.520***	0.513***		0.518***	0.535***	
	(0.005)	(0.004)		(0.006)	(0.003)	
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Property FE			Yes			Yes
County by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Black buyer FE				Yes	Yes	Yes
Other buyer FE				Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,743,104	34,313,433	19,791,941	1,907,273	20,045,603	10,159,730
First-stage F (KP)	249.262	180.324	130.190	80.710	181.471	267.230

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). Column 1 reproduces Column 3 from Table 2 using observed seller race. Column 2 uses predicted instead of observed seller race and therefore also includes transactions without links to the mortgage data resulting in a larger sample size. Column 3 adds property fixed effects using predicted race. Column 4-6 add buyer race interacted with $PM_{2.5}$ and buyer race fixed effects. Column 4 is based on observed seller and observed buyer race, with a resulting lower sample size. Columns 5-6 are based on predicted seller and buyer race, with Column 6 adding property fixed effects. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table 4: Capitalization rate differences: seller by buyer race

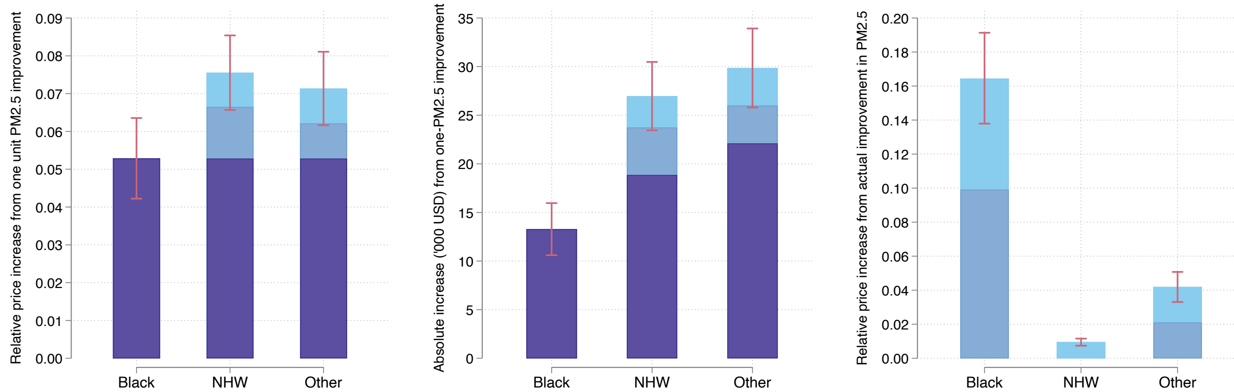
Seller/Buyer	Black Buyer	NHW Buyer	Other Buyer	Weighted average for seller
Black Seller	-0.014	-0.033	-0.029	-0.025
NHW Seller	0.020	0.000	0.004	0.001
Other Seller	0.004	-0.016	-0.011	-0.013
Weighted average for buyer	0.008	-0.002	-0.003	0.000

Notes: The table is based on Column 5 of Table 3 and shows the impact of a one-unit decrease in $PM_{2.5}$ on capitalization rate differences in percentage points (appropriately exponentiated coefficients), all compared to a NHW seller with a NHW buyer. For example, a Black seller with NHW buyer has a 3.3 percentage point lower capitalization rate (around 25% lower). The last column shows the weighted average across rows with shares based on sales going to respective buyer groups, applied before exponentiating to get percentage points.

only to those holding properties at the time of the air quality shock. Alternatively, if Black buyers do not receive a discount, the welfare implications from the capitalization rate differences would be magnified along the chain of sales. We begin by showing the degree of homophily in transactions in Appendix Table A.3. While the share of Black sellers selling to Black buyers is much higher than for other sellers selling to Black buyers, it is still only 40% with 60% selling to non-Black buyers, using the observations where we observe both seller and buyer race. To formally test how much seller vs. buyer race drives our result, we include buyer race interacted with $PM_{2.5}$ fully instrumented in Columns 4-6 of Table 3. Column 4 is based on observed seller and buyer race, while Column 5 is based on predicted seller and buyer race with Column 6 adding property fixed effects. In all three columns, the coefficient on Black seller remains positive and statistically significant, while the interaction coefficient with Black buyer is negative. Note that the coefficients in Column 4 for seller race interactions are somewhat smaller than in the rest of the table, but are also based on a much smaller subsample. They are in line with a version without buyer race interactions run on the same smaller subsample where we observe both races.[36] Importantly, we highlight that the OLS estimates are highly robust throughout the table for either observed or predicted seller race and with or without including buyers. This is reassuring as there is no a-priori reason why the OLS bias in the interaction of $PM_{2.5}$ with seller race should vary significantly *across* these specifications.

These results demonstrate that the disparity in capitalization rates is indeed driven by seller race. In fact, Black buyers pay a premium for air quality capitalizations, ceteris paribus, echoing results from other papers focused on buyers (Bayer et al. 2017, Higgins 2023). While this may suggest arbitrage opportunities, they may remain in part because of discriminatory bias itself, and in part because of high transaction costs in the housing market (Christensen & Timmins 2023). Based on Column 5 of Table 3 the capitalization rate for a NHW seller is around one-third higher compared with a Black seller who both sell to the same buyer race (e.g. a Black buyer). For a Black seller

[36] Appendix Table A.9 Column 3 shows that when running a specification without buyer race interactions, but on the same smaller sample where we observe both seller and buyer race, the coefficient on seller race interactions is smaller as well. Columns 4 and 5 of Table A.9 show results without using seller race at all, based on observed or predicted buyer race respectively. Due to (partial) homophily and omitting seller race, these estimates are positive, capturing some of the seller race effect.



(a) Relative effect of one-PM2.5 decrease (b) Absolute effect of (a) (c) Counterfactual with actual ΔPM2.5

Figure 4: Visualization of main effects and counterfactual

Notes: Panel (a) visualizes the main effect from Table 5 Column 1 showing relative price increases by race, where the violet bars represent the price increase common to all racial groups, the dark blue represents additional price increases from the seller-race component and the light blue from the neighborhood-composition component using the average Census block composition for a seller from the respective group in our sample. Panel (b) shows corresponding absolute price increases using the average home values by group. Panel (c) shows the relative price increases by racial groups in the absence of seller-race and neighborhood-composition differences, setting the interaction coefficients to zero and using the actual PM2.5 changes over the sample period by racial group. Red bars indicate 95% confidence intervals based on clustered Standard Errors at the Census tract level.

selling to a Black buyer compared to a NHW seller selling to a NHW buyer the capitalization rate is only 10% lower. Table 4 shows a three-by-three table for all combinations of seller and buyer race indicating the difference in capitalization rates in percentage points. Our main results capture the disparities across sellers averaged by the respective buyer shares in the data as shown in the last column of Table 4. Finally, we also include buyer race in our analysis of the relationship between seller race and sales price premium above local market rates within the same properties in Appendix Table A.6. Column 4 shows that the Black buyers pay a premium if buying from NHW sellers, and Black sellers receive a discount mainly when selling to NHW buyers, echoing the literature and results above.

E. *People vs. Places: Seller Race and Neighborhood Composition*

We now turn to decomposing the seller-group capitalization-rate gap into a seller-race component and a neighborhood-composition component. Black homeowners tend to live in blocks with different racial compositions than NHW homeowners as shown in Table 1. Note that we have already shown that the disparity is unchanged when controlling for interacted seller income or other property characteristics in Table 2. To separate seller race from place composition, we include both observed seller race and neighborhood racial shares interacted with $PM_{2.5}$ as in Equation 7.[37] We first show our baseline results and then ask whether observable place characteristics account for either

[37] While we defined neighborhoods here as blocks, our results are robust to defining them as tracts instead, as Column 1 of Table A.11 shows.

Table 5: Capitalization rates: Seller race and neighborhood composition

	Transaction Price (log)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel (a) OLS results:											
PM2.5	-0.018***	0.022***	0.068***	-0.007***	-0.021***	-0.018***	-0.013***	-0.017***	0.021***	-0.018***	-0.042***
	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)
PM2.5 * Black seller	0.024***	0.024***	0.023***	0.024***	0.024***	0.024***	0.024***	0.024***	0.021***	0.023***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * Other seller	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PM2.5 * Black share	0.038***	0.036***	0.033***	0.038***	0.038***	0.036***	0.037***	0.039***	0.022***	0.039***	0.036***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
PM2.5 * Other share	0.012***	0.010***	0.007***	0.011***	0.011***	0.010***	0.011***	0.013***	0.005***	0.012***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * log(Med. HH income)		-0.004***									
		(0.000)									
PM2.5 * log(Med. Hval)			-0.007***								
			(0.000)								
PM2.5 * log(Med. Rent)				-0.002***							
				(0.000)							
PM2.5 * Urban share					0.004***						
					(0.001)						
PM2.5 * Pov. rate						0.013***					
						(0.002)					
PM2.5 * NDVI							-0.010***				
							(0.002)				
PM2.5 * Built highmed								-0.000***			
								(0.000)			
PM2.5 * Opportunity									-0.063***		
									(0.003)		
PM2.5 * Supply Ela.										0.001	
										(0.001)	
PM2.5 * Arrest rate											0.534***
											(0.057)
log(SQFT)	0.520***	0.520***	0.520***	0.521***	0.520***	0.520***	0.519***	0.520***	0.521***	0.522***	0.520***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Panel (b) IV results:											
PM2.5	-0.075***	-0.055***	0.009	-0.068***	-0.073***	-0.075***	-0.083***	-0.071***	-0.051***	-0.062***	-0.071***
	(0.005)	(0.008)	(0.009)	(0.005)	(0.005)	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)	(0.010)
PM2.5 * Black seller	0.013***	0.013***	0.013***	0.012***	0.013***	0.013***	0.013***	0.013***	0.015***	0.013***	0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
PM2.5 * Other seller	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PM2.5 * Black share	0.027***	0.026***	0.022***	0.026***	0.028***	0.030***	0.030***	0.030***	0.025***	0.028***	0.026***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
PM2.5 * Other share	0.005***	0.004***	-0.000	0.005***	0.005***	0.007***	0.007***	0.008***	0.002	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * log(Med. HH income)		-0.002***									
		(0.000)									
PM2.5 * log(Med. Hval)			-0.006***								
			(0.000)								
PM2.5 * log(Med. Rent)				-0.001**							
				(0.000)							
PM2.5 * Urban share					0.005*						
					(0.003)						
PM2.5 * Pov. rate						-0.019***					
						(0.003)					
PM2.5 * NDVI							0.022***				
							(0.004)				
PM2.5 * Built highmed								-0.000***			
								(0.000)			
PM2.5 * Opportunity									-0.031***		
									(0.005)		
PM2.5 * Supply Ela.										-0.004*	
										(0.002)	
PM2.5 * Arrest rate											-0.013
											(0.323)
log(SQFT)	0.520***	0.520***	0.520***	0.521***	0.520***	0.520***	0.519***	0.520***	0.521***	0.522***	0.520***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,743,104	7,736,534	7,708,698	7,546,254	7,743,104	7,736,545	7,727,118	7,741,938	7,658,970	7,302,613	6,747,729
First-stage F (KP)	147.191	101.076	92.543	99.917	58.897	35.107	60.498	23.712	37.036	47.719	12.516

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). Seller race is observed and additional interactions with neighborhood racial composition are included throughout. Columns 2-11 add various neighborhood interactions with $PM_{2.5}$, as described in the text, and are fully instrumented in Panel (b). The non-interacted characteristics are absorbed by the fixed effects as they are time invariant. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

component.

Column 1 in Table 5 shows that our estimate for the interaction with Black seller drops by almost a half, and that the interaction with the share of Black residents in the Census block is highly significant, so at least part of the disparity is driven by the area where sellers live rather than seller race itself. Average racial neighborhood shares vary on the continuum between zero and one for different seller races, so we cannot directly read off the average split between the seller-race and neighborhood-composition components from the table.[38] Instead, we visualize our results from Column 1 in Figure 4a for a one-unit reduction in $PM_{2.5}$ and using average Census block compositions by respective seller race in our sample. The violet bars represent the price increase common to all racial groups, the dark blue bars represent additional price increases from the seller-race component and the light blue bars reflect price increases due to the neighborhood-composition component, with confidence intervals calculated from the covariance matrix of our estimates. The first thing to note is that the total capitalization rate by seller race is almost identical to that from Column 3 of Table 2, which omits neighborhood racial composition, and is also lower by around one-third for Black sellers compared with NHW sellers.[39] Second, the component of the Black-NHW capitalization rate disparity that is driven by seller race is around 60%, with the remaining 40% driven by the neighborhood racial composition. This, of course is the average, and varies depending on the racial composition of a specific block. At the extreme, when comparing a Black seller in an entirely Black community with a NHW seller in an entirely NHW community, the capitalization rate for the NHW seller would be 120% larger (7.8 vs 3.5 percent per μg),[40] and the split would reverse to be 32% seller race and 68% neighborhood racial composition. The capitalization rate for a NHW seller compared to a Black seller in a block with the same average composition would be 22% higher (7.5 vs 6.1 percent per μg), with the difference corresponding to the seller-race component alone, as both sellers incur the same neighborhood-composition component. As previously, the capitalization rates between NHW and Other sellers are similar, and the components in the visualization for Other sellers are with respect to Black sellers.

We next ask whether neighborhood racial composition can be accounted for by other amenities. Differences in amenities may change the value of clean air by being complements (or substitutes), such as green outdoor spaces, playgrounds, sports facilities, crime rates, walkability, or school quality. For example, Black Americans tend to live in areas with fewer green spaces, i.e. more impervious surfaces, and tend to be lower income (Table A.1). They also live in neighborhoods with fewer economic opportunities (Table A.1 and Chetty et al. (2018)). Housing supply elasticities may also vary with neighborhood composition, and areas with higher supply elasticities may see muted price

[38]Column 2 in Table A.10 shows that our results remain robust to re-centering our interaction instruments as in Borusyak & Hull (2023, 2025).

[39]More precisely, it is lower by 30% here and by 35% based on Table 2.

[40]Recall that the difference is 42% between Black and NHW homeowners based on average neighborhood compositions.

effects as housing stock expands more in response to air quality improvements (Chakma & Krause 2024).[41]

To formally test these, we begin by including median household income or median home value at the block group level at the beginning of our sample fully interacted with $PM_{2.5}$ and instrumented in Columns 2 and 3 of Table 5. Median income and home value themselves increase the value of air quality improvements, but including them hardly changes our main estimates based on race, both for the seller-race and the neighborhood-composition component (similar to including seller income in Table 2). While median home value arguably captures a summary value of amenities, we examine the role of several other neighborhood characteristics at the beginning of our sample in the remainder of Table 5 and visualize differences in capitalization rates stemming from the two components in Figure 5, where the labels indicate the relevant controls. The table and figure show that our results are largely unchanged when including interactions with median rent, urban share, poverty, median income, median home value, local vegetation index (NDVI), measures of imperviousness (BuiltHigh), economic opportunities, housing supply elasticities, or arrest rates.[42] These neighborhood observables decrease the neighborhood-composition component by at most 20% and the sum of the two components by at most 9%, with baseline median house values being the most impactful measure in our analysis. This yields a similar split of 35% to 65% between the neighborhood-composition and seller-race components. Overall, both the seller-race and the neighborhood composition effects are robust to including these controls and imply that lower capitalization of clean air for Black homeowners operates through both seller-race transaction channels and racialized place-based channels tied to where properties are located.

As a robustness check, we restrict our sample to blocks that see little change in racial composition (e.g. little gentrification) over time from 2000-2020 in Figure 6a, with a maximum 10 percentage point racial share difference in the right column (corresponding to Table A.10). The difference in capitalization rate becomes slightly larger in our restricted samples, with a roughly unchanged split between the seller-race and neighborhood-composition components, implying that, if anything, our results may be too conservative and underestimate the disparity in pollution capitalization rates. Table A.11 shows that these patterns are robust when defining neighborhoods and gentrification at the tract level, which captures changes in neighboring blocks as well.

Finally, in Table A.12, we use a long difference approach using data from two separate time periods, the 2000 Census and the 2015-2019 5-year averaged American Community Survey, with information on the median home value and median rent for each block group. For this analysis we only include block group composition interacted with pollution, not seller race, as we do not observe transactions here. First, it is reassuring that the results on home values without any Zillow

[41] Note, however, that this is unlikely to explain disparities in our setting since housing supply elasticities are, if anything, slightly *lower* in Black neighborhoods (see Table A.1).

[42] All of our results are also robust to using Census tract level racial composition instead of block level composition.

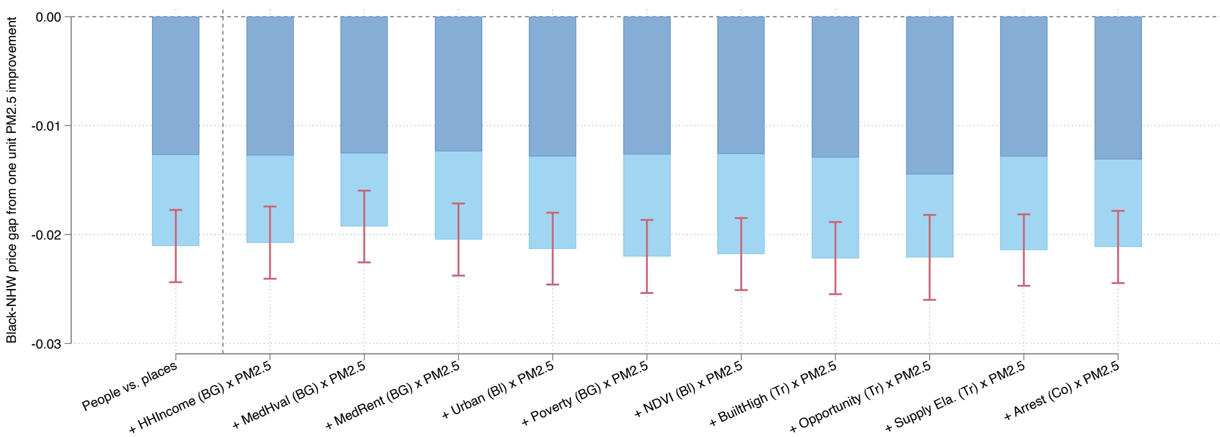


Figure 5: Visualizing difference in capitalization rates: Controlling for neighborhood characteristics

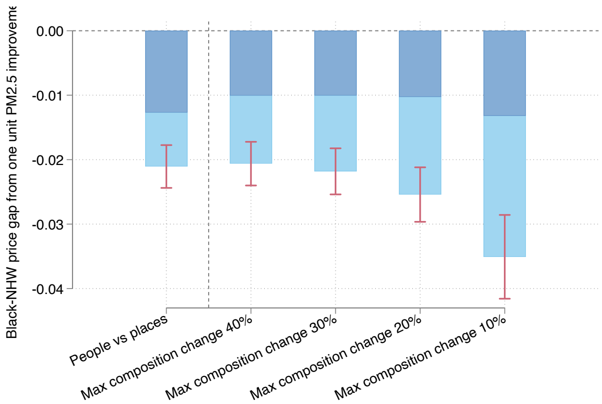
Notes: The figure visualizes the difference in capitalization rates between Black and NHW by seller-race component (blue) and neighborhood-composition component (light blue) from Table 5, varying the included controls that are interacted with PM2.5 as indicated and analogously to Table 5. BL, BG, Tr, and Co indicate controls measured at Census block, block group, tract, or county level. Red bars indicate 95% confidence intervals based on clustered Standard Errors at the Census tract level.

data show a similar pattern as our main analysis, with air quality capitalizing significantly less in areas with a higher Black population share. Second, rents are affected similarly as home values. That is, rental values increase relatively less per air quality improvement in neighborhoods with a higher Black population share. Unlike in our main analysis, the outcomes here are median values within neighborhoods. They may therefore mask variation within neighborhoods depending on race of renter and race of landlord, but the results suggest that some renters may benefit from lower capitalization rates.

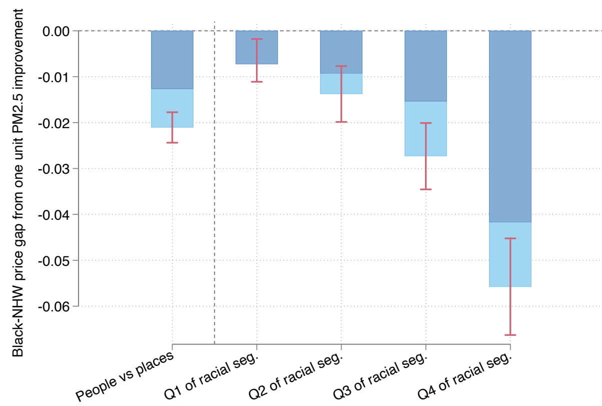
F. *Size of Effect and Counterfactual Analysis*

Our relative effects in Figure 4a, based on Column 1 of Table 5, imply that while a one-unit reduction in $PM_{2.5}$ increases house prices by 5.3% for Black homeowners, it increases them by 7.5% for NHW homeowners, a pollution capitalization rate that is 42% larger. We provide estimates in absolute terms in Figure 4b using average house prices by racial groups from Table 1. A one-unit improvement in $PM_{2.5}$ (around 10%) increases house prices by US$ 13,000 for average Black sellers, while the figure for average NHW sellers is roughly double that at US$ 27,000. This is larger than the difference in relative estimates, as baseline house prices are higher for NHW sellers.

On average, the NHW-Black house price gap in levels is 42% (see Table 1). While air quality improvements helped to improve home values for all groups, our striking result is that air quality improvements have actually widened the gap of home values between Black and NHW homeowners by 2 percentage points despite the shrinking pollution exposure gap, entirely due to the sizable



(a) Blocks with little compositional change over time



(b) Effect by quartile of residential segregation

Figure 6: Restricting to neighborhoods with little change, and results by residential segregation

Notes: The figure visualizes the difference in capitalization rates between Black and NHW by seller-race component (blue) and neighborhood-composition component (light blue). The leftmost baseline in both panels corresponds to Column 1 of Table 5. Panel (a) restricts the sample to blocks with a maximum compositional change in racial shares by the indicated percentage points (see Table A.10). Panel (b) shows results for sample splits by quartile of residential racial segregation within tracts (see Table 6). Red bars indicate 95% confidence intervals based on clustered Standard Errors at the Census tract level.

differences in pollution capitalization rates.[43] This can easily be seen as the reduction in pollution is only 35% greater for Black homeowners (6.6 vs 4.9 $\mu g/m^3$), but the capitalization rate is 42% greater for NHW homeowners.

We next ask what counterfactual prices would have prevailed if Black and Other homeowners had the same measured pollution reductions during our sample period, but experienced a capitalization rate at the level of NHW homeowners, i.e. setting β_j and γ_j to zero in Equation 7. Figure 4c shows that in this case house prices would be 16% higher for Black homeowners (equivalent to \$40,000 per homeowner) and 4% higher for Other homeowners, and higher by 1% for NHW homeowners (through neighborhood shares). This would have implied that the Black-NHW house price gap would have reduced by 14 percentage points from the air quality improvements over two decades, instead of the 2 percentage point increase.

Finally, we illustrate the loss from the seller and buyer double jeopardy where a Black seller sells an average home after a one-unit improvement in $PM_{2.5}$ and subsequently buys a home based on the results in Table 4. On average across all counter-parties, the loss is 2.5% from the sale discount and 0.8% from the purchase premium for a total of 3.3% of the home value, all compared to a NHW seller/buyer. If a Black seller sells to a Black buyer and buys from a Black seller, the loss is zero as the discount and premium cancel each other out. If on the other hand a Black seller sells to a NHW buyer and then buys from a NHW seller the loss is 3.3% from the sale discount and 2% from the purchase premium for a double jeopardy loss of 5.3%.

[43] This is calculated by combining the realized pollution improvements by racial groups with the capitalization rate and house price levels by group.

VI. Interpretation: Discrimination, Preferences, and Intermediation

We now turn to interpreting our findings. Before doing so, recall that our estimates are capitalization slopes: the pass-through of air quality improvements into sale prices in an intermediated, two-sided market. Slopes can reflect the marginal buyer's tastes for the bundle, including neighborhood composition and seller identity, as well as market frictions including search/steering, appraisal/lending, or information. We therefore interpret our results as outcomes, not as structural marginal willingness-to-pay (MWTP) measures as in the classical cross-sectional hedonic model of Rosen (1974). Because both sides of the market and intermediaries can shape capitalization, we first focus on whether the variation is indexed by the seller's identity or by neighborhood composition, and then discuss whether these patterns are more consistent with preferences, amenities, information, or intermediation.

A. Direct and Systemic Discrimination

Given the extensive evidence of past and present racial discrimination in housing markets, we ask whether either part of our decomposition, seller-race or racial neighborhood composition, may be related to discrimination. This distinction is useful because seller race and neighborhood composition enter housing transactions in conceptually different ways.

We begin with the seller-race component of the capitalization-rate gap. In the housing context, direct discrimination is defined as a race effect on a transaction outcome conditional on all relevant nonracial characteristics. The source of this discrimination can be taste-based, but could also be statistical (Phelps 1972), or paternalistic (Buchmann et al. 2024). Importantly for the housing context, it may also be mediated by agents, lenders, appraisers, and other intermediaries (Munnell et al. 1996, Charles & Hurst 2002, Avenancio-León & Howard 2022, Christensen et al. 2022, Christensen & Timmins 2022). As discussed in the conceptual framework in Section II., the ideal experiment to elicit forms of direct discrimination would vary race exogenously while holding fixed the house, neighborhood, and transaction environment. Such experiments exist for renters and buyers in correspondence and audit studies (e.g. Christensen et al. 2021, Christensen & Timmins 2022), which find significant discrimination in housing-market interactions, but it is not possible to experimentally vary seller race in actual housing transactions.

In the absence of such experimental variation, we assess whether the seller-race estimate is absorbed by seller or property characteristics that plausibly affect capitalization rates. Table 2 shows that property size, inferred property quality, and seller income, each fully interacted with pollution, do not affect the Black seller-race estimate. This does not rule out omitted seller traits that are correlated with race and could affect capitalization.[44] Seller income is nevertheless informative because

[44] For example, if Black sellers have fewer years of education on average, and education is correlated with the ability to market a property or navigate the transaction process, then controlling for education could reduce the seller-race estimate.

it is closely related to reservation values, credit conditions, bargaining position, and other dimensions of socioeconomic status that may shape the transaction process. The fact that interacted seller income does not move the estimate makes it less likely that omitted seller-side characteristics alone account for the seller-race component. The only factor that changes our seller-race estimate is neighborhood racial composition (Table 5). Neighborhood racial composition, as explained in Footnote 13, can in turn control for other unobserved amenities correlated with seller race, and Table 5 shows other amenity controls do not affect the seller-race estimate.

The racial neighborhood-composition component raises a different set of interpretive issues. Unlike seller race, neighborhood racial composition is a feature of place and may reflect both current amenities and past constraints on residential choice, investment, credit access, appraisal practices, and perceptions of neighborhood quality. This is the setting in which the distinction between direct and systemic discrimination becomes useful. Direct discrimination is a sharp but relatively narrow concept focused on treatment at the point of transaction. As Bohren, Hull & Imas (2025) show, a systemic perspective instead allows some variables typically treated as controls to embody accumulated effects of past and ongoing discrimination. In labor-market settings, for example, controlling for work experience may be important to elicit direct discrimination at the point of the transaction, but may understate systemic gender discrimination if labor-force attachment itself is shaped by childbearing norms or caregiving responsibilities. More broadly, characteristics often treated as fixed or exogenous may themselves reflect institutional, historical, or cultural forces that differ systematically across groups (Gneezy et al. 2009). In housing markets, neighborhood racial composition may similarly proxy for accumulated differences in investment, infrastructure, credit access, information, or residential sorting that affect capitalization rates.

In our context, there are reasons to believe that at least part of our estimate based on neighborhood racial composition, conditional on seller race, may capture such systemic discrimination. Buyers may have preferences over racial composition itself (Davis et al. 2024, Bayer et al. 2025, Almagro et al. 2024), which may affect neighborhood valuations conditional on all other factors, due to Becker taste-based or statistical discrimination. More importantly, neighborhood composition in the U.S. housing market is not driven by non-discriminatory preferences alone, but also reflects a long and dynamic history of constraints on residential choice, unequal public and private investment, racial steering, appraisal practices, credit-market frictions, and perceptions of neighborhood quality. For example, Sood & Ehrman-Solberg (2026) show that racial covenants shaped neighborhood formation and infrastructure and left persistent effects on segregation and housing price differentials long after the covenants ceased to be enforceable. More generally, neighborhoods are often shaped by persistent historical, institutional and cultural practices that unfairly privilege one group over another, with disadvantages that can accumulate over time (Schell et al. 2020).

As a result, racial neighborhood composition can, for example, be correlated with differences in access to complementary amenities that impact the capitalized value of clean air across communities (e.g. green outdoor spaces), but could also capture hard-to-measure objective or subjective factors that could affect the value of clean air (e.g. opportunities). Indeed, historical explicit discriminatory practices of "redlining" that limited finance based on the racial composition of neighborhoods led to underinvestment in these communities that persists today (Aaronson et al. 2021).[45] While home sellers of all races can live in all types of neighborhoods, Black sellers tend to be, on average, in neighborhoods with a larger share of Black residents, in part due to such historical and persistent practices that constrained options for Black homeowners with dynamic implications on neighborhood composition as well as investments in those communities (Ahmed & Hammarstedt 2008, Ewens et al. 2014, Akbar et al. 2022, Avenancio-León & Howard 2022, Christensen et al. 2022).

Thus, the neighborhood-composition component is informative about systemic, place-based channels to the extent that racial composition proxies for these historically shaped features of neighborhoods. This interpretation does not require that every effect operating through neighborhood composition be labeled systemic discrimination. Racial composition may also proxy for amenities, information, or expectations that affect the value of clean air. The previous Section and Table 5 show, however, that controlling for other fully interacted amenities reduces the neighborhood composition component by at most 20 percent, with hardly any impact on the seller-race estimate, but it is important to note that some neighborhood characteristics may proxy for amenities unrelated to discrimination, while others may themselves be part of the place-based channels of interest.[46]

The distinction between neighborhood racial composition and seller race also sheds light on possible mechanisms. In the terminology of Bohren, Hull & Imas (2025), systemic channels may operate through technological differences or informational distortions. In our setting, a technological channel would mean that racial composition is correlated with tangible complementary amenities that change the value of clean air. An informational channel would mean that racial composition affects perceptions, signals, search sets, stigma, or beliefs about neighborhood quality. The robustness of the neighborhood-composition component to interacted observable amenities in Table 5 makes the observable technological channel less compelling, although we cannot rule out the role of more granular amenities we do not measure. The majority of evidence points naturally toward information and intermediation: real estate agents, steering, listing and marketing choices, search frictions, and restricted buyer pools may all shape the racialized capitalization gaps we estimate.[47] We return

[45] Ward et al. (2025) show that intergenerational grandparent and great-grandparent effects for economic status in the US are driven by racial inequality and are non-existent within White families.

[46] In the framework of Bohren, Hull & Imas (2025), this is tantamount to assuming an alternative reference point treating those amenities as non-systemic controls, while leaving them in the place component treats them as possible manifestations of past legacies. In either case, the neighborhood-composition component remains large and significant.

[47] Neighborhoods may subjectively be valued based on racial composition including taste-based racial bias, stigmatized from outdated perceptions that were driven by direct discrimination itself, or from distorted and noisy signals about neighborhood quality. See also Fogli et al. (2024) for subjective biases from historical correlations driven by discrimination. As a result, beliefs about neighborhoods may be incorrect and distorted leading to inaccurate statistical discrimination as in Bohren, Haggag, Imas & Pope (2025), although in our setting these inaccurate beliefs can also be about neighborhoods irrespective of seller race.

to these channels after considering alternative explanations.

B. Alternative Explanations

We next consider whether the capitalization-rate differences could instead be explained by non-discriminatory differences in preferences, reservation prices, or timing. This amounts to asking whose behavior or preferences vary in the market, and whether such variation can account for the seller-race and neighborhood-composition components documented above. We explore several possible explanations.

First, it is possible that the seller-race and neighborhood-composition components simply capture differential marginal MWTP for clean air by racial groups. If Black buyers have a lower MWTP for clean air, for example due to income and information constraints, some degree of homophily could result in lower capitalization rates for Black sellers. The evidence in Table 3 and 4 shows that Black buyers pay roughly the same per unit of air quality improvement, and, if anything, pay slightly more than NHW buyers, making this explanation difficult to reconcile with the evidence.

Second, perhaps it is Black sellers, rather than buyers, that have a lower MWTP for clean air and therefore accept lower reservation prices. In this case, much of this lower MWTP likely stems from lower income, and including interacted income should attenuate our estimate. Table 2 shows the Black seller discount is, however, not driven by seller income, presenting evidence against this channel. Moreover, even if MWTP for clean air was lower for Black sellers, MWTP of the seller is arguably not the relevant primitive compared to the average MWTP in the market, since the capitalization accrues to incumbent property owners regardless of their MWTP, as in Greenstone & Gallagher (2008).

Third, NHW sellers might time the market better than Black sellers. Our inclusion of county by year fixed effects and our evidence in Figure 3b and Table A.6 where we net out the prevailing local market conditions from property sales imply that any timing advantage would have to be within years. While we cannot rule out short-run timing advantages, they are generally less important than the sort of longer-run advantages that we can rule out. Moreover, the source of short run timing advantages by race, to the extent that they exist, may themselves be due to types of discrimination, including through discriminatory intermediation.

Fourth, unobserved property quality could drive the seller-race component. This does not appear to be the case in our setting. We show that our results persist in a within-property analysis using repeat sales spells as well as when controlling for a fully interacted proxy of unobserved property quality in Table 2. This is further corroborated non-parametrically by evidence of a clear Black seller discount when selling the exact same property compared to a NHW seller in Figure 3b (see also Table A.6).

Finally, MWTP for clean air might depend on local purchasing power and therefore be driven by neighborhood income. In this case, MWTP for clean air could be lower for predominately Black neighborhoods that tend to be lower income. We test this directly by interacting neighborhood income (or home values or rent) with pollution in Table 5, showing that lower capitalization rates for Black neighborhoods are not driven by neighborhood income or wealth (although they do affect capitalization rates directly), and we show that they are also robust to including full interactions of key amenities including greenness or crime in Table 5.

Together, our empirical findings appear inconsistent with non-discriminatory preferences as the mechanism driving racial differences in capitalization rates, pointing toward discriminatory practices. Preferences of buyers (or agents) over transacting with a particular seller race conditional on the property and neighborhood likely captures forms of direct discrimination (including statistical or taste-based). Buyers may have preferences for neighborhood racial composition, in part, because composition is correlated with other neighborhood amenities (Davis et al. 2024, Bayer et al. 2025, Almagro et al. 2024). However, preferences of buyers (or agents) for neighborhood racial composition conditional on other amenities and seller race would be consistent with a systemic discrimination interpretation, again, whether taste-based, statistical, or informational. The apparent "arbitrage" opportunities both for properties sold by Black sellers and for properties in predominately Black neighborhoods can continue to exist precisely because of discriminatory preferences and practices.[48]

C. The Role of Intermediation and Experimental Evidence from Paired Testers

For both the direct and systemic discrimination components, the bias may come from the potential buyer pool but could also arise in the market through intermediation by real estate agents, for example through steering and restricted choice sets via discrimination (Christensen & Timmins 2023). This would lead to higher search frictions for Black homeowners or homeowners from predominately Black neighborhoods resulting in a thinner potential buyer pool or a pool with lower willingness-to-pay buyers. This smaller and selected pool of buyers will generally yield lower market clearing prices, which would drive both components in our setting.[49] There is experimental evidence on such differential search costs and steering by race in studies that focus on the buyer or renter side (Turner et al. 2013, Christensen & Timmins 2022, 2023), so it is plausible that there are similar distortions that affect search on the seller side.[50] Indeed, there is evidence that Black sellers are steered towards listing and advertising practices that result in lower transaction prices

[48] Transaction costs, which can be substantial for residential property purchases, can additionally help sustain apparent arbitrage gaps.

[49] Classical models of search predict a lower transaction price if sellers face higher search costs (Mortensen & Pissarides 1994).

[50] Furthermore, in Christensen & Timmins (2022) or Christensen & Timmins (2023), discriminatory search distortions for renters and buyers are larger in areas with higher environmental amenities such as air quality. This is consistent with our findings of increased discrimination on the seller side when environmental amenities improve. Note that our above results on the racial premium for Black buyers are also consistent with Christensen & Timmins (2023) as they find higher search constraints for Black renters leading to higher costs for them.

(Zillow Research 2025*b*,*a*). Results from Drukker & Ma (2025), which show that the homes of Black sellers spend more time on the market and realize lower housing returns, are also consistent with the notion of search frictions.

To shed more light on the role of search frictions and intermediation in the housing market, we use experimental data from Turner et al. (2013). They conduct paired tester trials where equally qualified renters or home buyers from two different racial groups inquire about the same available housing with the same real estate agent, and measure the responses and degree of unequal treatment from real estate agents.[51] We focus on the prospective home buyers and Black-White tester pairs to construct a measure analogous to Turner et al. (2013) to capture the difference in overall availability of homes by racial group. That is, we take the number of available properties the real estate agent said are available (intensive margin) multiplied by the local rate of obtaining a response from the agent at all (extensive margin), both by the racial group of the testers.[52] We construct the average Black-White gap in this measure for each of the 28 metro areas covered in Turner et al. (2013), and split our sample into high and low discrimination areas based on a median split (Figure A.2 maps these metro areas and overlays the nonattainment areas). Our presumption is that areas with more discrimination against Black buyers also have more discrimination against Black sellers, either directly based on seller race, or by steering buyers to neighborhoods with different racial composition. Columns 1 and 2 of Table 6 show that our results are indeed driven by the measured extent of housing discrimination. Both the seller-race and neighborhood-composition components are much higher in the above median sample than our overall estimates, with the former being insignificant and close to zero in our below median sample.

D. *Residential Segregation*

We know from the literature that racial bias and discrimination are generally higher in areas with more residential racial segregation (Enos & Celaya 2018, Ananat & Washington 2009, Ananat 2011, Avenancio-León & Howard 2022, Christensen et al. 2021), i.e. in Census tracts where different racial groups mainly live in separate blocks.[53] If our mechanisms are forms of discrimination, we would expect that the disparities in capitalization rates are larger in areas that are more residentially segregated. We test this formally by splitting our sample into quartiles based on our index of residential segregation by race (Figure A.3 maps our tract level index). Indeed, as the ascending Columns 3-6 in Table 6 show, and as visualized in Figure 6b, the disparity in pollution capitalization rates is much larger in tracts with more residential segregation.[54] Finally, for our analysis of the aver-

[51] Turner et al. (2013) find that Black home buyers are told or shown about 17% fewer homes on average.

[52] Our results are also robust to using the number of available properties alone.

[53] There is also a literature showing that underinvestment is higher and complementary amenities are lower in such areas (Trounstine 2016, Alesina et al. 1999).

[54] We find a similar result when using segregation within counties instead of segregation within tracts.

Table 6: Capitalization rates by homebuyer racial discrimination and by residential segregation

	Transaction Price (log)					
	Discrimination		Residential segregation			
	Low	High	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a) OLS results:						
PM2.5	-0.016***	-0.005**	-0.023***	-0.017***	-0.013***	-0.006**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
PM2.5 * Black seller	0.010***	0.032***	0.012***	0.019***	0.028***	0.048***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
PM2.5 * Other seller	0.001***	0.006***	0.002***	0.005***	0.007***	0.008***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)
PM2.5 * Black share	0.013***	0.051***	0.039***	0.035***	0.040***	0.026***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
PM2.5 * Other share	0.010***	0.007***	0.014***	0.012***	0.010***	0.011***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
log(SQFT)	0.540***	0.472***	0.527***	0.520***	0.519***	0.515***
	(0.005)	(0.012)	(0.008)	(0.007)	(0.012)	(0.013)
Panel (b) IV results:						
PM2.5	-0.145***	-0.091***	-0.095***	-0.053***	-0.071***	-0.064***
	(0.025)	(0.024)	(0.007)	(0.007)	(0.012)	(0.019)
PM2.5 * Black seller	-0.002	0.025***	0.007***	0.009***	0.016***	0.043***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)	(0.007)
PM2.5 * Other seller	-0.000	0.003***	0.002***	0.002**	0.003**	0.005*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
PM2.5 * Black share	-0.019***	0.055***	-0.003	0.015***	0.038***	0.045***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)	(0.007)
PM2.5 * Other share	-0.005**	0.007**	-0.002	0.000	0.016***	0.020***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.005)
log(SQFT)	0.540***	0.472***	0.527***	0.520***	0.519***	0.515***
	(0.005)	(0.012)	(0.008)	(0.007)	(0.012)	(0.013)
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes
County by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,201,377	1,305,404	2,100,872	2,064,784	1,945,040	1,622,392
First-stage F (KP)	21.375	5.300	67.875	68.701	16.930	6.888

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). The specification is the same as in Column 1 of Table 5. In Columns 1-2 we split the sample by median measured housing market discrimination based on Turner et al. (2013), indicated by low and high. In Columns 3-6, we split the sample by quartile of within-tract segregation, and report estimates by quartile, indicated by Q1-Q4. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

age seller race premium above local market rates within the same properties, Figure A.1 shows an increasing NHW premium for more segregated areas, in line with these findings.

VII. Conclusion

The environmental justice movement has its roots in the Civil Rights Movement of the 1960s, but its prominence in national priority setting and policy making is much more recent. Indeed, in an effort to address "...the disproportionate health, environmental, and economic impacts that have been borne primarily by communities of color..." President Biden issued Executive Order 14008 aimed at providing 40 percent of the benefits from Federal investments in the environment to marginalized communities, a commitment that was reinforced with Executive Order 14096 in 2023, but subsequently rescinded by President Trump in 2025. Our analysis underscores the complexity of this effort. Despite improvements in the pollution exposure gap, Black homeowners in the US benefited substantially less from pollution reductions than NHW homeowners. Indeed, had Black homeowners experienced the same capitalization rates as NHW homeowners, each would have gained \$40 thousand on average over our study period, equivalent to \$223 billion when extrapolating to all Black homeowners in the US.[55] These differential impacts are consistent with both seller-race transaction channels and racialized place-based channels, and highlight the need for research that moves beyond exposure analysis to better understand the marginal damages and benefits from that exposure.[56] They also underscore the inextricable link between various forms of inequality across communities such that environmental justice policies designed to overcome environmental disparities must also address social justice questions including forms of discrimination, search frictions, and unequal access to complementary amenities that help define the impacts of those disparities.

REFERENCES

Aaronson, D., Hartley, D. & Mazumder, B. (2021), 'The effects of the 1930s holc redlining maps', *American Economic Journal: Economic Policy* **13**(4), 355–92.

Ahlfeldt, G. M., Heblich, S. & Seidel, T. (2023), 'Micro-geographic property price and rent indices', *Regional Science and Urban Economics* **98**, 103836.

Ahmed, A. M. & Hammarstedt, M. (2008), 'Discrimination in the rental housing market: A field experiment on the internet', *Journal of urban economics* **64**(2), 362–372.

Akbar, P. A., Hickly, S. L., Shertzer, A. & Walsh, R. P. (2022), 'Racial segregation in housing markets and the erosion of black wealth', *Review of Economics and Statistics* pp. 1–45.

Alesina, A., Baqir, R. & Easterly, W. (1999), 'Public goods and ethnic divisions', *The Quarterly journal of economics* **114**(4), 1243–1284.

[55] Calculated using the coefficients of our main result in Table 5 Column 1, as explained in our section on counterfactual analyses. We multiply the relative 16% gain that Black homeowners would have gained with the average level of home values for Black homeowners to arrive at the per homeowner gain. The national figure comes from applying this to the total number of Black homeowners in the 2000 Census.

[56] While we study the distribution of housing capitalization changes resulting from pollution reductions, this insight is likely important for other realms of public policy such as health or education. Graff Zivin, Neidell, Sanders & Singer (2023), for example, find heterogeneous marginal pollution damages on health by vaccination status.

Almagro, M., Chyn, E. & Stuart, B. A. (2024), Neighborhood revitalization and inequality: Evidence from chicago's public housing demolitions, Technical report, Technical Report. Working paper.

Ambrose, B. W., Conklin, J. N. & Lopez, L. A. (2021), 'Does borrower and broker race affect the cost of mortgage credit?', *The Review of Financial Studies* **34**(2), 790–826.

Ananat, E. O. (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**(2), 34–66.

Ananat, E. O. & Washington, E. (2009), 'Segregation and black political efficacy', *Journal of Public Economics* **93**(5-6), 807–822.

Auffhammer, M., Bento, A. M. & Lowe, S. E. (2009), 'Measuring the effects of the clean air act amendments on ambient concentrations: The critical importance of a spatially disaggregated analysis', *Journal of Environmental Economics and Management* **58(1)**, 15–26.

Avenancio-León, C. F. & Howard, T. (2022), 'The assessment gap: Racial inequalities in property taxation', *The Quarterly Journal of Economics* **137**(3), 1383–1434.

Bajari, P., Fruehwirth, J. C., Kim, K. I. & Timmins, C. (2012), 'A rational expectations approach to hedonic price regressions with time-varying unobserved product attributes: The price of pollution', *American Economic Review* **102**(5), 1898–1926.

Banzhaf, H. S. (2021), 'Difference-in-differences hedonics', *Journal of Political Economy* **129**(8), 2385–2414.

Banzhaf, H. S. & Walsh, R. P. (2008), 'Do people vote with their feet? an empirical test of tiebout's mechanism', *American economic review* **98**(3), 843–863.

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

Barnes, S. & Stein, L. C. (2024), 'Racial discrimination in asset prices: Evidence from horse betting', *SSRN* **4904779**.

Baum-Snow, N. & Han, L. (2024), 'The microgeography of housing supply', *Journal of Political Economy* **132**(6), 1897–1946.

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

Bayer, P., Casey, M., McCartney, W. B., Orellana-Li, J. & Zhang, C. (2025), 'Distinguishing causes of neighborhood racial change: A nearest-neighbor design', *American Economic Review* **115**(11), 3999–4039.

Bayer, P., Keohane, N. & Timmins, C. (2009), 'Migration and hedonic valuation: The case of air quality', *Journal of Environmental Economics and Management* **58**(1), 1–14.

Bayer, P., McMillan, R., Murphy, A. & Timmins, C. (2016), 'A dynamic model of demand for houses and neighborhoods', *Econometrica* **84**(3), 893–942.

Bento, A., Freedman, M. & Lang, C. (2015), 'Who benefits from environmental regulation? evidence from the clean air act amendments', *Review of Economics and Statistics* **97**(3), 610–622.

Bhutta, N., Hizmo, A. & Ringo, D. (2022), 'How much does racial bias affect mortgage lending? evidence from human and algorithmic credit decisions', *FEDS Working Paper* .

Bishop, K. C., Ketcham, J. D. & Kuminoff, N. V. (2023), 'Hazed and confused: the effect of air pollution on dementia', *Review of Economic Studies* **90**(5), 2188–2214.

Bohren, J. A., Haggag, K., Imas, A. & Pope, D. G. (2025), 'Inaccurate statistical discrimination: An identification problem', *Review of Economics and Statistics* **107**(3), 605–620.

Bohren, J. A., Hull, P. & Imas, A. (2025), 'Systemic discrimination: Theory and measurement', *The Quarterly Journal of Economics* **140**(3), 1743–1799.

Borusyak, K. & Hull, P. (2023), 'Nonrandom exposure to exogenous shocks', *Econometrica* **91**(6), 2155–2185.

Borusyak, K. & Hull, P. (2025), Optimal formula instruments, Technical report, National Bureau of Economic Research.

Buchmann, N., Meyer, C. & Sullivan, C. D. (2024), 'Paternalistic discrimination', *Mimeo* .

Cassidy, A. W., Hill, E. L. & Ma, L. (2024), 'Who benefits from hazardous waste cleanups? evidence from the housing market', *NBER Working Paper* **30661**.

Chakma, T. & Krause, E. (2024), 'Where does air quality matter? new evidence from the housing market', *Mimeo* .

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

Chay, K. Y. & Greenstone, M. (2005), 'Does air quality matter? evidence from the housing market', *Journal of Political Economy* **113(2)**, 376–424.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R. & Porter, S. R. (2018), 'The opportunity atlas: Mapping the childhood roots of social mobility', *NBER working paper* **25147**, https://opportunityinsights.org/wp–content/uploads/2018/12/cty_covariates.dta (accessed on March 4th 2020).

Christensen, P., Sarmiento-Barbieri, I. & Timmins, C. (2021), Racial discrimination and housing outcomes in the united states rental market, Technical report, National Bureau of Economic Research.

Christensen, P., Sarmiento-Barbieri, I. & Timmins, C. (2022), 'Housing discrimination and the toxics exposure gap in the united states: Evidence from the rental market', *Review of Economics and Statistics* **104**(4), 807–818.

Christensen, P. & Timmins, C. (2022), 'Sorting or steering: The effects of housing discrimination on neighborhood choice', *Journal of Political Economy* **130(8)**.

Christensen, P. & Timmins, C. (2023), 'The damages and distortions from discrimination in the rental housing market', *The Quarterly Journal of Economics* **138**(4), 2505–2557.

Colmer, J., Hardman, I., Shimshack, J. & Voorheis, J. (2020), 'Disparities in pm2. 5 air pollution in the united states', *Science* **369**(6503), 575–578.

Colmer, J. M., Qin, S., Voorheis, J. L. & Walker, R. (2024), Income, wealth, and environmental inequality in the united states, Technical report, National Bureau of Economic Research.

Consumer Financial Protection Bureau (2018), *Data Point: 2017 Mortgage Market Activity and Trends: A First Look at the 2017 HMDA Data*, Washington, D.C.

Currie, J., Davis, L., Greenstone, M. & Walker, R. (2015), 'Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings', *American Economic Review* **105**(2), 678–709.

Currie, J., Voorheis, J. & Walker, R. (2023), 'What caused racial disparities in particulate exposure to fall? new evidence from the clean air act and satellite-based measures of air quality', *American Economic Review* **113**(1), 71–97.

Davis, M. A., Gregory, J. & Hartley, D. A. (2024), 'Preferences over the racial composition of neighborhoods: Estimates and implications', *Available at SSRN 4495735* .

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

Diamond, R. & Diamond, W. F. (2024), 'Racial differences in the total rate of return on owner-occupied housing'.

Didan, K. (2021), 'Modis/terra vegetation indices monthly l3 global 1km sin grid v061 [data set]'.
URL: *https://doi.org/10.5067/MODIS/MOD13A3.061*

Doleac, J. L. & Stein, L. C. (2013), 'The visible hand: Race and online market outcomes', *The Economic Journal* **123**(572), F469–F492.

Drukker, L. D. & Ma, L. (2025), 'Racial disparities in home selling', *SSRN* (4736956).

Enos, R. D. & Celaya, C. (2018), 'The effect of segregation on intergroup relations', *Journal of Experimental Political Science* **5**(1), 26–38.

EPA (2005), 'Air quality designations and classifications for the fine particles (pm2.5) national ambient air quality standards; final rule', *Federal Register* **70**(3).

Ewens, M., Tomlin, B. & Wang, L. C. (2014), 'Statistical discrimination or prejudice? a large sample field experiment', *Review of Economics and Statistics* **96**(1), 119–134.

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

FBI (2006), 'Uniform crime reporting program data [united states]: County-level detailed arrest and offense data, 2000'.
URL: *https://doi.org/10.3886/ICPSR03451.v4*

Fogli, A., Garcia Vazquez, M., Guerrieri, V. & Prato, M. (2024), 'Neighborhood segregation and endogenous racial bias', *Mimeo* .

Gneezy, U., Leonard, K. L. & List, J. A. (2009), 'Gender differences in competition: Evidence from a matrilineal and a patriarchal society', *Econometrica* **77**(5), 1637–1664.

Graff Zivin, J., Liao, Y. & Panassie, Y. (2023), 'How hurricanes sweep up housing markets: Evidence from florida', *Journal of Environmental Economics and Management* **118**, 102770.

Graff Zivin, J. & Neidell, M. (2012), 'The impact of pollution on worker productivity', *American Economic Review* **102**(7), 3652–73.

Graff Zivin, J., Neidell, M. J., Sanders, N. J. & Singer, G. (2023), 'When externalities collide: Influenza and pollution', *American Economic Journal: Applied Economics* **15**(2), 320–351.

Grainger, C. A. (2012), 'The distributional effects of pollution regulations: Do renters fully pay for cleaner air?', *Journal of Public Economics* **96**, 840–852.

Greenstone, M. (2002), 'The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufactures', *Journal of Political Economy* **110(6)**, 1175–1219.

Greenstone, M. & Gallagher, J. (2008), 'Does hazardous waste matter? evidence from the housing market and the superfund program', *The Quarterly Journal of Economics* **123**(3), 951–1003.

Higgins, B. E. (2023), 'Racial segmentation in the us housing market', *Mimeo* .

HMDA (2022), 'Loan/application register code sheets', *Home Mortgage Disclosure Act* .

Hsiang, S., Oliva, P. & Walker, R. (2019), 'The distribution of environmental damages', *Review of Environmental Economics and Policy* **13**(1), 83–103.

Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S. & Dominici, F. (2022), 'Air pollution exposure disparities across us population and income groups', *Nature* **601**(7892), 228–233.

Kahn, M. E. (2021), 'Racial and ethnic differences in the financial returns to home purchases from 2007 to 2020', *NBER Working Paper* .

Kermani, A. & Wong, F. (2021), 'Racial disparities in housing returns', *NBER Working Paper* .

Kuminoff, N. V. & Pope, J. C. (2014), 'Do "capitalization effects" for public goods reveal the public's willingness to pay?', *International Economic Review* **55**(4), 1227–1250.

List, J. A. (2004), 'The nature and extent of discrimination in the marketplace: Evidence from the field', *The Quarterly Journal of Economics* **119**(1), 49–89.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T. & Ruggles, S. (2022), 'Ipums national historical geographic information system: Version 17.0 [dataset]'. Accessed August 19, 2022.
URL: *http://doi.org/10.18128/D050.V17.0*

Mortensen, D. T. & Pissarides, C. A. (1994), 'Job creation and job destruction in the theory of unemployment', *The review of economic studies* **61**(3), 397–415.

Munnell, A. H., Tootell, G. M., Browne, L. E. & McEneaney, J. (1996), 'Mortgage lending in boston: Interpreting hmda data', *The American Economic Review* **86**(1), 25.

Myers, C. K. (2004), 'Discrimination and neighborhood effects: Understanding racial differentials in us housing prices', *Journal of urban economics* **56**(2), 279–302.

Perry, A., Rothwell, J. & Harshbarger, D. (2018), 'The devaluation of assets in black neighborhoods', *Library Catalog: www.brookings.edu* .

Phelps, E. S. (1972), 'The statistical theory of racism and sexism', *The american economic review* **62**(4), 659–661.

Reardon, S. F. & Firebaugh, G. (2002), 'Measures of multigroup segregation', *Sociological methodology* **32**(1), 33–67.

Ridker, R. G. & Henning, J. A. (1967), 'The determinants of residential property values with special reference to air pollution', *The Review of Economics and Statistics* **49**(2), 246.

Rosen, S. (1974), 'Hedonic prices and implicit markets: product differentiation in pure competition', *Journal of political economy* **82**(1), 34–55.

Sager, L. & Singer, G. (2025), 'Clean identification? the effects of the clean air act on air pollution, exposure disparities, and house prices', *American Economic Journal: Economic Policy* **17**(1), 1–36.

Schell, C. J., Dyson, K., Fuentes, T. L., Des Roches, S., Harris, N. C., Miller, D. S., Woelfle-Erskine, C. A. & Lambert, M. R. (2020), 'The ecological and evolutionary consequences of systemic racism in urban environments', *Science* **369**(6510), eaay4497.

Sood, A. & Ehrman-Solberg, K. (2026), 'The long shadow of housing discrimination: Evidence from racial covenants', *SSRN Working Paper* .
URL: *https://ssrn.com/abstract=4606234*

Trounstine, J. (2016), 'Segregation and inequality in public goods', *American Journal of Political Science* **60**(3), 709–725.

Tsivanidis, N. (Forthcoming), 'Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio', *American Economic Review* .

Turner, M. A., Santos, R., Levy, D. K., Wissoker, D. A., Aranda, C. & Pitingolo, R. (2013), *Housing discrimination against racial and ethnic minorities 2012*, U.S. Department of Housing and Urban Development.

US District Court (2022), 'Connolly et al v. lanham et al', *Complaint and Jury Demand* **1:2022cv02048**.

van der Loo, M. P. (2014), 'The stringdist Package for Approximate String Matching', *The R Journal* **6**(1), 111–122.
URL: *https://doi.org/10.32614/RJ-2014-011*

van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C. et al. (2021), 'Monthly global estimates of fine particulate matter and their uncertainty', *Environmental Science & Technology* **55**(22), 15287–15300.

Walker, W. R. (2013), 'The transitional costs of sectoral reallocation: Evidence from the clean air act and the workforce', *Quarterly Journal of Economics* **128(4)**, 1787–1835.

Ward, Z., Buckles, K. & Price, J. (2025), 'Like great-grandparent, like great-grandchild? multigenerational mobility in american history', *NBER working paper* .

Xie, F. (2022), 'rethnicity: An r package for predicting ethnicity from names', *SoftwareX* **17**, 100965.

Zillow (2020), *Transaction and Assessment Dataset (ZTRAX)*, Zillow. http://www.zillow.com/ztrax (accessed on November 20th 2020).

Zillow Research (2025*a*), 'Home sellers and private listing networks: Insights from a recent survey'. Research blog post.
URL: *https://www.zillow.com/research/pln-sellers-survey-34755/*

Zillow Research (2025*b*), 'Privately listed homes disproportionately harm sellers of color'. Press release.
URL: *https://zillow.mediaroom.com/2025-03-24-Privately-listed-homes-disproportionately-harm-sellers-of-color*

APPENDIX FOR ONLINE PUBLICATION

Disparities in Pollution Capitalization Rates

by Josh Graff Zivin[1,*] and Gregor Singer[2,*]

[1]UCSD and NBER

[2]London School of Economics

[*]Correspondence to: Josh Graff Zivin: jgraffzivin@ucsd.edu, Gregor Singer: g.a.singer@lse.ac.uk

A.1 Additional Descriptive Statistics

Table A.1: Summary statistics

	Mean	SD		Mean	
	Overall		Black	Non-Hisp. White	Other
Count	9186277	-	393483	7972109	820685
Shares	1	-	0.04	0.87	0.09
Transaction level					
Price in thousands (constant 2012 USD)	357.9	1731.3	251.2	357	418.4
SQFT	1900.3	42797.9	1797.8	1917.6	1781.2
Price per SQFT (constant 2012 USD)	201.4	1091.3	150.8	198.9	250.4
Predicted seller race: Black	0.09	0.28	0.75	0.07	0.01
Predicted seller race: Non-Hispanic White	0.59	0.49	0.11	0.74	0.03
Predicted seller race: Other	0.32	0.47	0.14	0.19	0.96
Income of seller in thousands (constant 2012 USD)	121	169.9	98.6	122.5	116.9
Observed buyer race: Black	0.05	0.21	0.4	0.03	0.05
Observed buyer race: Non-Hispanic White	0.84	0.37	0.47	0.89	0.49
Observed buyer race: Other	0.12	0.32	0.12	0.08	0.46
Income of buyer in thousands (constant 2012 USD)	121.4	160.6	96.2	122.4	122.4
Census block level					
Black population share of Census block (2000)	0.06	0.15	0.35	0.05	0.07
NHW population share of Census block (2000)	0.77	0.25	0.47	0.81	0.53
Other population share of Census block (2000)	0.17	0.2	0.18	0.14	0.4
Urban population share of Census block (2000)	0.86	0.35	0.92	0.85	0.94
PM2.5 concentration of Census block	9.3	3	10.1	9.1	11
Normalized Difference Vegetation Index of Census block	0.5	0.18	0.5	0.51	0.4
Census block group level					
Med. HH income in block group (2000, const. th. 2012 USD)	55.8	32.6	48.1	56.2	56.1
Share in poverty in Census block group (2000)	0.08	0.08	0.12	0.07	0.1
Med. home value in block group (2000, const. th. 2012 USD)	181.3	151.8	132.5	181	207.4
Med monthly rent in block group (2000, const. th. 2012 USD)	0.74	0.47	0.7	0.73	0.86
Census tract level					
Share high intensity built-up env. of Census tract (2001)	24.9	25.8	31.4	22.8	42.3
Opportunities in Census tract (1978-1983)	0.56	0.08	0.49	0.56	0.54
Housing supply elasticity of Census tract (2001)	0.37	0.26	0.33	0.38	0.26
Racial segregation within Census tract (2000)	0.28	0.1	0.28	0.28	0.24
Black population share of Census tract (2000)	0.07	0.13	0.33	0.05	0.07
NHW population share of Census tract (2000)	0.75	0.22	0.48	0.79	0.53
Other population share of Census tract (2000)	0.18	0.18	0.19	0.15	0.4
Urban population share of Census tract (2000)	0.86	0.29	0.93	0.85	0.94
County level					
Racial segregation within county (2000)	0.45	0.09	0.5	0.45	0.46
Arrest rate within county (2000)	0.05	0.02	0.05	0.05	0.05
PM2.5 Nonattainment county	0.3	0.46	0.45	0.29	0.34

Notes: The table shows the mean and standard deviation of indicated variables in the overall sample, and the mean by seller group.

Table A.2: Change in neighborhood composition and pollution changes

	(1)	(2)	(3)	(4)
Change in PM2.5	.00056	.001		
	(.00065)	(.00073)		
Log change in PM2.5			.0095	.012
			(.0073)	(.0096)
Observations	5904673	2979638	5904673	2979638
Incl. zero change blocks	Yes	No	Yes	No

Notes: The table shows regressions where the dependent variable is the change in the share of Black people in a given Census block between 2000 and 2020. All regressions include state fixed effects. The independent variable is based on the change from 2000 to 2020 of block level pollution concentrations. Column (2) and (4) exclude Census blocks with zero change in the dependent variable. Standard errors in parentheses are clustered at the county level.

Table A.3: Homophily: Percent of transactions going to specific buyer race, by seller race

Seller:	Black	NHW	Other
Black	40%	47%	12%
NHW	3%	89%	8%
Other	5%	49%	46%

Notes: The table shows the percent of transactions per seller racial group that go to specific buyer racial groups in our sample. This is based on observed seller and buyer race. See also Table A.1.

A.2 Additional Results, and Results in Table Form – Main Analysis

Table A.4: First Stage for Column 2 of Table 2

	PM2.5	PM2.5 * Black seller	PM2.5 * Other seller
	(1)	(2)	(3)
Nonattainment	0.021	-0.091***	-0.077***
	(0.143)	(0.007)	(0.027)
Nonattainment * Base PM2.5	-0.075***	0.009***	0.025***
	(0.010)	(0.001)	(0.002)
Nonattainment * Black seller	1.533***	3.262***	-0.078*
	(0.149)	(0.249)	(0.044)
Nonattainment * Black seller * Base PM2.5	-0.100***	-0.426***	-0.003
	(0.010)	(0.016)	(0.003)
Nonattainment * Other seller	1.938***	-0.020***	3.529***
	(0.135)	(0.006)	(0.185)
Nonattainment * Other seller * Base PM2.5	-0.141***	0.001*	-0.497***
	(0.009)	(0.000)	(0.011)
log(SQFT)	0.001	0.001*	-0.001**
	(0.001)	(0.000)	(0.001)
Census block by Black seller FE	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes
State by year FE	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes
Observations	7,745,962	7,745,962	7,745,962

Notes: The table shows first stage regression estimates of our three endogenous variables in Column 2 of Table 2, using our instruments. The combined Kleibergen-Paap F-stat of 279.6 is indicated in Table 2. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.5: Controlling for individual income or property size in a simple analysis

	Transaction Price (log)		
	(1)	(2)	(3)
Black seller	-0.378***	-0.279***	-0.330***
	(0.004)	(0.003)	(0.003)
Other seller	-0.072***	-0.019***	-0.041***
	(0.004)	(0.003)	(0.003)
log(HH income)		0.504***	
		(0.002)	
log(SQFT)			0.792***
			(0.008)
State by year FE	Yes	Yes	Yes
Observations	9,186,262	8,779,922	9,186,262

Notes: The table shows regression estimates from a simple analysis of how transaction price is correlated with seller race, and how this estimated relationship changes by including seller income or property size. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

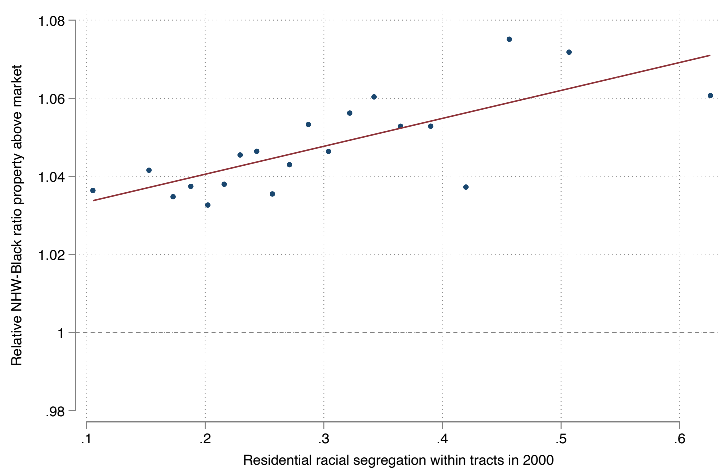


Figure A.1: Within-property NHW-Black premium ratio by residential segregation

Notes: The figure plots a binscatter based on repeat-sale properties that were each sold at least once by a Black seller and once by a NHW seller. We compare transaction prices to highly local price indices for PSQFT using locally weighted regressions as in Ahlfeldt et al. (2023). The figure plots the within-property premium ratio (NHW-Black) against the tract-level racial-segregation index.

Table A.6: Within-property sales price difference in PSQFT compared to local market average

	lDpsqft		Dpsqft	lDpsqft	
	(1)	(2)	(3)	(4)	(5)
Black seller	-0.045***	-0.044***	-5.091***	-0.024***	-0.025***
	(0.002)	(0.002)	(1.681)	(0.004)	(0.005)
Other seller	-0.022***	-0.023***	-4.967***	-0.017***	-0.016***
	(0.001)	(0.001)	(0.791)	(0.002)	(0.003)
Black buyer				0.021***	
				(0.003)	
Other buyer				0.002	
				(0.002)	
Black seller * RSEG					-0.063***
					(0.018)
Other seller * RSEG					-0.026**
					(0.010)
Property FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
County by year FE		Yes	Yes	Yes	Yes
Urban by year FE		Yes	Yes	Yes	Yes
Observations	2,347,298	2,127,829	2,127,369	389,022	2,126,120

Notes: The table shows regressions using repeat-sale properties, where the outcome variable is the difference of transaction PSQFT to the concurrent local price index. We construct these highly local price indices for PSQFT using locally weighted regressions as in Ahlfeldt et al. (2023). The columns vary by the inclusion of fixed effects or variables as indicated. All outcomes are in logs except for Column 3. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.7: Additional robustness for Table 2: fixed effects and baseline pollution

	Transaction Price (log)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel (a) OLS results:								
PM2.5	0.068***	-0.006***	0.014***	0.002	0.006***	-0.005**	-0.006**	-0.014***
	(0.004)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
PM2.5 * Black seller	0.042***	0.035***	0.038***	0.037***	0.039***	0.036***	0.037***	0.038***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
PM2.5 * Other seller	0.006***	-0.003***	0.000	0.007***	0.008***	0.001**	-0.002**	0.007***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
log(SQFT)				0.498***	0.520***	0.519***		0.520***
				(0.005)	(0.005)	(0.005)		(0.005)
log(HH income)		0.029***		0.090***				
		(0.002)		(0.001)				
PM2.5 * log(SQFT)	-0.011***							
	(0.001)							
PM2.5 * log(HH income)		-0.002***		-0.003***				
		(0.000)		(0.000)				
PM2.5 * PM2.5			-0.001***		-0.001***			
			(0.000)		(0.000)			
PM2.5 * Basline PM2.5						-0.001***	-0.001***	
						(0.000)	(0.000)	
Panel (b) IV results:								
PM2.5	0.020*	-0.118***	-0.005	-0.086***	0.091*	-0.027***	-0.013*	-0.065***
	(0.011)	(0.007)	(0.009)	(0.005)	(0.050)	(0.006)	(0.008)	(0.005)
PM2.5 * Black seller	0.021***	0.021***	0.018***	0.026***	0.027***	0.023***	0.017***	0.020***
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)	(0.005)	(0.002)
PM2.5 * Other seller	0.002**	0.003**	-0.000	0.005***	0.006***	0.002***	-0.000	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(SQFT)				0.498***	0.521***	0.520***		0.520***
				(0.005)	(0.005)	(0.005)		(0.005)
log(HH income)		-0.140***		0.032***				
		(0.014)		(0.004)				
PM2.5 * log(SQFT)	-0.010***							
	(0.001)							
PM2.5 * log(HH income)		0.016***		0.003***				
		(0.001)		(0.000)				
PM2.5 * PM2.5			-0.001***		-0.003***			
			(0.000)		(0.001)			
PM2.5 * Basline PM2.5						-0.002***	-0.002***	
						(0.000)	(0.000)	
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Property FE	Yes	Yes	Yes				Yes	
State by year FE		Yes	Yes			Yes	Yes	
County by year FE	Yes			Yes	Yes			Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,130,549	2,133,403	2,303,046	7,376,157	7,743,104	7,745,962	2,303,046	7,743,104
First-stage F (KP)	4.880	200.683	20.153	0.199	0.004	2.101	6.589	224.737

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). Columns 1-3 show the robustness of Columns 6, 7 and 9 of Table 2 to including property fixed effects. Columns 4-5 show results with county by year fixed effects instead of state by year fixed effects corresponding to Columns 7 and 9 in Table 2. Columns 6-7 show robustness to including the interaction of $PM_{2.5}$ with baseline pollution instead of pollution squared. Column 8 shows results when using re-centered IVs as in Borusyak & Hull (2023, 2025). Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.8: Robustness to data matching thresholds between HMDA and Zillow for observed race, and prediction thresholds for predicted race

	Transaction Price (log)					
	Lender match HMDA-Zillow			Race prediction threshold		
	60%	75%	90%	70%	60%	80%
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a) OLS results:						
PM2.5	-0.014***	-0.014***	-0.014***	-0.021***	-0.019***	-0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * Black seller	0.038***	0.040***	0.041***	0.025***	0.017***	0.033***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
PM2.5 * Other seller	0.007***	0.007***	0.007***	0.011***	0.011***	0.011***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
log(SQFT)	0.520***	0.521***	0.527***	0.513***	0.507***	0.519***
	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)
Panel (b) IV results:						
PM2.5	-0.070***	-0.071***	-0.069***	-0.118***	-0.108***	-0.127***
	(0.005)	(0.005)	(0.006)	(0.008)	(0.008)	(0.008)
PM2.5 * Black seller	0.024***	0.025***	0.025***	0.027***	0.020***	0.035***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)
PM2.5 * Other seller	0.004***	0.004***	0.003***	0.016***	0.016***	0.016***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(SQFT)	0.520***	0.521***	0.527***	0.513***	0.507***	0.519***
	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes
County by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,743,104	6,755,258	4,142,871	34,313,433	41,571,561	28,217,750
First-stage F (KP)	249.262	241.896	199.987	180.324	193.575	164.017

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). Columns 1-3 show robustness to different thresholds for the match quality of lender names between Zillow and HMDA data, using a 60% vs 75% vs 90% threshold as indicated. Columns 4-6 rely on predicted rather than observed race, where Column 4 replicates Column 2 of Table 3 with a 70% prediction threshold of a name belonging to a race. Columns 5 and 6 use a threshold of 60% and 80% respectively. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.9: Additional robustness for Table 3: Observed or predicted race of seller and buyer

	Transaction Price (log)				
	(1)	(2)	(3)	(4)	(5)
Panel (a) OLS results:					
PM2.5	-0.014***	-0.019***	-0.016***	-0.013***	-0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * Black seller	0.038***	0.023***	0.023***		
	(0.001)	(0.001)	(0.001)		
PM2.5 * Other seller	0.007***	0.006***	0.005***		
	(0.000)	(0.000)	(0.001)		
PM2.5 * Black buyer				0.019***	0.015***
				(0.001)	(0.001)
PM2.5 * Other buyer				0.003***	0.005***
				(0.000)	(0.000)
log(SQFT)	0.520***	0.523***	0.518***	0.506***	0.512***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Panel (b) IV results:					
PM2.5	-0.070***	-0.090***	-0.063***	-0.072***	-0.091***
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
PM2.5 * Black seller	0.024***	0.017***	0.004		
	(0.002)	(0.002)	(0.003)		
PM2.5 * Other seller	0.004***	0.004***	0.001		
	(0.000)	(0.001)	(0.001)		
PM2.5 * Black buyer				0.015***	0.011***
				(0.001)	(0.001)
PM2.5 * Other buyer				0.002***	0.004***
				(0.000)	(0.000)
log(SQFT)	0.520***	0.523***	0.518***	0.506***	0.512***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Census block by Black seller FE	Yes	Yes	Yes		
Census block by Other seller FE	Yes	Yes	Yes		
County by year FE	Yes	Yes	Yes	Yes	Yes
Census block by Black buyer FE				Yes	Yes
Census block by Other buyer FE				Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes
Observations	7,743,104	3,107,203	1,907,273	10,564,102	4,962,901
First-stage F (KP)	249.262	219.635	190.436	284.281	242.845

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). Column 1 repeats Column 1 of Table 3 using observed seller race, and Column 2 shows results when using predicted seller race but on the sample where we also observe seller race. Column 3 shows results with observed seller race but on the sample where we also observe buyer race. Column 4 uses observed buyer race, and Column 5 uses predicted buyer race, but on the sample where we also observe buyer race. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.10: Recentered IV, and capitalization rates when restricting the sample to neighborhood blocks with little compositional change over time

	Transaction Price (log)					
	All	Re-centered IV	$\leq$ 40 pp	$\leq$ 30 pp	$\leq$ 20 pp	$\leq$ 10 pp
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a) OLS results:						
PM2.5	-0.018***	-0.018***	-0.018***	-0.018***	-0.016***	-0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM2.5 * Black seller	0.024***	0.024***	0.019***	0.019***	0.019***	0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
PM2.5 * Other seller	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
PM2.5 * Black share	0.038***	0.038***	0.052***	0.057***	0.065***	0.080***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
PM2.5 * Other share	0.012***	0.012***	0.014***	0.014***	0.015***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(SQFT)	0.520***	0.520***	0.523***	0.528***	0.537***	0.545***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
Panel (b) IV results:						
PM2.5	-0.075***	-0.067***	-0.073***	-0.070***	-0.067***	-0.064***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
PM2.5 * Black seller	0.013***	0.010***	0.010***	0.010***	0.010***	0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)
PM2.5 * Other seller	0.002***	0.002***	0.002***	0.002***	0.002***	0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
PM2.5 * Black share	0.027***	0.023***	0.034***	0.038***	0.049***	0.072***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)
PM2.5 * Other share	0.005***	0.004***	0.006***	0.007***	0.008***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
log(SQFT)	0.520***	0.520***	0.523***	0.528***	0.537***	0.545***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes	Yes
County by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,743,104	7,743,104	6,781,568	5,988,388	4,573,743	2,400,413
First-stage F (KP)	147.191	92.363	130.109	109.310	86.013	83.382

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b). Column 1 reproduces our baseline result from Column 1 of Table 5. Column 2 shows results when using re-centered IVs as in Borusyak & Hull (2023, 2025). Columns 3-6 restrict our sample to blocks that experienced a maximum 40 (30, 20, 10) percentage point change in racial shares respectively. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.11: Capitalization rates using *tracts* composition, and when restricting the sample to neighborhood *tracts* with little compositional change over time

	Transaction Price (log)				
	All	$\leq$ 40 pp	$\leq$ 30 pp	$\leq$ 20 pp	$\leq$ 10 pp
	(1)	(2)	(3)	(4)	(5)
Panel (a) OLS results:					
PM2.5	-0.019***	-0.019***	-0.018***	-0.014***	-0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
PM2.5 * Black seller	0.021***	0.019***	0.019***	0.022***	0.032***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
PM2.5 * Other seller	0.004***	0.004***	0.004***	0.005***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
PM2.5 * Black share	0.053***	0.056***	0.059***	0.062***	0.080***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)
PM2.5 * Other share	0.014***	0.014***	0.015***	0.014***	0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
log(SQFT)	0.520***	0.521***	0.525***	0.537***	0.546***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.010)
Panel (b) IV results:					
PM2.5	-0.076***	-0.076***	-0.069***	-0.056***	-0.050***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)
PM2.5 * Black seller	0.011***	0.010***	0.009***	0.009***	0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)
PM2.5 * Other seller	0.002***	0.002***	0.002***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
PM2.5 * Black share	0.037***	0.038***	0.038***	0.046***	0.088***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)
PM2.5 * Other share	0.005***	0.005***	0.005***	0.005***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log(SQFT)	0.520***	0.521***	0.525***	0.537***	0.546***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.010)
Census block by Black seller FE	Yes	Yes	Yes	Yes	Yes
Census block by Other seller FE	Yes	Yes	Yes	Yes	Yes
County by year FE	Yes	Yes	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes	Yes
Observations	7,743,104	7,597,832	7,189,612	5,815,294	2,499,790
First-stage F (KP)	153.261	138.046	132.592	96.597	54.798

Notes: The table shows regression estimates from our transaction level approach using OLS in Panel (a) and IV in Panel (b), but using tract level racial population shares instead of block level shares. Column 1 is the equivalent of our baseline result from Column 1 of Table 5. Columns 2-5 restrict our sample to tracts that experienced a maximum 40 (30, 20, 10) percentage point change in racial shares respectively. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.12: Capitalization rates into local median house values and rents measured through Census and ACS data instead of Zillow data

	Levels of median		Logs of median	
	Home value	Rental rate	Home value	Rental rate
	(1)	(2)	(3)	(4)
Panel (a) OLS results:				
PM2.5	-28779.067***	-26.598***	0.008**	0.026***
	(1207.667)	(3.567)	(0.004)	(0.005)
PM2.5 * Black share	12312.405***	33.309***	0.023***	0.022***
	(288.472)	(0.795)	(0.001)	(0.001)
PM2.5 * Other share	21148.333***	46.713***	-0.012***	0.020***
	(688.275)	(1.455)	(0.002)	(0.002)
Panel (b) IV results:				
PM2.5	-42302.676***	-72.058***	-0.028*	-0.022
	(3974.689)	(12.098)	(0.015)	(0.019)
PM2.5 * Black share	11752.792***	32.326***	0.017***	0.018***
	(413.245)	(1.093)	(0.002)	(0.002)
PM2.5 * Other share	16513.462***	42.455***	-0.018***	0.016***
	(804.715)	(1.697)	(0.002)	(0.002)
Census block group FE	Yes	Yes	Yes	Yes
County by year FE	Yes	Yes	Yes	Yes
Urban by year FE	Yes	Yes	Yes	Yes
Baseline PM2.5 by year slopes	Yes	Yes	Yes	Yes
Observations	397,920	343,324	397,920	343,324
First-stage F (KP)	233.271	230.430	233.271	230.430

Notes: The table shows regression estimates from a long-difference approach using OLS in Panel (a) and IV in Panel (b). The outcome variable is the block-group level median house value or median rental rate as measured in two separate time periods, in the 2000 long form Census and from the 5-year 2015-2019 American Community Survey. The racial shares are measured at the block group level in 2000. Standard errors are clustered at the Census tract level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

A.3 Additional Descriptives – Maps

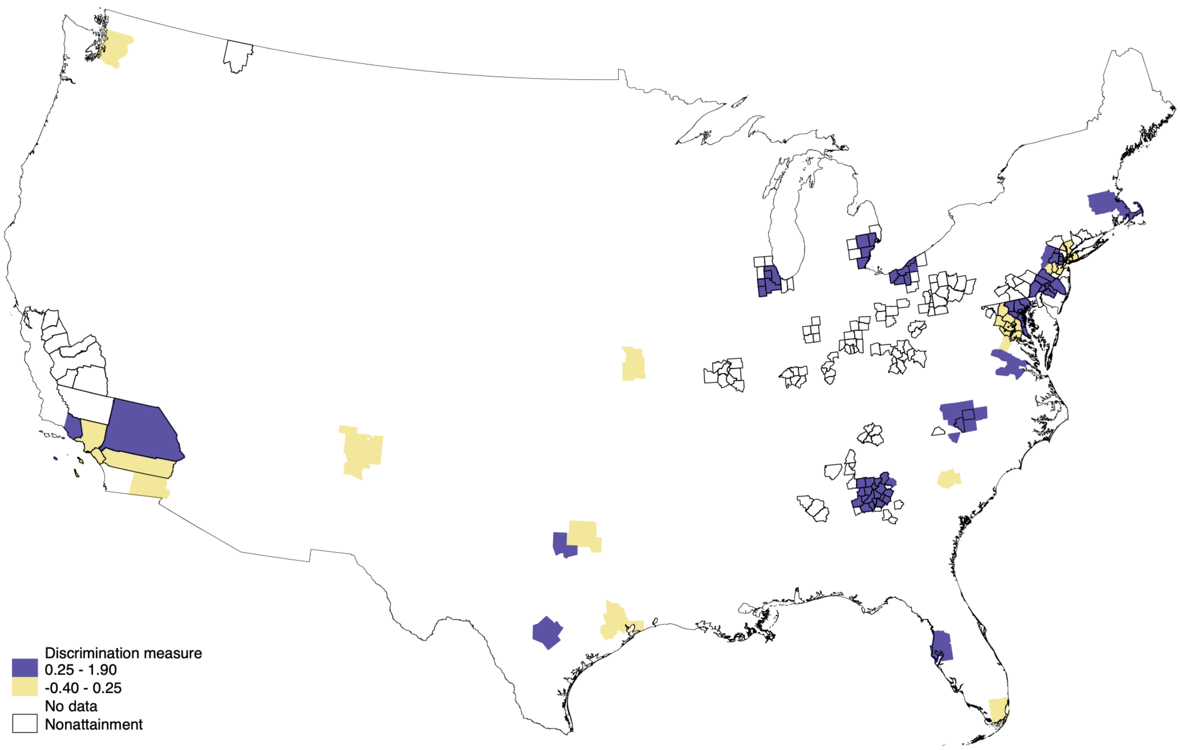


Figure A.2: Above and below median sample split discrimination measures and nonattainment areas

Notes: The map shows the above and below median sample split of buyer housing market discrimination based on data from Turner et al. (2013). The map also shows the overlap with nonattainment areas.

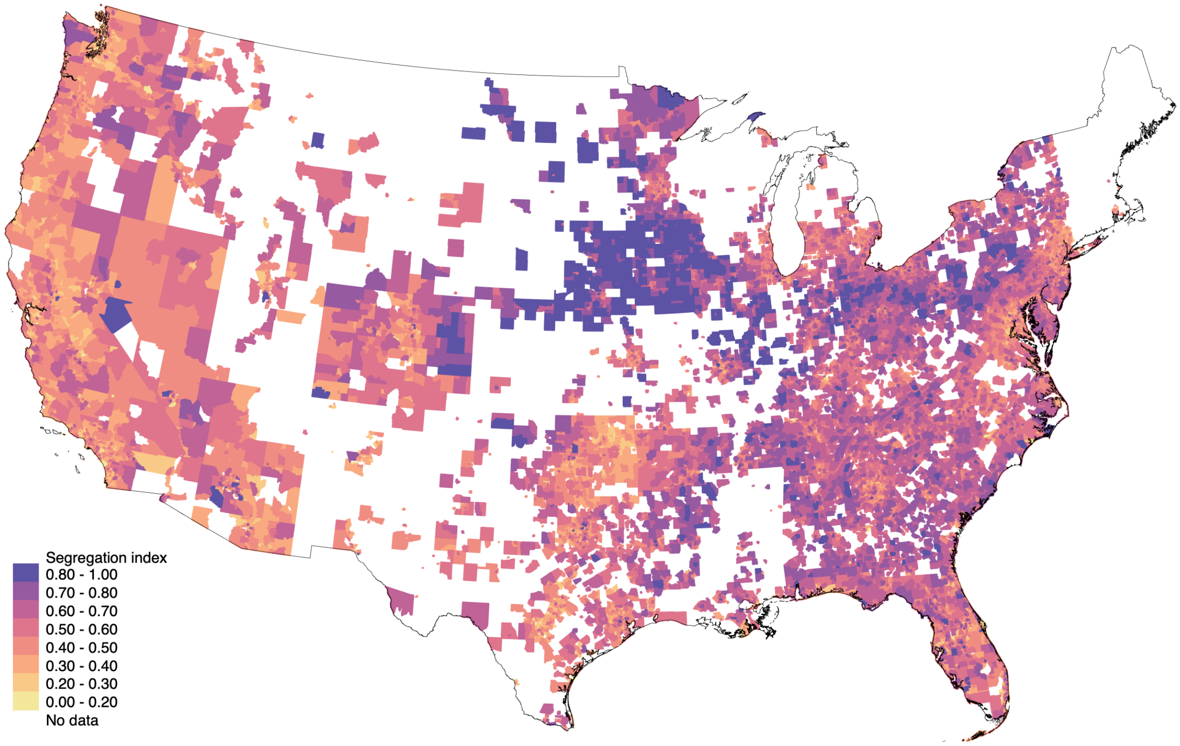


Figure A.3: Spatial distribution of within Census tract segregation in 2000

Notes: The map shows the spatial distribution of segregation within Census tracts across the contiguous US in 2000. We only use Census tracts which are in our final sample of transactions.

A.4 Details of Price, Location and Square Foot Data from Zillow

This section documents the transaction level data from Zillow. This has three building blocks which we discuss in turn: (A) identifying arm's length transactions for residential properties from the raw data, (B) identifying a property's location, and (C) identifying the property's area in square footage. Overall, we have both square footage and coordinates for 44,799,731 (84.3%) of our arm's length properties and for 80,544,782 (86.9%) of our arm's length transactions. Restricting this to the years used for analysis 2000-2019 and the contiguous US results in 64,737,270 (69.9%) transactions. The main text lists the number of observations of the subset used for analysis where we can observe or predict race.

A. Identification of residential arm's length transactions

The Zillow data contains a large number of transactions which are not arm's length housing transactions for residential property. These often have missing or zero prices, are foreclosures, intra-family transfers, pure loans, or refinancing transactions. This section documents how we identify arm's length transactions for residential properties. The raw transaction data contains 460.8 million observations which we reduce to 92.6 million transactions that are defined as arm's length, which are in turn based on 53.2 million properties.[57] The following sections document how we identify our set of residential arm's length transactions

1. Missing or low sales price (71.3% of total)

As a first step, we remove transactions with a missing or low sales price. This amounts to removing 328.7 million transactions bringing the count down to 132.1 million. This helps to address several other issues as well (as e.g. refinancing transactions are likely to have a missing sales price). We choose a threshold for low sales price of ≤ 1000 as there is a drop-off in density after 1000US$ and 5000US$. Out of all transactions (incl. Texas where prices are typically not recorded), 70% have a missing sales price, and 1.3% of transactions have a sales price of zero or $\leq$ 1000US$. Of the 1.3% of the last category ($\leq$ 1000US$), most prices are near zero. We drop all transactions with a sales price of ≤ 1000 or missing.

2. Foreclosures and distress sales (2.0% of total)

As a second step, all remaining identified foreclosures are removed. This includes all *data* types "Foreclosure". Since all of these observations have a missing or low sales prices, no additional observations are removed. We classify and remove additional *document* types associated with distressed

[57] This is the data downloaded from Zillow on April 7th 2020. This excludes observations with a transaction date before 1990. Since one transaction can contain multiple housing units, the number of units transacted in the raw data is slightly higher at 485 million.

sales and foreclosures. Some of them are not foreclosures in a strict legal sense, but practically very close to it (Receiver's Deed, for example). The Bargain and Sale Deed (BSDE) is one of the main deed types in Nevada and is therefore retained. BSDE deeds make up 69% of transactions with a sales price >US$1000 in Nevada, compared to 1.3% in all states. This removes an additional 9.3 million transactions bringing the count down to 122.8 million.

3. Intra-family and gift transfers (0.7% of total)

The third step is to remove intra-family and gift transactions. There is an intra-family flag coded by the Zillow team, which predominately corresponds to the INTR document type. We identify and remove additional intra-family or gift document types, such as Gift Deeds or Affidavit - Surviving Spouse.[58] In total, 3.4 million intra-family and gift transfer transactions are removed, bringing the count down to 119.5 million transactions.[59]

4. Credit lines, refinancing and pure mortgages (1.4% of total)

While Zillow defaults deed transfer documents to DEED, pure loan documents default to MTGE. All document types MTGE are removed (only 9 at this stage). The default value for loan types is empty. There are 6.5 million observations with recorded loan types, mainly commercial loans and seller take back loans. All transactions with a recorded loan type are removed, which includes refinancing transactions and new credit lines (e.g. HELOCs). In total this brings the observations down to 113.0 million transactions.[60]

5. Non-residential property types (1.2% of total)

We next remove non-residential property types. Table A.13 lists the retained and removed property types. Transactions with missing property types (45% of the remaining transactions) are also retained. Most of the non-missing property types are single family residences. This step removes 5.4 million transactions bringing the count to 107.6 million transactions.

6. Multiple properties per transaction and missing panel ids (2.3% of total)

A particular transaction can contain multiple units. All transactions with multiple units are dropped as the sales price cannot be assigned to a particular property. This first step removes 2.3 million transactions (0.5% of total). We also remove the transactions with missing property (panel) ids,

[58] There is an additional code in the data types (GT: No Consideration - Gift), which does not remove any additional observations, however.

[59] Quitclaim deeds are around 3 million transactions. Some may be used for intra-family transfer, but not necessarily, so they are retained in the data.

[60] It is possible that some of these loan types are actual transactions, although it is unlikely. One way to further refine this could be to use the information on buyers and sellers.

Table A.13: Retained and removed property types

Retained property types	Removed property types
AP – Apartment Building	AG – Agricultural
CD – Condominium	CI – Commercial & Industrial
MF – Multi-Family Dwelling (2-4 Units)	CM – Commercial
MH – Manufactured Home	CP – Cooperative
MX – Mixed Use	EX – Exempt
NW – New Construction	GV – Government
PD – Planned Unit Development	IM – Improved Land
RR – Residential	IN – Industrial
SR – Single Family Residence	MB – Mobile Home
	RC – Recreational
	UL – Unimproved Land/Lot
	VL – Vacant Land/Lot

Notes: The table lists the retained and removed property types. Missing property types are also retained.

which constitutes 8.4 million transactions (1.8% of the total). For a handful of small states, missing property ids are more than or near 50%, but in most states it is missing less than 20% of cases. In total, this step removes 10.7 million observations bringing the count to 96.8 million transactions.

7. Repeated sales (0.7% of total)

Next, a subset of transactions within a short period of time are removed. Specifically, if there are multiple transactions of the same property within a 90 day rolling window, only the last of these transactions is retained.[61] The last transaction of a spell of repeated sales is typically higher than the removed previous sales as Figure A.4 shows. Overall the last transaction in a spell is larger than the first in around 80% of spells. This step removes 3.4 million transactions bringing the total count to 93.5 million transactions.

8. Multiple unit properties (0.2% of total)

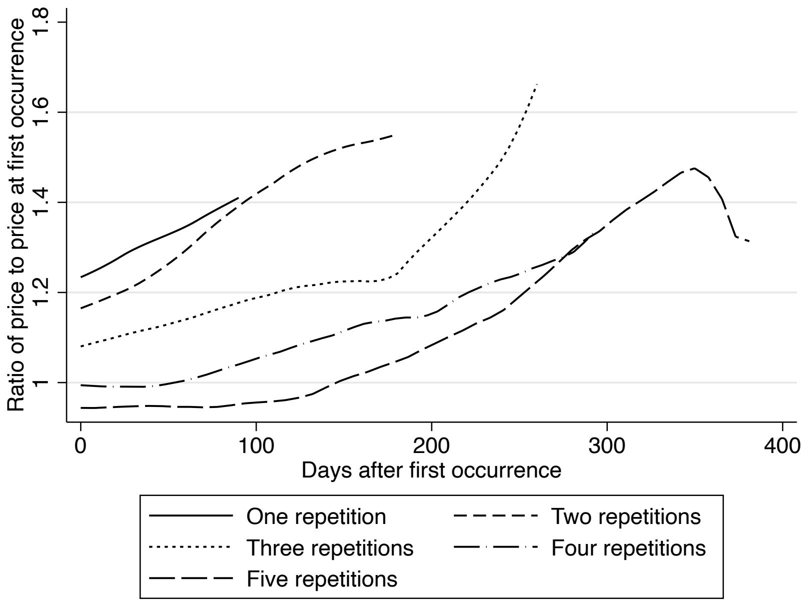
Finally, after merging the properties to the assessment data, there are a few properties with multiple units per unique property ID. These are for example two individual apartments treated as one. Since it is not clear how to aggregate hedonic variables across these, they are dropped. This removes 0.8 million transactions bringing the total count to 92.6 million.

B. Identifying Property Locations

We next describe how we define property locations. The 92,639,072 transactions that are defined as arm's length above are based on 53,164,562 properties.

[61]Since the window is rolling, if there are multiple repeated sales a spell of repeated sales can extend beyond 90 days, and the last observation of the entire spell is retained. A property can have multiple spells through time.

Figure A.4: Repeated sales within a 90 days rolling window and sales price



Notes: The figure plots the smoothed average ratio of the sales price to the sales price at the first occurrence of a spell of repeated sales. The average ratio is the exponentiated average of the log ratios to account for the non-linear scale of ratios. One particular property can have multiple spells of repeated sales. A transaction belongs to a spell of repeated sales if the previous transaction was up to 90 days ago. It is a rolling window, so a spell can extend beyond 90 days from the first transaction if there are multiple transactions in a spell. The figure plots five separate graphs by the number of repeated sales in a spell, e.g. "one repetition" indicates a spell of two transactions, and can therefore only be up to 90 days. Plotted is a kernel smoother with a triangular kernel and a bandwidth of 30 days. The graphs show that the later transactions in a spell of transactions are on average higher than the previous transactions. The graphs for more than five repetitions look similar and become more noisy due to fewer spells with highly repetitive sales.

The property geolocation and address is provided in the transaction, assessment and historical assessment tables. To evaluate the quality of the provided latitudes/longitudes and addresses, we have drawn a sample of 10,000 properties and geocoded the provided addresses with ESRI based on the provided address. For 95%, the newly geocoded location is less than 160 meters away from the original lats/lons, and for 99% it is less than 1400 meters away. One discrepancy, for example, arises in rural areas, where the geocoded ESRI coordinates are at the street entrance of the property, while the Zillow coordinates are sometimes on the property itself. In the few cases with a large distance between original lats/lons and the geocoded ones, the original lats/lons are closer to a third set of coordinates derived from Google Maps. The ESRI coordinates are slightly closer to the lats/lons from the transaction tables than to those in the assessment tables for the 3.1% when they do not match exactly. Furthermore, in the cases where the zip code from the transaction and assessment tables disagrees (0.8% of times), the transaction zip code matches the Google Maps zip code much more frequently (85%).[62]

[62]The disagreement between assessment and transaction coordinates and zips is scattered across all states and years.

We construct the set of lats/lons, zip codes and street addresses in five steps. First we take the lats/lons from the transaction tables, which are available in 97.5% of the cases (we do the same steps for zip codes and addresses).[63] Second, we complement missing ones from the assessment tables which adds 0.4 percentage points to the lats/lons. As a third step, we complement the missing values with the historic information, preferring the most recent non-missing values which adds another 0.7 percentage points to the lats/lons. Of the 53,164,562 properties with arm's length transactions, there are non-missing coordinates for 98.6% (52,443,223), non-missing zip codes for 99.8% (53,066,792), non-missing addresses for 97.3% (51,737,628).

As a fourth step, we ensure the quality of the existing lats/lons by calculating the distance to the official TIGER county boundaries. If the counties in the Zillow data match the TIGER counties (distance is zero) they pass our quality test. The existing lats/lons also pass the quality test if the distance to the matching counties is less than 1km. Manual inspection shows that the shape files at the county boundaries can be imprecise (i.e. in the case of a winding road at the border), and that the lats/lons are actually in the correct county. For the lats/lons that do not pass our quality test, we set them to missing and pass them to the next geocoding step. This adds 47,720 properties to the 721,339 properties with missing coordinates. In total, for the 769,059 properties with missing coordinates, we have 51.8% (398,378) with non-missing address and zip code, 3.9% and 38.5% with only address and zip respectively, and 5.8% without address or zip code.

To ensure a high quality of geocoding, we only geocode the properties with existing addresses and zip codes in the fifth step using ESRI Streetmap Premium.[64] A few of the geocoded properties have non-matching geocoded counties and original Zillow counties. We only use the geocoded coordinates for matching counties and where the ESRI score is high (≥80%), which is 91.2% of the 398,378 properties. With reverse geocoding, we retrieve missing addresses and zip codes from existing coordinates. We set the location of 912 properties to missing where the reverse geocoded counties do not match existing Zillow counties.

The final share of properties with non-missing coordinates is 99.0% (52,612,606), corresponding to 92,006,045 transactions. The share of properties with non-missing addresses is 98.7%, and the share with non-missing zip codes is 99.8%.

C. *Identifying Square Footage of the Property*

We next identify the size of the property in square footage and link it to our 92,006,045 transactions based on our 52,612,606 properties for that we also have coordinates from the previous section.[65]

[63] For multiple addresses per property for different transactions, we keep the longer street addresses, after cleaning upper/lower cases and spaces.

[64] We feed in the addresses and county names as the county identification should be the most reliable data because the raw data is obtained from the individual counties.

[65] Around 0.1% of these cannot be matched to the assessment tables. These missing properties are missing across states and years, and are not just concentrated in recent years. The ca. 50 million properties are a third of the 150 million

Due to the last step of identifying the arm's length transactions, all properties are single unit properties.[66]

There are several different types of building areas that define the size of the property. Some refer to total areas such as "Living Building Area" (BAL), "Gross Building Area" (BAG) or "Total Building Area" (BAT), and others refer to parts, such as "Balcony/Overhang", "Basement", "Porch". The coverage on the total areas is much better than on the individual parts. Each property can have multiple building area types, referring e.g. to the balcony area and the total area. According to Zillow, the "Living Building Area" is usually taken as the property area. While it has the lowest number of missing observations of all types of areas, it is still only available for 66.2% of arm's length properties in the assessment tables.

Before proceeding, we ask whether the missing data comes from particular counties or states. We calculate the share of properties with non-missing "Living Building Area" (BAL) information both within counties or within states. There are many counties that do not report the BAL for any property, so there seems to be little selection within counties. This is the main driver for the missing information. There are some states (e.g. Illinois) in which less than 40% of properties have information on the BAL.

We next supplement the 66.2% of nonmissing observations of BAL. As a first step we complement this data with information from historical versions of the assessment tables. This increases the share of non-missing "Living Building Area" to 73.6%.[67]

As a second step, we further impute the missing values of BAL by taking the other total area types into account – Total, Base, Finished, and Gross Building Area. Importantly, for 84.4% of properties, we have at least one area type reported. We impute the missing BALs, by taking one of the other codes adjusted by the median ratio between BAL and the other code.[68] We therefore recover square footage for 84.4% of the properties, corresponding to 80,618,103 of our arm's length transactions. Overall, we have both square footage and coordinates for 44,799,731 (84.3%) of our arm's length properties and 80,544,782 (86.9%) of our arm's length transactions.

properties in the raw assessment tables. For the other 100 million properties, there are no arm's length transactions recorded.

[66] There are sales prices available in the assessment tables as well, but it is recommended to avoid them, as Zillow notes: "Generally, you can think of the data in ZAssessment tables as data sourced ultimately from county's assessor's offices and ZTransaction tables as data ultimately sourced from legal recordings processed by each county recorder's offices. These are usually two separate agencies in the county administration. The Assessor's office tracks many things, like property attributes, completely independently from the County Recorder's office. However, when the County Assessor reports sale prices on homes (the SalesPriceAmount variable in the ZAssessment tables), this is data that the county assessor's office has taken from the recorder's office and blended into their data set before they sent it to us. Some counties will do this to use the most recent sales prices in their assessment amount models. That being said, we've found that the transaction data we get through assessors tends to be marginal and not always up to date, so when available, use the transaction data reported in the ZTransaction tables."

[67] The most recent available historical information is used for each individual property and area type to replace missing values.

[68] We use the other codes sequentially in the following order: BAT, BAG, BAF, BAB. The median and interquartile ranges of the ratios are unity except for BAG, where the median is 1.2.