

NBER WORKING PAPER SERIES

RACIAL DIFFERENCES IN THE TOTAL RATE OF RETURN ON OWNER-OCCUPIED
HOUSING

Rebecca Diamond
William F. Diamond

Working Paper 32916
<http://www.nber.org/papers/w32916>

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

We thank our research assistants Alex Almeida, Sam Asher, Marie Kaul, Weldon Lin, Alexis Payne, and Gustav Yang for outstanding work. We also thank Amir Kermani, Matthijs Korevaar, Mark Garmaise, and Kelly Shue for excellent discussions. Rebecca Diamond acknowledges support from the Stanford Graduate School of Business, the National Science Foundation (CAREER Grant 1848036). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Rebecca Diamond and William F. Diamond. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Racial Differences in the Total Rate of Return on Owner-Occupied Housing
Rebecca Diamond and William F. Diamond
NBER Working Paper No. 32916
September 2024
JEL No. G50, R30

ABSTRACT

We quantify racial differences in the total rate of return on housing in a nationally representative sample of homeowners from 1974-2021. We develop a new method to estimate the rental value of each owner-occupied house, using a house's resale value to proxy for unobservable quality. Black and Hispanic homeowners earn higher but more volatile rates of return than White homeowners, due in part to higher rental yields. These differences are largely explained by household income and education differences. Including returns on properties that do not transact in sample is crucial to accurately measure racial differences in total rates of return.

Rebecca Diamond
Graduate School of Business
Stanford University
655 Knight Way
Stanford, CA 94305
and NBER
diamondr@stanford.edu

William F. Diamond
The Wharton School
University of Pennsylvania
3620 Locust Walk
Philadelphia, PA 19104
diamondw@wharton.upenn.edu

A data appendix is available at <http://www.nber.org/data-appendix/w32916>

The largest single component of most households’ wealth is their home [Campbell, 2006]. The rate of return earned on an investment in housing is therefore a key determinant of a household’s ability to accumulate wealth. Unlike purely financial assets, an investment in housing provides a flow of consumption services (the ability to live in or rent out a home), financial costs such as taxes, and potential capital gains from the increase in a home’s price. This paper develops a new method to infer the rental value of owner-occupied housing and uses this method to measure the total rate of return homeowners get from a housing investment, inclusive of a home’s rental value, property taxes, and maintenance and home improvement costs. We apply this method to studying disparities in the total returns on housing earned by homeowners of different racial groups. We find that Black and Hispanic homeowners earn higher but more volatile returns than White homeowners, primarily due to higher rental yields on the homes they buy.

Racial minorities have a disproportionate share of their wealth in housing (59% for Black vs 38% for White), so the return on housing plays a key role in the evolution of the racial wealth gap [Derenoncourt et al., 2023, Barsky et al., 2002, Kuhn et al., 2020]. In addition, the Black homeownership rate is roughly half the White homeownership rate [Gupta et al., 2023, Charles and Hurst, 2002], motivating interest from policymakers in increasing minority homeownership. Consistent with previous work [Kermani and Wong, 2022, Akbar et al., 2022, Wolff, 2022, Kahn, 2024], we find a lower rate of house price growth for Black homeowners than White homeowners. However, we also find that Black homeowners earn higher rental yields. On net, average racial disparities in the total return on owner-occupied housing push towards narrowing the racial wealth gap.

Using data on a panel of homes from the American Housing Survey (AHS) from 1974 to 2021, we observe a subset of homes switching between the rented and owner-occupied segments of the housing market. This allows us to observe the rental value and sale price of the same home and estimate an empirical relationship between rents and house prices. We therefore use a house’s resale price as an additional predictor of its rental value beyond the information contained in the house’s observable characteristics. We develop a method that uses the information in house prices to accurately predict the rents of both low-value and high-value homes. We also apply a version of our method to a much larger administrative database of housing transactions and rentals from Corelogic, finding similar results.

Our rental imputation method integrates two sub-literatures on measuring the return on housing. One branch attempts to measure the total return to landlords by jointly observing rents and prices on the same housing units [Chambers et al., 2021, Eichholtz et al., 2021], without attempting to extrapolate to the owner-occupied housing stock. Another branch [Demers and Eisfeldt, 2001, Gorback and Schubert, 2024, Halket et al., 2024, Gilbukh et al., 2023, Colonnello et al., 2024] imputes rental yields using hedonic house characteristics and/or aggregates data at a local level. Our innovation is to use information on an individual house’s price to more accurately impute its rental value. As well as providing a new approach to measuring the total return on owner-occupied housing, our method could also be applied in other rent imputation contexts, such as the role of owners’ imputed rent in computing CPI and GDP statistics [Crone et al., 2009, Garner and Short, 2009].

We apply our method to measuring differences in the rental yield of owner-occupied housing across racial groups using our AHS data. We find Black homeowners’ average rent-to-price ratio is 8.86%, Hispanic homeowners’ is 8.36%, and White homeowners’ rent-to-price ratio is 7.83%. In a regression with year fixed effects, we find a Black-White rent-to-price ratio gap of 1.2 percentage points(pp) and Hispanic-White gap of 0.79 pp. These rental yields are inferred from a model fit to AHS data on houses that switch between being rented and owner-occupied. In the raw data on these “switchers,” we find an average rent-to-price ratio of 9.91% for Black homeowners, 9.42% for Hispanic homeowners, and 9.04% for White homeowners. The lower yields we find for the overall owner-occupied housing stock than these switcher homes in particular reflects the fact that a house’s rent-to-price ratio tends to decrease with its price. Because owner-occupied homes tend to be more expensive than observably equivalent rental homes, our rent imputation procedure is ideal for capturing these quality differences reflected in a home’s price.

We find independent evidence of higher rental yields for Black and Hispanic homeowners than White homeowners in two additional data sources. First, we examine direct evidence in the Consumer Expenditure Survey (CEX), which surveys homeowners both about the value of their homes and about the rent they believe they could earn on their home. We find qualitatively similar but somewhat larger racial disparities in rental yields in the CEX than the AHS. In addition, we independently estimate the rental yield of owner-occupied housing using administrative data from Corelogic, covering 2007-2017. Here, we jointly use data on housing purchases and sales matched to rents from Multiple Listing Services (MLS) to observe a population of “switchers” with both rent and price observations in the subset of regions where the data is available. After imputing rental yields for owner-occupied homes in this much larger dataset, we find similar racial disparities in rental yields as a subsample from the AHS that matches the geography and time period of our administrative data. In regressions with equivalent year/time fixed effects in our two datasets, we find a 1.67 pp rental yield gap between Black and White homeowners in Corelogic and a 1.47 pp gap in the AHS. We similarly find a 1.0 pp Hispanic-White gap in Corelogic and 0.92 pp Hispanic-White gap in the AHS. In addition, we show that roughly 30% of these racial disparities cannot be accounted for by zipcode-year fixed effects and would be missed if we aggregated our rental yield estimates at a zipcode level [Demers and Eisfeldt, 2001].

Next, we use our rental yield estimates to analyze racial disparities in total rates of return on housing. We find in our benchmark result with AHS data that the average log return on owner-occupied housing is 7.2% for Black homeowners, 7.9% for Hispanic homeowners, and 6.5% for White homeowners from 1974 to 2021. This gap is not driven by differences in house price appreciation. We find White homeowners earn 0.9% annual real appreciation rates, while Black homeowners earn 0.7%, consistent with prior results of a negative Black-White house price appreciation gap. We find Hispanics earn a higher average appreciation rate of 1.6%. The overall rate of return for Black homeowners is higher due to a 0.7 pp higher dividend yield than for White homeowners. Hispanic homeowners earn a 0.6 pp higher dividend yield than Whites. These dividend yield differences are lower than the rental yield differences since we find that Black and Hispanic homeowners spend more on home maintenance relative to the value of their

homes than White homeowners. We verify that the racial disparities in property taxes we measure are consistent with [Avenancio-Leon and Howard \[2022\]](#), and these differences have a small impact on the total rate of return on housing. Our average rate of return is similar to the housing return in [Jorda et al. \[2019\]](#).

We use our rich AHS microdata to analyze the determinants of these rate of return differences at the household level. We regress the log rate of return on each house on the homeowner’s race together with a growing list of control variables. After controlling for homeowner education and income, we are able to fully explain the average racial return differences with observable characteristics. We find that income/education is the most useful explainer of the Black-White total return gap, and both geography and income/education help to explain the Hispanic-White total return gap. Our evidence is consistent with lower income homeowners buying lower value homes, which in turn have higher average rates of return due to their higher rental yields. The racial disparities in housing returns we measure can mostly be accounted for by disparities in economic status across racial groups.

A second key difference in the total return on housing for Black and Hispanic homeowners relative to White homeowners is greater volatility and greater sensitivity to the business cycle. We find that both Black and Hispanic homeowners have particularly volatile and cyclical returns. Unlike the disparities in average returns, we find that observable differences between racial groups do not entirely explain the excess exposure of minority housing returns to the business cycle. Controlling for geography and income/education largely explains the White-Hispanic difference in return cyclicalities, but observables poorly account for lower returns earned by Black homeowners in recessions. We consider the contribution of both rental yields and price changes to the cyclicalities of returns. Even after accounting for observables, we find that Black and Hispanic homeowners have both rental yields and price appreciation that are more cyclical than for White homeowners. This is broadly consistent with the finding [[Kermani and Wong, 2022](#)] that the risk in minority housing returns is largely due to distressed sales during downturns, based on more severe financial constraints [[Gupta et al., 2023](#), [Bhutta and Hizmo, 2021](#)] for Black and Hispanic homeowners.

Our paper contributes to the literature on disparities in housing returns by the use of nationally representative, household-level data that allows us to measure the key determinants of housing returns over 45 years. Much of the previous literature on racial disparities in housing returns uses geographically aggregated data, with [Kermani and Wong \[2022\]](#) as a notable exception. Our AHS data has costs and benefits compared to the large transaction-level data used in [Kermani and Wong \[2022\]](#), and the final section of our paper reconciles our results. Their baseline analysis uses realized housing transactions, which more accurately measures the returns on transacted homes but does not include returns for the majority of houses that do not transact within their sample. Our data provides a housing value even if there is not a transaction, but it may miss some of the financial distress they observe in sales prices. We explore the robustness of our results to potential mismeasurement of distressed sale prices in section 7 and find these effects to be small. The fact that we can measure the return on non-transacted properties, instead of conditioning on a selected sample of transactions which must include all foreclosures, decreases

the sensitivity of our results to foreclosures. When we attempt to fill in the return on non-transacted properties in Corelogic transaction data, we find similar racial disparities in house price appreciation as we do in the AHS. If we make the conservative assumption that our results in the AHS entirely miss the impact of foreclosures, we show that Black and Hispanic homeowners still earn a higher total rate of return than White homeowners after we subtract a measure of foreclosure discounts from [Kermani and Wong \[2022\]](#).

In total, our paper demonstrates that it is crucial to account for the total return on housing, not just changes in house prices, when measuring racial disparities in housing returns. Our new method for imputing the rental value of owner-occupied housing may therefore be a broadly useful tool for this literature. Because lower value homes tend to have the highest rental yields, any method of measuring rates of return that ignores rents will understate the returns on low-value homes and overstate the returns on high-value homes. Because the average value of homeowners' residences varies by racial group, an accurate rental yield calculation is crucial for understanding racial disparities in the rate of return on housing. Together with other evidence on racial disparities in liquidity constraints [[Ganong et al., 2024](#)], home-ownership [[Charles and Hurst \[2002\]](#)], and credit market outcomes [[Indarte et al., 2023](#)], an accurate total rate of return measure for housing is key for understanding the future of the racial wealth gap.

1 Conceptual Framework

Unlike financial assets such as stocks or bonds, housing is both an investment good and a consumption good. Because a homeowner can live in their house, investing in housing generates a flow of consumption services as well as a possible increase in the value of the investment. To calculate the financial return to owning a home, we compare the financial benefit of homeownership to renting an identical home in a given year t . If a household were to rent house j in year t , it would pay r_{jt} . If the household were to purchase the home, it would pay price P_{jt} , spend $main_{jt}$ on maintenance, and tax_{jt} on property taxes and then receive P_{jt+1} , ignoring transaction costs.

The log financial return to owning house j , relative to renting it is:¹

$$\log(P_{jt+1}) - \log(P_{jt} - r_{jt} + tax_{jt} + maint_{jt}). \quad (1)$$

This return quantifies the *financial* benefit a household receives by owning a given house, instead of renting it. It is possible that if a homeowner switched to renting, it would prefer to rent a different home. While it would be more complicated to explicitly quantify homeowners' utility or welfare, the financial return we measure in equation 1 tells us the impact of homeownership on a household's wealth, compared to a counterfactual of paying rent for an identical home.

Taxes and maintenance (for which we include both routine maintenance and home improvement) in equation 1 are cashflows that can directly be observed in our AHS data. However, the rent r_{jt} for an

¹The "dividend" component of the housing return, $r_{jt} - tax_{jt} - maint_{jt}$, we will assume is received at time t . One could also write the rate of return calculation with dividends paid at time $t + 1$. While both of these are conceptually very similar, we have much better data coverage for all of these variables at time t than $t + 1$, since some houses that are owned at time t are not still owner occupied at $t + 1$.

owner-occupied house is never explicitly observed. As a result, much of the literature on inequality in returns has simply used the price appreciation $\log(\frac{P_{t+1}}{P_t})$ to measure housing returns. To include all of the relevant cashflows to measure housing returns, we develop and apply a new method to impute the rental value of owner-occupied housing.

2 Data

The bulk of our analysis uses the American Housing Survey (AHS) from 1974-2021. This is a nationally representative survey that is a panel at the housing unit level. From 1974-1981, the survey was done annually, but it switched to every other year in 1983. The AHS interviews the residents of housing units about many aspects of the house itself, many details of their housing costs, and about the housing residents themselves. This survey is ideal for calculating rates of return on housing since it collects raw microdata on housing maintenance, home improvements, property taxes, property values, property rents, and property characteristics. Resident race, which papers using housing transaction data often have to impute from external sources, is also directly reported. The main concern with using survey data is the possibility of measurement error in these variables, as compared to administrative data on property transactions. We investigate these differences in Section 7. Two clear benefits of performing this analysis in the AHS are the very long time series back to 1974 and the fact that the sample is nationally representative. Most datasets of administrative housing transactions begin in 1990s or 2000s [Kermani and Wong, 2022, Goldsmith-Pinkham and Shue, 2023] and exclude regions that do not publicly report housing transaction prices. In addition, the AHS allows us to calculate rates of return for houses that do not sell in sample, while administrative transaction data is restricted to properties that sell in sample. We will investigate all of these differences between the AHS and administrative house transactions data in section 7.

Table 2.1 reports summary statistics on key housing costs in 2019 dollars in the AHS by race. Black homeowners have lower average property values, rents, home improvement spending, and property taxes than White homeowners, but they have similar spending levels on routine maintenance. Hispanic homeowners have similar housing values to Whites, but they have higher spending on routine maintenance and lower spending on home improvement and property taxes.

Table 2.1: Summary Statistics for Owners and Renters in the AHS by Racial Group, 1974-2021

	Num. Obs.	Mean	Median	Std. Dev.
White				
Property value	711,651	239,470	175,780	258,446
Rent (monthly)	285,090	885	757	669
Routine maintenance (annual)	710,935	767	568	1,121
Home improvement (annual)	710,934	2,380	982	8,095
Property tax (annual)	710,933	2,782	1,949	3,435
Black				
Property value	69,374	168,834	119,530	207,478
Rent (monthly)	87,059	729	658	523
Routine maintenance (annual)	69,225	765	556	1,136
Home improvement (annual)	69,225	1,693	521	8,843
Property tax (annual)	69,225	1,985	1,294	2,744
Hispanic				
Property value	54,813	242,293	175,264	258,819
Rent (monthly)	65,810	930	830	628
Routine maintenance (annual)	54,695	880	570	1,296
Home improvement (annual)	54,695	2,117	384	7,428
Property tax (annual)	54,695	2,663	1,857	3,353

Notes: Data from the 1974-2021 AHS. Summary statistics for all data except rent are from owner-occupied homes. Rent is from rented homes.

3 Imputing the Rental Value of Owner-Occupied Housing

One of the most important components of the return on investing in housing is the rental income that a house generates. For a landlord, this is a cash flow paid explicitly by the tenant to the landlord. For a homeowner, the “owner’s equivalent rent” measures the market value of the housing services the household consumes by living in that home for a period of time. Because homeowners do not explicitly pay or receive rent, it is necessary to impute this owner’s equivalent rent to accurately measure their investment returns.

We develop a new method of imputing owner’s equivalent rent that allows for differences in quality across houses that are not captured by their observable characteristics. A house’s quality depends on its characteristics X_h , including both observable components X_h^o and unobservable components X_h^u . As shown in [Bajari and Benkard, 2005, Bajari et al., 2005], even when preferences are heterogeneous across households, there exists an equilibrium pricing function that maps house characteristics X_h (this includes physical and location characteristics) to equilibrium prices in the market. For house h with characteristics X_h at time t , the rent r_{ht} and purchase price p_{ht} can be written as $r_{ht} = r_t(X_h)$ and $p_{ht} = p_t(X_h)$, respectively. While we assume households can observe all characteristics X_h of a house, our data only include the subset X_h^o . Without loss of generality, we can write our rent and price functions as $r_{ht} = r_t(X_h^o, X_h^u)$ and $p_{ht} = p_t(X_h^o, X_h^u)$.

The key insight to our method is that the price a house would sell for in the purchase market is a proxy for the unobserved house characteristics, given the observed house characteristics. For example, consider two different 1400 square foot, 2 bed, 2 bath houses built in 1974 located in a given neighborhood. If one would sell for 10% more than the other, despite the observed characteristics being identical, the house

with the higher sale price would likely also rent for more, although the gap in rental rates might be more or less than 10%. To formalize this intuition, we make the following key single index assumption.

Assumption 1 (*Single Index*)

There exists at each time t a function f_t that maps the set of unobserved property characteristics X_h^u to a scalar such that the rent and price functions r_t and p_t can be written as a function of the observed house characteristics X_h^o and $f_t(X_h^u)$ where p_t is strictly monotone in $f_t(X_h^u)$, given X_h^o . E.g.:

$$\begin{aligned} r_{ht} &= r_t(X_h^o, X_h^u) \\ &= r_t(X_h^o, f_t(X_h^u)) \\ p_{ht} &= p_t(X_h^o, X_h^u) \\ &= p_t(X_h^o, f_t(X_h^u)). \end{aligned}$$

Under Assumption 1, we can invert the p_t function: $f_t(X_h^u) = p_t^{-1}(X_h^o, p_{ht})$ and plug the inverted price function into the rent function:

$$r_{ht} = r_t(X_h^o, X_h^u) \tag{2}$$

$$= r_t(X_h^o, p_t^{-1}(X_h^o, p_{ht})) \tag{3}$$

$$= r_t^*(X_h^o, p_{ht}). \tag{4}$$

Equation 4 shows that r_{ht} can be written as a flexible function of observed house characteristics, X_h^o and the purchase price of the home, p_{ht} .

To the extent that the purchase price of a house is predictive of its rental value even after controlling for observables X_h^o , a clear bias appears in any procedure that only uses X_h^o to predict rent. For houses with sufficiently low prices, their rent will generally be below $E(r_h|X_h^o)$, so the procedure overestimates the house's rent. Similarly, for houses with high enough prices, $E(r_h|X_h^o)$ is an underestimate of the houses' rent.

Because minority homeowners tend to own lower-price houses than White homeowners, this argument suggests that using only observables to predict a house's rent would induce bias into our calculation of race-specific rates of return on housing. As we show empirically in appendix G, the rental yield earned by minority homeowners would be artificially inflated compared to the rental yield earned by White homeowners if we ignored unobservable differences between homes. As we show empirically in section 4, the procedure we develop next accurately measures average rents for both low- and high- price homes and is therefore ideal for comparing rates of return on housing across racial groups.

3.1 Inferring rental yields using switchers in the AHS

Our baseline rent imputation procedure requires data on a panel of houses, as is provided by the AHS. Each house h at time t is either vacant, rented, or owner-occupied. For rented houses, we observe the rent r_{ht} . For owner-occupied houses, we observed the price $p_{h,t}$. The key difficulty we face is that in each

period at most one of r_{ht} and p_{ht} is observed. In addition, we see a vector of house characteristics X_h^o . Our goal is to produce an estimate $\hat{r}_t(X_h^o, p_{ht})$, the rent that could have been obtained for owner-occupied house h at time t .

Equation 4 implies that if we observed the sale price of a house along with observed property characteristics, we should be able to *perfectly* predict its rental price. Since the data are unlikely to provide a perfect rent forecast, we augment equation 4 to include an error term:

$$r_{ht} = r_t^*(X_h^o, p_{ht}) + \epsilon_{ht}. \quad (5)$$

Equation 5 reflects the possibility of measurement error in rents, or that our single index assumption above does not hold perfectly. See appendix A for a precise statement of our necessary assumptions to consistently impute rents, which do not require a perfect single index. For a given set of houses with identical (X_h^o, p_{ht}) , we should expect to observe a distribution of rents if ϵ_{ht} has nonzero variance.

Our rent imputation procedure relies on the population of “switcher” houses that switch between owned and rented status between times t and $t + 1$. We only can see data for both the price and rent of a single house if that house is a switcher. However, only a small fraction of the houses in our sample at any given time are switchers. In theory, our model suggests that we should estimate the relationship between rents and prices year-by-year, since the hedonic price of a given house changes over time, but we would under-powered due to the low sample size. We therefore develop methods that enable us to pool the sample of switcher houses across many years of our data.

To do so, we transform the relationship our single index provides between rents and prices into one between rent and price ranks. Let ρ_{ht}^r and ρ_{ht}^p respectively be the ranks of house h ’s rent and price at time t in the nationwide rent and price distributions. If R_t and P_t are respectively the CDFs of the national rent and distributions at time t , we have without loss of generality

$$r_{ht} = r_t^*(X_{ht}^o, p_{ht}) \quad (6)$$

$$\rho_{ht}^r = R_t(r_t^*(X_{ht}^o, P_t^{-1}(\rho_{ht}^p))) \quad (7)$$

$$\rho_{ht}^r = r_t^{*,\rho}(X_{ht}^o, \rho_{ht}^p). \quad (8)$$

This gives us a mapping between the rent and price rank of a given house h at a time t . This mapping is constructed by composing the CDF of the time t rent distribution, our function r_t^* at a given value of observables X_h^o , and the inverse CDF of the time t house price distribution. We then assume that this $r_t^{*,\rho}$ is the same function $r^{*,\rho}$ at all times t . This allows us to estimate a relationship

$$\rho_{ht}^r = r^{*,\rho}(X_h^o, \rho_{ht}^p) \quad (9)$$

between rent and price ranks using data from all times t .

This assumption requires that among houses with the same observables, a given price rank ρ_{ht}^p maps to the same rent rank ρ_{ht}^r , even if the houses are observed in different years. In this relation, these ranks

are computed relative to year-specific nationwide distributions. For example, consider all 2-bed 2-bath houses in the years 2000 and 2010. If a house in 2000 has the median price rank and 70th percentile rent rank relative to homes in 2000, it must also be true that the house in 2010 with the median price rank has the 70th percentile rent rank relative to homes in 2010. However, if prices increase and rents fall from 2000 to 2010, a home in 2010 could have a higher price and lower rent than another home in 2000.

We believe this assumption that ranks have a stable relationship over time is more likely to hold than a stable relationship between rents and prices themselves. First, the overall rent-to-price ratio for houses has decreased over time [Davis et al., 2008], similar to the decline in interest rates on bonds since the 1980s. Second, we provide evidence in section 3.3 that the empirical relationship between rank ranks and price ranks is similar across years, while the empirical relationship between dollar rents and dollar prices is not.²

3.2 Estimation

Before turning to our rent estimation procedure, we present some descriptive statistics in Table 3.1 on houses that switch between the owner-occupied and rental markets and compare them to units in the overall rental market and owner-occupied market by race. Among White-, Black-, and Hispanic-occupied units, we see the average units that switch between markets are the 42nd, 33rd, and 43rd percentile of the property value distribution, respectively. These switcher units are negatively selected from the overall owner-occupied distribution, where the average property values for White, Black, and Hispanic homeowners are at the 51st, 36th, and 48th percentile of the property value distribution. However, these switchers are positively selected from the rental market distribution. The average White-, Black-, and Hispanic-occupied rentals are at the 51st, 41st, and 52nd percentiles, respectively, while the switcher units are higher at the 56th, 48th, and 56th percentiles. Taken together, this implies that the average rental unit is of lower value than the average owner-occupied unit, highlighting the importance of using property prices to proxy for unobserved quality differences between owned and rented units. These relative differences in quality between owned, rented, and switcher units are also apparent when looking at differences in number of rooms, number of bedrooms, number of bathrooms, and age of the unit.

²One additional detail we abstract from here is that for a given house, we observe its rent and its price at different points in time. We state the precise assumption we need to cover this case, together with the possibility of measurement error, in appendix A.

Table 3.1: Property Characteristics by Switcher, Owner-occupied, and Renter-occupied

(1) White	Switchers	Owners	Renters
Number of Rooms	5.53 (1.59)	6.24 (1.66)	4.36 (1.49)
Number of Bedrooms	2.66 (0.91)	3.00 (0.87)	1.96 (0.95)
Number of Bathrooms	1.46 (0.60)	1.68 (0.67)	1.22 (0.46)
Age of the Building	36.06 (20.52)	31.76 (20.10)	34.66 (20.80)
Number of Units in the Building	1.81 (2.30)	1.23 (1.28)	4.54 (3.83)
Price Quantile	41.56 (28.78)	50.94 (28.53)	
Rent Quantile	55.67 (30.69)		51.48 (28.98)
Num. Obs.	31,468	751,606	305,684
(2) Black	Switchers	Owners	Renters
Number of Rooms	5.60 (1.50)	6.14 (1.53)	4.38 (1.36)
Number of Bedrooms	2.79 (0.91)	3.05 (0.84)	2.03 (0.93)
Number of Bathrooms	1.40 (0.58)	1.55 (0.64)	1.18 (0.42)
Age of the Building	41.99 (20.74)	36.34 (20.62)	39.66 (20.26)
Number of Units in the Building	1.82 (2.29)	1.23 (1.23)	4.98 (3.84)
Price Quantile	33.42 (26.74)	36.19 (26.65)	
Rent Quantile	48.27 (29.80)		41.39 (27.74)
Num. Obs.	5,081	68,305	88,456
(3) Hispanic	Switchers	Owners	Renters
Number of Rooms	5.42 (1.45)	5.97 (1.51)	4.31 (1.31)
Number of Bedrooms	2.75 (0.93)	3.04 (0.86)	2.00 (0.92)
Number of Bathrooms	1.51 (0.59)	1.69 (0.64)	1.22 (0.46)
Age of the Building	41.27 (21.09)	36.59 (21.70)	42.55 (20.90)
Number of Units in the Building	1.87 (2.38)	1.29 (1.43)	5.02 (3.86)
Price Quantile	43.20 (29.99)	47.64 (29.32)	
Rent Quantile	56.05 (29.45)		52.37 (27.06)
Num. Obs.	5,128	53,159	67,957

Note: Summary stats of AHS data from 1974-2021. Switchers are housing units that switch between being owner-occupied and renter-occupied between consecutive waves of the AHS. Owners are all owner-occupied units. Renters are all rent-occupied units. We drop observations occupied by other racial groups than Black, White, and Hispanic. Price quantile is the percentile of the house's price within the nationwide distribution of property values observed in the same survey wave. Rent quantile is the percentile the house's rent within the national distribution of property rents observed in the same survey wave.

To implement our rent prediction method, we first need to predict a house's rent rank given its observable characteristics and price rank as in equation 9. To do so, we estimate the regression equation

$$\rho_{ht}^r = \alpha + \beta_1 \rho_{ht}^p + \beta_2 (\rho_{ht}^p)^2 + \beta^o X_{ht}^o + \epsilon_i. \quad (10)$$

Because our predicted rent at a time t plugs this rank prediction into the non-linear function R_t , a mean-unbiased rent rank prediction need not result in a mean-unbiased prediction of rent. As a result, we aim instead to construct a median-unbiased rent prediction. To do so, we estimate equation 10 using a median regression.

Our characteristics X_{ht}^o include building age and its square, number of rooms, bedrooms, and bathrooms, units in the structure and its square, dummies for a townhouse and mobile home and the metro-census region level to account for the finding in Demers and Eisfeldt [2001] that rent-to-price ratios vary systematically across regions at a given time.³ We also include a dummy indicating whether the property was rented in the prior period (and is now owner-occupied), versus previously owner occupied, but now rented. To allow for the possibility that unobserved housing characteristics not captured by price impact rents in a way that is correlated with race, we estimate these regressions separately for each racial group.

Finally, we use a non-parametric estimate of $\hat{R}_t^{-1}$, the CDF R_t of rents at time t . We then simply use the empirical CDF for $p_{h,t}$ as our estimate of the CDF P_t of house prices in a given time t . The imputed rent for each owner-occupied house h is:

$$\hat{r}_{ht} = \hat{R}_t^{-1}(\hat{\rho}_{ht}^r(\hat{P}_t(p_{ht}), X_{ht}^o)). \quad (11)$$

In equation 11, we take a house’s realized price and first compute its empirical rank $\hat{P}_t(p_{ht})$ using the empirical house price CDF $\hat{P}_t$. Next, we use this price rank and the house’s observables to predict its rent rank from our estimate of equation 10. Finally, we plug this estimate of the house’s rent rank into the inverse CDF of rents $\hat{R}_t^{-1}$ to obtain an estimate of the house’s rent. This rent prediction uses both a house’s observables as well as unobservable quality proxied for by its price and is potentially an improvement over hedonic prediction methods that use observables alone.

3.3 Verifying the rank stability assumption

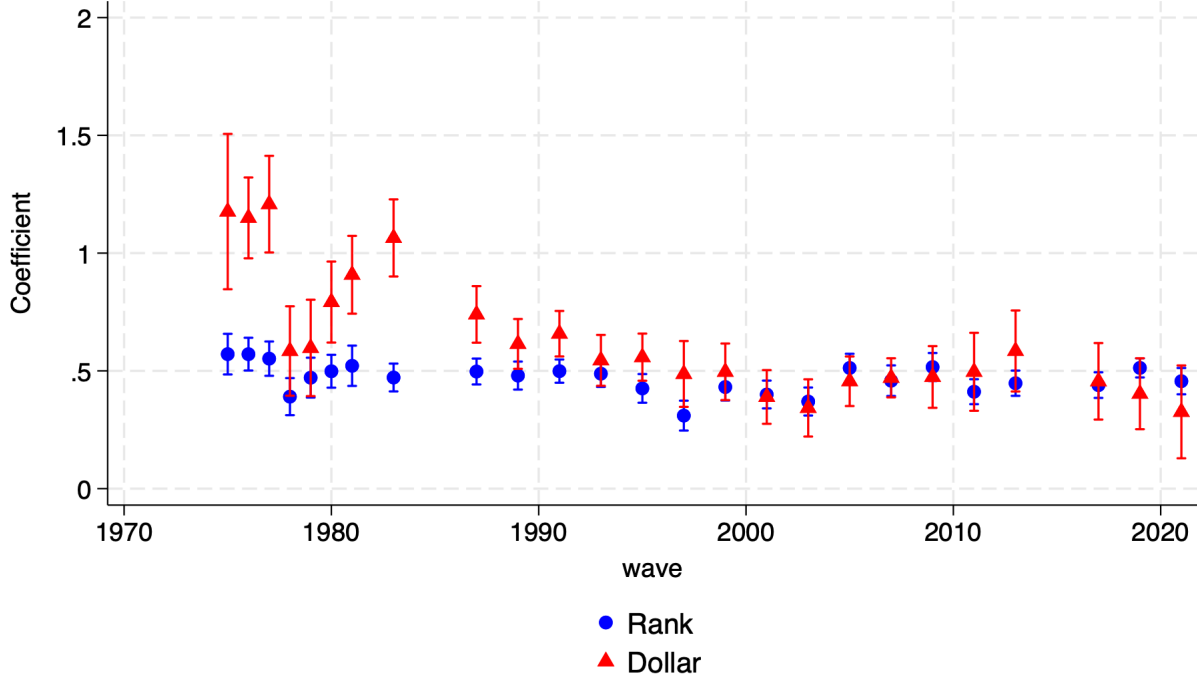
To investigate whether the rank stability assumption behind our rent imputation method is empirically plausible, we estimate the rank-rank regression in equation 10. We run this regression separately year by year on the subset of houses where we see both rent and price. We drop the quadratic term on price ranks and only focus on the linear coefficient as a single summary metric for the relationship between price ranks and rent ranks in each year. For comparison, we also run the same regression, except we regress a house’s dollar rent directly on its dollar house price, instead of using ranks. In doing so, we convert all prices into 2019 dollars. To make the units of these regression coefficients comparable between these methods, we replace each rent and price observation by its Z-score across the entire sample to put things in comparable units. We do this both for the rank measures and the dollar measures.

We plot the coefficients on price (both for ranks and dollars) in Figure 3.1. The relationship between dollar rents and dollar prices is unstable over time, with our regression coefficient decreasing over the sample. This coefficient decrease roughly tracks the aggregate decrease in rent-to-price ratios of houses

³We also include dummies for each variable if it is equal to the top or bottom code of a given continuous variable (e.g. 10+ bedrooms) and dummies for missing values of each of these variables.

over time. In contrast, our regression coefficient in the rank-rank regressions is quite stable over time, suggesting that our rank stability assumption is plausible.

Figure 3.1: Stability of rank-rank versus dollar-dollar relationship between rents and prices



Notes: The figure plots the coefficient and 95% confidence interval of price in equation 10. The blue dots come from a median regression of rent ranks on price ranks, along with controls. The red triangles come from a median regression of rent dollars on price dollars, in units of 2019 dollars, along with controls. Both measures of rents and prices have been z-scored to make units comparable.

3.4 Implied Rental Yields by Racial Group

Table 3.2 reports summary statistics on the annual rent-to-price ratio by racial group from 1974 to 2021. We find Black homeowners' rents are equal to 8.86% of their property value on average. Hispanic homeowners earn rents equal to 8.36% of their property value, and White homeowners earn 7.83% of their property values. Both racial minority group homeowners earn higher rents as a share of their property values than White homeowners. This is also apparent in the raw data on property that switch between the owner and renter markets, as shown in Table 3.2 Black-owned properties that switch into/out of the rental market earn rents equal to 9.91% of their property value, Hispanic-owned switcher properties earn 9.42% of their value, and White-owned switchers earn 9.04% of their value. All of these rent-to-price ratios are higher for the switcher properties than the overall owner occupied housing stock. This reflects the fact that the properties that switch into/out of the rental market are on average lower value than the typical owner-occupied property.

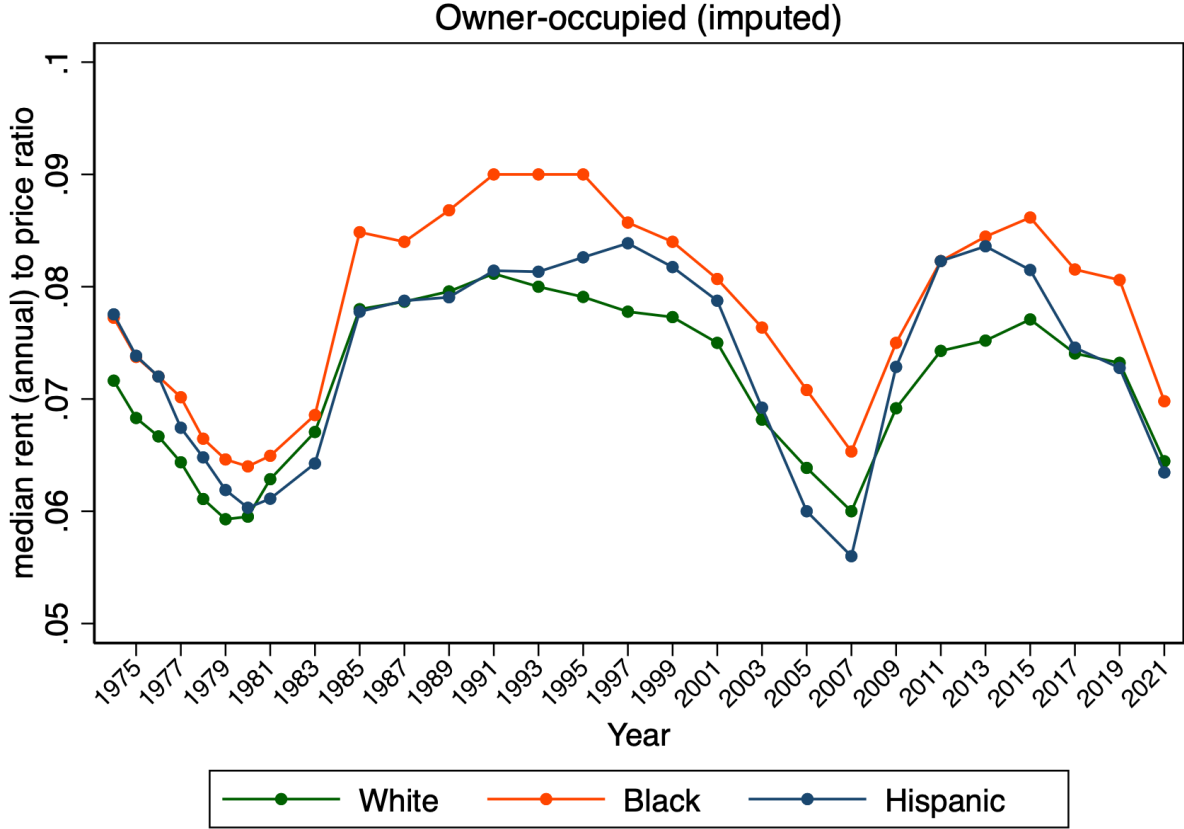
Table 3.2: Rent-to-Price Ratio in the AHS by Owner Race, 1974-2021

Panel A: Switchers	Mean	Median	SD	25th Pct.	75th Pct.	Num Obs.
White	9.04	7.58	6.85	4.90	10.60	21,259
Black	9.91	8.23	7.37	5.10	12.11	4,100
Hispanic	9.42	7.44	7.61	4.56	11.32	4,586
Panel B: Owners	Mean	Median	SD	25th Pct.	75th Pct.	Num Obs.
White	7.83	7.06	4.11	5.52	8.82	697,025
Black	8.86	7.79	4.74	5.82	10.33	66,548
Hispanic	8.36	7.27	4.84	5.26	9.72	52,079

Note: Summary stats on rent-to-price ratios by race. Panel A reports raw data from our switcher sample. Panel B reports our imputed rent-to-price ratios for the owner-occupied sample using our switcher method in the AHS. Data are winsorized at the 2.5th and 97.5th percentiles to limit the effects of outliers.

Figure 3.2 plots the median rent-to-price ratio for homeowners by race over time. We see rent-to-price ratios falling in the 1970s, rising in the 1980s, staying flat in the 1990s, falling in the 2000s, rising in the early 2010s and then falling again until 2021. These patterns are broadly similar for all racial groups. Consistently over time, the Black rent-to-price ratios are about a percentage point higher than the White levels. Hispanic rent-to-price ratios are about 0.5 percentage points higher than White levels on average but are more cyclical and sometimes fall below White rent-to-price ratios.

Figure 3.2: Plot of Median Rent-to-Price Ratio by Owner Race and Year, Imputed Rent for all Homeowners



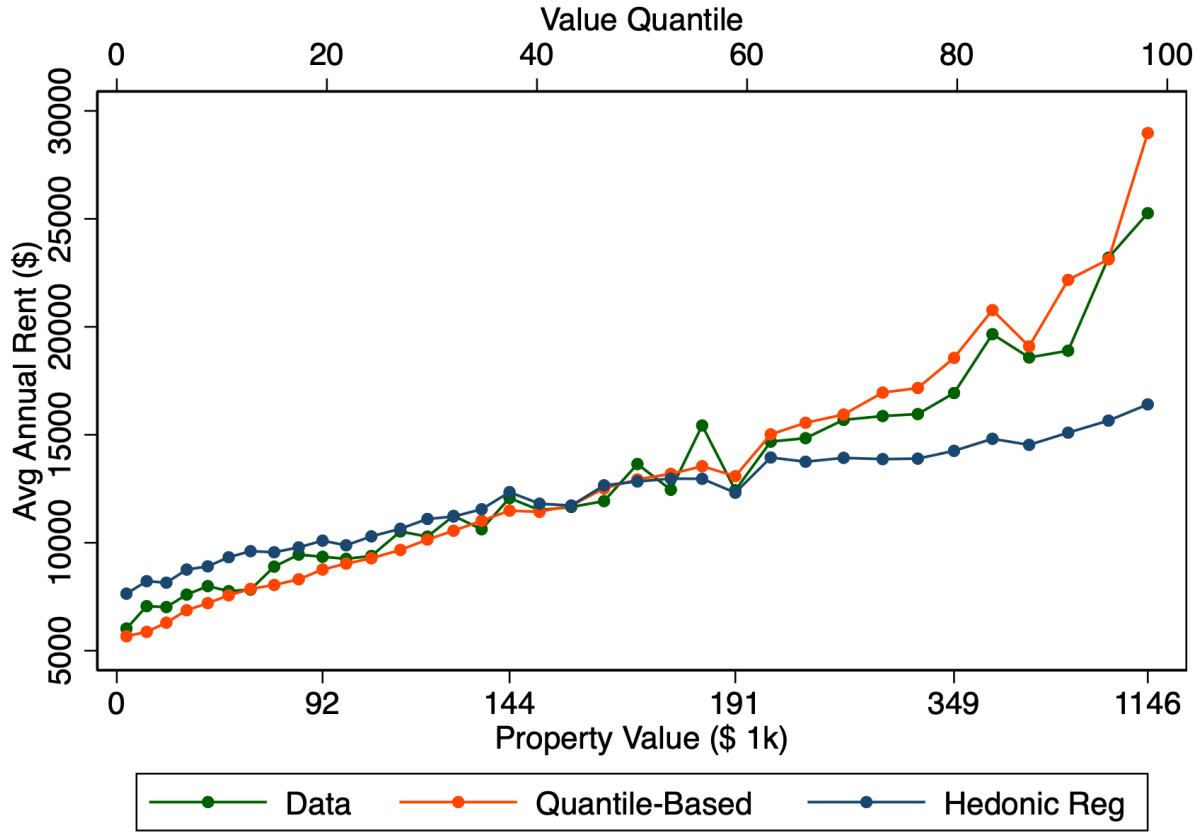
Note: Median annual rent-to-price ratios by race over time for the owner-occupied housing stock, as measured from our switcher imputation method in the AHS.

4 Comparing Rent imputation methods

We now compare our rent imputation method to simpler methods that only use the observable characteristics of a property. If a property has a rent r , observable characteristics X and a market price p , then we expect $E(r|X, p)$ to be increasing in the price p . As a result, $E(r|X)$ should overestimate the rent of low-price properties and underestimate the rent of high-price properties.

To impute a property's rent using only observable characteristics, we run the same median regression as above that we used to study rent ranks, but we switch the dependent variable to rents measured in dollars. We also drop the house sale price as a covariate. This standard hedonic price regression is then estimated on the entire sample of rental properties in the AHS, not just the subset of switchers.

Figure 4.1: Performance of Quantile-based and Hedonic Rent Imputation Methods on AHS Data



Note: This figure tests the results of our imputation methods against a 20% holdout sample of “switcher” houses with both rent and price data in the AHS from 1980 to 2019. All values are in 2019 dollars. The scale on the x-axis is equally spaced by quantiles of property value.

In each year, we use all rented properties in the AHS as our sample except for a random holdout sample of 20% of those properties that switch between being owned and rented. To test our paper’s quantile-based method which uses a property’s price to help predict its rent, we run our specification on the 80% of switcher properties that are not included in our holdout sample. We then test its predictions on our 20% holdout sample and present the results in figure 4.1. This graph shows the mean level of rent for homes of a given price in the raw switcher holdout data, based on the predictions of our quantile-based procedure that uses prices, and based on a hedonic regression that ignores prices.

Unlike our quantile-based procedure, the hedonic rent imputation over-predicts the rents of low-price properties and under-predicts the rent of high-price properties. The orange line representing the average predicted rent from our quantile-based procedure stays close to the green line that presents the average rent in the raw data of our holdout sample. For the bottom quartile of property prices the hedonic method predicts rents that are 25% too high, while our switcher method predicts rents that are within 5% of the true rents in the holdout sample. Similarly, in the top quartile of property values, the hedonic method predicts rents that are 25% too low, while our switcher method predicts rents within 5% of the true rents.

Using a method that accurately estimates the expected rent $E(r|p)$ of a property with a given price p is crucial for our application to understanding racial differences in housing returns. Because minority

homeowners tend to live in lower-price houses on average than White homeowners, the hedonic method would artificially increase the rental yield of minority homeowners relative to White homeowners. As a result, the return on housing investment for minority homeowners would be artificially increased relative to that of White homeowners, which we show empirically in appendix G.

5 Evidence on rent-to-price ratios by race from other sources

This section compares the rent-to-price ratios implied by our new rent imputation method to evidence from other data sources. We first examine raw data from the Consumer Expenditure Survey (CEX), which asks homeowners both about what they believe they could sell their homes for as well as what they think they could rent their homes for. As far as we are aware, this is the only dataset that directly measures rents and sale prices for owner-occupied properties in the US, though it relies heavily on homeowners' subjective beliefs. Second, we use large administrative datasets on housing transactions from Corelogic and on rents from Multiple Listing Services to estimate rent-to-price ratios using a method similar to our approach in the AHS.

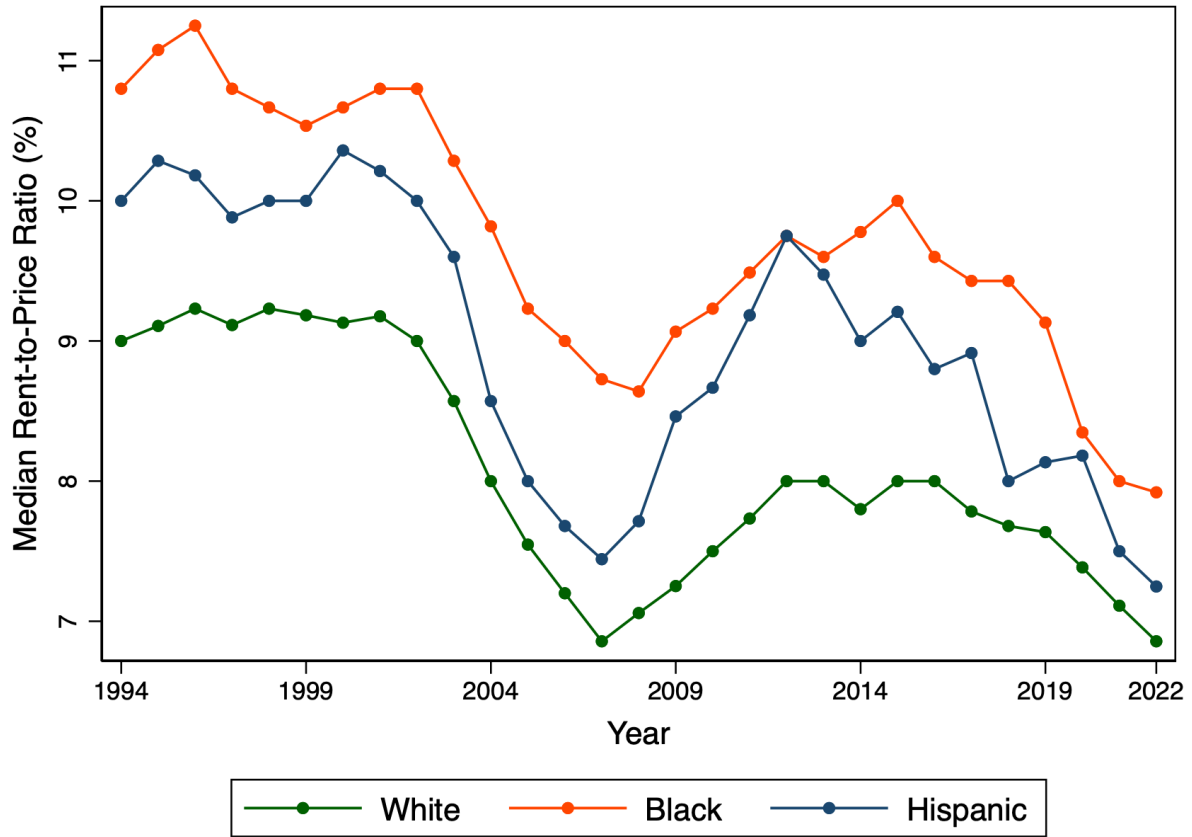
5.1 Consumer Expenditure Survey

Figure 5.1 presents a time series plot of the median rent-to-price ratio in the CEX from 1994-2022 by racial group. This figure looks qualitatively quite similar to our imputed AHS rents in Figure 3.2 with a flat period in the 1990s, declines in the 2000s, growth in the early 2010s and then declines to 2022. The CEX data also shows Black rent-to-price ratios above Hispanic and Hispanic rent-to-price ratios above White. However, the magnitudes are a bit higher than what we find in the AHS. Table 5.1 reports our AHS imputed rent-to-price ratios by race for the same sample period as our CEX data. In the CEX, the average rent-to-price ratio for White homeowners is 9.40%, while we find 7.96% in the AHS. For Black homeowners the CEX shows 11.51%, while the AHS shows 8.93%. For Hispanic homeowners, the rent-to-price ratio is 11.21% in the CEX and 8.47% in the AHS. The racial gaps in rent-to-price ratios are somewhat wider in the CEX than what we find in the AHS. While these results broadly confirm the rental yield disparities we found in the AHS, the quantitative difference between the findings in the two data sources raises the question of which is most correct.

One possibility is that most homeowners who do not rent their homes out are overly optimistic about the rent they could charge. This bias is likely worse for rent than for price data, biasing rental yields upwards. While there is always some concern with bias and measurement error in survey-reported variables, owner occupants did purchase their house, making them at least somewhat informed on property values. However, most owners will never rent out their house, possibly making them very uninformed about the rental value of their property. Alternatively, it could be that the raw data in the CEX is more accurate since our rent imputation approach in the AHS requires us to make statistical assumptions which could be violated. While our estimation procedure allows for unobservable differences in quality between the switchers we use for estimation and other owner-occupied homes, we have to assume that

these differences are captured by a home's price. To adjudicate this discrepancy with a third data source, we now turn to administrative data on property transaction prices, both in the rental market and the purchase market, where we believe that rents and prices are most accurately measured.

Figure 5.1: Median Rent-to-Price Ratios for Homeowners by Racial Group over Time, 1994-2022 in the Consumer Expenditure Survey



Note: Data from the Consumer Expenditure Survey. The CEX asks homeowners the property value and what their property could rent for. We use these to calculate rent-to-price ratios.

Table 5.1: Rent-to-Price Ratio for Owner-occupied units in the CEX and AHS by Owner Race, 1994-2021

CEX	Mean	Median	SD	25th Pct.	75th Pct.	Num Obs.
White	9.40	7.60	7.58	5.76	10.06	127,222
Black	11.51	9.29	8.47	6.96	12.51	12,962
Hispanic	11.21	8.47	9.65	6.17	12.00	14,757
AHS	Mean	Median	SD	25th Pct.	75th Pct.	Num Obs.
White	7.96	7.20	4.49	5.43	9.00	335,161
Black	8.93	7.91	5.00	5.77	10.44	37,204
Hispanic	8.47	7.34	5.22	5.14	9.85	39,126

Note: The sample includes all homeowners and is winsorized at the 2.5th and 97.5th percentiles.

5.2 Rental yields from administrative data

To complement our evidence on racial disparities in rental yields from the AHS and the CEX, this section reports evidence from rental yields inferred from data on actual transaction prices. We turn to administrative deeds records on property transactions, assessor records for property characteristics, and MLS rental listings from Corelogic. We also use data on property resident characteristics from Infutor. The deeds records on property transactions prices are linked to assessor records and MLS rental listing through unique parcel id numbers (APNs).⁴ These data are then linked to the property resident listed in the Infutor data through a link using address strings and spatial linking through GPS coordinates. See Appendix B for more details.

A key upside of using these administrative data sources is the very large sample size. This linking procedure provides us with 2,845,590 observations that have a reported rent and a property transaction which we use in estimation. However, our data only start in 2007 and are not geographically representative, as we show in appendix B. Over 50 percent of our sample comes from 5 MSAs: Las Vegas, NV; Los Angeles, CA; Phoenix, AZ; Washington DC, and West Palm Beach, FL.

Since our sample is so much larger than the AHS switcher sample, we run our rent estimation regression separately by metropolitan statistical area and by year. For our baseline specification, we run the following regression of housing rents on housing prices and other covariates

$$rent_{ht} = \alpha_{jt} + \beta_{jt}X_{ht} + \epsilon_{ht}. \quad (12)$$

The regression includes the house’s transaction price and a rich set of controls and fixed effects described in appendix B. We run this regression separately by MSA-year if we have at least a sample size of 1,000 observations. If this threshold fails, we run the same regression pooled across years with year fixed effects if there is a minimum of 2,000 observations. Regions that do not have enough data are excluded. One issue we must account for that was absent in the AHS is that the year at which the housing unit is rented need not be the same year in which the housing unit transacts in the purchase market. To deal with this, we compute the quantile of each housing unit’s price in a county-year-specific distribution of house values from the American Community Survey. We then assign the house an imputed purchase price with the same quantile as the house value distribution which matches the county and year of the rental transaction. We use this imputed price as a regressor to help predict rents.

We use our estimates of equation 12 to construct fitted values for the rental value of owner-occupied housing units. We then compute the rental yield of an owner-occupied home by dividing our predicted rent value by the house’s resale value $price_{ht}$. In years where the house does not transact, we use the approach described above using price quantiles to impute a price. The final sample we analyze covers 192 million observations of property values and imputed rents for owner-occupied properties between 2007 and 2016.

Table 5.2 presents regressions that analyze racial disparities in the log rent-to-price ratios we esti-

⁴We only study MLS rental listings that indicate that the rental listing was actually rented and lists the “closing” rent, which may differ from the original post rent in the listing.

mate.⁵ Column 1 controls for year fixed effects and finds Black homeowners earn a 1.6 percentage point higher rental yield than White homeowners. Hispanic homeowners earn 0.39 percentage points more than White homeowners. To handle our non-representative geographic sample, column 2 adds MSA fixed effects and column 3 adds MSA-year fixed effects, which look similar. When looking within MSAs, Hispanics now earn 1.02 percentage points higher rent-to-price ratios than Whites. This substantial widening comes from the fact that our sample over-weights California, where rent-to-price ratios are low and many Hispanics live.

To validate our AHS rent imputation method against this administrative data, columns 5 and 6 present regressions in our AHS sample restricted to the same time period and MSAs as covered in our administrative data. Just including year fixed effects, we find Black homeowners earn 1.49 percentage points higher rent-to-price ratios than White homeowners and Hispanic homeowners earn 0.19 percentage points more than White homeowners. This is extremely economically close to the Black-White gap we find of 1.61pp and Hispanic-White gap of 0.39pp in the administrative and easily within our 95% confidence intervals. Column 6 adds MSA-year fixed effects to our AHS sample and we find a within-MSA Black-White gap of 1.47pp and Hispanic-White gap of 0.92pp. These are again very economically and statistically close to our administrative data findings of 1.67pp and 1.02pp, respectively. This suggests that the racial disparities we found in our AHS data are the right size, while those in the CEX survey data are potentially too large.

Finally, we add zip-year fixed effects into our administrative data regression in column 4 of Table 5.2. The within zip-year Black-White gap falls to 0.48pp, and the Hispanic-White gap falls to 0.35pp. This implies that 29% (0.48/1.67) of the within-MSA Black-White and 34% (0.35/1.02) of the Hispanic-White rental yield gap occurs *within* zipcodes. This suggests that imputing rent-to-price ratios at the zipcode level using standard hedonic methods, as done in [Kermani and Wong \[2022\]](#) and [Demers and Eisfeldt \[2001\]](#), misses a substantial portion of racial inequality in rent-to-price ratios that happens within zipcodes.

Table 5.2: Racial Gaps in $\text{Log}(1 + \text{Rent-to-Price Ratio})$

	CoreLogic				AHS	
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.01607*** (0.00037)	0.01664*** (0.00031)	0.01671*** (0.00029)	0.00480*** (0.00015)	0.0149*** (0.00121)	0.0147*** (0.00107)
Hispanic	0.00392*** (0.00030)	0.01025*** (0.00025)	0.01026*** (0.00024)	0.00346*** (0.00007)	0.00190 (0.00138)	0.00925*** (0.00129)
Observations	192,329,864	192,329,864	192,329,864	192,329,864	26,857	26,857
Year FEs	Yes	Yes	-	-	Yes	-
MSA FEs	No	Yes	-	-	No	-
MSA $\times$ Year FEs	No	No	Yes	-	No	Yes
Zip Code FEs	No	No	No	-	No	No
Zip Code $\times$ Year FEs	No	No	No	Yes	No	No

Note: This table documents racial gaps in log rent-to-price ratios in MSAs where we have observations both in Corelogic and the AHS from 2007-2016. For consistency with the AHS, Corelogic data are weighted per household rather than per person. Standard errors (clustered at the Zip Code level for Corelogic data and the unit level for AHS data) are in parentheses.

⁵We use $\log(1 + \text{rent}/\text{price})$ as our rent-to-price metric as it will be identical to the metric used in the next section where we decompose total rates of return into their individual sources. This metric will measure the contribution of rents to total rates of return.

6 Racial Disparities in the Return on Owner-Occupied Housing in the AHS

We now apply our quantile-based rent imputation to compare the total rates of return on owner-occupied housing earned by different racial groups. Our analysis so far has shown that our rent imputation method provides accurate rent predictions both for low- and high- price homes and that rental yields vary significantly between racial groups. This section now shows that the higher rent-to-price ratios for minority homeowners are reflected in a higher average return on housing investment as well.

For each homeowner, we calculate the log-return on its housing to account for their home’s rental value, property taxes, and maintenance. First, we compute the total “dividend” of house i at time t as

$$dividend_{it} = rent_{it} - proptax_{it} - maint_{it}, \quad (13)$$

where the rent is imputed using our quantile-based methodology and property taxes and maintenance costs are directly reported in the AHS. Our maintenance measure includes both routine maintenance and occasional home improvement projects. For years when the AHS is reported annually, we measure log-returns as

$$\log(P_{it+1}) - \log(P_{it} - dividend_{it}). \quad (14)$$

After 1981, when the AHS becomes only bi-annual, we annualize our rate of return measure by dividing the expression in equation 14 by 2. This yields a house-time level panel of rates of returns. We CPI-adjust all dollars, so rates of return are real.

Table 6.1 presents summary statistics on our return measure by race. The average return on housing is 6.5% for White homeowners, 7.2% for Black homeowners, and 7.9% for Hispanic homeowners. We find that the standard deviation of returns is highest for Black and then for Hispanic homeowners. Figure 6.1 plots these rates of return by race over time. We find the rates of returns of all three racial groups track macroeconomic housing conditions, with lower return rates during recessions. This figure suggests that the higher returns earned by Black and Hispanic homeowners come in non-recession times, while it appears these groups frequently earn lower returns during recessions than White homeowners. This figure highlights the importance of studying these returns over a long time period, since shorter run business cycle dynamics impact the overall level of returns, as well as the racial disparities.

We decompose racial rate of return differences into property appreciation rates and dividend yields in tables 6.2 and 6.3. To measure the contribution of dividends to the rate of return, we set the property appreciation rate to zero and recalculate the rate of return. Similarly, we set the dividend to zero to measure the contribution of appreciation rates. We find slightly higher appreciation rates for White (0.9%) than Black (0.7%) homeowners but a substantially higher appreciation rate of 1.6% for Hispanic homeowners. In addition, we find that Hispanic and particularly Black homeowners have higher average dividend yields (5.9% and 6.0%, respectively) than White homeowners (5.3%). The high 29% standard deviation of total returns for Black homeowners is explained primarily by their property appreciation

rate standard deviation being 25.9%, the highest of all racial groups. This is broadly consistent with the [Kermani and Wong \[2022\]](#) finding that a disproportionate share of distressed sales for minority homeowners raises their probability of a large negative investment return.

Table 6.1: Log rate of return

	White	Black	Hispanic	All groups
Mean	0.065	0.072	0.079	0.066
Median	0.052	0.058	0.068	0.053
StdDev	0.245	0.291	0.270	0.251
Observations	626,636	59,142	44,810	758,828

Note: Data from 1974-2021 AHS home owners. Summary statistics of annual log rates of return by race.

Table 6.2: Log property value rate (excl. dividends)

	White	Black	Hispanic	All groups
Mean	0.009	0.007	0.016	0.009
Median	0.001	0.001	0.014	0.002
StdDev	0.220	0.259	0.237	0.225
Observations	626,636	59,142	44,810	758,828

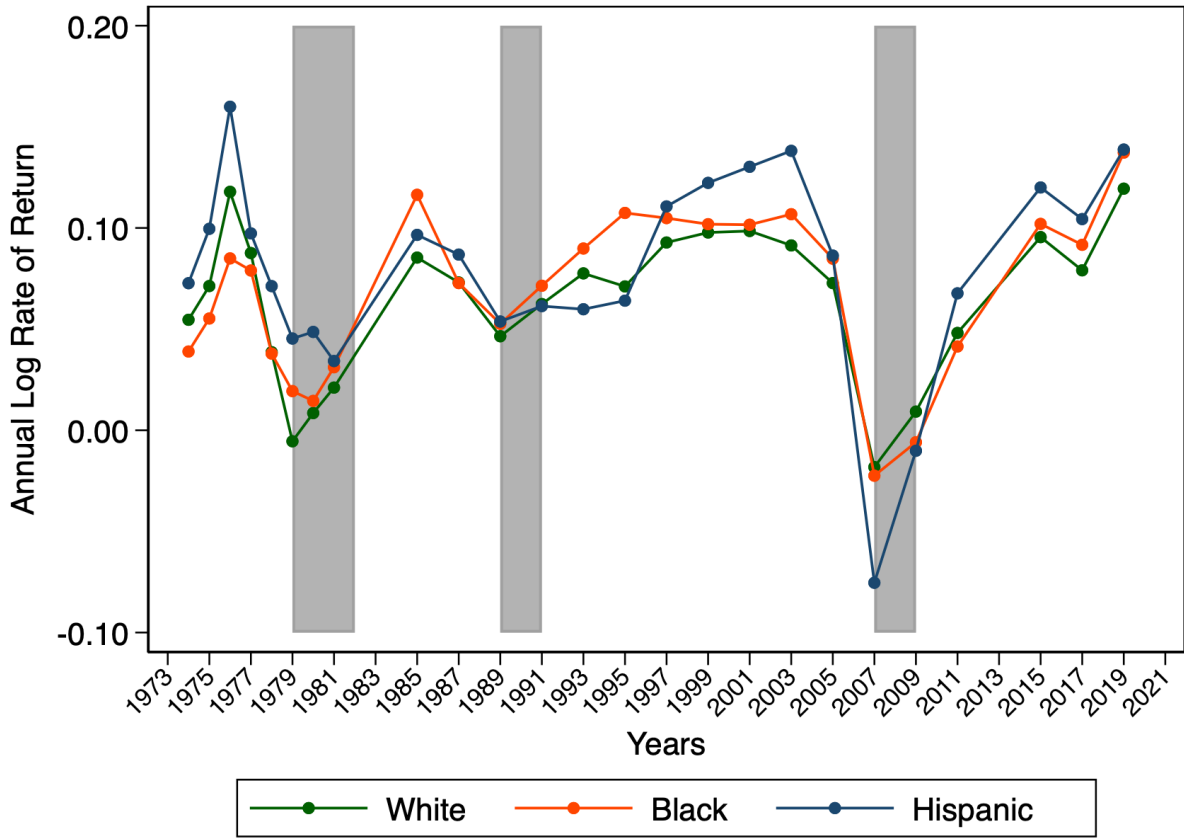
Note: Data from 1974-2021 AHS home owners. Summary statistics of annual log property value appreciation by race.

Table 6.3: Log dividend rate

	White	Black	Hispanic	All groups
Mean	0.053	0.060	0.059	0.053
Median	0.045	0.052	0.049	0.045
StdDev	0.065	0.075	0.076	0.067
Observations	626,636	59,142	44,810	758,828

Note: Data from 1974-2021 AHS home owners. Summary statistics of annual log dividend rate by race.

Figure 6.1: Average Annual Rate of Return by Racial Group, 1974-2021

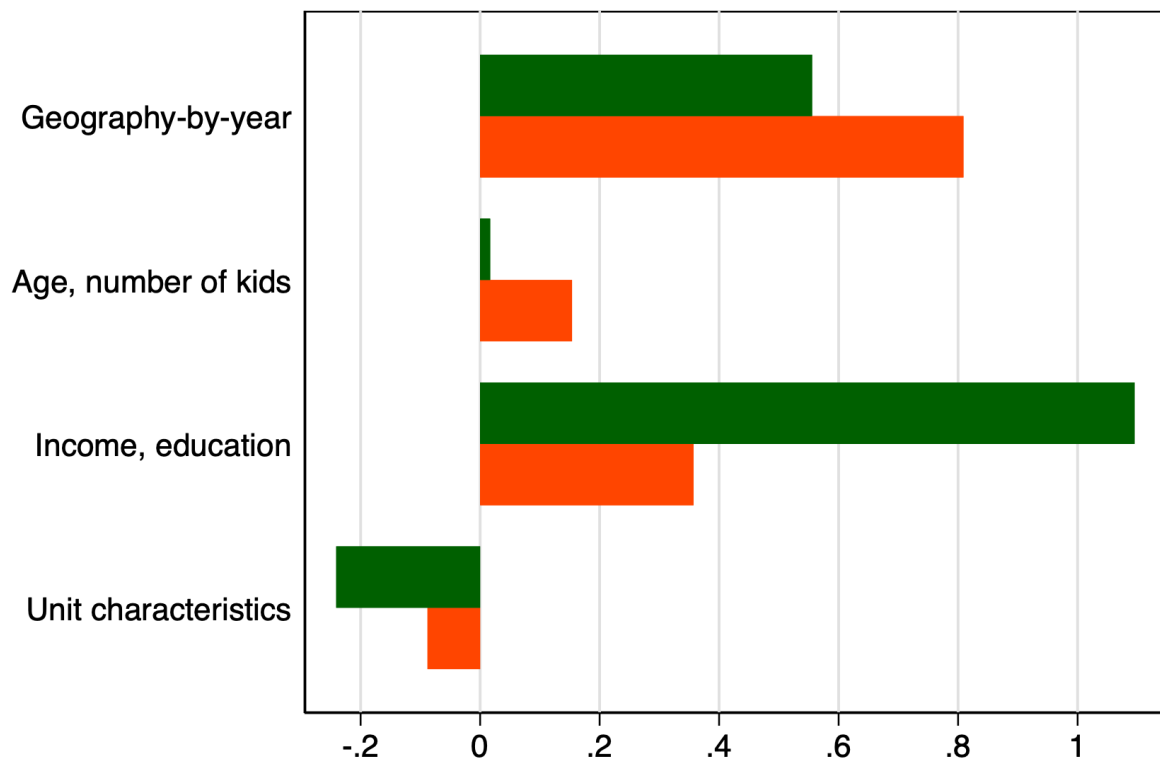


Note: Data from 1974-2021 AHS home owners. Plots average annual log rates of return by race. Shaded gray areas indicate periods of recession.

Table 6.4 investigates the extent to which we can account for disparities in housing rates of returns between racial groups with observable characteristics. To do so, we regress our individual-level log rates of return on a dummy variable for each racial group together with a growing list of control variables. Controlling only for year fixed effects indicates a 0.62pp higher rate of return for Black homeowners and 1.19pp higher rate of return for Hispanic homeowners relative to White homeowners. Column 2 adds MSA-year fixed effects. This does little to the Black-White gap, but it narrows the Hispanic-White gap a bit to 0.94 pp. Column 3 controls for household demographics: household head age, number of children in the household. These controls slightly increase the Black-White gap to 0.68pp and the Hispanic-White gap to 0.99pp. Controlling for household economic characteristics (household income, household head education, and years since household head moved into the housing unit) more than entirely accounts for the entire Black-White and Hispanic-White difference, leading to -0.36pp and -0.33pp respectively. This suggests that houses bought by lower income/education homeowners tend to have higher average returns and that these houses are disproportionately bought by minority homeowners. Adding additional controls for housing characteristics fully closes the Black-White and Hispanic-White gaps, leading to statistically and economically insignificant gaps of -0.16pp (Black-White) and -0.009pp (Hispanic-White).

We use a Gelbach [2016] decomposition in figure 6.2 to quantify our ability to explain differences in

Figure 6.2: [Gelbach \[2016\]](#) Decomposition of Explainable Share of Racial Disparities in Returns



Note: Data from 1974-2021 AHS home owners. Each bar reports the fraction of the total return disparity explained by each set of control variables using the [Gelbach \[2016\]](#) method. The green bar represents the Black-White total rate of return gap, while the orange bar represents the Hispanic-White total rate of return gap.

average returns across homeowners of different races. This decomposition compares a regression with all covariates in column 5 of table 6.4 to the one in the first column. The decomposition uses a multivariate version of the formula for omitted variable bias to account for how much each variable helps to explain the change in our regression coefficient when all are included jointly. Unlike the standard approach of adding regressors one-by-one, this decomposition does not require us to arbitrarily choose the order in which regressors are added. We find consistent with table 6.4 that the income/education controls are by far the most important for explaining Black-White differences in rates of return and play an important role for the Hispanic-White gap. However, geography also plays a large role, especially for the Hispanic-White gap.

We examine the observable contribution of rental yield differences to total rates of return across racial groups in table 6.5. To do this, we examine what returns homeowners would earn in equation 14 if all terms other than rental yields were set to zero. Similar to our overall summary statistics, we find log rates of return due only to rental yields that are 1.2pp higher for Black homeowners and 0.79pp higher for Hispanic homeowners relative to White homeowners, controlling for year fixed effects. Few of our controls help to explain this disparity except the household economic characteristics added in column 4. This control explains 70% of the Black-White rental yield gap and essentially all of the Hispanic-White gap. This suggests that lower income/education homeowners live in cheaper homes, which tend to have higher rental yields. Adding in property characteristic controls widens the gap, leaving a Black-White gap of 0.53pp and 0.30pp Hispanic-White gap, even with all controls.

Potential differences in property taxes also contribute to dividend rates. Table 6.6 compares racial differences in log rates of return driven only by property taxes. Note that since taxes are a cost, all rates of return here will be negative and that a negative racial minority gap implies higher racial minority property tax rates. Controlling only for year fixed effects, column 1 shows property taxes contribute to a small, positive Black-White return gap of 0.029%, while the Hispanic-White gap is even closer to zero at -0.004%. This implies Black homeowners face slightly lower property tax rates than White homeowners, while Hispanics face a tiny bit higher. However, both are economically very small contributors to rates of return. Adding MSA-year fixed effects now shows slightly negative (i.e. higher tax rates) rates of return for both Black-White and Hispanic-White gaps. Adding in additional controls for demographics, household economic conditions, and property characteristics does little to mitigate the Hispanic-White gap, but fully explains the Black-White gap in returns driven by property taxes. Overall, since property taxes are small in magnitude relative to other forces driving rates of return, racial gaps in property taxes have negligible effects on racial differences in total rates of return.

This may seem at odds with recent work on racial inequality in property taxes [Avenancio-León and Howard, 2022] that shows Black and Hispanic homeowners paying higher property tax rates than White homeowners. Column 2 of Table 6.6, which includes only MSAs-year fixed effects, is the most comparable specification to analysis done in [Avenancio-León and Howard, 2022]. This specification shows both Black and Hispanic homeowners paying higher tax rates within MSA-years compared to Whites, although our magnitudes are smaller than those in Avenancio-León and Howard [2022]. This smaller magnitude

appears to be driven by differences in time period and geographic span of the data in [Avenancio-León and Howard \[2022\]](#) and AHS. In Appendix C, we replicate the main findings of [Avenancio-León and Howard \[2022\]](#) in administrative Corelogic property tax data from 2007 to 2016. We then show how changes in sample definitions to align the time period and geography examined in the AHS and in Corelogic yield similar results across both datasets. Using only the years 2007-2016 and MSAs that are observed in the administrative property tax data, we find in the AHS that property taxes lower Black homeowner returns by 0.17 percent relative to White homeowner returns. This is substantially more negative than what we find for the overall nation from 1974-2021. This quantitatively small impact on housing returns is consistent with the magnitudes in [Avenancio-León and Howard \[2022\]](#). A back of the envelope calculation using numbers from [Avenancio-León and Howard \[2022\]](#) indicates the median home worth \$207,000 is on average taxed at 1.4 percent. Assuming their headline number that Black homeowners pay 12% higher tax rates within an MSA, this tax rate would be $1.4 * (1.12) = 1.568$ percent. This would imply a Black-White rate-of-return gap of 0.19%, very close to what we find.

Looking at differences in rates of return driven by spending on housing maintenance and home improvement, Table 6.7 shows that Black homeowners spend a greater fraction of their home’s value on maintenance and home improvement than White homeowners, leading to a 0.38% lower rate of return. Hispanic homeowners spend more than White homeowners as well, lowering their rate of return by 0.07% relative to White homeowners. This offsets some but not all of the higher rental yields Black and Hispanic homeowners earn. The Black-White and Hispanic-White gaps become more similar to each other, -0.28% and -0.24%, respectively, when adding MSA-year controls. Controlling for demographics has little impact on these numbers. Adding in household economic characteristics shrinks these gaps to -0.21% for Black homeowners and -0.11% for Hispanic homeowners. Adding in housing characteristics further narrows these gaps to -0.19% and -0.10% for Black and Hispanic homeowners, respectively, relative to White homeowners.

Next, we consider in Table 6.8 the determinants of racial disparities in property appreciation rates. Including only year fixed effects, we find a Black-White appreciation gap of -0.26% and a Hispanic-White gap of 0.48%. The fact that Hispanic homeowners earn higher appreciation rates is driven by their differences in geographic exposure to local appreciation rates. Column 2 adds MSA-year fixed effects, which leads to a Black-White gap of -0.18% and a Hispanic-White gap of 0.17%. Adding demographic controls does little to these gaps. Adding controls for household economic characteristics makes these gaps more negative. The Black-White gap is now -0.46% and Hispanic-White is -0.06%. Adding house characteristics makes these a bit more negative at -0.47% and -0.10%, respectively.

While we are able to replicate the previous literature’s finding that Black and Hispanic homeowners earn lower appreciation rates than White homeowners, our magnitudes are less negative than those found in [Kermani and Wong \[2022\]](#) using administrative data. We return to this point and reconcile the returns we find in the AHS and in Corelogic administrative data with [Kermani and Wong \[2022\]](#) in Section 7.

Table 6.4: Log rate of return

	(1)	(2)	(3)	(4)	(5)
Black	0.00619*** (0.00116)	0.00604*** (0.00118)	0.00682*** (0.00118)	-0.00355** (0.00118)	-0.00165 (0.00118)
Hispanic	0.0119*** (0.00137)	0.00940*** (0.00137)	0.00994*** (0.00138)	-0.00329* (0.00135)	-0.000942 (0.00132)
Observations	742,745	742,745	742,745	742,745	742,745
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time.

Table 6.5: Log rate of owner-occupied rent

	(1)	(2)	(3)	(4)	(5)
Black	0.0124*** (0.000747)	0.0109*** (0.000699)	0.0114*** (0.000697)	0.00372*** (0.000684)	0.00527*** (0.000672)
Hispanic	0.00794*** (0.000962)	0.0118*** (0.000897)	0.0118*** (0.000899)	0.000391 (0.000852)	0.00297*** (0.000760)
Observations	742,745	742,745	742,745	742,745	742,745
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time.

Table 6.6: Log rate of annual property taxes

	(1)	(2)	(3)	(4)	(5)
Black	0.000287* (0.000143)	-0.000251* (0.000126)	-0.000260* (0.000126)	0.000121 (0.000127)	-0.0000404 (0.000127)
Hispanic	-0.0000432 (0.000177)	-0.00153*** (0.000170)	-0.00149*** (0.000170)	-0.000963*** (0.000172)	-0.00119*** (0.000168)
Observations	742,745	742,745	742,745	742,745	742,745
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. The dependent variable measures the rate-of-return due only to property tax payments. A negative racial gap indicates higher property tax rates for the racial minority group. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time.

Table 6.7: Log rate of maintenance cost

	(1)	(2)	(3)	(4)	(5)
Black	-0.00375*** (0.000290)	-0.00278*** (0.000259)	-0.00263*** (0.000260)	-0.00214*** (0.000260)	-0.00189*** (0.000252)
Hispanic	-0.000731** (0.000261)	-0.00244*** (0.000273)	-0.00198*** (0.000275)	-0.00110*** (0.000279)	-0.000994*** (0.000267)
Observations	742,745	742,745	742,745	742,745	742,745
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. The dependent variable is the rate of return driven only by maintenance spending. A negative racial gap indicates higher spending on maintenance and home improvement as a share of the property value by the racial minority group. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time.

Table 6.8: Log rate of property values (excluding dividends)

	(1)	(2)	(3)	(4)	(5)
Black	-0.00261** (0.000813)	-0.00176* (0.000841)	-0.00169* (0.000844)	-0.00461*** (0.000853)	-0.00467*** (0.000854)
Hispanic	0.00478*** (0.000944)	0.00168 (0.000978)	0.00178 (0.000983)	-0.000560 (0.000996)	-0.00110 (0.000996)
Observations	742,745	742,745	742,745	742,745	742,745
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. The Dependent variable is the rate of return driven only by price appreciation. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time.

6.1 Exposure to Business Cycles

We explore racial disparities in the total rate of return on housing over the business cycle in Table 6.9. In this table, we regress the realized log total return homeowners' earn on dummy variables for race as well as interactions between race and an indicator for recession years with a growing list of control variables. In all of our specifications, the interaction effect between a recession indicator and an indicator for either a Black or Hispanic homeowner is negative. This implies that Black and Hispanic homeowners have returns on housing that are more sensitive to the business cycle than White homeowners. In column 1, the Black-White return gap is 0.82% in non-recession times, but it falls by 0.67 percentage points during recessions. The Hispanic-White gap is very sensitive to the business cycle. In non-recession times the gap is 2.05%, but it falls by 3.32 percentage points during recessions and becomes negative. A substantial share of Hispanic homeowners' excess exposure to boom-bust cycles appears to be due to the MSAs that they live in. Controlling for MSA-year effects in column 2 reduces the Hispanic-White gap to 1.14% in non-recessions, which now falls by 0.80 percentage points in recessions. MSA-year fixed effects have little effect on the Black-White gap, and if any widen the gap a bit, both during booms and busts. Adding controls for household demographics does little to these gaps for either racial group.

However, column 4 shows that controlling for household economic characteristics largely accounts for the greater return cyclicity of Hispanic homeowners and explains the excess returns of the Black-White gap during boom times. However, we find that these observables have almost no ability to account for the excessively negative returns of Black homeowners during recessions. Column 5 adds house characteristics controls. The Hispanic-White gap was already statistically indistinguishable from zero without these controls in both boom and bust times, and this specification pushes it even closer to zero. However, the Black-White return gap remains negative and significant in recession times at -0.68%. One possibility is that segregation between Black and White homeowners results in Black homeowners living in more cyclical areas. Indeed, geographic racial segregation indices show that Blacks are twice as geographically segregated from Whites as Hispanics are [Elbers, 2021]. Another possibility is Black homeowners are more likely to sell their homes as part of a foreclosure [Kermani and Wong, 2022], particularly in recessions.

We decompose this excess cyclicity into price appreciation and rental yields in tables 6.10 and 6.11. For both Black and Hispanic homeowners, we find that their rental yields only fall a bit in recessions relative to White homeowners. Even with all of our observable controls the Black-White and Hispanic-White return gap from rents remains positive and significant in boom and bust times at 0.56% and 0.31%, respectively during boom times and 0.44% and 0.25% during bust times. Turning to price appreciation rates, table 6.11 shows Blacks have slightly lower price appreciation in boom times at -0.14% and this falls to -0.55% in bust times. This excess cyclicity is even larger for the Hispanic-White gap, with a positive gap of 1.25% in boom times and -1.74% during bust times. Controlling for geography-year fixed effects narrows these gaps for Hispanics to 0.29% during boom times and -0.18% during bust times. However, this does little to explain the excess exposure of Black homeowners to the business cycle.

Table 6.9: Log rate of return

	(1)	(2)	(3)	(4)	(5)
Black	0.00815*** (0.00147)	0.00856*** (0.00150)	0.00924*** (0.00150)	-0.00143 (0.00149)	0.000310 (0.00147)
Black=1 $\times$ Recession=1	-0.00667* (0.00284)	-0.00884** (0.00289)	-0.00853** (0.00290)	-0.00731* (0.00291)	-0.00677* (0.00292)
Hispanic	0.0205*** (0.00159)	0.0114*** (0.00166)	0.0115*** (0.00166)	-0.00263 (0.00163)	-0.0000337 (0.00161)
Hispanic=1 $\times$ Recession=1	-0.0332*** (0.00342)	-0.00799* (0.00346)	-0.00643 (0.00346)	-0.00318 (0.00348)	-0.00396 (0.00343)
Observations	742,745	742,745	742,745	742,745	742,745
Recession-by-race FE (fully interacted)					
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. The dependent variable is the total log rate of return. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time. Recession year dummies (1975,1979,1980,1981,1989,1991,2007,2009) come from [FRED](#).

Table 6.10: Log rate of owner-occupied rent

	(1)	(2)	(3)	(4)	(5)
Black	0.0129*** (0.000814)	0.0114*** (0.000775)	0.0118*** (0.000772)	0.00416*** (0.000759)	0.00564*** (0.000752)
Black=1 $\times$ Recession=1	-0.00153 (0.000814)	-0.00164* (0.000810)	-0.00158 (0.000810)	-0.00144 (0.000819)	-0.00121 (0.000817)
Hispanic	0.00903*** (0.00103)	0.0124*** (0.000977)	0.0121*** (0.000979)	0.000223 (0.000929)	0.00310*** (0.000831)
Hispanic=1 $\times$ Recession=1	-0.00424*** (0.00112)	-0.00202 (0.00118)	-0.00152 (0.00118)	0.000320 (0.00116)	-0.000629 (0.00104)
Observations	742,745	742,745	742,745	742,745	742,745
Recession-by-race FE (fully interacted)					
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time. Recession year dummies (1975,1979,1980,1981,1989,1991,2007,2009) come from [FRED](#).

Table 6.11: Log rate of property values (excluding dividends)

	(1)	(2)	(3)	(4)	(5)
Black	-0.00137 (0.00112)	-0.0000327 (0.00115)	-0.0000783 (0.00115)	-0.00318** (0.00116)	-0.00323** (0.00116)
Black=1 $\times$ Recession=1	-0.00417 (0.00252)	-0.00607* (0.00257)	-0.00568* (0.00258)	-0.00494 (0.00259)	-0.00502 (0.00259)
Hispanic	0.0125*** (0.00123)	0.00327* (0.00129)	0.00292* (0.00130)	0.000332 (0.00131)	-0.000209 (0.00131)
Hispanic=1 $\times$ Recession=1	-0.0299*** (0.00295)	-0.00622* (0.00294)	-0.00468 (0.00295)	-0.00369 (0.00299)	-0.00369 (0.00298)
Observations	742,745	742,745	742,745	742,745	742,745
Recession-by-race FE (fully interacted)					
Dummy for other racial group	Yes	Yes	Yes	Yes	Yes
Year FE (not interacted)	Yes	-	-	-	-
Geography-by-year FE	No	Yes	Yes	Yes	Yes
Demographics (age, #kids)	No	No	Yes	Yes	Yes
Income, education, tenure	No	No	No	Yes	Yes
Housing characteristics	No	No	No	No	Yes

Note: Data from 1974-2021 owner-occupied properties in the AHS. The dependent variable is the rate of return driven only by price appreciation. Standard errors are clustered at the unit level. The excluded race group is White owner-occupants. Geography FEs include the most granular-defined geography variables in the AHS for each year, which change some over time. Recession year dummies (1975,1979,1980,1981,1989,1991,2007,2009) come from [FRED](#).

7 Reconciling Estimates of House Price Appreciation Rates

This section attempts to reconcile the measured rate of house price appreciation observed in the AHS to estimates from administrative datasets such as Corelogic. As documented by [Kermani and Wong \[2022\]](#), if we measure the return on housing by looking only at housing spells with a measured buy and sell price, White homeowners earn higher realized returns than Black or Hispanic homeowners. They demonstrate that racial disparities in foreclosure rates drive much of this overall return difference, and we find similar results to them when examining housing spells with measured buy and sell prices. However, we also show that in our Corelogic data, a significant share of housing spells either begin at the start of the sample or end at the conclusion of the sample and are therefore missing a buy price and/or a sell price. Our analysis is restricted to four counties (Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA), where we study foreclosure discounts in detail in appendix [D](#). Racial disparities in house price appreciation look much closer to our AHS results once we account for the entire housing stock including houses that do not transact. In addition, under the very conservative assumption that our AHS returns entirely miss foreclosure discounts, we use a measure of foreclosure rates and discounts from [Kermani and Wong \[2022\]](#) and show that minorities still earn a higher rate of return after accounting for racial disparities in foreclosures.

We first show in table [7.1](#) that a significant share of housing spells either begin at the start of our sample or are still continuing at the end, particularly for minorities. If we weigh these spells proportional to their length, we find that 74.9% of the years of overall homeownership and 86.1% of the years of Black homeownership are part of a spell that begins and/or ends outside of our sample. To make the rates of

return measured in AHS and Corelogic comparable, we need to account for the fact that the AHS tries to measure returns on the entire housing stock while Corelogic only reports data on housing units that transact.

Table 7.1: Summary Statistics: Completeness of Housing Spells by Race

	White		Black		Hispanic		All	
	Num.	Pct.	Num.	Pct.	Num.	Pct.	Num.	Pct.
Panel A: At the Spell Level								
Purchased and Sold Only	2,370,788	40.78	231,359	24.93	294,445	38.99	2,896,592	38.63
Purchased but Unsold Only	1,851,049	31.84	372,765	40.16	365,627	48.41	2,589,441	34.54
Not Purchased Only	1,592,150	27.38	324,090	34.92	95,177	12.60	2,011,417	26.83
Not Purchased or Unsold Only	3,443,199	59.22	696,855	75.07	460,804	61.01	4,600,858	61.37
Overall	5,813,987	-	928,214	-	755,249	-	7,497,450	-
Panel B: At the Spell-Year Level								
Purchased and Sold Only	29,684,541	25.08	3,301,150	13.90	3,522,036	23.88	36,507,727	23.27
Purchased but Unsold Only	38,725,323	32.72	8,390,092	35.32	8,004,645	54.28	55,120,060	35.14
Not Purchased Only	49,956,200	42.20	12,066,557	50.79	3,220,805	21.84	65,243,562	41.59
Not Purchased or Unsold Only	88,681,523	74.92	20,456,649	86.10	11,225,450	76.12	120,363,622	76.73
Overall	118,366,064	-	23,757,799	-	14,747,486	-	156,871,349	-

Note: This table presents summary statistics on the fraction of housing spells that begin with a purchase and/or terminate with a sale in sample, by race. Data is from 1975-2016 across four counties: Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA.

Table 7.2 analyzes how racial disparities in house price appreciation vary with the sample used to compute average returns. In each column, we regress the log of house price appreciation on dummy variables for each White/Black/Hispanic racial group and change our included sample and/or fixed effect controls across columns. We use data from four counties: Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA. In appendix D, we discuss how we have carefully measured the returns on foreclosed homes specifically in these counties. The two panels A and B of the table measure the return on foreclosed properties in different ways. Panel A uses the price at which houses sold in a foreclosure auction, while Panel B uses a measure of the value of debt which a homeowner expunges when it “sells” its home for the face value of its outstanding mortgage, as described in detail in appendix D. In the columns on the left of this table, we see that racial return disparities are sensitive to foreclosure measurement, while they become much less so in the columns on the right hand side of the table that use more and more expansive sets of homes.

In the table’s first specification in panel A, we find a 5.3% lower rate of house price appreciation for Hispanic and 9.9% lower rate of house price appreciation for Black homeowners relative to White homeowners. This column looks only at complete housing spells with a buy and a sell, restricts to a 2000-2016 sample consistent with [Kermani and Wong \[2022\]](#), and only examines homes bought with a mortgage (since [Kermani and Wong \[2022\]](#) use HMDA mortgage data to infer the race of a homeowner). We then include in column 2 the set of homeowners who buy houses in sample but do not sell. To measure the rate of appreciation on houses that do not have both a buy and a sell, we follow [Kermani and Wong \[2022\]](#) in using a year-county specific return on an FHFA local house price index as a proxy for the return on the individual house. We find a 3.0% Hispanic-White and 4.3% Black-White return gap in column 2, somewhat larger than the comparable 2.26% and 2.32% gaps Kermani and Wong report. This difference is likely due to the fact that our sample of 4 counties contains Chicago and Philadelphia,

poor/segregated cities with larger than usual racial disparities. The disparities we measure shrink as we expand our sample to 1975-2016 matching the AHS, include houses purchased without a mortgage, and re-weight our average over property ownership spells proportional to the length of the spell. This re-weighting is consistent with how we measure returns in the AHS, since each house-year observation gets equal weight and long ownership spells would have several observations.

In the final column, after we have controlled for time/county fixed effects and included the entire housing stock, even those that never transact, we find much smaller racial disparities in property appreciation rates than in column 1. The Black-White gap is now only 50 basis points, while the Hispanic-White gap is only 18 basis points. These disparities are very close to the 47 basis point and 11 basis point return disparities we see in the last column of table 6.8 with the most exhaustive set of controls. This suggests that the racial disparities in house price appreciation rates we observe in the AHS are consistent with data in Corelogic and with the findings in [Kermani and Wong \[2022\]](#) once differences in sample selection are accounted for.

Table 7.2: Racial Gaps in Log House Price Appreciation Rates: Reconciling with [Kermani and Wong \[2022\]](#)

	Purchased during 2000-2016	Never Sold	Purchase Year × County FEs	Purchased during 1975-2016	Cash Purchased	Tenure Length Weighted	Purchase Year × Sale Year × County FEs	Not Purchased
Panel A: Auction Price-based Return for Foreclosed Spells								
Black	-0.09913*** (0.00154)	-0.04313*** (0.00051)	-0.03234*** (0.00050)	-0.02845*** (0.00033)	-0.02207*** (0.00028)	-0.01348*** (0.00009)	-0.00930*** (0.00008)	-0.00505*** (0.00004)
Hispanic	-0.05285*** (0.00104)	-0.03037*** (0.00042)	-0.02019*** (0.00040)	-0.01479*** (0.00031)	-0.00965*** (0.00026)	-0.00655*** (0.00008)	-0.00291*** (0.00007)	-0.00180*** (0.00005)
Panel B: Outstanding Mortgage-based Return for Foreclosed Spells								
Black	-0.06741*** (0.00123)	-0.03322*** (0.00041)	-0.02501*** (0.00041)	-0.02382*** (0.00027)	-0.01922*** (0.00025)	-0.01303*** (0.00008)	-0.00812*** (0.00007)	-0.00441*** (0.00004)
Hispanic	-0.04349*** (0.00093)	-0.02675*** (0.00038)	-0.01689*** (0.00036)	-0.01290*** (0.00029)	-0.00839*** (0.00024)	-0.00633*** (0.00008)	-0.00255*** (0.00006)	-0.00157*** (0.00005)
Observations	476,115	1,129,740	1,129,740	3,011,164	5,486,006	5,486,006	5,486,006	7,497,430

Note: This table documents racial gaps in log house price appreciation for our Corelogic housing spell sample from 1975-2016 in four counties: Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA. Data are at the individual-spell level, with observations weighted by household each year. All numbers are adjusted for inflation. Standard errors (clustered at the household-spell level) are in parentheses.

As additional robustness, we explore the sensitivity of our average rate of return estimates in the AHS to potential mis-measurement of the impact of foreclosure sales. [Kermani and Wong \[2022\]](#) show that Black and Hispanic homeowners have a higher probability of foreclosure than White homeowners, especially during downturns. They provided us with a time series of estimates of the probability of foreclosure and size of the foreclosure discount for Black, White, and Hispanic homeowners, by year. Since our house price data is a survey-based measure of owners' perceived value of their house, it is possible that these values don't fully include the effect of foreclosure on home values. For example, if a house flipper buys a foreclosed home and then sells it to a new homeowner at a higher price between waves of our survey, we may miss the impact of this foreclosure. If, however, the new homeowner also buys at a distressed price, our data should likely reflect the impact of this foreclosure. Under the very conservative assumption that foreclosure discounts are entirely ignored in our estimates from the AHS, we show that our measure of the rates of return is still higher for Black and Hispanic homeowners than White homeowners with this discount added.

Table 7.3: AHS Log Rate of Return: 2001-2021

	Num. Obs.	Mean	Median	Std. Dev.
White	218,688	0.065	0.055	0.256
Black	24,792	0.071	0.061	0.287
Hispanic	27,490	0.078	0.071	0.276

Note: Data from 2001-2021 owner-occupied properties in the AHS. The sample period is restricted to 2001-2021 since we only have data from [Kermani and Wong \[2022\]](#) on foreclosure discounts in this period. The table reports our baseline log rate of return estimates in this restricted sample period, without the foreclosure adjustment from [Kermani and Wong \[2022\]](#).

Table 7.4: AHS Log Rate of Return with added foreclosure discount from [Kermani and Wong \[2022\]](#): 2001-2021

	Num. Obs.	Mean	Median	Std. Dev.
White	218,688	0.064	0.053	0.256
Black	24,792	0.065	0.056	0.288
Hispanic	27,490	0.072	0.066	0.277

Note: Data from 2001-2021 owner-occupied properties in the AHS and augmented with data shared by [Kermani and Wong \[2022\]](#) on foreclosure rates and foreclosure discounts by race. These rates of returns make the conservative assumption that the AHS completely misses the effects of foreclosures on property values. See text for details of foreclosure adjustment.

We present our results in tables 7.3-7.4. The average return for White homeowners decreases by 0.1 pp, while Black and Hispanic homeowners respectively have a decrease of 0.6 pp and 0.8 pp in their average rates of return. These differences in returns are far smaller than the 1.37 pp, 3.30 pp, and 5.14 pp decreases in returns [Kermani and Wong \[2022\]](#) respectively report for White, Black and Hispanic homeowners when including foreclosure sales relative to a non-distressed only subsample. One key reason why our average rate of return estimates are less sensitive to foreclosure discounts than those in [Kermani and Wong \[2022\]](#) is that we average over homeowners at a given time, while they average over realized transactions. The average of the one-period returns in equation 14 is over the I_t houses at time $t+1$, indexed by i , is:

$$\frac{1}{I_t} \sum_i [\log(P_{it+1}) - \log(P_{it} - \text{dividend}_{it})]. \quad (15)$$

Suppose each house has a probability $pr_{for,t}$ of foreclosure and is then sold for a fraction f_t of its non-foreclosure value P_{it+1}^* . This average return can be written as

$$\frac{1}{I_t} \sum_i [\log(P_{it+1}^*) - \log(P_{it} - \text{dividend}_{it})] + pr_{for,t} \log(1 - f_t), \quad (16)$$

so the reduction in the annual rate of return is $pr_{for,t} \log(1 - f_t)$. We use the annual series of foreclosure rates $pr_{for,t}$ ⁶ and discounts f_t provided by [Kermani and Wong \[2022\]](#) for each racial group. To get a sense of magnitudes, for an annual foreclosure rate of 2%, close to the peak after the 2008 financial crisis, and a large 30% foreclosure discount, this would equal $.02 * \log(.7)$, a relatively modest 31 basis point decrease in the average rate of return.

In contrast, [Kermani and Wong \[2022\]](#) compute average returns by considering pairs of transactions where a house i is bought at a time $t_{1,i}$ for a price $P_{b,i,t_{1,i}}$ and then later sold at time t_2 for a price

⁶We sum the authors' foreclosure probabilities across the quarters in a year, which slightly overestimates the odds of foreclosure due to double counting of houses foreclosed twice.

$P_{s,i,t_2,i}$, with I total transactions. If we ignore dividend payments for this calculation, their average rate of return can be computed as

$$\frac{1}{I} \sum_i \left(\frac{P_{s,i,t_2,i}}{P_{b,i,t_1,i}} \right)^{\frac{1}{t_{2,i}-t_{1,i}}}. \quad (17)$$

Unlike our estimate, which uses the entire housing stock at a given time including homes that do not transact, this summary statistic considers only realized transactions. Suppose that a randomly selected fraction of transactions with probability $pr(\text{foreclose}|\text{transact})$ are foreclosures. The probability $pr(\text{foreclose}|\text{transact})$ is higher than the foreclosure probability we use, since most houses do not transact in any given year and all foreclosures occur among houses that are sold. If the non-distressed sale price of a house would be $P_{s,i,t_2,i}^*$ and all foreclosures have the same discount f , then this sales-based average return equals

$$pr(\text{foreclose}|\text{transact}) \frac{1}{I} \sum_i f^{\frac{1}{t_{2,i}-t_{1,i}}} \left(\frac{P_{s,i,t_2,i}^*}{P_{b,i,t_1,i}} \right)^{\frac{1}{t_{2,i}-t_{1,i}}}. \quad (18)$$

From this expression, we see that a foreclosure on a short tenure transaction has a large impact on this average rate of return statistic, since the discount f is taken to the power $\frac{1}{t_{2,i}-t_{1,i}}$, which can be large. Because only examining transactions increases the weight on foreclosures in this statistic, and because foreclosures on short holding periods can dramatically impact the implied rate of return, the return summary statistic in [Kermani and Wong \[2022\]](#) is more sensitive to foreclosure discounts than ours. While we believe that our notion of average return, which averages equally over all homeowners of a racial group in a given year without requiring them to transact, is an intuitive measure of expected return, there is no contradiction between our results and the greater impact of a foreclosure discount in [Kermani and Wong \[2022\]](#).

8 Conclusion

This paper quantifies racial differences in the total rate of return on owner-occupied housing. The total rate of return of buying a house equals the appreciation in its price plus the rental value of its housing services minus taxes and costs. Our results rely crucially on a new method we develop to impute the rental value of owner-occupied housing that performs well for both low-value and high-value homes. Unlike existing hedonic methods, our estimator accurately predicts rents of both low- and high-value homes by using price information to account for house quality. This accuracy across the value distribution is crucial for our application to racial inequality, since White, Black, and Hispanic homeowners vary in the average value of their homes. We document similar racial disparities in the rental yield of owner-occupied housing across multiple data sources, from raw survey data in the CEX to a more sophisticated imputation approach using actual transactions from Corelogic and MLS.

With our total return measure, we document a new stylized fact that Black and Hispanic homeowners earn higher average total returns on housing than White homeowners. This racial disparity in returns is primarily due to differences in rental yields across racial groups. Our new method of rent imputation that accurately matches rental yields across the house price distribution is crucial for measuring this

precisely. This rental yield disparity is largely driven by a higher rent-to-price ratio for lower value homes, and controlling for homeowner income/education largely explains the return disparity it creates. In our nationally representative sample from 1974-2021, including the many homes that do not transact, differences in house price appreciation rates across racial groups are somewhat smaller. We also find that Black and Hispanic homeowners face riskier and more cyclical returns than White homeowners, and that the excess negative returns Black homeowners face in recessions cannot be explained by our observables. However, our results imply that at least on average, racial disparities in the total rate of return on owner-occupied housing reduce the racial wealth gap over time, once we account for rental yields and all other relevant cashflows.

References

- John Y. Campbell. Household Finance. *Journal of Finance*, 61(4):1553–1604, 2006.
- Ellora Derenoncourt, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. Wealth of two nations: The U.S. racial wealth gap, 1860-2020. *Quarterly Journal of Economics*, forthcoming, 2023.
- Robert Barsky, John Bound, Kerwin Ko’ Charles, and Joseph P Lupton. Accounting for the black–white wealth gap: a nonparametric approach. *Journal of the American statistical Association*, 97(459):663–673, 2002.
- Moritz Kuhn, Moritz Schularick, and Ulrike I Steins. Income and wealth inequality in america, 1949–2016. *Journal of Political Economy*, 128(9):3469–3519, 2020.
- Arpit Gupta, Christopher Hansman, and Pierre Mabilie. Financial Constraints and the Racial Housing Gap. *working paper*, 2023.
- Kerwin Kofi Charles and Erik Hurst. The Transition to Home Ownership and the Black/White Wealth Gap . *Review of Economics and Statistics*, 84(2):281–297, 2002.
- Amir Kermani and Francis Wong. Racial Disparities in Housing Returns. *working paper*, 2022.
- Prottoy A. Akbar, Sijie Li Hickly, Allison Shertzer, and Randall P. Walsh. Racial Segregation in Housing Markets and the Erosion of Black Wealth. *The Review of Economics and Statistics*, pages 1–45, 11 2022. ISSN 0034-6535. doi: 10.1162/rest_a_01276. URL https://doi.org/10.1162/rest_a_01276.
- Edward N. Wolff. Heterogenous rates of return on homes and other real estate: Do the rich do better? Do Black households do worse? *working paper*, 2022.
- Matthew E. Kahn. Racial and ethnic differences in the financial returns to home purchases. *Real Estate Economics*, 52(3):908–927, 2024. doi: <https://doi.org/10.1111/1540-6229.12475>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6229.12475>.
- David Chambers, Christophe Spaenjers, and Eva Steiner. The rate of return on real estate: Long-run micro-level evidence. *The Review of Financial Studies*, 34(8):3572–3607, 2021.

- Piet Eichholtz, Matthijs Korevaar, Thies Lindenthal, and Ronan Tallec. The total return and risk to residential real estate. *The Review of Financial Studies*, 34(8):3608–3646, 2021.
- Andrew Demers and Andrea Eisfeldt. Total Returns to Single Family Rentals. *working paper*, 2001.
- Caitlin Gorback and Gregor Schubert. The Impact of Cultural Preferences on Homeownership. *working paper*, 2024.
- Jonathan Halket, Lara Loewenstein, and Paul S. Willen. The cross-section of housing returns. 2024.
- Sonia Gilbukh, Andrew Haughwout, Rebecca J Landau, and Joseph Tracy. The price-to-rent ratio: A macroprudential application. *Real Estate Economics*, 51(2):503–532, 2023.
- Stefano Colonnello, Roberto Marfè, and Qizhou Xiong. Housing yields. In *Proceedings of Paris December 2021 Finance Meeting EUROFIDAI-ESSEC*, 2024.
- Theodore M. Crone, Leonard I. Nakamura, and Richard P. Voith. Hedonic estimates of the cost of housing services: Rental and owner-occupied units. *Price and Productivity Measurement*, 4:51–68, 2009.
- Thesia I Garner and Kathleen Short. Accounting for owner-occupied dwelling services: Aggregates and distributions. *Journal of Housing Economics*, 18(3):233–248, 2009.
- Carlos Avenancio-Leon and Troup Howard. The Assessment Gap: Racial Inequalities in Property Taxation . *Quarterly Journal of Economics*, 137(3):1383–1434, 2022.
- Oscar Jorda, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan Taylor. The Rate of Return on Everything, 1870–2015. *Quarterly Journal of Economics*, 2019.
- Neil Bhutta and Aurel Hizmo. Do Minorities Pay More for Mortgages? . *Review of Financial Studies*, 34(2):763–789, 2021.
- Peter Ganong, Damon Jones, Pascal J Noel, Fiona E Greig, Diana Farrell, and Chris Wheat. Wealth, race, and consumption smoothing of typical labor income shocks. *working paper*, 2024.
- Sasha Indarte, Bronsyn Argyle, Ben Iverson, and Christopher Palmer. Explaining Racial Disparities in Personal Bankruptcy Outcomes . *working paper*, 2023.
- Paul Goldsmith-Pinkham and Kelly Shue. The gender gap in housing returns. *The Journal of Finance*, 78(2):1097–1145, 2023.
- Patrick Bajari and C. Lanier Benkard. Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy*, 113(6), 2005.
- Patrick Bajari, C. Lanier Benkard, and John Krainer. House Prices and Consumer Welfare. *Journal of Urban Economics*, 58(3), 2005.

- Morris A. Davis, Andreas Lehnert, and Robert F. Martin. The rent-price ratio for the aggregate stock of owner-occupied housing. *The Review of Income and Wealth*, 54(2):279–284, 2008.
- Jonah B. Gelbach. When Do Covariates Matter? And Which Ones, and How Much? *Journal of Labor Economics*, 34(2), 2016.
- Carlos F Avenancio-León and Troup Howard. The assessment gap: Racial inequalities in property taxation. *The Quarterly Journal of Economics*, 137(3):1383–1434, 2022.
- Benjamin Elbers. Trends in u.s. residential racial segregation, 1990 to 2020. *Socius*, 7:23780231211053982, 2021. doi: 10.1177/23780231211053982. URL <https://doi.org/10.1177/23780231211053982>.
- Junting Ye, Shuchu Han, Yifan Hu, Baris Coskun, Meizhu Liu, Hong Qin, and Steven Skiena. Nationality Classification Using Name Embeddings. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, page 1897–1906, 2017.
- Rebecca Diamond, Tim McQuade, and Franklin Qian. The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco. *American Economic Review*, 109(9):3365–3394, 2019.