Appendix

A Formal Statement of Rank Stability Assumption

We state a version of the rank stability assumption behind our rent imputation method that accounts for the fact that rent and price observations on a given home cannot be seen at the same time. Our rent imputation procedure uses "switcher" homes that either have a rent observation at time t and a house price observation at time t+1, or a house price observation at time t and a rent observation at time t+1. In addition, it is possible that the observed rent and price ranks we see do not perfectly satisfy our single index assumption. Allowing for possible violations of the single index assumption, we assume that the joint distribution of rent and price ranks is stable over time.

Assumption 2 *(Rank stability). Let* $(\rho_{rh,t}, \rho_{rp,t})$ *be the rank of house h's rent r and price p within the distribution of all observed rents and prices at time t. These ranks are normalized to be uniformly distributed on* $[0, 1]$

Let $F_{switcher,t,t+1}(\rho_{rh}, \rho_{ph}, X_h^o)$ *be the joint CDF of rent ranks and price ranks observed for homes that switch between being rented and being owner-occupied between times t and t+1, conditional on observables* X_h^o*. We assume that this joint CDF is the same function* $F_{switcher}(\rho_{rh}, \rho_{ph}, X_h^o)$ *at all times t.*

Under the idealized version of our single index assumption, we showed in equation 9 that one can exactly solve for the price rank of a house by knowing its observable characteristics and its rent rank. Assumption 2 allows for there to be a non-degenerate distribution of price ranks given an observed rent rank but requires this distribution not to vary over time. If, for example, renters care more about the (unobserved) quality of restaurants and coffee shops near a home, while owners care more about the (unobserved) quality of local schools, this relative difference in preferences could lead to small violations of our single index assumption in the main text.

Even if our single index assumption does not hold, the joint distribution $F_{switcher}(\rho_{rh}, \rho_{ph}, X_h^o)$ can be observed directly from data. As a result, the median regression we run in equation 10 still provides a median-unbiased prediction of a house's rent rank given its price rank and observables without any additional assumptions. However, the purpose of the regressions we run on the switcher population is to impute rent estimates for the population of owner-occupied houses. We therefore have to assume that the rents we observe for switchers are representative of the rent of owner-occupied homes, conditional on home observables and prices.

Assumption 3 *(Representativeness) Let* $F_{owner}(\rho_{rh}, \rho_{ph}, X_h^o)$ *be the joint distribution of rent ranks between what an owner-occupied house would hypothetically rent for and the price rank of its observed price, conditional on observables* X_h^o*. We assume that this is the same joint distribution as the object* $F_{switcher}(\rho_{rh}, \rho_{ph}, X_h^o)$ *we can estimate from the switcher population*

This representativeness assumption trivially holds under our single index assumption, since one can perfectly solve for the rent value of a home by knowing its observables and purchase price. However, it is strictly more general. The assumption allows for some degree of selection between owned and rented

properties, if for example unobservably higher quality homes tend to be owned rather than rented. However, all of the selection on unobservables must be due to unobservables which are reflected in the house's price. Under this assumption, a median-unbiased prediction of the rent rank of a switcher also yields a median-unbiased prediction of the rent rank of an owner-occupied home with the same observables and price rank.

B Imputing Rents Using Corelogic and MLS

This appendix describes our methodology for estimating the rental yields of owner-occupied housing in the Corelogic database. Unlike the AHS, Corelogic is a large database of realized property transactions rather than a year-by-year survey of houses regardless of their transaction status. Corelogic has the benefit of being a considerably larger database than the AHS, but it has the downsides of not being geographically representative and only reporting property values at the time of a transaction. We therefore develop a modified rent imputation procedure tailored to the strengths and weaknesses of this database.

Our analysis begins with a database of rents from Multiple Listing Services (MLS), where real estate agents list properties for sale and for rent. We use all rental observations from the MLS where the rent paid by an actual tenant who occupies the property is reported and merge these rent observations to our property transaction data using each property's APN. This gives us a database of "switcher" properties for which we jointly observe rents and sales prices, allowing us to use the sales price of a property to proxy for its unobserved quality. Because this database is considerably larger than the AHS, we are able to run regressions predicting the rent of switcher properties separately by MSA-Year.

With each MSA-Year, we collect all MLS rental observations and merge them to all housing purchases on the same housing unit that occur within at most 10 years. We restrict this data to only include transactions where either the buyer or seller is an owner occupant, based on matching the name of the resident in Infutor to the Corelogic deed transaction. We then use our name-location-based measure of the owner's race, housing unit covariates from Corelogic's tax history file, and the price of the property to predict our rental yields. In doing so, we must account for the fact 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 do so, we compute the quantile of each housing unit's price in the distribution of house values in the American Community Survey in the same year and county.[7] We then assign the house an imputed purchase price with the same quantile during the year of the rental transaction as the house had during the year of the purchase transaction.

For our baseline specification, we run the following regression of housing rents on housing prices and

[7]For some observations in the ACS, we observe the PUMA but not the county FIPS of a home for data privacy reasons. For such homes, we compute the fraction of residents in the PUMA that reside in each neighboring county. We then treat the observation as if a fraction of it was seen in each neighboring county, weighted by the population weights of the PUMA. Our marginal distributions of house prices are computed after these fractional assignments of units which are missing a county FIPS.

other covariates in a given county and a given year

$$rent_{ht} = \alpha_{jt} + \beta_{jt} X_{ht} + \epsilon_{ht}. \tag{19}$$

The regression includes the following covariates X_{ht}. First, it includes the full set of interactions between the price and its square, the absolute value of the difference in years between the year t at which the unit is rented and the year of the housing transaction from which the price $price_{ht}$ is derived, and an indicator function for whether this difference in years is positive. Any term composed of multiplying together one, two, three, or all four of these four variables is included in the regression. In addition, we include linear and quadratic terms in the square footage of the housing unit, a linear term in the age of the housing unit, a dummy variable for whether the housing unit is a single family residence rather than an apartment, a dummy for each racial group of the property owner (Black/Hispanic/White), and zipcode fixed effects. On top of this, we include when possible the number of bedrooms, bathrooms, and total rooms in a housing unit. For each of these three remaining variables, we only include them in the specification when they are missing in no more than 10% of observations. In addition, we trim the top and bottom 1% of our rent and price variables each year to account for errors in data entry and outliers. We require a minimum sample size of 1,000 observations to run regression 19 in a given MSA-year. If this threshold fails, we run the same regression pooled across years with year fixed effect added when there is a minimum of 2,000 observations, in which case we replace our zipcode fixed effects with zipcode-year fixed effects.

After estimating the coefficients in equation 19, we use them to construct fitted values for the rental value of owner-occupied housing units. We do this for any housing unit for which we see a transaction in 1990 or later and for which the covariates needed to impute a rent are not missing. Finally, we compute the rental yield of an owner-occupied housing using by dividing our predicted rent value by the house's resale value $price_{ht}$. In years where the house does not transact, we construct this resale value in the same manner as we did to construct our house price covariates for equation 19. We first compute the quantile of the house's value within the county-year specific house price distribution in the ACS when it transacted. We then find the price level at the year for which we are imputing a rent which would have the same quantile in that county-year price distribution. The final data objects we analyze are the ratio of our rent fitted values divided by this time-matched value of the resale value of a home.

Geographical Representativeness of Corelogic-MLS Sample

To construct our rental yield estimates in Corelogic, we need accurate data on a house's characteristics, the race of its owner, and both a resale value of the house and data on its rent. This restricts us to a sample of 2007-2016, since our data on a house's characteristics from Corelogic tax history begin in 2007. In addition, MLS data exists only in certain subsets of the country, meaning that the unconditional means of our estimated rental yields cannot directly be compared to the nationally representative results in the AHS.

Table B.1 documents how our sample differs from the overall country. Our data comes from places that are richer, less White, and more concentrated in the West than the country as a whole. The fact

that the West tends to have both a large Hispanic population as well as high house prices and low rent-to-price ratios makes aggregate statistics from this data difficult to compare to the AHS. We therefore compare our Corelogic/MLS data on rental yields to the AHS only in a sample that is matched both on time and on geography.

Table B.1: Summary Statistics Based on County-Level ACS Data

	Whole US (Population Weighted)	Counties in Sample (Sample Weighted)
2012-2016 White Share (%)	61.96 [22.07]	56.31 [17.79]
2012-2016 Black Share (%)	12.28 [12.65]	13.39 [11.03]
2012-2016 Hispanic Share (%)	17.3 [16.90]	20.00 [14.70]
2012-2016 Median Housing Value ($)	231,282.74 [146,813.42]	273,331.84 [146,434.20]
2012-2016 Mean Household Income ($)	78,182.99 [20,647.29]	87,492.76 [21,114.44]
2012-2016 Northeast Share (%)	17.60	12.29
2012-2016 Midwest Share (%)	21.24	21.31
2012-2016 South Share (%)	37.59	33.28
2012-2016 West Share (%)	23.56	33.13

Note: This table presents summary statistics for the counties included in our Corelogic-MLS sample and compares them to summary statistics on the overall country. We compute these statistics by aggregating county-level data from the American Community Survey (ACS) in 2012-2016. 38.41% of the country's population lives within counties where we can construct rental yields using the Corelogic-MLS data. Standard deviations are in brackets.

Summary Statistics in our Corelogic-MLS Sample

Table B.2 shows the summary statistics of unit attributes and other covariates in the rental regression for switchers, owned, and rented homes in the Corelogic-MLS sample. We find that switcher homes are somewhat less expensive than the overall population of owner-occupied homes, despite having similar numbers of rooms, bedrooms, and bathrooms. Owner-occupied homes are also somewhat older and more likely to be single family homes. We find that the average monthly rent we imputed for owner-occupied homes is $2045.98, higher than the average $1626.57 we see for switchers. MLS renters have an average monthly rent of $1646.33, very close to the average for switchers. Rentals are less likely to be single family homes than either switchers or owner-occupied homes and have slightly lower square footage. Overall, MLS switcher homes seem to be of similar quality to other rental homes, but owner-occupied homes seem to be of higher quality. The average owner-occupied home sells for $487,928, well above the average $318,759 price of a switcher. Because our rent imputation procedure uses a home's price to help measure quality, it is ideally suited to our sample, where the price differences between switcher and owner-occupied homes seem larger than the differences in their hedonic characteristics. These patterns in relative quality across owners/renters/switchers are similar to those presented in table 3.1 for the AHS.

Table B.2: Summary Statistics of Corelogic-MLS sample

	MLS Switchers	Owner-Occupied Sample	MLS Rentals
Monthly Rent ($)	1,626.57	2,045.98	1,646.33
	[908.38]	[1,270.25]	[965.16]
Monthly Rent ($, in 2019)	1,826.69	2,298.12	1,837.75
	[1,012.53]	[1,423.88]	[1,077.82]
Year of Rental Contract	2,012	2,012	2,012
	[2.57]	[2.75]	[2.59]
Building Square Footage	1,919.88	2,210.11	1,672.86
	[2,008.39]	[1,165.03]	[746.43]
Total Number of Rooms	6.28	6.63	6.00
	[1.84]	[1.69]	[1.70]
Number of Bathrooms	2.51	2.43	2.36
	[1.66]	[1.67]	[0.88]
Number of Bedrooms	3.12	3.20	2.93
	[1.16]	[0.98]	[1.05]
Building Age	24.30	36.27	27.62
	[24.10]	[25.09]	[23.44]
Share of Single Family Residence (%)	80.29	91.20	65.88
Transaction Price ($)	262,606.96	366,449.89	-
	[213,587.63]	[363,719.55]	-
Transformed Price ($)	261,200.44	343,419.64	-
	[227,919.37]	[366,145.06]	-
Transaction Price ($, in 2019)	318,759.17	487,927.87	-
	[252,927.95]	[499,419.48]	-
Transformed Price ($, in 2019)	294,605.59	386,524.52	-
	[256,667.19]	[410,566.70]	-
Price Quantile (%, County Level)	48.53	52.23	-
	[24.95]	[27.14]	-
Year of Transaction	2,009	2,006	-
	[5.09]	[7.24]	-
Diff.: Transaction Year - Rental Year	-3.47	-6.35	-
	[5.01]	[7.50]	-
Absolute Diff.: Transaction Year - Rental Year	5.29	7.94	-
	[3.03]	[6.25]	-
Transaction Before Rental	0.72	0.70	-
	[0.45]	[0.41]	-
White Share (%)	87.85	82.88	-
Black Share (%)	4.70	6.29	-
Hispanic Share (%)	7.45	10.83	-
Observations	2,845,590	192,329,864	2,516,997

Note: This table presents summary statistics on rents and the main covariates we use to impute rents using equation 12. We have three populations: 1. "MLS Switchers" that have both rent observations from MLS and transaction prices from Corelogic deed. 2. "Owner-Occupied Sample" of all homes for which we are able to successfully impute a rent using equation 12. 3. "MLS Rentals" of all closed rental contracts in MLS, excluding advertised rents that do not result in a tenancy. The sample includes properties (single-family residences, condos, and townhouses) transacted or rented from 2006-2017 in the counties where the Corelogic-MLS data is available. The switchers' sample is at the transaction-rental pair level. The owner-occupied sample is at the housing unit-individual-year level. The rental sample is at the transaction level. We weight our data to ensure comparability. For switchers, data are weighted by the inverse of the number of transaction-rental pairs per housing unit each year. For the owner-occupied sample, data are weighted by the inverse of the number of residents in a housing unit each year. For the rental sample, data are weighted by the inverse of the number of rentals per housing unit each year. We topcode the total number of rooms at 10 for both the owner-occupied and rental samples. Additionally, for the MLS rental sample, we winsorize rent and unit square footage at the 1st and 99th percentiles and topcode the number of bathrooms at 5 and the number of bedrooms at 10. Standard deviations are in brackets.

C Property Tax Rates in Corelogic

This section shows that our results on racial disparities in property taxes are broadly consistent with the analysis of property taxes in Avenancio-León and Howard [2022]. Their headline result considers racial

disparities in the "assessment gap," the ratio between the assessed value of a home which determines property tax payments and the market value of a home in transactions. They find that Black homeowners have an assessment gap that is 12% higher than that of White homeowners. On the surface, this result seems to differ from our finding in table 6.6, where we find nearly the same property tax rate paid by each racial group. However, we show in table C.1 that our results can be reconciled and are internally consistent.

Their headline number of a 12% Black-White gap in assessment ratios comes from comparisons within taxation jurisdictions. We do not have data directly on the universe of tax jurisdictions, however they then add census tract-year fixed effects to this regression, which shrinks their Black-White assessment gap to 6.4%. We can run this specification in our data. The first column of table C.1 runs a regression that is nearly identical to the second regression in column 2 of table 2 of Avenancio-León and Howard [2022]. They find that in their sample from 2003-2016, Black homeowners have a 6.4% higher assessment ratio than White homeowners, after controlling for census-tract-year fixed effects. With these same controls, we find that Black homeowners have a 5.6% higher assessment ratio in our sample from 2007-2016.[8] Our 95% confidence interval of [.046, 0.067] contains their estimate.

In our next 3 columns, we first add Hispanic homeowners to our regression. Next, we reduce the geographic granularity of our fixed effects to MSA by year to match our analysis in the AHS. Finally, we add California back to our sample, which was dropped previously, and drop outlier property tax rates above 10% and assessment ratios that are not between 0.5 and 2. Column 4 shows that leaves us with a similar Black-White assessment gap of 6.5%.

In column 5 of table C.1, we switch the dependent variable of our regression from a homeowner's assessment ratio to the rate of return the household would get due only to property tax payments (setting appreciation, rents, and maintenance to zero), holding the sample fixed. Here we find a 38 basis point lower rate of return due to taxes for Black homeowners and 17.5 basis point lower return for Hispanic homeowners than for White homeowners. In the next two columns, we then add back in home transactions which were not preceded by a purchase mortgage and then finally include the entire housing stock inclusive of non-transacting homes. For non-transacting homes, we observe the nearest transaction in time and assign a property value at the time of the property tax observation that has the same value quantile, following the same procedure described in section B. In this final specification, we observe a 16.5 basis point lower rate of return from taxes for Black homeowners and 8.1 basis point lower rate of return for Hispanic homeowners relative to White homeowners.

In the final column of the table, we run an identical regression in the AHS, restricting ourselves to the MSAs that exist in our Corelogic data and matching our 2007-2016 time sample. Here we find a 16.6 basis point lower rate of return for Black homeowners and 21 basis point lower rate of return for Hispanic howeowners relative to White homeowners. This almost exactly matches the equivalent Corelogic specification for Black homeowners and if anything has a slightly lower rate of return for Hispanic

[8] Both our regression and that of Avenancio-León and Howard [2022] exclude California due to the unique impact of proposition 13 on its property tax assessments. We also restirct our sample to properties that transact in the sample year as the assessment measure and had a prior transaction in sample for the sample property that also originated a mortgage. We follow these sample restrictions from Avenancio-León and Howard [2022].

homeowners. This shows that within a geography- and time-matched sample, the racial disparities in property tax rates in the AHS are consistent with our data in Corelogic and with Avenancio-León and Howard [2022]. However, in the overall, nationally representative AHS sample from 1974 to 2021, racial disparities in property tax rates are close to zero on average.

Table C.1: Racial Gaps in Property Assessments and Rate of Return due to Taxes

	CoreLogic: Log(Assessment Ratio)				CoreLogic: Log(1 + Property Tax Rate)			AHS: Log(1 + Property Tax Rate)	
	Census Tract × Year FEs	Black and Hispanic	MSA × Year FEs	Add California/Drop Outliers	MSA × Year FEs	All Transacted	Non-transacted	Year FEs	MSA × Year FEs
Black	0.05646***	0.05795***	0.10589***	0.06473***	0.00381***	0.00336***	0.00165***	0.00248***	0.00166***
	(0.00536)	(0.00526)	(0.00795)	(0.00303)	(0.00023)	(0.00017)	(0.00014)	(0.000511)	(0.000485)
Hispanic		0.01921***	0.01811***	0.05013***	0.00175***	0.00159***	0.00081***	-0.000566	0.00211***
		(0.00261)	(0.00456)	(0.00233)	(0.00008)	(0.00006)	(0.00006)	(0.000458)	(0.000429)
Observations	5,942,075	6,088,651	6,088,651	5,410,397	5,410,397	17,921,862	200,452,594	26,857	26,857

Note: This table documents racial gaps in property assessment ratios and in effective property tax rates for the counties in our Corelogic sample from 2007-2016. The sample includes any owner-occupied housing units (single-family residences, condos and townhouses) with transactions recorded from 1990 onward and information on the owner's race. Standard errors (clustered at the Zip Code level for CoreLogic sample and the unit level for AHS sample) are presented in parentheses.

D Foreclosure Discount Measurement

This appendix considers the robustness of our measure of the rate of return on owner-occupied housing to different ways of measuring foreclosure discounts. In a foreclosure, borrowers who do not pay an existing mortgage have their homes auctioned off in order to recoup losses for the lender. We consider two ways of measuring the price a borrower receives in a foreclosure. First, we examine their unlevered return on housing (ignoring dividends) as if they sold their home at the foreclosure auction price. Second, we compare this to a rate of return (ignoring dividends) where they are considered to sell their home to their lender at the balance of their outstanding mortgage, with any remaining losses on the foreclosure being born by the lender rather than the borrower.

We estimate the outstanding balance on a property's mortgage as follows. First, we assume that borrowers pay a mortgage interest rate equal to the average 30-year fixed mortgage interest rate at the month of mortgage origination. Then, we use the following expression to calculate the current outstanding mortgage balance:

$$\text{Outstanding Mortgage Balance} = P \times \frac{(1+\frac{r}{12})^n - (1+\frac{r}{12})^m}{(1+\frac{r}{12})^n - 1}. \tag{20}$$

Here P is the principal amount of the mortgage loan, r is the annual mortgage interest rate, n is the total number of payments (e.g., n equals 360 for a 30-year fixed-rate mortgage), and m is the number of payments the homeowner has made. For foreclosure cases, we assume the default occurs 6 months before an observed foreclosure auction, resulting in 6 months of unpaid interest. Our final expression for the amount the borrower owes at the time of foreclosure is then

$$\text{Outstanding Mortgage Balance Amount} \times [1 + (\frac{r}{12} \times 6)].$$

For each foreclosed-upon homeowner, we calculate the annual house price appreciation rate as

$$\text{Annualized Log Appreciation Rate} = \frac{log(P_t) - log(P_0)}{\text{Num. Years of Occupation}},$$

where P_t is the sale price and P_0 is the purchase price. We examine how our average rate of house price appreciation varies with how we compute the sale prices in foreclosures.

We compare our data in Corelogic to a more detailed dataset, RIS, with information on all foreclosures in Cook County, IL. We match these two datasets based on a property's APN and the year and month of a foreclosure transaction from 2000 to 2016.[9] In each foreclosure, we observe a "complaint amount" which represents the amount of remaining money the lender believes is owed by the borrower.[10] We show that this complaint amount is approximated well by our method using outstanding mortgage balances, at least for the set of households who never refinance their mortgages.

Table D.1 presents regressions that quantify the relationship between different measures of a home's value in foreclosure in Cook County, IL. For households which never refinance their mortgages, our outstanding mortgage balance imputation predicts the complaint amount with a slope of 1.037 and an R-squared of .896. This shows that our mortgage balance imputation procedure is an accurate proxy for the true complaint amount in a foreclosure for homeowners who do not refinance. This R-squared falls to .599 when we separately examine homeowners who refinance, since we are not able to tell which refinanced mortgages did or did not involve prepaying old ones. The relationships we find between complaint amount and foreclosure auction prices, as well as auction prices and imputed mortgage balances, are weaker. This shows that a foreclosure auction price is not a particularly accurate measure of the amount of debt a homeowner erases in a foreclosure. These results suggest that we accurately estimate the value of debt a non-refinancing homeowner erases in a foreclosure with our imputed outstanding mortgage balance. However, auction prices seem only moderately related to the complaint amounts reported in RIS.

Table D.1: Regressions showing the relationship between foreclosure price measures in Cook County, IL using RIS and Corelogic Data

Sample	Y-variable	X-variable	Coefficient	R-squared
All Foreclosures	Complaint/Judgement Amount	Foreclosure Auction Price	0.839 (0.009)	0.567
Non-Refinancer	Complaint/Judgement Amount	Outstanding Mortgage Balance	1.037 (0.002)	0.896
Non-Refinancer	Foreclosure Auction Price	Outstanding Mortgage Balance	0.719 (0.016)	0.485
Refinancer	Complaint/Judgement Amount	Outstanding Mortgage Balance	1.255 (0.005)	0.599
Refinancer	Foreclosure Auction Price	Outstanding Mortgage Balance	0.993 (0.014)	0.543

Note: This table presents regression results of each price measure listed as the Y-variable on the alternative price measure listed as the X-variable. We control for interacted fixed effects both for the year of foreclosure and the year of mortgage origination. The data are at the foreclosure case level, with multiple foreclosures on a home in the same month weighted as a single observation. Complaint/Judgement Amount comes from RIS, while Outstanding Mortgage Balance and Foreclosure Auction Price are from Corelogic. Standard errors are in parentheses.

Next, we consider in table D.2 how our choice of foreclosure price measurement impacts the rate of

[9] When we see foreclosures that match on APN and year but not month, we include these as well but weigh any multiple matches for an APN-year as a single observation.

[10] We drop matched cases with missing complaint amounts. We also limit our analysis to foreclosures that report a purchase price and at least one of a purchase mortgage amount or foreclosure auction price in Corelogic. We restrict our analysis to complaint amounts, purchase mortgages, and auction prices between $10,000 and $2,000,000. Additionally, for cases with non-missing purchase mortgages, the initial LTV must be less than 1.05 at purchase.

house price appreciation earned by homeowners who experience foreclosure. We find that using auction prices results in considerably lower rates of house price appreciation than either our mortgage balance imputation or using RIS complaint amounts. For homeowners who never refinance, the mortgage balance yields a rate of house price appreciation that is only around 1-2 percentage points lower than the rate implied by using the complaint amount, while the gap is larger for homeowners who refinance. In all specifications, we find that the RIS complaint amount results in foreclosed homeowners earning a higher rate of house price appreciation than either using a mortgage balance imputation or an auction price. If we believe that the return earned by a homeowner should be measured by them "selling" their house for the amount of outstanding debt they erase in the foreclosure, then our mortgage imputation method should yield a somewhat more negative rate of return than the baseline truth provided in RIS.

Finally, in table D.3, we analyze homeowners' rates of house price appreciation in the 4 counties where we have carefully measured foreclosure sales. In Cook County and especially Philadelphia County, we find that foreclosure auctions imply considerably lower rates of house price appreciation earned by foreclosed homeowners than our method that imputes outstanding mortgage balances. For King County and Santa Clara County, we find if anything that foreclosure auction prices suggest a slightly higher rate of house price appreciation. In panel B of this table, we find that overall rates of house price appreciation are only minimally sensitive to our method of computing foreclosure discounts, except in Philadelphia where overall rates are 1.3% higher using the mortgage-based method than the auction price in foreclosure cases.

Table D.2: Log Housing Appreciation Rates For Foreclosed Properties Based on Different Measures of Foreclosure Discounts in Cook County, IL

	Complaint Amount-based (RIS)				Outstanding Mortgage-based (Corelogic)				Foreclosure Auction Price-based (Corelogic)			
	Mean	Median	Std. Dev.	Num. Obs.	Mean	Median	Std. Dev.	Num. Obs.	Mean	Median	Std. Dev.	Num. Obs.
Panel A: All Foreclosures in Corelogic/RIS Match												
Non-Refinancer	-0.050	-0.038	0.094	43,914	-0.069	-0.053	0.086	43,914	-	-	-	-
Refinancer	-0.010	-0.010	0.062	45,286	-0.054	-0.046	0.049	45,286	-	-	-	-
All Foreclosures	-0.029	-0.024	0.082	89,200	-0.061	-0.048	0.070	89,200	-	-	-	-
Panel B: Only Foreclosures with Non-Missing Auction Price												
Non-Refinancer	-0.038	-0.040	0.192	2,383	-0.051	-0.053	0.170	2,383	-0.168	-0.142	0.250	2,383
Refinancer	-0.011	-0.017	0.118	3,985	-0.046	-0.048	0.110	3,985	-0.077	-0.069	0.143	3,985
All Foreclosures	-0.021	-0.026	0.151	6,368	-0.048	-0.050	0.135	6,368	-0.111	-0.092	0.195	6,368

Note: This table presents summary statistics on log housing appreciation rates based on different measures of foreclosure discounts in Cook County, IL. The data includes cases with observed foreclosure auctions from 2000 to 2016 in Corelogic deeds which can be matched to RIS. All numbers are adjusted for inflation.

Table D.3: Overall Log Housing Appreciation Rates Based on Different Measures of Foreclosure Discounts

	Outstanding Mortgage-based				Foreclosure Auction Price-based			
	Mean	Median	Std. Dev.	Num. Obs.	Mean	Median	Std. Dev.	Num. Obs.
Panel A: Foreclosure Sales Only								
Cook	-0.026	-0.019	0.042	36,249	-0.041	-0.016	0.134	36,249
King	-0.035	-0.029	0.057	24,619	0.009	0.029	0.105	24,619
Santa Clara	-0.051	-0.042	0.078	10,862	-0.038	-0.006	0.200	10,862
Philadelphia	-0.017	-0.013	0.031	31,101	-0.174	-0.086	0.294	31,101
Panel B: All Sales								
Cook	0.072	0.073	0.140	1,408,185	0.072	0.073	0.142	1,408,185
King	0.088	0.086	0.111	806,746	0.090	0.086	0.111	806,746
Santa Clara	0.093	0.096	0.182	391,749	0.093	0.096	0.185	391,749
Philadelphia	0.080	0.067	0.156	310,625	0.063	0.067	0.197	310,625

Note: This table presents summary statistics on log housing appreciation rates for all complete housing spells observed in Corelogic in four counties: Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA. Panel A presents only spells that end with foreclosure sales, while Panel B presents all completed spells, including regular sales. All numbers are adjusted for inflation.

E Data Section: AHS

The data come from the American Housing Survey (AHS). The AHS is a nationally representative panel survey of housing units that has been conducted by the U.S. Census Bureau since 1973. The latest survey we use comes from 2021.

Housing costs

All monetary values are CPI-adjusted to 2019.

Property values

Property values in the AHS are directly reported by respondents, but the way they are recorded changes over time.

- Between 1973 and 1983, property values are categorical with bracket sizes that increase from $2,500 for low-value properties to $50,000 for high-value properties. We assign the mean value in each bracket to all properties in the bracket. To determine which value to assign to the top-coded bracket we utilize information on the distribution of property values in 1985 – when property values are first reported on a continuous scale. Specifically, we determine the percentile of properties in the top-bracket in each year before 1985 and compare this percentile to the distribution of property values in 1985. We then adjust the value for the top-coded bracket in each year before 1985 to match the 1985 distribution. Since property values in 1985 are top-coded, in some cases, i.e. when the percentile of properties prior to 1985 surpasses the top-coded percentile in 1985, there is no adjustment and the value assigned to units in the top-bracket remains unchanged.
- After 1985, property values are reported on a continuous scale and top-coded at levels that vary by metropolitan area.

There are some property values that are most likely misreported. First, we flag any property value below $10,000 in 2019 dollars are bad. This flags 1.5 percent of property value reports. Next, we flag the top and bottom one percent of annualized appreciation rates between survey waves as unreliable. This corresponds to annual returns less than -65 percent or greater than 195 percent. This flags extreme appreciation rates, but does not indicate whether the house value at time t or $t+1$ is the outlier value. To identify the outlier property value, we first look at adjacent appreciation rates for these houses with previous/subsequent survey ways. If the appreciation rate from $t-1$ to t is not an outlier, then we flag the property value at time $t+1$ as the bad value. If the appreciation rate from time $t+1$ to $t+2$ is not an outlier, then we identify the value at time t to be driving the outlier appreciation. For the remaining property values that cannot be flagged as causing the outlier appreciation rates from these methods, we use a hedonic regression of property values on property characteristics. Among the property values contributing to the outlier appreciation rates, we flag the property value as bad as the one which is further from the hedonic predicted property value.

We then impute missing property values. Property values are essentially never missing in the raw data for owned units since 1985. Prior to 1985, properties that on 10+ acres of land or were mobile homes were not asked their property values. We will exclude these properties from our data, as imputing them would be very hard since there are no observed property values for similar housing in the AHS during that time period.

We also need to impute property values for units that are owned in year t, but are not observed in the subsequent survey wave as still owner-occupied. If they are vacant or rented, the property value is not asked of the survey respondent, so we need to impute it. Alternatively, the unit could have exited the sample.

To impute missing property values, we run a hedonic regression of observed appreciation rates on property and household characteristics. We use the same covariates in our appreciation imputation as those included in our regressions decomposing our racial differences in total rates of return. We then impute the missing property value by taking the prior observed property value of the unit and multiplying it by the predicted appreciation rate from the hedonic regression.

Property taxes

Property taxes are reported on a continuous scale prior to 1985 and reported in steps of $50-$100 thereafter. For reports after 1985, we assign the mean value of each bracket to all observations in the bracket. Property taxes are also top-coded. We predict the property tax in the top-coded bracket using a regression of property taxes on property values and region-by-metro fixed effects. To impute missing reports and impute property taxes for units that are owner-occupied in t but not t+1 we regress property taxes on property value, race, year, building year cohort, and region-by-metro fixed effects.

Maintenance cost

Our variable for maintenance cost includes routine maintenance cost as well as home improvement projects. The AHS contains information on maintenance costs after 1985. Routine maintenance is asked for directly and we determine home improvement costs by aggregating information from a number of variables on cost for home improvement jobs (e.g. for kitchen, bath, etc.), alterations, and additions, over the past two years and halving the aggregate to obtain annual maintenance cost.

To impute missing maintenance cost for units that are owned in t, we regress maintenance cost on property value, property tax, building age (current year - built year), building age squared, unit size (in square feet), geographic region, and race dummies. The regression equation is also fully interacted with the race dummies. The maintenance cost variable is only available after 1985 so we also impute maintenance cost prior to 1985. To do so, we predict maintenance cost in the last period the unit is owned, using the same regression as described above, and linearly extrapolate backwards using the average change in maintenance cost across the reporting horizon (1985 to 2021). Last, we predict maintenance costs for units that are owner-occupied in t but no longer in t+1 using the same regression of maintenance cost on observables as above.

F Data Section: Corelogic

This section presents details on how we assembled our large database of housing transactions in Corelogic. Corelogic data comes from 4 separate sources: tax data, deed data, tax history data, and MLS data. Tax data provides information about the names of residents and the characteristics of individual homes at the end of our sample in 2017. Tax history provides analogous data going back each year until 2007. Deed data provides a systematic list of all housing purchases and mortgage issuances, providing information about the names of people active in each transaction, the price of any purchase or sale, the quantity lent in any mortgage, and a unique house identifier. Finally, the subset of MLS data we use provides information about the monthly rent advertised to live in a property at a given point in time. All of these datasets are easily merged to each other based on the Assessor's Parcel Number (APN) assigned to each property.

Corelogic-Infutor Merge

We merge all of these Corelogic databases to another database, Infutor, which provides information about the residents of each property. Infutor provides individual-level information about the name of residents, the sequence of addresses where they have lived over time, and a reported time at which they moved in and out. Merging Infutor to Corelogic allows us to compare the names of a home's owner listed in deed or tax data to the name of the resident, so that we are able to identify owner-occupied housing.

Our merge of Infutor to Corelogic proceeds in two steps. First, we attempt to merge the data based on the listed zipcode and address string of each property. The string merge takes place in several steps. First, we try to match addresses on entire strings, which yields 74% of our entire set of matched observations. We then look at additional matches with either have different zip codes, different city/town

names but identical zip codes, or missing apartment numbers or matches only on the numeric portion of apartment numbers (e.g. apartment 3 instead of 3A).

After our string merge, we attempt to use a spatial join to match additional addresses. To do so, we take the list of all our addresses from Infutor and feed them to the Geocodio geocoding service, who provided us with detailed latitudes and longitudes for each address. Next, we use parcel Polygons from a data provider Landgrid to map each APN in Corelogic to an associated geographic area. Finally, we observe all of the matches of latitudes and longitudes within each APN-associated area to construct matches from Infutor addresses to Corelogic APNs.

Table F.1 describes, of those rows of Infutor who are (A) not PO-Boxes, (B) whose cleaned addresses are not singletons (so that we see a nontrivial address history for the person), and (C) which successfully map to CoreLogic addresses, what proportion of mappings derive from each stage. There are of the 812 million rows that meet conditions (A) and (B), 84.0% can be successfully merged.

Table F.1: Success Rate of Merge between Infutor Person-Address Observations and Corelogic Addresses

Overall Match Rate	
Number of Total Infutor Observations	812 million
Fraction of Infutor Observations Merged	84.0 %
Fraction of Match Types Among Merged	
Complete String Matches	74.3 %
Incomplete String Matches	16.4 %
Spatial Join Matches	9.3 %

Determining if a Home is Owner-Occupied

Our Infutor-Corelogic match allows us to observe both the names of a housing unit's residents in Infutor and the names of its owners using deed and tax data from Corelogic. We determine whether a home is owner-occupied based on comparing these resident and owner names. For each housing unit, we obtain the most recent owner's name (or possibly multiple owner names) from Corelogic tax data and the names of all buyers and sellers of the property listed in Corelogic deed data in our sample. Since CoreLogic owner names sometimes include e.g. trusts (which sometimes contain owner last names i.e. "JOHN SMITH TRUST"), we remove a set of irrelevant strings that are unlikely to be named. We then employ two methods for identifying congruence of CoreLogic and Infutor names.

First, we calculate the Jaro-Winkler distance between the Infutor last name and each CoreLogic last name, which is a distance metric on the similarity of two strings. We say that a match is obtained if the string similarity is above 0.9. Second, we run a regular expression check for Infutor names being contained in each CoreLogic name (including searching for space- and hyphen-delimited tokens). If either of these matches, we say that the resident of a unit is an owner-occupant. Because we believe the Corelogic's buy and sell dates are more accurate than Infutor's dates when a resident moves in or out of a property, we use Corelogic's purchase and sale data to infer when an owner-occupant moves in or out of a property.

Identifying Homeowner Race

We impute homeowner race following [Ye et al., 2017, Diamond et al., 2019]. The first step uses the NamePrism machine learning algorithm to predict race given a homeowner's name. Next, the racial demographics of each census block are used to reweight the probabilistic race predictions of NamePrism using Bayes' rule. We then classify any homeowner with greater than a .8 chance of being White, Black or Hispanic as part of that racial group. This leads to a modest overestimate of the White share of the population, since the larger White population in the US will be reflected in the Bayesian weights coming from population shares. Moreover, Black and Hispanic homeowners are more likely to be identified in minority-dominant areas, leading to a greater degree of geographic segregation than in the true data. That said, name and geography are jointly a highly accurate predictor of race, and our large sample size allows for precise statistical inference even if we underestimate the size of the minority population.

For the set of owner-occupied homes we have identified, we use our name-based method of identifying race to infer the race of the homeowner. When there is a single homeowner whose name we can match between Infutor and Corelogic, this is straightforward. When there are several matching names in the same ownership spell, we assign the home the most common race of all of the house's matched owners. When we have the same number of homeowner name matches from two different racial groups in the same ownership spell, we choose of the two racial groups randomly and i.i.d. across such homes.

G Robustness: Rate of Return Using Alternative Rent Imputation Methods

This appendix considers the impact of three robustness checks on our measured racial disparities in the total rate of return on housing in the AHS. First, we replace our new method for rent imputation with a standard hedonic method and find that the disparities we measure would be even larger. Second, we run our quantile regressions for rent imputation separately across three distinct periods when the AHS refreshes its panel of homes and find similar results to the main text. Third, we separately estimate our rental yields first using only homes that switch from being rented to being owned and then only using homes that switch from being owned to being rented, whereas we use both directions of switchers together in our baseline approach.

Our hedonic rent imputation procedure is the same as used in section 4, using data on all AHS rentals to estimate the regression.

We present the results using hedonic rent imputation in tables G.1-G.2, which are analogous to tables 6.4 and 6.9 in the main text. In table G.1, the left column begins with a 2.85% higher rate of return for Black and 2.28% higher rate of return for Hispanic homeowners relative to White homeowners. In the main text, these disparities are only 0.62% and 1.19%. After accounting for our full set of controls in the fifth column of table G.1, we still have a remaining 1.8% Black-White return gap and 1.1% Hispanic-White return gap, while there is no remaining gap in the fifth column of table 6.4. This shows that if we had used hedonic methods to impute rental yields, we would have found an even higher rate of return for Black and Hispanic homeowners relative to White homeowners.

Table G.1: Log rate of return (Hedonic rent)

	(1)	(2)	(3)	(4)	(5)
Black	0.0285***	0.0280***	0.0290***	0.0158***	0.0184***
	(0.00142)	(0.00140)	(0.00139)	(0.00137)	(0.00134)
Hispanic	0.0228***	0.0240***	0.0252***	0.00826***	0.0108***
	(0.00163)	(0.00163)	(0.00163)	(0.00156)	(0.00155)
Observations	768,356	768,356	768,356	768,356	768,356
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. Rent is imputed using a hedonic method. 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 G.2: Log rate of return (Hedonic rent)

	(1)	(2)	(3)	(4)	(5)
Black	0.0294***	0.0294***	0.0304***	0.0168***	0.0193***
	(0.00170)	(0.00171)	(0.00171)	(0.00169)	(0.00163)
Black × Recession=1	-0.00294	-0.00488	-0.00473	-0.00350	-0.00318
	(0.00310)	(0.00314)	(0.00315)	(0.00316)	(0.00316)
Hispanic	0.0318***	0.0263***	0.0272***	0.00937***	0.0121***
	(0.00183)	(0.00190)	(0.00191)	(0.00184)	(0.00183)
Hispanic × Recession=1	-0.0350***	-0.00933*	-0.00812*	-0.00502	-0.00575
	(0.00359)	(0.00364)	(0.00367)	(0.00367)	(0.00361)
Observations	768,356	768,356	768,356	768,356	768,356
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. Rent is imputed using a hedonic method. 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.

In table G.2, we show that using a hedonic rent measure, we would have found that Hispanic but not Black homeowners have a greater return exposure to the business cycle than White homeowners. The Black/ recession interaction coefficient is not significant in any specification, while in table 6.9 this coefficient is significant at either the 90% or 95% confidence level in all specifications. In both tables 6.9 and G.2, the excess exposure to the business cycle of Hispanics in table can entirely be accounted for with observable controls used in the 4th and 5th columns, where the income/education of the homeowner is included.

Table G.3: Log rate of return (Separating Panels)

	(1)	(2)	(3)	(4)	(5)
Black	0.00490***	0.00478***	0.00548***	-0.00463***	-0.00211
	(0.00118)	(0.00119)	(0.00119)	(0.00119)	(0.00119)
Hispanic	0.0146***	0.0122***	0.0124***	-0.000131	0.00225
	(0.00135)	(0.00138)	(0.00138)	(0.00136)	(0.00133)
Observations	742,629	742,629	742,629	742,629	742,629
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. Rent is imputed using our switcher method and allows coefficients in switcher regression to change across the three distinct panels in the AHS: 1974-1983, 1985-2013, and 2015-2021. 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.

Tables G.3 and G.4 examine how our results change when we impute rental yields with separate regressions across 3 sub-periods when the AHS follows distinct panels of residents: 1974-1983, 1985-2013, and 2015-2021. Here, we find a 0.49% basis point rate of return for Black and 1.46% for Hispanic homeowners in the first column, close to the 0.62% and 1.19% we found in the first column of table 6.4. By column 5, we find that observable controls can account for effectively all of these racial disparities and slightly over-explain the Black-White return gap. In table G.4, we find like in table 6.9 that both Black and Hispanic homeowners have more cyclical returns than White homeowners. Like in the main text, observable differences by column 5 can account for most of the Hispanic-White gap but not the Black-White gap in exposure of housing returns to the business cycle.

Finally, tables G.5-G.8 use data constructed by running our switcher regressions for imputing rental yields by breaking the switcher population in two. Tables G.5-G.6 uses rental yields estimated only from properties that switch from owned to rented, while tables G.7-G.8 use rental yields estimated only from properties that switch from being rented to owned. We find that the estimated racial disparities in returns are somewhat higher when we use rent-to-own switchers than when we use own-to-rent switchers. However, even using the own-to-rent switchers, we find a 0.54% higher rate of return for Black homeowners and 1.04% for Hispanic homeowners relative to White homeowners in the first column of table G.5. These disparities increase to a Black-White return gap of 0.78% and a Hispanic-White return gap of 1.49% in the first column of table G.7. In both cases, we find by column 5 that observables can

Table G.4: Log rate of return (Separating Panels)

	(1)	(2)	(3)	(4)	(5)
Black	0.00793***	0.00799***	0.00862***	-0.00213	0.000189
	(0.00148)	(0.00149)	(0.00149)	(0.00149)	(0.00146)
Black × Recession=1	-0.0104***	-0.0112***	-0.0110***	-0.00859**	-0.00797**
	(0.00286)	(0.00291)	(0.00292)	(0.00293)	(0.00293)
Hispanic	0.0237***	0.0151***	0.0150***	0.00103	0.00359*
	(0.00159)	(0.00167)	(0.00168)	(0.00166)	(0.00163)
Hispanic × Recession=1	-0.0352***	-0.0111**	-0.0101**	-0.00537	-0.00590
	(0.00339)	(0.00343)	(0.00344)	(0.00346)	(0.00341)
Observations	742,629	742,629	742,629	742,629	742,629
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. Rent is imputed using our switcher method and allows coefficients in switcher regression to change across the three distinct panels in the AHS: 1974-1983, 1985-2013, and 2015-2021. 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 G.5: Log rate of return (Using only own-to-rent switchers)

	(1)	(2)	(3)	(4)	(5)
Black	0.00540***	0.00515***	0.00598***	-0.00465***	-0.00282*
	(0.00115)	(0.00116)	(0.00116)	(0.00116)	(0.00116)
Hispanic	0.0104***	0.00811***	0.00888***	-0.00462***	-0.00252
	(0.00135)	(0.00135)	(0.00135)	(0.00132)	(0.00129)
Observations	742,742	742,742	742,742	742,742	742,742
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. Rent is imputed using our switcher method only using switchers that switch from owner-occupied status in the prior period to rented in the current period. 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 G.6: Log rate of return (Using only own-to-rent switchers)

	(1)	(2)	(3)	(4)	(5)
Black	0.00722***	0.00752***	0.00826***	-0.00267	-0.000965
	(0.00145)	(0.00148)	(0.00148)	(0.00147)	(0.00144)
Black × Recession=1	-0.00618*	-0.00832**	-0.00804**	-0.00681*	-0.00641*
	(0.00283)	(0.00288)	(0.00289)	(0.00291)	(0.00291)
Hispanic	0.0190***	0.0102***	0.0106***	-0.00391*	-0.00158
	(0.00159)	(0.00165)	(0.00165)	(0.00162)	(0.00160)
Hispanic × Recession=1	-0.0335***	-0.00827*	-0.00674*	-0.00340	-0.00413
	(0.00335)	(0.00338)	(0.00338)	(0.00340)	(0.00334)
Observations	742,742	742,742	742,742	742,742	742,742
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. Rent is imputed using our switcher method only using switchers that switch from owner-occupied status in the prior period to rented in the current period. 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.

either entirely account for or more than account for this racial disparity in returns.

We find in tables G.6 and G.8 nearly identical results for how much more cyclical the returns of Black and Hispanic homeowners are compared to White homeowners. In both tables, we find a 3.3-3.4% greater drop in returns for Hispanic homeowners relative to White homeowners during recessions, which can be almost entirely explained by observable controls in column 5. In addition, we find a 0.62-0.68% greater drop in Black homeowner returns relative to White homeowners during recessions in column 1 of both tables, and this disparity cannot be explained by controlling for additional observables.

The fact that we find smaller racial disparities when using own-to-rent switchers mitigates a potential worry that racial disparities in foreclosure rates could be biasing our results. Because only own-to-rent switchers can follow a homeowner's foreclosure, we would expect a greater degree of foreclosures among minorities to depress their house prices relative to the rents earned on such houses after a foreclosure. This would result in higher measured rental yields for minorities, increasing our measured racial disparities in returns. We find the opposite result that estimating rental yields using rent-to-own switchers instead results in greater racial disparities in measured returns on housing.

Table G.7: Log rate of return (Using only rent-to-own switchers)

	(1)	(2)	(3)	(4)	(5)
Black	0.00782***	0.00755***	0.00829***	-0.00160	0.000246
	(0.00114)	(0.00115)	(0.00115)	(0.00115)	(0.00116)
Hispanic	0.0149***	0.0123***	0.0126***	0.0000725	0.00239
	(0.00138)	(0.00140)	(0.00140)	(0.00137)	(0.00133)
Observations	742,817	742,817	742,817	742,817	742,817
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. Rent is imputed using our switcher method only using switchers that switch from rented in the prior period to owner-occupied in the current period. 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 G.8: Log rate of return (Using only rent-to-own switchers)

	(1)	(2)	(3)	(4)	(5)
Black	0.00981***	0.0101***	0.0108***	0.000541	0.00220
	(0.00144)	(0.00146)	(0.00146)	(0.00146)	(0.00145)
Black × Recession=1	-0.00680*	-0.00900**	-0.00868**	-0.00738*	-0.00677*
	(0.00283)	(0.00286)	(0.00288)	(0.00289)	(0.00290)
Hispanic	0.0234***	0.0145***	0.0144***	0.000866	0.00343*
	(0.00161)	(0.00168)	(0.00168)	(0.00166)	(0.00162)
Hispanic × Recession=1	-0.0331***	-0.00849*	-0.00699*	-0.00370	-0.00445
	(0.00344)	(0.00348)	(0.00349)	(0.00350)	(0.00345)
Observations	742,817	742,817	742,817	742,817	742,817
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. Rent is imputed using our switcher method only using switchers that switch from rented in the prior period to owner occupied in the current period. 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.