

Supplemental Appendix to “Racial Disparities in Housing Returns”

Amir Kermani and Francis Wong

February 11, 2025

A Appendix Tables

Table A1: Decomposition of Levered Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Raw Differences								
Black	-2.325 (0.0739)	-2.207 (0.0745)	-1.908 (0.0767)	-8.077 (0.443)	-5.011 (0.394)	-2.062 (0.378)	-2.170 (0.378)	-2.413 (0.340)
Hispanic	-2.264 (0.194)	-2.284 (0.196)	-2.890 (0.197)	-9.301 (0.967)	-9.618 (0.762)	-6.786 (0.754)	-7.750 (0.749)	-8.580 (0.596)
Panel B. Fixed Effect Estimates								
Black	-1.278 (0.0318)	-1.155 (0.0313)	-1.047 (0.0293)	-6.970 (0.188)	-2.464 (0.108)	-0.229 (0.109)	-0.298 (0.109)	1.252 (0.0880)
Hispanic	-1.295 (0.0359)	-1.232 (0.0360)	-1.182 (0.0353)	-6.330 (0.163)	-2.670 (0.115)	-1.221 (0.112)	-2.204 (0.112)	-1.283 (0.117)
Outcome Mean	2.355	1.452	6.169	8.762	4.898	5.183	5.043	0.805
Outcome SD	8.161	7.938	7.855	51.57	43.48	43.39	43.35	30.47
N	13,630,629	13,630,629	13,430,665	13,427,973	13,430,541	13,430,561	13,430,560	13,430,560
Sale Costs		X	X	X	X	X	X	X
Closing Costs			X	X	X	X	X	X
Rent Minus Upkeep			X		X	X	X	X
Leverage				X	X	X	X	X
Mortgage Payment					X	X	X	X
Minority Rent Adj.						X	X	X
Minority Costs Adj.							X	X
Cost-Weighted								X

Notes: This table presents estimates of the racial gap in housing returns using eight different measures of housing returns to illustrate the contribution of each component of levered returns. Within each panel, columns correspond to separate regressions estimating Equation 4. Panel A presents raw estimates, while Panel B applies fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, fixed effects that interact county, purchase year, and deciles of income at home purchase, and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. Each column corresponds to different assumptions imposed on annualized housing returns. *Sale Costs* indicates that a 5% transaction cost upon property sale has been assumed. *Closing Costs* indicates that closing costs at purchase have been imputed. In Column 3, these have been imputed under the assumption of no leverage (i.e., half the closing costs associated with a mortgage with 80% LTV). *Rent Minus Upkeep* factors in implicit rents, property taxes, and maintenance costs. *Leverage* incorporates loan balances that offset up-front costs and repaid at sale. *Mortgage Payment* factors in monthly mortgage payments. *Minority Costs Adj.* factors in racial differences in maintenance, interest payments, and property taxes. *Minority Rent Adj.* factors in racial differences in implicit rents. *Cost-Weighted* indicates that observations have been weighted by (strictly positive) net present value of costs (i.e., upfront costs, mortgage payment, taxes and insurance, and maintenance). Data are from baseline sample of ownership spells described in Section I. Standard errors are clustered within purchase year, sale year, and county cells.

Table A2: Unlevered Returns and Distressed Sales By Self-Reported and Imputed Race

	(1)	(2)	(3)	(4)	(5)
Panel A. Unlevered Returns (% Annualized)					
Hispanic	0.39 (11.95)	0.48 (11.90)	0.62 (11.64)	0.84 (12.83)	2.60 (12.99)
Black	0.38 (9.34)	0.68 (10.08)	0.65 (10.30)	0.69 (11.93)	2.53 (13.39)
White	2.70 (7.11)	2.66 (7.13)	2.80 (7.41)	2.90 (8.75)	4.14 (10.32)
Asian	2.62 (7.77)	2.76 (7.30)	2.61 (7.32)	2.86 (8.48)	4.01 (9.16)
Overall	2.24 (8.17)	2.25 (8.18)	2.39 (8.24)	2.48 (9.62)	3.86 (10.84)
Missing Race	2.40 (8.24)	2.66 (8.07)	2.88 (8.05)	2.58 (9.65)	3.78 (10.91)
B-W Gap	-2.31	-1.99	-2.15	-2.21	-1.61
H-W Gap	-2.31	-2.18	-2.18	-2.06	-1.54
Panel B. Distressed Sales (%)					
Hispanic	27.19	27.08	26.96	26.84	22.06
Black	22.40	22.71	23.61	23.92	20.05
White	11.97	12.21	12.78	12.54	9.58
Asian	13.31	11.86	11.99	12.74	9.94
Overall	14.78	14.82	15.06	15.13	11.69
Missing Race	15.03	13.00	12.95	13.72	9.93
B-W Gap	10.43	10.49	10.83	11.38	10.47
H-W Gap	15.22	14.87	14.18	14.30	12.48
N	14,111,108	15,354,155	1,415,112	25,794,573	40,570,537
Sample	Baseline (HMDA)	Baseline (HMDA)	Missing HMDA Race	All Mortgaged Purchases	All Purchases (Incl. Cash)
Purchase Years	2000–2014	2000–2014	2000–2014	2000–2014	1990–2014
Race Assignment	Self-Report	Imputed	Imputed	Imputed	Imputed

Notes: This table presents means and standard deviations of annualized unlevered returns by self-reported and imputed race/ethnicity. Self-reported race is derived from HMDA records, while imputed race is imputed using homeowner names and Census block. *Baseline* denotes baseline sample of ownership spells used for main analysis. *Missing HMDA Race* denotes baseline sample restricted to observations that are merged to HMDA but where self-reported race/ethnicity is missing. *All Mortgaged Purchases* denotes sample of purchases made with a mortgage but not restricted to merge with HMDA data. *All Purchases (Incl. Cash)* denotes sample of all purchases in the property data, not restricted to merge with the HMDA data. Data come from property records and HMDA records described in Section I.

Table A3: Returns and Distressed Sales with Alternative Adjustment for Censoring

	Mean (SD) (1)	Mean (SD) (2)
<i>Panel A. Unlevered Returns (% Annualized)</i>		
Black	0.41 (9.24)	1.28 (8.79)
Hispanic	0.72 (11.48)	2.01 (10.89)
White	2.81 (7.02)	3.15 (6.68)
Overall	2.34 (7.99)	2.84 (7.55)
Black-White	-2.39	-1.87
Hispanic-White	-2.09	-1.15
<i>Panel B. Total Returns (% Annualized)</i>		
Black	6.04 (8.85)	6.78 (8.32)
Hispanic	5.10 (11.25)	6.30 (10.54)
White	7.10 (6.65)	7.42 (6.30)
Overall	6.72 (7.64)	7.18 (7.17)
Black-White	-1.06	-0.64
Hispanic-White	-2.00	-1.12
<i>Panel C. Levered Returns (% Annualized)</i>		
Black	0.23 (43.47)	2.21 (38.80)
Hispanic	-2.75 (51.98)	0.29 (45.95)
White	4.29 (39.83)	5.27 (34.96)
Overall	3.00 (42.03)	4.33 (36.96)
Black-White	-4.06	-3.06
Hispanic-White	-7.03	-4.98
<i>Panel D. Distressed Sale (%)</i>		
Black	30.00	23.58
Hispanic	30.58	23.46
White	15.99	12.23
Overall	19.03	14.60
Adjustment	Censoring	Cash+censoring

Notes: This table presents estimates of means and standard deviations of housing returns and rates of distressed home sales by race and ethnicity. Panel A presents statistics for annual unlevered returns. Panel B presents statistics for annual total returns. Panel C presents statistics for annual levered returns. Panel D presents statistics for distressed home sales. Column 1 adjusts for censoring due to our finite sample window using the alternative method described in Section F.3. Column 2 adjusts for both censoring and cash purchases. See Appendix F.3 for more details.

Table A4: Cross-Tabulation of Self-Reported and Imputed Race

	Self-Reported Race						N
	Hispanic	Black	White	Asian	Other	Missing	
Imputed Race							
Hispanic	72.7	1.5	14.4	3.4	0.6	7.4	2,102,238
Black	3.3	54.2	28.6	2.3	1.2	10.5	1,072,438
White	2.3	3.5	82.2	2.0	0.8	9.2	12,161,201
Asian	3.0	1.5	15.2	68.3	2.1	10.0	945,778
Other	6.3	8.1	37.3	32.8	5.1	10.5	39,292
Any Non-Missing	11.5	6.4	65.9	6.1	0.9	9.1	16,320,947
Missing	10.4	6.8	65.0	7.6	0.8	9.4	199,232

Notes: This table compares the sample distribution of self-reported and imputed race. Each number corresponds to the percent of homeowners self-reporting a given race/ethnicity, conditional on being categorized to a given race/ethnicity based on their name and Census block. Percentages are reported separately by race/ethnicity imputed from name and Census Block. Data are from sample of property records linked to HMDA data described in Section I.

Table A5: Reweighted Returns and Distressed Sales

	Unlevered Returns (1)	Total Returns (2)	Levered Returns (3)	Distressed Sales (4)
<i>Panel A. Property Data Weights</i>				
Black	0.60 (9.40)	5.39 (9.06)	3.80 (46.70)	23.1
Hispanic	0.82 (11.74)	4.33 (11.53)	-1.69 (53.52)	27.1
White	3.04 (7.32)	6.77 (7.01)	6.38 (40.99)	12.3
Overall	2.54 (8.31)	6.32 (8.02)	5.04 (43.52)	15.2
Black-White	-2.43	-1.39	-2.58	10.76
Hispanic-White	-2.21	-2.44	-8.07	14.75
<i>Panel B. Geography Weights</i>				
Black	0.09 (9.10)	5.42 (8.64)	8.26 (45.54)	20.6
Hispanic	1.06 (10.51)	4.88 (10.19)	2.02 (50.05)	22.0
White	2.59 (6.49)	6.72 (6.15)	10.16 (39.16)	9.8
Overall	2.22 (7.32)	6.41 (6.96)	9.13 (41.10)	12.0
Black-White	-2.50	-1.30	-1.90	10.85
Hispanic-White	-1.53	-1.85	-8.14	12.23
<i>Panel C. Demographic Weights</i>				
Black	0.60 (9.25)	5.32 (8.88)	4.64 (47.32)	21.9
Hispanic	0.49 (11.82)	3.96 (11.50)	-1.93 (54.46)	26.9
White	2.89 (7.13)	6.62 (6.80)	6.74 (40.77)	11.8
Overall	2.34 (8.21)	6.13 (7.88)	5.35 (43.66)	14.8
Black-White	-2.29	-1.30	-2.10	10.17
Hispanic-White	-2.40	-2.66	-8.68	15.16

Notes: This table presents reweighted estimates of means and standard deviations of housing returns and rates of distressed home sales by race and ethnicity. The outcomes are annual unlevered returns (Column 1), annual total returns (Column 2), annual levered returns (Column 3), and rates of distressed sales (Column 4). All outcomes are expressed in annual percentage terms. Each panel corresponds to a different set of weights. Panel A reweights the baseline sample using the larger sample of ownership spells in the property data that does not require the merge with the HMDA data, matching on ZIP code, purchase year, purchase price bins, CLTV bins, and property type. Panel B reweights the sample to match the national distribution of owner-occupied housing units across Census tracts from the 2010 ACS. Panel C reweights the sample to match demographics from the 2013 to 2017 waves of the ACS, matching on race/ethnicity, gender, bins of income, and number of bedrooms.

Table A6: Differences in Home Improvement Expenditures by Race

	Home Expenditures		Number of Permits		Permit Costs		Non-Home Durables	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	−0.214 (0.0232)	−0.201 (0.0274)	−0.0207 (0.000486)	−0.0103 (0.000483)	−0.0529 (0.00166)	−0.0215 (0.00187)	−165.5 (8.815)	−110.6 (11.83)
Hispanic	−0.0642 (0.0183)	−0.134 (0.0257)	−0.0138 (0.000684)	−0.00807 (0.000424)	−0.0457 (0.00179)	−0.0151 (0.00148)	2.094 (8.107)	−47.99 (11.66)
Outcome Mean	0.836	0.852	0.0646	0.0641	0.141	0.140	395.1	406.2
N	640,027	526,872	5,416,456	5,029,784	5,416,456	5,029,784	646,210	526,872
Controls	X		X		X		X	

Notes: This table presents estimates of differences in home improvement expenditures by race/ethnicity. Each column corresponds to a separate regression. The outcome in Columns 1 and 2 is total annual home improvement expenditures as a percentage of home value, and *Controls* indicates regression including fixed effects that interact state, purchase year, and current year. The outcome in Columns 7 and 8 is total monthly non-home durable expenditures, and *Controls* indicates regression controlling for log home value and including fixed effects that interact state, purchase year, and current year. The data for Columns 1, 2, 7, and 8 are from the 2001 to 2013 waves of the Consumer Expenditure Survey, used in Benmelech et al. (2017). The outcome in Columns 3 and 4 is the number of building permits for home improvement filed per year of ownership, and *Controls* denotes fixed effects that interact purchase year, sale year, and state. The outcome in Columns 5 and 6 is the annual reported job costs filed with building permits, normalized as a percentage of purchase price per year of ownership. Permits are measured in BuildFax data. The data in Columns 3 through 6 are from baseline sample of ownership spells restricted to properties merged with Buildfax data. Standard errors are clustered at the level of the household/ownership spell.

Table A7: Racial Disparities in NPV and Sharpe Ratio

Outcome	NPV	NPV	NPV	NPV	Sharpe Ratio (Un- levered)	Sharpe Ratio (Total)	Sharpe Ratio (Levered)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Raw Differences							
Black	-2.083 (0.444)	-1.147 (0.555)	-7.914 (0.512)	-3.296 (0.562)	-0.318 (0.0113)	-0.221 (0.0161)	-0.0631 (0.0118)
Hispanic	-2.725 (0.757)	-5.002 (0.757)	-7.962 (0.930)	-7.711 (0.815)	-0.150 (0.0220)	-0.371 (0.0280)	-0.198 (0.0190)
N	13,430,637	13,430,637	13,430,637	13,430,637	13,629,356	13,429,650	13,429,573
Panel B. Fixed Effect Estimates							
Black	-0.0785 (0.0916)	1.649 (0.108)	-3.234 (0.159)	0.673 (0.125)	-0.177 (0.00404)	-0.0633 (0.00406)	0.00340 (0.00291)
Hispanic	-1.371 (0.0971)	-0.311 (0.0999)	-3.620 (0.131)	-1.492 (0.131)	-0.150 (0.00362)	-0.111 (0.00374)	-0.0503 (0.00242)
N	12,841,103	12,841,103	12,841,103	12,841,103	13,026,345	12,841,103	12,841,027
Outcome Mean	10.87	2.802	6.482	1.284	0.0404	0.948	0.128
Outcome SD	45.64	38.08	52.08	40.19	1.219	1.435	1.200
Forecl. Costs	\$0	\$0	\$50,000	\$50,000	-	-	-
Cost-Weighted		X		X			

Notes: This table presents estimates of the racial gap in housing returns in net present value (NPV) and Sharpe ratios. Columns 1 through 4 present results for NPV, which is scaled as a percentage of the net present value of cash flows (i.e., upfront costs, mortgage payment, property taxes, insurance, and maintenance costs discounted by the 30-year fixed mortgage rate). Column 5 presents results for Sharpe ratio computed using unlevered returns. Column 6 presents results for Sharpe ratio computed using total returns. Column 7 presents results for Sharpe ratio computed using levered returns. For computing both Sharpe ratios, reference groups are properties purchased in the same county and year; risk-free rate is 10-year Treasury yield. Within each panel, columns correspond to separate regressions estimating Equation 4. Panel A presents raw estimates. Panel B applies fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant; fixed effects that interact county, purchase year, and deciles of income at home purchase; and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Cost-Weighted* indicates that observations have been weighted by (strictly positive) net present value of costs (i.e., upfront costs, mortgage payment, taxes and insurance, and maintenance). *Forecl. Costs* denotes additional dollar cost of foreclosure paid in final month of homeownership spell. Data come from baseline sample of ownership spells described in Section I. Standard errors are clustered within purchase year, sale year, and county cells.

Table A8: Racial Gaps in Unlevered Returns: Reconciling with Diamond and Diamond (2024)

	Never Sold	Purchase Year-County FEs	Cash Purchased	Tenure Length Weighted	Purchase Year-Sale Year-County FEs
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Real Annual Unlevered Return (Percent)</i>					
Black	-1.906 (0.0662)	-1.195 (0.0287)	-0.989 (0.0252)	-0.515 (0.0152)	-0.591 (0.0143)
Hispanic	-1.875 (0.166)	-1.264 (0.0439)	-0.956 (0.0383)	-0.557 (0.0234)	-0.341 (0.0121)
Outcome Mean	0.869	0.869	1.272	0.651	0.653
N	15,599,040	15,597,892	15,597,892	15,597,892	15,580,220
<i>Panel B. Outstanding Mortgage-based Return for Foreclosed Spells</i>					
Black	-0.794 (0.0401)	-0.338 (0.0103)	-0.293 (0.00986)	-0.127 (0.00539)	-0.182 (0.00471)
Hispanic	0.101 (0.0738)	-0.0767 (0.0125)	-0.00866 (0.0133)	-0.0122 (0.00561)	0.00815 (0.00393)
Outcome Mean	1.695	1.695	1.913	0.972	0.974
N	15,599,040	15,597,892	15,597,892	15,597,892	15,580,220
<i>Panel C. Real Annual Unlevered Return, Restricted to DD Counties</i>					
Black	-3.110 (0.391)	-1.843 (0.266)	-1.526 (0.230)	-0.864 (0.156)	-0.994 (0.165)
Hispanic	-2.866 (0.414)	-1.512 (0.221)	-1.169 (0.188)	-0.732 (0.131)	-0.605 (0.100)
Outcome Mean	0.447	0.447	0.792	0.171	0.171
N	713,604	713,603	713,603	713,603	713,603
<i>Panel D. Outstanding Mortgage-based Return, Restricted to DD Counties</i>					
Black	-1.575 (0.192)	-0.465 (0.0721)	-0.420 (0.0678)	-0.205 (0.0404)	-0.222 (0.0362)
Hispanic	-1.210 (0.170)	-0.218 (0.0439)	-0.143 (0.0430)	-0.135 (0.0255)	-0.0694 (0.0175)
Outcome Mean	1.210	1.210	1.387	0.498	0.498
N	713,604	713,603	713,603	713,603	713,603

Notes: This table presents estimates of racial gaps in annual unlevered housing returns. The outcome in Panels A and C is real annual unlevered returns. The outcome in Panels B and D is real unlevered housing returns, in which returns for distressed sales are computed by replacing the observed sale price with the estimated principal balance at sale. Race indicators pertain to self-reported race/ethnicity in HMDA. To maximize comparability with Diamond and Diamond (2024b), we restrict and extend our baseline sample to ownership spells beginning in 2000-2016. For further comparability, Panels C and D restrict the sample to the four counties studied in Diamond and Diamond (2024b): Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA. The specification and sample in Column 1 is most comparable to our baseline approach, and includes ownership spells with an observed purchase but no observed sale. Column 2 includes purchase year-by-county fixed effects. Column 3 adjusts for cash purchases by reweighting the sample (i.e., increasing the weights on non-distressed sales, which serve as a proxy for cash purchases). Column 4 additionally weights the sample by ownership spell duration in months. Column 5 applies purchase year-by-sale year-by county fixed effects.

Table A9: Impacts of Tenure Weights and Inflation Adjustment on Gaps in Unlevered Returns

	Unlevered Return (1)	Unlevered Return (2)	Unlevered Return (3)	Unlevered Return (4)
Black	-2.325 (0.0739)	-1.227 (0.0447)	-2.199 (0.0704)	-1.169 (0.0436)
Hispanic	-2.264 (0.194)	-0.863 (0.101)	-2.217 (0.183)	-0.852 (0.0991)
Outcome Mean	2.355	2.123	0.441	0.256
N	13,630,629	13,630,629	13,630,629	13,630,629
Tenure Length Weights		X		X
Inflation Adjusted			X	X

Notes: This table presents estimates of racial gaps in annual unlevered housing returns. *Tenure Length Weights* denotes weights for duration of ownership spell in months. *Inflation Adjusted* denotes that unlevered returns are in real terms. Race indicators pertain to self-reported race/ethnicity in HMDA. Data come from baseline sample of ownership spells described in Section I.

Table A10: Ownership Transitions by Race

	<i>Next Owner</i>				
	White	Black	Hispanic	Asian	Institutional
<i>Panel A. Owner Race/Ethnicity (All Sales)</i>					
White	80.96	3.91	8.62	6.19	6.03
Black	43.44	33.02	15.24	7.99	17.06
Hispanic	40.33	5.44	44.82	9.11	12.16
Asian	46.80	5.31	13.14	34.42	6.81
<i>Panel B. Owner Race/Ethnicity (Distressed Sales)</i>					
White	78.53	4.38	10.24	6.55	16.44
Black	44.94	31.32	14.60	8.91	27.26
Hispanic	44.16	5.47	38.68	11.44	20.82
Asian	47.93	5.90	14.75	31.08	16.21
<i>Panel C. Sale Type</i>					
Non-Distressed	90.63	88.63	85.83	87.86	
Distressed	9.37	11.37	14.17	12.14	

Notes: This table presents the transition matrix of homeowner characteristics across subsequent homeownership spells. Each number corresponds to the percentages of subsequent owners falling into a given category, conditional on original owner race/ethnicity (Panels A and B) or next owner race/ethnicity (Panel C). Panel A presents statistics for all sales, while Panel B presents statistics for distressed home sales. Panel C presents the share of purchases that are distressed or non-distressed by buyer race/ethnicity. Data are from subsample of baseline sample of ownership spells (described in Section I), restricted to ownership spells with observed sale within sample window.

Table A11: Racial Gaps in Large Income Losses in Current Population Survey

	(1)	(2)	(3)
Black	6.037 (1.125)	6.231 (1.455)	5.749 (1.909)
Hispanic	4.714 (1.179)	2.718 (1.608)	4.149 (2.270)
Outcome Mean	19.05	17.62	17.09
Outcome SD	39.27	38.10	37.65
N	39129	22583	13238
Job Tenure	Any Tenure	5+ Years	10+ Years

Notes: This table presents regressions of an indicator that a worker experiences a 30 percent loss in wage and salary income over 12 months on race indicators. All specifications control for fixed effects that interact year, state, and highest level of educational attainment; fixed effects for sex, marital status, income deciles; linear and quadratic terms in age; and fixed effects that interact 4-digit occupation code and year. All controls are measured at beginning of 12-month period. Column 1 presents estimates for all homeowners. Column 2 restricts to subsample with 5 or more years of experience in their current job. Column 3 restricts to subsample with 10 or more years of experience. Data come from sample of homeowners in CPS described in Section C.5. Robust standard errors in parentheses.

Table A12: Racial Disparities in Housing Returns (All Racial/Ethnic Groups)

Outcome	Unlevered Return (1)	Unlevered Return (2)	Total Return (3)	Total Return (4)	Levered Return (5)	Levered Return (6)	Distressed Sale (7)	Distressed Sale (8)
Black	-2.325 (0.0739)	-1.254 (0.0312)	-1.310 (0.0758)	-0.613 (0.0295)	-2.170 (0.378)	-0.293 (0.107)	0.105 (0.00369)	0.0395 (0.00143)
Hispanic	-2.264 (0.194)	-1.283 (0.0363)	-2.465 (0.196)	-1.009 (0.0352)	-7.750 (0.749)	-2.145 (0.112)	0.150 (0.00713)	0.0398 (0.00127)
Asian	-0.0691 (0.0781)	-0.335 (0.0119)	-0.703 (0.0767)	-0.215 (0.0112)	-3.986 (0.363)	-0.626 (0.0580)	0.0133 (0.00326)	0.00959 (0.000552)
Other	1.045 (0.0551)	-0.321 (0.0205)	1.202 (0.0497)	-0.184 (0.0184)	5.018 (0.275)	-0.155 (0.108)	0.0258 (0.00186)	0.0130 (0.000943)
Missing	-0.344 (0.0407)	-0.314 (0.0105)	-0.480 (0.0411)	-0.242 (0.00968)	-2.350 (0.183)	-0.555 (0.0510)	0.0311 (0.00157)	0.0150 (0.000494)
N	16,246,855	14,838,027	16,015,602	14,626,144	16,015,514	14,626,059	16,246,855	14,838,027
Outcome Mean	2.408	2.418	6.124	6.129	4.826	4.922	0.145	0.145
Outcome SD	8.148	8.119	7.797	7.769	43.17	43.32	0.352	0.352
Fixed Effects		X		X		X		X

Notes: This table presents estimates of racial gaps in housing returns and distressed sales. Columns 1 and 2 present results for unlevered returns. Columns 3 and 4 present results for total returns. Columns 5 and 6 present results for levered returns. Columns 7 and 8 present results for an indicator that a property is sold in a distressed sale. *Fixed Effects* indicates the inclusion of fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant; fixed effects that interact county, purchase year, and deciles of income at home purchase; and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. Race indicators pertain to self-reported race/ethnicity in HMDA. *Missing* denotes that no race/ethnicity was recorded in HMDA. Data come from sample of ownership spells described in Section I, expanded to include additional racial/ethnic groups. Standard errors are clustered within purchase year, sale year, and county cells.

Table A13: Racial Disparities in Housing Returns (All Racial/Ethnic Groups)

Outcome	Unlevered Return (1)	Unlevered Return (2)	Total Return (3)	Total Return (4)	Levered Return (5)	Levered Return (6)
Black $\times$ Reg.	-0.388 (0.0527)	-0.296 (0.0230)	0.517 (0.0515)	0.276 (0.0223)	4.175 (0.430)	0.952 (0.115)
Hispanic $\times$ Reg.	1.508 (0.0995)	0.189 (0.0436)	1.251 (0.0922)	0.405 (0.0408)	6.338 (0.572)	2.057 (0.225)
Asian $\times$ Reg.	0.405 (0.0540)	-0.117 (0.0135)	-0.181 (0.0502)	0.0281 (0.0128)	-1.362 (0.228)	0.948 (0.0792)
Other $\times$ Reg.	0.743 (0.0489)	0.0876 (0.0303)	1.032 (0.0411)	0.186 (0.0272)	6.165 (0.320)	1.731 (0.202)
Missing $\times$ Reg.	0.285 (0.0242)	0.0597 (0.0112)	0.123 (0.0222)	0.114 (0.0103)	0.267 (0.126)	1.064 (0.0700)
White $\times$ Distr.	-11.66 (0.112)	-9.435 (0.0804)	-10.80 (0.124)	-8.492 (0.0767)	-56.91 (0.428)	-44.71 (0.431)
Black $\times$ Distr.	-15.26 (0.151)	-12.65 (0.111)	-13.42 (0.173)	-10.99 (0.112)	-54.51 (0.486)	-42.18 (0.525)
Hispanic $\times$ Distr.	-17.72 (0.287)	-13.91 (0.165)	-17.34 (0.319)	-12.66 (0.165)	-71.16 (0.756)	-53.08 (0.839)
Asian $\times$ Distr.	-13.70 (0.211)	-10.66 (0.122)	-13.88 (0.239)	-9.839 (0.120)	-72.59 (0.690)	-52.65 (0.674)
Other $\times$ Distr.	-12.51 (0.188)	-10.81 (0.146)	-11.52 (0.211)	-9.563 (0.139)	-58.60 (0.733)	-48.73 (0.782)
Missing $\times$ Distr.	-13.17 (0.171)	-11.04 (0.113)	-12.50 (0.196)	-10.03 (0.109)	-62.49 (0.596)	-49.84 (0.631)
N	16,246,855	14,838,027	16,015,602	14,626,144	16,015,514	14,626,059
Outcome Mean	2.408	2.418	6.124	6.129	4.826	4.922
Outcome SD	8.148	8.119	7.797	7.769	43.17	43.32
Fixed Effects		X		X		X

Notes: This table presents estimates of racial gaps in housing returns and distressed sales. Columns 1 and 2 present results for unlevered returns. Columns 3 and 4 present results for total returns. Columns 5 and 6 present results for levered returns. *Distr.* is an indicator that a sale is distressed sale, and *Reg.* is an indicator that a sale is a non-distressed sale or is not sold in our sample window. *Fixed Effects* indicates the inclusion of fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant; fixed effects that interact county, purchase year, and deciles of income at home purchase; and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. Race indicators pertain to self-reported race/ethnicity in HMDA. *Missing* denotes that no race/ethnicity was recorded in HMDA. Data come from sample of ownership spells described in Section I, expanded to include additional racial/ethnic groups. Standard errors are clustered within purchase year, sale year, and county cells.

Table A14: Comparison with External Representative Sample

	Baseline Sample (1)	AHS 2015 (2)	AHS 2015 (3)
Hispanic (%)	11.40	9.88	11.21
Non-Hispanic Black (%)	7.66	8.54	8.31
Single-Family Home (%)	84.85	90.23	88.93
Current Value			
Mean	305,113	273,585	280,952
SD	229,794	232,774	232,339
p10	107,954	79,117	84,673
p90	579,249	548,102	553,100
Tenure (Years)*			
Mean	6.89	15.43	7.38
SD	4.49	12.54	4.40
p10	1.00	2.00	2.00
p90	13.00	34.00	14.00
Down Payment (%) [‡]			
Mean	14.58	17.72	16.86
SD	18.45	20.72	20.51
p10	0.00	0.00	0.00
p90	30.50	30.50	30.50
Census Division (%)			
New England	6.89	5.50	4.80
Middle Atlantic	13.19	13.50	11.94
East North Central	15.31	17.24	16.00
West North Central	3.76	7.90	8.18
South Atlantic	24.73	18.42	19.49
East South Central	3.93	5.31	5.61
West South Central	2.48	10.45	11.49
Mountain	7.68	7.27	8.21
Pacific	22.03	14.41	14.29
N	9,533,818	24,869	15,140
Moved Since 2000			X

Notes: This table compares characteristics of our baseline analysis sample to those of a representative external sample from the 2015 American Housing Survey (AHS). Column 1 presents characteristics of our baseline sample, which we restrict to properties that were purchased before 2015 and remained unsold as of 2015 in order to maximize comparability with the AHS sample. Column 2 presents statistics for the sample of homeowners in AHS. Column 3 presents statistics for homeowners in AHS who have moved into their current properties since 2000. See Appendix C.5 for more details on the AHS sample.

* In the AHS, tenure length is measured by year the householder moved into the property. In the baseline sample, tenure is measured using the year the property was purchased.

‡Down payment is measured as a percentage of purchase price and is discretized into eight bins. To compute a numerical value, we take the midpoint of each bin, and apply the same procedure to construct an analogous variable in the baseline sample.

Table A15: Geographical Coverage Relative to External Representative Sample

CBSA (%)	Baseline Sample (1)	AHS 2015 (2)	AHS 2015 (3)
Atlanta-Sandy Springs-Roswell, GA	1.84	1.93	2.26
Boston-Cambridge-Newton, MA-NH	2.45	1.74	1.52
Chicago-Naperville-Elgin, IL-IN-WI	5.00	3.29	3.09
Dallas-Fort Worth-Arlington, TX	0.07	2.07	2.43
Detroit-Warren-Dearborn, MI	1.58	1.61	1.40
Houston-The Woodlands-Sugar Land, TX	0.00	1.85	2.19
Los Angeles-Long Beach-Anaheim, CA	4.64	2.96	2.56
Miami-Fort Lauderdale-West Palm Beach, FL	2.81	1.56	1.59
New York-Newark-Jersey City, NY-NJ-PA	5.57	5.40	4.80
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	3.38	2.26	2.06
Phoenix-Mesa-Scottsdale, AZ	2.36	1.38	1.66
Riverside-San Bernardino-Ontario, CA	2.14	1.11	1.24
San Francisco-Oakland-Hayward, CA	1.85	1.40	1.31
Seattle-Tacoma-Bellevue, WA	2.52	1.29	1.31
Washington-Arlington-Alexandria, DC-VA-MD-WV	3.79	2.25	2.39
All other metropolitan areas	56.76	53.58	54.25
Not in a metropolitan area	3.22	14.31	13.93
N	9,533,818	24,869	15,140
Moved Since 2000			X

Notes: This table compares the distribution of our baseline analysis sample distribution across CBSAs to that of a representative external sample from the 2015 American Housing Survey (AHS). Column 1 presents characteristics of our baseline sample, which we restrict to properties that were purchased before 2015 and remained unsold as of 2015 in order to maximize comparability with the AHS sample. Column 2 presents statistics for the sample of homeowners in AHS. Column 3 presents statistics for homeowners in AHS who have moved into their current properties since 2000. See Appendix C.5 for more details on the AHS sample.

Table A16: Characteristics of Next Ownership Spell

	Institutional Next Owner (1)	Occupied by Next Owner (2)	Next Owner Tenure (3)	White Next Owner (4)
Black	0.0184 (0.000708)	-0.00417 (0.00115)	-1.388 (0.121)	-0.253 (0.00270)
Hispanic	-0.00231 (0.000443)	0.0118 (0.000882)	-1.439 (0.119)	-0.235 (0.00214)
Distressed	0.130 (0.000730)	-0.0429 (0.000738)	-6.101 (0.0697)	-0.0352 (0.00103)
Black $\times$ Distressed	0.0708 (0.00160)	-0.0204 (0.00208)	-0.460 (0.167)	0.0188 (0.00367)
Hisp. $\times$ Distressed	0.0474 (0.00124)	-0.0178 (0.00153)	0.712 (0.168)	0.0325 (0.00310)
Outcome Mean	0.0765	0.909	75.04	0.728
N	6,663,541	3,054,278	6,663,541	2,744,976

Notes: This table presents estimates of Equation 4 for outcomes pertaining to the ownership spell immediately following the current spell, controlling for purchase year, sale year, county, leverage, family composition, and income. Specifically, these specifications include fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, fixed effects that interact county, purchase year, and deciles of income at home purchase, and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Distressed* denotes an indicator that the current spell ends in a foreclosure or short sale. Each column corresponds to a separate regression and outcome. *Institutional Next Owner* is an indicator that the next owner is as a non-trust institution. Purchases by non-trust institutions are defined as those in which the buyer name in the property data is not that of a person or trust. *Occupied by Next Owner* is an indicator that the next owner lives in the property, conditional on a non-institutional next owner (i.e., the property is labeled as owner-occupied in the HMDA data). *Next Owner Tenure* is the number of months that the next owner holds the property. *White Next Owner* is an indicator that the next owner identifies as non-Hispanic White. Data are from subsample of baseline sample of ownership spells (described in Section I), restricted to spells with an observed sale within sample window. Standard errors are clustered within purchase year, sale year, and county cells.

Table A17: Differential Modification and Foreclosure Rates by Race

	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Modified</i>					
Black	6.733 (0.565)	5.772 (0.395)	5.047 (0.362)	3.897 (0.528)	2.545 (0.956)
Hispanic	-0.541 (0.782)	0.829 (0.547)	1.570 (0.427)	0.178 (0.343)	-0.133 (0.481)
<i>Panel B. Foreclosed</i>					
Black	-3.958 (0.928)	-4.714 (0.481)	-4.140 (0.524)	-4.064 (0.592)	-2.780 (0.986)
Hispanic	6.744 (0.923)	2.011 (0.559)	0.660 (0.450)	1.503 (0.370)	2.041 (0.642)
<i>Panel C. Self-Cured</i>					
Black	-3.223 (0.574)	-1.442 (0.235)	-1.116 (0.225)	0.147 (0.213)	0.562 (0.578)
Hispanic	-5.543 (0.631)	-1.907 (0.173)	-1.553 (0.135)	-0.861 (0.134)	-1.026 (0.346)
N	1,117,521	1,025,022	957,788	335,551	55,341
Controls	Baseline	Borrower, Mortgage	County-Time, Servicer FE	Tract-Time FE	Servicer-Tract- Time FE

Notes: This table presents regressions in which the outcomes capture the events following a homeowner becoming 90 days delinquent on their mortgage, and shows that minorities are more likely to receive a loan modification after becoming delinquent. The three outcomes are *Modified*, an indicator that the delinquency resulted in the homeowner's loan terms being modified ($\mu = 20.0\%$); *Foreclosed*, an indicator that the delinquency resulted in a foreclosure ($\mu = 63.8\%$); and *Self-Cured*, an indicator that the delinquency resulted in the homeowner becoming current or paying off the loan ($\mu = 14.0\%$). The specification in Column 1 includes fixed effects for the quarter of default. Column 2 interacts quarter of default with origination year, adds fixed effects for number of years remaining in term, indicators for the mortgage being an ARM, interest-only, or negative-amortization, and adds decile fixed effects for current LTV, original credit score, original income, and original loan amount. Column 3 applies fixed effects that interact county, quarter of default, and origination year, and adds a servicer fixed effect to the controls in (2). Column 4 applies fixed effects that interact tract, quarter of default, and origination year, in addition to the controls in (2). Column 5 applies fixed effects that interact tract, servicer, quarter of default, and origination year in addition to the controls in (2). Data comprised of homeowners who have become delinquent and are observed in the Fannie Mae, Freddie Mac, and ABSNet data, described in Section IV. Coefficients are scaled by 100 and are interpretable as the percentage point differences in the likelihood of each outcome. Standard errors are two-way clustered at the servicer level and within origination year, county, and year of default cells.

Table A18: Impacts of Modifications, Robustness and Heterogeneity Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Unlevered Returns (% Annualized)</i>						
Modification	11.24 (2.536)	9.411 (2.348)	13.30 (2.982)	10.58 (2.420)	9.844 (2.469)	10.22 (2.547)
Black $\times$ Mod.	-2.825 (2.144)	-3.675 (2.845)	-3.339 (2.963)	-1.805 (2.457)		
Hispanic $\times$ Mod.	3.223 (1.830)	4.122 (2.643)	2.592 (2.269)	4.467 (2.298)		
Distressed Tract $\times$ Mod.					5.139 (1.670)	
Single Applicant $\times$ Mod.						2.939 (1.460)
Outcome Mean	-12.58	-13.62	-13.63	-13.60	-12.59	-12.58
N	130354	71888	102048	80417	130298	130354
<i>Panel B. Distressed Sale</i>						
Modification	-0.362 (0.0981)	-0.309 (0.109)	-0.400 (0.130)	-0.317 (0.102)	-0.295 (0.0914)	-0.337 (0.0967)
Black $\times$ Mod.	0.103 (0.104)	0.197 (0.157)	0.107 (0.130)	0.133 (0.113)		
Hispanic $\times$ Mod.	0.0150 (0.0688)	0.0255 (0.0921)	0.0827 (0.0850)	0.0334 (0.0842)		
Distressed Tract $\times$ Mod.					-0.0965 (0.0515)	
Single Applicant $\times$ Mod.						-0.00794 (0.0643)
Outcome Mean	0.759	0.773	0.776	0.775	0.759	0.759
N	130354	71888	102048	80417	130298	130354
Controls	Baseline	Score	LTV	Income	Baseline	Baseline

Notes: This table presents robustness exercises for the analysis of the impacts of mortgage modifications along with heterogeneous impacts by neighborhood and household characteristics. The outcomes are the unlevered rate of return (Panel A) and an indicator that the ownership spell ends in a distressed sale (Panel B). Column 1 presents the baseline specification. Columns 2 through 4 interact baseline fixed effects with terciles of credit score at origination, LTV in the month of default, and income at origination, respectively. Column 5 presents heterogeneity results for distressed Census tracts. The share of sales that are distressed sales is computed for each tract-year, and distressed tracts are defined as those in the highest quartile. Column 6 presents heterogeneity results for an indicator that the household listed a single individual on their mortgage application. The baseline specification includes interacted fixed effects for purchase year, default year, tract, indicator for negative amortization loan, and indicator for interest-only loan. Data comprised of homeowners who have become delinquent and are observed in the Fannie Mae, Freddie Mac, and ABSNet data, described in Section I. Standard errors are two-way clustered at the servicer level and within origination year, county, and year of default cells.

Table A19: Modification Treatment Effects Placebo Outcomes

	Distressed Sale (1)	Unlevered Return (2)	Total Return (3)	Levered Return (4)
$\hat{\gamma}_{s(i),t}$	0.00510 (0.0100)	-0.0222 (0.360)	0.137 (0.394)	1.395 (1.326)
Black $\times \hat{\gamma}_{s(i),t}$	-0.0167 (0.0116)	-0.204 (0.466)	-0.280 (0.491)	-0.214 (1.468)
Hispanic $\times \hat{\gamma}_{s(i),t}$	-0.00101 (0.00847)	-0.234 (0.347)	-0.469 (0.381)	-2.106 (1.180)
Outcome Mean	0.708	-10.23	-6.238	-46.03
N	130,354	130,354	130,246	130,246

Notes: This table presents placebo exercises for the analysis of the impacts of mortgage modifications. Each column estimates the reduced-form impact of the server instrument and its interactions with race/ethnicity indicators on placebo outcomes. The placebo outcomes are defined using the predicted values from a regression of the true outcome (e.g. indicator that ownership spell ends in a distressed sale) on a vector of individual characteristics measured prior to the realization of the true outcome. The vector of characteristics includes loan type (i.e. conventional, FHA, VA); loan purpose; indicators for ARM, interest-only, and negative amortization; term; deciles of credit score, income, interest rate, and amount at origination; current year, and data source (i.e. Fannie Mae, Freddie Mac, ABSNet). Each placebo regression includes interacted fixed effects for purchase year, default year, tract, indicator for negative amortization loan, and an indicator for interest-only loan. Data comprised of homeowners who have become delinquent and are observed in the Fannie Mae, Freddie Mac, and ABSNet data, described in Section IV. Standard errors are two-way clustered at the servicer level and within origination year, county, and year of default cells.

Table A20: Annualized Housing Returns by Race and Ethnicity

	Baseline (1)	Regular Sales Only (2)	Distressed Sales Only (3)	Cash Adjustment (4)	Distress Adjustment (5)
<i>Panel A. Unlevered Returns (% Annualized)</i>					
Black	0.76 (9.80)	4.12 (6.25)	-11.12 (10.80)	1.48 (9.26)	0.54 (10.19)
Hispanic	1.05 (12.44)	6.25 (8.03)	-13.52 (10.95)	2.26 (11.77)	1.02 (13.02)
White	3.08 (7.65)	4.47 (6.20)	-7.49 (9.21)	3.41 (7.36)	3.01 (8.17)
Overall	2.63 (8.68)	4.66 (6.48)	-9.43 (10.21)	3.10 (8.26)	2.54 (9.19)
Black-White	-2.33	-0.35	-3.63	-1.93	-2.47
Hispanic-White	-2.03	1.78	-6.03	-1.14	-1.98
<i>Panel B. Total Returns (% Annualized)</i>					
Black	5.39 (9.25)	8.53 (5.53)	-5.68 (11.06)	6.06 (8.69)	5.12 (9.17)
Hispanic	4.43 (11.86)	9.43 (7.15)	-9.52 (11.24)	5.59 (11.15)	4.24 (11.80)
White	6.68 (7.16)	7.96 (5.69)	-3.03 (9.35)	6.98 (6.87)	6.52 (7.14)
Overall	6.27 (8.18)	8.18 (5.89)	-4.97 (10.44)	6.71 (7.75)	6.10 (8.14)
Black-White	-1.29	0.56	-2.65	-0.92	-1.40
Hispanic-White	-2.26	1.46	-6.49	-1.39	-2.28
<i>Panel C. Levered Returns (% Annualized)</i>					
Black	5.24 (50.10)	18.60 (43.56)	-41.82 (42.70)	5.94 (44.51)	3.55 (50.09)
Hispanic	0.92 (58.79)	22.14 (49.76)	-58.26 (37.83)	2.90 (51.73)	0.03 (58.63)
White	7.20 (43.81)	13.98 (38.98)	-44.13 (44.29)	7.38 (38.42)	6.53 (43.74)
Overall	6.18 (46.70)	15.27 (40.83)	-47.39 (43.04)	6.65 (41.01)	5.40 (46.63)
Black-White	-1.96	4.62	2.30	-1.44	-2.98
Hispanic-White	-6.28	8.16	-14.13	-4.48	-6.50
<i>Panel D. Distressed Sale (%)</i>					
Black	22.1			17.3	31.0
Hispanic	26.3			20.2	31.6
White	11.6			8.9	16.4
Overall	14.4			11.1	19.6

Notes: This table presents estimates of means and standard deviations of housing returns and rates of distressed home sales by race and ethnicity, which include ownership spells lasting less than 12 months. Panel A presents statistics for annual unlevered returns. Panel A presents statistics for annual unlevered returns. Panel B presents statistics for annual total returns. Panel C presents statistics for annual levered returns. Panel D presents rates of distressed home sales. Column 1 presents estimates from baseline sample of ownership spells described in Section I. Column 2 presents estimates for the subsample of properties not sold in a distressed sale, including properties not sold within the sample period. Column 3 present estimates for the subsample of properties sold in a distressed sale. Column 4 adjusts for cash purchases. Column 5 adjusts for bias from not observing future distressed sales among properties that remain unsold by the end of our sample window.

Table A21: Racial Disparities in Housing Returns

Outcome	Unlevered Re- turns (1)	Unlevered Re- turns (2)	Unlevered Re- turns (3)	Unlevered Re- turns (4)	Unlevered Re- turns (5)	Total Re- turns (6)	Total Re- turns (7)	Total Re- turns (8)	Total Re- turns (9)	Total Re- turns (10)	Levered Re- turns (11)	Levered Re- turns (12)	Levered Re- turns (13)	Levered Re- turns (14)	Levered Re- turns (15)	Distressed Sale (16)	Distressed Sale (17)	Distressed Sale (18)	Distressed Sale (19)	Distressed Sale (20)
<i>Panel A. Baseline</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.105	0.0819	0.0720	0.0407	0.0208
Black	2.325 (0.0739)	1.539 (0.0366)	1.479 (0.0332)	1.278 (0.0318)	0.706 (0.0201)	1.310 (0.0758)	0.783 (0.0351)	0.789 (0.0318)	0.627 (0.0301)	0.408 (0.0186)	2.170 (0.378)	0.0305 (0.135)	0.323 (0.125)	0.298 (0.109)	1.059 (0.0895)	0.00369 (0.00369)	0.00181 (0.00181)	0.00161 (0.00161)	0.00146 (0.00146)	0.00104 (0.00104)
Hispanic	2.264 (0.194)	1.598 (0.0552)	1.568 (0.0467)	1.295 (0.0359)	0.928 (0.0224)	2.465 (0.196)	1.196 (0.0546)	1.271 (0.0463)	1.019 (0.0348)	0.905 (0.0219)	7.750 (0.749)	2.221 (0.182)	3.053 (0.156)	2.204 (0.112)	3.025 (0.0910)	0.00713 (0.00713)	0.00213 (0.00213)	0.00182 (0.00182)	0.00129 (0.00129)	0.000934 (0.000934)
<i>Panel B. Interacted</i>	—	—	—	—	—	0.517	0.557	0.441	0.280	0.242	4.175	3.646	2.873	1.069	—	—	—	—	—	—
Black × Regular	0.388 (0.0527)	0.0912 (0.0208)	0.152 (0.0207)	0.301 (0.0233)	0.0201 (0.0240)	0.0515 (0.0515)	0.0202 (0.0202)	0.0202 (0.0202)	0.0226 (0.0226)	0.0229 (0.0229)	0.430 (0.430)	0.161 (0.161)	0.152 (0.152)	0.117 (0.117)	0.384 (0.113)	—	—	—	—	—
Hispanic × Regular	1.508 (0.0995)	0.492 (0.0459)	0.332 (0.0424)	0.189 (0.0438)	0.349 (0.0410)	1.251 (0.0922)	0.815 (0.0414)	0.554 (0.0393)	0.409 (0.0410)	0.337 (0.0388)	6.338 (0.572)	4.578 (0.261)	3.111 (0.235)	2.150 (0.227)	1.004 (0.194)	—	—	—	—	—
Black × Distr.	15.26 (0.151)	12.68 (0.110)	12.62 (0.110)	12.66 (0.111)	11.92 (0.109)	13.42 (0.173)	10.94 (0.114)	10.95 (0.113)	10.99 (0.112)	10.71 (0.109)	54.51 (0.486)	41.26 (0.434)	42.02 (0.442)	41.95 (0.525)	44.86 (0.546)	—	—	—	—	—
Hispanic × Distr.	17.72 (0.287)	14.24 (0.166)	14.08 (0.166)	13.92 (0.167)	13.31 (0.171)	17.34 (0.319)	12.94 (0.174)	12.87 (0.171)	12.66 (0.167)	12.41 (0.167)	71.16 (0.756)	52.40 (0.690)	53.49 (0.722)	52.86 (0.845)	55.26 (0.875)	—	—	—	—	—
White × Distr.	11.66 (0.112)	9.369 (0.0770)	9.338 (0.0768)	9.433 (0.0805)	9.336 (0.0817)	10.80 (0.124)	8.424 (0.0759)	8.406 (0.0752)	8.489 (0.0768)	8.480 (0.0770)	56.91 (0.428)	44.47 (0.377)	44.73 (0.377)	44.56 (0.432)	44.92 (0.437)	—	—	—	—	—
Outcome Mean	2.355	2.355	2.396	2.396	2.389	6.129	6.129	6.172	6.161	6.150	5.043	5.042	5.259	5.213	5.169	0.146	0.146	0.143	0.145	0.145
Outcome SD	8.161	8.161	8.088	8.118	8.122	7.807	7.807	7.736	7.764	7.765	43.35	43.35	43.18	43.42	43.44	0.353	0.353	0.350	0.352	0.352
N	13,630,629	13,629,573	13,628,762	13,026,380	12,942,301	13,430,637	13,429,650	13,032,746	12,841,103	12,777,531	13,430,560	13,429,573	13,032,670	12,841,027	12,777,456	13,630,629	13,629,573	13,628,762	13,026,380	12,942,301
Controls	Raw	+ Year, County	+ Fam- ily, In- come	+ LTV	+ Tract	Raw	+ Year, County	+ Fam- ily, In- come	+ LTV	+ Tract	Raw	+ Year, County	+ Fam- ily, In- come	+ LTV	+ Tract	Raw	+ Year, County	+ Fam- ily, In- come	+ LTV	+ Tract

Notes: This figure provides numerical values for the estimates used to construct Figure 1 in the main paper. Within each panel, columns correspond separate regressions. Labels correspond to those described in the notes of Figure 1. Standard errors clustered at the county-by-purchase year-by-sale year level.

Table A22: Heterogeneous Racial Gaps in Housing Returns and Distressed Sales

	Unlevered Returns		Total Returns		Levered Returns		Distressed Sales	
	Black (1)	Hispanic (2)	Black (3)	Hispanic (4)	Black (5)	Hispanic (6)	Black (7)	Hispanic (8)
Main Effect	−0.735 (0.0308)	−0.471 (0.0324)	−0.222 (0.0292)	−0.318 (0.0308)	−1.369 (0.130)	−2.190 (0.139)	0.0533 (0.00172)	0.0404 (0.00188)
Middle Income Tercile	−0.218 (0.0239)	−0.148 (0.0234)	−0.0658 (0.0223)	−0.0335 (0.0218)	−0.232 (0.125)	−0.227 (0.114)	0.00907 (0.00119)	0.00481 (0.00104)
Lowest Income Tercile	−0.449 (0.0314)	0.0866 (0.0322)	−0.225 (0.0291)	0.280 (0.0302)	−1.499 (0.146)	0.161 (0.145)	0.00664 (0.00141)	−0.0185 (0.00128)
CLTV in [95, 102.5)	−0.135 (0.0321)	−0.762 (0.0349)	−0.121 (0.0304)	−0.714 (0.0338)	2.401 (0.130)	−0.670 (0.150)	−0.0227 (0.00148)	−0.00316 (0.00156)
CLTV in [82, 95)	−0.405 (0.0355)	−0.489 (0.0294)	−0.335 (0.0336)	−0.430 (0.0283)	0.509 (0.135)	−0.0518 (0.129)	0.00218 (0.00163)	0.00643 (0.00136)
Single Female	−0.0587 (0.0271)	−0.372 (0.0381)	−0.0475 (0.0255)	−0.453 (0.0354)	0.984 (0.120)	0.850 (0.160)	−0.00941 (0.00123)	0.0184 (0.00136)
Single Male	−0.443 (0.0309)	−0.520 (0.0299)	−0.497 (0.0296)	−0.533 (0.0280)	−0.0550 (0.125)	0.0206 (0.119)	0.0122 (0.00122)	0.0185 (0.00104)

Notes: This table presents estimates of racial disparities in housing returns corresponding to Figure 2. Number of observations for each regression are as follows: 13199982 (unlevered returns), 13005723 (total returns) 13005650 (levered returns) 13199982 (distressed sales). Standard errors clustered at the county-by-purchase year-by-sale year level.

Table A23: Racial Gaps by Neighborhood Demographics and Sale Type

	Quintile 1 (1)	Quintile 2 (2)	Quintile 3 (3)	Quintile 4 (4)	Quintile 5 (5)
<i>Panel A. Unlevered Returns</i>					
Black	-1.919 (0.0494)	-1.314 (0.0364)	-1.083 (0.0324)	-0.976 (0.0327)	-0.862 (0.0408)
Hispanic	-1.999 (0.0665)	-1.461 (0.0417)	-1.252 (0.0442)	-1.080 (0.0351)	-0.764 (0.0423)
White	-0.453 (0.0166)	-0.288 (0.0119)	-0.209 (0.00968)	-0.147 (0.00739)	
<i>Panel B. Unlevered Returns by Sale Type</i>					
Black × Regular	-0.323 (0.0295)	-0.388 (0.0254)	-0.346 (0.0255)	-0.358 (0.0289)	-0.293 (0.0332)
Hispanic × Regular	0.208 (0.0583)	0.0803 (0.0458)	0.129 (0.0439)	0.0904 (0.0456)	0.271 (0.0455)
White × Regular	0.0769 (0.0112)	-0.0684 (0.00899)	-0.0863 (0.00803)	-0.107 (0.00636)	
Black × Distressed	-13.95 (0.125)	-12.05 (0.117)	-11.44 (0.133)	-11.02 (0.142)	-10.76 (0.167)
Hispanic × Distressed	-15.31 (0.176)	-13.57 (0.164)	-13.16 (0.199)	-12.58 (0.206)	-12.24 (0.240)
White × Distressed	-10.86 (0.0798)	-9.698 (0.0870)	-9.356 (0.0889)	-8.927 (0.0855)	-8.644 (0.0853)
<i>Panel C. Levered Returns</i>					
Black	1.368 (0.147)	-0.440 (0.142)	-2.110 (0.143)	-2.949 (0.158)	-3.292 (0.186)
Hispanic	-0.455 (0.205)	-2.465 (0.142)	-3.302 (0.146)	-3.921 (0.150)	-3.419 (0.169)
White	1.566 (0.0683)	0.0594 (0.0545)	-0.383 (0.0474)	-0.621 (0.0407)	
<i>Panel D. Levered Returns by Sale Type</i>					
Black × Regular	3.343 (0.161)	1.019 (0.139)	-0.376 (0.138)	-1.157 (0.155)	-1.289 (0.183)
Hispanic × Regular	5.053 (0.333)	2.237 (0.241)	1.340 (0.235)	0.149 (0.213)	0.0368 (0.195)
White × Regular	2.563 (0.0684)	0.632 (0.0544)	0.0427 (0.0477)	-0.401 (0.0386)	
Black × Distressed	-38.72 (0.512)	-41.59 (0.554)	-45.46 (0.639)	-47.09 (0.692)	-48.34 (0.749)
Hispanic × Distressed	-50.99 (0.836)	-52.76 (0.847)	-54.32 (0.953)	-54.03 (0.927)	-53.08 (1.004)
White × Distressed	-41.38 (0.381)	-44.05 (0.440)	-45.39 (0.471)	-45.47 (0.461)	-44.26 (0.452)
<i>Panel E. Distressed Sales</i>					
Black	0.0595 (0.00196)	0.0443 (0.00175)	0.0393 (0.00159)	0.0373 (0.00160)	0.0376 (0.00189)
Hispanic	0.0567 (0.00199)	0.0487 (0.00165)	0.0453 (0.00156)	0.0430 (0.00153)	0.0382 (0.00175)
White	0.0245 (0.000569)	0.0124 (0.000454)	0.00704 (0.000380)	0.00319 (0.000322)	

Notes: This table presents estimates of racial disparities in housing returns, corresponding to Figure 3. Each panel presents a separate regression. The outcome is the unlevered rate of return ($\mu =$). In Panel A, race indicators are interacted with quintiles of the 2010 non-Hispanic White share of homeowners in the Census tract. In Panel B, race indicators are interacted with tract non-Hispanic White share quintiles and indicators for sale type. Data from baseline sample of ownership spells. N=12,835,533. Standard errors clustered at the county-by-purchase year-by-sale year level.

Table A24: Decomposing Differences in Financial Distress

Specification	Mortgage 90 Days Past Due		Non-Mortgage 90 Days Past Due	
	Black	Hispanic	Black	Hispanic
Raw	2.62 (.019)	1.81 (.0118)	18.5 (.0647)	9.56 (.0427)
+ Purch. Year, County	2.19 (.019)	1.61 (.0123)	16.6 (.0641)	9.47 (.0435)
+ Family, Income	2.09 (.0191)	1.52 (.0125)	14.9 (.0637)	7.63 (.0437)
+ Leverage	1.79 (.0189)	1.26 (.0125)	12.9 (.0631)	6.07 (.0434)
+ Credit Score	.977 (.0183)	.937 (.0123)	5.81 (.0548)	2.74 (.0388)
+ Mortgage Chars.	.983 (.0182)	.939 (.0122)	5.66 (.0547)	2.61 (.0388)
+ Tract	.565 (.0196)	.697 (.0129)	4.8 (.0596)	2.27 (.0415)
+ Lender, Block	.502 (.021)	.621 (.0139)	4.5 (.0622)	1.95 (.0425)

Notes: This table presents a statistical decomposition of financial distress. *Raw* denotes a regression of the outcome on race/ethnicity indicators and current year fixed effects. *Purch. Year, County* adds fixed effects that interact current year, origination year, and county. *Family, Income* adds fixed effects that interact gender, the presence of a co-applicant on the mortgage, and current year, and fixed effects that interact deciles of income and current year. *Leverage* adds fixed effects that interact 1 percentage point bins of original loan-to-value and current year. *Credit Score* adds fixed effects that interact 10 point credit score bins and current year. *Mortgage Chars.* controls for a spline in current loan-to-value, log monthly payment, log interest rate, an indicators for interest-only loan, refinance, adjustable rate mortgage, and GSE loan, and adds fixed effects for mortgage term, deciles of original loan value, and deciles of debt-to-income. *Tract* adds Census tract fixed effects. *Lender, Block* adds lender and Census block fixed effects. Data come from panel of homeowners in credit bureau and mortgage servicing data, restricted to homeowners with CLTV less than or equal to 120%. Standard errors clustered at the loan level. N=37.4 million.

Table A25: Disparities in Wealth, Liquidity, and Income

Age			Liquid Assets			Net Wealth			Annual Income		
Hispanic	Black	White	Hispanic	Black	White	Hispanic	Black	White	Hispanic	Black	White
27	27	27	5.24	2.35	8.10	109.58	109.90	95.57	52,589	44,907	59,309
32	32	32	5.45	5.43	11.74	144.25	93.92	128.41	59,471	55,776	73,558
35	35	35	8.95	4.14	14.27	166.00	108.39	157.59	57,976	58,819	78,059
37	38	38	7.74	4.71	14.71	172.96	110.45	178.86	56,817	64,337	78,976
39	40	40	7.22	4.61	16.31	196.72	117.62	187.41	61,379	61,956	81,279
42	42	42	6.62	5.23	18.05	190.26	125.16	201.02	61,982	64,597	83,451
44	45	45	8.95	4.83	19.74	210.89	137.84	221.02	69,079	62,636	85,845
46	47	46	9.59	5.07	20.34	218.41	162.58	234.70	70,967	65,621	87,728
48	49	48	9.72	4.82	21.60	221.80	164.20	246.96	68,302	62,617	86,865
51	51	51	9.75	5.35	24.70	222.42	181.77	281.58	65,781	62,629	86,155
54	54	53	10.12	5.78	25.64	250.61	208.14	311.50	64,178	59,531	83,293
55	55	55	11.98	6.23	28.41	300.23	216.40	339.88	62,144	58,308	79,197
58	58	58	13.00	6.20	34.59	286.08	225.99	403.30	62,883	54,730	72,930
61	61	61	14.82	5.35	42.20	324.93	241.95	456.41	56,773	49,454	65,273
64	64	64	16.44	5.18	51.68	386.36	273.96	510.71	46,709	42,839	57,720
67	67	67	13.51	4.32	65.10	387.99	307.72	578.13	38,788	38,908	50,813
70	70	70	17.59	5.22	79.82	442.60	338.19	635.70	36,833	35,210	44,950
74	74	74	17.07	5.71	93.91	500.57	352.65	661.13	32,989	32,566	39,011
79	79	79	13.57	6.61	101.98	568.58	373.60	689.27	29,944	28,212	34,373
84	84	84	22.97	6.46	126.49	567.20	431.79	777.78	26,988	23,145	28,224

Notes: This table presents median values of wealth, liquidity, and income as a function of age and race, binned into ventiles of age. Liquidity and wealth scaled as a percentage of annual income. Data come from sample of homeowners in the Survey of Income and Program Participation (1990-2017). Race and age are assigned according to head of household. Dollar values are adjusted to 2016 levels. N=646,323.

Table A26: Disparities in Wealth, Liquidity, and Income, Continued

Annual Income			Layoff Probability		
Hispanic	Black	White	Hispanic	Black	White
18,553	17,165	17,404	13.47	11.29	8.54
34,631	34,719	34,955	13.95	12.21	8.99
47,396	47,630	47,754	12.83	10.27	8.08
59,185	59,220	59,413	13.36	11.95	7.81
70,814	71,179	71,064	11.23	10.95	7.50
83,745	83,691	83,771	12.42	9.62	7.79
98,361	98,099	98,675	10.23	10.04	7.80
117,129	116,823	117,209	9.47	10.49	7.23
143,894	144,173	145,408	12.02	12.18	7.73
208,073	209,519	214,390	9.32	10.30	7.03

Notes: This table presents the probability of experiencing a layoff as a function of income and race, binned into deciles of annual income. Data come from sample of homeowners in the Survey of Income and Program Participation (1990-2017). Race assigned according to head of household. Dollar values are adjusted to 2016 levels. N=270,994.

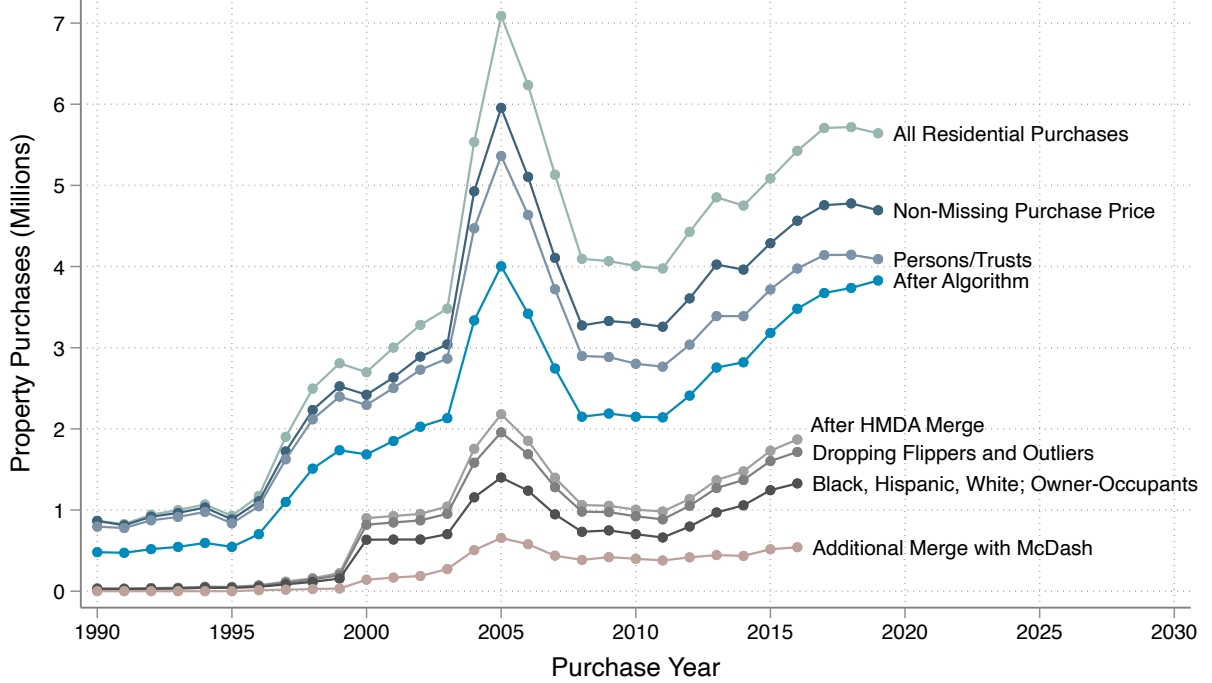
Table A27: Effects of Liquidity Shock by Race/Ethnicity

Event time	Hispanic (30 DPD)	Black (30 DPD)	White (30 DPD)	Hispanic (90 DPD)	Black (90 DPD)	White (90 DPD)
-12	10.60 (1.24)	1.79 (2.27)	-0.64 (0.48)	10.80 (0.93)	4.65 (1.76)	-0.38 (0.31)
-11	10.80 (1.23)	1.80 (2.31)	-0.68 (0.47)	10.90 (0.91)	5.86 (1.71)	-0.09 (0.31)
-10	9.63 (1.18)	1.70 (2.23)	-0.37 (0.46)	10.30 (0.88)	5.09 (1.67)	0.04 (0.30)
-9	9.25 (1.16)	0.11 (2.17)	0.04 (0.46)	10.10 (0.86)	5.67 (1.63)	0.26 (0.30)
-8	8.18 (1.12)	4.60 (2.10)	0.54 (0.44)	8.76 (0.81)	5.35 (1.57)	0.26 (0.29)
-7	7.94 (1.07)	2.17 (2.04)	0.64 (0.43)	8.39 (0.77)	4.91 (1.50)	0.23 (0.28)
-6	7.05 (1.04)	3.69 (2.00)	1.18 (0.43)	8.12 (0.73)	3.67 (1.43)	0.55 (0.27)
-5	6.02 (1.00)	3.90 (1.88)	0.72 (0.40)	6.83 (0.68)	3.93 (1.37)	0.43 (0.25)
-4	5.00 (0.94)	4.02 (1.76)	1.19 (0.38)	5.35 (0.61)	3.81 (1.24)	0.41 (0.23)
-3	3.95 (0.83)	1.28 (1.67)	0.55 (0.35)	3.76 (0.54)	2.41 (1.03)	0.50 (0.20)
-2	1.83 (0.72)	1.89 (1.40)	0.49 (0.31)	2.34 (0.44)	1.05 (0.76)	0.34 (0.16)
0	0.60 (0.76)	1.40 (1.43)	1.02 (0.33)	-0.25 (0.49)	-0.38 (0.90)	0.15 (0.17)
1	4.75 (0.94)	5.36 (1.63)	1.91 (0.39)	-1.22 (0.64)	-1.02 (1.18)	0.48 (0.23)
2	23.90 (1.14)	14.60 (1.96)	9.71 (0.45)	-1.67 (0.77)	0.56 (1.40)	0.39 (0.27)
3	36.70 (1.25)	23.50 (2.15)	14.40 (0.48)	-0.09 (0.93)	1.46 (1.61)	1.32 (0.31)
4	47.00 (1.31)	29.60 (2.19)	18.50 (0.52)	14.40 (1.10)	10.10 (1.79)	6.00 (0.37)
5	55.50 (1.37)	37.70 (2.26)	22.50 (0.54)	29.60 (1.19)	19.90 (1.93)	11.60 (0.41)
6	64.10 (1.43)	41.50 (2.35)	26.20 (0.56)	41.60 (1.25)	28.00 (2.04)	16.80 (0.45)
7	67.50 (1.43)	44.50 (2.38)	27.40 (0.57)	49.70 (1.27)	34.00 (2.13)	20.30 (0.47)
8	72.00 (1.43)	46.50 (2.38)	28.90 (0.58)	56.80 (1.29)	36.90 (2.19)	22.80 (0.49)
9	73.20 (1.46)	46.80 (2.40)	30.30 (0.58)	61.10 (1.31)	39.30 (2.23)	25.10 (0.51)
10	75.50 (1.47)	47.30 (2.44)	30.60 (0.60)	65.50 (1.32)	42.20 (2.24)	27.10 (0.52)
11	77.20 (1.47)	47.70 (2.44)	31.70 (0.60)	68.70 (1.34)	44.80 (2.26)	28.70 (0.53)
12	78.60 (1.47)	48.70 (2.50)	32.10 (0.61)	71.10 (1.34)	45.40 (2.29)	30.10 (0.54)
13	79.10 (1.50)	49.40 (2.50)	33.10 (0.61)	73.40 (1.35)	46.50 (2.31)	31.40 (0.55)
14	80.20 (1.51)	48.10 (2.53)	33.60 (0.62)	75.00 (1.35)	48.00 (2.33)	32.10 (0.55)
15	80.60 (1.51)	48.40 (2.53)	34.40 (0.62)	76.20 (1.37)	47.80 (2.34)	33.30 (0.56)
16	82.10 (1.52)	48.00 (2.54)	34.80 (0.63)	77.90 (1.37)	48.30 (2.39)	34.10 (0.57)
17	80.70 (1.52)	48.50 (2.58)	35.30 (0.63)	79.10 (1.37)	48.70 (2.43)	34.70 (0.57)
18	81.20 (1.53)	50.10 (2.53)	35.90 (0.63)	79.20 (1.38)	50.20 (2.42)	35.90 (0.57)

Notes: This table presents numerical values for Figure 6 in the main body. Coefficients pertain to estimates of the time path of 30-day and 90-day mortgage default rates around a change in monthly mortgage payments from adjustable rate mortgage resets. Standard errors clustered at the loan level.

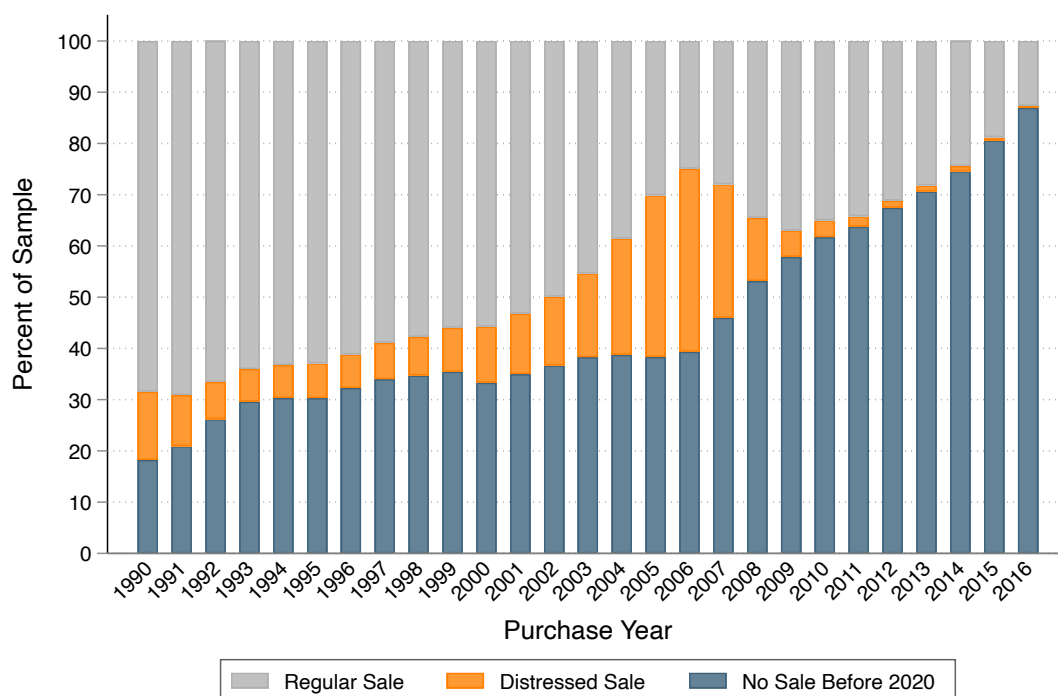
B Appendix Figures

Figure A1: Changes in Sample Size Due to Sample Restrictions



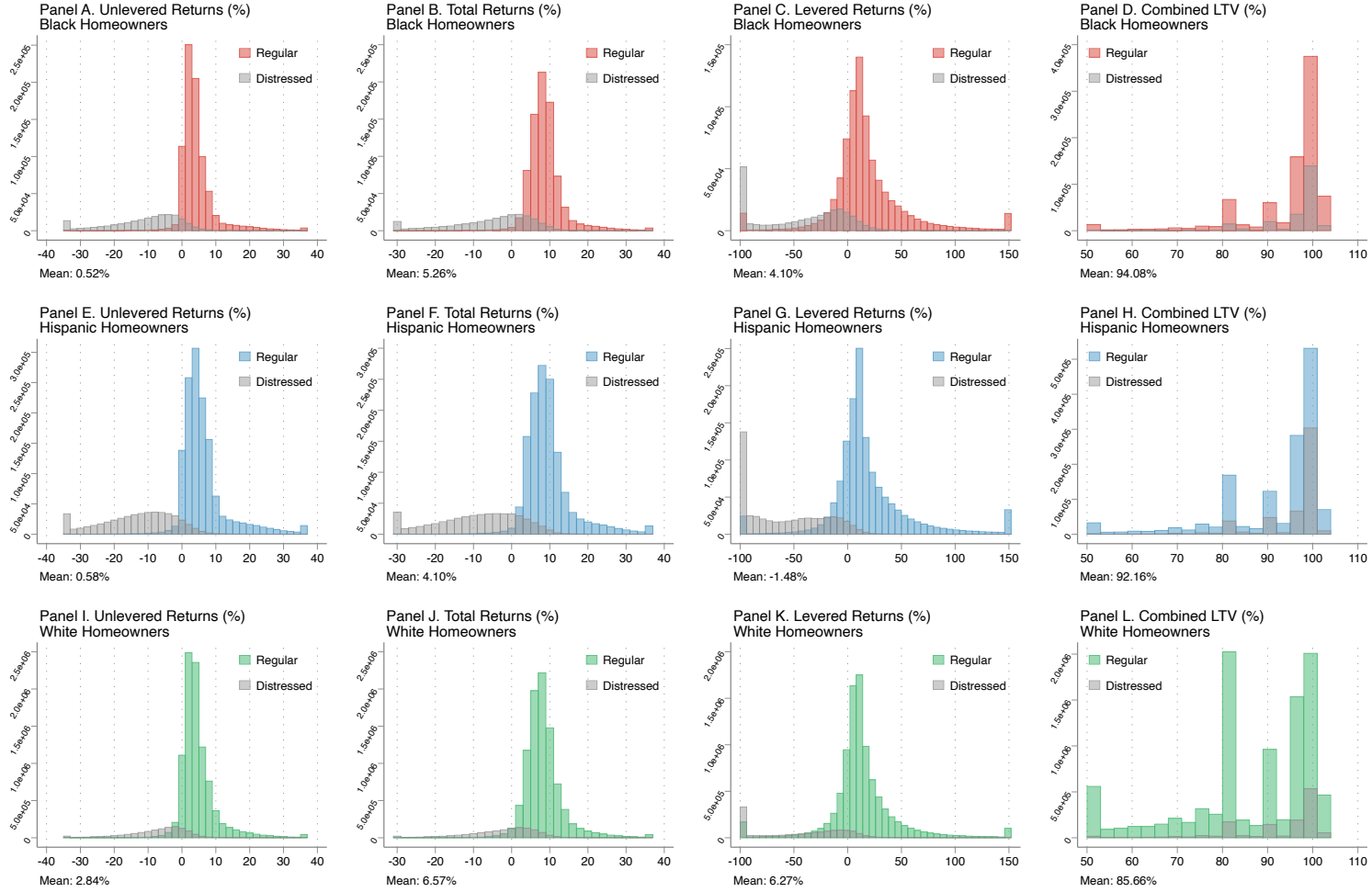
Notes: This figure illustrates the changes in sample size created by the sample restrictions we make to arrive at our baseline sample of ownership spells. Each point represents the number of properties that were purchased in a given year in a given subsample of the raw property data. *All Residential Purchases* denotes all arm's length, full consideration purchases in the ATTOM property data, for which the property use code is labeled as residential. *Non-Missing Purchase Price* restricts to purchases where a transaction price has been recorded. *Persons/Trusts* restricts to purchases made by persons or trusts. *After Algorithm* restricts to (i) properties with no subsequent sale by March 2020 and (ii) properties for which the repeat sales algorithm has identified the subsequent sale, but drops properties for which the repeat sales algorithm cannot be implemented (e.g., due to a missing transaction between consecutive ownership spells). *After HMDA Merge* restricts to mortgaged purchases where the mortgage can be linked to a record in the HMDA data. *Dropping Flippers and Outliers* drops ownership spells lasting less than 12 months, purchases with prices that are less than \$10,000 or with combined loan-to-value ratios of more than 102.5%. *Black, Hispanic, White; Owner-Occupants* restricts to purchases made by owner-occupants identifying as Hispanic, non-Hispanic Black, or non-Hispanic White. This sample represents our baseline analysis sample used to estimate racial gaps in housing returns. *Additional Merge with McDash* restricts to properties where the mortgage has been linked to the McDash servicing records. See Appendix C for more details.

Figure A2: Composition of Baseline Sample by Purchase Year and Sale Type



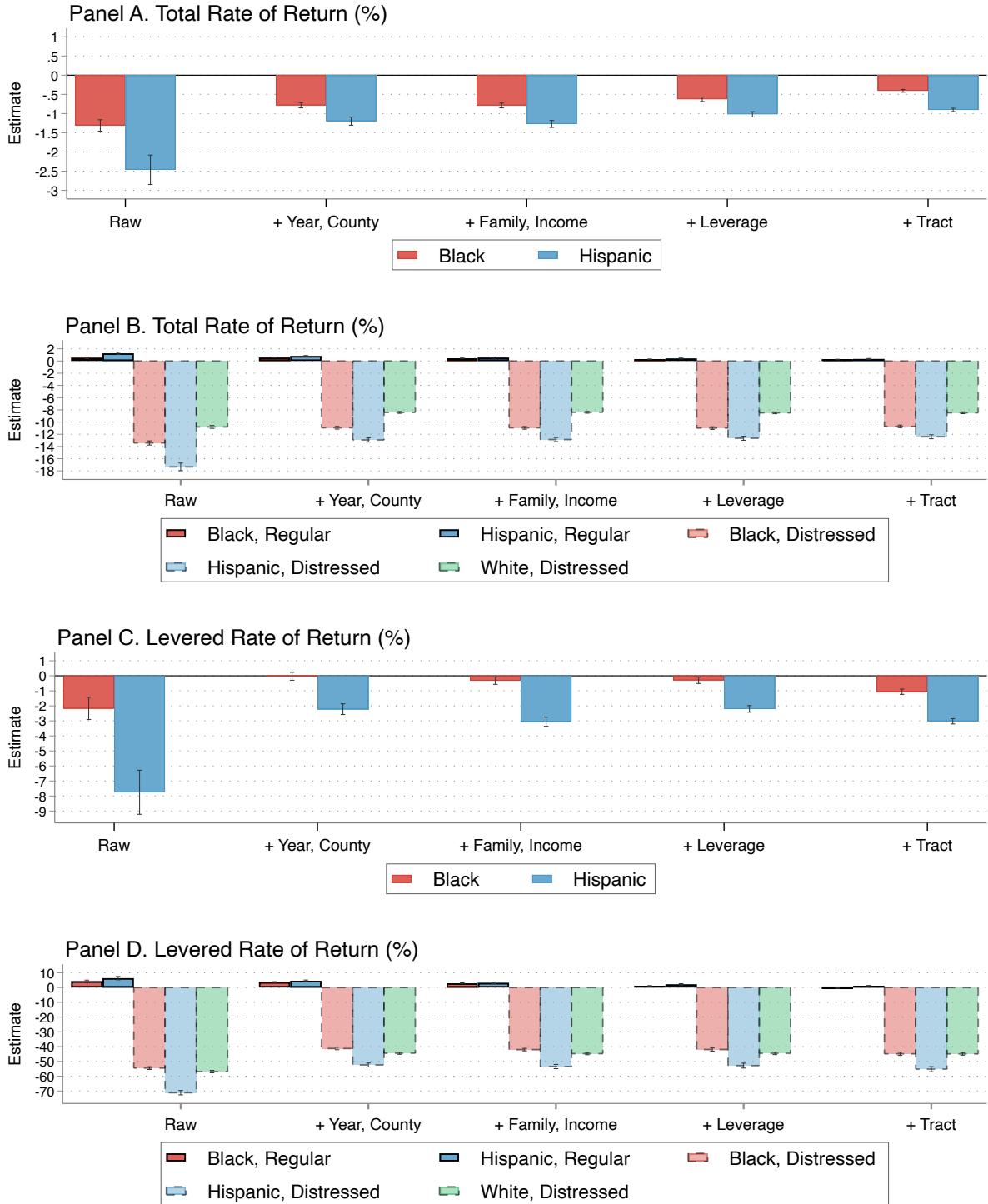
Notes: This figure illustrates the composition of sale types in our baseline sample of ownership spells. Each bar plots the percent of ownership spells belonging to a given type beginning in a given year. *Regular Sale* denotes properties that were purchased in a given year and sold by March 2020 in a non-distressed sale. *Distressed Sale* denotes properties that were sold by March 2020 in a distressed sale. *No Sale Before 2020* denotes properties that were not sold by March 2020. See Section I for more details on the baseline sample.

Figure A3: Distribution of Annualized Returns and Combined Loan-to-Value Ratio



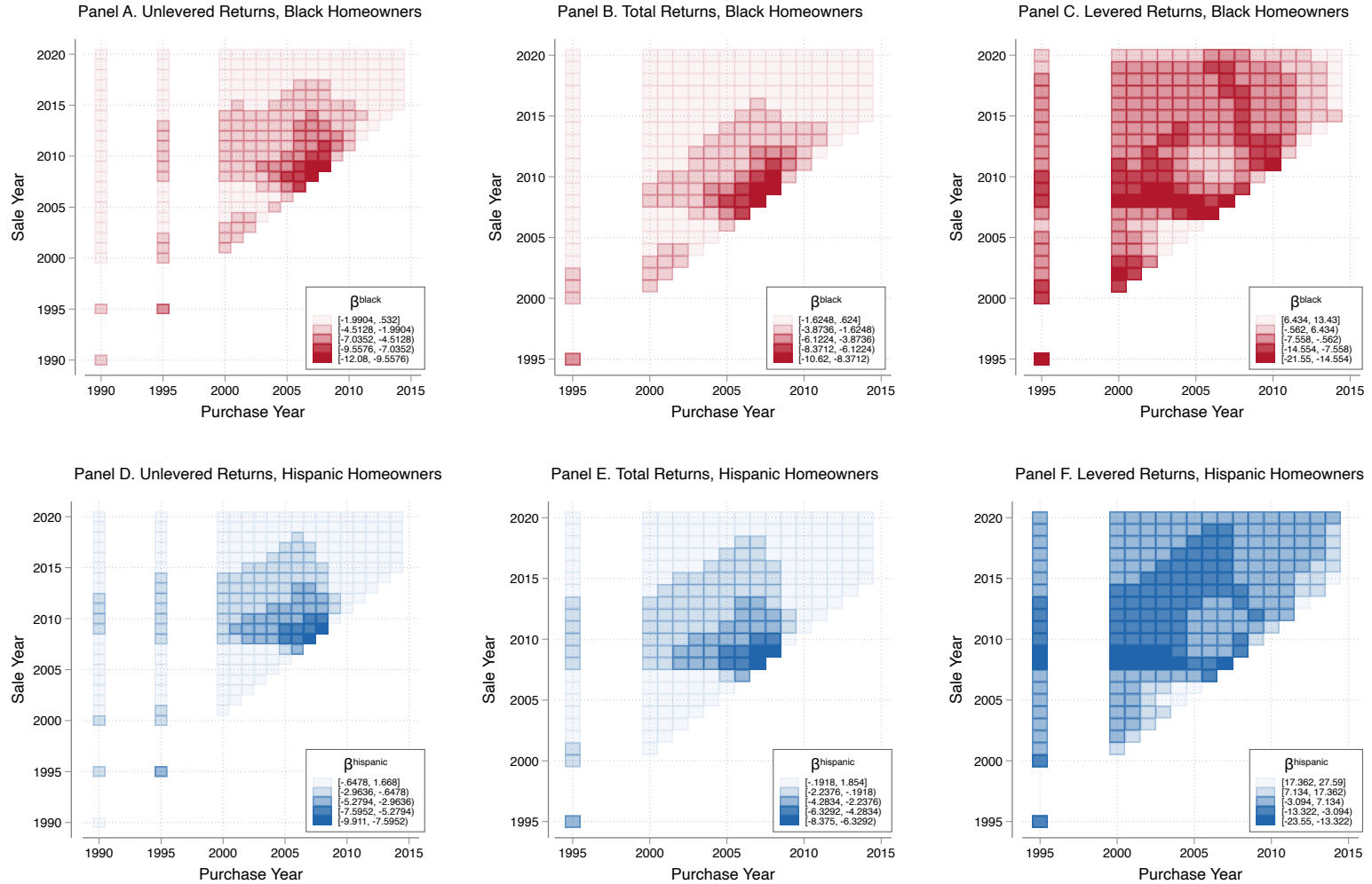
Notes: This figure plots the distribution in frequencies of annualized unlevered returns, total returns, levered returns, and combined loan-to-value ratio at origination. Distributions are plotted in separate panels for Black, Hispanic and White homeowners. For each race/ethnicity, the distributions for regular sales and distressed sales are presented separately within each panel. Distributions are winsorized for clarity. Data are from baseline sample of ownership spells described in Section I.

Figure A4: Racial Gaps in Total and Levered Housing Returns



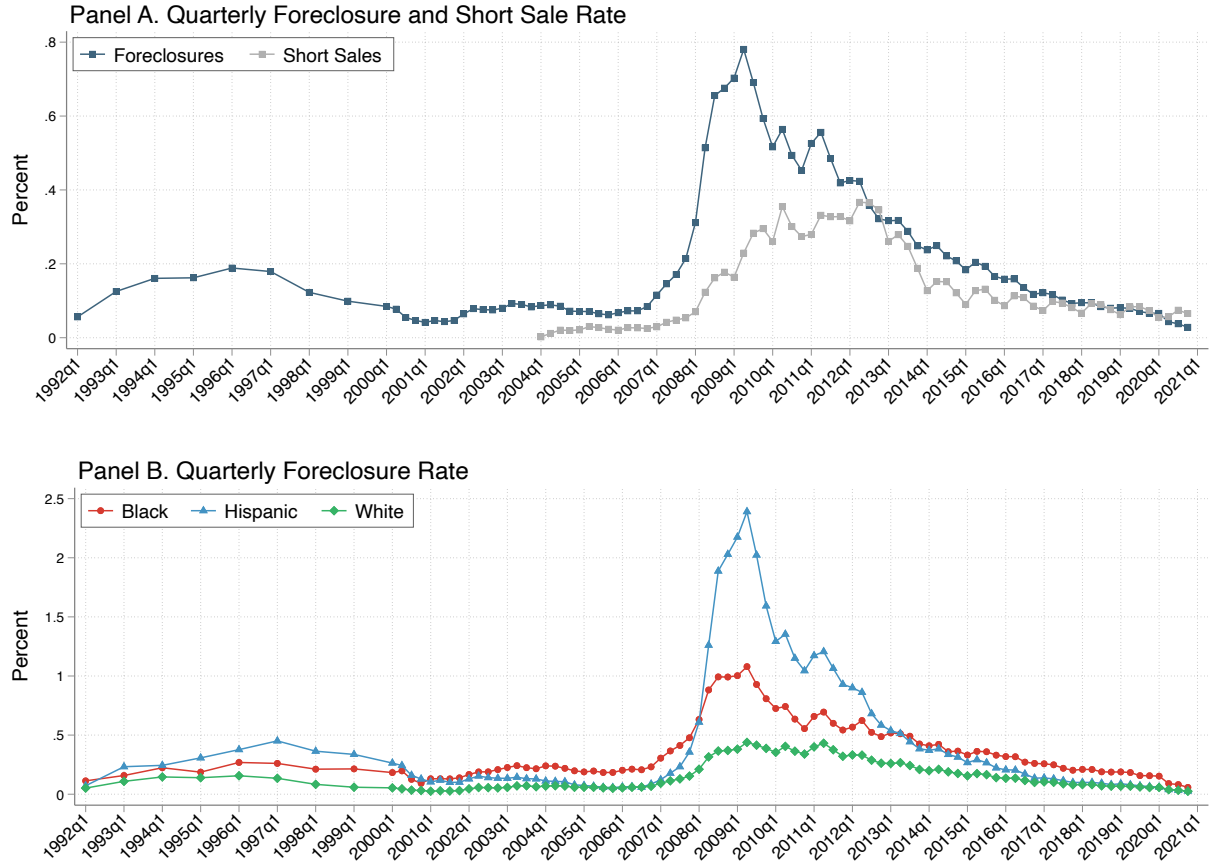
Notes: This figure presents estimates of racial gaps in levered and total housing returns. In Panels A and C, each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. In Panels B and D, each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in purchase year-by-county fixed effects. *+ Family, Income* adds fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, and fixed effects that interact county, purchase year, and deciles of income at home purchase. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year and Census tract. Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I.

Figure A5: Heat Map of Returns Gap by Purchase and Sale Year



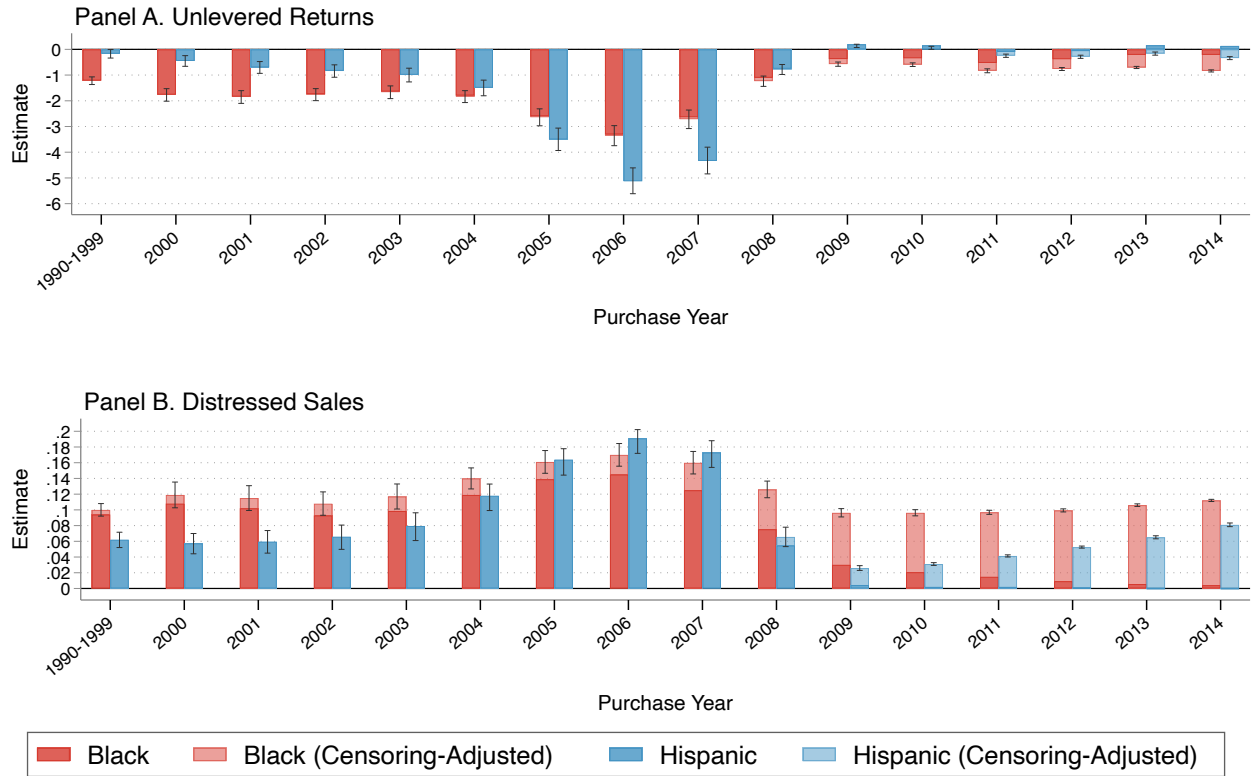
Notes: This figure presents estimates of racial gaps split by year of purchase and year of sale for unlevered returns (Panels A and D), total returns (Panels B and E), and levered returns (Panels C and F). Within each panel, the color of a square indicates the size of the estimated coefficient, with each square corresponding to a coefficient for race/ethnicity indicators in separate regressions estimated within purchase year-by-sale year cells. Regressions are estimated as in Equation 4 with county fixed effects. For purchase year, 1990 denotes period 1990-1994 and 1995 denotes period 1995-1999. For sale year, 1990 denotes 1990-1994 and 1995 denotes 1995-2001. Data are from baseline sample of ownership spells described in Section I.

Figure A6: Time Series of Aggregate Distressed Sales



Notes: This figure plots aggregate trends in distressed sales and housing returns from 1992 to 2020, and illustrates higher rates of distressed sales among minority homeowners throughout this period. Panel A plots quarterly foreclosure and short sale rates, defined as the percent of ownership spells beginning prior to a given quarter and ending in a foreclosure or short sale in that quarter. Panel B plots the quarterly foreclosure rate by race/ethnicity. Data are from baseline sample of ownership spells described in Section I, and includes sales that occurred after March 2020.

Figure A7: Gaps in Returns and Distress by Purchase Year



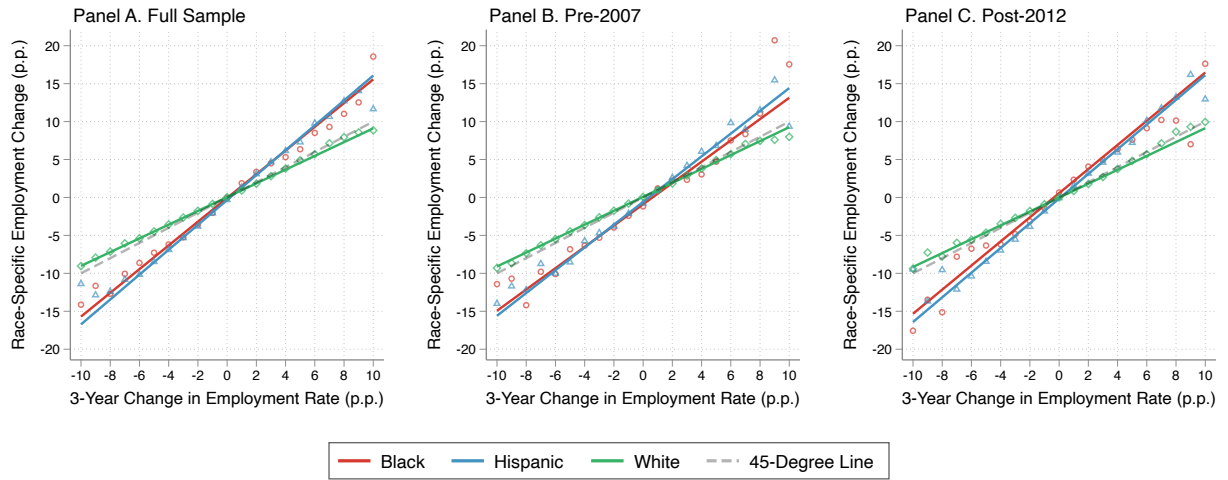
Notes: This figure presents estimates of racial gaps in annualized unlevered housing returns and distressed sale rates from regression specifications that compare homeowners living in the same county (Equation 4 estimated with county fixed effects). Each pair of bars corresponds to a separate regression, and each bar denotes an estimated coefficient corresponding to a race/ethnicity indicator. The outcome in Panel A is annualized unlevered housing returns, and the outcome in Panel B is the distressed sale rate. In both panels, outcomes are adjusted to account for distressed sales that occur out of our sample window using the procedure described in Section II.A. Darkly shaded region denotes size of unadjusted gaps. Data are from baseline sample of ownership spells described in Section I. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A8: Racial Gap in Foreclosure Rate Under Different Local Economic Conditions



Notes: This figure plots racial differences in foreclosure rates as a function of local economic conditions. The foreclosure rate is computed as the share of properties in our baseline sample that end in a foreclosure in a given year and county. We compute the Black-White and Hispanic-White difference in foreclosure rates for each county and year. Each panel presents a binned scatterplot of racial differences in foreclosure rates against a measure of local economic conditions. In Panels A and C, local economic conditions are measured as the county unemployment rate from BLS. In Panels B and D, local conditions are measured as house price growth in the preceding 3 years. Panels C and D restrict to county-years in 2012 and later. β denotes slope of estimated linear fit. Data are from baseline sample of ownership spells described in Section I.

Figure A9: Racial Differences in Employment Cyclicity



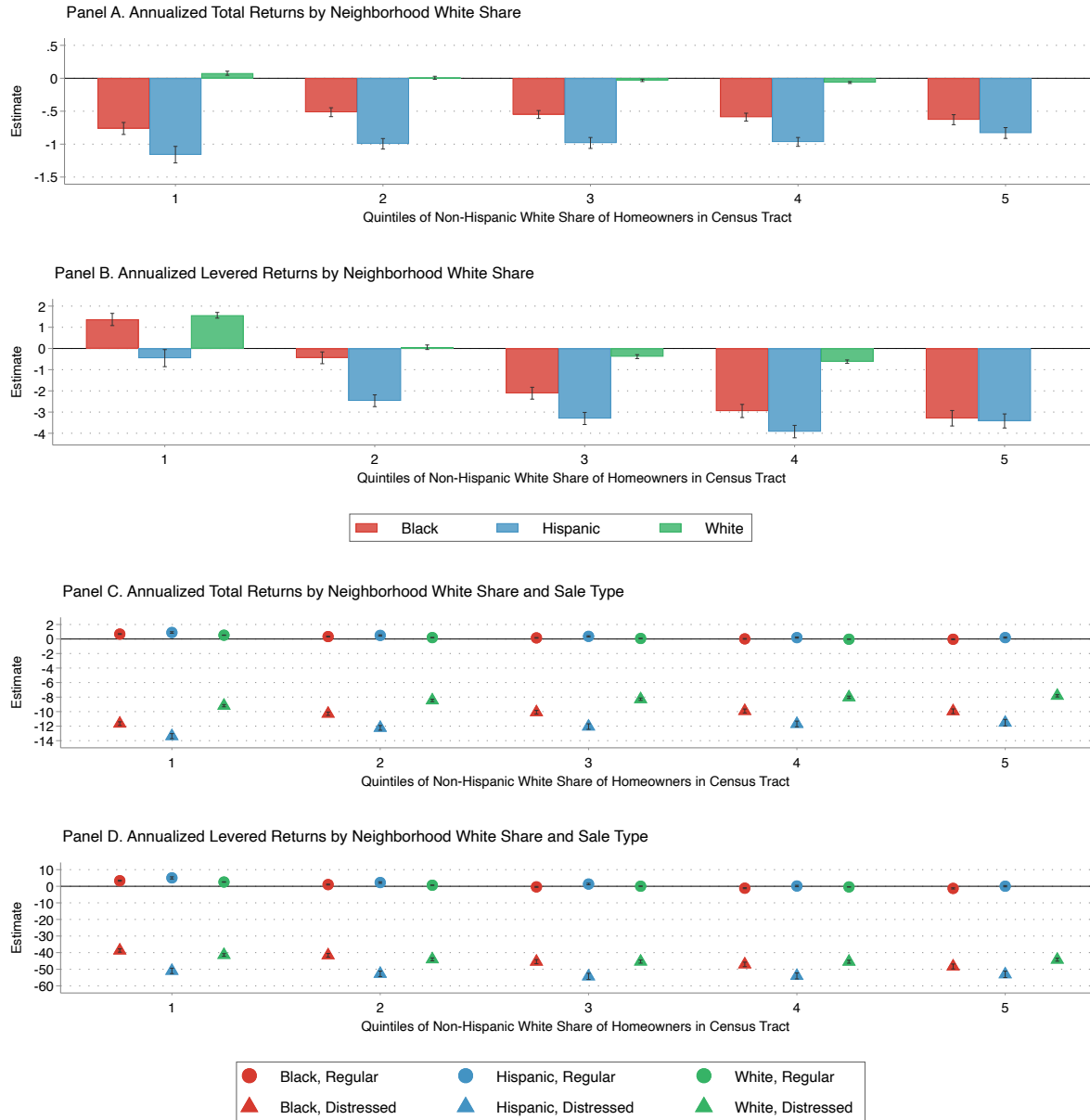
Notes: This figure plots the relationship between race-specific changes in county-level employment rates and changes in employment rates pooling the three racial groups. Panel A presents results including data from 2000 to 2019. Panel B restricts to years in 2007 and before. Panel C restricts to years in 2012 and after. Yearly employment counts are measured using Quarterly Workforce Indicator data from the US Census (United States Bureau of the Census, 2020). Yearly working-age population counts are derived from Census intercensal population estimates. Binned scatterplot and linear fit restrict to changes in employment rate between -10 and 10 percentage points. County-race-year observations are weighted by working age population.

Figure A10: Heterogeneous Racial Gaps



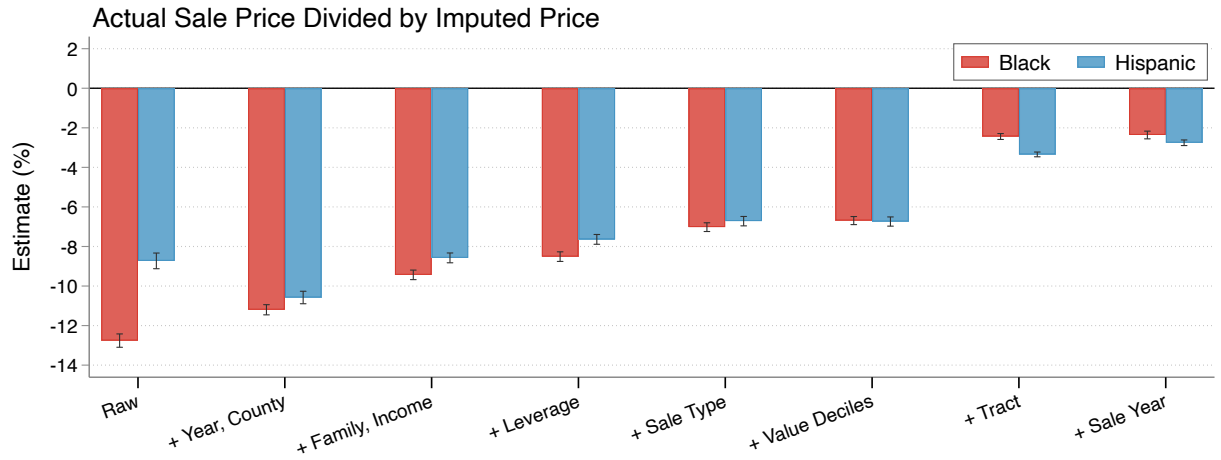
Notes: This figure documents heterogeneity in the racial gap in annual unlevered housing returns (Panel A), total housing returns (Panel B), levered housing returns (Panel C), and distressed sales (Panel D). Each dimension of heterogeneity provides estimates from a separate regression (Equation 4). Regressions include fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, fixed effects that interact county, purchase year, and deciles of income at home purchase, and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. Points denote estimated coefficients of race/ethnicity indicators interacted with homeowner characteristics (e.g., indicators for income tercile). *Baseline* denotes the full analysis sample. *Age* refers to a proxy of age defined as the age of the homeowner's oldest trade in the Equifax data plus 18 years. *FHA* denotes indicates the mortgage is identified as a loan from the Federal Housing Administration in the property data. *Home Value Tercile* denotes tercile of purchase price within county-year. *SFR* and *Condo* indicate that the property is a single-family residence and condominium, respectively. *Sq. Ft.* denotes interior square footage. *Bedrooms* refers to the number of bedrooms in the property. *DTI Tercile* denotes tercile of back-end debt-to-income ratio, measured at origination in mortgage servicing records provided by McDash. *Urban* denotes tracts in which all constituent Census blocks are urban, according to 2010 Census definitions, while *Rural* denotes tracts with at least one rural block. Data are from baseline sample of ownership spells described in Section 1. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A11: Racial Gaps by Neighborhood Demographics and Sale Type



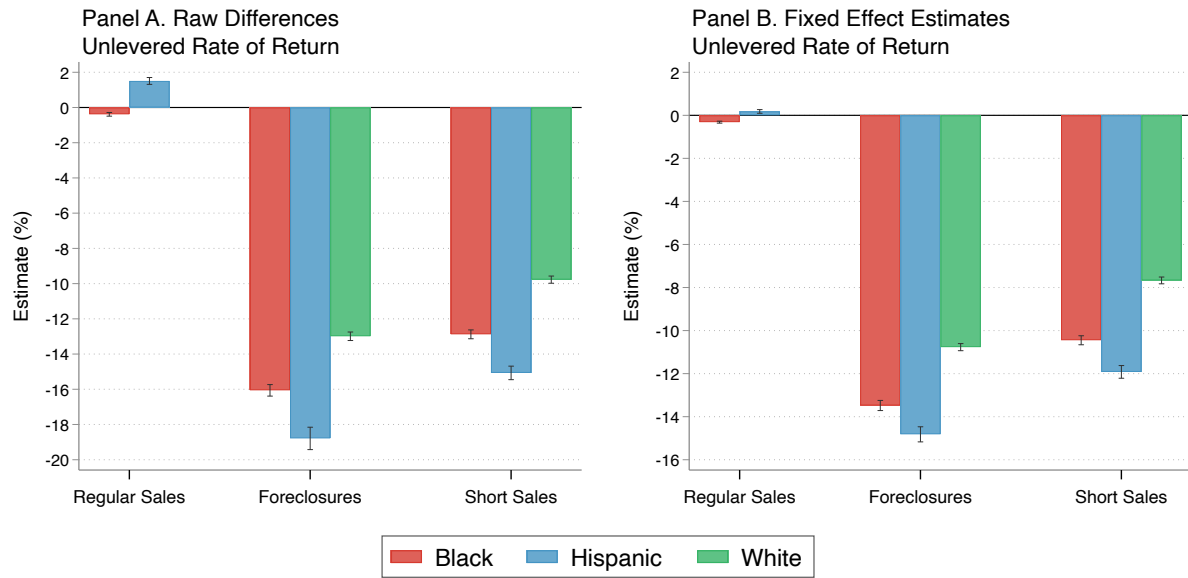
Notes: This figure presents estimates of racial gaps in annualized total housing returns (Panels A and C) and levered returns (Panels B and D) from regression specifications that compare homeowners with similar incomes, family structures, and leverage, living in the same county and buying their homes in the same year (Equation 4). Panels A and B present regression coefficients that interact individual race/ethnicity with quintiles of the non-Hispanic White share of homeowners in the individual's Census tract. The omitted category is non-Hispanic White homeowners in neighborhoods with the highest non-Hispanic White share. Panels C and D present regression coefficients that interact homeowner race/ethnicity with quintiles of the White share and homeowner's sale type (regular vs. distressed). Regular sales include properties that were not sold within our sample period. The omitted category in Panels C and D is White homeowners in neighborhoods with the highest White share whose property sale is not distressed. In all panels, quintiles are assigned within each county, such that higher quintiles contain neighborhoods in each county with the highest share of White homeowners. Data are from baseline sample of ownership spells described in Section I. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A12: Racial Gaps in Distressed Sale Discounts



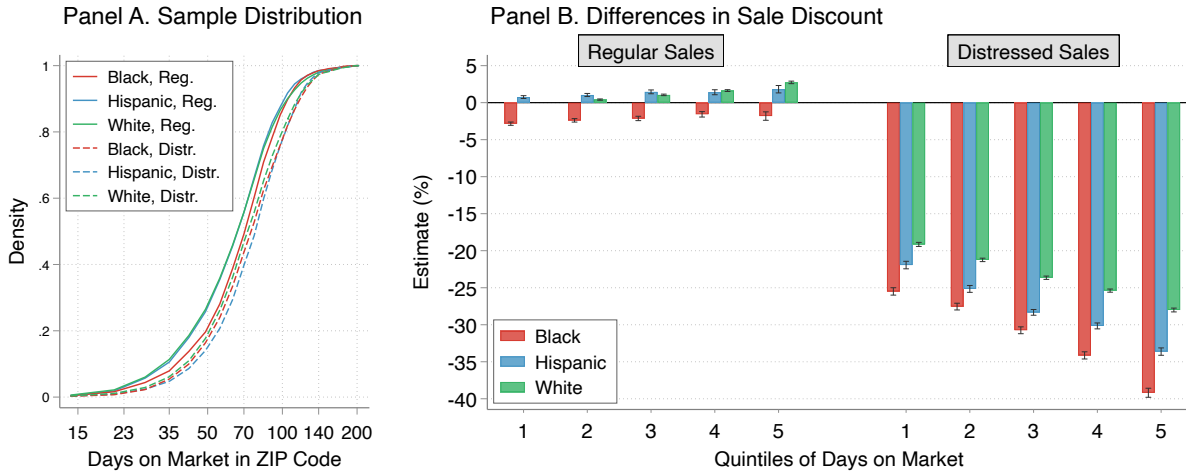
Notes: This figure presents estimates of racial gaps in distressed sale discounts. Each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. The outcome is a property's actual sale price divided by its imputed value, computed by inflating purchase price by the FHFA county-level home price index. Sample is restricted to properties sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in fixed effects that interact purchase year, sale year, and county. *+ Family, Income* adds fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, and fixed effects that interact county, purchase year, and deciles of income at home purchase. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Sale Type* adds fixed effects that interact purchase year, county, and an indicator that a sale is a short sale. *Value Deciles* adds fixed effects that interact purchase year, county, and deciles of property purchase price within county and purchase year. *Tract* adds fixed effects that interact purchase year and Census tract. *Sale Year* adds fixed effects that interact purchase year, sale year, and Census tract. Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I, restricted to distressed sales.

Figure A13: Gap in Housing Returns by Distressed Sale Type



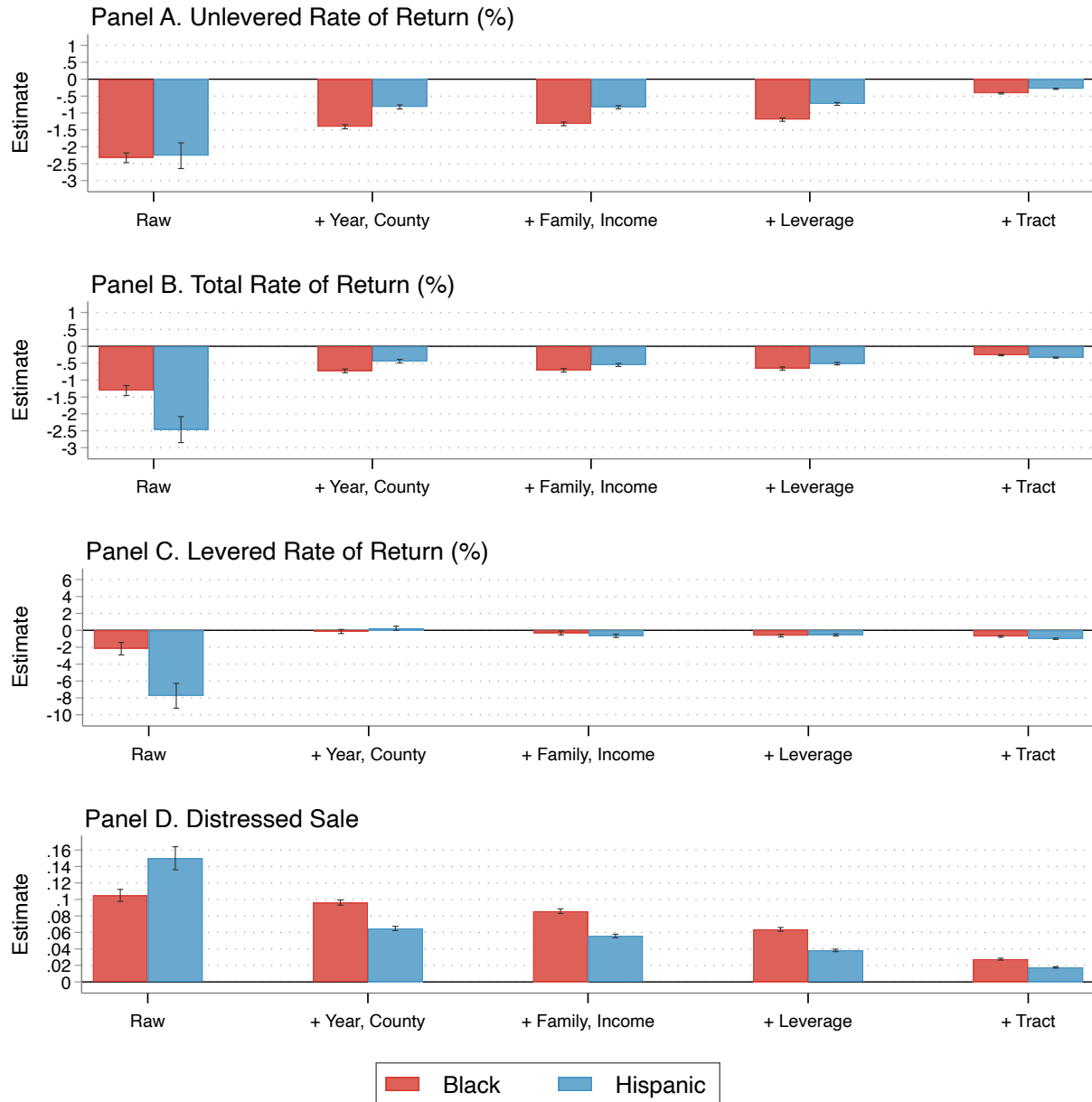
Notes: This figure presents estimates of racial gaps in housing returns. Each panel contains estimates from a separate regression (Equation 4). Bars depict estimated coefficients of Black and Hispanic indicators interacted with the sale type (regular, foreclosure, short sale). Panel A presents raw differences. Panel B includes fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, fixed effects that interact county, purchase year, and deciles of income at home purchase, and fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. The outcome in both panels is the unlevered rate of return (Equation 1). Data are from baseline sample of ownership spells described in Section I. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A14: Heterogeneity in Sale Discounts by Market Thickness



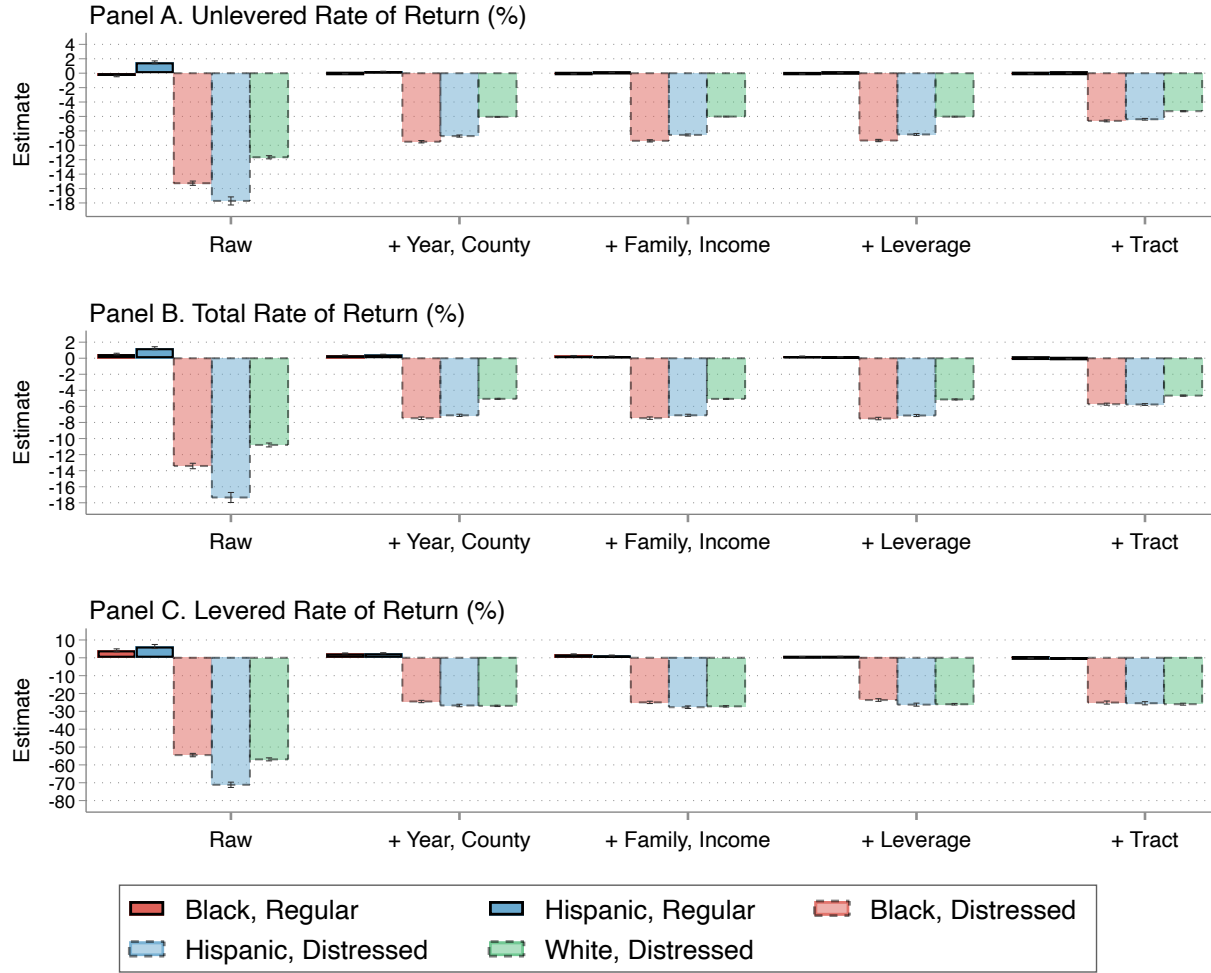
Notes: This figure presents racial differences in exposure to housing market thickness and estimates of racial gaps in annualized unlevered housing returns split by housing market thickness and sale type. Housing market thickness is measured by the median days on market of homes sold in a ZIP code, computed using MLS data. Panel A plots the CDF of days on market, split by racial group and sale type (distressed vs. regular). Panel B presents coefficients from a regression that interacts race indicators, sale type, and quintiles of days on market (Equation 4). The omitted category is White homeowners living in the quintile of ZIP codes in the thickest housing markets. The outcome is a property's actual sale price divided by its imputed value, computed by inflating purchase price by the FHFA ZIP-level home price index. The specification includes fixed effects that interact county, purchase year, and sale year. Data are from baseline sample of ownership spells described in Section I, restricted to properties that are sold within sample window and ZIP-years merged with the MLS data. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A15: Racial Gaps Controlling for Sale Year



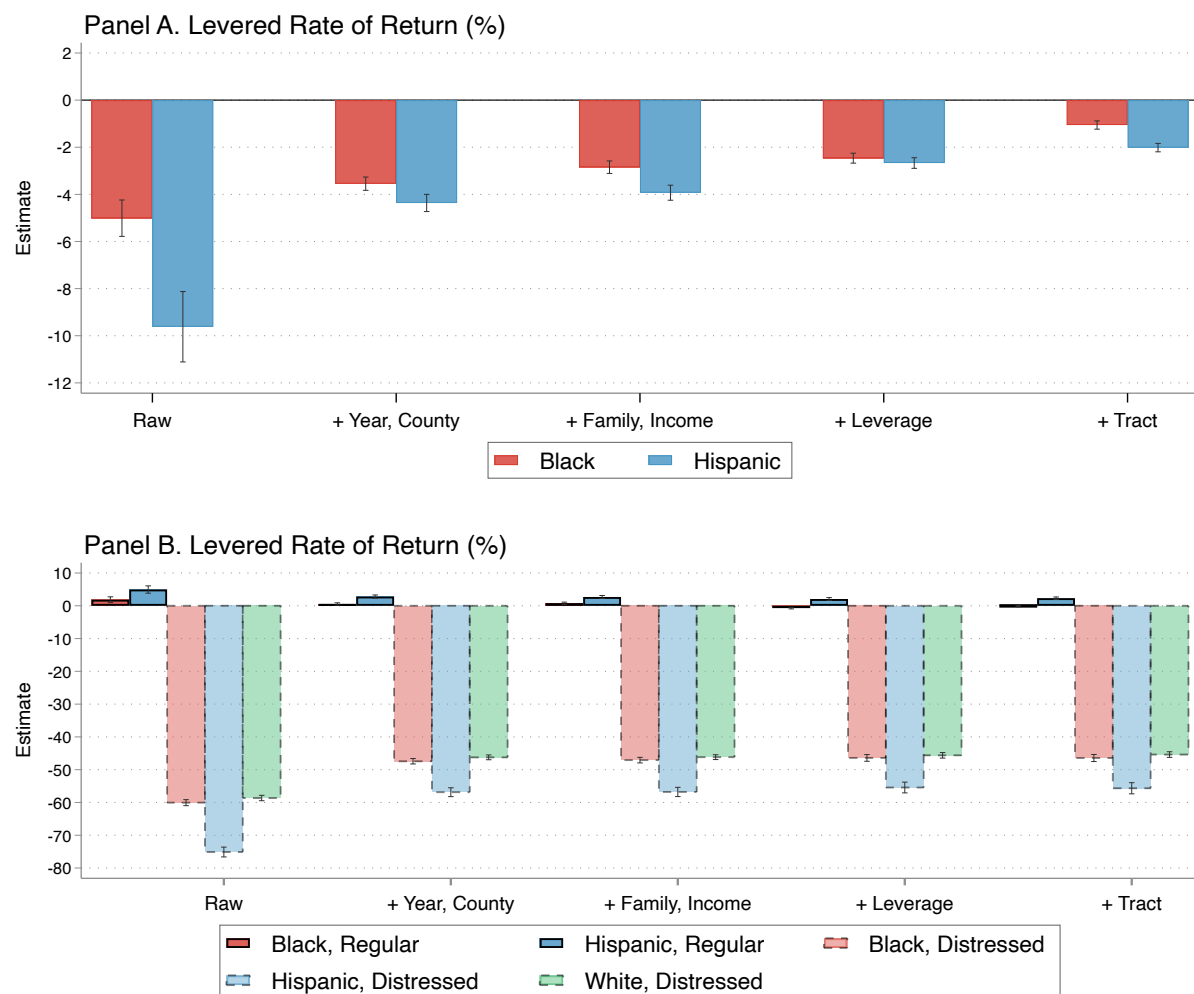
Notes: This figure presents estimates of racial gaps in housing returns and distressed sales. Each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. *Raw* denotes estimates without controls. *+ Year, County* adds in fixed effects that interact purchase year, sale year, and county. *+ Family, Income* adds fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, and fixed effects that interact county, purchase year, and deciles of income at home purchase. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year, sale year, and Census tract. The outcomes are the annualized unlevered rate of return (Panel A), the annualized total rate of return (Panel B), the annualized levered rate of return (Panel C), and an indicator that a property is sold in a distressed sale (Panel D). Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I.

Figure A16: Racial Gaps Controlling for Sale Year, Continued



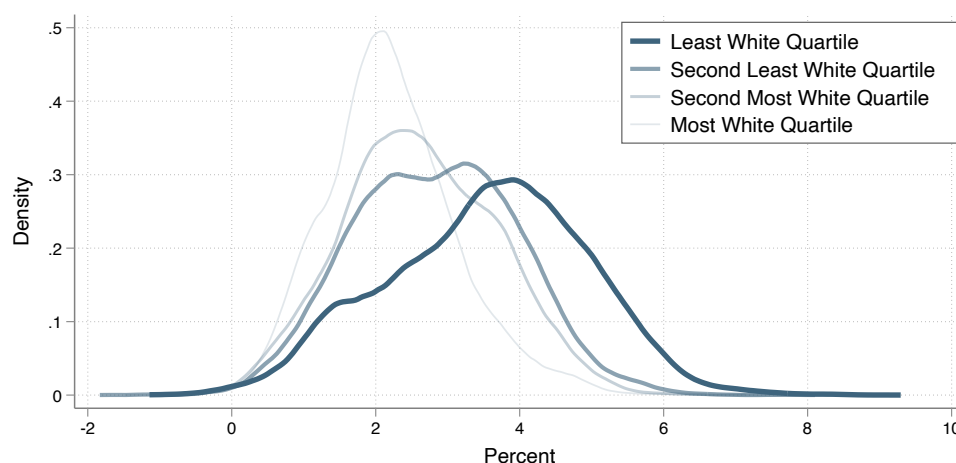
Notes: This figure presents estimates of racial gaps in housing returns by sale type. Each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in fixed effects that interact purchase year, sale year, and county. *+ Family, Income* adds fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, and fixed effects that interact county, purchase year, and deciles of income at home purchase. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year, sale year, and Census tract. The outcomes are the annualized unlevered rate of return (Panel A), the annualized total rate of return (Panel B), and the annualized levered rate of return (Panel C). Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I.

Figure A17: Racial Gaps in Levered Housing Returns Without Race-Specific Adjustment for Rents and Housing Costs



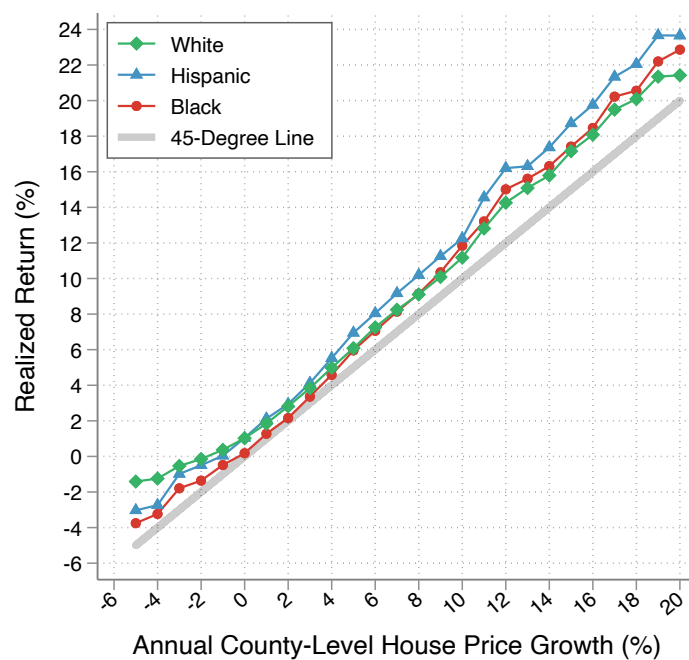
Notes: This figure presents estimates of racial gaps in levered housing returns that are not adjusted for race-specific differences in rents and housing costs. In Panel A, each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. In Panel B, each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in purchase year-by-county fixed effects. *+ Family, Income* adds fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, and fixed effects that interact county, purchase year, and deciles of income at home purchase. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year and Census tract. Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I.

Figure A18: Average House Price Growth 2001-2020 by Census Tract Demographics



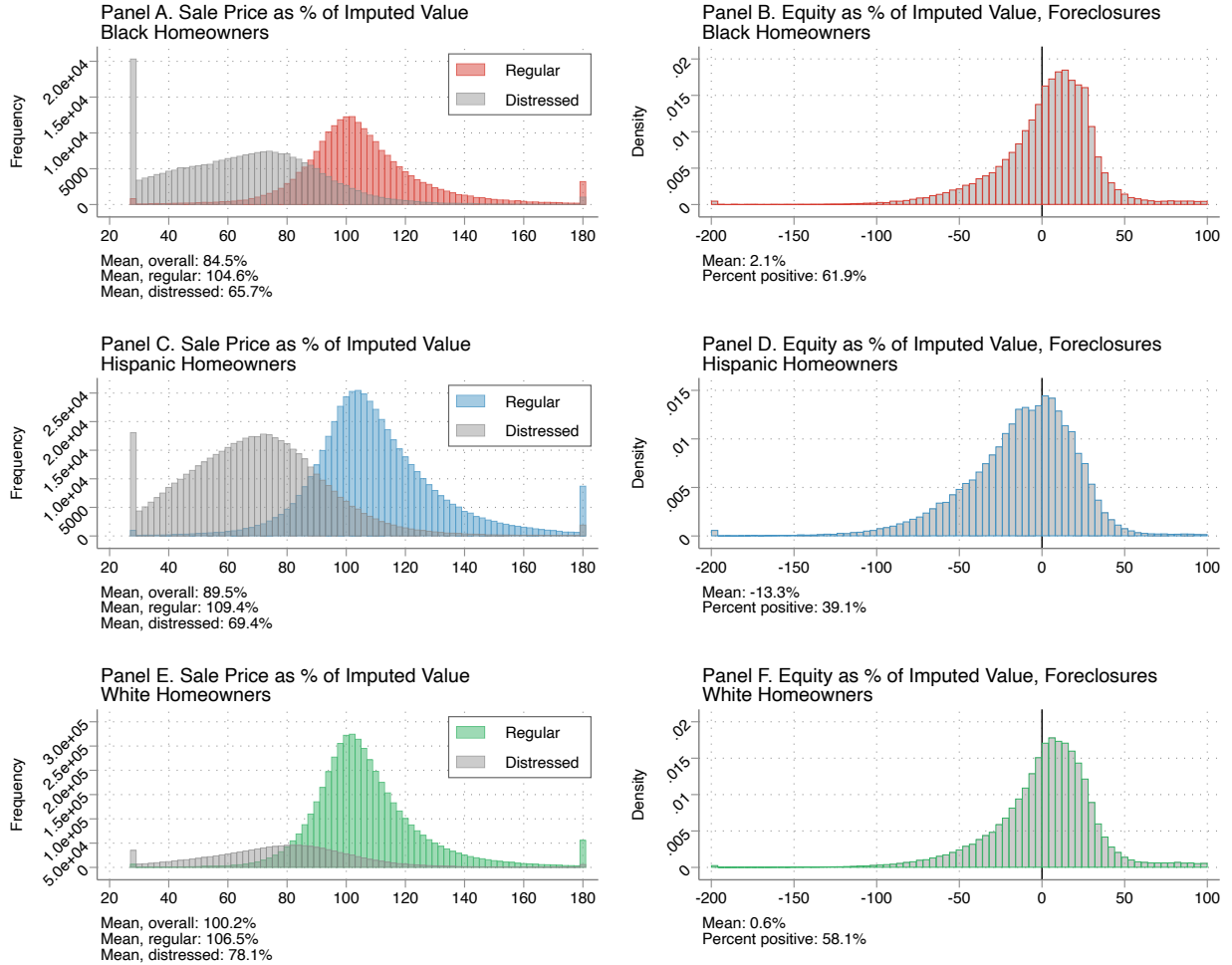
Notes: This figure presents the distribution of annual house price growth between 2001 and 2020 in kernel density form for US Census tracts. Tracts are categorized into quartiles of the share of homeowners in each tract identifying as non-Hispanic White in the 2010 Census (United States Bureau of the Census, 2022). Tract house price growth is measured using tract-level FHFA house price index (Federal Housing Finance Agency, 2021). These distributions indicate that tracts with more minority homeowners were more exposed to rapid levels of house price growth between 2001 and 2020.

Figure A19: Realized Returns vs. Local House Price Growth Among Non-Distressed Sales



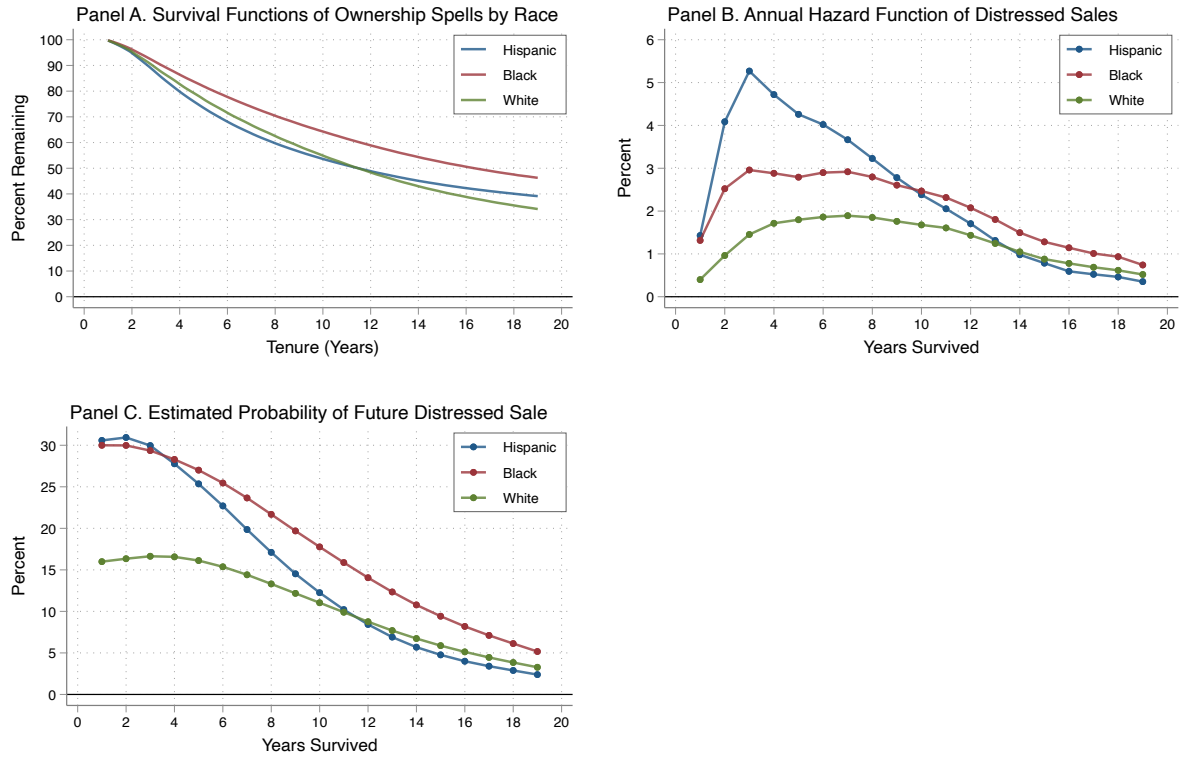
Notes: This figure presents a binned scatterplot of realized annual unlevered returns (Equation 1), plotted against annual local house price growth, computed for the years between property purchase and sale using the FHFA county-level house price index (Bogin et al., 2019). Data come from baseline sample of ownership spells described in Section I, restricted to properties that are sold within our sample window in a non-distressed sale.

Figure A20: Sale Price and Equity as Percentage of Imputed Value



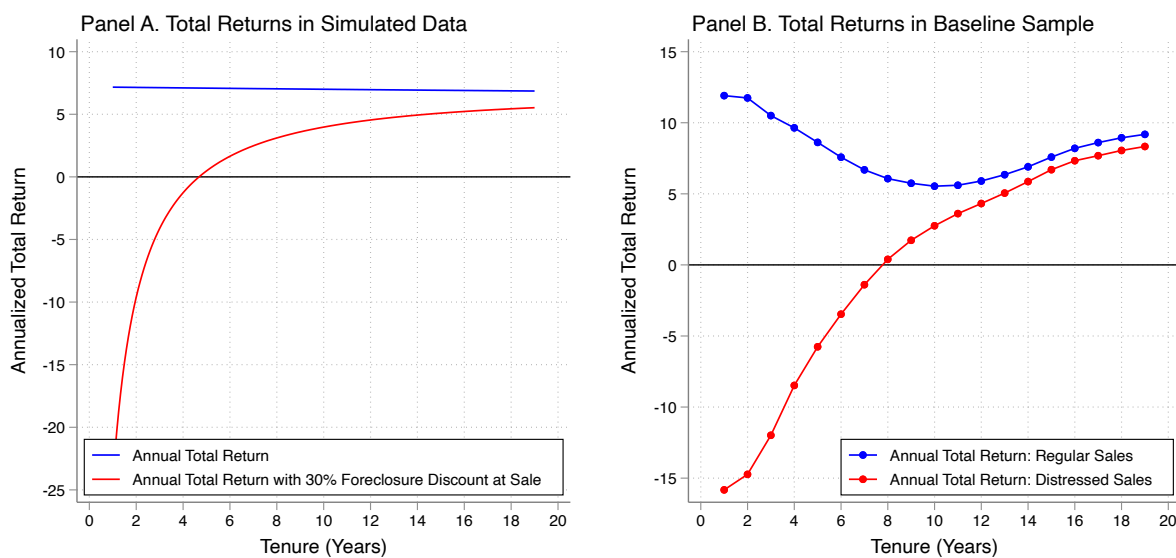
Notes: This figure plots the distribution of sale prices (Panels A, C, and E), and home equity at sale (Panels B, D, and F) as a percentage of the property's imputed value at time of sale. In Panels A, C, and E, imputed value at the time of sale is computed by inflating purchase prices using the FHFA county-level house price index (Bogin et al. 2019). Distributions are presented separately for regular sales and distressed sales. Data are from sample of homeowners with observed purchase and sale prices in baseline sample of ownership spells described in Section I. Data for Panels B, D, and F come from mortgage records from McDash, which contain unpaid principal balance at sale, and are restricted to foreclosed homes. Equity is computed by subtracting unpaid principal from imputed property value, calculated by inflating value at origination using the FHFA house price index. Distributions are plotted separately for Black (Panels A and B), Hispanic (Panels C and D), and White (Panels E and F) homeowners.

Figure A21: Survival and Hazard Functions of Distressed Sales



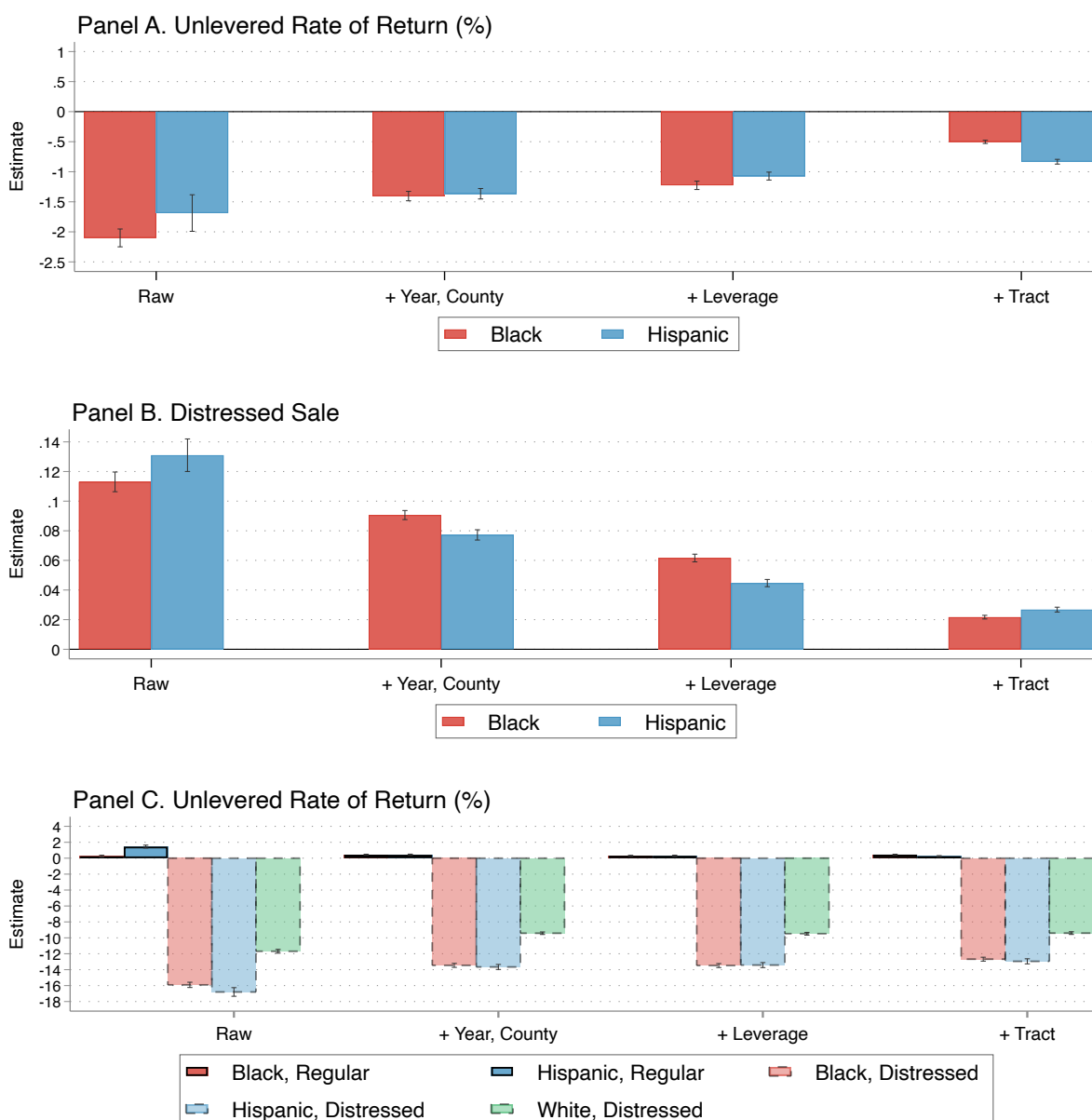
Notes: This figure illustrates the concentration of distressed sales early in tenure spells. Panel A plots Kaplan-Meier survival functions by race/ethnicity. Panel B plots the empirically-observed annual hazard function of distressed sales in our baseline sample (i.e., the percent of ownership spells ending in a distressed sale in the following year among the set of ownership spells lasting at least y years). Panel C plots the estimated probability of a future distressed sale conditional on an ownership spell lasting a given number of years.

Figure A22: Impacts of Distressed Sales on Returns at Longer Tenure Lengths



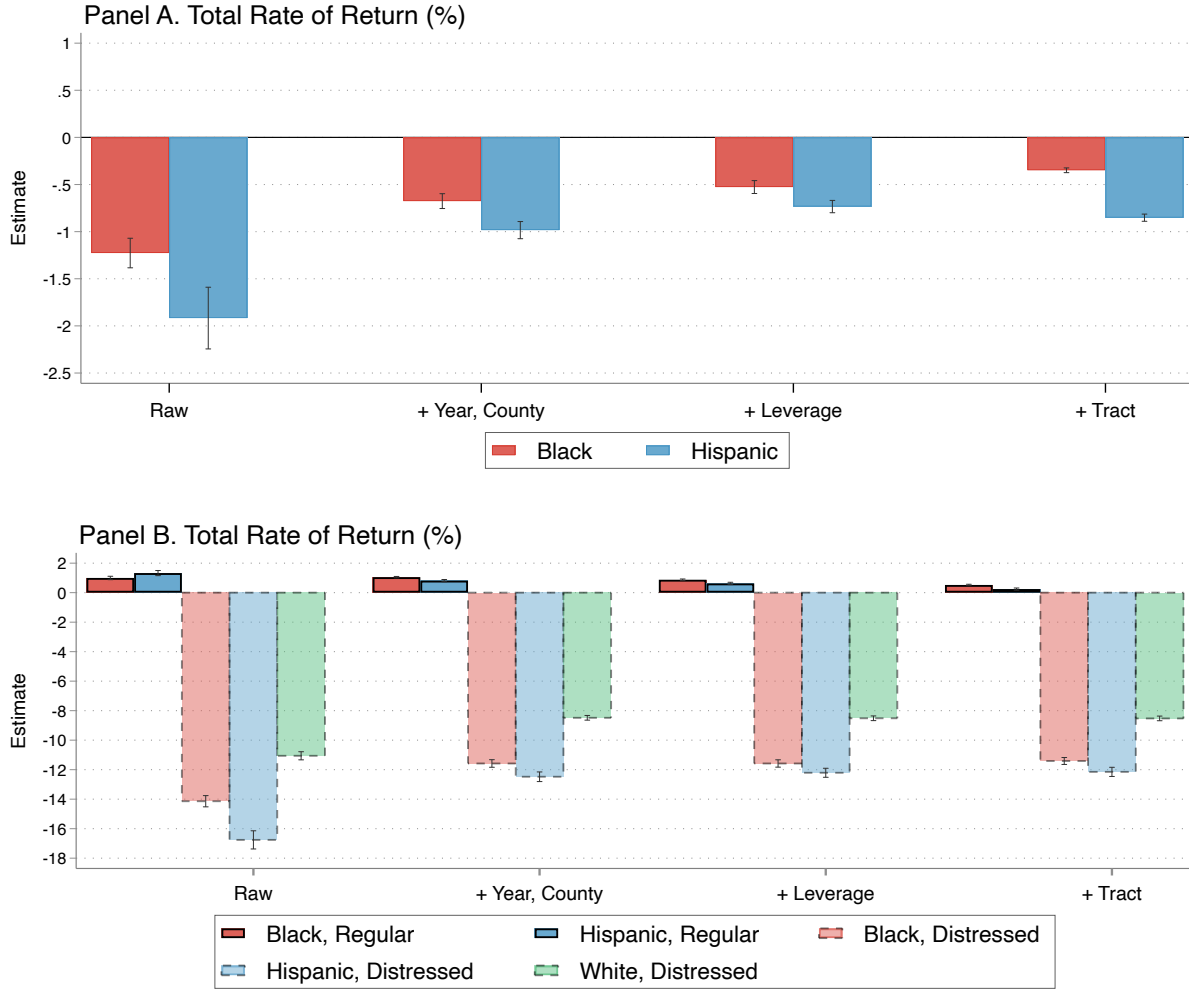
Notes: This figure illustrates the impact of foreclosure discounts on total returns as a function of tenure length. Panel A plots simulated annual total returns for a hypothetical property that is purchased for \$300,000 and appreciates at 3% per year. The property also yields a \$1,000 net rent dividend each month, which grows at 2% per year. Annual total returns are computed for each tenure length between 12 months and 40 years. Panel B plots analogous average returns derived from our baseline sample of ownership spells, restricted to properties where we observe both a purchase and a sale. In both panels, total returns are computed using formula for internal return, defined as the return that equates discounted cash flows.

Figure A23: Racial Gaps in Unlevered Housing Returns and Distressed Sales, Imputed Race/Ethnicity



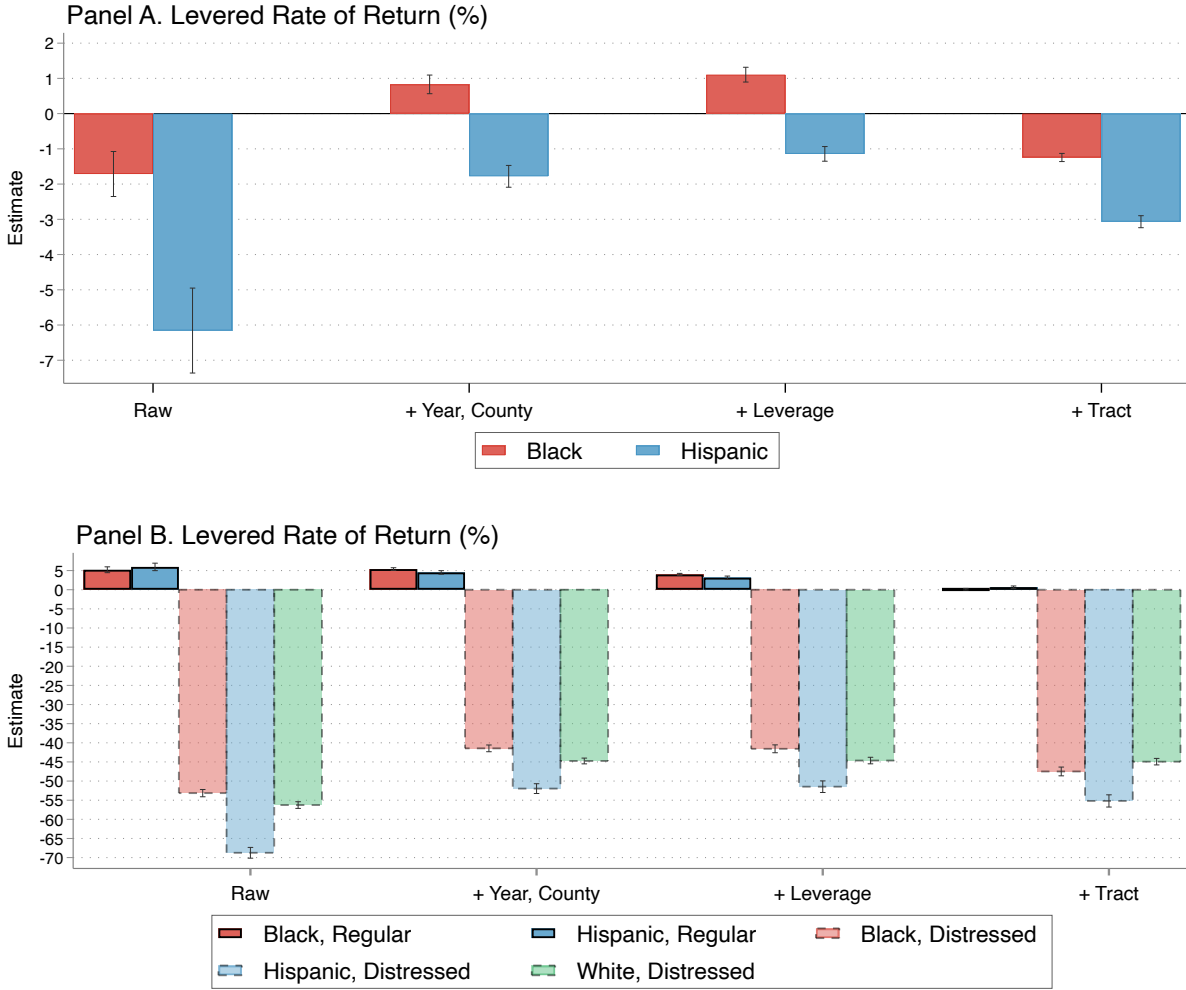
Notes: This figure presents estimates of racial gaps in unlevered housing returns and rates of distressed sales. Race/ethnicity has been imputed using homeowner name and location. In Panels A and B, each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. In Panel C, each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in purchase year-by-county fixed effects. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year and Census tract. The outcome in Panels A and C is the annualized unlevered rate of return (Equation 1), and the outcome in Panel B is an indicator that a homeowner experiences a distressed sale. Data are from baseline sample of ownership spells described in Section I. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A24: Racial Gaps in Total Housing Returns, Imputed Race/Ethnicity



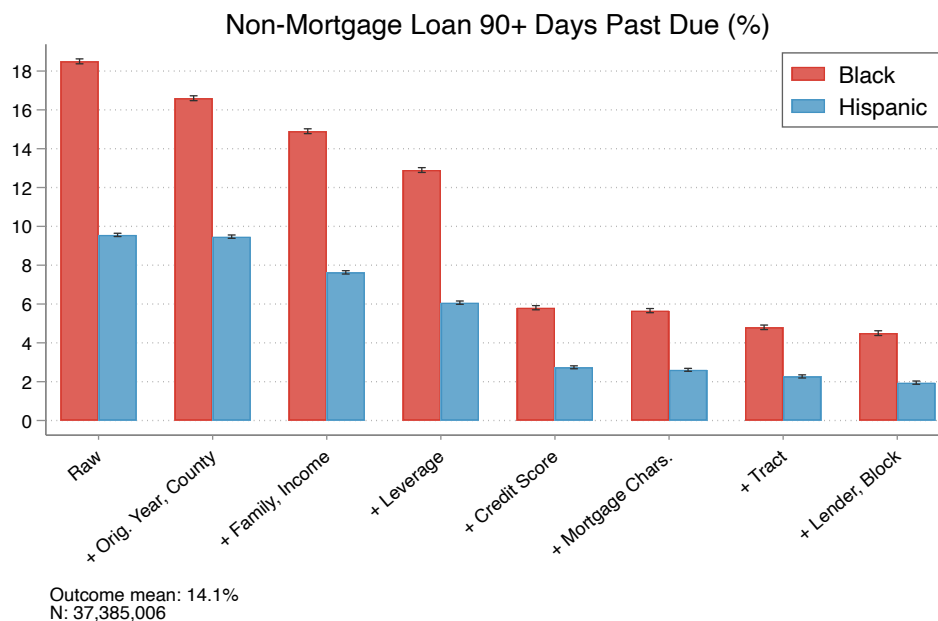
Notes: This figure presents estimates of racial gaps in total housing returns. Race/ethnicity has been imputed using homeowner name and location. In Panel A, each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. In Panel B, each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in purchase year-by-county fixed effects. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year and Census tract. Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I.

Figure A25: Racial Gaps in Levered Housing Returns, Imputed Race/Ethnicity



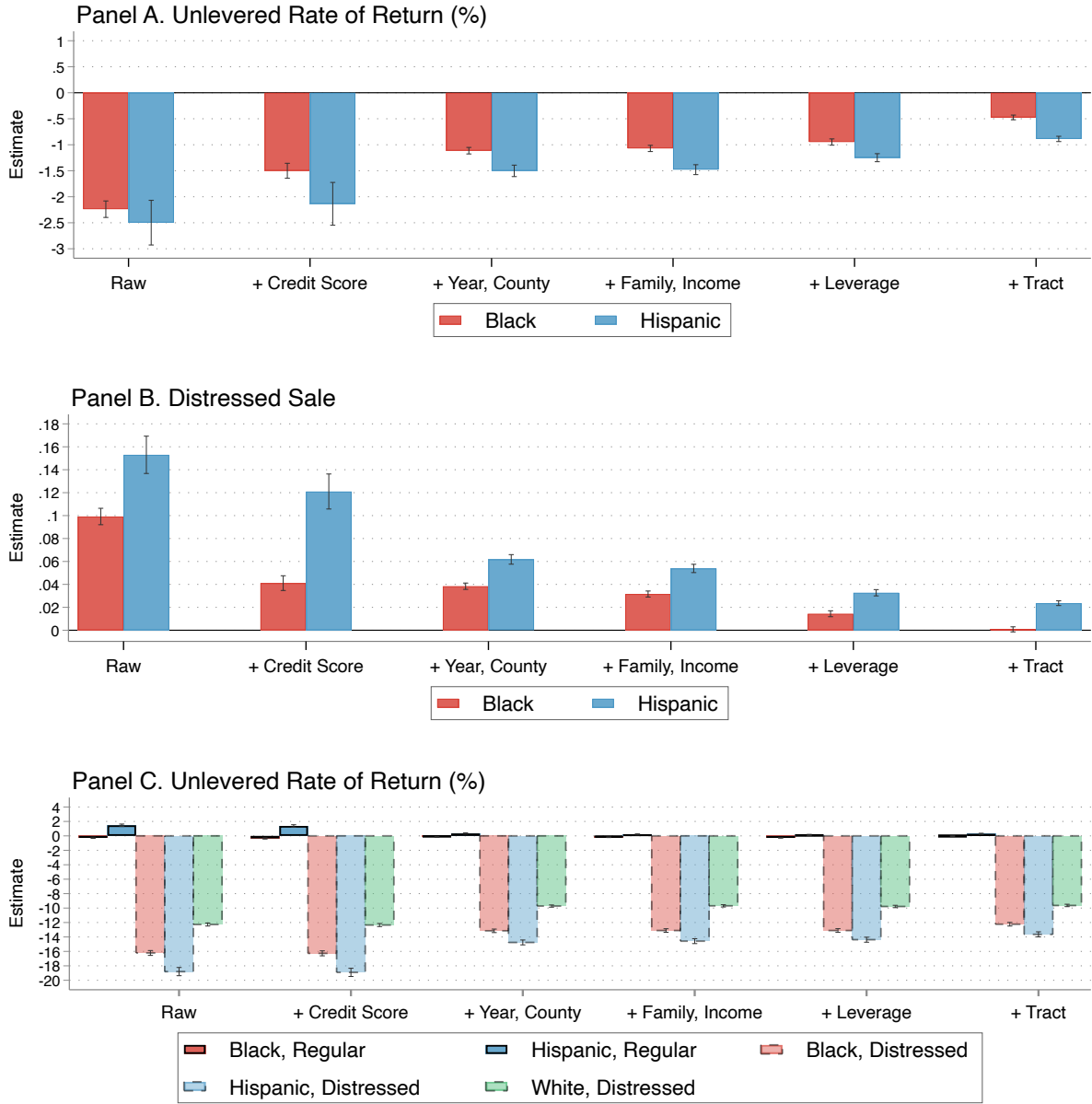
Notes: This figure presents estimates of racial gaps in levered housing returns (Equation 3). Race/ethnicity has been imputed using homeowner name and location. In Panel A, each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. In Panel B, each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Year, County* adds in purchase year-by-county fixed effects. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year and Census tract. Standard errors are clustered within purchase year, sale year, and county cells. Data are from baseline sample of ownership spells described in Section I.

Figure A26: Racial Disparities in Non-Mortgage Default



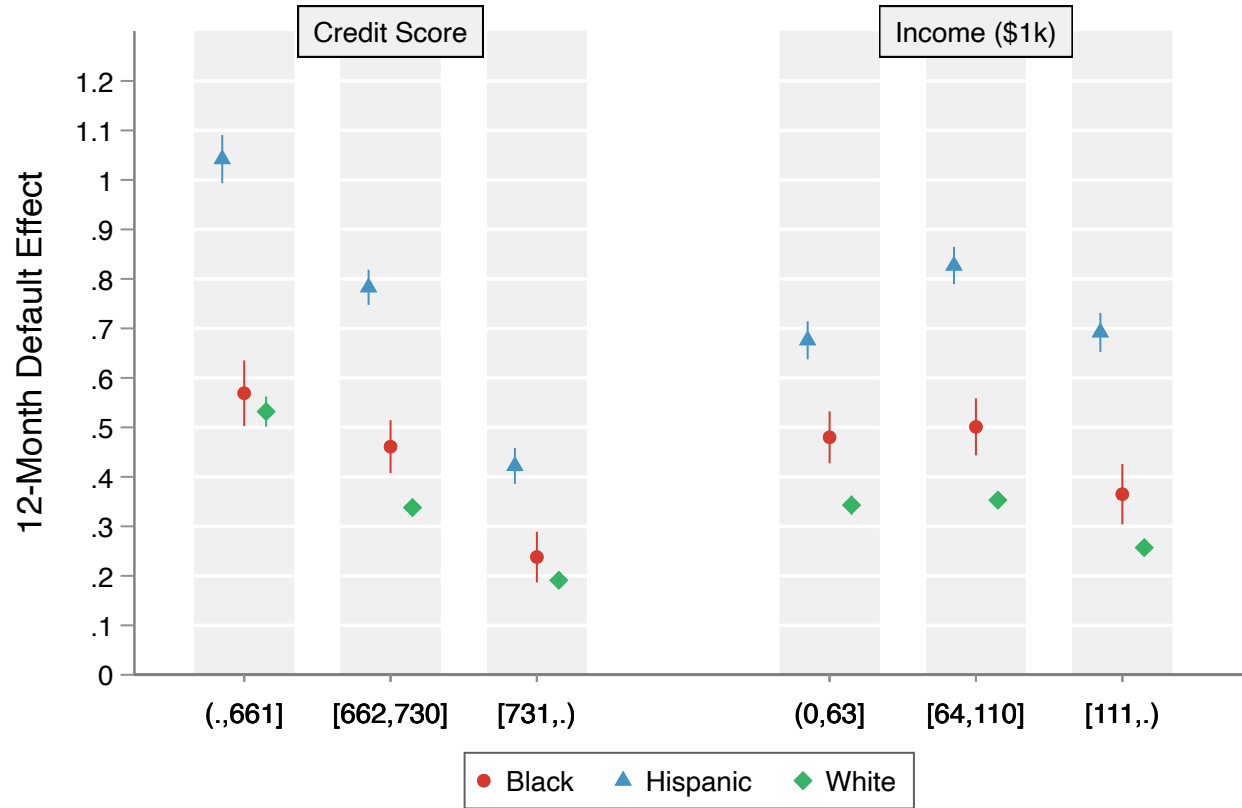
Notes: This figure presents estimates of racial differences in financial distress controlling for a range of observable homeowner characteristics (Equation 4). These estimates decompose racial differences in 90-day non-mortgage default into components that correspond to household, loan, and location characteristics. Each bar corresponds to the coefficient on a race/ethnicity indicator. Each pair of bars correspond to a separate regression with a particular set of covariates. *Raw* denotes a regression of the outcome on race/ethnicity indicators and current year fixed effects. *Orig. Year, County* adds fixed effects that interact current year, origination year, and county. *Family, Income* adds fixed effects that interact gender, the presence of a co-applicant on the mortgage, and current year, and fixed effects that interact deciles of income and current year. *Leverage* adds fixed effects that interact 1 percentage point bins of original loan-to-value and current year. *Credit Score* adds fixed effects that interact 10 point credit score bins and current year. *Mortgage Chars.* controls for a spline in current loan-to-value, log monthly payment, log interest rate, an indicators for interest-only loan, refinance, adjustable rate mortgage, and GSE loan, and adds fixed effects for mortgage term, deciles of original loan value, and deciles of debt-to-income. *Tract* adds Census tract fixed effects. *Lender, Block* adds Census block fixed effects. Standard errors are clustered at the loan level. Data are from a panel of homeowners with linked credit bureau and mortgage servicing records described in Section III.

Figure A27: Racial Gap in Housing Returns, Merged Credit Score Sample



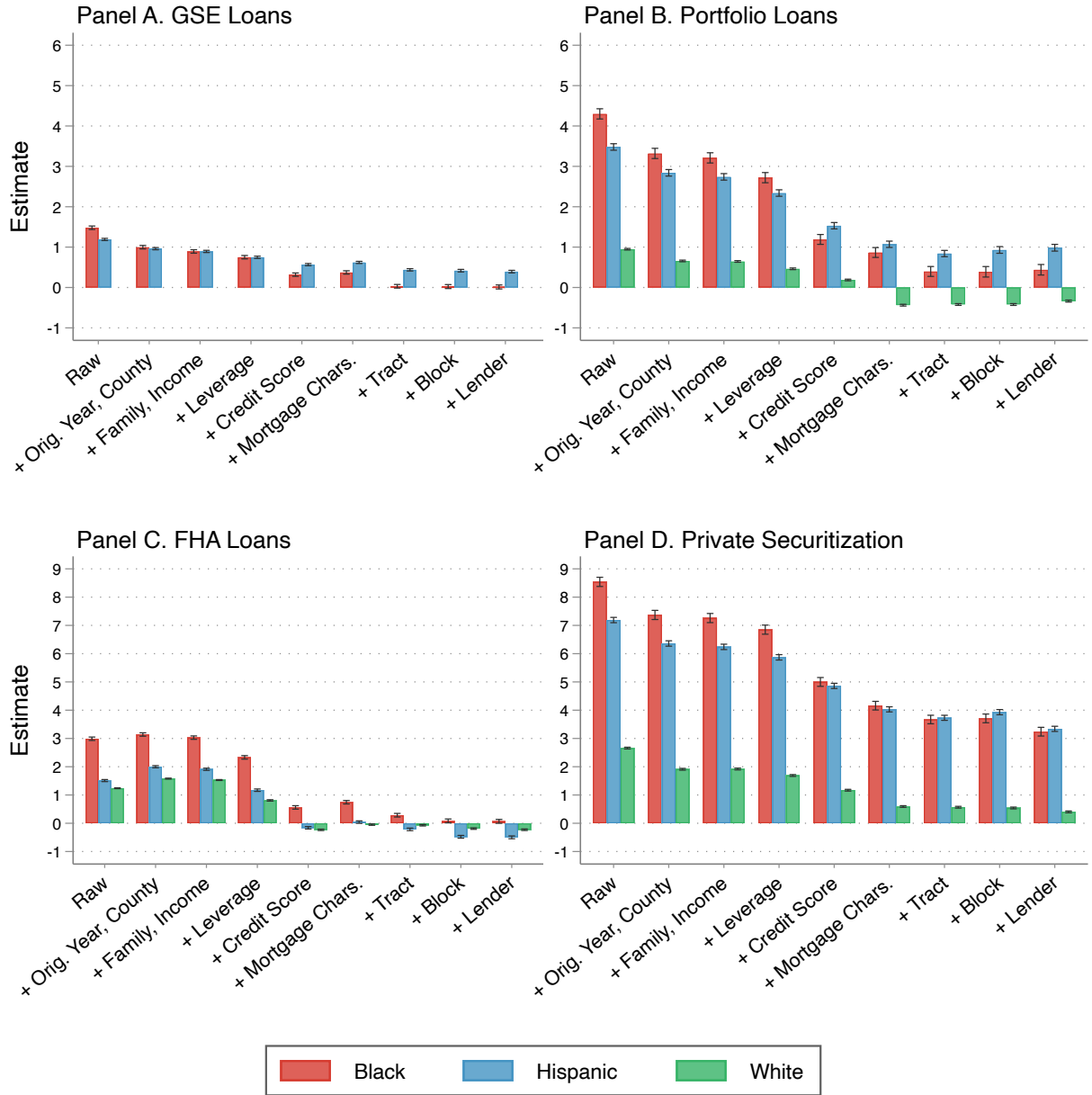
Notes: This figure presents estimates of racial gaps in housing returns and distressed sales. In Panels A and B, each pair of bars corresponds to a separate regression and indicates estimated coefficients of Black and Hispanic indicators from Equation 4 using a particular set of fixed effects. In Panel C, each set of bars corresponds to coefficients of race indicators interacted with an indicator that a property is sold in a distressed sale. *Raw* denotes estimates without controls. *+ Credit Score* adds ten point bins of credit score at home purchase. *+ Year, County* adds in purchase year-by-county fixed effects. *+ Family, Income* adds fixed effects that interact county, purchase year, gender, and an indicator for the presence of a mortgage co-applicant, and fixed effects that interact county, purchase year, and deciles of income at home purchase. *Leverage* adds fixed effects that interact county, purchase year, and 1 percentage point bins of combined-loan-to-value at purchase. *Tract* adds fixed effects that interact purchase year and Census tract. The outcome in Panels A and C is the annualized unlevered rate of return (Equation 1), and the outcome in Panel B is an indicator that a homeowner experiences a distressed sale. Data are from sample of ownership spells described in Section I, restricted to spells that are merged with mortgage servicer data. Standard errors are clustered within purchase year, sale year, and county cells.

Figure A28: Heterogeneous Effects of Payment Shock on 90-Day Default



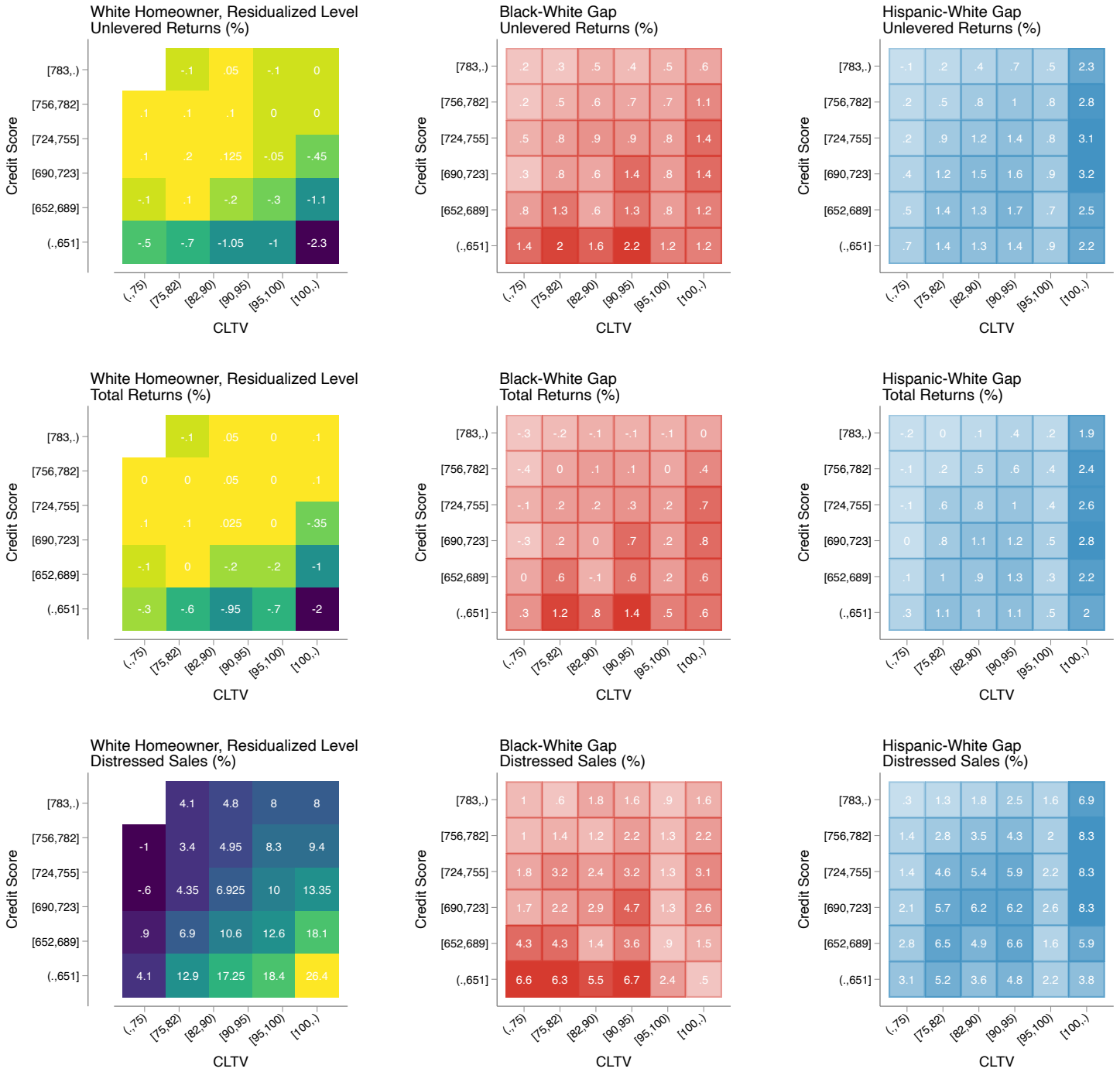
Notes: This figure documents heterogeneous impacts of monthly payment shocks on monthly 90-day mortgage default rates following the reset of an adjustable rate mortgage. Default rates are measured as an indicator that the homeowner's primary mortgage is 90 or more days past due. Coefficients correspond to effects after 12 months and are scaled to reflect percentage point impacts of a 1 percent increase in monthly payments. *Credit Score* splits the sample by terciles of credit score at origination. *Income (\$1k)* splits the sample by terciles of annual income at origination. Standard errors are clustered at loan level. Data come from panel of homeowners with linked credit bureau and mortgage servicing records with adjustable rate mortgages who experience a rate reset. See Section III for more details.

Figure A29: Racial Differences in Mortgage Default Interacted with Mortgage Investor Channel



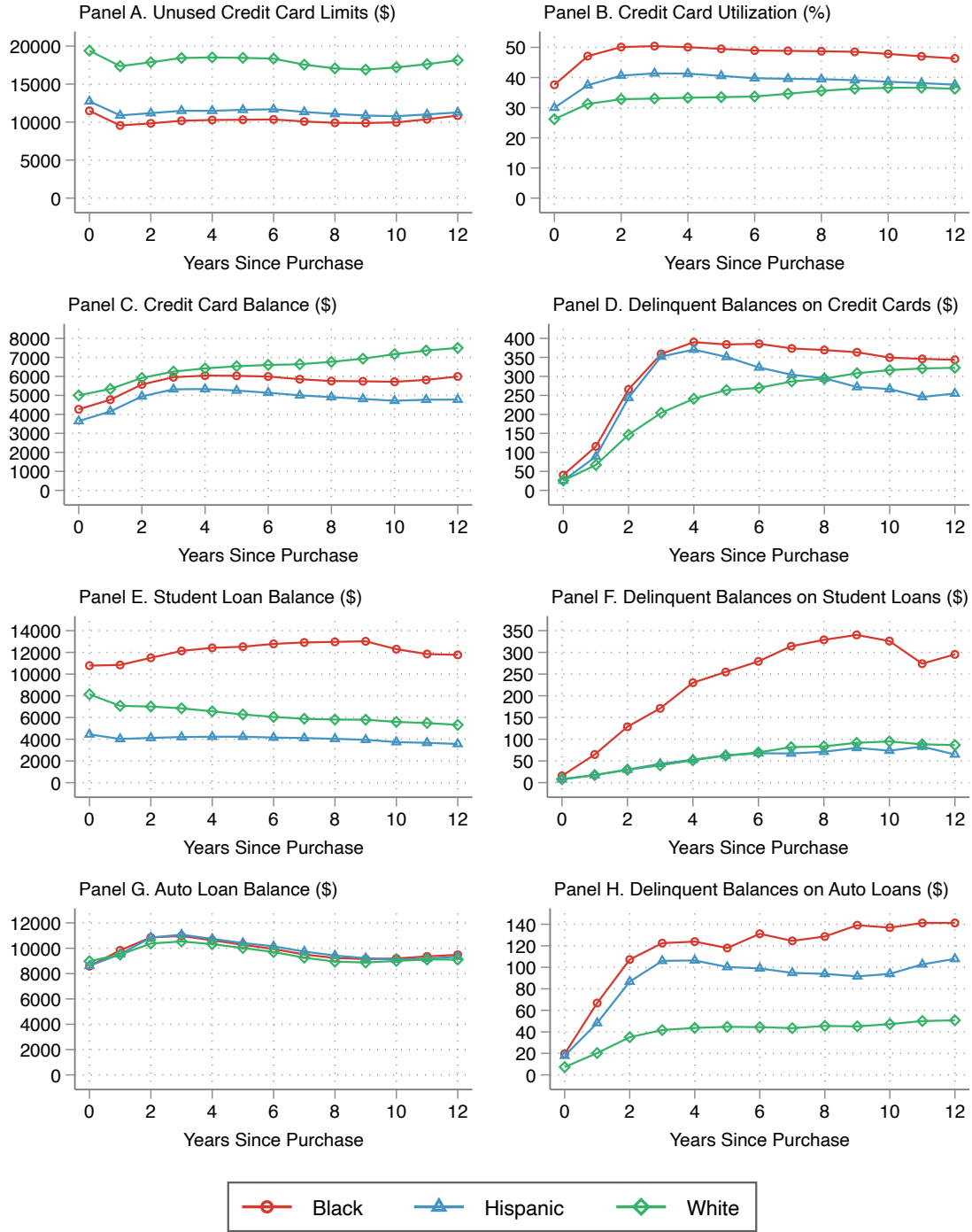
Notes: This figure presents estimates of racial differences in financial distress across mortgage investor channel, controlling for a range of observable homeowner characteristics. Estimates are derived from nine separate regressions, with eleven coefficients for each regression split across four panels, such that each panel presents coefficients for race indicators interacted with a given investor channel. Panel A corresponds to mortgages purchased by a government-sponsored enterprise (GSE); Panel B to mortgages held on a lender portfolio; Panel C to FHA mortgages; Panel D to privately securitized mortgages. The omitted category is mortgages held by White homeowners that were purchased by a GSE. *Raw* denotes a regression of the outcome on race/ethnicity indicators and current year fixed effects. *Orig. Year, County* adds fixed effects that interact current year, origination year, and county. *Family, Income* adds fixed effects that interact gender, the presence of a co-applicant on the mortgage, and current year, and fixed effects that interact deciles of income and current year. *Leverage* adds fixed effects that interact 1 percentage point bins of original loan-to-value and current year. *Credit Score* adds fixed effects that interact 10 point credit score bins and current year. *Mortgage Chars.* controls for a spline in current loan-to-value, log monthly payment, log interest rate, an indicators for interest-only loan, refinance, and adjustable rate mortgage, and adds fixed effects for mortgage term, deciles of original loan value, and deciles of debt-to-income. *Tract* adds Census tract fixed effects. *Block* adds Census block fixed effects. *Lender* adds lender fixed effects. Standard errors are clustered at the loan level. Data are from a panel of homeowners with linked credit bureau and mortgage servicing records described in Section III.

Figure A30: Heat Map by Credit Score and Leverage



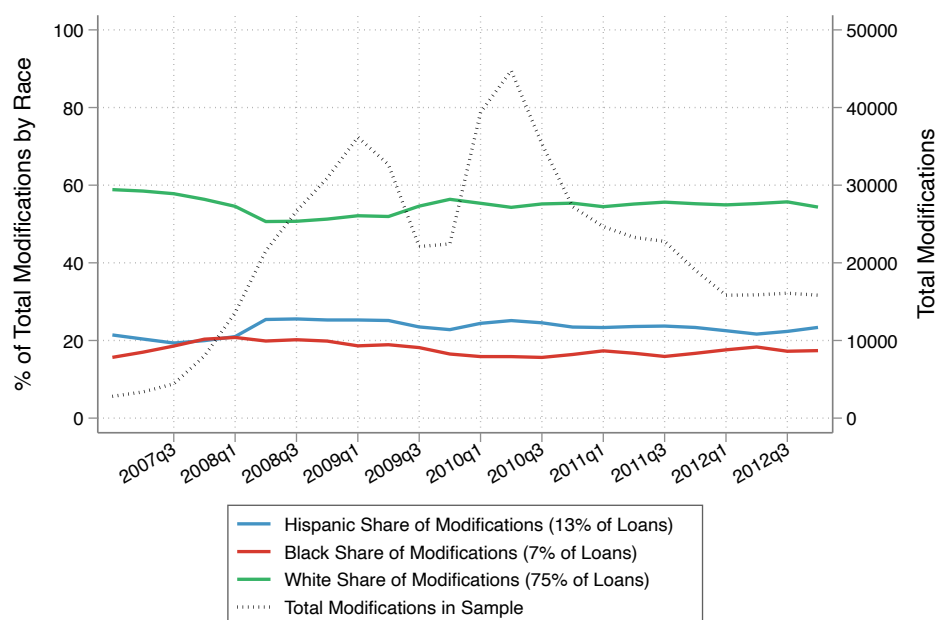
Notes: This figure presents estimates of racial differences in housing returns and distressed sales among homeowners by credit score and initial leverage. Credit score is measured at home purchase, and leverage is defined as combined loan-to-value (CLTV) at purchase. The first row presents results for annual unlevered returns. The second row presents results for annual total returns. The third row presents results for distressed sales. The first column presents average outcomes for White homeowners, residualizing by county and purchase year and normalizing to the average value for the highest-score and lowest-CLTV cell. The second and third columns present Black-White and Hispanic-White differences, respectively. Within each panel, the color of a square indicates the size of the estimated coefficient, with each square corresponding to a coefficient for race/ethnicity indicators estimated from a separate regression within each score-CLTV cell. Regressions in Columns 2 and 3 are estimated as in Equation 4 with fixed effects that interact county and purchase year. Data are from baseline sample of ownership spells described in Section I, restricted to properties merged to the McDash servicing records.

Figure A31: Racial Disparities in Financial Distress by Tenure Length



Notes: This figure presents financial outcomes by race/ethnicity as a function of the number of months since home purchase. The financial outcomes are dollar amount of unused credit card limits (Panel A), credit card utilization in percent (Panel B), credit card balances (Panel C), balances 30 or more days past due on credit cards (Panel D), student loan balances (Panel E), student loan balances 30 or more days past due (Panel F), auto loan balances (Panel G), auto loan balances 30 or more days past due (Panel H). Credit card utilization is conditional on having an open credit card. Homeowners without credit cards are coded as having \$0 in unused limits. Data are from a panel of homeowners with linked credit bureau and mortgage servicing records described in Section III.

Figure A32: Modifications During the Great Recession



Notes: This figure plots the aggregate quarterly time series of modification rates before, during, and after the Great Recession. The dashed line plots the total number of modifications observed in the sample. The solid lines plot the percent of total modifications received by race/ethnicity. These patterns indicate that Black and Hispanic homeowners received a disproportionately large and approximately constant share of modifications throughout this time period, despite only accounting for 7% and 13% of owner-occupied mortgages, respectively. The data come from a sample of 6.4 million first-lien owner-occupied mortgages originated in or before 2008, contained in the Fannie Mae, Freddie Mac, and ABSNet Loan databases in which mortgage modifications can be observed.

C Data Appendix

This section provides additional details on the various data sources we analyze, as well as the algorithm that identifies consecutive sales of properties used to construct the baseline sample of ownership spells. All administrative data linkages were conducted by the Fisher Center for Real Estate and Urban Economics at UC Berkeley.

C.1 Property Records

Our baseline sample is constructed from a linkage between the HMDA mortgage origination records and the ATTOM property transaction records (ATTOM Data Solutions 2020; FFIEC 2016). The ATTOM property data, which link mortgages to properties, are merged with the HMDA records by matching on transaction year, Census tract, dollar amount, and lender name, and allowing December transactions to match with January transactions. While the ATTOM data cover purchases made through December 2020, the merge between ATTOM and HMDA only covers ownership spells starting in 2016 and earlier. This means that while our baseline sample captures ownership spells in which the property was purchased in, for instance, December 2016 and sold in March 2020, it does not capture ownership spells in which the property was originally purchased in January 2017 and sold at a later date.

We rely on a proprietary algorithm from ATTOM to identify short sales, which appears to closely track external survey-based measures of short sales. Appendix Figure C1, Panel A shows that we can closely replicate this algorithm by defining short sales as those that are likely to have yielded proceeds below the outstanding balance of the mortgage. Panel B plots this percentage over time and shows that the ATTOM categorization closely tracks short sales measured using a monthly survey of real estate agents, reported in Campbell Communications (2011). These surveys are widely referenced by industry professionals (e.g., Mahon 2010). Since short sales by definition take place at prices below the outstanding principal balance, the patterns in Appendix Figure C1 suggest that the algorithm accurately identifies short sales. In addition, Zhang (2019) uses a sample of property transactions provided by DataQuick and classifies 36% of distressed sales as short sales, essentially identical to the 36% classified as short sales in our data. Ferreira and Gyourko (2015) take an alternative approach, categorizing short sales as those with sales proceeds below 90% of the unpaid principal balance. Replicating our analysis following this approach yields very similar results. Since our baseline sample restricts to arm’s length transactions, it excludes deeds in lieu of foreclosure, which represent only about 0.09% of transfers in the property data.

Our baseline approach assigns race and ethnicity using the self-reported race and ethnicity of the borrower in the HMDA records. While this offers a nearly ideal measure of homeowner race/ethnicity, this information is missing for about 10% of our main analysis sample, as reported in Appendix Table C1. To probe the extent of potential bias created by this missing information, we impute race/ethnicity using the methods developed by Imai and Khanna (2016) to impute homeowner race/ethnicity from the last name and Census block of the homeowner in the property records, although it classifies a relatively high share of self-reported Black homeowners as White. Appendix Table A4 shows that race/ethnicity imputed from homeowner name aligns closely with self-reported race/ethnicity. Reassuringly, racial/ethnic composition of homeowners with observed self-reported race/ethnicity appears to be similar to that with missing self-reported race/ethnicity, as shown in Appendix Table C1.

To compare our baseline estimates of racial differences in housing returns to those estimated using homeowner name, we construct a sample of ownership spells in the property data that does not restrict to properties merged with the HMDA, but retains the other sample restrictions. Specifically,

Table C1: Cross-Tabulation of Self-Reported and Imputed Race

	Imputed Race						N
	Hispanic	Black	White	Asian	Other	Missing	
Self-Reported Race							
Hispanic	80.4	1.9	15.0	1.5	0.1	1.1	1,900,047
Black	3.0	54.5	39.7	1.3	0.3	1.3	1,066,134
White	2.8	2.8	91.8	1.3	0.1	1.2	10,888,521
Asian	7.0	2.4	24.5	63.4	1.3	1.5	1,018,691
Other	9.3	8.5	66.1	13.6	1.4	1.1	145,599
Any Non-Missing	13.0	6.4	73.5	5.7	0.2	1.2	15,018,992
Missing	10.3	7.5	74.4	6.3	0.3	1.3	1,501,187

Notes: This table compares the sample distribution by self-reported and imputed race. Each number corresponds to the percent of homeowners assigned to a given race/ethnicity based on their name and Census Block. Percentages are reported separately by homeowner self-reported race/ethnicity measured in the HMDA data.

we drop ownership spells lasting less than 12 months, transactions of less than \$10,000, buyers that are neither persons nor trusts, and purchases with combined loan-to-value of more than 102.5%.

Estimates of racial differences in housing returns and distress using homeowner name are presented in Appendix Table A2. Comparing the average returns among homeowners with missing vs. non-missing self-reported race/ethnicity indicates that the distribution of returns is similar in both samples. This is also the case when computing the racial gap in housing returns by imputing race/ethnicity from names, within the subsample of homeowners with missing self-reported race/ethnicity. Consequently, estimates of Equation 4 which impute race/ethnicity from names when self-reported estimates are missing yield very similar to the baseline estimates.

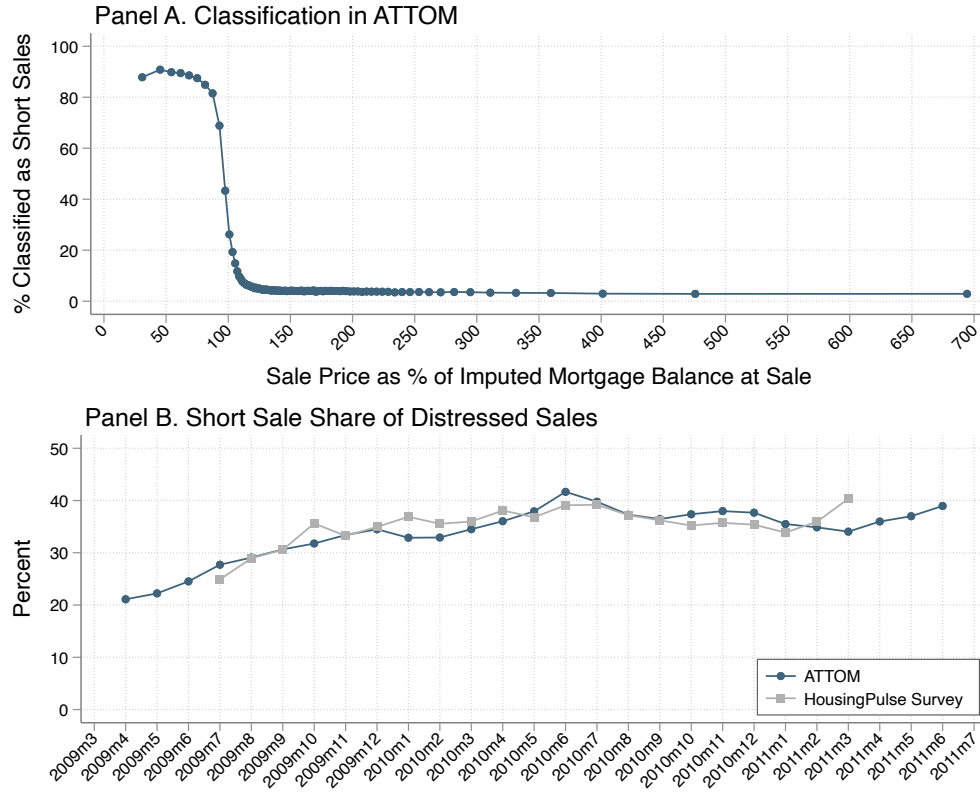
Although our primary focus is on comparing Black and Hispanic homeowners to White homeowners, differences among other racial/ethnic groups are also of scientific interest. We present estimates of returns and rates of distressed sales for these groups in Appendix Table A12 and Table A13. Average unlevered returns for Asian homeowners are similar for White homeowners (Appendix Table A12, Column 1); however, when controlling for county, purchase year, and demographics, returns are markedly lower (Column 2). Total and levered returns are also lower for Asian homeowners (Columns 3 through 6).¹ Distressed sales appear to play an important role in explaining these differences: for example, average unlevered returns among non-distressed sales are higher for Asian homeowners (Appendix Table A13, Column 1). However, controlling for timing, demographics, and neighborhood, returns among Asian homeowners not realizing a distressed sale are lower than those of White homeowners (Column 2). These differences indicate ample scope for future research to uncover mechanisms determining realized housing returns among Asian homeowners.

C.2 Comparing Sample to External Benchmarks

Our primary results rely on a merge between the HMDA records and property transaction records. In light of the fact that this sample of 13.6 million ownership spells represents a relatively small

¹Note that in the absence of external estimates of differences in property tax payments for Asian homeowners, our estimates of total and levered returns for racial groups other than White, Black, and Hispanic do not incorporate racial differences in housing costs (taxes, interest rates, and maintenance). Since these adjustments have a relatively modest impact on racial gaps for Black and Hispanic homeowners (Appendix Table A1), it is likely that analogous adjustments would also have modest impacts on estimates for other racial groups.

Figure C1: Short Sale Classification



Notes: This figure illustrates the algorithm used to identify short sales and compares the results of that algorithm to an external measure of short sales. Panel A plots the percent of property sales that are classified by ATTOM as short sales, as a function of the sale price as a percentage of the imputed mortgage balance at sale. Sample excludes sales classified as foreclosures based on the sale documentation. Data are from a sample of homeowners with observed purchase and sale prices in baseline sample of ownership spells described in Section I. Panel B plots the percentage of distressed sales that are classified as short sales by ATTOM in the baseline sample, along with the percentage classified as short sales in the HousingPulse Survey as reported in Campbell Communications (2011). See Appendix D for discussion of imputation of mortgage balance at sale.

share of the total number of purchases that occurred during this time period, we assess the representativeness of our sample by comparing it to an external representative sample from the 2015 American Housing Survey (AHS).

Reassuringly, our baseline sample is similar to the AHS sample along many dimensions. To ensure comparability of our spell-level sample to the AHS cross-section, we compare the subset of properties in our sample that were purchased before 2015 and remained unsold as of 2015, to homeowners in the AHS who were surveyed in 2015 and who purchased their homes with a mortgage. Section C.5 provides more details of the AHS sample. Appendix Table A14 shows that share of Black and Hispanic homeowners are similar in both samples, as is the share of properties that are single-family homes. The distribution of down payments also appears similar to the sample in the AHS, particularly after restricting to homeowners in the AHS who have moved since 2000, which more closely reflects the temporal coverage of our sample.

Given the way our sample was constructed, we expect certain modest differences in geographic coverage and homeowner characteristics. In terms of geographic coverage, the property data have

very little coverage in states that do not publicly disclose real estate purchase prices (e.g., Texas). This is most clearly illustrated in a comparison of sample coverage across CBSAs (Appendix Table A15). Nonetheless, the distribution of our sample across large geographic areas (i.e., Census Division) is similar to the AHS, with modest differences largely due to limited coverage of non-disclosure states. However, the property data, even without the merge with HMDA, have less coverage in rural areas: the non-metropolitan share of the sample is 5% in the property data (3% after the merge to HMDA), compared to 13% in the AHS. Likely as a result, the average property in our baseline sample is valued at about 10% more than the national average in the AHS.

Another expected difference relative to the external sample is the tenure length of homeowners. Since our sample begins in 1990, and the property data have limited coverage before 2000, the distribution of tenure length in our sample is lower than the national average. In line with this difference in coverage, our sample appears very similar to the AHS after restricting to homeowners in the AHS that moved into their current residence in 2000 or later (Appendix Table A14, Column 4).

Overall, comparing our sample to the AHS confirms the dimensions on which we expect our sample to differ from a representative cross-section. We therefore view our results as being externally valid for homeowners who purchased their homes within the temporal window which is covered by our sample.

Given the importance of distressed sales to our analysis, it is also valuable to confirm that the rates of distressed sales in our sample are similar to external estimates. Appendix Figure A6 depicts the aggregate foreclosure rate over time in our data, which is similar to the trends reported in Corbae and Quintin (2015) using data from the Mortgage Bankers Association National Delinquency Survey.²

C.3 Algorithm for Identifying Repeat Sales of Properties

We develop an algorithm for identifying repeat sales of properties, allowing us to calculate realized housing returns for each ownership spell. This algorithm distinguishes transaction records that represent mortgages used to purchase a property, transfers of ownership, and mortgage refinances. We identify property purchases by restricting to arm’s length, full-consideration transactions that are recorded as home purchases in HMDA. To identify the future sale of a given purchased property among the set of all future transactions of that property, we drop transactions in which the new buyer’s name is similar to the original buyer name, and select the first subsequent arm’s length full-consideration transaction. A natural language processing algorithm classifies names as individuals, trusts, and non-trust institutions (e.g., banks, governments). We restrict to purchases in which the buyer is a person or trust, and to sales in which the seller in the second transaction is the same as the buyer in the first transaction, excluding distressed sales from the requirement because distressed sales are typically executed by institutions rather than individuals. We drop transactions with prices that are less than \$10,000. For the purposes of computing housing returns, we drop purchases with combined loan-to-value ratios of more than 102.5% and ownership spells that last less than 12 months. This algorithm yields our baseline sample of ownership spells, consisting of 13.6 million ownership spells purchased by Black, Hispanic, or White owner-occupants between 1990 and 2014. Note that this sample includes spells where the property remained unsold by the end of March 2020.

²Our quarterly foreclosure rates are lower than those in Corbae and Quintin (2015) due to differences in the definition of foreclosure. The National Delinquency Survey measures foreclosure starts, whereas our ownership spell data measures foreclosure sales. Since not all foreclosure starts end in a foreclosure sale, aggregate counts of foreclosure starts are naturally higher (ATTOM, 2023).

Ownership spells in our baseline sample fall into one of two categories: those where the purchaser retains ownership through March 2020 and those where the owner sells or loses ownership of the property. For the former category, we impute sales using house price indices. For the latter category, we require the spell to end in an arm’s length, full consideration transaction. Two consequences of this requirement are worth highlighting. First, this requirement excludes many transactions that are categorized as deeds in lieu of foreclosure. Specifically, 90% of transactions that are labeled deed-in-lieu are not labeled as arm’s length, full consideration transactions. These are therefore dropped from our main analysis sample. Transactions labeled deed-in-lieu comprise only about 0.09% of all residential transactions in the property data. Second, the restriction to arm’s length transactions drops transfers that occur within-family.

C.4 Linkages to Additional Administrative Data Sources

The HMDA and ATTOM datasets lack a number of variables of interest, such as measures of underlying financial well-being, characteristics of the mortgage contract, and loan modifications. In order to observe these variables, we turn to linkages with the Equifax, McDash, Fannie Mae, Freddie Mac, and ABSNet datasets (Equifax 2022; Intercontinental Exchange, Inc. 2022; Lewtan 2016; Freddie Mac 2020; Fannie Mae 2020). These additional datasets allow us to observe important information on our study sample. However, the linkages between the ATTOM property data and these datasets are different than between the ATTOM and HMDA. In particular, they likely entail non-negligible amounts of measurement error generated by imprecision in the linkages. In the remainder of this section, we provide more details on these linkages and the strategies used to minimize measurement error.

The Equifax and McDash data allow us to observe a wide range of financial outcomes and behaviors. These data are known as the Equifax Credit Risk Insight Servicing McDash Database (CRISM). The CRISM data contain two components: mortgage servicing records from McDash and credit bureau records from Equifax. The McDash data contain information on both mortgage characteristics measured at origination and loan performance. The Equifax data are comprised of information from the Equifax credit bureau records of borrowers of mortgages captured in the McDash data. The Equifax data are at the monthly level and capture a broad range of financial outcomes and behaviors, including balances and delinquencies on credit cards, auto loans, and mortgages, as well as accounts in collections.

We rely on a k-nearest neighbors algorithm developed by the Fisher Center to link the ATTOM property records with the CRISM, Fannie Mae, Freddie Mac, and ABSNet datasets. To create the linkage to the CRISM dataset, the algorithm proceeds as follows. Within each US county, the algorithm creates a stable linkage between transactions in ATTOM and loans in McDash, matching records (“neighbors”) along a vector of attributes. These attributes include the loan amount, the value of the property, the origination date, the purpose of the loan (e.g., purchase or refinance), whether the loan ended in distress (e.g., foreclosure), the loan lien type, the interest rate, and the date the loan was paid off. The data merge we use is similar to merges between the CRISM and HMDA datasets used in the previous work (e.g., Gerardi et al. 2020).

We apply a number of sample restrictions to mitigate the potential impact of measurement error. First, we restrict to matches for which the algorithm chose the nearest neighbor (e.g., as opposed to the second-nearest neighbor). Second, we restrict to matches for which there are no other close matches. Our measure of match closeness comes from a score generated by the algorithm that denotes the closeness of the match along the vector of matched attributes. For the merge with the CRISM data, the 10th, 25th, 50th, and 75th percentiles of this score are 270, 639, 919, and 1,850, respectively (lower scores denote closer matches). We drop any match with a neighbor with

a score within 200 of the score of the chosen match. Third, we restrict our analysis to matches with scores at or below 2,000 (slightly above the 75th percentile of the score).

We use the Fannie Mae, Freddie Mac, and ABSNet datasets allow us to study mortgage modifications. Fannie Mae and Freddie Mac publish publicly-available mortgage databases containing subsets of purchased or guaranteed single-family conventional fixed rate mortgages originated since 2000 and 1999, respectively. To complement these databases, we include loans in the ABSNet Loan database. The ABSNet data are sourced from reports to securitization trustees and cover over 90% of loans collateralized through private-label residential mortgage backed securities. In addition to observing loan modifications, these data sources also record the identity of the mortgage servicer. To focus on a sample of homeowners who are eligible to receive a mortgage modification, we construct a sample of loans that become 90 or more days past due.

As with the credit bureau and mortgage servicer records, we rely on a k-nearest neighbors algorithm developed by the Fisher Center to link mortgage modifications in the Fannie Mae, Freddie Mac, and ABSNet data to our main study sample. This algorithm is the same as that used for the CRISM datasets and uses a similar vector of matched attributes (e.g., loan amount, property value, origination date). We apply analogous restrictions to minimize measurement error, restricting to nearest-neighbor matches with no other close match and low distance scores. For the merge with the Fannie Mae, Freddie Mac, and ABSNet data, the 10th, 25th, 50th, and 75th percentiles of this score are 435, 499, 824, and 1,929, respectively.

The merge score provides a convenient way to test whether measurement error significantly impacts our results. In unreported results, we replicate our analyses that use the CRISM, Fannie Mae, Freddie Mac, and ABSNet datasets using more stringent restrictions on merge quality, restricting to matches with a score of 700 or less (approximately the median score), which leaves our findings largely unchanged but entails some loss in precision due to the use of a smaller sample. The robustness to these sample choices indicates that measurement error is not likely to significantly affect the conclusions we draw from the analysis using the CRISM, Fannie Mae, Freddie Mac, and ABSNet data sources.

C.5 External Data Sources

This section provides additional details on the construction of the external data sources from the Survey of Income and Program Participation (SIPP), the Current Population Survey (CPS), and the American Housing Survey (AHS).

For the SIPP sample (United States Bureau of the Census, 2025b), we draw on waves between 1991 and 2018. All results use person weights, and dollar-denominated values are deflated to 2016 dollars. In Table 3, we restrict the sample to homeowners with positive liquid wealth. Our final sample includes 1.15 million observations corresponding to 427 thousand households, with households defined as the unique combination of SIPP sample units/address IDs, and family IDs. To construct a liquid wealth variable, we follow Chetty et al. (2017) and define liquid wealth as the sum of assets held in stocks, bonds, checking accounts, and savings accounts, excluding retirement, accounting for changes to variable construction and variable names across panels in some years. We winsorize this variable at the 1% level. Our annual unemployment variable is an indicator that measures whether in the prior 12 months, any earner in the household had no job in a month, was on layoff, or was looking for work in all weeks. Income is the annualized monthly income for the household, and delinquency corresponds to the question, “Was there any time in the past 12 months when (you/your household) did not pay the full amount of the rent or mortgage?”

For the CPS sample (Flood et al., 2022), we construct a dataset at the individual level using interviews between 2000 and 2023. We restrict to Black, White, and Hispanic individuals aged 20 to

55 and who were employed working at least 30 hours per week as of their first interview. We restrict to individuals who report being homeowners two months after their first interview. Unemployment is measured 12 months after first interview, whereas income loss is measured by comparing income measured 2 and 14 months after the first interview. Applying these restrictions narrows the sample window to 2002 to 2018. For about two-thirds of the sample, we compare interview months 0 to 12, and for the remaining one-third we compare interview months 2 to 14. All results use Annual Social and Economic Supplement Weights. Our combined sample includes 39,129 individuals.

For the AHS sample (United States Bureau of the Census, 2025a), we use the 2015 survey wave to construct a sample that can be compared to the primary analysis sample. We restrict to owner-occupied properties surveyed in 2015, occupied by homeowners who have lived in their properties for at least one year. To restrict to properties purchased with a mortgage, we keep properties where the homeowner reported making a down payment less than 100% of the purchase price, or who have a current mortgage open. We also drop mobile homes and vacant properties. The resulting sample includes 24,869 households. Continuous outcomes are winsorized at the 1% level. All results use included AHS sampling weights.

D Constructing Total and Levered Returns

This section describes the calculations used to estimate the total and levered rate of return (Equations 2 and 3). The discounted sum of cash flows can be written as follows:

$$-DownPay_{i0} + \sum_{t=1}^{T_i-1} \frac{rent_{it} - pymt_{it}}{(1 + r_i)^t} + \frac{\max\{0.01, rent_{iT} - pymt_{iT} + 0.95P_{iT} - UPB_{iT}\}}{(1 + r_i)^{T_i}}$$

Recall that the discounted cash flows implicitly define the levered rate of return. Moreover, the total rate of return can be computed as the levered rate of return with zero borrowing.

We draw on a variety of data sources to compute each component of this equation. $DownPay_{i0}$ denotes the homeowner’s down payment and is the sum of initial equity and closing costs. Equity is measured directly as the difference between purchase price and loan values in the ATTOM dataset. To estimate closing costs, we use the 2018 to 2019 HMDA data, which contain information about closing costs paid for originated mortgages, including points. We restrict to owner-occupied purchase mortgages. For primary (secondary) mortgages, we restrict to loans with LTV less than or equal to 102.5% (35%). We compute total closing costs as total loan costs minus lender credits. For primary mortgages, we regress closing costs as a share of the loan amount on indicators corresponding to five LTV bins: (0,82%), [82%,90%), [90,95%), [95%,100%), and [100%, 102.5%]. These categories allow our calculations to reflect higher closing costs for higher-LTV mortgages (e.g., due to higher mortgage insurance costs). We also include log loan amount and the interaction of the LTV bins with LTV. For secondary mortgages, we estimate analogous regressions but using two LTV bins, corresponding to a cutoff of 20%. We then impute closing costs for each transaction in ATTOM using the coefficients from these regressions. For the calculation of total returns, we assume closing costs that are half of the imputed value for a loan with an LTV of 80%, since a large share of closing costs are associated with borrowing. When we use these imputations to compute statistics weighted by total upfront costs, we winsorize total upfront costs at the 99.9% level.

Table D1: Summary Statistics and Validation for Imputed Values

<i>Panel A. Summary Statistics</i>	Mean	SD	p10	p90
Principal and Interest Payment	1315	993.9	498	2378
Escrow Payment	386	752.6	174	647
Primary Balance at Sale	183815	141198.4	67880	330836
Interest Rate, First Lien	0.057	0.0116	0.040	0.070
Interest Rate, Second Lien	0.084	0.0125	0.071	0.096
Monthly Rent	1490	6435.6	626	2513
Rent-to-Price Ratio	0.070	0.0156	0.050	0.091
Annual Maintenance (% of Home Value)	0.817	0.1097	0.676	0.955
Closing Costs	4881	1846.3	2823	7203
<i>Panel B. Validation with McDash</i>	β	Constant	SE	R2
Primary Principal and Interest Payment	1.0338	84.799	0.0005	0.7404
Escrow Payment	0.9645	19.398	0.0007	0.5983
Interest Rate, First Lien	1.0090	-0.001	0.0002	0.7625
Primary Balance at Sale	1.0198	-3010.133	0.0002	0.9740
Combined Loan-to-Value	0.8586	10.701	0.0003	0.7284
Primary Loan-to-Value	0.9171	5.525	0.0002	0.8184

Notes: This table presents summary statistics (Panel A) and validation (Panel B) of the imputed values used to construct the internal rate of return (Equation 3). Panel A plots means, standard deviations, and 10th and 90th percentiles. Panel B plots regression coefficients derived from regressing the value observed in the merged McDash data on its imputed value, trimming variables at the 1% level. Note that in the ATTOM data, loan-to-value and combined loan-to-value are measured from recorder documents and thus not imputed using external sources. SE refers to the standard error of the slope coefficient. Note that all variables have been calculated without race-specific adjustments. See Appendix Section D for more details.

When imputing monthly net rents $rent_{it}$ (i.e., gross rent minus property taxes, insurance, and maintenance costs), we must calculate rents for properties that are actually owner-occupied. A well-known measurement challenge is that the stock of rental properties differs from that of owner-occupied properties, such that simply comparing local rents to local home values yields biased estimates of rent-to-price ratios (Demers and Eisfeldt, 2022).

To impute gross rents, we construct a debiased measure of local rent-to-price ratios (henceforth, R:P). We begin by using US Census data to construct a yearly series of county-level rents and home values for each year 1996-2020. Data from the Decennial Census and the American Community Surveys yield median rents and home values for 2000 and 2010-2020. We interpolate house prices using the FHFA house price index and rents using HUD fair market rents. Dividing the two series and winsorizing at the 1% level yields a county-by-year R:P.

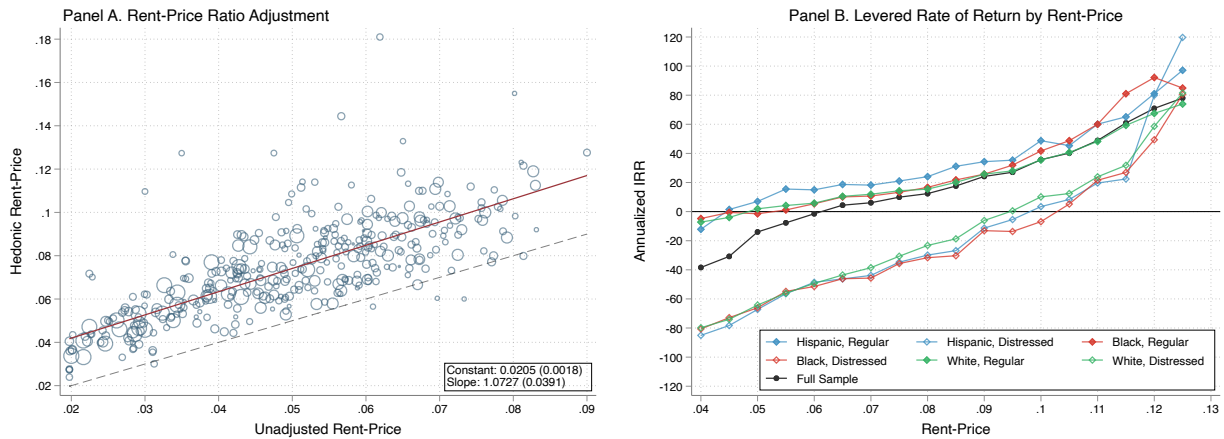
As previously stated, this measure of R:P is biased due to compositional differences between rented and owner-occupied properties (i.e., differences in property characteristics and quality). To address this bias, we follow the procedures in Demers and Eisfeldt (2022) and Gilbukh et al. (2017) to construct MSA-level R:P using hedonic regressions. We use a sample of rented and owner-occupied properties in the 1999-2013 waves of the American Housing Survey (AHS). For each rented property, we regress log rents on interactions of MSA and property age, availability of air conditioning, presence of a basement, bathrooms, bedrooms, all rooms, number of floors, urban/rural, presence of a garage, structure type, lot size, unit square footage, and tenure length. We then use the predicted values from this regression to impute rents for owner-occupied properties, which yields an estimate of R:P for each owner-occupied observation. Following the procedure in Gilbukh et al. (2017), we restrict to a sample of rental units with a similar owner-occupied property

in that MSA (and vice versa), and adjust our estimates to account for the transformation of the prediction from logs to levels. Lastly, we restrict to observations with tenures of five years or less.

We collapse our data to the MSA-year level, allowing us to compare the hedonic R:P constructed from AHS to the unadjusted R:P constructed from county-level Census data (i.e., computed by dividing median rent by median house price). Figure D1, Panel A compares the two series. The unadjusted R:P is highly correlated with the hedonic R:P, and the relationship between the two is well-summarized by a level increase in the unadjusted R:P. To construct our measure of imputed rents, we therefore apply the linear adjustment in Panel A to our county-level measure. This allows us to compute R:P for all county-years in the data while also adjusting for compositional bias.

To incorporate racial differences in imputed gross rents, we rely on estimates in Demers and Eisfeldt (2022) of differences across ZIP codes within MSAs in R:P. Because ZIP codes are a much smaller level of geography than MSAs (the smallest level of geography in the AHS), the adjustment using AHS data is not well-suited to removing the compositional bias in ZIP-level R:P. Since Demers and Eisfeldt (2022) use ZIP-level data on R:P ratios from Corelogic that were not accessible for this study, we apply the estimated cross-ZIP differences from Demers and Eisfeldt (2022), Figure 6. Specifically, we assign each Census tract to a quintile of house prices within each county based off of observed purchases in the property records. We then assume that R:P is 1.25 percentage points higher in the lowest-price quintile, 0.5 percentage points higher in the second quintile, 0.5 percentage points lower in the fourth quintile, and 1.25 percentage points lower in the highest-price quintile. This adjustment calculates relatively higher implicit rents for minorities.

Figure D1: Rent-Price Adjustment



Notes: This figure illustrates the imputation of rent-price ratios and the relationship between rent-price and average realized returns. Panel A plots the hedonically-estimated rent-price ratio (measured in the American Housing Survey) as a function of the unadjusted rent-price ratio (measured using Census median county house prices and rents) by MSA. Solid line denotes linear fit. Panel B plots realized annual levered housing returns as a function of the imputed price-to-rent ratio. Levered housing returns are calculated before applying additional tract-specific adjustment to embed racial differences in rents, and omits adjustments that embed racial differences in housing costs. See Appendix D for more details.

We impute monthly mortgage payments $pymt_{it}$ (and UPB_{iT}), we apply standard amortization formulas assuming a 30-year term. We impute interest rates using a sample of fixed interest mortgages contained in McDash. We impute interest rates separately for first- and second-lien mortgages. For first-lien mortgages, we regress interest rates on the full interaction of LTV and the

five LTV bins used to calculate closing costs as well as closing quarter-by-county fixed effects and log loan amount. For second-lien mortgages, we regress interest rate on LTV and the interaction of the two LTV bins used to compute closing costs, as well as state-by-closing year fixed effects. Coarser fixed effects for second liens are used to compensate for relatively fewer observations of second lien mortgages. For missing county-quarter cells, we impute using the mean value of the fixed effects in that quarter. We then impute interest rates for each transaction in ATTOM using the coefficients and fixed effects from these regressions. Due to limited temporal coverage in the McDash data, if a home was purchased in 1996 or 1997, we impute using the 1998 values and adjust using changes in the average US 30-year fixed interest mortgage rate.

To incorporate racial differences in interest rates into our calculation of levered returns, we require an estimate that nets out differences due to characteristics like location and leverage that are already included in our imputation. To do so, we construct a yearly panel of interest rates for first-lien loans linked to our baseline sample of ownership spells. Interest rates are observed in the McDash mortgage servicing data. In principle, interest rates can vary within an ownership spell either because the homeowner refinances the loan or because the rate changes within a loan (e.g., for adjustable-rate mortgages). The panel constructed using the mortgage servicing data captures both types of changes. We then regress the time-varying interest rate on race indicators and year fixed effects, while also including the same set of controls previously used to impute interest rates (i.e., the interaction of LTV and the five LTV bins, closing quarter-by-county fixed effects, and log loan amount). This estimate yields annual interest rates that are 7.7 and 6.9 basis points for Black and Hispanic homeowners, respectively.³

To impute monthly tax and insurance payments, we use a sample of first-lien mortgages where escrow is observed in the McDash data. We measure escrow payments 18 months after the closing month and regress escrow payment as a share of property value on the full interaction of LTV and the five LTV bins used previously, as well as log loan amount and county-by-closing year fixed effects. We use the predicted values from this regression to impute escrow payments in the main data. If a home was purchased in 1996 or 1997, we impute using the 1998 values due to limited temporal coverage in the McDash data. When we incorporate property taxes that are 12% higher for minorities, we do so by assuming that property taxes comprise two-thirds of monthly escrow payments. When incorporating racial differences in tax payments into our calculation of levered returns, we draw on Avenancio-León and Howard (2019) and assume that minorities pay property taxes that are 12% higher. For the purposes of computing the total return, we conduct this imputation as if a property was purchased with an LTV of 80%, since insurance payments can include mortgage insurance, which is more likely to apply to high-LTV loans.

To impute maintenance costs, we analyze a sample of homeowners in the 2001-2013 waves of the Consumer Expenditure Survey (CEX). We regress total annual home expenditures as a percentage of property value on log property value and state fixed effects. We use the predicted values of this regression to impute maintenance costs. To incorporate racial differences in expenditures in our calculation of levered returns, we regress home expenditures as a percentage of home value on indicators for Black and Hispanic homeowners, and log property value, and estimate that expenditures are 22 and 7 percentage points lower for Black and Hispanic homeowners, respectively. Note that these differences are smaller than those in Appendix Table A6 because our baseline imputation procedure controls for state and home value, meaning that part of these racial differences in home expenditures are already embedded in our baseline imputations.

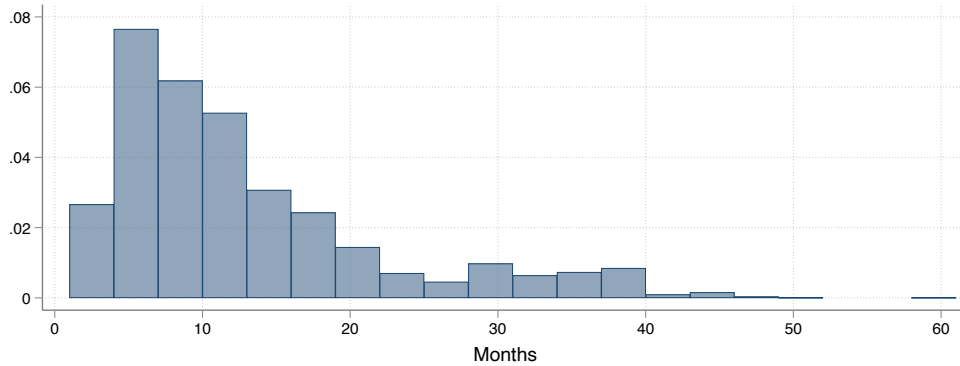
³Note that this approach does not capture racial differences in the cost of credit through channels other than interest rates, such as mortgage fees. Since prior research consistently finds higher costs among minorities, (e.g., Bayer et al. 2018; Clarke and Rothenberg 2018), the omission of these channels likely leads us to underestimate racial gaps in housing returns, making our approach relatively conservative.

Our calculations also factor in growth over time in each of these imputed components. While monthly mortgage payments stay fixed over time, we allow imputed rents to change according to rent growth in our county-level rent series. All other user costs are assumed to grow at the rate of inflation.

Table D1, Panel A provides summary statistics on the imputed variables described in this section. To assess the accuracy of our imputations, we compare the imputed values to the actual values as reported in the McDash data. Table D1, Panel A presents statistics from regressions of actual values on imputed values. The regression coefficients are close to 1, and the R-squared values indicate that our imputations capture a large share of the observed variation.

In the months leading up to a foreclosure, homeowners by definition do not make required mortgage payments, and likely do not make tax payments and home improvement expenditures. To incorporate this behavior, we measure the number of months between the last payment and the REO month for foreclosures in the McDash data. We impute the number of months in a given year that a foreclosed homeowner lives in their home without making payment using the median of this difference by state and year, minus four months. We subtract four months because not all homeowners stay in their homes until the REO month. Figure D2, Panel A plots the distribution of the number of months in our sample. This measure is highly correlated with those used by Mian et al. (2015) and Ghent and Kudlyak (2011) (results available upon request).

Figure D2: Distribution of Imputed Months Without Payment in Foreclosure



Notes: This figure presents the distribution of the imputed number of months that a foreclosed homeowner lives in their home without making mortgage, tax, or home improvement payments. Data consist of ownership spells in baseline sample of ownership spells (described in Section I) ending in a foreclosure.

We assume that in the event of a foreclosure, there is no surplus revenue to be redistributed to the property owner (i.e., the final term evaluates to 0.01). This latter assumption is justified by the fact that in the Freddie Mac single-family loan database, only about 3% of foreclosures have net sales proceeds listed as covered (i.e., there exist surplus funds that could potentially be returned to the homeowner). Imposing a floor of \$0.01 ensures that r^l is well-defined. Lastly, we assume that 5% of the sale price goes towards paying closing costs.

Our baseline sample restrictions (i.e., dropping purchases with combined loan-to-value ratios of more than 102.5% and ownership spells of less than 12 months) are designed to minimize the potential for measurement error in both of our measures of housing returns. We also winsorize annual returns at the 1% level. For 1.5% of the sample, levered returns cannot be computed because of missing imputed components. The levered rate of return is undefined for less than

0.01% of the remaining sample.

E Comparison to Diamond and Diamond (2024)

In this section, we present additional details of our comparison to the results in Diamond and Diamond (2024b), henceforth DD.

For the purpose of illustrating the different impacts of distressed sales on returns in our sample of ownership spells, relative to the DD sample of owner-occupied property-years, we present the Kaplan-Meier survival function of ownership spells in Appendix Figure A21, Panel A. This survival function is computed using our baseline sample (i.e., properties where we observe a purchase but not necessarily a sale). Panel B of Appendix Figure A21 presents the annual hazard function of distressed sales, computed as the percent of properties in our baseline sample surviving to a given year of tenure length that end in a distressed sale in the following twelve months. Panel C plots the estimated probability of a future distressed sale using the same approach used to adjust for unobserved distressed sales in Table 2, Column 5. Specifically, we compute this probability as $p_{rt} = \sum_{i=t}^{30} Pr[\text{distressed sale} | \text{realized tenure length } i]_r \times Pr[\text{sale in year } i | \text{tenure length } \geq t]_r$. We assume that ownership spells lasting 30 years or more do not end in distress. We estimate these components using the observed tenure length survival curve and rates of distressed sales by race and tenure length. The estimated probabilities are presented in Appendix Figure F2.

In addition, Appendix Table A8 presents results that are designed to be as comparable as possible to Table 7.2 in Diamond and Diamond (2024b), which estimates racial differences in log appreciation rates (comparable to unlevered returns), using a sample of four counties where property sales are recorded in Corelogic property records. To maximize comparability, we restrict to properties purchased between 2000 and 2016. In Panels C and D of Appendix Table A8, we further restrict to the four counties of interest: Cook County, IL, Philadelphia County, PA, King County, WA, and Santa Clara County, CA. Reassuringly, our estimates are quantitatively similar to those in DD. This similarity underscores the interpretation that differences in our estimates are driven by differences in the object of interest, sample window, and approach to imputing rents.

Note that in Panels B and D of Appendix Table A8, we compute unlevered returns by replacing the sale price with the imputed unpaid principle balance for ownership spells that end in a distressed sale. Although this enables a more direct comparison with the results in Diamond and Diamond (2024b), Table 7.2, we note that these estimates must be interpreted with caution. Intuitively, this approach to measuring distressed sale proceeds is intended to offer a closer approximation of the financial impacts on homeowners. However, this quantity is unlikely to reflect the entirety of the financial costs borne by homeowners experiencing distressed sales, as indicated by the findings in Diamond et al. (2020).

E.1 Approach to Imputing Rents

Our baseline approach to addressing bias in imputing rents using rent-to-price ratios is to apply an adjustment based on Demers and Eisfeldt (2022), which we discuss in the previous subsection. In this subsection, we take an alternative approach to imputing rents based on the methods in Diamond and Diamond (2024a) and Diamond and Diamond (2024b), henceforth DD. This approach uses properties that switch between occupancy status (i.e., owner-occupied vs. rented) in the American Housing Survey (AHS) to impute rents for owner-occupied properties. This subsection briefly summarizes this approach and presents the accompanying results. To apply the approach in DD, we proceed in three steps.

First, for each property purchase i , we compute its quantile in the national distribution of property sales in year t , denoted by $q_{p,i,t} = F_{p,t}(P_{it})$. In this notation, the subscript p denotes the distribution of purchases. We compute $q_{p,i,t}$ among the sample of all arm’s length, full-consideration transactions of residential property in the ATTOM property data.

Second, we estimate a quantile regression to predict median rents conditional on $q_{p,i}$:

$$q_{r,i} = \alpha + \beta_1 q_{p,i} + \beta_2 q_{p,i}^2 + \beta_3 q_{p,i}^3 + \gamma_{g(i)} + \varepsilon_i \quad (1)$$

In the above equation, γ denotes fixed effects for Census region-by-metro area. $q_{r,i}$ denotes the quantile in the national distribution of monthly rent for property i (computed for each year). Under the rank stability assumption in Diamond and Diamond (2024b), time subscripts can be omitted and this equation can be estimated pooling time periods. Note that this regression requires observing both $q_{r,i}$ and $q_{p,i}$ for each property, which is the reason that it is estimated on a sample of properties that switch occupancy status.

Third, we map quantiles of the distribution of rents to dollar values by computing the inverse CDF of rents for each year (i.e., $F_{r,t}^{-1}(q_{r,t})$). We compute this inverse CDF by estimating a flexible polynomial in $q_{r,t}$, separately for each year. Together, these three steps allow us to map the value of property i in year t (P_{it}) to its rental value (r_{it}).

We conduct each of these three steps using a panel of properties surveyed in the 1997-2013 waves of the American Housing Survey. We restrict to occupied properties, drop mobile homes, and drop owner-occupied properties with self-reported home values that appear atypical using the criteria applied in DD. We use the empirical distribution of self-reported home values to estimate $F_{p,t}$, then estimate Equation 1 on a sample of 4,660 properties that switch occupancy status. We estimate the inverse CDF of rents by fitting the flexible polynomial to the full sample of renter-occupied properties.

To apply this method to our baseline sample of ownership spells, for each purchased property, we compute its place in the distribution of purchase prices in the purchase year. We then apply the coefficients from Equation 1 (estimated in the AHS data) to get an estimated rent quantile, and finally apply the coefficients from the flexible polynomial (also estimated in the AHS data), which yields an estimate of its rent in dollars at the time of purchase. We then inflate rent throughout the ownership spell as in our primary approach.

Appendix Table E2 presents rent-to-price ratios (RtP) at home purchase, imputed using different methods. Our baseline approach to imputing rent (Column 2), which applies an adjustment based on Demers and Eislefeldt (2022), yields higher RtP for Black and Hispanic homeowners relative to the unadjusted imputation (Column 1). Computed within the AHS, the DD adjustment yields even higher RtP for minorities (Column 3). However, when applied to our baseline sample of ownership spells (Column 4), which uses actual sale prices and conditions on observing a property purchase, the DD rent imputation yields racial differences in RtP that are more similar to those using our baseline approach.

The difference between the DD method when applied to the AHS versus when applied to our baseline data (Appendix Table E2, Column 3 vs. Column 4) is due to the difference between self-reported home value and actual house price. To show this, Appendix Table E3 compares the distribution of self-reported home value in the AHS survey data to the distribution of purchase price in our baseline administrative data. Columns 1 and 2 report average self-reported home value and purchase prices in dollar terms. Gaps in dollar terms are larger for purchase prices than for self-reported home value, mostly because the level of purchase prices is higher overall. However, when computed in terms of quantiles (Columns 3 and 4), racial gaps are significantly smaller. For the

Black-White gap, this appears to be in large part due to smaller differences among homeowners who have purchased in more recent years: Column 5 restricts to homeowners in the AHS who purchased their homes in or after 2000, which results in a smaller estimated gap. Moreover, this difference is not due to geographic differences in sample composition: Column 6 re-weights the baseline data to match the cross-sectional distribution of owner-occupied units across Census tracts from the 2010 ACS 5-year estimates, which also results in smaller differences relative to Column 3.

We find that the DD method yields racial gaps in both levered returns and total returns that are similar to those estimated using our baseline approach. Appendix Table E4 presents mean returns by race/ethnicity. Columns 1 and 3 reproduce our baseline results for total returns and levered returns, respectively. Columns 2 and 4 present estimates that use the DD method to impute rents. Although levels of returns are lower for all groups (consistent with lower RtP imputed using the DD method), racial gaps are similar between the two methods. This similarity is at least in part due to the lower levels of rents imputed by the DD method, meaning that racial differences in rents have a smaller impact on racial differences in returns. Overall, this finding offers reassurance that our analysis is robust to alternative approaches to imputing rents.

Table E2: Comparison of Rent Imputation Methods

	Rent-to-Price (%) (1)	Rent-to-Price (%) (2)	Rent-to-Price (%) (3)	Rent-to-Price (%) (4)
Total	6.98	6.94	7.25	6.06
White	7.07	6.94	7.10	5.99
Black	7.17	7.48	8.32	6.78
Hispanic	6.33	6.64	7.66	6.04
Black-White	0.10	0.54	1.22	0.79
Hispanic-White	-0.74	-0.30	0.56	0.05
Sample Adjustment	Baseline Unadjusted	Baseline Baseline	AHS DD (2024)	Baseline DD (2024)
N	13,657,987	13,657,958	104,646	13,343,602

Notes: This table presents average imputed rent-to-price ratios (RtP) by race and ethnicity, and compares different imputation methods. Column 1 presents unadjusted RtP. Column 2 presents our baseline RtP imputation, which applies the adjustment in Demers and Eisfeldt (2022). Column 3 presents RtP imputed using the imputation methods in Diamond and Diamond (2024a) and Diamond and Diamond (2024b), applied to American Housing Survey data. Column 4 presents the imputation method in Column 3, applied to the baseline dataset See Appendix E.1 for more details.

F Additional Estimates of Racial Gaps in Housing Returns

F.1 Alternative Measures of Housing Returns

Our definition of the internal levered rate of return in Equation 3 factors in many relevant cash flows, but it does not take into account substantial indirect costs associated with foreclosure that have been documented in previous work (Diamond et al. 2020; Ganong and Noel 2020). Moreover, imposing a positive floor on terminal cash flows mechanically limits the losses of homeowners who voluntarily choose to sell an underwater property at a loss. As a result, we may underestimate the size of the racial gap in returns. To relax these assumptions, we define measures of the net present value (NPV) of the home purchase, calculated using the same cash flows in Equation 3, but relaxing

Table E3: Comparison of Rent Imputation Methods, Continued

	Self- Reported Home Value (1)	Purchase Price (2)	Self- Reported Value Quantile (3)	Purchase Price Quantile (4)	Self- Reported Value Quantile (5)	Purchase Price Quantile (6)
Total	197851	275302	50.291	57.516	53.721	52.598
White	204093	283997	51.903	58.535	55.121	53.519
Black	145551	219622	37.477	49.112	44.104	42.703
Hispanic	187578	256715	46.934	56.389	50.012	53.208
Black-White	-58542	-64375	-14.426	-9.423	-11.017	-10.816
Hispanic-White	-16515	-27282	-4.969	-2.146	-5.109	-0.311
Sample	AHS	Baseline	AHS	Baseline	AHS Purchased After 2000	Baseline
Weights	AHS Sample Weights	None	AHS Sample Weights	None	AHS Sample Weights	Geographic
N	104,646	13,854,702	104,646	13,797,085	43,335	13,796,308

Notes: This table presents estimates of mean home values and quantiles of home values by race and ethnicity using data from the AHS and from our baseline sample of ownership spells. Column 1 presents average self-reported home value in the AHS sample. Column 2 presents average purchase price in the baseline sample of ownership spells. Column 3 presents quantiles of self-reported home value in the AHS. Column 4 presents quantiles of purchase price from the ownership spell sample. Column 5 restricts Column 3 to homes purchased in or after 2000. Column 6 reweights Column 4 to match the national distribution of owner-occupied housing units across Census tracts in 2010 from the ACS 5-year estimates.

the floor on cash flows in the final period for non-distressed sales. We use the US average 30-year fixed mortgage rate in the month of home purchase to discount cash flows in our NPV calculation.

In addition, we compute Sharpe ratios for ownership spells in our study sample. For this calculation, we use 10-year Treasury yields as the risk-free rate, and compute the standard deviation of excess returns within county-purchase year cells. We calculate Sharpe ratios using unlevered, total, and levered returns.

We provide estimates of racial differences in NPV scaled as a percentage of the net present value of cash flows (i.e., upfront costs, mortgage payment, property taxes, insurance, and maintenance costs discounted by the 30-year fixed mortgage rate) and Sharpe ratios in Appendix Table A7. In addition to our baseline NPV calculations, we also provide two alternative variations of NPV. The first variation calculates NPV imposing a \$50,000 non-pecuniary cost paid at foreclosure, which captures a consumption-equivalent utility cost of foreclosure as in Ganong and Noel (2020). The second variation of NPV weights observations by strictly positive NPV of cash flows. The results of the NPV regressions (weighted and unweighted, with and without consumption-equivalent utility costs), as well as those for Sharpe Ratios are qualitatively similar to results for annual returns.⁴

In addition, we compute measures of returns that only include some of the components of

⁴The analysis of Sharpe ratios also confirms the previous findings in Giacomelli (2021) that investment in individual houses suffers from relatively large idiosyncratic shocks, and on a risk-adjusted basis, its financial performance in recent years has been inferior to investments in the stock market. Of course, households may derive other non-financial benefits from homeownership that compensate them for the excess volatility of their housing investment returns.

Table E4: Robustness to Alternative Rent Imputation Methods

	Total Returns (%) (1)	Total Returns (%) (2)	Levered Returns (%) (3)	Levered Returns (%) (4)
Total	6.13	5.06	5.04	1.50
White	6.57	5.43	6.27	2.61
Black	5.26	4.38	4.10	0.48
Hispanic	4.10	3.29	-1.48	-4.27
Black-White	-1.31	-1.05	-2.17	-2.13
Hispanic-White	-2.47	-2.14	-7.75	-6.88
Adjustment	Baseline	DD (2024)	Baseline	DD (2024)
N	13,430,637	13,122,730	13,430,560	13,122,715

Notes: This table presents estimates of mean housing returns by race and ethnicity using alternative approaches to imputing rents. Columns 1 and 3 replicate the baseline estimates of racial gaps in annual total returns and levered returns, respectively. Columns 2 and 4 present estimates using the alternative rent imputation method in Diamond and Diamond (2024a) and Diamond and Diamond (2024b). See Appendix E.1 for more details.

Equation 3. Appendix Table A1 provides estimates of differences in housing returns, sequentially incorporating components of levered returns. This exercise offers a quantitative comparison of how differences in unlevered returns are affected by incorporating one-time transaction costs, monthly rent and upkeep costs, and leverage. These comparisons illustrate that the most important difference between our measures of levered and unlevered returns is the calculation of leverage itself, rather than monthly costs or one-time transaction costs. This table also elucidates the influence of our adjustments that embed racial differences in rents, property taxes, interest rates, and maintenance.

All of the measures of housing returns that we have examined are financial, and therefore do not capture any non-financial benefits that, in theory, could compensate minorities for lower financial returns. Homeownership offers many benefits that are not captured by financial returns, such as the opportunity to locate in neighborhoods with desirable amenities, such as high-quality schools. If buying a home helps households locate in desirable neighborhoods (Bergman et al., 2019), such non-financial benefits have the potential to contribute to household wealth accumulation.

F.2 Reweighted Estimates

To probe the robustness of our results to sample selection, we conduct three reweighting exercises, which are presented in Appendix Table A5. First, we reweight our baseline sample to match characteristics of the larger sample of ownership spells that does not make use of the merge with the HMDA data. This exercise most directly assesses the impact of potential sample composition bias arising from the merge with the HMDA data, and carries the advantage that both datasets are at the ownership spell level. Specifically, we reweight to match on property ZIP code, purchase year, 0.25 log point bins of log purchase price, and 5 point bins of combined loan-to-value at origination.

Second, we reweight to match the geographic distribution of owner-occupied properties across the US, which serves to impact of the geographic coverage of the property data. Specifically, we reweight to match the distribution of owner-occupied housing units across Census tracts in 2010 from the ACS 5-year estimates.

Third, we reweight the sample to match homeowner demographic characteristics from the ACS.

We construct weights using the 2013-2017 waves of the ACS. We reweight on race, gender, quartiles of income, and number of bedrooms. In the baseline data, quartiles of income are defined within purchase year and state. In the ACS, quartiles of income are defined within state, survey wave, and time since moved into current residence. When number of bedrooms are missing in the baseline data, we impute number of bedrooms using the median number of bedrooms in that property’s decile in the national distribution of interior square footage.

F.3 Alternative Adjustment for Censoring

We estimate racial returns gaps using a sample of property transactions between January 1990 and March 2020. The finite sample window means that certain ownership spells are censored, in that some ownership spells remain unsold by the end of our sample period. The share of censored spells by purchase year is illustrated in Appendix Figure A2. Our baseline approach to addressing censoring is to impute sale prices for unsold properties using local house price indices. In this section, we present an alternative approach to adjusting our estimates of annual returns to address sample censoring. While this approach avoids the bias in our baseline approach associated with not observing distressed sales that occur outside of our sample window (discussed in Section II), our baseline approach has the advantage of being simpler to implement and therefore more transparent. Reassuringly, both approaches yield similar estimates.

Our alternative adjustment for censoring decomposes average returns for homeowners of race r as follows:

$$E[R_i | \text{race}_i = r] = \sum_{t=1}^T (E[R_i | \text{tenure}_i = t, \text{race}_i = r] \times \text{Pr}[\text{tenure}_i = t, \text{race}_i = r]) \quad (2)$$

In the above equation, the expected return for a home purchase for a homeowner i of race r can be decomposed into the weighted sum of expected returns by tenure lengths t .

We compute each component in Equation 2 using our baseline sample of ownership spells, restricted to properties for which we observe a sale by March 2020. Since we no longer rely on house price indices, we additionally include properties purchased in 2015 and 2016, but retain the remaining sample restrictions in our baseline analysis sample described in Section I.

To estimate $\text{Pr}[\text{tenure}_i = t, \text{race}_i = r]$, we use standard tools from survival analysis. Using the sample of properties purchased with mortgages for which we can observe the race of the buyer, Figure F1, Panel A plots non-parametric Kaplan-Meier estimates of the survival functions for our sample of ownership spells in the baseline sample of ownership spells separately by race. The survival curve for Black homeowners is somewhat higher than that of White homeowners throughout the first 30 years of tenure lengths, while the curve for Hispanic homeowners is lower for the first 12 years. Panels B through D plot the distribution of tenure length conditional on sale, and reveal that average tenure lengths for non-distressed sales are somewhat shorter than for distressed sales.

Given that ownership spells in our sample occur between 1990 and 2020 (and mostly between 2000 and 2020), estimating $E[R_i | \text{tenure}_i = t, \text{race}_i = r]$ requires extrapolating returns outside of the 30 year window. Figure F2 plots unlevered and levered returns and the distressed share of sales by tenure length. Plotted points correspond to the realized values in the data. This figure shows that in our sample, annual housing returns stabilize between years 15 and 20. In unreported results, we find that this plateau occurs even among properties purchased in the 1990s, suggesting that the plateau is not merely an artifact of changing sample composition. Accordingly, we extrapolate returns from year 20 onward using the sample-weighted average of housing returns among properties

held for 19 years and longer. To extrapolate distressed sales, we assume that the distressed share of sales declines linearly to 0 between years 19 and 30. These extrapolated values are illustrated in Figure F2.

Combining the non-parametric estimates of the distribution of tenure lengths ($Pr[\text{tenure}_i = t, \text{race}_i = r]$) with the extrapolated values of returns by tenure length ($E[R_i | \text{tenure}_i = t, \text{race}_i = r]$), we can estimate returns corrected for censoring by race. These are presented in Appendix Table A3, Column 1.

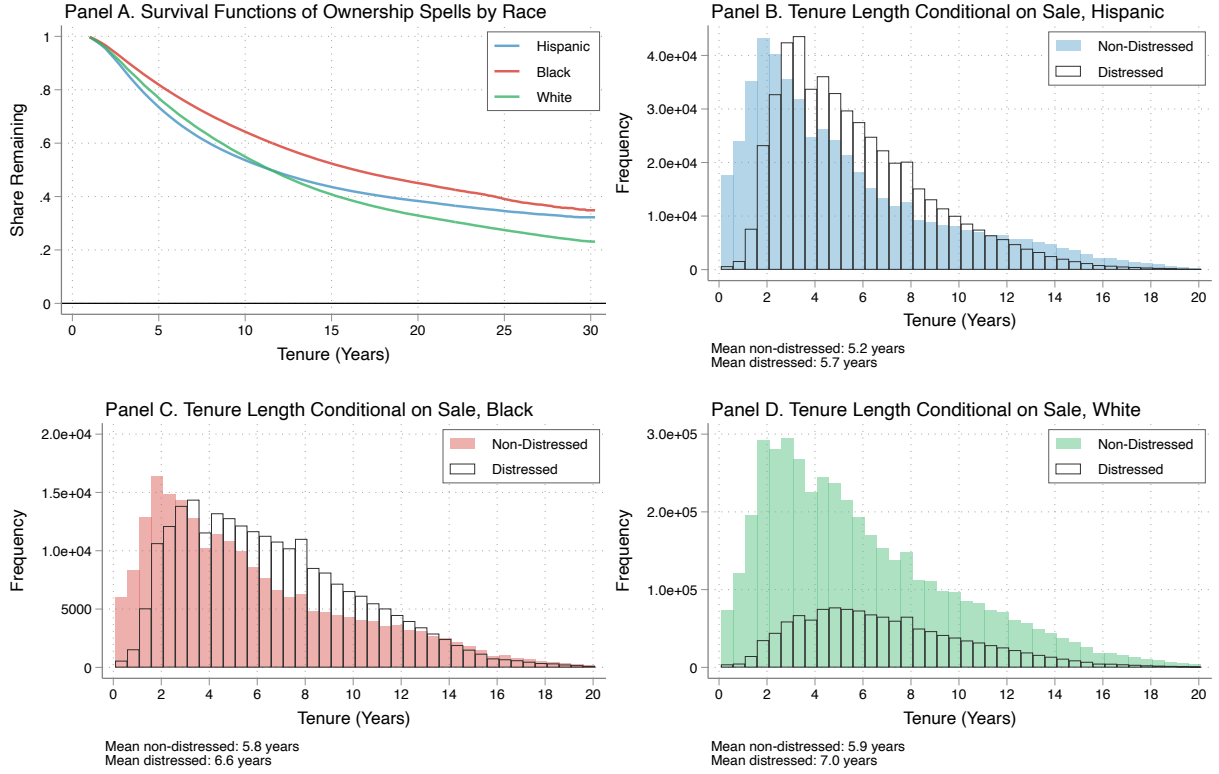
Reassuringly, racial differences in housing returns are similar when we adjust for censoring using our baseline approach (Table 2) and our alternative approach (Appendix Table A3, Column 1). In particular, average unlevered returns are quantitatively very similar by race, yielding similar racial gaps. In addition, racial differences in levered returns are also quantitatively similar. Although average levels of race-specific levered returns are somewhat different, this is partly attributable to much larger variance in levered returns, as well as to the conservative bias in our baseline estimates associated with coding unsold properties as not ending in a distressed sale (comparing Columns 1 and 5 of Table 2). Lastly, the share of properties ending in distressed sale is markedly lower using our baseline adjustment for censoring; however, this also reflects the conservative bias in our baseline estimates. After adjusting our baseline estimates for this downward bias, average rates of distressed sales are very similar across the two approaches (Table 2, Column 5).

F.4 Alternative Adjustment for Cash Purchases

We apply an adjustment for cash purchases to our alternative adjustment for censoring. We take an approach that is conceptually similar to our baseline adjustment for cash purchases, which uses the returns of mortgaged properties not sold in a distressed sale as a proxy for the returns of cash purchases. As discussed in Section II, we observe home purchases made in cash in the ATTOM data but cannot measure the self-reported race and ethnicity of cash buyers. For non-cash buyers in our baseline approach, we measure homeowner race and ethnicity using the HMDA mortgage origination data, which do not cover home purchases made in cash. To construct a race-specific measure of housing returns that is adjusted for cash purchases, we begin by using the American Community Survey (ACS) to construct a race-specific measure of the share of owner-occupied home purchases that are made in cash. In the 2013-2017 ACS data, we restrict to homeowners that have been living in their current residences for less than two years. Among this subsample, 76.5%, 78.6%, and 76.7% of White, Black, and Hispanic homeowners have an outstanding mortgage, respectively. We take this as our measure of the share of homeowners buying a house with a mortgage.

To construct a race-specific measure of housing returns for cash purchases, we observe that the unlevered housing returns associated with properties purchased in cash are very similar to those purchased with a mortgage and sold in a non-distressed sale. Figure F3, Panel A indicates that these returns track each other closely throughout the distribution of tenure length. While returns among cash properties appear to be higher for short tenure lengths, this is likely to be at least in part driven by non-owner-occupant investment properties that are not the focus of our analysis. Therefore, we view the race-specific returns associated with properties bought with mortgages and sold in non-distressed sales as a reasonable proxy for the race-specific returns associated with cash properties. Figure F3, Panel C further validates this assumption by imputing race/ethnicity from homeowner last name and Census block using the methods developed by Imai and Khanna (2016), and shows that average unlevered returns among cash purchases are very similar by imputed race/ethnicity. Our race-specific estimates adjusted for cash purchases are computed as the weighted average of returns for mortgaged and cash properties, with weights corresponding to the race-specific estimates of the mortgaged purchase shares from the ACS. We proxy for the annual unlevered rate of return

Figure F1: Survival Function by Race



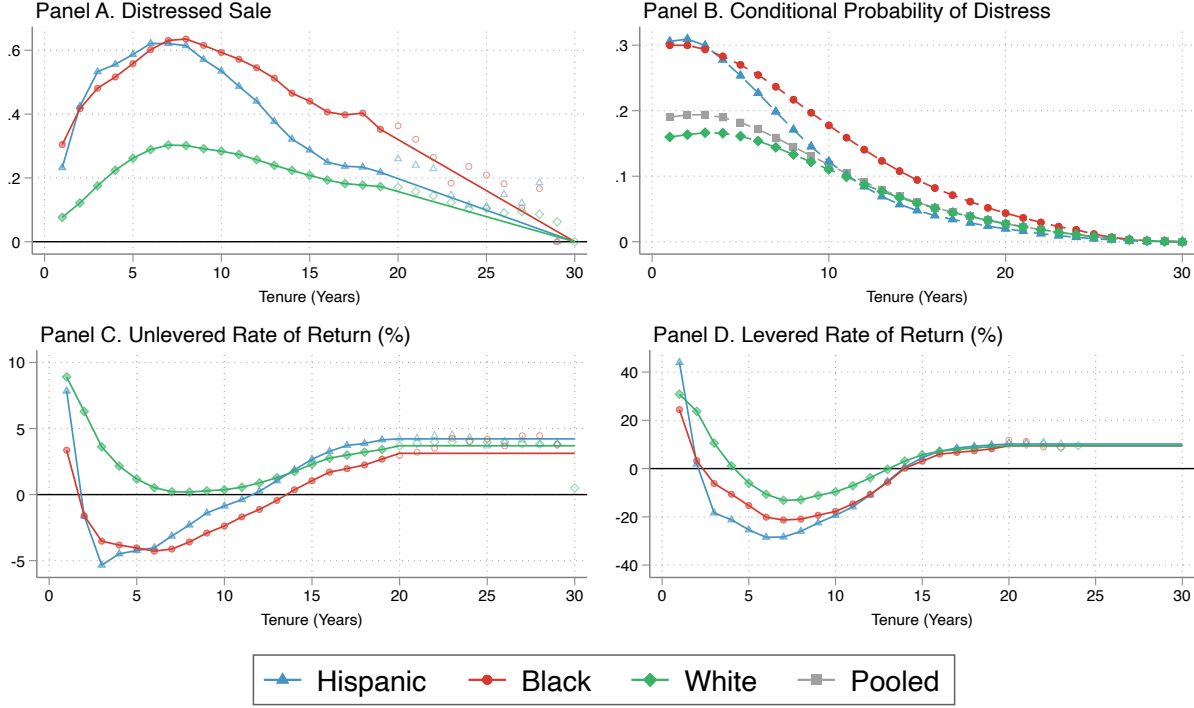
Notes: This figure illustrates the distribution of tenure length associated with ownership spells in which the property was purchased with a mortgage, by race/ethnicity. Panel A plots Kaplan-Meier survival functions by race/ethnicity. Panels B through D plot the distribution of tenure length conditional on sale, separately for distressed and non-distressed sales by race/ethnicity. Data for Panel A come from sample of properties with observed purchase prices, including those that have not yet been sold by the end of March 2020. Data for Panels B through D restrict to properties with observed sale prices.

for cash purchases using the unlevered returns for mortgaged properties sold in non-distressed sales.

To adjust levered returns for cash purchases, we use our computation of the internal rate of return without leverage, namely the total return. That is, we include implicit rents, tax and insurance payments, maintenance costs, a 5% transaction cost, and assume closing costs that are half of the imputed value for a loan with an LTV of 80% (given that a large share of closing costs are not associated with borrowing). We then use these total returns among the subset of properties that were not sold in a distressed sale as a race-specific proxy of internal returns associated with cash purchases. Lastly, we assume that homes bought without leverage do not result in distressed home sales.

To adjust cash purchases for censoring, we compute an analogous version of Equation 2 for cash purchases. First, we estimate race-specific survival functions for cash purchases. Figure F3, Panel B presents Kaplan-Meier survivor functions among both cash and mortgaged purchases. Since the survival curves differ slightly, we adjust the race-specific survival functions for mortgaged properties by the ratio of the survival functions of cash and mortgaged properties at each tenure length. This adjustment yields race-specific tenure probabilities for cash purchases ($\Pr[\text{tenure}_i = t, \text{race}_i = r]$).

Figure F2: Extrapolated Returns and Rates of Distressed Sales

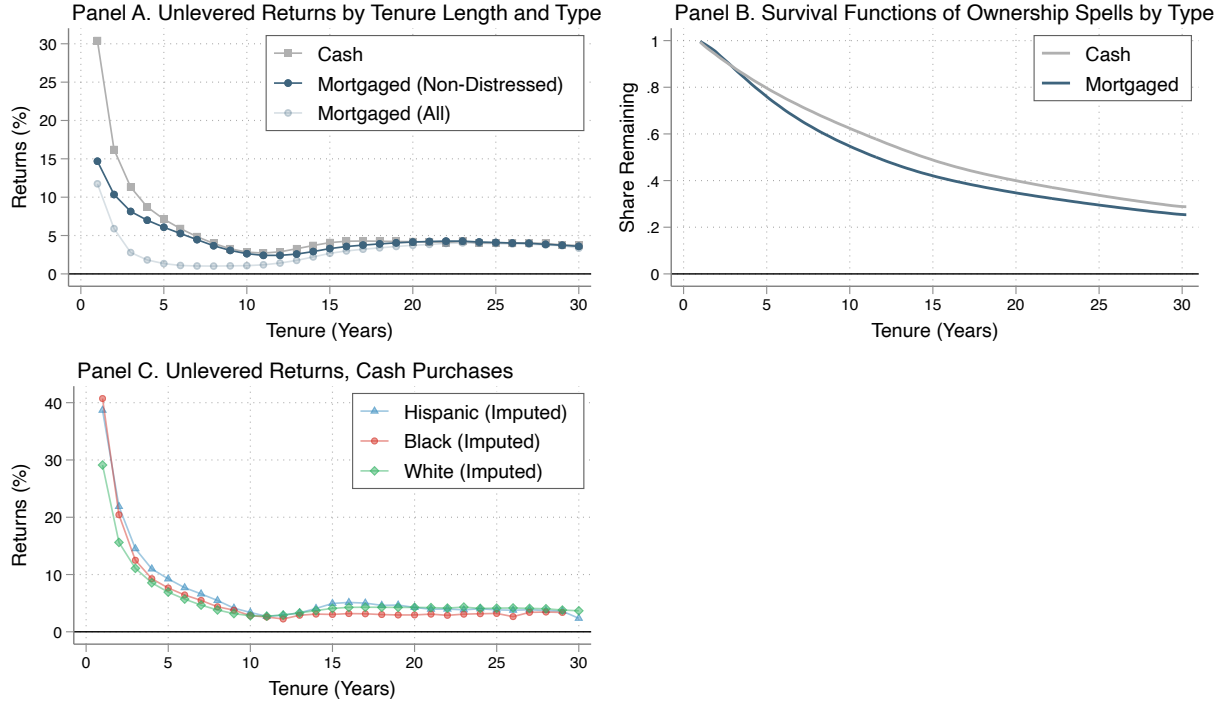


Notes: This figure plots the extrapolated values of rates of distressed sales (Panel A), conditional likelihoods of an ownership ending in distress given survival to a given number of years (Panel B), annual unlevered housing returns (Panel C), and annual levered returns (Panel D) as a function of tenure length, separately by race/ethnicity. Points correspond to average values observed in the data. Lines correspond to extrapolated values. Since outcomes are only extrapolated for tenure lengths longer than 19 years, extrapolated values are equal to observed values at earlier tenure lengths.

Second, we use the race- and tenure-specific returns associated with non-distressed cash purchases as proxies for the returns of cash purchases ($E[R_i | \text{tenure}_i = t, \text{race}_i = r]$). Combining the two using Equation 2 yields race-specific estimates of returns associated with cash purchases adjusted for censoring. To combine censoring-adjusted returns for cash purchases and distressed home sales, we take the weighted average of estimates for cash and mortgaged purchases that have been adjusted for censoring, with weights corresponding to the estimates of cash purchases by race from the ACS.

The estimates of race-specific housing returns using our alternative adjustment for censoring and cash purchases are presented in Table A3, Column 2. These results can be compared to our baseline adjustment for cash purchases, which is presented in Table 2, Column 4. As with results that have not been adjusted for cash purchases, estimated racial gaps in both housing returns and distressed sales are very similar between the two approaches. Moderate differences in race-specific average levered returns and rates of distressed sales appear to be caused by the anticipated bias in our baseline approach. In our baseline approach, we treat properties that remain unsold as of the end of our sample window as if they are not sold in a distressed sale. Since some of these properties will eventually be sold in a distressed sale, this approach leads to an upward bias in returns, and since minorities are more likely to realize a distressed sale, a downward bias in the racial gap in housing returns. Consequently, our baseline approach should be viewed as conservative. Given

Figure F3: Annual Housing Returns and Tenure Lengths by Purchase Type



Notes: This figure compares housing returns and tenure lengths between home purchases made in cash and those made with a mortgage. Panel A indicates that housing returns among cash purchases are similar to the non-distressed returns realized by mortgaged properties not sold in distressed sales. Panel B plots Kaplan-Meier survival functions for cash and mortgaged purchases. Panel C plots average annual unlevered returns by tenure length and race/ethnicity imputed from homeowner name and Census block. Data for Panel A come from sample of properties with observed purchase and sale prices, including properties purchased in cash. Data for Panel B include properties that have not yet been by the end of March 2020. Data for Panel C come from sample of cash purchases with observed purchase and sale prices.

that the estimates of the gap are of a very similar magnitude, and that our baseline approach is simpler to implement and therefore more transparent, we present results using the baseline approach throughout the paper.

G Racial Differences in Ownership Transitions

Are minority homeowners living in distressed neighborhoods able to take advantage of the availability of nearby discounted homes being sold at foreclosure? In theory, if all real estate transactions occur within race (e.g., Hispanic sellers only selling to Hispanic buyers) and foreclosures do not entail substantial property depreciation, then higher rates of distressed sales need not depress aggregate housing returns for minority homeowners. We analyze racial transitions by analyzing the ownership spell that occurs following a given spell in our baseline sample. Using the subsample for which race and ethnicity is observed in the HMDA data for two consecutive ownership spells, Appendix Table A10 shows that the majority of distressed minority home sales involve buyers of a different race/ethnicity. Specifically, 79% of distressed White homeowners sell to a White buyer, 31% of Black homeowners sell to a Black buyer, and 39% of Hispanic homeowners sell to a Hispanic buyer. While minority homebuyers are disproportionately likely to purchase distressed homes owned by minorities, most individual buyers of distressed minority homes are not minorities. Nonetheless, minority homebuyers are slightly more likely than non-minorities to purchase distressed homes. Appendix Table A10, Panel C indicates that Black (Hispanic) homeowners are 2 (5) percentage points more likely to purchase a distressed home, relative to White homeowners.

What role do buyers outside of the neighborhood play in purchasing distressed sales? Appendix Table A10 indicates that homes owned by minorities are on average more likely to be purchased by institutional buyers, especially in distressed sales. In Appendix Table A16, we examine patterns using granular fixed effects (i.e., county, purchase and sale year, income, gender, family composition, and leverage). Distressed homes are substantially more likely to be purchased by an institutional buyer or as an investment property (i.e., not owner-occupied). Institutional and investment buyers appear to have an even larger presence in buying minority-owned distressed homes. In addition, buyers of distressed homes own the home for less time than buyers of non-distressed homes, suggesting that distressed properties are more likely to be flipped.

Purchasing a distressed home can require substantial additional investment in order to counteract recent property depreciation. Since we do not observe this investment, we are unable to measure the net discount associated with distressed home sales. Nonetheless, our results suggest that net discounts associated with distressed properties owned by minorities, to the extent that they exist, may disproportionately benefit outside investors and buyers of other racial groups. By the same reasoning, since minority homebuyers are slightly more likely to purchase distressed homes (Appendix Table A10, Panel C), it is possible that these homebuyers may be more able to benefit from these net discounts.

H Analysis of Mortgage Modifications

In this section, we document that minority homeowners are more likely to receive modifications than observationally similar White homeowners. We then estimate the causal impacts of mortgage modifications on housing returns.

H.1 Racial Differences in Receipt of Mortgage Modifications

We analyze a de-identified sample of homeowners who have become 90 or more days past due in the Fannie Mae, Freddie Mac, and ABSNet datasets. We classify the outcome of a homeowner’s first 90-day delinquency into one of three categories: modified, foreclosed, or self-cured. We directly observe modifications and foreclosures, and we define a loan as self-cured if the borrower makes three consecutive payments or pays off the loan. The outcome of the delinquency is defined as

whichever of these three events occurs first. In this sample, 62% of 90-day delinquencies result in foreclosure, 16% are resolved by the borrower self-curing, and 21% result in a modification.⁵

We estimate the following equation:

$$\mathbb{1}\{\text{Delinquency Outcome}_i\} = \alpha_0 \mathbb{1}\{\text{Black}_i\} + \alpha_1 \mathbb{1}\{\text{Hispanic}_i\} + \mu_{f(i)} + \varepsilon_i \quad (3)$$

The outcome is defined as an indicator that the 90 day delinquency of homeowner i ended in modification, foreclosure, or self-cure. $\mu_{f(i)}$ denotes fixed effects which vary by specification. The values of $\hat{\alpha}_0$ and $\hat{\alpha}_1$ capture the extent to which delinquent Black and Hispanic homeowners are more or less likely than White homeowners to end their delinquency in a given outcome.

Appendix Table A17, Column 1 estimates a baseline version of Equation 3 with fixed effects for quarter of default, and shows that Black homeowners are about 7 percentage points more likely to receive a modification relative to White homeowners. The Black-White difference is driven by relatively lower foreclosure and self-cure rates. Hispanic homeowners are slightly less likely to receive a modification, and appear to be much less likely to self-cure. Adding in granular fixed effects that capture borrower characteristics (e.g., credit score, income) and mortgage characteristics (e.g., current LTV, origination year, interest type) results in a 5.0 and 1.6 percentage point higher modification rates for Black and Hispanic homeowners, respectively (Column 3). Notably, even when comparing homeowners within the same servicer, Census tract, and time period, Black homeowners are 2.5 percentage points more likely to receive a modification than White homeowners (with no statistically significant difference between Hispanic and White homeowners).

These results indicate that even conditional on defaulting on payments as well as a wide range of observables, Black (and to some extent Hispanic) homeowners appear to receive favorable treatment from mortgage servicers. These findings are similar to those in Collins et al. (2015), who document similar patterns in a sample of subprime loans originated between 2004 and 2006. This behavior is consistent with mortgage servicers internalizing the higher house price penalties associated with foreclosures on minority-owned properties. Such discrimination could arise from an equilibrium in which mortgage servicers attempt to maximize value for investors and anticipate that the value of avoiding a foreclosure for a Black or Hispanic homeowner is higher than for a White homeowner.

H.2 Empirical Strategy: Variation from Mortgage Servicers

To estimate the causal impacts of mortgage modifications on housing returns, we use data from Fannie Mae, Freddie Mac, and ABSNet to measure modifications for properties in our baseline sample of ownership spells. Fannie Mae and Freddie Mac publish publicly-available data on single-family conventional mortgages originated since 2000 and 1999, respectively. These data contain mortgage characteristics, monthly payment information, the identity of the mortgage servicer, and records of loan modifications. ABSNet provides similar data for over 90% of residential mortgages collateralized through private-label mortgage-backed securities.

We restrict attention to a de-identified sample of 1.2 million loans in which the homeowner has become 90 days delinquent on their mortgage and whose homes are observed in the property data. After first becoming 90 days delinquent, 62% of homeowners subsequently experience a foreclosure, 21% receive a modification before a foreclosure occurs, and 16% catch up on their payments without a modification. Appendix C provides additional details on the mortgage modifications sample.

We leverage quasi-experimental variation in servicers' propensities to modify mortgages, which has been shown in prior work to vary both across servicers and within servicers over time (Piskorski

⁵Note that a small fraction of loans (less than 2%) end in bankruptcy or repurchase. These observations are kept in the analysis dataset but not captured in the three main categories of outcomes.

et al. 2010; Agarwal et al. 2017; Aiello 2019; Korgaonkar 2020). To construct a measure of servicer modification propensity, we estimate equations of the following form on the sample of homeowners who have become 90 days delinquent:

$$\mathbb{1}\{\text{Mod}_{it}\} = \mu_{f(i)} + \gamma_{s(i),t} + \varepsilon_i \quad (4)$$

In Equation 4, i denotes homeowner, t denotes year, $\mu_{f(i)}$ denotes a vector of fixed effects that includes fixed effects for the Census tract interacted with origination year and current year; the source of the data (i.e., Fannie Mae, Freddie Mac, or ABSNet); deciles of credit score at origination; an indicator that the loan is interest-only; and an indicator that the loan is a negative amortization loan. The vector also includes fixed effects capturing deciles of the original loan amount, current LTV, and years remaining in the loan term, and an indicator that the loan is an adjustable-rate mortgage. $\gamma_{s(i),t}$ denotes servicer-by-year fixed effects. The outcome is defined as an indicator that the loan received a modification within 12 months of first becoming 90 days delinquent. We estimate a separate $\gamma_{s(i),t}$ for each state, restricting the sample to loans outside of that state and outside of any ZIP codes and commuting zones that overlap with that state. The estimated $\hat{\gamma}_{s(i),t}$ provide a plausibly exogenous measure of servicer propensities to modify loans.

The estimated servicer propensities serve as an instrument for modifications. We estimate the following specification by 2SLS, using $\hat{\gamma}_{s(i),t}$ as an instrument for $\mathbb{1}\{\text{Mod}_{it}\}$:

$$r_i = \alpha_0 \mathbb{1}\{\text{Mod}_{it}\} + \mu_{f(i)} + \varepsilon_i \quad (5)$$

Equation 5 regresses an outcome r_i (e.g., the rate of return realized by homeowner i) on indicators that i received a modification within 12 months of default interacted with race indicators. $\mu_{f(i)}$ denotes a vector of fixed effects. In our baseline specification, this vector includes interacted fixed effects for Census tract, purchase year, year of default, and indicators for interest-only loan and negative amortization loan, as well as servicer fixed effects. We two-way cluster standard errors at the level of the servicer and within county, origination year, and default year cells. Table 5, Column 1 presents the first stage, indicating that the instrument has a large and statistically significant impact on modifications.

Under the exclusion assumption that the servicer modification propensity affects realized returns only through receipt of modification, estimating Equation 5 by 2SLS recovers the causal impacts of modification receipt by race. The exclusion assumption is plausibly satisfied in this setting because homeowners are unlikely to be aware of their mortgage servicer’s propensity to modify loans. Moreover, the inclusion of servicer fixed effects controls for potential bias from systematic sorting into servicers and leverages within-servicer over-time differences in modification propensities.

We find that modifications reduce the likelihood of experiencing a distressed home sale for all racial groups. Table 5, Column 2 presents naive OLS estimates, which indicate that a modification for a White homeowner is associated with a 5.5 percentage point increase in housing returns (and with slightly higher returns for Black and Hispanic homeowners). 2SLS estimates in Column 3 report that modification increases annual returns by 11.2 percentage points for White homeowners, indicating that the OLS estimates are downward biased. A possible reason why the OLS estimates are downward biased is if servicers allocate modifications to distressed neighborhoods where foreclosures are particularly costly. While the 2SLS point estimates are smaller for Black homeowners, the difference between Black and White homeowners is relatively modest compared to the overall effect and not statistically significant. Moreover, point estimates suggest modestly larger impacts for Hispanic homeowners, but these are also not statistically significant. Analyzing impacts on levered returns (Column 4) yields similar results. These findings imply that modifications have economically large impacts on housing returns for minorities.

Modifications increase housing returns by helping homeowners avoid distressed sales. 2SLS estimates in Table 5, Column 5 indicate that modifications reduce the share of ownership spells ending in a distressed sale by 36 percentage points, with statistically insignificant differences across racial groups. These findings are consistent with Collins et al. (2015), who find similar associations between modification and foreclosure across racial and ethnic groups in a sample of subprime loans originated between 2004 and 2006.

We conduct two robustness exercises in order to validate our research design. First, we probe the sensitivity of our estimates to alternative sets of controls. In Appendix Table A18, we interact our baseline fixed effects with terciles of credit score at origination (Column 2), loan-to-value ratio at default (Column 3), and income at loan origination (Column 4). The estimated impacts of modifications are quantitatively similar across specifications. Second, we conduct a placebo exercise in which we regress the outcome of interest (e.g., indicator for a distressed sale) on a vector of characteristics measured prior to default and use the predicted values to define an index.⁶ Intuitively, if our exclusion restriction holds, then the modification instrument should have no impacts on these outcomes measured prior to default. Forming an index using these predicted values concisely summarizes the relationship between the vector of characteristics and the outcome. In support of our empirical design, Appendix Table A19 regresses the index on the modification instrument and shows that the estimated coefficients are small and statistically insignificant.

To evaluate whether an expansion of modifications could be made more cost-effective by targeting specific types of homeowners, we analyze heterogeneous effects of modifications. Appendix Table A18, Columns 5 and 6 interact the modification indicator with measures of market distress and family size, respectively. For each homeowner i who is first delinquent in year t , we compute the share of property sales in year t (excluding the sale of homeowner i) that are distressed sales. We define homeowners in distressed tracts as those in the top quartile of this measure. Similarly, we define single applicants as those whose loan application registered in the HMDA data does not include a co-applicant. The impact of modifications on unlevered returns is 5.1 percentage points larger in distressed tracts and 2.9 percentage points larger for single-applicant households. These results suggest that targeting distressed neighborhoods and households may be particularly cost-effective.

⁶The characteristics include loan type (i.e., conventional, FHA, VA) fixed effects; loan purpose (i.e., purchase or refinance) fixed effects; indicators for adjustable-rate, interest-only, and negative amortization; fixed effects for deciles of credit score, income, interest rate, and loan amount at origination; current year fixed effects, and data source fixed effects.

References

- Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., Piskorski, T., and Seru, A. (2017). Policy Intervention in Debt Renegotiation: Evidence From the Home Affordable Modification Program. *Journal of Political Economy*, 125(3):654–712.
- Aiello, D. (2019). Financially Constrained Mortgage Servicers. *Available at SSRN 3063513*.
- ATTOM (2023). U.s. foreclosure activity shows continued rise in third quarter. <https://www.attomdata.com/news/market-trends/foreclosures/attom-q3-and-september-2023-u-s-foreclosure-market-report/>. 2023-10-12.
- ATTOM Data Solutions (2020). Assessor and Recorder Data. 3.0 Bulk Data Product Line, Version 2.04.
- Avenancio-León, C. and Howard, T. (2019). The Assessment Gap: Racial Inequalities in Property Taxation. *Available at SSRN 3465010*.
- Bayer, P., Ferreira, F., and Ross, S. L. (2018). What drives racial and ethnic differences in high-cost mortgages? the role of high-risk lenders. *The Review of Financial Studies*, 31(1):175–205.
- Benmelech, E., Guren, A., and Melzer, B. T. (2017). Making the house a home: The stimulative effect of home purchases on consumption and investment. Technical report, National Bureau of Economic Research.
- Bergman, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L. F., and Palmer, C. (2019). Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice. Technical report, National Bureau of Economic Research.
- Bogin, A., Doerner, W., and Larson, W. (2019). Local House Price Dynamics: New Indices and Stylized Facts. *Real Estate Economics*, 47(2):365–398.
- Campbell Communications (2011). Tracking Real Estate Market Conditions Using the Housing-Pulse Survey. [Link](#). Accessed 2021-09-01.
- Chetty, R., Sándor, L., and Szeidl, A. (2017). The Effect of Housing on Portfolio Choice. *The Journal of Finance*, 72(3):1171–1212.
- Clarke, K. A. and Rothenberg, L. S. (2018). Mortgage pricing and race: Evidence from the northeast. *American Law and Economics Review*, 20(1):138–167.
- Collins, J. M., Reid, C. K., and Urban, C. (2015). Sustaining Homeownership After Delinquency: The Effectiveness of Loan Modifications by Race and Ethnicity. *Cityscape*, 17(1):163–188.
- Corbae, D. and Quintin, E. (2015). Leverage and the foreclosure crisis. *Journal of Political Economy*, 123(1):1–65.
- Demers, A. and Eisefeldt, A. L. (2022). Total returns to single-family rentals. *Real Estate Economics*, 50(1):7–32.
- Diamond, R. and Diamond, W. (2024a). Affordability of owner-occupied housing across us cities. In *AEA Papers and Proceedings*, volume 114, pages 340–345. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.
- Diamond, R. and Diamond, W. F. (2024b). Racial differences in the total rate of return on owner-occupied housing. Technical report, National Bureau of Economic Research.
- Diamond, R., Guren, A., and Tan, R. (2020). The Effect of Foreclosures on Homeowners, Tenants, and Landlords. Technical report, National Bureau of Economic Research.
- Equifax (2022). Credit Risk Insight Servicing McDash. ABS Credit Risk Insight Version 4.0.
- Fannie Mae (2020). Single-Family Loan Performance Data. <https://capitalmarkets.fanniemae.com/credit-risk-transfer/single-family-credit-risk-transfer/fannie-mae-single-family-loan-performance-data>.
- Federal Housing Finance Agency (2021). Annual House Price Index, Census Tracts. <https://www.fhfa.gov/data/hpi/datasets?tab=additional-data>. Accessed: 2021-12-29.
- Ferreira, F. and Gyourko, J. (2015). A new look at the us foreclosure crisis: Panel data evidence of prime and subprime borrowers from 1997 to 2012. Technical report, National Bureau of Economic Research.
- FFIEC (2016). Home Mortgage Disclosure Act. <https://www.ffiec.gov/hmda/hmdaproducts.htm>.
- Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J. R., and Westberry, M. (2022). Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D030.V10.0>.
- Freddie Mac (2020). Single Family Loan Level Data. <https://www.freddiemac.com/research/datasets/sf-loanlevel-dataset>.

- Ganong, P. and Noel, P. J. (2020). Why Do Borrowers Default on Mortgages? A New Method for Causal Attribution. Technical report, National Bureau of Economic Research.
- Gerardi, K., Willen, P., Zhang, D. H., et al. (2020). Mortgage Prepayment, Race, and Monetary Policy. Technical report.
- Ghent, A. C. and Kudlyak, M. (2011). Recourse and residential mortgage default: evidence from us states. *The Review of Financial Studies*, 24(9):3139–3186.
- Giacoletti, M. (2021). Idiosyncratic risk in housing markets. *The Review of Financial Studies*, 34(8):3695–3741.
- Gilbukh, S., Haughwout, A., and Tracy, J. (2017). The price to rent ratio: a macroprudential application. *November*, 24:2017.
- Imai, K. and Khanna, K. (2016). Improving ecological inference by predicting individual ethnicity from voter registration records. *Political Analysis*, 24(2):263–272.
- Intercontinental Exchange, Inc. (2022). McDash Primary Data Set.
- Korgaonkar, S. (2020). The Limited Benefits of Mortgage Renegotiation. *Available at SSRN 2924981*.
- Lewtan (2016). ABSNet Loan RMBS. Version 52.
- Mahon, J. (2010). Short Sales Stand Tall. Federal Reserve Bank of Minneapolis [Link](#). Accessed 2021-09-01.
- Mian, A., Sufi, A., and Trebbi, F. (2015). Foreclosures, House Prices, and the Real Economy. *The Journal of Finance*, 70(6):2587–2634.
- Piskorski, T., Seru, A., and Vig, V. (2010). Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis. *Journal of Financial Economics*, 97(3):369–397.
- United States Bureau of the Census (2020). Quarterly Workforce Indicators. <https://lehd.ces.census.gov/data/>. Accessed: 2020-09-01.
- United States Bureau of the Census (2022). 2010 Census, Tract-Level Statistics. data.census.gov. Accessed: 2022-07-05.
- United States Bureau of the Census (2025a). American Housing Survey (AHS) 1997-2015 Panels. <https://www.census.gov/programs-surveys/ahs.html>. Accessed: 2025-01-13.
- United States Bureau of the Census (2025b). Survey of Income and Program Participation (SIPP) 1991-2018 Panels. <https://www.census.gov/programs-surveys/sipp.html>. Accessed: 2025-01-13.
- Zhang, C. (2019). A shortage of short sales: Explaining the underutilization of a foreclosure alternative.