Returns to Homeownership and Inequality: Evidence from the First-Time Homebuyer Tax Credit

Marina Gindelsky[1]
U.S. Bureau of Economic Analysis

Jeremy Moulton[2]
University of North Carolina – Chapel Hill

Kelly Wentland[3]
George Mason University

Scott Wentland[4]
U.S. Bureau of Economic Analysis

December 17, 2022

Abstract

Homeownership is often promoted as a key to building wealth and a pathway to reducing wealth inequality across racial groups in the U.S. However, the prior two decades saw a housing bubble form and burst with significant impacts on homeownership rates and returns, which varied by race. In this paper, we examine 1) how returns to homeownership differ by race using address-level microdata over this period, and 2) whether a policy intended to expand homeownership among more liquidity constrained buyers, the 2008-10 First-time Homebuyer Tax Credit (FHTC), altered this broader dynamic. To answer these questions, we employ a unique national dataset linking internal American Community Survey (ACS) households to transactions from Zillow's ZTRAX database for 2000-2016. We exploit the FHTC setting using a difference-in-differences framework, finding that income-eligible homebuyers had substantially higher gross returns to homeownership compared to ineligible households. In contrast to our initial findings, where minority householders realized lower returns to housing over the broader boom-bust-recovery period, the FHTC results also show eligible Black and Hispanic householders who purchased a home during the relevant policy period significantly outperformed White householders.

Keywords: inequality, housing, home prices, homeownership, wealth, race, wealth gap
JEL Classifications: D31, G50, J15, R30, R38

Disclaimer: Any views expressed here are those of the authors and not necessarily those of the Bureau of Economic Analysis, U.S. Census Bureau, or the U.S. Department of Commerce. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: 7507311, sub-project number 7517751, Disclosure Review Board (DRB) approval number: CBDRB-FY23-033). Data was provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group. We would like to thank Christine Dobridge for help comments on a prior draft. Any errors are our own.

[1] Office of the Chief Economist, 4600 Silver Hill Rd, Suitland, MD 20746; Marina.Gindelsky@bea.gov
[2] Department of Public Policy, Abernethy Hall, CB 3435, Chapel Hill, NC 27599; moulton@email.unc.edu
[3] Department of Accounting, School of Business, 4400 University Drive, Fairfax, VA 22030; kwentlan@gmu.edu
[4] *Contact author*: Office of the Chief Economist, 4600 Silver Hill Rd, Suitland, MD 20746; scott.wentland@bea.gov

1. Introduction

Housing is a key sector of the economy, making up a large proportion of investment,[1] personal consumption expenditures,[2] and fixed assets[3] in the United States. A home is also the single largest asset for most U.S. households (Wolff 2017), where the median homeowner had $225K of housing equity in their primary residence in 2019 (Bhutta et al. 2020). Yet, like other assets, neither homeownership nor housing wealth are equally distributed across all racial and demographic groups. Conditional on owning a home, White households had a median asset value of $230K, while Black and Hispanic households had $150K and $200K, respectively.[4] Black households are far less likely (45%) to own a home than White households (74%) according to the 2019's Survey of Consumer Finances (SCF) and housing dominates their portfolios. In contrast, White households had more diversified portfolios with significantly higher proportions in equities and other assets, which yielded a higher return than housing since 1950. Recent literature has attributed these significant differences in homeownership rates and returns as key contributors to both the present and historic wealth gap between Black and White households (Derenoncourt et al. 2022, Kermani and Wong 2021, Wolff 2022, and Xavier 2021).[5]

Given the importance of the housing sector and relevance of this asset's returns to the wealth gap, we investigate two empirical questions in this paper. First, we examine whether the gross rate of return (ROR: the unlevered return generated from the difference in the initial purchase

[1] Historically, residential investment is often between 3 to 6% of GDP, and in 2021 it was 4.8% of GDP. It is large share of gross private domestic investment (U.S. BEA 2022a).

[2] Expenditure on housing (i.e., housing services) typically constitute around 10-12% of GDP, and was about 10% of GDP and 15% of Personal Consumption Expenditures in 2021 (U.S. BEA 2022b).

[3] Residential housing made up approximately 49% of private fixed assets in 2021 (U.S. BEA, Table 1.1. Current-Cost Net Stock of Fixed Assets and Consumer Durable Goods - FAAt101)

[4] More data from the Federal Reserve's *Survey of Consumer Finances* is available here, which is the source of the SCF figures above.

[5] Derenoncourt et al. (2022) construct a new data series and show that the wealth gap is more pronounced than the numbers above initially imply, underscoring the key role of differential returns on assets.

price and the eventual sale price) to buying and selling one's home is relatively homogenous, or whether it differs significantly by demographics during the boom-bust-recovery period of the 2000s and 2010s. Using a unique address-level linked dataset, we are able to account for both household-level and location characteristics in our analysis by matching internal Census records with property-level data from Zillow's Transaction and Assessment Dataset (ZTRAX). Second, we focus on non-distressed single-family homes (excluding short sales and foreclosures) and use a set of federal policies that expanded homeownership, the 2009-2010 First-Time Homebuyer Tax Credit (FHTC), as an empirical setting to investigate how returns to housing evolve across homebuyers when homeownership is expanded among this lower-income cohort of more marginal homebuyers.

Our initial descriptive analysis and multivariate regression results suggest that the gross ROR to buying and selling a single-family house which transacted from 2000 through 2016 was significantly lower for Black and Hispanic householders; and, returns for Asian householders were not meaningfully different than White householders. While the mean annualized ROR for White homeowners was about 1%, this relatively modest return does not outpace inflation over a period of American history where inflation was quite low. Conversely, Black householders with similar demographics, home characteristics, and in similar locations had a 1.2 to 2 percentage point (pp) lower annualized ROR, yielding a negative *nominal* return over this boom-bust-recovery period, on average.

While informative, the differential returns during this period do not necessarily provide a policy prescription about the effectiveness of promoting new homeownership. They do, however, provoke further thought about assumptions in the literature on the housing sector and motivate further analysis. The bulk of these incentives motivate homeownership on the margin. So, while *overall* returns to homeownership for minority households may have been lackluster or even negative over this period, a natural question is: do more *marginal* homeowners entering the market have a higher, lower, or equal return?

To answer this next empirical question, we examine the effects of the FHTC policies on rates of return to housing in our main set of analyses. The FHTC was implemented in phases over approximately two years (2008-10), providing lower-income first-time buyers with as much as $8,000 in up-front capital in the form of a tax credit (with the initial phase providing an interest-free loan rather than what effectively became a grant). While Hembre (2018) and Berger et al. (2020) shed light on the credit's impact on homeownership rates, a missing link in this literature is determining whether the low-to-middle income homebuyers were able to build equity in a way that might help narrow the wealth gap. The FHTC offers a quasi-experimental policy setting to investigate this question.

On one hand, the timing of the FHTC near the bottom of the housing bust may have helped liquidity constrained buyers gain access to an asset market that would subsequently rise in price and help build wealth through capital gains. On the other hand, there are a number of plausible scenarios consistent with the prior evidence in the literature which suggest opposing outcomes in wealth accumulation. First, if the FHTC induced marginal homeowners into the market, they may have bid against others in the same income cohort on similar homes, inflating prices for homes (consistent with the evidence in Berger et al. 2020), which may not eventually sell for the same premium in a later period. Hence, one possible scenario might be that the gross return to this asset could be smaller for those eligible than for those outside the income-eligibility policy threshold policy. Second, those who actually took up the tax credit may have bid against those in the same income cohort who were not eligible as "first-time buyers," securing winning bids for better homes, where their greater gross return to homeownership could be offset by lower returns among the losers, making the predicted overall impact ambiguous. Third, even if the eligible cohort had greater returns overall, these returns may not have been universally shared by all; in particular, if minorities which are underrepresented in the housing market (e.g., Hispanic and Black homeowners) earned lower returns than White homeowners, this could potentially exacerbate the racial wealth gap.[6] Given the possible scenarios above that generate ambiguous predictions *ex ante*, it is thus an empirical question whether low-to-middle income households eligible for this policy outpaced higher income ineligible cohorts, and whether minority homeowners experienced relatively favorable gains.

Using a difference-in-differences (DiD) design, our results show that the income cohort eligible for the FHTC had higher gross ROR (1%, annualized) to housing than those ineligible (by income). This is robust to a number of relevant considerations, including comparing the "treated" income cohort to two different "control" groups not eligible for the policy (1 - a "control" cohort just above the eligibility threshold, who would later be eligible; 2 - a "control" cohort well-above the threshold that would never be eligible for the policy). When we limit the sample such that the DiD design compares a treatment group and control group closer to the income-eligibility threshold, we find the Black and Hispanic households in the treated cohort (i.e., those who purchased a home during the relevant FHTC policy period) experienced much higher returns than those ineligible. Thus, our analysis suggests that while the FHTC may have helped all eligible

[6] Statistics from the Federal Housing Administration (FHA), with whom the vast majority of loans are for first-time homebuyers, suggest that the agency has "long been known to serve a disproportionately larger number and share of minority homebuyers, particularly African-American and Hispanic buyers" (Comeau et al. 2012, p. 2). The FHA attributes this to the fact that many low-income minority households do not have traditional loans provided by conventional lenders and instead tend to be served either by Government-Sponsored Enterprises (GSEs) like Fannie Mae, FHA, or subprime or other nontraditional conventional loans.

participants build wealth, the greater impact on minority households may have helped reduce the wealth gap.

These results contribute to multiple large literatures and provide useful insights into how returns to these assets evolved over an important period of the U.S. housing market. First, we add to a growing body of evidence that returns to homeownership are not homogenous across races, exploiting a unique national dataset that allows for direct, address-level linkages. Differential returns are an integral portion of more recent literature on between-race inequality levels for White and Black households (see Akbar et al. 2022 and Derenoncourt et al. 2022 for excellent discussions of the historical background). Thus, not only does the existence of differential returns in housing provide some insight into the origins of existing wealth inequality disparities, but this nationally representative analysis also suggests heterogeneity in returns (here by race, rather than income) is a necessary consideration when constructing estimates of wealth inequality (see Fagereng et al. (2016, 2020) for an illustration of heterogeneity by income for Norway). Finally, by shedding new light on how experiences and outcomes in the housing market differ by race, this paper also adds to extensive research in real estate and urban economics evaluating racial differences in a variety of related contexts, including homeownership,[7] lending,[8] appraisals,[9] tax assessments,[10] and sorting and location decisions.[11]

---

[7] See Carrillo and Yezer (2009), Gabriel and Rosenthal (2005), Deng, Ross, and Wachter (2003), Charles and Hurst (2002), Haurin and Rosenthal (2007), Gyourko, Linneman, and Wachter (1999), and Coulson (1999).

[8] For example, see Bartlett et al. (2022), Zhang and Willen (2021), Ambrose, Conklin, Lopez (2021), Bhutta and Hizmo (2021), Bayer, Ferreira, and Ross (2018), Cheng, Lin, and Liu (2015), Kau, Keenan, and Munneke (2012), and Ladd (1998).

[9] See, for example, LaCour-Little and Green (1998).

[10] See Avenancio-León and Howard (2022), Hodge, McMillen, Sands, and Skidmore (2017), and McMillen, D.P. and Weber (2008).

[11] See, for example, Shertzer and Walsh (2019), Christensen and Timmons (2018), Bayer and McMillan (2012), and Boustan (2012), among many others.

2. Background

In this section, we discuss the context of our contribution in greater depth by providing additional background and discussion of: (A) prior literature related to racial disparities in homeownership and returns (and subsequent implications for wealth inequality), and (B) homeownership incentive policies and the FHTC. While we touch on some of this literature in the introduction above, the subsections below provide greater context for our results in the literature and further motivation for examining the FHTC policy in particular.

A. Homeownership Disparities

Though a number of papers have looked at returns to homeownership over time, many of these exercises have either focused on a subset of cities or states, or else assessed returns on a coarser geographic level (typically zip code or county). Moreover, there is significant variation in the results found in these papers. One reason for the variation is the time period considered.[12] For example, analyses of pre-1990s housing returns, such as Akbar et al. (2022), Blau and Graham (1990), Dawkins (2005), and Chambers (1991), find significantly lower returns to housing for Black households due to factors for Black households such as (1) segregation policies leading to households paying a premium, (2) a negative correlation between the share of Black homes and house values in the neighborhood, (3) household demographic and socioeconomic characteristics associated with lower permanent incomes (including less inheritance income), (4) slower first time homeownership transitions, and (5) liquidity constraints. While many of these factors still contribute to gaps in returns in the present-day, recent literature has argued the dynamics have changed in important ways. Bayer et al. (2017) consider the next two decades (1990-2008) and

[12] See also Collins and Margo (2011), Kollmann and Fishback (2011), and Rothstein (2017) for a discussion of homeownership gaps.

discuss the concept of "decentralized discrimination" (vs. centralized racism) wherein a transaction tax occurs as participants in the transactions drive up the prices for Black and Hispanic buyers (a 1-3% premium).

Though White households have the highest rates of homeownership, and possibly the highest returns, there are significant differences among nonwhite homeowners as well. Kahn (2021) examines the more recent period and finds that while Black households have a lower rate of return than Whites, Asian and Hispanic households have higher rates of return. They attribute part of this to the fact that Asians and Hispanics are more likely (than Blacks) to be located in areas which have high home price appreciation (such as California). Kermani and Wong (2021) find that the lower housing returns for Blacks and Hispanics in the recent period are due to their higher share of distressed sales and foreclosures.[13]

One limitation in this literature has been the data. Ideally, to better understand returns to housing at a very micro-level, researchers would be able to account for confounding variation at the household-level, like socioeconomic and demographic characteristics of the household purchasing/selling the home. Inferences made from more aggregated data require additional assumptions about the underlying data in this regard. One innovative approach in recent years has been for researchers to link housing microdata at the property transaction-level with anonymized data that contains demographic information, like the Home Mortgage Disclosure Act (HMDA) data, which contain some demographic information like race and income. Although the

[13] The policy setting in their paper is the expansion of the availability of mortgage modifications for distressed homeowners. We view this as complementary to our policy setting of the FHTC Tax Credit, as they explore the impact of addressing one of the drivers of returns (distressed sales) that target *existing* homeowners (after they have already purchased a home), while we investigate how expanding *new* homeownership drives returns (excluding foreclosures and distressed sales from our analysis). We view understanding the consequences of policies that expand *new* homeownership as critical, given the numerous federal policies since the 1968 Fair Housing Act promoting homeownership more generally, and policies emphasizing expanding minority homeownership in particular.

anonymized HMDA data do not disclose the names and addresses of the home transactions, they do contain geographic information that allow researchers to link aggregated data at some fine level of geography (e.g., census tract or zip code) or to probabilistically link a given transaction in the housing dataset by linking HMDA transactions with the same characteristics in the same location. In this study, we evade this limitation by using internal ACS data from the Census Bureau, which is a nationally representative sample of households that is directly linkable to ZTRAX at the address-level. We discuss the details of this linkage in Section 3 below; however, in the context of the literature, it is important to highlight this novel approach and the data's unique capacity to incorporate a rich set of household and property characteristics into our analyses of the rate of return to homeownership. We return to this point in Sections 3 and 4.

One of the implications of lower homeownership rates and returns for minorities is the persistence of wealth inequality.[14] Portfolio composition and the role of differential returns has been explored as a key reason for the persistence of wealth inequality in recent papers (Xavier (2021), Derenoncourt et al. (2022), Wolff (2021), and Wolff (2022)). Outside of very important geographical factors, the primary household determinants of wealth inequality can be thought of as those which drive income inequality (namely, demographic and human capital differences) and those that drive asset (or liability) accumulation (intergenerational patterns and portfolios). Minority homeownership growth led to significant convergence with White households from 1989-2007, but renewed divergence during the Great Recession (Wolff 2022).[15] Both Hispanics

[14] Overall, the large literature on this topic presents many possible explanations for persistent racial disparities. See also Altonji et al. (2000), Charles and Hurst (2002), Barsky et al. (2002), and Killewald (2013) for further discussion of racial wealth gaps.

[15] Wolff (2022) mentions that, "though the substantial differences by race in asset allocation that are documented below are well known, the evidence on rate of return by asset type is rather scanty." He summarizes some exceptions to this from decades ago, which look at differences in the housing market. For example, see: Blau and Graham (1990) and Denton (2001).

and Blacks experienced sharper falls in wealth during the Great Recession than white households (McKernan et al. 2014). From 2007 to 2010, median wealth fell by 44% as wealth inequality rose (Wolff 2017; Piketty, Saez and Zucman 2018). Depending on the nature of the shock, those at the bottom (or top) of the distribution may be disproportionately affected, causing rises (or falls) in wealth inequality. This was particularly evident during the run-up and fallout of the housing crash in 2007, which our sample captures. Moreover, although most assets declined during this period, the negative rate of return to housing was nearly double that of financial assets (held mostly by those in the top 10% of the distribution) (Wolff 2017).

B. Incentivizing homeownership, First-Time Homebuyer Tax Credit, and the Wealth Gap

Given that a fundamental difference in the asset position of Whites and non-Whites is the share of homeownership (and related returns), policies promoting homeownership have been advocated as a way to reduce this racial gap. In a recent survey article on homeownership, Goodman and Mayer (2018) remark that certain assumptions about the financial benefits of homeownership were widespread: "for decades, it was taken as a given that an increased homeownership rate was a desirable goal" (p. 31). A host of policies and market dynamics had long promoted or incentivized homeownership.[16] On the banking and credit side, this included increased supply of credit that fueled spending (e.g., Mian and Sufi (2009, 2011)) and expansions of subprime lending (Mayer, Pence, and Sherlund 2009). On the demand side, there was price

[16] On the policy side, for example, the Housing Act of 1968, HOME Investment Partnerships Program of 1990, Mortgage Revenue Bond Program, and Community Development Block Grant Program of 1974 promote homeownership by subsidizing low-income homeownership in the U.S.. This list is far from exhaustive.

appreciation that begat more market entrants expecting further price appreciation (Foote, Gerardi, and Willen 2012) and lower perceived risk of default (Gerardi et al. 2008).[17]

Another such policy that incentivized homeownership was the FHTC. After the housing market crash, Congress authorized the FHTC in three pieces of legislation (Housing and Economic Recovery Act of 2008, American Recovery and Reinvestment Act (ARRA) of 2009, and the Worker, Homeownership, and Business Assistance Act of 2009) that would allow first-time buyers (defined as not having owned a home within the last three years) a tax credit of up to $8,000. While the initial phase of the program was slightly less ($7,500), it had to be repaid over the next 15 years, effectively functioning as a small interest-free loan to help new homeowners with a down-payment or closing costs. For buyers to be eligible for this credit, the legislation required a modified adjusted gross income to be below $75,000 for single-filer households and $150,000 for joint-filer households.[18] Analyses from Hembre (2018) and Berger et al. (2020) found that this initial phase was less effective at stimulating homeownership than the later phases (January 2009 through July 2010), where the statutes then stipulated that the amount up to $8,000 would *not* have to be paid back.[19] We return to the details of how we use this policy setting as a source of quasi-experimental variation in Section 4; but, it is important to note here in the context of the broader literature that this policy shares a common thread with numerous policies promoting homeownership among more marginal buyers. Thus, a better understanding of the causal effects or consequences of policies incentivizing homeownership is one of the goals of this paper.

[17] Prior to the housing bust, a number of studies had shown that homeownership was associated with long-term wealth accumulation in a variety of contexts. See, for example: Belsky and Duda 2002; Haurin and Rosenthal 2004; Herbert and Belsky 2008.

[18] The subsequent legislation would later expand eligibility to $225,000 for joint and $125,000 for single filers, starting in November 2009. The initial (repayable) phase included homes purchased from April 9, 2008 through December 31, 2008, while the non-repayable credit phase ran from January 1, 2009 through July 2010.

[19] Both Hembre 2018 and Berger et al. 2020 document causal evidence that the FHTC increased homeownership, and Berger et al. (2020) found it increased median home prices by about $2,400 (or 1.1%).

Though the FHTC did not promote homeownership for racial minorities explicitly, there are several reasons to expect marginal first-time homebuyers would be particularly important for low-income minority households.[20] First, while the tax credit is provided to low-income first-time homebuyers in general, this group of homebuyers tends to include a higher share of minority households than higher-income first-time homebuyer groups (Herbert and Belsky 2008).[21] Second, loan service providers report that constraints with traditional loans from conventional lenders are more binding for minority households, especially African-American and Hispanic buyers, and the tax incentive was more likely to be relevant where traditional market incentives (including lending opportunities) were not sufficient to induce a home purchase. Finally, a number of interest groups advocated that the FHTC would be critical to the economic development and welfare of particular minority groups. Hence, given all of the above, it may not be surprising that Goodwin and Zumpano (2011) found that Black, Hispanic, and Asian homebuyers were more likely than White homebuyers to be making their purchase as a result of the FHTC using data from surveys of homebuyers during the policy period. Thus, this policy becomes a particularly germane setting to understanding the effects of more general policies promoting homeownership that are also specifically relevant for minority homebuyers.

3. Data

A key contribution of this paper is the novelty of the dataset we assembled and its utility in answering the two empirical questions of this study. Specifically, we constructed a unique address-level dataset from internal American Community Survey (ACS) microdata matched with

[20] For example, in a discussion of the ARRA of 2009, the Congressional Black Caucus outlines the FHTC as one of the credits critical to economic development and welfare of African American communities during the policy period (Hackshaw 2009, p. 6).
[21] Herbert and Belsky (2008) make this point in their discussion of the income and racial distribution of first-time homebuyers using tabulations of first-time homebuyers over 1991-2003 using responses from the American Housing Survey (see the discussion of their Exhibit 3 on p. 15 of their review).

Zillow Transaction and Assessment Dataset (ZTRAX) data. To do this, Census Bureau staff matched ZTRAX assessment and transaction data on an address-level to an internal household-level address identifier used by the ACS.[22] This allowed us to link parcel-level ZTRAX housing data provided by Zillow, which is drawn from public tax assessor records at local municipalities and is organized by state. We began by merging the provided assessment and transaction data so that each parcel was associated with its sale history. The data was then cleaned and collapsed such that each address (parcel) and its associated property characteristics (e.g., square footage, lot size, bedrooms, bathrooms, etc.) and transaction information (sale price, date, type, and mortgage) could be merged to household data from the ACS. For this analysis, we restricted our sample to single family residences and dropped transactions outside the scope of our analysis like commercial transactions, vacant land, and agricultural sales.

Though the Census Bureau has multiple surveys which collect information on housing and households, the 5-year pooled samples from the ACS have the largest samples (3.5 million households per year), which include many questions on housing, in addition to demographic and labor market questions that are widely used across the social sciences. Moreover, it is representative of the U.S. population (with survey weights applied).[23] While the ACS collects both individual and household-level data, for this exercise we focus on only household(er) level data. Thus, any individual-level characteristics are those of the household head, allowing for joining our data at the property-level.

---

[22] Specifically, the Census Bureau assigned the ZTRAX records a Master Address File Identifier (MAFID) at an address-level, which is an identifier unique to each address and is used internally at Census to link household-level data. The MAFID was then used to link the datasets. Though imperfect, this matching process is preferable to any matching on observables. As even address-matching can have some mismatches (i.e., some "fuzzy matches"), we take additional steps in culling the data to discard bad matches (described below).

[23] There have been many studies which utilize the American Housing Survey (AHS), since it contains many detailed housing questions. However, it is a much smaller sample and lacks some of the demographic information available in the ACS.

As this is a new linked dataset, we take additional steps to undertake a pre-analysis to ensure the quality of the match. To do so, we compared characteristics of the home as reported by respondents in the ACS and by localities in the ZTRAX data. We not only limited the sample to single-family residences in both data sets,[24] but also ensured the observations were similar in terms of home characteristics. Following approaches similar to Nolte et al. 2021 and others using ZTRAX,[25] we sought to create a "Zillow-consistent" sample and reduce the possibility of matching-error by dropping observations if (1) the ACS and ZTRAX counties were different, (2) the difference in the number of bedrooms reported is more than two, (3) the home has likely been flipped or changes hands multiple times within a short period of time (i.e., bought and sold within the same year), or (4) the homeowners had moved into the home less than one year before the survey year to ensure the respondents are those who bought the home (since the ACS survey is conducted 12 months a year). Further, we exclude observations when: (1) the home was foreclosed (from information provided by Zillow) or (2) the difference in the reported year built is greater than 10 prior to 2002 (ACS year built is in buckets) or greater than 5 after 2002 (finer buckets).

Given the possible years for our matched sample (2008-2016) and our research question, we only included transactions that took place after 2000. Not only were this transaction data more complete in coverage, but they were closer to the surveyed years (which is important when matching household characteristics to the right buyers and sellers). We further culled the sample by dropping the following "outlier" observations: (1) the householder is younger than 20, (2) household income is less than $1,000, (3) sale price is less than $10,000 or greater than

[24] Single family residences are by far the most common type of domicile in the United States; however, non-single family residences are much more likely to have a lower quality match due to messiness in unit number reporting (e.g., APT vs. Unit vs. #, or even missing unit numbers entirely). Out of this practical consideration to maximize the quality of our dataset, we chose not to include these in our sample and instead focus on single family residences for this research question, while acknowledging the limitation for extrapolation of our results.

[25] For example, our data cleaning considerations are similar to those in Gindelsky et al. 2020, Moulton and Wentland (2018), and Chen et al. 2022, which use the same vintage of ZTRAX data.

$10,000,000, (4), the aggregate rate of return is greater than 500% or the annualized rate of return is greater than 50%, or (5) the ratio of the sale price (in ZTRAX) to home value (in ACS) is greater than 300% or less than 30%. The resulting dataset yields about 140,000 observations that are consistent with the above criteria. This sample will be used for our default specification, regressing the rate of return on economic, geographic, and demographic characteristics (a full list of those variables is in Table 1).

The largest restriction on the dataset is matching ACS respondents who purchased *and* sold a home within our sample period. There are generally sizable periods of time between when a household purchases and sells a home. According to a recent survey by the National Association of Realtors (NAR), the typical home seller has been in their home for about 8 years,[26] with a much lower tenure among younger homebuyers of 3-5 years for sellers under 40.[27] To calculate a return based on the initial purchase price and the eventual sale price, we can only use records from ACS respondents who were surveyed in between those transaction dates of a linked property in the ZTRAX sample.[28] Thus, because we are (by necessity) excluding many long-tenure homeowners, our results should be interpreted as reflecting the returns for homeowners closer to the average tenure, which is likely to be (mechanically) skewed younger than the total population. In our linked sample, the average tenure for a household selling a home is 5 years. While this sample skews somewhat younger, these transactions nonetheless represent an important step on the path of long-time homeownership. Repeat-buyers tend to roll over equity from their prior home sale into their

[26] NAR 2021 Profile of Home Buyers and Sellers. Prior to the Great Recession, NAR reports that average home tenure was historically 6-7 years, but in recent years it has typically fallen within 8 to 10 years.
[27] See NAR's 2020 Home Buyer and Seller Generational Trends report: https://www.nar.realtor/sites/default/files/documents/2020-generational-trends-report-03-05-2020.pdf
[28] One limitation of this approach is that it must drop observations for which we have a purchase price and no eventual sale price (i.e., they are still living in their home by the time our sample period ends). Or, we may have a sale price and not an initial purchase price (i.e., they had purchased the home prior to our sample period beginning – 2000).

next home, leading to a larger down payment, lower likelihood of having to pay private mortgage insurance (PMI), and lower likelihood of needing a loan/gift from a family member.[29]

From this data, we measure the gross rate of return to homeownership in three different ways. We first use the simplest calculation, the nominal or gross rate of return (ROR), which is the difference between the initial purchase price and the eventual sale price and is then scaled by the initial purchase price. Our second measure is an inflation-adjusted ROR (IAROR), which we calculate by deflating the ROR using the Personal Consumption Expenditure (PCE) Index produced by BEA (cited as the Federal Reserve's preferred measure of inflation).[30] Given households bought and sold these assets after holding them for different time periods, we then simplify comparisons of the dependent variable across households and to provide a more standard measure consistent with the asset returns literature with our third measure: an annualized ROR (defined in Table 1 using a standard compound annual growth rate formula). The latter is our default specification for most of our analysis when we refer to the "return," unless otherwise stated. One limitation of this dataset is that it does not allow for a "net rate of return," as it does not include a host of potentially important idiosyncratic costs for the household, property-specific investments, and a variety of other costs associated with homeownership.[31] We thus emphasize that the results should be interpreted as gross returns and continue to refer to them as such throughout the paper. Goldsmith-Pinkham and Shue (2022) focus on a similar measure as the outcome of interest in studying the gender gap of housing returns.

---

[29] For additional details on sources of down payments, see NAR's 2020 report on Downpayment Expectations & Hurdles to Homeownership.

[30] See, for example: https://www.federalreserve.gov/faqs/economy_14400.htm

[31] Kermani and Wong (2021) evaluate a sample from the Panel Study of Income Dynamics over 2001 through 2017, finding that expenditures on repairs (as a percentage of the home's value) is very similar across races. There is a moderate difference in home improvement expenditures, where White households spend a bit more on home improvements. We leave further analysis of a "net" measure of housing returns to future research.

Table 2 provides descriptive statistics for the broader sample used for our initial analysis. The average gross ROR in our sample is 12.5% (annualized 0.7%). This average ROR had not outpaced inflation over this period, thus the inflation adjusted rate of return (IAROR) was negative, or about -2.5%. There are also large differences in annualized ROR by race. While White and Asian householders had an average annualized ROR of about 1%, Hispanic and Black householders had an average annualized return of -0.8% and -2.14%, respectively.

Those with the most expensive homes (highest quintile, as measured by the respondent's value of their own home in their state at the time they were surveyed) had very high rates of return (23.0%), which is about 8 times the returns of those with the least expensive homes (3.2%). Table 2 also shows that there is significantly less heterogeneity by education and age, though there is some. The average home in the sample is a 3 bedroom, owned about 5 years, built in 1980, with a value of about $300,000; the average household income in the sample is $108,600. Half of the households surveyed in our linked sample were surveyed between 2008-2010. This is particularly helpful in examining the FHTC. This sample of homeowners is mostly White (80%) and fairly young, with half of the householders under 40 (though the average age is 44). The householders are also fairly educated, with half having at least a college degree.

4. Methods

A. Baseline Multivariate Analysis of Differential Returns by Householder Characteristics

We begin exploring a couple broader questions, conditioning on a variety of household-level factors in a similar multivariate regression analysis. First, are gross returns to buying and selling a home similar across demographic groups? And second, more specific to the racial wealth gap literature, do racial minority householders realize lower, higher, or the same returns? Answering these initial questions is important for establishing some baseline facts about observable differences in returns across groups over an extended period of time (2000 to 2016), conditioning on potentially relevant factors. Given that these differences could arise from numerous potential sources, mechanisms, and policies, this initial set of analyses should be interpreted with caution and not necessarily viewed as causal. The answers do, however, motivate our subsequent analysis of the FHTC policy. That is, if there are significant observable (non-causal) differences in returns to housing, do policies like the FHTC (causally) exacerbate or mitigate such differences?

To answer these questions, we begin by estimating the following multivariate regression:

$$\mathrm{ror}_{ht} = \beta_0 + \sum \beta\, race_h + \sum \gamma\, Household\ Demographics_h + \sum \delta\, Property\ Characteristics_h + \propto_h^{location} + \propto_t^{year} + \varepsilon \qquad (1)$$

where *ror* is the gross rate of return (ROR) for a given household *h* when selling their home purchased in year *t*, which we define a few different ways in our initial analysis as we discussed in the prior section (and we use annualized ROR for the subsequent analysis of the FHTC). Our primary variables of interest are the race/ethnicity and other socioeconomic and demographic characteristics of the householder. Here *race* is defined by the ACS as the race (ethnicity) of householder *h*, which we code into mutually exclusive categorical dummy variables: White, Black (primary racial response), Asian (primary racial response), Hispanic (regardless of race), Other (all other racial responses).[32] When the sample size is large in our initial sample period (2000-2016), we include the "Other" category; but, in later analysis when we narrow the sample substantially, we will omit this category when observation counts become sufficiently low. We

[32] Though there are numerous ways to characterize race and ethnicity, we chose this common breakout of the data, as institutions like the Census Bureau and Federal Reserve use a similar White/Black/Hispanic/Other categories when they report data on household assets in the SCF, for example.

incorporate the following *Household Demographics* from the ACS and examine their impact on returns for the household: age, education, marital status, household size, and income. See Table 1 for more precise definitions of these variables as we use them in the regression. In untabulated tests, we explored multiple functional forms of these variables, but we include the categorical and dummy variables for ease of interpretation and consistency of later exploring sample cuts or stratifications by these categorical thresholds.[33]

The remaining variables in equation (1) account for variation in returns to housing specific to the property characteristics, location, and time of the purchase/sale, which are commonly used in hedonic pricing models. Specifically, *Property Characteristics* come from both ACS and Zillow ZTRAX datasets, which include the following (with the corresponding dataset in parentheses): square footage of the living area of the home (ZTRAX), lot size (ZTRAX), number of years owned (ZTRAX), number of bedrooms (ACS), year built category (ACS), mortgage status (ACS), home value category (ACS). These variables and categorical thresholds are defined in more detail in Table 1. We include purchase year fixed effects for households who purchase a home in year t. To account for location-specific time-invariant heterogeneity, we include county fixed effects in our default specification. For robustness, we alter this in various specifications to include different location and location-by-year interactions in later analyses (i.e., state-by-year, county-by-year, census tract, block group, and zip code fixed effects). We cluster robust standard errors by county.

Overall, the richness of this data provides a unique opportunity for interpreting our variables of interest. Conditioning on these characteristics, the β coefficients for each race indicator

[33] We explore the sensitivity to using categories versus other forms of these demographic variables in another study using this linked data (Gindelsky et al. 2022), albeit covering a different topic. Our results are not sensitive to using categories (e.g., age buckets or a college education dummy) as opposed to a linear form of age or income, or a broader categorization of education. Given that these household characteristics are obtained via survey over an extended period of time, where the survey also changes subtly over time, we include survey year fixed effects.

variable become a comparison of individual householders of different races who purchased a similar home (along key observable dimensions), in the same general location, within the same year, with similar socioeconomic characteristics as White householders, who also held the property for a similar number of years. When they subsequently sold their properties, the difference in their rate of return is β (relative to the return of the similar White household as the omitted group), on average. Of course, there are likely a number of potential confounders that are not available in our data, which we again emphasize caution in too strong of a causal interpretation. However, it is the first analysis to our knowledge with nationally representative detailed microdata, which include a broad range of key covariates over 15 years of transactions. We discuss new insights drawn from this initial analysis in the results section below.

B. FHTC within a Difference-in-Differences (DiD) Framework

Prior to evaluating whether the FHTC policy had a differential impact by race, we begin by first examining its overall impact on a household's gross return to selling their home (regardless of race). As we briefly discussed earlier in the paper, the FHTC provides a quasi-experimental policy shock to the U.S. housing market, as this temporary policy contained a discrete, arbitrary cutoff for eligibility by income (and other qualifications discussed in more detail below). Conceptually, this allows us to compare returns of those eligible as the "treatment group" to those who were *never* eligible as the "control group" before and after the policy was implemented. As an alternative control group, we also compare the treatment group to those who fell just outside the threshold, but would *eventually* be treated in a later iteration of the policy. During the initial phase of the policy, before it was known that the policy would be expanded to include these households, this group represents a plausible alternative counterfactual.

For this analysis, we adapt the multivariate regression from the prior subsection to include the following difference-in-differences setup for ACS respondents who both purchased and sold a home over our broader sample period (2000-2015):

$$\mathrm{ror}_{ht} = \beta_0 + \beta_1\, Post + \beta_2\, Eligible + \beta_3\, (Post * Eligible) + \sum \beta\, race_h + \sum \gamma\, Household\ Demographics_h + \sum \delta\, Property\ Characteristics_h + \propto_h^{location} + \propto_t^{year} + \varepsilon \quad (2)$$

where *Post* for our default specification includes households *h* who purchased a home between January 1, 2009 and June 30, 2010, when the tax credit did not have to be paid back;[34] the *Treat* group includes those eligible for this policy by income and marital status who had a gross income less than $150,000 for joint-filers (or $75,000 for single and other filers) at the initial phase;[35] and, the remaining variables are defined as discussed in the prior subsection (with county fixed effects and standard errors clustered by county in our default specification). In this regression, β_3 is the diff-in-diff estimator on the homeowner's annualized gross rate of return, representing the marginal "post" effect of the policy treatment for the income-eligible cohort who purchased a home during the qualifying window relative to the control group's effect.

[34] The policy was passed and implemented in three phases. When the FHTC tax credit initially passed, the first phase (April 9, 2008-July 2009) included a tax credit up to $7,500 that had to be repaid over 15 years, which functioned as effectively a small interest free loan. Hembre (2018) and Berger, Turner, and Zwick (2020) found that this phase of the policy had little effect on incentivizing potential homeowners into homeownership. We exclude the period from our initial analysis where the repayable credit was the only option for new homebuyers (April 9-December 31 2008). The second and third phases made this credit effectively a grant up to $8,000 (or 10% of the purchase price), which did not have to be repaid unless the property was sold within 3 years and the seller made a capital gain on the home (but if the householder had died or sold the home within the 3 years without a capital gain, it also did not have to be repaid). The second phase was enacted on February 17, 2009 and included eligible first-time homebuyer purchases from January 1, 2009 through November 2009 (retroactively excluding first-time homebuyers from having to pay back the credit if they purchased the home after January 1, 2009), and the third phase extended this period through June 30, 2010 while expanding eligibility to higher earners.

[35] The First-time Homebuyer Tax Credit did not literally require a household to be a homebuyer for the "first time." In addition to income-eligibility requirements, it required that the homebuyer not be a homeowner in the previous three years. We discuss below the limitation of not being able to identify who is a first-time homebuyer in the data.

At first glance, the specification above may appear oversimplified, given that the FHTC policy had multiple phases and expanded eligibility. For illustration purposes, we first estimate a regression with all the moving parts, so-to-speak, which include all policy windows, all treated groups, and their corresponding interactions in an expanded form of the specification above. However, while the results are similar (as we discuss in more detail in the next section), the control group comparison is less intuitive and less clean. Hence, we opt for a more straightforward approach by separately estimating two different stratifications throughout the proceeding analysis consistent with the specification (2) above. Specifically, we separately compare the treatment group with two different control groups: 1) never eligible (income > \$225K for married households, >\$125K for single/other filers), or 2) not initially eligible (married households making between \$150-225K, singles making between 75-125K) for the full credit. In the first specification (i.e., "never eligible" control), we drop all observations from the second control group, whereas the only control group comparison is among homebuyers who would not have had an income eligible for the credit in any phase of the policy.[36]

One concern with this "never eligible" control group is that married households making more than \$225K, for example, might be different than those making under \$150K (in ways unobservable to us even after accounting for a rich set of household controls). Therefore, these groups may have fundamentally different experiences in the housing market and thus trend differently prior to the policy. While we examine common trends below to explore this possibility, we also consider a second, alternative control group that is more comparable in terms of income.

[36] There is also a phase-out portion of the credit, where households just outside the eligible cutoff can receive a partial credit. For simplicity, since these households do not receive the full credit, these households are lump them in with the control. We acknowledge that this may attenuate differences between the treatment and control group with part of the control group receiving a partial "dosage" of the treatment. However, this is not an issue with how we define the "never eligible" control cohort; and, our results are generally similar however we define the partial credit group.

In the second specification ("not initially eligible"), we compare the treated to those in the income range just over the initial treatment cutoffs (but under the "never eligible" cutoffs), excluding observations in the first ("never eligible") control group. Since this alternative control group would eventually become eligible, we exclude all observations during the policy window in which the second group became eligible, shortening the *Post* period to be from January through November 2009 in variation of equation (2) above. If results stemming from the comparisons of the treatment to these control groups are similar, it may be indirect evidence that the unobservables associated with income (and potential differential trends) are not driving the underlying treatment effect. We explore other possibilities below.

Finally, we follow Hembre (2018) and Berger, Turner, and Zwick (2020) by focusing on the policy period that actually motivated marginal homebuyers with the credit that did not have to be paid back. Thus, in most specifications, we limit the default sample to homes that were initially purchased from January 1, 2007 through July 31, 2010, dropping those purchased in the repayment period (April 9, 2008 through December 31, 2008) so that we are comparing returns on home purchases just prior to the policy window to those during the window. We vary the size of these windows in the Appendix, which we discuss in the results section. To clarify all of the above, we illustrate the research design and corresponding diff-in-diff comparisons in Figure 1 below.[37]

C. Parallel Trends, Data Limitations, and Interpretations from this Approach

As we alluded to above, one concern about comparing the cohort directly affected by the policy (i.e., within the income-eligibility threshold) with that not eligible is that these groups may have empirically relevant differences, even beyond the household demographic factors for which

[37] While some in the second group were eligible for a partial credit (a linear $20,000 income phase-out) in the initial phase, we exclude them in the first specification so the comparison is a cleaner "full eligible" versus "never eligible" interpretation.

we control in the regression. For example, there may be differences in wealth (rather than income), tastes in homes/amenities, or other unobservable factors that could also lead to differences in gross returns to homeownership across these groups. However, for the interpretation to be causal, the diff-in-diff approach does not require treatment and control groups to be identical across observables or unobservables; rather, a key identifying assumption is that treatment and control group exhibit parallel trends (or these factors merely *shift* the comparison group and remain relatively constant over time). We examine this assumption directly by plotting the annualized rate of return on a home purchased over time for the treatment group and both control groups.

Figure 2 shows monthly averages of annualized ROR for homes purchased either a few years prior to the FHTC policy (since 2005 – left column) or nearly a year and a half prior to the FHTC policy (since 2007 – right column). The treatment group is depicted as the orange time series and the control groups are shown in blue (where "not initially eligible" is in the top row and the "never eligible" group is in the bottom row). Overall, Figure 2 illustrates both the treatment group and control groups following similar trends (since 2005 and since 2007). As the boom and bust in the U.S. housing market occurred around this time, households who purchased a home during this infamous period generally had a negative annualized ROR, which dipped more negative around the peak of the housing bubble in 2007. Homebuyers in both groups who purchased during the subsequent downturn had returns moving in the positive direction prior to the FHTC policy. While the treated group's time series is clearly shifted lower than both control groups, it follows a similar trend over time. This is easier to see when we consider linear trend lines in both of the panels in the right column, showing an approximately parallel path.

Of course, an "eye test" of parallel trends may not be sufficient, motiving a more formal statistical test of common trends. Specifically, we test whether these trends are statistically different from one another in untabulated regressions,[38] where we regress these trends on the dependent variable (monthly average annualized ROR) along with group identifiers and an interaction term. While there may be slight visible differences in trends (perhaps due to the scale of the graph), results show no statistically significant difference in the slope of these trends in any of these time series. Thus, even before we incorporate relevant controls, the raw (monthly) annualized return data provides initial, direct evidence that the control groups, while not identical, serve as reasonable counterfactuals for the treated group.[39]

As we discussed in more depth in the data section above, one contribution of this study is the data. The data allows us to control for a variety of relevant household characteristics (like income, education, marital status, size of household, age, and race), which could contribute to a variety of differences (either directly or indirectly) in a household's return to housing, ultimately reducing model specification error. However, our data does not indicate whether the household had owned a home within the prior 3 years; so, the treatment group is determined by income eligibility.[40] Unlike an average treatment effect on the treated (ATT) estimator or the local average

---

[38] This untabulated set of tests was approved by Census's Disclosure Review Board for the general public, but is omitted here for brevity and is available upon request.

[39] We should also note that the post-2007 trend is more linear than the post-2005 time series, which is one advantage of the former being a better candidate for the pre-period in our default regression. For robustness, we also vary which time period to use in our analysis, which we discuss in more detail later in the paper.

[40] We use household income reported in the ACS, not taxable income reported to the IRS, which exposes the estimates from this data to measurement error in two ways. First, there are documented differences between survey reported income and income claimed on taxes (Bee and Rothbaum 2019), however, AGI and total money income are correlated 99% (Gindelsky 2016) and measurement error at the median is classical (Bee and Rothbaum 2019)..Second, the survey year does not necessarily coincide with the year of the home purchase, as the survey must have occurred *after* the initial purchase but *before* the eventual sale. We set the year of the home's purchase as our base year and adjust income based on PCE inflation (or average income growth for taxpayers in a similar strata of income during that period). For example, for an individual surveyed in 2010 that had their home purchase in 2009, we would adjust their income for inflation (or the average growth in that strata) to 2009 dollars to represent income relevant for assessing the credit threshold limits in the year of purchase. On average, nominal income growth was slow during the period we analyze (e.g., nominal median household personal income grew 7% from 2007-2010 (BEA Distributional Accounts), but for an individual this might not be the case, as householders may have experienced a job change or promotion/demotion in a year that does not align with the year of the home purchase. This should be the case for both the treatment and control group, and thus the measurement error likely yield noisier results than if we had more precise income data.

treatment effect (LATE), the interpretation of the treatment in our case is subtly different than a scenario where we knew with certainty that the individual household qualified *and* took-up the credit. Instead, the interpretation of this treatment is simply that they belong to the income-eligible cohort, whether or not they actually claimed the credit.[41]

One benefit of this approach is that the interpretation of the treatment group speaks to the more general policy objective concerning the effect on the broader group. That is, the broader measure can be interpreted as an estimate of the effect for entire income cohort, whereas the policy eligibility and take-up is a narrower subset of that group.[42] On one hand, it is possible that the benefits/costs of those who took-up the credit were broadly offset by those who could not, where a null effect may represent this zero-sum outcome. After all, there are some zero-sum aspects of the housing market, where one winning bid on a home represents another potential buyer's losing bid (and thus buying elsewhere). On the other hand, an observed positive effect could represent an additive combination of: 1) a positive impact on those who received the credit, 2) a negative impact on those who did not receive the credit, but not enough to offset the positive impact on the former subset, and/or 3) a null or positive impact on those who did not receive the credit. While our measure of the treatment group does not disentangle these effects individually, it does help us answer a variation of the core question raised earlier in the paper: does a policy targeted at a subset of low/medium income homebuyers help the overall cohort build (gross) home equity at a clip outpacing the ineligible higher-income homebuyer? Thus, for home equity to ameliorate the wealth

[41] We also use age of the householder in our default specifications, which we use to exclude those from the sample who are least likely to be eligible, approximately 80% of those 60 are already homeowners.
[42] For example, an analysis of a change in the earned income tax credit (EITC) or minimum wage policy may not study the policy's effect exclusively on those who took up the credit or made minimum wage, rather it may include all low wage earners regardless of whether the credit or wage was binding.

gap among this cohort of earners, a broader measure of the policy effect on the low/middle-income cohort may be more suited to this purpose.

D. FHTC and Differential Returns by Race

To investigate whether the FHTC policy has a differential effect by race, we modify the specification from equation (2) to estimate potential heterogeneity in the policy effect on rate of return by race:

$$\mathrm{ror}_{ht} = \beta_0 + \beta_1\, Post + \beta_2\, Eligible + \beta_3\, (Eligible * Post) + \beta_4\, (Race * Post) + \beta_5\, (Race * Eligible) + \beta_6\, (Race * Eligible * Post) + \sum \gamma\, Household\ Demographics_h + \sum \delta\, Property\ Characteristics_h + \propto_h^{location} + \propto_t^{year} + \varepsilon \quad (3)$$

where the variables are defined the same as above, but we separately interact Race (i.e., a binary indicator for either nonwhite, Black, Hispanic, or Asian) with our diff-in-diff parameters. In the specifications above we omit other racial/ethnic minority groups from the sample, separately comparing the estimated effect against Whites as the control group. Therefore, in this specification, β_6 is the diff-in-diff-in-diff (triple diff) estimator of interest, which estimates the relative rate of return on a home for a given racial category and purchased a home during the policy treatment window. If these interactions are statistically significant and positive (negative), then the minority group in the treated income cohort has a higher (lower) return than comparable White households who have the same eligibility criteria and purchased a home during the policy window. If the β_6 coefficient is null and/or not statistically significant, then we interpret the policy as having a homogenous impact on a given racial minority household. In the context of the racial wealth gap, we interpret a positive and significant interaction coefficients as potentially narrowing this wealth gap in gross terms by yielding a relatively higher nominal return on this particular asset.[43]

One objection to this setup might be that our treatment group in both sets of analyses above (i.e., overall or by race) is unbounded and includes the lowest income households. From an empirical standpoint, this may be too coarse of a grouping, particularly if we are comparing these treatment and control groups across race, where income differences (and distributional differences) are well-documented (see Semega and Kollar (2022) for the most recent Census estimates). We thus limit the sample by trimming the lowest income households (under 75K for married households; under 37.5K for single). For example, for married households, we are comparing those making \$75K-150K to those making \$150K-225K in our "not initially eligible" specification. In the Appendix, we show the full results without making this cut, which illustrates the predicted attenuation bias.

From a parallel trends standpoint, we investigate whether this is a more apples-to-apples comparison in Figure 3 below. As with our overall sample, a key identifying assumption for each minority subgroup is whether the control groups serve as reasonable counterfactuals by trending similarly prior to the policy shock. In Figure 3, we now report the monthly time series for annualized ROR for this more limited sample that excludes lower income households. Figure 3 separately illustrates a time series for all races (labeled "Overall" in the top-left panel) and for White (top-right), Black (bottom-left), or Hispanic (bottom-right) subsamples, respectively.[44] A

[43] As we discussed above, perhaps not on net, if there are asymmetries in home improvement investments, transactions costs, or other costs that outweigh the gross return differences. While we have not yet seen evidence to suggest that the asymmetries would be large enough to substantially offset the effect sizes we observe, we leave it to future work to estimate racial differences in these costs.
[44] We exclude Asian and Other races in this analysis because we do not find significant differences in our diff-in-diff-in-diff analysis for these subsamples. We discuss further in the results section.

key takeaway from all four figures is that the time series are less "shifted," as they are more or less on top of one another. By this simple change to the sample, the raw monthly data generally appear to the naked eye as trending similarly. The White and Black subsamples fluctuate somewhat noisily around their linear trend lines, which appear to have slight differences in trends; but, like our analysis above, we also test whether these trends are statistically distinguishable from one another. In untabulated tests, we find that none of these trends of the treatment group are statistically significantly different from their corresponding control group at any conventional threshold. Taken together, the evidence suggests that we have reasonable treatment and control groups for our difference-in-differences analysis, both overall and by race.

5. Results

A. Baseline Analysis of Gross Returns to Homeownership – 2000 to 2016

In our initial analysis, we find gross returns to homeownership vary significantly across race and demographic characteristics for households who transacted single-family homes over the 2000-2016 sample period. Table 3 provides the results of estimating equation (1) in a number of different ways. Specifically, we estimate equation (1) on ROR, IAROR, and annualized ROR in columns (1) through (3), respectively. For the rest of the analysis, we choose annualized ROR, and in columns (3) through (6), we vary geographic fixed-effects from state-by-year (3), to county (4), to census tract (5), to census block groups (6) (with survey year fixed effects not interacted in the latter specifications). In the final set of analysis, we narrow the period of analysis to exclude homes in the early part of the decade (prior to the peak years of the boom).

Overall, the coefficient estimates for race on returns tabulated in Table 3 tell a similar story as the unconditioned means from Table 2, Panel B, albeit with different magnitudes and some nuanced caveats. Columns (1) and (2) of Table 3 show that nonwhite households (Black, Hispanic, and Other) have significantly lower ROR and IAROR relative to White householders, even after we account for a number of key socioeconomic household characteristics and property characteristics. The exception is that Asian households did not experience a meaningfully different return (where the results are generally not statistically significant, with the exception of columns 4 and 5 where there is an economically small difference, but it is not robust to using alternative fixed effects). While we control for how long the household owned the home, for the remainder of the analysis we look at annualized ROR for ease of comparisons to the broader literature of returns on assets.

When we compared the raw means, recall that there was about a 3 percentage point difference between the mean annualized ROR when we compared White to Black homeowners. The conditioned estimates from columns (3) through (6) suggest that Black householders have a 1.2 to 2 percentage point lower annualized ROR than otherwise similar (across the observable variables in our dataset) White households. Hispanic and Other householders have a *less bad* return of about 1 percentage point lower than similar White householders' annualized returns.

Importantly, the results from Table 3 illustrate that the unconditioned mean differences are not primarily driven by differences in observables *or* location. Census tracts and block groups are relatively small geographic units; so, while we lose some data due to dropping singleton observations within these geographies in these specifications, the "within geography" interpretation here is that these differences persist even after accounting for "neighborhood differences" in returns.[45] Meyers (2004) found that racial differences in home prices depended on

[45] Tracts roughly approximate the size of what one might think of as a broader neighborhood, and block groups are subsets of that, although the exact geographic size varies depending on the population density of the area. Tracts, block groups, and blocks are commonly used in hedonic real estate literature to account for location-specific (dis)amenities and neighborhood heterogeneity (see, for example, Bian et al. 2021, Turnbull et al. 2019, Brastow et

the level of location controls used. In our sample, accounting for finer location differences are more important for Black returns in terms of economic magnitude when going from county (-1.76%) to block group (-1.26%) fixed effects, as compared to Hispanic or Other, which do not change much.[46] Overall, the boom-bust-recovery period of 2000-2016 did not produce great returns to housing, on average, but it was particularly lackluster for nonwhite (and non-Asian) racial minority householders who experienced significantly lower returns over this period.

Other demographics and property characteristics also had statistically significant coefficient estimates on housing returns. Demographically, the results from Table 3 suggest that if the householder was married, younger (<=30), college educated, higher income, or had a smaller number of people living in the household, then they would have had significantly higher rates of return. Properties that are larger (i.e., more bedrooms, larger lot size), newer (built 2011+), have higher value to the survey respondent,[47] and do not have a mortgage, all have significantly higher returns.

B. FHTC Analysis – Overall Diff-in-Diff Results

Before we discuss the DiD regression results, it is useful to highlight some differences and similarities of the treatment and control groups tabulated in Table 4. The unconditioned, raw means of the annualized rate of return for the control groups and the treatment group are similar for the

al. 2018, or Moulton et al. 2018). For more information about the size, scope, and construction of these geographic units, see: https://www2.census.gov/geo/pdfs/reference/GARM/Ch11GARM.pdf

[46] We lose nearly 20% of the sample observations when we incorporate block group fixed effects, which is an inherent tradeoff when incorporating finer fixed effects and requiring non-singleton fixed effects. We are reluctant to over-interpret magnitude differences across specifications with different samples and potential over-fitting issues with finer fixed effects. However, given the sample loss, we chose county fixed effects for the remainder of the analysis to balance these tradeoffs.

[47] These value categories are quite broad, so we are not particularly worried about a mechanical relationship with ROR. While the results are similar when we omit the value categories, one reason to have this variable in the regression is that, controlling for the other characteristics, a higher value category may be capturing some combination of unobserved quality the householder's willingness to accept (WTA) threshold when they ultimately sell the home.

overall sample (*note*: this does not compare the Post period to the Pre period). Other variables like year built and the number of years the house was owned are similar across groups, too. However, there are significant differences across many of the remaining variables. For example, there are moderate differences in household size, number of bedrooms, square footage, and age of the householder, while there are large difference in marriage rates, income, and home value. These differences provide motivation for why we control for these variables in the multivariate regression, so the DiD estimator can be interpreted as a change in β while holding these factors constant. As we discussed in the methodology section above, there may be concern about unobservables associated with some of these differences (e.g., income). We return to this point in the next subsection when we conduct additional analysis by excluding low income homebuyers from the sample to maintain more comparable treatment and control subsamples.

Table 5 reports the DiD estimates from the specification laid out in equation (2) above in columns (2) through (5), where the initial column (1) draws from a broader sample and breaks out the policy period into four different windows. The initial column includes homes purchased in 2005 and after, which is a sample equal to the final regression in Table 3. Effectively, we have a DiD estimator for each policy period (A through D), which is an interaction between each respective Post variable and the treatment eligibility variable. We find that the initial period (A – April 2008 through December 2008), where the tax credit had to be paid back in subsequent periods, had no statistically significant effect on the annualized rate of return. The next three phases have statistically significant effects (or, at least marginally significant with a $p<0.10$ in the case of Post C interaction), which we subsequently collapse into a single policy phase for the remainder of the analysis. Note that we have an interaction for the "not initially eligible" group during the window they were eligible (Post D – November 2009 through June 2010), which is statistically significant relative to the control group ("never eligible"). As we discussed in the methodology section, having both control groups in this regression will likely create some confusion about the treatment and control group comparisons.

To simplify, we separately estimate the DiD relative to each control group in the remaining columns in Table 5, and we cull the sample to include the most relevant observations for the research design. Specifically, in columns (2) and (3) we dropped all observations from the Post A period (where the credit had to be repaid), all homes purchased for over \$800K (since they would later be ineligible for the credit), household incomes of greater than \$500K, and other ineligible sales (i.e., when the home was sold within 3 years of the initial purchase and the household main a capital gain). We initially shorten up the pre-period to compare only transactions beginning in April 2007, although we alter this date in later tables and the Appendix for robustness. We end the sample with homes initially purchased by the end of June 2010. In columns (4) and (5), we have all of these restrictions and we further restrict the sample to householders under 60 years old, which is our default set of sample restrictions going forward. Finally, in the Control 1 columns, we exclude all Control 2 observations, and vice versa for Control 2. We have the additional restriction in the Control 2 sample that we omit control observations during the policy window in which they were eligible for the credit. Across all four columns, regardless of the varying sample restrictions, we find that income-eligible households who purchased a home during the main FHTC credit policy window (January 2009 through June 2010) had about a 1% higher annualized return on their home compared to either control group.

In the Appendix, we alter these sample restrictions and specifications for robustness, finding broadly similar results for the policy effect, albeit with slightly different magnitudes for the DiD estimator. In Appendix Table A1, we alter the pre-period length, allowing purchases in the pre-period to go back through 2005 or back to January 2007, where in the latter case the pre-period is approximately the same as the post-policy implementation period. The DiD estimator remains statistically significant, and if anything, modestly larger in Table A1. Next, we examine the robustness of the results when we alter the geographic fixed effects in Table A2. We find the results are qualitatively similar when we instead include: county-by-year effects, zip code, census tract, and block group fixed effects, albeit somewhat larger in magnitude in the block group specification (as high as 1.6%).

C. FHTC Analysis – Diff-in-Diff Results by Race

In our final set of analysis, we consider differential returns by race for those eligible (by income) for the FHTC during the policy window. In Table 6, we estimate both control samples with the same restrictions as the previous table, but instead we estimate an interacted DiD model as shown in equation (3) where each race indicator is interacted with the corresponding DiD design variables. All columns have White householders as the reference group (excluding all other race categories). Out of concern for thin cells, we do not include the Other race interaction and we drop these households from all samples, too.

As we noted in the prior section, we exclude all households with an adjusted household income less than \$75K for married households (\$37.5K for single) or a more apples-to-apples comparison of the treatment and control group. We find the main result for the effect of the FHTC is virtually unchanged when we exclude low income households from the sample in column (1). However, the results from Table 6 show the nonwhite, Black, and Hispanic interacted DiD estimators are statistically significant and positive in columns (2), (3), and (4), respectively. The evidence using these more comparable treatment and control groups (along the dimension of income) suggests that income-eligible Nonwhite, Black, and Hispanic households had realized

2.75, 5.73, and 4 percentage points higher annualized return from purchasing a home during the tax credit period compared to the control group. When we use the broader sample that includes the bottom end of the income distribution in the Appendix Table A3, there are not significant differences in the DiD estimator in the interaction effects, consistent with the attenuation bias described above. In the Table A3 specifications, nonwhite households had returns that were not statistically different from the returns of White households for income-eligible groups purchasing a home during the policy window (relative to either control group).[48] Overall, the evidence from Table 6, which compares more similar treatment and control groups, suggests that not only did all races realize a higher gross annualized return to housing; but, minority homeowners benefited more, likely supporting a reduction in the racial wealth gap.

6. Conclusion

The collective experience in the U.S. housing market during the boom-bust period in the first decade of the 21st century led many Americans to sour on the prospect of housing as a vehicle to building wealth. In their assessment of the recent literature and trends in the U.S. housing market, Goodman and Mayer (2018) concluded that, "while two decades of policies in the 1990s and early 2000s may have put too much faith in the benefits of homeownership, the pendulum seems to have swung too far the other way, and many now may have too little faith in homeownership" (p. 32-33). Taken together, the results from both sets of analyses in our paper

[48] This could reflect a number of other attenuating possibilities correlated with the bottom end of the income distribution. For example, prior to this policy, there have been a variety of low income programs, including some state-administered FHTC program for those at much lower income thresholds than the 2008-2010 policies. One such policy, the American Dream Downpayment Assistance Act of 2003, was a limited federal program administered by states that required FHTCs to have an income lower than 80% of their local's median income. The thresholds for the policy we examine exceed the median income (about $52K in 2008) of the U.S. generally as well as the vast majority of locales in the U.S. during this period.

support a more balanced, nuanced view of the housing market by knocking down strawmen on both ends of the pendulum.

On one end of the pendulum, the results from our initial descriptive analysis of the broader sample period (containing sales from 2000 to 2016) stand in stark contrast to the notion that homeownership is a universal panacea for building wealth. Using a unique sample of internal Census records linked with property-level data from Zillow's ZTRAX, the evidence from this rich data overwhelmingly suggests that the benefits of homeownership were not universally shared across households. Specifically, we found Black and Hispanic households experienced significantly lower rates of returns to buying and selling single-family homes (relative to otherwise White households with similar observables). And, this return may have even been negative in nominal terms, on average, for Black households. While our initial descriptive analysis is not causal in nature, these results would be difficult to reconcile with a view that housing is always a vehicle for building wealth, and one that would thus always reduce the racial wealth gap. Strawmen notwithstanding, the more practical takeaway from the initial descriptive analysis of differential returns to housing by race is that the magnitude of these differences can be quite large and meaningful, motivating future work on exploring the implications of differential returns to housing for wealth inequality and housing sector dynamics more generally. For example, non-financial returns to homeownership may provide offsetting benefits (or exacerbating costs) over this period, which could be weighed against financial returns to homeownership when evaluating the net impact on wealth accumulation.

On the other end of the pendulum, the results from our second set of analysis provide evidence that not all homeownership policies of that era were disastrous for wealth accumulation among the low to middle end of the income distribution. Far from it – the evidence from our difference-in-differences analysis of the FHTC shows that the income-eligible cohort realized higher gross rates of returns to housing than non-eligible, high-income cohorts. More strikingly, when we compared cohorts of more similar income just above and below the eligibility threshold, Black and Hispanic households realized substantially higher returns from this policy than White householders. Provided that other factors do not swamp these gross returns, the evidence from our analysis of the FHTC suggests that expanding homeownership among eligible, marginal homeowners may have helped to reduce the racial wealth gap for low to medium income households during this era.

References

Akbar, P., Li, S., Shertzer, A. and R.P. Walsh. 2022. "Racial Segregation in Housing Markets and the Erosion of Black Wealth." *The Review of Economics and Statistics*. pp.1-45. Nov. 15.

Altonji, J., Doraszelski, U. and L. Segal. 2000. "Black/White Differences in Wealth", *Economic Perspectives*, Federal Reserve Bank of Chicago. Winter.

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

Avenancio-León, C.F. and Howard, T., 2022. "The assessment gap: Racial inequalities in property taxation." *Quarterly Journal of Economics*, 137(3), pp.1383-1434.

Barsky, R., Bound, J., Charles KK., and J.P. Lupton. 2002. "Accounting for the black-white wealth gap: a nonparametric approach." *Journal of the American Statistical Association*. Vo. 97. No. 459. pp. 663-673.

Bartlett, R., Morse, A., Stanton, R. and Wallace, N. 2022. "Consumer-lending discrimination in the FinTech era." *Journal of Financial Economics*, 143(1), pp.30-56.

Bayer, P. and McMillan, R., 2012. Tiebout sorting and neighborhood stratification. *Journal of Public Economics*, 96(11-12), pp.1129-1143.

Bayer, P., Casey, M., Ferreria, F. and R. McMillan. 2017. "Racial and Ethnic Price Differentials in the Housing Market." *Journal of Urban Economics*, 102, pp. 91-105.

Bayer, P., Ferreira, F. and S.L. Ross, 2018. "What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders." *Review of Financial Studies*, *31*(1), pp.175-205.

Bee, A. and J. Rothbaum 2019. "The Administrative Income Statistics (AIS) Project: Research on the Use of Administrative Records to Improve Income and Resource Estimates." U.S. Census Bureau SEHSD Working Paper. Vol. 36

Belsky, E. and M. Duda. 2002. <u>Low-income Homeownership: Examining the Unexamined Goal</u>. Joint Center for Housing Studies. Brookings Institution.

Berger, D., Turner, N. and E. Zwick. 2020. "Stimulating Housing Markets". *Journal of Finance*. Vol. 75, No. 1, pp. 277-321.

Bhutta, N., Chang, A., Dettling, L., and J. Hsu. 2020. "Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances." FEDS Notes, September 28.

Bhutta, N. and A. Hizmo. 2021. "Do minorities pay more for mortgages?". *Review of Financial Studies*, *34*(2), pp.763-789.

Bian, X., Contat, J.C., Waller, B.D. and S.A. Wentland. 2021. "Why disclose less information? Toward resolving a disclosure puzzle in the housing market." *The Journal of Real Estate Finance and Economics*, pp.1-44.

Blau, F.D. and J.W. Graham. 1990. Black-white differences in wealth and asset composition. *Quarterly Journal of Economics*, 105(2), pp.321-339.

Brastow, R.T., Waller, B.D. and S.A. Wentland. 2018. "Temporally dynamic externalities and real estate liquidity." *Journal of Real Estate Research*, 40(2), pp.199-240.

Boustan, L.P., 2012. "Racial Residential Segregation in American Cities." In *The Oxford Handbook of Urban Economics and Planning*. Oxford University Press.

Carrillo, P. and A. Yezer, 2009. "Alternative measures of homeownership gaps across segregated neighborhoods." *Regional Science and Urban Economics*, 39(5), pp.542-552.

Chambers, D.N. 1992. "The racial housing price differential and racially transitional neighborhoods." *Journal of Urban Economics*, 32(2), pp. 214-232.

Charles, KK. And E. Hurst. 2002. “The Transition to Home Ownership and the Black-white Wealth Gap”. *The Review of Economics and Statistics*. 84(2). pp. 281-297.

Chen, J., Cornwall, G. and S.A. Wentland. 2022. “It’s the Smell: How Resolving Uncertainty about Local Disamenties Affects the Housing Market.” Working Paper. Available at SSRN 4123386.

Cheng, P., Lin, Z. and Y. Liu. 2015. “Racial discrepancy in mortgage interest rates.” *Journal of Real Estate Finance and Economics*, 51(1), pp.101-120.

Christensen, P. and C. Timmins. 2018. “Sorting or steering: Experimental evidence on the economic effects of housing discrimination.” NBER Working Paper.

Collins, J. and R. Margo. 2011. “Race and Home Ownership from the End of the Civil War to the Present.” *American Economic Review*. 101(3), pp. 355-359. May.

Comeau, J., Padmasini, R., William, R., and E.J. Szymanoski. 2012. “The FHA Single-Family Insurance Program: Performing a Needed Role in the Housing Finance Market.” HUD Office of Policy Development and Research. Working Paper. December.

Coulson, N.E., 1999. “Why are Hispanic-and Asian-American homeownership rates so low?: Immigration and other factors.” *Journal of Urban Economics*, 45(2), pp.209-227.

Dawkins, C.J. 2005. “Racial Gaps in the transition to first-time homeownership: The role of residential location.” *Journal of Urban Economics*. vol. 58, pp. 537-554.

Deng, Y., Ross, S.L. and S.M. Wachter. 2003. “Racial differences in homeownership: the effect of residential location.” *Regional Science and Urban Economics*, 33(5), pp.517-556.

Derenoncourt, E., Kim, C.H., Kuhn, M, and M. Schularick. 2022. “Wealth of two nations: The U.S. racial wealth gap, 1860-2020.” NBER Working Paper #30101, June.

Denton, N. A. 2001. “Housing as a Means of Asset Accumulation: A Good Strategy for the Poor?” in Shapiro, T. M., and E. N. Wolff, (eds), *Assets for the Poor: The Benefits of Spreading Asset Ownership*. Russell Sage Foundation, New York.

Fagereng, A., Guiso, L., Malacrino, D., and L. Pistaferri. 2016. “Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality,” *American Economic Review*, 106(5), pp. 651-655.

Fagereng, A., Guiso, L., Malacrino, D. and L. Pistaferri. 2020. “Heterogeneity and Persistence in Returns to Wealth,” *Econometrica*, 88(1): pp. 115–170.

Foote, C.L., Gerardi, K.S., and P.S. Willen. 2012. “Why did so many people make so many ex post bad decisions? The causes of the foreclosure crisis.” Federal Reserve Bank of Atlanta. Working Paper. 2012-07.

Gabriel, S.A. and S.S. Rosenthal. 2005. “Homeownership in the 1980s and 1990s: aggregate trends and racial gaps.” *Journal of Urban Economics*, 57(1), pp.101-127.

Gerardi, K.S., Lehnert, A., and S. Sherlund. 2008. “Making Sense of the Subprime Crisis.” *Brookings Papers on Economic Activity.* Vol. 39. No. 2. pp. 69-145.

Gindelsky, M. 2016. “Will Inequality Continue to Rise? Forecasting Income Inequality in the United states.” Working Paper. IARIW 2016. Dresden, Germany.

Gindelsky, M., Moulton, J., Wentland, K., and S. Wentland. 2022. “When do Property Taxes Matter? Tax Salience and Heterogeneous Policy Effects.” Working Paper.

Gindelsky, M., Moulton, J.G. and S.A. Wentland. 2020. “Valuing housing services in the era of big data: A user cost approach leveraging Zillow microdata” (No. c14274). National Bureau of Economic Research.

Goodman, L. and C. Mayer. 2018. “Homeownership and the American Dream.” *Journal of Economic Perspectives*. 32(1). Winter.

Goodwin, K. and L. Zumpano. 2011. “The Home Buyer Tax Credit of 2009 and the Transition to Homeownership.” *Journal of Housing Research*. 20(2). pp. 211-224.

Gyourko, J., Linneman, P. and S. Wachter. 1999. “Analyzing the relationships among race, wealth, and home ownership in America.” *Journal of Housing Economics*, 8(2), pp.63-89.

Hackshaw, Alana. 2009. A Resource Guide for African Americans – The American Recovery & Reinvestment Act of 2009. Congressional Black Caucus Foundation Inc. Available at: < https://www.abfe.org/wp-content/uploads/2014/02/The-American-Recovery-and-ReinvestmentAct-of-2009.pdf>

Haurin, D. and S. Rosenthal. 2004. “The Sustainability of Homeownership: Factors affecting the Duration of Homeownership and Rental Spells.” December. Washington, D.C.: Office of Policy Development and Research. U.S. Department of Housing and Urban Development.

Haurin, D.R. and S.S. Rosenthal. 2007. The influence of household formation on homeownership rates across time and race. *Real Estate Economics*, *35*(4), pp.411-450.

Hembre, E. 2018. “An Examination of the First-Time Homebuyer Tax Credit.” *Regional Science and Urban Economics*. Vol. 73, pp. 196-216.

Herbert, C.E., and E. Belsky. 2008. “The Homeownership Experience of Low-Income and Minority Households: A Review and Synthesis of the Literature.” *Cityscape*. 10(2)

Hodge, T.R., McMillen, D.P., Sands, G. and M. Skidmore. 2017. “Assessment inequity in a declining housing market: The case of Detroit.” *Real Estate Economics*, *45*(2), pp.237-258.

Kau, J.B., Keenan, D.C. and H.J. Munneke. 2012. Racial discrimination and mortgage lending. *Journal of Real Estate Finance and Economics*, *45*(2), pp.289-304.

Kermani, A., and F. Wong. 2021. “Racial Disparities in Housing Returns.” NBER Working Paper #29306. September.

Killewald, A. 2013. “Return to Being Black, Living in the Red: a race gap in wealth that goes beyond social origins.” *Demography*. 50(4), pp. 1177-1195.

Kollmann, T. and P.V. Fishback. 2011. “The New Deal, Race, and Home Ownership in the 1920s and 1930s.” *American Economic Review*. 101(3), pp. 366-70. May.

LaCour-Little, M. and Green, R.K., 1998. “Are minorities or minority neighborhoods more likely to get low appraisals?”. *Journal of Real Estate Finance and Economics*, 16(3), pp.301-315.

Ladd, H.F., 1998. “Evidence on discrimination in mortgage lending.” *Journal of Economic Perspectives*, 12(2), pp.41-62.

Mayer, C., Pence, K., and S.M. Sherlund. 2009. “The Rise in Mortgage Defaults.” *Journal of Economic Perspectives.* 23(1). pp. 27-50.

McKernan, S.M., Ratcliffe, C., Simms M., and S. Zhang. 2014. “Do Racial Disparities in Private Transfers Help Explain the Racial Wealth Gap? New Evidence from Longitudinal Data.” *Demography*.51(3). pp. 949-974.

McMillen, D.P. and R.N. Weber. 2008. “Thin markets and property tax inequities: A multinomial logit approach.” *National Tax Journal*, 61(4), pp.653-671.

Mian, A. and A. Sufi. 2009. “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis.” *Quarterly Journal of Economics*. 24(4). November.

Mian, Atif and Amir Sufi, 2011. "House Prices, Home Equity Based Borrowing, and the U.S. Household Leverage Crisis," *American Economic Review* 101: 2132-2156.

Moulton, J.G., Waller, B.D. and S.A. Wentland. 2018. “Who benefits from targeted property tax relief? Evidence from Virginia elections.” *Journal of Policy Analysis and Management,* 37(2), pp.240-264.

Myers, C. 2004. "Discrimination and neighborhood effects: understanding racial differentials in US housing prices." *Journal of Urban Economics*. vol. 56, pp. 279-302.

Nolte, C., Boyle, K.J., Chaudhry, A.M., Clapp, C.M., Guignet, D., Hennighausen, H., Kushner, I., Liao, Y., Mamun, S., Pollack, A. and J. Richardson. 2021. "Studying the impacts of environmental amenities and hazards with nationwide property data: best data practices for interpretable and reproducible analyses." *Available at SSRN*.

Piketty, T, Saez, E., and G. Zucman. 2018 "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics*. Vol. 162. pp. 89-100.

Rothstein, R. 2017. The Color of Law: A Forgotten History of How Our Government Segregated America. Liveright Publishing: New York.

Shertzer, A. and R.P. Walsh. 2019. "Racial sorting and the emergence of segregation in American cities." *Review of Economics and Statistics*, 101(3), pp.415-427.

Semega, J. and M. Kollar. 2020. "Income in the United States: 2021." Census Bureau. Report No. P60-276. September.

Turnbull, G.K., Waller, B.D., Wentland, S.A., Witschey, W.R. and V. Zahirovic-Herbert. 2019. "This old house: historical restoration as a neighborhood amenity." *Land Economics*, 95(2), pp.193-210.

Wolff, E. 2017. "Household Wealth Trends in the United States, 1962 to 2016: Has Middle Class Wealth Recovered?" NBER Working Paper #24085. November.

Wolff, E. 2022. "African-American and Hispanic Income, Wealth, and Homeownership since 1989." *Review of Income and Wealth*. 68(1). pp. 189-233.

Xavier, I. 2021. "Wealth Inequality in the U.S.: The Role of Heterogenous Returns." SSRN #3915439. Working Paper.

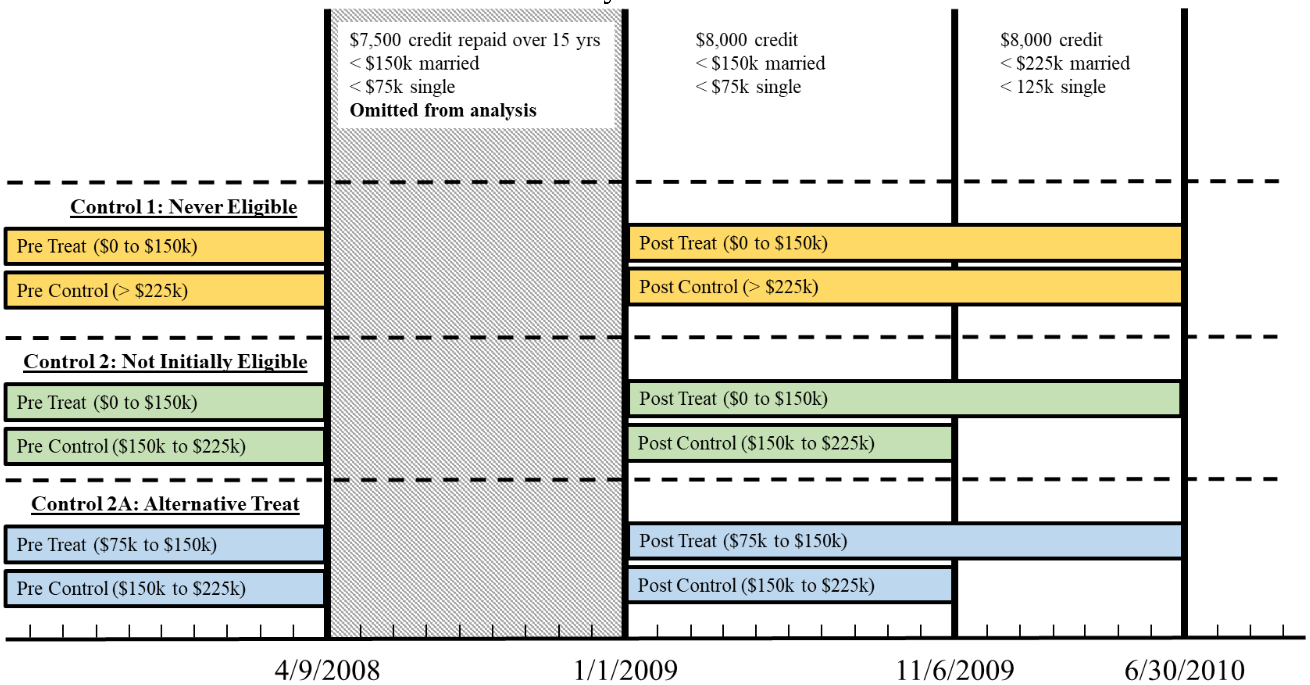
U.S. Bureau of Economic Analysis, Shares of gross domestic product: Gross private domestic investment: Fixed investment: Residential [A011RE1Q156NBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/A011RE1Q156NBEA, November 3, 2022a.

U.S. Bureau of Economic Analysis, Personal Consumption Expenditures by Type of Product: Services: Household Consumption Expenditures: Housing [DHSGRC0], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DHSGRC0, November 3, 2022b.

Zhang, D.H. and Willen, P.S., 2021. "Do Lenders Still Discriminate? A Robust Approach for Assessing Differences in Menus" (No. w29142). National Bureau of Economic Research Working Paper.

Figures

Figure 1 – Treatment and Control Groups for a Diff-in-Diff Analysis of the First-time Homebuyer Credit



Treat and Control shown for married. Single groups can be determined using single income thresholds.

Figure 2 – Difference-in-differences Pre-Trend Analysis for the First-time Homebuyer Credit

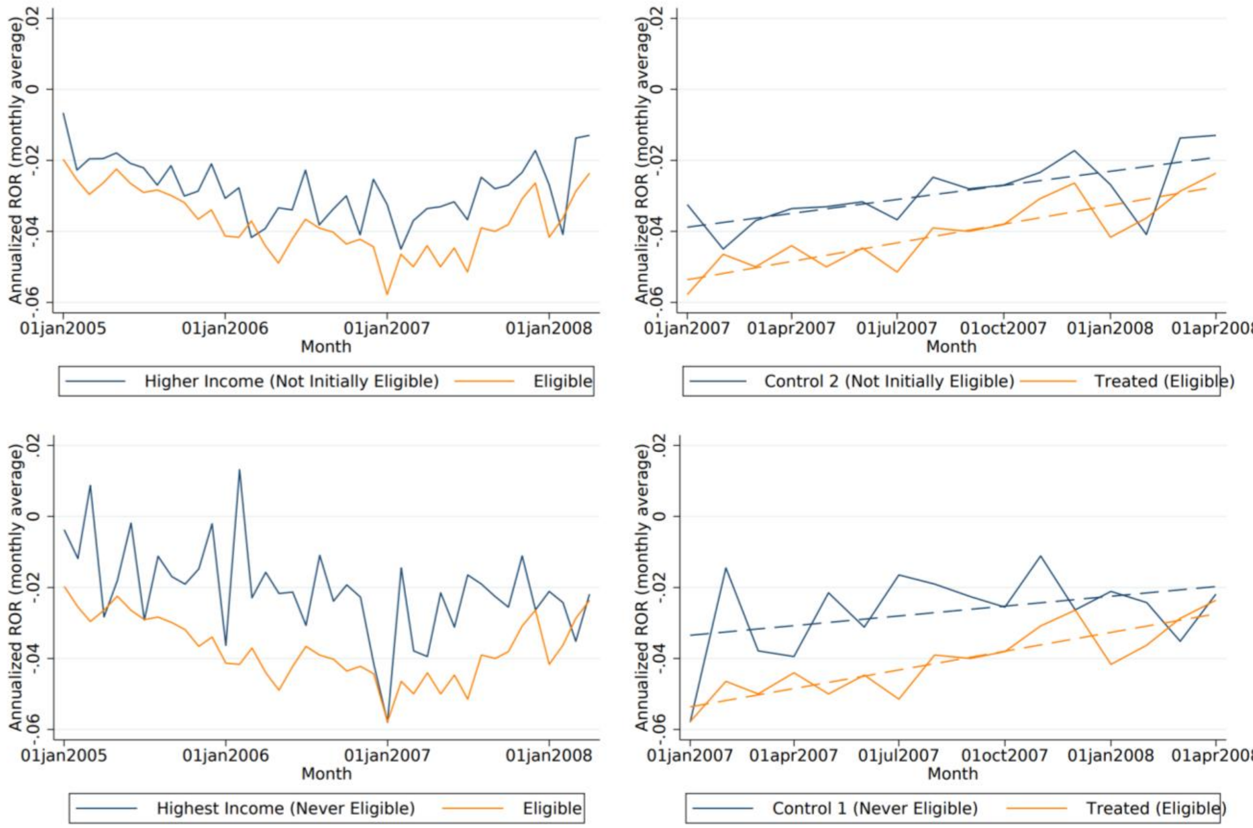


Figure 2 presents trends in annualized rates of return for households that would meet income threshold eligibility criteria for the First-Time Homebuyer Credit (eligible households) relative to alternative control groups for our difference-in-differences analysis in the periods leading up to credit policy period.

Figure 3 – Diff-in-Diff Pre-Trend Analysis for the First-time Homebuyer Credit by Race

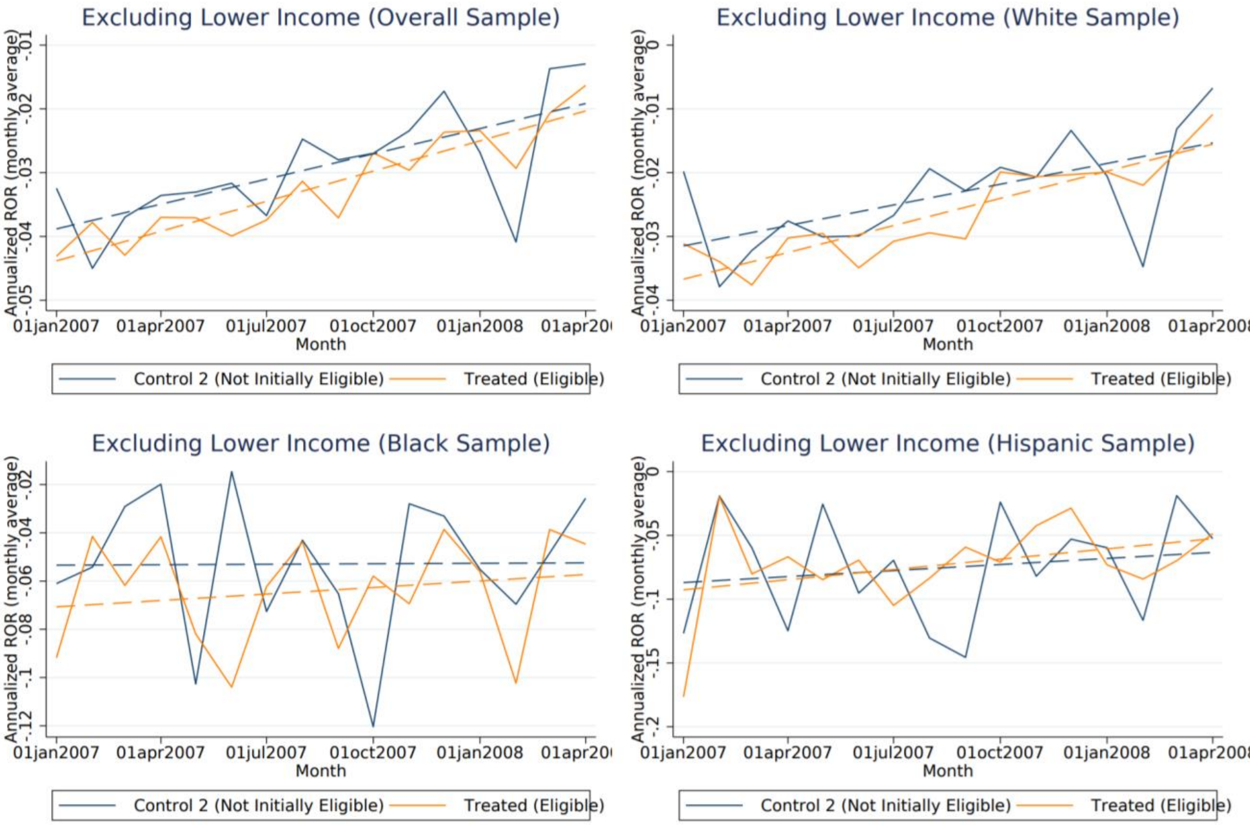


Figure 3 reports trends by race in annualized rates of return for households that would meet income threshold eligibility criteria for the First-Time Homebuyer Credit (eligible households) relative to those that would not initially meet eligibility thresholds in the periods leading up to credit policy period.

Tables

Table 1

Variable Descriptions

Variable	Description
Dependent Variables	
Rate of Return (ROR)	(Sale Price – Buy Price)/Buy Price
Inflation-Adjusted ROR	ROR deflated by the Personal Consumption Expenditures (PCE) Index
Annualized ROR	(Sale Price/Buy Price)^(1/(Sale Year-Buy Year)) -1
Treatment Variables	
Eligible	Borrower is eligible for the FHTC based on adjusted household income (Dummy variable, Yes=1). Included as dummy and interacted with the credit window variables.
Credit Window	Period of time for which credit could be claimed (Dummy variables, Yes=1). Included as dummy and interacted with eligibility. **A:** April 9, 2008 – December 31, 2008 **B:** January 1, 2009 – June 30, 2009 **C:** July 1, 2009 – November 5, 2009 **D:** November 6, 2009 – June 30, 2010
Explanatory Variables (ACS)	
Variable	**Description**
Race	Race (ethnicity) of householder, as reported in ACS, coded into mutually exclusive categories (Dummy variables, Yes=1) White only Black (primary racial response) Asian (primary racial response) Hispanic (regardless of race) Other (all other racial responses)
Age	Age of the household head, as reported in ACS, and then collapsed into buckets as below (Dummy variables, Yes=1) Age <=30 (or 20-30) Age 31-60 Age 61+
Education	Dummy Variables: Yes=1 if at least college. Derives from the years of education of the householder, as reported in ACS.
Married	Dummy variable: Yes=1 if household head is married, as reported in ACS
Household Size	Number of household members, as reported in ACS. (*hhsize*)
Years Owned	Years since the householder owned the home, as reported in Zillow.
Year Built	Year the home was built, as reported in the ACS, and then collapsed into the categories below (Dummy variables, Yes=1) <=1970 1971-1990 1991-2010 2011+
Bedrooms	Number of bedrooms, as reported in ACS.

Table 1 – Continued

Income	Inflation-adjusted household income, as reported in ACS. Then, households are ranked (within state and survey year) and quintiles are constructed and used as below (Dummy variables, Yes=1) Low income (quintile 1) Medium income (quintiles 2, 3, and 4) High income (quintile 5)
Home Value	Inflation-adjusted home value, as reported in ACS. Then, households are ranked (within state and survey year) and quintiles are constructed and used as below (Dummy variables, Yes=1) Low value (quintile 1) Medium value (quintiles 2, 3, and 4) High value (quintile 5)
Mortgage status	Mortgage status of each household, as reported in ACS (Dummy variables, Yes=1, if household has a mortgage)
Explanatory Variables (Zillow)	
Variable	**Description**
Lot Size	Natural log of property lot size (in acres), as reported in ZTRAX
Square Footage	Home square footage of the living area, as reported in ZTRAX

Table 2 - Summary Statistics for the Baseline Analysis Sample (2000-2016)

Panel A Descriptive Statistics (N = 140,000)

	Mean	Median	St. Dev.
Annualized ROR	0.0068	0.0108	0.067
IAROR	-0.0256	-0.0526	0.4092
ROR	0.1252	0.0822	0.4832
Married	0.6615	1	0.4732
Household Size	2.769	2	1.45
Bedrooms	3.297	3	0.8042
Year Built	1980	1984	19.48
Years Owned	5.069	5	3.133
Square Footage	2,219	1,775	8,204
Age	44.35	42.00	13.75
Income	108,600	87,100	105,800
Home Value in ACS	299,000	226,000	294,100
Purchase Price	288,800	220,000	268,200
Sale Price	313,100	238,000	315,700

Panel B Rates of Return for Single-Family Homes by Race (N = 140,000)

	ROR (Mean)	ROR (Median)	IAROR (Mean)	IAROR (Median)	Ann. ROR (Mean)	Ann. ROR (Median)
White	0.142	0.0875	-0.0115	-0.0466	0.0102	0.0115
Black	-0.0473	-0.0735	-0.1806	-0.2061	-0.0214	-0.0092
Asian	0.1517	0.112	-0.0014	-0.0269	0.0108	0.0145
Hispanic	0.0642	0.0435	-0.0763	-0.0974	-0.0079	0.0053

Panel C Representation by Category (N = 140,000)

Income Quintiles	% of the Sample in the Category
1	11.92
2	18.04
3	21.30
4	23.81
5	24.92

Race Categories	% of the Sample in the Category
White	79.54
Black	4.16
Asian	4.54
Hispanic	7.34
Other	4.43

Table 2 presents summary statistics for the sample of households in our baseline analysis examining determinants of ROR. Panel A reports the distribution of values for different ROR variables, a series of determinants we examine, and several component variables. Panel B provides means and medians for our measure of ROR by race over the 2000-2016 sample period. Panel C outlines the distribution of households in two categories relevant to our later difference-in-differences analysis. Values are rounded in accordance with Census disclosure requirements. Because values are rounded, the %s in Panel C per category do not sum to exactly 100%.

Table 3 - OLS Regressions of Rates of Return for Single-Family Homes on Household and Property Characteristics

Dependent Variable:	ROR (1)	IAROR (2)	Ann. ROR (3)	Ann. ROR (4)	Ann. ROR (5)	Ann. ROR (6)	Ann. ROR (7)
Black	-0.1115***	-0.0953***	-0.0186***	-0.0176***	-0.0120***	-0.0126***	-0.0199***
Asian	-0.0100	-0.0079	-0.0020	-0.0034**	-0.0033*	-0.0026	-0.0034
Hispanic	-0.0364***	-0.0296***	-0.0096***	-0.0118***	-0.0097***	-0.0094***	-0.0126***
Other Race	-0.0451***	-0.0362***	-0.0091***	-0.0105***	-0.0086***	-0.0083***	-0.0108***
Married	0.0322***	0.0271***	0.0055***	0.0058***	0.0051***	0.0052***	0.0066***
Household Size	-0.0139***	-0.0117***	-0.0024***	-0.0025***	-0.0023***	-0.0023***	-0.0032***
Bedrooms	-0.0022	-0.0013	0.0005	0.0008*	0.0013**	0.0013*	0.0017**
Built 1971-1990	-0.0607***	-0.0509***	-0.0046***	-0.0031***	-0.0015	-0.0015	-0.0029**
Built 1991-2000	-0.0426***	-0.0369***	-0.0054***	-0.0032***	-0.0012	-0.0011	-0.0051***
Built 2011+	0.6141***	0.5635***	0.0402*	0.0422**	0.0268	0.0149	0.0449**
Ln(Lot Size)	0.0187***	0.0152***	0.0003	0.0007*	0.0016***	0.0022***	0.0007
Ln (Square Footage)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Age <=30	0.0613***	0.0540***	0.0071***	0.0064***	0.0044***	0.0032**	0.0062***
Age 31-60	0.0172**	0.0165***	0.0011	0.0001	-0.0015	-0.0019*	0.0005
At least college	0.0197***	0.0172***	0.0046***	0.0041***	0.0024***	0.0022***	0.0042***
Medium Income	0.0145**	0.0099*	0.0027***	0.0026***	0.0015	0.0019	0.0024*
High Income	0.0148*	0.0097	0.0038***	0.0033***	0.0012	0.0013	0.0022
Medium Home Val	0.0808***	0.0633***	0.0124***	0.0110***	0.0057***	0.0055***	0.0060***
High Home Val	0.1837***	0.1477***	0.0204***	0.0172***	0.0066***	0.0061***	0.0096***
Has Mortgage	-0.0402***	-0.0364***	-0.0074***	-0.0079***	-0.0070***	-0.0062***	-0.0121***
Constant	0.4856***	0.116	0.0157**	0.0141**	0.0108	0.0099	-0.0144**
Survey Year x State FE	Yes	Yes	Yes	No	No	No	No
Survey Year FE	No	No	No	Yes	Yes	Yes	Yes
Purchase Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years Owned FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	No	No	No	Yes	No	No	Yes
Census Tract FE	No	No	No	No	Yes	No	No
Block Group FE	No	No	No	No	No	Yes	No
Transaction Years	2000-2016	2000-2016	2000-2016	2000-2016	2000-2016	2000-2016	2005-2016
Adj. R-squared	0.2301	0.2345	0.3475	0.3393	0.4071	0.4381	0.3622
N	140,000	140,000	140,000	140,000	132,000	113,000	85,500

Table 3 reports the estimation of eq. (1) at the household level. Columns differ based on dependent variable used (col. 1 ROR, col. 2 IAROR, and cols. 3-7 annualized ROR), fixed effects, and the sample period. Standard errors are clustered by county. All coefficients and N (obs) are rounded according to Census disclosure guidelines. Note that Sqft has a coefficient estimate of zero across all columns due to Census rounding rules. We omit reporting standard errors for brevity in the table. We define the variables in Table 1. The symbols ***, **, and * denote statistical significance at the 5%, 1% and 0.1% levels (two-tailed), respectively.

Table 4 - Summary Statistics for the Difference-in-Differences Analysis Sample

Panel A Descriptive Statistics for Different Groups						
	FHTC Income-Eligible Group (N=14,000)		**Control Group 1 – "Never Eligible" (N=1,200)**		**Control Group 2 – "Not Initially Eligible" (N=2,600)**	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Annualized ROR	-0.0117	0.075	-0.0096	0.0623	-0.0101	0.0615
IAROR	-0.0686	0.3574	-0.0767	0.3029	-0.0808	0.2847
ROR	0.0214	0.3939	0.0113	0.3361	0.0104	0.3199
Married	0.7179	0.4501	0.4139	0.4927	0.448	0.4996
Household Size	2.936	1.425	2.367	1.432	2.493	1.391
Bedrooms	3.246	0.749	3.608	0.8849	3.4	0.7931
Year Built	1982	20.32	1983	21.25	1982	20.73
Years Owned	3.443	1.76	3.337	1.853	3.51	1.869
Square Footage	1,884	2,513	2,481	1,152	2153	2,859
Age	37.28	9.451	41.49	9.216	39.34	9.454
Income	77,320	37,350	244,300	90,000	145,400	52,130
Median Home Value	214,400	120,000	380,100	175,200	300,900	156,100
Panel B Race Category Breakouts by Group						
Race Categories (% of the Group Subsample)						
White		78.86		82.84		83.19
Black		4.58		3.61		3.43
Asian		4.04		6.71		4.47
Hispanic		7.33		3.87		4.82
Other		5.19		2.97		4.09

Table 4 presents summary statistics for the sample of households in our difference-in-differences analysis with the First-time Homebuyer Credit. Panel A reports the distribution of values for different ROR variables, a series of determinants we examine, and several component variables. Panel B outlines the distribution of households by race. Values are rounded in accordance with Census disclosure requirements.

Table 5
Difference-in-differences (DiD) Analysis with the First-time Homebuyer Credit

Dependent Variable: Annualized ROR [Eq. (2)]					
	All Treatment Periods	**Control 1 Sample "Never Eligible"**	**Control 2 Sample "Not Initially Eligible"**	**Control 1 Sample "Never Eligible," Under 60 Only**	**Control 2 Sample "Not Initially Eligible," Under 60 Only**
	(1)	(2)	(3)	(4)	(5)
Eligible	-0.0007	-0.0016	-0.0050**	0.0012	-0.0038*
	(0.0014)	(0.0026)	(0.0016)	(0.0026)	(0.0017)
Post		-0.0102	-0.0082*	-0.0114	-0.0096*
		(0.0060)	(0.0041)	(0.0066)	(0.0047)
Post x Eligible		0.0122**	0.0099**	0.0125**	0.0100**
		(0.0040)	(0.0031)	(0.0039)	(0.0035)
Post A	0.0165***				
	(0.0028)				
Post A x Eligible	0.0026				
	(0.0019)				
Post B	-0.0055*				
	(0.0027)				
Post B x Eligible	0.0091**				
	(0.0029)				
Post C	-0.0072*				
	(0.0032)				
Post C x Eligible	0.0060				
	(0.0032)				
Post D	-0.0203***				
	(0.0038)				
Post D x Eligible	0.0176***				
	(0.0043)				
Eligible 2	-0.0004				
	(0.0010)				
Post D x Eligible 2	0.0131**				
	(0.0040)				
Adj. R-Squared	0.3643	0.4379	0.4381	0.4445	0.4468
N	85,500	17,500	19,000	17,000	18,500
Controls	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Purchase Year FE	Yes	Yes	Yes	Yes	Yes
Years Owned FE	Yes	Yes	Yes	Yes	Yes

Table 5 report estimates of the eq. (2) difference-in-differences using the First-time Homebuyer Credit. Col. (1) examines all treatment periods within the same regression whereas Col. (2)-(4) differ based on the control group used and the sample restrictions listed in the text. Survey year, county, purchase year, and years owned fixed effects are used with all estimates. Standard errors are clustered by county and reported in parentheses. Coefficients and N (obs) are rounded according to Census disclosure guidelines. Because N is rounded, the sum of N across Col. (2)-(4) will not equal N in Col. (1). We define the variables in Table 1. The symbols ***, **, and * denote statistical significance at the 5%, 1% and 0.1% levels (two-tailed), respectively.

Table 6 - Difference-in-differences (DiD) Analysis – By Race (Excluding Low Income Households by Limiting the Sample to Households around the Eligibility Threshold)

Dependent Variable: Annualized ROR [Eq. (3)]					
	Control 2a Sample "Not Initially Eligible"	Control 2a Sample "Not Initially Eligible"	Control 2a Sample "Not Initially Eligible"	Control 2a Sample "Not Initially Eligible"	Control 2a Sample "Not Initially Eligible"
	(1)	(2)	(3)	(4)	(5)
	Default	**Race=Nonwhite**	**Race=Black**	**Race=Hispanic**	**Race=Asian**
Post	-0.0157**	-0.0150**	-0.0110	-0.0148*	-0.0100
	(0.0056)	(0.0057)	(0.0060)	(0.0061)	(0.0061)
Eligible	-0.0040	-0.0020	-0.0030	-0.0020	-0.0020
	(0.0026)	(0.0026)	(0.0027)	(0.0026)	(0.0027)
Post X Eligible	0.0100*	0.0056	0.0057	0.0052	0.0057
	(0.0044)	(0.0043)	(0.0045)	(0.0043)	(0.0045)
Race		-0.0040	-0.0130	-0.0020	-0.0050
		(0.0074)	(0.0193)	(0.0140)	(0.0061)
Race X Post		-0.004	-0.0320	-0.0020	0.0011
		(0.0085)	(0.0237)	(0.0136)	(0.0073)
Race X Eligible		-0.008	0.0016	-0.0170	-0.0006
		(0.0083)	(0.0219)	(0.0159)	(0.0075)
Race X Post X Eligible		0.0275*	0.0573*	0.0407*	0.0035
		(0.0111)	(0.0262)	(0.0170)	(0.0092)
Adj. R-Squared	0.4565	0.4586	0.4425	0.4536	0.4391
N	8,600	8,300	7,500	7,700	7,700
Controls	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Purchase Year FE	Yes	Yes	Yes	Yes	Yes
Years Owned FE	Yes	Yes	Yes	Yes	Yes

Table 6 reports coefficients from estimating eq. (3), which represents the difference-in-differences by race. The specifications in this table exclude low income households from the analysis, keeping households within a symmetric range around the income eligibility threshold for the FHTC. All columns compare the new treated cohort to the second control group ("not initially eligible"). Columns differ based which race is being compared to White households (and White households are the omitted group in all regressions in this table). Sample restrictions are described in the text. Standard errors are clustered by county and reported in parentheses. All coefficients and N (obs) are rounded according to Census disclosure guidelines. We omit controls for brevity. We define the variables in Table 1. The symbols ***, **, and * denote statistical significance at the 5%, 1% and 0.1% levels (two-tailed), respectively.

Appendix

Table A1
DiD Analysis with the FHTC - Table 5 with Alternative Sample Periods

Dependent Variable: Annualized ROR [Eq. (2)]						
	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"
	(1)	(2)	(3)	(4)	(5)	(6)
Eligible	-0.0014	-0.0012	-0.0013	-0.0052**	0.0016	-0.0037*
	(0.0020)	(0.0012)	(0.0028)	(0.0018)	(0.0027)	(0.0018)
Post	-0.0186*	-0.0114*	-0.0121	-0.0096*	-0.0139*	-0.0107*
	(0.0074)	(0.0055)	(0.0067)	(0.0046)	(0.0070)	(0.0049)
Post x Eligible	0.0165***	0.0081*	0.0135**	0.0102**	0.0141**	0.0101**
	(0.0042)	(0.0032)	(0.0043)	(0.0033)	(0.0046)	(0.0036)
Period Start	1/2005	1/2005	4/2007	4/2007	1/2007	1/2007
Period End	6/2010	6/2010	6/2010	6/2010	11/2009	11/2009
Adj. R-Squared	0.4099	0.4138	0.4414	0.4426	0.4377	0.4418
N	35,000	38,600	15,000	16,500	14,500	16,500
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Purchase Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Years Owned FE	Yes	Yes	Yes	Yes	Yes	Yes

Table A1 reports Table 5 estimates of eq. (2) with alternative sample periods. Columns differ based on the sample period examined and the control sample used. Col. (1) and (2) use the period 1/1/2005-6/30/2010. Col. (3) and (4) use the period 4/1/2007-6/30/2010. Col. (5) and (6) use the period 1/1/2007-11/30/2009. Odd columns use the Control 1 Sample "Never Eligible," and even columns use the Control 2 Sample "Not Initially Eligible." All columns are limited to householders under 60. Survey year, county, purchase year, and years owned fixed effects are used with all estimates. Standard errors are clustered by county and reported in parentheses. Coefficients and N (obs) are rounded according to Census disclosure guidelines. We define the variables in Table 1. The symbols ***, **, and * denote statistical significance at the 5%, 1% and 0.1% levels (two-tailed), respectively.

Table A2

DiD Analysis with the FHTC - Table 5 with Alternative Geographic Fixed Effects

Dependent Variable: Annualized ROR [Eq. (2)]

	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Eligible	0.0012	-0.0038*	-0.0001	-0.0048**	-0.0014	-0.0004	0.0007	0.0003	0.0001	-0.0013
	(0.0026)	(0.0017)	(0.0028)	(0.0017)	(0.0022)	(0.0014)	(0.0033)	(0.0017)	(0.0047)	(0.0023)
Post	-0.0114	-0.0096*	-0.0095	-0.0080	-0.0244**	-0.0172**	-0.0137	-0.0112	-0.0111	-0.0141
	(0.0066)	(0.0047)	(0.0060)	(0.0044)	(0.0081)	(0.0060)	(0.0097)	(0.0068)	(0.0132)	(0.0093)
Post x Eligible	0.0125**	0.0100**	0.0146***	0.0115***	0.0164***	0.0089*	0.0164*	0.0087	0.0166	0.0159**
	(0.0039)	(0.0035)	(0.0039)	(0.0033)	(0.0048)	(0.0036)	(0.0067)	(0.0047)	(0.0096)	(0.0059)
Adj. R-Squared	0.4445	0.4468	0.5339	0.5355	0.4681	0.4714	0.5110	0.5128	0.5272	0.5288
N	16,500	18,500	16,000	17,500	32,700	36,500	25,000	28,500	18,000	19,000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	No	No	No	No	No	No	No	No
Zip Code FE	No	No	No	No	Yes	Yes	No	No	No	No
Census Tract FE	No	No	No	No	No	No	Yes	Yes	No	No
Block Group FE	No	No	No	No	No	No	No	No	Yes	Yes
Purchase Year FE	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Years Owned FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County x Purchase Year FE	No	No	Yes	Yes	No	No	No	No	No	No

Table A2 reports Table 5 estimates of eq. (2) with alternative geographic fixed effects. Columns differ based on the fixed effects examined and the control sample used. Odd columns use the Control 1 Sample "Never Eligible," and even columns use the Control 2 Sample "Not Initially Eligible." All columns are limited to householders under 60. Survey year, county, purchase year, and years owned fixed effects are used with all estimates. Standard errors are clustered by county and reported in parentheses. Coefficients and N (obs) are rounded according to Census disclosure guidelines. Because N is rounded, even when columns use the same control group and period, their N can differ. We define the variables in Table 1. The symbols ***, **, and * denote statistical significance at the 5%, 1% and 0.1% levels (two-tailed), respectively.

Table A3 - Difference-in-differences (DiD) Analysis with the First-time Homebuyer Credit – By Race

Dependent Variable: Annualized ROR	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"	Control 1 Sample "Never Eligible"	Control 2 Sample "Not Initially Eligible"
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nonwhite	-0.0100	-0.0227***						
	(0.0099)	(0.0045)						
Post	-0.0183*	-0.0163***	-0.0120	-0.0103*	-0.0161*	-0.0143**	-0.0110	-0.0095*
	(0.0073)	(0.0048)	(0.0075)	(0.0047)	(0.0076)	(0.0051)	(0.0072)	(0.0045)
Nonwhite X Post	0.0234	0.0285***						
	(0.0128)	(0.0071)						
Eligible	0.0041	-0.003	0.0029	-0.0030	0.0039	-0.0030	0.0027	-0.0030
	(0.0027)	(0.0018)	(0.0028)	(0.0018)	(0.0028)	(0.0018)	(0.0027)	(0.0018)
Nonwhite X Eligible	-0.0221**	-0.008						
	(0.0076)	(0.0049)						
Post X Eligible	0.0111**	0.0085*	0.0111**	0.0086*	0.0112**	0.0085*	0.0114**	0.0085*
	(0.0042)	(0.0034)	(0.0041)	(0.0036)	(0.0042)	(0.0034)	(0.0041)	(0.0036)
Nonwhite X Post X Eligible	0.0149	0.0099						
	(0.0124)	(0.0080)						
Black			0.0025	-0.0223*				
			(0.0106)	(0.0101)				
Black X Post			0.0256	0.0131				
			(0.0212)	(0.0156)				
Black x Eligible			-0.0255*	-0.002				
			(0.0114)	(0.0120)				
Black x Post x Eligible			-0.0110	0.0030				
			(0.0223)	(0.0198)				
Hispanic					-0.0340	-0.0205***		
					(0.0182)	(0.0062)		
Hispanic X Post					0.0371	0.0489***		
					(0.0208)	(0.0100)		
Hispanic x Eligible					-0.008	-0.0217**		
					(0.0147)	(0.0082)		
Hispanic x Post x Eligible					0.0274	0.0153		
					(0.0178)	(0.0115)		
Asian							0.0077	-0.0196*
							(0.0070)	(0.0079)
Asian X Post							0.0176	0.0201*
							(0.0162)	(0.0083)

Asian x Eligible							-0.0211**	0.0060
							(0.0079)	(0.0091)
Asian x Post x Eligible							-0.0010	-0.0050
							(0.0169)	(0.0109)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Purchase Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years Owned FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.4476	0.4488	0.4229	0.4234	0.4448	0.4449	0.4219	0.4207
N	18,000	17,500	14,500	15,500	14,500	16,000	14,500	16,000

Table A3 report coefficients from estimating eq. (3), which interacts the difference-in-differences design by a race indicator. Columns differ based on which race is being compared to White households (and White households are the omitted group in all regressions in this table). Sample restrictions are described in the text. Standard errors are clustered by county and reported in parentheses. All coefficients and N (obs) are rounded according to Census disclosure guidelines. We omit controls for brevity. We define the variables in Table 1. The symbols ***, **, and * denote statistical significance at the 5%, 1% and 0.1% levels (two-tailed), respectively.