The Cross-section of Housing Returns*

Jonathan Halket†

Mays School of Business, Texas A&M University

May 5, 2025

Abstract

I document large neighborhood spreads in the return to single-family residential property. Houses in neighborhoods with lower household income or credit scores or higher shares of black residents have higher yields and returns. As these spreads may also be caused by measurement error, I exploit time-series changes in the spreads to document a role for segmented housing markets where local discount rates may price local assets. Aggregate shocks to mortgage supply change neighborhood expected return spreads and, in sample, imply poor credit neighborhoods have higher returns and higher house price and return volatility. Since this channel for differential risk exposure is through intertemporal marginal rates of substitution, return spreads are not purely compensation for bearing extra risk.

Keywords: Housing Returns, Segmented Markets, Mortgages, Credit

*I am grateful for feedback from Morris Davis, Pedro Gete and Athena Tsouderou, and especially Siddhartha Biswas, Marcus Casey, John Clapp, Antonio Gargano and Daniel Greenwald for insightful conference discussions and to seminar participants at Baruch College, Federaal Reserve Bank of Chicago, Federal Reserve Board, Office for Financial Research, Rutgers University and the University of Connecticut, and conference participants at the 2022 Urban Economics Association Conference, the 2022 University of Florida Real Estate Finance and Investment Symposium, AREUEA's 2023 International Conference, the Echoppe Conference 2023, the 2023 CREDA Real Estate Symposium, the 2024 AFA meetings, ITAM Finance Conference, the Boca Finance and Real Estate Conference and the Easter Finance Association Conference. Data and research were supported by grants from Unidos/Urban Institute and the Texas A&M Center for Hispanic Housing Studies under a grant from the Department of Housing and Urban Development. All errors are attributable to me.

†Email: `jhalket@tamu.edu`. Address: Finance Department; Wehner Building; Mays School of Business; Texas A&M University; College Station, TX 77845

1 Introduction

In this paper, I combine several large data sets of individual rental and sale transactions, local property tax information and household economic and demographic characteristics to create estimates of the variation in the total returns to single family housing across zip codes and time within many U.S. CBSAs. I show that the returns to holding single family real estate vary considerably and predictably even within urban areas. Ex-ante differences in local demographic and economic characteristics, such as neighborhood median income, average credit score and racial composition, predict future returns mainly by predicting yields. Yields and future returns for properties in areas with ex-ante lower median income or credit scores or a larger share of Black or Hispanic residents increase more than similar houses located in other neighborhoods in the CBSA when credit contracts.

The estimated spreads in average returns across local areas within CBSAs are generally large: a property in a zip code with a one standard deviation higher median income or average credit score as compared to a property in another zip code in the same CBSA earns returns around 1 to 1.5 percentage points less per year. However, these estimates may be influenced by measurement error. For this reason, I also exploit time-series changes in yield and returns spreads to argue that some of the predictable dispersion in returns is neither measurement error *nor* compensation for bearing risk. The spreads in future yields and returns predicted by current area characteristics varies considerably over time in the data. The spread is narrow prior to the Great Recession, then widens during the recession before slowly narrowing after 2016.[1]

In Section 2, I build a model to show how different effective discount rates may price different houses within the same housing market, even if there exists a deep pocketed landlord with a low discount rate. My model combines the approaches of Eisfeldt and Rampini (2009) and Piazzesi and Schneider (2016). As in Landvoigt et al. (2015) and Piazzesi and Schneider (2016), housing may be traded in segmented markets since housing investments tend to be lumpy and geographically differentiated. I show that when households vary in their opportunity cost of credit (OCC), they may choose to sort into different market segments based on this heterogeneity, so that the intertemporal marginal rate of substitution (IMRS) for the marginal property owner that "prices" housing in one part of the market may vary from the IMRS of the marginal owner in another part of the market. Similar to Eisfeldt and Rampini (2009), the dollar flow value of owning a real asset can depend on who owns the asset. In the case of housing, even though landlords may have a low opportunity cost of funds, they may be inefficient relative to owner-occupiers at converting a house into housing services and thus potentially unwilling to bid-up the price of a house that may otherwise be valued by a high OCC owner-occupier. Put together, this can mean that heterogeneity in IMRSs (perhaps also due to differently binding borrowing constraints) among owners can imply that the expected

[1] Predicted returns here are at the asset level, not the owner level, and thus do not account for differences in owner leverage or default rates. The data suggest that leverage is much higher for low credit borrowers and numerous other studies find higher default rates for low credit, low income borrowers (for example, Davis et al. (2022)). Accounting for each of these, if desired, would likely imply an even wider spread in owner-level returns as compared to asset-level returns.

return to holding a house can vary cross-sectionally within a market for reasons other than risk.

I then show that households in the data do indeed sort by economic and demographic characteristics that are known to be linked to their OCC and that these characteristics do predict differences in the cross-section of returns. Because the cross-sectional estimates may be contaminated by measurement error, I exploit a series of quasi-natural experiments to show that aggregate shocks to OCC spreads change local areas' average yields and returns differently depending on each location's characteristics. Properties in areas with lower credit scores have larger changes in yields and returns in response to changes in the access to or cost of credit. This seems theoretically reasonable: for instance, changes in macroprudential mortgage regulations may likely matter more for areas where households were *ex ante* more likely to be credit constrained.

An implication of this, though, is that house prices move around more in my sample (during which there was a large expansion and then large contraction in the OCC for many households) in low credit areas. So, in sample, properties in areas with higher average returns also have more volatile prices and thus may be perceived by both an agent and the econometrician as riskier. In other words, much of the higher price volatility in low credit areas is caused by the higher volatility of area IMRSs due to those areas' extra sensitivity to aggregate shocks that affect OCCs.

Differential sensitivity to these sorts of shocks may be a risk-factor that causes differences in expected returns across properties. However, market segmentation is crucial for this channel. When markets are segmented, risk and return may be correlated but risk does not need to *cause all* of the extra return. So even though properties with (in some cases much) higher returns also have higher risk, there need be no version of the "equity premium puzzle" in single-family housing.[2]

The results here also connect to the large and growing literature on racial differences in housing market outcomes and wealth building (e.g. Ambrose et al. (2020); Bayer et al. (2016, 2017); Begley and Purnanandam (2021); Bhutta and Hizmo (2020); Higgins (2023)). My finding that areas with high shares of Black or Hispanic households have higher yields and returns, even after controlling for other OCC measures, is consistent with the common notion that housing in minority neighborhoods is undervalued. The fact that the differences in returns is driven mainly by yields implies that owners in minority neighborhoods are realizing a higher utility flow or rent (relative to the house value) from owning homes. However, because ownership rates may also vary with race, my results do not necessarily imply that minority households enjoy much of the extra returns that their sorting may be causing in equilibrium.

I use novel data with a large sample of detailed rental and sales transactions from a panel of U.S. CBSAs to contribute to the growing literature that estimates and analyzes returns to property at granular levels. This data permits me to estimate the rental price and for-sale value of most single-family housing at zip code levels within 21 U.S. CBSAs using hedonic methods. Hedonic methods allow me to focus on how returns to holding the same observable structure vary across locations within a CBSA, thereby ameliorating concerns discussed in Halket et al. (2020) about

[2] In other words, we do not necessarily need to worry about whether there is *enough* variation in risk to explain predictable return variations across different types of housing.

selection across rentals and owner-occupied houses on unobserved maintenance costs.[3]

My estimates reveal substantial variation in returns to owning the same structure across locations within a CBSA. The differences in returns across locations come predominantly from differences in the net yield. Simply put, a high average net rent-to-price ratio[4] in a particular location within a CBSA relative to other locations within the CBSA predicts higher relative returns in that location and not lower capital gains.

Eisfeldt and Demers (2015), using different data and methods, finds that yields and returns are correlated with "pricing tiers" within single-family rentals; zip codes with lower housing prices tend to have higher returns. Here, I build on their findings by using more granular rental data to estimate the returns to both rentals and owner-occupied housing[5] and by showing that a location's average yield and return are highly correlated with many of the location's ex-ante economic and demographic characteristics. Within CBSAs, land yields and my demographic/economic factors are generally all statistically significantly correlated in the same direction. All 21 CBSAs in our sample have higher yields in areas with lower average credit scores, 16 of the CBSAs have significantly higher yields in low-income zip codes and 15 (14) of the CBSAs have significantly higher yields in zips with a higher share of Black (Hispanic) residents.

In concurrent work, Diamond and Diamond (2024), using an alternative approach to control for selection bias, finds similar patterns in returns; in particular that houses where Blacks live tend to earn higher yields than Whites. My paper contains similar results on race and shows also how important credit is for understanding the yield patterns that this growing literature finds and provides an equilibrium theory that can explain these emerging results.

It is not possible to observe all components of housing returns in my data. While I observe the agreed rent, property taxes and can estimate changes in the value of the house, I do not observe potentially important elements of operating expensing such as maintenance, rental non-payment and vacancy costs. These unobserved elements of yields and returns may be correlated with the factors used in the papers.[6] Despite this non-classical measurement error in the cross-section, I use several methods to argue that one can still conclude that markets are segmented.

I use hedonic methods to estimate how yields and returns vary holding across locations, holding

[3]Furthermore, the heterogeneity in the return to structures due to sorting is likely small because, unlike locations, returns to installing new structures are generally pinned down by marginal costs in the construction sector.

[4]The net rent-to-price ratio, or "cap rate," is rent less any operating costs in the numerator (approximately "Net Operating Income") and the stock value of the asset in the denominator.

[5]Better rental data allows me to hedonically estimate rents and prices using a richer set of location controls than Eisfeldt and Demers (2015). This allows me to separately estimate the variation in yields and returns to properties across location, holding fixed their structural characteristics, which in turn allows me, when combined with the findings in Halket et al. (2020), to plausibly conclude that owner-occupied returns vary similarly to rental returns. In addition, Eisfeldt and Demers (2015) moots that lower priced zip codes' higher correlation with city-wide house prices may be a risk factor which explains these zip codes' higher returns. I find the same correlation but show that there are still large excess returns in low credit zip codes even after controlling for this risk factor.

[6]For instance, Humphries et al. (2024) finds sizable eviction and rental non-payment costs in low income rentals.

fixed the observed structural characteristics of the properties. So unobserved heterogeneity in maintenance that is correlated with the structure of the house do not cause the estimated spreads. Furthermore, local area credit scores, income and race remain significant in regressions of yield and return spreads after also including a measure of vacancy rates.

The relationship between ex-ante household credit scores and subsequent returns varies over my sample. At the tail end of the housing boom in 2006, the estimated yield spread across zip codes was relatively narrow, consistent with evidence in Lewellen and Williams (2021). During the subsequent housing bust, house prices tended to fall further in areas where low credit score households lived, thereby causing yield spreads to widen considerably within most CBSAs (since rents did not fall as much prices). In this period, low credit areas strongly predict higher yields and thus higher future returns. Home ownership rates also fall in these same neighborhoods. The negative yield spread narrows late in the recovery but remains negative throughout the sample.

I argue that unobserved rental non-payment and eviction is unlikely to explain this joint pattern of credit shocks, yields, returns and home ownership during the bust. Eviction and rental non-payment pertain to landlords only and not owner-occupiers. For these costs to explain the changes in the spreads of yields, capital gains and returns during the housing bust, the spread in these costs with respect to credit scores needs to go up during this period and this change in costs needs to be mostly capitalized into prices and not rents. Absent any other changes though, a rise in landlord operating expenses(relative to owner-occupied expenses) should push more housing into owner-occupancy.[7] This is counter-factual to this particular episode: home ownership rates decreased most in low credit neighborhoods during this period. So the expected value of a house being owner-occupied must have decreased relative to the expected value of it being rented out. The most likely channel for this is that owner-occupiers in low credit areas must also have experienced a significant relative increase in their OCC.

To further test the hypothesis that differences in expected returns are related to ex ante differences in OCC and are not due to omitted unobserved variables, I follow Loutskina and Strahan (2015), Adelino et al. (2022) and Greenwald and Guren (2021) by using changes in conforming loan limits (CLLs) as a set of natural experiments that changed the cost of credit for some locations more than others. Such changes are not likely correlated with changes in unobserved operating costs. Starting with the Housing and Economic Recovery Act of 2008, CLLs were changed annually at the CBSA rather than national level. I find that locations within CBSAs that previously had relatively many mortgage originations near the old CLL experienced greater falls in yields as compared to other locations within the CBSA and especially so when the national spread between conforming and jumbo mortgages was high. I also show that this same treatment does not forecast future net rent growth and thus likely points to changes in the local IMRS as the cause for the change in yields.

In the cross-section, the relationships between the factors and total returns across locations are very strong; all 21 CBSAs in our sample have a statistically and economically significant relationship

[7] See Halket and Pignatti Morano di Custoza (2015) for an example.

with credit score and all but one CBSAs has two or more other significant factors. In a multivariate "horse race" with all our economic factors, credit scores are the most important by far. However double sorting on both credit and race reveals that the share of Black or Hispanic households in an area can have a large effect on returns even conditional on area credit scores.

These results have many potentially important economic and econometric implications. Higher returns in low income areas may imply that owner-occupancy in these areas is a potent way to build wealth, while segmented markets may also explain the high rent-to-price ratios in low-income neighborhoods. Also, if the high cost of borrowing for a subset of households suppresses prices for the types of houses that these households live in, this could lower the incentive for developers to build houses for these households. Finally, if properties in different locations respond differently to changes in the cost of and access to credit, then the effects of monetary policy may have important intra-city heterogeneity; and the response of house prices to changes in monetary policy may vary within markets.

Econometrically, the results imply that different houses can have different long-run risk-adjusted expected returns. Time-series studies that follow Campbell-Shiller decompositions (e.g. Campbell et al. (2009) should be wary of estimating models where this returns is restricted to be identical for all properties.

Finally, when markets are segmented, the same factor can be both a risk-factor and a discount-rate factor. This can lead to invalid conclusions based on widely used measures, such as CAPM coefficients or Sharpe Ratios.

1.1 Related Literature

Structural dynamic models of housing and home ownership (for example Ríos-Rull and Sánchez-Marcos (2008) Landvoigt et al. (2015), Garriga et al. (2019), Kaplan et al. (2020)) often feature binding borrowing constraints that affect the relative equilibrium price of housing across different parts of their models' housing markets. My results provide novel evidence for the mechanisms in these models and also serve as alternative evidence of incomplete regional risk sharing (as in Lustig and Van Nieuwerburgh (2010)).

Low credit areas' higher house price sensitivity to changes in the access to or cost of credit is at the core of studies of the 2000's housing boom and bust (for example, Kuminoff and Pope (2012) and Landvoigt et al. (2015)) and the growing literature on the heterogeneous effects of monetary policy and macroprudential policies (for example, Adelino et al. (2022),Bosshardt et al. (2023) and Gorea et al. (2022)). A near necessary condition for this excess sensitivity is that housing markets are segmented such that different IMRSs are pricing different properties.[8]

A huge literature looks at the time-varying relationship between risk and returns in housing markets, with an eye to understanding the market or macro level factors that may be driving

[8]While theoretically possible that changes in OCC affect house prices through changes in the net expected future value of housing services (the dividend flow of housing), I find no evidence for this channel and the intuitions/narratives in most research assume a discount rate channel.

them (see Goetzmann et al. (2021) for a recent summary). Some of this literature finds dispersion in housing returns (variously measured) within metropolitan areas (”markets”) and attempts to explain it by using differences in risk (variously measured). Housing is both an asset and a consumption good, so the relationship between expected returns and risk may be non-trivial. For example, Sinai and Souleles (2005) finds a positive relationship between price-to-rent ratios and risk across markets. Han (2013) finds a positive relationship between housing returns (measured using only capital gains) and risk within some markets and a negative relationship within others. Giacoletti (2021) finds that idiosyncratic capital gains risk is a significant part of housing risk, particularly over short horizons.

Measuring returns for real estate, particularly single-family residential housing, is complicated because many components of cash flows are not observed in most data sets. For this reason, historically, many studies of risk and return in housing focus on capital gains. My findings contribute to the growing number of studies which show yields contain important information on the cross-section of returns (e.g. Eichholtz et al. (2021)). Demers and Eisfeldt (2022) finds that yields and, thus, returns are higher in the lowest priced zip codes within markets and that price appreciation is more correlated with city-level risk in these same zip codes. Amaral et al. (2021), using a long panel of city-level property returns, finds that larger cities have lower returns and yields and also lower correlations with income shocks. Damen et al. (2025) finds comparable yield and return spreads to ours using multi-family data from the U.S. and other countries. They argue that risk alone is unable to explain these spreads and moot market segmentation as an important channel for explaining the results. Plazzi et al. (2010) studies risk and return among CRE properties using Campbell-Shiller decompositions.

The effect of housing wealth on IMRSs has been studied at least since Campbell and Cocco (2003, 2007). Lustig and Van Nieuwerburgh (2005) and Lustig and Van Nieuwerburgh (2010) study how time-series variation in IMRSs can explain various features of market returns (i.e. the return to wealth) and impart predictability to excess returns. A large literature going back to Case and Shiller (1989) and Case and Shiller (1990) finds predictability in the excess returns to housing. Cochrane (2011) discusses how variation in discount rates can perhaps explain this predictability. Campbell et al. (2009) finds that variation in a “housing premia” over the risk-free rate is important for understanding changes in housing yields.

Models with segmentation are fundamental to urban economics (Muth (1966), Sweeney (1974)) and, more recently, in dynamic models used in macroeconomics and asset pricing (Piazzesi and Schneider (2016)). Landvoigt et al. (2015) finds evidence for non-linear house prices and differential capital gains in the San Diego market. Piazzesi et al. (2020) finds evidence of housing market segmentation in the search behavior of households and Bernstein et al. (2019) argues that housing segmentation may be important for our understanding of how climate risk is priced. Nathanson (2020) uses segmented markets in a model of inter-city trade. Higgins (2023) examines the equilibrium housing implications of racism when markets are segmented.

There is a similarly long literature on the time-series variation in the OCC of mortgages and its relationship to house prices (e.g. Demyanyk and Hemert (2009); Justiniano et al. (2022); Davis et al. (2022) and citations within). In this paper, I show that the nexus of market segmentation and time-varying spreads in the OCC leads to predictable differences in the returns to housing. In Section 2.1, I use a simple two period model to demonstrate how heterogeneous IMRSs can lead to segmented housing markets and spreads in user-costs and returns, even when there is deep pocketed landlord that may enter any part of the market freely. Section 2.2 builds on the preceding section and constructs an econometric model that is brought to the data. Section 3 discusses the novel data and Section 4 the empirical methods. Sections 5 and 6 present the results and Section 7 concludes with some suggestions for future research.

2 Analytical Framework

2.1 A Simple Model

To illustrate how housing markets can segment in equilibrium, I adapt a simple setting from Piazzesi and Schneider (2016) and add heterogeneity in discount rates. This heterogeneity alone, when coupled with a heterogeneous housing stock and indivisible housing, is sufficient to generate variation in the expected return of housing even though housing is riskless in the simple model here. When later in this subsection, homogeneous landlords are added to the economy, a simple technological wedge between owner-occupiers and landlords is sufficient to preserve some heterogeneity in returns.

In the economy, there are two goods, consumption and housing. The economy consists of over-lapping generations of two period-lived households. Each generation has unit mass and all households are identically endowed with wealth $w_1 > 1$ in consumption goods when born.

Households obtain housing services by living in exactly one house. Houses come in different qualities h distributed uniformly over $[0, 1)$ and quality is indivisible. Owner-occupied houses require maintenance of δ^o per unit quality.

Households receive utility over consumption and housing quality in their first year of life and from their terminal wealth w_T in year two. I assume that their utilities are linear in each of consumption, housing services and terminal wealth and that there are no assets available to trade.[9]

Households are heterogeneous only in their discount rates β, which I assume are distributed uniformly over $[0, 1)$ each generation.

A household in generation t with discount rate β that chooses to own its own house solves the following problem:

$$
\begin{aligned}
&\max_{c_t, h_t} c_t + h_t + \beta w_{t+1} \\
s.t. \qquad & c_t + p^o_t(h_t) + \delta^o h_t = w_1 \\
& w_{t+1} = p^o_{t+1}(h_t)
\end{aligned}
$$

[9]As there is no risk in this economy, the economy only lacks a risk-free asset. I could allow for one and instead impose borrowing constraints on households and qualitatively similar results as found below can be attained.

The first-order condition for housing is:

$$\frac{dp_t^o}{dh} = \frac{1-\delta^o}{1-\beta} \tag{1}$$

In an equilibrium where all houses are owner-occupied, households with discount rates β live in $h = \beta$ quality houses. The equilibrium price of housing (we have dropped the time subscripts for convenience) is

$$p^o(h) = \int_0^h \frac{dp^o}{dh}(\tilde{h})d\tilde{h} = (1-\delta^o)\ln(\frac{1}{1-h}) \tag{2}$$

and the gross return to holding a house is

$$E[R(h)] = 1 + \frac{h(1-\delta^o)}{p^o(h)} = 1 + \frac{h}{\ln(\frac{1}{1-h})}. \tag{3}$$

Even with linear utilities, house prices are non-linear and expected returns decrease with quality. The characteristics of the marginal owner at any particular house quality matter for the relative price of that house.

If I assume that houses may also be rented out instead of just owner-occupied, it is clear from households' preferences that the equilibrium rental cost per unit quality $r(h) = h$. In keeping with the large literature on moral hazard problems in renter markets (Halket et al. (2020) and citations therein), I assume landlords have a higher maintenance costs δ^l per unit quality and that there is an elastic supply of landlords that maximize wealth and discount at some homogeneous rate $\beta^l \in (0,1)$.[10] The willingness to pay of a landlord is then:

$$p^l(h) = \frac{1-\delta^l}{1-\beta^l} r(h) \tag{4}$$

In an equilibrium with both landlords and owner-occupiers, landlords' value of investing a dollar into a house of quality h is

$$V^l(h) = \beta^l + \frac{h(1-\delta^l)}{p(h)} \tag{5}$$

while the value for an owner-occupier with discount rate β is

$$V^o(h,\beta) = \beta + \frac{h(1-\delta^o)}{p(h)} \tag{6}$$

where $p(h)$ is the equilibrium price of housing.

The equilibrium will then have landlords owning a house wherever $V^l > V^o(h,h)$. Using the

[10] I could instead assume that the set of potential landlords is simply the set of living households. In equilibrium then, high discount rate households will rent to lower discount rate households. The equilibrium would be richer with more complicated pricing.

fact that $p(h) = p^l(h) = \frac{1-\delta^l}{1-\beta^l}h$ wherever landlords own, there exists a cutoff,

$$h^* = \beta^l + \frac{\delta^l - \delta^o}{1-\delta^l}(\beta^l - 1) \tag{7}$$

such that landlords (households) own all housing below (above) h^*. The home ownership rate, $1 - h^*$, decreases when landlords' maintenance differential $\delta^l - \delta^o$ increases.

Below h^*, $p(h) = \frac{1-\delta^l}{1-\beta^l}h$. Above h^*,

$$p(h) = p(h^*) + \int_{h^*}^{h} \frac{dp^o}{dh}(\tilde{h})d\tilde{h} = p(h^*) + (1-\delta^o)\ln(\frac{1-h^*}{1-h}) \tag{8}$$

In an equilibrium with both landlords and owner-occupiers, landlords' expected gross return for a house of quality h is

$$E[R^l(h)] = 1 + \frac{h(1-\delta^l)}{p(h)} \tag{9}$$

while an owner-occupiers' is

$$E[R^o(h)] = 1 + \frac{h(1-\delta^o)}{p(h)} \tag{10}$$

Remark 1 *Equations 7-10 display some key features of the equilibrium:*

1. *A property may be owner-occupied even if the owner-occupier has a higher discount rate than a landlord if landlords have sufficiently higher maintenance costs for the property than the owner-occupier.*

2. *Conditional on h, landlords will have uniformly lower (higher) expected returns if their cost of maintenance is uniformly higher (lower) than owner-occupiers. Lower returns do not mean landlords do not own property in equilibrium; what matters is their expected discounted returns, V^o and V^l.*

3. *Different properties may have different expected returns depending on the opportunity cost of credit of the owners who own them in equilibrium.*

Real estate is a real asset. A change in the owner of a property affects returns not just by changing the discount rate applied to the cash flows that the property generates but also potentially changes the cash flows themselves. Therefore, the owner of the asset is not necessarily the agent with the lowest discount rate. A calibrated example is plotted in Figure 1.

Indexing all owners by i, denoting the equilibrium correspondence which maps an owner i to a set of houses $T(i)$ and indexing their expected returns similarly, the economy has a set of Euler equations (one of each i):

$$E[\beta^i R^{i,T(i)}] = 1 \tag{11}$$

Despite the heterogeneity in riskless returns, Euler equations hold with equality in this economy as long as one uses the correct, marginal investor's IMRS for each house and account for that investor's technology for operating the house.

2.2 Returns in a two sector model with added heterogeneity

In this section I generalize the intuition from subsection 2.1 in order to obtain a model that can be estimated with my data. I allow for additional heterogeneity in both household and property characteristics that will later prove important in the data. The cost of the additional heterogeneity is that I am no longer able to easily describe the complete equilibrium of the economy; the model does not analytically explain why particular agents own particular houses.[11] Despite this, I am able to describe several interesting features of equilibria in this economy which I will argue are important for understanding the evolution of the cross-section of observed returns in the data.

Time is discrete. Each property is vector of characteristics $z^{\varepsilon} \in Z^{\varepsilon}$ a compact, convex subset of $\mathbb{R}^{n_{\varepsilon}}$. These characteristics may or may not be observable to the econometrician. Examples of characteristics include location, lot size, floor space, etc.... As each property has a unique location, z^{ε} uniquely identifies a property. All agents are a vector of characteristics (state variables) $s \in S$. For convenience, I assume that I can partition the state into two parts $S = S^h \cup S^l$. As will be made clear below, S^h is the set of characteristics relevant to a household's enjoyment of the property, whereas S^l is the set of characteristics relevant to an agent's management of a property.[12] Owners, which may be households or not, have a vector of characteristics $s_l \in S^l$. Examples of household characteristics that may be in S^h are income, martial and family status, age, etc.... Examples of owner characteristics could include any of S^h as well as measures of managerial ability, etc....[13] S^l includes an indicator as to whether the owner is a landlord (i.e. rents to another agent) or an owner-occupier. As with property characteristics, household and owner characteristics may or may not be observable to the econometrician.

Define $U(z^{\varepsilon}, s^h)$ as the flow value (in non-durable consumption units) from a resident-owner pair in state $s = \{s_h, s_l\} \in S$ of a property of type z^{ε} given price and rent functions, $P : Z^{\varepsilon} \to \mathbb{R}$ and $r :: Z^{\varepsilon} \to \mathbb{R}$. Assume maintenance costs (including property taxes) are $c(z^{\varepsilon}, s^l)P(z^{\varepsilon})$ and the opportunity cost of capital is $\rho(s^l)$.[14] To simplify notation below, I assume maintenance costs are paid at the end of each time period. Let $g(z^{\varepsilon}, s^l)$ be the expected after-tax capital gains.

I assume that the willingness to pay to own a property z^{ε} by of an owner s^l matched with a resident (which could be the owner) s^h is

$$\pi(z^{\varepsilon}; s) = U(z^{\varepsilon}, s^h) - \frac{c(z^{\varepsilon}, s^l)P(z^{\varepsilon})}{1 + \rho\left(s^l\right)} + \frac{\left(1 + g(z^{\varepsilon}, s^l)\right) P(z^{\varepsilon})}{1 + \rho(s^l)}. \tag{12}$$

Thus, the willingness to pay equals the current net utility flow plus the discounted expected future

[11]One could build a computational general equilibrium model that would, in part, allocate agents to houses, as in Landvoigt et al. (2015) and Higgins (2023). Due to the curse of dimensionality, such models do not as of now permit high-degrees of heterogeneity in property characteristics.

[12]Characteristics may appear in both sets.

[13]Without loss of generality, if a certain characteristic, like location of a corporate landlord's headquarters, is relevant only to corporate owners we can assume households have a 0 value for it.

[14]The functions themselves may be time-dependent (i.e. dependent on some macro-state variables) and owners may change states over time. I suppress time-dependent notation for ease of reading.

value of the property.[15][16]

I further assume that in equilibrium there is a correspondence mapping properties to residents and owners $T : Z^{\varepsilon} \rightrightarrows S$. In equilibrium, if an agent in state $s^l \in S^l$ buys a property z^{ε} occupied by a household with $s^h \in S^h$ then $\pi(z^{\varepsilon}; T(z^{\varepsilon})) = P(z^{\varepsilon})$.[17] T can itself be partitioned into two correspondences $T^h : Z^{\varepsilon} \rightrightarrows S^h$ and $T^l : Z^{\varepsilon} \rightrightarrows S^l$ such that equation (12) can be rewritten as

$$U(z^{\varepsilon}, T^h(z^{\varepsilon})) = \left[\frac{c(z^{\varepsilon}, T^l(z^{\varepsilon})) + \rho(T^l(z^{\varepsilon})) - g(z^{\varepsilon}, T^l(z^{\varepsilon}))}{1 + \rho(T^l(z^{\varepsilon}))}\right] P(z^{\varepsilon}) \tag{13}$$

$$\approx \left[c(z^{\varepsilon}, T^l(z^{\varepsilon})) + \rho(T^l(z^{\varepsilon})) - g(z^{\varepsilon}, T^l(z^{\varepsilon}))\right] P(z^{\varepsilon}) \tag{14}$$

The approximation becomes exact as the duration of the time period shrinks. The term in brackets is usually referred to as the user-cost for house. Here, equation 14 reveals a user-cost $uc : Z^{\varepsilon} \times T^l(Z^{\varepsilon}) \to \mathbb{R}$ for properties that is both property and owner dependent. I assume a competitive rental market such that, rents, r, equals the gross flow value of occupancy so that in equilibrium, $r(z^{\varepsilon}) = U(z^{\varepsilon}, T^h(z^{\varepsilon}))$ and:

$$\frac{r(z^{\varepsilon})}{P(z^{\varepsilon})} = uc(z^{\varepsilon}, T^l(z^{\varepsilon})) = c(z^{\varepsilon}, T^l(z^{\varepsilon})) + \rho(T^l(z^{\varepsilon})) - g(z^{\varepsilon}, T^l(z^{\varepsilon})) \tag{15}$$

The expected return for a property is the net yield, $\frac{\tilde{r}(z^{\varepsilon})}{P(z^{\varepsilon})} \equiv \frac{r(z^{\varepsilon})}{P(z^{\varepsilon})} - c(z^{\varepsilon}, T^l(z^{\varepsilon}))$, plus expected capital gains. Using equation 15, the expected returns for a property z^{ε} owned by $T(z^{\varepsilon})$ is:

$$E[R(z^{\varepsilon}, T^l(z^{\varepsilon}))] = \frac{r(z^{\varepsilon})}{P(z^{\varepsilon})} - c(z^{\varepsilon}, T^l(z^{\varepsilon})) + g(z^{\varepsilon}, T^l(z^{\varepsilon})) = \rho(T^l(z^{\varepsilon})) \tag{16}$$

Equations 15-16 yield similar features as in Remark 1. In addition, the following is evident:

Remark 2 *Holding fixed $T^l(z^{\varepsilon})$, cross-sectional differences in the time-series variances of different owners' discount rates will lead to cross-sectional differences in the time-series variance of expected returns of the houses they own.*

Furthermore, if shocks disproportionately affect certain owners' discount rates more than others', then the former's houses may have higher return variances as compared to the latter's. For instance, take two owners, A and B. Suppose that (i) A has a lower IMRS than B, perhaps because A is borrowing unconstrained with a lot of liquid wealth and B is constrained (and therefore A's property has lower expected returns than B's) and (ii) there is factor that causes borrowing constraints to tighten exogenously (perhaps from a change in government policy) in such a way that

[15] A similar expression can be found in Piazzesi and Schneider (2016). The focus there is only on the equilibrium price using the characteristics of the marginal owner, whereas here I characterize the willingness to pay of any potential owner-resident pair.

[16] One can readily extend the model to include features like adjustment costs for properties switching between owners. See Halket et al. (2020) for an example.

[17] See, e.g. Nesheim (2006) for formal proof for the general hedonic case.

the IMRS of owner A is unaffected but owner B's IMRS goes up when the policy tightens. Absent other effects, when the policy tightens, expected future returns in A will remain unchanged[18] but expected returns in B will go up and current prices and returns in B will go down. Landvoigt et al. (2015) discuss an instance of this in San Diego. Ex-post, realized price and return volatility for owner B will likely be higher than A. Ex-ante, B may then have more exposure to this credit risk-factor. This could lead it to have still higher returns in equilibrium if households are averse to this risk, *ceteris paribus*. Given a long enough sample, the property owned by B will have higher mean returns and higher variances, so Sharpe Ratios for property B may be either lower or higher than those for property A.

Continuing from above, I assume that net rents, $\tilde{r}(z^{\varepsilon})$, and prices are each well-approximated by a semi-log specification:

$$\log \tilde{r}(z^{\varepsilon}) = \alpha z + \varepsilon_r \tag{17}$$

$$\log P(z^{\varepsilon}) = \beta z + \varepsilon_p \tag{18}$$

where z is a vector of observable characteristics in $Z \subset Z^{\varepsilon}$, $\left(\varepsilon_r, \varepsilon_p\right) \sim N(0, \Sigma)$ and $\Sigma = \begin{pmatrix} \sigma_r^2 & \rho_{rp} \\ \rho_{rp} & \sigma_p^2 \end{pmatrix}$. Predicted user costs or gross property yields are

$$E\left[\frac{\tilde{r}(z^{\varepsilon})}{P(z^{\varepsilon})}|z\right] = \exp\left((\alpha - \beta)z + \frac{\sigma_r^2}{2} + \frac{\sigma_p^2}{2} + \rho_{rp}\right) \tag{19}$$

Using (17), (18) and (19) and assuming one can partition z into elements which are "structure," z_s, and elements which are location, z_l, predicted yields (or user-costs) are the product of three components: $\exp((\alpha_s - \beta_s)z_s)$, $\exp((\alpha_l - \beta_l)z_l)$, and Jensen inequality terms. I will build on this specification further in Section 4 but in the next section I will first introduce the data.

3 Data

My data on rents and prices come from the CoreLogic Multiple Listing Service (MLS) data, which is collected from participating regional boards of realtors that contribute their data to a centralized database. Over 90 boards participate, providing coverage for approximately 56 percent of all active listings nationwide. The data includes both listing and closing prices and rents, as well as property information including street address, square footage of living space, number of bedrooms, bathrooms, and the square footage of the plot of land. The main data used in this paper is the full set of closed sales and rental listings on single family homes and condos.

In addition, I identify a set of properties for which there is both a rental and sale transaction within one year of each other. This provides a direct measure of property-level gross yields and is the basis for estimating some Jensen inequality terms and as a robustness check. I find matching

[18] Though, using the logic of 2.1, A's prices may change.

sale prices for about 21 percent of rental listings from MLS. There are no significant differences in rental rates between properties for which I did and did not find a match.

Historical coverage in the data varies by market. The main analysis is limited to CBSAs for which the data have at least 500 rent transactions without missing information in each year between 2009 and 2021 and for which there is also a sufficient matched sample of property-level yields. Within each CBSA, I only consider a balanced panel of zip codes, and drop zip code-year combinations for which the standard deviation of the log sale price or rental rate is greater than one. Finally, I only consider CBSAs for which I have a balanced panel of at least 23 zip codes. This leaves a sample of 21 CBSAs.

I perform some data cleaning. I remove any properties whose listings comments indicate contain an accessory dwelling. The distributions of building and land square footage contain some outliers. I winsorize the distributions of building square footage at 300 square feet at the lower end and 15,000 at the higher end, and similarly at 500 square feet and 500,000 square feet for land parcel sizes and the number of bedrooms and bathrooms at five.

My main measure of property-level rent is the annual rental income net of property taxes. The MLS data often includes information on property taxes in the listing. In addition, Corelogic has matched the MLS data with data collected from local tax assessors. Whenever possible, I net out the actual dollar amount of property taxes associated with a given property from the annual rental income. For properties for which I do not have property tax information from Corelogic, I estimate property taxes using the average implied property tax rate in that county.

Properties only transact intermittently. Both to reduce noise and to reduce concerns related to sample selection, I expand our sample of sale prices by using estimated sale prices for properties in years in which they did not transact. I do this in two ways. First, I interpolate sale prices for any properties that transact more than once. Second, we estimate sale prices for properties that only transact once, or for years outside the first and last transaction of a property that transacts more than once using annual tract-level house prices indices from the FHFA.

The resulting data set is then combined with a variety of other data sources. I obtain information about the credit scores of people in a given zip code from a major credit bureau using an an anonymous, random panel of households with credit reports. This data also includes mortgage loan amounts for house purchases. Demographic information, such as race, age, and income comes from Decennial Census and American Community Survey. Housing vacancy rates are based on USPS administrative data and made available by the US Department of Housing and Urban Development (HUD).[19]

4 Estimating Returns

I estimate yields, capital gains, and total returns in each zip code for each CBSA in my sample using a hedonic approach and the full sample of sale prices and rents. My methodology builds on

[19] Available here: `https://www.huduser.gov/portal/datasets/usps.html`

that of Kuminoff and Pope (2012), who use the market values of properties to estimate the values of the underlying land and structure.

I mainly focus on estimating variations in various objects (like prices and rents) for properties with constant structural characteristics across different location. For convenience, I will often refer to this as the "location value of a house." I do not estimate explicit land values or rents. This approach has a number of benefits relative to other approaches. For example, one could estimate land values from the sales of empty land parcels. However, the sample size of empty land parcels is small and not random in the sense that they may only be available for sale in certain parts of each city and little-to-no available rent data for land parcels; thus precluding estimates of yields and total returns. Another approach is to estimate structure values from their replacement cost and then attribute the remainder of the market value of the house to land. However, again, this approach would not provide the rental values of land or structure.[20].

I run the following regression year-by-year using our sample of sale and rental transactions:

$$\begin{aligned}\ln(\text{price}_{ijkt}) = &\beta_{0,k,t} + \beta_{1,k,t}\text{Sq. Ft.}_i + \beta_{2,k,t}\text{Sq. Ft.}_i^2 + \beta_{3,k,t}\text{Bedrooms}_i \\ &+ \beta_{4,k,t}\text{Bathrooms}_i + \beta_{5,k,t}\text{Building Age}_{i,k,t} + \beta_{6,k,t}\text{Building Age}_{i,t}^2 \\ &+ \gamma_{jt}\text{Land SqFt}_i + \delta_{jt} + \epsilon_{ijkt}, \end{aligned} \tag{20}$$

where i, j, k, t indexes the property, the zip code, the CBSA and the year, respectively and the dependent variable is either the log of the transaction price in the case of a sale or the log of the annual net rent for rental transactions. The $\gamma_{j,t}$ are separate coefficients on the size of the land plot for each zip code j. The $\delta_{j,t}$ are zip code fixed effects.

To account for the fact that the MLS data is not necessarily a representative sample of rental properties, I construct weights $w_{i,j,k,t}$ for use in the hedonic regressions for log rents using the relative likelihood of a one-unit property built in a given year in zip code j appearing in the MLS data in year t relative to its share of the one-unit renter-occupied housing stock according to the American Community Survey (ACS):

$$w_{i,j,k,t} = \frac{S_{t,j,\text{year built}_i,ACS}}{S_{t,j,\text{year built}_i,MLS}} \tag{21}$$

where $S_{t,j,\text{year built},ACS}$ is the share of all one-unit renter-occupied housing units in the ACS in zip code j that are built in a given year and $S_{t,j,\text{year built},MLS}$ is the corresponding share of rental units in the MLS. For the year 2000, the shares in the numerator are from the 2000 decennial census. I then linearly interpolate the shares between 2000 and 2011 (the first year for which the ACS is available) for each zip code. I use the value of the shares in 2000 for any years pre-2000 and the values of the shares in 2020 for any years post-2020. Any zip code-year-year built combination that is missing a weight is given a weight of one.

[20]The former approach, using vacant land, estimates the land value as "vacant," whereas the latter approach using the "land residual" typically estimates the value of "land as improved." These two measures of land value need not be equal.

I then calculate the market value (and rent) of each property assuming constant-structure characteristics and constant-structure prices. I calculate the predicted value (and rent) both in- and out-of-sample (that is, I predict sale prices for rental properties and rental rates for owner-occupied properties) assuming that each house is a two bedroom, two bath, 2,000 square foot, 10 year old house on a 2,000 square foot plot of land, with the values for β_1–β_6 equal to those estimated using Equation (20) for the year 2015. Thus the only hedonic coefficients that change in the predicted location values over time are the sets of $\beta_{0,k,t}$, γ_{jt} and δ_{jt}. I call the average log of the predicted values (and rents) in a given (j, t): $\ln(\text{price}_{L,j,t})$ $(\ln(\text{rent}_{L,j,t}))$. While the levels of these prices also contain the values of the constant-characteristic, constant-price structure, differences in the log values are attributable to differences in location values or rents over space and/or time.

I compare my estimates of the value of location per square foot to the estimates in Davis et al. (2021), who estimate land values for land used for single family residential purposes using appraisal data from the GSEs. Their approach is to calculate the value of land as the value of the house minus a depreciated replacement cost for the structure. The results of the comparison are in Figure A.2 in the appendix. As the two approaches are estimating different things, the two measures will differ in levels. But, as can be seen in the figures, their correlation is extremely high; the median CBSA has a correlation of 0.81 between my zip code level measure of location value and the land value measure in Davis et al. (2021).

As I will discuss further later, structure and location tend to have different gross yields and different capital gains. Structure requires more periodic maintenance (which in equilibrium raises gross yields) and tends to depreciate (due to age effects), whereas location value has tended to appreciate over long samples. Differences in land share within CBSAs could bias any inference on the causes of differences in returns at the property level. This is another reason why for much of the remainder of the paper I focus on the returns to location, holding structure constant. Using the data from Davis et al. (2021), Figure A.4 shows that land values vary considerably both across and within CBSAs. Higher income areas have higher land shares: Higher income areas have higher structure values and higher land values, but the latter grows with income more. As I will show later, the estimated location values tend to be much more volatile than structure values. Anticipating results below, Figure A.4 shows that the greater returns that I estimate in low income areas are not likely compensation for higher land leverage.

Using the net rents and the values over time for each location, I can then form a panel of returns of properties with the same structure characteristics but different locations. The estimated level of the total return to the entire property may be biased slightly because I do not have a good measures of certain costs, like maintenance. However, since property taxes are well measured, most of the poorly measured (or unobserved costs) likely vary with differences in structure. So though the level of returns may be biased, the cross-sectional variation in returns (and its components) across locations, holding structure fixed, is hopefully not.

Location yields and capital gains are defined by:

$$\text{Yield}_{L,j,t} = \exp\left\{\ln(\text{rent}_{L,j,t}) - \ln(\text{price}_{L,j,t}) + \frac{\sigma^2_{r,k,t} + \sigma^2_{p,k,t} - 2\text{cov}_k(\epsilon_r, \epsilon_p)}{2}\right\} \tag{22}$$

$$\text{Capital Gains}_{L,j,t} = \ln(\text{price}_{L,j,t}) - \ln(\text{price}_{L,j,t-1}). \tag{23}$$

where $\text{cov}_k(\epsilon_r, \epsilon_p)$ is the covariance of the residuals from a single simultaneous regression system using our full sample of properties with matched prices and rents, where both regressions take the form of Equation (20), and $\sigma_{r,k,t}$ and $\sigma_{p,k,t}$ are the standard deviations of the residuals from the full-sample regression for CBSA k, and year t.

The total return to location is calculated as:

$$\text{Total Return}_{L,j,t} = \text{Yield}_{L,j,t-1} + \text{Capital Gains}_{L,j,t}.$$

Though they are not the main focus of the paper, I estimate yields, capital gains and total returns to structures by CBSA by holding location values constant across time, but allowing the estimated value of the structure to vary. Specifically, I take the estimated location price or rent of the zip code with the highest number of housing units in each CBSA. The price or rental value of a structure in any year is then the price or rental value of location in that zip code plus the estimated value based on coefficients $\beta_1, ..., \beta_6$ from the annual rent and price regressions and the same constant characteristics used in the location estimates. Similar to the estimates of location value, the levels of these values are not solely attributed to the structure[21], but any differences are solely attributable to the structure.

Last, I estimate yields, capital gains, and total returns to housing (both structure and location) for a property with the characteristics above by taking the average predicted value from the rent and price regressions, holding all characteristics constant but using the variation in all the hedonic coefficients.

One concern with estimating rents and prices using MLS data is that the selection into listing on the MLS servers may vary across locations. While selection into MLS is not likely an issue for the for-sale sample[22], one may worry that there is selection into MLS for the rental sector. This could cause the unobservable heterogeneity of rentals in the sample to vary by location in a way that the unobservable heterogeneity of prices does not, thus biasing estimates of how yields vary across location.

In order to examine this potential bias, I compare the estimated location yields to the implied zip code-level location yields from the matched sample of property-level rent-price ratios. The matched sample has the same sample for both the rent and price hedonic regressions by construction. The yield estimates from the matched sample are from a single regression of the sample specification

[21] Owing to the log additive specification, one cannot, for instance, add the structure return to the location return to get overall returns

[22] I believe nearly every single-family, non-REO property sold at arms length will be posted on an MLSs in my set of CBSAs.

as in Equation 20, but with property-level yields as the dependent variable. The estimated value holding all characteristics constant as described above are the estimated location-yields. The results are in Figure A.1 in the appendix. While the levels of the two yield estimates are different (owning to the different methods of computing them), the correlation with economic characteristics across areas within each city are very similar.

5 Cross-Sectional Results

In this section, after presenting some summary statistics on yields, capital gains and returns, I examine how my measures of risk and return are correlated in the cross-section with economic characteristics. In the following section, I then explore how the cross-sectional relationship changes over time, including using a series of quasi-natural experiment exploiting changes in CBSA conforming loan limits.

5.1 Summary statistics

Average location yields, capital gains, and total returns for each of the 21 CBSAs in our sample are in Table 1. The unconditional standard deviation of each is calculated as the average time-series standard deviation across zip codes:

$$\sigma_{x,k} = \frac{\sum_j \left(\sqrt{\sum_t (x_{j,k,t} - \mu_j)^2 / N}\right)}{M}$$

where N the number observations for each zip code and is always equal to 13 since I limit our analysis to 2009–2021 (unless otherwise specified), M is the number of zip codes in the CBSA, and μ_j is the average value over time in zip code j. Similar summary statistics for housing and structure returns are in Tables A.1 and A.2 respectively. Information on the variation in structure and location returns across CBSAs are in Table A.3.

The tables validate some priors. Markets in the sunbelt have seen higher capital gains on average over the sample, while other cities (for example, Bridgeport and Hartford) saw lower average capital gains. Structure capital gains are negative, which is consistent with likely aging effects on structure, while location values tend to appreciate. Variation also exists in yields, but there are fewer priors on what to expect.

5.2 Location returns within cities

Figure 2 features binned scatter plots of the estimated average zip code-level average of total returns to location for 2010–2021 against the 2009 average credit score in the zip code for 20 of the 21 CBSAs in our sample[23]. The binned scatter plots are weighted by the number of households

[23]For space reasons, we omit St. Louis but this CBSA is available from the author by request.

living in single unit housing units in the bin according to the 2011 5-year American Community Survey. Figure A.3 contains similar binned scatter plots for yields.

These graphs illustrate a striking pattern. Yields and total returns vary across zip codes within CBSAs. Specifically, they are higher in low credit score zip codes. This is not just true by credit score but also income and race.[24]

To more formally explore the correlations between location returns and local demographic and economic factors, I run a series of univariate regressions of the following form:

$$y_{j,k} = \beta_{0,k} + \beta_1 x_{j,k} + \varepsilon_{j,k} \tag{24}$$

where $y_{j,k}$ is the average annual log yield, average annual capital gain, or average annual total return for zip code j in CBSA k over the years 2010-2020 and $x_{j,k}$ is either the log of median household income, the shares of the population that is Black and Hispanic, the share of properties that are vacant, or the average credit score of the resident population in 2009. Each regressor is normalized using its CBSA mean and the within- CBSA standard deviation. The regression is weighted by the number of single family housing units in the ACS in 2011.

The results, in Table 2, show that yields and returns are higher in low income, low credit and high Black or Hispanic household share zip codes. For example, using the point estimates, a zip code with an average credit score 1 standard deviation larger than its CBSA mean[25] has on average a 1.52 percentage point lower total return to its house's location value. The results are yields are similar to what Damen et al. (2025) finds for multi-family housing in the U.S. in their concurrent work.[26]

Tables A.5 (yields), A.6 (capital gains), and A.7 (total returns) repeat these regressions CBSA by CBSA. Of the 21 CBSAs in our sample, lower income implies significantly[27] higher yields in 16 CBSAs (and negative point estimates for all 21 CBSAs), higher Black (Hispanic) population shares implies higher yields in 15 (13) CBSAs (with 21 (18) having positive point estimates), and all 21 CBSAs have higher yields in zips with lower average credit scores. With two exceptions for Hispanic shares, no CBSA has a significant relationship with these factors in the opposite direction. By contrast, there is only slight evidence in-sample that higher vacancy rates lead to significantly higher yields in zip codes.[28]

The systematic relationships between capital gains and income, race and credit score in an area

[24]Similar figures for these characteristics are omitted here for space but are available from the author by request.

[25]The inter-quartile range relative to each CBSA mean for credit scores for 2009 was 56 points.

[26]Damen et al. (2025) looks at yield spreads by rent deciles and finds about 10–15% yield spread across 90–10 rent deciles. Here, I find a 12.3% yield spread due across zip codes with a one standard deviation difference in area average credit scores.

[27]Here and elsewhere, I use the 5% significance level as the threshold for statistical significance.

[28]This is not surprising as, in short samples, it can be difficult to detect vacancy patterns. For one, there is likely a lot of measurement error in this vacancy rate data. For another, in the "short-run" there may be a negative relationship between yields and vacancies, while in the "long-run" there may be a positive rate (higher average vacancy rates could be compensated for with higher yields gross of vacancy as in Halket and Pignatti Morano di Custoza (2015).)

is slightly noisier. Still, it is clear that the relationship between yields and the economic factors is not counterbalanced by capital gains; zip codes with higher yields do not have lower capital gains. If anything, many of the point estimates have the same sign as their counterparts for yields; just as in Eisfeldt and Demers (2015), zip codes with higher yields often have higher average capital gains in-sample.

Putting these two results together, the relationships between our factors and total returns across zips is very strong; nearly all 21 CBSAs have statistically and economically significant relationships for the income and credit factors. Race is mixed: the share of Black households remains a strong predictor while the share of Hispanics' ability to predict total returns is muddled by the noise in capital gains. A zip code with a one standard deviation higher median income or average credit score than another within the same CBSA can expect anywhere between roughly 0.5 and 3 percentage point lower returns on their property per year. The results on the share of black residents are similarly striking. Except for a handful of cities, areas with high shares of black residents pay higher rents relative to prices so that an area with a one standard deviation higher share of black residents has roughly between 0.5 and 6.3 percentage point higher returns.

Of course, area income, credit and race are all correlated. So I run a series of horse races in Appendix Tables A.9 (yields), A.10 (capital gains), and A.11 (total returns). Credit score remains a very strong predictor of yields in 14 out of the 21 CBSAs even after controlling for income, race and vacancy, and the point estimates are generally much larger than their univariate counterparts. Income and race become less important after controlling for credit and vacancy, though in several of the CBSAs where higher credit does not significantly predict lower yields, race and/or income do. Results for total returns are similar, albeit noisier.

The horse race results do not rule in or out any factor as causing the differences in yields and returns across locations. Credit score may just be a better measure of access to credit or household discount rates credit in our data. Of course, race and income may affect access to credit through a household's credit score (Bayer et al. (2016)) and also race and income may affect a household's idiosyncratic return to real estate within zip codes (see Bayer et al. (2017), Begley and Purnanandam (2021), Ambrose et al. (2020) and Bhutta and Hizmo (2020) for many varying results on this question).

In any case, I am not mainly interested in which factor is the best predictor but rather in what all of these results are potentially signaling about the more fundamental reasons why yields and returns vary predictably. I will pick up with this in Section 6.

5.2.1 Race and returns

To better examine whether race has a separate effect on housing returns, I double sort zip codes by race and credit score, comparing the average return for zip codes in the top and bottom terciles of the shares of Black or Hispanic residents for their CBSA conditional on being in the either the bottom or top tercile of credit scores for their CBSA. I do a similar double sort on share of owner-occupants and race.

Table 3 shows that high Black shares within zip codes affects returns in both low and high credit score zip codes. Differences in Hispanic shares matters for returns as well, though mostly only for high credit score zip codes.[29] Race matters for returns, even once credit score has been accounted for. When we condition instead on share of owner-occupancy, we get similar results: race (i.e. the share of Black households) matters for returns both in high homeownership and low homeownership areas. Interestingly, the effects of race are much larger in low credit (and also low ownership) areas whereas those for ethnicity (shares of Hispanic residence) matters more in high credit and homeowership areas. One possible reason for this may be that, on the margin, race and discrimination in mortgage markets matters more for ”marginal” borrowers based on credit Bayer et al. (2017). Ethnicity may be strongly correlated with other important factors like wealth or unobservable factors, even conditional on credit score, that are more important in high credit/ownership areas whereas race could be correlated with factors that matter more in low homeownership rate locations.[30]

Whether homeownership builds wealth depends on the returns realized by homeowners on their investments. The lower house values in high minority share areas can be seen as a positive since they allow lower-income, lower-wealth (or indeed all) households to afford to purchase better homes. House price levels matter in so far as high prices may restrict some households' ability to become homeowners due to binding borrowing constraints.[31] Our results indicate that homeownership may be a particularly potent way for Black households who are able to buy to build wealth, in part (though not exclusively) by earning an especially high implied yield on their foregone rental payments.

Additionally higher credit score neighborhoods have lower LTV loans on average.[32] This is in large part because of the presence of the Federal Housing Administration (FHA), which insures the credit risk of low-down payment loans for low-income households with the express purpose of increasing access to homeownership. The upshot is that the strong pattern in unlevered returns to location that I document above is not undone by mortgage leverage: households in high-income or high credit areas are less levered (have lower LTV mortgage loans) than households in low-income neighborhoods.

[29] All reported values for the returns in the table are averages from 2010–2021 for each zip code, centered relative to the weighted mean of the CBSA. So the reported returns levels are not all that informative. More relevant is the difference across bottom and top terciles in returns.

[30] Aliprantis et al. (2022) finds no evidence that wealth explains sorting across neighborhoods by race and instead argues for preference-based explanations for sorting. Bhutta and Hizmo (2020) finds that, conditional on other characteristics, different races receive comparable mortgage offers but that minority applicants are more likely to choose offers with higher discount point. This may indicate that minority borrows have higher OCCs, even after conditioning on these other factors.

[31] Indeed Amior and Halket (2014) finds a strong relationship at the MSA level between price levels and homeownership rates.

[32] Results from data on this available by request.

5.3 Risk and Returns

I calculate several measures to see if differences in returns across zip codes are correlated with differences in risk. My sample has a wide panel of returns but a relatively short one. This makes the estimated time-series standard deviations of returns, which themselves are fitted from estimates, quite noisy. Nevertheless one can discern some patterns in the results. Two measures of risk are the standard deviations in the year-on-year log differences in location rents and in capital gains. Univariate regressions of these measures on credit score are in Table 4. Results using income and race are similar.

Point estimates indicate that for many CBSAs there were higher realized rent and capital gains volatilities in areas with lower credit scores. Putting these results together, higher credit score areas within CBSAs tend to have lower returns but less volatile rents and capital gains. These results are consistent with changes in the way credit affects expected returns over time, particularly in low credit areas, leading to higher realized volatility and lower Sharpe Ratios[33]). Higher return volatility could cause some of the observed return premium in low credit areas, however the point estimates seem relatively small. Below I will delve further into a possible joint cause of higher returns and higher capital gains volatility after first looking at the risk-return relationship through the lens of a standard CAPM regression.

5.3.1 CAPM

To understand how much location returns in each CBSA vary with market returns and risk, I run the following regression separately for each CBSA:

$$\begin{aligned}\text{Total Return}_{L,j,k,t} =& \beta_{0,k} + \beta_{1,k} R_m - R_f + \beta_{2,k} \text{Credit Score}_{j,k,2010} \\ &+ \beta_{3,k}(R_m - R_f) \times \text{Credit Score}_{j,k,2010} + \beta_{4,k} \text{CBSA Return}_{L,k,t} \\ &+ \beta_{5,k} \text{CBSA Return}_{L,k,t} \times \text{Credit Score}_{j,k,2010} + \epsilon_{j,k,t}\end{aligned} \tag{25}$$

where Total $\text{Return}_{L,j,k,t}$ is the total return to location in zip code j and CBSA k in year t. The net market return $(R_m - R_f)$ is from the Fama-French data library.[34] The CBSA return is the residual of the average total return to location in the CBSA regressed on national house price growth and the credit variable is normalized to have zero mean and unit standard deviations for each CBSA. It is conventional to include metro area housing returns in CAPM regressions of local returns. I remove national housing returns from the metro return measure so as not to confound the detection of a relationship between credit and returns if changes in the relationship between credit and expected returns are national.

The results are in Table A.12. β_2 measures whether ex-ante area credit scores can be used

[33]Eisfeldt and Demers (2015) use a different sample and also finds that Sharpe ratios are higher for properties with higher rental yields.

[34]Available here.

to predict average returns to location, after controlling for potential differences in some risks (i.e. "α"). Using the point estimate, credit score negatively affects "α" in 19 out of 21 CBSAs. In 14 of these CBSAs the relationship is significant at the five percent (or better) level. In these, a one standard deviation higher average local credit score implies 1.0 to 2.3 percentage points in returns.

Meanwhile market betas (β_3) are low. The betas on CBSA net returns are much higher and universally significant, with (in many CBSAs) zip codes with lower credit having higher betas. In summary, consistent with our findings above and below, low credit areas tend to load on aggregate (city or national) shocks more, leading to more volatile returns, and to have higher average returns as well as well. As I discuss below, aggregate shocks that differentially affect how different households access credit or otherwise discount the future can generate this pattern.

6 Cross-Sectional Results Over Time

My hypothesis is that differences in the discount rates of households caused perhaps by differences in their opportunity costs of credit cause a difference in expected housing returns and yields within markets. My best proxy for borrowing costs is lagged credit score.

In this section, I examine how the credit channel of housing returns varies over time. I do this for two reasons. Firstly, it connects the results with narratives around the housing boom and bust circa 2005-2015 and related macroprudential policy decisions. Secondly, I exploit the time-series dimension to show that the results are robust to plausible measurement error concerns.

The first set of results shows that the measured yield spreads are largest after the 2009 housing bust when credit was especially scarce for low credit households. Meanwhile proxies for spreads in costs, such as spreads in rental vacancy rates, show no such increase. However, because one cannot fully observe all potential changes in cost spreads, the second set of results exploits a series of quasi-natural experiments in order to create an instrumental variable that is unlikely to be correlated with cost spreads but is tightly connected to opportunity costs of credit.

6.1 Changes in credit and risk and returns

Historically, particularly in the last 20 years, the relationship between credit score and OCC has likely varied a lot over time. Figure A.6 shows the share of mortgage originations (not weighted by dollar value) that went to households with a credit score lower than 680 in any given year. During the boom period from 2003-2007, this share rose in all CBSAs in our sample, usually by more than 20 percentage points, highlighting the relatively weak relationship between credit score and OCC then. Around the onset of the Great Recession, credit standards tightened (Goodman et al. (2018)) and the share of mortgages going to lower credit households fell dramatically. Indeed in no year since 2010 has any CBSA in our sample had a share higher than 20 percent.

To explore how relationship between the ex-ante characteristics of a zip code is related to realized returns, yields and capital gains over time, I repeat the regressions in equation 24 of returns, log yields and capital gains on credit score but allow the effects to vary with time and include year fixed effects as well:

$$y_{j,k,t} = \beta_{0,k} + \beta_{1,t} x_{j,k,t-2} + \delta_t + \varepsilon_{j,k,t} \tag{26}$$

Figure 3 shows the time-series profile of the univariate relationship between location yields and lagged credit scores. Prior to 2007, just before the Great Recession, marginal effects are slightly negative. In this period, credit score has only a small affect on yields to housing across submarkets, consistent with findings in Justiniano et al. (2022) and Davis et al. (2022) that differences in mortgage costs across borrowers were historically low from 2004-2006.

After 2007, yield spreads widen in most CBSAs as access to credit (at least as proxied by mortgage origination data) narrows and risk-based mortgage pricing returns. Though the spreads eventually narrow again in some CBSAs, like Atlanta, Tucson and Tampa, in others the spread persists throughout the sample. Regardless, for most CBSAs, the marginal effect remained significantly negative up to at least 2021.

Figure 4 shows the effect of the boom and bust from another angle. Here we plot the marginal effect of credit score on the yearly capital gain to location in the zip code for each year. There are striking large positive marginal effects due to the onset of the Great Recession. In the bust years, zip codes with low credit scores saw much larger falls in house prices than their higher credit score counterparts. Landvoigt et al. (2015) finds that low quality houses rose more during the boom and fell more during the bust in San Diego. Since 2016, the patter has reversed, with low credit areas growing faster. This may or may not be related to the entry of institutional investors (e.g. Garriga et al. (2022) and many others).

The role of mortgage lending in the housing boom and bust of the 2000s has been discussed extensively (e.g. Favara and Imbs (2015), Justiniano et al. (2015), Landvoigt (2017), Favilukis and Van Nieuwerburgh (2021) and Griffin et al. (2021)) but not conclusively (e.g. recently Conklin et al. (2020)). Some of the debates around the causes of the boom revolve around whether the exogenous expansion in credit supply was concentrated in particular areas (e.g. lending to subprime borrowers as in Mian and Sufi (2009) or were more widespread expansions (e.g. Conklin et al. (2020)). I do not take a stand here on this debate as I do not need to for my hypothesis. Instead I simply propose that, should differences in the OCC across borrowers decline, changes in house prices would likely be greater in areas where households were more likely to have ex ante high OCCs. My results show that different areas' house prices, rents and returns may respond differently when hit with potentially the same shock. This may be true even if the areas have the same house supply elasticities, which may lead to questions about the validity of some instruments commonly used in the literature to disentangle the causal direction of lending and property prices in the boom.[35]

[35] For example, Guren et al. (2020) suggests using historical differences in local price sensitivities to regional house demand shocks as a potential instrument for current changes in house prices. As discussed there and also in Conklin et al. (2020), the validity of the instrument depends on controlling for the other channels that may cause differences in local price variation.

6.2 Robustness to measurement error

The results heretofore show that credit score (among other variables) is correlated with our measured yields and returns in the cross-section. One concern may be that credit scores (and other variables) may be correlated with the measurement error in the yields and returns measures. In this respect, the paramount concern is that unobserved spreads in costs may be correlated with credit score in such as way as to potentially explain the spread in my measure of yields: for instance, if expected future vacancy or rental non-payment rates are higher in low credit zip codes[36].

6.2.1 Using vacancies

While I do not have good data on non-payment of rents, I do control for vacancy rates in Table A.9 and still find the negative yield spread with respect to local economic and demographic variables. This should handle any measurement error that is correlated with vacancy rates.

6.2.2 Using equilibrium logic

Figure 3 shows that the spread in our measured yields was widest from about 2010 to 2013. We have argued that at least some of the this change in spreads is driven by the likely change in OCC around the same time. Suppose instead that our measured change in spreads was driven by a change in the spread in the unobserved cost of maintenance during this same time period. The chief culprit could perhaps be expected landlord-related costs that go up more in low credit areas: for instance, expected future rental non-payment rates or tenant property abuse going up more in low credit areas than in high credit area[37]. Figure 4 shows that this mooted change was capitalized into prices and not compensated for by higher rents.

Following the logic of Henderson and Ioannides (1983) and Halket et al. (2020), if landlord-related costs go up, then home ownership rates should go up, everything else equal. In other words, if increases in spreads in expected rental non-payment explained the yield spreads found in this paper, we should also see a relative increase in home ownership rates in precisely the areas where rental non-payment rates were going up most. Instead we see the opposite. Figure 5 plots the spread in home ownership rates using the American Community Survey using the same regressions setup 24. The data come from the ACS 5-year survey where the plotted year reflects that year and the previous four years' surveys, so there is considerable time-series correlation across the cross-sectional regressions. Nevertheless, it is clear that ownership spreads if anything widened from the first survey in 2011 (which reflects data from 2007-11) to 2016 (which includes data from 2012-16). This should be surprising: home ownership rates fell in low credit areas following the bust. This fact is far more consistent with low credit households having trouble attaining a mortgage than these households further troubling their landlords.

[36]Halket and Pignatti Morano di Custoza (2015) shows that this may be the case in the cross-section.

[37]In order for, e.g., changes in non-payment rates to affect current prices, the changes must come in the future and be currently forecastable.

6.3 A series of quasi-natural experiments

Finally, I explore causally how changes in the access to credit affects yields and returns by following Loutskina and Strahan (2015), Adelino et al. (2022) and Greenwald and Guren (2021) in using the differential impact of changes in conforming loan limits (CLL). Interest rates on "conforming" mortgages backed by Fannie Mae and Freddie Mac are typically lower than the rates on non-conforming mortgages due to various subsidies. The CLLs, which generally vary over time and across CBSAs, dictate the maximum size a mortgage may have and still potentially qualify as conforming. An increase in the CLL in a CBSA thereby lowers the cost of credit for households that can newly access conforming mortgages. This change will tend to be more valuable when the national spread in mortgage rates between conforming and non-conforming mortgages is relatively high.

My hypothesis is that within a CBSA, locations (zip codes) where many mortgages were recently originated near the CLL should see a relative decrease in their yields when the CLL goes up if the spread between non-conforming mortgage rates and conforming mortgage rates is relatively large.

To test this hypothesis I use the two-year lagged share of loan originations (by number) within 5 percent (on either side) of the new county-level conforming loan limit according to the data in the credit panel as my measure of treated mortgages. A measure of the difference in conforming and non-conforming mortgage costs is the jumbo-conforming spread, calculated using the difference in the national annual average 30-year fixed-rate jumbo rate according to Bank Rate and the average annual 30-year fixed-rate mortgage rate from Freddie Mac. I standardize the spread to have mean zero and unit standard-deviation over the sample.

Table 5 shows the results from an OLS regression of changes in zip code location log yields on our interacted variables of interest as well as a host of controls. Zip codes with 1 percent of their (lagged) originations near the new CLL for their county have about 0.6 percentage point lower yields when the jumbo-conforming spread is one standard deviation above its mean. The results are qualitatively similar if we use the (lagged) total share of non-conforming mortgage origination instead of just those originations near the CLL.[38] The effect is robust to controlling for local variation in lagged average credit scores, race and household income. The effect is also fairly robust across CBSAs.

Following the logic of Campbell and Shiller (1988) and Campbell et al. (2009):

$$\log yield_{j,k,t} = q_{j,k} + \mathcal{I}_{j,k,t} - \mathcal{G}_{j,k,t} \tag{27}$$

where q is a constant that can vary over locations, and $\mathcal{I}_{j,k,t}$ and $\mathcal{G}_{j,k,t}$ are the expected present values of the sums of future discount rate premia for housing and future rent growth, respectively. Lower yields can be caused by lower discount rates or higher expected future rent growth. To rule

[38]The estimated effect using all non-conforming mortgages is likely biased towards 0 as not all non-conforming mortgages would be treated by the change in CLL. All results are similar if we instead use log yields (and not changes of log yields) as the regressand and include zip code fixed effects as regressors. These are available from the author upon request.

out the latter, I also regress the one year growth in rents on the same explanatory variables and report those results in Table 5. The treatment has no statistically significant explanatory power on future rents. Therefore it seems likely that the treatment variable affects yields through changes in the discount rate applied to housing in the area.

7 Conclusion

I measure the returns to housing and land in a large set of metropolitan areas in the United States. I find large dispersions in the average returns and yields to land that are correlated with many important demographic and economic characteristics. Variables which may proxy for the opportunity cost of credit are especially correlated with returns: areas where residents may high OCCs have higher average returns. Return and yield spreads widen during periods when measured of the cost or access to credit widened as well. While some measures of risk are also correlated with returns, the return spread is not likely explained as compensation for bearing extra risk but rather as evidence of segmented housing markets.

In this paper, I have shown that changes in OCC lead to changes in the dispersion of yields and returns across areas. The degree to which different owners with perhaps lower discount rates are willing to enter areas with higher yields can be another driver of dispersion (or convergence) in yields and returns. Some high yield areas may see low cost of credit households move in, ”gentrifying” the area (e.g. Guerrieri et al. (2013)). Small landlords ”searching for yield” may enter when and where yields are high too (e.g. Garriga et al. (2022)). Or landlords’ technology for operating single-family rentals may improves (e.g. ”prop tech” landlords) so that they find it sufficiently profitable to purchase more housing in high yield areas, driving prices up and yields down. To the extent that these factors may partly explain the slight convergence in yield spreads in the latter half of the 2010s seen in Figure 3 remains an interesting avenue of future research.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino (2022). “Credit supply and house prices: Evidence from mortgage market segmentation.” *SSRN Electronic Journal*.

Aliprantis, Dionissi, Daniel R. Carroll, and Eric R. Young (2022). “What explains neighborhood sorting by income and race?” *Journal of Urban Economics*, p. 103508.

Amaral, Francisco, Martin Dohmen, Sebastian Kohl, and Moritz Schularick (2021). “Superstar returns.” Technical Report 999, Federal Reserve Bank of New York. URL `https://ideas.repec.org/p/fip/fednsr/93542.html`.

Ambrose, Brent W., James N. Conklin, and Luis A. Lopez (2020). “Does borrower and broker race affect the cost of mortgage credit?” *The Review of Financial Studies*, 34(2), pp. 790–826.

Amior, Michael and Jonathan Halket (2014). "Do households use homeownership to insure themselves? evidence across u.s. cities." *Quantitative Economics*, 5(3), pp. 631–674. doi: 10.3982/qe346.

Bayer, Patrick, Fernando Ferreira, and Stephen L. Ross (2016). "The vulnerability of minority homeowners in the housing boom and bust." *American Economic Journal: Economic Policy*, 8(1), pp. 1–27.

Bayer, Patrick, Fernando Ferreira, and Stephen L. Ross (2017). "What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders." *The Review of Financial Studies*, 31(1), pp. 175–205.

Begley, Taylor A. and Amiyatosh Purnanandam (2021). "Color and credit: Race, regulation, and the quality of financial services." *Journal of Financial Economics*, 141(1), pp. 48–65.

Bernstein, Asaf, Matthew T. Gustafson, and Ryan Lewis (2019). "Disaster on the horizon: The price effect of sea level rise." *Journal of Financial Economics*, 134(2), pp. 253–272.

Bhutta, Neil and Aurel Hizmo (2020). "Do minorities pay more for mortgages?" *The Review of Financial Studies*, 34(2), pp. 763–789.

Bosshardt, Joshua, Marco Di Maggio, Ali Kakhbod, and Amir Kermani (2023). "The credit supply channel of monetary policy tightening and its distributional impacts." Technical report, FHFA Staff Working Paper Series.

Campbell, John and Joao Cocco (2003). "Household risk management and optimal mortgage choice." *The Quarterly Journal of Economics*, 118(4), pp. 1449–1494.

Campbell, John Y. and Joao F. Cocco (2007). "How do house prices affect consumption? Evidence from micro data." *Journal of Monetary Economics*, 54(3), pp. 591–621.

Campbell, John Y. and Robert J. Shiller (1988). "The dividend-price ratio and expectations of future dividends and discount factors." *The Review of Financial Studies*, 1(3), pp. 195–228.

Campbell, Sean D., Morris A. Davis, Joshua Gallin, and Robert F. Martin (2009). "What moves housing markets: A variance decomposition of the rent-price ratio." *Journal of Urban Economics*, 66(2), pp. 90–102.

Case, Karl E. and Robert J. Shiller (1989). "The efficiency of the market for single-family homes." *American Economic Review*, 79(1), pp. 125–137.

Case, Karl E. and Robert J. Shiller (1990). "Forecasting prices and excess returns in the housing market." *Real Estate Economics*, 18(3), pp. 253–273.

Cochrane, John H. (2011). "Presidential address: Discount rates." *The Journal of Finance*, 66(4), pp. 1047–1108.

Conklin, James W., Scott Frame, Kristopher Gerardi, and Haoyang Liu (2020). "Villains or scapegoats? The role of subprime borrowers in driving the U.S. housing boom." Working paper 2013, Federal Reserve Bank of Dallas.

Damen, Sven, Matthijs Korevaar, and Stijn Van Nieuwerburgh (2025). "An alpha in affordable housing?"

Davis, Morris A., William D. Larson, Stephen D. Oliner, and Jessica Shui (2021). "The price of residential land for counties, ZIP codes, and census tracts in the United States." *Journal of Monetary Economics*, 118, pp. 413–431.

Davis, Morris A, William D Larson, Stephen D Oliner, and Benjamin R Smith (2022). "A quarter century of mortgage risk." *Review of Finance*, 27(2), pp. 581–618.

Demers, Andrew and Andrea L. Eisfeldt (2022). "Total returns to single-family rentals." *Real Estate Economics*, 50(1), pp. 7–32.

Demyanyk, Yuliya and Otto Van Hemert (2009). "Understanding the subprime mortgage crisis." *Review of Financial Studies*, 24(6), pp. 1848–1880.

Diamond, Rebecca and William F Diamond (2024). "Racial differences in the total rate of return on owner-occupied housing."

Eichholtz, Piet, Matthijs Korevaar, Thies Lindenthal, and Ronan Tallec (2021). "The total return and risk to residential real estate." *The Review of Financial Studies*, 34(8), pp. 3608–3646.

Eisfeldt, Andrea and Andrew Demers (2015). "Total returns to single family rentals." Working paper 21804, National Bureau of Economic Research.

Eisfeldt, Andrea L and Adriano A Rampini (2009). "Leasing, ability to repossess, and debt capacity." *The Review of Financial Studies*, 22(4), pp. 1621–1657.

Favara, Giovanni and Jean Imbs (2015). "Credit supply and the price of housing." *American Economic Review*, 105(3), pp. 958–992.

Favilukis, Jack and Stijn G. Van Nieuwerburgh (2021). "Out-of-town home buyers and city welfare." *The Journal of Finance*, 76(5), pp. 2577–2638.

Garriga, Carlos, Pedro Gete, and Athena Tsouderou (2022). "The economic effects of real estate investors."

Garriga, Carlos, Rodolfo Manuelli, and Adrian Peralta-Alva (2019). "A macroeconomic model of price swings in the housing market." *American Economic Review*, 109(6), pp. 2036–2072.

Giacoletti, Marco (2021). "Idiosyncratic risk in housing markets." *The Review of Financial Studies*, 34(8), pp. 3695–3741.

Goetzmann, William N., Christophe Spaenjers, and Stijn G. Van Nieuwerburgh (2021). "Real and private-value assets." *The Review of Financial Studies*, 34(8), pp. 3497–3526.

Goodman, Laurie S., Bing Bai, and Wei Li (2018). "Real denial rates: A new tool to look at who is receiving mortgage credit." *Housing Policy Debate*, 29(5), pp. 795–819.

Gorea, Denis, Oleksiy Kryvtsov, and Marianna Kudlyak (2022). "House price responses to monetary policy surprises: Evidence from the u.s. listings data." *SSRN Electronic Journal*.

Greenwald, Daniel and Adam Guren (2021). "Do credit conditions move house prices?" Working paper 29391, National Bureau of Economic Research.

Griffin, John M., Samuel Kruger, and Gonzalo Maturana (2021). "What drove the 2003-2006 house price boom and subsequent collapse? Disentangling competing explanations." *Journal of Financial Economics*, 141(3), pp. 1007–1035.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2013). "Endogenous gentrification and housing price dynamics." *Journal of Public Economics*, 100, pp. 45–60.

Guren, Adam M., Alisdair McKay, Emi Nakamura, and Jón Steinsson (2020). "Housing wealth effects: The long view." *The Review of Economic Studies*, 88(2), pp. 669–707.

Halket, Jonathan, Lars Nesheim, and Florian Oswald (2020). "The housing stock, housing prices, and user costs: the roles of location, structure, and unobserved quality." *International Economic Review*, 61(4), pp. 1777–1814.

Halket, Jonathan and Matteo Pignatti Morano di Custoza (2015). "Homeownership and the scarcity of rentals." *Journal of Monetary Economics*, 76(C), pp. 107–123.

Han, Lu (2013). "Understanding the Puzzling Risk-Return Relationship for Housing." *The Review of Financial Studies*, 26(4), pp. 877–928. Publisher: Oxford University Press.

Henderson, J. V. and Y. M. Ioannides (1983). "A model of housing tenure choice." *American Economic Review*, 73(1), pp. 98–113.

Higgins, Brian E. (2023). "Racial segmentation in the us housing market." Technical report, Stanford University. URL `https://web.stanford.edu/~higginsb/Higgins_JMP_RacialSegmentation.pdf`.

Humphries, John Eric, Scott Nelson, Dam Linh Nguyen, Winnie van Dijk, and Daniel Waldinger (2024). *Nonpayment and Eviction in the Rental Housing Market*. doi:10.3386/w33155.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti (2015). "Household leveraging and deleveraging." *Review of Economic Dynamics*, 18(1), pp. 3–20.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti (2022). "The mortgage rate conundrum." *Journal of Political Economy*, 130(1), pp. 121–156.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante (2020). "The housing boom and bust: Model meets evidence." *Journal of Political Economy*, 128(9), pp. 3285–3345.

Kuminoff, Nicolai V. and Jaren C. Pope (2012). "The value of residential land and structures during the great housing boom and bust." *Land Economics*, 89(1), pp. 1–29.

Landvoigt, Tim (2017). "Housing demand during the boom: The role of expectations and credit constraints." *The Review of Financial Studies*, 30(6), pp. 1865–1902.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2015). "The housing market(s) of san diego." *American Economic Review*, 105(4), pp. 1371–1407.

Lewellen, Stefan and Emily Williams (2021). "Did technology contribute to the housing boom? evidence from MERS." *Journal of Financial Economics*, 141(3), pp. 1244–1261.

Loutskina, Elena and Philip E. Strahan (2015). "Financial integration, housing, and economic volatility." *Journal of Financial Economics*, 115(1), pp. 25–41.

Lustig, Hanno and Stijn G. Van Nieuwerburgh (2010). "How much does household collateral constrain regional risk sharing?" *Review of Economic Dynamics*, 13(2), pp. 265–294.

Lustig, Hanno N. and Stijn G. Van Nieuwerburgh (2005). "Housing collateral, consumption insurance, and risk premia: An empirical perspective." *The Journal of Finance*, 60(3), pp. 1167–1219.

Mian, Atif and Amir Sufi (2009). "The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis." *The Quarterly Journal of Economics*, 124(4), pp. 1449–1496.

Muth, Richard F (1966). *Urban residential land and housing markets*. Institute for Urban and Regional Studies, Washington University.

Nathanson, Charles (2020). "Trickle-Down Housing Economics." URL `https://charlesnathanson.com/s/TrickleDownHousingEconomics.pdf`.

Nesheim, Lars (2006). "Hedonic price functions." Technical Report 18/06, Institute for Fiscal Studies, Centre for Microdata Methods and Practice. URL `https://ideas.repec.org/p/ifs/cemmap/18-06.html`.

Piazzesi, Monika and Martin Schneider (2016). "Chapter 19 - Housing and macroeconomics." In *Handbook of Macroeconomics*, volume 2, pp. 1547–1640. Elsevier.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel (2020). "Segmented housing search." *American Economic Review*, 110(3), pp. 720–759.

Plazzi, Alberto, Walter Torous, and Rossen Valkanov (2010). "Expected returns and expected growth in rents of commercial real estate." *The Review of Financial Studies*, 23(9), pp. 3469–3519.

Ríos-Rull, José-Víctor and Virginia Sánchez-Marcos (2008). "An aggregate economy with different size houses." *Journal of the European Economic Association*, 6(2-3), pp. 705–714.

Sinai, Todd and Nicholas S. Souleles (2005). "Owner-occupied housing as a hedge against rent risk." *The Quarterly Journal of Economics*, 120(2), pp. 763–789.

Sweeney, James L. (1974). "A commodity hierarchy model of the rental housing market." *Journal of Urban Economics*, 1(3), pp. 288–323.

	Average			Std. Dev.		
	Yield	Cap Gains	Tot Return	Yield	Cap Gains	Tot Return
Atlanta, GA	7.05	5.94	12.99	1.40	8.51	8.91
Boston, MA-NH	6.42	4.13	10.55	1.12	3.16	2.69
Bridgeport, CT	4.52	1.03	5.54	0.62	4.12	4.22
Charlotte, NC-SC	7.01	6.01	13.02	1.85	5.10	6.02
Chicago, IL-IN-WI	6.32	1.07	7.39	0.89	5.06	4.58
Dallas, TX	7.04	6.80	13.85	1.07	4.33	4.35
Detroit, MI	8.50	4.60	13.11	1.08	7.69	7.55
Hartford, CT	6.84	1.31	8.15	1.42	2.67	2.72
Houston, TX	6.74	5.40	12.14	0.90	3.82	3.82
Jacksonville, FL	6.46	3.70	10.15	0.96	7.38	6.85
Los Angeles, CA	5.03	4.25	9.28	0.66	4.60	3.41
Miami, FL	6.37	4.67	11.04	1.18	9.21	8.67
Orlando, FL	7.91	3.92	11.83	1.75	13.50	12.93
Phoenix, AZ	5.50	5.07	10.57	0.99	8.16	8.05
Riverside, CA	6.50	3.77	10.26	0.96	7.37	7.16
San Diego, CA	5.77	4.37	10.15	0.72	3.52	3.26
San Francisco, CA	4.32	2.85	7.17	0.88	6.04	5.85
St. Louis, MO-IL	6.26	2.71	8.96	1.09	3.46	3.59
Tampa, FL	7.92	4.36	12.28	2.13	10.59	9.02
Tucson, AZ	4.93	3.27	8.20	0.95	5.81	5.79
Virginia Beach, VA-NC	8.01	2.37	10.37	0.59	3.96	3.20

Table 1: Summary Statistics for Returns to Location by CBSA. *Note:* Values are weighted averages of the average and standard deviation (over time) of zip code-level yields, capital gains, and total returns between 2009 and 2021, where the weights are the number of households that are living in single-unit structures. *Source:* Corelogic MLS and the American Community Survey.

	Total Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Median HH Income) (Normalized)	-0.0109***					-0.000726
	(0.00177)					(0.00204)
Share Black (Normalized)		0.0111***				0.00153
		(0.00121)				(0.00139)
Share Hispanic (Normalized)			0.00762***			-0.00233
			(0.00193)			(0.00154)
Average Credit Score (Normalized)				-0.0152***		-0.0150***
				(0.00148)		(0.00200)
Vacancy Rate (Normalized)					0.00497*	-0.000361
					(0.00242)	(0.000910)
Constant	0.119***	0.119***	0.119***	0.119***	0.119***	0.119***
	(0.000103)	(0.0000114)	(0.0000386)	(0.00000499)	(0.000171)	(0.000127)
N	1556	1556	1556	1556	1556	1556
R2	0.70	0.71	0.63	0.80	0.61	0.81
CBSA FE	✓	✓	✓	✓	✓	✓

	Log Yields					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Median HH Income) (Normalized)	-0.0795***					0.00394
	(0.0183)					(0.0222)
Share Black (Normalized)		0.0937***				0.00310
		(0.0147)				(0.0180)
Share Hispanic (Normalized)			0.0452***			-0.0497***
			(0.0149)			(0.0151)
Average Credit Score (Normalized)				-0.123***		-0.154***
				(0.0171)		(0.0315)
Vacancy Rate (Normalized)					0.0342	-0.00901
					(0.0295)	(0.0167)
Constant	-2.758***	-2.761***	-2.763***	-2.763***	-2.760***	-2.763***
	(0.00107)	(0.000137)	(0.000299)	(0.0000577)	(0.00209)	(0.00136)
N	1556	1556	1556	1556	1556	1556
R2	0.57	0.60	0.51	0.68	0.50	0.70
CBSA FE	✓	✓	✓	✓	✓	✓

	Capital Gains					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Median HH Income) (Normalized)	-0.00481***					-0.00163
	(0.000748)					(0.00101)
Share Black (Normalized)		0.00301***				-0.000367
		(0.000943)				(0.000841)
Share Hispanic (Normalized)			0.00480***			0.000923
			(0.00109)			(0.000992)
Average Credit Score (Normalized)				-0.00592***		-0.00480***
				(0.000982)		(0.000920)
Vacancy Rate (Normalized)					0.00110	-0.00108
					(0.000826)	(0.00113)
Constant	0.0528***	0.0525***	0.0524***	0.0525***	0.0526***	0.0525***
	(0.0000436)	(0.00000884)	(0.0000219)	(0.00000331)	(0.0000585)	(0.0000747)
N	1556	1556	1556	1556	1556	1556
R2	0.79	0.76	0.78	0.82	0.74	0.82
CBSA FE	✓	✓	✓	✓	✓	✓

Table 2: Determinants of Log Yields, Capital Gains and Total Returns. *Note:* Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1. Standard errors are clustered by CBSA. *,**,and *** indicate significance at the 10%, 5%, and 1% levels, respectively. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

	Bottom Tercile		Top Tercile			
	N	Avg. Return	N	Avg. Return	P	P_L
Share Black w/n Low Credit Score	168	1.06	156	1.84	0.00	0.00
Share Black w/n High Credit Score	216	-1.65	176	-1.25	0.01	0.00
Share Black w/n Low Owner-Occupancy	139	-0.93	245	-0.42	0.04	0.02
Share Black w/n High Owner-Occupancy	232	-0.28	128	0.43	0.00	0.00
Share Hispanic w/n Low Credit Score	186	1.34	136	1.38	0.87	0.44
Share Hispanic w/n High Credit Score	227	-1.58	167	-1.36	0.13	0.06
Share Hispanic w/n Low Owner-Occupancy	200	-0.70	206	-0.50	0.40	0.20
Share Hispanic w/n High Owner-Occupancy	248	-0.16	88	0.32	0.01	0.01

Table 3: Double Sort of Total Location Returns. *Note:* All sorts are weighted terciles where the weight is the number of households in each zip code in 2010, so the top row is comparing the top tercile by share of black residents within the bottom tercile by credit score, where the terciles are within CBSA. Total Returns are in %-form. All values for the total returns are averages from 2010–2021 for each zip code, centered relative to the weighted mean of the CBSA. P is the two-sided P value from a standard t test for differences in means. P_L is the one sided P value testing whether the bottom tercile for the second sort has a lower total return than the top tercile. *Source:* Authors' calculations using data from a major credit bureau; 2010 Decennial Census; Corelogic MLS.

	Sharpe Ratio	Rent Volatility	Capital Gain Volatility
	(1)	(2)	(3)
Average Credit Score (Normalized)	0.167***	-0.0129***	-0.0116***
	(0.0374)	(0.00216)	(0.000588)
Constant	2.277***	0.131***	0.0627***
	(0.0317)	(0.00184)	(0.000500)
N	1440	1513	1513
R2	0.37	0.21	0.65
CBSA FE	✓	✓	✓

Table 4: Regressions of Zip Code Measures of Risk on Zip Code Average Credit Score. *Note:* Values are coefficients from regressions of each risk measure on the average zip code-level Credit Score as of 2009 normalized so that the score is CBSA-mean zero and has within-CBSA standard deviations equal to 1. The Sharpe ratio for location is calculated as the average total return holding structure constant in each zip code over the standard deviation of that return between 2010 and 2020. Location rent volatility is the standard deviation of the annual log difference in rents holding the rent due to structure constant. Location capital gains volatility is the standard deviation of the property capital gains holding the value of structure constant. Regressions are weighted by the number households residing in one-unit housing units in 2011. *Source:* Authors' calculations using the Corelogic MLS data and data from a major credit bureau.

	Δ ln(Yield$_t$)				Δ ln(Rent$_{t+1}$)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated Loan Originations (%)	0.00101	-0.000200	-0.000135		-0.0852	-0.170	-0.0978	
	(0.00148)	(0.00156)	(0.00162)		(0.149)	(0.157)	(0.163)	
Treated Loan Originations (%) $\times$ Jumbo Conforming Spread	-0.00666***	-0.00659***	-0.00660***		0.159	0.164	0.162	
	(0.00132)	(0.00132)	(0.00132)		(0.133)	(0.133)	(0.133)	
Non-Conforming Originations (%)				-0.000259				-0.00371
				(0.000537)				(0.0543)
Non-Conforming Originations (%) $\times$ Jumbo Conforming Spread				-0.00256***				0.0715*
				(0.000407)				(0.0412)
Lagged Average Risk Score (Normalized)		0.00225**	0.00239*	0.00245*		0.157*	0.311**	0.308**
		(0.000922)	(0.00133)	(0.00133)		(0.0930)	(0.134)	(0.134)
ln(Average Household Income)			-0.000515	0.000191			-0.565	-0.623*
			(0.00351)	(0.00360)			(0.354)	(0.364)
N	12731	12731	12731	12731	12887	12887	12887	12887
R2	0.56	0.56	0.56	0.56	0.46	0.46	0.46	0.46
CBSA $\times$ Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var. (%)	-.01	-.01	-.01	-.01	3.14	3.14	3.14	3.14
Std. Dev. of Dep. Var.	.14	.14	.14	.14	13.	13.	13.	13.
Mean of Treated Share (%)	.29	.29	.29	.29	.29	.29	.29	.29
Std. Dev. Treated Share	.68	.68	.68	.68	.69	.69	.69	.69
Sample	2004–2018	2004–2018	2004–2018	2004–2018	2004–2018	2004–2018	2004–2018	2004–2018

Table 5: Effect of Credit Constraints on Change in Location Yields and Future Rent Growth. *Note:* Treated loan originations are measured as the two year lagged share of loan originations (by number) within 5 percent (on either side) of the new county-level conforming loan limit according to data from a major credit bureau. The share non-conforming is the two year lagged share by number of non-conforming loan originations according to data from a major credit bureau. The jumbo conforming spread is calculated using the difference in the annual average 30-year fixed-rate jumbo rate according to Bank Rate and the average annual 30-year fixed-rate mortgage rate from Freddie Mac. *Source:* Authors' calculations using data from Corelogic, the Decennial Census, BankRate, data from a major credit bureau and Freddie Mac.

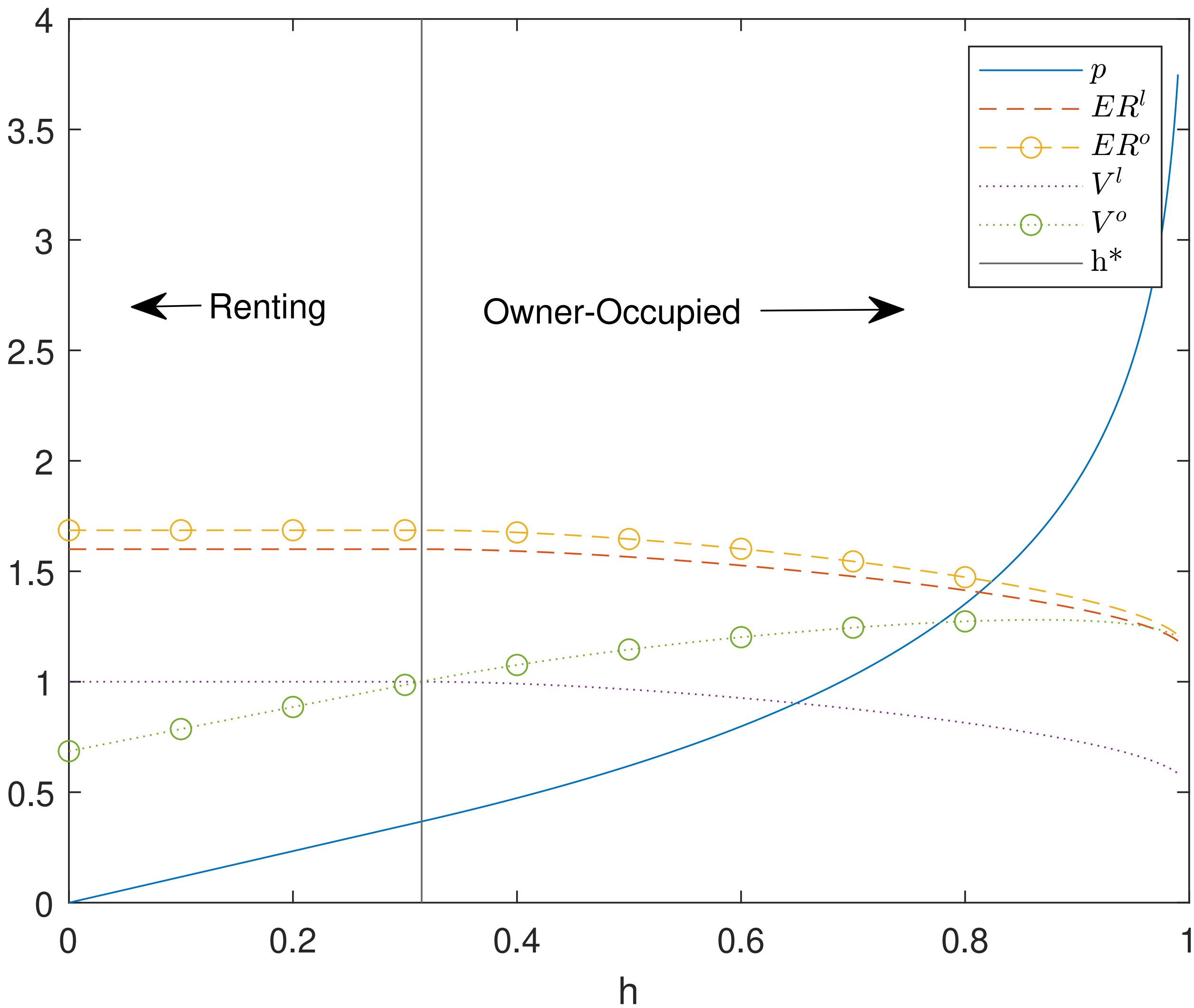

Figure 1: Two period example with $\delta^l = .3, \delta^o = .2, \beta^l = .4$.

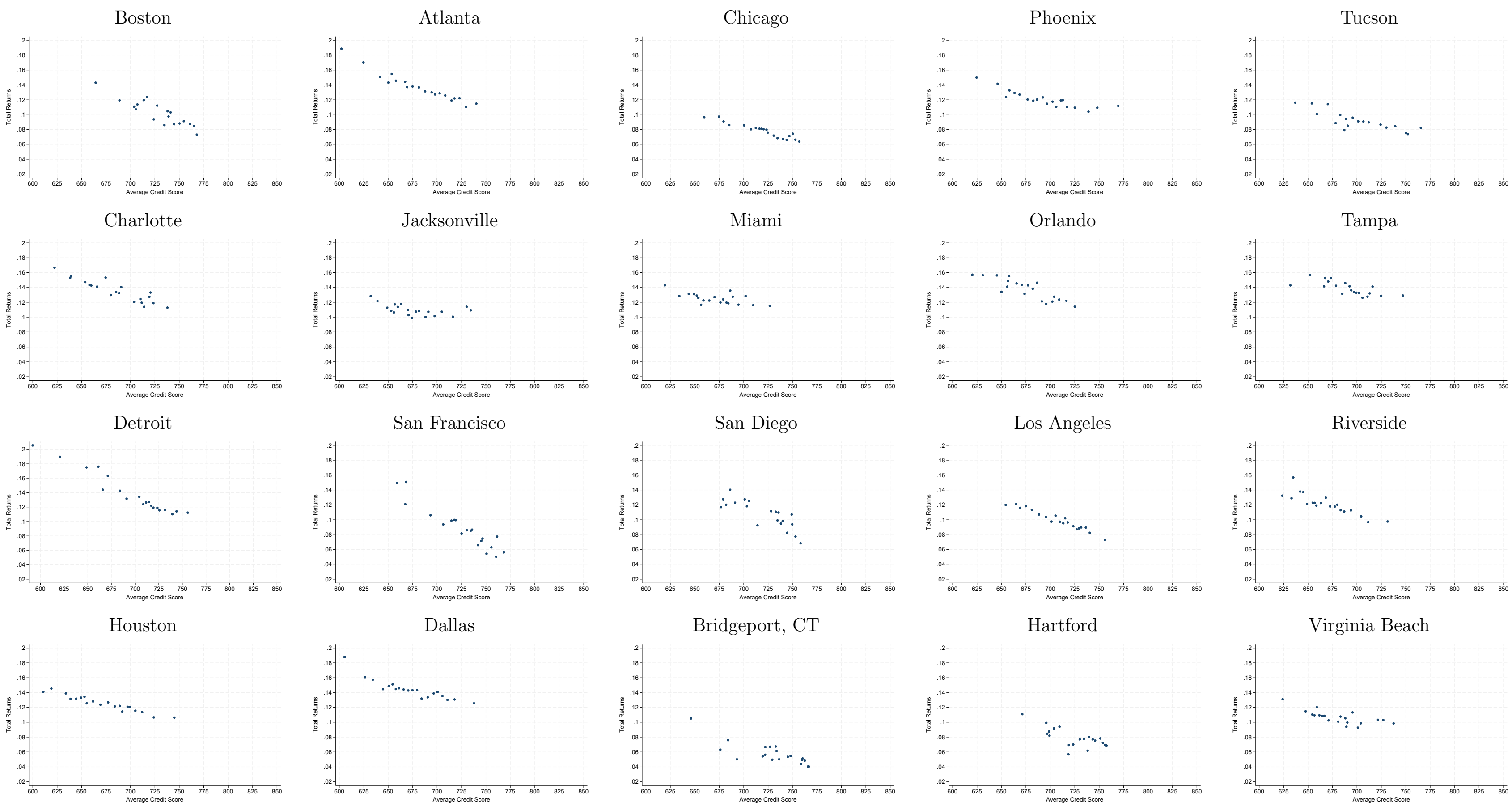

Figure 2: Total Returns to Location by Average Credit Score of Population. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011. The average credit score is measured in 2009. Data is limited to 2010–2021. *Source:* Corelogic MLS data, the American Community Survey, and a major credit bureau.

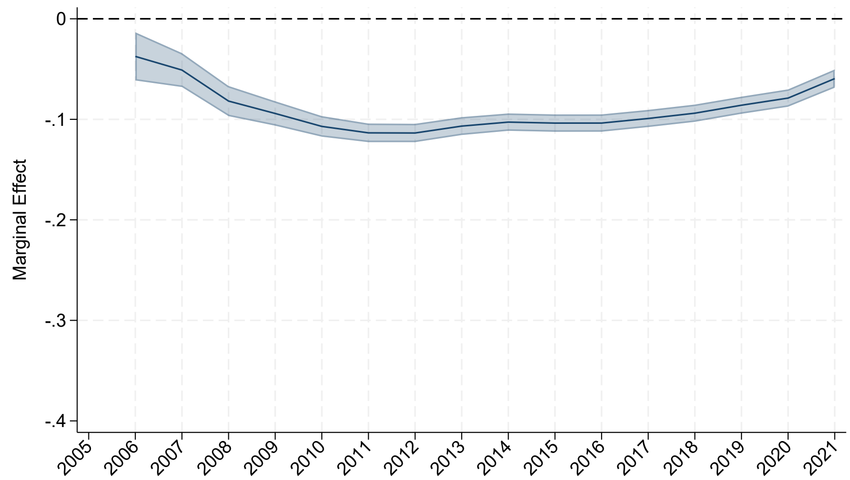

Figure 3: Relationship between Lagged Credit Scores and Yields Over time. *Note:* Marginal effect of two-year lagged credit score on log yields *Source:* Authors' calculations using MLS and data from a major credit bureau.

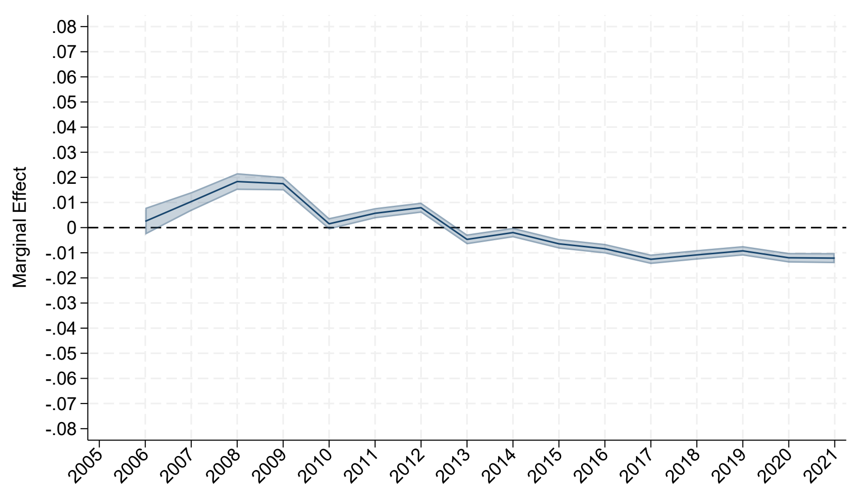

Figure 4: Relationship Between Location Capital Gains and Credit Scores Over Time. Marginal effect of two year lagged credit score on capital gains. *Source:* Authors' calculations using Corelogic MLS and data from a major credit bureau.

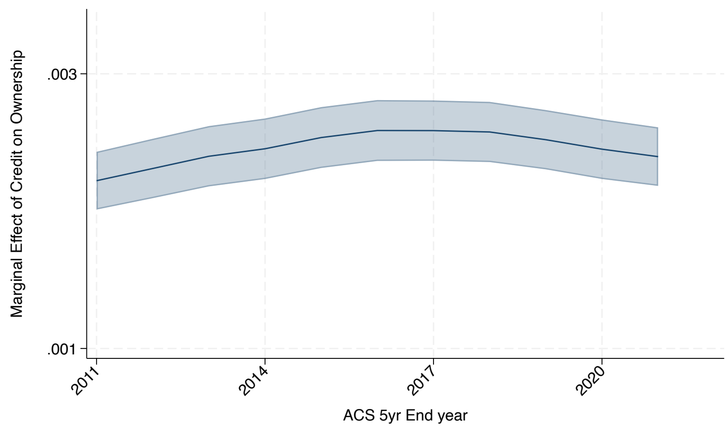

Figure 5: Effect of Credit on Ownership Rates. *Note:* Marginal effect of credit score on home ownership rates *Source:* Data from a major credit bureau and American Community Survey.

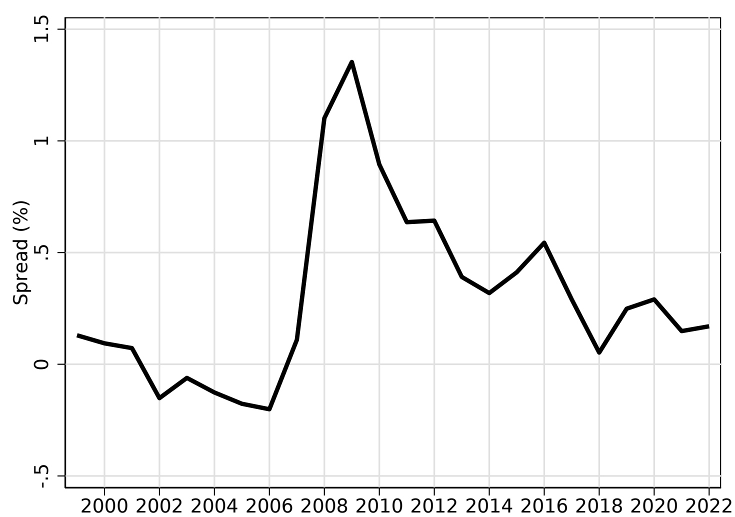

Figure 6: Jumbo-Conforming Spread. *Note:* Spread is calculated as the average annual 30-year fixed-rate jumbo rate according to BankRate.com minus the 30-year fixed-rate conforming rate according to Freddie Mac. *Source:* BankRate.com and Freddie Mac.

A Appendix

	Average			Std. Dev.		
	Yield	Cap Gains	Tot Return	Yield	Cap Gains	Tot Return
Atlanta, GA	7.35	3.66	11.02	1.18	7.19	7.12
Boston, MA-NH	5.08	3.44	8.52	0.67	3.25	3.01
Bridgeport, CT	5.23	0.21	5.44	0.71	4.42	4.07
Charlotte, NC-SC	6.59	3.67	10.26	1.28	4.93	4.90
Chicago, IL-IN-WI	6.36	0.57	6.93	1.25	4.75	4.49
Dallas, TX	6.77	4.92	11.69	1.06	4.20	3.92
Detroit, MI	7.86	3.60	11.45	1.36	7.60	7.68
Hartford, CT	6.05	0.57	6.62	0.86	2.54	2.87
Houston, TX	6.99	3.76	10.75	0.90	3.84	3.55
Jacksonville, FL	6.73	2.09	8.83	1.03	6.34	5.57
Los Angeles, CA	4.92	3.30	8.23	0.36	4.38	3.27
Miami, FL	6.65	3.12	9.78	0.89	7.65	6.61
Orlando, FL	7.89	2.76	10.65	1.59	12.32	11.56
Phoenix, AZ	5.10	3.14	8.24	0.64	8.37	7.94
Riverside, CA	5.86	2.92	8.78	0.72	7.14	5.26
San Diego, CA	5.35	3.15	8.50	0.84	3.46	3.05
San Francisco, CA	4.50	2.23	6.73	0.55	6.18	6.03
St. Louis, MO-IL	7.77	2.00	9.78	1.47	3.33	3.60
Tampa, FL	7.90	3.22	11.12	1.48	10.31	9.98
Tucson, AZ	5.33	1.58	6.91	0.59	5.42	4.79
Virginia Beach, VA-NC	6.90	1.34	8.23	0.82	3.68	4.50

Table A.1: Summary Statistics for Housing Returns by CBSA. *Note:* Yields, Capital Gains and Total Returns are in %-form. Values are weighted averages of the average and standard deviation (over time) of zip code-level yields, capital gains, and total returns between 2009 and 2021, where the weights are the number of households that are living in single-unit structures. *Source:* Corelogic MLS and the American Community Survey.

	Average			Std. Dev.		
	Yield	Cap Gains	Tot Return	Yield	Cap Gains	Tot Return
Atlanta, GA	10.23	–2.28	7.95	0.69	1.86	2.32
Boston, MA-NH	5.70	–0.69	5.01	1.16	0.46	1.28
Bridgeport, CT	4.40	–0.81	3.59	0.51	0.89	0.90
Charlotte, NC-SC	6.94	–2.34	4.61	0.77	1.32	1.43
Chicago, IL-IN-WI	5.18	–0.50	4.68	0.42	0.52	0.80
Dallas, TX	7.89	–1.89	6.01	1.37	0.80	1.57
Detroit, MI	8.62	–1.00	7.62	0.70	0.86	0.78
Hartford, CT	6.77	–0.74	6.03	1.83	0.65	1.86
Houston, TX	8.04	–1.64	6.41	0.33	1.45	1.30
Jacksonville, FL	7.67	–1.60	6.07	0.90	1.56	2.22
Los Angeles, CA	5.15	–0.94	4.20	0.71	0.78	1.32
Miami, FL	5.03	–1.55	3.49	0.62	2.12	2.41
Orlando, FL	8.18	–1.16	7.02	0.82	1.99	2.21
Phoenix, AZ	4.85	–1.93	2.92	0.78	0.83	1.01
Riverside, CA	5.08	–0.85	4.24	0.85	0.60	1.70
San Diego, CA	5.72	–1.22	4.50	1.26	0.89	1.35
San Francisco, CA	6.96	–0.61	6.35	1.46	0.97	1.78
St. Louis, MO-IL	9.35	–0.70	8.65	2.97	0.66	3.18
Tampa, FL	10.42	–1.13	9.29	1.62	1.08	2.00
Tucson, AZ	6.38	–1.69	4.69	1.60	1.03	1.86
Virginia Beach, VA-NC	6.53	–1.03	5.50	0.62	0.83	0.69

Table A.2: Summary Statistics for Structure Returns by CBSA. *Note:* Yields, Capital Gains and Total Returns are in %-form. Values are weighted averages of the average and standard deviation (over time) of zip code-level yields, capital gains, and total returns between 2009 and 2021, where the weights are the number of households that are living in single-unit structures. *Source:* Corelogic MLS and the American Community Survey.

	Mean	Std. Dev.	Min	Max
Capital Gains...				
Structure	–0.01	0.01	–0.02	–0.00
Land	0.04	0.02	0.01	0.07
Log Yields...				
Structure	–2.76	0.26	–3.20	–2.29
Land	–2.79	0.19	–3.19	–2.51
Jensen	0.02	0.02	–0.03	0.08

Table A.3: Variation in Structure and Location Returns Across CBSAs. *Note:* Values are summary statistics for CBSA-level average structure and location capital gains and yields. *Source:* Authors' calculations using MLS data.

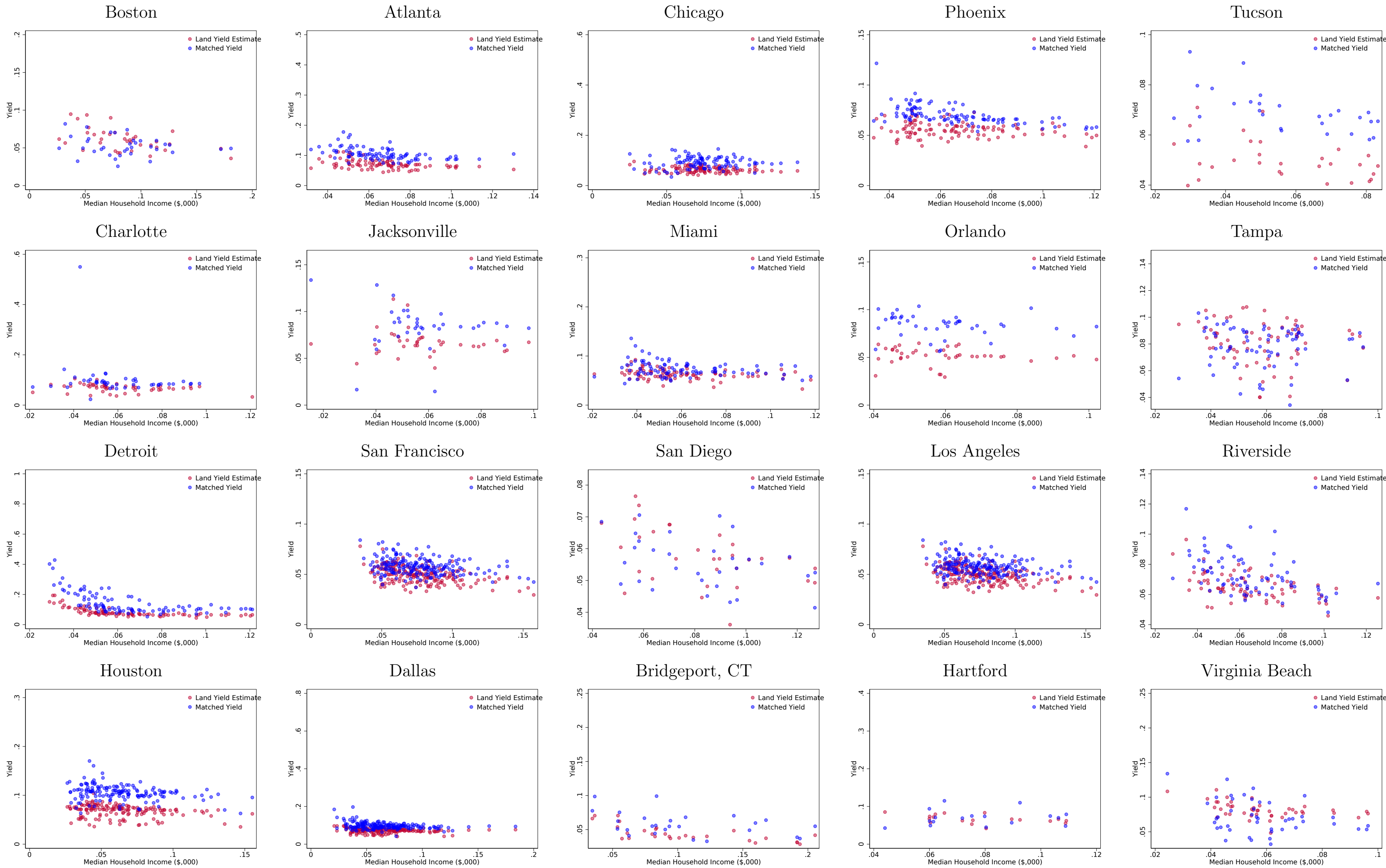

Figure A.1: Average matched yields and estimated land yields by zip code. *Note:* Average matched yields are averages of predicted values based on a single regression of property-level price-rent ratios on hedonics. The estimated land yields are estimated as described in Section 4. The x-axis is 2010 median household income from the decennial census. *Source:* Corelogic MLS data and the Decennial Census.

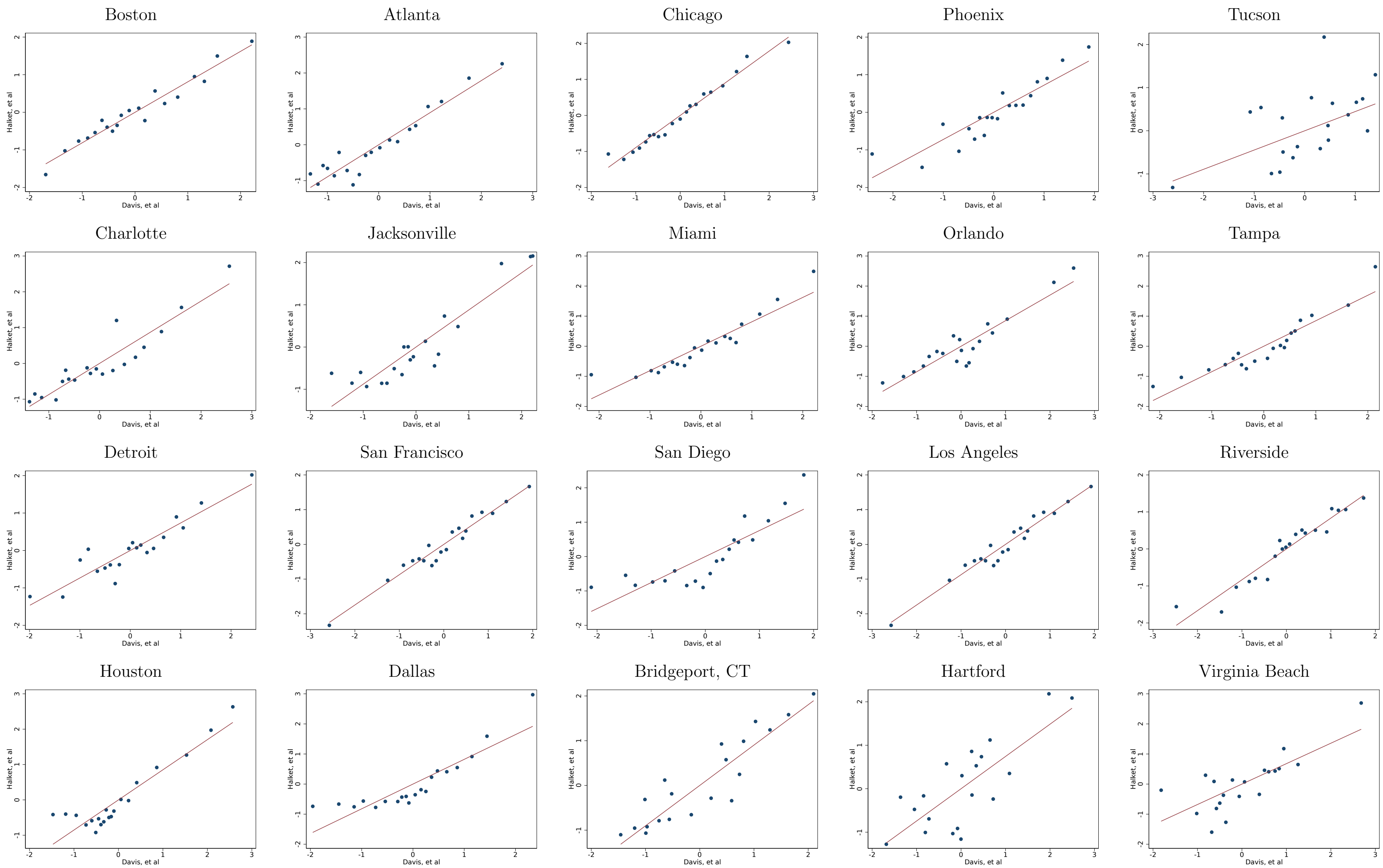

Figure A.2: Price of Land Per Sq. Ft., Comparison with Davis et al. (2021) *Note:* Our prices per square foot of land are estimated as described in Section 4. Both our estimates, and the estimates from Davis et al. (2021) are normalized to be mean zero and have a standard deviation of one. *Source:* Corelogic MLS data and Davis et al. (2021)

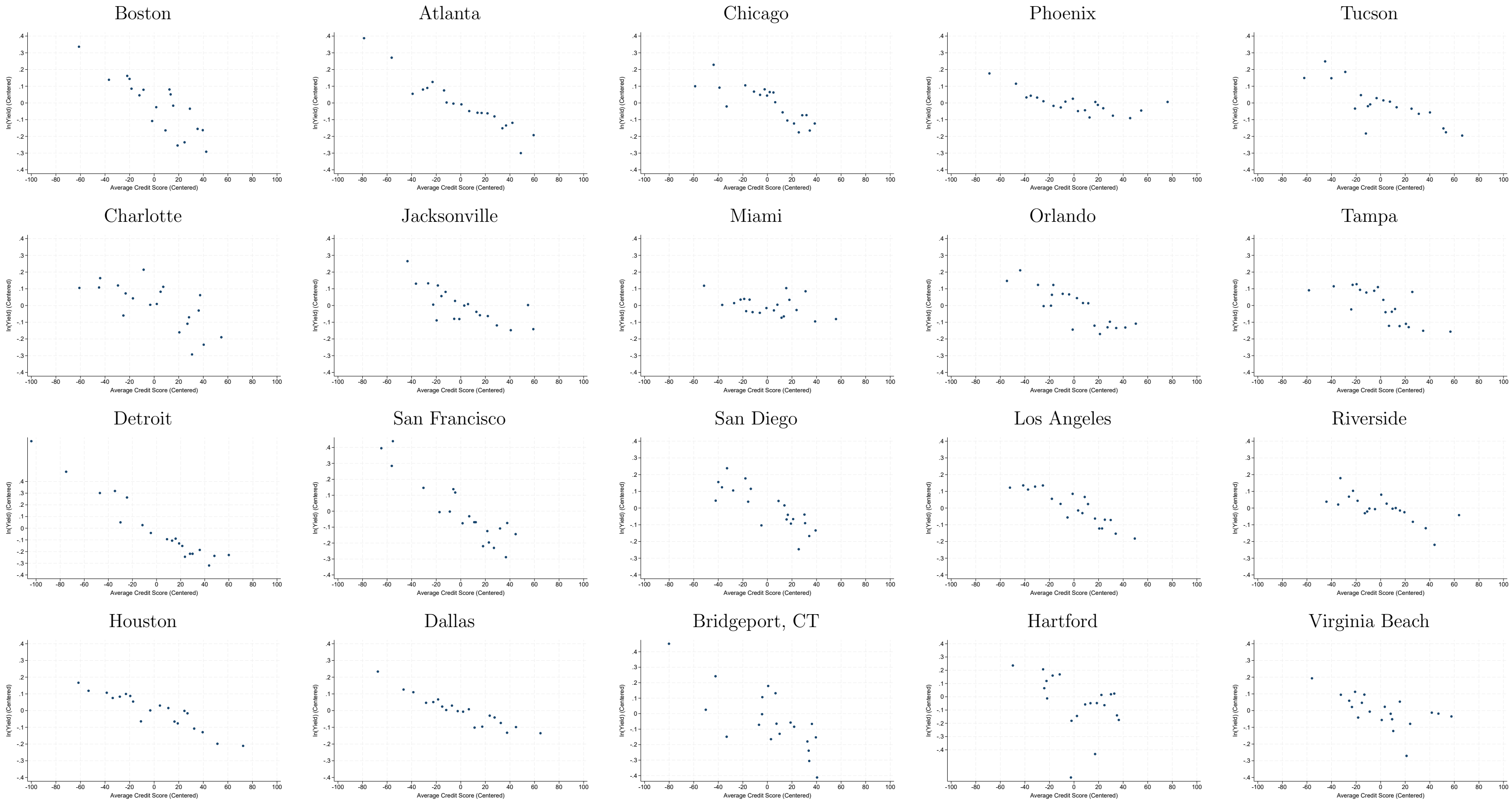

Figure A.3: Log Location Yields by Average Credit Score of Population. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011 and all values are centered relative to their weighted mean. The average credit score is measured in 2009. Data is limited to 2010–2021. *Source:* Authors' calculations using Corelogic MLS data, the American Community Survey, and data from a major credit bureau.

	Average					Std. Dev.		
	Yield Base House Val	Yield House	Yield Base Land Val	Yield Land	Jensen	Yield House	Yield Land	Jensen
Atlanta, GA	–0.41	–0.39	–1.91	–2.24	–0.01	0.10	0.19	0.02
Boston, MA-NH	0.47	0.19	–3.14	–3.22	0.02	0.17	0.17	0.00
Bridgeport, CT	–0.00	0.13	–3.32	–3.19	0.05	0.09	0.14	0.01
Charlotte, NC-SC	–0.50	–0.56	–2.14	–2.21	0.03	0.13	0.22	0.02
Chicago, IL-IN-WI	0.04	0.01	–3.02	–2.87	0.04	0.09	0.19	0.01
Dallas, TX	–0.15	–0.22	–2.35	–2.47	–0.02	0.08	0.11	0.00
Detroit, MI	–0.20	–0.30	–2.17	–2.32	0.02	0.08	0.12	0.01
Hartford, CT	0.21	0.08	–2.89	–3.00	0.08	0.22	0.21	0.01
Houston, TX	–0.40	–0.38	–2.17	–2.31	0.01	0.04	0.12	0.01
Jacksonville, FL	–0.46	–0.42	–2.19	–2.35	0.04	0.13	0.16	0.02
Los Angeles, CA	0.17	0.13	–3.15	–3.18	0.02	0.13	0.13	0.00
Miami, FL	–0.55	–0.54	–2.52	–2.25	0.03	0.08	0.19	0.02
Orlando, FL	–0.53	–0.54	–2.00	–2.07	0.02	0.11	0.23	0.03
Phoenix, AZ	–0.53	–0.61	–2.42	–2.37	–0.03	0.21	0.15	0.03
Riverside, CA	–0.03	–0.17	–2.88	–2.72	0.02	0.14	0.12	0.01
San Diego, CA	0.14	0.02	–2.95	–2.99	0.01	0.22	0.11	0.00
San Francisco, CA	0.04	0.05	–2.80	–3.22	0.02	0.13	0.16	0.01
St. Louis, MO-IL	–0.42	–0.26	–2.26	–2.41	0.06	0.17	0.14	0.02
Tampa, FL	–0.34	–0.37	–1.97	–2.25	0.03	0.19	0.23	0.02
Tucson, AZ	–0.36	–0.30	–2.51	–2.67	0.00	0.23	0.18	0.01
Virginia Beach, VA-NC	–0.28	–0.44	–2.32	–2.27	0.02	0.10	0.07	0.01

Table A.4: Summary Statistics for Log Yield Components by CBSA. *Note:* Values are weighted averages and standard deviations of mean zip code-level returns from 2009–2019. Weights are the number of housing units in the zip code in 2010. *Source:* Corelogic MLS.

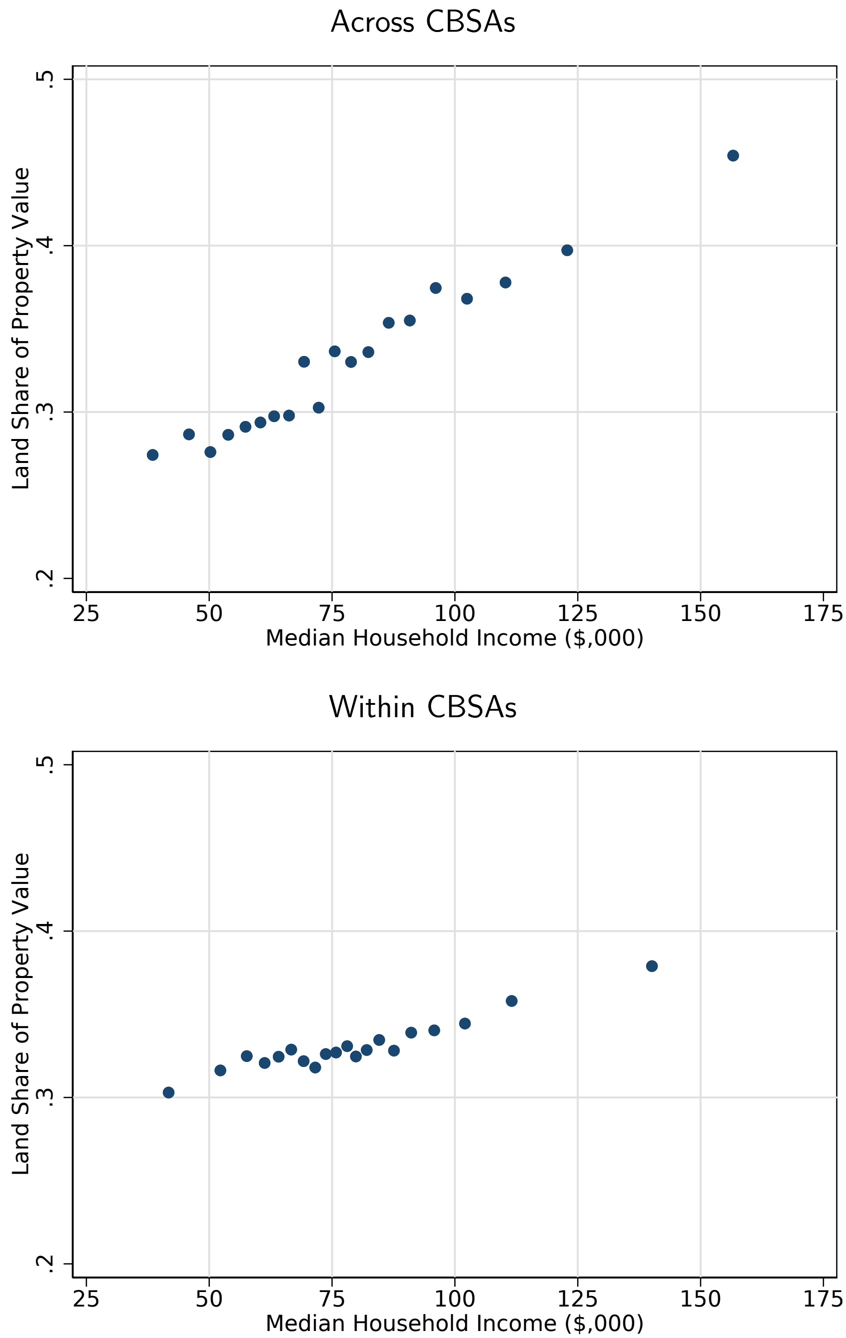

Figure A.4: THE LAND SHARE OF PROPERTY VALUES BY MEDIAN HOUSEHOLD INCOME. *Note:* Values are by census-tract. The land share of property values is measured as of 2012. Median household income is measured as of 2010. *Source:* Davis et al. (2021) and the Decennial Census.

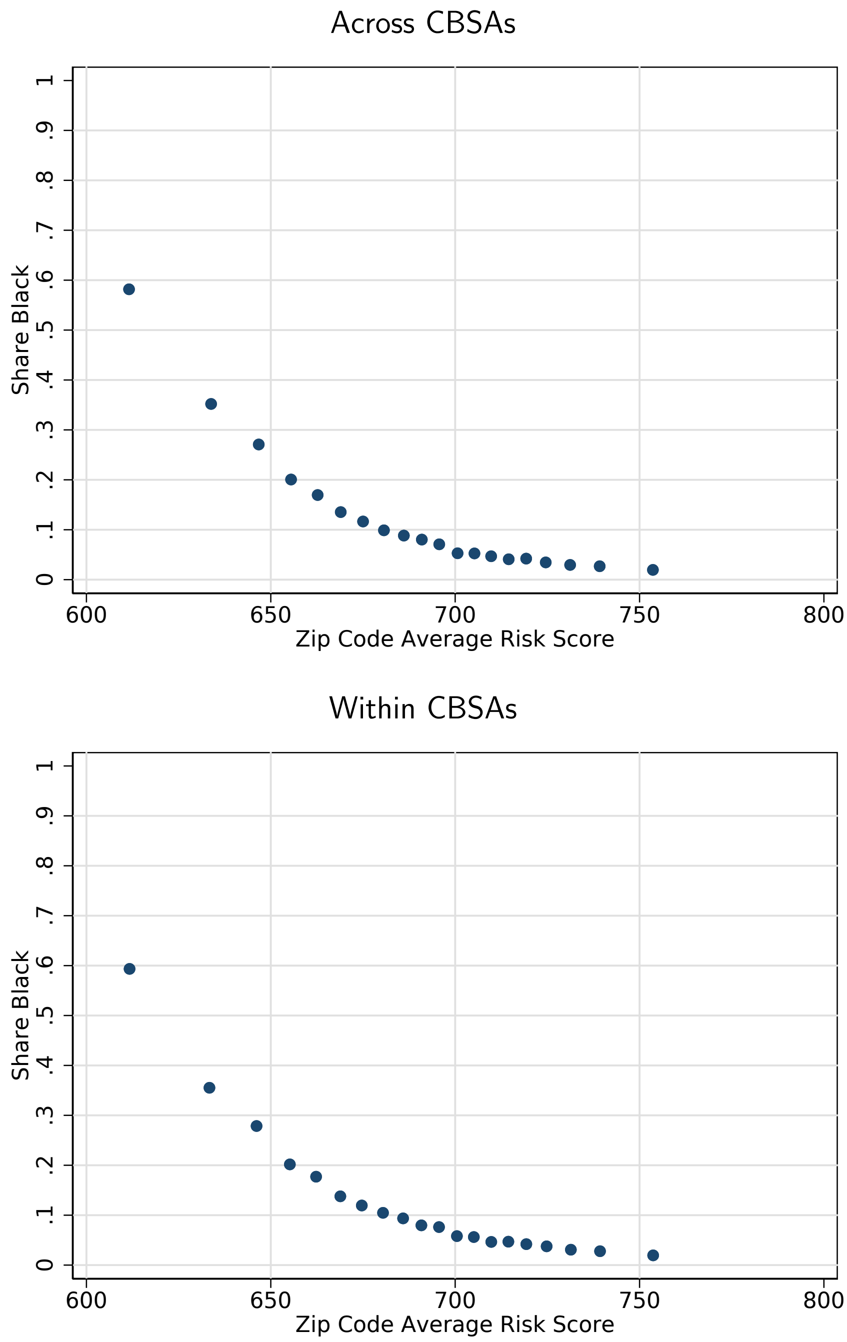

Figure A.5: Black Population Share Vs. Average Credit Score. *Note:* Values are for the calendar year 2009. The share black is from the 2010 Decennial Census, for which the survey is conducted in 2009. The average credit score is the average value by zip code in 2009. *Source:* Data from a major credit bureau and the Decennial Census.

	ln(Median Income)	Share Black (%)	Share Hispanic (%)	Average Credit Score	Share Vacant (%)
Atlanta, GA	-0.135***	0.0870***	-0.0103	-0.153***	0.118***
	(0.0225)	(0.00901)	(0.0408)	(0.0101)	(0.0327)
Boston, MA	-0.150***	0.394***	0.208***	-0.195***	-0.0463
	(0.0239)	(0.126)	(0.0563)	(0.0256)	(0.0830)
Bridgeport, CT	-0.136***	0.231***	0.201***	-0.178***	0.0902
	(0.0219)	(0.0509)	(0.0597)	(0.0294)	(0.139)
Charlotte, NC	-0.0513	0.0425*	0.0574	-0.0968***	-0.0148
	(0.0340)	(0.0232)	(0.0875)	(0.0284)	(0.0351)
Chicago, IL	-0.0393*	0.0388	0.0684***	-0.107***	-0.0386
	(0.0220)	(0.0260)	(0.0225)	(0.0192)	(0.0266)
Dallas, TX	-0.0350***	0.0724***	0.0104	-0.0991***	0.0129
	(0.00991)	(0.00935)	(0.0125)	(0.0100)	(0.0139)
Detroit, MI	-0.247***	0.146***	0.228	-0.241***	0.146***
	(0.0197)	(0.0111)	(0.207)	(0.0108)	(0.0113)
Hartford, CT	-0.0908*	0.146**	0.288**	-0.142***	-0.0276
	(0.0523)	(0.0612)	(0.119)	(0.0472)	(0.0790)
Houston, TX	-0.0372***	0.0542***	0.0335**	-0.104***	-0.00299
	(0.0135)	(0.0143)	(0.0152)	(0.0134)	(0.0204)
Jacksonville, FL	-0.0475	0.104***	0.0799	-0.111***	0.0141
	(0.0337)	(0.0233)	(0.198)	(0.0266)	(0.0343)
Los Angeles, CA	-0.0788***	0.190***	0.0730***	-0.122***	-0.00406
	(0.0129)	(0.0465)	(0.00858)	(0.0121)	(0.0415)
Miami, FL	-0.0539***	0.0633***	-0.0367***	-0.0387**	0.0138
	(0.0152)	(0.0144)	(0.00923)	(0.0182)	(0.0146)
Orlando, FL	-0.142***	0.112***	0.0695**	-0.128***	-0.00280
	(0.0300)	(0.0265)	(0.0291)	(0.0241)	(0.0276)
Phoenix, AZ	-0.0299**	0.216***	0.0609***	-0.0495***	-0.0132*
	(0.0128)	(0.0574)	(0.0123)	(0.00955)	(0.00762)
Riverside, CA	-0.0737***	0.0849*	0.0312**	-0.0762***	0.0254*
	(0.0146)	(0.0457)	(0.0150)	(0.0164)	(0.0154)
St. Louis, MO	-0.107*	0.0883	0.366	-0.179***	-0.0424
	(0.0651)	(0.0582)	(1.239)	(0.0569)	(0.0600)
San Diego, CA	-0.105***	0.366**	0.119***	-0.136***	-0.0467
	(0.0233)	(0.147)	(0.0201)	(0.0227)	(0.111)
San Francisco, CA	-0.209***	0.456***	0.239***	-0.208***	0.233***
	(0.0285)	(0.0650)	(0.0400)	(0.0199)	(0.0883)
Tampa, FL	-0.0595**	0.0548*	0.109***	-0.0942***	-0.0359**
	(0.0237)	(0.0325)	(0.0377)	(0.0250)	(0.0142)
Tucson, AZ	-0.0492**	0.353	0.0958***	-0.108***	0.00907
	(0.0229)	(0.216)	(0.0152)	(0.0183)	(0.0182)
Virginia Beach, VA	-0.0728***	0.0854***	-0.504***	-0.0820***	0.0639
	(0.0253)	(0.0123)	(0.177)	(0.0224)	(0.0491)

Table A.5: UNIVARIATE DETERMINANTS OF LOG YIELDS. *Note:* Coefficient estimates of univariate regressions of log location yields on factors. Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1.*Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

	ln(Median Income)	Share Black (%)	Share Hispanic (%)	Average Credit Score	Share Vacant (%)
Atlanta, GA	-0.00709***	0.00393***	0.00499**	-0.00674***	0.00601***
	(0.00113)	(0.000511)	(0.00202)	(0.000669)	(0.00167)
Boston, MA	-0.00698***	0.0279***	0.0121***	-0.00771***	0.00209
	(0.00143)	(0.00596)	(0.00290)	(0.00179)	(0.00443)
Bridgeport, CT	-0.00424***	0.00907***	0.00776***	-0.00648***	0.00818*
	(0.000849)	(0.00154)	(0.00194)	(0.000929)	(0.00457)
Charlotte, NC	-0.00832***	0.00747***	0.0180***	-0.00875***	0.00766***
	(0.00138)	(0.000666)	(0.00381)	(0.00108)	(0.00148)
Chicago, IL	-0.00312***	0.00336***	0.00382***	-0.00470***	0.00181**
	(0.000698)	(0.000833)	(0.000708)	(0.000587)	(0.000899)
Dallas, TX	-0.00384***	0.00502***	0.00310***	-0.00635***	0.00467***
	(0.000558)	(0.000545)	(0.000733)	(0.000597)	(0.000779)
Detroit, MI	-0.00159*	-0.000157	-0.000155	-0.000522	-0.00161***
	(0.000913)	(0.000539)	(0.00597)	(0.000780)	(0.000517)
Hartford, CT	-0.00318***	0.00254*	0.00852***	-0.00340***	0.00351**
	(0.000983)	(0.00142)	(0.00223)	(0.000979)	(0.00152)
Houston, TX	-0.00257***	0.00192***	0.00357***	-0.00405***	-0.0000462
	(0.000534)	(0.000598)	(0.000572)	(0.000565)	(0.000844)
Jacksonville, FL	0.00101	-0.00178*	-0.0124*	0.00146	0.00141
	(0.00126)	(0.00103)	(0.00696)	(0.00117)	(0.00123)
Los Angeles, CA	-0.00872***	0.0129***	0.00685***	-0.0109***	-0.00340
	(0.000631)	(0.00293)	(0.000402)	(0.000525)	(0.00262)
Miami, FL	-0.00504***	0.00406***	-0.00122*	-0.00371***	0.00266**
	(0.00108)	(0.00110)	(0.000736)	(0.00133)	(0.00106)
Orlando, FL	-0.00532***	0.00124	0.00549***	-0.00566***	-0.00214
	(0.00189)	(0.00174)	(0.00149)	(0.00149)	(0.00149)
Phoenix, AZ	-0.00589***	0.0232***	0.00849***	-0.00705***	0.000973
	(0.00104)	(0.00507)	(0.000921)	(0.000681)	(0.000696)
Riverside, CA	-0.00782***	0.0126***	0.00581***	-0.00989***	-0.000144
	(0.00134)	(0.00422)	(0.00131)	(0.00136)	(0.00151)
St. Louis, MO	0.00222	-0.00400***	0.00309	0.00188	-0.00365***
	(0.00142)	(0.00103)	(0.0269)	(0.00143)	(0.00107)
San Diego, CA	-0.00746***	0.0266*	0.00855***	-0.0107***	-0.00938
	(0.00241)	(0.0139)	(0.00223)	(0.00239)	(0.00999)
San Francisco, CA	-0.0171***	0.0385***	0.0228***	-0.0186***	0.0144*
	(0.00281)	(0.00615)	(0.00326)	(0.00173)	(0.00823)
Tampa, FL	-0.00290*	-0.000885	-0.00165	-0.000814	0.00325***
	(0.00151)	(0.00208)	(0.00249)	(0.00170)	(0.000853)
Tucson, AZ	-0.00261**	0.0141	0.00545***	-0.00601***	-0.0000587
	(0.00119)	(0.0115)	(0.000679)	(0.000869)	(0.000955)
Virginia Beach, VA	-0.000736	0.000819	-0.00425	-0.00183**	0.00197
	(0.000981)	(0.000629)	(0.00686)	(0.000874)	(0.00176)

Table A.6: UNIVARIATE DETERMINANTS OF CAPITAL GAINS. *Note:* Coefficient estimates of univariate regressions of location capital gains on factors. Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1.*Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

	ln(Median Income)	Share Black (%)	Share Hispanic (%)	Average Credit Score	Share Vacant (%)
Atlanta, GA	-0.0179***	0.0113***	0.00340	-0.0189***	0.0156***
	(0.00257)	(0.000971)	(0.00489)	(0.00112)	(0.00386)
Boston, MA	-0.0171***	0.0569***	0.0268***	-0.0208***	-0.000803
	(0.00203)	(0.0108)	(0.00498)	(0.00234)	(0.00842)
Bridgeport, CT	-0.0108***	0.0210***	0.0178***	-0.0154***	0.0127
	(0.00170)	(0.00343)	(0.00440)	(0.00193)	(0.0107)
Charlotte, NC	-0.0129***	0.0106***	0.0228***	-0.0157***	0.00890***
	(0.00261)	(0.00161)	(0.00739)	(0.00185)	(0.00297)
Chicago, IL	-0.00622***	0.00658***	0.00870***	-0.0121***	0.000264
	(0.00159)	(0.00189)	(0.00157)	(0.00117)	(0.00204)
Dallas, TX	-0.00677***	0.0108***	0.00411***	-0.0138***	0.00642***
	(0.00110)	(0.000997)	(0.00146)	(0.00106)	(0.00158)
Detroit, MI	-0.0252***	0.0149***	0.0134	-0.0241***	0.0138***
	(0.00185)	(0.00103)	(0.0206)	(0.000998)	(0.00120)
Hartford, CT	-0.00981***	0.0126***	0.0290***	-0.0133***	0.00307
	(0.00321)	(0.00391)	(0.00674)	(0.00256)	(0.00550)
Houston, TX	-0.00540***	0.00562***	0.00613***	-0.0110***	0.000205
	(0.00111)	(0.00121)	(0.00124)	(0.00104)	(0.00176)
Jacksonville, FL	-0.00249	0.00563***	-0.0123	-0.00632***	0.00244
	(0.00225)	(0.00167)	(0.0129)	(0.00186)	(0.00222)
Los Angeles, CA	-0.0131***	0.0225***	0.0108***	-0.0174***	-0.00356
	(0.00115)	(0.00486)	(0.000724)	(0.000962)	(0.00438)
Miami, FL	-0.00885***	0.00855***	-0.00378***	-0.00648***	0.00392**
	(0.00173)	(0.00171)	(0.00115)	(0.00215)	(0.00174)
Orlando, FL	-0.0166***	0.0108***	0.0110***	-0.0161***	-0.00201
	(0.00338)	(0.00320)	(0.00310)	(0.00255)	(0.00313)
Phoenix, AZ	-0.00733***	0.0373***	0.0125***	-0.0102***	-0.0000286
	(0.00164)	(0.00750)	(0.00140)	(0.00106)	(0.00106)
Riverside, CA	-0.0131***	0.0196***	0.00827***	-0.0156***	0.00123
	(0.00205)	(0.00667)	(0.00212)	(0.00215)	(0.00237)
St. Louis, MO	-0.00651	0.00375	0.00515	-0.0111***	-0.00470
	(0.00398)	(0.00365)	(0.0759)	(0.00346)	(0.00358)
San Diego, CA	-0.0141***	0.0510**	0.0163***	-0.0195***	-0.0119
	(0.00366)	(0.0220)	(0.00325)	(0.00350)	(0.0163)
San Francisco, CA	-0.0279***	0.0635***	0.0356***	-0.0296***	0.0269**
	(0.00400)	(0.00859)	(0.00487)	(0.00232)	(0.0124)
Tampa, FL	-0.00802***	0.00377	0.00775**	-0.00871***	0.000958
	(0.00223)	(0.00321)	(0.00377)	(0.00246)	(0.00144)
Tucson, AZ	-0.00542**	0.0326	0.0108***	-0.0119***	0.000581
	(0.00227)	(0.0220)	(0.00121)	(0.00159)	(0.00184)
Virginia Beach, VA	-0.00703***	0.00785***	-0.0480***	-0.00879***	0.00911**
	(0.00225)	(0.00108)	(0.0158)	(0.00188)	(0.00429)

Table A.7: Univariate determinants of Total Returns. *Note:* Coefficient estimates of univariate regressions of location total returns on factors. Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1.*Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

	Average		Std. Dev.	
	Cap Gain House	Cap Gain Land	Cap Gain House	Cap Gain Land
Atlanta, GA	–0.02	0.06	0.02	0.09
Boston, MA-NH	–0.01	0.04	0.00	0.03
Bridgeport, CT	–0.01	0.01	0.01	0.04
Charlotte, NC-SC	–0.02	0.06	0.01	0.05
Chicago, IL-IN-WI	–0.00	0.01	0.01	0.05
Dallas, TX	–0.02	0.07	0.01	0.04
Detroit, MI	–0.01	0.05	0.01	0.08
Hartford, CT	–0.01	0.01	0.01	0.03
Houston, TX	–0.02	0.05	0.01	0.04
Jacksonville, FL	–0.02	0.04	0.02	0.07
Los Angeles, CA	–0.01	0.04	0.01	0.05
Miami, FL	–0.02	0.05	0.02	0.09
Orlando, FL	–0.01	0.04	0.02	0.14
Phoenix, AZ	–0.02	0.05	0.01	0.08
Riverside, CA	–0.01	0.04	0.01	0.07
San Diego, CA	–0.01	0.04	0.01	0.04
San Francisco, CA	–0.01	0.03	0.01	0.06
St. Louis, MO-IL	–0.01	0.03	0.01	0.03
Tampa, FL	–0.01	0.04	0.01	0.11
Tucson, AZ	–0.02	0.03	0.01	0.06
Virginia Beach, VA-NC	–0.01	0.02	0.01	0.04

Table A.8: Summary Statistics for Cap Gains Components by CBSA. *Note:* Values are weighted averages and standard deviations of mean zip code-level returns from 2009–2019. Weights are the number of housing units in the zip code in 2010. *Source:* Corelogic MLS.

	ln(Median Income)	Share Black (%)	Share Hispanic (%)	Average Credit Score	Share Vacant (%)	Constant
Atlanta, GA	0.121***	-0.00126	0.0234	-0.230***	-0.0161	-2.726***
	(0.0280)	(0.0112)	(0.0242)	(0.0226)	(0.0215)	(0.0139)
Boston, MA	-0.00513	-0.0266	0.00563	-0.195***	-0.0828	-2.660***
	(0.0573)	(0.132)	(0.0636)	(0.0697)	(0.0531)	(0.0686)
Bridgeport, CT	-0.0396	-0.111	-0.361***	-0.406**	0.0214	-2.842***
	(0.0627)	(0.133)	(0.104)	(0.163)	(0.0904)	(0.0647)
Charlotte, NC	0.157*	-0.142***	-0.167**	-0.432***	-0.0259	-2.753***
	(0.0814)	(0.0459)	(0.0844)	(0.0893)	(0.0499)	(0.0723)
Chicago, IL	0.0409	-0.0378	-0.0475	-0.220***	-0.0893***	-2.670***
	(0.0297)	(0.0416)	(0.0415)	(0.0456)	(0.0310)	(0.0192)
Dallas, TX	0.0420***	-0.00966	-0.0536***	-0.185***	-0.0163	-2.749***
	(0.0153)	(0.0122)	(0.0141)	(0.0193)	(0.0137)	(0.0108)
Detroit, MI	-0.00890	0.0216	0.115	-0.178***	0.0233	-2.420***
	(0.0329)	(0.0200)	(0.0954)	(0.0445)	(0.0153)	(0.0888)
Hartford, CT	0.252*	-0.0994	0.147	-0.418**	-0.137*	-2.505***
	(0.150)	(0.101)	(0.163)	(0.183)	(0.0726)	(0.110)
Houston, TX	0.0745***	-0.0837***	-0.0845***	-0.298***	-0.0418**	-2.811***
	(0.0235)	(0.0187)	(0.0228)	(0.0272)	(0.0187)	(0.0167)
Jacksonville, FL	0.170***	0.0573	0.0316	-0.176***	0.0437	-2.758***
	(0.0497)	(0.0424)	(0.171)	(0.0612)	(0.0308)	(0.126)
Los Angeles, CA	0.0309	0.0423	0.00931	-0.130***	0.00543	-2.931***
	(0.0191)	(0.0460)	(0.0187)	(0.0333)	(0.0350)	(0.0370)
Miami, FL	-0.0779***	-0.0188	-0.0787***	-0.0462*	-0.0627***	-2.690***
	(0.0170)	(0.0212)	(0.0136)	(0.0257)	(0.0156)	(0.0194)
Orlando, FL	-0.109***	0.0653*	0.00725	-0.0369	-0.0505**	-2.599***
	(0.0358)	(0.0395)	(0.0454)	(0.0574)	(0.0219)	(0.0254)
Phoenix, AZ	-0.0410***	-0.0302	0.0269	-0.0278	-0.0331***	-2.923***
	(0.0155)	(0.0802)	(0.0246)	(0.0181)	(0.00833)	(0.0416)
Riverside, CA	-0.0504***	-0.0811*	-0.0875***	-0.162***	0.0147	-2.772***
	(0.0185)	(0.0484)	(0.0219)	(0.0316)	(0.0156)	(0.0269)
St. Louis, MO	-0.0696	0.284*	0.906	0.0249	-0.287**	-1.890
	(0.124)	(0.166)	(1.269)	(0.173)	(0.135)	(1.209)
San Diego, CA	-0.0435	-0.108	0.0685	-0.0466	-0.107	-2.922***
	(0.0308)	(0.155)	(0.0438)	(0.0653)	(0.0784)	(0.0984)
San Francisco, CA	-0.121***	0.127	-0.123**	-0.157***	0.00143	-2.840***
	(0.0385)	(0.0877)	(0.0584)	(0.0577)	(0.0492)	(0.0500)
Tampa, FL	-0.116***	-0.00111	0.0138	-0.0473	-0.0752***	-2.570***
	(0.0273)	(0.0362)	(0.0462)	(0.0398)	(0.0154)	(0.0270)
Tucson, AZ	0.0688	0.152	0.0534*	-0.101**	0.00650	-2.909***
	(0.0426)	(0.177)	(0.0308)	(0.0498)	(0.0208)	(0.104)
Virginia Beach, VA	-0.000854	0.0942***	-0.345*	0.00984	-0.0620	-2.971***
	(0.0427)	(0.0283)	(0.177)	(0.0451)	(0.0474)	(0.151)

Table A.9: MULTIVARIATE DETERMINANTS OF LOG YIELDS. *Note:* Coefficient estimates of multivariate regressions of log location yields on factors. Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1.*Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

	ln(Median Income)	Share Black (%)	Share Hispanic (%)	Average Credit Score	Share Vacant (%)	Constant
Atlanta, GA	0.00548***	0.00159**	0.00769***	-0.00770***	0.000117	0.0660***
	(0.00189)	(0.000760)	(0.00164)	(0.00153)	(0.00146)	(0.000940)
Boston, MA	-0.00491	0.0122	0.00424	0.00163	0.000966	0.0518***
	(0.00366)	(0.00844)	(0.00406)	(0.00446)	(0.00339)	(0.00438)
Bridgeport, CT	0.00294	-0.00323	-0.0131***	-0.0191***	0.00567**	0.0269***
	(0.00186)	(0.00395)	(0.00307)	(0.00483)	(0.00268)	(0.00192)
Charlotte, NC	-0.000762	0.00737***	0.00690**	0.00298	0.00222	0.0652***
	(0.00280)	(0.00158)	(0.00291)	(0.00307)	(0.00172)	(0.00249)
Chicago, IL	0.000883	0.00234*	0.00217	-0.00336**	-0.00164	0.0201***
	(0.00101)	(0.00141)	(0.00141)	(0.00155)	(0.00106)	(0.000653)
Dallas, TX	0.00188*	0.00226***	0.000260	-0.00505***	0.00260***	0.0693***
	(0.00100)	(0.000798)	(0.000922)	(0.00127)	(0.000900)	(0.000706)
Detroit, MI	-0.00398**	0.00244**	-0.0106**	-0.00323	-0.00713***	0.0470***
	(0.00171)	(0.00104)	(0.00495)	(0.00231)	(0.000793)	(0.00461)
Hartford, CT	-0.00223	0.00289	0.00583	0.00269	0.00246	0.0208***
	(0.00331)	(0.00222)	(0.00360)	(0.00403)	(0.00160)	(0.00242)
Houston, TX	0.00222*	-0.0000385	0.00247**	-0.00469***	-0.000717	0.0526***
	(0.00115)	(0.000910)	(0.00111)	(0.00132)	(0.000910)	(0.000816)
Jacksonville, FL	0.000515	-0.00354*	-0.0128	-0.00273	0.00123	0.0359***
	(0.00243)	(0.00207)	(0.00833)	(0.00299)	(0.00150)	(0.00615)
Los Angeles, CA	-0.00241***	-0.000751	0.00168**	-0.00682***	-0.00516***	0.0476***
	(0.000790)	(0.00190)	(0.000773)	(0.00137)	(0.00145)	(0.00153)
Miami, FL	-0.00405***	0.000267	-0.00186	-0.00256	0.000152	0.0598***
	(0.00148)	(0.00185)	(0.00119)	(0.00224)	(0.00136)	(0.00169)
Orlando, FL	-0.00498**	-0.00239	0.000118	-0.00478	-0.00362**	0.0551***
	(0.00230)	(0.00254)	(0.00293)	(0.00370)	(0.00141)	(0.00164)
Phoenix, AZ	-0.00372***	-0.0161***	0.00323*	-0.00597***	-0.00119**	0.0580***
	(0.00107)	(0.00550)	(0.00169)	(0.00124)	(0.000571)	(0.00286)
Riverside, CA	-0.00600***	-0.00515	-0.00513***	-0.0128***	-0.00122	0.0514***
	(0.00164)	(0.00428)	(0.00194)	(0.00280)	(0.00138)	(0.00238)
St. Louis, MO	-0.000443	-0.00621*	-0.0180	-0.00269	0.000299	0.0126
	(0.00271)	(0.00363)	(0.0277)	(0.00378)	(0.00294)	(0.0264)
San Diego, CA	-0.00193	-0.0193	0.00142	-0.0104	-0.0124	0.0385***
	(0.00338)	(0.0170)	(0.00481)	(0.00717)	(0.00861)	(0.0108)
San Francisco, CA	-0.00164	0.00129	-0.000909	-0.0193***	-0.00869*	0.0587***
	(0.00371)	(0.00846)	(0.00564)	(0.00557)	(0.00475)	(0.00483)
Tampa, FL	0.00121	-0.00423	-0.000294	-0.00348	0.00378***	0.0549***
	(0.00198)	(0.00263)	(0.00335)	(0.00289)	(0.00112)	(0.00196)
Tucson, AZ	0.000785	0.000207	0.00357***	-0.00387**	-0.00139*	0.0437***
	(0.00166)	(0.00689)	(0.00120)	(0.00194)	(0.000811)	(0.00405)
Virginia Beach, VA	0.00475**	-0.00218	-0.0133	-0.00786***	0.00301	0.0185**
	(0.00213)	(0.00141)	(0.00885)	(0.00225)	(0.00237)	(0.00753)

Table A.10: Multivariate determinants of Capital Gains. *Note:* Coefficient estimates of multivariate regressions of location capital gains on factors. Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

	ln(Median Income)	Share Black (%)	Share Hispanic (%)	Average Credit Score	Share Vacant (%)	Constant
Atlanta, GA	0.0136***	0.00307**	0.00912***	-0.0232***	-0.00137	0.133***
	(0.00300)	(0.00120)	(0.00260)	(0.00242)	(0.00231)	(0.00149)
Boston, MA	-0.00500	0.0141	0.00521	-0.0109*	-0.00437	0.126***
	(0.00488)	(0.0112)	(0.00541)	(0.00593)	(0.00452)	(0.00584)
Bridgeport, CT	0.00313	-0.00898	-0.0317***	-0.0426***	0.00657	0.0854***
	(0.00350)	(0.00742)	(0.00577)	(0.00908)	(0.00504)	(0.00361)
Charlotte, NC	0.0115*	-0.00320	-0.00360	-0.0269***	0.00416	0.134***
	(0.00634)	(0.00358)	(0.00658)	(0.00695)	(0.00388)	(0.00564)
Chicago, IL	0.00361**	0.000533	-0.000370	-0.0172***	-0.00682***	0.0890***
	(0.00176)	(0.00246)	(0.00245)	(0.00270)	(0.00184)	(0.00114)
Dallas, TX	0.00476***	0.00245*	-0.00332**	-0.0174***	0.00218	0.137***
	(0.00171)	(0.00136)	(0.00157)	(0.00216)	(0.00153)	(0.00120)
Detroit, MI	-0.00489*	0.00562***	0.000742	-0.0159***	-0.00198	0.140***
	(0.00296)	(0.00180)	(0.00858)	(0.00400)	(0.00137)	(0.00799)
Hartford, CT	0.0122	-0.00170	0.0167*	-0.0216**	-0.00488	0.103***
	(0.00791)	(0.00532)	(0.00861)	(0.00964)	(0.00383)	(0.00580)
Houston, TX	0.00658***	-0.00469***	-0.00216	-0.0225***	-0.00297*	0.116***
	(0.00194)	(0.00154)	(0.00188)	(0.00224)	(0.00154)	(0.00138)
Jacksonville, FL	0.0112***	0.000238	-0.0185	-0.0150***	0.00361*	0.0932***
	(0.00336)	(0.00286)	(0.0115)	(0.00414)	(0.00208)	(0.00851)
Los Angeles, CA	-0.00130	0.00124	0.00248*	-0.0129***	-0.00513*	0.102***
	(0.00151)	(0.00363)	(0.00148)	(0.00263)	(0.00276)	(0.00292)
Miami, FL	-0.00918***	-0.000961	-0.00715***	-0.00595*	-0.00366*	0.130***
	(0.00208)	(0.00261)	(0.00168)	(0.00316)	(0.00191)	(0.00238)
Orlando, FL	-0.0129***	0.00408	0.00159	-0.00689	-0.00714***	0.133***
	(0.00375)	(0.00413)	(0.00476)	(0.00601)	(0.00229)	(0.00266)
Phoenix, AZ	-0.00562***	-0.0187**	0.00579**	-0.00748***	-0.00327***	0.113***
	(0.00167)	(0.00864)	(0.00265)	(0.00195)	(0.000897)	(0.00448)
Riverside, CA	-0.00984***	-0.0100	-0.0111***	-0.0239***	-0.000660	0.116***
	(0.00238)	(0.00623)	(0.00282)	(0.00407)	(0.00201)	(0.00347)
St. Louis, MO	-0.00628	0.0129	0.0128	-0.00100	-0.0176**	0.109
	(0.00708)	(0.00948)	(0.0724)	(0.00988)	(0.00768)	(0.0689)
San Diego, CA	-0.00444	-0.0238	0.00617	-0.0129	-0.0186	0.0947***
	(0.00490)	(0.0246)	(0.00697)	(0.0104)	(0.0125)	(0.0157)
San Francisco, CA	-0.00738	0.0114	-0.00518	-0.0255***	-0.00843	0.119***
	(0.00488)	(0.0111)	(0.00742)	(0.00732)	(0.00625)	(0.00635)
Tampa, FL	-0.00767**	-0.00462	0.00228	-0.00723*	-0.00135	0.134***
	(0.00298)	(0.00396)	(0.00505)	(0.00435)	(0.00169)	(0.00295)
Tucson, AZ	0.00425	0.00756	0.00704***	-0.00862**	-0.000980	0.0988***
	(0.00306)	(0.0127)	(0.00221)	(0.00358)	(0.00150)	(0.00748)
Virginia Beach, VA	0.00564	0.00461*	-0.0453***	-0.00869**	0.000537	0.0621***
	(0.00356)	(0.00235)	(0.0148)	(0.00376)	(0.00395)	(0.0126)

Table A.11: Multivariate determinants of Total Returns. *Note:* Coefficient estimates of multivariate regressions of location total returns on factors. Each regressor is normalized to be CBSA-mean zero and have within-CBSA standard deviations equal to 1. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and data from a major credit bureau.

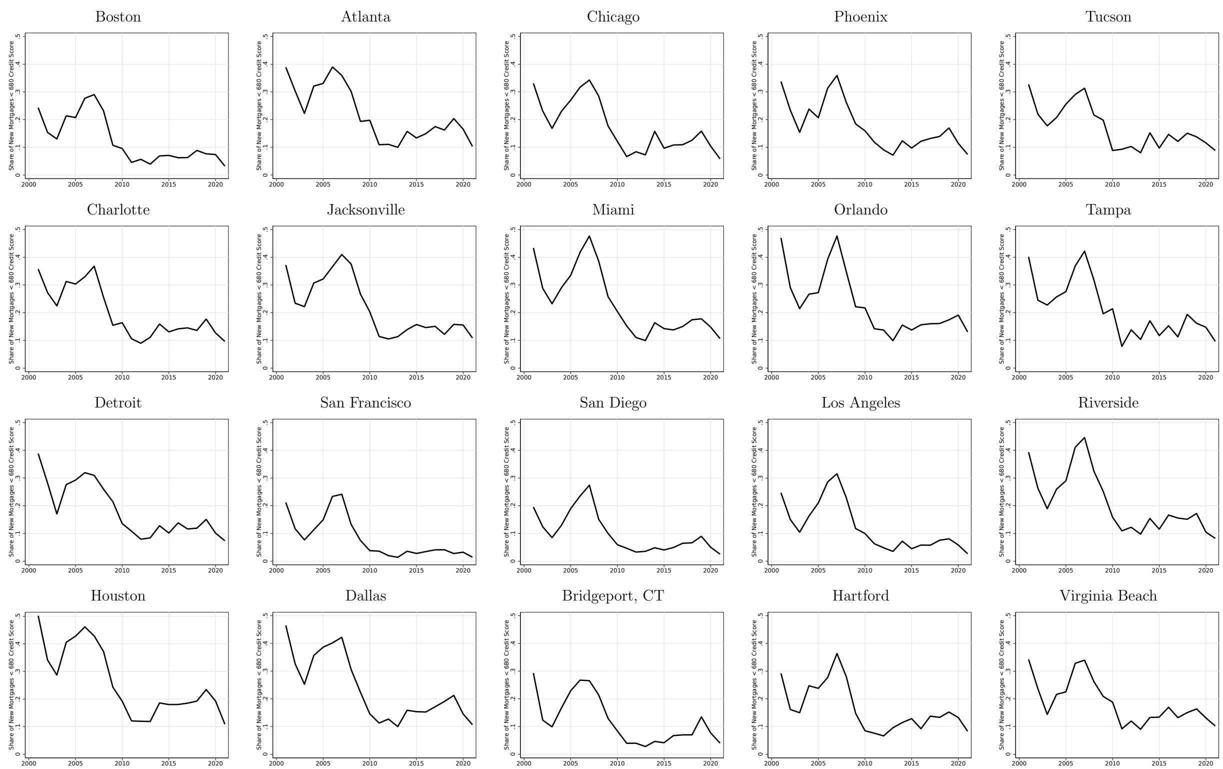

Figure A.6: Share of Mortgages Originated to People with a 680 Credit Score or Lower. Values are share of number, not value. *Source:* Authors' calculations using data from a major credit bureau.

	$R_m - R_f$	Credit Score	$R_m - R_f \times$ Credit Score	CBSA Net Return	CBSA Net Return $\times$ Credit Score	Constant
Atlanta, GA	0.00186***	-1.316***	-0.000354**	0.962***	-0.143***	9.386***
	(0.000158)	(0.300)	(0.000155)	(0.0354)	(0.0347)	(0.306)
Boston, MA	0.00103**	-2.316***	0.000131	1.122***	-0.113	9.447***
	(0.000518)	(0.845)	(0.000435)	(0.266)	(0.224)	(1.005)
Bridgeport, CT	0.00134***	-1.449***	-0.0000765	1.342***	-0.282	4.818***
	(0.000329)	(0.466)	(0.000242)	(0.261)	(0.192)	(0.633)
Charlotte, NC	0.000976***	-1.370***	-0.000191	0.976***	-0.131*	10.70***
	(0.000141)	(0.282)	(0.000147)	(0.0688)	(0.0720)	(0.270)
Chicago, IL	0.00146***	-1.019***	0.0000287	0.907***	-0.0334	5.415***
	(0.000149)	(0.283)	(0.000146)	(0.0541)	(0.0530)	(0.289)
Dallas, TX	0.000973***	-1.531***	0.000106	1.136***	0.000921	11.41***
	(0.0000915)	(0.172)	(0.0000914)	(0.0408)	(0.0407)	(0.172)
Detroit, MI	0.00145***	-2.323***	-0.0000187	0.662***	0.00197	10.93***
	(0.000190)	(0.340)	(0.000171)	(0.0573)	(0.0517)	(0.378)
Hartford, CT	0.00100***	-1.496***	0.0000673	1.274***	-0.315**	7.245***
	(0.000299)	(0.521)	(0.000268)	(0.157)	(0.141)	(0.582)
Houston, TX	0.000489***	-1.216***	0.00000741	1.109***	0.124**	10.34***
	(0.000100)	(0.173)	(0.0000893)	(0.0704)	(0.0627)	(0.194)
Jacksonville, FL	0.00244***	-0.613	0.0000291	1.150***	-0.188	6.181***
	(0.000331)	(0.767)	(0.000407)	(0.115)	(0.142)	(0.624)
Los Angeles, CA	0.00121***	-1.024***	-0.000421***	0.865***	-0.0686	7.975***
	(0.0000776)	(0.176)	(0.0000896)	(0.0365)	(0.0422)	(0.152)
Miami, FL	0.00192***	-0.579	0.0000253	0.323***	-0.0900	8.178***
	(0.000217)	(0.502)	(0.000248)	(0.0616)	(0.0706)	(0.438)
Orlando, FL	0.00235***	0.366	-0.00105*	0.496***	-0.233**	8.187***
	(0.000470)	(1.088)	(0.000543)	(0.0866)	(0.100)	(0.941)
Phoenix, AZ	0.00140***	-1.070***	-0.0000718	-1.486***	0.102	10.49***
	(0.000128)	(0.275)	(0.000131)	(0.0979)	(0.101)	(0.267)
Riverside, CA	0.00135***	-1.249***	-0.0000660	0.195**	-0.149*	8.485***
	(0.000185)	(0.388)	(0.000191)	(0.0760)	(0.0785)	(0.376)
St. Louis, MO	0.00137***	-0.0116	-0.000620***	0.842***	-0.0324	7.198***
	(0.000247)	(0.426)	(0.000220)	(0.205)	(0.182)	(0.480)
San Diego, CA	0.00118***	-0.874*	-0.000628***	0.572***	0.217	9.270***
	(0.000242)	(0.456)	(0.000231)	(0.220)	(0.210)	(0.478)
San Francisco, CA	0.00159***	-1.537***	-0.000897***	0.748***	0.150**	7.948***
	(0.000335)	(0.489)	(0.000250)	(0.0986)	(0.0737)	(0.654)
Tampa, FL	0.00220***	0.0668	-0.000237	0.522***	-0.323***	9.249***
	(0.000298)	(0.833)	(0.000416)	(0.0895)	(0.125)	(0.597)
Tucson, AZ	0.00156***	-0.699	-0.000293	0.978***	-0.155	6.411***
	(0.000243)	(0.481)	(0.000249)	(0.152)	(0.156)	(0.470)
Virginia Beach, VA	0.00108***	-0.982***	-0.000137	1.159***	-0.134	8.131***
	(0.000166)	(0.346)	(0.000181)	(0.125)	(0.135)	(0.319)

Table A.12: DETERMINANTS OF β AND α. *Note:* Credit Score is the average credit score in 2009 normalized to have mean zero and standard deviation equal to 1. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, data from a major credit bureau, and Fama-French factors.