The Cross-Section of Housing Returns*

Jonathan Halket†
Mays School of Business, Texas A&M University

Lara Loewenstein‡
Federal Reserve Bank of Cleveland

Paul S. Willen§
Federal Reserve Bank of Boston and NBER

March 9, 2023

Abstract

We document large systematic variations in the return to single-family residential property within U.S. metropolitan areas. Areas with low income, low credit scores or high shares of black residents have higher yields and therefore higher returns. Yield spreads between low credit areas and high credit areas widened considerably during periods when credit availability was low. This causes the areas with higher returns to also have higher risk, in sample. However we argue that the excess return that some areas earn is not purely compensation for bearing extra risk but is rather evidence for segmented housing markets where different local discount rates price local assets.

*The views in this paper are those of the authors' and do not reflect the opinions of the Federal Reserve Bank of Boston, the Federal Reserve Bank of Cleveland, or the Federal Reserve System.

†jonathan@halket.com
‡lara.loewenstein@clev.frb.org
§paul.willen@bos.frb.org

1 Introduction

Housing investments tend to be lumpy and geographically differentiated. The majority of residential properties are owner-occupied and different landlords and households sort into different houses and locations, meaning that the marginal property owner that "prices" housing in one part of the market may vary from the marginal owner in another part of the market. If these owners have heterogeneous intertemporal marginal rates of substitution (IMRS) or heterogeneous access to credit (perhaps due to differently binding borrowing constraints), then the expected return to holding a house can vary cross-sectionally within a market for reasons other than risk. In this paper, we show that household heterogeneity translates into predictable differences in the returns to housing, one of the most important household assets.

Specifically, we show that the returns to holding single family real estate vary considerably and predictably even within urban areas. Ex-ante differences in neighborhood demographics and economic characteristics, such as the median income of a neighborhood, average credit scores and racial composition, predict future returns mainly by predicting yields. Properties in areas with ex-ante lower median income or credit scores or a larger share of black residents have higher yields. The spreads in average returns across local areas within cities are generally large: a property in a zip code with double the median income or a 100 point higher average credit score as compared to a different property in another zip code in the same CBSA can expect anywhere between roughly 1 and 6 percent lower returns.

We do not directly observe discount rates in our data, but the link between, e.g. credit scores, and the opportunity cost of borrowing is straightforward, moderated potentially by the entry of landlords from outside the area. Using both repeated cross-sectional regressions as well as a series of quasi-natural experiments, we show that different local areas' average yields and returns respond differently to aggregate shocks to the cost of credit. Properties in areas with lower credit scores have larger changes in yields and returns in response to changes in access to credit. This seems theoretically reasonable: changes in access to credit likely matter more for areas where households were more likely to be ex ante credit constrained.

An implication of this, though, is that house prices move around more in our sample (during which there was a large expansion and then large contraction in credit) 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, the higher price volatility in low credit areas is caused by the higher volatility of area IMRSs due to those areas' extra sensitivity to change in access to credit.

Higher price volatility in low credit areas is at the core of studies of the 2000's housing boom and bust in Kuminoff and Pope (2012) and Landvoigt et al. (2015). A necessary condition for this is that housing markets are segmented such that different IMRSs are pricing different properties. Differential sensitivity to changes to the access of credit may be a risk-factor that causes differences in expected returns across properties. However, when markets are segmented, excess risk and excess return may be correlated but excess risk may not *cause all* of the extra return. We elaborate on this key point further towards the end of the section. But the upshot is that 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.

Estimating the returns to single-family housing is complicated: the gross dividend for owner-occupied housing (under some assumptions, the implied rent) is not observed in any data set and data on rents are sometimes scarce; data on the costs of owning a home (maintenance, depreciation, taxes, etc...) may be limited; and houses are traded infrequently, which means values are not always observed.

We use novel data with a large sample of detailed rental and sales transactions for a panel of U.S. CBSAs to contribute to the growing literature that estimates and analyzes returns to property at granular levels. Our date permit us 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 enable us to focus on the how returns to holding the same observable structure vary across locations within a CBSA, thereby ameliorating concerns discussed in Halket et al. (2020) about selection across rentals and owner-occupied houses on unobserved maintenance costs.[1]

Our 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[2] in a particular location within a CBSA relative to other locations within the CBSA predicts higher returns in that location relative to other locations in that CBSA (and not lower capital gains).

Eisfeldt and Demers (2015) using different data and methods finds that yields and returns are correlated with pricing tiers; zip codes with lower housing prices tend to have higher returns. In this paper, we further show 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 our demographic/economic factors are generally all statistically significantly correlated in the same direction. 20 out of 21 CBSAs in our sample have higher yields with lower average credit scores and 11 (17) of the CBSAs have significantly

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

[2] 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.

higher yields in low-income (a higher share of black residents) zip codes.

The relationship between ex-ante household credit scores and subsequent returns varies over our sample. During the housing boom prior to 2007, the yield spread from credit across locations within a CBSA narrowed considerably and in many of the CBSAs in our sample for which we have data during this period, the spread was statistically indistinguishable from zero. During the housing bust, house prices tended to fall further in areas where low credit score households lived causing yields spreads to widen considerably in most CBSAs (since rents did not fall as much prices). In this period low credit areas predicted higher yields and thus higher returns. This negative yield spread tended to narrow late in the recovery but in most CBSAs the spread remains negative throughout the rest of the sample (until 2022).

To further test the hypothesis that differences in expected returns are related to ex ante differences in access to credit, we follow Loutskina and Strahan (2015) and Greenwald and Guren (2021) by using changes in conforming loan limits as a set of natural experiments that changed the cost of credit for some locations more than others. We find that locations within CBSAs that previously had many mortgage originations near a new, higher limit experienced greater falls in yields than other locations within the CBSA. We also show that this same treatment does not forecast future net rent growth and thus likely points to changes in the local discount rate as the cause for the change in yields.

The relationship between yields and the economic factors is not counterbalanced by capital gains; zip codes with higher average yields do not have lower average capital gains. If anything, most but not all of the point estimates have the same sign as their counterparts for yields in our balanced sample; zip codes with higher average yields often have higher average capital gains in sample. Unlike the relationship between our factors and yields, the relationship between the factors and capital gains appears to be due to the entirety housing bust but not the entirety of the boom being in our balanced sample; in most years besides the period around the onset of the Great Recession, the ability of our factors to predict capital gains is zero.

Putting these two results together, the relationships between our factors and total returns across location is very strong; nearly all 21 CBSAs have statistically and economically significant relationships for all three factors (besides vacancy). A zip code with double the median income or a 100 point higher average credit score than another within the same CBSA can expect anywhere between roughly 1 and 6 percent lower returns. The results on race are similarly striking. In many CBSAs, locations with high shares of black residents pay higher net rents relative to prices so that an area with a 10 percent higher share of black residents has roughly between 0.3 and 4.8 percentage point higher returns.

When we run a multivariate "horse race" with all economic factors, credit scores are the most important by far. However when we double sort on both credit and race, we find that race has a large effect in locations where many owners are owner-occupiers or where households have high credit. This finding is novel but intuitive: race may impair relative access to credit more for households that otherwise should have had easy access to credit rather than for households that would likely anyway have difficulty obtaining a mortgage given their low credit score.

1.1 Implications of our findings

Our results have many potentially important implications both economically and econometrically.

In Sections 2 and 3, we 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. 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. The essence of the result is that housing is a real asset and the dollar flow value of owning a real asset can depend on who owns the asset. In this way, real estate has something in common with corporate and entrepreneurial finance.

We argue that spreads in IMRSs can cause spreads in returns across locations and therefore that the return spread we observe is evidence of segmented housing markets: different households are pricing housing in different parts of the market. So our results may serve as a measure of the incomplete smoothing of consumption across households (as in Lustig and Van Nieuwerburgh (2010).

Our results also point to the potentially large equilibrium pricing implications of discrimination in the housing market. Many studies (recent contributions include Bayer et al. (2017), Begley and Purnanandam (2021), Ambrose et al. (2020) and Bhutta and Hizmo (2020), Kermani and Wong (2022) for many varying results on this question) rightly search for an effect of discrimination *within* locations so as to better control for omitted variables. However, since households spatially sort along many dimensions, any discrimination would likely also manifest itself in returns. This can help shed light on the equilibrium mechanisms underlying many important questions related to housing.

Economically, higher returns in low income areas may imply that owner-occupancy in these areas is a potent way to build wealth and segmented markets may also explain the high rent-to-price ratios in low-income neighborhoods. Secondly, if the high cost of borrowing for certain households suppresses prices for the types of houses that these households live in, this could lower the incentive for developers to develop these houses. Finally, if properties in different locations respond differently to changes in the cost 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 as well.

Econometrically, our results imply that different houses can have different long-run expected returns. Time-series studies that follow Campbell-Shiller decompositions (e.g. Campbell et al. (2009) should be wary of estimating models where the return to all housing is restricted in the long-run to be the same.

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.

1.2 Related Literature

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 them (see Goetzmann et al. (2021) for a recent summary). Some of this literature looks for and 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. Our findings contribute to the growing number of studies which show yields contain a lot of 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. Plazzi et al. (2010) studies risk and return among CRE properties using Campbell-Shiller decompositions.

Structural dynamic models of housing and home ownership (e.g. 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. Our results provide novel evidence for this equilibrium.

The effect of housing wealth on IMRSs has been studied at least as far back as Campbell and Cocco (2003), Lustig and Van Nieuwerburgh (2005), Campbell and Cocco (2007) and Lustig and Van Nieuwerburgh (2010). 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.

Ours is not the first study to moot housing market-segmentation. Piazzesi and Schneider (2016) discusses it at length. 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.

In Section 2 we 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 3 builds on the preceding section and constructs an econometric model we can bring to the data. Section 4 discusses our novel data and Section 5 our empirical methods. Section 6 presents the results and Section 7 concludes with some suggestions for future research.

2 A Simple Model

To illustrate how housing markets can segment in equilibrium, we adapt a simple setting in Piazzesi and Schneider (2016).

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 in consumption goods when born.

Households obtain housing services by living in exactly one house. Houses come in different qualities $h \in [0, 1-\rho]$. We assume the economy is endowed with a mass θ of $h = 0$ quality houses and a mass $1-\theta$ of higher quality houses distributed uniformly over $[0, 1-\rho]$. The technology converting quality into services s is linear. Owner-occupied houses require maintenance of δ^o per unit quality.

Households receive utility over consumption and housing services in their first year of life and from their terminal wealth w_T in year two. We assume that their utilities are linear in each of consumption, housing services and terminal wealth and that there are no assets available to trade.[3]

Households are heterogeneous only in their discount rates β. We assume discount rates are distributed uniformly over $[0, 1]$ each generation and that $w_1 > [1 - \rho]$.

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

$$\begin{aligned} & \max_{c_t, h_t} c_t + h_t + \beta W_T \\ s.t. \qquad & c_t + p_t(h_t) + \delta^o h_t = w_1 \\ & w_T = p_{t+1}(h_t) \end{aligned}$$

The first-order condition for housing for this household implies:

$$p_t'(h_t) = 1 - \delta^o + \beta \tag{1}$$

In an equilibrium where all houses are owner-occupied, households with $\beta < \rho$ live in $h = 0$ houses while households with $\beta \geq \rho$ live in $h = \beta - \rho$ quality houses. In this equilibrium, the price of housing is

$$p_t(h) = \int_0^h p_t'(\tilde{h}) d\tilde{h} = (1 - \delta^o + \rho)h + \frac{h^2}{2} \tag{2}$$

and the return to holding a house is

$$ER(h)] = 1 - \delta^o + h + \rho. \tag{3}$$

Even with linear utilities, house prices are non-linear and expected returns vary with quality.

If we assume that houses may also be rented out instead of just owner-occupied, it is clear from households linear preferences that the rental cost $R(h) = h$. In keeping with the large literature on moral hazard problems in renter markets (Halket et al. (2020) and citations therein), we allow landlords to have 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 $\beta^l \in (\rho, 1]$.[4]

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

[4]We 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.

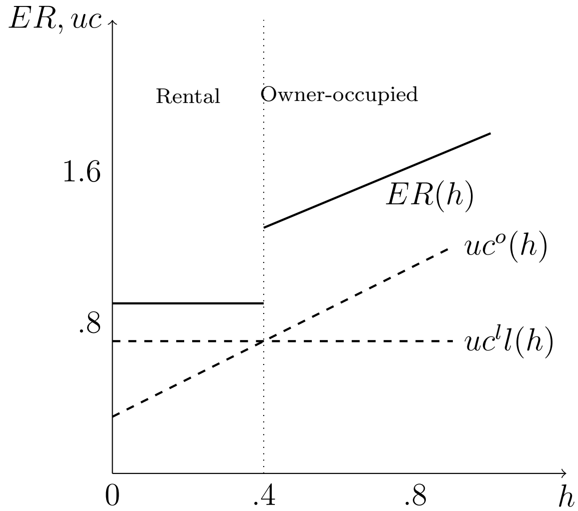

Figure 1: Two period example with $\delta^l = .3, \delta^o = .2, \beta^l = .4, \rho = .1$.

In equilibrium, the willingness to pay of a landlord is

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

so that houses with quality

$$h < h_r^* = 2[\beta^l - \rho - (\delta^l - \delta^o)] \tag{5}$$

are rented and the rest are owned. Households with discount rates $\beta < \beta^l - (\delta^l - \delta^o)$ will rent (as long as $h_r^* > 0$) and are indifferent over $h \in [0, h_r^*]$ and households with higher discount rates will owner-occupy.

Equivalently we can express things in terms of user-costs. Capital gains are 0 here, so user costs are just discount rates plus the cost of maintenance. So for any house owned by an agent with discount rate β, the user-cost of that agent is $uc = \beta + \delta$. Equation (5) is equivalent to a condition on user-costs: houses are rented whenever the user-cost of the landlord for that house is less than the user-cost of the equilibrium owner-occupier for that house. An example of this is shown in Figure 1.

3 Returns in a two sector model

In this section we generalize the intuition from Section 2 in order to obtain a model we can estimate with our 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. Households are a vector of characteristics (state variables) $s \in S$. For convenience we assume that we can partition the state into two parts $S = S^h \cup S^l$. As will be made more 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 the characteristics relevant to its management.[5] 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 measures of managerial ability, location of headquarters, etc....[6] S^l includes an indicator as to whether the owner is 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 the flow value (in non-durable consumption units) from an resident-owner pair in state $s = s_h, s_l \in S$ of a property of type z^{ε} given the price function P as $U(z^{\varepsilon}, s^h)$. Assume maintenance costs (including property taxes) are $c(z^{\varepsilon}, s^l)P(z^{\varepsilon})$ and the opportunity cost of capital is $r(s^l)$.[7] To simplify notation below, we 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.

We 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 itself) s^h is given by

$$\pi(z^{\varepsilon}; s, P, R) = U(z^{\varepsilon}, s^h) - \frac{c(z^{\varepsilon}, s^l)P(z^{\varepsilon})}{1 + r\left(s^l\right)} + \frac{\left(1 + g(z^{\varepsilon}, s^l)\right) P(z^{\varepsilon})}{1 + r(s^l)}. \tag{6}$$

Thus, the willingness to pay equals the current net utility flow plus the discounted expected future value of the property.[8][9]

We assume that in equilibrium there is a correspondence mapping properties to residents and owners $T : Z^{\varepsilon} \Rightarrow S$. T can itself be partition into two correspondences $T^h : Z^{\varepsilon} \Rightarrow S^h$ and $T^l : Z^{\varepsilon} \Rightarrow S^l$ such that, 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 $\pi(z^{\varepsilon}; T(z^{\varepsilon}), P, R) = P(z^{\varepsilon})$.[10] Equation (6) can be rewritten as

[5] Characteristics may appear in both sets.

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

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

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

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

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

$$U(z^\varepsilon, T^h(z^\varepsilon)) = \left[\frac{c(z^\varepsilon, T^l(z^\varepsilon)) + r(T^l(z^\varepsilon)) - g(z^\varepsilon, T^l(z^\varepsilon))}{1 + r(T^l(z^\varepsilon))}\right] P(z^\varepsilon) \tag{7}$$

$$\approx \left[c(z^\varepsilon, T^l(z^\varepsilon)) + r(T^l(z^\varepsilon)) - g(z^\varepsilon, T^l(z^\varepsilon))\right] P(z^\varepsilon) \tag{8}$$

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 8 reveals a user-cost $uc : Z^\varepsilon \times T^l(Z^\varepsilon) \to \mathbb{R}$ for properties that is both property and owner dependent. Holding fixed U, the agents with the lowest user-costs will have highest willingness-to-pay and thus will own the property in equilibrium. We 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:

$$\frac{R(z^\varepsilon)}{P(z^\varepsilon)} = uc(z^\varepsilon, T^l(z^\varepsilon)) = c(z^\varepsilon, T^l(z^\varepsilon)) + r(T^l(z^\varepsilon)) - g(z^\varepsilon, T^l(z^\varepsilon)) \tag{9}$$

The expected return for a property is the net yield plus expected capital gains. Using equation 9, the expected returns for a property z^ε owned by $T(z^\varepsilon)$ is:

$$E[R(z^\varepsilon, T^l(z^\varepsilon))] = \frac{R(z^\varepsilon)}{P(z^\varepsilon)} - c(z^\varepsilon, T^l(z^\varepsilon)) + g(z^\varepsilon, T^l(z^\varepsilon)) = r(T^l(z^\varepsilon)) \tag{10}$$

Remark 1 *Equations 9-10 display a key feature of the equilibrium:*

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

2. *A property may be owner-occupied even if the owner-occupier (with $s^l_1 \in S^l$) is borrowing constrained and has a higher effective discount rate than a landlord (with $s^l_2 \in S^l$), $r(s^l_1) > r(s^l_2)$, if, for instance, landlords have sufficiently higher maintenance costs for the property than an owner-occupier, $c(z^\varepsilon, s^l_2) + r(s^l_2) > c(z^\varepsilon, s^l_1) + r(s^l_1)$.*

3. *Holding fixed $T^l(z^\varepsilon)$, the variance of returns for different houses equals the variance of their respective discount rates.*

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.

Furthermore, if shocks disproportionately affect certain owners' discount rates over 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 (a) 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 (b) there is factor that causes borrowing constraints to tighten exogenously (perhaps from a changes in government policy) in such a way that 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 returns and prices in A will remain unchanged but expected returns in B will go up and prices in B will go down (and likewise only B's returns will go down when the policy loosens.) Landvoigt et al. (2015) discuss an instance of this in San Diego. Ex-post, realized price and return volatility for owner B will be higher than A. Ex-ante, B will 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 (which is what we see in our data).

Continuing from above, if we assume that the user-cost relationship in equation 9 is well-approximated by a semi-log specification where z is a vector of observable characteristics in $Z \subset Z^{\varepsilon}$ then:

$$\log uc(z^{\varepsilon}, T(z^{\varepsilon})) = (\alpha - \beta)z + \varepsilon_r - \varepsilon_p \tag{11}$$

with $\left(\varepsilon_r, \varepsilon_p\right) \sim N(0, \Sigma)$ and $\Sigma = \begin{pmatrix} \sigma_r^2 & \rho_{rp} \\ \rho_{rp} & \sigma_p^2 \end{pmatrix}$. Then we can express rents and prices as

$$\log R(z^{\varepsilon}) = \alpha z + \varepsilon_r \tag{12}$$

$$\log P(z^{\varepsilon}) = \beta z + \varepsilon_p \tag{13}$$

and predicted user costs or gross property yields are

$$E\left[\frac{R(z^{\varepsilon})}{P(z^{\varepsilon})}|z\right] = \exp\left((\alpha - \beta)z + \frac{\sigma_r^2}{2} + \frac{\sigma_p^2}{2} + \rho_{rp}\right) \tag{14}$$

Note, using (12), (13) and (14), if we assume we can partition z into elements which are "structure," z_s, and elements which are location, z_l, then 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 the Jensen inequality terms.

We will build on this specification further in Section 5 but in the next section we will first introduce the data.

4 Data

Our data on rents and prices comes 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. Our main data set is the full set of closed sale and rental listings on single family homes and condos.

In addition, we identify a set of properties for which we there is both a rental and sale transaction within one year of each other. This provides a direct measure of property-level yields. We use this to adjust our measure of yields for our full sample and as a robustness check. We find matching sale prices for about 21 percent of rental listings from MLS. There are no significant differences in rental rates between properties for which we did and did not find a match.

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

We perform some data cleaning. We cap the number of bedrooms and bathrooms at 5 and remove any properties for which the comments indicate either are an accessory dwelling, or contain an accessory dwelling. The distributions of building and land square footage contain some outliers. We 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.

Our 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, we net out the actual dollar amount of property taxes associated with a given property from the annual rental income. For properties for which we do not have property tax information, we 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, we expand our sample of sale prices by using estimated sale prices for properties in years in which they did not transact. We do this in two ways. First, we 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.

We also use a variety of other data sources. We obtain information about the credit scores of people in a given zip code from the New York Fed Equifax Consumer Credit Panel, which is an anonymous, representative panel of households with Equifax credit reports. Information on LTVs and credit scores on newly originated mortgages comes from Black Knight Analytics. Demographic information, such as race, age, and income comes from Decennial Census and American Community Survey. We also use a measure of the housing vacancy rate based on USPS administrative data and made available by the US Department of Housing and Urban Development (HUD).[11]

5 Estimating Returns

We estimate yields, capital gains, and total returns in each zip code for each CBSA in our sample using a hedonic approach and our full sample of sale prices and rents. Our methodology builds on that of Kuminoff and Pope (2012), who use the market values of properties to estimate the values of the underlying land and structure.

We mainly focus on estimating variations in various objects (like prices and rents) for properties with constant structural characteristics across different location. For convenience, we will often then refer to, e.g., the "location value of a house." We do not estimate explicit land values or rents. This approach has 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. Furthermore, there is no available rent data for land parcels that we are aware of, which would preclude us from estimating yields and total returns. Another approach is to estimate structure values from the 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.[12].

We run the following regression year-by-year using our sample of sale and rental trans-

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

[12] 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

actions:

$$
\begin{aligned}
\ln(\text{price}_{ijkt}) = {} & \beta_{0,k,t} + \beta_{1,k,t}\text{Sq. Ft.}_i + \beta_{2,k,t}\text{Sq. Ft.}_i^2 + \beta_{3,k,t}\text{Bedrooms}_i \\
& + \beta_{4,k,t}\text{Bathrooms}_i + \beta_{5,k,t}\text{Building Age}_{i,k,t} + \beta_{6,k,t}\text{Building Age}_{i,t}^2 \\
& + \gamma_{jt}\text{Land SqFt}_i + \delta_{jt} + \epsilon_{ijkt},
\end{aligned}
\tag{15}
$$

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 a representative sample of rental properties (see Table A.1 in the appendix), we weight the hedonics 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). Specifically, we use the following weights:

$$
w = \frac{S_{t,j,\text{year built},ACS}}{S_{t,j,year,MLS}} \tag{16}
$$

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. We then linearly interpolate the shares between 2000 and 2011 (the first year for which the ACS is available) for each zip code. We then 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.

We then calculate the market value (and rent) of each property assuming constant-structure characteristics and constant-structure prices. We calculate the predicted value (and rent) both in- and out-of-sample (that is, we 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 (15) for the year 2015. Thus the only hedonic coefficients that change in our predicted location values over time are the sets of $\beta_{0,k,t}$, γ_{jt} and δ_{jt}. We call the average log value (and rent) 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 location.

We can compare how our 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 our comparison are in Figure A.2 in the appendix. Due to the difference in measurement, the two measures will differ in levels. But, as can be seen in the figures, the correlation is extremely high; the median CBSA has a correlation of 0.81 between our zip code level measure of location value and the land value measure in Davis et al. (2021).

As we 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 tends to appreciate. 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 we focus on the returns to location, holding structure constant. Using the data from Davis et al. (2021), Figure 2 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 we will show later, our estimated location values tend to be much more volatile than structure values. Anticipating results below, Figure 2 shows that the greater returns that we 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, we 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 we do not have a good measure of certain costs, like maintenance. However, since we do observe property taxes, 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.

We calculate the capital gains to location as the annual log difference in predicted price

$$\text{Capital Gains}_{L,j,t} = \ln(\text{price}_{L,j,t}) - \ln(\text{price}_{L,j,t-1}). \tag{17}$$

We calculate the predicted yield of each location using the following equation:

$$\text{yield}_{j,t} = \exp\left\{\ln(\text{rent}_{L,j,t}) - \ln(\text{price}_{L,j,t}) + \frac{\sigma^2_{r,k,t} + \sigma^2_{p,k,t} - 2\text{cov}_k(\epsilon_r, \epsilon_p)}{2}\right\} \tag{18}$$

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 (15), 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:

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

We 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, we 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 our annual rent and price regressions and the same constant characteristics used in our location estimates. Similar to our estimates of location value, the levels of these values are not solely attributed to the structure[13], but any differences are solely attributable to the structure.

Last, we 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 our rent and price regressions, holding all characteristics constant but using the variation in all the hedonic coefficients.

We perform a check to validate our methodology. We compare our estimated location yields to the implied zip code-level location yields from our matched sample of property-level rent-price ratios. The yield estimates from the matched sample are from a single regression of the sample specification as in Equation 15, 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 pattern across areas within each city are very similar.

6 Cross-Sectional Results

In this section, after presenting some summary statistics on yields, capital gains and returns, we examine how our measures of risk and return are correlated in the cross-section with economic characteristics. We 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.

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

6.1 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 11 since we limit to our analysis to 2010–2020 (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.[14] Similar summary statistics for housing and structure returns are in Tables 2 and 3 respectively. Information on the variation in structure and location returns across CBSAs are in Table 4.

The tables validate some priors. Markets in the sunbelt have seen higher capital gains on average over our sample, while other cities (for example, St. Louis and Chicago) saw lower average capital gains. Variation also exists in yields, but we have fewer priors on what to expect.

6.2 Location returns within cities

Figures A.3, A.4, and A.5 feature binned scatter plots of the zip code-level average of each of capital gains, yields and total returns to location respectively against 2010 median household income for 20 of the 21 CBSAs in our sample[15]. The binned scatter plots are weighted by the number of households living in single unit housing units in according to the 2011 5-year American Community Survey.

These graphs illustrate a striking pattern. Capital gains, yields and total returns vary across zip codes within CBSAs. Specifically, they are higher in low-income zip codes. This is not just true by income. In Figure 4 we produce a similar set of graphs with the average Equifax Risk Score (a credit score) of the residents on the x-axis. Yields are higher in low-credit score zip codes.

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

$$y_{j,k} = \beta_{0,k} + \beta_{1,k} x_{j,k} + \varepsilon_{j,k} \tag{19}$$

where $y_{j,k}$ is the average annual yield, average annual capital gain, or average annual total

[14]Since we use the yield at time $t-1$ value, the earliest dates for which we have yields and total returns for our 21 CBSAs is 2010.

[15]For space reasons, we omit St. Louis.

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 share of the population that is black, the share of properties that are vacant, or the average credit score of the resident population in 2009. The regression is weighted by the number of single family housing units units in the ACS in 2011.

The results are in Tables 5 (yields), 6 (capital gains), and 7 (total returns). These indicate that across CBSAs location yields and our demographic/economic factors are generally all statistically significantly correlated in the same direction. Of the 21 CBSAs in our sample, lower income implies significantly higher yields in 11 CBSAs (and negative point estimates for all 21 CBSAs), higher black population shares implies higher yields in 17 CBSAs (with 20 positive point estimates), and 20 CBSA have higher yields in zips with lower average credit scores (all 21 have negative point estimates). No CBSA has a significant relationship with these factors in the opposite direction. By contrast, there is little evidence in-sample that significantly higher yields in zip codes with higher vacancy rates.[16]

The systematic relationship between capital gains and the income, race and credit score in an area is slightly noisier. 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, most but not all 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 all three factors (besides vacancy).[17] A zip code with double the median income or a 100 point higher average credit score than another within the same CBSA can expect anywhere between roughly 1 and 7 percent lower returns on their property per year. The results on the share of black residents are similarly striking. Besides St. Louis and Detroit, areas with high shares of black residents pay higher rents relative to prices so that an area with a 10 percentage point higher share has roughly between 0.3 and 4.8 percentage point higher returns.

Of course, area income, credit and race are all correlated. So we run a series of horse races in Appendix Tables A.4 (yields), A.5 (capital gains), and A.6 (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. Income and race become less important after controlling for credit and vacancy, though in several of the CBSAs where higher credit does not significantly

[16]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 our vacancy rates. 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).)

[17]Due to its capital gains patterns, Detroit is the only outlier here.

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).

To better examine whether race has a separate effect on housing returns, we perform a double sort of zip codes by race and credit score. We compare the average return for zip codes in the top and bottom terciles of the share of black residents for their CBSA conditional on being in the either the bottom or top tercile of credit scores for its CBSA. We also double sort on share of owner-occupants and race.

Table 8 shows that high black shares within zip codes affects returns primarily in *high* credit and high owner-occupied locations. Zip codes within the top terciles of credit score or ownership shares that also have high black population shares have approximately 0.6 percent higher annual returns on average than the same zip codes with low black population shares. Within the bottom tercile of credit or ownership, return differences across black population share are economically much smaller and statistically insignificant.[18] One possible reason for this may be that low credit score may suffice to curtail households' access to credit (particularly post-2008) and that, on the margin, race, conditional on low credit, does not further impair access to credit. Or it may be that, even if race further affects a households' access to credit conditional on low credit scores, landlords are the marginal owners in low credit locations and thus yields are unaffected by changes in households' access to credit in these areas.

6.2.1 Leverage

Figure 7 is a plot of the average FICO score of first-lien purchase mortgage originated in 2010 against the average LTV of those loans in that zip code. Higher FICO score neighborhoods have lower LTV loans on average. 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 we have

[18] 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.

documented 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.

6.3 Risk and Returns

We calculate several measures to see if differences in returns across zip codes are correlated with differences in risk. We have 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 we can discern some patterns in our results. The two measures of risk we estimate are the standard deviations in the year-on-year log differences in location yields and in capital gains. In Table 9 we show results of univariate regressions of these measures on credit score. Results using income and race are similar.

Point estimates indicate that for many CBSAs there were higher realized rent volatilities in areas with lower credit scores. For seven CBSAs, this relationship is significant at the five percent level. The relationship between capital gains volatility and credit scores is clearer: 14 CBSAs have significantly negative estimates of the effect of credit on volatility; only the California CBSAs and Boston report positive (but very much insignificant) point estimates.

Putting these results together, higher credit 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 (with indeterminate effects on Sharpe Ratios[19]). We develop this further below after first looking at the risk-return relationship through the lens of a standard CAPM regression.

6.3.1 CAPM

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

$$\begin{aligned}\text{Total Return}_{L,j,k,t} =& \beta_{0,k} + \beta_{1,k} R_m - R_f + \beta_{2,k}\text{Credit Score}_{j,k,2010} \\ & + \beta_{3,k}(R_m - R_f) \times \text{Credit Score}_{j,k,2010} + \beta_{4,k}\text{CBSA Return}_{L,k,t} \\ & + \beta_{5,k}\text{CBSA Return}_{L,k,t} \times \text{Credit Score}_{j,k,2010} + \epsilon_{j,k,t}\end{aligned} \tag{20}$$

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

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.[20] The CBSA return is the residual of the average total return to location in the CBSA regressed on national house price growth. It is conventional to include metro area housing returns in CAPM regressions of local returns. We 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 (thereby causing a national downturn in house values).

The results are in Table 10. β_2 measures whether ex-ante area credit scores can be used 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 20 out of 21 CBSAs. In 11 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 0.6 to 1.3 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 most 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 we discuss below, aggregate shocks that differentially affect how different households access credit or otherwise discount the future can generate this pattern.

6.4 Changes in credit and risk and returns

Our hypothesis is that differences in the discount rates caused perhaps by differences in the opportunity cost of credit can explain the differences in returns and yields within markets. Our best proxy for borrowing costs is lagged credit score. The local discount rate that prices a particular house can change over time either because of a change to the discount rate of local investors who *had* priced a home or because different owners (either landlords or households) with different discount rates are new marginal investors in the local area.

Historically, particularly in the last 20 years, the relationship between credit score and the access to credit has likely varied a lot over time. Figure 11 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 access to credit then. Around the onset of

[20] Available here.

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 an capital gains over time, we repeat our regressions in 21 of returns, yields and capital gains on credit score but allow the effects to vary with time:

$$y_{j,k,t} = \beta_{0,k,t} + \beta_{1,k,t} x_{j,k,t-2} + \varepsilon_{j,k,t} \tag{21}$$

In Figure 8 we plot the time-series profile of the univariate relationship between location yields and lagged credit scores. The sample varies by CBSA depending on data availability. For CBSA with data going far enough back, a pattern usually emerges: around 2007, just before the Great Recession, marginal effects are barely negative and in many cases statistically 0. In this period, credit score has virtually no affect on yields to housing across submarkets. After 2007, yield spreads widen in most CBSAs as access to credit (at least as proxied by our mortgage origination data) narrows. 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 2020.

Figure 9 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 CBSA and year. Though realized capital gains tend to be noisier than yields, many of the plots (e.g. Phoenix, Houston, all of California) show striking large positive marginal effects due to the onset of the Great Recession. In those CBSAs, sub-markets 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. Our plot from this CBSA shows a similar higher rise and steeper fall in areas with lower credit scores. While a few other CBSAs share this boom-bust dynamic with San Diego, many others, like Phoenix and San Francisco had similar booms across their zip codes but differential busts.

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)). We do not take a stand here on this debate, however it is natural to suppose that, should credit constraints be relaxed, the changes in house prices would likely be greater in areas where households were more likely to be ex ante constrained. Our 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.[21]

6.5 A series of quasi-natural experiments

To explore causally how changes in the access to credit affects yields and returns, we follow Loutskina and Strahan (2015) 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.

Our 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 we 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 NY CCP as our measure of treated mortgages. Our 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. We standardize the spread to have mean zero and unit standard-deviation over the sample.

Table 11 shows the results from 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 0.86 percentage point lower yields when the jumbo-conforming spread is one standard deviation above

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

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.[22] 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):

$$\log yield_{j,k,t} = q_{j,k} + \mathcal{I}_{j,k,t} - \mathcal{G}_{j,k,t} \tag{22}$$

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 out the latter, we also regress the one year growth in rents on the same explanatory variables and report those results in Table 11. 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

We measure the returns to housing and land in a large set of metropolitan areas in the United States. We find a very large dispersion in the average returns and yields to land that are correlated with many important demographic and economic characteristics. Variables which may proxy for access to credit are especially correlated with returns: areas wheres residents may have limited access have higher average returns. This return spread widens during periods when access to credit was particularly difficult. 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, we have shown that changes in access to credit leads 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

[22] 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.

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 8 remains an interesting avenue of future research.

References

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. doi:10.1093/rfs/hhaa087.

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. doi:10.1257/pol.20140074.

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. doi:10.1093/rfs/hhx035.

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. doi:10.1016/j.jfineco.2021.03.001.

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. doi:10.1016/j.jfineco.2019.03.013.

Bhutta, Neil and Aurel Hizmo (2020). "Do minorities pay more for mortgages?" *The Review of Financial Studies*, 34(2), pp. 763–789. doi:10.1093/rfs/hhaa047.

Campbell, John and Joao Cocco (2003). "Household risk management and optimal mortgage choice." *The Quarterly Journal of Economics*, 118(4), pp. 1449–1494. doi:10.1162/003355303322552847.

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. doi:10.1016/j.jmoneco.2005.10.016.

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. doi:10.1093/rfs/1.3.195.

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. doi:10.1016/j.jue.2009.06.002.

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. URL `https://www.jstor.org/stable/1804778`.

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. doi:10.1111/1540-6229.00521.

Cochrane, John H. (2011). "Presidential address: Discount rates." *The Journal of Finance*, 66(4), pp. 1047–1108. doi:10.1111/j.1540-6261.2011.01671.x.

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. doi:10.24149/wp2013.

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. doi:10.1016/j.jmoneco.2020.12.005.

Demers, Andrew and Andrea L. Eisfeldt (2022). "Total returns to single-family rentals." *Real Estate Economics*, 50(1), pp. 7–32. doi:10.1111/1540-6229.12353.

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. doi:10.1093/rfs/hhab042.

Eisfeldt, Andrea and Andrew Demers (2015). "Total returns to single family rentals." Working paper 21804, National Bureau of Economic Research. doi:10.3386/w21804.

Favara, Giovanni and Jean Imbs (2015). "Credit supply and the price of housing." *American Economic Review*, 105(3), pp. 958–992. doi:10.1257/aer.20121416.

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. doi:10.1111/jofi.13057.

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. doi:10.1257/aer.20140193.

Giacoletti, Marco (2021). “Idiosyncratic risk in housing markets.” *The Review of Financial Studies*, 34(8), pp. 3695–3741. doi:10.1093/rfs/hhab033.

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. doi:10.1093/rfs/hhab035.

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. doi:10.1080/10511482.2018.1524441.

Greenwald, Daniel and Adam Guren (2021). “Do credit conditions move house prices?” Working paper 29391, National Bureau of Economic Research. doi:10.3386/w29391.

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. doi:10.1016/j.jfineco.2020.06.014.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2013). “Endogenous gentrification and housing price dynamics.” *Journal of Public Economics*, 100, pp. 45–60. doi:10.1016/j.jpubeco.2013.02.001.

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. doi:10.1093/restud/rdaa018.

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. doi:10.1111/iere.12475.

Halket, Jonathan and Matteo Pignatti Morano di Custoza (2015). “Homeownership and the scarcity of rentals.” *Journal of Monetary Economics*, 76(C), pp. 107–123. doi:10.1016/j.jmoneco.2015.08.003.

Han, Lu (2013). “Understanding the Puzzling Risk-Return Relationship for Housing.” *The Review of Financial Studies*, 26(4), pp. 877–928. URL `https://www.jstor.org/stable/23355384`, publisher: Oxford University Press.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti (2015). "Household leveraging and deleveraging." *Review of Economic Dynamics*, 18(1), pp. 3–20. doi: 10.1016/j.red.2014.10.003.

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. doi: 10.1086/708816.

Kermani, Amir and Francis Wong (2022). "Racial disparities in housing returns." Working paper 29306, National Bureau of Economic Research. doi:10.3386/w29306.

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. doi:10.3368/le.89.1.1.

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. doi:10.1093/rfs/hhx026.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2015). "The housing market(s) of san diego." *American Economic Review*, 105(4), pp. 1371–1407. doi:10.1257/aer.20111662.

Loutskina, Elena and Philip E. Strahan (2015). "Financial integration, housing, and economic volatility." *Journal of Financial Economics*, 115(1), pp. 25–41. doi:10.1016/j.jfineco.2014.09.009.

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. doi:10.1016/j.red.2009.09.005.

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. doi:10.1111/j.1540-6261.2005.00759.x.

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. doi:10.1162/qjec.2009.124.4.1449.

Nesheim, Lars (2006). "Hedonic price functions." Technical Report 18/06, Institute for Fiscal Studies, Centre for Microdata Methods and Practice. doi:10.1920/wp.cem.2006.1806. 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. doi:10.1016/bs.hesmac.2016.06.003.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel (2020). “Segmented housing search.” *American Economic Review*, 110(3), pp. 720–759. doi:10.1257/aer.20141772.

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. doi:10.1093/rfs/hhq069.

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. doi:10.1162/JEEA.2008.6.2-3.705.

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. doi:10.1093/qje/120.2.763.

	Average			Std. Dev.		
	Yield	Cap Gains	Tot Return	Yield	Cap Gains	Tot Return
Atlanta, GA	8.01	6.34	14.36	2.14	7.71	8.78
Boston, MA-NH	6.16	4.88	11.04	0.94	3.76	3.55
Bridgeport, CT	4.57	1.22	5.79	0.56	4.01	4.22
Charlotte, NC-SC	8.41	6.20	14.61	3.01	4.90	6.83
Chicago, IL-IN-WI	7.67	1.25	8.92	1.98	4.78	4.75
Dallas, TX	7.71	6.86	14.57	1.21	4.08	4.65
Detroit, MI	8.76	5.53	14.29	1.27	6.69	6.44
Hartford, CT	8.31	1.20	9.50	1.24	2.75	2.92
Houston, TX	6.91	5.64	12.55	0.92	3.65	3.77
Jacksonville, FL	6.70	4.41	11.11	1.96	7.04	5.45
Los Angeles, CA	4.09	4.76	8.84	0.38	4.55	3.88
Miami, FL	6.98	5.86	12.84	1.45	6.96	6.22
Orlando, FL	5.24	5.29	10.52	0.92	7.29	6.72
Phoenix, AZ	5.65	6.41	12.06	0.89	6.58	6.32
Riverside, CA	5.06	4.91	9.97	0.64	4.79	4.28
San Diego, CA	5.75	4.83	10.58	0.60	3.29	2.91
San Francisco, CA	3.39	3.62	7.01	0.54	5.30	5.08
St. Louis, MO-IL	6.66	3.07	9.73	1.57	3.25	2.77
Tampa, FL	8.30	5.16	13.46	3.17	6.24	4.79
Tucson, AZ	5.42	3.81	9.22	1.66	5.20	5.19
Virginia Beach, VA-NC	8.21	2.52	10.73	0.99	3.88	3.44

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.

	Average			Std. Dev.		
	Yield	Cap Gains	Tot Return	Yield	Cap Gains	Tot Return
Atlanta, GA	7.76	4.02	11.78	1.79	6.42	6.89
Boston, MA-NH	6.08	3.98	10.06	1.40	3.79	3.93
Bridgeport, CT	5.08	0.52	5.61	0.59	4.04	3.73
Charlotte, NC-SC	6.70	3.95	10.65	1.19	4.73	5.31
Chicago, IL-IN-WI	7.81	0.73	8.54	2.27	4.50	5.02
Dallas, TX	6.76	4.97	11.73	1.12	4.01	4.06
Detroit, MI	8.29	4.46	12.75	1.66	6.74	6.81
Hartford, CT	6.06	0.47	6.53	0.83	2.62	2.94
Houston, TX	7.08	3.85	10.93	0.97	3.88	3.41
Jacksonville, FL	5.83	2.66	8.49	1.14	5.86	6.16
Los Angeles, CA	3.94	3.77	7.71	0.29	4.24	3.39
Miami, FL	6.62	4.08	10.70	0.88	5.56	4.29
Orlando, FL	5.96	3.45	9.41	1.04	6.50	5.43
Phoenix, AZ	5.25	4.48	9.73	0.69	6.90	6.36
Riverside, CA	5.62	4.07	9.69	0.75	4.75	3.09
San Diego, CA	5.22	3.86	9.08	0.86	3.38	3.00
San Francisco, CA	4.11	3.05	7.17	0.59	5.64	5.42
St. Louis, MO-IL	7.43	2.35	9.78	1.02	3.05	2.46
Tampa, FL	6.52	3.76	10.28	1.15	5.90	6.27
Tucson, AZ	4.94	2.12	7.06	1.12	5.04	4.79
Virginia Beach, VA-NC	7.00	1.57	8.57	0.78	3.52	4.42

Table 2: Summary Statistics for Housing Returns 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.

	Average			Std. Dev.		
	Yield	Cap Gains	Tot Return	Yield	Cap Gains	Tot Return
Atlanta, GA	8.13	–2.33	5.80	0.57	1.89	2.19
Boston, MA-NH	6.84	–0.90	5.95	1.61	0.43	1.45
Bridgeport, CT	4.12	–0.69	3.43	0.40	0.76	0.75
Charlotte, NC-SC	6.33	–2.26	4.07	1.19	1.45	2.31
Chicago, IL-IN-WI	5.57	–0.52	5.05	0.52	0.52	0.77
Dallas, TX	7.26	–1.89	5.37	0.50	0.88	1.14
Detroit, MI	8.81	–1.07	7.74	1.05	0.90	0.72
Hartford, CT	5.78	–0.73	5.05	1.16	0.68	1.36
Houston, TX	7.74	–1.79	5.95	0.18	1.53	1.18
Jacksonville, FL	6.42	–1.75	4.67	2.04	1.86	3.00
Los Angeles, CA	4.04	–0.99	3.06	0.28	0.83	1.03
Miami, FL	5.04	–1.78	3.26	0.51	1.97	2.26
Orlando, FL	6.45	–1.84	4.61	0.66	1.93	2.15
Phoenix, AZ	4.95	–1.93	3.02	0.72	0.83	1.00
Riverside, CA	5.70	–0.84	4.86	0.92	0.40	1.83
San Diego, CA	5.84	–0.97	4.87	1.32	0.82	1.27
San Francisco, CA	6.59	–0.57	6.02	1.01	0.99	1.26
St. Louis, MO-IL	8.51	–0.72	7.79	1.58	0.73	1.52
Tampa, FL	6.60	–1.40	5.20	1.86	1.43	2.47
Tucson, AZ	5.33	–1.69	3.64	2.02	0.98	2.14
Virginia Beach, VA-NC	6.65	–0.95	5.70	0.69	0.88	0.73

Table 3: Summary Statistics for Structure Returns 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.

	Mean	Std. Dev.	Min	Max
Capital Gains…				
Structure	–0.01	0.01	–0.02	–0.01
Land	0.04	0.02	0.01	0.07
Log Yields…				
Structure	–2.83	0.21	–3.23	–2.47
Land	–2.83	0.24	–3.44	–2.53
Jensen	0.04	0.03	–0.00	0.10

Table 4: 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.

	ln(Median Income)	Share Black (%)	Share Vacant (%)	Average Credit Score
Atlanta, GA	-0.194***	0.00401***	0.0182*	-0.00299***
	(0.0610)	(0.00101)	(0.0105)	(0.000360)
Boston, MA	-0.365***	0.0201***	-0.0143	-0.00453***
	(0.0533)	(0.00775)	(0.0250)	(0.000514)
Bridgeport, CT	-0.341***	0.0113***	0.0205	-0.00398***
	(0.0549)	(0.00318)	(0.0443)	(0.000754)
Charlotte, NC	-0.107	0.00468***	-0.0428**	-0.00323***
	(0.124)	(0.00176)	(0.0182)	(0.000870)
Chicago, IL	-0.115**	0.00398	-0.0175*	-0.00280***
	(0.0551)	(0.00283)	(0.00943)	(0.000459)
Dallas, TX	-0.000764	0.00399***	-0.0155***	-0.00205***
	(0.0300)	(0.000775)	(0.00562)	(0.000308)
Detroit, MI	-0.504***	0.00614***	0.0363***	-0.00468***
	(0.0415)	(0.000637)	(0.00398)	(0.000253)
Hartford, CT	-0.229**	0.00805***	-0.0107	-0.00322***
	(0.114)	(0.00310)	(0.0233)	(0.000969)
Houston, TX	-0.0546	0.00306***	-0.0113	-0.00232***
	(0.0375)	(0.000954)	(0.00724)	(0.000350)
Jacksonville, FL	-0.173	0.00777***	0.0188	-0.00308***
	(0.119)	(0.00188)	(0.0152)	(0.000905)
Los Angeles, CA	-0.224***	0.0265***	-0.0195	-0.00313***
	(0.0466)	(0.00724)	(0.0150)	(0.000507)
Miami, FL	-0.0742**	0.00256***	-0.000610	-0.000423
	(0.0366)	(0.000881)	(0.00444)	(0.000434)
Orlando, FL	-0.0862	0.0111***	-0.0304***	-0.00242***
	(0.0908)	(0.00344)	(0.00733)	(0.000610)
Phoenix, AZ	-0.0426	0.00964***	-0.00588**	-0.000925***
	(0.0307)	(0.00322)	(0.00238)	(0.000225)
Riverside, CA	-0.107**	0.00406	0.00213	-0.00191***
	(0.0535)	(0.00509)	(0.00532)	(0.000536)
St. Louis, MO	-0.0897	-0.0156	-0.0589***	-0.00324**
	(0.145)	(0.0101)	(0.0172)	(0.00143)
San Diego, CA	-0.272***	0.0297***	-0.0332	-0.00357***
	(0.0656)	(0.00831)	(0.0403)	(0.000590)
San Francisco, CA	-0.454***	0.0244***	0.0635**	-0.00475***
	(0.0741)	(0.00382)	(0.0286)	(0.000475)
Tampa, FL	-0.0531	0.0116***	-0.0326***	-0.00344***
	(0.0833)	(0.00375)	(0.00495)	(0.000754)
Tucson, AZ	0.0126	0.00758	-0.00780*	-0.000933**
	(0.0461)	(0.00891)	(0.00420)	(0.000471)
Virginia Beach, VA	-0.155*	0.00548***	0.0505**	-0.00179**
	(0.0869)	(0.00110)	(0.0236)	(0.000735)

Table 5: Univariate determinants of Log Yields. *Note:* Coefficient estimates of univariate regressions of log location yields on factors. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and the FRBNY/Equifax CCP.

	ln(Median Income)	Share Black (%)	Share Vacant (%)	Average Credit Score
Atlanta, GA	-0.0281***	0.000577***	0.00458***	-0.000303***
	(0.00653)	(0.000105)	(0.00108)	(0.0000437)
Boston, MA	-0.0125*	0.00173***	0.000336	-0.000118
	(0.00701)	(0.000639)	(0.00210)	(0.0000822)
Bridgeport, CT	-0.0128***	0.000694***	0.00296	-0.000197***
	(0.00285)	(0.000109)	(0.00189)	(0.0000293)
Charlotte, NC	-0.0210***	0.000449***	0.00265***	-0.000207***
	(0.00463)	(0.0000491)	(0.000779)	(0.0000321)
Chicago, IL	-0.00780**	0.0000465	0.00161***	-0.0000948***
	(0.00348)	(0.000181)	(0.000587)	(0.0000325)
Dallas, TX	-0.00595***	0.000245***	0.00125***	-0.000126***
	(0.00178)	(0.0000474)	(0.000338)	(0.0000188)
Detroit, MI	0.0416***	-0.000976***	-0.00908***	0.000515***
	(0.0140)	(0.000170)	(0.000759)	(0.000107)
Hartford, CT	-0.00801	-0.0000729	0.00113	-0.0000470
	(0.00811)	(0.000237)	(0.00154)	(0.0000793)
Houston, TX	-0.00895***	0.000214***	0.000265	-0.000113***
	(0.00211)	(0.0000559)	(0.000434)	(0.0000218)
Jacksonville, FL	0.000706	-0.0000392	0.00106*	0.0000180
	(0.00495)	(0.0000948)	(0.000593)	(0.0000428)
Los Angeles, CA	-0.0127***	0.00170**	-0.00397***	-0.000316***
	(0.00474)	(0.000713)	(0.00139)	(0.0000471)
Miami, FL	-0.0224***	0.000605***	0.00154**	-0.000202***
	(0.00504)	(0.000121)	(0.000643)	(0.0000614)
Orlando, FL	-0.0206**	0.000752*	0.00261***	-0.0000448
	(0.00910)	(0.000389)	(0.000826)	(0.0000755)
Phoenix, AZ	-0.0229***	0.00217***	0.000807**	-0.000248***
	(0.00424)	(0.000474)	(0.000371)	(0.0000286)
Riverside, CA	-0.0156**	0.00110*	-0.0000337	-0.000299***
	(0.00691)	(0.000646)	(0.000698)	(0.0000642)
St. Louis, MO	-0.00269	-0.0000149	-0.00129***	-0.0000908***
	(0.00333)	(0.000246)	(0.000409)	(0.0000309)
San Diego, CA	-0.0200**	0.00292***	0.00149	-0.000266***
	(0.00807)	(0.000919)	(0.00435)	(0.0000815)
San Francisco, CA	-0.0539***	0.00324***	0.00667*	-0.000596***
	(0.00879)	(0.000390)	(0.00344)	(0.0000475)
Tampa, FL	-0.0163***	-0.000106	0.000282	-0.0000462
	(0.00383)	(0.000207)	(0.000328)	(0.0000444)
Tucson, AZ	-0.00758	0.00225*	-0.000414	-0.000284***
	(0.00692)	(0.00131)	(0.000687)	(0.0000511)
Virginia Beach, VA	-0.000319	0.0000456	-0.000393	-0.0000435*
	(0.00283)	(0.0000436)	(0.000780)	(0.0000238)

Table 6: Univariate determinants of Capital Gains. *Note:* Coefficient estimates of univariate regressions of location capital gains on factors. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and the FRBNY/Equifax CCP.

	ln(Median Income)	Share Black (%)	Share Vacant (%)	Average Credit Score
Atlanta, GA	-0.0427***	0.000884***	0.00598***	-0.000522***
	(0.00878)	(0.000138)	(0.00151)	(0.0000483)
Boston, MA	-0.0343***	0.00304***	-0.000441	-0.000392***
	(0.00709)	(0.000744)	(0.00276)	(0.0000821)
Bridgeport, CT	-0.0278***	0.00123***	0.00393	-0.000378***
	(0.00439)	(0.000208)	(0.00352)	(0.0000485)
Charlotte, NC	-0.0321***	0.000853***	0.000373	-0.000467***
	(0.0113)	(0.000130)	(0.00192)	(0.0000661)
Chicago, IL	-0.0160***	0.000349	0.000518	-0.000292***
	(0.00476)	(0.000252)	(0.000851)	(0.0000381)
Dallas, TX	-0.00642**	0.000551***	0.000256	-0.000277***
	(0.00311)	(0.0000763)	(0.000603)	(0.0000293)
Detroit, MI	-0.000430	-0.000438**	-0.00581***	0.000116
	(0.0133)	(0.000174)	(0.000918)	(0.000108)
Hartford, CT	-0.0270***	0.000584**	0.000667	-0.000310***
	(0.00863)	(0.000272)	(0.00197)	(0.0000741)
Houston, TX	-0.0130***	0.000412***	-0.000372	-0.000262***
	(0.00314)	(0.0000799)	(0.000643)	(0.0000276)
Jacksonville, FL	-0.0118	0.000482***	0.00241**	-0.000185**
	(0.00926)	(0.000159)	(0.00111)	(0.0000750)
Los Angeles, CA	-0.0221***	0.00278***	-0.00464***	-0.000442***
	(0.00525)	(0.000803)	(0.00161)	(0.0000487)
Miami, FL	-0.0272***	0.000776***	0.00156**	-0.000226***
	(0.00619)	(0.000147)	(0.000797)	(0.0000761)
Orlando, FL	-0.0251***	0.00129***	0.00125	-0.000158**
	(0.00951)	(0.000382)	(0.000964)	(0.0000769)
Phoenix, AZ	-0.0254***	0.00271***	0.000499	-0.000297***
	(0.00487)	(0.000530)	(0.000430)	(0.0000313)
Riverside, CA	-0.0207**	0.00127	0.0000721	-0.000390***
	(0.00833)	(0.000793)	(0.000853)	(0.0000750)
St. Louis, MO	-0.00870	-0.000878	-0.00478***	-0.000290***
	(0.0105)	(0.000755)	(0.00117)	(0.0000974)
San Diego, CA	-0.0359***	0.00467***	-0.000121	-0.000471***
	(0.0100)	(0.00116)	(0.00592)	(0.0000950)
San Francisco, CA	-0.0696***	0.00412***	0.00901**	-0.000762***
	(0.0105)	(0.000467)	(0.00424)	(0.0000518)
Tampa, FL	-0.0199***	0.000726*	-0.00207***	-0.000296***
	(0.00756)	(0.000371)	(0.000549)	(0.0000738)
Tucson, AZ	-0.00708	0.00264	-0.000797	-0.000333***
	(0.00867)	(0.00163)	(0.000842)	(0.0000679)
Virginia Beach, VA	-0.0148	0.000543***	0.00466*	-0.000208***
	(0.00972)	(0.000129)	(0.00266)	(0.0000807)

Table 7: UNIVARIATE DETERMINANTS OF TOTAL RETURNS. *Note:* Coefficient estimates of univariate regressions of location total returns on factors. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and the FRBNY/Equifax CCP.

	Bottom Tercile		Top Tercile			
	N	Avg. Return	N	Avg. Return	P	P_L
Share Black w/n Low Credit Score	151	0.81	132	0.90	0.79	0.40
Share Black w/n High Credit Score	198	-1.38	150	-0.81	0.00	0.00
Share Black w/n Low Owner-Occupancy	125	-0.89	207	-0.77	0.71	0.35
Share Black w/n High Owner-Occupancy	186	-0.19	122	0.55	0.00	0.00

Table 8: Double Sort of Total Location Returns. *Note:* Both 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. 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 FRBNY/Equifax CCP; 2010 Decennial Census; Corelogic MLS.

	Location Sharpe Ratio	Location Rent Volatility	Location Cap Gains Volatility
Atlanta, GA	-0.000733	-0.00108***	-0.000318***
	(0.000780)	(0.000329)	(0.0000465)
Boston, MA	-0.0810***	0.000515	0.00000550
	(0.0247)	(0.000863)	(0.000209)
Bridgeport, CT	-0.00154	0.0000324	-0.000264***
	(0.00228)	(0.000263)	(0.0000871)
Charlotte, NC	-0.00146*	-0.000573	-0.000171***
	(0.000833)	(0.000373)	(0.0000280)
Chicago, IL	0.0140**	-0.00287***	-0.000290***
	(0.00595)	(0.000802)	(0.0000415)
Dallas, TX	0.00113	-0.000704***	-0.000150***
	(0.00113)	(0.000169)	(0.0000244)
Detroit, MI	0.00937	-0.000298***	-0.000610***
	(0.00927)	(0.0000942)	(0.0000760)
Houston, TX	0.00443	-0.000674***	-0.000145***
	(0.00278)	(0.000221)	(0.0000380)
Jacksonville, FL	0.0218***	-0.000732*	-0.000309***
	(0.00597)	(0.000409)	(0.0000692)
Los Angeles, CA	-0.00818*	0.000335*	0.0000648
	(0.00471)	(0.000180)	(0.0000545)
Miami, FL	0.0108***	-0.000682*	-0.000335***
	(0.00265)	(0.000391)	(0.0000617)
New Haven, CT	0.000904	0.000328	-0.000414***
	(0.00226)	(0.000200)	(0.000110)
Orlando, FL	0.00103	-0.000341	-0.000248***
	(0.00306)	(0.000604)	(0.0000897)
Phoenix, AZ	0.00198*	-0.000315**	-0.000157***
	(0.00110)	(0.000139)	(0.0000286)
Riverside, CA	-0.00916***	-0.0000183	0.0000128
	(0.00325)	(0.000224)	(0.0000509)
San Diego, CA	-0.00840	0.000310	0.0000396
	(0.00864)	(0.000278)	(0.0000883)
San Francisco, CA	-0.0118***	0.000297	0.000179
	(0.00450)	(0.000248)	(0.000150)
Tampa, FL	0.135*	-0.00121***	-0.000232***
	(0.0817)	(0.000319)	(0.0000701)
Tucson, AZ	0.00219	-0.00176**	-0.000216***
	(0.00217)	(0.000889)	(0.0000526)
Virginia Beach, VA	-0.0439	-0.000407**	-0.000269***
	(0.0531)	(0.000187)	(0.0000366)

Table 9: Regressions of Zip Code Measures of Risk on Zip Code Average Equifax Risk Score. *Note:* Values are coefficients from regressions of each risk measure on the average zip code-level Equifax Risk Score as of 2010. 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 the FRBNY/Equifax CCP.

	$R_m - R_f$	Credit Score	$R_m - R_f \times$ Credit Score	CBSA Net Return	CBSA Net Return $\times$ Credit Score	Constant
Atlanta, GA	0.123***	-1.170***	-0.0319*	1.021***	-0.118***	11.03***
	(0.0146)	(0.331)	(0.0176)	(0.0365)	(0.0449)	(0.279)
Boston, MA	0.121*	-0.158	-0.0963	1.203***	-0.158	9.631***
	(0.0704)	(1.320)	(0.0674)	(0.373)	(0.352)	(1.368)
Bridgeport, CT	0.110***	-1.164**	0.00619	1.715***	-0.480**	4.422***
	(0.0320)	(0.458)	(0.0240)	(0.271)	(0.202)	(0.613)
Charlotte, NC	0.0612***	-1.282***	-0.0242	0.966***	-0.0660	12.34***
	(0.0144)	(0.305)	(0.0156)	(0.0954)	(0.105)	(0.276)
Chicago, IL	0.110***	-1.075***	0.00180	1.066***	-0.137**	6.125***
	(0.0150)	(0.298)	(0.0155)	(0.0603)	(0.0616)	(0.286)
Dallas, TX	0.0599***	-0.811***	-0.000170	1.054***	-0.0496	12.88***
	(0.00838)	(0.162)	(0.00854)	(0.0379)	(0.0379)	(0.159)
Detroit, MI	0.113***	-1.154***	-0.0190	1.076***	-0.186***	11.16***
	(0.0144)	(0.280)	(0.0148)	(0.0492)	(0.0504)	(0.274)
Hartford, CT	0.0909***	-1.170**	0.0105	1.342***	-0.330*	7.495***
	(0.0326)	(0.572)	(0.0299)	(0.189)	(0.174)	(0.616)
Houston, TX	0.0393***	-0.828***	-0.00533	1.014***	0.0524	10.62***
	(0.00993)	(0.162)	(0.00834)	(0.0719)	(0.0599)	(0.192)
Jacksonville, FL	0.185***	-0.590	0.00339	1.169***	-0.240**	7.372***
	(0.0276)	(0.649)	(0.0340)	(0.0945)	(0.121)	(0.522)
Los Angeles, CA	0.124***	-0.502*	-0.0627***	0.990***	-0.0298	7.632***
	(0.0124)	(0.262)	(0.0136)	(0.0506)	(0.0551)	(0.240)
Miami, FL	0.145***	-0.677*	0.0146	0.878***	-0.142**	9.548***
	(0.0198)	(0.410)	(0.0210)	(0.0650)	(0.0686)	(0.385)
Orlando, FL	0.116***	-0.0119	-0.0463	0.975***	-0.0388	8.066***
	(0.0298)	(0.633)	(0.0321)	(0.0915)	(0.0937)	(0.598)
Phoenix, AZ	0.190***	-0.891***	-0.0117	0.761***	-0.0199	8.816***
	(0.0135)	(0.260)	(0.0133)	(0.139)	(0.138)	(0.263)
Riverside, CA	0.129***	-1.280***	0.00658	0.955***	-0.119	7.202***
	(0.0196)	(0.440)	(0.0227)	(0.0861)	(0.104)	(0.380)
St. Louis, MO	0.222***	0.476	-0.120***	1.472***	0.0849	6.392***
	(0.0392)	(0.637)	(0.0330)	(0.433)	(0.366)	(0.755)
San Diego, CA	0.0955***	-0.595	-0.0499**	0.974***	0.0299	9.502***
	(0.0202)	(0.390)	(0.0199)	(0.146)	(0.144)	(0.396)
San Francisco, CA	0.0999***	-1.249***	-0.0741***	0.679***	0.137**	7.584***
	(0.0266)	(0.392)	(0.0203)	(0.0759)	(0.0582)	(0.516)
Tampa, FL	0.161***	-0.544	-0.0218	1.155***	-0.184	9.936***
	(0.0161)	(0.431)	(0.0221)	(0.108)	(0.148)	(0.312)
Tucson, AZ	0.175***	-0.104	-0.0467*	2.194***	-0.843***	6.156***
	(0.0238)	(0.451)	(0.0249)	(0.256)	(0.258)	(0.441)
Virginia Beach, VA	0.0884***	-0.382	-0.0214	0.942***	-0.320***	8.494***
	(0.0170)	(0.347)	(0.0178)	(0.103)	(0.111)	(0.331)

Table 10: DETERMINANTS OF β AND α. *Note:* The average credit score is normalized to have mean zero and standard deviation equal to 1. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the FRBNY/Equifax CCP, and Fama-French factors.

	$\Delta \ln(\text{Yield}_t)$				$\Delta \ln(\text{Rent}_{t+1})$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated Loan Originations (%)	0.00397*	0.00312	0.00121		0.0735	0.0473	-0.00942	
	(0.00205)	(0.00215)	(0.00237)		(0.198)	(0.208)	(0.230)	
Treated Loan Originations (%) × Jumbo Conforming Spread	-0.00874***	-0.00875***	-0.00861***		0.0531	0.0522	0.0549	
	(0.00160)	(0.00160)	(0.00160)		(0.152)	(0.152)	(0.152)	
Non-Conforming Originations (%)				-0.000170				-0.0101
				(0.000759)				(0.0734)
Non-Conforming Originations (%) × Jumbo Conforming Spread				-0.00288***				0.0152
				(0.000437)				(0.0417)
Average Credit Score (Normalized)		0.00175	-0.00153	-0.00210		0.0557	-0.0445	-0.0517
		(0.00133)	(0.00217)	(0.00223)		(0.134)	(0.219)	(0.225)
ln(Average Household Income)			0.00923*	0.0125**			0.285	0.318
			(0.00483)	(0.00531)			(0.491)	(0.539)
N	13225	13225	13225	13225	14439	14439	14439	14439
R2	0.48	0.48	0.48	0.48	0.34	0.34	0.34	0.34
CBSA × Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var. (%)	-.01	-.01	-.01	-.01	2.61	2.61	2.61	2.61
Std. Dev. of Dep. Var.	.18	.18	.18	.18	18.	18.	18.	18.
Mean of Treated Share (%)	.35	.35	.35	.35	.36	.36	.36	.36
Std. Dev. Treated Share	.78	.78	.78	.78	.80	.80	.80	.80
Sample	2002–2018	2002–2018	2002–2018	2002–2018	2001–2018	2001–2018	2001–2018	2001–2018

Table 11: 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 the NY CCP. The share non-conforming is the two year lagged share by number of non-conforming loan originations according to the NY CCP. 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, the FRBNY/Equifax CCP, and Freddie Mac.

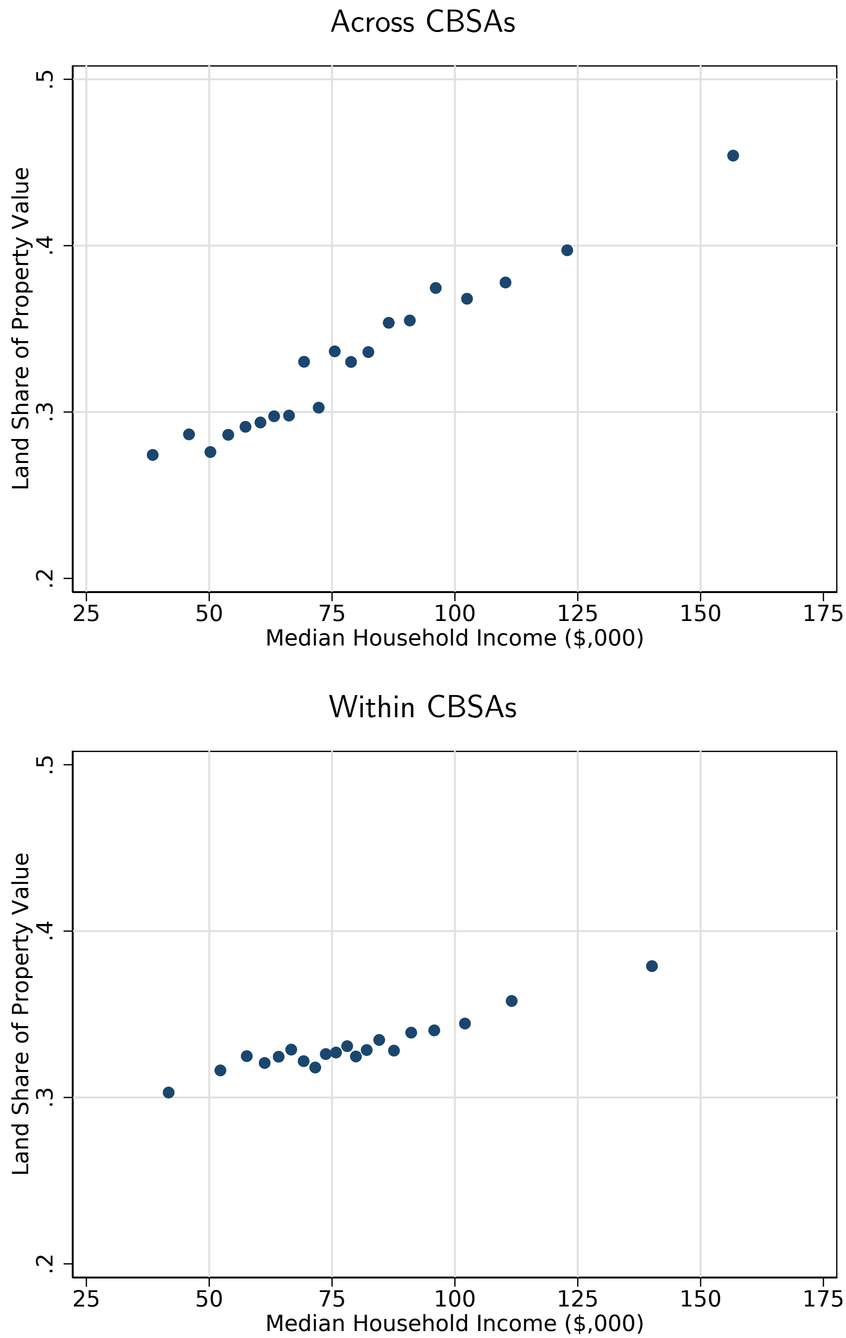

Figure 2: The Land Share of Property Values by Median Household Income. *Note:* Values are by census-tract. The land share of property values is measured as of 2012. Median household income is measured as of 2010. *Source:* Davis et al. (2021) and the Decennial Census.

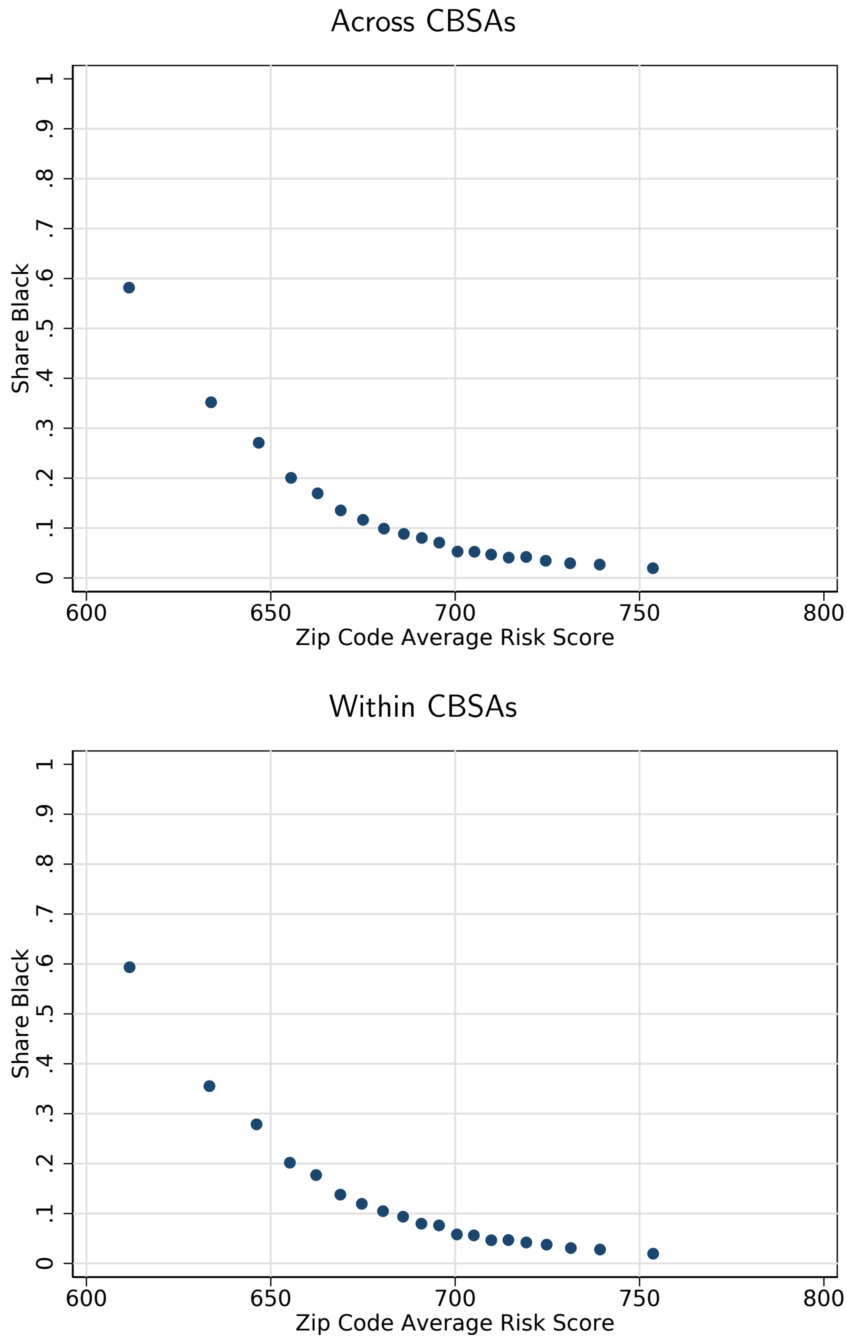

Figure 3: Black Population Share Vs. Average Risk Score. *Note:* Values are for the calendar year 2009. The share black is from the 2010 Decennial Census, for which the survey is conducted in 2009. The average credit score is the average values by zip code according the FRBNY/Equifax CCP in 2009. *Source:* The NY CCP and the Decennial Census.

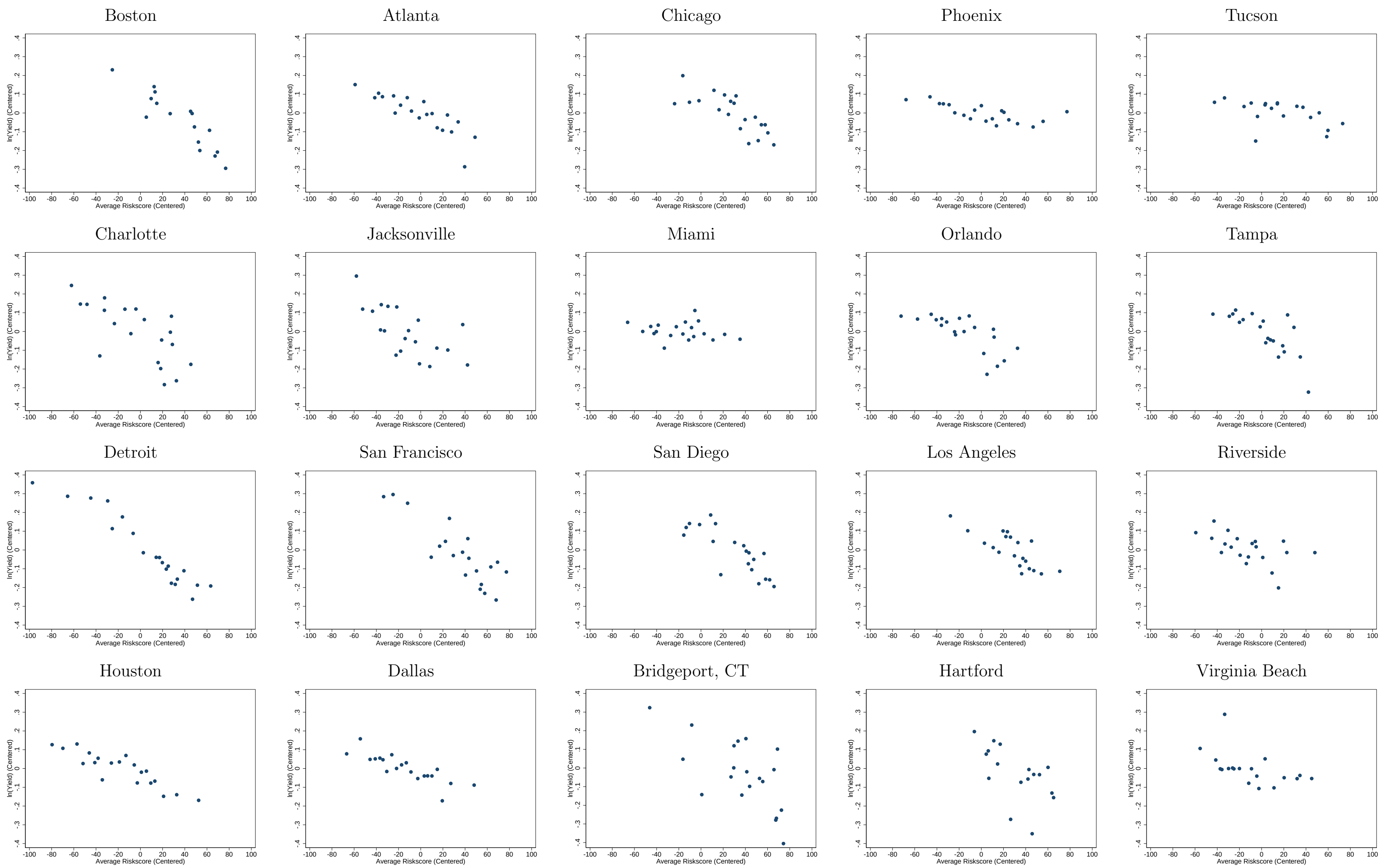

Figure 4: Log Location Yields by Average Equifax Riskscore of Population. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011 and all values are centered relative to their weighted mean. The average riskscore is measured in 2009. Data is limited to 2010–2021. *Source:* Authors' calculations using Corelogic MLS data, the American Community Survey, and and the FRBNY/Equifax CCP.

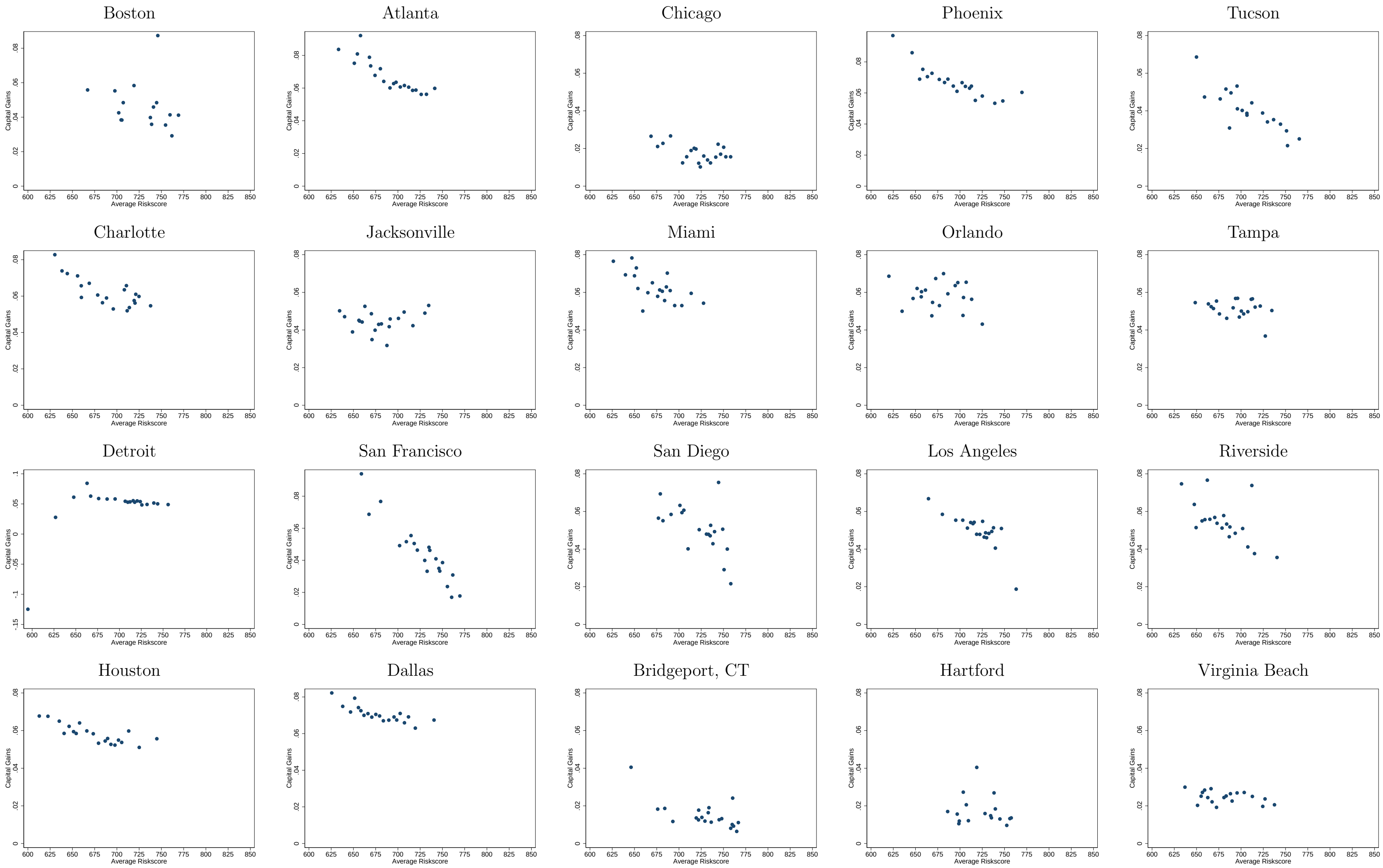

Figure 5: Capital Gains by Location and Average Equifax Riskscore of Population. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011. The average riskscore is measured in 2009. Data is limited to 2010–2021. *Source:* Authors' calculations using the Corelogic MLS data, the American Community Survey, and and the FRBNY/Equifax CCP.

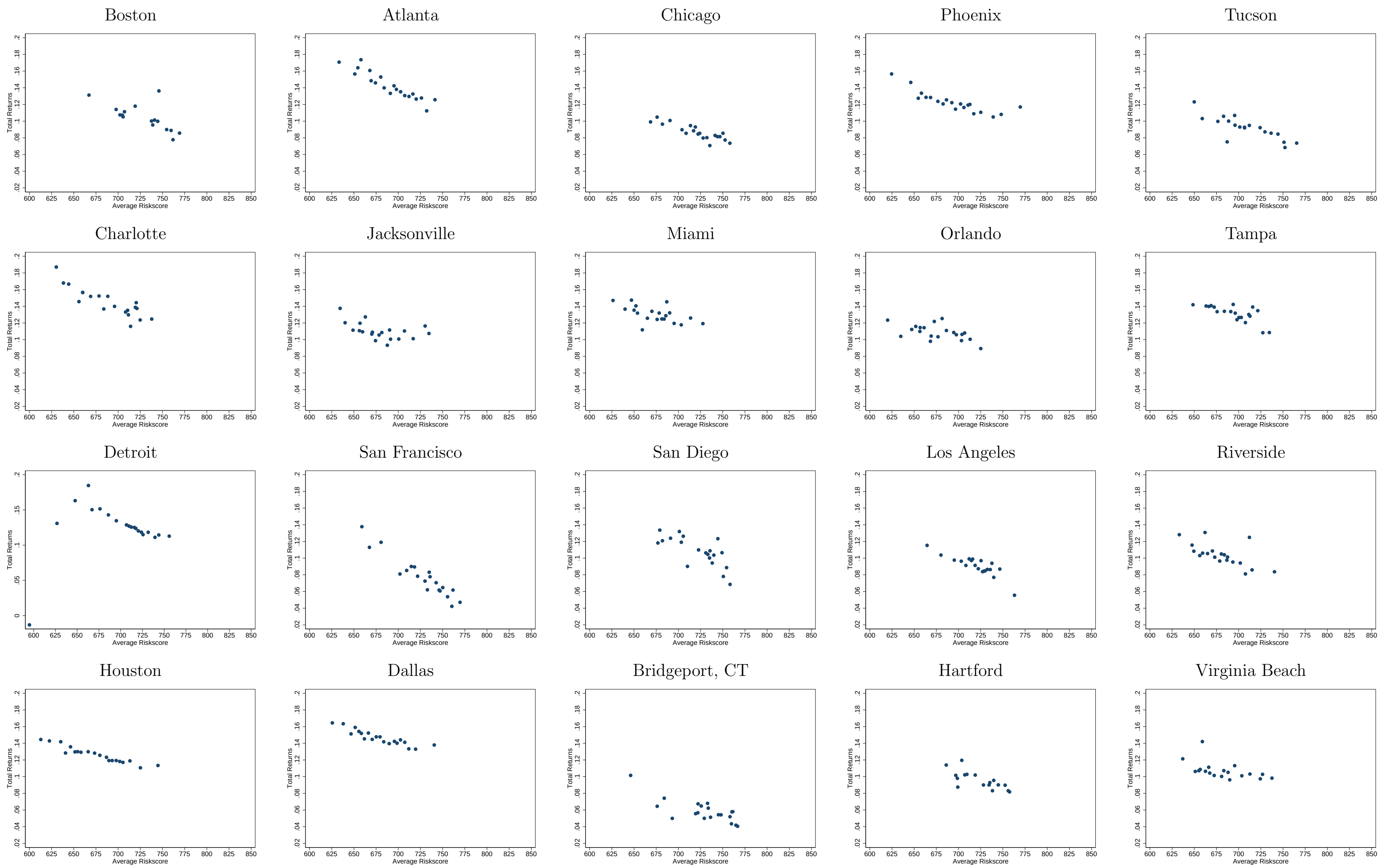

Figure 6: Total Returns to Location by Average Equifax Riskscore of Population. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011. The average riskscore is measured in 2009. Data is limited to 2010–2021. *Source:* Corelogic MLS data, the American Community Survey, and and the FRBNY/Equifax CCP.

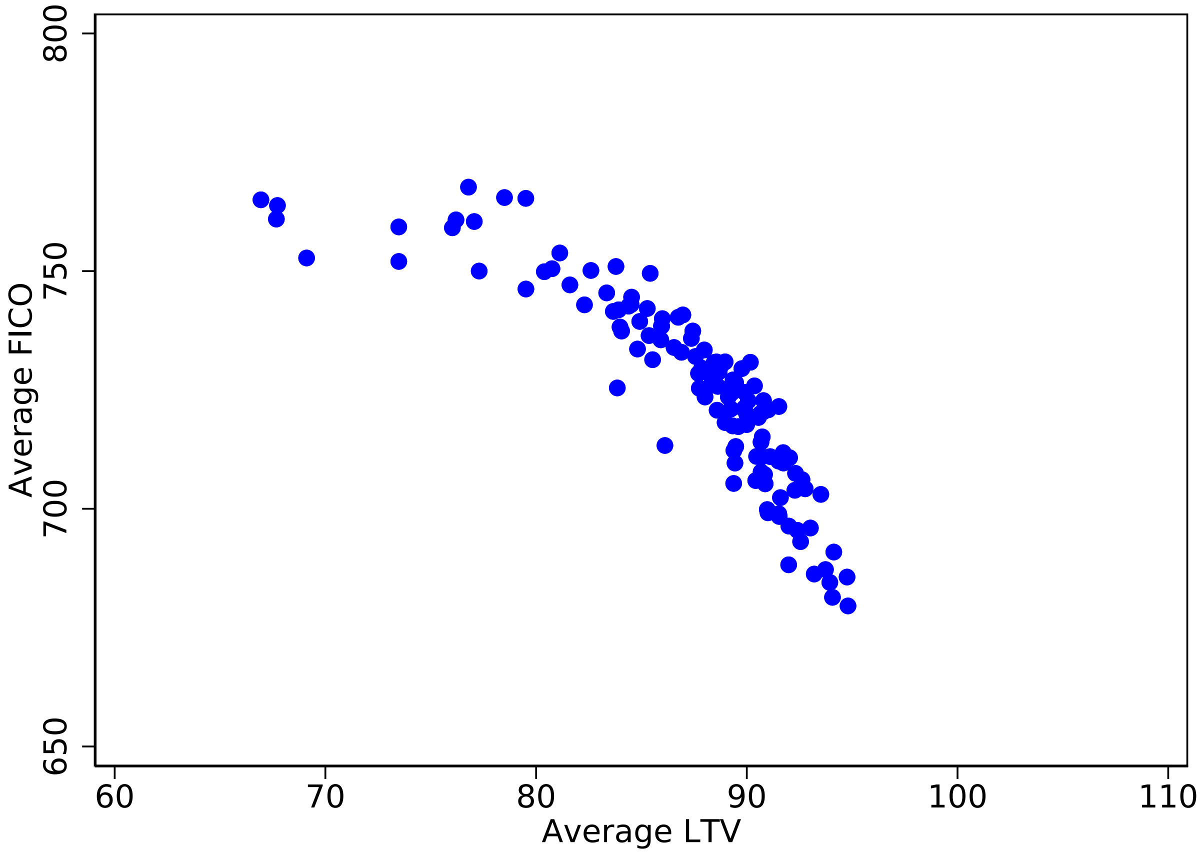

Figure 7: LTVs and FICO Scores. *Note:* Average values by zip code for first-lien purchase mortgage originations in 2010. *Source:* Authors' calculations using Black Knight Analytics servicing data.

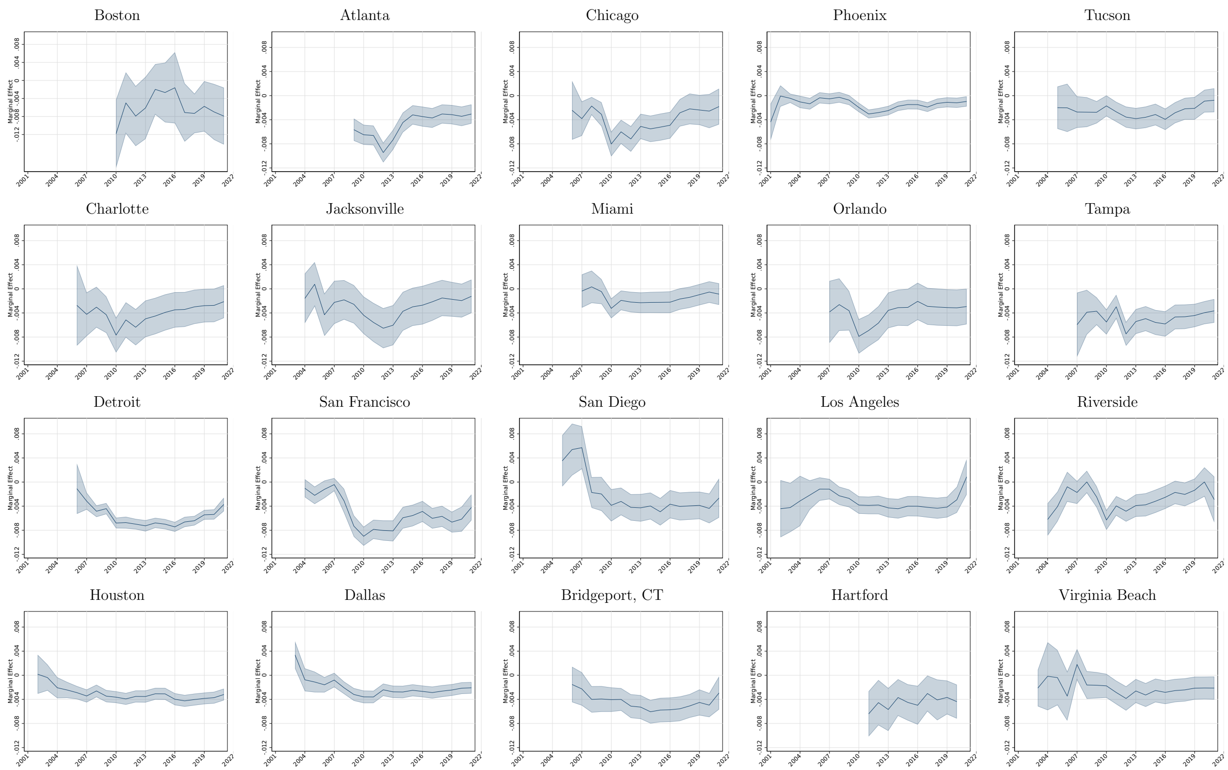

Figure 8: Relationship Between Log Location Yields and Credit Scores Over Time. Values are slope coefficients from CBSA-level regressions of zip code-level location yields on the two year lagged zip code average Equifax Risk Score. *Source:* Authors' calculations using Corelogic MLS and the FRBNY/Equifax CCP.

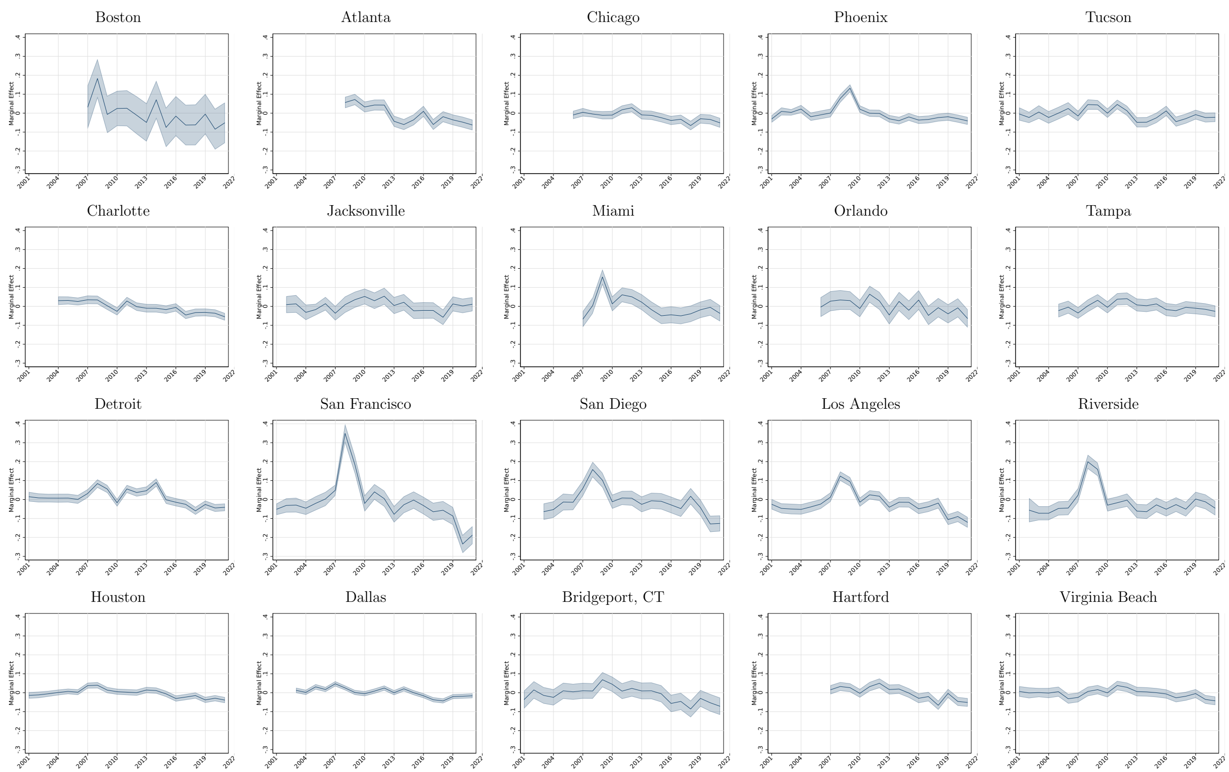

Figure 9: Relationship Between Location Capital Gains and Credit Scores Over Time. Values are slope coefficients from CBSA-level regressions of zip code-level capital gains on the two year lagged zip code average Equifax Risk Score. *Source:* Authors' calculations using Corelogic MLS and the FRBNY/Equifax CCP.

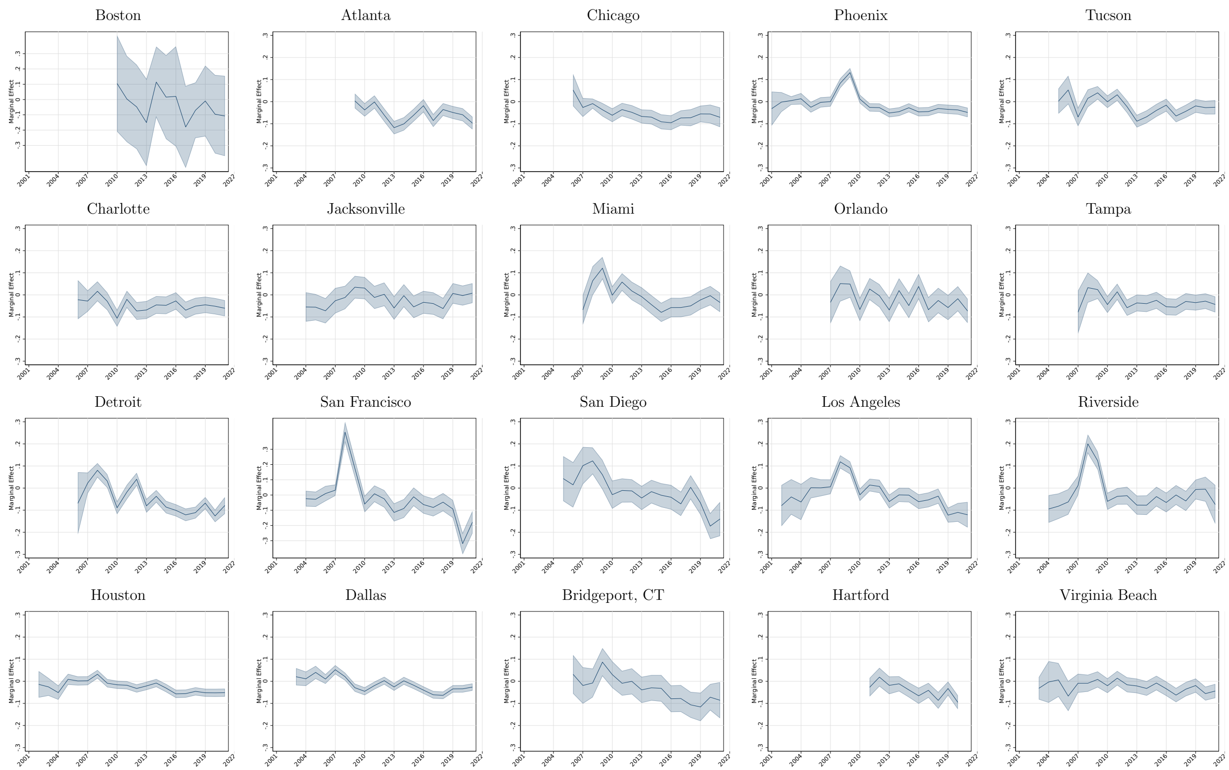

Figure 10: Relationship Between Total Returns to Location and Credit Scores Over Time. Values are slope coefficients from CBSA-level regressions of zip code-level total returns to location on the two year lagged zip code average Equifax Risk Score. *Source:* Authors' calculations using Corelogic MLS and the FRBNY/Equifax CCP.

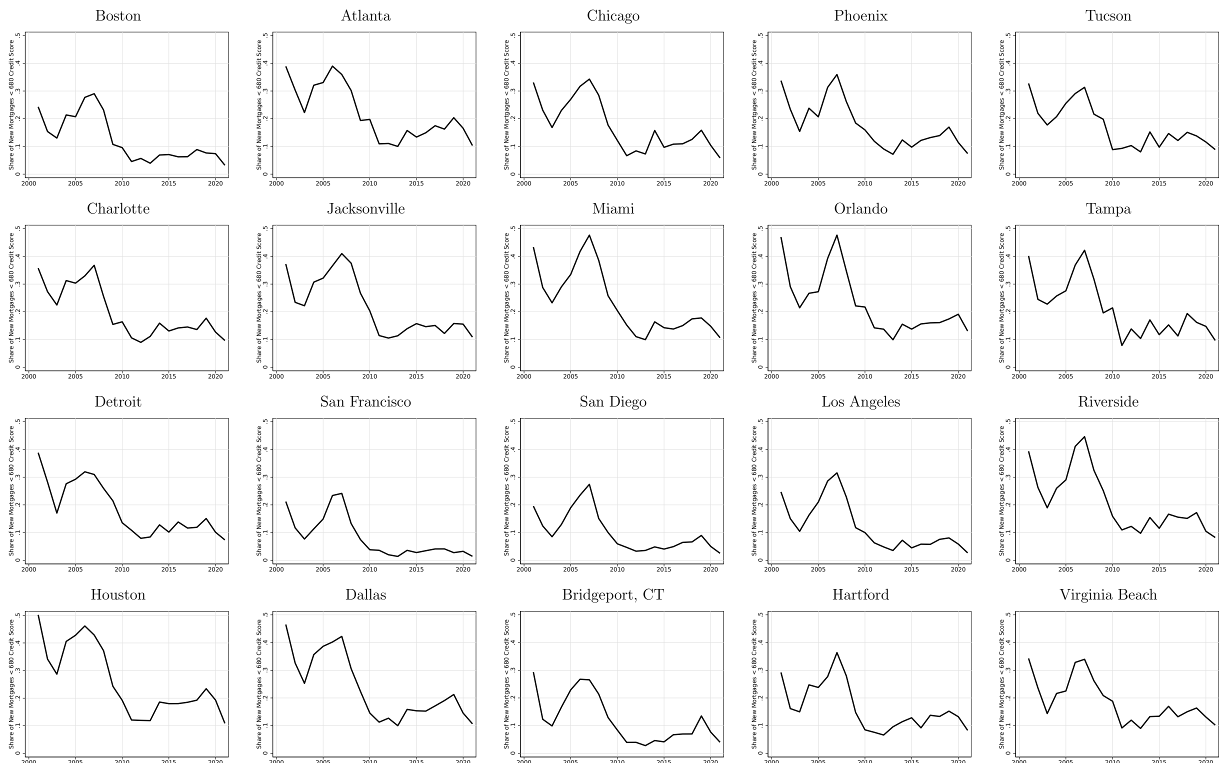

Figure 11: Share of Mortgages Originated to People with a 680 Credit Score or Lower. Values are share of number, not value. *Source:* Authors' calculations using the FRBNY/Equifax CCP.

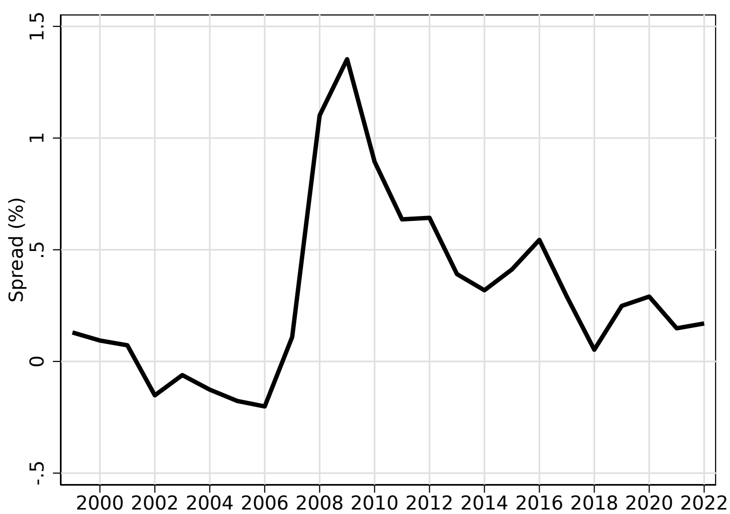

Figure 12: Jumbo-Conforming Spread. *Note:* Spread is calculated as the average annual 30-year fixed-rate jumbo rate according to BankRate.com minus the 30-year fixed-rate conforming rate according to Freddie Mac. *Source:* BankRate.com and Freddie Mac.

A Appendices

A.1 MLS Rental Listings

A potential issue with using information on rents from MLS is that properties listed for rent on MLS are higher quality the average rental unit in the US. This can be seen in Table A.1, in which we compare rents on properties listed in MLS between 1999 and 2019 to rents on housing units in the 1999-2019 waves of the American Housing Survey (AHS). We limit the AHS sample to market-rate units[23] in which the household moved since the previous survey and convert all prices to 2010 dollars. Rental listings in MLS are higher priced, on newer buildings, and for larger units than the average rental unit in the AHS. By comparison, sale transactions in MLS are representative, matching closely statistics for newly occupied owner-occupied units in the AHS. In the third column, we weight the MLS data to match the AHS size distribution, which results in moving our MLS sample somewhat closer to the AHS.

[23] We remove all rent-controlled and subsidized housing units.

	Renter Occupied			Owner-Occupied	
	AHS	MLS Unweighted	MLS Weighted	AHS	MLS Unweighted
Characteristics					
Rent or Price (2010 $)	919	1,875	1,795	225,648	261,043
Year Built	1967	1979	1974	1976	1976
Bedrooms (#)	2	3	2	3	3
Bathrooms (#)	2	2	2	3	2
Size					
Share < 500 Sq. Ft.	7	1	7	1	8
Share 500–750 Sq. Ft.	19	4	19	2	2
Share 750–1,000 Sq. Ft.	27	9	27	7	7
Share 1,000–1,500 Sq. Ft.	29	30	30	23	28
Share 1500+ Sq. Ft.	17	54	17	61	56

Table A.1: COMPARISON OF AHS AND MLS. *Note:* Values for the AHS are weighted averages from the 1999–2019 surveys and are limited to households that moved since the previous survey. Rental units from the AHS exclude all rent controlled and subsidized housing units. Values from MLS are from listings closed in 1999–2019. *Source:* AHS and MLS.

A.2 Supplemental Exhibits

This section contains additional figures and tables referenced in the main text.

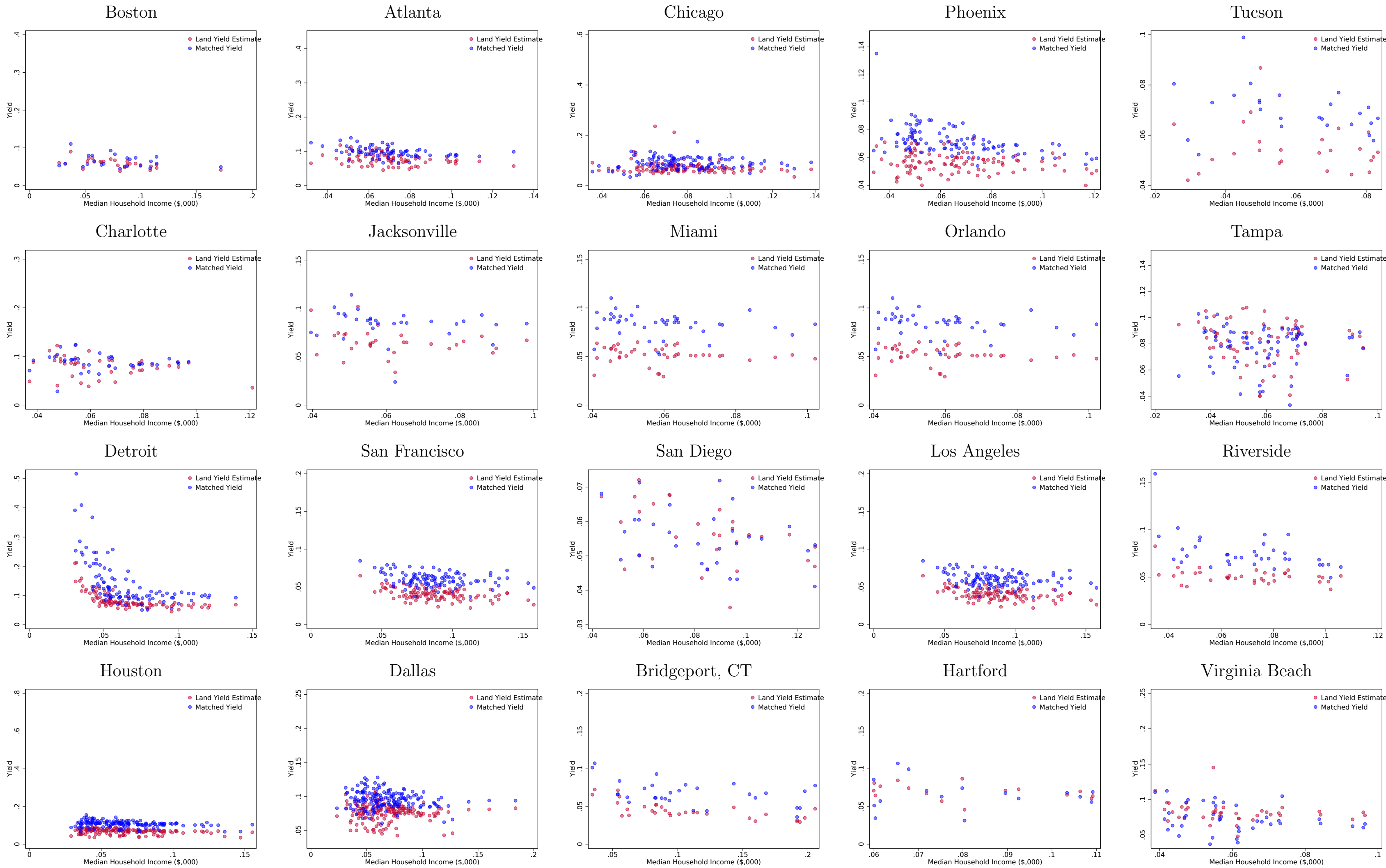

Figure A.1: Average matched yields and estimated land yields by zip code. *Note:* Average matched yields are averages of predicted values based on a single regression of property-level price-rent ratios on hedonics. The estimated land yields are estimated as described in Section 5. The x-axis is 2010 median household income from the decennial census. *Source:* Corelogic MLS data and the Decennial Census.

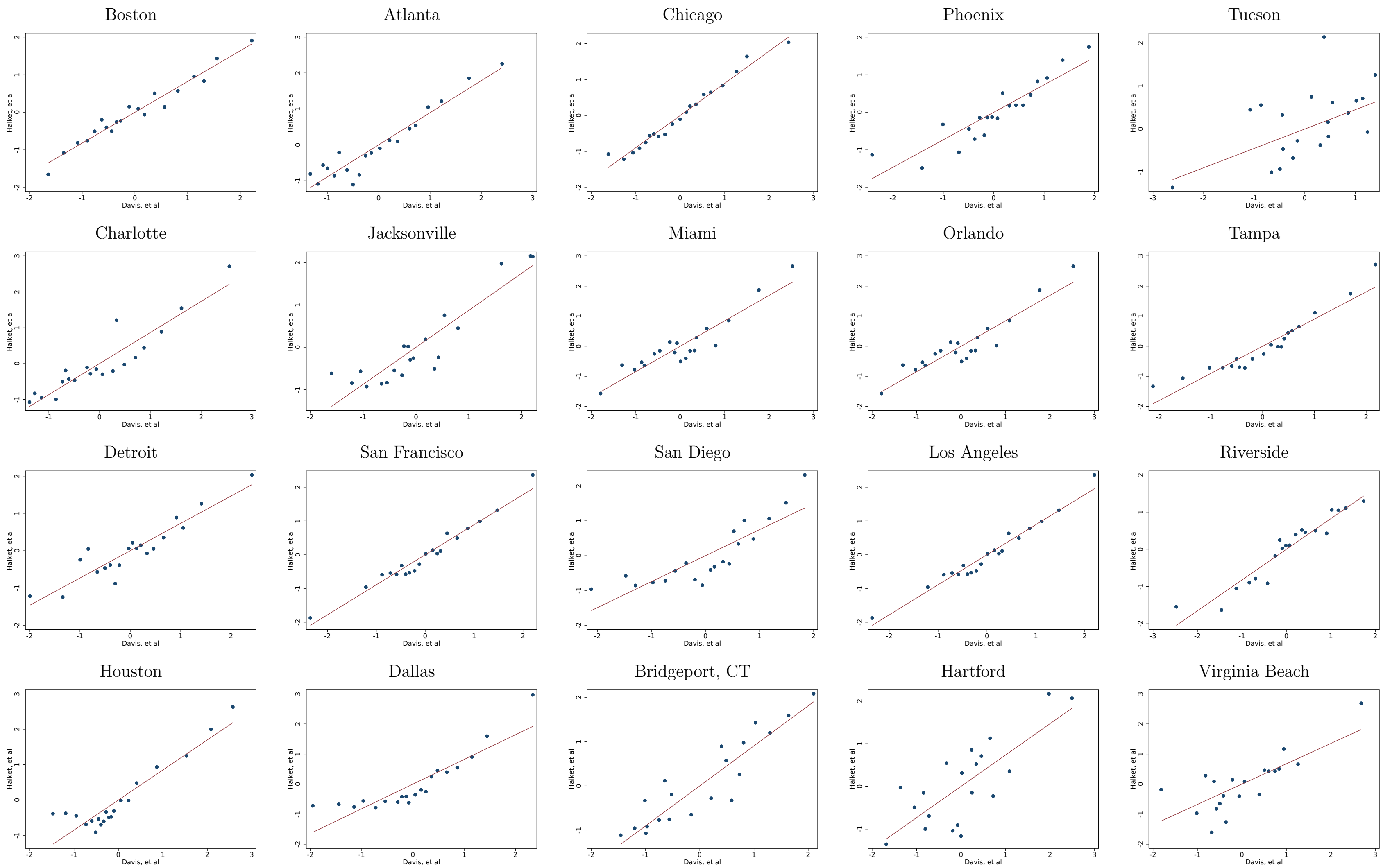

Figure A.2: Price of Land Per Sq. Ft., Comparison with Davis et al. (2021) *Note:* Our prices per square foot of land are estimated as described in Section 5. Both our estimates, and the estimates from Davis et al. (2021) are normalized to be mean zero and have a standard deviation of one. *Source:* Corelogic MLS data and Davis et al. (2021)

	Average					Std. Dev.				
	Yield Base House Val	Yield House	Yield Base Land Val	Yield Land	Jensen	Yield Base House Val	Yield House	Yield Base Land Val	Yield Land	Jensen
Atlanta, GA	–0.37	–0.40	–2.14	–2.23	0.03	0.00	0.09	0.00	0.22	0.04
Boston, MA-NH	0.34	0.30	–3.03	–3.17	0.02	0.00	0.23	0.00	0.17	0.00
Bridgeport, CT	0.02	0.12	–3.35	–3.17	0.04	0.00	0.09	0.00	0.13	0.01
Charlotte, NC-SC	–0.48	–0.67	–2.16	–2.12	0.05	0.00	0.25	0.00	0.29	0.04
Chicago, IL-IN-WI	0.18	0.18	–3.13	–2.89	0.06	0.00	0.08	0.00	0.22	0.02
Dallas, TX	–0.18	–0.31	–2.32	–2.41	0.00	0.00	0.07	0.00	0.15	0.01
Detroit, MI	–0.26	–0.35	–2.13	–2.27	0.04	0.00	0.11	0.00	0.14	0.01
Hartford, CT	0.26	–0.06	–2.91	–2.86	0.10	0.00	0.20	0.00	0.16	0.01
Houston, TX	–0.41	–0.38	–2.19	–2.31	0.02	0.00	0.03	0.00	0.12	0.02
Jacksonville, FL	–0.81	–0.92	–1.95	–2.03	0.08	0.00	0.36	0.00	0.24	0.04
Los Angeles, CA	0.04	0.01	–3.21	–3.25	–0.00	0.00	0.06	0.00	0.10	0.00
Miami, FL	–0.54	–0.59	–2.46	–2.20	0.06	0.00	0.12	0.00	0.19	0.03
Orlando, FL	–0.78	–0.66	–2.13	–2.24	0.04	0.00	0.14	0.00	0.19	0.06
Phoenix, AZ	–0.53	–0.61	–2.41	–2.35	–0.00	0.00	0.19	0.00	0.13	0.03
Riverside, CA	–0.29	–0.19	–2.72	–2.74	0.03	0.00	0.14	0.00	0.13	0.02
San Diego, CA	0.11	0.01	–2.89	–3.00	0.01	0.00	0.21	0.00	0.10	0.01
San Francisco, CA	–0.20	–0.00	–2.74	–3.24	0.01	0.00	0.15	0.00	0.16	0.01
St. Louis, MO-IL	–0.43	–0.30	–2.23	–2.37	0.06	0.00	0.19	0.00	0.20	0.02
Tampa, FL	–0.43	–0.64	–2.17	–2.17	0.05	0.00	0.37	0.00	0.29	0.04
Tucson, AZ	–0.38	–0.47	–2.58	–2.61	0.03	0.00	0.48	0.00	0.27	0.02
Virginia Beach, VA-NC	–0.32	–0.48	–2.26	–2.22	0.02	0.00	0.10	0.00	0.11	0.01

Table A.2: 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.

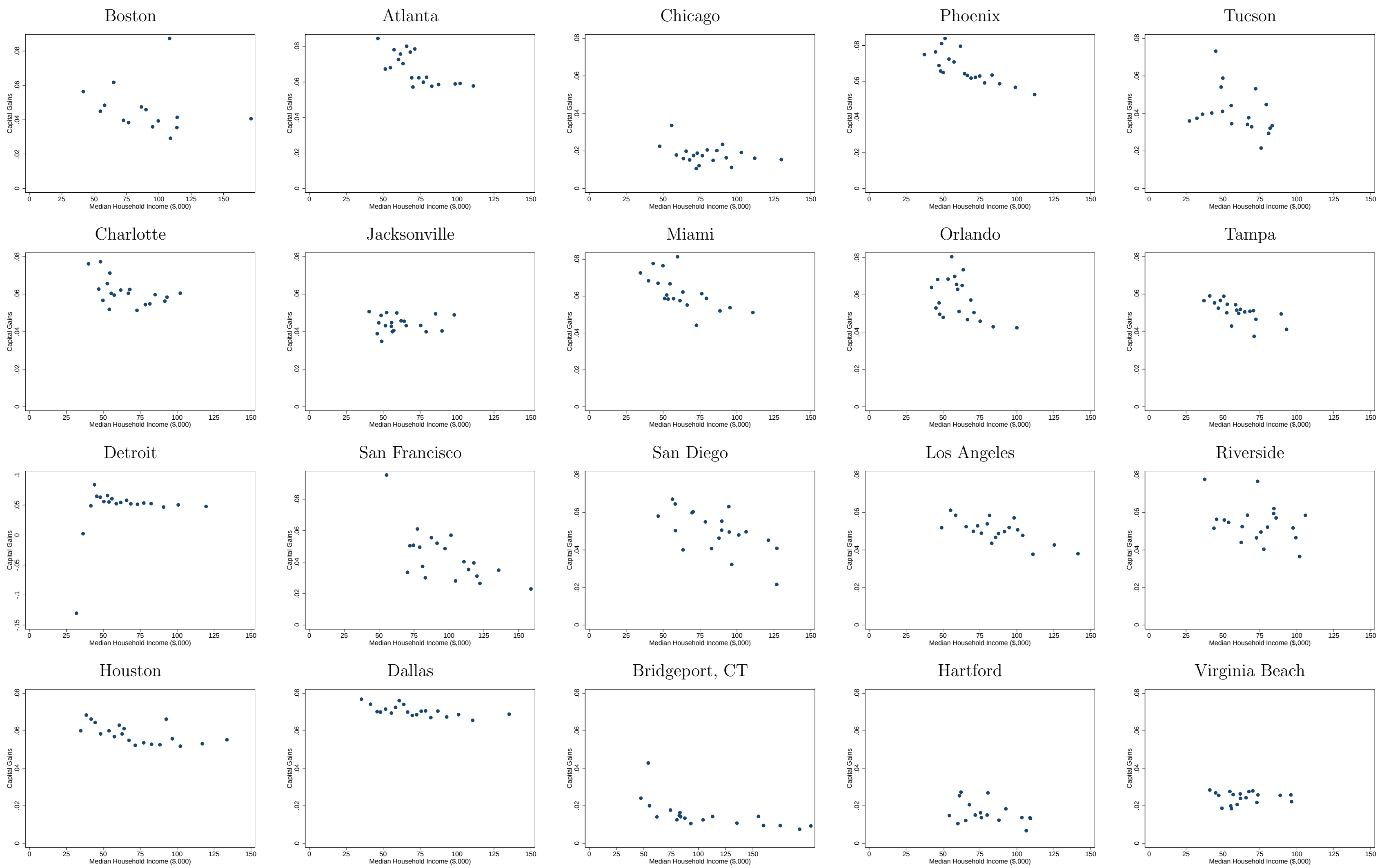

Figure A.3: Variation in Capital Gains by Location and Median Household Income. *Note:* Zip codes are weighted by the number of housholds in single-unit structures in 2011. *Source:* Corelogic MLS data and the American Community Survey.

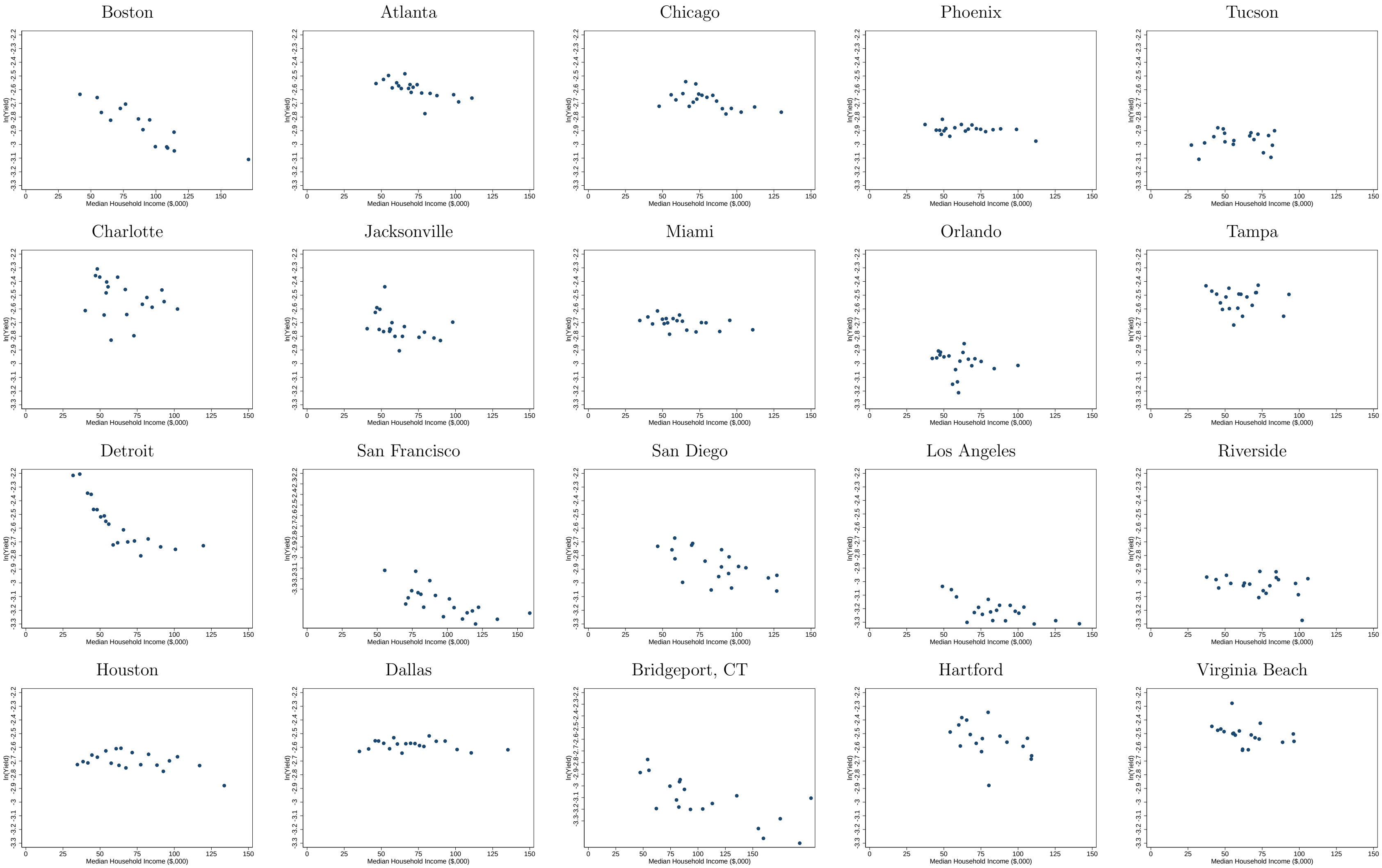

Figure A.4: Log Location Yields by Median Household Income. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011. *Source:* Authors' Calculations using the Corelogic MLS data and the American Community Survey.

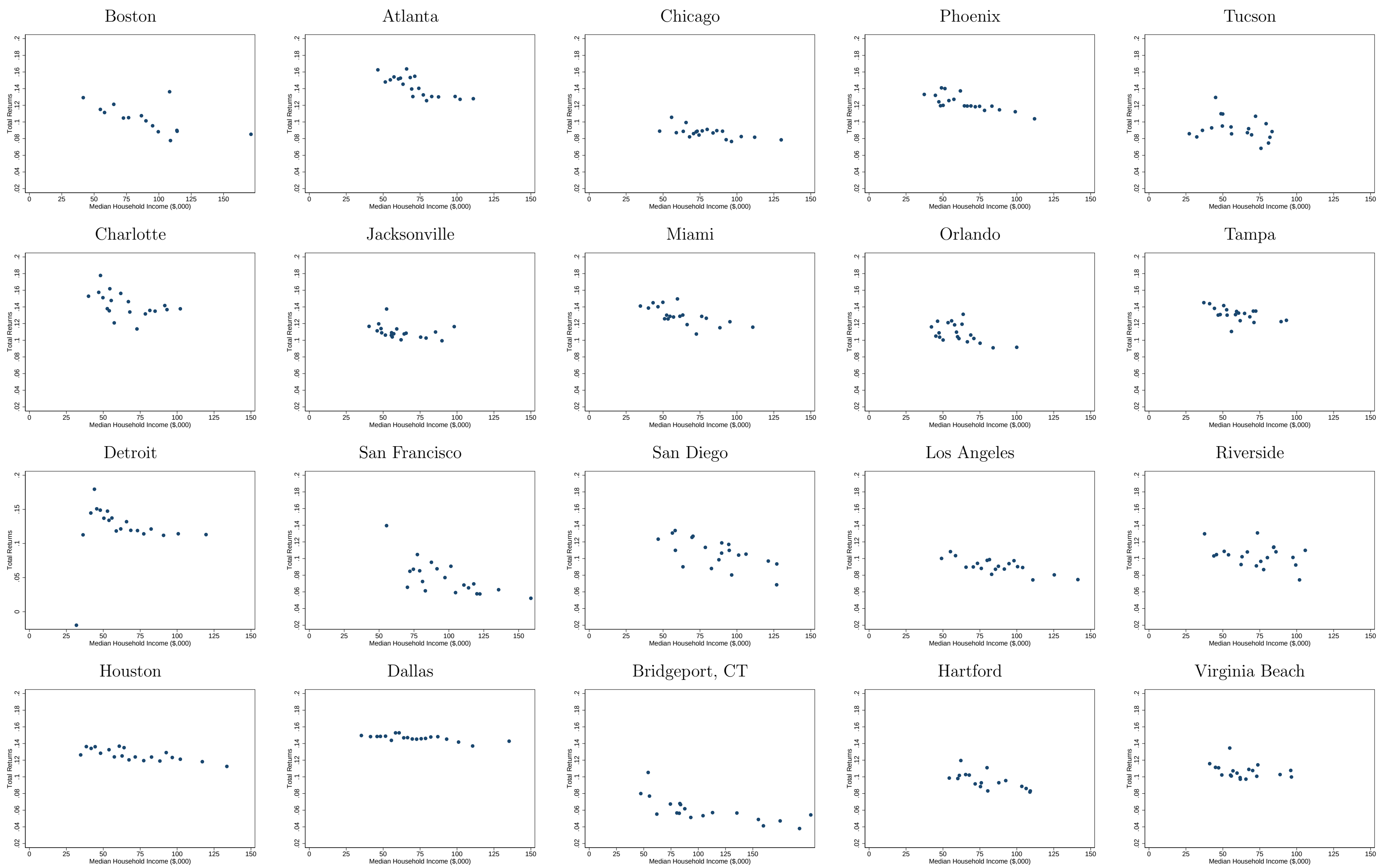

Figure A.5: Total Returns to Location by Median Household Income. *Note:* Zip codes are weighted by the number of households in single-unit structures in 2011. *Source:* Corelogic MLS data and the American Community Survey.

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

Table A.3: Summary Statistics for Log 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 Vacant (%)	Average Credit Score	Constant
Atlanta, GA	0.260***	-0.00241**	-0.00483	-0.00546***	-1.665**
	(0.0728)	(0.00109)	(0.00886)	(0.000639)	(0.668)
Boston, MA	-0.114	-0.0135***	-0.0424***	-0.00473***	1.985***
	(0.0816)	(0.00453)	(0.00953)	(0.000886)	(0.554)
Bridgeport, CT	-0.210	-0.00847	-0.0320	-0.00393	2.255
	(0.181)	(0.00832)	(0.0324)	(0.00390)	(1.504)
Charlotte, NC	0.316	-0.00205	-0.0592***	-0.00754***	-0.696
	(0.212)	(0.00309)	(0.0204)	(0.00230)	(1.407)
Chicago, IL	0.0400	-0.00496*	-0.0385***	-0.00434***	0.107
	(0.0647)	(0.00261)	(0.00872)	(0.000615)	(0.611)
Dallas, TX	0.206***	-0.000618	-0.0162***	-0.00425***	-1.955***
	(0.0382)	(0.000822)	(0.00529)	(0.000471)	(0.328)
Detroit, MI	0.00648	-0.000808	-0.00217	-0.00531***	1.095**
	(0.0773)	(0.00101)	(0.00481)	(0.000981)	(0.441)
Hartford, CT	0.280	-0.00176	-0.0389*	-0.00678*	-0.712
	(0.310)	(0.00486)	(0.0209)	(0.00350)	(1.467)
Houston, TX	0.384***	-0.00320***	-0.0103*	-0.00662***	-2.429***
	(0.0507)	(0.000934)	(0.00574)	(0.000600)	(0.370)
Jacksonville, FL	0.532***	0.00586*	0.0361***	-0.00436**	-5.842***
	(0.170)	(0.00307)	(0.0125)	(0.00191)	(1.392)
Los Angeles, CA	-0.0536	0.0124*	-0.0204	-0.00229***	-0.961*
	(0.0688)	(0.00702)	(0.0131)	(0.000789)	(0.540)
Miami, FL	-0.111**	0.00331***	-0.0116**	0.00112*	-2.234***
	(0.0470)	(0.00104)	(0.00537)	(0.000594)	(0.437)
Orlando, FL	0.129	0.00544	-0.0206***	-0.00178*	-3.193***
	(0.0978)	(0.00395)	(0.00754)	(0.000998)	(0.856)
Phoenix, AZ	-0.0908**	-0.000853	-0.0112***	-0.000826**	-1.266***
	(0.0370)	(0.00427)	(0.00273)	(0.000324)	(0.421)
Riverside, CA	-0.00236	-0.00451	0.00651	-0.00248***	-1.303*
	(0.0727)	(0.00600)	(0.00732)	(0.000784)	(0.708)
St. Louis, MO	-0.101	-0.0104	-0.0433	-0.00137	-0.477
	(0.267)	(0.0111)	(0.0332)	(0.00331)	(1.509)
San Diego, CA	-0.0806	-0.0000831	-0.0254	-0.00290***	0.160
	(0.0854)	(0.0100)	(0.0289)	(0.00110)	(0.836)
San Francisco, CA	-0.0695	0.00235	-0.0154	-0.00417***	0.413
	(0.0970)	(0.00575)	(0.0189)	(0.00131)	(0.794)
Tampa, FL	-0.272***	0.00402	-0.0392***	-0.000770	1.062
	(0.0710)	(0.00377)	(0.00514)	(0.000907)	(0.713)
Tucson, AZ	0.105	0.00660	-0.00552	-0.00171**	-2.900***
	(0.0939)	(0.00983)	(0.00633)	(0.000727)	(0.793)
Virginia Beach, VA	0.270*	0.00852***	0.0432*	0.000643	-6.207***
	(0.141)	(0.00180)	(0.0234)	(0.00131)	(1.220)

Table A.4: Multivariate determinants of Log Yields. *Note:* Coefficient estimates of multivariate regressions of log location yields on factors. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and the FRBNY/Equifax CCP.

	ln(Median Income)	Share Black (%)	Share Vacant (%)	Average Credit Score	Constant
Atlanta, GA	0.0134	0.000120	0.00229*	-0.000288***	0.108
	(0.00976)	(0.000146)	(0.00119)	(0.0000857)	(0.0895)
Boston, MA	-0.00596	0.00178*	0.000431	0.0000842	0.0437
	(0.0174)	(0.000964)	(0.00203)	(0.000189)	(0.118)
Bridgeport, CT	0.00979	0.0000130	0.000351	-0.000304*	0.124**
	(0.00748)	(0.000344)	(0.00134)	(0.000162)	(0.0622)
Charlotte, NC	-0.00330	0.000472***	0.000435	0.0000456	0.0572
	(0.00908)	(0.000133)	(0.000875)	(0.0000987)	(0.0604)
Chicago, IL	0.00243	-0.000351*	0.00134*	-0.000118**	0.0741
	(0.00508)	(0.000205)	(0.000684)	(0.0000483)	(0.0480)
Dallas, TX	0.00725***	0.0000817	0.00124***	-0.000139***	0.0814***
	(0.00272)	(0.0000584)	(0.000376)	(0.0000335)	(0.0233)
Detroit, MI	0.00204	0.000260	-0.0146***	-0.000533**	0.451***
	(0.0202)	(0.000265)	(0.00126)	(0.000257)	(0.116)
Hartford, CT	-0.0162	-0.0000114	0.000352	0.0000976	0.127
	(0.0289)	(0.000454)	(0.00195)	(0.000326)	(0.137)
Houston, TX	-0.00262	0.0000653	-0.000312	-0.0000777*	0.139***
	(0.00396)	(0.0000728)	(0.000448)	(0.0000468)	(0.0289)
Jacksonville, FL	0.00623	-0.0000123	0.00132*	-0.0000216	-0.0130
	(0.00976)	(0.000176)	(0.000718)	(0.000109)	(0.0797)
Los Angeles, CA	0.0154**	0.000346	-0.00316***	-0.000439***	0.194***
	(0.00604)	(0.000616)	(0.00115)	(0.0000693)	(0.0474)
Miami, FL	-0.0123**	0.000412***	0.000709	-0.0000666	0.235***
	(0.00628)	(0.000138)	(0.000717)	(0.0000794)	(0.0585)
Orlando, FL	-0.0205*	0.000942**	0.00259***	0.000101	0.198**
	(0.0111)	(0.000446)	(0.000853)	(0.000113)	(0.0968)
Phoenix, AZ	-0.0133***	-0.000554	-0.000112	-0.000238***	0.382***
	(0.00481)	(0.000554)	(0.000355)	(0.0000421)	(0.0546)
Riverside, CA	-0.00459	-0.0000695	0.000562	-0.000313***	0.316***
	(0.00882)	(0.000729)	(0.000889)	(0.0000952)	(0.0859)
St. Louis, MO	0.00278	0.000207	-0.000992	-0.0000766	0.0558
	(0.00605)	(0.000252)	(0.000753)	(0.0000748)	(0.0342)
San Diego, CA	-0.00158	0.00174	0.00314	-0.000147	0.166
	(0.0116)	(0.00137)	(0.00393)	(0.000150)	(0.114)
San Francisco, CA	-0.00335	0.000792	-0.00373**	-0.000517***	0.460***
	(0.00910)	(0.000540)	(0.00177)	(0.000123)	(0.0745)
Tampa, FL	-0.0185***	-0.000150	-0.000413	-0.00000364	0.260***
	(0.00512)	(0.000272)	(0.000371)	(0.0000655)	(0.0515)
Tucson, AZ	0.0282***	0.000436	0.000381	-0.000483***	0.0713
	(0.00848)	(0.000888)	(0.000572)	(0.0000657)	(0.0716)
Virginia Beach, VA	0.0128**	-0.0000349	0.000285	-0.000149***	-0.0140
	(0.00559)	(0.0000717)	(0.000931)	(0.0000520)	(0.0485)

Table A.5: MULTIVARIATE DETERMINANTS OF CAPITAL GAINS. *Note:* Coefficient estimates of multivariate regressions of location capital gains on factors. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and the FRBNY/Equifax CCP.

	ln(Median Income)	Share Black (%)	Share Vacant (%)	Average Credit Score	Constant
Atlanta, GA	0.0313***	-0.0000243	0.00195	-0.000670***	0.251***
	(0.0106)	(0.000158)	(0.00128)	(0.0000927)	(0.0968)
Boston, MA	-0.0103	0.00118	-0.00199	-0.000211	0.371***
	(0.0175)	(0.000969)	(0.00204)	(0.000190)	(0.119)
Bridgeport, CT	0.00415	-0.000369	-0.00118	-0.000528*	0.401***
	(0.0129)	(0.000596)	(0.00232)	(0.000279)	(0.108)
Charlotte, NC	0.0236	0.000357	-0.00291	-0.000517**	0.238*
	(0.0193)	(0.000282)	(0.00186)	(0.000209)	(0.128)
Chicago, IL	0.00569	-0.000673***	-0.00118	-0.000420***	0.331***
	(0.00575)	(0.000232)	(0.000775)	(0.0000547)	(0.0543)
Dallas, TX	0.0217***	0.0000676	0.000203	-0.000435***	0.199***
	(0.00388)	(0.0000835)	(0.000537)	(0.0000478)	(0.0333)
Detroit, MI	0.00404	0.000190	-0.0142***	-0.000955***	0.803***
	(0.0212)	(0.000277)	(0.00132)	(0.000269)	(0.121)
Hartford, CT	0.00426	-0.0000940	-0.00227	-0.000413	0.350***
	(0.0263)	(0.000413)	(0.00178)	(0.000297)	(0.125)
Houston, TX	0.0203***	-0.000128	-0.000935*	-0.000485***	0.231***
	(0.00452)	(0.0000831)	(0.000511)	(0.0000534)	(0.0330)
Jacksonville, FL	0.0378***	0.000415	0.00371***	-0.000269*	-0.141
	(0.0141)	(0.000254)	(0.00104)	(0.000158)	(0.115)
Los Angeles, CA	0.0125**	0.000848	-0.00390***	-0.000525***	0.329***
	(0.00619)	(0.000632)	(0.00118)	(0.0000710)	(0.0486)
Miami, FL	-0.0196**	0.000633***	-0.00000471	0.00000963	0.330***
	(0.00766)	(0.000169)	(0.000875)	(0.0000969)	(0.0713)
Orlando, FL	-0.0152	0.00122**	0.00168*	0.0000232	0.243**
	(0.0124)	(0.000498)	(0.000953)	(0.000126)	(0.108)
Phoenix, AZ	-0.0185***	-0.000567	-0.000719*	-0.000280***	0.527***
	(0.00522)	(0.000602)	(0.000385)	(0.0000457)	(0.0593)
Riverside, CA	-0.00469	-0.000311	0.000865	-0.000433***	0.448***
	(0.0101)	(0.000832)	(0.00102)	(0.000109)	(0.0982)
St. Louis, MO	-0.00359	-0.000346	-0.00362*	-0.000158	0.259***
	(0.0177)	(0.000737)	(0.00220)	(0.000219)	(0.0998)
San Diego, CA	-0.00656	0.00187	0.00200	-0.000302*	0.392***
	(0.0136)	(0.00160)	(0.00461)	(0.000175)	(0.133)
San Francisco, CA	-0.00575	0.000941	-0.00413**	-0.000652***	0.619***
	(0.00986)	(0.000584)	(0.00192)	(0.000133)	(0.0806)
Tampa, FL	-0.0377***	0.000111	-0.00323***	-0.0000663	0.600***
	(0.00754)	(0.000401)	(0.000547)	(0.0000965)	(0.0758)
Tucson, AZ	0.0334***	0.000760	0.000108	-0.000571***	0.128
	(0.0113)	(0.00118)	(0.000760)	(0.0000874)	(0.0952)
Virginia Beach, VA	0.0387**	0.000718***	0.00488*	-0.000105	-0.274*
	(0.0172)	(0.000220)	(0.00286)	(0.000160)	(0.149)

Table A.6: MULTIVARIATE DETERMINANTS OF TOTAL RETURNS. *Note:* Coefficient estimates of multivariate regressions of location total returns on factors. *Source:* Authors' calculations using Corelogic MLS data, FHFA house price indices, the decennial census, HUD vacancy rates, and the FRBNY/Equifax CCP.