Spatial Dispersion in Returns to Rental Housing: A Decomposition of Local Rent-Price Ratios

Peleg Samuels[‡]

December 2024

Abstract

What explains the significant dispersion in the rent-price ratio of rental housing units across and within US cities? Conventional urban models attribute this heterogeneity to differences in rent growth expectations, although it could also reflect a dispersion in returns earned by investors. In this paper, I construct local average ratios using administrative data on rental properties that reflect the investment decision of household real estate investors in a broad cross section of US cities and neighborhoods. A strong negative covariance between ratios and prices, along with a novel application of the Campbell-Shiller decomposition, can be used to reject the null that investors' rent growth expectations plausibly account for this dispersion. To quantify the relative importance of returns, I decompose past rent-price ratio dispersion into realized dispersion in rent growth and returns. I find that 80% of within-city and 90% of cross-city rent-price dispersion is explained by differential returns, implying large risk differentials, or barriers to arbitrage. These findings underscore the importance of integrating financial mechanisms in understanding the geography of housing markets and household finance.

I am deeply grateful to John Campbell, Edward Glaeser, Adam Guren and Sam Hanson for their invaluable advice, support, and insight before and throughout this project. I thank Adi Sunderam, Gabriel Kreindler, Winnie Van Dijk, Naomi Hausman, Denise DiPasquale, Luis Viceira and participants at Harvard's DUP and Finance graduate student seminars, as well as participants at seminars at Hebrew University, Tel Aviv University, University of Wisconsin Madison, and the Alrov Symposium for helpful comments and advice. Henry Weiland provided excellent research assistance. Any mistakes are my own. This work also benefited from the generous support of the Bradely Foundation.

‡ Tel Aviv University. Email: pelegsamuels@tauex.tau.ac.il

1 Introduction

In 2013, comparable apartments rented in San Jose, CA and Dallas, TX would earn their owners 4% and 12% of their prices annually in rent net of taxes, respectively. This gap reflects significant spatial heterogeneity in the ratio of rents to prices across US metro areas, paid to and by the approximately 11.7 million household real estate investors, who collectively own 50% of all US rental units. This heterogeneity has persisted over time and extends within cities and across housing segments, with spreads of up to 7 percentage points being common.[1] Identifying the source of this dispersion—whether reflecting differences in rent growth or returns[2]—is critical for understanding real estate pricing and valuation. Such dispersion has far-reaching implications for household finance, the supply of housing, diagnostics for identifying mispricing, and macroeconomic stability, particularly given the widespread ownership of rental assets by households, whose portfolios are less diversified and less understood compared to institutional investors.

This paper makes two contributions to the literature. First, it documents large dispersion in rent-price ratios, occasionally referred to below as "yields", across a rich set of US geographies, using unit-level rent and price estimates to address some key concerns raised in the literature (Eichholtz et al., 2021; Glaeser and Gyourko, 2007). Estimated yields exhibit distinct spatial patterns, suggesting local mechanisms that cause variation in either rent growth or returns. Second, I use the Campbell-Shiller decomposition (Campbell and Shiller, 1988) to show that this dispersion primarily reflects differences in returns, rather than growth rates, both ex ante and ex post. Contemporaneously high-price areas tend to experience slightly faster rent growth, but lower returns. In particular, I argue that as of 2013, the dispersion was too high to plausibly be explained solely by differential rent growth expectations. I then quantify the relative contribution of each channel in realized data, and return differences account for 80-90% of past cross-sectional differences in returns over the last 40 years.

While differential returns in a cross section of assets are generally expected, in the case of housing, this dispersion has significant implications for households, the housing market, and the broader economy. Spatial dispersion suggests that households face unequal financial opportunities and incentives to homeownership depending on their location, affecting wealth

[1] See Figure 1. These differences have existed since at least 1980, see Figure 10 in Amaral et al. (2022), and until at least 2018 (Appendix Figure 6). Similar dispersion has also been identified in other developed countries (Colonnello et al., 2023; Vanhapelto, 2022). For the importance of household real estate investors, see Samuels (2024).

[2] Returns are also referred to as 'discount rates,' 'capitalization rates,' or 'interest rates' in the real estate and urban economics literature.

accumulation and raising questions about whether the decision to own versus rent (Sinai and Souleles, 2005; Lockwood et al., 2023), varies geographically, and how policy affects homeownership rates (Greenwald and Guren, 2021).

Whereas rent growth reinforces concerns about land- and regulation-based supply constraints contributing to the affordability crisis in high-cost cities, dispersion in returns points to broader implications for capital allocation, previously underexplored in urban models. Low returns can reflect low risk (or low betas) in high-cost areas, as suggested by Demers and Eisfeldt (2022); Amaral et al. (2021), which could either encourage or discourage new investment, potentially exacerbating housing supply shortages. Alternatively, low returns can reflect barriers to arbitrage (Samuels, 2024; Levy, 2021), which lock in capital and drive return dispersion, potentially increasing housing supply in high-capital, low-yield areas. This dynamic, in turn, shapes the effectiveness of policies intended to create housing. Such barriers to arbitrage could also rationalize differences in the predictability of returns, preserving market efficiency even though yields only moderately predict rent growth, with implications for the use of rent-price ratios (and their derivatives) as diagnostic tools for mispricing (Himmelberg et al., 2005). Finally, because housing is a large component in household portfolios and debt obligations, these findings may impact monetary policy transmission and macro stability (e.g., Garriga et al. (2017)).

I begin with a careful construction of local rent-price ratios as averages of unit-level ratios, taking into account the heterogeneity of housing across space. Rent and price estimates for each unit are derived from hedonic regressions that leverage detailed information on location and physical characteristics, using DataQuick data, which provide near-comprehensive coverage of US real estate properties. This approach is consistent with recent studies that employ granular data and is critical for producing consistent estimates of local ratios (Colonnello et al., 2023) and returns in real estate, as recently discussed in the context of time series analysis by Eichholtz et al. (2021); Chambers et al. (2021).

I construct this dataset for each decade between 1980 and 2010, as well as for 2013. I also account for differences in real estate taxes and show that spatial dispersion is not driven by housing type, quality, or age of the housing stock, which might confound local average yields (Glaeser and Gyourko, 2007). I document significant dispersion both within and across cities for all types of housing, in line with recent research on other countries, (Amaral et al., 2021; Vanhapelto, 2022; Colonnello et al., 2023); across small samples of cities within the US (Fama and French, 2023; Campbell et al., 2009); and for single-family housing within them (Demers and Eisfeldt, 2022)) The mean rent-price ratio across metros is 12%, with a standard deviation of 5%. Within city dispersion is similarly high, with a mean of 11% and standard deviation of 6%, variation that is not captured by city fixed effects or other

unit-level controls.

Throughout this analysis, I focus on properties held by investors, which I identify using the mailing addresses listed on each property's deed. The baseline dataset includes 19.7 million units in 14.9 million properties, primarily owned by small, household investors who manage their properties from their own residence.[3] For homeowners, the focus of much of the literature (e.g. Diamond and Diamond (2024)), rent-price ratios reflect a mix of financial and consumption-based decisions (e.g., preferences for homeownership and location) considerations. In contrast, the rents earned and prices paid by investors identify differences in expected rental streams or their valuations. This obviates the need to impute renter-equivalent rents and focuses the analysis on financial fundamentals, rather than non-pecuniary preferences, which can both vary over space and imply average higher valuations (for any given rental stream), decoupling returns from subjective factors.[4]

To motivate a cross-sectional decomposition of yields, I document patterns in their spatial dispersion across and within US cities. Geographically, yields are low in metros on the West coast, Northeast, and Upper Midwest and high in metros in the South, Sunbelt, and Eastern Midwest, though within-region heterogeneity exists. Within cities, ratios fall with distance to the CBD, but are highest at a distance of 2-3 km from the center, not at the center itself. Ratios are not well explained by variables commonly used in the urban economics literature, as found by Colonnello et al. (2023) for Germany. A relationship to proxies of maintenance costs exists, but vanishes when controlling for building age. These patterns suggest local yields reflect underlying local economic differences, not measurement error or differences in the housing stock.

Building on these spatial patterns, I focus on a striking pattern: rent-price ratios systematically decline as prices rise, with rents failing to keep pace in high-price areas, implying a large negative covariance between rent-price ratios and prices.[5] Using a cross-sectional variation of the Campbell-Shiller decomposition (Campbell and Shiller, 1988) and under the null assumption of spatially arbitraged returns, I show that this covariance quantifies relative growth expectations as a function of contemporaneous prices. The results suggest implausibly high expected growth rates in expensive areas and implausibly low (even negative) rates in cheap areas. For example, in San Jose, where rent-price ratios are low and prices are exceptionally high, investors would need to expect annual average growth rates to exceed

[3] I also observe properties owned by larger commercial institution, but data coverage necessary for the hedonic regressions is lacking in some larger buildings, and so I also consider sub-samples inclusive of only household investors and find similar results.

[4] For example, among homeowners, rent-price dispersion might reflect segmentation between renters and homeowners due to differing local marginal buyers and renters, spatial or temporal preferences, or variations in preferences for homeownership as a consumption good.

[5] Demers and Eisfeldt (2022), focusing on single family rentals, document similar patterns.

the national rates by 10.8 percentage points indefinitely.

Next, I apply the empirical Campbell-Shiller decomposition, developed to identify the predictability of equity returns from dividend yields (Cohen et al., 2003; Cochrane, 2008), to the cross section of local rent-price ratios. I use high-quality data from DataQuick for each decade between 1980 and 2020, and an additional dataset with annual frequency and at the metro-level, which uses FHFA appreciation and BLS rent growth data, both benchmarked to rent-price levels computed using the DataQuick data. I regress realized rent growth, returns, and discounted future rent-price ratios on each period's returns, at different horizons. I control for time fixed effects, capturing macroeconomic factors and aggregate forcastability of returns.[6]

Both across and within cities, a higher rent-price ratio signals lower subsequent rent growth, higher returns, and higher discounted future rent-price ratios. These effects, by definition, sum to one, and the relative sizes of the coefficients signify the importance of each channel. Over time, a significant share of dispersion is realized as dispersion in growth rates and returns. Of these, returns consistently play a larger role: across cities, it is 90% of the realized dispersion, a result which is stable over all tested horizons between one and 15 years. Within cities, returns account for 82.6% of yield dispersion, a similarly high share.

This paper builds on a growing literature examining the cross section of rent-price ratios and housing returns. Several studies explore either the role of rent growth or returns in explaining yields. Fama and French (2023) show that yields predict rent growth more effectively in a small cross section of metro areas than in the time series but do not consider returns. Vanhapelto (2022) decomposes yield dispersion in Finland using a structural model and finds the implied residuals suggest a significant role for returns. Focusing on 'superstar' cities, Amaral et al. (2021) and Gyourko et al. (2013) document lower rent-price ratios. Amaral et al. also compute returns and attribute lower yields and returns to lower housing-income covariances, while Gyourko et al. emphasize supply constraints, linking lower yields to higher price appreciation.[7]

[6] A sizable literature has focused on the temporal evolution of rent-price ratios, as well as the predictability and characteristics of returns in real estate over time, generally finding that returns, both to single family housing and institutionally owned apartment buildings are somewhat predictable, see Campbell et al. (2009); Plazzi et al. (2010).Amaral et al. (2022) analyze the time series dynamics of the cross section of rent-price ratios, arguing low interest rates cause polarization in yields. Jordà et al. (2019) find surprisingly high returns to real estate and Sharpe-ratios over long-horizons. Subsequent papers (Eichholtz et al., 2021; Chambers et al., 2021), using unit-level data, have found lower returns, suggesting measurement error in aggregate data is significant. Loewenstein and Willen (2023) decompose price changes in the time series, using changes to rent-price ratios, and as in this paper employ a high-quality data on rents and prices for rental units.

[7] Recent studies on supply constraints also suggest that rent growth expectations explain rent-price dispersion (Büchler et al., 2021; Hilber and Mense, 2021; Molloy et al., 2022). My results indicate these mechanisms play a smaller role, as rent growth accounts for little of rent-price dispersion, and capital gains in constrained cities would counteract lower returns.

Relative to these studies, my paper explicitly estimates the relative importance of both rent growth and returns, considers a broader cross section of metro areas and within-city variation, and constructs rent-price ratios using rent and price estimates for the same units, following advances in the time series real estate literature (Eichholtz et al., 2021).

This paper also complements recent findings on housing returns. Diamond and Diamond (2024) document differential housing returns by racial group, and Goldsmith-Pinkham and Shue (2023) document higher returns for men. Demers and Eisfeldt (2022) construct rent-price ratios and returns across and within metro areas, focusing on single-family rentals. They find only slightly higher returns in low-yield cities, partially at odds with my finding. While this could be partly explained by highly persistent metro yields, this is unlikely to fully account for the discrepancy. Instead, differences in data and methodology may play a larger role: Demers and Eisfeldt use aggregate housing appreciation at the city level, whereas I compute capital appreciation at the rental property level, providing a more granular measure of returns. Within cities, their finding of higher returns in low-yield zip codes aligns with my results.[8] My analysis complements Demers and Eisfeldt by directly estimating the contribution of rent growth and returns to contemporaneous rent-price ratios directly estimating their relative importance.

The paper proceeds as follows: Section 2 discusses measurement of local rent-price ratios and the data. Section 3 presents spatial patterns in rent-price ratios across and within cities. Section 4 discusses excess price dispersion, and its implications for expected growth rates, and Section 5 conducts the ex-post Campbell Shiller decomposition.

2 Measurement

In this section, I describe how I estimate local average rent-price ratios and the details of the sample. The goal is to construct local measures that are informative on spatial differences in investors' expectations for growth or returns. I begin with a discussion of the difficulty in estimating these local ratios, and how I use administrative data to estimate hedonic unit-level price and rent data to alleviate some these concerns. I describe the sample construction process and the subset of privately owned rental properties it reflects. Finally, I provide the descriptive statistics of derived rent, price, and rent-price ratios.

[8]Within cities, where aggregate appreciation rates better reflect the local rental stock, the failure of higher appreciation to offset higher yields suggests that metro-level aggregate growth may overstate the capital growth component of returns. Differences in accounting for maintenance costs could also contribute to the discrepancy. Over the long run, for appreciation to offset yields, either the transversality condition must fail, or buyers must anticipate future purchasers with fundamentally different valuations, such as in increasingly segmented renter and owner markets.

2.1 Estimating Local Rent-Price Ratios

Local average rent-price ratios are difficult to estimate and to compare across space, for a few reasons. Chief among these are: *(1)* the average ratio is not the ratio of averages; *(2)* rents and prices are difficult to observe jointly; *(3)* housing is a highly differentiated good, and *(4)* prices paid by different buyers might have different implications (i.e. investors may maximize returns while owner occupiers may have nonpecuniary considerations).

To address these concerns, I use administrative data on the near universe of US real-estate properties to estimate unit-level rent and price data, compute unit-level rent-price ratios, adjust ratios for quality and real estate taxes, and limit my focus to investor owned properties. I discuss these challenges and my approach below.

1. The Average Ratio. Most local real estate price data is available as summary statistics, including as averages or appreciation series. Unfortunately, computing local rent-price ratios from such averages is problematic, as the average ratio is not the same as the ratio of averages. The bias introduced by ignoring this will generally lead to a downward bias in estimated average ratios. The bias in log rent-price ratios can be approximated as the covariance of units' *relative* deviation from the mean rent and their *relative* deviation from the mean price.[9]

This bias is potentially consequential, as the purpose of studying rent-price ratios is often directly linked to the way prices reflect, are informed by, or discount future rents; processes which all relate to the correlation between rents and prices, and not just their averages.

A consistent estimate of rent-price ratio must therefore start with an average of ratios, and not a ratio of averages.

2. Observing Rents & Prices. Computing the local average of ratios requires both rent and price data for the same units, at a high spatial resolution. Rent data in particular are rarely observed outside of Census administered surveys which deliver aggregate statistics.[10]

Rent microdata is observed in some datasets, and can be used to compute consistent rent-

[9]Formally, define a unit's price as P_i and rent as D_i, average prices as $\bar{P}$ and average rents as $\bar{R}$, and a unit's relative deviation from either as $\tilde{P}_i, \tilde{D}_i$. The average rent-price ratio is equal to

$$\frac{\overline{D}}{P} = \frac{1}{N}\sum_i \frac{D_i}{P_i} \approx \frac{\bar{D}}{\bar{P}} + \frac{1}{\bar{P}^2} E[(P_i - \bar{P})(D_i - \bar{D})] = \frac{\bar{D}}{\bar{P}}\left(1 - \text{Cov}(\tilde{P}_i, \tilde{D}_i)\right)$$

Where in the approximation equality I employ a first-order Taylor approximation. Such that the logarithmic average rent-price ratio is approximately $\overline{(d-p)} \approx \bar{d} - \bar{p} - \text{Cov}(\tilde{P}_i, \tilde{D}_i)$. (Small letters denote log versions of the corresponding capital letter.) Higher-order effects mitigate this bias, but are smaller by an order of magnitude.

[10]The underlying survey microdata is also not useful in this setting, as renters are not asked, and likely would not know, what an investor has or would pay for their unit – see discussion below.

price ratios (see, e.g. Greenwald and Guren (2021)). However, these data are often limited to large multifamily structures operated by large corporate landlords, which operate in a specialized and geographically concentrated segment of the market, or small cross sections, as in Eichholtz et al. (2021).

To construct a dataset with high spatial resolution and national coverage, I estimate unit level rents and prices on administrative data. These estimates rely on standard hedonic assumptions, and are based on detailed location and physical characteristics. To do this, I utilize a comprehensive dataset of residential real estate properties across the US, derived from DataQuick, estimate each unit's rent and price, and compute averages across units' ratios.[11]

3. Heterogeneous Housing & Real Estate Tax Rates. Housing is highly heterogeneous in its physical qualities, in ways that are well known to affect prices within and across cities. This does not necessarily pose a challenge for estimating local rent-price ratios, as both rents and prices reflect these different qualities. Interpreting spatial differences, however, depends on whether physical differences reflect segmentation of the marketplace. If housing is segmented along physical characteristics in ways that effect the object of interest (e.g. returns or growth rates), then the large spatial variation in the housing stock (age, size, ownership type) could drive or amplify spatial differences in rent-price ratios.

The data allow me to control for variables that address this concern: I residualize rent-price ratios on variables that reflect the physical properties of housing, including type of structure and age. I then compute local averages on these residualized ratios. Generally, these adjustments have a small impact on results.

Local real estate taxes, which vary significantly over space, could also drive differences in rent-price ratios. Real estate taxes, computed as a percentage of the contemporaneous price of properties, are costs that vary over space and so naturally drive differences in nominal rent-price ratios.

To control for this, I always account for these taxes by computing the average zip-specific tax ratio in the ACS, and deducting the implied tax from the annual rent payment. (This is isomorphic to deducting the tax rate from the log rent-price ratio.) I cannot control for various tax breaks, though these are generally not available to investment properties.

4. Buyer Types. Heterogeneous buyers, and particularly buyer types, may affect local rent-price ratios. One such key dimension is the difference between investors and consumers who buy for their own consumption.

[11] Another approach developed recently to estimate rents and price for the same units, utilizes rent data from the AHS and requires mapping either rents or prices between the rental and owner-occupied setting. See, e.g. Demers and Eisfeldt (2022), and Diamond and Diamond (2024).

Investors likely purchase units with the hope of maximizing returns, and subject to their information set, are likely less effected by non-pecuniary concerns. Owners have non-pecuniary considerations that are either idiosyncratic or otherwise deviate from the path of expected rents, which would have been paid by renters.

Investors may be willing to pay more for properties, if they are less risk averse, or less, if owner-occupiers' place a premium on ownership. (The consistently show owner-occupiers pay more than investors.) Observed sales are made to the highest bidder, implying that wherever an owner occupier purchases a housing unit, they outbid an investor, lowering the rent-price ratio that would have reflected an investor's expected return or growth rate.

The choice of sample over which ratios are computed is therefore an important one. Studying the home ownership margin is beyond the scope of this paper, and instead I focus only on investors: by constructing rent-price ratios only using properties that were bought and are held by investors.[12] In some analyses, I further exclude markets that are likely to have second homes, which I discuss further in Section 2.3.

2.2 Estimating Rent and Prices

The principal building block of local average rent-price ratios is unit-specific rents and prices. I estimate these using a series of hedonic regressions on administrative data, which include information on the location and physical characteristics of the property. Rent estimates aggregate to observed ACS zip averages by construction, and price estimates utilize county assessed value, to proxy for property-level unobserved relative quality.

Rents. I estimate rents using a hedonic in characteristics and a log-additively separable location premium. Each unit i's rent at time t is equal to the location premium adjusted for that unit's size, age, and building characteristics. The value of these physical characteristics is assumed constant across space. I allow the location premium and the valuation of physical characteristics to change with t.

Equation (1) details the specification: for a survey vintage conducted at time t I regress unit i's log rents R_{it} on the following 'textbook' hedonic characteristics (DiPasquale and Wheaton, 1996): number of rooms, bedrooms, and bathrooms, log square footage and a quadratic building age effect. (In some specifications I use decade fixed effects.) I also include fixed effects for the type of building (e.g. single-family house, 5-10 units, etc.) and, for each type, a linear number-of-units effect.

[12] This is the inverse of the process used by Diamond and Diamond (2024), who use transitions of rental properties into the owner-occupied segment to predict the value of the rental flow to owners.

$$\log(R_{it}) = \beta_t X_{it} + \mathcal{L}_{mt} \tag{1}$$

I assume that conditional on metro fixed effects $\mathcal{L}_{mt}$ and the physical characteristics X_{it}, the unobserved local premium is uncorrelated with the error in the regression. The fixed effect $\mathcal{L}_{mt}$ can then be thought of as a weighted average of more local (e.g. zip) location premium of the sampled units, which under the hedonic model's assumptions identifies β_t. This is an implicit assumption often made in hedonic regressions, and as applied here is weaker given the use of granular rent averages used to compute final rents.

I denote the estimated β_t vector as $\hat{\beta}_t^{\text{AHS}}$. Given its extensive coverage, the average rent across DataQuick units in each zip j should be equal to the average zip rent as reported by the ACS, as in Equation (2).

$$\bar{R}_{it}^{\text{ACS}} = \sum_{i \in j} \exp(\mathcal{L}_{jt} + \hat{\beta}_t^{\text{AHS}} X_{it}^{\text{DQ}}) \tag{2}$$

I invert this relationship to back out $\{\hat{\mathcal{L}}_{jt}\}_{jt}$, for each zip and period. It is then straightforward to reconstruct unit-level rents from the unit and building characteristics and the zip location premium for each property.

I estimate rents for five time periods - 1980, 1990, 2000, 2010, and 2013. In each period, I use either the census or 5-year ACS for average rents, and the closest preceding AHS sample. I do not use units constructed after each period when imputing the L_js, but I cannot account for demolished or significantly altered properties.

Prices. Estimating prices follows a simpler procedure and uses only DataQuick data. In addition to information on each property's location and characteristics, the data include sales information for transacted units. I run a *county level* hedonic regression for all units transacted in the two years preceding t, where t is each decade between 1980 and 2010, as well as 2013. Equation (3) details the specification.

$$\log(P_{it}^{\text{Transact}}) = \delta_c + \delta_{tc} + f_{ct}(P_{it}^{\text{Assessed}}) + \gamma_{ct} X_{it} + \nu_{it} \tag{3}$$

I include fixed effects for the year of sale (δ_{tc}), and the intercept δ_c serves as a county fixed effect. Data include unit characteristics (X_{it}, shared with the rent hedonic regression), as well as a third-degree polynomial in the assessed value of the unit, whose coefficients vary by county and over time, and denoted $f_{ct}(\cdot)$.

Running this regression separately for each county allows each county's assessment process to be arbitrarily inaccurate (some municipalities under-value average prices by as much

as 50% relative to transaction prices), but still uses the within-county variation as a signal of unobserved quality. Assessed price therefore serves as a local rank-ordering of unobserved quality, as in Diamond and Diamond (2024) (although I can only use this information to estimate prices). These models are then projected onto the stock of housing in each county at each time period.

Aggregation. I compute the local average rent-price ratio as the average rent-price ratio across all units within each locality. Since the hedonic rent estimates for each unit are the average ACS zip rent adjusted for the value of each unit's physical characteristics, this average ratio can be thought of as the ratio of the rent and price averages, adjusted for the correlation between unit-specific rent and prices induced by the observed physical characteristics.

If the unobserved unit-level rent and price residuals are uncorrelated, or if price residuals are small, the result is an unbiased estimate of local average rent-price ratios.[13] I discuss measurement error further below, but note on the high R^2 found for most units in the price regressions.

2.3 Data & Sample

The main dataset I use is DataQuick for 2013, which includes the near-universe of real estate properties across the US, and transaction prior to 2013. The data include the location of the property, number of units, age of structure. In most counties the data also include physical characteristics, such as number of rooms and square footage. Transactions include information on the sale price and sale date.

The data is at the property level, which is conducive to study the rent-price ratios faced by investors, as multifamily units are usually purchased as a whole building. However, for such properties room counts and square footage are reported for the entire property, and so for these I use the average (per-unit) counts/size. Crucially, the data include the owner's name (or the name of a holding company), and the owner's mailing address. I use this information to construct a sample of likely rental properties, that are privately owned. I filter out properties that are owner-occupied or owned by public institutions and banks, so as to focus on return-maximizing investors.

[13] The local average rent-price ratios is unbiased if:

$$\overline{D/P}_j = E_{i\in j}\left[\frac{\hat{D}_i \exp\epsilon_i}{\hat{P}_i \exp\nu_i}\right] = E\left[\frac{\hat{D}_i}{\hat{P}_i}\exp\varepsilon_i\right] \approx E\left[\frac{\hat{D}_i}{\hat{P}_i}\right] E\left[\exp\varepsilon_i\right] =^{*} E\left[\frac{\hat{D}_i}{\hat{P}_i}\right]$$

Where I denote estimated variables with hats. In the second equality, I define: $\varepsilon_i = \frac{\exp\epsilon_i}{\exp\nu_i}$, which is the ratio of the unobserved rents and prices. These residuals are derived from a hedonic with the same variables, and so ε_i is uncorrelated with the estimated rent-price ratio, implying the third equality. Note that by construction $E[\nu_i] = 0$, $E[\epsilon_i] = 0$. The last equality ($=^*$) holds under one of two conditions: $\epsilon_i \perp \nu_i$, or $\nu \approx 0$.

I construct a variable for corporate ownership based on the name of the owner. Corporate owners, particularly large towers with hundreds of units, sometimes lack size information on units. As a result, I also construct a sample that focuses only on household investors, which I discuss in more detail in Samuels (2024). This sample includes only investors whose mailing address is itself a residential address.

For each vintage, I exclude zip codes with fewer than 10 observed properties, or counties with fewer than 25 recent transactions. I also exclude block sales and those that appear to not be at-arms-length transactions. I filter out properties with outlier per-unit price values of more than 5 times the standard deviation across properties, and winsowrize rent-price ratios at the 1^{st} and 99^{th} percentiles within each metro area.

The baseline dataset includes 14.9 million properties, and 19.7 million units, representing about 50% of the rental national stock at the time. The sample covers urban and suburban areas, and housing segments owned by smaller investors - single family houses, condos, and smaller multifamily properties. These data include the vast majority of urban and suburban zips (10,730) and almost all metro areas (319).

The limitations imposed by the coverage of the final dataset imply that spatial differences in local rent-price ratios should be interpreted as those purchased by household investors in urban and suburban areas, and all segments of the market, except for large multifamily structures.

2.4 Summary Statistics

Table 1 provides summary statistics for the rents, prices and rent-price ratios estimated for 2013. Panel A details statistics at the unit level, Panel B at property level, Panel C summarizes the aggregates at the zip level, and Panel D at the metro level.

The average monthly rent for units in the sample is \$1,031. This is slightly higher than most reported rent statistics at the time, likely reflecting the exclusion of voucher and other subsidized housing, as well as the sample which weighjts more expensive urban and suburban areas. The average price of a unit in the sample is \$153,920, with a large standard deviation (\$160,120)[14]. At the property level, the average total rent in the sample is \$1,396, which reflects a right tail of larger buildings that generate high monthly cash flows. Similarly, the distribution of property prices reflects large differences in quality and in the number of units in each, with a high standard deviation.

The average rent-price ratio is 0.12, which implies that the average investor earned 12%

[14]The high standard error is likely driven by the combination of higher quality units that resemble the average price of an owner-occupied unit at the time (e.g. the 75^{th} percentile, \$192,830) and lower quality units and small and mid-sized multifamily structures (e.g. the 25^{th} percentile, \$58.9)

of their property's value every year, as of 2013, net of real estate taxes.[15] This is significantly higher than the ratio of the average rent and price, demonstrating the bias induced by the covariance of property-level rent and prices discussed above. The distribution of rent-price ratios is also condensed relative to rents or prices, but is still significant, with a standard deviation of 8pp.

The average rent-price ratio at the zip and metro level is similar to the overall average, 0.11 and 0.12 respectively. Despite the aggregation across units, the standard deviation at both the zip and metro level remains significant - an indication of the importance of geography. This variation is not explained by differences in the housing stock across space - when residualizing property level rent-price ratios by type, age and owner status the standard deviation and 25-75th percentile range remains similar. The residualized zip average also include metro fixed effects, meaning that zip variation is not driven by metro variation.

Table 1: Summary Statistics, 2013 Estimates

	Num Obs	Mean	Median	Sd	25th Pctile	75th Pctile
			Panel A: Units			
Rent ($; monthly)	13,192,572	1,031	946	448	729	1236
Price ($1,000)	13,192,572	154	109	160	59	193
			Panel B: Properties			
Rent ($; monthly)	13,192,572	1,396	1,002	8,712	759	1342
Price ($1,000)	13,192,572	194.98	116.97	1,276.18	64.03	209.48
Units per Property	13,192,572	1.37	1.00	5.65	1.00	1.00
D/P	13,192,572	0.12	0.10	0.08	0.06	0.15
			Panel C: Zip			
Avg D/P	10,730	0.11	0.10	0.06	0.07	0.14
Avg Resid' D/P	10,730	-0.01	-0.02	0.06	-0.05	0.02
			Panel D: Metro			
Avg D/P	319	0.12	0.12	0.05	0.08	0.15
Avg Resid' D/P	319	-0.00	-0.01	0.05	-0.03	0.02

Note: Summary statistics of rents, prices, and rent price ratios, as derived using the DataQuick data and hedonic regressions. Reisdualized rent price ratios are residualized on owner type (corporate yes/no), building type, and construction decade, zip residualized ratios also residualized on metro fixed effects.

[15]This figure is 2-3 pp higher than that reported for this period in Diamond and Diamond (2024) and Demers and Eisfeldt (2022) which exclude maintenance costs. Diamond and Diamond's sample is also of rental units transitioning to owner-occupied, which tend to have higher quality and prices, lower rent-price ratios.

3 Spatial Rent-Price Dispersion & Patterns

In this Section I motivate the analysis of cross-sectional dispersion in rent-price ratios by identifying distinct regional and geographic patterns. These patterns align with some of those for rents and prices individually, but are also distinct.

Regional Patterns Figure 1 maps city level yields across US metropolitan areas. The Figure presents clear spatial patterns in rent-price ratios, with a tendency for more expensive cities to have lower ratios, and less expensive cities to have higher ratios.

Ratios are lowest on the Pacific coast and at a number of cities in the Northeast (NYC has the lowest estimated ratio among large cities.) Yields are also low in the Mountain states such as Colorado and Utah. Ratios are high in the South, Sunbelt, and across the Plains. Large cities in Texas, Georgia, and Florida have ratios that are almost three times higher than comparably large cities on the Coasts.

Other regions have a mix of high and low ratio cities. The Upper Midwest, and particularly the states of Minnesota, Iowa, and Wisconsin, have low ratios, whereas ratios are high through the rest of the Rust Belt; Cities like Detroit and Cleveland have ratios double those of Minneapolis and Milwaukee. Similarly, California's Central Valley ratios are higher by 20-30% relative to its Coastal Cities.

High-ratio areas reflect a combination of growing Sun-Belt cities, with higher supply elasticities, and shrinking Rust-Belt cities, characterized by low demand. Low-ratio cities include both 'superstar', large, cities with diverse workforces, and small mountain or coastal towns.

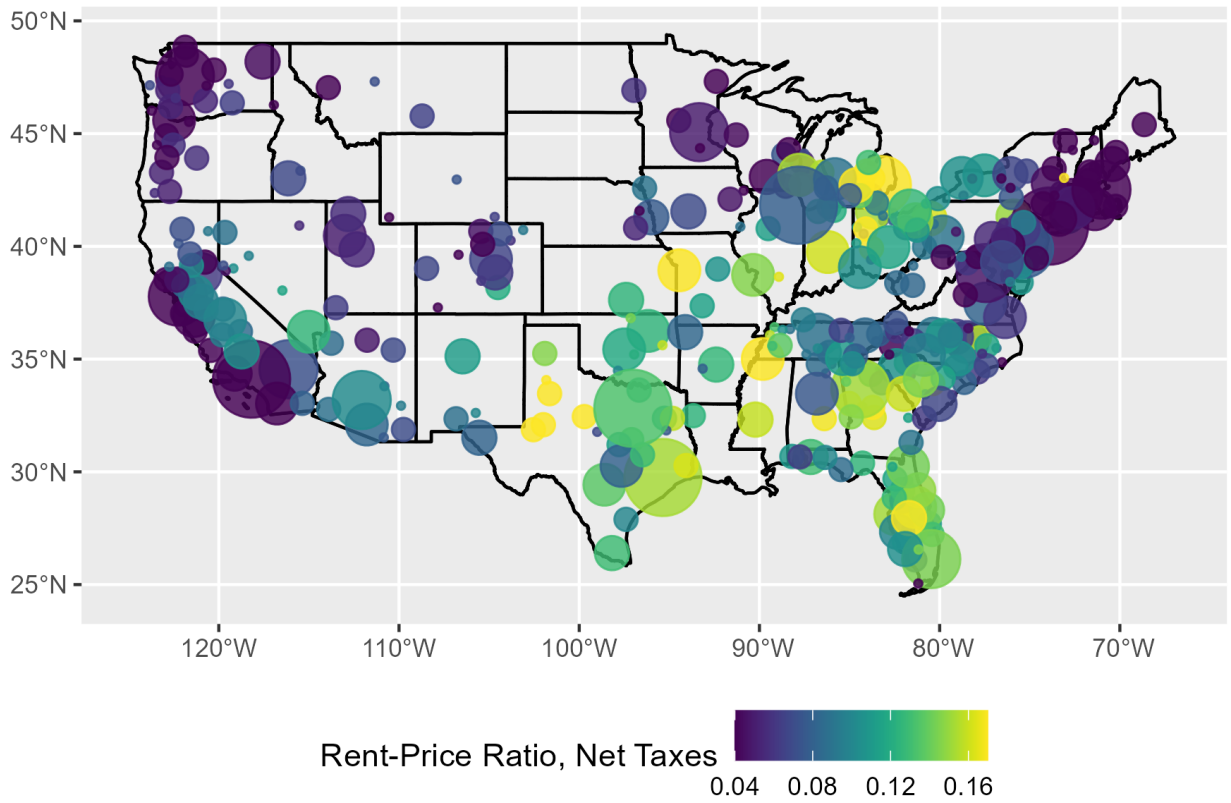
Within City Patterns. Rent-price ratios tend to be higher near city centers, and decline with distance from the CBD. Figure 2 presents a bin-scatter and non-parametric kernel mean estimate of the rent-price ratio-distance curve.[16] I use rent-price ratios residualized on metro fixed effects, as well as building characteristics including type and age. Figure **??** in the Appendix presents a similar plot, without residualization, showing the same pattern.

Ratios appear to be the highest in neighborhoods located 1 to 3 miles from city centers, where they are about 1pp higher than average - higher than at the city centers. Neighborhoods in this area tend to be cheaper – they do not have the urban amenities and low commute costs of city centers, nor the open spaces and greenery of suburbs, and tend to serve working class and racially diverse residents.

Ratios decline precipitously with distance - they are 2pp lower than average 20 miles

[16]Rent-price ratios are at the zip code level, and I compute distances using each zip code's centroid to the CBD as determined by Hausman et al. (2023).

Figure 1: Map of Metro-area Average Rent-Price Ratios



2013 Metro area average rent-price ratios, net of real estate taxes and residualized on unit characteristics including building decade of construction, building type (SFH/condominium/Duplex, etc), and size category (1 unit, 2-4, etc.). Rent and prices are estimated for each unit observed in DataQuick based on location and a rich set of physical characteristics for units not occupied by their owners. Dot size corresponds to city size. The average ratio across cities is 12.3% and the standard deviation is 6.4%.

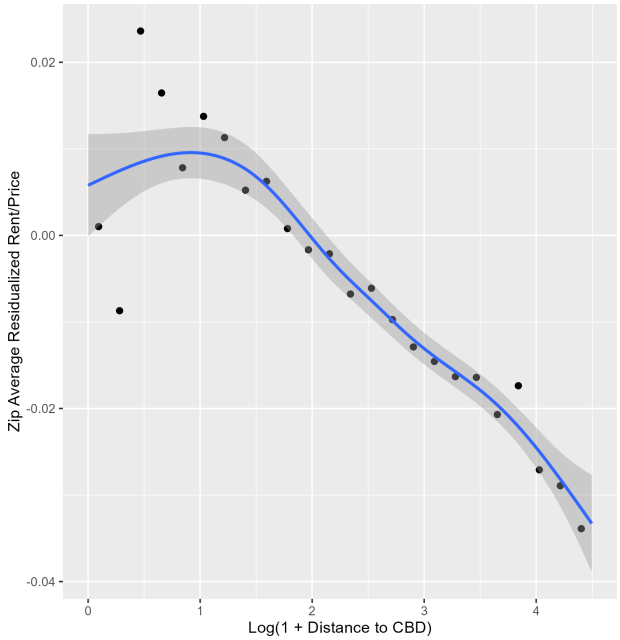
from the city center, and 3.5pp lower at the 55 mile mark. This pattern suggests suburbs tend to have lower rent-price ratios, with exurbs displaying the lowest ratios.

Comparison of Correlations with Local Variables Table 2 presents regressions of rent, prices and rent-price ratios for metros and zip codes. The regressions include variables that are known to be important determinants of rent, which are all standardized as z-scores to compare their magnitudes. Overall, rent-price ratios follow similar, though distinct, patterns to rents and prices individually, and are not as well explained.

Panel A presents regressions at the metro level; variables include population, income, the share of college educated residents, building regulation (WRULRI index), as well as the average wage of HVAC/Plumbers.[17] Consistent with the urban literature, these variables

[17]Inclusion of other tradesmen wages does not add explanatory power – HVAC/Plumbers seems to capture the full effect.

Figure 2: Within-City Local Rent-price Ratios Over Distance From CBD



Bin-scatter plot of local average rent-price ratios at the zip level, by log distance from the CBD. Ratios are residualized by metro fixed effects, building type (SFH, MFH size category) and age. Confidence intervals are of the non-parameteric average, not of the data.

explain a large share of the variation - 51% and 45.5% across cities, but the R^2 of ratios is only 15.6%. Ratios follow the same patterns for income as rents and prices, but seem unaffected by building regulation. In an encouraging sign, tradesmen wages are positively correlated with ratios, while they only slightly affect rents and prices.[18]

Panel B presents regressions within cities; I similarly include income and education, but replace population and building regulation with density and distance to CBD. Ratios are effected less by income and education, but go in the same direction as prices - suggesting the price effect dominates. Interestingly, ratios are much more sensitive to density and distance to CBD than rents or prices independently.

4 Excess Price Dispersion and Growth Expectations

The previous Section established that local rent-price ratios vary across space in distinct patterns, that hint at locally operating mechanisms. In this section, I focus on a particularly strong pattern: ratio dispersion are primarily driven by large dispersion in prices – rents

[18]The tradesmen effect disappears when using residualized rent-price ratios, suggesting this correlation is driven by differences in the housing stocks' type and age.

Table 2: Correlations Table

	Panel A: Metro			Panel B: Zip		
	Log(Rent) (1)	log(Price) (2)	log(Rent/Price) (3)	Log(Rent) (4)	log(Price) (5)	log(Rent/Price) (6)
log(Pop) (z-score)	0.029***	-0.033	0.068*			
	(0.008)	(0.033)	(0.036)			
log(Income) (z-score)	0.168***	0.335***	-0.223***	0.118***	0.162***	-0.125***
	(0.021)	(0.037)	(0.040)	(0.007)	(0.008)	(0.013)
Share BA+ (z-score)	0.014	0.041	-0.028	0.037***	0.101***	-0.066***
	(0.016)	(0.030)	(0.036)	(0.006)	(0.012)	(0.013)
WRULRI Index (z-score)	0.041***	0.053***	-0.015			
	(0.012)	(0.020)	(0.024)			
HVAC/Plumbing Wage (z-score)	-0.018	-0.057**	0.113***			
	(0.015)	(0.027)	(0.033)			
log(Density)				0.013*	-0.006	0.053***
				(0.007)	(0.004)	(0.009)
log(Dist' CBD) (z-score)				-0.051***	-0.003	-0.090***
				(0.005)	(0.008)	(0.011)
Unit Characteristics	✓	✓	✓	✓	✓	✓
Metro FE	-	-	-	✓	✓	✓
Fit statistics						
Observations	319	319	319	10,710	10,710	10,710
R^2	0.511	0.454	0.156	0.737	0.873	0.678
Within R^2				0.352	0.497	0.293

Note: This table presents regressions of rents, prices, and rent price ratios on variables that are know to drive the former two. Panel A is at the metro level, and panel B at the zip level. The table compares the correlation of mechanisms know to effects rents and prices, to their effect on rent price ratios, finding similar, yet distinct patterns, that do not explain as much of the variation. Population, density, income, and education are measured in the ACS; land use regulation is based on the WRULRI Index, of Gyourko et al. (2021); HVAC/Plumbers wages are from the County Business Patterns data. Rent, Prices and Rent Price ratios are derived from hedonic regressions using DataQuick, ACS and AHS data, as described in Section 2.2. The regressors are normalized as z-scores, to make correlations comparable across variables. Standard errors in Panel B are clustered at the metro area. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

'dont keep up'. This pattern was first documented by Demers and Eisfeldt (2022), for single family rentals, below I expand this to a larger cross-section of additional housing segments, metro and zip areas, and analyze the potential for it being driven by measurement error.

I then I quantify the magnitude of this rent-price-price relationship, and argue that it is informative on the magnitude of price-rent growth expectation relationship, under the null that returns are arbitraged across space. The implied growth expectations in high (low) price areas are implausibly high (low), suggesting that expected returns vary significantly over space.

Intuitively, prices in high-price areas are so high relative to local rents and average rent-price ratios that, for returns to equate over space, investors in these areas would need to have impossibly high growth expectations that are counterfactual to observed patterns.The magnitude of the ratio-price covariance can be used, under the null, to quantify investors'

expected growth rates in high- and low-price areas, which appear implausible and counterfactual.

4.1 Excess Price Dispersion

Figure 3 visualizes the extent to which price variation drives rent-price dispersion in the data. Rent-price ratios are consistently lower where prices are higher, suggesting rents 'don't keep up' with prices in the cross section. This pattern holds across the price distribution, with a non-parametric kernel estimate suggesting average log ratios falling almost linearly in log prices. It also holds with city-fixed effects and controls for building characteristics.

To emphasize the magnitude of the negative covariance, Figure 3 also plots two benchmark scenarios: (*1*) rent and prices co-move across space - as would be the case if prices are the NPV of rents, discounted at the same rate (dotted line); (*2*) rents are statistically independent of prices (dashed line). The data are far closer to the second scenario than the, often assumed and far more intuitive, first scenario.

These patterns are also evident at the local level, controlling for differences in building characteristics. Column 1 and 2 of Table 3 present regressions of prices on rent-price ratios at the local level: at the metro level in Panel A, and at the zip level in Panel B. I control for building construction decade, type (single family, condo, type of multifamily), as well as metro fixed effects for regerssions at the zip level.

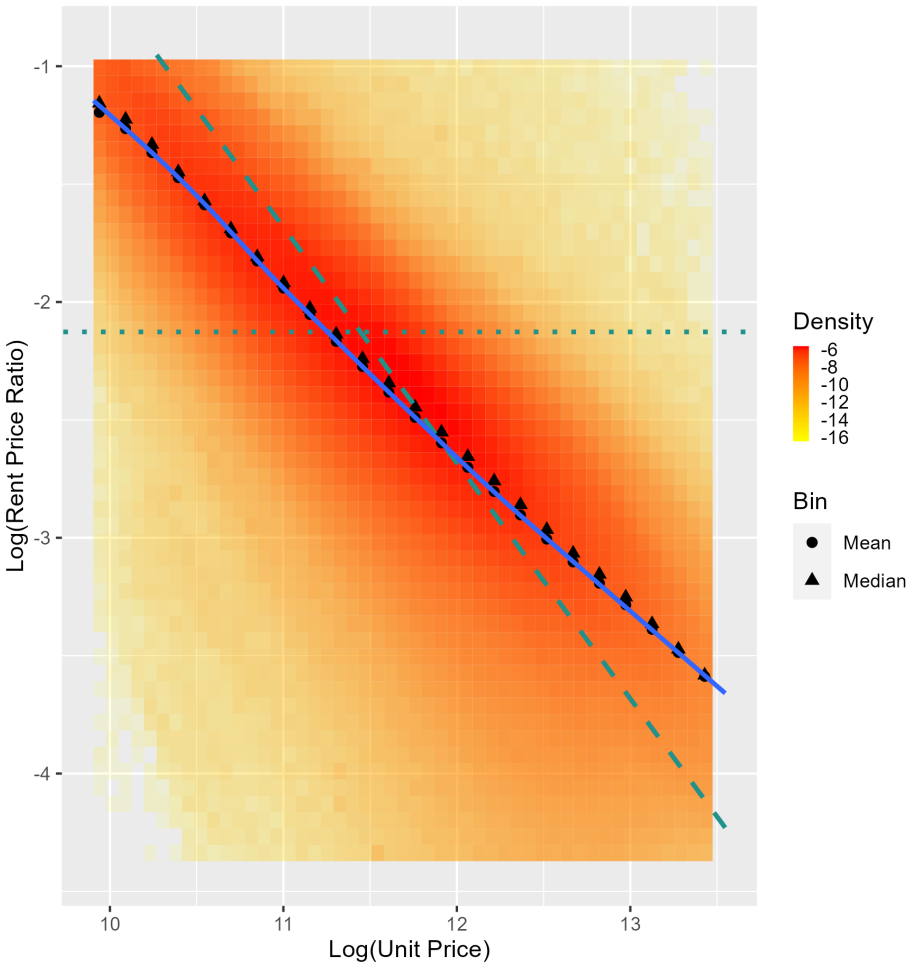
The results suggest a ratio elasticity of -0.82 with respect to prices across cities, and -0.76 across zips within cities. This implies that rents only offset 0.18-0.24% of every 1% difference in prices across space. This relationship also explains a large portion of spatial dispersion in ratios, with an R^2 of 0.597 and 0.565.

This pattern could be driven by three primary sources: (*1*) measurement error, (*2*) dispersion in expected growth rates, and (*3*) dispersion in expected returns. I argue that measurement error is unlikely to be a key driver, and then turn to analyzing the plausibility of growth rates as a driver of rent-price dispersion, as opposed to returns.

Bounding Measurement Error. The negative covariance between prices and rent-price ratios can be spurious due to measurement error, particularly if the error in prices dominates and there is no correlation between the errors in rents and prices. While I argue that this is unlikely to be the case, I also run a series of regressions with alternative measures for prices as an IV, which estimate a consistent yield-price covariance.

To fix ideas, consider a log-liner model of measurement error, as implied by the hedonic

Figure 3: Excess Cross-sectional Price Dispersion



This figure visualizes the relationship between log rent-price ratios and log prices. Color implies distribution density. Dots show the bin-scatter mean and median, and the blue solid line displays a kernel-smoothed mean. The dashed and dotted lines each plot a theoretical benchmark; in the former rents are uncorrelated with prices, and in the second they move in tandem. Plot area is truncated at the 1^{st} and 99^{th} percentiles of the x and y axes.

price model of Section 3, consider a log-additive error in both rents d and prices p:

$$\tilde{p} = \log(P \cdot e^{\epsilon})$$
$$\tilde{d} = \log(D \cdot e^{\nu})$$

I assume, as in the hedonic regressions, that the errors are independent of the true prices and rents. The observed covariance between $\tilde{d} - \tilde{p}$ (log rent-price ratios) and $\tilde{p}$ is:

$$\text{Cov}(\tilde{d} - \tilde{p}, \tilde{p}) = \text{Cov}(d - p, p) - \text{Var}(\epsilon) + \text{Cov}(\nu, p) + \text{Cov}(\nu, \epsilon) \tag{4}$$

The existence of a measurement error in prices, ϵ, biases the estimated covariance downward, and the bias depends on the variance of the error. Measurement error in rents only impacts this covariance if it is correlated with either (true) prices or the error in prices. If measurement error is driven by unobserved quality, these covariances are likely to be positive, attenuating the effect of measurement error in prices.

We can put an upper bound on the bias induced through measurement error in prices by estimating the variance of the residuals in the hedonic price regressions. Assuming that the hedonic model holds, we can then estimate $\text{Var}(\epsilon)_c$ for each county c, using the distribution of errors in each county's hedonic regression, $\sum \hat{\epsilon}^2_{i \in c} = \widehat{\text{Var}(\epsilon)}_c$. To compute a 'national' error variance I average over counties, weighting by the number of transactions in each. I similarly compute standard errors, and use the 10^{th} and 90^{th} percentiles of county variance estimates to compute a 10% confidence interval. I similarly compute a weighted standard deviation of county variances.

The 'national' error variance is 0.19, its standard deviation is 0.09, and the 10% confidence interval is $(0.1, 0.34)$. These variance estimates are significantly smaller than the observed yield-price covariance ($\text{Cov}(\tilde{d} - \tilde{p}, \tilde{p})$), which is -0.52. If transacted and non-transacted properties are similar, then the mean upper bound of this covariance would be -0.33 at the mean variance, and -0.18 at the upper limit of the confidence interval.[19] These are upper bounds on the covariance, since rent and price errors are likely to be correlated, attenuating the effect of error in prices.

Instrumenting for Prices. To estimate the magnitude of the yield-price covariance, without the bias induced by measurement error, I run a series of regressions at the zip-code

[19] If measurement error among non-transacted properties is greater than that among transacted properties, one could instead calculate how much higher this error would need to be to fully explain the negative observed covariance. In this case, the measurement error for nontransacted properties would have to exceed that observed among transacted properties in almost every county. Specifically, it would need to be at least 3.6 standard deviations higher than the cross-county average for transacted properties. This, too, is a conservative estimate.

and metro level with alternative measures for prices as instruments for baseline prices. The results indicate that measurement error contributes little to the magnitude of the negative covariance.

Table 3 displays the results of these regressions. Panel A displays the regressions at the metro level and panel B at the zip level. Columns 1 and 2 present the OLS specifications discussed above, and Columns 3 through 5 present the IV regressions. In particular, Column 5 uses the price of owner-occupied units, measured by the ACS; Column 6 uses the average observed sale price (only based on transacted properties); and Column 7 uses the average log square footage of local units as a shifter of prices that by construction is exogenous to the hedonic measurement error.

Table 3: Rent Price Ratio - Price Correlation

	OLS		IV		
Instrument			log(Owner P)	log(Sales P)	log(Sqft)
	Panel A: Metro level; dep' Variable: resid' log(Rent/Price)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
log(Price)	-0.819***		-0.695***	-0.761***	-1.05***
	(0.038)		(0.043)	(0.049)	(0.226)
	-	-	[-0.778,-0.608]	[-0.856,-0.664]	[-1.774,-0.611]
log(Price) (z-score)		-0.402***			
		(0.018)			
Statistics					
1st Stage F	-	-	66.36	24.34	0.58
Within R^2	0.597	0.597			
Observations	319	319	319	301	319
	Panel B: Zip level; dep' Variable: resid' log(Rent/Price)				
log(Price)	-0.762***		-0.717***	-0.691***	-0.771***
	(0.006)		(0.007)	(0.008)	(0.023)
	-	-	[-0.732,-0.703]	[-0.706,-0.675]	[-0.817,-0.725]
log(Price) (z-score)		-0.451***			
		(0.004)			
Statistics					
1st Stage F	-	-	330.06	189.75	8.5
Within R^2	0.565	0.565			
Observations	10,730	10,730	10,730	10,392	10,730

Note: This table presents regressions of local (metro and zip) average rent price ratios on local prices. Variables are residualized by building type, decade of construction and corporate ownership. Zip level regressions are also residualized by metro area. Column 4 instruments for prices with average unit level square footage, Column 5 instruments with owner-occupied prices in the same areas, from the ACS, and Column 6 instruments with sale prices in the same zip, for properties transacted between 2011 and 2013. Rents and Prices are estimated as a hedonic based on unit characteristics, location averages, and in the case of prices also assessed values. Standard errors and F-stats are clustered by metro for zip level regressions. Standard errors for metro regressions are assumed iid. Anderson-Rubin confidence intervals are provided in square brackets for IV regressions. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The IV regressions are similar to the OLS specification, implying that measurement error is not driving the estimated negative yield-price covariance, and the implied excess dispersion

in prices. Specifically, using owner-occupier prices and transacted prices yields coefficients of -0.695 and -0.761 across cities, and -0.717 and -0.691 within them. The square footage instrument yields a negative coefficient, whose confidence intervals, although large, include the OLS coefficient.

I report robust Kleibergen and Paap (2006) first stage F-stat[20] The first two instruments are not weak, but I cannot reject that the square foot instrument is. Since one of the instruments is weak, I compute Anderson-Rubin (Anderson and Rubin, 1949) confidence intervals, and report these in square brackets. In all cases, confidence set includes the point estimates for the OLS and is bounded away from zero.

4.2 Implied Expected Growth

Since the large negative ratio-price covariance is not driven by measurement error, it must be some combination of spatial dispersion in: (*1*) returns, and (*2*) expected growth rates. In this section, I argue that it is highly implausible that all or even most of this relationship can be attributed to different growth expectations using the magnitude of the yield-price covariance itself. To do so, I will make a simplifying assumption on how expected growth rates relate to current prices. Section 5 will make a similar argument, using realized data on returns and growth, without imposing assumptions on the growth process.

I show that in high-price areas, prices are so high relative to rents that justifying them solely with growth rates would require local investors to have unrealistically high growth expectations. For example, in Boston's Back Bay, where prices are extremely high and rent-to-price ratios are very low compared to other neighborhoods, investors would need to anticipate growth rates 10 percentage points above the city average to justify these disparities.

Cross-sectional Campbell-Shiller Decomposition. I turn to quantifying the implications of the large negative ratio-price covariance on expected growth, under the 'null' that cross-sectional expected returns are equal. Consider a cross-sectional variant of the Campbell and Shiller (1988) decomposition, in which rent growth is partially determined (or proxied) by a contemporaneous variable x. This model, under a linearity assumption on the x-growth relationship, can be easily estimated and used to quantify differences in growth expectations based on differences in x. Taking x as current prices, we use the estimates in Table 3 to estimate the implied difference in growth expectations in high- and low-price areas.

The standard Campbell-Shiller decomposition can be applied in the cross section (across

[20]Kleibergen and Paap (2006) are also equivalent to the 'efficient' Olea and Pflueger (2013) F-stat in the case of a single regressor.

spatial markets denoted i) as:

$$(d-p)_{t,i} = \frac{-\kappa}{1-\rho} + \sum_{j=0}^{\infty}(\rho^j \mathbb{E}\left[r_{t+1+j} - \Delta d_{t+1+j,i}\right]), \tag{5}$$

where ρ and κ are log-linearization parameters defined around a 'steady state' rent-price ratio.[21] I follow the standard notation for the cash-flow component, $dp_{iF,t,i} = -\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j,i}$ and $dp_{DR,t} = \sum_{j=0}^{\infty}\rho^j r_{t+1+j}$. Note that the returns component, $dp_{DR,t}$ is constant across markets, consistent with the null hypothesis.

The key simplifying assumption is that rent growth is a linear function of the contemporaneous local x (denoted $x_{t,i}$). Growth can otherwise follow any path, which is denoted by $\epsilon_{t,j,i}$. I assume that the linear term $x_{t,i}$ captures all the effects of it on growth, formally $\epsilon_{t,j,i} \perp x_{t,i}$.

$$dp_{iF,t,i} = -\sum_{j=0}^{\infty}\rho^j \mathbb{E}\left[\Delta d_{t+1+j,i}\right] = -\sum_{j=0}^{\infty}\rho^j(\alpha_{0,i} + \alpha_1 x_{t,i} + \mathbb{E}[\epsilon_{t,j,i}]) \tag{6}$$

We can now quantify the effect of $x_{t,i}$ on rent growth expectations:

$$dp_{iF,t,i}(x_{t,i}) - dp_{iF,t,i}(\bar{x}_t) = \frac{\alpha_1}{1-\rho}(x_{t,i} - \bar{x}_t)$$

which captures how much higher is the aggregate, discounted, rent-growth component in market i, due to x_i, relative to some average $\bar{x}$, at time t. α_1 can also be thought of as the difference in annual growth rates among higher $x_{t,i}$ locals.

Estimating Growth Expectations in Low vs High Price Areas. Taking x as current prices, the above model suggests the following empirical specification (omitting time subscript for brevity):

$$(d-p)_i = \beta_0 + \beta_1 x_i + \nu_i \tag{7}$$

β_0 and ν_i capture the return component $dp_{DR,i}$ and any growth process in $dp_{iF,i}$ that is orthogonal to x_i. β_1 is identified under the model's assumptions, and can be interpreted as the difference in cash flow $dp_{iF,i}$ for a log-point difference in x_i. We can further compute α_1 as $\alpha_1 = -\beta_1(1-\rho)$.

Setting $x_i = p_i$ and estimating Equation (7) results in the coefficients reported in Table 3. This relationship can be interpreted as either: (1) prices reflecting a set of local fundamentals

[21] The cross-sectional linearization parameters are equal to:

$$\rho = \frac{1}{1+\exp(\overline{d-p})}, \qquad \kappa = -\log(\rho) - (1-\rho)\log(1/\rho - 1)$$

that determine growth, or as investors forming rent growth expectations based on observed prices. In the former interpretation, current prices only capture a set of fundamentals, not set them, and the model estimates the reduced-form relationship between rent-price ratios and these fundamentals. Regardless, under the null of equal expected returns, the estimated coefficient is informative on the degree to which investors expect local growth to vary with contemporaneous prices.

We can therefore use the regressions in Table 3, adjusted for $1-\rho$[22], to estimate investors' growth expectations from current prices. Column 3, which uses owner-occupier prices as an IV, suggests that an increase of 1 log points in metro prices, relative to the national average, implies investors expect 6.95 pp annual higher growth rates. Within cities, investors expect a similar 7.2 pp higher growth for every log-point difference in current prices. We can repeat this exercise in z-scores: a one standard deviation increase in local prices across (within) metro areas, results in 4pp (4.5 pp) higher expected growth per year.

Figure 4 uses the contemporaneous price distribution to plot the density of implied relative expected growth rates, across and within metro areas. These relative growth rates seem implausible for two reasons: (1) they imply extreme growth rates, rarely seen in reality, and (2) the dispersion of growth rates across space is implausibly large.

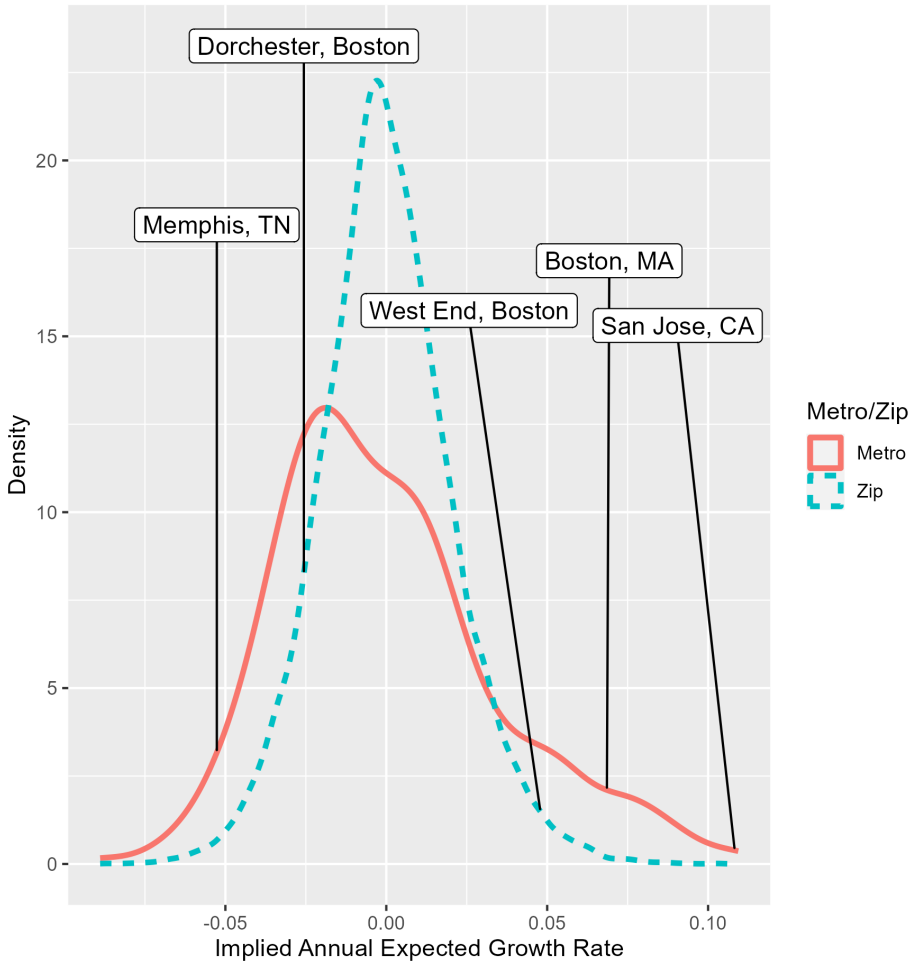
For example, rents in expensive metros such as Boston, MA, and San Jose, CA, are implied to grow annually at 6.8pp and 10.8pp above the national average. These rates imply an expectation that the cities' nominal rents double every 8 and 5.67 years (respectively), and double relative to the national average every 10.5 and 6.75 years. Similarly unlikely are the results for cheaper cities, such as Memphis, TN, where low prices imply a relative average growth rate of -5.3pp, which would result in a nominal average rent decline. Such long-term declines are far rarer than would be implied by these results, and far from reality for cities such as Memphis.

Too-high and too-low rates are also implied within cities. Expensive zip-codes are implied to have growth much higher than their metro average, with 10% growing 2.5pp faster, and 5% growing 3.4pp faster, annually. These patterns extend to cheaper neighborhoods as well. For example, established areas around Downtown Boston are expected to grow 4.7pp faster than Boston's average, whereas in parts of the gentrifying neighborhood of Dorchester rent would be expected to decline by 2.5pp every year. Such high and low expectations, as well as their extreme dispersion, both across and within metros, are highly implausible, and contrary to observed patterns.

Comparison with 2010-2020 Growth. While it is impossible to rule out all potential

[22]Reasonable estimates for rho could include the average across all units, across zip codes, or across markets, all of which result in $\rho = 0.89$ and $(1-\rho) = 0.11$.

Figure 4: Implied Average Rent Growth Rate



This figure visualizes the average annual rent growth rate implied by contemporaneous prices, under the null that returns arbitrage across space. The solid line plots the distribution of metro growth rates, relative to the national average, and the dashed line plots the distrubtion of zip growth rates, relative to each zip's metro's average.

growth trajectories that could, on average, match the average rates described above, it is noteworthy—albeit unsurprising—that such extreme differences in rent growth did not materialize in the decade following the data's period.

Table 5 in the Appendix repeats the regressions of Panel B in Table 3, adding ex-post growth data from 2010 to 2020 to the left-hand side variable. If the ratio-price relationship was primarily driven by expected growth rates, the coefficients in these new regressions should decline. As a comparison, I also add three local variables from Table 2 - density, education, and distance to CBD. However, the price coefficients decrease only slightly, from approximately -0.06 to -0.05. Of the other variables, only the coefficient for the education variable attenuates considerably.

This indicates that while rent growth was somewhat higher in more expensive areas, it is far from sufficient to account for the strong relationship between ratios and prices. Of the correlations in Table 2, only education appears to signal rent growth[23] To further explore this point, I turn to a panel setting, using past growth realized and returns data to estimate the relative significance of each channel.

5 Ex-post Decomposition (Campbell-Shiller)

In this section I estimate the relative importance of rent growth and returns in explaining rent-price ratio dispersion, in ex-post data. I construct a panel dataset on rent-price ratios, rent growth rates, and returns between 1980 and 2040, for metros and zip codes. I consistently find that ex-post return dispersion explains a large, if not dominant share of realized rent-price dispersion - 82.6% within metros and 90% across metros. These results are therefore in line with those on growth expectations (Section 4), suggesting some role for growth expectations, but a larger one for returns, in explaining the cross-sectional dispersion in rent-price ratios.

The advantage of this analysis is that it does not make assumptions on the price- growth relationship, and that it estimates the relative importance of returns in ratio dispersion, rather than simply rejecting the null that they play no role. Its disadvantage is that it only relates to realized, not expected, returns.

[23]This finding is in line with the urban literature, which has found education to be a key driver in rent growth and the spatial polarization of the economy and housing markets, see Diamond (2016); Glaeser et al. (2018).

5.0.1 Cross-Sectional Campbell-Shiller Decomposition

The empirical Campbell-Shiller decomposition (Campbell and Shiller, 1988) starts from the log-linearization of the definition of returns, as in Equation (5). Differencing out this equation by one period, we can write:

$$r_{i,t+1} + g_{i,t} - (p_{i,t+1} - d_{i,t+1}) = (d_{i,t} - p_{i,t}) \tag{8}$$

Where $g_{i,t}$ denotes the one period growth rate at location i. (The future path of growth and returns, after time $t+1$, cancels out.) Over a finite period, the path of rent growth isn't fully realized, and so rent-prices partially persist. I refer to the non-persistent component as 'realized' rent-price ratio dispersion.

I follow Cohen et al. (2003) in iterating Equation (8) N periods forward, and derive a system of equations, Equation (9), that decomposes rent-price ratios at time t into three components: rent growth, returns, and future rent-price ratios, as in Cochrane (2008).

I add period-specific fixed effects to cleanse intertemporal variation, so that estimates are based on cross-sectional variation only.[24] These fixed effects capture nationwide shocks to rent-price ratios and aggregate growth and returns and absorb any nationwide shocks between the periods, such as changes in interest rates.

$$\begin{aligned}
\tilde{g}_{i,t+N} &\equiv \sum_{j=1}^{N} \rho^{j-1} g_{i,t+j} = \delta_t^g + b^d(N)(d-p)_{i,t} + \varepsilon_{i,t+1}^d \\
\tilde{r}_{i,t+N} &\equiv \sum_{j=1}^{N} \rho^{j-1} r_{i,t+j} = \delta_t^r + b^r(N)(d-p)_{i,t} + \varepsilon_{i,t+1}^r \\
\rho^N \tilde{dp}_{i,t+N} &\equiv \rho^N (d-p)_{i,t+1} = \delta_t^{dp} + \phi(N)(d-p)_{i,t} + \varepsilon_{i,t+1}^{dp}
\end{aligned} \tag{9}$$

Where the tilde variables denote the aggregation over the horizon N, and ρ denotes the fixed effects in time. The coefficients $b^d(N)$ and $b^r(N)$ capture the share of the cross-sectional variation in ratios at time t that is realized into growth and returns, by time $t+N$. $\phi(N)$ is the autocorrelation of ratios across the horizon N.

Equation 8 implies a tight relationship between the coefficients $b^r(N), b^d(N)$ and $\phi(N)$ - they must add up to one. Therefore, any realized (if $\phi < 1$) rent-price dispersion is realized into some combination of returns (b^r) or growth (b^d). The null hypothesis, that returns are

[24] Fama and French (2023) estimate the rent-growth component of Equation 9 on a small ten metro sample, with and without time fixed effects, and focus on a shorter time horizon of up to one year. The authors find rent growth to be more predictive in the cross section. Fama and French then repeat the analysis for price growth, finding no difference. Estimating Equation (9) achieves a different goal - estimating the relative importance of rent growth and returns in cross-sectional rent-price dispersion.

equal over space, can be expressed as $b^r(N) = 0$, which also implies $b^d(N) = \phi(N) - 1$.

Rather than simply rejecting this null, we can quantify the importance of dispersion in returns. A useful benchmark for this the ratio $\frac{b^r(N)}{b^r(N)-b^d(N)}$, which is a continuous measure of the importance of return differences to spatial dispersion in rent-price ratios, and equals zero under the null. More precisely, as contemporaneous rent-price ratios signal a path of future returns and rent growth, this ratio captures the relative importance of returns within the realized part of that path at horizon N. If one were to adopt Cochrane's 'infinite horizon' decomposition (by assuming $b^r(N), b^d(N)$ are constant over N,) then this ratio would simply equal the share of rent-price dispersion attributed to returns.[25]

5.0.2 Constructing Panel Datasets

To implement the analysis I construct two panel datasets: a zip and a metro panel. The zip dataset builds on the DataQuick data, and extends it to four decade points, from 1980 to 2010, with additional data for 2020 based on HMDA data. The metro panel is based on CBSA-level rent and price indices reported by the BLS and FHFA and benchmarked using the DataQuick data. This dataset is at the annual level and begins in 1975. Despite their differences the two datasets lead to a similar qualitative result.

Decadal DataQuick Panel. I construct a zip and metro panel from unit-level rents and prices as described in Section 2, but apply it separately for each decade starting from 1980 until 2010.[26] Based on owners' mailing address I restrict the sample to household investors - classified as those whose mailing address is a residential address. I expand the data to 2020, using DataQuick information on the housing stock, the 2020 census, and 2019 AHS for rents and HMDA reported prices paid by investors within each zip code.

I average unit rent and price at each decade and zip code. I then repeat this process at the metro level. These values represent those earned for the 'average unit' held be small investors. (For more details on the rent and price estimation, see Section 2.) I construct 10-year-horizon aggregates by splitting up decadal changes into implied average annual changes, and discounting as described in Equation (9).

Annual BLS/FHFA Panel. I construct the metro level panel using average rents and price growth, based on annual metro BLS and FHFA indices, as in Chodorow-Reich et al.

[25] I show below that, consistent with this interpretation, the variables $b^r(N), b^d(N)$ are remarkably constant over different horizons N, although the short time dimension of the panel limit $N \leq 10$.

[26] Local zip mean rents are only available at a decadel frequency until 2007, and thereafter only available as part of the 5-year ACS waves. This is the key limiting factor that forces the decadal structure of the data. I use each year's zip-level average rent (as reported by the Census or ACS) and that year's AHS survey. Since the AHS is not conducted every year, I use the 1989 wave for 1990, the 1999 for 2000, and the 2011 for 2010.

(2021). I use my rent and price estimates to benchmark each series to 2013 and derive dollar values and then compute annual returns and rent growth. The cross section of the BLS sample is limited, with only 24 cities in an unbalanced panel. I compute annual rent growth and returns for each year and aggregate these across different horizons. These values correspond to a compounded, discounted, growth, and returns for a 'metro average unit' over each horizon N.

Comparing the Datasets The DataQuick and BLS/FHFA datasets complement each other: the former has a rich spatial dimension but short temporal frequency, and the latter a longer time series but smaller cross section.

The advantages of the DataQuick sample are its within-city variation and its use of only rental properties (as of 2013) to construct rent-price ratios and price changes. The sample is appropriate for estimating the role of returns both within and across cities, accounting for granular spatial heteroskedasticity (as in Conley (1999)) and alleviates measurement concerns. Its disadvantage is the decadal frequency, which implies only a 10-year horizon, and does not allow accounting for time-series heteroskedasticity in computing standard errors.

The key advantage of the BLS/FHFA metro data is its annual frequency, which allows testing of multiple time horizons, and a more robust treatment of serial correlation as in Newey and West (1987) and Driscoll and Kraay (1998). I focus on annual and 10-year horizons, but also present results for 3-, 5- and 15-year horizons in Figure 5.

5.0.3 Decomposition Results

Rent-price ratios are somewhat persistent, particularly across cities, but higher ratios signal significantly lower rent growth and higher returns. Of these, returns explain the vast majority of realized rent-price dispersion, both across and within cities. This pattern does not seem to depend on the horizon over which returns are computed.

Table 4 details these results. Panel A presents results at the 1-year horizon and panel B at a 10-year horizon. Columns 1-3 estimate the system in equation (9) across metros for the BLS/FHFA metro, and columns 4-6 do the same within cities for the DataQuick sample. Below each set of regressions, I report the ratio $\frac{b^r(N)}{b^r(N)-b^d(N)}$ and its estimated standard deviation.

Differential returns appears to be significantly more important than rent growth in explaining contemporaneous cross-sectional dispersion in rent-price ratios. Across metros, 90% of realized rent-price ratio dispersion is realized into different returns, and only 10% into rent growth. Within metros, 82.6% of ratio dispersion is realized into differential returns, suggesting investors are somewhat better at forecasting returns within cities and/or that more

of return dispersion is can be arbitraged. The importance of returns is not unique to the chosen horizons, Figure 5 shows the ratio $b^r(N)/(B^r(N) + b^d(N))$ across N to be stable. Differences across and within cities are not driven by the different samples, Table 6 in the Appendix provides a cross-metro analysis using DataQuick, showing returns account for 87.6% of realized dispersion. Among the same 24 metro areas, the DataQuick effects are qualitatively similar, where returns explain 75% of dispersion.

Table 4: Campbell-Shiller Decomposition of Cross-Sectional Rent-Price Dispersion

	Across Metros			Within Metros		
Dependent Variables:	(1) $g_{i,t+N}$	(2) $r_{i,t+N}$	(3) $\rho^N dp_{i,t+N}$	(4) $g_{i,t+N}$	(5) $r_{i,t+N}$	(6) $\rho^N dp_{i,t+1}$
	Panel A: 1-Year Horizon					
$dp_{i,t}$	-0.008***	0.070***	0.930***	-	-	-
	(0.002)	(0.007)	(0.006)			
Statistics						
Within R^2	0.03	0.31	0.99	-	-	-
Observations	1,037	1,037	1,037	-	-	-
$\frac{b^r}{b^r - b^d}$	0.9		-	-		-
	(0.035)					
	Panel B: 10-Year Horizon					
$dp_{i,t}$	-0.056***	0.504***	0.475***	-0.120***	0.571***	0.312***
	(0.010)	(0.015)	(0.015)	(0.016)	(0.056)	(0.013)
	-	-	-	[0.01]	[0.029]	[0.008]
Statistics						
Within R^2	0.09	0.73	0.88	0.07	0.58	0.38
Observations	821	821	821	16,896	16,896	16,896
$\frac{b^r}{b^r - b^d}$	0.9		-	0.826		-
	(0.021)			(0.035)		
Fixed-effects & Panel Size						
Year/Decade	Yes	Yes	Yes	Yes	Yes	Yes
Metro				Yes	Yes	Yes
T	47	47	47	4	4	4
J	24	24	24	5,390	5,390	5,390

Metro level data is based on BLS rent and FHFA price indices, benchmarked as of 2013. Zip level data is based on zip-level aggregations from DataQuick for the housing stock available at each period and contemporaneous zip-level averages. Standard errors across metros are Driscoll-Kraay with an optimal bandwidth selected as in Lazarus et al. (2018). For the zip level analysis CBSA clustered standard errors are reported in parentheses, and Conley standard errors, with a spatial threshold of 20 miles, are reported in square brackets. The ratio of $b^r(N)/(b^r(N) - b^d(N))$ captures the relative importance of different returns across space. I calculate its standard error using the delta method and the Driscoll-Kraay (for metro sample) and clustered (for zip sample) covariance matrices. Reported stars use t-statistics adjusted for fixed-b asymptotics as in Kiefer and Vogelsang (2005). *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Standard Errors. For the cross-metro sample, I calculate Driscoll-Kraay standard errors, which are robust to time series and cross-sectional heteroskedasticity, optimally choosing the auto-correlation bandwidth (Driscoll and Kraay, 1998; Lazarus et al., 2018). Due to the short panel, I cannot do the same for the DataQuick sample. Instead, I report two sets of standard errors: (a) clustered errors along CBSAs in parentheses; and (b) Conley spatial standard errors (Conley, 1999) in square brackets, with a spatial threshold of 20 miles. These approaches span a metro-level and highly localized heteroskedasticity error structure. Estimates coefficients are relatively precise, and differences are highly significant even when using fixed-b critical values as in Kiefer and Vogelsang (2005).

I estimate standard errors for the ratio $\frac{b^r(N)}{b^r(N)-b^d(N)}$ using the delta method and the robust estimator covariance matrix (Driscoll-Kraay for the metro sample, clustered for the zip sample).[27] The standard errors around these ratios are small, such that returns explain more than half the rent-price dispersion in all specifications and for horizons in a statistically significant way.

Stability Across the Horizon N. Figure 5 presents estimates for the ratio $\frac{b^r(N)}{b^r(N)-b^d(N)}$ across metros (BLS/FHFA dataset), estimated over different horizons N. These estimates span intermediate ranges between Panels A and B in Table 4, as well as adding an estimate over 15 years. (The horizon is limited by the overall duration of the panel, 47 years.)

Across horizons, the importance of returns is flat, at about 90% of realized dispersion, at horizons from 1-year to 15 years. None of the estimated coefficients are statistically different from one another. The figure suggests a possible downward trend at longer horizons, but a longer panel will be required to further investigate this.

6 Conclusion

This paper analyzes the cross-sectional dispersion of rent-price ratios in the U.S., utilizing property-level data to estimate local rent-price ratios across metro areas and neighborhoods. These estimates address key measurement concerns raised in the literature (Eichholtz et al., 2021; Glaeser and Gyourko, 2007). I document significant dispersion, both across and within cities, organized in spatial patterns that partly align with known economic factors affecting prices and rents. Differences in yields imply either differential expectations for rent growth or differential returns on similar assets, both with distinct implications for local economies.

[27] To estimate the variance-covariance matrix for b^r, b^d, ϕ across all three regressions I stack the OLS moment conditions. This is equivalent to running a regression with a single dependent variable equal to the stacked $(\tilde{g}_{i,t+N}, \tilde{r}_{i,t+N}, \rho^N \tilde{dp}_{i,t+N})$ vector on a matrix of regressors equal to $dp_{i,t} \otimes \mathcal{I}_{3\times 3}$ I add fixed effects by de-meaning each moment by each time period. The result of this regression is identical to running three regressions, but estimates the full variance co-variance matrix for the estimators.

Using the Campbell-Shiller decomposition, I show that differential returns are a key component of yield dispersion, both ex-ante and ex-post. Expected returns must vary considerably over space to explain contemporaneous yield patterns. Realized data from 1980 to 2020 confirm this: returns explain 90

Differential returns across geographic markets have significant implications for household finance, real estate investment, and housing supply. Real estate, beyond owner-occupied housing, represents a key asset class held by 7–9% of U.S. households, often supported by inexpensive financing. Evidence from Samuels (2024) reveals that these investors are predominantly upper-middle-class but, conditional on wealth, are less educated and less White. The higher returns in some markets may explain these patterns and counterbalance the general tendency of wealthier households to achieve higher returns in other asset classes Xavier (2021).[28]

If rental markets influence or reflect owner-occupier markets, as suggested by Ater et al. (2021), differential returns driven by higher relative rents could have profound implications for household wealth accumulation—impacts not fully captured by analyses focused solely on price growth (e.g., Goldsmith-Pinkham and Shue (2023)). Future research is essential to explore the motivations and broader implications of household real estate investment.

These findings challenge a canonical assumption in urban economic models: that returns are arbitraged across space and primarily accrued by "absentee" landlords DiPasquale and Wheaton (1996). While simplifying these models, the observed spatial variation in yields and returns suggests this assumption may overlook key mechanisms. Risk, as proposed by Demers and Eisfeldt (2022) and Amaral et al. (2021), may partly explain these patterns, but evidence is limited, often to aggregate data Lockwood et al. (2023) or residualized proxies Eiling et al. (2019).

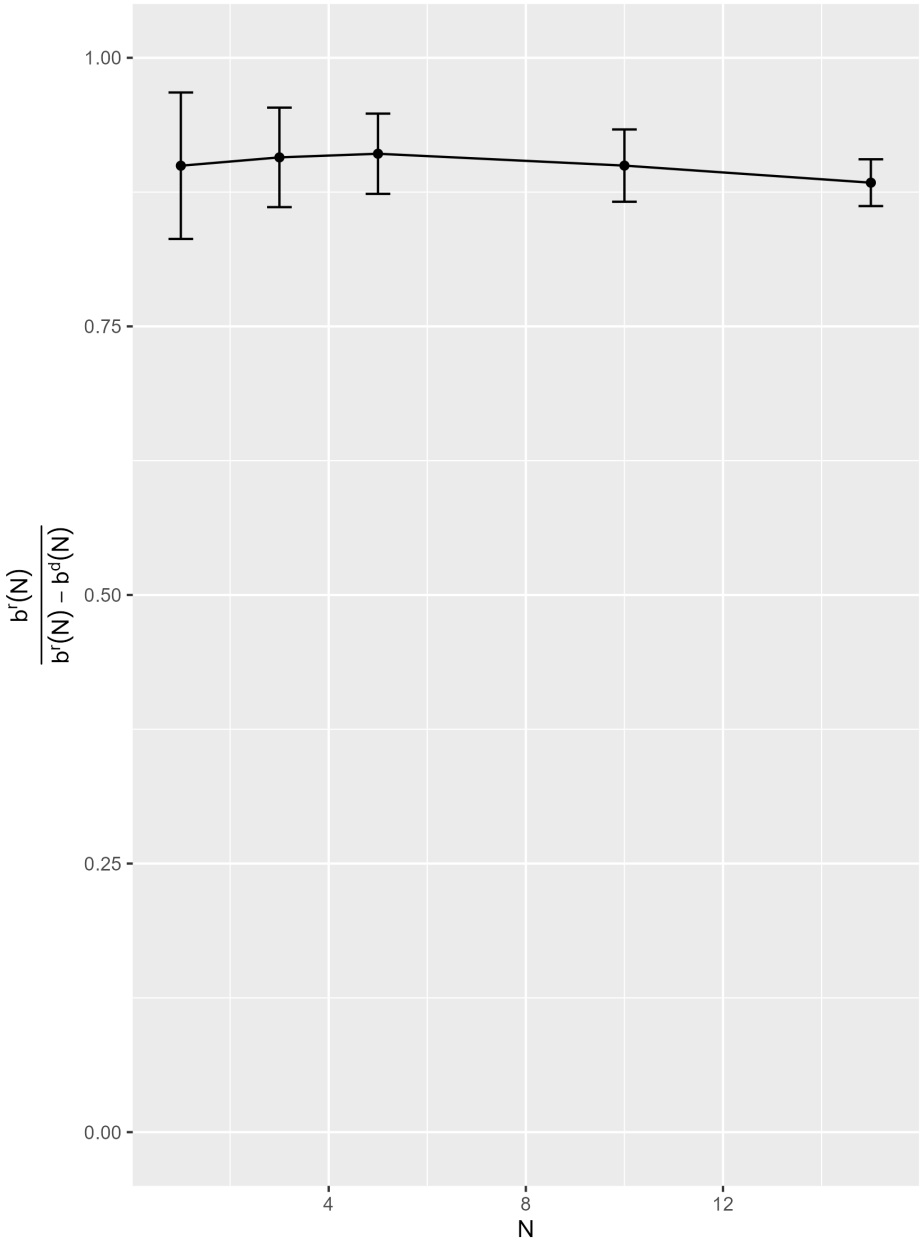
In addition to risk, mechanisms rooted in local market dynamics are likely important. Levy (2021) demonstrates how home-biased, wealthy landlords increase housing capital supply, while Samuels (2024) extends this by endogenizing returns as a function of local fundamentals, including the intersection of risk and investor characteristics. These mechanisms offer a promising avenue for understanding housing markets and local economic dynamics, yet remain underexplored in the literature.

Together, these findings underscore the need for further research on real estate investors, their motivations, and their broader impacts on local markets. Financial tools like the Campbell-Shiller decomposition, combined with frameworks that account for risk aversion

[28]Differential returns may also shed light on the rise of institutional single-family rentals, as argued by Demers and Eisfeldt (2022), a finding this paper extends to a broader cross-section of markets and housing segments.

and portfolio decision-making, could deepen our understanding of local economies, housing markets, and household finance.

Figure 5: Importance of Returns Out of Realized Dispersion, at Different Horizons



This figure presents the ratio of the share of cross-sectional rent-price ratio dispersion that can be explained by different returns, b^r, relative to that of both returns and rent-growth, $b^r - b^d$. The error bars are delta method derived 95% confidence intervals around this ratio.

References

F. Amaral, M. Dohmen, S. Kohl, and M. Schularick. Superstar returns. 2021.

F. Amaral, M. Dohmen, S. Kohl, and M. Schularick. Interest rates and the spatial polarization of housing markets. 2022.

T. W. Anderson and H. Rubin. Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of mathematical statistics*, 20(1): 46–63, 1949.

I. Ater, Y. Elster, and E. B. Hoffmann. Real-estate investors, house prices and rents: Evidence from capital-gains tax changes. *Maurice Falk Institute for Economic Research in Israel. Discussion paper series.*, (1):1–33, 2021.

S. Büchler, M. v. Ehrlich, and O. Schöni. The amplifying effect of capitalization rates on housing supply. *Journal of urban economics*, 126:103370, 2021.

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

S. D. Campbell, M. A. Davis, J. Gallin, and R. F. Martin. What moves housing markets: A variance decomposition of the rent–price ratio. *Journal of Urban Economics*, 66(2): 90–102, 2009.

D. Chambers, C. Spaenjers, and E. Steiner. The rate of return on real estate: Long-run micro-level evidence. *The Review of Financial Studies*, 34(8):3572–3607, 2021.

G. Chodorow-Reich, A. M. Guren, and T. J. McQuade. The 2000s housing cycle with 2020 hindsight: A neo-kindlebergerian view. Technical report, National Bureau of Economic Research, 2021.

J. H. Cochrane. The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, 21(4):1533–1575, 2008.

R. B. Cohen, C. Polk, and T. Vuolteenaho. The value spread. *The Journal of Finance*, 58 (2):609–641, 2003.

S. Colonnello, R. Marfè, Q. Xiong, et al. Housing yields. *Carlo Alberto Notebooks*, pages 1–86, 2023.

T. G. Conley. Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45, 1999.

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

R. Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3):479–524, 2016.

R. Diamond and W. F. Diamond. Racial differences in the total rate of return on owner-occupied housing. Technical report, National Bureau of Economic Research, 2024.

D. DiPasquale and W. C. Wheaton. Urban economics and real estate markets. *(No Title)*, 1996.

J. C. Driscoll and A. C. Kraay. Consistent covariance matrix estimation with spatially dependent panel data. *Review of economics and statistics*, 80(4):549–560, 1998.

P. Eichholtz, M. Korevaar, T. Lindenthal, and R. Tallec. The total return and risk to residential real estate. *The Review of Financial Studies*, 34(8):3608–3646, 2021.

E. Eiling, E. Giambona, R. Lopez Aliouchkin, and P. Tuijp. Homeowners' risk premia: Evidence from zip code housing returns. *Available at SSRN 3312391*, 2019.

E. F. Fama and K. R. French. House prices and rents. *Chicago Booth Research Paper*, (20-48), 2023.

C. Garriga, F. E. Kydland, and R. Šustek. Mortgages and monetary policy. *The Review of Financial Studies*, 30(10):3337–3375, 2017.

E. L. Glaeser and J. Gyourko. Arbitrage in housing markets. Technical report, National Bureau of Economic Research, 2007.

E. L. Glaeser, H. Kim, and M. Luca. Nowcasting gentrification: using yelp data to quantify neighborhood change. In *AEA Papers and Proceedings*, volume 108, pages 77–82. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2018.

P. Goldsmith-Pinkham and K. Shue. The gender gap in housing returns. *The Journal of Finance*, 78(2):1097–1145, 2023.

D. L. Greenwald and A. Guren. Do credit conditions move house prices? Technical report, National Bureau of Economic Research, 2021.

J. Gyourko, C. Mayer, and T. Sinai. Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–199, 2013.

J. Gyourko, J. S. Hartley, and J. Krimmel. The local residential land use regulatory environment across us housing markets: Evidence from a new wharton index. *Journal of Urban Economics*, 124:103337, 2021.

N. Hausman, P. Samuels, M. C. Cohen, and R. Sasson. Urban pull: The roles of amenities and employment. *Available at SSRN 3670974*, 2023.

C. A. Hilber and A. Mense. Why have house prices risen so much more than rents in superstar cities? 2021.

C. Himmelberg, C. Mayer, and T. Sinai. Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, 19(4):67–92, 2005.

Ò. Jordà, K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor. The rate of return on everything, 1870–2015. *The Quarterly Journal of Economics*, 134(3):1225–1298, 2019.

N. M. Kiefer and T. J. Vogelsang. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, 21(6):1130–1164, 2005.

F. Kleibergen and R. Paap. Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics*, 133(1):97–126, 2006.

E. Lazarus, D. J. Lewis, J. H. Stock, and M. W. Watson. Har inference: Recommendations for practice. *Journal of Business & Economic Statistics*, 36(4):541–559, 2018.

A. Levy. Housing policy with home-biased landlords: Evidence from french rental markets. 2021.

L. M. Lockwood, L. Kueng, P. Ong, and S. R. Baker. Renting hedges wage risk. Preliminary and incomplete: Please do not circulate without permission., March 2023.

L. Loewenstein and P. S. Willen. House prices and rents in the 21st century. Technical report, National Bureau of Economic Research, 2023.

R. Molloy, C. G. Nathanson, and A. Paciorek. Housing supply and affordability: Evidence from rents, housing consumption and household location. *Journal of Urban Economics*, 129:103427, 2022.

W. K. Newey and K. D. West. Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pages 777–787, 1987.

J. L. M. Olea and C. Pflueger. A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369, 2013.

A. Plazzi, W. Torous, and R. Valkanov. Expected returns and expected growth in rents of commercial real estate. *The Review of Financial Studies*, 23(9):3469–3519, 2010.

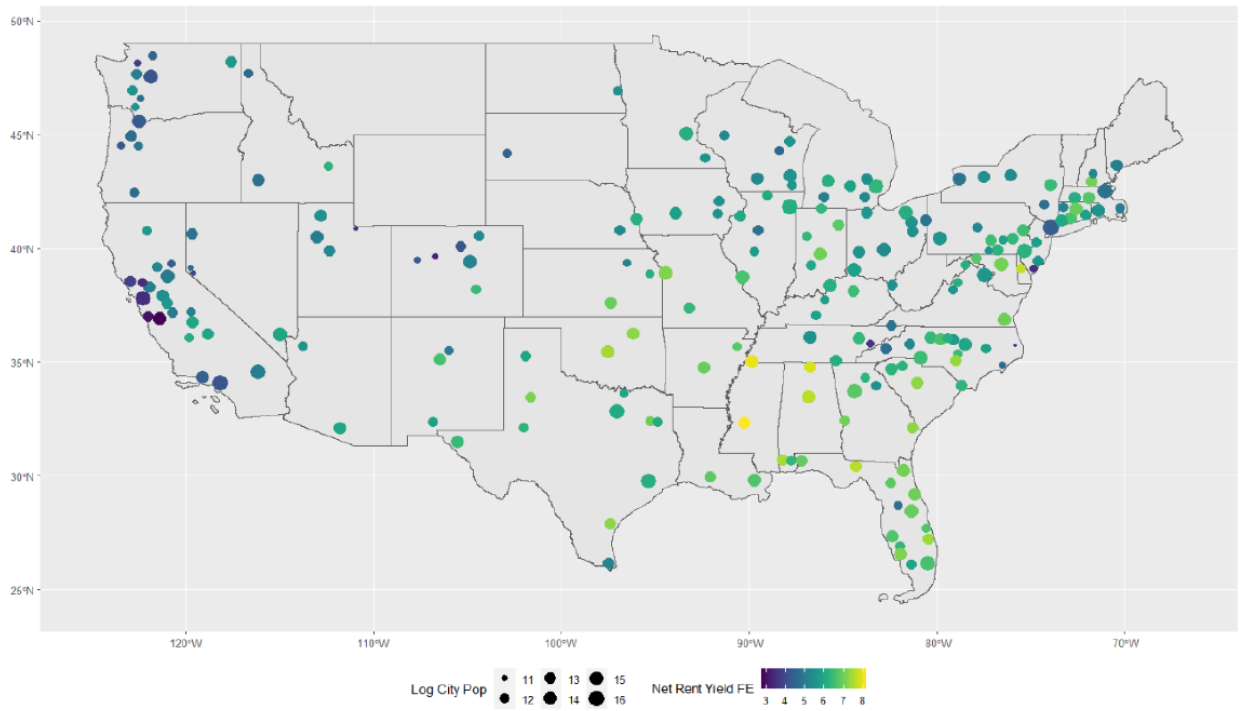
P. Samuels. Real estate investors, their wealth and local housing prices and returns. *Working Paper*, 2024.

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

T. Vanhapelto. House prices and rents in a dynamic spatial equilibrium. *Job Market Paper*, 2022.

I. Xavier. Wealth inequality in the us: the role of heterogeneous returns. *Available at SSRN 3915439*, 2021.

Figure 6: Estimated Metro Rent-Price Ratios, 2018



Estimated CBSA level Rent-Price ratio in 2018. Data is based on tract level rents and prices paid for investor properties in HMDA. Unlike the DQ sample used throughout this paper, rents are not adjusted for unit-level characteristics.

7 Appendix

7.1 Figures

7.2 Tables

Table 5: Rent-Price Ratios Patterns and Implied Growth

	Panel A: Within Metro (Zip); Dep' Var: D/P_i						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(Density)	0.007*						
	(0.003)						
Share BA+ (z-score)		-0.014***					
		(0.002)					
log(Dist' CBD) (z-score)			-0.013***				
			(0.002)				
log(Rental Price) (z-score)				-0.067***	-0.058***	-0.062***	-0.062***
				(0.006)	(0.005)	(0.006)	(0.007)
1st Stage F	-	-	-	-	1261.79	9581.91	5287.77
Observations	10,699	10,699	10,699	10,699	10,699	10,699	10,380
Within R^2	0.015	0.064	0.051	0.486			
Dependent Variable:	*Panel B:* Within Metro (Zip); Dep' Var: $D/P_i + \hat{g}_{10\to20,i}$						
log(Density)	0.009**						
	(0.004)						
Share BA+ (z-score)		-0.008***					
		(0.002)					
log(Dist' CBD) (z-score)			-0.017***				
			(0.002)				
log(Rental Price) (z-score)				-0.062***	-0.057***	-0.047***	-0.047***
				(0.005)	(0.006)	(0.007)	(0.007)
Instrument	-	-	-	-	log(Sqft)	log(owner P)	log(Sales P)
1st Stage F	-	-	-	-	1261.79	9581.91	5287.77
Fit statistics							
Observations	10,413	10,413	10,413	10,413	10,413	10,413	10,129
Within R^2	0.024	0.018	0.085	0.378	0.376	0.356	0.364
Instrument	-	-	-	-	log(Sqft)	log(owner P)	log(Sales P)
Unit Characteristics	✓	✓	✓	✓	✓	✓	✓

Note: This table presents regressions of zip area (CBSA) yields on local characteristics. All variables are at the zip level. Yields are residualized by building type, decade of construction and corporate ownership. The left hand side variables are standardized as z-scores across zip-codes. Column 4 instruments for prices with unit level (residualized) square feet and number of bedrooms, Column 5 instruments with owner-occupied prices in the same zip from the ACS, and Column 6 instruments with sale prices in the same zip. Rents and Prices are estimated as a hedonic based on unit characteristics, location averages, and in the case of prices also assessed values. *Note:* *p<0.1; **p<0.05; ***p<0.01.

Table 6: Metro-level Campbell-Shiller Decomposition, DataQuick Sample

	All Metro (DQ)			Metros in BLS Sample (DQ)		
	Panel C: 10-Year Horizon					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables:	$\tilde{g}_{i,t+N}$	$\tilde{r}_{i,t+N}$	$\rho^N \tilde{dp}_{i,t+N}$	$\tilde{g}_{i,t+N}$	$\tilde{r}_{i,t+N}$	$\rho^N \tilde{dp}_{i,t+N}$
$dp_{i,t}$	-0.085***	0.597***	0.292***	-0.157***	0.488***	0.306***
	(0.012)	(0.032)	(0.015)	(0.021)	(0.071)	(0.039)
$\frac{b^r}{b^r - b^d}$	0.876		-	0.757		-
	(0.018)		-	(0.04)		-
T	4	4	4	4	4	4
N	207	207	207	19	19	19
Decade	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	676	676	676	72	72	72

*Clustered (CBSAFP) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*