The Effect of Local Landlords on Rental Housing Prices and Returns

Peleg Samuels[‡]

Aptil 2025

Abstract

What is the role of real estate investors in determining housing prices? Household investors, who own small, undiversified portfolios and are highly home-biased, own a large share of the U.S. rental market: 53.5% of all rental housing units and higher shares in older and smaller properties. I propose a framework to understand the effect of household investors' behavior and characteristics on the price, returns, and rent-price ratio of rental housing. In the model, risk-averse investors self-select from the local population and supply capital in a local 'capital market' for rental units. Areas in which residents have a higher propensity to invest in real estate, or higher wealth, have higher rental prices, and lower returns and rent-price ratios. To test the model, I construct a novel dataset of local rent-price ratios, returns, and local investors' characteristics. At the neighborhood level, I utilize an IV derived from local ancestry composition, and find that when the share of investors out of the population, or their wealth, is high, prices are higher relative to rents, signaling lower returns. At the metro level, I construct a shift-share instrument from nation-wide equity shocks and metro exposure to the equity market, and find that following positive wealth shocks, local prices rise and future returns decline. These novel mechanisms drive spatial variation in prices and wealth accumulation, and potentially affect the cost of home-ownership and the incentives for housing supply.

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

Housing prices are well understood to be a function of local demand for housing, stemming from local labor market conditions and amenities, and of local housing supply, which depends on local geography, regulation, and other costs, as well as the expectations thereof. The availability of capital and the characteristics of landlords, however, have rarely been considered as a determinant of local prices. This is due, in part, to reasonable expectations that return differences should be arbitraged across space, and to the convenience of the popular 'absentee landlord' assumption[1]. However, a growing literature documents persistent spatial differences in returns to housing, (Jordà et al., 2019; Eisfeldt and Demers, 2021; Samuels, 2024; Damen et al., 2025), and US household real estate investors, in particular, are highly home-biased - 80% live in the same metro area and 30% live in the zip-code of the units they own. These landlords also own most rental housing units - 82% of properties and 53.5% of units, and hold even higher shares in some market segments, including older and smaller properties, suggesting their behavior may impact local rental prices and returns. However, neither the behavior nor the impact of these investors on housing markets is well understood.

In this paper, I propose a framework for understanding and testing the role of household landlords who invest in rental estate in determining local prices, returns, and rent-price ratios. In particular, I find that areas in which local residents have higher wealth, or a higher propensity to invest in real estate, have higher rental prices, and lower rental returns and rent-price ratios. I consider risk-averse investors who choose whether to become investors, and how much of their wealth to invest in real estate. My model nests an asset pricing framework in a spatial equilibrium setting and generalizes a number of mechanisms through which investor behavior can drive local returns, including rent growth expectations (e.g. Hilber and Mense (2021)), risk (as in (Eisfeldt and Demers, 2021; Amaral et al., 2021)) and investor characteristics, such as wealth (Levy (2021)), race (as in Diamond and Diamond (2024)), and a 'proclivity' for real estate, to drive local differences in prices, returns, and rent-price ratios. In the data, tests of both the wealth and proclivity channels suggest a substantial role for investor wealth and investment proclivity in shaping price variables. These mechanisms could affect real outcomes such as tenure choice, housing supply, renter welfare, the home ownership rate, and the entrance of institutional investors (Greenwald and Guren, 2021; Büchler et al., 2021; Vanhapelto, 2022; Coven, 2024; Chang, 2024).

This paper makes three contributions. First, it documents new facts about US household real estate investors, showing that they are heavily home-biased, tend to be less educated,

[1]For an exception, see Ater et al. (2021).

and are influenced by cultural norms, distinguishing them from households investing in other assets (Campbell, 2006). Second, it develops and tests a novel model in which residents endogenously become local real estate investors and their characteristics shape local returns and prices. The reduced form evidence supports the model's predictions, revealing significant effects of investor wealth and residents' propensity to invest on rent-price ratios and returns. By embedding an asset pricing framework within an urban setting, the model bridges the gap between urban economics, which treats housing as a consumption good, and real estate finance, which views it as an asset (Glaeser and Gyourko, 2007). In doing so, it relaxes the 'absentee landlord' assumption and explores its implications. Finally, the paper contributes to the asset pricing literature by analyzing a setting with extreme market segmentation, where both returns and portfolio shares are endogenous, demonstrating the critical role of investor wealth. The model aligns with demand-based asset pricing (Koijen and Yogo, 2019) but is micro-founded in the canonical portfolio allocation framework.

I begin with a discussion of the households that invest in real estate, and document novel facts about them. These investors are an important part of the market, and are particularly important in supplying rentals of single-family houses and small and medium multifamily – structures ranging from duplexes to 20 units. They are also very home-biased – about 30% live in the same zip code as their property, 50% within 10km, and more than 80% within 75km. Owners of multiple properties, or multifamily properties, are even more home-biased. Landlords are therefore both local and important for housing markets, both across and within cities.

Household landlords do not follow traditional investor patterns of participation. Participation is not as concentrated among the wealthiest, relative to other private businesses, and many investors come from the middle and upper-middle classes. Conditional on wealth, these investors tend to be *less* White, and significantly *less* educated, in sharp contrast to equity investors. Household investors are also more likely to report high-risk aversion compared to equity investors, despite owning undiversified assets with significant idiosyncratic risk. Investment also seems influenced by ancestry, independent of race and socioeconomic controls, at least at the local level. I show that local ancestry-based real estate affinity, estimated from local ancestries' probability of *working* in real estate, predicts local investor shares.These results suggest investors are a nontraditional group, operating locally, and driven by various idiosyncratic considerations that could be informed by an affinity for real estate and an under-appreciation of, or ability to manage, its risk.

I next develop a model of housing markets in which investors self-select out of the local population, based on their wealth and an idiosyncratic 'ability' to operate real estate. The supply of housing is fixed, and rent and the price of owner-occupied housing are set in spatial

equilibrium. Financial friction creates demand for rental housing and drives a wedge between the price of owner-occupied and rental properties. Investors solve a canonical portfolio choice model (Campbell, 2017),[2] given expected returns and their variance, and price rental housing. In equilibrium, enough local residents must choose to become investors and supply enough capital to provide these rental units. The model is static and features uncertainty in the form of a local income shock. that affects rent, and hence investors' returns. As investors are also exposed to this income shock, returns are correlated with investors' non-wealth consumption.

The combination of local investors and inelastic housing (asset) supply creates implicit local capital markets. In these markets, returns adjust to clear the level of capital such that enough rental units are supplied, and there is no excess capital. Equilibrium returns are therefore a function of the intersection of capital demand (driven by the share of renters) and of capital supply, which is driven by market and investor characteristics: residents' propensity to become investors, those investors' wealth, and rent variance. Wealth inequality matters both through the share of renters and of investors, and can either increase or decrease local returns and investor participation.

To my knowledge, my model is the first to nest risk-averse local investors, which price local returns within an urban framework, where rents and prices must obey spatial equilibrium constraints, and proposes a novel way to solve models with home-biased investors and fixed-supply assets.[3] As a result, and contrary to most urban models (in which landlords are assumed absentee), in this model investors' characteristics depend on the local residents and are important for local prices, particularly of rental property, and local investors' return vary across space. This setting can also generalize mechanisms from both the real estate literature, such as the market risk suggested (Eisfeldt and Demers, 2021; Amaral et al., 2022), as well as the average investor wealth, as in Levy (2021).[4] It also suggests a novel mechanism, the propensity of the local population to become investors, and wealth inequality.

To test the model's predictions I construct a dataset with-property-level rent and price estimates based on the 2013 'DataQuick' dataset, encompassing the near universe of US residential properties. I focus on properties that are likely rental properties and adjust for quality differences across space and across the rental-owner occupied market segments (Glaeser and Gyourko, 2007). These data include the location and physical characteristics (such as building type, square footage, etc.) of housing units, as well as sale price and

[2]While some of the investors characteristics shown above could be explained by behavioral, or information, or trust mechanisms, I propose a model of rational investors, to show that such a model can still capture the investors' behavior and their effect on local markets.

[3]See also contemporaneous work in Johnstone and Lin (2023).

[4]The model can allow for any source of return risk, but microfounds rent growth volatility as a mechanism for such risk, and broadens the effect of wealth discussed in Levy (2021) to depend on its full distribution across residents, and the propensity of investors to deploy their wealth in the local market.

assessed values. I estimate rents and prices for the same units using hedonic regressions, deduct real estate taxes, and calculate ratios for all units that are not occupied by their owners. I then compute average rent price ratios at the zip and metro levels.[5]

The data also allow for linking properties to their owners' location and own housing unit, based on the name and mailing address of the owner, listed on each property's deed. I use these data in two ways. First, I calculate the distance between each property and their owner's mailing address. Second, I identify that property in the original DataQuick data, which allows me to determine if that property is a residential or commercial address, filtering the sample to owners who operate real estate from their own residence - a novel and direct measure of 'household' investors, without relying on textual analysis of owner names. Third, I estimate the value of investors' own homes, which I argue can be used to proxy for investor wealth. I confirm the tight relationship between home value and wealth in the Survey of Consumer Finances, where the value of the home explains 58% the net wealth of housing. The result is a novel dataset that links characteristics and rent-price ratios of properties owned by household investors and a proxy of the investors' wealth and their home's distance to the property.

I test the effect of landlord capital supply in a static setting on rent-price ratios[6]. At the property level, I find that investor wealth is strongly negatively correlated with rent-price ratios, but this is driven by the average wealth of nearby investors, not an investor's own wealth. In an appendix, I further confirm this result by including investor fixed effects or controlling for the investor's own wealth, and by using a rich set of property-level controls, such as property age. I find consistently large negative effects of ;pca; average wealth that dominate the investor's own wealth effect.

Second, I test the model's comparative statics with market-level regressions. I define local sub-metro markets as clusters of zip codes, which exploit within-city variation but include a significant share of local investors (these clusters contain, on average, 50% of investors owning property within them). I regress local rent-price ratios on investor wealth, the shares of renters and investors, and a set of local controls, including homeowner prices that capture differences in current and future local amenities, while controlling for city fixed effects.

I instrument for the share of investors using ancestry-based for real estate.This instrument

[5]In estimating prices I use property transactions and run regressions at the county level, and control for polynomials of the property's assessed value, which accounts for a property's value rank, as in Diamond and Diamond (2024). Rent estimates aggregate back to ACS zip-level averages by construction. The aggregation of unit level ratios of rents and prices, rather than dividing averages, removes a downward bias that depends on within zip variation. My procedure therefore maintains consistent estimates for rents and prices, and minimizes this bias to the extent possible through observed data. I discuss this further below.

[6]Rent-price ratios are static measure of returns whose spatial dispersion signals return differences, see Samuels (2024)

builds on recent research showing that cultural heritage influences attitudes toward real estate ownership (Gorback and Schubert, 2023; Deng et al., 2021). I also instrument for investor wealth using local average equity wealth calculated from IRS data (Chodorow-Reich et al., 2021), to purge potential mechanical correlation of the wealth measure with local housing prices while ensuring that the regressor reflects investor wealth.[7]

I find strong evidence supporting the model's predictions in both the OLS and IV specifications: investors' wealth and the share of investors negatively predict rent-price ratios, while the share of renters positively predicts them. The elasticity of local rent-price ratios with respect to the share of investors, whose causal effect is identified using the ancestry-based IV, is estimated at -0.665. The coefficients for the share of renters and investor wealth are 0.383 and -0.836, respectively. These findings align with the model's comparative statics and confirm the important role of capital markets in shaping the valuation of local rental properties.

My second analysis uses a metro level panel, to estimate the causal effects of wealth shocks on local returns. I benchmark FHFA price and BLS rent series with DataQuick rent and rent-price ratios to compute returns. I use a shift-share design with equity shocks, based on metro equity position data from Chodorow-Reich et al. (2021) and national equity shocks to construct wealth shocks. This generates wealth shocks that are exogenous to local circumstances, especially when I control for time and city fixed effects. The results consistently show that local exposure to equity wealth lower subsequent rent-price ratios and returns, though the former is not statistically significant with robust standard errors. Return effects are statistically significant.

The paper proceeds as follows: Section 2 documents novel facts about US investors, Section 3 discusses the data, Section 4 presents the model, and Section 5 conducts the reduced form test.

2 Household Real Estate Investors

An overwhelming share (81.8%) of US residential rental properties, and a majority (53.5%) of units are held by household investors, the majority of which tend to live near the property they own. In fact, more than 75% live within 50 km of their units, and slightly more

[7]From the view of the model, these IVs identify the comparative statics for investor wealth, and share and renter share. Wealth is assumed to be exogenous, and its IV addresses measurement error. The share of investors, endogenous in the model, is identified by the ancestry-based IV, while the share of renters, determined by spatial equilibrium and liquidity constraints, is exogenous to the asset market. Homeowner prices capture differences in expected income and amenities, while risk and model misspecification are captured as residuals, given the linear regression framework.

than 40% live within the same zip-code as their units. The household investor population, which is made up of approximately 6.2-9% of households, tends to be concentrated among populations that are less likely to invest in other assets, with a higher prevalence of immigrants, non-white and less educated households. The propensity to invest increases with wealth, but also appears to depend on demographic and idiosyncratic factors, which together create spatial differences in the potential investor pool. The emerging picture is that of a specialized, local class of investors that play an important role in housing markets. In this section, I provide statistics on the importance, location, and attributes of these investors.

Defining Household Investors. 'Household' investors are defined in various ways in the literature and media, but generally refer to households using their own equity and consumer credit to purchase and manage rental properties. Typically, these investors are not real estate professionals and hold moderate real estate portfolios. (Some may employ professional property managers.) In each dataset, I attempt to construct a definition that aligns with this group.

I identify such household investors in DataQuick based on owners' mailing address. I consider a nonowner-occupied property to be held by an investor household if the mailing address listed on the property's deed is itself a residential address, and the listed owner name for that mailing address does not indicate a corporate entity.[8] Excluded are properties operated from an office, and is likely a conservative definition, as some household investors own both their rental and their property under tax savings vehicles, such as an LLC, which receive mail at a lawyer's office, or utilize PO boxes, which cannot be identified as residential addresses. Investor households are the set of households that own at least one such property. The data construction process and sample are detailed in Section 3, where I also discuss vacation homes.

For some analyses, I use the Survey of Consumer Finances (SCF) and Rental Housing Finance Survey (RHFS). In the SCF, which surveys households, I define participants' real estate holdings as 'household investor' only if they are held directly, and not through REITs or an interest in a private company. In the survey, households are explicitly asked about such property.

The RHFS, which samples properties, classifies owner entity. I consider properties as held by household investors if their owners are classified as either 'individuals' or 'real estate trusts', which are typically tax savings vehicles employed by nonprofessional investors. I exclude LLCs, which are classified in the RHFS together with LCs and LPs, even though

[8] I classify a corporate name based on textual permutations that are common for corporate or non-profit holders of real estate, including 'Bank', 'Agency', 'LLC', 'Foreclosure', 'Fund', etc.

some are likely tax-saving vehicles of household investors.[9]

2.1 The Importance of Household Landlords

Table 2.1 demonstrates the importance of household ownership of rental real estate, both for rental markets and for the households themselves. Panel A focuses on the housing market, demonstrating that (as of 2015) such households own 81.8% of all rental properties and 53.5% of units, and own even larger shares of smaller and older properties. Panel B documents that investment in these assets is widespread, about 8.8-11 million households participate, and is on par with the share of households who own private businesses. Conditional on participation, investors hold on average 2.3-2.7 units worth $450,000-490,000, which represent on average 25% of their wealth – in line with private business assets, but lower than the portfolio share owner-occupiers allocate to their own residences.

Panel A utilizes the 2015 Rental Housing Finance Survey (RHFS), which surveys a representative sample of rental properties and includes information on the property and its owner. (I use the 2015 vintage since the 2012 vintage includes only multi-family properties.) Households own 17.77 million properties out of a total stock of 21.72 million (81.8%), and 25 million out of 46.7 million units (53.5%), suggesting that their investment is a critical component of capital supply to the rental sector.

The importance of household investors is even greater among smaller properties and older units. Households own 84.1% of single units (these are primarily single family homes, though some are condominiums), and 79.2% of duplexes. Households are an important part of small multifamily properties as well: they own most units in the 3-4 and 4-9 units categories, and more than a third of the 9-19 and 20-49 units categories. Only in the largest properties (50+ units) are household investors a small minority of owners.[10] Households are similarly important in older properties. Building construction year is only sparsely available in the 2015 RHFS, but is available for 2012 multifamily properties. Among these properties,

[9]In DataQuick, I observe properties whose listed owner's name includes 'LLC', but the listed mailing address is a residential address, who's owner's name is a not an LLC, or other name that would indicate a corporation. I observe 345,552 properties held by household investors (based on mailing address) who's owner's name includes 'LLC'. Based on a manual review of a sample of the data, many of these LLCs are often similarly to the name of the mailing address's owner. For example, the mailing address of a property owned by 'Smith Properties LLC' will be owned by 'John and Joe Smith', suggesting that some LLCs are in fact household investors.

[10]Units in property should not be conflated with units in structure. 50+ unit properties are common (30.7% of units) – the same is not true for 50+ structures (11.9% of units in the ACS, 9.1% in the AHS). This suggests that most large properties consist of multiple smaller buildings – buildings that, were they their own property, households are much more likely to own. This suggests a sorting of similarly sized buildings into 'groups' (properties) that are differentially owned by households and corporate investors, which is left for further research.

households mainly hold older units: 59.9% of units built before 1920, but only 10.1% of properties built between 2007 and the end of 2012. These patterns reflect a tendency of household investors to either own older and larger properties, or smaller and newer properties.

Table 1: Importance of Household Rental Real Estate Ownership

Panel A: Importance of Hh Investors to Rental Markets (RHFS)

Segment	Properties (1,000s)		Units (1,000s)	
	Stock	% Hh' Owned	Stock	% Hh' Owned
All Rental	21,724	81.8%	46,711	53.5%
Units in Property (2015)				
SFH/Condo	18,981	84.1%	18,981	84.1%
Duplex	1,282	79.2%	2,564	79.2%
3-4	702	57.4%	2,340	57.8%
4-9	388	61.3%	2,452	62.0%
9-19	138	47.7%	1,946	45.8%
20-49	128	38.5%	4,081	36.8%
50+	70	11.9%	14,345	12.3%
Construction Year (MFH Only; 2012)				
Before 1920	353	77.8%	1,781	59.9%
1920 - 1940	374	83.0%	1,875	55.4%
1940 - 1960	374	88.3%	1,844	57.0%
1960 - 1970	159	67.0%	3,761	21.6%
1970 - 1980	207	69.3%	3,668	23.0%
1980 - 1990	169	65.6%	3,044	21.4%
1990 - 2000	87	57.2%	1,793	17.1%
2000 - 2007	80	53.6%	1,840	12.1%
2007 - 2013	34	83.2%	762	10.1%

Panel B: Importance to Investors (SCF)

Asset Class	Participation	Cond' Mean ($k)	Cond' Portfolio %
Real Estate Assets			
Own Housing	0.65	$288	0.57
Vacation Home	0.05	$171	0.16
Other Res' RE	0.09	$442	0.26

Continued on next page...

(Continued from previous page.)

Segment	Properties (1,000s)		Units (1,000s)	
	Stock	% Hh' Owned	Stock	% Hh' Owned
Construction Year				
Rental Res' RE	0.06	\$490		0.24
(Lower Bound)				
Alternative Assets				
Equity	0.48	\$261		0.18
Private Business	0.12	\$1,034		0.25

Note: This table reports statistics on the role of rental real estate in housing markets and investor portfolios. Panel A uses the Rental Finance Housing Survey (RFHS) to show household participation in rental segments, using 2015 data for unit counts and 2012 data for construction year. Property and units counts reflect all properties, not just those owned by households. The 2012 RFHS excludes single-unit properties, while the 2015 RFHS has limited construction year data. Households are defined as individual or estate owners (excluding LLCs). Units in property refers to the number of units in each property rather than individual buildings (this differs from ACS definitions), and construction year refers to the oldest building in a property, and is unavailable for single-unit properties. Panel B uses the Survey of Consumer Finances (SCF) to analyze household rental investors' participation and portfolios, focusing on their first two properties. Additional properties (held by 15.8% of investors) are not detailed. Imputed rental real estate, defined as multifamily properties or properties with rental income, provides a lower bound for participation.

Panel B presents statistics on the importance of rental real estate to household finance. I use the SCF to compute the participation rate in this asset class, as well as the average holding value and potfolio weight, and then compare it to other important assets. The table documents a 6.2-9% participation rate, equivalent to 8.8-11 million investor households, and comparable to the share of households that own private businesses but far lower than participation in equity markets of home-ownership. The portfolio weight of rental real estate for households who invest in it is 24%, also similar to private businesses, but is smaller in nominal terms. This suggests that for investors in it, rental real estate is an important asset class, but on average one of multiple assets held in the portfolio.

The range in participation share stems from an apparent inconsistency in reporting rental payments. While 9% of households reports owning residential real estate other than their home and vacation/recreation homes, only half of these report a positive rent, even among owners of multifamily housing. In Appendix C, I argue that this is likey due to under-reporting of rent payments. The 6.2% is a lower bound for participation, defined as either reporting positive rent or owning a multi-family property. However, this indicator is conservative and likely underestimates the number of investors, as some owners of condominiums or single-family homes may also under-report rent.

2.2 Home-bias Among Household Landlords

Real estate investors are extremely segmented geographically - overwhelmingly investing in their own cities and neighborhoods. Owners of multiple properties, or of properties with multiple units, are even more segmented than owners of single family houses and condominiums.

To measure segmentation I use DataQuick's mailing address records for the property-investor matched sample of properties. I calculate distances between the property's zip code's centroid and that of its mailing address.

Figure 1 plots the CDF of this distance. The plot demonstrates that real estate investing is considerably segmented with investors having a significant home bias. The CDF rises very sharply with distance until about the 50km mark before reaching an inflection point and flattening out. Owners of all property types tend to live very close to their properties - across most of the US 80% of owners live within 50km of their units. Segmentation holds even within neighborhoods - about 30% of owners live in the same zip code as their units, and 45%-60% live within 10km. Single-family residences and condominium owners are somewhat less segmented, with 75% and 60% of them, respectively, living within 50km of their units.[11] Table 8 in Appendix A presents the CDF for household investors at various thresholds.

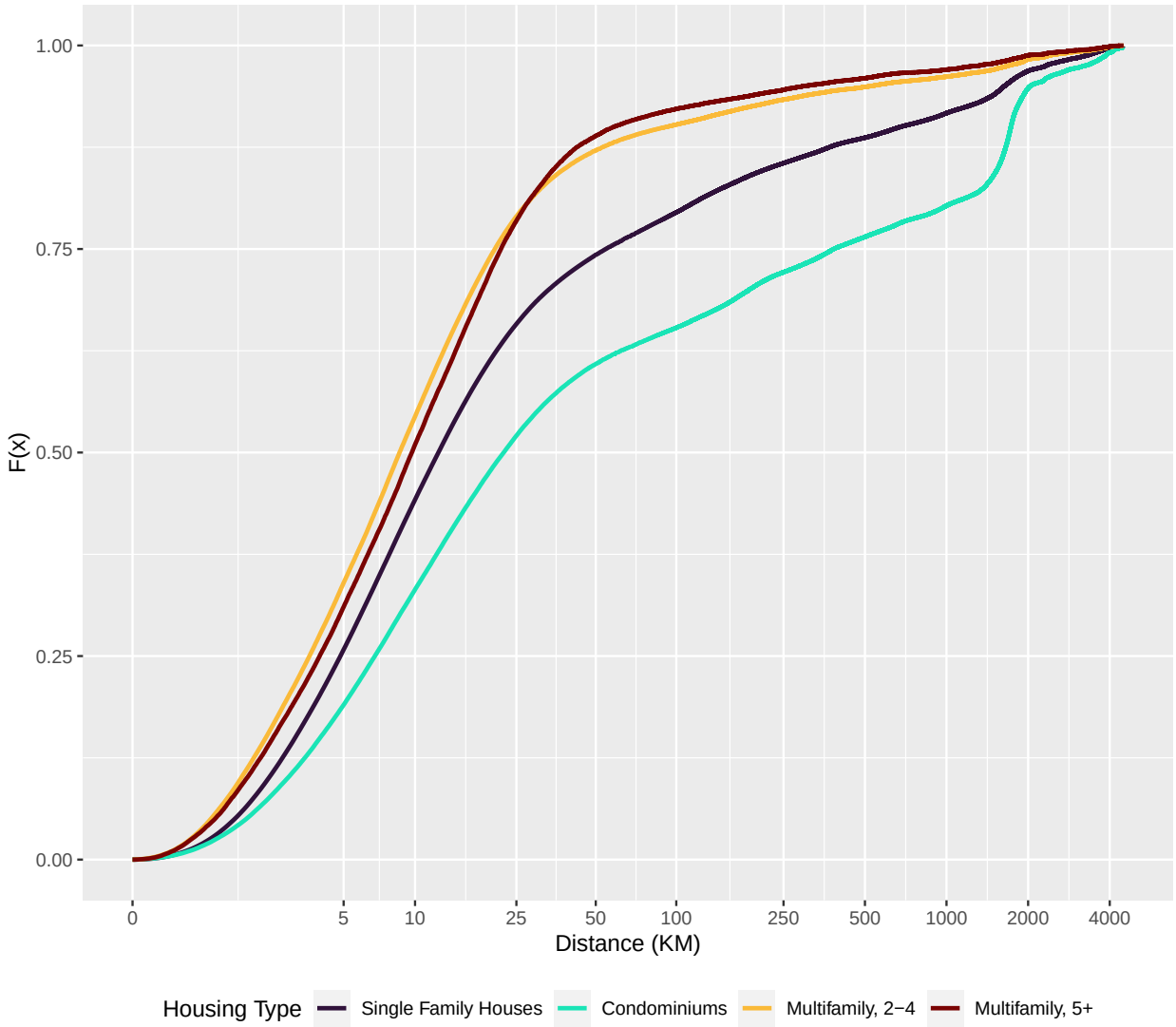
2.3 Characteristics of Real Estate Investors

Unsurprisingly, real estate investors tend to be wealthier than the general population, though a sizable share, about a third, come from below the 85^{th} wealth percentile. Conditional on wealth, these investors differ from those in other asset classes: they are generally *less* white and *less* well-educated, and more likely to report higher risk-aversion than equity investors. Additionally, areas with more immigrants, and descendants of immigrants from particular regions, tend to have more such investors.

Wealth Distribution. I turn to analyzing where investors fit in the wealth distribution. Using the SCF I reproduce Figures 2 and 3 in Campbell (2006) to show the participation and portfolio decisions of households with different levels of wealth. Figure 2 shows that participation in real estate - both in the original and refined definitions - rises steadily from the 40^{th}, and more rapidly from the 85^{th} wealth percentiles. However, the participation of middle class Americans is not negligible - about a third of rental real estate investors come from below the 85^{th}. Notably, whereas homeownership for one's own residence rises very

[11] There's a sharp uptick in condominium ownership between 1,500 and 2,000 km, which is the distance between Florida and major population centers in the Northeast and Midwest. Florida's market, particularly condominiums, appear to be less segmented than the rest of the nation.

Figure 1: CDF of Distance Between Investor and Property



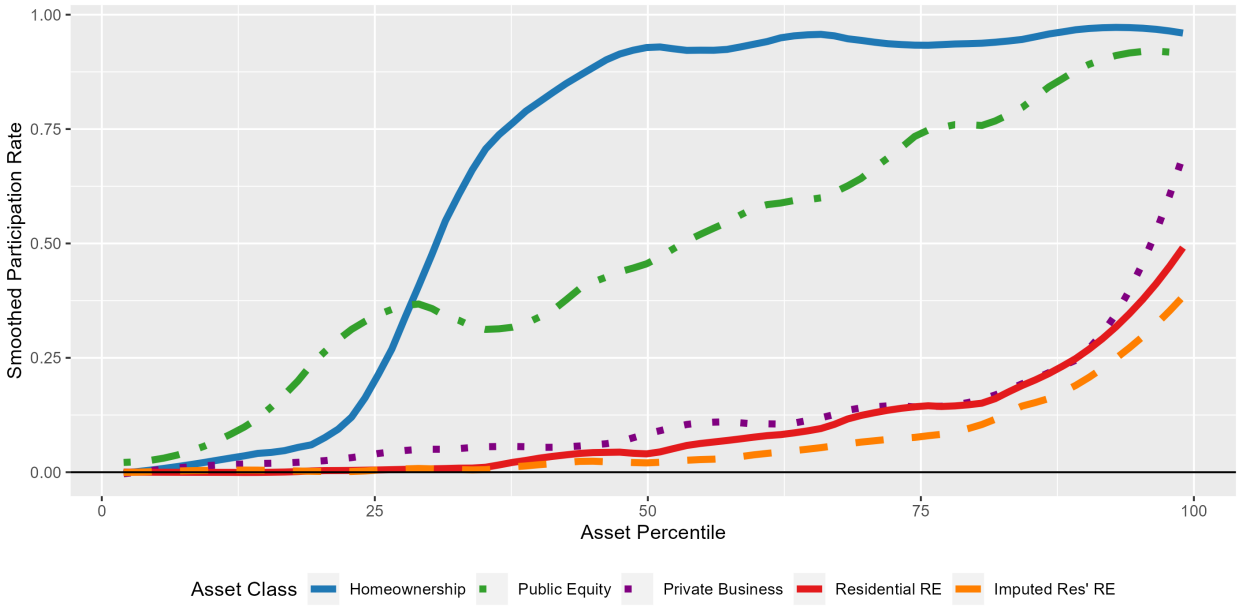
This figure plots the CDF of distance between each property its investor's mailing addresses, in log scale. Color represents CDF for four types of buildings: single family houses, condominiums, small multifamily (2-4 units), and larger multifamily (5+ units). Distance is calculated using zip code centroids. Sample is limited to non-corporate, private investors. .

rapidly, starting at the 20^{th} percentile, participation in real estate investing starts only in the 40^{th} percentile and rises very gradually.

These findings suggest that while the propensity to become a landlord increases with wealth and requires a certain wealth threshold, participation includes a broader, more diverse, swath of society, which may be driven by non-pecuniary considerations.

Demographic Characteristics. Controlling for wealth and income, real estate investors tend to be less White and less educated. Table 2.3 presents regressions of participation

Figure 2: Participation Rates for Real Estate, Other Assets, by Wealth



Participation rates for select assets across the wealth distribution. Residential RE corresponds to any household reporting residential real estate in addition to their own home, or for recreational purposes. Imputed Res' RE corresponds to the imputed rental real estate investor and includes households who received rent for residential properties, or own a 2+ unit. Participation kernel smoothed using local polynomial regression, span=0.3

and portfolio shares for investors in three asset classes - equities, private business, and real estate; following the structure of Table 1 in Campbell (2006). The left panel of Table 2.3 reports logit regressions of participation on demographic, education, and income and wealth variables. The second panel reports OLS regressions of portfolio shares on the same variables. The first column in each panel reports equity, the second private business (as in ibid,) and the third reports real estate.

Table 2: Demographic Determinants of Real Estate Investment

	Binomial Model			OLS		
	Whether Hh Owns			Cond' Portfolio Share		
	Dependent variable:					
	Equity	Private Bus'	Res' RE	Equity	Bus'	Res' RE
	(1)	(2)	(3)	(4)	(5)	(6)
	Demographic Variables					
Married	−0.04	−0.20	−0.13	0.06***	0.06*	0.09*
	(0.12)	(0.16)	(0.24)	(0.01)	(0.03)	(0.04)
Female	0.46***	−0.26	0.02	−0.04***	−0.09**	−0.11**
	(0.13)	(0.19)	(0.27)	(0.01)	(0.04)	(0.05)
Age	0.01	0.08***	0.06*	0.002	−0.01*	−0.02***
	(0.02)	(0.03)	(0.03)	(0.001)	(0.005)	(0.01)
Age^2	−0.0003**	−0.001***	−0.001*	−0.0000	0.0001	0.0002***
	(0.0001)	(0.0002)	(0.0003)	(0.0000)	(0.0000)	(0.0001)
# Kids	−0.13***	0.002	−0.25***	−0.01***	0.01	−0.001
	(0.04)	(0.05)	(0.08)	(0.004)	(0.01)	(0.01)
Race - Black	−0.13	0.09	1.05***	−0.03**	−0.004	−0.06
	(0.13)	(0.20)	(0.27)	(0.01)	(0.04)	(0.04)
Race - Latino	−0.66***	−0.25	0.27	−0.01	0.05	−0.11*
	(0.16)	(0.22)	(0.37)	(0.02)	(0.04)	(0.06)
Race - Other	−0.40***	−0.08	−0.03	−0.02	−0.005	0.03
	(0.14)	(0.18)	(0.23)	(0.01)	(0.03)	(0.04)
No Risk Tolerance	−0.87***	−0.35***	−0.27	−0.05***	−0.01	0.05*
	(0.08)	(0.13)	(0.21)	(0.01)	(0.03)	(0.03)
	Education Variables					
High School	0.46***	0.32	−0.48	0.03*	−0.03	−0.07
	(0.14)	(0.24)	(0.35)	(0.02)	(0.05)	(0.05)
College	0.93***	0.06	−0.84**	0.05**	−0.09*	−0.04
	(0.17)	(0.26)	(0.37)	(0.02)	(0.05)	(0.06)
Graduate	0.77***	0.17	−0.72*	0.06***	−0.16***	−0.06

Continued on next page...

(Continued from previous page.)

	Whether Hh Owns Asset			Cond' Portfolio Share		
	Dependent variable:					
	Equity	Private Bus'	Res' RE	Equity	Bus'	Res' RE
	(1)	(2)	(3)	(4)	(5)	(6)
	(0.19)	(0.28)	(0.38)	(0.02)	(0.05)	(0.06)
	Income & Wealth Variables					
Income	4.50^{***}	-1.76	3.33^{*}	0.11	0.01	0.12
	(1.29)	(1.14)	(1.92)	(0.11)	(0.21)	(0.37)
Income^2	-0.17^{***}	0.06	-0.14^{*}	-0.004	-0.002	-0.01
	(0.06)	(0.05)	(0.08)	(0.005)	(0.01)	(0.02)
Wealth	0.53^{**}	-0.16	5.82^{***}	-0.26^{***}	-0.56^{***}	0.20
	(0.25)	(0.25)	(1.55)	(0.03)	(0.09)	(0.26)
Wealth^2	-0.003	0.04^{***}	-0.17^{***}	0.01^{***}	0.02^{***}	-0.01
	(0.01)	(0.01)	(0.06)	(0.001)	(0.003)	(0.01)
Constant	-34.85^{***}	4.47	-68.84^{***}	1.07^{*}	4.11^{***}	-0.62
	(6.77)	(6.09)	(12.60)	(0.57)	(1.20)	(2.54)
Observations	5,046	5,046	5,046	2,639	728	361
Adjusted R^2				0.08	0.13	0.14
Log Likelihood	$-11,493$	$-7,013$	$-3,643$			

Note: This table presents regressions of participation and portfolio shares in three assets: equities, private business, and residential real estate. All data based on the SCF. Columns (1) through (3) regress participation using Binomial regressions. Columns (4) through (6) regress portfolio shares, among participants, using OLS. All implicates in the data are included and the number of observations indicates unique respondents. Standard errors are adjusted to correct for the five implicates. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Black and Latino households are much more likely to invest in real estate, with the former almost three times as likely to do so. *Less* educated Americans are also *more* likely to invest in real estate: college education lowers the odds-ratio of investment by 56.8%, and college education by 51.3%. Even high school graduates are less likely to be investors, relative to high school dropouts, by 38%, although this effect is not statistically significant. This contrasts with well-established patterns in equity and private business investments, where generally White and better educated Americans are more likely to participate.

Another interesting difference between investment in equities and residential real estate is that individuals reporting a high degree of risk aversion are significantly less likely to participate in the former, but not much less likely to participate in the latter. Specifically, reporting no risk tolerance reduces the odds-ratio of investing in equities by 58% and in real

estate by only 26%, the latter being statistically insignificant. Furthermore, those reporting higher risk aversion allocate a *larger* share of their portfolios to real estate but less to equities, which is remarkable given the significant idiosyncratic risk in owning any particular real and illiquid asset.

Ancestry. Recent research has highlighted that ancestry significantly influences the investment decisions of U.S. individuals, with studies by Burchardi et al. (2019) and Badarinza et al. (2022) establishing a connection between ancestral ties and economic behavior. Furthermore, Gorback and Schubert (2023) find that ancestry affects U.S. residents' propensity to become homeowners, at the expense of other assets, with patterns linked to homeownership rates in their countries of origin.

Building on these insights, my analysis demonstrates that ancestry, independent of race and being Hispanic, plays a distinct role in shaping participation in real estate markets. This relationship motivates an instrumental variable, on which I elaborate in Section 5.1.2. Since detailed data on both ancestry and investment in real estate are not available, I first estimate the proclivity of different ancestries to work in the real estate licensing industry (NAICS code 531), controlling for individual characteristics and metro fixed effects. [12] I then show that neighborhoods where the average resident has a higher ancestry-induced proclivity for real estate, also have higher shares of investors out of the local populace.

I pool the 1% Census and 1-year ACS samples from 2000 to 2019, and estimate a regression of employment in real estate leasing on year, metro, race, education, and ancestry fixed effects. Confidence intervals for the ancestry effects are calculated using bootstrapped estimates based on successive 20% samples. U.S.-derived ancestries (e.g., 'American', 'United States', or various States) are grouped and used as the reference category. [13] The estimated fixed effects capture each ancestry's proclivity of working in real estate.

I find that ancestry is highly statistically significant in predicting employment in real estate, with an F-statistic of 42 when all control fixed effects are included. Estimated fixed effects are economically important, and range from -0.004 to 0.015, with a standard deviation of 0.002, relative to a baseline probability of 0.009. Out of 150 ancestries, 27 are statistically associated with a lower probability of engaging in real estate compared to individuals reporting U.S.-derived ancestry, while 30 predict a higher probability statistically significantly. These results remain qualitatively robust across various ancestry classifications, including

[12] While employment in this industry involves a different risk and payoff structure compared to property ownership, the underlying assumption is that cultural tendencies toward real estate would influence both forms of participation.

[13] This group, which comprises 21% of the sample, serves as a baseline for comparison, reflecting a more assimilated U.S. culture. Ancestries such as 'African American' are not included in this group. In alternative specifications, I also aggregate ancestries into broader categories (e.g., Bavarian into German, Sicilian into Italian). More details on data construction and estimation are available in Appendix X.

sub-national categories.

Figure 4 in the Appendix maps those fixed effects that correspond to modern nations, showing that certain regions, such as Southeast Europe and the Balkans, exhibit large and statistically significant positive estimates, while others, like the Arab world, are associated with negative estimates. These regional trends are reinforced by non-mapped ancestries which align with the broader regional patterns, such as 'Yugoslavian' (positive effect) and 'Other Arab' (negative effect).

Ancestry's effect on proclivity to work in real estate extends to owning it as an investment. I compute zip level average affinity for real estate, by summing the estimated fixed effects across residents of different ancestries, and dividing by the total number of residents.

Tables 3 presents regressions of the resulting measure, called average ancestry real estate proclivity (AREP), on the share of residents who invest in real estate in each zip. Across specifications, AREP positively and significantly predicts the investor share. This positive association remains when controlling for local density, racial make-up, rents, prices and rent-price ratios. Controlling for all these, a one-standard deviation increase in AREP predicts a 0.003 pp increase in the investor share, relative to a baseline average of 0.062.

These findings suggest that ancestry influences the likelihood of individuals working and investing in real estate, likely through cultural or historical factors that vary across ancestry-regions. Importantly, this effect is independent of race, education, or geographic clustering of ancestries in specific cities. Furthermore, the statistical significance, robustness of the results, as well as their regional patterns, make it highly unlikely that they are driven by noise.

3 Data

In this section, I describe the data and sample construction, including how I estimate investor and market wealth. I use micro-data on US properties and property transactions to estimate the value and rent of each unit. I then estimate investor and local wealth by matching each investor to their own residence and using the tight relationship in the SCF between housing value and wealth. This matching also allows for calculating the distance between investors and their properties. I augment this data with county-level equity wealth estimates based on Chodorow-Reich et al. (2021).

The process is as follows: I begin by estimating unit-level prices and rents based on the location and physical characteristics of these units. I then match properties to their owners, based on mailing address. I restrict the sample to 'small' investors - household owners of real estate that is less likely to be recreational second homes. Finally, I aggregate these variables

Table 3: Net Ancestry Effect & Investor Share

Dependent Variable:	Share Investors						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average AREP	0.024	0.107***	0.103***	0.087***	0.049***	0.046***	
	(0.018)	(0.020)	(0.016)	(0.017)	(0.017)	(0.017)	
Average AREP z-score							0.003***
							(0.0009)
log(Rent/Price)				-0.122***		0.060*	0.060*
				(0.016)		(0.031)	(0.031)
log(Rent)					-0.0005	-0.004	-0.004
					(0.003)	(0.003)	(0.003)
log(Price, Rental Units)					0.021***	0.027***	0.027***
					(0.002)	(0.005)	(0.005)
Controls							
Metro FE	-	✓	✓	✓	✓	✓	✓
Race & log Density	-	-	✓	✓	✓	✓	✓
Fit statistics							
Observations	9,646	9,646	9,646	9,646	9,495	9,495	9,495
R^2	0.002	0.244	0.355	0.377	0.402	0.405	0.405
Within R^2		0.018	0.162	0.191	0.226	0.229	0.229

Note: This table presents regressions of the share of residents who are real estate investors, on ancestry-derived proclivity to work in real estate, controlling for local characteristics and metro fixed-effects. A unit of observation is a zip-code. Average AREP (ancestry real estate proclivity) is each zip's average ancestry induced probability of working in the real estate licensing industry. Each ancestry's probability is computed as fixed effect in a regression of a real estate industry dummy on individual, metro, time and ancestry fixed effects in ACS and census micro-data. Zip averages are computed across all residents, based on tallies of their ancestries. The unit of the Average AREP variable is pp probability, and has a cross-zip average of 0.034 and standard deviation of 0.05. Across all specifications, the dependent variable is the the investor share, computed as the share of residents in each zip code who own investment real estate, based on DataQuick data. The average of the dependent variable is 0.062. Specifications progressively add controls - metro fixed effects (column 2); race and log-density (3); rent-price ratios (4,6); log rents and rental unit prices (5,6). Column 7 regresses z-scores of the average AREP. Standard errors are clustered at the metro area level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

to various geographic units.

Estimating Rent and Prices

The main data set I use is the 2013 DataQuick dataset, which includes the near-universe of US real estate properties at property level. The includes information on the location and physical characteristics of the property, as well as the owner's name and mailing address.

Throughout, I filter out units with missing data for the number of units in the building, for which there is no rooms, bedroom, or square footage information, or for which there is no average zip code rent. When large multi-family structure include only aggregate information (e.g. total bedrooms) I divide these equally across units within the building.

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 average rent in i's zip-code at t adjusted for that unit's size, age, and building characteristics. The value of these physical characteristics is assumed constant across space, and is estimated for each t based on the American Housing survey.

Equation (1) details the specification: for a survey 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. 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.

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

I assume that conditional on city 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 an amalgamation of the location premium of the sampled units, which under the model in Equation (1) 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 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 equal the average zip rent as reported by the ACS, as in Equation (2). I invert this relationship to back out $\{\mathcal{L}_{jt}$, 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. These estimates can be thought of as the average zip rent, adjusted for the average value of the unit's physical characteristics.

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

I estimates 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 property removed or significantly altered.

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 hedonic regression for all transacted units in the two years preceding t, where t is each decade between 1980 and 2010, as well as 2013. The data include the unit's characteristics X_{it}, building type, building age (as of the relevant period), and the unit's assessed price for real estate tax purposes, as of 2013.

I run these regression separately for each county and each time period - to maximize predictive power and impose the fewest parametric assumption. In particular, this allows for each county's value 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 to as a signal of unobserved quality. These models are then projected onto all housing units, in each county at each time period.

Sample Construction

I estimate rents and prices for 14,351,906 properties and 21,457,592 units, about 55% of the total rental housing stock of the nation at the time. In this section, I describe how I build my sample of 'small' real estate investors, as well as my estimates for wealth.

I define a small or household investor based on the attributes of the mailing address listed on each property's deed. If the mailing address is in a commercial property or the name of the owner of that address, as listed on the deed, is identified as a commercial entity by having e.g. 'LLC' or 'LP' in its name, then the owner is *not* a small investor. [14] I identify these details for each mailing address by matching the address to the full DataQuick dataset (which includes commercial properties.) This procedure identifies small investors even if they place their properties under LLCs or trusts for tax purposes.

I identify 5,754,558 unique investor households, slightly below the 7.5 million implied by Lee (2017), and the imputed participation rate of 7.2% discussed in Section 2. These investors hold a total of 9,110,025 properties. The gap is partially explained by two limitation in the matching algorithm: (a) about 1.3 million properties have a P.O. box mailing address; (b) mailing addresses in rural counties appear to systematically not be matched, and these counties tend to not have key data, such as building construction year and number of rooms.

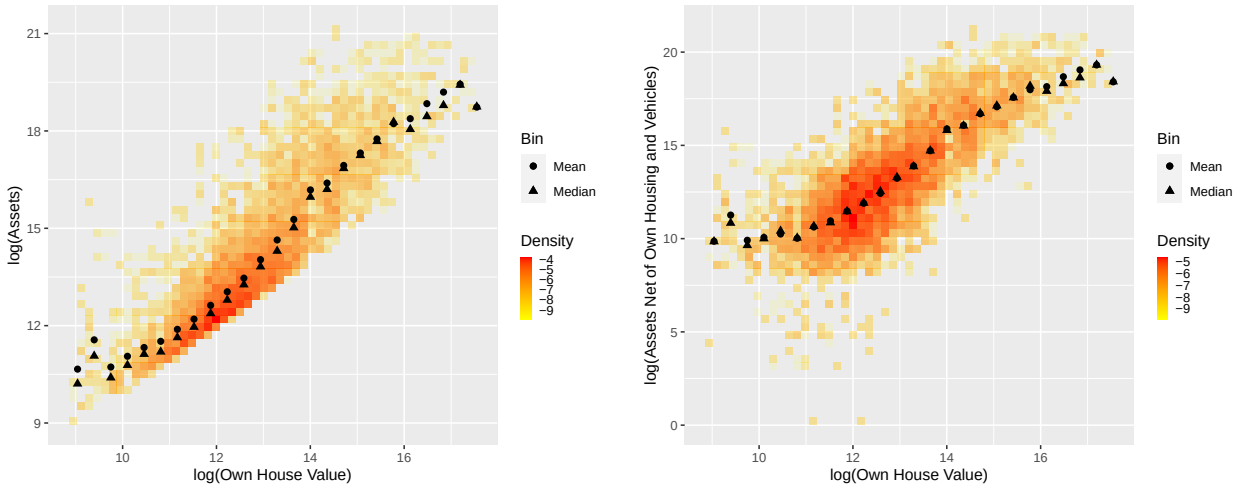
Wealth Estimates

To estimate investor wealth I use the estimated price of the investor's own residence. In the SCF the relationship between one's house and their assets is surprisingly tight - own

[14] I do not match on owner name for a few reasons: names are susceptible to different spellings, partners may put different properties under different first names, and some small investors hold their properties under.

home value explains 75% of log total assets, and 58% of log assets net of own home value. 3 visualizes these relationships. In the appendix, Table (to be added) provides more details on the regressions.

Based on these relationships, I construct two measures of investor wealth. The first predicts total wealth based on the relationship estimated in panel (a) of Figure 3. The second uses panel (b) of the same Figure to estimate net of all housing wealth, and then adds the investors' observed real estate holdings, including their own house. The advantage of the first is that it is less susceptible to outliers in real-estate wealth, whereas the second guarantees that investors' portfolio shares are always less than one. Results are robust to both, with reported results using definition (b)



(a) All Wealth

(b) Wealth Excluding Own House

Figure 3: This figure plots the relationship between wealth and own-house value using the SCF for all homeowners. Panel (a) includes all assets in the y-axis, and panel (b) includes all assets net of own-housing. The respective R^2 are 75% and 58%.

To validate whether own home value can be used as a proxy for wealth, I compare its correlation with real estate investment outcomes in the SCF and DataQuick datasets. Table 4 regresses two investment outcomes observed both in the SCF and DQ data on own-house value, wealth, and wealth net of own-house value. Panel A focuses on the log of rental value (reported for SCF, assessed for SQ), and Panel B similarly focuses on rent-price ratios.

Own-housing wealth are similarly correlated in the SCF (which directly asks measures wealth) and DataQuick sample. There are two quantitative differences. The relationship between wealth and the value of held real estate is attenuated using the implied wealth measures, possibly due to measurement error induced by the prediction exercise. The correlation between rent-price ratios and wealth is slightly higher in DataQuick, potentially because of under-reporting of rent in the SCF, and see Appendix **??**.

Table 4: Housing, Wealth, and Investment Characteristics in SCF and DataQuick

	SCF			DQ		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A Dep' Variable: log(P)					
log(Own Housing)	0.68***			0.681***		
	(0.03)			(0.006)		
log(Assets)†		0.75***			0.518***	
		(0.03)			(0.005)	
log(Assets - Housing)‡			0.59***			0.360***
			(0.02)			(0.003)
Constant	2.81***	1.02**	3.58***	3.070***	4.561***	7.201***
	(0.37)	(0.46)	(0.25)	(0.071)	(0.059)	(0.036)
R^2	0.41	0.23	0.45	0.380	0.380	0.380
	Panel B Dep' Variable: log(D/P)					
log(Own Housing)	−0.35***			−0.54***		
	(0.07)			(0.008)		
log(Assets)†		−0.30***			−0.41***	
		(0.07)			(0.005)	
log(Assets - Housing)‡			−0.23***			−0.287***
			(0.06)			(0.003)
R^2	0.05	0.05	0.04	0.297	0.297	0.297
Observations	1,129	1,129	1,129	7,009,714	7,009,714	7,009,714

Note: This table compares the relationship between measures of investor wealth and characteristics of these investors' rental property. An observation is an investor-property pair. Panel A presents coefficients of a regression of each measure of wealth on the investor's rental property. Panel B presents similar coefficients, with log rent-price ratios on the right hand side. Prices and rents are self-reported in SCF sample, and estimated for the DataQuick sample. Constants are omitted for brevity. Standard errors for DQ clustered at rental and investor level zips. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

4 Model

In this section I present a steady-state model of housing markets in which rental housing is owned by local risk averse residents with heterogeneous wealth and idiosyncratic propensity to become real estate investors. The model reconciles two popular approaches to pricing real estate - as a consumption good and as an investment good, and demonstrates that in such a setting real estate returns (prices) decline (increase) in investor wealth and investment propensity and increase (decrease) in rent volatility.

Risk-averse residents purchase housing as a consumption and as an investment good, and sort into renters, owner-occupiers, and owner-occupiers that invest in real estate. Rents and the prices of owner-occupied housing are determined in spatial equilibrium.[15] Home ownership is rationed by a liquidity constraint, creating a demand for rental units which are financed only by local residents-turned-investors. Real estate returns clear the local market for rental housing capital, and depend on wealth, propensity to invest, and risk.

4.1 Set Up

Residents. There is a continuum H of residents, indexed by i, that can purchase real estate in the market. Residents are heterogeneous and endowed with wealth W_i and an idiosyncratic ability to manage rental properties κ_i. Residents also have CRRA utility over consumption C_i with risk aversion parameter γ. Residents can purchase a home for their own consumption (termed homeowners), denoted $o_i = 1$, and which earns them a premium ψ. Finally, residents can also invest a continuous share of their wealth in other housing units, $\alpha_i \geq 0$, as an investment.

$$U = \mathbb{E}\left[\frac{(C_i)^{1-\gamma}}{1-\gamma}\right] + \psi \cdot o_i \tag{3}$$

Timing & Budget Constraint There are two periods: in the first, residents make all real estate purchases - both for their own consumption, and for investment. In the second period a market level income shock is revealed, rents are determined and paid. Households that do not buy are renters. Renters must pay rent, which is endogenously determined in the second period, or they can leave the market. Households consume in the second period subject to

[15] Appendix E provides a version of the model with an exogenous rent setting process, but with multiple risky assets. The comparative statics are qualitatively the same, with the effect of local investors being either attenuated or extenuated by the alternative risky assets, depending on their correlation with local real estate.

their budget constraint, with the exception that the homeowners premium ψ is consumed in the first period.

The budget constraint is given by:

$$C_i = Y + (W_i - P_o \cdot o_i)(1 + R_i) - (1 - o_i)D \tag{4}$$

Household consumption is equal to the common wage Y net of rent D for renters, plus their wealth net of the price of housing units for owner occupiers. Households also receive returns R_i on their wealth, or alternatively must pay interest on their mortgage debt. This is a steady-state model, and so in the second period housing does not have additional value.

Real Estate. There are H identical houses available for purchase, implying that in equilibrium all residents live in the city.[16,17] For tractability, I abstract from ex-ante differences in housing quality.

Residents can only purchase a single house for their own consumption, and doing makes them immobile in the second period. Purchasing a home for consumption is subject to a liquidity constraint LC modeled as an LTV limit and which I assume binds on the margin. Otherwise home buyers may borrow at a rate R_f.

$$W_i/P_o \geq LC \cdot o_i$$

Investment in housing is done via resident's portfolio share α_i in housing ($\alpha > 0$ implies an active investor). The return on housing is denoted R. Alternatively, investors can purchase, or borrow against, a risk-free asset that pays R_f. κ_i can be thought of as a household's aptitude in maintenance, tenant management, monitoring costs, or even a distaste parameter from the nuisance of managing rental real estate.

Investor i's return thus depends on their investment share in local real estate and their κ_i term. In equilibrium, only owner-occupiers will choose to invest in real estate, and R_i for owners will simply be R_f, reflecting mortgage interest payments.

$$R_i = \alpha_i(R/\exp(\kappa_i)) + (1 - \alpha_i)R_f$$

[16]While renters can leave the city, in equilibrium they do not do so. Alternatively, as many renters come to the city as renters that left.

[17]I do not explicitly model the previous owners of the housing stock. I assume these owners have no outside option, and to the extent they are local these values are already included in residents' wealth distribution. Housing can also be thought of as elastically supplied up to a maximum H housing units.

4.2 Rents and Owner Occupier Prices

Rents are determined in spatial equilibrium, and the price of owner-occupied housing is determined by an indifference condition for the marginal buyer. The LTV liquidity constraints rations home-ownership, driving a wedge between the price paid by owner-occupiers and investors, where the latter is lower in equilibrium than the former.

4.2.1 Rent

Rents are determined ex-post after uncertainty is resolved. Spatial equilibrium requires that the utility of mobile agents be equal, and so rents offset differences in utility across space. Equation 5 must therefore hold for all renters.

$$V + W_i(1 + R_f) = Y + W_i(1 + R_f) - D \tag{5}$$

V Denotes the reservation utility outside the market.

The ex-ante distribution of rents thus reflects the ex-ante distribution a location's value. To make the model more tractable I define a unit's value as a shock at the market level ν relative relative to the reservation utility in the broader economy V. The value of living in the market is thus $Y = V\nu$. I assume $\log(\nu - 1) \sim \mathcal{N}(\mu_d, \sigma_d)$. It is straightforward to rewrite the expected logarithmic rents $\mu_d + \nu$ as current rents plus a stochastic growth term.

Corollary 1

Inter-city and within-city spatial equilibrium implies that rent for any unit j is distributed log-normal:

$$d_j \sim \mathcal{N}(\mu_d + v, \sigma_d) \tag{6}$$

See proof in Appendix D.4. Rents follow a log-normal distribution, which in expectation is equal to the expected city premium over the reservation utility.

4.2.2 Owner-occupier Prices

There are two groups, owners and investors, who simultaneously bid for housing units. Prices are equal for all buyers in each group, and are set to satisfy the indifference condition for the buyers with the lowest valuation. The supply of housing facing both groups is perfectly inelastic, implying prices are bid down by residents with different valuations.

Prices for owner occupiers are not tractable in this model. Instead, I argue that under reasonable conditions, that fit patterns in the data, both home prices and the wealth of the marginal buyer are independent of rental housing returns. This allows solving for returns, taking owner-occupier prices as given.

Willingness-to-pay and the Marginal Buyer. Homeowners purchase houses as long as their utility from doing so is higher than their expected utility as renters. In equilibrium, prices are set to make the marginal buyer, W_o, indifferent to renting. (In equilibrium the liquidity constraints binds for the marginal buyer and so $LC \cdot P_o = W_o$.)

$$E\left[\frac{(W_i - P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = \frac{(V + W_i(1+R_f))^{1-\gamma}}{1-\gamma} \tag{7}$$

The utility of purchasing a home stems from the ownership premium ψ and from the value of the expected positive wage payment. Ownership has three costs to households: (a) the house price P_o, (b) foregone return or interest payments, and (c) the implicit insurance offered by remaining mobile.[18] Intuitively, home prices are 'bid-down' the willingness-to-pay curve, until the liquidity constraint binds and prices and the marginal buyer is determined. Critically, the resulting price does not depend on the return to rental real estate, and all owner-occupiers bid more for their own homes than investors do for rentals.

Independence of Returns. Proposition 2 shows that under reasonable conditions home prices for owner-occupiers, and the wealth of the marginal buyer do not depend on returns. These variables are still a function of the local labor market Y, the outside option V, and mortgage terms: the rate R_f and the liquidity constraint LC. (Net rates are denoted with a lower-case r.)

Proposition 1

Assuming the liquidity constraint binds at the margin, and under certain regularity conditions, for any return r there exists a non-empty interval $\left[\underline{\psi}, \overline{\psi}\right]$ such that if $\psi \in \left[\underline{\psi}, \overline{\psi}\right]$ then $P_o > P_I(r)$, and home-ownership is limited to $W_i > W_o$, Where both P_o and W_o are independent of returns r.

See proof in Appendix D.5.

Discussion of Proposition 1. Proposition 1 focuses the model on a specific region of the parameter space, which is significantly more tractable: first, it allows for solving for local returns and portfolio shares taking owner-occupier prices as given, and second, it implies that the marginal home buyer never invests in real-estate, limiting the group of investors to only a subset of owner-occupiers.

In addition to the tractability benefits of proposition 2, its assumptions and implications fit patterns in the data. As discussed in Section 2.3 and shown in Figure 2 home-ownership increases rapidly from the 25$^{\text{th}}$ wealth percentile, and real estate investment, only starts

[18] This trade-off of rent risk vs. local risk is contrary to the 'ownership as a hedge against rent' intuition of Sinai and Souleles (2005), but in-line with the 'renting as a hedge against local risk' of Lockwood et al. (2023).

at the 40^{th} percentile. Appendix Figure 5 shows significant bunching at institutional LTV limits, suggesting a binding of a liquidity constraint in the US. In my data, the average rental unit is cheaper than an owner-occupier unit, conditional on location and unit controls.

Lastly, there is a robust literature showing that owner-occupied home prices are priced by owner-occupiers, and not investors.[19]

The intuition for the proposition is as follows: Owner-occupiers trade-off the positive mean, but risky, city shock and ownership benefit ψ against the costs of ownership. The value of the city shock is smaller for poorer households, for whom it is a bigger risk relative to their consumption, and for whom the value of the insurance offered by mobility is higher. Therefore, willingness-to-pay for homes declines in wealth and poorer households near the margin only buy due to ψ. ψ must be high enough relative to the risk for the marginal liquidity constrained buyer to value the house more than the investor, which only receives discounted rents; but such a ψ necessarily exists for any r.

4.2.3 Renter and Owner Populations

Given P_o and the liquidity constraint, the resident population is split into a renter group H^{rent} and owner group H^{owner} where

$$\begin{aligned} H^{\text{rental}} &= H \cdot F_w(lc - p_o) \\ H^{\text{own}} &= H \cdot (1 - F_w(lc - p_o)) \end{aligned} \tag{8}$$

This follows immediately from Proposition 1.

4.3 Real Estate Investors

Residents solve a portfolio choice problem, choosing the optimal portfolio share (occasionally referred to as 'exposure') devoted to real estate, α_i. I first show how the FOC from this problem yields an Eurler equation, familiar in the portfolio choice literature. I then present proposition 3, which identifies and characterizes the set of investors and their portfolio shares, as well as pins down the price of rental units.

4.3.1 Investor Capital Supply

Since P_o does not depend on returns, we can write $\tilde{W}_i = W_i - P_o$, as well as define $R_i = R/K_i$. I solve the investor's problem, as in Campbell et al. (2002) and Campbell (2017),

[19] An exception is Ater et al. (2021), who show that prices and rents in Israel are sensitive to constraining the participation of investors.

by log-linearizing the budget constraint, around a value summarized by the the elasticity of consumption with respect to wealth, denoted ρ_i, and solving for the investor's first order condition.[20]

The investors' Euler equation implies

$$E[r - \kappa_i - r_f] + \frac{1}{2}\sigma_r^2 = \gamma\sigma_r^2\alpha\rho_i + \gamma(1-\rho_i)\sigma_{yr} \tag{9}$$

Equation (9) can also be thought of as an investor specific capital supply function which is upwards sloping in returns. Each investor's capital supply decreases in the investor's κ_i - investors that are better at operating real estate offer more capital, at any level of returns, creating heterogeneity in who owns rental real estate, and how much of it. Finally, investor capital also depends on market level variables: risk aversion γ, return risk σ_r, and the covariance of returns with city level income Y.[21]

4.3.2 Identifying Investors

We can define a set of marginal investors as those for which the optinal α_i in equation (9) equals zero. Proposition 2 characterizes this set, uses it to find the price of rental units, and describe how portfolio shares adjust along the wealth and κ_i dimensions for the set of investors.

Proposition 2

If the conditions in corollary 1 and proposition 1 hold, then participation is limited to $W_i > W_m$, and investment shares α_i increase wealth and decline in κ_i; specifically, the set of investors is defined as the double integral over wealth on the interval $[W_m, \infty)$ and κ on the interval $[0, \kappa^*(W)]$, where $\kappa^*(W)$ is a concave increasing function in W.

See proof in Appendix D.6. Proposition 2 allows for characterizing the group of investors using only wealth, and the participation rate within each wealth level. Total local capital supply is the integral over the portfolio share for participating investors. Portfolio share increases in wealth, and decreases in κ on both the extensive and intensive. $\kappa^*(w_i)$ delineates

[20] Log-linearization the budget constraint implies

$$c - y \approx C + \rho_i(\tilde{w}_i + r_{p,i} - y)$$

where C is a constant that will drop, and where

$$\rho_i \equiv \frac{\exp(\tilde{w}_i - \mu_\nu - v - \alpha_i\kappa_i + \bar{r}_p)}{1 + \exp(\tilde{w}_i - \mu_\nu - v - \alpha_i\kappa_i\bar{r}_p)}$$

[21] σ_{yr} is positive as Proposition 1 tightly relates the rent payment to the city shock, and so investors require an additional premium because their own consumption of the city's shock correlates with the rent they will receive on their assets.

the participation rate within each wealth level w_i and identifies the lowest wealth investor w_m that participates. Proposition 3 also caps the set of investors even at very high levels of wealth, a pattern found in the data, as seen in Section 2.

4.4 Equilibrium & Comparative Statics

Local returns are set in equilibrium by clearing the market for local real estate. Investor capital supply is integrated over all residents who chose to invest, and must equal the value of rental real estate. Equation (10) describes this equilibrium, using Propositions 1 through 3, and substituting in each investor's capital supply, Equation (9) as well as the definition of prices as discounted expected rent payment.

$$\begin{aligned} &\underbrace{\int_w \int_\kappa \alpha(w,\kappa) \cdot (W_i - P_o) \cdot \mathbb{1}\{\alpha(w,\kappa) \geq 0\} \cdot dF_{w,\kappa}}_{\text{capital supply}} = \underbrace{P_I(r) \cdot H^{\text{rental}}}_{\text{capital 'demand'}} \\ &\int_{w_m}^{\infty} \int_0^{\kappa^*(w)} \left(\frac{1}{\rho_i} \frac{E[r - r_f] - \kappa_i + \frac{1}{2}\sigma_r^2}{\gamma \sigma_r^2} - \frac{1-\rho_i}{\rho_i} \beta_{yr} \right) \cdot (W_i - P_o)\, dF_{w,\kappa} = \\ &\qquad \exp(\mu + v - E[r]) \cdot F_w(lc - p_o), \end{aligned} \tag{10}$$

which only depends on r. Return risk and covariance are determined by rent (growth) volatility:

$$\begin{aligned} \sigma_r &= Var(r) = Var(d-p) = Var(d) = \sigma_d \\ \sigma_{yr} &= Cov(r,y) = Cov(\log(\nu - 1), \log(\nu)) \end{aligned} \tag{11}$$

Comparative Statics on Returns. While the model is not analytically solveable, we can characterize certain comparative statics. Proposition 3 summarizes these effects.

Proposition 3

Assuming the conditions of Proposition 1 and 2 hold, equilibrium returns decline in: (1) an increase in wealth among any subgroup of investors, (2) a decrease in the distribution of κ_i, (3) an increase in rent volatility and (4) an increase in the risk free rate.

Any one of these variables shifts out the capital supply, lowering equilibrium returns to preserve equilibrium. An increase in investors' wealth increases the supply of capital on the intensive and extensive margins: each investor supply more capital, and more investors in each wealth group supply capital. These effects are somewhat attenuated by the positive covariance of returns with investors' own consumption, but since α_i increases with wealth this secondary effect is always smaller than the main effect.

Note that an increase in inequality has an ambiguous effect on returns: higher wealth lowers returns, but a higher share of renters creates additional demanding for housing, this increasing returns.

Comparative Statics on Prices and Rent-Price Ratios. The comparative statics on log rent-price ratios mostly follow from those on returns, and can be read from the identity

$$E[\log(D/P)] = E[r] + \sigma_d^2/2$$

Comparative statics on prices are also straightforward, and are the inverse of those on rent-price ratios, with the exception of expected rent (growth), which has an additional, direct, positive effect.

In sum, returns and rent-price ratios decrease and rental stock prices increase in investor wealth and in the propensity of investors to invest (through κ_i). Returns and rent-price ratios increase and prices decrease in the number of renter households. Returns increase in risk, but the effect on rent-price ratios and prices is ambiguous.

4.5 Simplified Case: Discrete Groups Solution

Assuming simplified heterogeneity allows solving the model analytically, and deriving straightforward comparative statics. Suppose that wealth takes three values $\tilde{w}_i \in \{0, w^L, w^H\}$ and that κ can be either high or low, denoted $\bar{\kappa}$ and $\underline{\kappa}$. Correlation across these groups may be arbitrary.

Households with zero wealth are renters such that $0 < lc - p_o < w^L$, and that w^L, w^H are owners. The marginal buyer is w^L and owner prices are set to make w^L indifferent between buying and renting.[22] Both the w^L and $w = 0$ groups do not invest even if $\kappa = \underline{\kappa}$, and w^H invest only if $\kappa = \underline{\kappa}$.

We can solve equation (10) under these assumptions, and find equilibrium returns and portfolio shares.

$$\begin{aligned} E[r - r_f] + \sigma_r^2/2 &= \underline{\kappa} + (1-\rho)\gamma\sigma_{yr} \\ &\quad + \mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(\mu + v - r_f - \tilde{w}^H - \underline{\kappa} + \sigma_r^2 - (1-\rho^I)\gamma\sigma_{yr})\right) \quad (12) \\ \alpha &= \frac{1}{\rho^I\gamma\sigma_r^2}\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(\mu + v - r_f - \tilde{w}^H - \underline{\kappa} + \sigma_r^2 - (1-\rho^I)\gamma\sigma_{yr}))\right) \end{aligned}$$

[22] In the discrete case the liquidity constraint may not bind for the marginal group, it just needs to be violated for the renter group; the free parameter ψ can guarantee the requirements of Proposition 2 that prices for owner occupiers and investment properties differ. For simplicity I assume it does.

Where $\mathcal{W}$ is the Lambert function, defined as $ae^a = b \rightarrow a = \mathcal{W}(b)$. See Appendix D.10 for derivation and discussion of the assumptions on the parameters $\psi, \underline{\kappa}, \bar{\kappa}$ and σ_r.

Deriving comparative statics is straightforward, and the results align with Proposition 3 and 4: return and the portfolio share of investors decline in wealth and risk. These solution imply additional prediction on the magnitude and shape of the effect of wealth - its elasticity should be less than unit, and it should be concave - implying that the impact of wealth should be lower in wealthier markets and higher in markets.

One advantage of this simplified model is that it summarizes the effects of participation rates into a single term - the share of residents that invest: ϕ^I. Returns and rent-price ratios decline in the share of investor, and interestingly, in the *relative* share of investors and renters: both are set by the relative shares of these groups - how many renters must each investor finance.

5 Empirical Evidence

In this section, I provide evidence that the wealth of local residents, as well as their propensity to become investors, affect local real estate prices; in particular that they increase prices more than rents, and lower returns. I conduct two sets of reduced-form exercises: in the cross-section (Section 5.1), utilizing unit-level and neighborhood level data, which allows for detailed controls and an IV to identify the effect of the share of investors; and in a metro-level panel dataset (Section 5.2), which implements shift-share design for wealth shocks, based on national equity market shocks, to credibly identify the effect the of investors' wealth.

5.1 Cross-Sectional Regressions

I conduct four cross-sectional tests of the model's implications, which offer progressively more explicit, less naive, and more parametric, tests of the role of wealth and the propensity to invest. The first test shows that the wealth-rent-price relationship is explained by investor location and the second test shows that property-level rent price ratios fall in local wealth, conditional on a rich set of investor, property, and local controls. The third includes investor fixed effects, showing properties held by the same investor, but located in neighborhoods where others' investors are wealthier have lower rent-price ratios.

The fourth analysis is at the neighborhood level, at which investors are still segmented, but the spatial variation is rich enough to test both the effect of wealth and the effect of the propensity of residents to become investors (investment propensity) using 2SLS.[23] I use

[23]In the model, different local propensities would be modeled by different distributions of κ

equity holdings as an IV to better measure wealth, addressing inherent measurement errors specific to active investors, and local ethnic composition as a more standard IV to plausibly identify the effect of the propensity of residents to invest. This dual approach leverages the narrower focus of the first IV on wealth while the second IV broadly captures investment tendencies.

5.1.1 Property-level Regressions

In this section, I report property level regressions that utilize property and investor specific controls, such as building age and investor wealth. I measure the average wealth of investors investing in each property's zip, excluding the property's own investor's wealth. This allows me to avoid defining markets and the potential investor set within them, using only actual investor markets.[24] The results indicate a significant and robust negative association between average investor wealth and rent price ratios, driven by investor location, and not an investor's own wealth. I defer discussion of the causality of this effect for the next, more parametric, analyses.

Wealth and Investor location Table 5 reports regressions of log rent-price ratios on their investor's own home value which, as shown above, is a reasonable proxy for investor wealth. Both in the SCF and in DataQuick there's a negative covariance between investor wealth and their property's rent price ratio. However, this covariance is primarily explained by investor location fixed effect, suggesting wealthier investors are 'forced into' low rent-price properties, rather than chose them for unobserved reasons.

In the first column, I use data from the SCF on the 1,246 real estate investors who report rent payments, and find that an elasticity of -0.29 between own-home value and log rent-price ratios. In column (2) I repeat this regression in the investor-property linked dataset derived from the DataQuick data, and find a similar coefficient of -0.2, with 4.8 million observations.[25] Columns (3) adds controls for the building type, construction decade property-investor distances; controls which leave the coefficient and (within) R^2 unchanged.

The main results are reported in Columns (4) and (5). Adding metro and investor fixed effects reduces the coefficient to -0.14 and -0.13 respectively. More importantly, the explanatory power of own-home value declines precipitously - from 9.8% to 4% and then to 2%. The relationship between investor wealth and yields is almost fully explained by an

[24] I.e. if markets are also segmented based on cultural or generational lines: descendent of residents of a neighborhood might be more likely to invest in it, regardless of their physical proximity to it; or former students at a college might be more likely to invest around it.

[25] The difference between the coefficients does not seem statistically significant, but could nonetheless be driven by under-reporting of rent in the SCF, as discussed in Appendix 2.

Table 5: Investor Wealth, Rent-Price Ratio, and Location

Dependent Variable:	log(Average D/P)						
Sample:	SCF	DataQuick					
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
log(Inv' House Value)	-0.29**	-0.196***	-0.193***	-0.143***	-0.132***	-0.067***	-0.055***
	(0.10)	(0.030)	(0.034)	(0.022)	(0.006)	(0.024)	(0.014)
log(Inv'-property Distance)			-0.016	0.005	0.016***	0.018*	0.022**
			(0.013)	(0.006)	(0.005)	(0.011)	(0.010)
log(Mean Wealth at Inv' Zip)						-0.488***	
						(0.044)	
log(Mean D/P at Inv' Zip)							0.951***
							(0.043)
Fit statistics							
Observations	1,246	4,776,117	4,681,113	4,681,113	4,681,113	4,681,113	4,681,113
R^2	0.03	0.098	0.164	0.509	0.567	0.313	0.463
Within R^2			0.101	0.04	0.02	0.261	0.422
Fixed-effects							
Construction Decade			Yes	Yes	Yes	Yes	Yes
SFH/Condo/Apt Bldg			Yes	Yes	Yes	Yes	Yes
Bldg Size Category			Yes	Yes	Yes	Yes	Yes
Metro				Yes	Yes		
Investor Zip					Yes		

Note: This table presents regressions of rent-price ratios on investor wealth and investor location fixed effects. An observation is an investor-property pair. The unit of observation is an owner-property pair. The first column uses data from the SCF, with self reported housing, real estate, and rent data. Columns (2) through (5) use the investor-property linked DataQuick derived dataset, with unit-level estimates for these variables. Standard Errors for the last five columns are clustered by CBSA. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

investor's location. Columns (6) and (7) show that when controlling for market averages, the relationship declines even further.

Average Local Wealth. Table 11 in the Appendix reports regressions at the investor-property level, controlling for unit, investor, and local characteristics. These regression directly test the effect of the average wealth of investors who invest around each unit, calculated without each unit's owner's wealth. To address concerns that the measurement error in wealth is spuriously correlated with the units' prices (due to their proximity) I instrument for average local wealth with an IRS derived measure of local equity wealth. I include a rich set of unit controls, including the unit's age, structural characteristics, as well as the owner's own wealth and distance to the property, and for local amenities with local racial composition, income, BA share, and distance to CBD. I also control for the average price of owner-occupied housing which, in most models, capitalizes differences in both current and future amenities. I find the elasticity of a unit's rent-price ratio with respect to average wealth of local investors to be a significant -0.39, much larger than the elasticity with respect to an investor's own wealth (-0.01). See Appendix F for further discussion.

Investor Fixed Effects. In Appendix Table 9 I also include investor fixed effects, estimating the effect of neighboring wealth for investors who hold at least two properties in multiple locations. The model does not allow or account for the existence of such investors, but in reality, some investors hold properties in various markets, likely due to idiosyncratic reasons - such as having lived in an area in the past or inheritance. These 'out-of-the-model' specifications are still useful to estimate wealth effects not driven by selection on investors' characteristics.

5.1.2 Market-level Regressions

In this section, I regress market-level average rent-price ratios on market aggregates that capture local capital supply: average wealth and the propensity of residents to become investors (investor share). These regressions directly test the comparative statics of the model, using well-defined markets based on the spatial segmentation of investors. Local wealth, investor share, and others are easily aggregated within each spatial market. I confirm many of the model's predictions, including the shape and magnitude of the effects: local wealth and the share of investors lower rent price ratios, with an elasticity of less than unity, and the share of investors has the same magnitude of an effect as the share of renters.

I use a combination of IVs and controls to identify the effect of wealth and, particularly, the share of investors. I control for common market level measures of amenities - including BA share and distance to CBD, as well as the average price of owner-occupied units, which

should capitalize differences in amenities across space. I also instrument with local equity wealth to purge measurement error. Finally, I use a novel IV for investment propensity, based on ethnicity-induced differences in preferences for real estate, to recover a causal effect for this variable.

Market Definition and Specification. I calculate market attributes as averages for all rental properties and investors that live within a 5-km radius of each focal zip. Distance is calculated using centroids, and zip clusters do not cross metro lines.[26]

Within each market i in city c I aggregate variables for rental properties, and for the investors living in that market, regardless of where these investors hold property. I construct average rent price ratios (D/P_{ic}), average investor wealth (W_{ic}), and the share of investors ϕ_{ic} out of local residents. I add data from the ACS, including the share of renters (ϕ^R_{ic}), average wages, BA-shares, and the price of owner-occupied units. To form market averages of these variables, I weight each zip by the number of rental units.[27] The result is the average, e.g. wage, that each rental unit is exposed to (and thus the average wage effecting that unit's rent, price, and growth prospects.)

The preferred specification is as follows:

$$\log(D/P_{ic}) = \alpha \log(W_{ic}) + \gamma_1 \log(\phi^I_{ic}) + \gamma_2 \log(\phi^R_{ic}) + \beta X_{ic} + \delta_c + \epsilon_{ic} \tag{13}$$

X is a vector of controls, including race, income, and the average price of owner occupied housing.

This specification is based on the simplified heterogeneity model result in Equation (12), if one were to replace the Lambert function with a log function, and assume we are in steady state, such that rent-price ratios equal returns.

The city fixed effects control for differences in risk, as well as the labor market differences in expected growth μ. The control variables, X, capture market-specific differences in growth expectations (particularly owner-occupied housing).

The error term, ϵ_{ic} can be interpreted in light of the model as a combination of (1) sub-city rent volatility risk, (2) rent-investor consumption covariance, and (3) growth expectations not captured by the controls X_{ic}. These variables, even within the model, are not exogenous to either investor wealth or the share of investors - as the marginal investor depends on risk and growth expectations. The error is likely also a function of amenities that co-determine the location of different population groups.

α and γ_1 test the sign of effect of capital supply. The model predicts the following

[26] As the market areas partially overlap by design, I cluster standard errors at the metro area. This is conservative, but also accounts for likely correlation within labor and metro housing markets.

[27] In the case of the share of renters, I add all renters and divide by the total population.

comparative statics: $\alpha_0 < 0, \gamma_1 < 0, \gamma_2 > 0$. The model further predicts $\gamma_1 \approx -\gamma_2$, but this relationship cannot be evaluated in an IV specification, as these estimate the LATE.

Discussion of the Instruments. To rigorously address the endogeneity and simultaneity concerns identified within the model, I instrument for the share of investors with measures of local ancestry. I also instrument for local average investor wealth with local average equity wealth, to purge measurement error in the former that is potentially correlated with housing prices.

The ancestry IV identifies the share of investors and serves a dual purpose: it addresses simultaneity and reverse causality — key concerns when considering the effect of capital on rent-price ratios — but also acts as a shifter for the demand for rental real estate. This IV is based on the discussion in Section 2.3 and Table 3: that ancestry, and particularly average AREP, predicts participation in real estate, driven by certain regions, and conditional on individuals' education. By shifting the demand for real estate, and conditional on race and other observables, average AREP allows for the identification of the supply effects of capital, through the share of residents who become investors.

The exclusion restriction assumes that the estimated AREP effect is not correlated with local omitted variables, among them unobserved growth expectations. To support this assumption, I point to the fact that AREP is estimated nationally and controls for individual characteristics, and to the fact that the literature has generally not emphasized ancestry, controlling for race, as an important predictor of any housing outcome. Furthermore, such residual correlation [28] should be soaked by controlling for local owner-occupied prices.[29]

I also employ an instrument for wealth to mitigate error stemming from the variable's measurement process. The challenge arises from the use of investors' own home values as a proxy for their overall wealth. If home prices are spatially correlated, and investors are local, measured wealth would mechanically correlate with local housing market residuals, potentially biasing the results.

To counteract this, I use equity wealth, measured for all residents, as an alternative measure that is not correlated with the residuals of the housing market. This approach leverages the idea of using distinct measures of the same underlying variable to effectively purge the wealth variable from measurement errors correlated with housing prices, thereby enhancing the reliability of the model's estimations.

In summary, these IVs provide a robust framework for isolating the causal impact of

[28] Through, perhaps, certain clustered ancestries drawing foreign direct investment Badarinza et al. (2022); Burchardi et al. (2019), or alternatively positive externalities to diversity Ehrl and Monasterio (2024).

[29] I also describe below alternative specifications of the IV, that directly utilize the ancestry composition in each market, either utilizing all ancestry shares, or the first few components of a PCA analysis of ancestry composition.

landlord capital, particularly the share of investors, ensuring that empirical findings are credible and well grounded within the theoretical structure of the model.

Results. Table 6 reports a set of the market level regressions. Specifications 1 through 3 are OLS, and specifications 4 through 6 report the IV specifications, where investor average wealth is instrumented for with equity wealth, and the share of investors is instrumented for with the ancestry and nativity variables. F-stats are reported separately for each variable.

Columns (1) and (4) regress log average rent-price ratios on just market mean wealth, and the controls. Columns (2) and (5) add the share of renters and investors. Columns (3) and (6) include a non-linear term for the effect of wealth. All specifications control for owner-occupied price, and a set of amenity and race controls as in Table 11, as well as population and density controls.

Throughout, the effect of wealth is negative and significant, and is less than unity when controlling for the share of renters and investors. The effects of the share renters and investors fits the models prediction: the former is positive and the latter is negative, and they are of a similar quantitative magnitude and the difference is statistically insignificant. The elasticity of rent-price ratios with respect to wealth for the poorer half of cities is -0.541, but only -0.126 among the wealthier cities. An increase of 1% in investor share lowers local rent-price ratios by -0.157% whereas an increase of 1% in renters increases local rent-price ratios by 0.143%.

Bias relative to the model. Appendix E.2 discusses two ways in which the reduced form estimates may be biased relative to the model: (1) omitted variable bias due to endogenous equity shares, and (2) specification error inherent in the linear functional form. I find that (1) omitting equity shares may downward bias the wealth coefficient but, under the null of no or positive wealth effect, cannot generate a negative sign; and (2) that the linear coefficient is similar in Monte Carlo simulation to the model's marginal effect of wealth evaluated at the mean.

5.2 Panel Data - Wealth Shocks

In this analysis I identify a causal effect of wealth shocks on returns and test the model's prediction that positive wealth shocks lower returns. I use a shift-share instrument, where plausibly exogenous shifts come from annual equity market shocks. The instruments constructs shifts to level of *local* wealth as a function of the exogenous shift and baseline exposure. I find a positive shock lowers local returns and rent-price ratios even when controlling for year and metro fixed effects, although the effect on the latter is not statistically significant when I use Driscoll-Kraay standard errors, which are both serially and spatially robust.

Table 6: Within City, Net Ancestry Effect IV, 5km Sample

Dependent Variable:	log(Average D/P)				
	OLS	IV		OLS	IV
Model:	(1)	(2)	(3)	(4)	(5)
log(Share Investor)	-0.128***	-0.441**	-0.930**	-0.057***	-0.665**
	(0.011)	(0.191)	(0.433)	(0.017)	(0.334)
log(Share Renter)	0.139***	0.258***	0.492***	0.094***	0.383***
	(0.015)	(0.074)	(0.184)	(0.011)	(0.143)
log(Mean Investor Wealth)				-0.286***	-0.836***
				(0.027)	(0.164)
log(Average Owner Price)			-0.153**	-0.011	0.163**
			(0.059)	(0.023)	(0.064)
Controls					
Metro Fixed Effects	✓	✓	✓	✓	✓
Amenities & Income	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓
Statistics					
Observations	10,108	10,108	9,962	9,962	9,962
Within R^2	0.315	-	-	0.407	-
Investor 1st F	-	62.31	46.17	-	24.21
Wealth 1st F	-	-	-	-	271.5
IVs	-	1	1	-	2

This table presents regressions of local rent-price ratios on local determinants of capital supply, as well local controls. The unit of observation is a cluster of zip codes, defined around each zip-code with a 5-km radius. Mean investor wealth is computed for investors that live within each market, and instrumented for with the average equity wealth, derived from IRS tax data by Chodorow-Reich et al. (2021). Investor and renter shares are computed out of total households. Investor share is instrumented for with the ancestry derived real estate proclivity (AREP) variable computed from participation in the real estate licensing industry (NAICS 503). Controls include the average price of owner-occupied properties, race, and amenities, including BA share, density, and distance from CBD, all from ACS, except the distance to CBD, which is measured from city centers from Hausman et al. (2023). Standard errors clustered at the metro level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Equation (14) presents the reduced form test. I regress changes to returns ($\Delta r_{c,t}$) as a function of changes in wealth ($\Delta W_{c,t}$), controlling for city and time fixed effects (δ_c and δ_t, respectively). The hypothesis is $\beta < 0$.

$$\Delta r_{c,t} = \delta_c + \delta_t + \beta \Delta W_{c,t} + \epsilon_{c,t} \tag{14}$$

A potential concern is the correlation between housing demand and returns and local wealth at the city level. To address this, a shift-share instrument for equity wealth shocks ($\Delta EW_{c,t}$) is constructed, which relies on the initial per-household exposure to the equity market, along with the return to the equity market at time t ($r_{eq,t}$). The instrumented changes in equity wealth are then regressed on changes in rent/price ratios or returns, allowing for the estimation of the causal effect of wealth shocks on returns.

Data. A disadvantage of this data is that it uses noisier measures for both rent price ratios, as well as wealth. Rent price ratios are noisier because price appreciation data is for all housing, not just rental housing. Wealth data is noisier since it is for all residents, not just real estate investors. Both of these should attenuate the estimates , and inflate their standard errors.

Instrument Construction. The instrument is given by Equation (15). Each city-year instrument is equal to the level shift in equity wealth, driven by that year's equity market surprise, the city's baseline marketbeta $\beta_{c,t=0}$ and the city's hypothetical, baseline based, equity position at time $t-1$. Specifically, this position is equal to city's baseline portfolio $EW_{c,t=0}$, compounded over realized returns and assuming the city's residents made no withdrawals over the period.

$$B_{c,t} = \underbrace{EW_{c,0} \cdot \beta_{c,t=0}}_{\text{Baseline Position}} \cdot \left(\underbrace{\prod_{\tau=0}^{t-1}(1 + r_{eq,\tau})}_{\text{Cum' return at } t-1} \cdot \underbrace{(r_{eq,t} - E[r_{eq}]}_{\text{Surprise at } t} \right) \tag{15}$$

The analysis focuses on a city-level panel dataset due to data constraints. I use a metro-level panel dataset, available from Samuels (2024), constructed from BLS and FHFA rent and price series, covering 22 metropolitan areas. The initial exposure to the equity market and the city betas are imputed from annual county total equity wealth and betas, based on annual IRS reported dividend taxes. These county-level data are aggregated to the metropolitan level, and annual per-person average equity wealth is computed. The equity market shocks are measured using S&P 500 returns. I use a baseline for the analysis set at either 1990 or 2000.

Results. Results are given in Table 7. Columns 1 through 3 regress changes to log rent-price ratios, and Columns 4 through 6 regress changes to returns. Columns 2 and 5 add metro fixed effects. Columns 3 and 6 add year fixed effects, narrowing the identifying variation to differences at the metro-year level, specifically driven by each city's exposure to annual equity market shock. I compute Driscoll-Kraay standard errors which flexibly control for

Table 7: Wealth Shock Impact on Returns

Dependent Variables:	$\Delta \log(D/P_{c,t+1})$			$\Delta r_{c,t+1}$		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔEW ($10,000)	-1.62*	-1.61	-0.684	-0.919**	-0.916**	-1.28*
	(0.971)	(1.10)	(1.21)	(0.398)	(0.402)	(0.764)
Fixed-effects						
Metro	-	✓	✓	-	✓	✓
Year	-		✓	-	-	✓
Fit statistics						
Observations	404	404	404	404	404	404
1st F	129.2	129.8	66.4	129.2	129.8	66.4

This table presents regressions of changes in rent-price ratios and returns (at time $t+1$) on wealth shocks at time t. Rent-price ratios and returns are computed at the metro level for each year, based on FHFA and BLS price and rent series, and benchmarked to DataQuick's averages for rental units, as of 2013. Wealth data are equity wealth from Chodorow-Reich et al. (2021). Wealth shocks are identified using a shift-share design, with national equity market shocks, and shares set at cumulative return on baseline equity position. Each city-year instrument is computed using only baseline shares and cumulative national shocks. Baseline year is 2000. All standard errors are Driscoll-Kraay, computed with optimal bandwidth as in Lazarus et al. (2018). *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

time-series and cross-sectional heteroskedasticity. The coefficients are also significant when clustering by year.

The results consistently show that positive wealth shocks to local equity wealth reduces local rent-price ratios and returns, with the former remaining statistically significant despite the limited power of the analysis.

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.

J. D. Angrist, G. W. Imbens, and D. B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455, 1996.

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.

C. Badarinza, T. Ramadorai, and C. Shimizu. Gravity, counterparties, and foreign investment. *Journal of financial economics*, 145(2):132–152, 2022.

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

K. B. Burchardi, T. Chaney, and T. A. Hassan. Migrants, ancestors, and foreign investments. *The Review of Economic Studies*, 86(4):1448–1486, 2019.

J. Y. Campbell. Household finance. *The journal of finance*, 61(4):1553–1604, 2006.

J. Y. Campbell. *Financial decisions and markets: a course in asset pricing*. Princeton University Press, 2017.

J. Y. Campbell, L. M. Viceira, L. M. Viceira, et al. *Strategic asset allocation: portfolio choice for long-term investors*. Clarendon Lectures in Economic, 2002.

K. Chang. Diversifying the suburbs: Rental supplyand spatial inequality. *University of California, Berkeley*, 2024.

G. Chodorow-Reich, P. T. Nenov, and A. Simsek. Stock market wealth and the real economy: A local labor market approach. *American Economic Review*, 111(5):1613–1657, 2021.

J. Coven. The impact of institutional investors on homeownership and neighborhood access. *Available at SSRN*, 2024.

S. Damen, M. Korevaar, and S. Van Nieuwerburgh. An alpha in affordable housing? *Available at SSRN*, 2025.

Y. Deng, M. R. Hu, and A. D. Lee. Melting pot or salad bowl: Cultural distance and housing investments. *Real Estate Economics*, 49(S1):235–267, 2021.

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.

P. Ehrl and L. Monasterio. Inherited cultural diversity and wages: surname-based evidence. *Journal of Economic Geography*, page lbae011, 2024.

A. Eisfeldt and A. Demers. Total returns to single family rentals. Technical report, National Bureau of Economic Research, 2021.

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

C. Gorback and G. Schubert. The impact of cultural preferences on homeownership. *Working Paper*, 2023.

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

D. Johnstone and Y.-X. Lin. A new equation for the price of a lognormal payoff under power utility. *Available at SSRN 4554777*, 2023.

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.

R. S. Koijen and M. Yogo. A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515, 2019.

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.

H. Lee. Who owns rental properties, and is it changing? *Joint Center for Housing Studies of Harvard University*, 18, 2017.

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.

P. Samuels. Spatial dispersion in rental returns: A decomposition of local rent-price ratios. Technical report, 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.

Appendix

A Appendix Tables

Table 8: Segmentation Table by Bins

# Units	Same Zip	$\leq$ 10km	$\leq$ 25km	$\leq$ 50km	$\leq$ 100km
1	0.29	0.44	0.62	0.70	0.74
2	0.36	0.60	0.80	0.86	0.89
(2,4]	0.34	0.59	0.78	0.85	0.89
(4,9]	0.23	0.58	0.81	0.89	0.93
(9,19]	0.20	0.54	0.77	0.86	0.90
(19,49]	0.31	0.58	0.76	0.85	0.89
(49,·]	0.42	0.60	0.74	0.83	0.86

Note: Share of properties that are located within distance thresholds. Distance is calculated using zip-code centroids. *p<0.1; **p<0.05; ***p<0.01

Table 9: Unit-level Impact of Local Wealth, Investor Fixed Effects

Dependent Variable:	log(D/P)	
Model:	(1)	(2)
Variables		
log(Mean Zip Wealth)	-0.383***	-0.266***
	(0.038)	(0.093)
log(Investor-property Distance)	-0.0002	-0.0005
	(0.001)	(0.002)
log(Mean Owner Occupier Value)	-0.437***	-0.481***
	(0.037)	(0.049)
Wealth 1st F	-	516745.33
Amenities & Income	✓	✓
Race	✓	✓
Fixed-effects		
Investor ID	Yes	Yes
Metro	Yes	Yes
Construction Decade	Yes	Yes
SFH/Condo/Apt Bldg	Yes	Yes
Bldg Size Category	Yes	Yes
Fit statistics		
Observations	2,799,971	2,799,971
R^2	0.827	0.826
Within R^2	0.093	0.092

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

B Appendix Figures

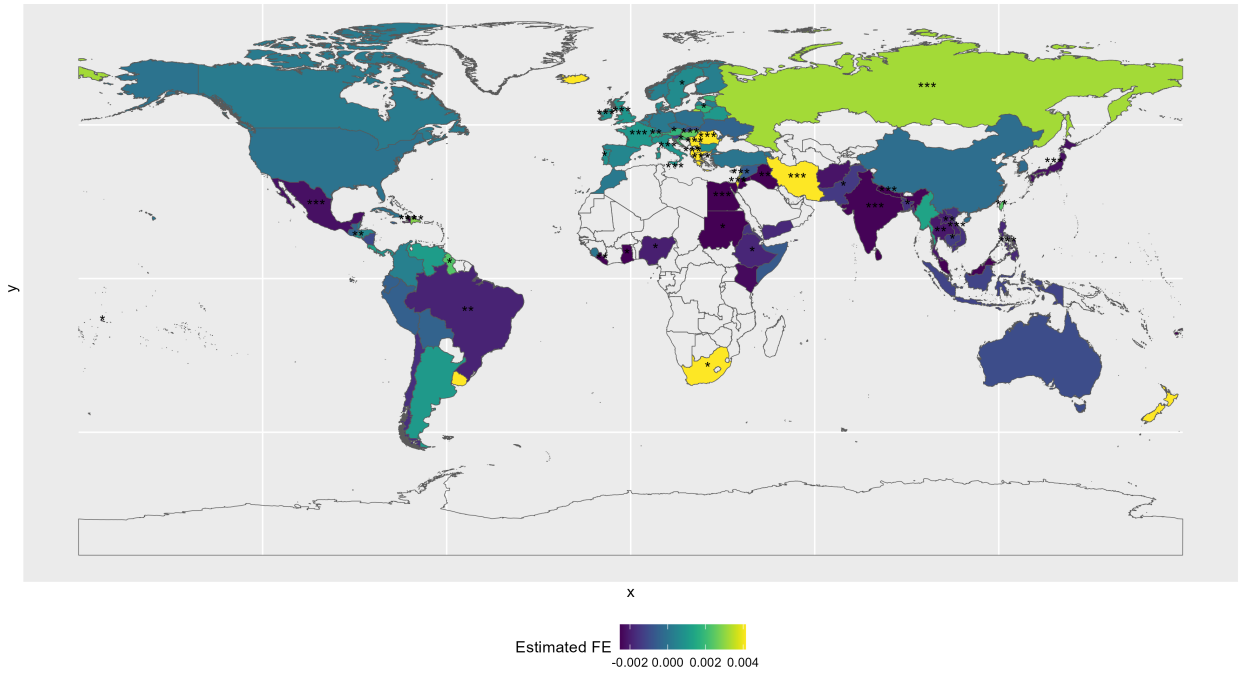


Figure 4: Estimated ancestry fixed effects; sub-national ancestries are mapped to national ones (e.g. Bavarian to German). Ancestries that cannot be mapped to existing nations (e.g. Kurds, Yugoslavians) are estimated but do not appear in this map. Blank countries have no or insufficient individuals reporting a corresponding ancestry. Stars denote statistical significance of estimated fixed effects. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

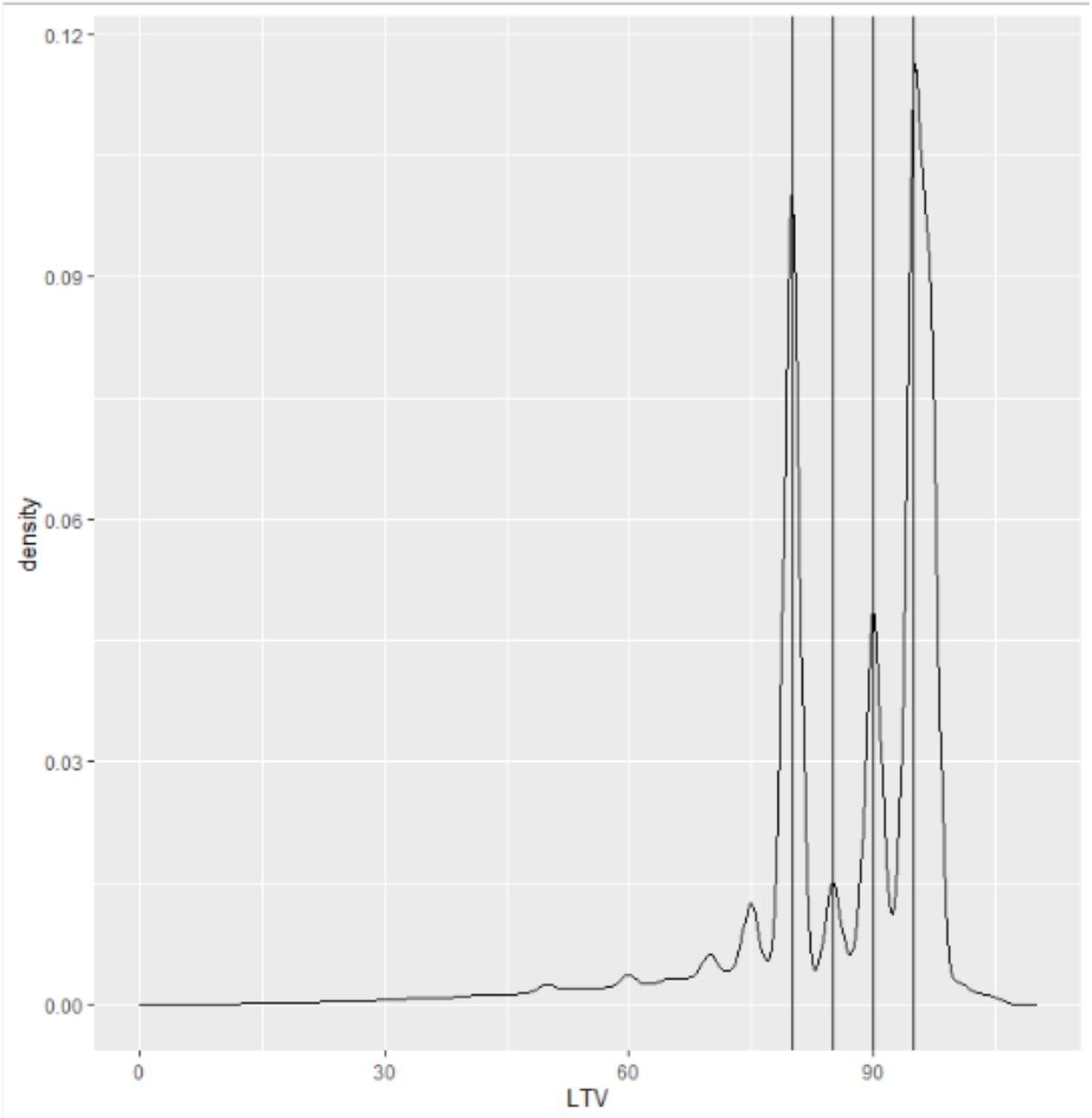


Figure 5: Distribution of LTV at mortgage origination for owner-occupiers 2018-2020, dashed lines are GSE limits; Data from HMDA mortgage microdata.

C Imputing Household Participation in Rental Real Estate in the SCF

This section discusses a concerning gap in the share of SCF respondents thare report owning residential real estate, not for recreation or vacation purposes (9%) and those that report receiving positive rent payments from such assets (4%). These differences lead to stark differences in the number of household residential RE investors. In fact, both numbers are at odds with other data on the number of rental units in the US, and the characteristics of these properties. A useful benchmark is available from 2015 Rental Housing Finance Survey - which samples rental units and properties. Lee (2017) uses these data to report that the average household investor holds 1.35 units.

Table 10 provides additional details on investors and their holdings. The average investor holds 2 properties, and the average rent collecting investor owns slightly more - 2.36. Given this, if participation in *rental* residential is 9%, an average investor would hold 0.94 units -*too low*. Yet if participation is 4%, each investor would have to hold 1.95 units - *too high*. Only 4% of investor report keeping their properties in legal entities, suggesting the definition of 'household' ownership is unlikely to be the issue. The problem appears to be under-reporting of rent: only 52% (77%) of investors that hold large multi-family properties with more than 5 units (2-4 unit properties) report any rent income, an implausibly low figure.

One solution is to impute participation as either reporting positive rent or owning at least one property with more than two units. This method includes only owners' of a single property if they report positive rent, but all owners' of multiple properties, under the logic that few households own a property with multiple units for leisure purposes. Under this definition, participation is 6.2% and the average investor holds 1.3 units - in line with Lee (2017).

Table 10: Characteristics of Investors and Holdings, by Rent Collection

	All Residential Investors		Rent Collecting Investors	
	All	of these, w' Rent	All	R/P
	Panel A: **Investor Has**			
Rent > 0	0.43	1.00	1.00	0.08
Avg # Properties	2.02	–	2.36	–
LP,LLC, etc.	0.04	0.61	0.06	–
	Panel B: **Investor Holds Any of**			
SFH	0.51	0.40	0.48	0.08
Condo	0.17	0.46	0.18	0.04
2-4 Unit	0.08	0.77	0.15	0.05
5+ Unit	0.22	0.52	0.27	0.04

Note: This table presents statistics on owners of non-recreational real estate in the SCF. Panal A tests whether investors' holdings have a certain variable: whether they report positive rent, the average number of properties, and whether they report holding the properties under an LP or LLC directly. Panel breaks the sample down by type of property own. (Note, the SCF only collects detailed information, including type, on the first two reported properties an investor owns, but asks for the existence, and rent paying status, of additional properties.) The first column includes all residential investors, and the second includes all the investors for which the variable is positive (i.e., reports an LLC, or has a single family property), and computes the share that report positive rent payments. The third column limits the sample to all investors reporting positive rent, and recomputes outcomes for that population. The fourth column reports the average rent-price ratio reported by rent-reporting investors, using total rents and the investors' assessment of their properties' value.

D Appendix - Proofs

D.1 Deriving α_i

$$
\begin{aligned}
&\log(\alpha_i) = z - \gamma\sigma_i^2 f(r_g, \Sigma) - \gamma\sigma_i^2 \Omega \alpha_i \\
&\log(\alpha_i) + \gamma\sigma_i^2 \Omega \alpha_i = z - \gamma\sigma_i^2 f(r_g, \Sigma) \\
&\alpha_i \exp\left(\alpha_i \gamma\sigma_i^2 \Omega\right) = \exp\left(z - \gamma\sigma_i^2 f(r_g, \Sigma)\right) \\
&\alpha_i \cdot \gamma\sigma_i^2 \Omega \exp\left(\alpha_i \gamma\sigma_i^2 \Omega\right) = \exp\left(z - \gamma\sigma_i^2 f(r_g, \Sigma)\right) \cdot \gamma\sigma_i^2 \Omega \\
&\alpha_i = \frac{1}{\gamma\sigma_i^2 \Omega} \cdot \mathcal{W}\left(\exp\left(z - \gamma\sigma_i^2 f(r_g, \Sigma)\right) \cdot \gamma\sigma_i^2 \Omega\right) \\
&\alpha_i = \frac{1}{\gamma\sigma_i^2} \frac{1}{\Omega} \mathcal{W}\left(\gamma\sigma_i^2 \Omega \cdot \exp\left(\underbrace{z - \gamma\sigma_i^2 f(r_g, \Sigma)}_{\equiv A}\right)\right)
\end{aligned}
\tag{16}
$$

D.2 Proposition 1

Proposition 1 follows immediately from the properties of the Lambert function: it is a concave function with $\mathcal{W}(b)' > 0, \mathcal{W}(b)'' < 0$. Since $b = \exp(z)$ increases in z, so does the Lambert function. Furthermore, since $\exp(z) > 0$ for any finite z, and $\gamma\sigma_i^2$ are positive the investor indeed always holds a positive share of housing in this model.

D.3 Proposition 2

A useful way to consider the concavity implied by the Lambert function is to define the sum of the log fundamentals:

$$
z = \underbrace{E[d] - r_f}_{\text{relative payoff}} + \underbrace{h - w_0}_{\text{relative stocks of housing and wealth}} + \sigma^2/2 \tag{17}
$$

and rewrite returns as:

$$
E[r - r_f] = \mathcal{W}\left(\gamma\sigma^2 \exp(z)\right) \tag{18}
$$

This function is clearly concave in (upper-case) Z, the product of housing fundamentals and wealth, with the following characteristics:

$$\begin{aligned} &E[r - r_f](Z = 0) = 0 \\ &\frac{\partial E[r - r_f]}{\partial Z}(Z) > 0 \\ &\frac{\partial^2 E[r - r_f]}{\partial^2 Z}(Z) < 0 \end{aligned} \tag{19}$$

This implies that investors' behavior is concave with respect to dollars - there is a diminishing marginal effect of rent and of wealth.

However, with respect to (lower-case) z, returns are *convex* and can be characterized as follows (where I've suppressed $\gamma\sigma^2$ for notational convenience)

$$\begin{aligned} &\lim_{z\to\infty} \frac{\mathcal{W}(\exp(z))}{z} \to 1 \\ &\lim_{z\to-\infty} \mathcal{W}(\exp(z)) \to 0 \\ &\frac{\partial \mathcal{W}(\exp(z))}{\partial z} > 0 \\ &\lim_{z\to\infty} \frac{\partial \mathcal{W}(\exp(z))}{\partial z} = 1 \\ &\frac{\partial^2 \mathcal{W}(\exp(z))}{\partial^2 z} \geq 0 \end{aligned} \tag{20}$$

This implies that the return function converges to a log-linear relationship at the limit where $z \to \infty$, and to a horizontal line at 0 when $z \to -\infty$. This implies that at high levels of z, when wealth is very low, or housing quantity plus rents are very high, the equilibrium function for portfolio share behaves as a log-linear relationship:

$$\begin{aligned} &\lim_{z\to\infty} E[r - r_f](z) \to \\ \log(\gamma\sigma^2) + z \quad &\\ &\log(\gamma\sigma^2) + E[d] - r_f + h - w_0 + \sigma^2/2 \end{aligned} \tag{21}$$

For these high values of z returns are **smaller** than this log-linear relationship, and only approach it as z increases. When z is small or negative, returns are bounded below at zero, making them *higher* than the log-linear relation in (21). This suggests returns can be characterized by a log-linear relationship plus an 'offset', which is negative when z is high and positive when z is low. Figure 9 visualizes the relationship of $\mathcal{W}(\exp(z))$ relative to z.

To make this argument precise, I rewrite equation (18), defining and utilizing the differ-

ence between a Lambert functions and log for any x as

$$O(x) = \mathcal{W}(x) - \log(x) \tag{22}$$

$O(x) > 0$ for any $x < e$ and negative otherwise. Log is more concave than the Lambert function, and so $O(x)$ is a downward sloping, convex function everywhere. See Figure 6. This implies returns can be written as:

$$\begin{aligned} E[r - r_f] = E[d] - r_f + h - w_0 + \sigma^2/2 + \log\left(\gamma\sigma_i^2\right) + \\ O\left(\gamma\sigma_i^2 \exp\left(E[d] - r_f + h - w_0 + \sigma^2/2\right)\right) \end{aligned} \tag{23}$$

Where $O(\cdot)$ is a non-linear offset, that works in the opposite direction to the linear term.

Here we see that returns are equal to the sum of:

- Rent fundamentals (here only $E[d]$), housing supply (h), the risk-free rate, and investor wealth.
- risk premium ($\gamma\sigma_i^2$) and σ^2.
- a term, $O(\cdot)$, that captures the excess (or dampened) concavity of the Lambert function relative to the log-linear function. This concavity term is the 'offset' described intuitively above. It depends on the risk premium, and on z, and is negative if z is positive and high, and positive if z is small and high.

We can pin down when the offset term switches from positive to negative: $O(\gamma\sigma^2, z) = 0$ when $z = 1 - \log(\gamma\sigma^2)$. At this threshold excess returns must be one, an exceedingly high figure:

$$E[r - r_f](z = 1 - \log(\gamma\sigma^2)) = \log(\gamma\sigma^2) + (1 - \log(\gamma\sigma^2)) = 1 \tag{24}$$

As a result we can bound $\gamma\sigma^2 \exp(z) < e$, which implies $r < \mathcal{W}(e) = 1$ and conclude that for empirically relevant ranges the concavity offset is positive and increases returns relative to the log-linear benchmark - keeping returns and participation higher than they would otherwise.

QED.

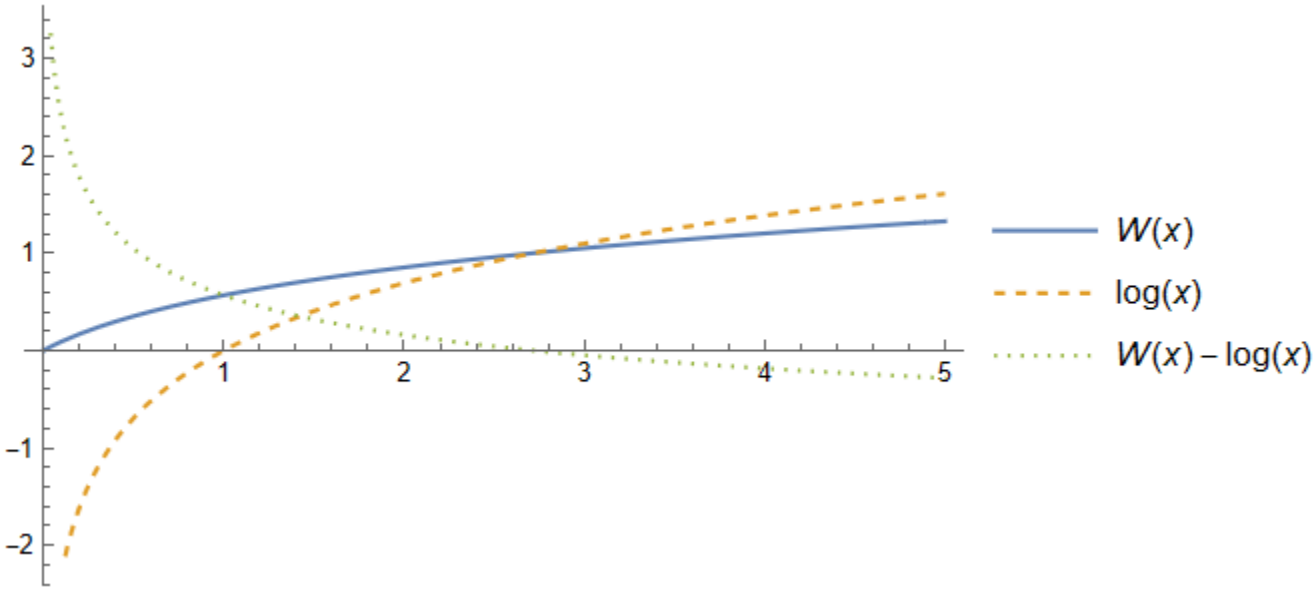


Figure 6: Visual Comparison of the Lambert function $\mathcal{W}(x)$ and $\log(x)$. $O(\cdot)$ as defined in equation (23) is the difference between the two functions.

D.4 Proposition 1

Recall that the reservation utility for renters is

$$V_i^{\text{Reservation}} = V + \boldsymbol{\alpha}_i(1 + \boldsymbol{R}_i) \tag{25}$$

Where V is constant across markets. To preserve spatial equilibrium renters in the city must not have a utility higher than the reservation utility. To clear the rental housing market, renters must not have a utility lower than the reservation utility. Utility in the city must thus make the renter indifferent to their outside option.

$$\begin{aligned} &\nu V + \boldsymbol{\alpha}_i(1 + \boldsymbol{R}_i) - D_j = \boldsymbol{\alpha}_i(1 + \boldsymbol{R}_i) + V \rightarrow \\ &d_j = v + \log(\nu - 1) \\ &d_j \sim \mathcal{N}\left(\mu_d + v, \sigma_d\right) \end{aligned} \tag{26}$$

D.5 Proposition 2

If $\gamma > 1$ and LC < 1 binds then for any r there exists a non-empty interval $\left[\underline{\psi}, \overline{\psi}\right]$ such that if $\psi \in \left[\underline{\psi}, \overline{\psi}\right]$ then $P_o > P_I(r)$, and home-ownership is limited to $W_i > W_o$, Where both P_o and W_o are independent of returns r.

Outline of proof. I first argue willingness to pay increases with wealth if $P_o < \frac{E[D]}{1+R_f} + \epsilon$ for some $\epsilon > 0$. I then show that $P_o > P_I$ if the liquidity constraint is to bind. I then argue that for a certain set of ψ a price P_o must exist in the range between $(P_I, \frac{E[D]}{1+R_f})$. To do this I show willingness to pay increases with with ψ, which allows me to write prices as a function of ψ. I then show that there necessarily exists a range $\left[\underline{\psi}, \overline{\psi}\right]$ such that if ψ is in it, then the $P_o > P_I(r)$ and the wealth of the marginal buyer do not depend on returns r.

First, note that $P_I < \frac{E[D]}{1+r_f}$ because investors must be compensated for risk ($r > r_f$.).

Lemma 1 If $P_o < \frac{E[D]}{1+R_f}$ then

Equation (27) describes the willingness to pay for each resident i.

$$E\left[\frac{((W_i - P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = \frac{(V + W_i(1+R_f))^{1-\gamma}}{1-\gamma} \tag{27}$$

Differentiating the implicit function $P_o(W_i)$ from the indifference condition

$$\begin{aligned}
&E\left[\frac{((W_i - P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = \frac{(V + W_i(1+R_f))^{1-\gamma}}{1-\gamma} \rightarrow \\
&(1+R_f)\frac{1-\gamma}{1-\gamma}(1-P_o')E[((W - P_o)(1+R_f)+Y)^{-\gamma}] = (1+R_f)\frac{1-\gamma}{1-\gamma}(V + W(1+R_f))^{-\gamma} \\
&1 - P_o' = \frac{(V + W(1+R_f))^{-\gamma}}{E[((W - P_o)(1+R_f)+Y)^{-\gamma}]} \\
&P_o' = 1 - \left(E\left[\left(\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right)^{\gamma}\right]\right)^{-1}
\end{aligned} \tag{28}$$

Where P_o' is the derivative of the willingness-to-pay with respect to wealth. To prove the lemma is equivalent to showing that

$$E\left[\left(\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right)^{\gamma}\right] > 1$$

Define $f(x)$ as x^{γ} and assume $\gamma > 1$. Since $f(x)$ is convex Jensen's inequality implies:

$$\begin{aligned}
E\left[\left(\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right)^{\gamma}\right] &> \left(E\left[\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right]\right)^{\gamma} \\
&= \left(\frac{W(1+R_f)+V}{W(1+R_f)+V} - \frac{P_o(1+R_f)+V}{W(1+R_f)+V} + E\left[\frac{Y}{V + W(1+R_f)}\right]\right)^{\gamma} \\
&= \left(1 - \frac{P_o(1+R_f)+V}{W(1+R_f)+V} + \frac{E[Y]}{V + W(1+R_f)}\right)^{\gamma}
\end{aligned} \tag{29}$$

Which is greater than 1 if

$$
\begin{aligned}
&1 - \frac{P_o(1+R_f)+V}{W(1+R_f)+V} + \frac{E\left[Y\right]}{V+W(1+R_f)} > 1 \\
&P_o(1+R_f) < E[Y-V] = E[V(\nu-1)] = E[D] \\
&P_o < \frac{E[D]}{1+R_f}
\end{aligned} \tag{30}
$$

Where in the the second line I use Proposition 1. QED.

Lemma 2

The price of owner-occupied housing is greater than the price for rental housing: $P_o > P_I$.

Prices are thus bid down by residents, down the wealth distribution. The existence of investors caps prices at a lower bound - the price that investors are willing to pay. This creates a kinked supply curve in the owner-occupied market - perfectly elastic for $P > P^I$ and perfectly inelastic otherwise. Define the implicit function $W_o^*(P_o)$ as the buyer W_i satisfying equation (27) for a given price level.

$$
E\left[\frac{((W_o^*(P_o)-P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = E\left[\frac{(V+W_o^*(P_o)(1+R_f))^{1-\gamma}}{1-\gamma}\right] \tag{31}
$$

Denote the price that both satisfies (27) and the liquidity constraint as P_o^{LC} and which price satisfies $W_o^*(P_o^{LC}) = P_o^{LC} \cdot LC$ Prices are thus $P_o = \max(P_o^{LC}, P_I)$.

Lemma 1 implies the function $W_o^*(P_o)$ is upwards sloping for $P_o < \frac{E[D]}{1+R_f}$.

Assume by way of contradiction that $P_o = P_I$, or that $P_o^{LC} < P_I$. Since $P_I < \frac{E[D]}{1+R_f}$ then $W_o^*(P_I) > W_o^*(P_o^{LC})$. This implies that liquidity constraint does not bind if prices are concave in wealth or that the liquidity constraint is violated, if prices are convex. QED. ($\rightarrow$ Show prices are concave in wealth to simplify this part)

Corrolary 1

It follows that residents with $W_i > W_o$ will purchase the unit, and otherwise will not.

It remains to be shown that for any r there exists P_o that is independent of prices paid by investors.

Lemma 3 As long as LC is not 'too high', where $LC < 1$ suffices, then $\partial P_o/\partial\psi > 0$.

Proof By implicitly differentiating $P_o(\psi)$ based on (31) we can show

$$
\frac{\partial P_o}{\partial \psi} = -\delta \frac{1}{(LC-1)E[(P_o(LC-1)(1+R_f)+Y)^{-\gamma}] - (V - P_o LC(1+R_f))^{-\gamma}} > 0
$$

Note that $E[(P_o(LC-1)(1+R_f)+Y)^{-\gamma}] > 0$, which is true by Jensen's inequality:

$$E[(P_o(LC-1)(1+R_f)+Y)^{-\gamma}] > (E[(P_o(LC-1)(1+R_f)+Y))^{-\gamma} = (E[Y]-P_o(1-LC)(1+R_f))^{-\gamma} > 0$$

In words: the sum left to be paid by the owner in the second period must be lower than the expected income in the second period. This is satisfied by the conditions in lemma 1.

The denominator of $\partial P_o/\partial\psi$ is negative as long as LC is not 'too high'. In particular, it is enough for $LC < 1$ to satisfy this requirement. Prices thus rise with ψ.

Intuitively prices increase with ψ because it increases the utility from ownership. Higher prices also increase the wealth of the marginal buyer, which increases the outside utility. As long as this increase is not too large (LC isn't too high) the net effect of ψ on prices is positive.

We can define $\underline{\psi}, \bar{\psi}$ such that for any investor price P_I the following holds:

$$P_o(\bar{\psi}) = \frac{E[D]}{1+R_f}$$
$$P_o(\underline{\psi}) = P^I$$

Following from $P^I < \frac{E[D]}{1+R_f}$ and lemma 3 that $\underline{\psi} < \bar{\psi}$, and that $\exists\psi \in (\bar{\psi}, \underline{\psi})$ such that P_o is higher than investors' prices but still high enough to satisfy lemma 1. QED

QED for proposition 2.

I next provide a graphical analysis of this proof. Intuitively, the parameter ψ guarantees owner prices can be set in a band between their risk-neutral discounted value (expected rent) and the price paid by investors. These investors would price the unit by less than the owners since they too fact the same risk, but do not benefit from ψ[30]

The reader will note that this proof is an 'if' statement. The use of Jensen's inequality in equation (29) guarantees the solution under the described conditions

Figure 7 visualizes this - the dashed line is a hypothetical reservation utility. Prices would be at the intersection of that line and the lowest wealth level, in this case $w_i = 0.6$ for $p \approx 1.1$.

This introduces a wedge between prices paid by owners to those paid by investors. To rationalize this I assume that homeowners purchase first. House prices are bid down to P_o, and all buyers that want to and can afford to purchase their house. After these sales are finalized, investors purchase houses for rent. Prices for remaining homes are bid down to P^I.

[30]since they are also residents they benefit from ψ in their own units, but do not do so again when buying a rental property.

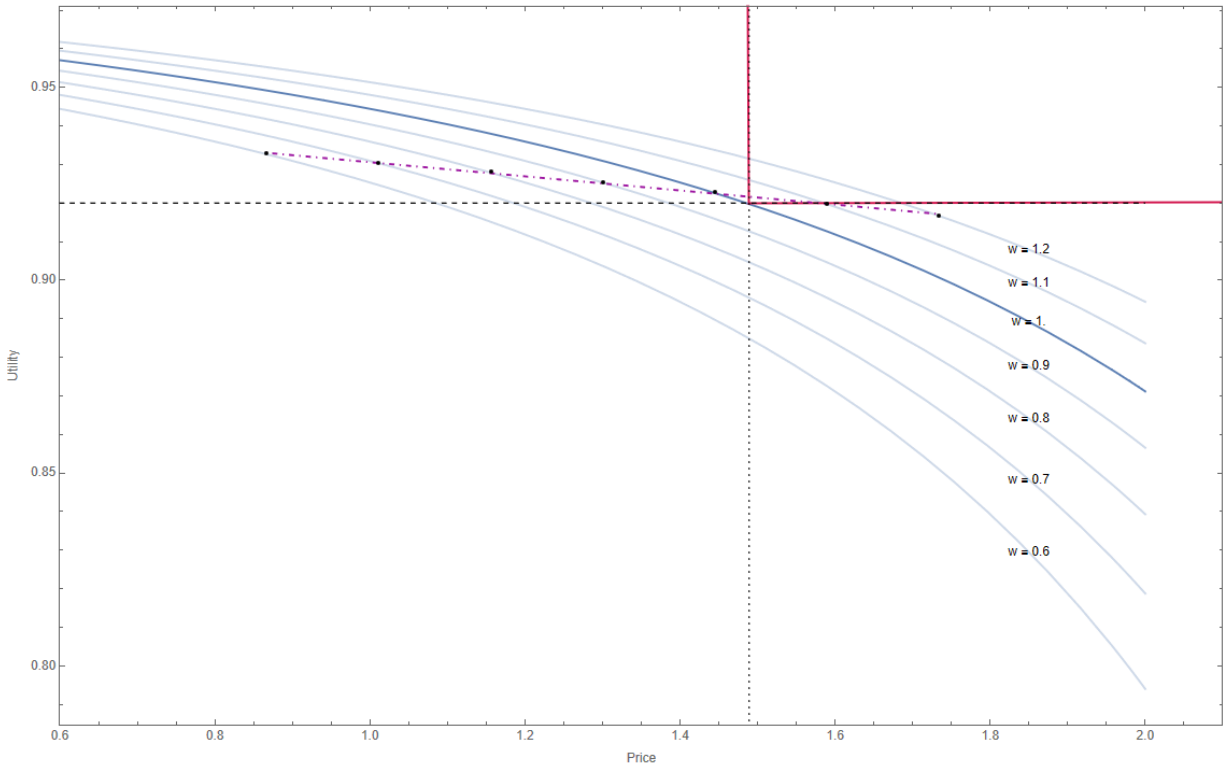


Figure 7: Visualization of residents' willingness to buy housing at different wealth levels. Each solid blue line is an iso-wealth line mapping prices to utility. The dashed line represent a hypothetical reservation utility, and the dotted line represents the price investors are willing to pay. The maroon lines denote the kinked supply curve facing buyers. The points represent combinations of price and wealth at which a liquidity constraint of $LC = 0.69$ binds.

The marginal buyer, denoted W_o satisfies:

$$W_o/P_o = LC \tag{32}$$

This condition can hold for any wealth group, depending on the price. This is visualized in Figure 7 as solid points along the iso-wealth utility curves. The liquidity constraint condition holds for any of these iso-wealth curves from those points and towards the left. This implies that the point $(U(w = 1, P^I), P^I)$ is no longer a possible equilibrium, as it is to the right of the solid point on the $w = 1$ line. Pinning down the price is the same bidding-down process: owner-occupied prices are set so that the marginal buyer W_o is indifferent to renting while satisfying the liquidity constraint. In Figure 7 this new price is at intersection of the dashed line and the liquidity constraint line (the dotted-dashed line).

$$E\left[\frac{(P_o(LC-1)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = E\left[\frac{(V+P_o\cdot LC(1+R_f))^{1-\gamma}}{1-\gamma}\right] \tag{33}$$

D.6 Proposition 3

If the conditions in propositions 1 and 2 hold, then participation is limited to $W_i > W_m$, and investment shares α_i rise in wealth and decline in κ_i; specifically, the set of investors is defined as the double integral over wealth on the interval $[W_m, \infty)$ and κ on the interval $[0, \kappa^*(W)]$, where $\kappa^*(W)$ is a monotonically increasing function in W.

Proof The proof proceeds in three steps. First, I find the set of marginal investors by describing a function $\kappa^*(w_i)$ which maps from log-wealth to the maximum κ that (potential) investors can sustain. Second, I show this function monotonically increases in wealth, has a single crossing with $\kappa = 0$ and conveniently converges to a limit for high values of w. Third, I use the function $\kappa^*(w_i)$ to describe the set of active investors

I identify the set of investors as the complement of the set of non-investors, for whom κ_i, ρ_i are such that the optimal position is negative, $\alpha_i < 0$.

$$\begin{aligned} &\alpha_i < 0 \\ &\left(E[r - r_f] + \frac{1}{2}\sigma_r^2 - \kappa_i\right)(1 + \exp(\tilde{w}_i - \mu_\nu - v - \kappa_i + \bar{r}_p)) < \gamma\sigma_{yr} \end{aligned} \tag{34}$$

Since ρ_i is a function of both w_i, κ_i this inequality defines a set $\{w, \kappa\}$ which makes up the non-investors.

$$\{w, \kappa\} = \{\tilde{w}_i, \kappa_i | \alpha(\tilde{w}_i, \kappa_i) < 0\} \tag{35}$$

Lemma 1 There exists a single-valued function $\kappa^*(\tilde{w}_i)$ for which $\alpha(\tilde{w}_i, \kappa^*(\tilde{w}_i)) = 0$, which defines the groups investors and non-investors for each wealth level.

I will show that α monotonically declines with κ_i for any given $\tilde{w}_i$, implying that $\kappa^*(\tilde{w}_i)$ is indeed a single-valued function.

$$\begin{aligned} \frac{\partial \alpha}{\partial \kappa_i} &= \frac{1}{\rho_i^2}\rho_i' \frac{1}{\gamma\sigma_r^2}\left(E[r - r_f] + \sigma_r^2/2 - \kappa_i - \gamma\sigma_{yr}\right) - \frac{1}{\rho}\frac{1}{\gamma\sigma_r^2} \\ &< \frac{1}{\rho_i^2}\rho_i' \frac{1}{\gamma\sigma_r^2}\left(E[r - r_f] + \sigma_r^2/2 - \kappa_i - \gamma\sigma_{yr} + \rho\sigma_{yr}\right) - \frac{1}{\rho}\frac{1}{\gamma\sigma_r^2} \\ &= \frac{1}{\rho_i^2}\rho_i'\left(\rho_i\alpha_i\right) - \frac{1}{\rho}\frac{1}{\gamma\sigma_r^2} \\ &= \frac{1}{\rho}\left(\rho_i'\alpha_i - 1\right) \end{aligned} \tag{36}$$

Where $0 < \rho_i' < 1$ is the change in the marginal consumption from wealth at the marginal dollar of wealth. This result implies that, intuitively, α_i declines with κ_i. This is true as long as α isn't greater than some threshold that we know to be larger than $\alpha = 1$. In particular,

this is true at the margin around non-investment ($\alpha = 0$).[31]

I thus treat $\kappa^*(\tilde{w}_i)$ as a function, mapping from wealth to the threshold of κ_i below which an individual invests in real estate.

It immediately follows that at any wealth level w_i the population is split into investors ($\kappa_i \leq \kappa^*(\tilde{w}_i)$) and non-investors ($\kappa_i > \kappa^*(\tilde{w}_i)$). QED.

Lemma 2 $\kappa^*(\tilde{w}_i)$ monotonically increases in wealth and has a single crossing with $\kappa = 0$.

Implicitly differentiating $k^*(\tilde{w}_i)$ at $\alpha_i = 0$

$$\begin{aligned} &\alpha(\tilde{w}_i, k^*(\tilde{w}_i)) = 0 \rightarrow \\ &E[r - r_f] + \frac{\sigma_r^2}{2} - k^*(\tilde{w}_i) - (1 - \rho_i)\gamma\sigma_{yr}^2 = 0 \rightarrow \\ &\frac{\partial k^*(\tilde{w}_i)}{\partial \tilde{w}_i} = \frac{\rho_i' \gamma \sigma_{yr}}{1 + \rho_i' \gamma \sigma_{yr}} > 0 \end{aligned} \tag{37}$$

To find $\kappa^*(\tilde{w}_i) = 0$ we set $\alpha(\tilde{w}_i,0)$

$$\begin{aligned} &E[r - r_f] + \sigma_r/2 = \gamma\sigma_{yr}(1 - \rho(w_i, 0) \rightarrow \\ &w_m = w_i(\alpha_i = 0, \kappa^*(w_i) = 0) = E[d] - r_f + \log\left(\frac{\gamma\sigma_{yr}}{E[r - r_f] + \sigma^2/2} - 1\right) \end{aligned} \tag{38}$$

Which exists as long as $2 \cdot \gamma\sigma_{yr} > E[r - r_f] + \sigma^2$. Since $\sigma_{yr} \approx \sigma$ this holds as long as returns aren't too high relative to risk. QED.

$k^*(\tilde{w}_i)$ increases with wealth as long as the correlation of income and returns σ_{yr} is positive, which is the case per Proposition 1.[32]

This positive slope is true regardless of the magnitude of σ_{yr} This is of note, because the standard result is that it is the magnitude of σ_{yr} that determines whether correlated income tilts investors to or away from risky assets. Generally, lower wealth means income is relatively higher, crowding out riskless investing and tilting portfolios towards risky assets. There is an offsetting effect due to the correlation of income and consumption, which if big enough, can cause households with higher income (or lower wealth) to tilt away from risky

[31] At very high exposure levels the indirect effect of κ_i on ρ offsets the direct effect, but this is never an issue for non-investors and those on the margin of investing.

[32] The stochastic properties of returns only depend on second period variables, as prices are known and set in the first period, implying the covariance between returns and income is necessarily positive.

$$\sigma_{yr} = \text{Cov}(r, y) = \text{Cov}(d - p, y) = \underbrace{\text{Cov}(d, y)}_{>0} - \underbrace{\text{Cov}(p, y)}_{=0} = \text{Cov}\left(\log(\nu - 1) + \epsilon_z + \epsilon_j, \log(\nu) + \epsilon_z + \epsilon_j\right) \approx \iota'\Sigma\iota$$

for reasonable values of μ, σ.

assets.

However, at $\alpha = 0$ the former effects cancels out[33] and we are left only with the latter offsetting effect. This implies that wealthier households *at the margin* require a lower rate of return and are willing to sustain higher monitoring costs.

$$\kappa^*(\tilde{w}_i) : E[r - r_f] - \kappa^*(\tilde{w}_i) + \sigma_r^2/2 = \gamma(1 - \rho(\tilde{w}_i, \kappa^*(\tilde{w}_i))\sigma_{yr} \tag{39}$$

Finally, the value $\kappa^*(\tilde{w}_i)$ is capped as $\tilde{w}_i$ goes to infinity. Since the function for α_i is smooth we can rely on limits algebra and find

$$\lim_{w_i \to \infty} k^*(w_i) = E[r - r_f] + \sigma^2/2 \tag{40}$$

The set of investors can thus be characterized as those residents with wealth $w_i > w_m$ and $\kappa \in [0, k^*(w_i)]$, which converges to $[0, E[r - r_f] + \sigma^2/2]$ at very high levels of wealth.

Figure 8 visualizes the set of investors, demonstrating how the set of investors increases with wealth and declines in κ_i. The function $\kappa^*(w_i)$ can be seen as the line separating the green-from the red. As the figure makes clear, we can therefore characterize the set of investors as a double integral: the inner integral over κ_i from 0 to $\kappa^*(\tilde{w}_i)$ and the outer integral over wealth w_i.

QED.

D.7 Proposition 4

The equilibrium condition in the capital market is that the total capital supplied by all investors equals the total value of the rental housing stock. Equation 41 imposes the assumptions of lemmas 1 and 2 above and uses proposition 3 to aggregate capital supply across the set of active investors investors

$$\underbrace{\int_w \int_\kappa \alpha(w,\kappa) \cdot \mathbb{1}\{\alpha(w,\kappa) \geq 0\} \cdot \tilde{W}_i dF_{w,\kappa}}_{\text{capital supply}} = \underbrace{P \cdot H^{\text{rental}}}_{\text{capital 'demand'}} \tag{41}$$

$$\int_{w_m}^{\infty} \int_0^{\kappa^*(w)} \alpha(w_i, \kappa_i, r) \cdot (\exp(\tilde{w}))\, dF_{w,\kappa} = \exp(\mu + v - E[r]) \cdot F_w(lc - p_o)$$

[33] This can be seen (at $\alpha_i = 0$) by multiplying both sides by ρ_i, cancelling out its effect.

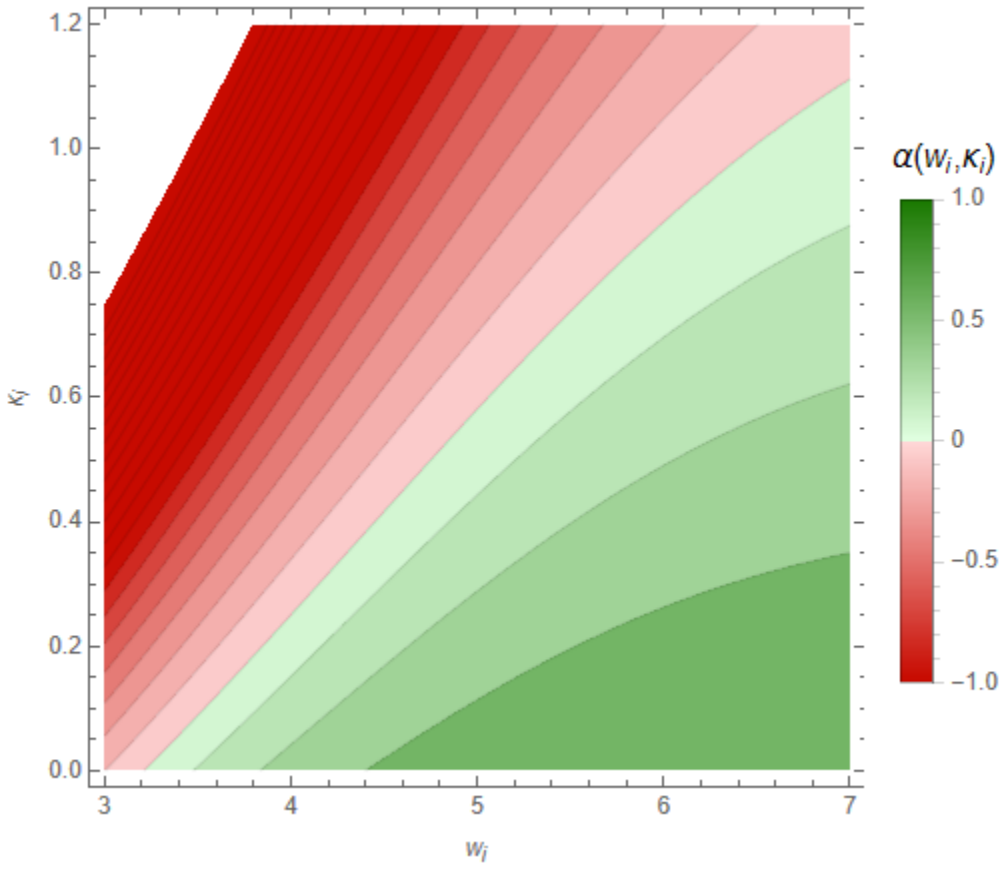


Figure 8: Simulation of $\alpha(\tilde{w}_i, \kappa_i)$, setting $E[r - r_f] = 0.11, \mu_\nu = v + \bar{r}_p = 4, \gamma = 5, \sigma_r^2 = 0.7, \sigma_{yr} = 0.4$

Proposition 3 also tells us that

$$\frac{\partial \alpha(w,k,r)}{\partial w} > 0 \qquad \frac{\partial \kappa^*(w)(\cdot)}{\partial w} > 0 \tag{42}$$

We want to show that r decreases if there is a rightward shift in the distribution F_w for the region $[w_m, \infty]$. Suppose we have a rightward shift in the wealth distribution of the investor group, i.e. holding the investor group constant. Assume first that wealth increases for only one group, w_i which satisfies $w_i > w_m$. Specifically, suppose that for each resident in w_i we increase wealth by $\epsilon > 0$: $w_i' = w_i + \epsilon$. For each i we can derive the change in the portfolio share for $\epsilon \to 0$

$$\frac{d\alpha(w_i,\kappa_i,r)}{dw_i} = \frac{\partial \alpha(w_i,\kappa_i,r)}{\partial w_i} + \frac{\partial \alpha(w_i,\kappa_i,r)}{\partial r}\frac{\partial r(w_i)}{\partial w_i} \tag{43}$$

Where $\frac{\partial r}{\partial w_i}$ is the change in the equilibrium returns for a marginal increase in wealth for the group i. Proposition 2 implies $\frac{\partial \alpha(w_i,\kappa_i,r(w_i))}{\partial w_i} > 0$ and $\frac{\partial \alpha(w_i,\kappa_i,r(w_i)}{\partial r} > 0$ Furthermore, from Lemma 2 above we know that

$$\frac{\kappa^*(w_i)}{\partial w_i} > 0$$

Which implies that the share of residents who belong to wealth group w_i increases with their wealth.

Assume, for the sake of contradiction that $\frac{\partial r}{\partial w_i}$. It follows that

$$\alpha(w_i', \kappa_i, r(w_i') \geq \alpha(w_i, \kappa_i, r(w_i)) \tag{44}$$

and strictly for $\kappa_i < \kappa^*(w_i')$ - the subset of resident who had wealth w_i (now w_i'). It follows that the total capital supplied by the group i increases at the new wealth level. Notice that the integral is condition on having w_i wealth to begin with.

$$\int_0^{\kappa^*(w_i')} \alpha(w_i', \kappa_i, r(w_i))\cdot(\exp(w_i' - P_O))\, dF_{\kappa|w_i} > \infty \int_0^{\kappa^*(w_i)} \alpha(w_i, \kappa_i, r(w_i))\cdot(\exp(w_i - P_O))\, dF_{\kappa|w_i}$$

Furthermore, since we assumed that $r(w_i') \geq r(w_i)$ all other investors weakly increase their capital supply. Because we can sum over original wealth distribution on any order we can

break down the :

$$\int_{w_m}^{\infty} \int_{0}^{\kappa^*(w)} \alpha(w_i, \kappa_i, r) \cdot (\exp(\tilde{w}))\, dF_{w,\kappa} = \int_{w \in [w_m, \infty] \setminus w_i} \int_{0}^{\kappa^*(w)} \alpha(w_i, \kappa_i, r) \cdot (\exp(w - p_O))\, dF_{\kappa,w} + \int_{0}^{\kappa^*(w_i)} \alpha(w_i, \kappa_i, r(w_i)) \cdot (\exp(w_i - P_O))\, dF_{\kappa|w=w_i} F_{w=w_i} \tag{45}$$

Where the second line is the total capital supplied by all the groups except for w_i and the third is the capital supplied by the w_i group. Since the former is weakly larger and the second is strictly larger for w_i' than for w_i it follows that the total capital supplied increases when wealth for group i increases.

Imposing equilibrium in both cases we can write:

$$\exp(\mu + \nu - r(w_i'))F_{w'}(lc - P_O) > \exp(\mu + \nu - r(w_i))F_w(lc - P_O) \tag{46}$$

μ is pinned down by the renters, and is unaffected by $w_i' > w_i$; ν is given; and since all investors are owners from Proposition 2 $F_{w'}(lc - P_O) = F_w(lc - P_O)$. It must be then that $r(w_i') < r(w_i)$ Contradictory to our assumption. It follows that if wealth increases for set of investors, equilibrium returns must strictly decline.

It follows that $\frac{\partial r(w_i)}{\partial w_i} < 0$.

A marginal increase in one group's wealth triggers a positive direct effect on that group's portfolio share and increase the share of investors out of that wealth group. However, by lowering returns, there's an indirect decrease in portfolio and participation shares across all groups.

Of these two effects, the direct one is larger. We know this by applying the same logic, in reverse: if the returns shrink with a marginal increase in wealth then capital demand increases. In equilibrium, total capital supply must *increase* as well.

Since the effects are aggregative, a marginal increase in wealth across *all* groups can be broken down into a series of increases across all groups. As we've shown, each such increase, regardless of the group's wealth, increases total supply and decreases returns. It follows that a marginal increase in the wealth of all investors reduces local returns.

QED.

D.8 Lemma 3

Assuming $LC < 1$ prices paid by investors are lower than those paid by owner occupiers if and only if the marginal homebuyer has lower wealth than the marginal investor. To show this, I first show that if $P_I < P_o$ then $W_m < W_o$. Lemma 2 concludes that under the assumption that $P_I < P_o$ the investor's marginal wealth is

$$W_m - P_o = E[d] - r_f + \log\left(\frac{\gamma\sigma_{yr}}{E[r - r_f] + \sigma_r^2/2} - 1\right) > 0$$

Where the inequality implies that the investor does not make negative returns. Substituting $W_o/P_o = LC \leq 1$ implies

$$W_m - P_o = W_m - W_o/LC > 0 \rightarrow W_m > W_o$$

Next, suppose that $W_m > W_o$. As with proposition 2, we first assume that P_o is independent of returns, and then prove that there is a range in which it is. Given this assumption, we can derive W_m as previously. Next,

D.9 Compariative Static of D/P

Consider $x \in \{E[d], h, w_0, r_f\}$

$$\begin{aligned}
\frac{\partial \log(D/P)}{\partial x} &= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x} + \frac{\partial O(\gamma\sigma_i^2 \exp(z))}{\partial x}\gamma\sigma_i^2 \exp(z)\frac{\partial z}{\partial x} \\
&= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x} + \left(-\frac{1}{\gamma\sigma_i^2 \exp(z)} + \frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{\gamma\sigma_i^2 \exp(z)\left(1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))\right)}\right)\gamma\sigma_i^2 \exp(z)\frac{\partial z}{\partial x} \\
&= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x} + \left(-1 + \frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))}\right)\frac{\partial z}{\partial x} \\
&= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x}\frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))}
\end{aligned} \tag{47}$$

So for example for $x = w_0$

$$\frac{\partial \log(D/P)}{\partial w_0} = -\frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))} < 0 \tag{48}$$

D.10 Discrete Case: Derivation & Parameters

Denote the probabilities of $\{0, w^L, w^H\}$ as $\{Pr_w^0, 1 - Pr_w^0 - Pr_w^H, Pr_w^H\}$ and of $\underline{\kappa}, \bar{\kappa}\}$ as $\{Pr_k^{\underline{\kappa}}, Pr_k^{\bar{\kappa}}\}$ The equilibrium condition for the discrete case model is

$$
\begin{aligned}
&\alpha(w^H, \underline{\kappa}) \cdot \exp(\tilde{w}^H) \cdot \underbrace{Pr_w^H Pr_k^0}_{\phi^I} = \exp(p) \cdot \underbrace{Pr_w^0}_{\phi^R} \\
&\alpha(W^H, \underline{\kappa}) = \frac{1}{\rho_I} \frac{E[r - r_f] - \underline{\kappa} + \frac{1}{2}\sigma_r^2}{\gamma \sigma_r^2} - \frac{1 - \rho_I}{\rho_I} \beta_{yr} \\
&p = \log(E[D]) - E[r] = \mu + v + \frac{1}{2}\sigma^2 - E[r] \\
&\left(\frac{1}{\rho_I} \frac{E[r - r_f] - \underline{\kappa} + \frac{1}{2}\sigma_r^2}{\gamma \sigma_r^2} - \frac{1 - \rho_I}{\rho_I} \beta_{yr} \right) \cdot \exp(\tilde{w}) \phi^I = \exp(\mu + v + \frac{1}{2}\sigma^2 - E[r]) \cdot \phi^R
\end{aligned}
\tag{49}
$$

Where $\phi^I = Pr_w^H Pr_k^{\underline{\kappa}}$ is the share of investors, and ϕ^R is the share of renters, $1 - Pr_w^H$ We can write the propensity to consume out of wealth of investors as ρ^I:

$$
\rho^I = \frac{\exp(w^H - \mu - v - \alpha(E[r] - \underline{\kappa}) - (1 - \alpha) r_f}{1 - \exp(w^H - \mu - v - \alpha(E[r] - \underline{\kappa}) - (1 - \alpha) r_f)}
$$

Note that these distributional assumptions implicitly shut down the marginal impact of r on the number of renters. This would be true until the point at which p is so high that group w^L also become renters. Simplified discrete wealth distribution miss this channel.

To solve this define

$$
x = E[r - r_f] - \underline{\kappa} + \sigma_r^2/2 - (1 - \rho^I)\gamma \sigma_{yr} \tag{50}
$$

We can rewrite equation (49):

$$
\begin{aligned}
&\frac{1}{\gamma \sigma_r^2} \left(E[r - r_f] - \underline{\kappa} + \sigma^2/2 - (1 - \rho^I)\gamma \sigma_{yr} \right) (W^H - P_o) \phi^I = \rho^I \exp(\mu + v + \sigma^2/2 - E[r]) \cdot \phi^R \\
&\frac{1}{\gamma \sigma_r^2} x \phi^I = \rho^I \phi^R \exp(\mu + v + \sigma^2/2 - \tilde{w^H} - E[r]) \\
&x \exp(x) = \gamma \sigma_r^2 \rho^I \cdot \frac{\phi^R}{\phi^I} \exp\left(\mu + v + \sigma^2 - \tilde{w^H} - r_f - \underline{\kappa} - \gamma \sigma_{yr}(1 - \rho^I) \right) \\
&x = \mathcal{W}\left(\gamma \sigma_r^2 \rho^I \cdot \frac{\phi^R}{\phi^I} \exp\left(\mu + v + \sigma^2 - \tilde{w^H} - r_f - \underline{\kappa} - \gamma \sigma_{yr}(1 - \rho^I) \right) \right)
\end{aligned}
\tag{51}
$$

From which we derive both returns and real estate portfolio share

$$E[r - r_f] + \sigma^2/2 = \underline{\kappa} + \gamma\sigma_{yr}(1-\rho^I) + \mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp\left(\mu + v + \sigma^2 - \tilde{w^H} - r_f - \underline{\kappa} - \gamma\sigma_{yr}(1-\rho^I)\right)\right)$$
$$\alpha = \frac{1}{\rho^I\gamma\sigma_r^2}\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp\left(\mu + v + \sigma^2 - \tilde{w^H} - r_f - \underline{\kappa} - \gamma\sigma_{yr}(1-\rho^I)\right)\right) \tag{52}$$

Prices for investors are thus

$$p = \mu + v + \sigma^2/2 - \underline{\kappa} - (1-\rho^I)\gamma\sigma_{yr} - \mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp\left(\mu + v + \sigma^2 - \tilde{w^H} - r_f - \underline{\kappa} - \gamma\sigma_{yr}(1-\rho^I)\right)\right) \tag{53}$$

This solution is of the same form as that of the first variant of the model. Wealth enters negatively in z, implying that higher wealth *lowers* returns and increases prices. This model highlights that this result is true even if investors self-sort, and in-fact this channel is reinforced if investors are a smaller group, or if renters are a larger share.

Finally, we specify the conditions for the various assumptions made above:

- w^L are home owners, but do not invest even if$\kappa = \underline{\kappa}$. This is a condition on lc, and is given by $p_o > p^I \rightarrow$

$$p(W^L, \psi) > \mu_\nu - \underline{\kappa} + \sigma_r^2/2 + (1-\rho^I)\sigma_{yr} - \mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right)$$

- w^L do not invest even if $\kappa = \underline{\kappa}$, this is a condition on w^L, as well as the return fundamentals (z, ϕ^R, ϕ^I, etc.)

$$\begin{aligned} &\alpha(w^L, \underline{\kappa}) < 0 \\ &\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right) < \gamma\sigma_{yr}(\sigma_r^2(1-\rho(w^L, \underline{\kappa})) - (1-\rho(w^H, \underline{\kappa})) \end{aligned} \tag{54}$$

  Note that $\rho(w^L, 0) < \rho(w^H, 0)$, and that this condition becomes easier to satisfy the higher w^H and the lower w^L are.

- w^H invest if $\kappa = \underline{\kappa}$, a condition which is satisfied as long as σ_r^2 isn't too big

$$\begin{aligned} &\alpha(w^H, \underline{\kappa}) > 0 \\ &\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right) > \gamma\sigma_{yr}(1-\rho(w^H, 0))(\sigma_r^2 - 1) \end{aligned} \tag{55}$$

- w^H do not invest if $\kappa = \bar{\kappa}$

$$\begin{aligned} &\alpha(w^L, \bar{\kappa}) < 0 \\ &\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right) < \bar{\kappa} + \gamma\sigma_{yr}(\sigma_r^2(1-\rho(w^H,\bar{\kappa})) - (1-\rho(w^H,0)) \end{aligned} \tag{56}$$

Note that $\rho(w^H, \bar{\kappa})$ shrinks as $\bar{\kappa}$ grows, making the condition easier to satisfy.

E Simplified Model - Exogenous Rent Setting; No Homeownership for Investors

This Appendix presents a simplified version of the model, which emphasizes the asset pricing components and their technical details, while downplaying some of the housing market details.

I collapse heterogeneity into two distinct groups, renters and investors. I further abstract away from the owner occupier market and do not specify the rent-setting process, other than a function of observables and a stochastic error. I also assume rent is independent from investors' consumption. More precisely, I allow rents D to follow an exogenous stochastic process, set $Y = 0$, and $W_i = W$.

This is a portfolio choice problem with a fixed supply of an asset in which investors price returns. There are also $n-1$ other risky assets, each with an exogenous return. All assets have an exogenous variance-covariance matrix Σ.

E.1 Model

The model demonstrates how in a setting where investors both price the asset, but also choose exposure levels, their wealth level become a critical component of returns and local risk matters. The concavity of the solution also delivers excess dispersion in prices relative to rents, as is observed in housing markets.

Investor's utility function is:

$$\begin{aligned} U &= E_t\left[\frac{W_1^{1-\gamma}-1}{1-\gamma}\right] \\ &s.t. \\ W_1 &= (1+R_p)W_0 \\ R_p &= \alpha R + (1-\alpha\cdot\iota)R_f \end{aligned} \tag{57}$$

Where α and R are $n \times 1$ vectors, and ι is a vector of 1's.

Rents are distributed log-Normal, with mean $E[d]$ and variance σ_g^2. Lower-case letters denote the log version of their upper-case counterparts.

The asset allocation problem is given as usual by the Euler equation, which implies returns, excess of the risk free rate, are a function of exposure (α):

$$E[r - r_f] + \frac{\sigma^2}{2} = \gamma \Sigma \alpha \tag{58}$$

I solve for α_1, r_1, the portfolio share and (endogenous) return in the local asset. Define the vectors $\tilde{\alpha}, \tilde{r}, \tilde{\sigma}$ as the $n-1 \times 1$ portfolio shares, returns, and variances excluding the local asset, and $\tilde{\Sigma}$ as the $n-1 \times n-1$ variance-covariance matrix excluding the first row and column of the local asset. I also define Σ_1 as a $n-1 \times 1$ vector with the covariance terms between the local and non-local assets (σ_{1j}; the first column of Σ excluding the first term, which is the return variance of the first asset, termed σ_1).

Solving for $\tilde{\alpha}$ we get the capital *supply* function, ie the amount of capital supplyied to the local capital market by investor:

$$\begin{aligned} &E[\tilde{r} - r_f] + \tilde{\sigma}^2/2 = \gamma \tilde{\Sigma} \tilde{\alpha} + \gamma \alpha_1 \Sigma_1 \rightarrow \\ &\tilde{\alpha} = \frac{1}{\gamma} \Sigma_1^{-1} \left(E[\tilde{r} - r_f + \tilde{\sigma}^2/2] \right) - \alpha_1 \tilde{\Sigma}^{-1} \Sigma_1, \end{aligned} \tag{59}$$

Which implies the relationship between the local asset's returns and its portfolio share is:

$$\begin{aligned} E[r_1 - r_f] + \sigma_1^2/2 &= \gamma \alpha_1 \left(\sigma_1^2 + \Sigma_1' \tilde{\Sigma}^{-1} \Sigma_1 \right) + \Sigma_1' \tilde{\Sigma}^{-1} E[\tilde{r} - r_f + \tilde{\sigma}^2/2] \\ &= \gamma \alpha_1 \left(\sigma_1^2 + \Omega \Sigma_1 \right) + \Omega \tilde{Er} \end{aligned} \tag{60}$$

The second line simplifies notation with the following definitions $\Omega = \Sigma_1' \tilde{\Sigma}^{-1}$ and $\tilde{Er} = E[\tilde{r} - r_f + \tilde{\sigma}^2/2]$

This function is the investors' demand for real estate, but can also be seen as the supply of capital: a higher return implies a higher quantity α.

E.1.1 Rental Property's Demand for Capital / Supply of the Asset

By definition, prices are equal to the expected discounted rent payment $p = E[d] - r$. To clear the market for rental units, the value of the investor's local investment must equal the value of the properties they hold.

$$\begin{aligned} \alpha_1 W_0 &= PH \\ \log(\alpha_1) + w_0 &= E[d] - r + h \end{aligned} \tag{61}$$

This function is the capital demand function: a higher return implies less demand for capital – as the value of the real estate sector declines, and demand for capital by the real estate

sector declines. (Note: this differs from the investors demand for the asset, demand is used here to imply the side of the market that utilizes capital)

E.1.2 Equilibrium in Capital

I solve for equilibrium in the capital market based on (58) and (61), which is analogous to solving for quantities in any demand-and-supply framework. See Appendix D.1 for full derivation.

$$\begin{aligned}\alpha_1 &= \frac{1}{\gamma\left(\sigma_1^2+\Omega\Sigma_1\right)}\mathcal{W}\left(\gamma\left(\sigma_1^2+\Omega\Sigma_1\right)\cdot\exp(E[d]-r_f+h-w_0+\sigma_1^2/2-\Omega\tilde{E}r)\right)\\ E[r_1-r_f] &= \mathcal{W}\left(\gamma\left(\sigma_1^2+\Omega\Sigma_1\right)\cdot\exp\left(E[d]-r_f+h-w_0+\sigma_1^2/2-\Omega\tilde{E}r\right)\right)+\Omega\tilde{E}r,\end{aligned} \tag{62}$$

where $\mathcal{W}$ is the Lambert function, defined as $ae^a = b \rightarrow a = \mathcal{W}(b)$.

A few key differences from standard asset pricing theory emerge. Returns and portfolio shares increase in expected rent payments ($E[d]$), as the capital supply increases less than 1 for 1 with increasing rents. If the asset supply is large, returns and portfolio shares must also be large, to incentivize the investors to purchase all of them. Return and portfolio shares fall in wealth, as a higher wealth implies additional capital for any level of portfolio share – implying that shares must be lower to preserve equilibrium. Returns must therefore also fall to incentivize lower shares.

Finally, other assets matter. Higher returns on other assets unambiguously lower portfolio shares in real estate, but increase returns, as the assets are still substitutes. The covariance of other assets acts to reinforce the effect of volatility.

I turn to characterizing the equilibrium return, and particularly its relationship to local fundamentals and how they compare to a log-linear benchmark which does not have the Lambert function.

Proposition 1

α_1 and r_1 are strictly positive, and monotonically increase in (a) rent expectations ($E[d]$), (b) the housing stock h; α and r decline in (c) local investor wealth w and (d) the risk free rate r_f. See proof in Appendix D.2

To interpret the equilibrium return it is useful to consider it relative to a more standard log-linear benchmark in which the Lambert function is replaced with a Log function.

Proposition 2

As long as $r_1 < \mathcal{W}(e) = 1$ then r_1 and α_1 are higher than the log-linear benchmark, and the

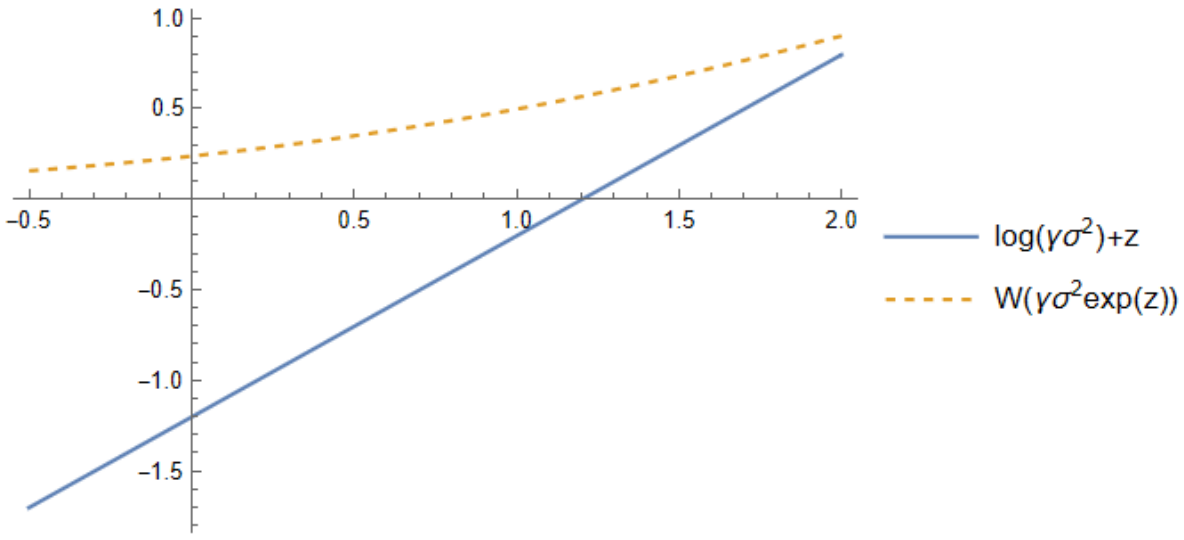


Figure 9: This figure compares returns as derived in equation (62) to a log-linear benchmark. Dashed line corresponds to $r_1 = W(\gamma\sigma_1^2 \exp(z))$ where $z = E[d] - r_f + h - w_0 + \sigma_1^2/2 - \Omega\tilde{E}r$ to a log-linear benchmark in z. Returns are convex and rise more slowly relative to the benchmark, but are always higher relative to it. The x-axis is z, a composite of expected payoff, asset stock, and wealth. The y-axis is the value of the Lambert and Log functions. $\gamma\sigma^2 = 0.3$. In this figure, I abstract away from other assets.

marginal effect of rent, housing stock, wealth, and the risk-free rate is smaller than 1, and approaches 1 as the sum of these variables goes to infinity.

See proof in Appendix D.3

Figure 3 visualizes the behavior of equilibrium returns relative to a log-linear benchmark. The elasticity between returns and fundamentals is driven by prices shifting both the value of the rental stock and the desirability of the asset. An increase in the price of the asset (p) lowers the return, decreasing the supply of capital, but simultaneously increases the value of rental stock driving up the demand for capital. A shift in the demand for capital is a movement along the (somewhat inelastic) log-linear supply curve, which corresponds to a less than one-for-one change in both equilibrium returns and exposure. This can be seen in the slope of equilibrium returns observed in Figure 9 being lower than 1.

To understand the convexity in Figure 9, note that the local shape of the non-linear demand function matters: it determines how changes in fundamentals translate to shits in demand and thus how inelastic returns are to these fundamentals. When demand is inelastic then then movements of the demand function (i.e. due to changes in wealth) and along the supply function result in a bigger change in returns and smaller change in exposure; whereas if demand is elastic then shifts in the demand function mostly translate to changes

in exposure. This implies, among other things, that more risky cities where supply is inelastic will have a higher elasticity of returns with respect to fundamentals, and safer cities with elastic supply will have lower such elasticities.

These results are best understood through the lens of segmentation. Segmentation is both 'blessing' and a 'curse' for investors: if fundamentals are strong investors increase their exposure, but their ability to do so is constrained by their limited wealth. As a result, investors enjoy higher returns. Returns must also be high than a standard log-linear solution, and particularly must be positive to make sure all the housing stock is purchased. But if investors have more wealth, they wish to invest more of it in the local asset, bidding up the price, and driving down its returns.

E.1.3 Comparative Statics on $\log(D/P)$

We can read comparative statics of the rent price ratio using our result for returns, since in a two period model the definition of net return equals log yields. For convenience denote $z = E[d] - r_f + h - w_0 + \sigma_1^2/2 - \Omega\tilde{E}r$, the sum of fundamentals in the exponent of (62). The comparative static with respect to some variable $x \in \{E[d], w_0, h, r_f, \Omega\tilde{E}r\}$ can be computed as

$$\begin{aligned}\frac{\partial E\left[\log(D/P)\right]}{\partial x} &= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x}\frac{\mathcal{W}((\sigma_1^2 + \Omega\Sigma_1)\exp(z))}{1 + \mathcal{W}((\sigma_1^2 + \Omega\Sigma_1)\exp(z))} + \frac{\partial \Omega\tilde{E}r}{\partial x} \\ &= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x}\cdot\frac{r_1 - \Omega\tilde{E}r}{1 + (r_1 - \Omega\tilde{E}r)} + \frac{\partial \Omega\tilde{E}r}{\partial x}\end{aligned} \tag{63}$$

See Appendix D.9 for full derivation. The sign of the comparative static of the expected log yield depends on the sign of $\partial z/\partial x$/[34] Yields rise with rent fundamentals and housing stock, and fall with wealth and the risk free rate. Furthermore, we can state that the elasticity of expected returns (and hence expected log yields) with respect to any of the fundamentals in x is always significantly than one.

Of particular note is the fact that this holds for rents: since yields have rents in the denominator, and the comparative static of returns to rents is less than one, it must be that prices increase by more (in percentage terms) than expected rents do. Holding all else constant, this introduces excess dispersion in prices, rationalizing the negative correlation of yields and prices.

Overall, we conclude that yields fall with wealth and increase in housing fundamentals,

[34]Under reasonable assumptions on $\Omega\tilde{E}r$; Except for $x = r_f$ where there is an additional positive term stemming from the fact that rent-price ratios are equal to returns, not excess returns. To account for this we add another r_f term. Further note that $\frac{\partial\Omega\tilde{E}r}{\partial x}$ equals 1 if $x = \Omega\tilde{E}r$ and zero otherwise.

that the elasticity of returns with respect to wealth and housing fundamentals is less than one, a prediction that is born out in the data in Section 2. In particular, prices move more with respect to fundamentals, and rents will have lower dispersion than prices, rationalizing the negative correlation observed in the data between the two. Finally, returns are bounded below from zero, and will tend to be high even if fundamentals are poor.

E.2 Model Misspecification Implications for OLS and IV Specifications

The simplified model shows local return are a function of local wealth:

$$E[r_1] = \mathcal{W}\left(\gamma(\sigma_1^2 + \Omega\Sigma_1) \cdot \exp\left(E[d] - r_f + h - w_0 + \sigma_1^2/2 + \Omega\tilde{E}r\right)\right) \tag{64}$$

We are interested in the implications of testing the comparative statics of the model, and particularly the sign of the effect of wealth, using linear regression. The model suggests two threats to such a test: (1) endogenous equity shares, invalidating the equity wealth IV; and (2) model specification error. I argue that while (1) biases the OLS estimate downward, it cannot induce a negative coefficient under the null that the true relationship is weakly positive and that (2) the linear approximation matches the marginal effect at the mean, both in sign and in magnitude.

E.2.1 Endogenous Equity Shares

In this section, I consider the possibility that a log-linear approximation of the true model is biased because of omitted variable bias, and that the equity wealth instrument is itself endogenously determined with the portfolio allocation to the real estate asset. I formally discuss the use of this IV in this case, and argue that it leads to a downward bias in the estimated effect of wealth, but cannot induce a negative estimated effect, under the null that wealth does not have a negative effect.

Consider the simplified model, approximated as log-linear:

$$\begin{aligned} r_{1ci} &= \phi \log\left(\gamma\left(\sigma_{1c}^2 + \Omega_c\Sigma_{1c}\right) \cdot \exp\left(E[d_{ci}] - r_f + h_{ci} - w_{ic} + \sigma_{1c}^2/2 + \Omega_c\tilde{E}r\right)\right) \\ &= \delta_c + \beta x_{ci} + \phi w_{ci} + \epsilon_{ci} \end{aligned} \tag{65}$$

Where i denotes a market within a city c and ϕ captures the sign and linear approximation error of the effect of wealth. δ_c are terms that do not vary within a city, and are thus captured by a city fixed effect. These include all the variance terms in Ω_c, Σ_{1c}, as well as the global

exogenous return vector $\tilde{Er}$. To generate a straight-forward omitted variable bias, suppose $E[d]_{ci} + h_{ci} = \beta x_{ci} + \epsilon_{ci}$ where x_{ci} are observed covariates that determine expected future rents, such as distance from CBD and BA share, and housing unit supply h. ϵ_{ci} will serve as the unobserved, omitted, variable which can be thought of as unobserved determinants of rent growth.

w_{ci}, local investor wealth, is not directly observed. Instead, I observe a measure of wealth imputed from investors' own house-value, which I denote $\hat{w_{ci}} = w_{ci} + \zeta_{ci}$ where ζ_{ci} is a market-specific measurement error, which could be correlated with housing fundamentals. In particular, I will assume that ζ_{ci} is arbitrarily correlated with ϵ_{ci}. As a result, to identify the sign of w_{ci} an IV is required.

One such IV is a measure for equity wealth. I assume such a measure is available and denote it $\omega = \log(\alpha_{Eci} W_{ci} \exp(\nu_{ci}))$ where α_{Eci} is the share of wealth invested in equity at i, and ν_{ci} is measurement error in the equity wealth measure.

The key assumption is that ν_{ci} is uncorrelated with ϵ_{ci}, and that α_{Eci} is determined as part of the full α portfolio weights vector. The exclusion restriction is that $\text{Cov}(\epsilon_{ci}, \omega_{ci}) = 0$. However, the portfolio share in real estate and equities is jointly determined, and if local returns depend on wealth, so will equity shares. The intuition for this covariance is simple: when returns are high in a certain market investors are incentivized to invest in housing, substituting away from equities. The model predicts a negative relationship between wealth and returns. If when returns are high wealth is biased downward, but less so when they are high, that would imply the estimated return-wealth trend is less steep then it is.

$$\begin{aligned}
\alpha_{Eci} &= \frac{1}{\gamma}\tilde{\Sigma}_E^{-1}\tilde{Er} - \alpha_{1ci}\Omega'_E \\
&= \frac{1}{\gamma}\tilde{\Sigma}_E^{-1}\tilde{Er} - \Omega'_E\left(\frac{E[r_{1ci} - r_f + \sigma_{1c}^2/2] - \Omega\tilde{Er}}{\gamma(\sigma_{1c}^2 + \Omega\Sigma_{1c})}\right) \\
&= \frac{1}{\gamma}\left(\tilde{\Sigma}_E^{-1} - \frac{\Omega'_E}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}\right)\tilde{Er} + \Delta_c - \frac{\Omega'_E(\beta X_{ci} + \phi w_{ci} + \epsilon_{ci})}{\sigma_{1c}^2 + \Omega\Sigma_{1c}},
\end{aligned} \tag{66}$$

where $\tilde{\Sigma}_E^{-1}, \Omega_E$ are the equity-relevant row and value in the the inverse return variance-covariance matrix and in Ω. Δ_c is a constant for c, and depends on δ_c and the variance-covariance of returns.

Consider the reduced-form regression of returns on equity wealth:

$$r_{1ic} = \delta_c + \beta x_{ci} + \phi\omega_{Eci} - \phi\log(\alpha_{Eci}(w_{ci}, \epsilon_{ci})) - \phi\nu_{ci} + \epsilon_{ci} \tag{67}$$

By assumption, ν_{ci} is classic measurement error and so only attenuates the estimate of ϕ. However, α_{Eci} is unobserved and is correlated with ϵ, and so could bias ϕ.

Consider the long, short, and auxiliary regressions, omitting for simplicity the variables x and the fixed effects:

$$\begin{aligned} r_{1ic} &= \phi^{\text{long}}\omega_{Eci} - \phi\log(\alpha_{Eci}(w_{ci},\epsilon_{ci})) - \phi\nu_{ci} + \epsilon_{ci} \qquad \text{(Long)} \\ r_{1ic} &= \phi^{\text{short}}\omega_{Eci} - \phi\nu_{ci} + \epsilon_{ci} \qquad \text{(Short)} \\ \omega_{Eci} &= \pi\log(\alpha_{Eci}(w_{ci},\epsilon_{ci}) + \zeta_{ci} \qquad \text{(Auxiliary)} \end{aligned} \tag{68}$$

Following the standard result:

$$\phi^{\text{short}} = \phi^{\text{long}}(1-\pi)$$

The hypothesis being tested is that $\phi < 0$, and the null hypothesis is that $\phi \geq 0$. The goal is to show that under the null the omitted variable bias in the short regression cannot induce a negative ϕ^{short}, and that $\phi^{\text{short}} < 0$ **only if** $\phi^{\text{long}} < 0$. To do so, I show that $\pi < 1$ under the null.

The auxiliary regression estimates:

$$\begin{aligned} \pi &= \frac{\text{Cov}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}),\omega_{ci})}{\text{Var}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}))} \\ &= \frac{\text{Cov}(\log(\alpha_{Eci}),\log(\alpha_{Eci}) + w_{ci} + \nu_{ci})}{\text{Var}(\log(\alpha_{Eci}))} \\ &= \frac{\text{Var}(\log(\alpha_{Eci})}{\text{Var}(\log(\alpha_{Eci})} + \frac{\text{Cov}(\log(\alpha_{Eci}),w_{ci})}{\text{Var}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}))} \\ &= 1 + \frac{\text{Cov}\left(\log\left(g(\Sigma) - \frac{\Omega'_E\phi w_{ci}}{\sigma^2_{1c}+\Omega\Sigma_{1c}}\right), w_{ci}\right)}{\text{Var}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}))} \end{aligned} \tag{69}$$

(Where $g(\Sigma)$ is shorthand for a function of Σ, which is used for notation brevity.)

I first prove that the covariance between wealth and the equity portfolio share is an upper bound for the covariance between wealth and log equity portfolio share. More generally, consider the covariance between some $x > 0$ and $f(y)$ where $f(y) < y$ for all y. (As is the case for log.)

$$\begin{aligned} \text{Cov}(x, f(y)) = E[(x-\mu_x)(f(y)-\mu_{f(y)}] = E[xf(y)] - \mu_x\mu_{f(y)} \\ \int_x\int_y xf(y)F_{xy}dydx - \mu_x\mu_{f(y)} < \int_x\int_y xyF_{xy}dydx - \mu_x\mu_y \end{aligned} \tag{70}$$

Where μ_z is the average of z, and the inequality follows directly from the application of

$f(y) < y, \quad \forall y$, for both the product $xf(y) < xy$ and the average $\mu_y < \mu_{f(y)}$.

It follows that (assuming $w_{ci} > 0$):

$$\text{Cov}\left(\log\left(g(\Sigma_c) - \frac{\Omega_E'\phi w_{ci}}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}\right), w_{ci}\right) < \text{Cov}\left(g(\Sigma_c) - \frac{\Omega_E'\phi w_{ci}}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}, w_{ci}\right) < -\frac{\Omega_E'\phi}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}\text{Var}\left(w_{ci}\right) \tag{71}$$

Under the null that $\gamma \geq 0$, the term on the right is negative, implying that $\pi < 1$, and that $\phi^{\text{short}} < 0$ if and only if $\phi^{\text{long}} < 0$.

Furthermore, the structure of the correlation between the equity portfolio share can be used to show that conditional on the real $\phi < 0$ the estimated ϕ in the IV regression is a lower limit (yet still negative).

The sign of the bias in the IV regression depends on $\text{Cov}(\epsilon_{ci}, \omega_{ci}|\text{Obs})$:

$$\begin{aligned}
\text{Cov}(\epsilon_{ci}, \omega_{ci}|\text{Obs}) &= \text{Cov}(E[r_{1ci} - \delta_c - \phi h_{ci} - \beta x_{ci} + \phi(\hat{w}_{ci} - \zeta_{ci}), \log(\alpha_{Eci} + w_{ci} + \nu_{ci})|\text{Obs}_{ci}) \\
&= \text{Cov}(E[r_{1ci}] - \phi\zeta_{ci}), \log(\alpha_{Eci} + \nu)|\text{Obs}) \\
&= \text{Cov}(E[r_{1ci}], \log(\alpha_{Eci})) + \underbrace{\text{Cov}(E[r_{1ci}], \nu_{ci}) - \phi\text{Cov}(\zeta_{ci}, \log(\alpha_{Eci})) - \phi\text{Cov}(\zeta_{ci}, \nu)}_{=0}
\end{aligned} \tag{72}$$

Where the last three terms are zero due to the assumptions on the measurement errors. The bias depends on the covariance of housing returns with the log equity wealth share. It turns out that this covariance is strictly negative, and smaller than the covariance between housing returns and equity wealth share:

$$\text{Cov}(E[r_{1ci}], \log(\alpha_{Eci})) < \text{Cov}(E[r_{1ci}], \alpha_{Eci}) < 0$$

Where recall $\tilde{E}r, \Omega, r_f$ are national and σ_{1c}, Σ_{1c} vary only across cities, not across markets within a city.

We can sign the bias:

$$\text{Cov}(\epsilon_{ci}, \omega_{ci}|\text{Obs}) = \text{Cov}(E[r_{1ci}], \log(\alpha_{Eci})) < \text{Cov}(E[r_{1ci}], \alpha_{Eci}) = -\frac{\text{Var}(E[r_{1c}])}{\gamma(\sigma_{1c}^2 + \Omega\Sigma_{1c})} < 0 \tag{73}$$

Where $\text{Var}(E[r_{1c}])$ is the variance of expected returns across markets.

E.2.2 Specification Error

Inherent in the log-linear approximation of the Lambert function is a specification error. Furthermore, in the model, the marginal effect of wealth on returns is not constant, and instead depends increases with returns. In this section, I discuss the effects of using OLS, and argue that the estimated log-linear coefficient is not significantly biased relative to the marginal effect of wealth on returns *at the mean return*. However, this coefficient masks considerable heterogeneity in the marginal effect of wealth.

Consider the following model:

$$y_i = \mathcal{W}(\gamma\sigma_i \exp(x + \phi w_i + \epsilon_i)), \tag{74}$$

where I simplify the variance term to σ_i, and reduced the model's variables (except for wealth, w_i) into an observable x_i component and unobservable ϵ_i residual. Adding the coefficient ϕ further generalize the functional form (in the model $\phi = -1$.)

As discussed in section E.1.3, the marginal effect of w_i on y_i is:

$$\rho_i(w_i) \equiv \frac{\partial y_i}{\partial w_i}(w_i) = -\phi \frac{y_i(w_i)}{1 + y_i(w_i)} \tag{75}$$

An OLS regression of w_i on y_i is misspecified in that it estimates a single marginal effect across all w_i and y_i. In the spirit of Angrist et al. (1996), the OLS (or IV) coefficient is some weighted average of the marginal effects ρ_i.[35]

To gain insight of the impact of the linear approximation and use of OLS, I run a Monte Carlo simulation, comparing the estimated OLS coefficient to the 'true' marginal effect at the mean w_i. Specifically, I generate data using the Lambert function and different values of ϕ, and run the following regression with the ensuing data:

$$y_i = x + \log(\sigma_i) + \rho w_i + \epsilon_i \tag{76}$$

Where ϵ_i, w_i are normally distributed and σ_i has a uniform distribution. (I postpone a discussion of omitted variable bias to the next section.) I compute y_i as in Equation (74). The distribution of the variables are set to match the first and second moments of zip rent-price ratios at $\phi = -1$.[36]

[35] As the source of heterogeneity is the misspecification of the model and not the propensity to receive treatment, it does not necessarily follow the specific weights in Angrist et al. (1996).

[36] I set $x = 6.8$, $\sigma_i \sim U(0.3, 1), \epsilon_i \sim \mathcal{N}(0, 1), w_i \sim \mathcal{N}(10, 0.5)$. The generated mean is 0.106 and standard deviation is 0.063, compared with the data mean of 0.108 and standard deviation of 0.062. Note that this calibration does not identify the model's parameters, rather they are simply chosen to replicate these statistics - other ways of achieving the same do not materially change the results.

Figure 10 plots the estimated OLS coefficient ρ, as well as the implied marginal effect ρ_i at the average y_i, denoted $\rho_i(\bar{y})$. The plot also visualizes the standard deviation of ρ_i around $\rho_i(\bar{y})$.

Both $\rho, \rho_i(\bar{y})$ track closely together, but vary significantly as a function of ϕ, and are generally lower than ϕ. The relationship is U-shaped as both measures of the marginal effect first decrease rapidly with ϕ, then rise and level out at zero. This relationship can be understood using Equation (75): the marginal effect increases with ϕ, but as ϕ increases, the distribution of y_i shifts toward zero. This shrinks the ratio $y_i/(1+y_i)$, which when ϕ is high approaches zero (and so $\rho_i \to 0$ for all i) and when ϕ is low approaches one (and so $\rho_i \to \phi$).

As long as $\bar{y}$ is on the order of 0.1, the OLS coefficient ρ and marginal effect $\rho_i(y_i)$ are significantly lower than ϕ. The exact relationship depends on the distribution of y_i, and so on the other parameters of the problem. Centering the distribution of y_i to match the empirical patterns suggests that, under the model $\phi = -1$, one should expect an OLS estimate of -0.18. Such an estimate would be very close to $\rho_i(\bar{y})$ (at the observed $\bar{y}$).[37]

F Reduced Form Additional Details

I compute the average wealth of investors that hold property in the zip code of each property, excluding the unit's owner's wealth.[38]

To account for potential differences in maintenance costs, I control for the investor's own wealth (capturing potential differences in the value of time) and unit characteristics, including building size, unit type, and construction decade. To account for local growth prospects, I control for relevant measures highlighted by the urban literature: local wages, BA-share, restaurant presence, racial composition, and distance to CBD. I also control for the price of owner-occupied housing, which under most models (and the model presented above) should capitalize amenity and growth differences across space.[39] Given these controls, the effect of wealth is identified in the model.

I instrument for average investor wealth with average local equity wealth, This instrument does not identify a causal affect, but rather purges correlation between investor wealth and housing prices induced by the use of investors' own houes as a proxy for their wealth. The

[37]Such an estimate is significantly lower than estimates in Table **??**, which (based on Equation (75)) would yield $\rho = -7.7$.

[38]Most of these investors also live near each unit, but I do not limit the sample to only these investors in these regressions. Section 5.1.2 calculates average local wealth as the wealth of investors who physically live in each market.

[39]An identification threat that remains is systematic differences in the valuation of amenities by different groups, which would imply owner-occupier prices may not capitalize the path of rent-growth (rather only the path of rent-equivalent growth)

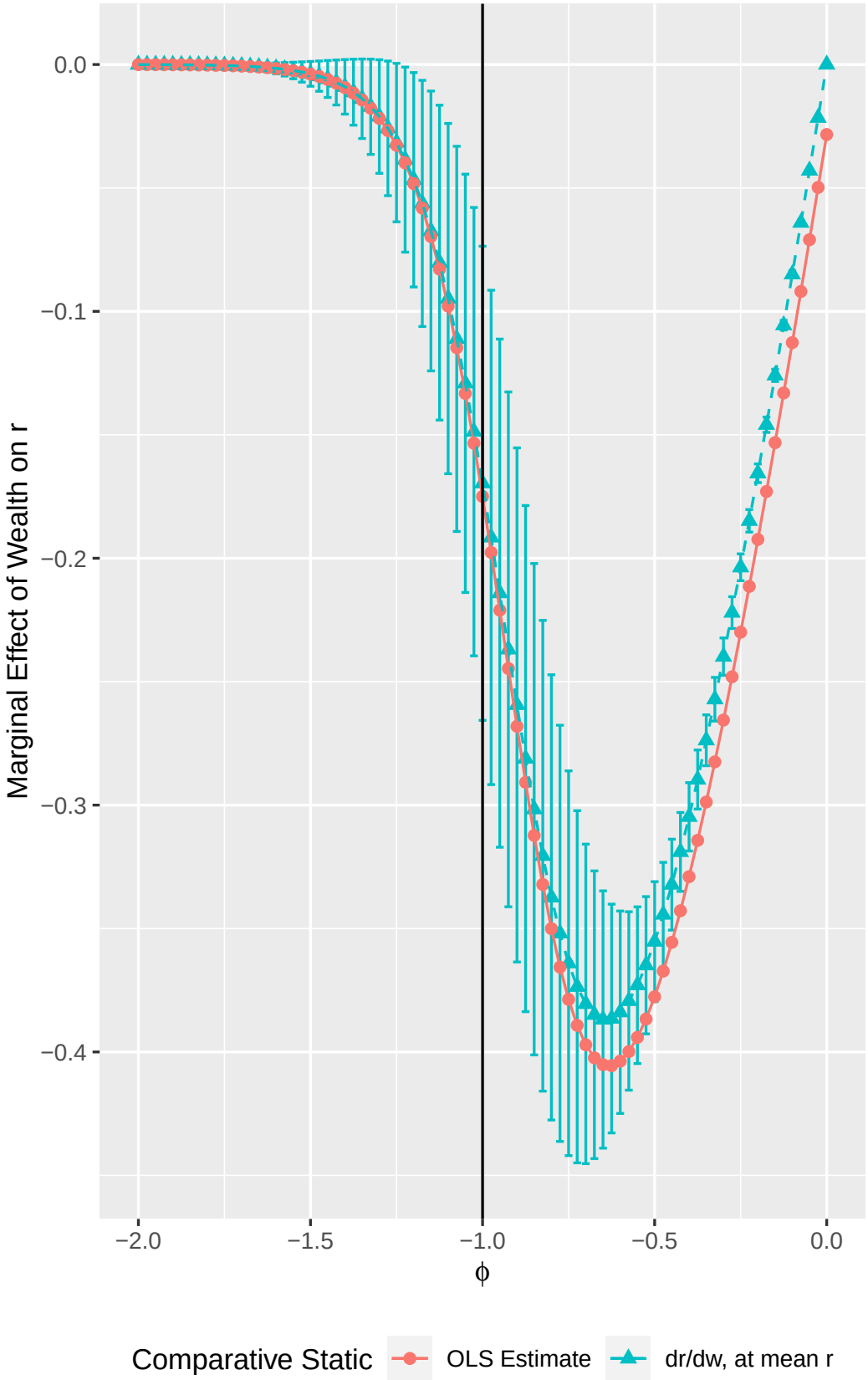


Figure 10: Comparison of the OLS coefficient ρ and marginal effect $\rho_i(\bar{y})$ at $\bar{y}$, the mean y_i. The x-axis presents different values of ϕ. The error bars show the standard deviation of ρ_i around $\rho_i(\bar{y})$. The distribution of returns changes as ϕ changes.

equity wealth measure is itself an imperfect measure of wealth - as it is measured for all residents, not just investors. By instrumenting for the former with the latter, the effect of the investors' wealth should be identified, except for reverse causality (which I address in the shift-share panel design below.)

Columns 1 through 3 report the effect of different measures of local log average investor wealth, in OLS specifications. The average house price of local investors and the implied average investor wealth negatively predict local rent-price ratios, with an elasticity of about one-half. This effect is much larger than that of the investor's own wealth. Using local average equity wealth results in a smaller, but still significant, elasticity of -0.167. This coefficient may be attenuated due to measurement error introduced by the equity wealth being measured for all households, and not just real estate investors. Column 4 utilizes equity wealth as an IV for local investor wealth . The effect of wealth remains negative and significant, with an elasticity of -0.392.

Table 11: Unit-level Impact of Local Wealth

Dependent Variable:	log(D/P)			
Model:	(1)	(2)	(3)	(4)
Variables				
log(Mean Neigh' Investor Value)	-0.463***			
	(0.056)			
log(Mean Zip Wealth)		-0.514***		-0.392***
		(0.042)		(0.120)
log(Mean County Equity Wealth)			-0.167***	
			(0.055)	
log(Investor-property Distance)	-0.011***	-0.011***	-0.013***	-0.012***
	(0.003)	(0.003)	(0.003)	(0.003)
log(Investor Own House Value)	-0.048***	-0.047***	-0.049***	-0.048***
	(0.006)	(0.006)	(0.006)	(0.006)
log(Mean Owner Occupier Value)	-0.625***	-0.487***	-0.696***	-0.540***
	(0.045)	(0.045)	(0.045)	(0.059)
Controls				
Amenities & Income	✓	✓	✓	✓
Race	✓	✓	✓	✓
Fixed-effects				
Metro	✓	✓	✓	✓
Construction Decade	✓	✓	✓	✓
SFH/Condo/Apt Bldg	✓	✓	✓	✓
Bldg Size Category	✓	✓	✓	✓
Statistics				
Observations	5,918,845	5,715,640	5,715,640	5,715,640
Within R^2	0.229	0.241	0.222	0.239
Wealth 1st F	-	-	-	1406016.23

Note: This table presents regressions of log rent-price ratio on measures of wealth, where wealth is averaged over investors who invest in the vicinity of the focal property. An observation is a property transacted between 2011 and 2013. Columns 1 through 4 are OLS specifications, and columns 5 and 6 are two stage least square, instrumenting for mean investor price and wealth with county mean equity wealth. Rent is estimated using hedonic and zip-level mean from the ACS, and are net of local real estate taxes. Rental and Investor unit prices are estimated via based transacted and estimated prices. Owner occupied prices are from the ACS. Properties are trimmed at the 2^{nd} and 98^{th} percentiles based on rent-price ratio. Home values are corrected-assessed values, as are prices of owner occupied homes. Rents are based on the zip-level means (ACS) and unit characteristics (AHS). All specifications include transaction year fixed effects. Standard errors clustered at metro level.