The Effect of Local Landlords on Rental Housing Prices and Returns

Peleg Samuels*

Tel Aviv University

August 2025

Abstract

What role do household, 'mom-and-pop' real estate investors play in determining housing prices? These small, undiversified, and highly home-biased investors collectively own a dominant share of rental housing. I develop a model in which risk-averse residents endogenously become landlords, supplying capital in a local market for rental units. Rent-price ratios and returns adjust to clear this market and depend on investor characteristics. To test the model, I construct a novel dataset on rent-price ratios, returns, and investor wealth. Using two shift-share instruments based on local ancestry and metro equity exposure, I show that increases in local wealth and investor participation rates lower rent-price ratios and future returns. These mechanisms help explain spatial variation in prices, returns, and wealth accumulation, with implications for homeownership costs and housing supply.

∗ *Email*: pelegsamuels@tauex.tau.ac.il I am deeply grateful to John Campbell, Edward Glaeser, Adam Guren and Sam Hanson for their invaluable advice and support. This paper also benefited from comments by Adi Sunderam, Gabriel Kreindler, Winnie Van Dijk, Naomi Hausman, Denise DiPasquale, Luis Viceira, Antoine Levy, and participants at Harvard's DUP and Finance graduate student seminars, as well as participants at seminars at the Hebrew University, Tel Aviv University, University of Wisconsin Madison, and the Alrov Symposium. Henry Weiland provided excellent research assistance. Any mistakes are my own. This paper was previously distributed as part of my doctoral dissertation, "Land, Landlords, and Local Capital in Housing Markets" (Harvard University, 2024). It benefited from the generous support of the Bradely Foundation.

1 Introduction

Housing prices are well understood to be a function of demand and supply for housing, stemming from local labor markets, amenities, geography, regulation, as well as the expectations thereof. While real estate is also often purchased by investors, these are often assumed to be 'absentee', with any local differences in returns, and hence in relative pricing, quickly arbitraged away. However, small household real estate investors, often called 'mom-and-pop landlords' own 82% of US rental properties and are highly home-biased, investing where they live. These facts suggest an important role for such investors in determining the price of- and return to local rental properties, yet neither their behavior nor impact is well understood. [1]

This paper develops and tests a framework in which the local supply of capital from household investors sets rental property prices, rent-price ratios, and returns. First, I document new facts: household investors own roughly four-fifths of US rental properties, are highly home-biased, and, conditional on wealth, are less-educated and more racially diverse than non-investors. Second, I build a tractable model in which risk-averse residents decide whether and how much to invest in nearby rental housing and, through their capital supply, shift local prices and returns. This model unifies urban spatial equilibrium and with finance models of housing and relaxes the standard absentee-landlord assumption. Third, I confirm the model's comparative statics in the data: neighborhoods with wealthier – or more – local investors exhibit significantly lower rent-price ratios, and positive, plausibly exogenous, wealth shocks lower future returns.

A growing literature documents substantial spatial dispersion in rent-price ratios and in housing returns, with the former often predicting the latter (Samuels, 2024).[2] Proposed explanations include risk premia (Eisfeldt and Demers, 2021), rent-growth expectations (Hilber and Mense, 2021), buyer characteristics such as wealth (Levy, 2021) and race (Diamond and Diamond, 2024). I nest these mechanisms in a unified framework that endogenizes both ratios and returns, while highlighting a largely overlooked driver: the local supply of household-landlord capital. A focus on spatial investor segmentation agrees with Damen et al. (2025), who show that risk differences alone cannot account for cross-sectional return gaps, and suggest a barrier-to-arbitrage explanation.

In a closely related analysis, Levy (2021) finds that home-biased French investors accept lower yields and expand local housing supply. I extend these insights by (i) endogenizing

[1] Returns, capitalization rates, and discount rates all refer to closely related concepts: they connect future payoffs to present values and reflect an investor's compensation. Following the asset pricing tradition, I refer to them simply as returns.

[2] For evidence on dispersion in rent-price ratios, see also (Amaral et al., 2022; Vanhapelto, 2022; Colonnello et al., 2023).

returns and adding a continuous investment margin with portfolio choice, (ii) subjecting investors to the shocks and prices faced by other residents, and (iii) introducing risk and risk-aversion. To my knowledge, my paper is the first to develop a tractable urban model in which risk-averse residents choose both real estate investment and price local returns in a spatial equilibrium setting.

My empirical results are consistent with the price effects of small-landlord entry documented after the Great Recession by Garriga et al. (2023) and of foreign buyers by Gorback and Keys (2020), suggesting that those effects persist in the cross section across cities and neighborhoods. They also suggest that some institutional investors targeting single family homes enter areas with high rent-price ratio, as noted by Mills et al. (2019), are trying to exploit high returns driven by relatively low local capital supply. The paper thus complements the growing literature on institutional ownership impacts (Coven, 2024; Chang, 2024) by analyzing the 'small investor' side of the market.

Finally, the paper contributes to asset-pricing and household finance literature by examining asset pricing under extreme home bias and offering a new solution to the canonical portfolio-choice problem—in which, unlike the usual solution, investor wealth and entry propensity strongly shape local returns.[3] It also sheds light on household investment behavior: unlike most asset classes, where richer investors earn higher returns, (Balloch and Richers, 2023; Bonaparte, 2021) higher housing returns tend to be earned in *poorer* neighborhoods by poorer investors.

To ground the analysis in investor behavior, the paper establishes three empirical facts about household landlords. First, household landlords dominate rental property ownership, holding roughly 82% of all rental structures, and supply a small majority of rental units (53.5%), including four-fifths of single-family rental units and most units in 2-to-20-unit buildings. Second, these investors are strikingly home-biased: about one-half of owners live within ten kilometers of their own homes, and more than four-fifths within 75 kilometers. The pattern is still stronger for multi-property and multifamily owners and mirrors French evidence in Levy (2021). Third, I document novel demographic facts. Participation draws on the middle and upper-middle class, and landlords differ systematically from investors in other assets: conditional on wealth landlords are less White, less educated, and report higher risk aversion than stockholders.Investment is also affected by cultural background, with some ancestries predicting higher participation rates in real estate licensing as a profession and as an investment, consistent with recent evidence on the effect of ancestry on homeownership (Gorback and Schubert, 2023).

[3]The solution method is related in spirit to Johnstone and Lin (2023) but focuses on different economic questions. On the impact of home-bias among equity investors see Hong et al. (2008) and subsequent work.

The paper then models such risk-averse, home-biased landlords in a spatial-equilibrium housing market with a fixed housing stock. Housing supply is inelastic and rents equalize utility across locations; a financing friction creates rental demand and opens a wedge between rental and owner-occupier prices, both of which are set endogenously. Local residents differ in wealth and in an idiosyncratic 'landlord ability,' and face a location-specific income shock that drives rent volatility and covaries with non-housing earnings. Potential landlords solve a continuous, static portfolio problem over rental housing and a risk-free asset, self-selecting from the local population.[4]

Equilibrium capital allocation is local: Rental prices and expected returns adjust so that the capital supplied by residents who choose to invest finances exactly the required number of rental units, forming an implicit local capital market. Participation patterns follow those observed in the data, and local returns decline with local wealth and with residents' propensity to invest in real estate (both of which raise capital supply) and rise with rent volatility and with the share of renters; wealth inequality can either raise or lower returns by shifting both rental demand and capital supply.

To test the model, I construct a property–level dataset that links prices, rents, and owner characteristics for nearly the entire US housing stock. The backbone is the 2013 *DataQuick* data, which reports details on structure, location, assessed value, and ownership for residential properties. Prices are imputed by regressing transaction values, within each county, on high-order polynomials in assessed value and a rich set of physical controls: a flexible hedonic regression that treats assessed value as a within-county rank, in the spirit of Diamond and Diamond (2024). Rents, by contrast, are estimated more parsimoniously and then scaled to match zip code means observed in the ACS. The resulting zip- and metro-level rent–price ratios are unbiased under standard hedonic assumptions.[5]

The data also allow for analyzing characteristics of owners, by matching owners' mailing addresses to the underlying property, providing three distinct benefits. First, the data offer a novel way to identify *household investors*: by determining whether the mailing address is itself a residential property, without requiring name parsing and correctly identifying properties held through trusts. Second, it yields distance between the rental and the landlord's residence, a precise measure of geographic segmentation. Third, it provides a proxy for investor wealth via the market value of the landlord's own home.[6] These links create novel nationwide data that simultaneously observe the local prices and rent-price ratios of rental

[4]In the appendix, I extend the framework to allow any number of risky, non-local assets, as an alternative to local real estate.

[5]I provide further details in a companion paper, Samuels (2024).

[6]in the Survey of Consumer Finances, own-home value explains 58 percent of homeowners' wealth, net of own-home value, and 75% of total wealth inclusive of own home.

properties, as well as the characteristics of the household investors who own them.

I test the model in three empirical steps corresponding to three geographic levels: the investor-property, a neighborhood cross section, and a metro panel. Beginning at the property-investor level, a regression of the rent–price ratio, which predicts spatial return differences (Samuels, 2024), on the landlord's own wealth yields a strong negative slope. However, this effect is largely explained by the average wealth of neighboring investors, which remains dominant after investor fixed effects and rich property controls are added. Differences in rent-price ratios reflect *local* characteristics of neighboring investors, rather than the individual investor's wealth.

The second test exploits within-city variation by clustering zip codes into 'sub-metro' markets that retain, on average, more than half of local landlords. Local rent-price ratios decline in mean investor wealth and the share of residents who are real estate investors, and rise with the share of renters. Identification comes from two instruments and a set of controls, which include homeowner prices, that capture local amenities and expectation, and city fixed effects.

This first instrument, which I term Ancestry-induced Real Estate Proclivity (AREP), uses plausibly exogenous ancestry weights interacted with local ancestry shares. The instrument, which captures culturally driven differences in the tendency to engage in real estate professionally, predicts local real estate investment and does not depend on whether the local ancestry composition is exogenous. [7] I also instrument for mean investor wealth with average household equity wealth from IRS capital-income data to purge mechanical measurement concerns between home values and the wealth proxy. The IV estimates match the signs and magnitudes of the model: a one-log point increase in the investor's share lowers the rent-price ratio by 7 pp, a one-log point increase in the renter's share increases it by 3.8 pp, and a one-log point increase in the mean investor wealth reduces it by 5.2 pp. OLS coefficients are similar but highly attenuated.

The final analysis uses a metro-year panel of prices and total returns. Combining FHFA house price indices with BLS rent growth and benchmarking with the DataQuick data yields a series of local realized total returns. Metro wealth shocks are constructed by interacting each metro's equity-wealth exposure (Chodorow-Reich et al., 2021b) with national equity-market surprises, producing plausibly exogenous variation. Controlling for city and year fixed effects, positive wealth shocks systematically depress subsequent rent–price ratios and realized returns, mirroring the cross-sectional elasticities and underscoring the causal role of

[7] I estimate ancestry weights from ancestry tendencies to work in real-estate licensing in the national ACS micro-data, purged of race, income, education and labor-market effects. Throughout, I control for race and income, both in estimating weights and the regressions that utilize them. Following Borusyak and Hull (2020) I control for variables that re-center the shift-share instrument on these dimensions.

local capital supply in pricing rental assets.

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. Section 7 concludes.

2 Household Real Estate Investors

An overwhelming share (81.8%) of US residential rental properties, and most units (53.5%), are held by household investors, the majority of whom tend to live near the property they own. In fact, more than 75% live within 50 km of their units, and slightly more 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 rise with wealth, but conditional on it is higher for non-White, less educated households, and those reporting higher risk aversion. These trends suggest demographic and idiosyncratic factors, which together create spatial differences in the potential investor pool. 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 in the media, generally refering to households using their own equity and consumer credit to purchase and manage rental properties (and often called 'mom-and-pop'). Typically, these investors are not real estate professionals and hold moderate real estate portfolios. (Some may employ professional property managers.)

I propose to identify such investors based on whether their business is conducted from an office or from their own residence. Specifically, I define a non-owner-occupied property to be held by an investor household if the mailing address listed on the property's deed is itself a residential address. As an additional check, I exclude properties whose mailing address is associated with an owner name that indicates a corporate entity.[8] The sample is conservative, as some household investors might own both their rental and their property under tax savings vehicles, such as an LLC and receive mail for both at a lawyer's office or a PO box, which cannot be identified as residential addresses. Investor households are the set of households that own at least one such property. I provide more details on the data and sample 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 as

[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.

'household investor' only if they hold real estate property other their main or vacation homes, and not through REITs or as an interest in a private company. In the survey, households are explicitly asked about such properties.

When using the RHFS, I consider properties as held by household investors if their owners are classified by the survey 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 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 investor households themselves.

Panel A utilizes the 2015 RHFS and focuses on the housing market, demonstrating that (as of 2015) 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%). [10] 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.[11] 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, 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.

[9] In DataQuick, I observe 345,552 properties under a name that includes 'LLC', but whose mailing address is a residential address. Based on a manual review of a sample of the data, many of these LLCs are clearly held by households and many overlap in their names: e.g. the name associated with the mailing address of a property owned by 'Smith Properties LLC' will itself be 'John and Joe Smith'.

[10] I use the 2015 vintage since the 2012 vintage includes only multi-family properties.

[11] 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 would be 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. Exploring this fact is left for further research.

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
Rent Paying Res' RE (Lower Bound)	0.06	$490	0.24
Alternative Assets			
Equity	0.48	$261	0.18

Continued on next page...

(Continued from previous page.)

Panel B: Importance to Investors (SCF)

Asset Class	Participation	Cond' Mean ($k)	Cond' Portfolio %
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 investors, using the SCF. I compute the participation rate, as well as the average holding value and portfolio weight, and then compare it to other important assets. The table documents a participation rate of 6.2-9%, equivalent to 8.8-11 million investor households, comparable to the share of households that own private businesses, but much lower than participation in equity markets of home-ownership. Conditional on participation, the portfolio weight of rental real estate 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 doesn't, on average, dominate investors' portfolios.

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 own residence or vacation homes, only half of them report positive rent, even among those who report ownership of multifamily housing. In Appendix C, I argue that this is likey due to under-reporting of rent payments. I consider 6.2% figure to be a lower bound, and includes households either reporting positive rent or owning a multi-family property.

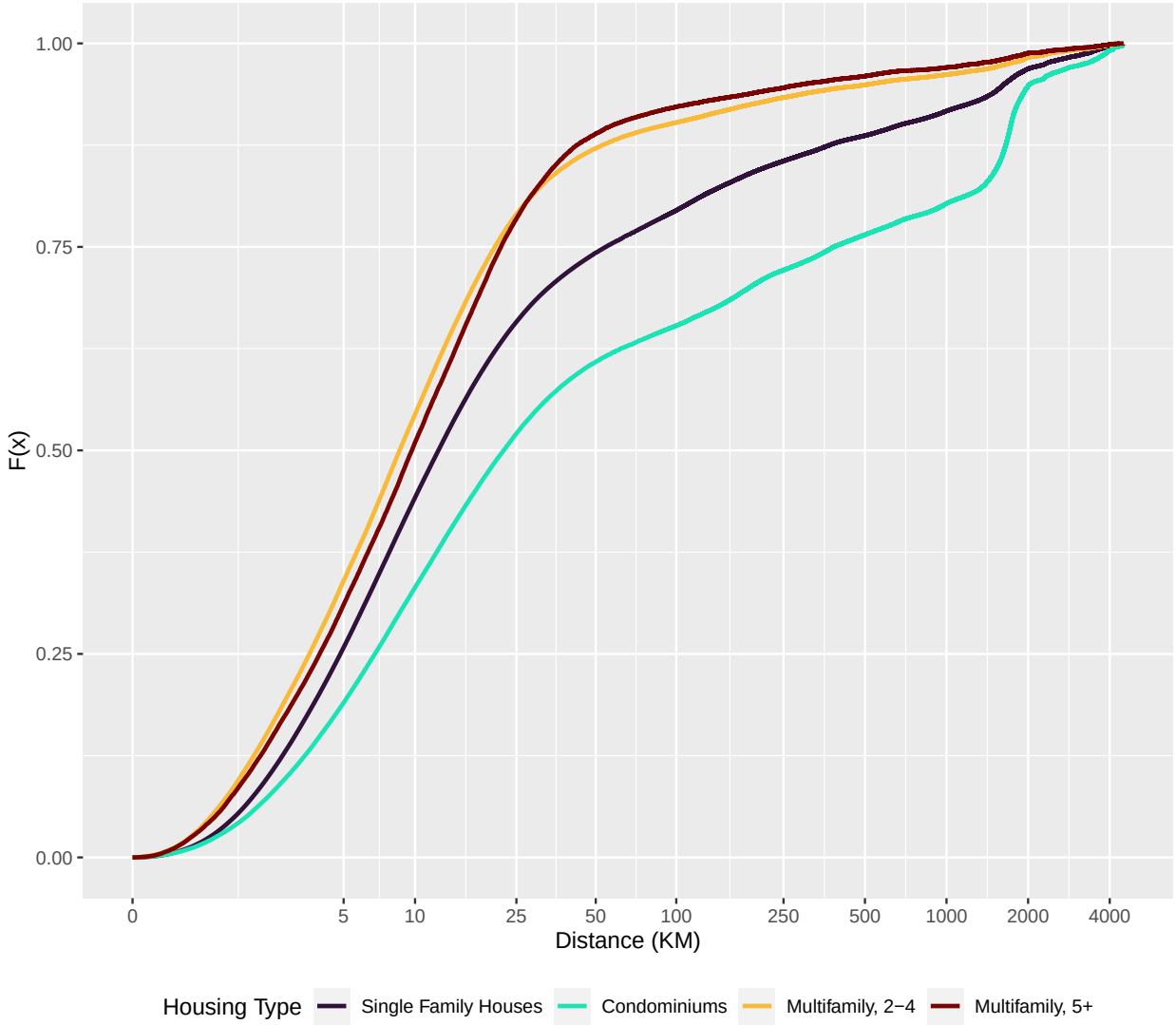
2.2 Home-bias Among Household Landlords

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: CDF of Distance Between Investor and Property



This figure plots the CDF of distance between each property and its owner's mailing addresses, in log scale. Colors represent the 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 household investors.

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.[12] Table A1 in Appendix A presents the CDF for household investors at various thresholds.

2.3 Characteristics of Real Estate Investors

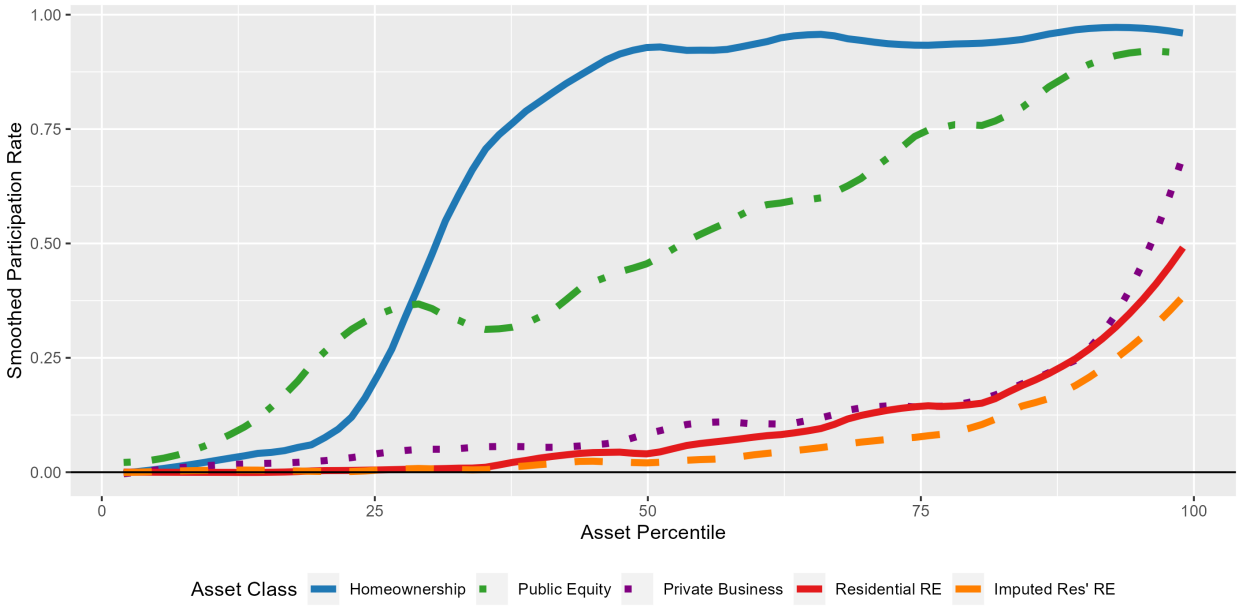
Unsurprisingly, real estate investors tend to be wealthier than the general population, though a sizable share, about a third, come from below the 85^{th} wealth percentile. Conditional on wealth, these investors differ from those in other asset classes: they are generally *less* white and *less* well-educated, 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. Figure 2 reproduces Figure 2 in Campbell (2006) with the addition of rental real estate, and shows that participation in real estate - using either the conservative or full definition - rises gradually from the 40^{th}, and more rapidly from the 85^{th} wealth percentiles. The participation of middle class Americans is not negligible - about a third of rental real estate investors come from below the 85^{th}. Notably, real estate investment patterns are very different to homeownership, which rises very rapidly at the 20^{th} percentile. Both assets have a 'minimum investment threshold', but rental real estate is far more idiosyncratic along the wealth distribution.

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 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.

[12] 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 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

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*

Conditional on wealth, 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 post-college education by 51.3%. Even high school graduates are less likely to be investors - 38% less relative to high school dropouts, although this effect is not statistically significant. This contrasts with well-established patterns in equity and private business investments, reproduced here, where generally White and better educated Americans are more likely to participate. Conditional on participation, these characteristics do not have a statistically significant effect on the rental real estate's portfolio share.

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.

3 Data

In this section, I describe the data and sample construction, including how I estimate investor wealth. I use micro-data on US properties and property transactions to estimate the value and rent of each unit. I then match properties to their owners, based on mailing address, and restrict the sample to household investors, and attempt to filter out vacation markets. Finally, I 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 data is augmented with zip, county, and metro data from public sources.

While the process estimates rent-price ratios for each property in the sample, the main empirical analyses use *local* averages, aggregating across many properties in each. Samuels (2024), a companion paper, utilizes the same process and discusses details of aggregation across properties, and the robustness of the procedure to noise in the estimation of property-specific prices or rents.

3.1 Estimating Rent and Prices

The main data I use is the 2013 DataQuick dataset, which includes the near-universe of US real estate properties at property level. This 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 on key variables: the number of units in the building, number of rooms, bedrooms, or square footage, or properties in zip codes for which the ACS does not report average rent. When large multi-family structure include only aggregate information (e.g. total bedrooms) I divide these equally across units within the building, abstracting away from within-building differences.

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

characteristics to change with t.

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

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

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

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

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

I invert this relationship to back out $\{\hat{\mathcal{L}}_{jt}\}_{jt}$, for each zip and period. It is then straightforward to reconstruct unit-level rents from the unit and building characteristics and the zip location premium for each property. These estimates can be thought of as equal to the average zip rent, adjusted for the value of each unit's physical characteristics.

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

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

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

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

Running this regression separately for each county allows each county's assessment process to be arbitrarily inaccurate (some municipalities under-value average prices by as much as 50% relative to transaction prices), but still uses the within-county variation as a signal of unobserved quality. Assessed price serves as a local rank-ordering of unobserved quality, as in Diamond and Diamond (2024).[13] These models are then projected onto the stock of housing in each county at each time period.

3.2 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 detail how I build my sample of household real estate investors.

I define a household investor based on the attributes of the mailing address listed on each property's deed. I identify properties whose owner does not live in them, and whose mailing address is itself a residential address. I also exclude properties whose owner's mailing address is itself owned by a commercial or public entity, based on its name, e.g. having 'Agency' in its name, or 'LP' in the end of the name. I identify these details for each mailing address by matching the mailing address back on 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, as long as they do not also do this for their own property, or otherwise send their mail to a lawyer's office or PO box.

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, implying most owners hold only one property. The sample covers urban and suburban areas, and primarily housing segments owned by smaller investors - single family houses, condos, and smaller multifamily properties.

Missing properties are on average larger properties, for which DataQuick data on unit

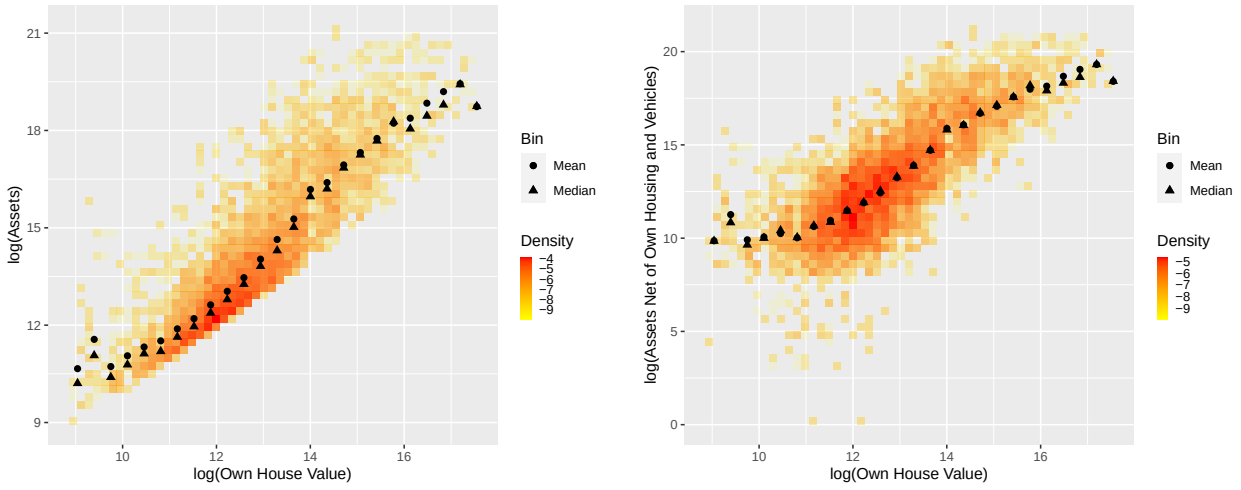
[13] Unlike Diamond and Diamond (2024) I cannot control for this 'rank-ordering' in computing rents. However, as detailed in Samuels (2024), minimizing the correlation of unobserved rent and price errors by better matching prices is sufficient to estimate local rent-price ratio averages.

details is more sparse. Additionally, 1.3 million properties have a P.O. box mailing address, and so the investors' own home cannot be identified. Finally, mailing addresses in rural counties appear to systematically have lower match rates – i.e. the mailing address is not found in Dat, and these counties tend to miss key data, such as properties' construction year and number of rooms, making it impossible to estimate rents and prices for properties in them.

3.3 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 home value explains 75% of log total assets, and 58% of log assets net of own home value. 3 visualizes these relationships.

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' real estate portfolio shares are always no larger 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 3 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 is similarly correlated with the value of the investors' rental real estate, as well as with its rent-price ratio, both in the SCF and in the DataQuick sample. Total wealth and wealth net of own housing are also both positively correlated, though the correlations for the imputed measures in DataQuick are attenuated. This attenuation could be due to measurement error or under-reporting of rent in the SCF, as discussed above.

3.4 Metro Panel Data

To augment the primary DataQuick data set, I construct a complementary panel using annual data on metro-level rent and home price appreciation from the Bureau of Labor Statistics (BLS) and the Federal Housing Finance Agency (FHFA), allowing a panel analysis of total returns across time and geography. Following the approach in Chodorow-Reich et al. (2021a), I use BLS rent and FHFA price indices for 24 metropolitan areas, forming an unbalanced panel with coverage beginning in 1975 for most cities.

To compute total returns and rent-price ratio levels, I benchmark each index to 2013 using the corresponding rental and price levels from the DataQuick sample, convert the indices to dollar values, and derive annual returns. The resulting panel reflects the returns and rent-price ratios for rental housing held by small investors, under the assumption that appreciation of these properties tracks broader metro-level trends captured in the BLS and FHFA data.

The primary advantage of the BLS/FHFA panel is its annual frequency, which allows for a flexible time-series framework and direct estimation of return dynamics. It also enables consistent treatment of serial correlation using heteroskedasticity- and autocorrelation-consistent (HAC) standard errors. Its main limitations are the limited cross-sectional coverage - relative to the broader spatial scope of the primary dataset - and the use of aggregate metro level indices, which may not fully capture the specific trends in properties owned by small investors.

Table 3: 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 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 rental 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.[14] 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 (called homeowners and denoted $o_i = 1$), which earns them a premium ψ, consumed immediately. Finally, residents can also invest a positive share of their wealth in other housing units as an investment, which I denote $\alpha_i \geq 0$.

$$U = \mathbb{E}\left[\frac{(C_i)^{1-\gamma}}{1-\gamma}\right] + \psi \cdot o_i. \tag{4}$$

Timing & Budget Constraint There are two periods: in the first, residents make all real estate purchases - for their own consumption, and as an investment. In the second period a market level income shock is revealed, rents are determined and paid by renters, and everyone consumes. Renters remain mobile and can move away and receive a reservation

[14] 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.

utility instead of paying rent. Households consume in the second period subject to their budget constraint.

The budget constraint is given by:

$$C_i = Y + (W_i - P_o \cdot o_i)(1 + R_i) - (1 - o_i)D. \tag{5}$$

Household consumption is equal to the common wage Y, less rent D for renters, plus households' wealth, net of the price of housing units for owner-occupants. Households also receive returns R_i on their wealth, or alternatively must pay interest on their mortgage debt. As this is a static model, housing does not have a resale value in the second period.

Real Estate. There are H identical houses available for purchase, implying that, in equilibrium, all residents live in the city.[15,16] For simplicity, 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. Appendix E presents a version of the model with multiple risky assets. 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.

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

[15]While renters can leave the city, in equilibrium they do not do so. Alternatively, the net flow of renters is equal to zero.

[16]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 ration 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 6 must therefore hold for all renters.

$$V + W_i(1 + R_f) = Y + W_i(1 + R_f) - D, \tag{6}$$

where V Denotes the reservation utility outside the market.

The ex ante distribution of rents thus reflects the ex ante distribution of a location's value. To make the model more tractable I define the market income shock Y in relative terms to the reservation utility in the broader economy V, such that $Y = V\nu$, where ν expresses the shock. I assume $\log(\nu - 1) \sim \mathcal{N}(\mu_d, \sigma_d)$.

Corollary 1

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

$$d_j \sim \mathcal{N}\left(\mu_d + v, \sigma_d\right). \tag{7}$$

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. It is straightforward to rewrite the expected logarithmic rents $\mu_d + \nu$ as equal to a 'current' rent level plus a stochastic growth term, expressing uncertainty in terms growth rather than levels.

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.. The supply of housing facing both groups is perfectly inelastic, implying prices are bid down by residents with different valuations. The price of owner-occupied housing is bid down along the indifference condition for buyers, until the liquidity constraint is met.

Analytically deriving prices for owner-occupants is not possible. 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 without an explicit solution for them.

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 $LC \cdot P_o = W_o$.)

$$E\left[\frac{(W_i - P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = \frac{(V + W_i(1+R_f))^{1-\gamma}}{1-\gamma}. \tag{8}$$

The utility of purchasing a home stems from the ownership premium ψ and from avoiding paying the expected rent payment (alternatively, it is a claim on the positive rent 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.[17]

Independence of Owner-occupied Prices and Rental 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. Both 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 nonempty 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 at the 40$^{\text{th}}$ percentile. Appendix Figure B1 shows significant bunching at institutional LTV

[17] 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 'renting as a hedge against local risk' of Lockwood et al. (2023).

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 even when controlling for location and a rich set of unit characteristics. Lastly, there is a robust literature showing that owner-occupied home prices are priced by owner-occupiers, and not investors.[18]

The intuition for the proposition is as follows: Owner-occupiers trade-off risky city shock and ownership benefit ψ against the costs of ownership. The value of the 'gamble' inherent in committing to the city 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

$$\begin{aligned} H^{\text{rental}} &= H \cdot F_w(lc - p_o). \\ H^{\text{own}} &= H \cdot (1 - F_w(lc - p_o)). \end{aligned} \tag{9}$$

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 2, 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), by log-linearizing the budget constraint, around a value summarized by the the elasticity

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

of consumption with respect to wealth, denoted ρ_i, and solving for the investor's first order condition.[19]

The investors' Euler equation is

$$E[r - \kappa_i - r_f] + \frac{1}{2}\sigma_r^2 = \gamma\sigma_r^2\alpha\rho_i + \gamma(1-\rho_i)\sigma_{yr}. \tag{10}$$

Equation (10) 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.[20]

4.3.2 Identifying Investors

We can define a set of marginal investors as those for which the optimal α_i in equation (10) 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 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 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 the participation rate within each wealth level w_i and identifies the lowest wealth investor

[19] 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

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

[20] 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.

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 (11) describes this equilibrium, using Propositions 1 and 2, and substituting in each investor's capital supply, Equation (10) as well as the definition of prices as discounted expected rent payment.

$$\begin{gathered}\underbrace{\int_w \int_\kappa \alpha(w,\kappa)\cdot(W_i - P_o)\cdot \mathbb{1}\{\alpha(w,\kappa)\geq 0\}\cdot dF_{w,\kappa}}_{\text{capital supply}} = \underbrace{P_I(r)\cdot H^{\text{rental}}}_{\text{capital 'demand'}} \\ \int_{w_m}^{\infty}\int_0^{\kappa^*(w)} \left(\frac{1}{\rho_i}\frac{E[r-r_f]-\kappa_i+\frac{1}{2}\sigma_r^2}{\gamma\sigma_r^2} - \frac{1-\rho_i}{\rho_i}\beta_{yr}\right)\cdot(W_i-P_o)\,dF_{w,\kappa} = \\ \exp(\mu+v-E[r])\cdot F_w(lc-p_o),\end{gathered} \tag{11}$$

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

$$\begin{aligned}\sigma_r &= Var(r) = Var(d-p) = Var(d) = \sigma_d. \\ \sigma_{yr} &= Cov(r,y) = Cov(\log(\nu-1),\log(\nu)).\end{aligned} \tag{12}$$

Comparative Statics on Returns. Although the model is not analytically solvable, we can characterize important comparative statics. Proposition 3 summarizes these effects.

Proposition 3

Assuming the conditions of Propositions 1 and 2 hold, equilibrium returns decline in: (1) an increase in wealth among any subgroup of investors, (2) a leftward shift 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 supplies 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 among the wealthy lowers returns, but a higher share of renters creates additional demand

for housing, 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 high or low, denoted $\bar{\kappa}$ and $\underline{\kappa}$. The correlation between these groups may be arbitrary.

Households with zero wealth are renters such that $0 < lc - p_o < w^L$, and 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.[21] 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 (11) under these assumptions and find equilibrium returns and portfolio shares.

$$\begin{aligned} E[r - r_f] + \sigma_r^2/2 &= \underline{\kappa} + (1-\rho)\gamma\sigma_{yr} \\ &\quad + \mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(\mu + v - r_f - \tilde{w}^H - \underline{\kappa} + \sigma_r^2 - (1-\rho^I)\gamma\sigma_{yr})\right), \\ \alpha &= \frac{1}{\rho^I\gamma\sigma_r^2}\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(\mu + v - r_f - \tilde{w}^H - \underline{\kappa} + \sigma_r^2 - (1-\rho^I)\gamma\sigma_{yr}))\right) \end{aligned} \tag{13}$$

[21] 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 that 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 reflect Proposition 3: return and the portfolio share of investors decline in wealth and risk. These solutions 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, they increase prices more than rents, and lower returns. I perform three sets of reduced-form exercises: two in the cross section, utilizing property (Section 5.1) and neighborhood-level (Section 5.2) data, which allow for detailed controls and an IV to identify the effect of the share of investors; and one in panel setting at the metro level (Section 5.3), which implements a shift-share design for wealth shocks, based on national equity market shocks, to credibly identify the effect of investors' wealth.

5.1 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. 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.

For each property, I measure the average wealth of other investors who invest in that property's zip, excluding that 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.[22] 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 a discussion of the causality of this effect for the next, more parametric analyses.

5.1.1 Wealth and Investor location

Table 4 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.[23] 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 investor's location. Columns (6) and (7) show that when controlling for market averages, the relationship declines even further.

5.1.2 Average Local Wealth.

Table A2 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 rental property's price (due to their proximity) I instrument for average

[22] I.e. markets may also be segmented based on cultural or generational lines, e.g. descendants of residents of a neighborhood might be more likely to invest in it, regardless of their physical proximity to it; alternatively, former students at a college might be more likely to invest in that town even after leaving it.

[23] 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 4: 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*

local wealth with an IRS derived measure of local equity wealth, taken from Chodorow-Reich et al. (2021b). 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.

5.1.3 Investor Fixed Effects.

In Appendix Table A3 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.

5.2 Market-level Regressions

This subsection turns from property-level evidence to market-level regressions that relate local capital supply to rent–price ratios. I regress the rent–price ratio in each market on two key variables: *average wealth* and the *investor share*—the fraction of residents who own rental property. Estimates support the model: Higher wealth and larger investor share both lower the rent–price ratio, with a similar magnitude.

Identification relies on rich controls and two instruments. Metro fixed effects and controls for BA share, log population, log density, and owner-occupied prices capture amenities. Wealth is instrumented with local average equity wealth to mitigate measurement error, whereas investor share is instrumented with an ancestry-based shift–share measure that exploits cultural differences in the proclivity to engage in real estate professionally.

5.2.1 Market Definition and Specification

Markets are clusters of zip codes whose centroids fall within 5 km of a focal zip; clusters never cross metro boundaries.[24] On average, a cluster contains 50% of an investor's properties

[24]Because clusters overlap, standard errors are conservatively clustered at the metro level, capturing correlation in labor and housing markets.

located within 10 km of her home, yet still exhibits within-city variation in rent–price ratios and in the ancestry-based instrument.[25]

For market i in city c I compute the rent–price ratio D/P_{ic}; mean investor wealth W_{ic}; investor share ϕ^I_{ic}; renter share ϕ^R_{ic}; average wages, BA share, and owner-occupied price. zip-level values are weighted by rental-unit counts to reflect the average rental unit, whereas ACS-based race and ancestry shares are unweighted to reflect the resident population.

$$\log(1 + D/P_{ic}) = \alpha \log W_{ic} + \gamma_1 \log \phi^I_{ic} + \gamma_2 \log \phi^R_{ic} + \beta X_{ic} + \delta_c + \varepsilon_{ic}, \tag{14}$$

where X_{ic} includes race composition, income, and owner-occupied prices, and δ_c is a metro fixed effect that absorbs metro-wide risk and growth expectations. The right-hand side of Equation 14 also mirrors the simplified heterogeneity result in Equation (13): replacing the Lambert term with a log and assuming a steady-state and spatially constant risk-free rates yields Equation 14. [26]

The model predicts $-1 < \alpha$, $\gamma_1 < 0$ and $0 < \gamma_2 < 1$. Of these coefficients, I focus on identifying the *capital-supply effects—α for wealth and γ_1 for investor share.*[27]

Within the model, the error term ε_{ic} aggregates (i) sub-metro rent-volatility risk, (ii) the covariance between rents and investors' consumption, and (iii) growth expectations not captured by X_{ic}. More generally, independent of the model, ε_{ic} represents all unobserved factors that affect rent–price ratios. Because these unobservables can correlate with wealth and investor share, the next subsection introduces an instrumental-variables strategy to obtain consistent estimates of α and γ_1.

5.2.2 Identification

To overcome endogeneity and measurement error in local investor capital supply, I use two instruments. The primary instrument, *Average Real-Estate Propensity* (AREP), is a cross-sectional shift–share IV for the investor share. The secondary instrument, introduced at the end of this subsection, purges measurement error in per-investor wealth.

AREP addresses simultaneity and reverse causality while also shifting demand for rental

[25] Appendix Table A6 reports robustness to zip clusters defined based on contiguity.

[26] In the model's static benchmark—where the entire payoff from the asset is the single rent flow ($D \geq P$)—net return simplifies to $r = \log(D/P)$. In actual housing markets, $D \ll P$ because the asset's resale value persists, and also includes an expected capital gain term; the theoretically correct empirical analogue of the model is therefore $\log(1 + D/P)$, which coincides with net returns r yet remains well defined when yields are small. Appendix Table A7 considers different functional forms for rent-price ratios, including $log(D/P)$. I also test regressions with investor and renter share in levels.

[27] The model further implies $\gamma_1 \approx -\gamma_2$, but IV estimates of γ_1 are LATEs and cannot be compared directly to OLS estimates of γ_2.

property, addressing omitted variable bias in the errors ϵ_{ic} due to unobserved risk or growth expectations. It interacts zip-level ancestry shares ('shares') with nationwide ancestry-specific propensities to invest in real estate ('shifts'). I argue that certain ancestries exhibit culturally driven investment tendencies that, conditional on race and other covariates, are orthogonal to local housing conditions, even if the spatial distribution of those ancestries is not. This exogenous variation in ϕ_{ic}^{I} identifies the effect of investor capital supply on rent-price ratios. The instrument is motivated by recent work by Gorback and Schubert (2023), that demonstrate cultural background has an impact home-ownership and subsequent portfolios, and Burchardi et al. (2019), that demonstrate that investment patterns follow ancestry even over very long horizons.

Ancestry IV Construction. Let s_{zl} denote the share of ancestry l in zip z and g_l the national 'shift' capturing ancestry-specific real-estate proclivity. The AREP instrument is

$$AREP_z \;=\; \sum_l s_{zl}\, g_l.$$

Because individual-level data on real-estate investment by ancestry are unavailable, I estimate g_l as each ancestry's probability of working in real-estate licensing (NAICS 531, which excludes brokers). This occupation is related to property ownership, yet, by excluding brokerage, is less geographically tied to specific markets, reducing concerns about reverse causality.

I pool the 1% Census and 1-year ACS micro-samples for 2000-2013 and run

$$Y_{jt} \;=\; \delta_t + \delta_{c(j)} + \beta X_j + g_{l(j)} + \varepsilon_{jt}, \tag{15}$$

where $Y_{jt} = 1$ if the individual j works in NAICS industry 531 in year t; X_j includes race, Hispanic status, log wages (plus a zero-wage indicator), and six education dummies; and $c(j)$ indexes the metro.[28] US-derived ancestries (e.g. responses such as American or State names) serve as the reference group ($g_l = 0$).

To aid interpretation, I rescale the estimated g_l to basis points: the mean is small $\bar{g} =$ 1.92bp, but with a significant standard deviation of $\sigma_g = 33.84$bp and a similar inter-quartile range. These figures are relative to a baseline employment probability in real estate licensing of 90bp. Bootstrapping 300 draws of 20% sub-samples yields 95% confidence intervals; roughly 32 of the 139 ancestry coefficients differ significantly from zero at the 5% confidence level.

Validation. The ancestry coefficients $\hat{g}_l$ capture the propensity of each group to work in real

[28]Education dummies include no schooling, less than high school, high school, some college, college, and post-college.

estate licensing relative to the pooled "American" reference group. If ancestry encodes deep-seated cultural attitudes toward real-estate activity, then exogenous cross-market variation in ancestry composition delivers a plausibly orthogonal first stage for the investor share. I validate this intuition in two steps: (i) showing that the geography of $\hat{g}_l$ traces recognizable cultural patterns, and (ii) confirming that the resulting shift–share measure (AREP) strongly predicts local investor participation.

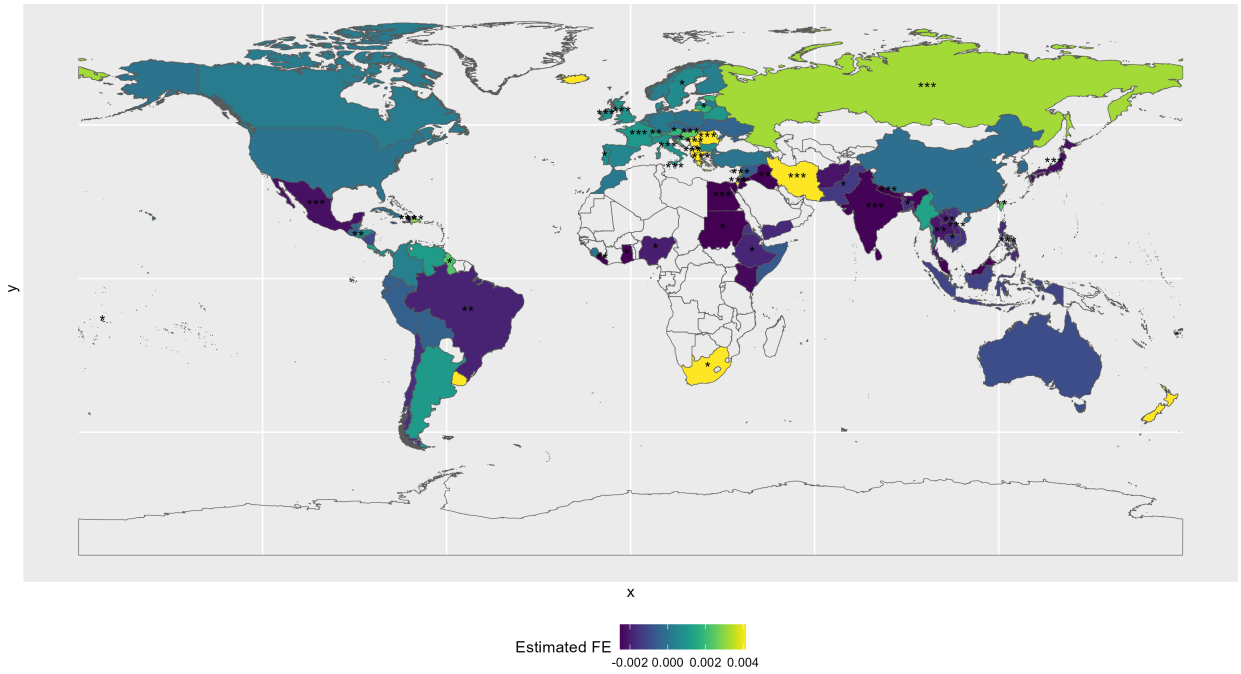


Figure 4: Estimated ancestry fixed effects (bp). Sub-national labels are mapped to modern nations (e.g. Bavarian→German). Stars indicate significance: *p< 0.10, **p< 0.05, ***p< 0.01. Blank countries lack sufficient observations.

Figure 4 reveals a striking cultural gradient. South-East European ancestries (e.g. Greek, Romanian, other Balkans) exhibit the largest positive proclivities, while Western European groups (e.g. UK, France, Italy) are modestly positive. Arabic-speaking Middle-Eastern ancestries post nearly uniformly large negative coefficients, whereas two non-Arab neighbors, Iranian and Israeli Americans, are strongly positive, underscoring a cultural and not merely regional channel. Asian groups tend to have negative coefficients with India, Japan, and the Philippines markedly so; China and Taiwan are higher and lie near zero. Unmapped ancestry follows the same pattern, for example, 'Yugoslavian' (positive) and 'Other Arab' (negative).[29]

[29]These patterns are robust to alternative ancestry aggregations, including the treatment of sub-national (e.g. Flemish) and supra-national (e.g. Scandinavian) identities.

Local validation comes from the first-stage regressions in Appendix Table A4. Sequentially adding the full set of housing market controls, a 1-bp increase in AREP increases investor share by 0.047 to 0.11 pp. Given a mean investor share of 6.3 pp, this large and precisely estimated elasticity confirms the economic relevance of the instrument and suggests that the variation it exploits is not driven by local housing dynamics.

Exclusion Restriction and IV Recentering. The exclusion restriction requires the shift vector g_l to be uncorrelated with the regression error ε_{ic}; the ancestry shares s_{zl} themselves may remain endogenous. Because g_l is estimated from nationwide employment data—purged of race, income, education, and metro fixed effects and reflecting country-of-origin cultural traits—it is plausibly orthogonal to local housing shocks. Two residual threats remain.

(A) Correlation with local fundamentals through race and amenities. If ancestries with high real-estate proclivity are disproportionately White, and White households cluster in fast-growing suburban zips, AREP could covary with growth expectations embedded in ε_{ic}. To break this link, the regression controls for ZIP racial shares, BA share, log income, distance to the CBD, and owner-occupied prices—absorbing systematic amenity and growth differences across ZIP types and their correlation with ancestry.

(B) Differential exposure to shocks through the shares themselves. Even after controlling for these covariates, some ZIPs may face higher average exposure to positive g_l simply because their ancestry mix is more "real-estate-prone." Following Borusyak and Hull (2020) (BH) I add three *recentering* terms that capture differences in exposure:[30]

(a) $RS_z = \sum_{l \neq \text{Am}} s_{zl}$: the total share of non-'American' ancestries, included because AREP omits the reference group.

(b) $RC_{zn} = \sum_l s_{zl} q_{ln}$: race exposure terms, where q_{ln} is the ACS share of ancestry l reporting race n. To avoid multicolinearity, I include RC_{nz} for $n =$ White, Black, and Asian.

(c) $RW_z = \sum_l s_{zl} \hat{\omega}_l$: income-exposure terms, where $\hat{\omega}_l$ is the ancestry-predicted (log) wage estimated with a similar regression as Equation 15, replacing Y_{jt} with log wages.

The local covariates address threat (A), while the BH recentering terms mitigate threat (B), ensuring that identification comes solely from differential ancestry composition rather than unobserved fundamentals or systematic exposure differences.

Robustness: Alternative Ancestry-Based IVs. AREP is estimated using employment in real-estate licensing, but other industry-specific propensities could, in principle, serve as

[30] BH show that, for linear shift–share instruments, controlling for these sources of differential exposure is sufficient; no explicit model of expected exposure is required.

ancestry shift–share instruments. I therefore construct AREP analogues for the 25% most common industries (in the ACS microdata) and compare their ability to predict local investor shares. Appendix Figure B2 shows that AREP ranks near the top in predicting local investor shares, but is not the only industry to do so. Most of the positive predictors are intuitive, and include financial services (plausibly through familiarity with investment, risk and leverage) and professional/technical services (architects, engineers); negative predictors include health care, vehicle manufacturing, and (perhaps surprisingly) construction.

Table A5 in the appendix repeats the second stage regression using four alternative industry IV bundles: financial services, construction, social / health services, and a mix of 'other' industries. The financial-services and construction instruments yield negative but statistically insignificant estimates of γ_1; the remaining bundles produce null effects. Although not a strict placebo, this contrast underscores the unique ability of *real-estate licensing* proclivity to shift investment in *rental real-estate* and supports the intuition of the AREP as reflecting true cultural proclivity, and not other ancestry correlations.

Wealth IV. To mitigate measurement error in average investor wealth W_{ic}, I exploit a second instrument based on *equity wealth*—liquid financial assets reported for all residents. Because W_{ic} is proxied by investors' own home values, spatially correlated house prices mechanically align measured wealth with local housing-price residuals, biasing α. Equity wealth, unrelated to local real-estate conditions, serves as an independent proxy of the same latent variable, purging it of correlated measurement error.

Conclusion. Together, AREP for the investor share and the equity wealth of W_{ic}, together with the Borusyak-Hull recentering terms, create a credible 2SLS design for estimating Equation (14). This framework isolates the causal effect of landlord capital, especially investor share, on rent-price ratios, firmly based on empirical results on the comparative statics of the model.

5.2.3 Results

Table 5 summarizes the market-level regressions. Column 1 reports OLS coefficients for investor and renter shares. Columns 2–3 replicate the specification using AREP as an instrument for the investor share; Column 3 additionally controls for owner-occupied house prices, absorbing unobserved amenities. Columns 4–5 introduce mean investor wealth, instrumented with equity wealth in Column 5.

Across all specifications the investor share has a negative sign, while the renter share is positive. The magnitude of the estimated effects is large: a 1-log point increase in the investor share decreases rent-price ratios by 7pp, or by 6.2pp when not controlling for investor

Table 5: Cross-Sectional Effects of Investor Capital Supply

Dependent Variable:	log(1+D/P)				
	OLS	IV		OLS	IV
	(1)	(2)	(3)	(4)	(5)
log(Investor Share)	-0.004***	-0.047	-0.062**	-0.003***	-0.070**
	(0.001)	(0.029)	(0.028)	(0.0008)	(0.027)
log(Renter Share)	0.006***	0.030	0.033**	0.006***	0.038**
	(0.001)	(0.020)	(0.017)	(0.0009)	(0.016)
log(Mean Investor Wealth)				-0.027***	-0.052**
				(0.002)	(0.021)
log(Mean Owner Price)			-0.019***		0.0001
			(0.004)		(0.010)
RW_z		0.426	0.279		-0.058
		(0.368)	(0.377)		(0.408)
Controls					
Metro Fixed Effects	✓	✓	✓	✓	✓
Amenities & Income	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓
Re-centering Variables	-	✓	✓	-	✓
Statistics					
Observations	11,103	11,103	10,629	11,103	10,629
Within R^2	0.456	-	-	0.509	-
Investor 1st F	-	76.25	65.41	-	34.16
Wealth 1st F	-	-	-	-	203.49
IVs	-	1	1	-	2

This table reports 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. 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). 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. (2021b). 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). Recentering variables include RC_z and RS_z, as defined in the main text. Standard errors clustered at the metro level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

wealth. Instrumenting for investor-share sharply increases its estimated coefficient sharply (from $-$ 0.4 pp to $-$ 4.7 pp), consistent with strong simultaneity bias attenuating the OLS

estimate. Mean investor wealth is also negative and statistically significant; the preferred 2SLS estimate in column 5 implies that a 1-log point increase in wealth lowers rent-price ratios by 5.2pp. These effects are statistically significant, particularly when controlling for owner prices, and the first-stage F statistics exceed 34 for every instrument, indicating strong relevance. These large effects imply that differences in investor capital supply could lead to the observed large spatial differences in rent-price ratios and returns.

Bias relative to the model. Appendix E.2 evaluates two ways the reduced-form coefficients might deviate from the structural model. (i) Omitting endogenous equity investment, used here as an instrument, can attenuate the instrumented wealth coefficient but cannot flip its sign as long as the true wealth effect is non-positive. (ii) Replacing the model's non-log-linear (Lambert) term with a log-linear specification produces 2SLS estimates that, in Monte Carlo simulations, closely follow the model's marginal effect of wealth on the observed mean return.[31] These checks indicate that the results in Table 5 can be read as valid local approximations of the model's comparative statics, despite the inherent misspecification. In terms of magnitude, the model predicts wealth and log investor share to lower the Lambert function's argument by -1; the empirical result suggest a similar, slightly larger, structural effect, of -1.25.

5.3 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 the 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.

Equation (16) 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 $\eta < 0$.

$$\Delta r_{c,t} = \delta_c + \delta_t + \eta \cdot \Delta W_{c,t} + \epsilon_{c,t} \tag{16}$$

A potential concern is the correlation between housing demand, returns, and local wealth

[31] Though the model's Lambert function suggests heterogeneous effects around this average, which the log-linear model does not.

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 city exposure (per-household) 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 and returns, allowing for the estimation of the causal effect of wealth shocks on returns. The inclusion of metro and year fixed effects soaks up any plausible re-centering variables, and so I omit explcit recentering.

Data. The dataset used in this analysis is the BLS/FHFA panel data described in Section 3.4. This data allows for an annual panel setting, where total returns can be directly tested. 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 estimates, and inflate their standard errors.

Instrument Construction. The instrument is given by Equation (17). 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 market beta $\beta_{c,t=0}$ and the city's hypothetical, baseline-derived, equity position at time $t-1$. Specifically, this position is equal to city's baseline portfolio $EW_{c,t=0}$, compounded over realized returns without allowing residents to make (endogenous) portfolio changes.

$$B_{c,t} = \underbrace{EW_{c,0} \cdot \beta_{c,t=0}}_{\text{Baseline Position}} \cdot \left(\underbrace{\prod_{\tau=0}^{t-1}(1 + r_{eq,\tau})}_{\text{Cum' return at } t-1} \cdot \underbrace{(r_{eq,t} - E[r_{eq}]}_{\text{Surprise at } t} \right) \tag{17}$$

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

Table 6: 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. (2021b). 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*

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.

6 Quantifying the Importance of Different Channels

This section is work in progress. I outline the current methodology below, but results are forthcoming and will be included in a future draft.

In this section I develop and estimate a structural model to quantify the importance of different observable channels. Investors self-select out of residents and select one of a discrete set of markets, subject to expected local return, their distance to the market, and an idiosyncratic shock, in the spirit of Berry et al. (1995) (BLP). Local returns are set as in the Model presented in Section 4, clearing the local rental capital market. The model identifies the preference parameters of investors, and the effect of investors' attributes', including average investor wealth and investor participation, on local returns and rent-price ratios. To identify investors' preferences, I utilize standard BLP moments and to identify the effect of local investor wealth and participation I use similar moments to those used in the reduced form section.

6.1 Structural Model

I use the simplified model presented in Section 4.5, in which the investor group is assumed to be homogeneous, characterized by average investor wealth $\tilde{w}$ and operating costs $\underline{\kappa}_c$. These variables are determined by the set of local investors who self-select out of residents and choose a market, using a BLP-style set up. For tractability, I allow investors to first select a market, and the treat portfolio share decisions and subsequent returns as being set by a representative decision, using the average wealth of investors who chose the market.

I allow growth expectations to equal a log-linear vector of covariates, such as housing supply restrictions and a structural error. To allow the model to quantitatively describe the importance of different channels, I generalize it by adding a parameter on that governs the effect of wealth.

The model implies that local equilibrium excess returns are equal to:

$$\begin{aligned} E[r_c - r_f] + \sigma^2_{r,c}/2 = \underline{\kappa}_c + (1-\rho)\gamma\sigma_{yr,c} \\ + \mathcal{W}\left(\gamma\sigma^2_{r,c}\rho \cdot \frac{\phi^R_c}{\phi^I_c}\exp(\mu_c + v - r_f - \tilde{w}_c - \underline{\kappa}_c + \sigma^2_{r,c} - (1-\rho)\gamma\sigma_{yr,c})\right), \end{aligned} \tag{18}$$

where $\underline{\kappa}_c, \tilde{w}, \phi^I_c$ are local investor characteristics of the local investor population, detailed below. c denotes a market, and clarifies which variables vary across cities. Risk aversion

γ, the elasticity of consumption of out wealth ρ^I are assumed parameters. The model also predicts a functional form for the investor portfolio share, α. However, due to difficulties in observing this variable I do not use it in the structural estimation.

6.1.1 Investor Participation

Investor i chooses a market c in which to invest, or the outside option which implies no investment. The investor faces a BLP-style decision, in which local (market-level) returns are weighed relative to a market-specific operating cost $\underline{\kappa_c}$, the investors distance from each market d_{ic}, and an investor-market specific shock, which represents an investor's own operating cost within each market.

Investor i's indirect utility from investing in market c equals:

$$V_{ic} = \psi_{1,i} r_c + \psi_{2,i} d_{ic} - \underline{\kappa_c} + \varepsilon_{ic}, \tag{19}$$

or receive V_0, an outside option, by not investing. $\psi_{1,i}$ and $\psi_{2,i}$ are random preference parameters that can depend on investor characteristics D_i, which include investor wealth, and d_{ic} is the log distance between investor i and market c, $\underline{\kappa_c}$ is the average operating cost in market c, and is treated as a market-specific constant that is unobserved by the econometrician. ε_{ic} is an investor-market specific shock, distributed with a Type-1 extreme value.

This 'demand system' for investors follows a BLP set-up, can be estimated with the usual two-step estimator, utilizing market shares and micro-data investor-market matches following Conlon and Gortmaker (2023). $\underline{\kappa_c}$ is treated as the unobserved constant econometric residual, the analogue of the product quality (usually denoted ξ in IO setting. I utilize instruments to identify the distribution of coefficients $\psi_{i,1}, \psi_{i,2}$, using 'BLP-instruments' - the characteristics of other markets.

6.1.2 Generalizing the Effect of Wealth

The model written in Equation (18) implies a tight relationship on the relative importance of different mechanisms. For example - the log of average wealth w_c and the log of the investor share $-\log(\phi_c^I)$ has the same quantitative effect.

To allow the model to capture richer patterns I add a parameter that governs the importance of investor wealth, multiplying the variable w_c by b_w. This can be thought of as a leverage or tax effect, changing the effective amount of capital supplied by residents in the local capital market in the market clearing condition $\alpha_c(W_c^H - P_c^O)^{b_w} \cdot \phi_c^I = P_c^I \cdot \phi_c^R$.

6.1.3 Growth Expectations

To take the model to the data I rewrite the expectation term, $\mu_c + \nu$, as a linear function of local covariates equal to $b_y y_c$, where y_c is a vector of length $k \times 1$ that includes variables known to predict rent growth - such as housing supply constraints and the share of college educated residents. (b_y is therefore a $1 \times k$ vector.)

6.1.4 Structural Errors

To rationalize the data, I add an additive error to Equation (18), denoted ζ_c. This error can be thought of as measurement error or an expectation term not captured by the model.

6.1.5 Measuring Risk

Estimating Rent Risk. Estimating the return variance faced by undiversified real estate investors requires addressing volatility in rent and capital gains. Due to limited data, the former is the key challenge. In Appendix G I develop a method to combine repeated ACS and AHS cross sections and vacancy data to estimate city-level rent growth variance. Rent is modeled using the hedonic specification used above, with zip-specific location premia, unit characteristics, and an exponential model of vacancy. Under simplifying assumptions, rent risk reduces to a function of three parameters: σ_c^L, the city-specific volatility of location premia across zip codes; $X_c' \Sigma^\beta X_c$, the exposure of the city's housing stock to changing hedonic valuations; and λ_c, the vacancy hazard rate, inferred from ACS vacancy data. Estimates of σ_c^L and Σ^β are constructed from repeated cross-sections of hedonic regressions between 1977 and 2013 in the ACS and AHS respectively. I find that rent risk averages 8.4% across cities, with most variation in the cross-section explained by differences in vacancy rates, which account for 85% of the observed dispersion.

Calibrating Return-Income Volatility. The covariance between income and returns, $\sigma_{yr,c}$, plays a key role in Equation (18), and cannot be directly computed even with estimates of $\sigma_{r,c}$. Instead, I approximate the value of $\sigma_{yr,c}$ using a second-order Taylor approximation in expected rent growth μ_c and risk $\sigma_{r,c}$. This approximation can be written as $\sigma_{yr,c}(\mu_c, \sigma_{r,c})$ and achieves an R^2 of 99.95% over the plausible range of both arguments.[32]

[32]The approximation is equal to $\sigma_{yr,c}(\mu_c, \sigma_{r,c}) = a_1\mu_c + a_2\mu_c^2 + a_3\sigma_{r,c} + a_4\sigma_{r,c}^2 + a_5\mu_c \cdot \sigma_{r,c}$. I estimate the parameters a_1 through a_6 using OLS on simulated data with μ_c ranging from 0 to 1 (0 to 100% growth) and $\sigma_{r,c}$ ranging from 0.01 to 0.9 (1% to 90%).

6.1.6 Structural Returns

Taking the above, excess returns can be written as a function $r(x_c, \zeta_c|\Theta)$, where x_c is data, $x_c = \{y_c, w_c^I, \phi_c^I, \phi_c^R, \sigma_{r,c}\}$, and Θ includes estimated parameters $\Theta = \{b_y, b_w, \underline{\kappa}, \gamma\} \in R^{k+3}$.[33] I calibrate r_f from average mortgage rates in 2013.

6.2 Estimation

I estimate the parameter vector Θ using moments similar to those used in the reduced form (Section 5 using standard GMM.

6.2.1 Moments

The first set of moments interact the errors ζ_c with instruments z_c, and require the expected value to equal zero:

$$\mathbb{E}[z_c \cdot \zeta_c] = \mathbb{E}\left[z_c \cdot (\hat{r}_c - r_c(x_c|\Theta))\right] = 0, \tag{20}$$

where $\hat{r}_c$ denotes observed returns. I use $z_c = \{1, y_c, \text{AREP}_c\}$, implying that (1) the average error is zero, (2) the error is uncorrelated with the predictors y_c, and (3) the error is uncorrelated with AREP. For notational convenience, I define the set of moments for which $z_c \in \{1, y_c\}$ as $Q_1(\Theta)$, and $z_c = \text{AREP}_c$ as $Q_2(\Theta)$.

These moments impose 'nationally' rational expectations and impose the same conditions as in the within-city reduced form regression (14) - that return residuals be uncorrelated with ancestry-based AREP and the growth predictors (utilized as controls in the reduced form). There are therefore $k + 2$ such moments – one for each variable in y_c, one for AREP, and one for nationally rational expectations.

I add a moment based on the cross-metro wealth shock instrument, requiring that the average shift in returns implied by a \$10,000 wealth shock match that estimated in the reduced form section:

$$\mathbb{E}[r_c(x_c|\Theta) - r_c(x_c'|\Theta)]\hat{\eta} = 0, \tag{21}$$

where x_c' equals x_c except that $\exp(w') = \exp(w + \log(10,000))$ and $\hat{\eta}$ is the estimated coefficient on wealth shock from Equation (16), estimated with year and metro fixed effects. I denote this moment as $Q_3(\Theta)$.

[33] ρ_c, the log-linearization parameter that is interpreted as the propensity to consume out of wealth, is computed using baseline variables for each city and held constant.

6.2.2 Estimation

I estimate Θ using two-stage GMM, following Newey and McFadden (1994):

$$\hat{\Theta} = \arg\min Q(\Theta)' \mathbb{W} Q(\Theta) \tag{22}$$

Where $Q(\Theta) = \{Q_1(\Theta), Q_2(\Theta), Q_3(\Theta)\}$ and $\mathbb{W}$ is a weighting matrix, equal to the identity matrix in the first stage and the optimal Chamberlain weighting matrix in the second.

6.2.3 Identification

The estimator is convex and globally unique and identifies Θ. The details and a formal proof are presented in Appendix H. Intuitively, $\underline{\kappa}$ is identified from the nationally rational expectations, and the parameter vector b_y is identified as in a non-linear least squares. γ is pinned down by Q_2, as long as there is a 'first stage' - that AREP and the function $\mathcal{W}(x_c|\Theta)$ are correlated. b_w is identified from Q_3 - the effect of wealth is intuitively set by the responsiveness of returns to a wealth shock.

6.3 Results

Results for this subsection are forthcoming and will be added in the next version of the paper.

7 Conclusion

This paper analyzed the behavior of small household investors in the US and their impact on the price of and return to rental housing. I begin by documenting bringing together key facts about these landlords: they own the majority of rental properties, especially in the single-family and small multifamily segments, and exhibit extreme home bias, with most investing close to their residence. Household investors also tend to hold small portfolios and come from the middle and upper middle class, and conditional on wealth, are less likely to be White or well educated. These features suggest that the characteristics and behavior of these local, under-diversified investors could play a central force in shaping large segments of the housing market.

To interpret these patterns, I developed a model in which risk-averse home-biased residents decide whether to supply capital to nearby rental housing. Local returns adjust to clear an implicit "capital market", so investor participation and capital supply endogenously respond to local returns. The framework nests existing explanations for rent-price and return dispersion: risk premia, rent-growth expectations, and demand shocks, and adds a novel one: larger local capital supply, driven by either higher investor wealth or a greater propensity to invest, lowers rent-price ratios and returns, whereas stronger rental demand raises them.

I tested the model in three empirical settings. At the property level, a negative covariance between investor wealth and that investor's property's rent–price ratio is explained by the wealth of nearby landlords, an effect that persists when controlling for investor fixed effects. At the neighborhood level, I showed that areas with more investors, or with wealthier investors, exhibit lower rent–price ratios. Using a novel instrument based on ancestry, I isolated plausibly exogenous variation in investor participation. The estimates match the model's predictions: a one-log-point increase in the local investor share reduces the rent–price ratio by 7 percentage points. Finally, at the metro level, I tested wealth shocks using a shift-share instrument derived from metro-level equity exposure and national shocks. Wealth shocks predict subsequent price increases and return declines.

This study helps bridge two perspectives in the housing literature: the housing-as-consumption view from urban economics and the asset-pricing view from finance. Standard models typically assume absentee landlords or treat returns as exogenous. In contrast, the framework developed here embeds investor behavior in a spatial equilibrium model with endogenous returns, allowing the same agents to consume housing, supply capital, and shape local price dynamics. This unified approach could better incorporate risk and forward-looking buyers in housing models.

A promising direction for future work is to examine how local capital supply shapes

housing production. Levy (2021) shows that household capital can expand supply; my findings suggest that it also affects returns, pointing to the need for a unified framework linking capital supply, prices, and construction. Household investors may also influence not just the quantity, but also the type and quality of the housing stock that gets built and maintained.

Although this paper focuses on household investors, institutional landlords are increasingly operating in the same segments (single-family and small multifamily) traditionally dominated by households (Mills et al., 2019; Coven, 2024; Chang, 2024). Future work should examine how these groups compete or specialize, what drives entry, and how these dynamics shape both efficiency and inequality. Since rental housing is an important asset for middle-class investors, the dispersion of returns can have significant distributional consequences: reallocating investor capital could raise returns for some, at others' expense. Finally, more work is needed to quantify the relative importance of capital supply versus risk, a distinction with both theoretical and policy relevance, but limited empirical clarity.

These findings also have policy implications. A low rent-price ratio, often cited as evidence of a bubble, need not reflect mispricing; in this framework, it may instead indicate a lower local required return, driven by abundant capital from wealthier or more investment-prone residents, especially when returns on other assets are depressed. Policymakers focused on housing supply should recognize that investor participation creates an additional source of demand, operating alongside fundamentals like household formation or migration. Accounting for this channel can help distinguish shifts in valuation—especially in prices rather than rents—and guide more targeted interventions for affordability or financial stability.

References

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. Balloch and J. Richers. *Asset Allocation and Returns in the Portfolios of the Wealthy.* LSE Financial Markets Group, 2023.

S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

Y. Bonaparte. Why do wealthy investors have a higher return on their stocks? *Available at SSRN 3809372*, 2021.

K. Borusyak and P. Hull. Non-random exposure to exogenous shocks: Theory and applications. Technical report, National Bureau of Economic Research, 2020.

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

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, 2021b.

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

C. Conlon and J. Gortmaker. Incorporating micro data into differentiated products demand estimation with pyblp. NBER Working Paper 31605, National Bureau of Economic Research, 2023. Available at NBER Working Paper No. 31605.

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.

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.

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

C. Garriga, P. Gete, and A. Tsouderou. The economic effects of real estate investors. *Real Estate Economics*, 51(3):655–685, 2023.

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

C. S. Gorback and B. J. Keys. Global capital and local assets: House prices, quantities, and elasticities. Technical report, National Bureau of Economic Research, 2020.

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.

H. Hong, J. D. Kubik, and J. C. Stein. The only game in town: Stock-price consequences of local bias. *Journal of Financial Economics*, 90(1):20–37, 2008.

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

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.

J. Mills, R. Molloy, and R. Zarutskie. Large-scale buy-to-rent investors in the single-family housing market: The emergence of a new asset class. *Real Estate Economics*, 47(2):399–430, 2019.

W. K. Newey and D. McFadden. Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245, 1994.

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 A1: 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 A2: 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.

Table A3: 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*

Table A4: AREP & Investor Share

Dependent Variable:	Investor Share					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
AREP	0.008	0.088***	0.113***	0.113***	0.095***	0.047***
	(0.017)	(0.019)	(0.017)	(0.016)	(0.017)	(0.016)
Ancestry Induced Wages				-0.003	-0.048	0.002
				(0.247)	(0.262)	(0.261)
log(D/P)					-0.130***	
					(0.020)	
log(P)						0.019***
						(0.003)
Controls						
Metro FE	-	✓	✓	✓	✓	✓
Race	-	-	✓	✓	✓	✓
log Pop' and Pop' Density	-	-	✓	✓	✓	✓
Fit statistics						
Observations	10,152	10,152	10,152	10,152	10,152	10,152
R^2	0.0002	0.231	0.338	0.338	0.356	0.386
Within R^2		0.011	0.148	0.148	0.172	0.211

Notes: The dependent variable is the share of residents in each zip who own at least one investment property, based on DataQuick and measured in percentage points (mean = 6.3 pp). The key regressor, *Ancestry Real-Estate Proclivity* (AREP), is a zip-level weighted average of nationwide probabilities that measure ancestries' relative probability of working in real-estate licensing, using local ancestry shares as weights. AREP is measured in basis points, so each coefficient gives the percentage-point change in investor share (in pp) associated with a 1 bp increase in ancestry-driven real-estate employment probability. Columns 2–6 add, cumulatively, metro fixed effects; race shares, log population, and log density; ancestry-predicted wages; rent-to-price ratios; and average rental prices. Standard errors are clustered by metro area. Standard errors are clustered at the metro area level throughout. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A5: Comparison of Different Ancestry-Industry Measures as IV

Dependent Variable:	log(1+Average D/P)							
Industries:	Finance, PST		Social & Healthcare		Others†		Construction, Mntnce	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
log(Share Investor)	-0.016	-0.013	0.002	-0.007	0.006	0.005	-0.052	-0.045
	(0.021)	(0.020)	(0.019)	(0.017)	(0.012)	(0.011)	(0.036)	(0.033)
log(Share Renter)	0.012	0.011	0.004	0.008	0.002	0.003	0.029	0.026
	(0.011)	(0.010)	(0.009)	(0.009)	(0.006)	(0.005)	(0.021)	(0.019)
log(Mean Investor Wealth)		-0.052***		-0.046***		-0.043***		-0.042***
		(0.014)		(0.013)		(0.012)		(0.016)
log(Average Owner Price)	-0.017***	0.004	-0.016***	0.001	-0.015***	0.001	-0.019***	-0.002
	(0.001)	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)	(0.007)
RW_z	0.170	-0.194	0.127	-0.165	0.119	-0.174	0.257	-0.051
	(0.292)	(0.280)	(0.276)	(0.270)	(0.280)	(0.267)	(0.346)	(0.335)
Controls								
Metro FE	✓	✓	✓	✓	✓	✓	✓	✓
Amenities & Income	✓	✓	✓	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓	✓	✓	✓
Rec-entering Vars	✓	✓	✓	✓	✓	✓	✓	✓
Statistics								
Observations	10,629	10,629	10,629	10,629	10,629	10,629	10,629	10,629
Within R^2	0.459	0.438	0.491	0.490	0.474	0.492	-0.231	-0.071
Investor 1st F	41.8	31.8	22.3	18.4	28.9	25.0	28.2	19.6
Wealth 1st F	-	94.14	-	85.42	-	58.22	-	134.32
IVs	1	2	1	2	1	2	1	2

This table reports regressions of the local rent–price ratio on the investor share and mean investor wealth, using *alternative* ancestry-industry shift–share instruments in place of AREP. Instruments are built identically to AREP but swap the industry that defines ancestry proclivity: Cols.1–2: Financial Services (NAICS 52) and Professional, Scientific & Technical Services (54); 3–4: Social Work (624) and Health Care (621-3); † 5–6: "Other" industries that negatively predict investing—Vehicle Mfg (336), Retail Trade (452, 441), Nondurable Wholesalers (424), Accommodation (721), and Food & Drink (722); 7–8: Construction (23) and Repair/Maintenance (811).

Observations are 5-km clusters of zips. Investor and renter shares are computed over households. All specifications control for owner-occupied house prices, race shares, BA share, log density, log income, and distance to the CBD (from Hausman et al., 2023); BH recentering terms RC_z and RS_z are included. Standard errors are clustered by metro. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A6: Effects of Investor Capital Supply, Contiguous Sample

Dependent Variable:	log(1+D/P)				
	OLS	IV		OLS	IV
	(1)	(2)	(3)	(4)	(5)
Variables					
log(Investor Share)	-0.005***	-0.059	-0.064*	-0.005***	-0.078**
	(0.001)	(0.037)	(0.036)	(0.0009)	(0.036)
log(Renter Share)	0.007***	0.048	0.051*	0.007***	0.062**
	(0.002)	(0.031)	(0.030)	(0.0010)	(0.030)
log(Mean Investor Wealth)				-0.029***	-0.064***
				(0.002)	(0.024)
log(Average Owner Price)			-0.014***		-0.004
			(0.003)		(0.004)
RW_z		0.314	0.261		-0.107
		(0.378)	(0.398)		(0.530)
Controls					
Metro Fixed Effects	✓	✓	✓	✓	✓
Amenities & Income	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓
Re-centering Variables	-	✓	✓	-	✓
Statistics					
Observations	12,785	12,785	12,060	12,785	12,060
Within R^2	0.470	-	-	0.528	-
Investor 1st F	-	44.8	45.18	-	25.6
Wealth 1st F	-	-	-	-	312.63
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 as the set of contiguous zip codes around each focal zip, and is a robustness check for the results presented in Table 5. 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). 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. (2021b). 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). Recentering variables include RC_z and RS_z, as defined in the main text. Standard errors clustered at the metro level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A7: Effects of Investor Capital Supply, Alternative Functional Forms

Dependent Variables:	log(D/P)				log(1+ D/P)	
	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
log(Investor Share)	-0.988***	-1.11***				
	(0.241)	(0.282)				
log(Renter Share)	0.527***	0.595***				
	(0.137)	(0.142)				
Share Investor pp			-0.153**	-0.183*	-0.010**	-0.012**
			(0.070)	(0.100)	(0.004)	(0.005)
Share Renter pp			0.064**	0.077*	0.004**	0.005**
			(0.028)	(0.040)	(0.002)	(0.002)
log(Mean Wealth)		-0.822***		-0.745*		-0.048*
		(0.249)		(0.424)		(0.026)
log(Mean Owner Price)	-0.161***	0.150	-0.348**	-0.107	-0.031***	-0.016
	(0.054)	(0.117)	(0.159)	(0.177)	(0.009)	(0.012)
	(0.386)	(0.609)	(0.785)	(1.56)	(0.045)	(0.071)
RW_z	4.09	-1.28	-9.51	-16.9	-0.611	-1.08*
	(5.49)	(5.63)	(6.73)	(11.3)	(0.400)	(0.569)
Controls						
Metro Fixed Effects	✓	✓	✓	✓	✓	
Amenities & Income	✓	✓	✓	✓	✓	
Race	✓	✓	✓	✓	✓	
Re-centering Variables	-	✓	✓	-	✓	
Statistics						
Observations	10,616	10,616	10,616	10,616	10,629	10,629
Investor 1st F	69.5	36.7 0	10.3	10.3	9.42	9.42
Wealth 1st F	-	202.74	-	197.43	-	197.95
IVs	1	2	1	2	1	2

This table presents 2SLS regressions of local rent-price ratios on local determinants of capital supply, as well local controls, using alternative functional forms for the rent-price ratios and investor and renter shares, as a robustness check to Table 5 The unit of observation is a cluster of zip codes, defined around each zip-code with a 5-km radius. 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). 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. (2021b). 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). Recentering variables include RC_z and RS_z, as defined in the main text. Standard errors clustered at the metro level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A8: Effects of Investor Capital Supply, Ancestry PCA IV, 5-km Sample

Dependent Variable:	log(1+D/P)				
	OLS	IV		OLS	IV
	(1)	(2)	(3)	(4)	(5)
Variables					
log(Share Investor)	-0.004***	-0.005	-0.015	-0.003***	-0.019*
	(0.001)	(0.012)	(0.012)	(0.0008)	(0.010)
log(Share Renter)	0.006***	0.007	0.011	0.006***	0.014**
	(0.001)	(0.008)	(0.007)	(0.0009)	(0.006)
log(Mean Investor Wealth)				-0.027***	-0.041***
				(0.002)	(0.009)
log(Average Owner Price)			-0.019***		-0.003
			(0.002)		(0.004)
Controls					
Metro Fixed Effects	✓	✓	✓	✓	✓
Amenities & Income	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓
Statistics					
Observations	10,629	10,629	10,629	10,629	10,629
Within R^2	0.482	-	-	0.516	-
Investor 1st F	-	12.32	12.32	-	12.23
Wealth 1st F	-	-	-	-	23.32
IVs	-	25	25	-	26

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 as the set of contiguous zip codes around each focal zip, and is a robustness check for the results presented in Table 5. Investor and renter shares are computed out of total households. Investor share is instrumented for with a set of principle components of ancestry shares, chosen based on the Kaiser criterion, and as a robustness check for the results presented in Table 5 . 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. (2021b). 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). Recentering variables include RC_z and RS_z, as defined in the main text. Standard errors clustered at the metro level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A9: Effects of Investor Capital Supply, All Ancestries IV, 5-km Sample

Dependent Variable:	log(1+D/P)				
	OLS	IV		OLS	IV
	(1)	(2)	(3)	(4)	(5)
Variables					
log(Share Investor)	-0.004***	-0.004	-0.008	-0.003***	-0.007*
	(0.001)	(0.007)	(0.006)	(0.0009)	(0.004)
log(Share Renter)	0.006***	0.006	0.008**	0.006***	0.009***
	(0.001)	(0.004)	(0.004)	(0.0009)	(0.003)
log(Mean Investor Wealth)				-0.024***	-0.048***
				(0.002)	(0.007)
log(Average Owner Price)			-0.019***	-0.008***	0.001
			(0.002)	(0.002)	(0.003)
Controls					
Metro Fixed Effects	✓	✓	✓	✓	✓
Amenities & Income	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓
Statistics					
Observations	11,103	11,103	10,629	10,629	10,629
Within R^2	0.456	-	-	0.516	-
Investor 1st F	-	9.71	8.95	-	8.97
Wealth 1st F	-	-	-	-	14.43
IVs	-	85	85	-	86

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. Investor and renter shares are computed out of total households. Investor share is instrumented for with local ancestry shares, as a robustness check for the results presented in Table 5. 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. (2021b). 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). Recentering variables include RC_z and RS_z, as defined in the main text. Standard errors clustered at the metro level. *Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B Appendix Figures

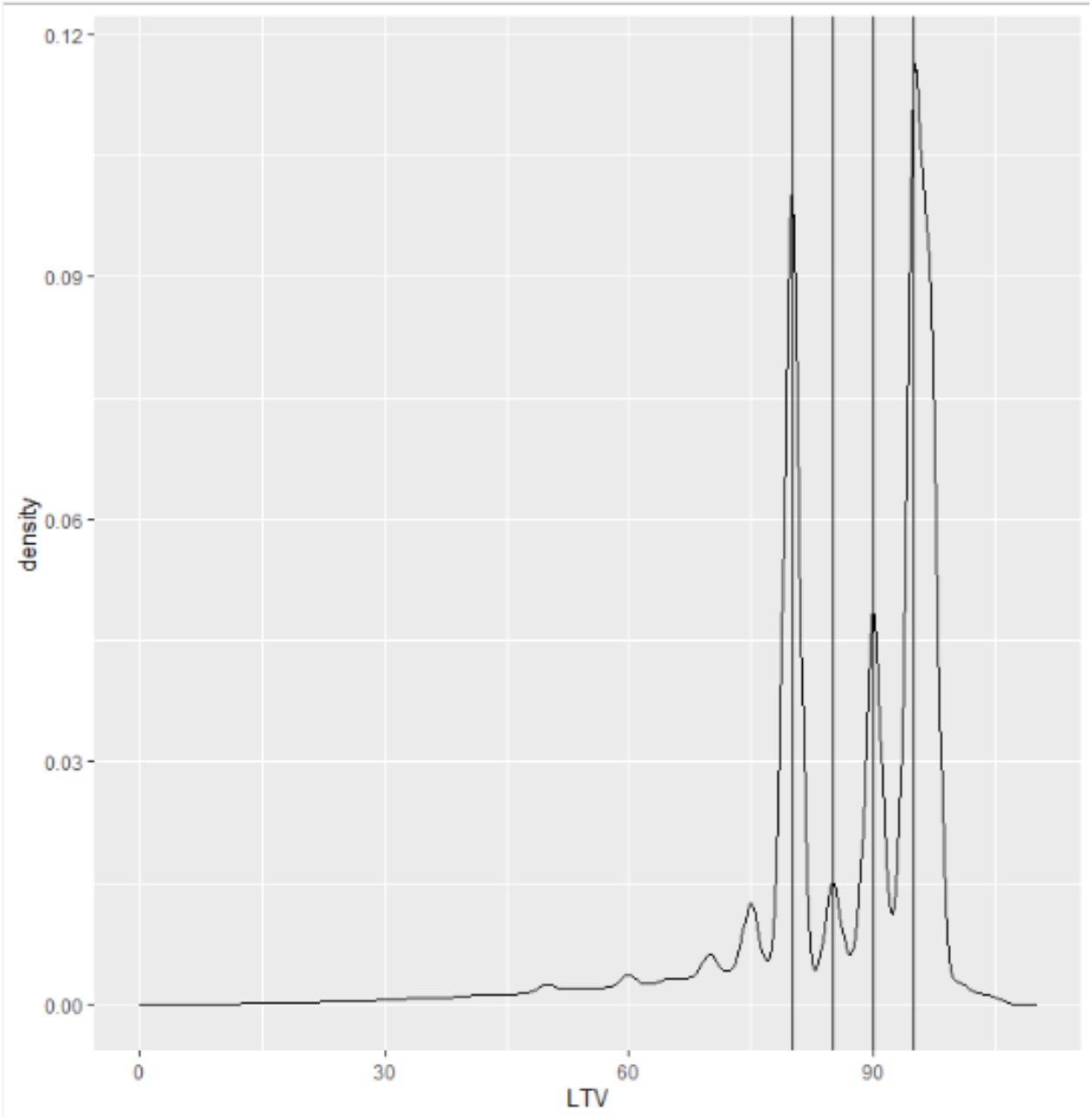


Figure B1: Distribution of LTV at mortgage origination for owner-occupiers 2018-2020, dashed lines are GSE limits; Data from HMDA mortgage microdata.

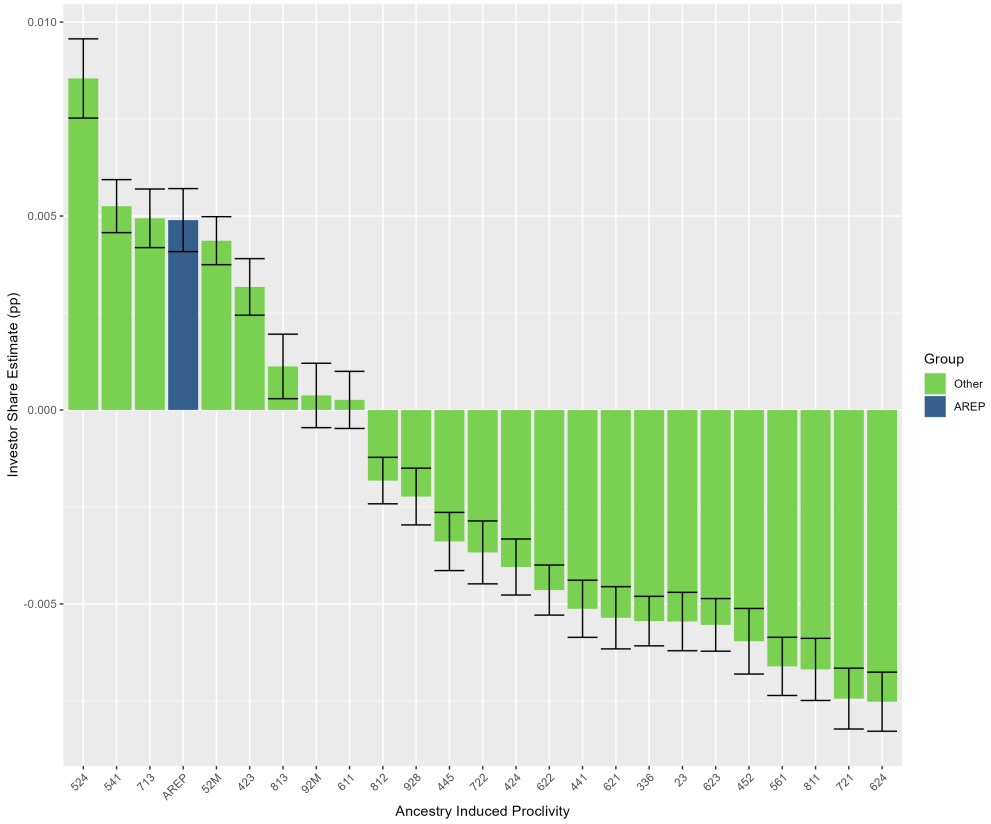


Figure B2: Comparison of relationships between the relationship of AREP and analogous local-ancestry-industry measures with local investor share.

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 C1 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 C1: 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

$$
\begin{aligned}
&\log(\alpha_i) = z - \gamma\sigma_i^2 f(r_g, \Sigma) - \gamma\sigma_i^2 \Omega \alpha_i \\
&\log(\alpha_i) + \gamma\sigma_i^2 \Omega \alpha_i = z - \gamma\sigma_i^2 f(r_g, \Sigma) \\
&\alpha_i \exp\left(\alpha_i \gamma\sigma_i^2 \Omega\right) = \exp\left(z - \gamma\sigma_i^2 f(r_g, \Sigma)\right) \\
&\alpha_i \cdot \gamma\sigma_i^2 \Omega \exp\left(\alpha_i \gamma\sigma_i^2 \Omega\right) = \exp\left(z - \gamma\sigma_i^2 f(r_g, \Sigma)\right) \cdot \gamma\sigma_i^2 \Omega \\
&\alpha_i = \frac{1}{\gamma\sigma_i^2 \Omega} \cdot \mathcal{W}\left(\exp\left(z - \gamma\sigma_i^2 f(r_g, \Sigma)\right) \cdot \gamma\sigma_i^2 \Omega\right) \\
&\alpha_i = \frac{1}{\gamma\sigma_i^2} \frac{1}{\Omega} \mathcal{W}\left(\gamma\sigma_i^2 \Omega \cdot \exp\left(\underbrace{z - \gamma\sigma_i^2 f(r_g, \Sigma)}_{\equiv A}\right)\right)
\end{aligned}
\tag{D1}
$$

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:

$$
z = \underbrace{E[d] - r_f}_{\text{relative payoff}} + \underbrace{h - w_0}_{\text{relative stocks of housing and wealth}} + \sigma^2/2 \tag{D2}
$$

and rewrite returns as:

$$
E[r - r_f] = \mathcal{W}\left(\gamma\sigma^2 \exp(z)\right) \tag{D3}
$$

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

$$\begin{aligned} &E[r - r_f](Z = 0) = 0 \\ &\frac{\partial E[r - r_f]}{\partial Z}(Z) > 0 \\ &\frac{\partial^2 E[r - r_f]}{\partial^2 Z}(Z) < 0 \end{aligned} \tag{D4}$$

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)

$$\begin{aligned} &\lim_{z\to\infty} \frac{\mathcal{W}(\exp(z))}{z} \to 1 \\ &\lim_{z\to-\infty} \mathcal{W}(\exp(z)) \to 0 \\ &\frac{\partial \mathcal{W}(\exp(z))}{\partial z} > 0 \\ &\lim_{z\to\infty} \frac{\partial \mathcal{W}(\exp(z))}{\partial z} = 1 \\ &\frac{\partial^2 \mathcal{W}(\exp(z))}{\partial^2 z} \geq 0 \end{aligned} \tag{D5}$$

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:

$$\begin{aligned} &\lim_{z\to\infty} E[r - r_f](z) \to \\ \log(\gamma\sigma^2) + z \quad &\\ &\log(\gamma\sigma^2) + E[d] - r_f + h - w_0 + \sigma^2/2 \end{aligned} \tag{D6}$$

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 (D6). 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 E1 visualizes the relationship of $\mathcal{W}(\exp(z))$ relative to z.

To make this argument precise, I rewrite equation (D3), defining and utilizing the differ-

ence between a Lambert functions and log for any x as

$$O(x) = \mathcal{W}(x) - \log(x) \tag{D7}$$

$O(x) > 0$ 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 D1. This implies returns can be written as:

$$\begin{aligned} E[r - r_f] = E[d] - r_f + h - w_0 + \sigma^2/2 + \log\left(\gamma\sigma_i^2\right) + \\ O\left(\gamma\sigma_i^2 \exp\left(E[d] - r_f + h - w_0 + \sigma^2/2\right)\right) \end{aligned} \tag{D8}$$

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:

$$E[r - r_f](z = 1 - \log(\gamma\sigma^2)) = \log(\gamma\sigma^2) + (1 - \log(\gamma\sigma^2)) = 1 \tag{D9}$$

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.

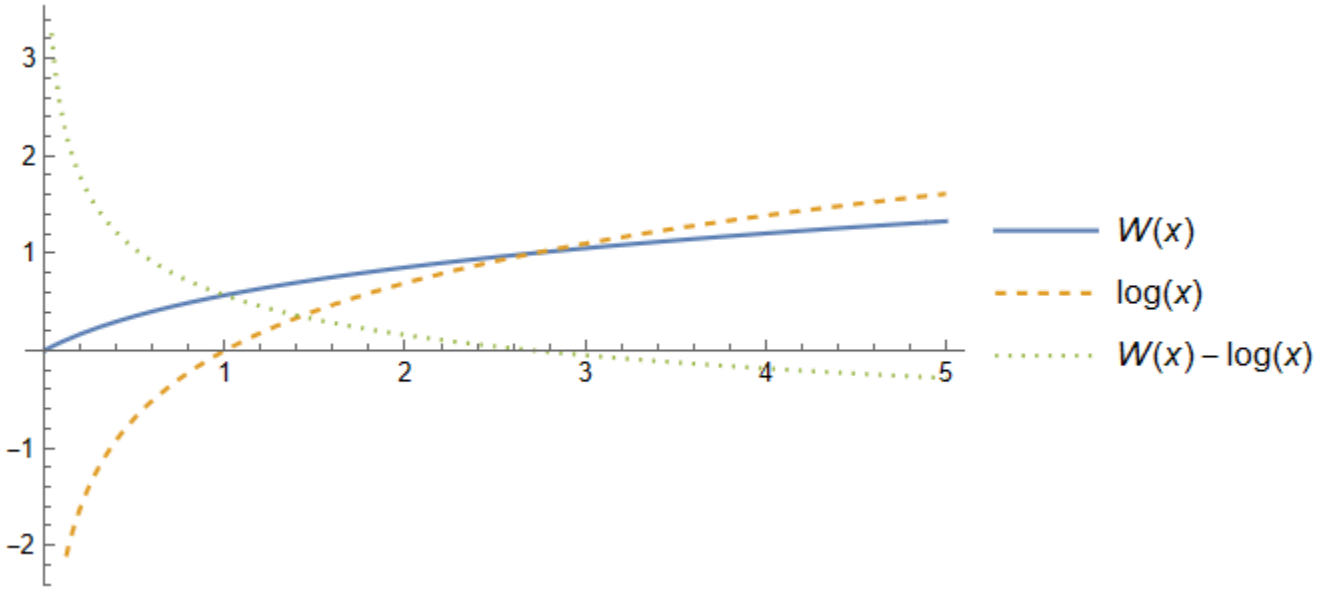


Figure D1: Visual Comparison of the Lambert function $\mathcal{W}(x)$ and $\log(x)$. $O(\cdot)$ as defined in equation (D8) is the difference between the two functions.

D.4 Corrolary 1

Recall that the reservation utility for renters is

$$V_i^{\text{Reservation}} = V + \boldsymbol{\alpha}_i(1 + \boldsymbol{R}_i) \tag{D10}$$

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.

$$\begin{aligned} &\nu V + \boldsymbol{\alpha}_i(1 + \boldsymbol{R}_i) - D_j = \boldsymbol{\alpha}_i(1 + \boldsymbol{R}_i) + V \rightarrow \\ &d_j = v + \log(\nu - 1) \\ &d_j \sim \mathcal{N}\left(\mu_d + v, \sigma_d\right) \end{aligned} \tag{D11}$$

D.5 Proposition 1

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 (D12) describes the willingness to pay for each resident i.

$$E\left[\frac{((W_i - P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = \frac{(V + W_i(1+R_f))^{1-\gamma}}{1-\gamma} \tag{D12}$$

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

$$\begin{aligned}
&E\left[\frac{((W_i - P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = \frac{(V + W_i(1+R_f))^{1-\gamma}}{1-\gamma} \rightarrow \\
&(1+R_f)\frac{1-\gamma}{1-\gamma}(1-P_o')E[((W - P_o)(1+R_f)+Y)^{-\gamma}] = (1+R_f)\frac{1-\gamma}{1-\gamma}(V + W(1+R_f))^{-\gamma} \\
&1 - P_o' = \frac{(V + W(1+R_f))^{-\gamma}}{E[((W - P_o)(1+R_f)+Y)^{-\gamma}]} \\
&P_o' = 1 - \left(E\left[\left(\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right)^{\gamma}\right]\right)^{-1}
\end{aligned} \tag{D13}$$

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:

$$\begin{aligned}
E\left[\left(\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right)^{\gamma}\right] &> \left(E\left[\frac{(W - P_o)(1+R_f)+Y}{V + W(1+R_f)}\right]\right)^{\gamma} \\
&= \left(\frac{W(1+R_f)+V}{W(1+R_f)+V} - \frac{P_o(1+R_f)+V}{W(1+R_f)+V} + E\left[\frac{Y}{V + W(1+R_f)}\right]\right)^{\gamma} \\
&= \left(1 - \frac{P_o(1+R_f)+V}{W(1+R_f)+V} + \frac{E[Y]}{V + W(1+R_f)}\right)^{\gamma}
\end{aligned} \tag{D14}$$

Which is greater than 1 if

$$
\begin{aligned}
&1 - \frac{P_o(1+R_f)+V}{W(1+R_f)+V} + \frac{E\left[Y\right]}{V+W(1+R_f)} > 1 \\
&P_o(1+R_f) < E[Y-V] = E[V(\nu-1)] = E[D] \\
&P_o < \frac{E[D]}{1+R_f}
\end{aligned}
\tag{D15}
$$

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 (D12) for a given price level.

$$
E\left[\frac{((W_o^*(P_o)-P_o)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right] + \delta\psi = E\left[\frac{(V+W_o^*(P_o)(1+R_f))^{1-\gamma}}{1-\gamma}\right] \tag{D16}
$$

Denote the price that both satisfies (D12) 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 (D16) 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[34]

The reader will note that this proof is an 'if' statement. The use of Jensen's inequality in equation (D14) guarantees the solution under the described conditions

Figure D2 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.

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

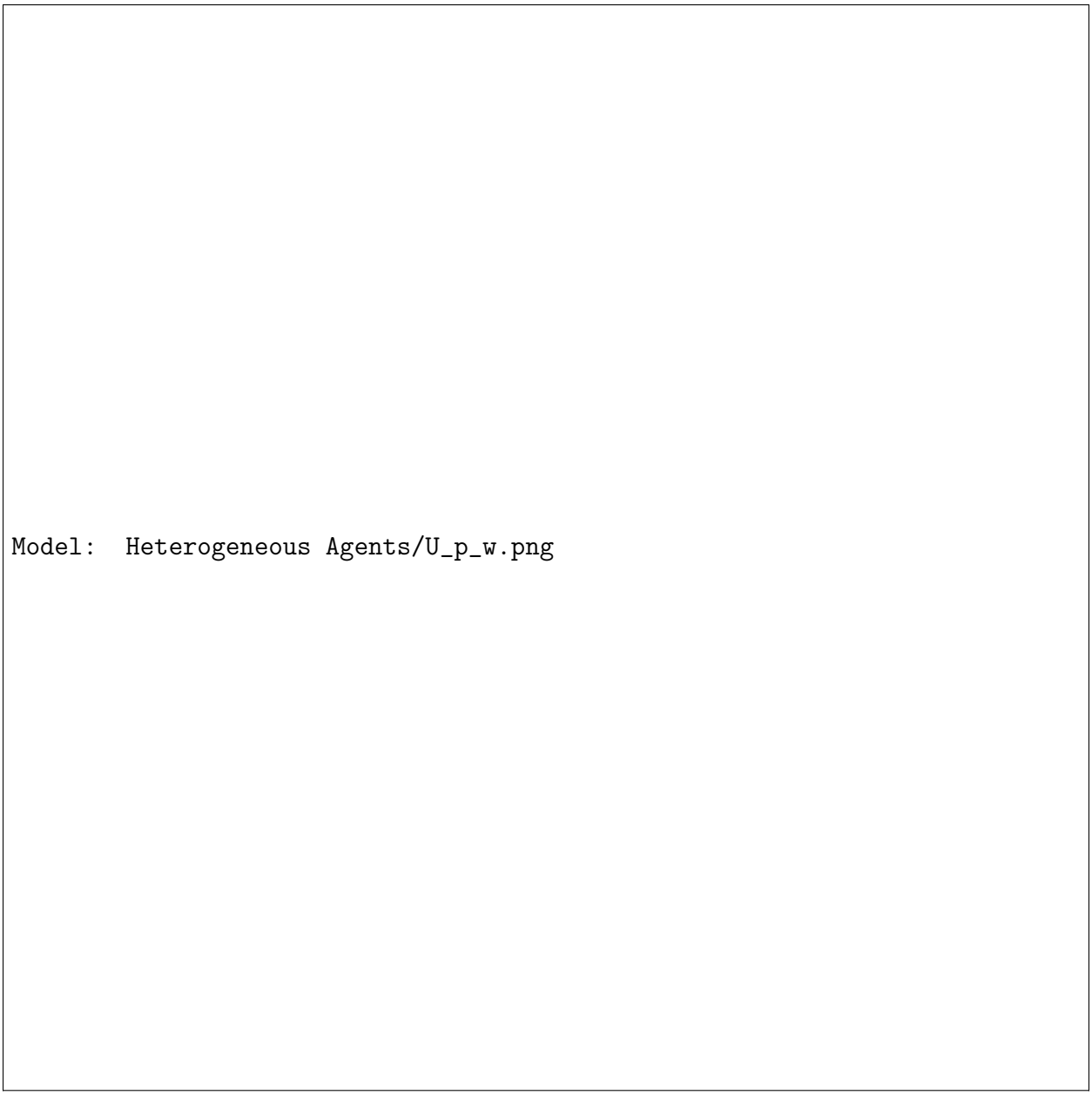


Figure D2: Visualization of residents' willingness to buy housing at different wealth levels. Each solid blue line is an iso-wealth line mapping prices to utility. The dashed line represent a hypothetical reservation utility, and the dotted line represents the price investors are willing to pay. The 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:

$$W_o/P_o = LC \tag{D17}$$

This condition can hold for any wealth group, depending on the price. This is visualized

in Figure D2 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 D2 this new price is at intersection of the dashed line and the liquidity constraint line (the dotted-dashed line).

$$E\left[\frac{(P_o(LC-1)(1+R_f)+Y)^{1-\gamma}}{1-\gamma}\right]+\delta\psi=E\left[\frac{(V+P_o\cdot LC(1+R_f))^{1-\gamma}}{1-\gamma}\right] \tag{D18}$$

D.6 Proposition 2

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, $\alpha_i<0$.

$$\begin{aligned}&\alpha_i<0\\&\left(E[r-r_f]+\frac{1}{2}\sigma_r^2-\kappa_i\right)(1+\exp(\tilde{w}_i-\mu_\nu-v-\kappa_i+\bar{r}_p))<\gamma\sigma_{yr}\end{aligned} \tag{D19}$$

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.

$$\{w,\kappa\}=\{\tilde{w}_i,\kappa_i|\alpha(\tilde{w}_i,\kappa_i)<0\} \tag{D20}$$

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.

$$\begin{aligned}\frac{\partial \alpha}{\partial \kappa_i} &= \frac{1}{\rho_i^2}\rho_i' \frac{1}{\gamma\sigma_r^2}\left(E[r-r_f]+\sigma_r^2/2-\kappa_i-\gamma\sigma_{yr}\right)-\frac{1}{\rho}\frac{1}{\gamma\sigma_r^2} \\ &< \frac{1}{\rho_i^2}\rho_i' \frac{1}{\gamma\sigma_r^2}\left(E[r-r_f]+\sigma_r^2/2-\kappa_i-\gamma\sigma_{yr}+\rho\sigma_{yr}\right)-\frac{1}{\rho}\frac{1}{\gamma\sigma_r^2} \\ &= \frac{1}{\rho_i^2}\rho_i'\left(\rho_i\alpha_i\right)-\frac{1}{\rho}\frac{1}{\gamma\sigma_r^2} \\ &= \frac{1}{\rho}\left(\rho_i'\alpha_i-1\right)\end{aligned} \tag{D21}$$

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 ($\alpha=0$).[35]

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 $\alpha_i=0$

$$\begin{aligned}&\alpha(\tilde{w}_i,k^*(\tilde{w}_i))=0 \rightarrow \\ &E[r-r_f]+\frac{\sigma_r^2}{2}-k^*(\tilde{w}_i)-(1-\rho_i)\gamma\sigma_{yr}^2=0 \rightarrow \\ &\frac{\partial k^*(\tilde{w}_i)}{\partial \tilde{w}_i}=\frac{\rho_i'\gamma\sigma_{yr}}{1+\rho_i'\gamma\sigma_{yr}}>0\end{aligned} \tag{D22}$$

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

$$\begin{aligned}&E[r-r_f]+\sigma_r/2=\gamma\sigma_{yr}(1-\rho(w_i,0)\rightarrow \\ &w_m=w_i(\alpha_i=0,\kappa^*(w_i)=0)=E[d]-r_f+\log\left(\frac{\gamma\sigma_{yr}}{E[r-r_f]+\sigma^2/2}-1\right)\end{aligned} \tag{D23}$$

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

[35] 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.

positive, which is the case per Proposition 1.[36]

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 assets.

However, at $\alpha = 0$ the former effects cancels out[37] 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.

$$\kappa^*(\tilde{w}_i) : E[r - r_f] - \kappa^*(\tilde{w}_i) + \sigma_r^2/2 = \gamma(1 - \rho(\tilde{w}_i, \kappa^*(\tilde{w}_i))\sigma_{yr} \tag{D24}$$

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

$$\lim_{w_i \to \infty} k^*(w_i) = E[r - r_f] + \sigma^2/2 \tag{D25}$$

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 D3 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.

[36]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.

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

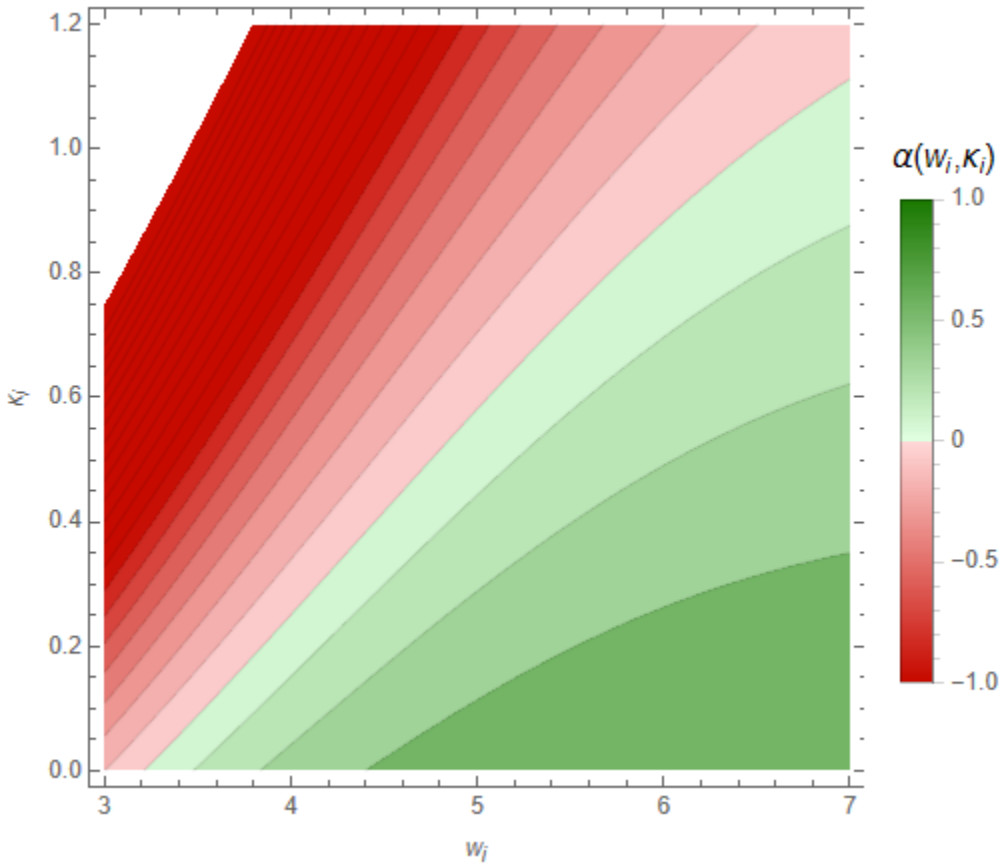


Figure D3: Simulation of $\alpha(\tilde{w}_i, \kappa_i)$, setting $E[r - r_f] = 0.11, \mu_\nu = v + \bar{r}_p = 4, \gamma = 5, \sigma_r^2 = 0.7, \sigma_{yr} = 0.4$

D.7 Proposition 3

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 D26 imposes the assumptions of lemmas 1 and 2 above and uses proposition 3 to aggregate capital supply across the set of active investors investors

$$\underbrace{\int_w \int_\kappa \alpha(w,\kappa) \cdot \mathbb{1}\{\alpha(w,\kappa) \geq 0\} \cdot \tilde{W}_i dF_{w,\kappa}}_{\text{capital supply}} = \underbrace{P \cdot H^{\text{rental}}}_{\text{capital 'demand'}} \tag{D26}$$

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

Proposition 3 also tells us that

$$\frac{\partial \alpha(w,k,r)}{\partial w} > 0 \qquad \frac{\partial \kappa^*(w)(\cdot)}{\partial w} > 0 \tag{D27}$$

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 $\epsilon \to 0$

$$\frac{d\alpha(w_i,\kappa_i,r)}{dw_i} = \frac{\partial \alpha(w_i,\kappa_i,r)}{\partial w_i} + \frac{\partial \alpha(w_i,\kappa_i,r)}{\partial r}\frac{\partial r(w_i)}{\partial w_i} \tag{D28}$$

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

$$\alpha(w_i', \kappa_i, r(w_i')) \geq \alpha(w_i, \kappa_i, r(w_i)) \tag{D29}$$

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 :

$$\begin{aligned} &\int_{w_m}^{\infty} \int_0^{\kappa^*(w)} \alpha(w_i, \kappa_i, r) \cdot (\exp(\tilde{w}))\, dF_{w,\kappa} = \\ &\qquad \int_{w \in [w_m, \infty] \setminus w_i} \int_0^{\kappa^*(w)} \alpha(w_i, \kappa_i, r) \cdot (\exp(w - p_O))\, dF_{\kappa,w} + \\ &\qquad \int_0^{\kappa^*(w_i)} \alpha(w_i, \kappa_i, r(w_i)) \cdot (\exp(w_i - P_O))\, dF_{\kappa|w=w_i} F_{w=w_i} \end{aligned} \tag{D30}$$

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:

$$\exp(\mu + \nu - r(w_i'))F_{w'}(lc - P_O) > \exp(\mu + \nu - r(w_i))F_w(lc - P_O) \tag{D31}$$

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

$$\begin{aligned}
\frac{\partial \log(D/P)}{\partial x} &= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x} + \frac{\partial O(\gamma\sigma_i^2 \exp(z))}{\partial x}\gamma\sigma_i^2 \exp(z)\frac{\partial z}{\partial x} \\
&= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x} + \left(-\frac{1}{\gamma\sigma_i^2 \exp(z)} + \frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{\gamma\sigma_i^2 \exp(z)\left(1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))\right)}\right)\gamma\sigma_i^2 \exp(z)\frac{\partial z}{\partial x} \\
&= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x} + \left(-1 + \frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))}\right)\frac{\partial z}{\partial x} \\
&= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x}\frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))}
\end{aligned} \tag{D32}$$

So for example for $x = w_0$

$$\frac{\partial \log(D/P)}{\partial w_0} = -\frac{\mathcal{W}(\gamma\sigma_i^2 \exp(z))}{1 + \mathcal{W}(\gamma\sigma_i^2 \exp(z))} < 0 \tag{D33}$$

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

$$\begin{aligned}
&\alpha(w^H, \underline{\kappa}) \cdot \exp(\tilde{w}^H) \cdot \underbrace{Pr_w^H Pr_k^0}_{\phi^I} = \exp(p) \cdot \underbrace{Pr_w^0}_{\phi^R} \\
&\alpha(W^H, \underline{\kappa}) = \frac{1}{\rho_I} \frac{E[r - r_f] - \underline{\kappa} + \frac{1}{2}\sigma_r^2}{\gamma\sigma_r^2} - \frac{1 - \rho_I}{\rho_I}\beta_{yr} \\
&p = \log(E[D]) - E[r] = \mu + v + \frac{1}{2}\sigma^2 - E[r] \\
&\left(\frac{1}{\rho_I} \frac{E[r - r_f] - \underline{\kappa} + \frac{1}{2}\sigma_r^2}{\gamma\sigma_r^2} - \frac{1 - \rho_I}{\rho_I}\beta_{yr}\right) \cdot \exp(\tilde{w})\phi^I = \exp(\mu + v + \frac{1}{2}\sigma^2 - E[r]) \cdot \phi^R
\end{aligned} \tag{D34}$$

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

$$x = E[r - r_f] - \underline{\kappa} + \sigma_r^2/2 - (1 - \rho^I)\gamma\sigma_{yr} \tag{D35}$$

We can rewrite equation (D34):

$$\begin{aligned}
&\frac{1}{\gamma\sigma_r^2}\left(E[r-r_f]-\underline{\kappa}+\sigma^2/2-(1-\rho^I)\gamma\sigma_{yr}\right)(W^H-P_o)\phi^I=\rho^I\exp(\mu+v+\sigma^2/2-E[r])\cdot\phi^R\\
&\frac{1}{\gamma\sigma_r^2}x\phi^I=\rho^I\phi^R\exp(\mu+v+\sigma^2/2-\tilde{w^H}-E[r])\\
&x\exp(x)=\gamma\sigma_r^2\rho^I\cdot\frac{\phi^R}{\phi^I}\exp\left(\mu+v+\sigma^2-\tilde{w^H}-r_f-\underline{\kappa}-\gamma\sigma_{yr}(1-\rho^I)\right)\\
&x=\mathcal{W}\left(\gamma\sigma_r^2\rho^I\cdot\frac{\phi^R}{\phi^I}\exp\left(\mu+v+\sigma^2-\tilde{w^H}-r_f-\underline{\kappa}-\gamma\sigma_{yr}(1-\rho^I)\right)\right)
\end{aligned}\tag{D36}$$

From which we derive both returns and real estate portfolio share

$$\begin{aligned}
&E[r-r_f]+\sigma^2/2=\underline{\kappa}+\gamma\sigma_{yr}(1-\rho^I)+\mathcal{W}\left(\gamma\sigma_r^2\rho^I\cdot\frac{\phi^R}{\phi^I}\exp\left(\mu+v+\sigma^2-\tilde{w^H}-r_f-\underline{\kappa}-\gamma\sigma_{yr}(1-\rho^I)\right)\right)\\
&\alpha=\frac{1}{\rho^I\gamma\sigma_r^2}\mathcal{W}\left(\gamma\sigma_r^2\rho^I\cdot\frac{\phi^R}{\phi^I}\exp\left(\mu+v+\sigma^2-\tilde{w^H}-r_f-\underline{\kappa}-\gamma\sigma_{yr}(1-\rho^I)\right)\right)
\end{aligned}\tag{D37}$$

Prices for investors are thus

$$p=\mu+v+\sigma^2/2-\underline{\kappa}-(1-\rho^I)\gamma\sigma_{yr}-\mathcal{W}\left(\gamma\sigma_r^2\rho^I\cdot\frac{\phi^R}{\phi^I}\exp\left(\mu+v+\sigma^2-\tilde{w^H}-r_f-\underline{\kappa}-\gamma\sigma_{yr}(1-\rho^I)\right)\right)\tag{D38}$$

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.)

$$\begin{aligned} &\alpha(w^L, \underline{\kappa}) < 0 \\ &\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right) < \gamma\sigma_{yr}(\sigma_r^2(1-\rho(w^L,\underline{\kappa})) - (1-\rho(w^H,\underline{\kappa})) \end{aligned} \tag{D39}$$

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

$$\begin{aligned} &\alpha(w^H, \underline{\kappa}) > 0 \\ &\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right) > \gamma\sigma_{yr}(1-\rho(w^H,0))(\sigma_r^2 - 1) \end{aligned} \tag{D40}$$

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

$$\begin{aligned} &\alpha(w^L, \bar{\kappa}) < 0 \\ &\mathcal{W}\left(\gamma\sigma_r^2\rho^I \cdot \frac{\phi^R}{\phi^I}\exp(z)\right) < \bar{\kappa} + \gamma\sigma_{yr}(\sigma_r^2(1-\rho(w^H,\bar{\kappa})) - (1-\rho(w^H,0)) \end{aligned} \tag{D41}$$

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:

$$\begin{aligned} U &= E_t\left[\frac{W_1^{1-\gamma}-1}{1-\gamma}\right] \\ &s.t. \\ W_1 &= (1+R_p)W_0 \\ R_p &= \alpha R + (1-\alpha\cdot\iota)R_f \end{aligned} \tag{E1}$$

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):

$$E[r - r_f] + \frac{\sigma^2}{2} = \gamma \Sigma \alpha \tag{E2}$$

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:

$$\begin{aligned} &E[\tilde{r} - r_f] + \tilde{\sigma}^2/2 = \gamma \tilde{\Sigma} \tilde{\alpha} + \gamma \alpha_1 \Sigma_1 \rightarrow \\ &\tilde{\alpha} = \frac{1}{\gamma} \Sigma_1^{-1} \left(E[\tilde{r} - r_f + \tilde{\sigma}^2/2] \right) - \alpha_1 \tilde{\Sigma}^{-1} \Sigma_1, \end{aligned} \tag{E3}$$

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

$$\begin{aligned} E[r_1 - r_f] + \sigma_1^2/2 &= \gamma \alpha_1 \left(\sigma_1^2 + \Sigma_1' \tilde{\Sigma}^{-1} \Sigma_1 \right) + \Sigma_1' \tilde{\Sigma}^{-1} E[\tilde{r} - r_f + \tilde{\sigma}^2/2] \\ &= \gamma \alpha_1 \left(\sigma_1^2 + \Omega \Sigma_1 \right) + \Omega \tilde{Er} \end{aligned} \tag{E4}$$

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.

$$\begin{aligned} \alpha_1 W_0 &= PH \\ \log(\alpha_1) + w_0 &= E[d] - r + h \end{aligned} \tag{E5}$$

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 (E2) and (E5), which is analogous to solving for quantities in any demand-and-supply framework. See Appendix D.1 for full derivation.

$$\begin{aligned}\alpha_1 &= \frac{1}{\gamma\left(\sigma_1^2+\Omega\Sigma_1\right)}\mathcal{W}\left(\gamma\left(\sigma_1^2+\Omega\Sigma_1\right)\cdot\exp(E[d]-r_f+h-w_0+\sigma_1^2/2-\Omega\tilde{E}r)\right)\\ E[r_1-r_f] &= \mathcal{W}\left(\gamma\left(\sigma_1^2+\Omega\Sigma_1\right)\cdot\exp\left(E[d]-r_f+h-w_0+\sigma_1^2/2-\Omega\tilde{E}r\right)\right)+\Omega\tilde{E}r,\end{aligned} \tag{E6}$$

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

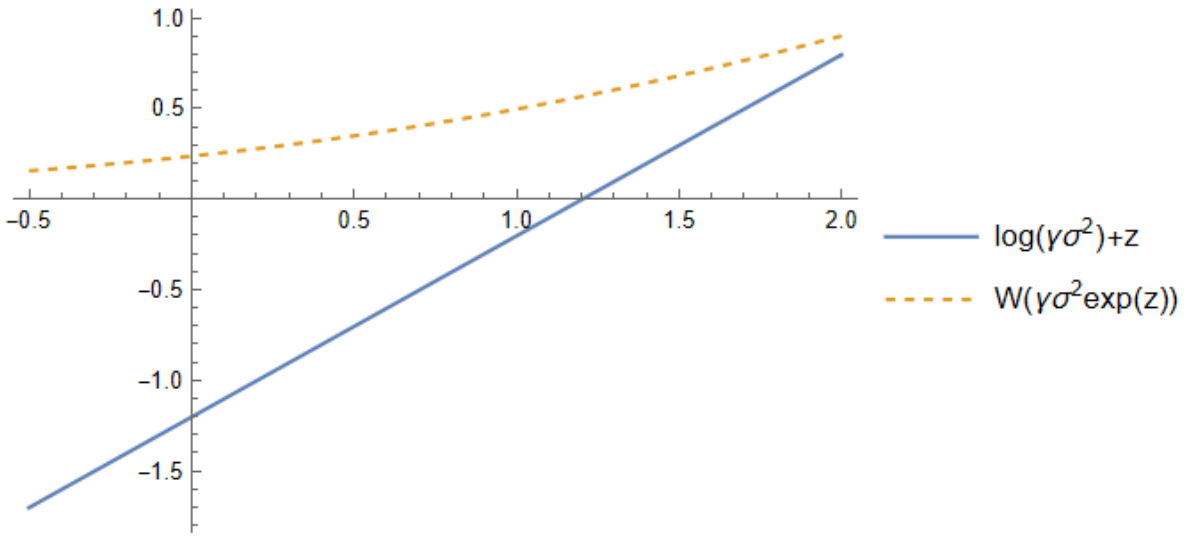


Figure E1: This figure compares returns as derived in equation (E6) to a log-linear benchmark. Dashed line corresponds to $r_1 = W(\gamma\sigma_1^2 \exp(z))$ where $z = E[d] - r_f + h - w_0 + \sigma_1^2/2 - \Omega\tilde{E}r$ to a log-linear benchmark in z. Returns are convex and rise more slowly relative 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 E1 being lower than 1.

To understand the convexity in Figure E1, 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 (E6). 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

$$\begin{aligned}\frac{\partial E\left[\log(D/P)\right]}{\partial x} &= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x}\frac{\mathcal{W}((\sigma_1^2 + \Omega\Sigma_1)\exp(z))}{1 + \mathcal{W}((\sigma_1^2 + \Omega\Sigma_1)\exp(z))} + \frac{\partial \Omega\tilde{E}r}{\partial x} \\ &= \frac{\partial r_f}{\partial x} + \frac{\partial z}{\partial x}\cdot\frac{r_1 - \Omega\tilde{E}r}{1 + (r_1 - \Omega\tilde{E}r)} + \frac{\partial \Omega\tilde{E}r}{\partial x}\end{aligned} \tag{E7}$$

See Appendix D.9 for full derivation. The sign of the comparative static of the expected log yield depends on the sign of $\partial z/\partial x$/[38] 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,

[38] 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:

$$E[r_1] = \mathcal{W}\left(\gamma(\sigma_1^2 + \Omega\Sigma_1) \cdot \exp\left(E[d] - r_f + h - w_0 + \sigma_1^2/2 + \Omega\tilde{E}r\right)\right) \tag{E8}$$

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:

$$\begin{aligned} r_{1ci} &= \phi \log\left(\gamma\left(\sigma_{1c}^2 + \Omega_c\Sigma_{1c}\right) \cdot \exp\left(E[d_{ci}] - r_f + h_{ci} - w_{ic} + \sigma_{1c}^2/2 + \Omega_c\tilde{E}r\right)\right) \\ &= \delta_c + \beta x_{ci} + \phi w_{ci} + \epsilon_{ci} \end{aligned} \tag{E9}$$

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.

$$\begin{aligned}
\alpha_{Eci} &= \frac{1}{\gamma}\tilde{\Sigma}_E^{-1}\tilde{Er} - \alpha_{1ci}\Omega_E' \\
&= \frac{1}{\gamma}\tilde{\Sigma}_E^{-1}\tilde{Er} - \Omega_E' \left(\frac{E[r_{1ci} - r_f + \sigma_{1c}^2/2] - \Omega\tilde{Er}}{\gamma(\sigma_{1c}^2 + \Omega\Sigma_{1c})} \right) \\
&= \frac{1}{\gamma}\left(\tilde{\Sigma}_E^{-1} - \frac{\Omega_E'}{\sigma_{1c}^2 + \Omega\Sigma_{1c}} \right)\tilde{Er} + \Delta_c - \frac{\Omega_E'(\beta X_{ci} + \phi w_{ci} + \epsilon_{ci})}{\sigma_{1c}^2 + \Omega\Sigma_{1c}},
\end{aligned} \tag{E10}$$

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:

$$r_{1ic} = \delta_c + \beta x_{ci} + \phi\omega_{Eci} - \phi\log(\alpha_{Eci}(w_{ci}, \epsilon_{ci})) - \phi\nu_{ci} + \epsilon_{ci} \tag{E11}$$

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:

$$\begin{aligned} r_{1ic} &= \phi^{\text{long}}\omega_{Eci} - \phi\log(\alpha_{Eci}(w_{ci},\epsilon_{ci})) - \phi\nu_{ci} + \epsilon_{ci} \qquad \text{(Long)} \\ r_{1ic} &= \phi^{\text{short}}\omega_{Eci} - \phi\nu_{ci} + \epsilon_{ci} \qquad \text{(Short)} \\ \omega_{Eci} &= \pi\log(\alpha_{Eci}(w_{ci},\epsilon_{ci}) + \zeta_{ci} \qquad \text{(Auxiliary)} \end{aligned} \tag{E12}$$

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:

$$\begin{aligned} \pi &= \frac{\text{Cov}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}),\omega_{ci})}{\text{Var}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}))} \\ &= \frac{\text{Cov}(\log(\alpha_{Eci}),\log(\alpha_{Eci}) + w_{ci} + \nu_{ci})}{\text{Var}(\log(\alpha_{Eci}))} \\ &= \frac{\text{Var}(\log(\alpha_{Eci})}{\text{Var}(\log(\alpha_{Eci})} + \frac{\text{Cov}(\log(\alpha_{Eci}),w_{ci})}{\text{Var}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}))} \\ &= 1 + \frac{\text{Cov}\left(\log\left(g(\Sigma) - \frac{\Omega'_E\phi w_{ci}}{\sigma^2_{1c}+\Omega\Sigma_{1c}}\right),w_{ci}\right)}{\text{Var}(\log(\alpha_{Eci}(w_{ci},\epsilon_{ci},\omega_{ci}))} \end{aligned} \tag{E13}$$

(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.)

$$\begin{aligned} \text{Cov}(x,f(y)) = E[(x-\mu_x)(f(y)-\mu_{f(y)}] = E[xf(y)] - \mu_x\mu_{f(y)} \\ \int_x\int_y xf(y)F_{xy}dydx - \mu_x\mu_{f(y)} < \int_x\int_y xyF_{xy}dydx - \mu_x\mu_y \end{aligned} \tag{E14}$$

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 $w_{ci} > 0$):

$$\text{Cov}\left(\log\left(g(\Sigma_c) - \frac{\Omega_E'\phi w_{ci}}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}\right), w_{ci}\right) < \text{Cov}\left(g(\Sigma_c) - \frac{\Omega_E'\phi w_{ci}}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}, w_{ci}\right) < -\frac{\Omega_E'\phi}{\sigma_{1c}^2 + \Omega\Sigma_{1c}}\text{Var}\left(w_{ci}\right) \tag{E15}$$

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 $\text{Cov}(\epsilon_{ci}, \omega_{ci}|\text{Obs})$:

$$\begin{aligned}
\text{Cov}(\epsilon_{ci}, \omega_{ci}|\text{Obs}) &= \text{Cov}(E[r_{1ci} - \delta_c - \phi h_{ci} - \beta x_{ci} + \phi(\hat{w}_{ci} - \zeta_{ci}), \log(\alpha_{Eci} + w_{ci} + \nu_{ci})|\text{Obs}_{ci}) \\
&= \text{Cov}(E[r_{1ci}] - \phi\zeta_{ci}), \log(\alpha_{Eci} + \nu)|\text{Obs}) \\
&= \text{Cov}(E[r_{1ci}], \log(\alpha_{Eci})) + \underbrace{\text{Cov}(E[r_{1ci}], \nu_{ci}) - \phi\text{Cov}(\zeta_{ci}, \log(\alpha_{Eci})) - \phi\text{Cov}(\zeta_{ci}, \nu)}_{=0}
\end{aligned} \tag{E16}$$

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:

$$\text{Cov}(\epsilon_{ci}, \omega_{ci}|\text{Obs}) = \text{Cov}(E[r_{1ci}], \log(\alpha_{Eci})) < \text{Cov}(E[r_{1ci}], \alpha_{Eci}) = -\frac{\text{Var}(E[r_{1c}])}{\gamma(\sigma_{1c}^2 + \Omega\Sigma_{1c})} < 0 \tag{E17}$$

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

E.2.2 Specification Error

Utilizing a log-linear OLS/2SLS approach implies a log approximation of the Lambert function, introducing potentially significant specification error. Furthermore, in the model, the marginal effect of log wealth on returns is not constant, and instead increases with returns. In this section, I discuss the using the coefficients estimating using OLS/2SLS to interpreting the model, 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:

$$y_i = \mathcal{W}(\gamma\sigma_i \exp(x + \phi w_i + \epsilon_i)), \tag{E18}$$

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:

$$\rho_i(w_i) \equiv \frac{\partial y_i}{\partial w_i}(w_i) = -\phi \frac{y_i(w_i)}{1 + y_i(w_i)} \tag{E19}$$

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.[39]

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:

$$y_i = x + \log(\sigma_i) + \rho w_i + \epsilon_i \tag{E20}$$

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 (E18). The distribution of the variables are set to match the first and second moments of zip rent-price ratios at $\phi = -1$.[40]

[39] 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).

[40] 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

Figure E2 plots the estimated OLS/2SLS 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 (E19): 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/2SLS estimate of $\rho = -0.18$. Such an estimate would be very close to average effect $\rho_i(\bar{y})$, at the observed average return $\bar{y}$. An estimate of $\rho = -0.18$ is larger than as the estimated 2SLS coefficient of -0.07 for log investor share and -0.05 for log wealth. However, these coefficients would only imply a slightly larger structural ϕ coefficient, equal to about -1.25 instead of -1.

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.

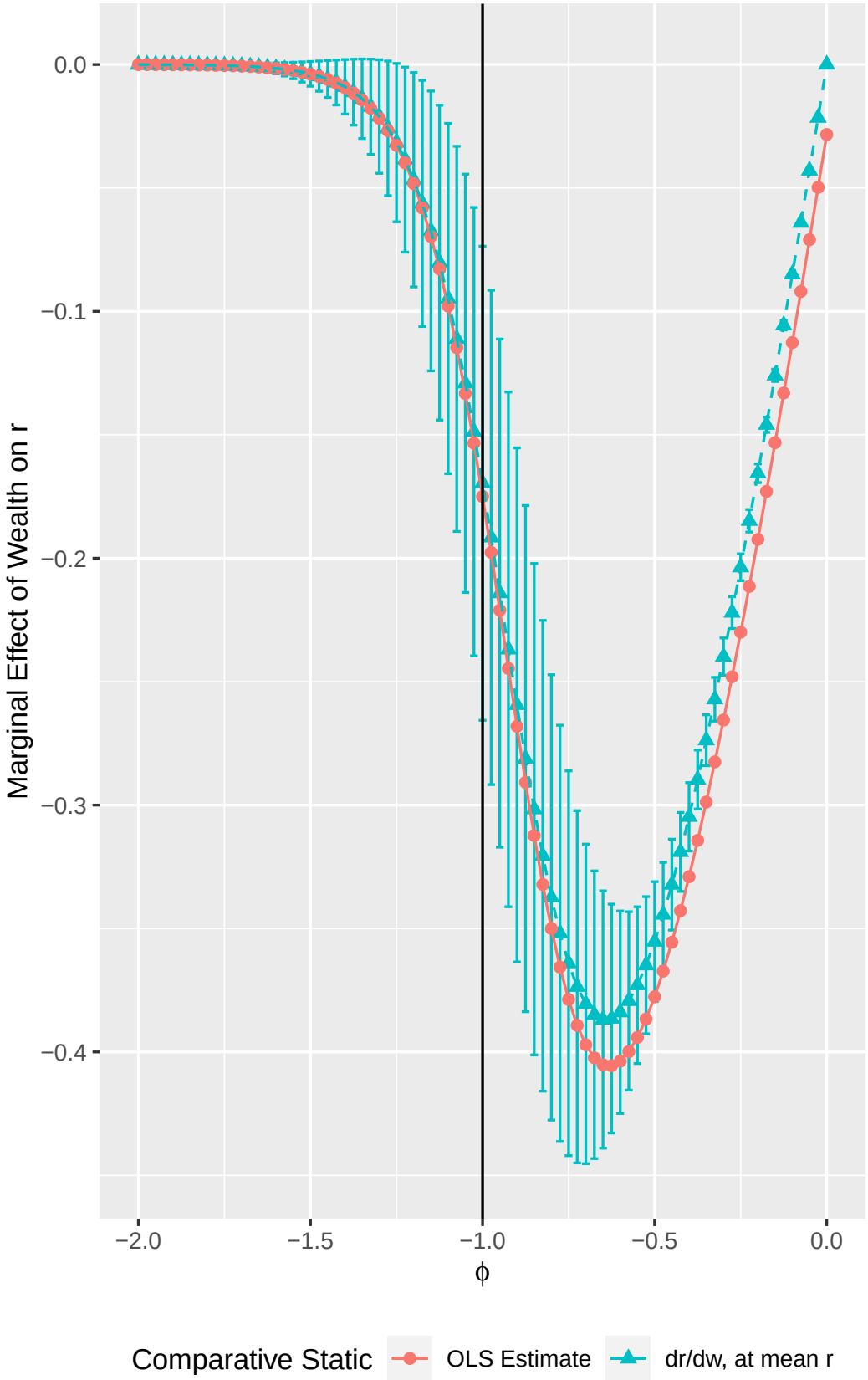


Figure E2: Comparison of the OLS coefficient ρ and marginal effect $\rho_i(\bar{y})$ at $\bar{y}$, the mean y_i. The x-axis presents different values of ϕ. The error bars show the standard deviation of ρ_i around $\rho_i(\bar{y})$. The distribution of returns changes as ϕ changes.

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.[41]

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.[42] 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 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 purged of spurious measurement error.

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.

[41] Most of these investors also live near each unit, but I do not limit the sample to only these investors in these regressions. Section 5.2 calculates average local wealth as the wealth of investors who physically live in each market.

[42] 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)

G Estimating City-level Risk

In this appendix, I describe the estimation procedure for return variance, and briefly discuss the estimation results. The process decomposes the return variance faced by an average real estate investor in city c, $\sigma_c = Var(r_c)$. The challenge is the lack of granular, high-frequency data on real estate returns, inclusive of both long-term rent growth risk and short-term vacancy risk.

Instead, I estimate city-level rent risk by relying on a hedonic rent model with added vacancy risk, and utilize both long-horizon zip-level rent variance and volatility in hedonic prefernce parameters. The model also suggests how to incorporate more easily estimable capital-gains risk.

Set Up

Suppose the 'market rate for a unit i at time t is determined by a location premium L_{zct}, where z is a zip code and a set of physical unit characteristics X_i:

$$d_{izct} = \log(D_{izc}) = L_{zct} + \beta_t X_i + \nu_{izct} \tag{G1}$$

Where ν_{izct} is an unobserved unit-level shock, which is mean-zero at the zip level.

In the hedonic model, two key assumptions are: (1) location characteristics are additively separable from the location premium, and (2) their value, β_t, is constant across space (though it may vary over time). Additionally, I assume that investors anticipate shocks to both β_t and L_{zct}. The expected distribution of shocks to β_t remains constant across both time and space. In contrast, the expected distribution of shocks to L_{zct} is constant over time and within cities but varies between cities. In simpler terms, investors expect a national shock to β_t and a series of local shocks to L_{zct}, with each city's shocks drawn from a unique, city-specific distribution.

To incorporate vacancy risk, I assume that the rent a landlord receives is the unit's fair market rate, adjusted by a 'vacancy tax' representing the time the unit is occupied. Net rent is given by $\tilde{d}_{izct} = d_{izct}\epsilon_{it}$, where $\epsilon_{it} = 1 - x_{it}$. Here, x is the proportion of time the unit is vacant, modeled as an exponentially distributed variable, $x_{it} \sim Exp(\lambda_c)$, with λ_c being a city-specific parameter. This distribution is assumed to be constant over time. [43]

[43] This assumption may lead to an underestimation of risk, as the vacancy-duration distribution likely shifts to the right when rent growth is low, introducing a positive covariance term. Vacancy rates tend to rise when growth is low or negative, suggesting $cov(g^D, \epsilon)$ is positive. Given ϵ represents the proportion of time the unit is occupied, the assumption $E[g^D, \epsilon] = E[g^D]E[\epsilon]$ might be inaccurate. If the positive covariance holds, $E[g^D \cdot \epsilon] > E[g^D]E[\epsilon]$, resulting in a downward bias in the estimate of σ_{dc} below.

I will consider the average one period return for an 'average property' at city c, which is characterized as having the average characteristics of a rental in the city: $X_{jc} = \frac{1}{N_c}\sum_{i \in c} X_i$.

$$P_{jzct} = \frac{\tilde{D}_{jzc,t+1} + P_{jzc,t+1}}{1 + R_{jzct}} \rightarrow$$
$$R_{jzct} = \frac{D_{jzct}}{P_{jzct}}\epsilon_{jt}\left(1 + g^D_{jzct}\right) + g^P_{jzct}$$

Where g^D_{jzct} and g^P_{jzct} are, respectively, the one period rent growth and price growth.

Return Variance

The one period variance for the average unit j, at the average rent-price ratio, can be decomposed into three components: (a) rent growth variance (σ_{dc}), (b) capital gains risk (σ_{pc}), and (c) the covariance between the two.[44]

$$\sigma_{ct} = Var(R_{cz}) = \left(\overline{\frac{D_{izct}}{P_{izct}}}\right)^2 \underbrace{Var(\epsilon(1 + g^D_{jzct}))}_{\sigma_{dc}} + \underbrace{Var(g^P_{jzct})}_{\sigma_{pc}} + \overline{\frac{D_{izct}}{P_{izct}}}\,\underbrace{Cov(\epsilon(1 + g^D_{jzct}), g^P_{jzct})}_{\nu_c} \tag{G2}$$

I start with rent risk, σ_{dc}, for which less data is usually available. I show it to be a function of city-level parameters under the assumption that within city variance in location premium is constant across neighborhoods, but differs across cities. This assumption subsumes two ways a city may be risky: all location premiums may vary (city-wide shocks), or specific neighborhoods may experience shocks.

I also assume that the vacancy risk is independent of rent growth, as well as higher order moments of these random variables. This assumption follows from the assumption that the distribution of vacancy is time-invariant, and the focus on an 'average unit', for which the hedonic unobserved errors (ν_{izct}) average out to zero.

[44] Note that due to the additively separable nature the average rent-price ratio corresponds to unit j at an average of zip-code location premiums.

$$\begin{aligned}\sigma_{dc} &= Var(\epsilon_{ct}) + Var(\epsilon g^D_{jzct}) + Cov(\epsilon_{ct}, \epsilon_{ct} g^D_{jzct}) \\ &= Var(\epsilon) + Var(\epsilon_{ct} g^D_{jzct}) + E[g^D] Var(\epsilon) \\ &= Var(\epsilon_{ct})(1 + E[g^D_{jzct}]) + Var(g^D_{jzct})(Var(\epsilon_{ct}) + E[\epsilon_{ct}]^2) + E[g^D_{jzct}](Var(\epsilon_{ct}) - E[\epsilon_{ct}]^2)]\end{aligned} \tag{G3}$$

Rent risk therefore depends on four variables: (a) vacancy risk, (b) expected vacancy risk, (c) rent growth risk, and (d) expected rent growth. We can identify each of these given the hedonic model and our parametric assumptions on vacancy.

1. **Vacancy risk**: $Var(\epsilon_{ct}) = Var(1 - x_{ct}) = \frac{1}{\lambda_c^2}$
2. **Expected vacancy**: $E[\epsilon_{ct}] = E[1 - x_{ct}] = 1 - \frac{1}{\lambda_c}$
3. **Growth risk**: $Var(g^D_{jzct}) = Var(\Delta L_{zct} + \Delta \beta_t X_j) = \sigma_c^L + \Sigma^\beta X_j + \sigma_{L\beta c}$
   Where σ^L is the variance in location premium for city c, and Σ^β is a variance-covariance matrix of the hedonic value of unit characteristics. $\sigma_{L\beta c}$ is a covariance term, capturing the possibility city c's location premium may vary with the hedonic values. [45]
4. **Expected growth**: $E(g^D_{jzct}) = E(\Delta L_{zct} + \Delta \beta_t X_j) = \mu_c$ where μ_c is the expected rent growth in c, including both expected location premium growth, as well as changes to how renters may value the city's average stock X_j.

Which allows us to write rent risk as follows:

$$\sigma_{dc} = \frac{1}{\lambda_c^2}(1 + \mu_c) + (\sigma_c^L + \Sigma^\beta X_j + \sigma_{L\beta c}) \left(\frac{1}{\lambda_c^2} + \left(1 - \frac{1}{\lambda_c}\right)^2 \right) + \mu_c \left(\frac{1}{\lambda_c^2} - \left(1 - \frac{1}{\lambda_c}\right)^2 \right) \tag{G4}$$

Estimation & Data

Estimating city return risk, σ_c, implies estimating three city-wide quantities: rent risk σ_{dc}, capital gains risk σ_{pc} and their covariance ν_c. Estimating σ_{pc} is relatively straightforward using standard transaction data, like DataQuick or Corelogic. Estimating the covariance term, ν_c, is beyond the scope of this exercise. Instead, I leave it to future work (with better data) and will treat it in my setting as a residual.

I focus on estimating rent risk, σ_{dc}, which involves estimating (1) λ_c, (2) σ_c^L, and (3) Σ^β.

[45] i.e., a work-from-home shock may effect cities with natural beauty while also increasing tastes for larger houses, resulting in a 'double' positive shocks for cities that have both. In practice β_t is quite stable over time, see below.

Measuring λ_c

We can use a city's vacancy rate to impute the parameter λ_c. Suppose there's a mass H of units. At each time τ a density f_τ of units come on the market. By time t a share $\exp -\lambda(t-\tau)$ of these units remain vacant. To compute the overall vacancy rate, we abstract away from within-year differences, and focus on a steady state, setting $f_t = 1/H$

The vacancy rate, measured at time t is therefore

$$V_c = H \int_{-\infty}^{t} \frac{1}{H} \exp(\lambda(t-\tau))d\tau = \frac{\exp(\lambda t - \lambda t)}{\lambda} = \frac{1}{\lambda}$$

λ_c can therefore be easily imputed from the city's vacancy rate.

The ACS measures the vacancy status of all housing units in a city, breaking vacancy down by type of vacancy. One of the categories includes vacant-for-rent, which, together with data on the number of rental units, allows calculation of the vacancy rate among rentals.

Measuring σ_c^L

I measure σ_c^L directly, using a panel on imputed zip-level location premium. This dataset is derived in Samuels (2024), using the hedonic model above and estimating β_t using the AHS.

These data are available for $T = \{1980, 1990, 2000, 2007, 2010, 2013\}$. I compute growth rates between each period, and adjust for the duration of the period, to deliver the implied annual growth rate. I then compute the city variance term $\sigma_c^L = Var(\{\tilde{L}_{zt}\}_{z \in c,t})$.

The use of decadal data, instead of annual data, almost certainly underestimates annual variance in location premium. However, while this is a significant issue in most economic time-series, it is likely less of a concern for rent, which tends to be much more stable, and is more affected by long-term shocks.

Figure G1 displays month-level rent index growth at the city level for select cities. [46] Most cities enjoy stable trends over time, with periods of faster and slower growth corresponding to cities' fundamentals changing. With few exceptions, cities have stable 5 to 10 year periods, which occasionally change or diverge from one another. These data, however, do not address concerns over annual shocks at the nieghborhood level which might cancel out at the city-level.

Measuring $\Sigma^\beta X_{jc}$

The goal is to estimate each city's exposure to the volatility in how American's value a city's housing stock, X_j. To do so, I first estimate Σ^β using a series of hedonic regressions for each

[46] The BLS collects metro-level detailed data for about 30 metro areas, used in CPI calculations.

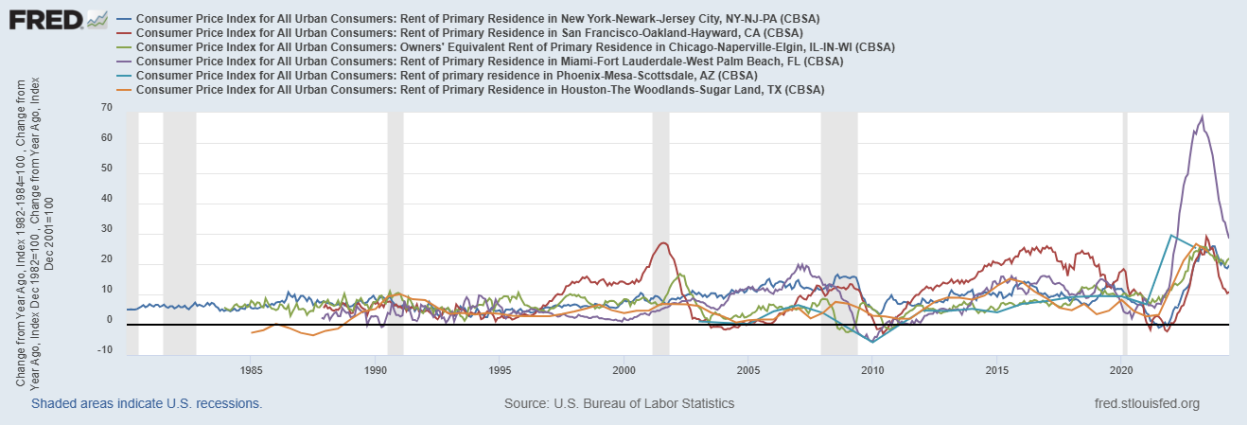


Figure G1: City-level month-level 12-month rent change, for the New York, San Francisco, Chicago, Miami, Phoenix, and Houston metropolitan areas.

survey wave of the American Housing Survey (AHS), conducted every two years from 1973 to 2021. This results in a variance-covariance matrix of the volatility in hedonic valuation, which I then multiply by each city's average housing stock (at each time). This delivers each city's exposure to volatility in hedonic valuation.

$$\log(R_{ic}) = L_c + \beta_t X_{it} + \epsilon_{it} \tag{G5}$$

I first construct a standardized AHS dataset, as detailed in Appendix A. I then estimate a series of regressions, as in Equation (G5), separately for each year. Finally, I compute bi-annual changes to β_t across consecutive waves, adjusting for the duration between each wave.

Many counties report either room counts or square-footage (or both) in their administrative records. As these are key variables in the value of a housing unit, I try three different specifications when running Equation(G5), which differ either include both, or just one of these variables.

A key result of these regressions is that changes in the hedonic coefficients over time are modest, though some secular trends do appear. Furthermore, the differences between specifications (including either or both of rooms and square footage) are small, with only a level shift in the value of square-footage when the rooms variable is omitted.

Figure G2 illustrates the evolution of various components of β_t over time. From 1980 to 2000, the building age effect on rents became more convex, a trend which reversed from 2000 to 2020. The value of additional bathrooms has declined over the past forty years, whereas the value of bedrooms (controlling for the number of rooms) has increased. The values of room count and square footage have remained mostly constant, except for the 2015 to 2021

cycles. This anomaly is likely due to a change in square footage reporting in 2015, from a continuous variable to a discretized and top-coded one. Therefore, I use survey waves only up to 2013 when computing Σ^{β}

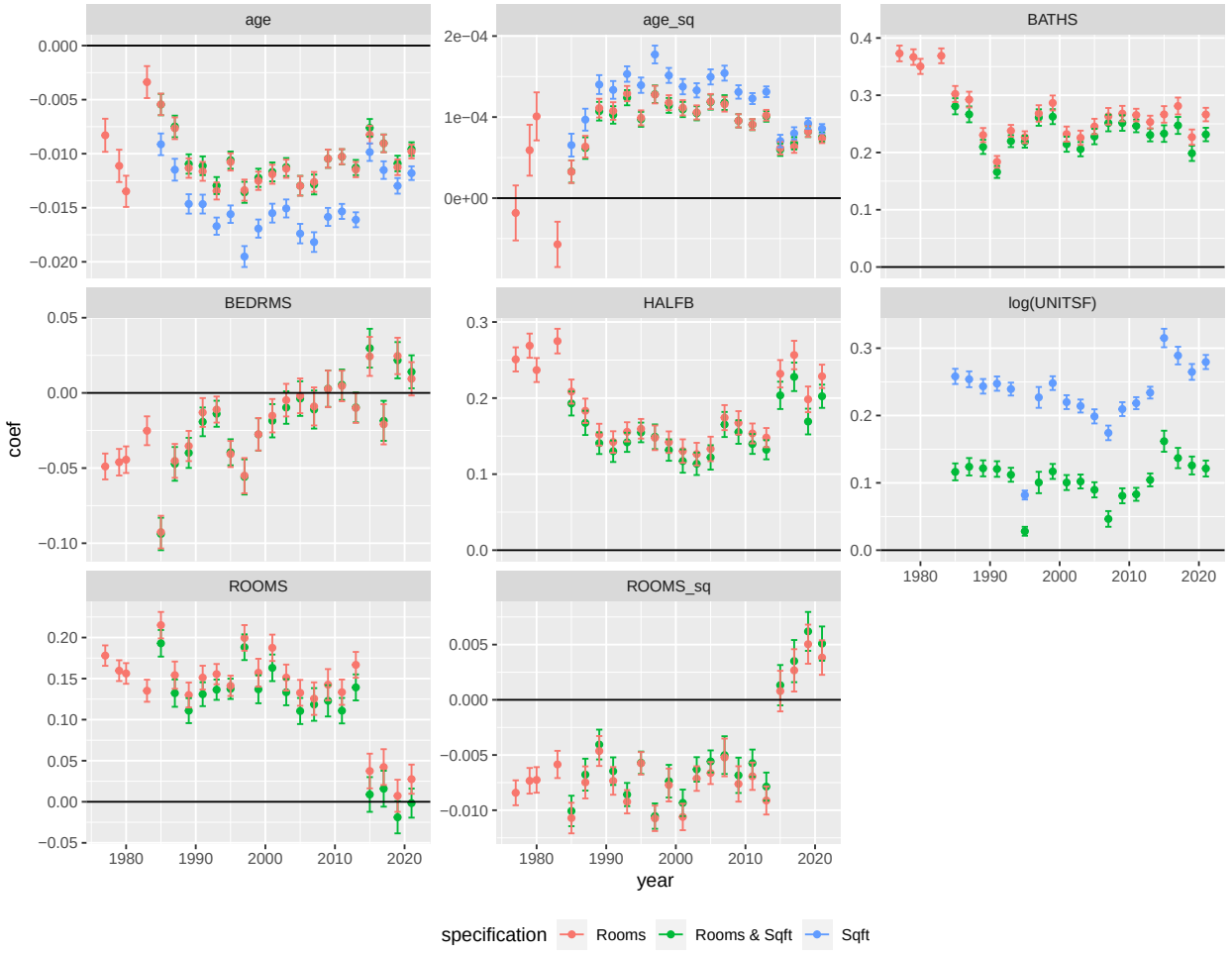


Figure G2: The hedonic value of select housing characteristics, estimated based on AHS data and including city and urban/suburban fixed effects.

Finally, I compute the variance-covariance matrix of changes in β_t across consecutive waves, from 1975 to 2013, separately for each specifications; (just rooms, just square footage, and both.)

Results

I estimate rent risk (σ_{dc}) using estimates for λ_c, $\Sigma^{\beta} X_{jc}$ and σ_c^L, by combining these terms as in Equation (G4). The average rent risk is 0.084 (8.4%), with most cities in a relatively tight band in the 0.075-0.084 (7.5%-8.4%) range. The magnitude can be thought of as relative

to growth rates, or to returns, and they imply a considerable variance in rental payments, relative to an average rent price ratio of 12.5%. A decomposition below suggests that most of this risk is due to vacancy risk, and not growth risk.

There are some regional differences. Figure G3 present a map of the estimated rent variance σ_{dc}, and table G1 linearly decomposes the variance into its three sources. Cities with higher variances appear to cluster in the south, and particularly in Florida. Several Northeastern cities, including New York, and some cities in the Easter Midwest, such as Detroit, have somewhat elevated variance, relative to cities on the West coast and the Upper Midwest. There does not appear to be large differences between small and large cities - with cities in both categories having both higher and lower risk.

According to a linear decomposition of the variance in σ_{dc} (an ANOVA decomposition), most cross-city differences in rent risk are due to different vacancy rates. About 85% of the inter-city variance in risk is explained by this variable. Land premium volatility accounts for 6% of the variance, while volatility in hedonic values accounts for 7.6%. The low importance of location premium might be a result of the use of decadal data, which may mask annual changes which average out.

Table G1: Relative Importance of Variances Sources

Source	**Relative Importance (lmg)**
σ_c^L	0.0598
$\Sigma^\beta X_{jc}$	0.0760
λ_c	0.8491

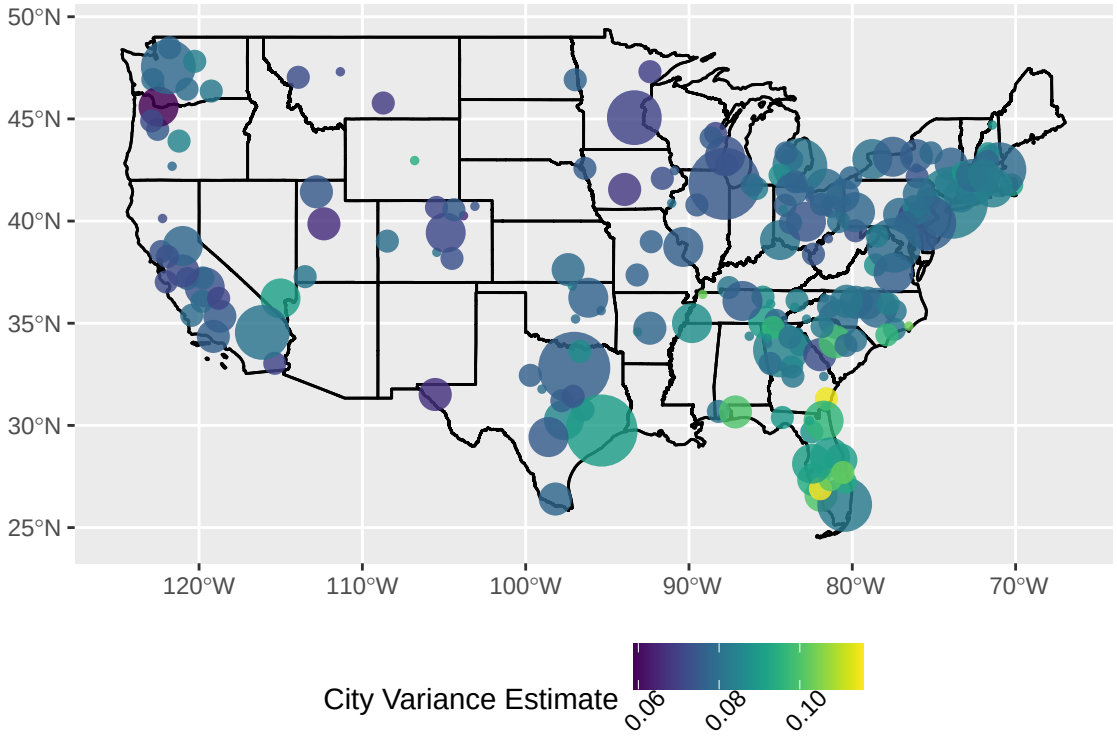


Figure G3: Estimates of σ_{dc}

H Identification of the Structural Estimation

This appendix establishes identification of the vector of structural parameters $\Theta = \{\beta, \gamma\} \in \mathbb{R}^{k+1}$, where β is a $k \times 1$ vector and γ is a scalar, from three moment conditions used in the analysis. Each moment condition generates a function that depends nonlinearly on Θ, through the Lambert function $\mathcal{W}$, which is interpreted as the principal branch. For ease of exposition I define $\beta_k = -\theta$, and treat it as one of the vector of covariates β.

$$\mathcal{W}(z(x, w, \phi \mid \Theta)) = \mathcal{W}(\gamma\phi \exp(\{x, w\}^\top \beta)).$$

Where x is a vector of growth-predicting covariates, w is wealth, and ϕ is the renter-investor ratio ϕ^R/ϕ^I. For brevity, I omit the 'constant' terms (w.r.t. to Θ) in the exponent.

Overall Objective Function

Let $m_j(\cdot \mid \Theta)$ denote the j-th moment function, for $j = 1, 2, 3$. The estimator minimizes the following GMM-type criterion function:

$$Q(\Theta) = \sum_{j=1}^{3} \left(\mathbb{E}[m_j(\cdot \mid \Theta)]\right)^2.$$

Identification of Θ requires showing that $Q(\Theta)$ is uniquely minimized at the true parameter Θ^*. For each moment, we demonstrate the following:

1. The moment condition is satisfied at the true value: $\mathbb{E}[m_j(\cdot \mid \Theta^*)] = 0$,
2. The function $\Theta \mapsto \mathbb{E}[m_j(\cdot \mid \Theta)]$ is strictly monotonic along any direction from Θ^*,
3. The squared moment $\left(\mathbb{E}[m_j(\cdot \mid \Theta)]\right)^2$ is convex in β, implying that $Q(\Theta)$ is globally convex and uniquely minimized at Θ^*.

We now present and analyze each of the three moments in turn.

Moment 1: Growth Predictors

The first moment is formed as the interaction of the $k-1$ growth-predicting covariates x with the residual between observed and predicted return, and intuitively identifies the vector $\beta \in \mathbb{R}^{k-1}$, not including $\beta_k = -\theta$. The residual is $\epsilon_i(\cdot) = r_i - \hat{r}(\mathcal{W}(z \mid \Theta))$ is the residual between the model's predicted return and the observed return. The moment condition thus a vector of length k, with each element equaling zero at the true parameter value.

$$\mathbb{E}[m_1(x \mid \Theta^*)] = \mathbb{E}\left[x \cdot (C - \mathcal{W}(x, \phi \mid \Theta^*))\right] = 0,$$

where C does not depend on x's or Θ), and equals the observed return minus the term κ and $(1-\rho)\sigma_{yr}$.

Assumptions:

A1. The growth-predicting covariate vector $x \in \mathbb{R}^{k-1}$ satisfies $x > 0$ componentwise with positive probability.

A2. The true parameter Θ^* satisfies the moment condition: $\mathbb{E}[m_1(x \mid \Theta^*)] = 0$.

Moment definition:

$$m_1(x, \mathcal{W}(z \mid \Theta)) = x \cdot \epsilon(\mathcal{W}(z \mid \Theta)) \in \mathbb{R}^k,$$

and the objective function:

$$Q_1(\Theta) = \|\mathbb{E}[m_1(x \mid \Theta)]\|^2 = \mathbb{E}[m_1(x \mid \Theta)]^\top \mathbb{E}[m_1(x \mid \Theta)].$$

1. Value at true Θ^*: By assumption **A2**, the moment condition is satisfied at the true parameter value:

$$\mathbb{E}[m_1(x \mid \Theta^*)] = 0.$$

Furthermore, the FOC is satisfied at this value, as

$$\nabla_\Theta Q_1(\Theta) = 2J(\Theta)^\top \mathbb{E}[m_1(z \mid \Theta])],$$

which equals zero through the moment condition.

2. Uniqueness & Convexity:

1. β: To show local uniqueness of β^* as minimizing the objective function Q_1, w.r.t. the vector β I show that the Jacobian implied by the derivative of the moment vector w.r.t to the vector β is negative semidefinite. I then show global uniqueness by arguing the function is monotonic in each β, and that the Hessian is positive semidefinite.

   Let $J(\beta) = \nabla_\beta \mathbb{E}[m(z \mid \Theta)] \in \mathbb{R}^{k \times k}$ be the Jacobian matrix. Each entry of $J(\beta)$ is given by:

$$\frac{\partial m_i(z(x \mid \beta))}{\partial \beta_j} = -x_i x_j \cdot \frac{\mathcal{W}(z(x \mid \beta))}{1 + \mathcal{W}(z(x \mid \beta))},$$

by the properties of the Lambert function. The Jacobian is:

$$J(\beta) = -\mathbb{E}\left[xx^\top \cdot \frac{\mathcal{W}(z(x \mid \beta))}{1 + \mathcal{W}(z(x \mid \beta))}.\right]$$

Since $x > 0$ component-wise and $\mathcal{W}(z(x \mid \beta))/(1 - \mathcal{W}(z(x \mid \beta))) > 0$ for all $z > 0$, and that $xx^\top$ is symmetric and positive definite (e.g., x has full support), we conclude:

(a) $J(\beta)$ is negative definite and full rank.

(b) The gradient of the objective is:

$$\nabla_\beta Q_1(\beta) = 2J(\beta)^\top E[m_1(z \mid \Theta)].$$

(c) At the true value Θ^*, we have $\mathbb{E}[m_1(z \mid \Theta^*) = 0$, so:

$$Q_1(\Theta^*) = 0, \quad \nabla_\beta Q_1(\beta^*) = 0.$$

To establish monotonicity and convexity, consider a deviation $v = \beta - \beta^* \geq 0$ (strict for at least one element in β). It follows that $x^\top\beta > x^\top\beta^*$ and

$$C - \mathcal{W}(z(x^\top\beta, \theta, \gamma)) < C - \mathcal{W}(z(x^\top\beta^*, \theta, \gamma))$$

Implying the following vector inequality:

$$m_1(z \mid \Theta) < m_1(z \mid \Theta^*)$$

And so

$$E[m_1(z \mid \Theta)] < E[m_1(x \mid \Theta^*)] = 0.$$

The opposite holds for $v \leq 0$. It therefore follows that:

(d) $\nabla_\beta Q_1(v > 0) = 2J(\beta)^\top E[m_1(x \mid \beta > \beta^*, \theta, \gamma)] > 0$ since $E[m_1(x \mid \beta > \beta^*, \theta, \gamma)] < 0$ for $v > 0$. That is, the derivative of the objective function is positive when $\beta > \beta^*$. Similarly $\nabla_\beta Q_1(v < 0) = 2J(\beta)^\top E[m_1(x \mid \beta < \beta^*)] < 0$ for $\beta < \beta^*$.

(e) The Hessian is:

$$\nabla^2_\beta Q_1(\beta) = 2J(\beta)^\top J(\beta) + 2\sum_{i=1}^{k} E[m_{1,i}(z \mid \beta, \theta, \gamma)) \cdot \nabla^2_\beta m_{1,i}(x \mid \beta, \theta, \gamma),$$

which is positive definite at $\Theta = \Theta^*$ since $\mathbb{E}[m_1(z \mid \Theta^*)] = 0$.

The objective function decreases with β for $\beta < \beta^*$ and increases for $\beta > \beta^*$, and is locally minimized at β^* holding w, γ at their true values.

2. γ: Similarly to the derivative with respect to θ, $Q_1(\gamma)$ is a vector of moments, that are each minimized at Θ^* at global minimum, and the function is monotonic and convex. Consider element j in the vector of derivatives $\partial Q_1(\gamma)/\partial\gamma$:

$$\frac{\partial Q_{1j}(\theta)}{\partial\theta} = -2E[m_1(x, \mathcal{W}(z \mid \Theta))] \cdot \mathbb{E}\left[\left(x_j \frac{1}{\gamma} \frac{\mathcal{W}(z \mid \Theta)}{1 - \mathcal{W}(z \mid \Theta)}\right)\right]$$

Which equals 0 at Θ^*, by A2. It further follows that for $\gamma > \gamma^*$ we have $\mathcal{W}(z \mid \beta^*, \theta^*, \gamma) > \mathcal{W}(z \mid \beta^*, \theta^*, \gamma^*)$ and $m_{1j}(z \mid \beta^*, \theta, \gamma^*) < \mathcal{W}(z \mid \beta^*, \theta^*, \gamma^*)$, and so $\frac{\partial Q_{1j}(\gamma)}{\partial\gamma}|_{\gamma>\gamma^*} > \frac{\partial Q_{1j}(\gamma)}{\partial\gamma}|_{\gamma=\gamma^*} = 0$. Vice versa for $\gamma < \gamma^*$. Finally the vector $Q_1(\theta)$ is concave, as each element of the objective function vector is convex:

$$\frac{\partial^2 Q_{1j}(\gamma)}{\partial^2\gamma} = \frac{2}{\gamma^2}\left(\mathbb{E}\left[\left(x_j \frac{\mathcal{W}(z \mid \Theta)}{1 - \mathcal{W}(z \mid \Theta)}\right)\right]\right)^2$$

is positive.

Moment 2: AREP Instrument

The second moment is formed as the interaction of the scalar instrument d with the residual between observed and predicted return, and intuitively identifies the parameter $\gamma \in \mathbb{R}_+$. As in the first moment the residual is $\epsilon_i(\cdot) = r_i - \hat{r}(\mathcal{W}(z \mid \Theta))$.

Assumptions:

A1. The growth-predicting covariate vector $x \in \mathbb{R}^k$ satisfies $x > 0$ componentwise with positive probability.

A2. The instrument is demeaned $\mathbb{E}[d] = 0$ and it's marginal expectation,$\mathbb{E}[d|\phi]$, is monotonic with respect to ϕ. (The same holds for any x instead of ϕ.)

A3. The true parameter Θ^* satisfies the moment condition: $\mathbb{E}[m_1(x \mid \Theta^*)] = 0$.

Moment definition:

$$m_2(x, \mathcal{W}(z \mid \Theta)) = d \cdot \epsilon(\mathcal{W}(z \mid \Theta)) \in \mathbb{R}^k,$$

and the objective function:

$$Q_2(\Theta) = \|\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta)]\|^2 = \mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta)]^\top \mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta)].$$

1. Moment equals zero at true Θ^*: By assumption **A2**, the moment condition is satisfied at the true parameter value:

$$\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] = 0. \tag{1}$$

The first order condition is satisfied at this value, as:

$$\frac{\partial Q}{\partial \Theta} = 2\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta)] \cdot \mathbb{E}[\frac{\partial}{\partial \Theta} m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta)] \tag{2}$$

which equals zero at Θ^* due to the moment condition.

2. Uniqueness & Convexity: I establish uniqueness and convexity by showing that $Q_2(\Theta)$ is strictly decreasing for any element in $\Theta < \Theta^*$ and increasing for $\Theta > \Theta^*$. Furthermore, the function is convex and the second order condition is satisfied with respect to each element in Θ. Since $Q_2(\Theta)$ is a scalar (as opposed to Q_1) we can consider only a vector of derivatives and show that the above result holds for each element.

1. β: For ease of notation, define

$$h(x, z, \Theta) := \frac{x \cdot \mathcal{W}(z(\Theta))}{1 + \mathcal{W}(z(\Theta))} > 0, \tag{3}$$

and increases in ϕ. Note that

$$\frac{\partial m_2}{\partial \beta}(d, x, z \mid \Theta) = -d \cdot h(x, z, \Theta).$$

The derivative of the objective function is

$$\frac{\partial Q_2(\Theta)}{\partial \beta} = -2\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta)] \cdot \mathbb{E}[d \cdot h(x, z, \Theta)] \tag{4}$$

By assumption A2 and by Chebyshev's sum inequality, the sign of the derivative depends on the sign of $\text{Cov}(d, \phi)$ (analogously for x):

$$\mathbb{E}[d \cdot h(x, \phi, \beta)] = \mathbb{E}[\mathbb{E}[d \mid \phi] \cdot h(x, \phi, \beta)] > \mathbb{E}[\mathbb{E}[d|\phi]] \cdot \mathbb{E}[h(x, \phi, \beta)] = 0, \tag{5}$$

and if $\text{Cov}(d, x) < 0$ then:

$$\mathbb{E}[d \cdot h(x, \phi, \beta)] = \mathbb{E}[\mathbb{E}[d \mid \phi] \cdot h(x, \phi, \beta)] < \mathbb{E}[\mathbb{E}[d|\phi]] \cdot \mathbb{E}[h(x, \phi, \beta)] = 0.$$

Consider first the case that $\text{Cov}(d, \phi) > 0$. We know (from A3) that at $\Theta = \Theta^*$,

$\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] = 0$, and that its derivative is positive. It follows that for any $\beta > \beta^*$ then $\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] > 0$ and for $\beta < \beta^*$ then $\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] < 0$. From Equation (4) it follows that the objective function decreases for $\beta < \beta^*$ and otherwise increases. The same holds if $\text{Cov}(d, \phi) < 0$, with the signs reversed, but canceling out.

To establish convexity, consider the second derivative at Θ^*

$$\frac{\partial^2 Q_2(\Theta)}{\partial^2 \beta}|_{\Theta=\Theta^*} = 2\left(\mathbb{E}[d \cdot h(x, z, \Theta)]\right)^2 > 0, \tag{6}$$

regardless of $\text{Cov}(d, \phi)$, impliying the objective function is convex.

2. γ: Identification of γ follows a similar argument. Note that

$$\frac{\partial m_2}{\partial \gamma}(d, x, z \mid \Theta) = d\frac{1}{\gamma}\frac{\mathcal{W}(z(\Theta))}{1 + \mathcal{W}(z(\Theta))}. \tag{7}$$

The sign of which also follows the sign of $\text{Cov}(d, \phi)$, by the same argument utilizing Chebyshev's inequality and for a slightly different $h(\cdot)$ function: $h(z, \mid \Theta) = \frac{\mathcal{W}(z(\Theta))}{1+\mathcal{W}(z(\Theta))}$. For any $\gamma > \gamma^*$ then $\mathbb{E}[m_2(z \mid \Theta^*)] > 0$ and for $\beta < \beta^*$ then $\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] < 0$. From Equation (**??**) it follows that the objective function decreases for $\beta < \beta^*$ and otherwise increases.

The second derivative equals:

$$\frac{\partial^2 Q_2(\Theta)}{\partial^2 \gamma}|_{\Theta=\Theta^*} = \frac{2}{\gamma^2}\left(\mathbb{E}[\, d \cdot \frac{\mathcal{W}(z(\Theta))}{1 + \mathcal{W}(z(\Theta))}\,]\right)^2 > 0, \tag{8}$$

which is always positive.

Note that for $\gamma \gg 0$ the first order is 'close' to zero around Θ^*, suggesting that while identification is unique, it might be difficult to implement.

Moment 3: Wealth Shock

The third moment equates the average difference in returns given an increase to wealth, ΔW, and intuitively identifies the effect of wealth, $\beta_k = -\theta$.

Assumptions:

A1. $\Delta W \ll \mathbb{E}[W]$.

A2. The true parameter Θ^* satisfies the moment condition: $\mathbb{E}[m_3(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] = 0$.

Moment definition:

$$m_3(x, \mathcal{W}(z \mid \Theta)) = \Delta r(\Delta W \mid \Theta) - D,$$

where $\Delta r(\Delta W)$ is shorthand for $r(x, \phi, W \mid \Theta) - r(x, \phi, W + \Delta W \mid \Theta)$ and the objective function:

$$Q_3(\Theta) = \left(\mathbb{E}[\Delta r(\Delta W \mid \Theta) - D]\right)^2.$$

1. Moment equals zero at true Θ^*: By assumption **A2**, the moment condition is satisfied at the true parameter value:

$$\mathbb{E}[m_2(x, \mathcal{W}(z \mid \Theta) \mid \Theta^*)] = 0. \tag{9}$$

The first order condition is satisfied at this value, as:

$$\frac{\partial Q_3}{\partial \Theta} = 2\mathbb{E}\left[m_3(x, \mathcal{W}(z \mid \Theta) \mid \Theta)\right] \cdot \mathbb{E}\left[\frac{\partial}{\partial \Theta} m_3(x, \mathcal{W}(z \mid \Theta) \mid \Theta)\right] \tag{10}$$

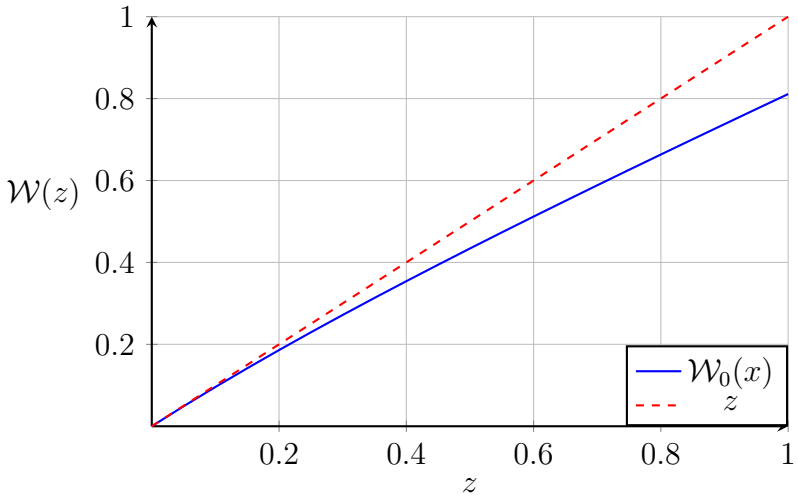
which equals zero at Θ^* due to the moment condition.

2. Uniqueness & Convexity: This moment only 'weakly identifies γ and β_1 through β_{k-1}, but identifies β_k.

1. β_1 through β_{k-1}: The derivative of the moment function with respect to any $\beta_j \in \{\beta_1, \ldots, \beta_{k-1}\}$ is:

$$\frac{\partial m_2}{\partial \theta}(d, x, z \mid \Theta) = x_j \cdot \left(\frac{\mathcal{W}(Z(w_1, \Theta))}{1 + \mathcal{W}(Z(w_1, \Theta))} - \frac{\mathcal{W}(Z(w_2, \Theta))}{1 + \mathcal{W}(Z(w_2, \Theta))}\right), \tag{11}$$

   Where $w_2 = \log(W_1 + \Delta)$. Equation 12 suggests that the objective function in Equation 10 is identified, as $\mathcal{W}(Z(w_1, \Theta)) < \mathcal{W}(Z(w_2, \Theta))$ as long $\beta_k^* < 0$, suggesting $Q_3(\Theta)$ declines for $\beta_j < \beta_j^*$). However, in practice Equation (12) is likely close to zero for any β_j close to β^*. This is because for reasonable values of $\mathcal{W}(z(\Theta))$, between 0 and 0.5 (so that maintenance costs + risk adjustment are not greater than 0.5), its derivative is flat, and the function is near linear.



2. γ: the derivative w.r.t. γ is

$$\frac{\partial m_3}{\partial \gamma}(d, x, z \mid \Theta) = \frac{1}{\gamma} \cdot \left(\frac{\mathcal{W}(Z(w_1, \Theta))}{1 + \mathcal{W}(Z(w_1, \Theta))} - \frac{\mathcal{W}(Z(w_2, \Theta))}{1 + \mathcal{W}(Z(w_2, \Theta))} \right). \tag{12}$$

which is near zero for the same reasons as above.

3. $\beta_k = -\theta$: The third moment identifies β_k. The derivative of the moment function is:

$$\frac{\partial m_3}{\partial \theta}(d, x, z \mid \Theta) = w_1 \frac{\mathcal{W}(z(w_1, \Theta))}{1 + \mathcal{W}(z(w_1, \Theta))} - w_2 \frac{\mathcal{W}(z(w_2, \Theta))}{1 + \mathcal{W}(z(w_2, \Theta))}. \tag{13}$$

The sign of Equation (17) is uncertain, as $w_1 < w_2$ but the derivative decreases in w. I test when the former effect dominates. Denote

$$g(w \mid \Theta) = \mathcal{W}(z(w \mid \Theta)). \tag{14}$$

This is a downwards sloping, convex function. Note that for any function $h(y) = f'(x)x$, we have $h'(y) < 0$ if

$$f''(x)x + f'(x) > 0 \tag{15}$$

Taking $f = g$ and utilizing the derivative of the Lambert function, we have that:

$$\begin{aligned} &\frac{\theta^2 g(w)}{(1 - g(w))^2(1 + g(w))} - \frac{\theta g(w)}{(1 - g(w))} < 0 \\ &\frac{\theta}{1 - g(w)^2} < 1 \\ &\theta < 1 - g(w)^2. \end{aligned} \tag{16}$$

Which implies that

$$\frac{\partial m_3}{\partial \theta}(d, x, z \mid \Theta) < 0 \iff \theta < 1 - \mathcal{W}(z \mid \Theta)^2 \approx 1. \tag{17}$$

If $\theta < 1$, then

$$\frac{\partial Q_3}{\partial \Theta} = 2\mathbb{E}[m_3(x, \mathcal{W}(z \mid \Theta))] \cdot \underbrace{\mathbb{E}\left[w_1 \frac{\mathcal{W}(z(w_1, \Theta))}{1 + \mathcal{W}(z(w_1, \theta))} - w_2 \frac{\mathcal{W}(z(w_2, \Theta))}{1 + \mathcal{W}(z(w_2, \Theta))}\right]}_{<0}, \tag{18}$$

which means that if $\theta > \theta^*$ then $m_3(\theta) < m_3(\theta^*)$, and $\mathbb{E}[m_3(x, \mathcal{W}(z \mid \beta^*, \theta < \theta^*, \gamma^*))] < 0$, leading to $\frac{\partial Q_3}{\partial \theta} > 0$. If $\theta < \theta^*$ then vice versa, and $\frac{\partial Q_3}{\partial \theta} < 0$ so that the objective function declines monotonically for $\theta < \theta^*$ and increases otherwise, so that θ^* is a global minimum.