The Commerical Real Estate Ecosystem

November 19, 2024

Abstract

We develop a new approach to understand the joint dynamics of transaction prices and trading volume in the market for commercial real estate. We start from a micro-founded model in which buyers and sellers differ in their private valuation of building characteristics, such as size, location, and quality. Consistent with the decentralized nature of the commercial real estate market, we model the probability that a seller meets a particular buyer, where the meeting probability depends on the characteristics of the buyer, the seller, and the building. In equilibrium, the mapping from building characteristics to observed transaction prices depends on the identity of the buyer and the seller, which is missed by traditional hedonic valuation models. We estimate the model using granular data on commercial real estate transactions, which contain detailed information on the identity of buyers and sellers. Our central finding is that the identity of buyers and sellers has a first-order effect on transaction prices and trading volume. We show how the model can be used for counterfactual analyses. As a concrete example, we find that the Manhattan office market would have seen 10% lower valuations if it had not been for a large inflow of foreign buyers in 2018–2021. Our methodology directly extends to other private markets, including private equity, private credit, and infrastructure.

1 Introduction

Financial markets have been undergoing a major shift from public to private markets over the past 25 years. For example, the number of publicly traded stocks in the U.S. decreased from around 8,000 in 1997 to around 4,000 in 2024. Over this same period, assets under management in the private equity market, broadly defined, have grown from a small base to $7.1 trillion in the U.S. and to $13 trillion globally.[1] Many large institutional investors have decreased the share of their portfolios away for public equity and fixed income markets into "alternatives." For example, public pension funds have increase the share in private and real assets from below 10% in the late 1990s to over 25% by 2020 (Mittal, 2024).[2] The recent surge in private debt markets is another manifestation of this trend.[3] Using a broad definition of private and real assets that includes privately-held businesses and collectibles, Goetzmann et al. (2021) estimate a total asset base of $84 trillion in 2020.

This "great rotation" to private and real assets has important implications for finance research and practice. Private and real assets differ in important ways from public assets. They trade infrequently, often in bilateral search and matching markets. Prices are observed infrequently, which poses challenges for risk management and creates scope for volatility laundering. Assets are lumpy and heterogeneous, where unique features such as asset location can be important. Finally, there is an ecosystem of specialized investors. This paper provides a new methodology for understanding the valuation of such assets.

Under the traditional rational asset pricing paradigm, the identity of buyers and sellers is irrelevant. If one buyer or one seller does not show up, another buyer or seller always stands ready to take her place. There are no implications for asset prices since the other buyer is assumed to have the same valuation for the asset in expectation. Asset characteristics, "hedonics", determine asset value, not the nature of the investor base. There are many reasons to believe that real world asset markets violate these assumptions, especially for private and real assets. This paper focuses on commercial real estate (CRE) assets, a large segment of the private and real asset space, and one for which we have excellent data on the identity of the buyers and sellers, as well as a diverse ecosystem with buyers and sellers of different types and sizes.

To motivate the value of considering investor composition, consider two CRE examples. A suburban shopping mall has seen its anchor tenant Sears go dark, triggering a major decline in foot traffic and in

[1]McKinsey Global Private Markets Review 2024. Private asset markets include buyout, venture capital, growth, and other private equity, private debt, real estate, and infrastructure and natural resources, and assets under management are measured as of the end of June 2023. AUM in private markets is predicted to double by 2029.

[2]This includes private equity, hedge funds, real estate, commodities, and miscellaneous alternative.

[3]Private credit refers to lending outside of the traditional banking system, in which lenders work directly with borrowers to negotiate and originate privately held loans that are not traded in public markets. Strategies include direct lending, subordinated debt (mezzanine, second lien debt and preferred equity), distressed debt, and special situations. The private credit market has grown from $1 trillion to $1.5 trillion between 2020 and 2024. AUM is predicted to double to $2.8 trillion by 2028 (Morgan Stanley, Understanding Private Credit, June 2024).

rents. This mall has been owned by a publicly-listed REIT that specializes in malls and is experiencing similar issues in other parts of its portfolio. This REIT is at a competitive disadvantage in turning the mall into an Amazon warehouse since its expertise is in owning and operating malls. Also, the REIT is unable to raise new equity capital to fund an asset repositioning since its stock price is trading below NAV. An opportunistic real estate private equity (REPE) fund, like Blackstone, which is looking for creative ways to deploy its ample dry powder, with geographic and sector flexibility and generalist managerial talent, may be better suited to reposition the asset. The asset is worth more in the hands of the REPE fund than in the hands of the REIT. The value of the mall is increasing in the amount of REPE capital. A second example considers the importance of large foreign institutional investors, mostly sovereign wealth funds and pension funds, in U.S. CRE markets. Portfolio diversification considerations and expectations of lower returns on bond or stock markets may lead them to increase allocations to alternative asset classes like CRE. Large institutional foreign capital is particularly fond of large, high-quality office and retail properties in "superstar cities," which they consider to be "bond-like" assets. How large an effect do these foreign buyers have on the prices of these properties when they start buying billions of dollars of such assets? What would these prices have looked like in the absence of the foreign investment surge these markets experienced in 2015 and 2018? What are the spillover effects of this foreign purchase activity on other lower-quality office and retail assets in the same city? In other, second-tier cities? These questions have taken on renewed urgency as cross-border investors turned from net buyers into net sellers after 2018.

Our paper is the first to systematically document the composition of buyers and sellers in commercial real estate and its evolution over time. We use a data set from MSCI Real Capital Analytics (henceforth RCA) that covers the near-universe of CRE transactions between the year 2001 and 2023. It contains about 470,000 transactions that combine for $10 trillion in traded value, covering the four traditional CRE sectors of apartment, office, retail, and industrial properties. RCA has spent enormous effort unraveling the true buyers and sellers involved in each transaction.[4] There are about 325,000 unique investors in our sample. Using the history of transactions, we build up the portfolios of these investors, and find that they account for about $8.6 trillion dollars in holdings at the end of 2023. We document new stylized facts on the composition of the investor base: who trades and holds what CRE assets. We organize investors into eight types, but also note important heterogeneity in terms of total portfolio size (and, related, the size of the typical asset traded) and geographic concentration. The resulting picture is one of partial segmentation where smaller investors tend to specialize in smaller assets and fewer geographies, while larger investors tend to be national in scope and trade larger ticket sizes.

[4]This is a difficult task since CRE properties are usually bought using a separate LLc for each building, typically with a non-descriptive name.

We develop a model that takes seriously the decentralized nature of trade of private assets. Each investor's valuation for each asset depends on a conditional mean, which is a function of asset characteristics, and a idiosyncratic stochastic component that could capture liquidity or funding shocks hitting the investor. The conditional mean component reflects that investor's valuation of the characteristics of that asset. Different investors have different valuations for the property's attributes such as asset size, location, cash flow generation, etc. The first main finding of the paper is that allowing for investor-specific feature valuation substantially increases the model's explanatory power for transaction prices. Compared to a model with asset hedonics, geography-specific macro variables, and time fixed effects, adding investor characteristics significantly improves model fit. Specifically, investor characteristics reduce the unexplained variation of the richest hedonic model by around 20%.

We then model the probability of listing a property for sale, meeting a buyer, and transacting with that buyer. The meeting probability lies at the core of the model. We model the likelihood that a buyer meets a building that is for sale as an increasing function of the gains from trade, the size of the buyer, the relative attractiveness of the market in the prior period, and the similarity of the asset to other assets the buyer has in her portfolio, where similarity is in terms of the asset's size, geography, and sector. Conditional on a meeting, the transaction goes through if the value of the buyer exceeds that of the seller. Finally, the listing probability depends on time in a way that allows the model to match the time-varying transaction volume, given all other parameters.

A key insight that makes the estimation of the meeting probability computationally feasible is to use negative sampling, a concept used in the natural language processing (NLP) computer science literature (Mikolov et al., 2013). Rather than assessing how a candidate buyer's meeting probability compares to that of *every* other potential buyer (B potential buyers) for every one of N buildings in every one of T periods, which involves computing $B \times N \times T$ numbers, we only need to sample a relative small number of $K << B$ potential buyers, $K - 1$ of which did not buy the building and one who did. Maximizing the likelihood ratio of the actual buyer relative to the other $K - 1$ non-buyers delivers consistent parameter estimates (Ma and Collins, 2018). The higher K, the more efficient the estimator. We show this procedure works well for $K = 100$, with diminishing efficiency gains for larger values.

Having estimated the model parameters, we turn to the main counterfactuals of interest. An important object in the computation of counterfactuals is the *potential price distribution* of a property that is listed for sale. The seller could have met any one of a large number of buyers besides the one she transacted with. By sampling (with replacement) many potential buyers, in proportion to their probability of meeting the seller, and discarding the ones where there are no gains from trade, we can construct a distribution of prices for which the property could have traded. This potential price distribution is a useful object. It can be used to assess the seller's luck or the skill of her broker (by seeing where in the distribution the realized sales price lies), for setting a future sales price target, or for risk management purposes. To quantify the impact of a group of buyers on the price, we remove that group from the potential matching partners, and recompute the potential price distribution. The difference between the benchmark potential price distribution and the counterfactual price distribution that removes one investor group quantifies the impact of that investor group on prices. Naturally, the average price is a key moment as is the inter-quartile range, which is a measure of price risk.

Our paper contributes to the literature on valuing private assets. Practitioner approaches often ignore risk, simply comparing total cash outflows from operations and asset disposition to the initial amount invested (TVPI) or taking into account the timing of the cash flows and time value of money (IRR). The finance literature has emphasized the importance of recognizing the systematic risk in the cash flows, with risk being measured in increasingly sophisticated ways. The public market equivalent (PME) approach Kaplan and Schoar (2005) considers the alternative investment in a stock market index fund with the same timing of cash flows. The generalized PME approach of Korteweg and Nagel (2016) takes into account that the private asset cash flows may have a risk exposure to the equity market different from one. The strip-by-strip approach of Gupta and Van Nieuwerburgh (2021) takes into account that private asset cash flows may be exposed to a richer menu of risk factors present in public markets, not just the equity index but several cross-sectional equity factors as well as interest rate risk priced in the cross-section of government bonds, and that exposures of operating cash flows may differ from those of disposition cash flows. [5] While Gupta et al. (2022) value CRE assets using a richer model for the cash flow dynamics, like the earlier approaches, they continue to assume that the Stochastic Discount Factor can be inferred from public markets data. This paper deviates from this tradition and does not assume that the market prices of risk for the private market-relevant risk exposures are determined in public securities markets. The valuation model allows for the investor composition in the private asset market to impact valuations.

We also contribute to the literature that studies asset prices using asset demand systems (Koijen and Yogo, 2019; Koijen et al., 2024). We share the same objective to jointly understand valuations, asset characteristics, portfolio holdings, and trades. This is the first paper to develop a demand system for private assets. Since these assets have fundamentally different features than public assets in that they trade in bilateral markets, the approach we pursue is very different. Rather than modeling portfolio shares for divisible shares of a public stock, we model the probability that an entire private asset changes hand in a meeting between one buyer and one seller. We provide a micro foundation of the valuation equation reflecting this unique feature of

[5] Other important contributions to the literature of measuring risk and return private equity performance literature include (Korteweg and Sørensen, 2010; Driessen et al., 2012; Ang et al., 2018).

private assets. Our approach extends to other private assets or assets that trade in over-the-counter markets: private credit, private equity, infrastructure and natural resource assets, collectibles (art), et cetera.

The standard hedonic regression approach for valuing private assets, which posits that the asset's value is a function of its physical characteristics each with a fixed price attached to them, has a long tradition in economics (Lancaster, 1966; Griliches, 1971; Rosen, 1974; Witte et al., 1979; Wallace, 1996). Our model adds the *composition of the investor base* drives the mapping from characteristics to transaction prices. Indeed, the transaction price reflect the average valuation of the buyer's and seller's valuation, and thus their valuation of the characteristics. Since that composition changes over time and in the cross-section, the standard hedonic model's residual variance would change over time and in the cross-section. If hedonic coefficients on the building characteristics are allowed to depend on time, then that time-variation may pick up valuation changes that are in fact due to a changing investor mix. For example, the characteristic "number of floors" may appear to have gained in importance (price) in the estimation of a standard hedonic model for Manhattan office. But instead, it could pick up the increased demand from foreign investors combined with their relative preference for tall office buildings. Controlling for investor mix can "cleanse" the hedonic price index from the impact of (time-varying) investor composition.

A large REIT literature uses standard empirical asset pricing tools to explain returns of publicly listed CRE companies (Van Nieuwerburgh, 2019, for a recent example). However, REITS represent only 12% of transaction value in our sample. As a large and growing share of CRE transactions moves from public to private markets, the task of understanding risk and return in private CRE becomes both more difficult and more important.[6] Peng (2016), Sagi (2021), and Giacoletti (2021) study risk and return at the individual property level and emphasize liquidity considerations. Our potential price distribution's IQR provides a complementary liquidity risk measure. Plazzi et al. (2008, 2010) study cross-sectional variation in expected returns and expected cash flow growth in CRE markets across U.S. cities.

Recent work has begun to analyze the importance of investor heterogeneity for CRE valuation. Ghent (2021) shows that institutional investors like REPE funds concentrate their CRE investments in markets with many other similar institutions. The liquidity generated by a large pool of buyers results in higher valuations for assets with otherwise identical cash-flows. Similarly,Cvijanović et al. (2022) finds that institutional investors in CRE are highly sensitive to liquidity considerations. Badarinza et al. (2022) emphasize the importance of same-nationality buyers and sellers in a CRE transaction as a mechanism to resolve the gravity puzzle. In earlier work, Badarinza and Ramadorai (2018) exploit differences in pre-existing concentration of foreigners in London neighborhoods to identify and estimate the effect of cross-border purchases on home

[6] Van Nieuwerburgh et al. (2015) contains a discussion of the challenges with measuring returns for infrequently traded CRE assets and the various approaches that have been proposed to make progress.

prices.

We contribute a new estimation technique to matching literature. We import insights from the Artifical Intelligence (AI) literature which has estimated large language models using word embeddings. Particularly, using negative sampling (Mikolov et al., 2013; Ma and Collins, 2018) is necessary to overcome the computational challenges when each of many assets could trade each period and be purchased by one of many buyers. Fox (2018) proposes a maximum score estimator to overcome the curse of dimensionality in many-to-many matching settings. Gabaix et al. (2024) uses techniques from the LLM literature to define similar stocks using investor portfolio data. Badarinza et al. (2024) estimates the matching function in the housing market, using unique data on online search for properties and broker meeting requests. We do not directly observe meetings and must estimate the meeting function.

The rest of the paper is organized as follows. Section 2 sets up the model, which contains a valuation and a matching model. Section 3 details the estimation procedure for the model. Section 4 describes the data and provides summary statistics. Section 5 shows results for the parameter estimates, and for the main counter-factuals. Section 6 concludes. The appendix contains additional details on the data and empirical methods.

2 Model

2.1 Notation

There are four building types, which we call sectors: office, retail, industrial, apartments. We estimate separate models for each sector. Time is indexed by $t, t = 1, \ldots, T$. Buildings are indexed by n, $n = 1, \ldots, N$, and investors by $i, i = 1, \ldots, I$. If a building is sold, then we refer to the seller as investor s and the buyer as investor b. The transaction price is $P_t(n)$.

Buildings are differentiated by a vector of characteristics, x_{nt}, where we assume the first element is one. We use several asset characteristics. The first seven are measured at the property level in the RCA data: log asset size (measured in terms of square feet for office, retail, and industrial or number of apartment units for multifamily), an indicator variable indicating whether the building is located in a superstar city, an indicator variable indicating whether the property is located in the central business district (CBD), the number of floors, log building age, the log renovation-adjusted age, and a property sub-type dummy.

Next, we consider five local macro-economic variables, first-four of which vary at the market-level. We define 60 markets.[7] We fold in the property characteristics x_{nt} the following market variables: the log

[7] Appendix A.2 shows our mapping from the RCA markets to our 60 markets.

population (for apartments, from the BEA) or log employment (for the other sectors, BEA) of the market, log real personal income per capita (BEA), the market occupancy rate (NCREIF), and the market NOI growth rate (NCREIF).[8]

The next asset characteristic is the net effective rent (NER) per square foot or the net operating income (NOI) per apartment unit in the block (small geographic area) where the property is located. For simplicity, we will refer to the NOI for multifamily as the NER, so we can use the same variable name for all four sectors. Appendix A.3 describes the details.

All market variable series vary over time. The first one is a measure of market size, the second one an indicator of the average income level in the market, the next two are key metrics of the overall health of the specific CRE sector in that market. The last one, the NER, is arguably the most important one, and captures the local neighborhood conditions.

The final two characteristics that enter into x_{nt} are indicator variable for whether the asset belongs to a distressed sale and an indicator variable for whether it belongs to an entity sale.

Investor heterogeneity is captured by a vector of four characteristics z_{it}. It includes information about the investor's type (e.g., REIT, PE fund, foreign investor), the log dollar value of the investor portfolio across all sectors, the share of the overall portfolio in superstar markets, whether the investor transacted as a part of a Joint Venture (JV), and a measure of the relative bargaining power of the buyer versus the seller, measured as the log ratio of the buyer's portfolio size to the seller's portfolio size.

2.2 Investors' Private Valuations

We first provide the main empirical specification of an investor's private valuation in Section 2.2.1, for which we provide a micro foundation in Section 2.2.2.

2.2.1 A characteristics-based model of investors' private valuations

We assume that investor i's valuation of building n is given by

$$v_{it}(n) \equiv \ln V_{it}(n) = h\left(z_{it}, x_{nt}; \gamma_t\right) + \epsilon_{it}(n),$$

where γ_t is a vector of parameters that may vary over time. For conciseness, we use will at times use the short-hand notation $h_{it}(n) = h\left(z_{it}, x_{nt}; \gamma_t\right)$. The error term, $\epsilon_{it}(n)$, captures liquidity or funding constraints, unobserved quality characteristics or differences in investors' beliefs. We assume that the valuation shocks

[8] The National Council for Real Estate Investment Fiduciaries (NCREIF) is a consortium of large institutional real estate investors that share data on their institutional-quality portfolios allowing for the construction of market-level cash flow and occupancy statistics. We thank NCREIF for making these data available to us. We build a bridge between the NCREIF markets and our 60 markets. Appendix A.2 provides the details.

are uncorrelated across investors and buildings, conditional on all characteristics that include investor-type and geography fixed effects, and normally distributed, $\epsilon_{it}(n) \sim N\left(0, \sigma_t^2\right)$.

In our empirical specification, we parameterize the valuation model as

$$h_{it}(n) = \beta_{x,i}' x_{n,t} + \gamma_t, \tag{1}$$

where γ_t is a time fixed effect that captures changing macroeconomic conditions or time-varying investor preferences or wealth and recall that $x_{n,t}$ includes a one as its first element. We model the heterogeneity in coefficients as

$$\beta_{x,i} = \beta_x z_{i,t}, \tag{2}$$

where $\beta_x \in \mathbb{R}^{\dim(x)\times\dim(z)}$. $z_{i,t}$ includes a one as a first element, investor-type fixed effects, and several more investor characteristics mentioned above, such as portfolio size.

2.2.2 A micro foundation of the private valuation model

We consider a two-period model, $t = 0, 1$. In period $t = 0$, the investor considers buying a building. In period $t = 1$, the investor receives the net cash flow and the resale value of the building. In mapping the model to the data, we think of $t = 1$ as the payoff of the building over a longer holding period.[9]

The building in question may be part of a broader property portfolio or, in case of a pension fund or insurer, part of a broader portfolio that may include equities and fixed income assets. Without the new building, the broader portfolio generates a payoff D_{1i}. If the investor has no other assets, $D_{1i} = 0$.

At $t = 0$, the investor has access to cash C_{0i} that can be used to purchase the building. Any remaining cash earn a gross risk-free rate that is normalized to one. This implies that, without purchasing the new building, the investor's wealth in $t = 1$ is $A_{1i} = D_{1i} + C_{0i}$.

The new building generates a time $t = 1$ payoff of N_{1i}. We assume that $(D_{1i}, N_{1i}) \sim N(\mu_i, \Sigma_i)$. We denote elements of μ_i and Σ_i by their logical counterparts, e.g., μ_{Ni} and Σ_{NDi}. We allow investors to have heterogeneous beliefs about future payoffs and index all moments by i. If the investor adds the building to her portfolio, period $t = 1$ wealth equals $A_{1i}^P = D_{1i} + C_{0i} - P_0 + N_{1i}$, where P_0 is the purchase price of the property.[10] To complete the model, we assume that the investor has mean-variance preferences over

[9] At the expense of some additional notation, this micro foundation can be extended to a multi-period model.

[10] We assume throughout that the investor has deep pockets and $C_{0i} > P_0$. The model can easily be extended so that the investor finances a fraction ϕ of the property with a mortgage at a rate $R > 1$, or the investor's valuation is bounded by C_{0i}, which caps the valuation at C_{0i}.

terminal wealth:

$$\mathbb{E}_i[A_{1i}] - \gamma_i \mathrm{Var}_i(A_{1i}),$$

where γ_i is risk aversion.

As the building is a discrete purchase (compared to buying shares in a company), we compute the investor's valuation as the $t = 0$ price $P_0 = V_{0i}$ that makes the investor indifferent between purchasing the building or not. This valuation then solves the following equation:[11]

$$\mathbb{E}_i[D_{1i} + C_{0i}] - \gamma_i \mathrm{Var}_i(D_{1i}) = \mathbb{E}_i[D_{1i} + C_{0i} - V_{0i} + N_{1i}] - \gamma_i \mathrm{Var}_i(D_{1i} + N_{1i}),$$

where the left-hand side is the investor's utility if she does not buy the building and the right-hand side is the utility when she does buy the building. We use this equation to solve for the investor's private valuation:

$$V_{0i} = \mathbb{E}_i[N_{1i}] - \gamma_i \mathrm{Var}_i(N_{1i}) - 2\gamma_i \mathrm{Cov}_i(D_{1i}, N_{1i}), \tag{3}$$

which depends on the expected payoff, $\mathbb{E}_i[N_{1i}]$, a discount for its variance, $\mathrm{Var}_i(N_{1i})$, and a further discount or premium that depends on the property's covariance with other assets in the investor's portfolio, $\mathrm{Cov}_i(D_{1i}, N_{1i})$.

The main valuation equation is (3). To obtain a characteristics-based model of investors' private valuations, we follow Koijen and Yogo (2019) and model the moments as functions of characteristics with investor-specific coefficients that reflect differences in beliefs:

$$\begin{aligned}
\mathbb{E}_i[N_{1i}] &= \beta_{i0}' x_n, \\
\gamma_i \mathrm{Var}_i(N_{1i}) &= \beta_{i1}' x_n, \\
\gamma_i \mathrm{Cov}_i(D_{1i}, N_{1i}) &= \beta_{2i}' x_n,
\end{aligned}$$

where we note that risk aversion and (certainly) D_{1i} are heterogeneous across investors. We model the heterogeneity across investors as a function on the size of their portfolio, location, institutional type, etc.:

$$\beta_{ki} = \beta_k' z_i,$$

for $k = 0, 1, 2$. This yields the valuation equation in (1)-(2).[12]

[11] To focus on the key economics, we abstract from transaction costs here.

[12] The difference between a valuation in levels versus logs arises because we assume mean-variance preferences. Alternatively, we can assume CRRA preferences and log-linearize the model.

2.3 Transaction Prices

There is a proportional cost of $cP_t(n)$ if investor i and j transact. The bargaining power of the seller is measured by $\beta \in [0,1]$, which is used to split the proportional transaction cost and in setting the price. Investor b is willing to buy the building when

$$V_{bt}(n) > P_t(n)(1+\beta c),$$

or, equivalently,

$$v_{bt}(n) - \chi_b > p_t(n),$$

where $\chi_b \equiv \ln(1+\beta c) \simeq \beta c$. Analogously, investor s is willing to sell the building when

$$V_{st}(n) < P_t(n) - (1-\beta)cP_t(n),$$

or, equivalently,

$$v_{st}(n) - \chi_s < p_t(n)$$

where $\chi_s \equiv \ln(1-(1-\beta)c) \simeq -(1-\beta)c$.

We assume that the price is set by Nash bargaining (in logs),[13] implying

$$p_t(n) = \beta v_{bt}(n) + (1-\beta)v_{st}(n) - c.$$

This implies that trade takes place when

$$v_{bt}(n) - v_{st}(n) > c.$$

In our empirical analysis, we set $c = 0$ and $\beta = 0.5$.

2.4 Transaction Probabilities

We now discuss the model of the listing probability, the meeting probability, and the probability of transacting conditional on a meeting taking place between a buyer and a seller. At this stage, we assume that we know the parameters of the valuation model. We therefore have the component of valuations explained by

[13] Formally, the price is the solution to

$$p^* = \arg\max\left((p-c) - v_s\right)^{\beta}\left(v_b - (p+c)\right)^{1-\beta}.$$

characteristics, h_b and h_s, for the buyer and seller.

We model the probability of transaction as:

$$\pi(b,s) = \pi_{l,t}\pi_m(b,s)\pi_\tau(b,s),$$

where the first term, $\pi_{l,t}$, is the probability of listing, π_m the probability of meeting, and π_τ the probability that the valuation of the buyer exceeds the valuation of the seller.

We initially assume that the listing probability is only a function of time, $\pi_{l,t}$. More advanced models can model the listing decision as a function of the price dynamics in a certain market, for example to capture the importance of reference points. In this case, the listing probability would increase in last year's price in that market relative to a running maximum of the price index up to that point. Indeed, if prices declined sharply last year, owners of buildings are less likely to list.

The second term is the meeting probability, which is a reduced-form approach to modeling directed search. We assume that the meeting probability depends on the gains from trade based on observable characteristics, $\Delta^h_{b,s} \equiv h_b - h_s$,

$$\pi_m(b,s) = \begin{cases} 0 & \text{if } b = s \\ \frac{\exp\left(\lambda_1 \Delta^h_{b,s} + \lambda_2 A_b + \lambda_3' \delta_{b,s} + \lambda_4 N_b\right)}{\sum_{c \neq s} \exp\left(\lambda_1 \Delta^h_{c,s} + \lambda_2 A_c + \lambda_3' \delta_{c,s} + \lambda_4 N_c\right)} & \text{otherwise,} \end{cases} \tag{4}$$

where $\lambda_i \geq 0$, $i = 1, \ldots, 4$. When $\lambda_i = 0$, $i = 1, \ldots, 4$, the matching probability is the inverse of the number of buyers and we have random matching. When $\lambda_1 > 0$, matches with larger gains from trade are more likely to happen. λ_2 captures the role of investors' propensity to trade as measured by the average number of transactions in the past five years, A_b. Active buyers with a higher transaction history are more likely to match when $\lambda_2 > 0$, reflecting the fact that some investors trade significantly more assets than others.

Next, $\delta_{b,s}$ is used to capture an investor's consideration set by measuring the similarity between the characteristics of seller s's building and the portfolio characteristics of buyer b. It has three dimensions: Price alignment, i.e., the similarity between the building's log price and the average log asset size in the buyer's portfolio, Geography (x_s is the building's market and x_b is share of the investor's portfolio in that market), and Sector (x_b is the share the buyer's portfolio in the sector of building s). Hence, λ_3 is a 3×1 vector. For a given characteristic x_s of the building and x_b, the average characteristic of the buyer's portfolio, similarity is derived from the distance $|x_b - x_s|$ by converting smaller distances into higher similarities. [14] Lastly, N_c is an indicator variable that equals one when the buyer owns more than two buildings and zero

[14] Specifically, we use 1 - $|x_b - x_s|$ for geography and sector characteristics and $\frac{1}{1+|x_b - x_s|}$ for price alignment.

otherwise.

The third term measures, conditional on a meeting, whether a transaction takes place, $\pi_\tau(b,s) = P(V_s < V_b)$. Under the assumptions made before, we have

$$
\begin{aligned}
P(V_s < V_b) &= P(\epsilon_s - \epsilon_b < h_b - h_s) \\
&= \Phi\left(\frac{h_b - h_s}{\sqrt{2}\sigma}\right).
\end{aligned} \tag{5}
$$

as $\epsilon_b, \epsilon_s \sim N(0,\sigma^2)$.

The probability that a transaction does not happen is given by

$$\pi_n(s) = (1-\pi_l) + \pi_l \sum_b \pi_m(b,s)(1-\pi_\tau(b,s)). \tag{6}$$

The first term, $1-\pi_l$, is the probability that the seller does not list the building. The second term is the probability that the building lists, and thus a match will be formed, but the valuation of the buyer does not exceed the valuation of the seller, as captured by $1-\pi_\tau(b,s)$. As $\sum_b \pi_m(b,s) = 1$, $\sum_b \pi_m(b,s)(1-\pi_\tau(b,s)) \leq 1$ and $\pi_n(b,s) \leq 1$.

Taken together, it holds

$$\pi_{l,t} \sum_b \pi_m(b,s)\pi_\tau(b,s) + \pi_n(b,s) = 1.$$

3 Estimation Procedure

The model is estimated recursively in two steps. In Section 3.1, we discuss the estimation of the valuation equation. We discuss the estimation of the listing and matching model in Section 3.2.

3.1 Valuation Model

We begin by estimating the coefficients of the valuation equation. We estimate the valuation model separately for each sector (A, O, I, R), i.e., we allow for the intercepts and characteristic betas to depend on sector in addition to having an investor-type i specific intercept. We rely on observable transaction prices to estimate the other parameters. The log price is given by

$$
\begin{aligned}
p_t(n) &= \frac{1}{2}(h_{bt}(n) + h_{st}(n)) + \frac{1}{2}(\epsilon_{bt}(n) + \epsilon_{st}(n)) \\
&= \frac{1}{2}(\beta_{x,b} + \beta_{x,s})' x_{n,t} + \frac{1}{2}(\epsilon_{bt}(n) + \epsilon_{st}(n)).
\end{aligned} \tag{7}
$$

The number of parameters to be estimated equals $\dim(x)+\dim(x)\times\dim(z)+T$, which is large. We therefore use ridge regression and and choose the regularization hyper-parameter using 10-fold cross-validation.

We note that prices are observed only when $v_b > v_s$, which is when $h_b - h_s > \epsilon_b - \epsilon_s$. The difference in shocks is uninformative about the sum of the shocks because

$$\mathbb{E}\left[\epsilon_b + \epsilon_s \mid \epsilon_b - \epsilon_s\right] = 0,$$

when $\epsilon_b, \epsilon_s \sim N(0, \sigma^2)$. This implies that estimating the coefficients of the valuation model based solely on the observed transactions does not create a bias.

3.2 Listing and Meeting Probabilities

Using the estimates of $(\alpha, \beta_x', \beta_z', \gamma_t)'$ from the valuation equation, we obtain estimates of $h_{bt}(n)$ and $h_{st}(n)$. In principle, we can then estimate the listing and meeting probabilities using maximum likelihood. The likelihood contribution of the building of seller s is

$$\ell_s = \sum_{b=1}^{B} y_{b,s} \ln \pi(b,s) + (1 - \sum_{b=1}^{B} y_{b,s}) \ln \pi_n(s),$$

where $y_{b,s}$ equals one when a transaction take place and zero otherwise.

Computing this likelihood is a high-dimensional problem due to the large number of buildings (and thus sellers) and buyers. We instead estimate the model using negative sampling, inspired by applications in the natural language processing (NLP) literature (Mikolov et al., 2013).

Intuitively, instead of using the full likelihood, we consider a classification problem. For each transaction, we sample another K potential buyers that did not purchase the building of seller s, $c \in \mathcal{N}_s$, $|\mathcal{N}_s| = K$. We then compute the probability that buyer b purchases the building, conditional on one of the $K+1$ potential buyers purchasing the property,

$$\pi_r(b,s) = \frac{\xi_{b,s}}{\xi_{b,s} + \sum_{k\in\mathcal{N}_s} \xi_{k,s}}, \tag{8}$$

where $\xi_{b,s} = \exp\left(\lambda_1 \Delta_{b,s}^h + \lambda_2 A_b + \lambda_3' \delta_{b,s} + \lambda_4 N_b\right) \pi_\tau(b,s)$.[15]

We then optimize the loss function over the observed transactions (i.e., (b,s) pairs): $-\sum_s \ln \pi_r(b,s)$. This estimator is the ranking estimator in Ma and Collins (2018), who also develop the asymptotic theory. In particular, they show that the estimator is consistent for $K \geq 1$, asymptotically normal, and it converges to the maximum-likelihood estimator as $K \to \infty$.

[15]The common denominator, $\sum_{c\neq s} \exp\left(\lambda \Delta_{c,s}^h + \lambda_2 A_c + \lambda_3' \delta_{c,s} + \lambda_4 N_c\right)$, drops out, simplifying the loss function.

Lastly, we estimate the listing probability, $\pi_{l,t}$. The total number of buildings that transact in a given year is denoted by T_t, which we match using the expected number of transactions in a year:

$$\sum_s \pi_{l,t} \sum_b \pi_m(b,s)\pi_\tau(b,s) = T_t, \tag{9}$$

which implies

$$\pi_{l,t} = \frac{T_t}{\sum_s \sum_b \pi_m(b,s)\pi_\tau(b,s)}. \tag{10}$$

3.3 Potential Price Distribution and Counterfactuals

We are interested in computing the distribution of prices $f(p)$ at which trade could have occurred for an asset listed for sale. We are particularly interested in the expected price, $\mathbb{E}[p]$, and its inter-quartile range, $IQR(p)$. Once the parameters of the valuation and matching models have been estimated, we implement the following algorithm to determine the potential price distribution.

For each asset n that we see transact at time t, we know the transaction price ($p_t(n)$), the characteristics of the asset $x_{n,t}$, and the valuation of the buyer ($h_{b,t}$) and the seller ($h_{b,t}$). This allows us to back out $\epsilon_{p,t} = \frac{1}{2}(\epsilon_{s,t} + \epsilon_{b,t})$. We use Bayes rule to estimate the most likely $\epsilon_{s,t}$: $\widehat{\epsilon}_{s,t} = \mathbb{E}[\epsilon_{s,t} \mid \epsilon_{p,t}]$. This gives us $v_{s,t} = h_{s,t} + \widehat{\epsilon}_{s,t}$. Next we compute the probabilities that this seller meets any possible buyer b', including the actual buyer who materialized: $\{\pi_m(b',s)\}_{b=1}^B$ from (4). We take C random draws with replacement from the distribution of candidate buyers candidate buyers with sampling weights given by the distribution $\{\pi_m(b',s)\}_{b'=1}^B$. For each potential buyer that is drawn, we draw a latent demand shock $\epsilon_{b',t} \sim N(0, \sigma_t^2)$. We construct that candidate buyer's valuation $v_{b',t} = h_{b',t} + \epsilon_{b',t}$. For each candidate buyer, we check that $v_{b',t} > v_{s,t}$. If this condition is satisfied, we record the price as $p_t(n) = \frac{1}{2}(v_{b',t} + v_{s,t})$. If the valuation of the buyer is too low, $v_{b',t} < v_{s,t}$, we set the price to missing. We then compute the mean $\mathbb{E}[p]$ and inter-quartile range $IQR(p)$ of the $C \times 1$-dimensional price distribution across the non-missing values. We refer to (the non-missing part of) this distribution as the potential price distribution.

The potential price distribution reflects the alternative buyers the seller could have met and the prices that would have resulted from those meetings, if indeed the deal was consummated. Seeing where in the potential price distribution the actual transaction price with the actual buyer b falls, provides a gauge of how lucky the seller and/or how skilled her broker was. The potential price distribution is a useful tool in private markets with bilateral price formation for price setting strategy and risk management.

The algorithm is also how we compute counterfactuals. To understand the importance of a subset G of investors, we compute the potential price distribution removing that group from the buyer pool. The sampling of C potential buyers is now from the set $B\backslash G$. We can compute the potential price distribution and its moments for every subset of buyers that makes up the universe B. Comparing means, we see the effect of each group of buyers on average prices, and comparing IQRs provides a metric of how the buyer pool affects the price risk of an asset.

4 Data

Our data consists of the universe of commercial real estate transactions in the four main property sectors (apartments, office, retail, industrial) between 2001 and 2023, about 370,000 transactions representing $10 trillion in transaction value. The data are from MSCI Real Capital Analytics (RCA). Its unique feature is that each transaction contains the name of the buyer(s) and seller(s), as well as an investor type classification.[16] Appendix A.1 shows our data cleaning algorithm that leads to the final data sample used in the estimation.

Figure 1 shows the number of transactions (left) and the dollar volume of transactions (right) in our final sample. All dollar values in the paper are expressed in 2023 real dollars. Henceforth, we use the word volume to describe the dollar value of transactions. The graph shows three boom periods, from 2001 until 2007, from 2009 until 2019, and in 2021-22. It also shows three busts, the Great Financial Crisis in 2008-09, the Covid-19 recession in 2020, and the interest rate hiking cycle in 2022-23. Appendix Table A.2 shows transaction volume by year in table format. There are additional transaction dynamics that are sector specific, such as the decline in retail due to the rise in e-commerce and the decline in office due to the rise in working-from-home in the post-2020 period (Gupta et al., 2022). Figures A.1 and A.2 show transaction volume by sector and property subtype, respectively.

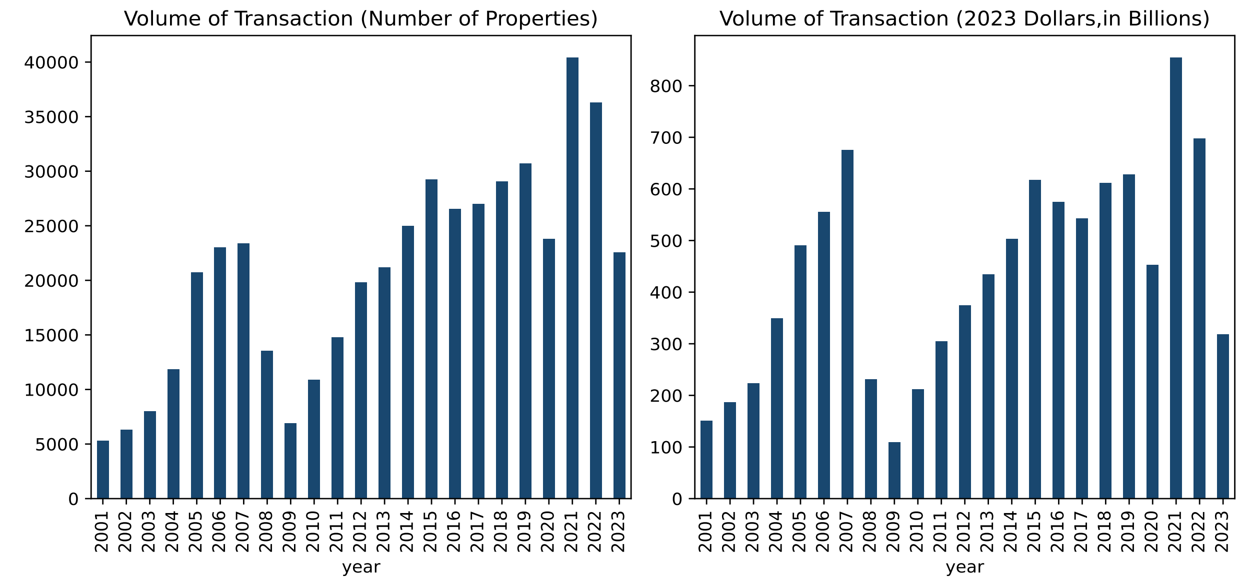
Property transactions are about evenly split between six size categories: >$250M, $100-250M, $50-100M, $25-50M, $10-25M, < $10M. Table A.3 shows the breakdown. Most investors typically stay within one of these size buckets.

Table A.4 reports transaction volume for the 16 largest markets, as well as all other 44 markets combined. Not surprisingly, Manhattan is the largest property market in the sample, followed by Los Angeles. There are large differences in the importance of each sector across markets. Office accounts for over 60% of volume in Manhattan, Washington DC, and San Francisco, but only for 34% of volume in Los Angeles and 24% in Dallas. About half of total transaction volume (and 40% of the number of transactions) is concentrated in the largest 16 markets.

About 75% of transactions accounting for 70% of volume are conventional sales, where one property changes hands. About 19% of transactions and 21% of volume are portfolio sales, where a group of properties

[16] Most property transactions take place using a property-specific LLC. The RCA team has spent years unraveling who is behind each LLC, often cross-checking each transaction using multiple sources.

Figure 1: Commercial Real Estate Transactions



Notes: The figure shows the aggregate number of transactions across all property sectors, i.e. Apartments, Office, Industrial and Retail, by year between 2001 and 2023. The left panel shows the transaction volume in terms of the number of properties traded while the right panel shows the total dollar value of properties traded, expressed in billions of 2023 U.S. dollars.

is sold together. Table A.5 shows that the remainder of the transactions are entity transactions, where an entire corporate entity changes hand, or distressed sales (foreclosures or debtor/trustee sales). In the case of portfolio and entity sales, RCA imputes the price of each property that is part of the transaction. We further improve on this imputation as explained in Appendix A.1.

Next, we turn to investor characteristics. Table A.6 shows how we map RCA's investor types in a smaller group of eight investor types. Our breakdown creates meaningful heterogeneity in terms of constraints, investment objectives, geographic reach, etc. Table 1 shows the transaction breakdown by investor type. About 600 Real Estate Private Equity funds (REPE) account for about 10% of volume. About 3,400 Institutional investors, made up of investment managers, pension funds, insurance, banks and finance companies, and open-ended funds, make up about 15% of volume. The two largest investment categories are local and national Owner-Operator-Developers (OOD_L and OOD_N, respectively), accounting for about 15% and 30% of volume, respectively. These are typically smaller, more geographically-focused, private firms, that are fully dedicated to commercial real estate. They tend to have substantial discretion in terms of investments, but subject to tighter financial constraints. We define local OODs as those whom we only see transact in one market, and national OODs as those we see transact in more than one market. Individuals are made up of high-net worth buyers as well as non-traded REITs who serve high-net worth households, and account for about 4% of volume. REITS are publicly-listed commercial real estate vehicles and account for 12% of volume. We include real estate operating companies and listed funds in this category as well. As listed entities, REITs are subject to much stricter oversight from the regulatory and financial analyst communities and face difficulty raising additional equity when their share price trades below net asset value. Foreign investors are made up of sovereign wealth funds and foreign OODs. They represent about 8% of volume.[17] Users consist of corporations, governments, non-profits, educational and religious institutions charities that own their own real estate. They account for about 5% of transactions. For about 1-2% of volume we do not know the investor type, and create an Unknown category.

Table A.7 shows that 87% of transactions by count and 73% by volume involve a single buyer and a single seller. A further 10% by transactions and 20% by volume involve one buyer and two sellers or two buyers and one seller. The remainder has multiple buyers and multiple sellers. In case of joint ventures (JVs), we observe up to four buyers and up to four sellers. Table A.8 provides additional detail on the JVs, in terms of which investor types tend to collaborate with each other.

Table 1: Transactions By Investor Type

	Buyer (#Trans)	Buyer ($ Vol)	Buyer (% $ Vol)	Seller (#Trans)	Seller ($ Vol)	Seller (% $ Vol)	Unique Investors
REPE	28,853	1241	12.30	22,058	1031	10.22	596
Institutional	38,066	1479	14.66	38,148	1371	13.59	3,435
OOD_L	147,030	1316	13.05	161,967	1656	16.41	238,140
OOD_N	150,678	3083	30.56	131,966	3081	30.54	25,385
Individual	19,453	406	4.02	19,399	285	2.82	15,811
REITS	33,518	1182	11.72	35,135	1254	12.43	389
Foreign	17,606	844	8.37	13,055	616	6.11	2,782
User	28,771	418	4.14	33,663	554	5.49	29,845
Unknown	12,044	119	1.18	20,627	241	2.39	7,802
Total	476,018	10,088	100	476,018	10,088	100	324,185

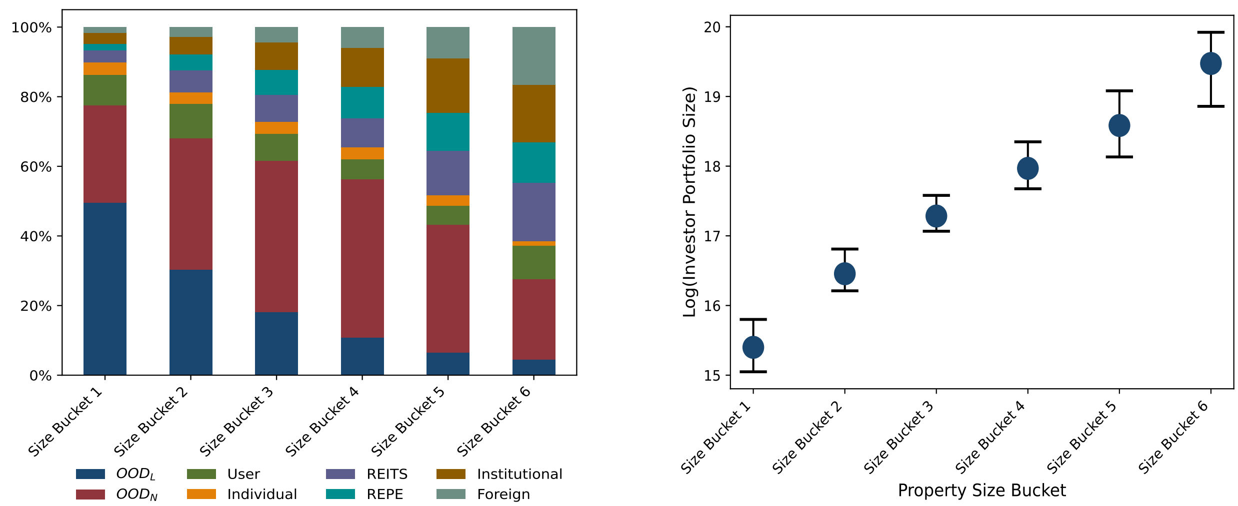
Notes: The top table shows the total value and share of value of properties bought for various investor types, i.e REPE (Real Estate Private Equity), Institutional Investors, OOD_L (Local Developers), OOD_N (National Developers), Individual investors, REITs, Foreign Investors and Users from 2001 to 2023. Bottom table shows the same stats by investor types defined by RCA. The third column in that table corresponds to our definition. Volume is expressed in billions of 2023 U.S. dollars.

Investors are also heterogeneous in terms of their portfolio size, the size of the typical asset they trade, and their geographic concentration. Figure A.3 shows the distribution of investor sizes at the end of 2023. The left panel of figure 2 shows that OOD_L predominantly hold smaller properties. As property size increases, the proportion of OOD_L holdings diminishes, giving way to a more diverse mix of investor types in higher size bucket. In contrast, Foreign investors, Institutions and REITs become increasingly prominent in larger

[17] As Table A.6 shows, there are some foreign entities among other investor types. For example, about 33% of pension fund buys and 21% of pension fund sells are by foreign pension funds. We keep those pension funds in the pension fund (Institutional) category rather than reclassifying them as Foreign. An alternative approach is to create a foreign investor flag for each investor type and eliminate Foreign as an investor category.

size buckets, indicating a preference or ability to invest in larger properties. The right panel illustrates the average log of investor portfolio size by property size bucket, with error bars showing the interquartile range. A clear upward trend in log portfolio size is evident, suggesting that investors associated with larger property buckets tend to have larger portfolio sizes, reflecting economies of scale and capital accessibility for larger investors.

Figure 2: Size Concentration



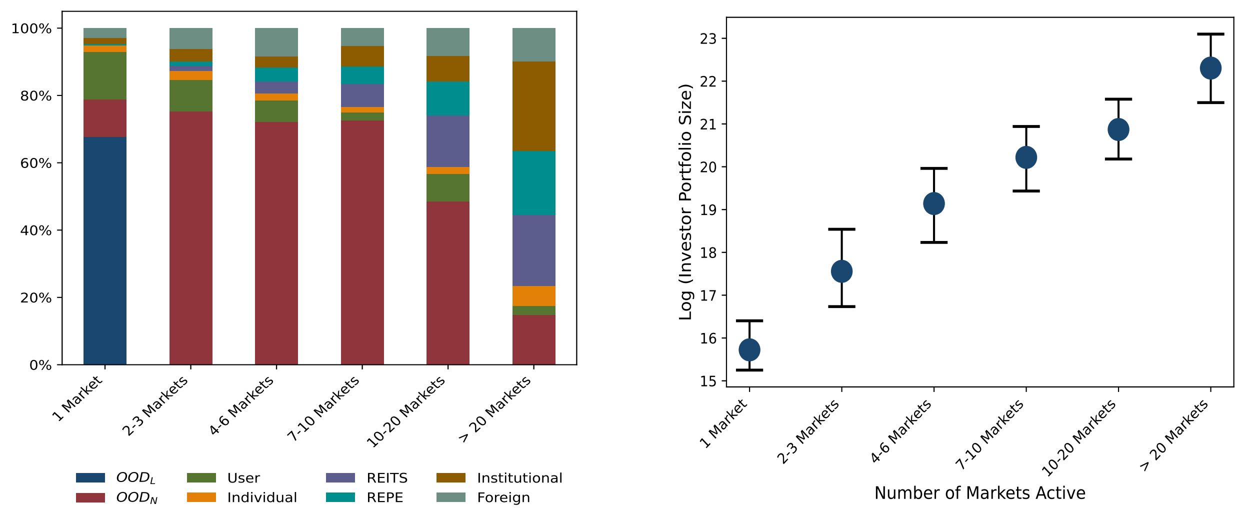
Notes: This figure shows investor concentration by the aforementioned property size buckets at the end of 2023 for investors that own atleast one property[18]. The left figure shows the distribution of investor types by the size buckets while the right figure shows average log investor size (and inter-quartile range) of average log(investor portfolio size) for each property size bucket.

Figure A.3 shows the distribution of investor sizes at the end of 2023. The left panel of figure 2 shows that OOD_L predominantly hold smaller properties. As property size increases, the proportion of OOD_L holdings diminishes, giving way to a more diverse mix of investor types in higher size bucket. In contrast, Foreign investors, Institutions and REITs become increasingly prominent in larger size buckets, indicating a preference or ability to invest in larger properties. The right panel illustrates the average log of investor portfolio size by property size bucket, with error bars showing the interquartile range. A clear upward trend in log portfolio size suggests that investors associated with larger property buckets tend to have larger portfolio sizes, reflecting economies of scale and capital accessibility for larger investors.

The left panel of figure 3 summarizes the distribution of investor types by the number of active markets for the investor. While OOD_L are active in a single market, by definition, OOD_N and Users also tend to be restricted to fewer markets. REPE, REITs and institutional investors, on the other hand are more likely to hold geographically diversified portfolios. The right figure shows that the larger investors are moire

[18] Of the 350,000 investors, 293,000 ever hold only one asset.

Figure 3: Geographic Concentration



Notes: This figure shows investor concentration by geography at the end of 2023 for investors that own atleast one property[17]. The left figure shows the distribution of investor types by the number of aggregated RCA markets they are active in while the right figure shows average log investor size (and inter-quartile range) of average log(investor portfolio size) by the number of markets the investor is active in.

likely to be geographically diversified. Overall, this suggests that there is a substantial amount of market segmentation among the investors in our sample

Finally, we document trading patterns between among types in Table A.9. We see significant trade within and across investor types. Notably, a fourth of all transaction volume involves buyers and sellers of the same type, reflecting segmentation and shared investment strategies among similar investor groups. There is substantial trading activity between local and national OODs. Additionally, REITs and Users exhibit a higher propensity to trade within their own investor type, indicating a degree of segmentation in trading preferences among these groups.

5 Results

5.1 Valuation Model Estimation

The first step of the estimation focuses on the valuation model (7), estimated from transaction prices. We estimate the coefficients $(\beta_{x,b}, \beta_{x,s})$. We impose that investors of the same type have the same valuation for each characteristic, regardless of whether they are buyers or sellers: $\beta_{x,b} = \beta_{x,s}$.[19] The left-hand side variable is the log real price per square foot for sectors O, I, and R and the log price per unit for sector A.

To illustrate the importance of investor characteristics, we proceed in steps. We first estimate the

[19]We can relax this assumption, but it will result in many more characteristics with diminishing incremental gains.

valuation model with only asset characteristics: log asset size, log age, log renovation-adjusted age, subtype indicator, number of floors, CBD indicator,[20] superstar city indicator,[21], entity sale indicator, and distressed sale indicator. In the second step, we add macro-economic variables: log population for A/log employment for O, I, R, log real per capita personal income, NOI growth, and occupancy rate. These macro variables vary by market and year. We also add Net Effective Rent, which varies by neighborhood and year. In the third step, we add year fixed effects. This third model is the benchmark hedonic model, against which we wish to assess the role of investor covariates.

Next, we add investor characteristics. Again, we proceed in steps. In the fourth row, we add investor type fixed effects, leaving out the indicator for the Unknown category: REPE, Institutional, OOD_L, OOD_N, Individual, REIT, Foreign, User. In the fifth row, we add the log portfolio value of the investor, the fraction of the number of properties that the investor has in superstar markets, the joint venture dummy, and the bargaining power variable. These portfolio variables are computed just prior to the transaction for the sellers and just after the transaction for the buyers, so that the transacted asset is always included. In the sixth row, we add the interaction effects between the 8 investor types and the 4 portfolio characteristics with the 7 property and 5 market characteristics. This potentially adds 144 interaction effects.[22] We use Ridge regression to regularize the coefficients, using 10-fold cross-validation. For consistency, every row of the table is estimated with Ridge regression.

The R^2 for these six models are reported in Table 2. We compute the comparative R-squared (R2C) for the alternative models in rows 4, 5, and 6, relative to the null model without investor characteristics in row 3. The R2C is the ratio of the difference in R-squared values between the alternative and null models to the unexplained variation in the dependent variable by the null model:

$$R2C = \frac{R^2_{alt} - R^2_{null}}{1 - R^2_{null}}$$

Kim (2024) develops the asymptotic distribution of R2C and a test of the null hypothesis that R2C equals zero. The stars on the R2C statistic reflect the p-values of this test. The table shows that investor characteristics explain 10.45% (A), 12.84% (I), 17.64% (O), and 13.67% (R) of the variation in property transaction prices, after including hedonics, macro variables, and time trends. These gains amount to between 22.81% and 31.28% of the unexplained variation of the model without investor characteristics. This gain is highly statistically significant in all four sectors. In sum, investor characteristics are a "new hedonic" in commercial

[20]Since the office subtypes are CBD and suburban office, we drop the CBD indicator from the model for office.

[21]Superstar cities are defined as those markets with office prices above $350 per square foot in 2023. The superstar dummy is one for the following markets (for all sectors): Washington DC, DC VA burbs, DC MD burbs, Chicago, Los Angeles, Orange Co, Manhattan, NYC Boroughs, San Jose, San Francisco, East Bay, Seattle, Boston, Austin, Miami/Dade Co, and San Diego.

[22]We do not interact the entity sale and distress sale dummies due to lack of heterogeneity.

real estate valuation that strongly adds to the valuation model's explanatory power.

Table 2: Valuation Model Estimation

Sector	Apartment		Industrial		Office		Retail	
	R^2	R2C	R^2	R2C	R^2	R2C	R^2	R2C
Property Chars	40.61		44.83		32.32		51.27	
+ Macro Vars	52.86		52.38		42.96		55.57	
+ Year Dummies	54.18		53.35		43.6		56.28	
+ Investor Types	56.47	5.00***	55.42	4.44***	46.76	5.60***	58.00	3.93***
+ Portfolio Vars	60.45	13.68***	58.57	11.19***	51.93	14.77***	60.76	10.25***
+ Interaction Vars	64.63	22.81***	66.19	27.52***	61.24	31.28***	69.95	31.27***

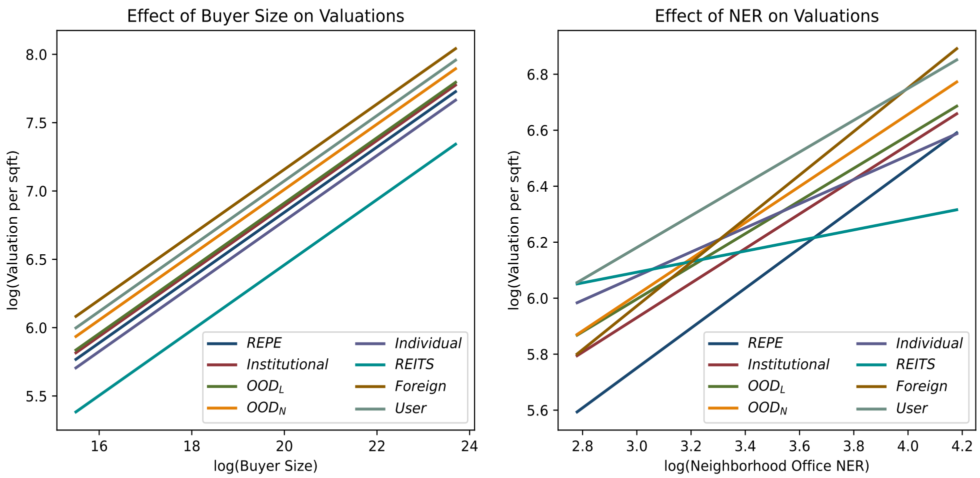
Notes: This table reports the R^2 from six valuation model estimations in equation (7) by Ridge regression. The first model contains 9 property characteristics (row 1). The second model adds five macro variables which differ by market (row 2). The third model adds year fixed effects (row 3). The fourth model adds 8 investor type indicator variables (row 4). The Unknown investor category is omitted. Row 5 adds the investor's portfolio value, the share of the portfolio located in superstar cities, Joint venture dummy, all averaged across the buyer and seller and the bargaining power. Row six allows for interactions between the 12 asset characteristics in rows 1 and 2 and the 12 investor types and characteristics in rows 4 and 5. This creates 144 interaction effects. Each row chooses the optimal Ridge regression penalty parameter based on 10-fold cross validation.

The improvement in the explanatory power of investor features comes to a large extent from the interaction effects (row 6 vs. row 5). These interaction effects indicate that investors who differ by type or size value different property and market characteristics differentially. The traditional hedonics are priced differently across investors. Appendix Figure A.4 highlights the most important features selected by the Ridge model. Many of the top determinants of valuations are interaction effects of investor variables with asset characteristics.

The valuation model also allows us to compute the residual standard deviation for buyers and sellers in each year. We allow for time variation but impose: $\sigma^2_{b,t} = \sigma^2_{s,t}$.

To further illustrate how investor attributes affect valuation, Figure 4 plots the impact of buyer size (left panel) and neighborhood NER (right panel) on valuations. It evaluates equation (7) at the asset characteristics of the average Manhattan office property in 2019, for each investor type. The left panel shows that valuations are increasing in investor portfolio size for all investor types. But even holding buyer size fixed, the same asset trades for substantially different values whether it gets bought by a foreign investor or by a REIT. The right panel shows a similar plot but varies the asset's net effective rent (in neighboring properties). This variable measures the cash-flow generating potential of the asset, which is correlated with the quality of the location. The right panel shows that different investors value locational quality differentially. End users have higher willingness to pay for a given NER than REPE funds. Foreign investors exhibit greater sensitivity to NER than REITS.

Figure 4: Manhattan Office Market 2019: Effect of Investor Size and NER



Notes: This figure illustrates the effect of buyer size and the neighborhood NER on valuations in Manhattan Office Market. The left panel shows how variations in buyer size, across different investor types, influence valuations, while the right panel explores the effect of NER on valuations. For this analysis, building characteristics are held constant at their mean values for a representative Manhattan office.

5.2 Listing and Meeting Model Estimation

With the coefficient estimates for the valuation model in hand, we can compute the private valuation gap between buyers and sellers $\Delta^h_{b,s}$. We also compute the transaction activity index A_b for every buyer using the log of average number of transaction in the last five years, the relative valuation of the market $RV_{b,s}$, the three consideration set variables in $\delta_{b,s}$ and N_b, the greater than two holdings indicator. We then estimate the coefficients $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ that minimize the loss function $-\sum_s \ln \pi_r(b,s)$, with π_r as defined in (8). We calibrate the models for each sector separately on an annual basis, using all transactions from that year. Since most investors in our sample transact infrequently, including all investors in the negative sample would introduce a substantial number of inactive investors, adding noise to the estimation. Therefore, we limit the negative sample to investors who have transacted at least once in the preceding five years. To enhance computational efficiency, we use a smaller negative set size, $K = 100$.[23] The estimated meeting model parameters are presented in Table 3.

The results in Table 3 highlight key drivers of meeting probabilities in the market. We find that a 1% higher valuation gap $\Delta_{b,s}$ between the seller for one potential buyer versus another increases the likelihood of a meeting of the first buyer by $\lambda_1 = 0.257\%$ relative to the second in the apartment sector and 0.346% in the office sector. This underscores the role of directed search in these markets, though the estimates for λ_1

[23] Model parameters stabilize around the average when $K = 100$, and increasing K beyond this does not significantly alter the results.

Table 3: Meeting Model Calibrations

	λ_1	λ_2	$\lambda_{3,1}$	$\lambda_{3,2}$	$\lambda_{3,3}$	λ_4
Apartment	0.257	1.398	5.241	8.407	3.548	3.365
	(0.741)	(0.186)	(0.517)	(0.464)	(0.479)	(0.375)
Industrial	-0.021	1.52	4.721	8.726	3.998	4.127
	(0.494)	(0.175)	(0.44)	(0.597)	(0.246)	(0.29)
Office	0.346	1.297	4.691	8.435	3.897	3.991
	(0.661)	(0.18)	(0.457)	(0.379)	(0.486)	(0.469)
Retail	-0.752	1.94	6.137	9.863	4.57	4.904
	(0.853)	(0.652)	(1.476)	(1.885)	(0.931)	(1.056)

Notes: This table reports the average calibrated values of the meeting model parameters along with their standard errors in bracket. Standard errors are calculated as the standard deviation of calibrated values for each year. Parameters are obtained by minimizing the loss function as described in (8).

exhibit some imprecision.

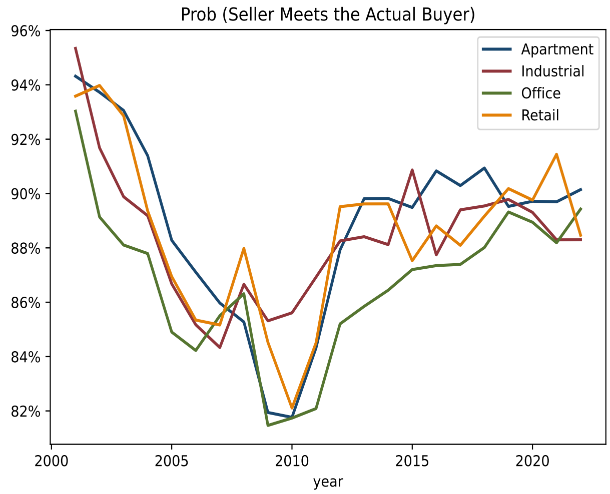
Transaction activity of the buyer, A_b, has a large positive effect on the meeting (and transaction) probability. A potential buyer with 1% more transactions than another has between 1.3% (for office) to 2% (for retail) higher likelihood of meeting the seller, indicating that more active buyers are significantly more likely to meet potential sellers.

The consideration set variables ($\lambda_{3'}$) all have the expected positive sign and are large and significant. Buyers are much more likely to meet assets that have similar size, geography, and sector to the assets they already own. Geographic alignment ($\lambda_{3,2}$) is of particular importance. Potential buyers that are 1% more invested in the same market are about 8% more likely to meet the seller. Price alignment ($\lambda_{3,3}$) also exhibits substantial effects across all sectors, indicating that buyers prefer assets with price characteristics similar to their portfolio. Finally, λ_4 is positive and significant, indicating that owning at least two assets makes a buyer more likely to meet the asset for sale.

Table A.10 successively leaves out each of the variables in the meeting model and performs a likelihood ratio test to better understand the importance of each of the variables. Unsurprisingly, the consideration set variables emerge as the key drivers of the meeting process. Buyers overwhelmingly prefer assets that are similar to their existing portfolios, which aligns with theoretical expectations of search efficiency and risk mitigation. Geography and sectoral alignment contribute the most to the model's fit, suggesting that buyers tend to focus on regions and sectors they are already familiar with. Buyer's transaction activity A_b and owning at least two properties also meaningfully increase the likelihood of the model.

Figure 5 displays the likelihood that the model identifies the true buyer from among the set of 100 potential buyers and the true buyer. That likelihood is consistently high, around 90%, underscoring the model's ability to capture the key determinants of buyer-seller matches and highlighting the model's accuracy in capturing

Figure 5: Meeting Probabilities



Notes: This figure shows the probability of meeting the actual buyer from among a set of $K = 100$ potential buyers (negative set) and the actual buyer, as implied by the benchmark model estimation.

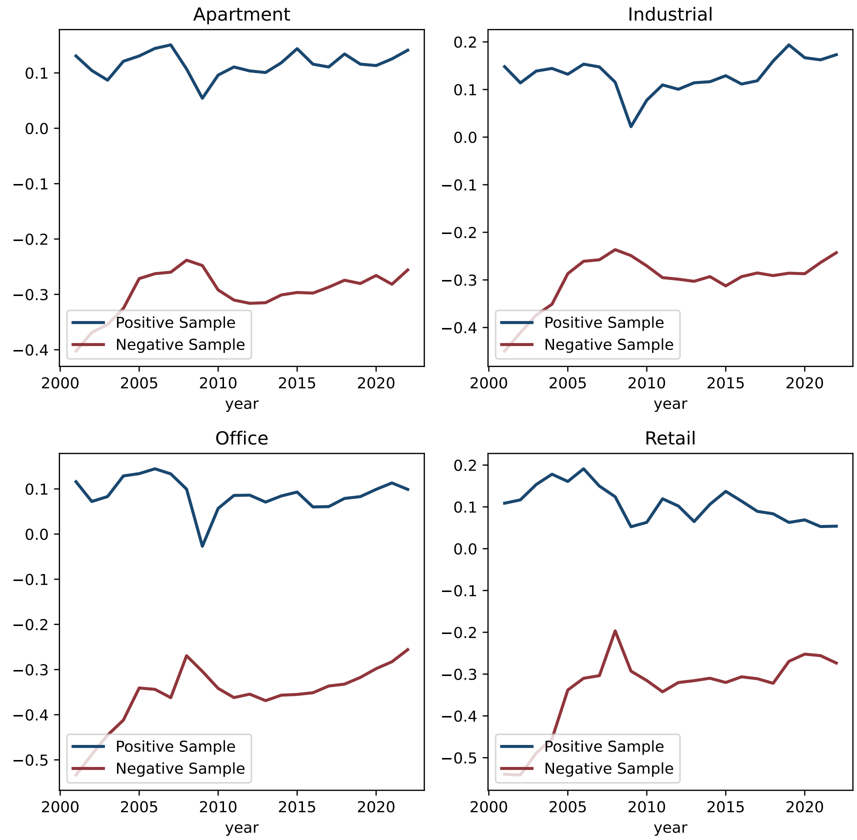
the directed search behavior inherent in CRE markets. The sharp drop in meeting probabilities during the financial crisis is evident across all sectors. This reflects heightened uncertainty and reduced liquidity in the market during the crisis, possibly disrupting the alignment between buyer preferences and available assets.

Figure 6 plots the valuation gap $\Delta_{b,s}$ for the positive set (the actual buyer) and the negative set (the average among the K potential buyers). Across all sectors, the true buyers exhibit consistently higher valuation gaps compared to the average of the potential buyers. This reinforces the model's assumption that matches occur where buyers perceive significant gains from trade. The valuation gap for true buyers typically lies in the range of 0.1 to 0.2 on the log scale, indicating a 10–20% higher perceived value relative to sellers. For the negative sample (the non-buyers), the valuation gap is large and negative, suggesting $\sim 75\%$ lower perceived value.

Figure 7 shows the model-implied probability that the actual buyer transacts with the seller. The left panel highlights that the probability of the seller transacting with the actual buyer is consistently high, reflecting the model's ability to accurately capture directed search and buyer-seller dynamics. The right panel reflects the aggregate probability of any transaction occurring, factoring in the likelihood of listing, meeting, and the valuation gap dynamics captured by π_{tau} in equation 5.

Finally, Figure 8 shows the listing probability, defined in (10). The listing probability reconciles the transaction probabilities with the observed transaction volumes in the data. It implies a boom-bust in listing probabilities in the 2000-2010 period, and a rise in the post-2010 period across all sectors and a dip again in 2020.

Figure 6: Valuation Gap



Notes: This figure shows the valuation gaps, i.e. difference between the log valuation of the buyer and the log valuation of the seller for the positive sample (the actual buyer) and the negative sample (the set of K potential buyers)

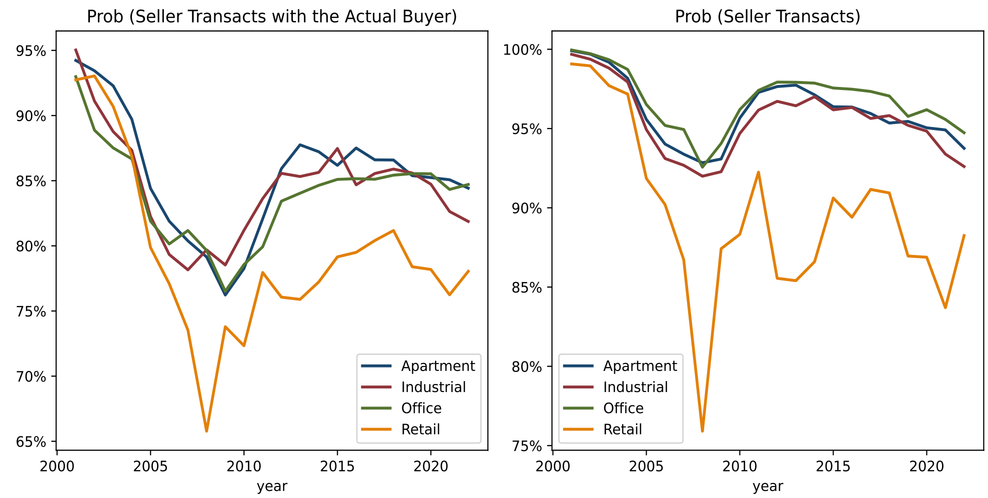
5.3 Price Distribution and Counterfactuals

Having estimated the model, we proceed to the main application, which is to quantify the importance of investor composition on prices. Our main tool for doing so is the potential price distribution explained in Section 3.3. We explore the role of investor types on valuations.

As a concrete application of this toolbox, we ask what would happen if there had been no foreign buyers in the Manhattan office market. As illustrated in Figure A.5, foreigners make up an important and time-varying share of property transactions. As noted in the previous section, they also have different valuations for features such as cash-flow potential/location quality as measured by NER.

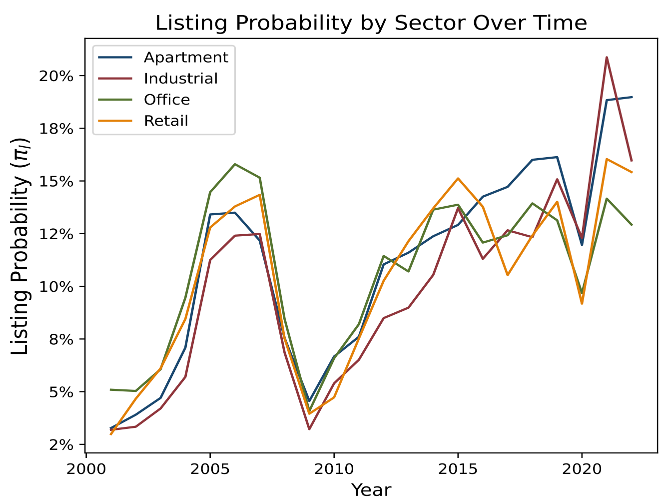
To assess the impact of foreign buyers on the Manhattan office market, we compute what prices would have looked like in their absence, and compare them to the prices in their presence. The top panel of Figure 9 shows the potential price distribution the baseline model, in the presence of foreign buyers, for each year from 2008 until 2022. The top and bottom whiskers indicate the 75th and 25th percentiles of the

Figure 7: Transaction Probabilities



Notes: This figure shows the model-implied probabilities of transactions. The left panel displays the probability that the seller transacts with the actual buyer, combining the meeting and transaction probabilities. The right panel shows the probability that the seller transacts with any buyer, reflecting aggregate transaction likelihoods across all potential buyers.

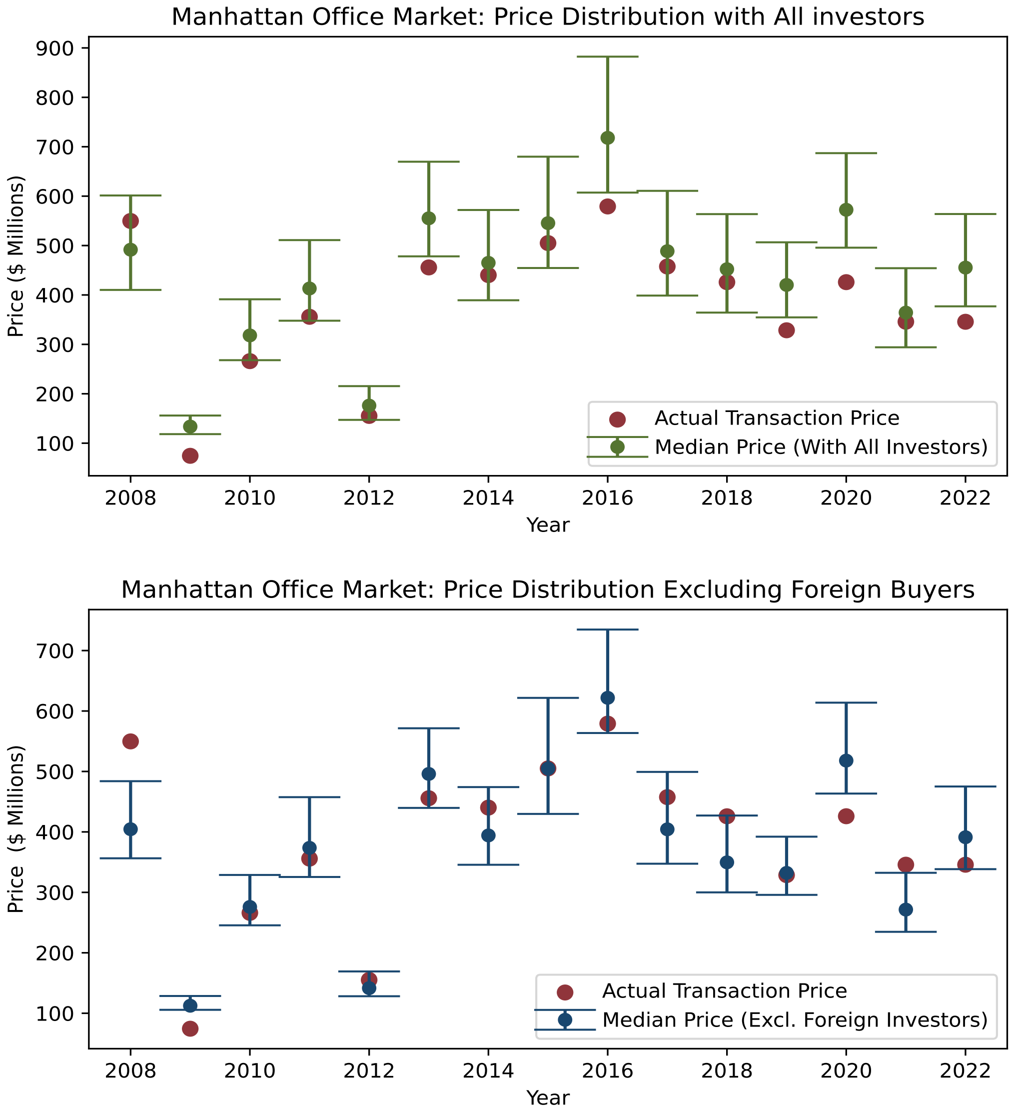
Figure 8: Listing Probabilities



Notes: This figure shows the listing probabilities π_l for the benchmark model over time across all sectors (Apartment, Industrial, Office, and Retail), as defined in (10).

distribution, the green dot in the middle the median. The solid red dot is the actual transaction price. The bottom panel plots the potential price distribution in a counter-factual world where we have removed all the foreign investors. The assets that in reality were bought by foreigners are now bought by other types of investors, sampled in proportion to their model-implied transaction probabilities. We see that there are several years, such as 2018 and 2021, where the counterfactual price distribution shifts down meaningfully

Figure 9: Manhattan Office Market: Counterfactuals



Notes: This figure illustrates the potential price distribution for Manhattan office in the benchmark model (top panel, green) and in a counter-factual without foreign investors (bottom panel, blue) in each year between 2008 and 2022. The top line segment indicates the 75th percentile of potential prices, while the actual price (red dot) typically falls within this range, demonstrating alignment with modeled transaction dynamics.

from the benchmark one. The years 2018 and 2021 were years with large foreign buyer activity, both in terms of dollars and share of purchases (see Figure A.5). That foreign purchase activity meaningfully pushed up the price of Manhattan office. Removing foreign buyers lowers average office prices by about 10% during 2008-2022, underscoring their pivotal role in sustaining price levels.

6 Conclusion

We provide a novel approach to determine the joint evolution of prices and volume in decentralized markets for private and real assets. We apply our approach to a granular data set of commercial real estate transactions containing detailed information on the identities of the buyers and sellers. We find that using investor composition is a critical hedonic for understanding asset prices in the commercial real estate ecosystem. The meeting model, inspired by directed search and matching theory, demonstrates that buyer-seller interactions are shaped by factors such as valuation gaps, buyer activity, portfolio similarity, and geographic and sectoral alignment. These insights underscore the role of matching frictions and heterogeneous preferences in shaping transaction dynamics. The counterfactual analysis highlights the pivotal role of investor composition in determining price outcomes. For instance, the absence of foreign buyers in Manhattan's office market would have led to substantial declines in prices, emphasizing how specific investor groups influence market equilibria. These insights extend beyond real estate, offering a powerful methodology for understanding price formation in other private asset classes, such as private equity, private credit, and infrastructure, where buyer-seller interactions are similarly crucial.

A Data Appendix

A.1 Data Cleaning Process

This appendix describes the process used to clean and filter the raw transaction data from RCA. The dataset includes transaction-level data for four core sectors—apartments, industrial, office, and retail—from 2001 to 2023.

We begin by removing transactions that are unlikely to be useful or may introduce noise. Using the transaction type variable, we retain only "Sale," "Construction," "Entity Level," and "Transfer" transactions, excluding land sales, air rights, and leases. Additionally, we drop incomplete transactions (e.g., contracts not completed) and those valued under \$500,000. These filters eliminate approximately 5% of all transactions.

For each transaction, RCA provides up to four buyers and sellers, though less than 1% involve all four. We consolidate investor names with inconsistent spellings or formats using the method in Arora and Dell (2023). Specifically, we use two independent models to generate 20 closest matches per name, apply a 90% similarity threshold, manually discard false positives, and replace names with the most frequent version. This standardization allows us to track investors over time, aiding in two key tasks: (1) identifying property owners for each year (critical for the seller role in our meeting model), and (2) constructing investor-level portfolios to study asset evolution. The latter supports variables like investor transaction activity, size, and consideration set characteristics.

Transactions suspected of being non-arms-length (e.g., multiple deals involving the same property within 30 days) are consolidated. We merge such transactions into a single record, using the price and characteristics of the final transaction and assigning buyers and sellers from the first and last transactions, respectively.

Some properties change sectors during the sample period due to conversions. To ensure accurate analysis, we assign separate property IDs for periods before and after such conversions.

Transactions where the buyer and seller are the same entity are removed. For cases where one party transacts on behalf of another, we retain the primary investor and drop intermediaries (e.g., investment managers). If there are multiple buyers or sellers, we allocate equal ownership shares unless specific ownership structures are known (e.g., in Joint Ventures, where we adjust shares accordingly).

We map RCA's raw investor types (20 groups) into eight core investor categories, as shown in Table A.6, to capture meaningful heterogeneity in investment objectives, size, and geographic focus

We track property ownership over time to construct investor portfolios. Challenges such as missing or incorrect transactions are addressed by inferring missing links from prior and subsequent transactions. For example, if investor A sells to B, and C later sells the same property to D, we estimate an intermediate transaction where B sells to C. We adjust buyer and seller shares to ensure total ownership sums to 100%

without negative values. Approximately 20% of properties require such adjustments during the sample period. These corrections are used solely for portfolio construction and are excluded from valuation and meeting model calibrations, which rely strictly on actual transaction data.

For entity sales, where RCA equally distributes transaction value across properties, we adjust property-level prices by assigning per-square-foot values proportional to RCA's sector-market price indices. This provides a more accurate estimate than assuming uniform prices across locations.

We construct property price panels for each quarter/year by combining RCA price indices with unadjusted CPI prices and adjusting them to 2023 dollars. CPI adjustments use the seasonally adjusted Consumer Price Index (CPIAUCSL) from FRED.

Market occupancy data is primarily sourced from NCREIF at the CBSA-sector level. For missing values, we calculate volume-weighted occupancy rates using RCA data, provided at least 15 properties in the sector-year contain valid occupancy information. Net Operating Income (NOI) data, also from NCREIF, is CPI-adjusted and annualized. NOI growth is calculated logarithmically if both the current and previous NOI values are positive; otherwise, we use a smoothed growth rate formula. Extreme outliers (top/bottom 1%) are replaced with median values.

Geographic variables (e.g., zip codes, FIPS codes) are filled where possible. For markets with intermittent missing data (e.g., at least 50% coverage), we use the Simple Python Recommendation System Engine (SURPRISE) Hug (2020) to impute values. Missing data is more prevalent for remote areas and for retail and industrial properties, which are often further from city centers. For geographies with no information, we impute values using sector-level national averages, though such cases are rare.

A.2 Mapping Geographies

The RCA dataset provides geo-codes for every property, which we leverage to map properties to various geographic indicators. This is essential because the macroeconomic time series data we use comes from multiple sources, each often organized by different geographic units. Using the 3-digit rounded latitude and longitude data provided by RCA, we generate ZIP codes for each property. This is done using standard geolocation algorithms that match coordinates to corresponding ZIP codes.

We use the county-level Federal Information Processing Standards (FIPS) codes to map properties to Core-Based Statistical Areas (CBSAs). This mapping is based on a crosswalk file obtained from the National Bureau of Economic Research (NBER). RCA provides its own market classification system, which consists of 165 individual markets. To align with other datasets and enhance interpretability, we aggregate RCA's 165 markets into 61 broader geographic categories based on logical groupings and economic characteristics.

The aggregated RCA markets enable direct mapping of properties to RCA price index files and cap rate data, both of which are reported at the level of these 61 aggregated markets. This standardization allows for consistent integration of property-level data with macroeconomic indicators. Table A.11 shows the mapping between RCA markets and the aggregate markets.

A.3 Neighborhood Net Effective Rent

This appendix describes how we construct net effective rent statistics in the neighborhood of each property that we see transact in the RCA data. Net effective rents are observed for office, retail, and industrial properties at the lease level in the Compstak database. We lack the information to merge the RCA and the Compstak databases at the property level. The reason is that property name, address, and geopoint information is noisy. The same office building can have different street addresses, latitude/longitude, and names in the two data sets and can even change name or address. Therefore, we create variables in the lease-level data at a fine level of geography (call it a “block”), and match each property we see transact in RCA to its block-level variables.

The finest level of geography we consider uses the first 2 digits of the latitude and longitude. For example, for a property with latitude and longitude (40.754932,-73.984016), we define a 2d area as [40.75,40.76;-73.99,-73.98]. This corresponds to an area of roughly 0.6 miles by 0.6 miles (1km by 1km). We collect all the leases in this block to calculate the average rent measures. Given that leasing activity is sparse, we use all leases signed in the past three years to calculate the block-year level average rent. We weigh leases by their transaction sqft, and assign a lower weight for leases signed in the more distant past. In the baseline, we assign weights of 1, 1/2, and 1/3 to leases signed one, two, and three years before, and then renormalize to make the weights sum to one.

When there are fewer than 5 buildings in the block, or when there are no leases in the past three years, we use the average rent calculated in a larger block. The next block we consider is defined based on the first decimal place of the latitude and longitude. For example, a property at latitude and longitude (40.754932,-73.984016) belongs to block [40.7,40.8;-74.0,-73.9].

When the average rent is still missing in this roughly 6 miles by 6 miles block, we fill it with the value in the block defined based on the unit digit of the latitude and longitude. In our example, this would be the area given by [40,41;-73,-74], an area of roughly 60 miles by 60 miles.

For the remaining missing values, we sequentially fill them with the average rent at the RCA market, average rent at our own market definition (which combines RCA markets, see appendix A.2), and if necessary, average rent measured at the national level.

For apartment properties, we observe the net operating income per apartment unit at the loan level in the Fannie Mae's multifamily mortgage origination data. We follow a similar approach to calculate the NOI (referred to as NER below for simplicity). Given that loan origination activity is also sparse, we use all originations in the past three years to calculate the block-year level average NER. We start with the 2-decimal place blocks, use the same weighting scheme explained above, and fill missing values sequentially with values computed on increasingly coarse geographic levels.

Table A.1 reports the proportion of the final NER observations for each sector (columns) that come from each level of geography (rows). Most of the NER are measured at the 1-decimal level (6 miles by 6 miles). For office, we have nearly 42% of observations measured at the even finer 2-decimal level (0.6 miles by 0.6 miles).

To understand how well the block-level NER measures captures property-level variation in NER, we estimate a linear regression of property-level NER in the Compstak/Fannie Mae data on the block-level NER variable, where we use the most spatially granular variable available for that property. The R^2 of those property-level regressions is reported in the last row of Table A.1.

We have explored several more neighborhood-level variables available in the Compstak data, such as occupancy, quality grade (class A, B, C), remaining lease duration, tenant size (number of employees of the tenant), tenant concentration (Herfindahl index computed form the leases in a building), but none of these variables add explanatory power to the valuation model, once the NER is included.

Level	**Office Sector**	**Industrial Sector**	**Retail Sector**	**Apartment Sector**
2-decimal lat/lon	41.9%	35.6%	21.2%	38.7%
1-decimal lat/lon	44.5%	43.8%	55.0%	51.0%
unit-digit lat/lon	9.8%	14.4%	18.3%	9.5%
RCA Market	2.2%	2.9%	3.8%	0.8%
Market	1.3%	2.4%	1.5%	0.0%
National	0.2%	1.0%	0.2%	0.0%
Regression R2	74.5%	69.2%	45.3%	76.8%

Table A.1: NER Information Source

A.4 Transaction Summary Statistics

Table A.2: Statistics by Transaction Year

Year	# Trans	% Trans	Cum. % Trans	Volume	% Vol	Cum. % Vol	% Foreclosures	% EntitySales
2001	5,272	1.11	1.11	150.71	1.49	1.49	0.14	6.82
2002	6,287	1.32	2.43	186.11	1.84	3.34	0.06	2.19
2003	7,977	1.68	4.10	223.00	2.21	5.55	0.21	2.96
2004	11,821	2.48	6.59	348.93	3.46	9.00	0.18	7.44
2005	20,725	4.35	10.94	490.05	4.86	13.86	0.15	3.14
2006	23,025	4.84	15.78	555.20	5.50	19.36	0.13	9.39
2007	23,385	4.91	20.69	675.31	6.69	26.05	0.34	21.01
2008	13,531	2.84	23.53	230.60	2.28	28.34	2.70	2.83
2009	6,884	1.45	24.98	109.09	1.08	29.42	17.25	0.38
2010	10,873	2.28	27.26	211.82	2.10	31.52	11.88	0.33
2011	14,749	3.10	30.36	304.41	3.02	34.53	6.64	8.58
2012	19,785	4.16	34.52	374.61	3.71	38.24	4.94	5.74
2013	21,200	4.45	38.97	434.17	4.30	42.54	2.30	5.71
2014	24,974	5.25	44.22	503.24	4.99	47.53	2.27	5.90
2015	29,252	6.15	50.36	616.83	6.11	53.64	0.48	10.14
2016	26,540	5.58	55.94	574.22	5.69	59.33	0.71	4.87
2017	26,996	5.67	61.61	542.27	5.37	64.70	0.73	4.33
2018	29,044	6.10	67.71	611.31	6.06	70.76	0.66	12.52
2019	30,679	6.44	74.16	627.41	6.22	76.98	0.36	1.26
2020	23,793	5.00	79.15	452.95	4.49	81.46	0.48	4.37
2021	40,410	8.49	87.64	854.42	8.47	89.93	0.55	4.27
2022	36,266	7.62	95.26	697.79	6.91	96.84	0.86	8.10
2023	22,554	4.74	100.00	318.54	3.16	100.00	1.68	2.95

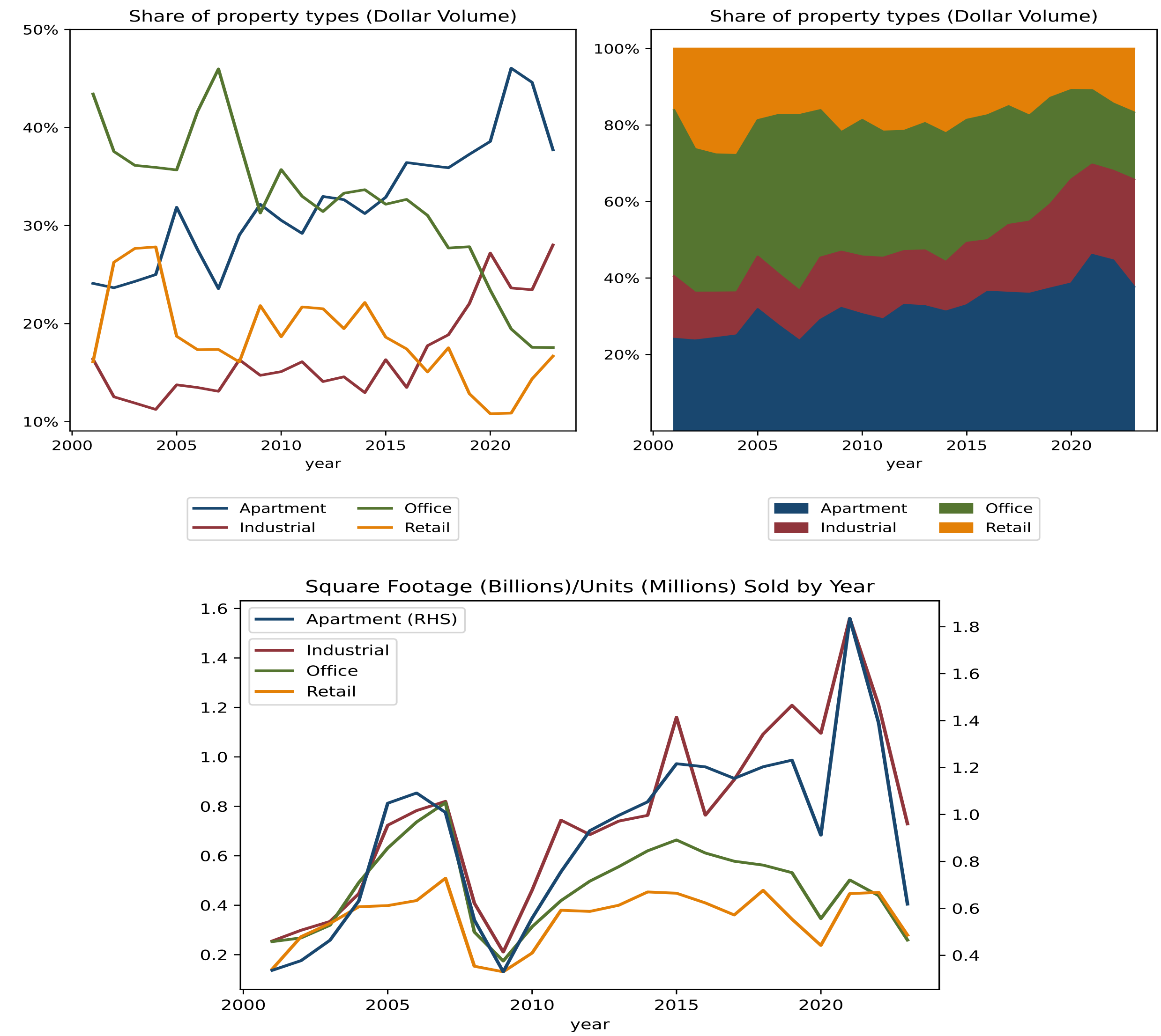
Notes: This table shows the number of transactions, the share and cumulative share of total transactions, the volume (dollar value), share and cumulative share of volume by transaction year. The last two columns displays the share of transactions that are foreclosures and the share of transactions that are entity sales. Volume is defined in billions of 2023 U.S. dollars.

Table A.3: Transactions By Transaction Size

	# Trans	% Trans	Cum. % Trans	Vol	% Vol	Cum. % Vol
Above 1 Bil	269	0.06	0.06	327	3.24	3.24
500 Mil - 1 Bil	701	0.15	0.20	374	3.71	6.95
250-500 Mil	2,368	0.50	0.70	704	6.97	13.92
100-250 Mil	12,525	2.63	3.33	1,726	17.10	31.02
75-100 Mil	9,301	1.95	5.29	772	7.65	38.68
50-75 Mil	19,926	4.19	9.47	1,181	11.71	50.38
25-50 Mil	52,693	11.07	20.54	1,814	17.97	68.35
20-25 Mil	22,517	4.73	25.27	496	4.91	73.26
15-20 Mil	33,779	7.10	32.37	578	5.72	78.99
10-15 Mil	57,414	12.06	44.43	695	6.89	85.87
5-10 Mil	135,100	28.38	72.81	951	9.42	95.30
Below 5 Mil	129,429	27.19	100.00	474	4.70	100.00

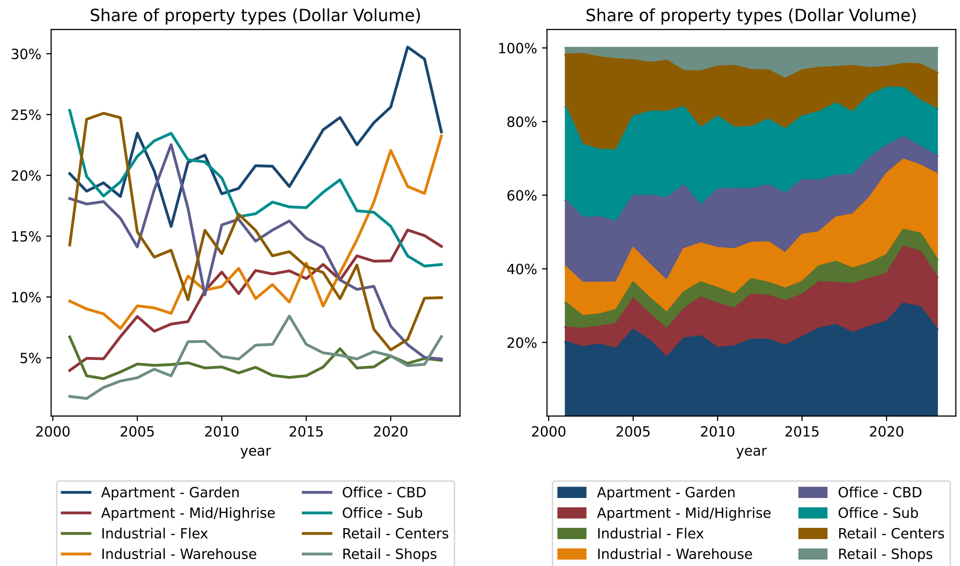
Notes: This table shows the number of transactions, the % of transactions, Value and % of Value for transactions by various transaction size buckets. Dollar volume is defined in terms of Billions of 2023 US Dollars.

Figure A.1: Transaction Volume by Property Type



Notes: The top panel shows share of properties transacted by property types, i.e. Apartments, Office, Industrial and Retail, from 2001 to 2023. Value is defined in terms of 2023 USD. The bottom panel shows the total square feet transacted in billions of Industrial, Office and Retail space, and the number of Apartment units bought in millions.

Figure A.2: Transaction Volume by Property Subtype



Notes: The top panel shows share of properties transacted by property subtypes., i.e. Apartments, Office, Industrial and Retail, from 2001 to 2023. Value is defined in terms of 2023 USD. The bottom Panel shows the total square footage bought (in billions) of Industrial, Office and Retail space, and the number of Apartment units bought (in millions) from 2001 to 2023.

Table A.4: Transactions By Major Market

	# Trans	% Trans	Vol	% Vol	%A	%I	%O	%R
Manhattan	12,617	2.65	733.41	7.27	26.15	0.85	63.82	9.18
Los Angeles	30,892	6.49	578.30	5.73	29.03	20.25	33.95	16.77
Dallas	18,720	3.93	448.09	4.44	44.79	18.30	24.20	12.72
Chicago	19,060	4.00	405.98	4.02	21.55	24.14	36.34	17.97
Atlanta	15,828	3.33	372.71	3.69	43.72	17.27	24.37	14.64
Houston	12,937	2.72	303.41	3.01	42.57	14.15	28.35	14.92
Boston	8,268	1.74	303.20	3.00	20.36	12.95	57.92	8.78
Seattle	10,744	2.26	279.32	2.77	34.78	14.64	39.27	11.30
Phoenix	13,512	2.84	277.81	2.75	46.14	16.21	22.14	15.51
San Francisco	7,561	1.59	242.48	2.40	21.18	8.48	60.49	9.85
DC VA burbs	5,051	1.06	236.12	2.34	36.26	10.96	42.38	10.40
Northern NJ	10,114	2.12	205.36	2.03	24.83	28.42	32.81	13.94
San Diego	9,332	1.96	199.01	1.97	31.34	19.69	33.75	15.22
San Jose	6,280	1.32	197.36	1.96	17.02	23.26	50.56	9.15
Washington DC	2,395	0.50	147.88	1.47	16.20	1.18	78.14	4.48
Miami	7,239	1.52	142.94	1.42	30.99	19.78	27.49	21.74
All Others	285,472	59.97	5,019.61	49.73	36.85	20.13	21.08	21.94

Notes: This table shows the number of transactions, the % of transactions,volume and share of volume for transactions for the largest 16 markets, ranked by volume, as well as in the remaining 44 markets (last row). The last four columns shows the percentage breakdown of volume by sector (Apartments, Industrial, Office, and Retail) within each market. Volume is expressed in billions of 2023 U.S. dollars. The cities in bold are 11 of the 16 superstar cities. The other five superstar cities that are not explicitly listed in the top of this table are: East Bay, DC MD burbs, Orange Co, NYC Boroughs, and Austin.

Table A.5: Transactions By Transaction Type

Type	Subtype	# Trans	% Trans	Vol	% Vol
Conventional Sale	Ind. Asset	354,712	74.53	7068.52	70.06
	Portfolio	89,256	18.75	2087.41	20.69
Entity Sale	Private Merger	1,852	0.39	61.10	0.61
	Privatization	5,832	1.23	289.44	2.87
	Public Merger	7,549	1.59	274.93	2.73
	Publicization	2,098	0.44	61.28	0.61
Foreclosure	Ind. Asset	7,895	1.66	128.55	1.27
	Portfolio	1,867	0.39	22.48	0.22
Debtor/Trustee Sale	Ind. Asset	3,486	0.73	74.64	0.74
	Portfolio	1,415	0.30	20.50	0.20

Notes: This table shows the number of transactions, the share of transactions,volume and share of volume by transaction type and subtype. Volume is expressed in billions of 2023 U.S. dollars.

Table A.6: Mapping Investor Types

RCA Group	RCA Type	Investor Type	%Trans Buy	% Vol Buy	%For Buy	%Trans Sell	% Vol Sell	%For Sell
CMBS	CMBS	Institutional	0.90	0.71	0.68	0.93	0.62	0.43
Equity Fund	Equity Fund	REPE	6.79	12.54	2.03	5.34	10.43	1.89
Institutional	Bank	Institutional	1.25	1.83	26.65	1.75	2.15	25.37
	Endowment	Institutional	0.09	0.12	0.00	0.09	0.12	0.00
	Finance	Institutional	0.33	0.32	22.93	0.32	0.30	13.16
	Insurance	Institutional	0.84	1.88	25.77	0.97	2.13	14.98
	Investment Manager	Institutional	5.90	11.36	19.58	5.00	8.85	17.31
	Open-Ended Fund	Institutional	0.01	0.08	98.74	0.00	0.02	94.23
	Pension Fund	Institutional	1.12	2.73	33.09	1.01	2.29	21.31
	Sovereign Wealth Fund	Foreign	0.57	0.89	100	0.22	0.30	100
Private	Developer/Owner/Operator	Foreign	0.67	0.98	100	0.45	0.68	100
		DOO_L	29.40	12.96	0.02	32.62	16.33	0.02
		DOO_N	32.07	30.43	0.00	28.24	30.44	0.00
	High Net Worth	Individual	1.81	1.20	22.71	2.97	1.41	7.14
	Non Traded REIT	Individual	2.37	3.19	0.06	1.22	1.64	5.09
Public	Listed Funds	REITS	0.01	0.02	89.01	0.01	0.03	78.36
	REIT	REITS	7.09	12.27	6.66	7.34	12.93	5.60
	REOC	REITS	0.61	0.94	72.49	0.49	1.05	64.98
Unknown	Unknown	Unknown	2.39	1.22	0.83	4.12	2.44	0.27
User/Other	Cooperative	User	0.01	0.01	3.68	0.03	0.04	0.20
	Corporate	User	4.17	3.04	6.49	5.84	4.85	6.59
	Educational	User	0.31	0.34	0.10	0.12	0.16	0.13
	Government	User	0.62	0.49	1.12	0.38	0.34	0.73
	Non-Profit	User	0.48	0.31	1.69	0.39	0.30	7.96
	Other	User	0.06	0.06	4.33	0.03	0.03	0.00
	Religious	User	0.14	0.07	2.10	0.11	0.10	2.54

Notes: The top table shows the mapping from RCA investor group (column 1) and RCA investor type (column 2) into our investor type (column 3). We show the share of buy transactions and buy volume for each type in columns 4 and 5 and the share of sell transactions and sell volume for each type in columns 7 and 8. Columns 6 and 9 show the share of buy and sell transactions that include a foreign entity. Volume is expressed in billions of 2023 U.S. dollars.

Table A.7: Transactions by Number of Investors

	1 Seller	2 Sellers	3 Sellers	4 Sellers
1 Buyer	87.30	5.723	0.584	0.154
2 Buyers	4.73	0.654	0.099	0.020
3 Buyers	0.50	0.076	0.004	0.003
4 Buyers	0.14	0.022	0.001	0.001

	1 Seller	2 Sellers	3 Sellers	4 Sellers
1 Buyer	73.34	10.033	1.270	0.367
2 Buyers	10.31	1.914	0.422	0.118
3 Buyers	1.38	0.424	0.026	0.014
4 Buyers	0.29	0.071	0.003	0.018

Notes: The left panel splits the number of transactions into sixteen groups, by the number of buyers and the number of sellers involved in each transactions. The numbers in the table add up to 100%. The right panel does the same based on transaction volume, expressed in billions of 2023 U.S. dollars.

Table A.8: Joint Ventures

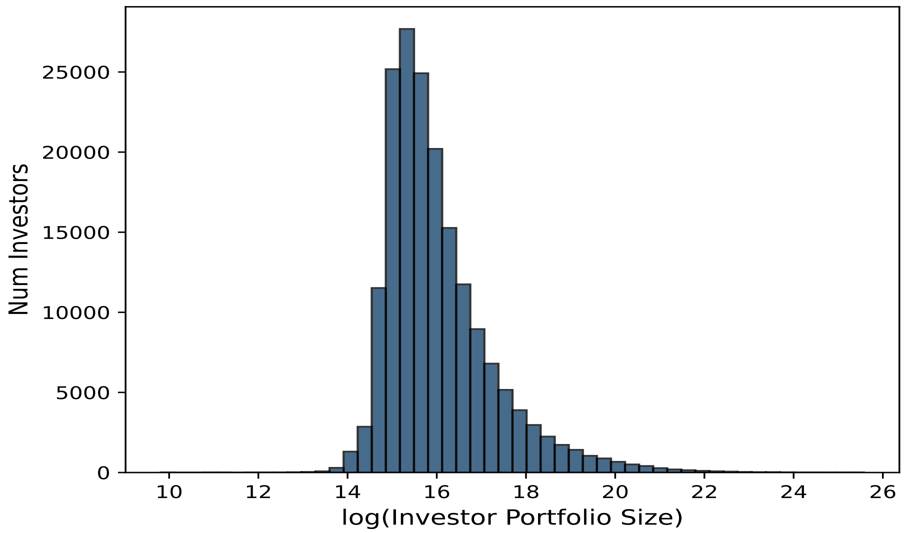
	REPE	Institutional	OOD_L	OOD_N	Individual	REITS	Foreign	User
REPE	2.35							
Institutional	5.03	2.51						
OOD_L	1.22	1.12	5.2					
OOD_N	13.31	14.06	7.45	13.63				
Individual	0.66	1.07	1.0	1.96	0.46			
REITS	1.46	2.75	0.23	2.37	1.09	0.07		
Foreign	1.62	3.41	0.64	6.26	0.72	1.83	4.49	
User	0.09	0.28	0.35	1.01	0.04	0.04	0.04	0.15

	REPE	Institutional	OOD_L	OOD_N	Individual	REITS	Foreign	User
REPE	9.12	16.63	7.10	22.16	9.40	14.84	8.54	4.68
Institutional	19.53	8.29	6.52	23.41	15.31	27.94	17.96	14.16
OOD_L	4.74	3.71	30.24	12.41	14.32	2.28	3.36	17.36
OOD_N	51.70	46.50	43.29	22.70	28.07	24.10	32.94	50.25
Individual	2.55	3.54	5.82	3.27	6.54	11.07	3.77	1.85
REITS	5.68	9.10	1.31	3.95	15.59	0.73	9.62	2.22
Foreign	6.31	11.29	3.71	10.43	10.25	18.58	23.59	1.97
User	0.36	0.94	2.03	1.68	0.53	0.45	0.21	7.51

	REPE	Institutional	OOD_L	OOD_N	Individual	REITS	Foreign	User
% Total Buys	27.34	24.77	4.43	13.99	11.66	10.73	34.16	2.42

Notes: This first panel breaks down the total number of buy transactions involving multiple buyers into who forms a JV with whom. The second panel shows this number as a share of JV property purchases by each investor type; every column sums to 100%. The third panel table shows the share of properties bought in JV with every other investor type, as a share of each investor's total purchases.

Figure A.3: Investor Size Distribution



Notes: Figure shows a histogram of log(investor portfolio size) at the end of 2023 for investors that own atleast one property. Log investor size varies from 11.8 (∼ $135K) to 25.5 (∼ $119 bi). Not that 16.2 = $10 mi, 17.7 = $50 mi, 20.7 = $1 bi.

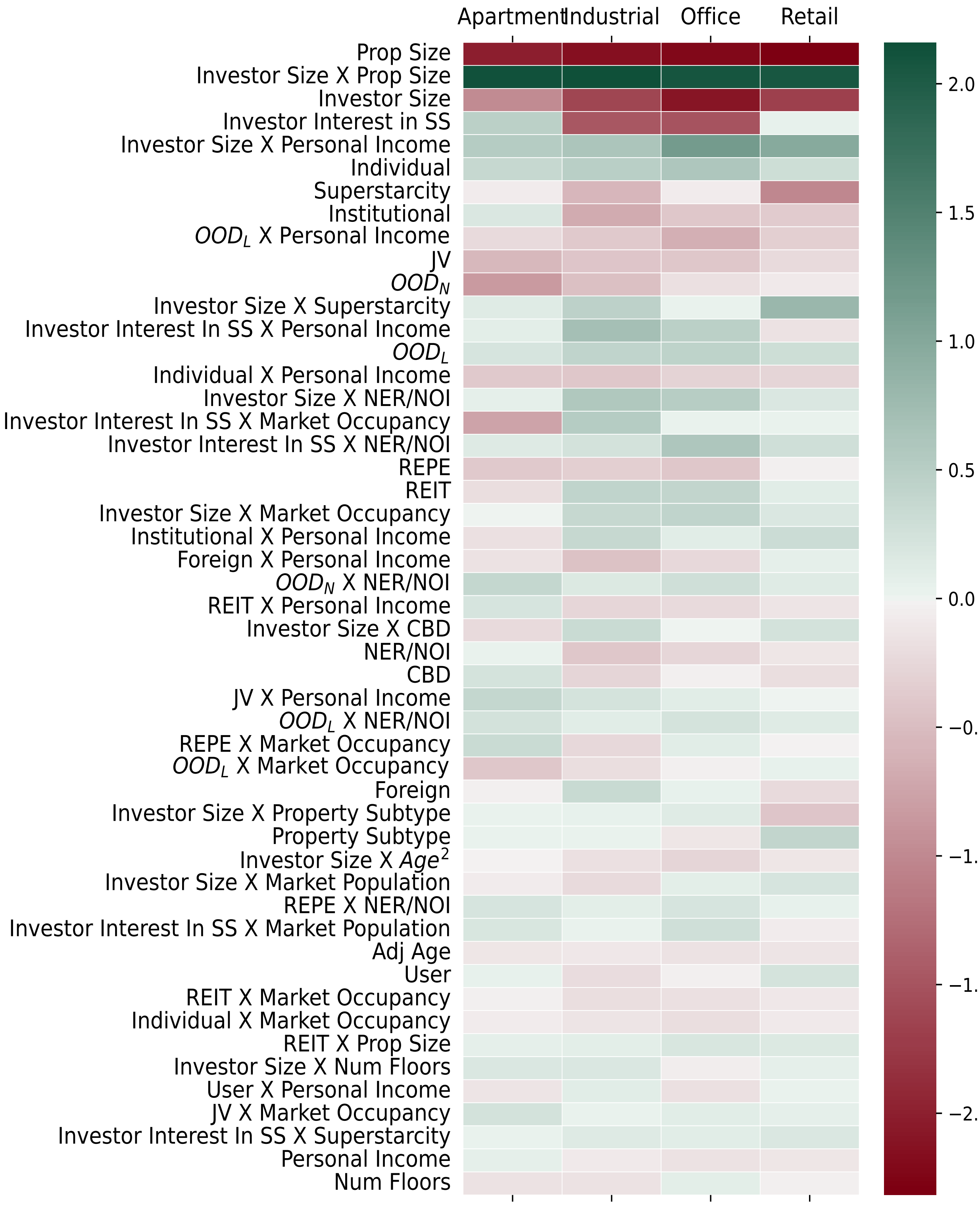
Table A.9: Trade Networks

	REPE	Institutional	OOD_L	OOD_N	Individual	REITS	Foreign	User	Unknown
REPE	205.87	241.76	55.18	334.05	44.70	132.03	159.68	28.33	4.71
Institutional	243.07	321.74	82.39	461.55	56.27	192.81	176.05	42.65	7.44
OOD_L	102.32	137.32	590.45	610.10	52.38	83.67	55.94	96.05	27.51
OOD_N	368.63	536.94	395.71	1263.38	152.55	309.66	235.54	111.02	29.29
Individual	27.06	36.34	52.15	82.97	17.98	62.32	31.71	9.40	2.23
REITS	318.93	243.37	39.15	278.62	51.98	324.63	186.06	30.75	8.51
Foreign	145.44	131.80	29.09	164.04	32.83	71.97	123.05	25.37	1.93
User	68.04	81.48	88.76	182.84	22.79	80.48	34.79	74.60	3.90
Unknown	9.67	23.89	48.21	76.17	9.65	27.93	8.43	12.82	37.86

	REPE	Institutional	OOD_L	OOD_N	Individual	REITS	Foreign	User	Unknown
REPE	1.81	2.13	0.49	2.94	0.39	1.16	1.40	0.25	0.04
Institutional	2.14	2.83	0.72	4.06	0.49	1.70	1.55	0.38	0.07
OOD_L	0.90	1.21	5.19	5.37	0.46	0.74	0.49	0.84	0.24
OOD_N	3.24	4.72	3.48	11.11	1.34	2.72	2.07	0.98	0.26
Individual	0.24	0.32	0.46	0.73	0.16	0.55	0.28	0.08	0.02
REITS	2.80	2.14	0.34	2.45	0.46	2.85	1.64	0.27	0.07
Foreign	1.28	1.16	0.26	1.44	0.29	0.63	1.08	0.22	0.02
User	0.60	0.72	0.78	1.61	0.20	0.71	0.31	0.66	0.03
Unknown	0.09	0.21	0.42	0.67	0.08	0.25	0.07	0.11	0.33

Notes: This table summarizes the trade networks among various investor types. We show buyers along the columns and sellers along the rows. Top Panel reports the $ Volume of transaction between different investor types, expressed in billions of 2023 U.S. dollars. If a transaction involves multiple buyers (sellers) of different types, transaction value is split among them proportionately. The bottom table reports the share of overall transaction $ Volume. Values in the entire table sum up to 100%

Figure A.4: Feature Importance



Notes: Figure shows the fifty most important features selected by the Ridge model. Variables are sorted by the average absolute coefficient across sectors.

Table A.10: Meeting Model Analysis

Apartments

Model	λ_1	λ_2	$\lambda_{3,1}$	$\lambda_{3,2}$	$\lambda_{3,3}$	λ_4	Avg. Log. Likelihood
Base Model	0.257	1.398	5.241	8.407	3.548	3.365	-0.204
Ex. λ_1		1.44	5.265	8.397	3.544	3.421	-0.255
Ex. λ_2	3.62		4.565	8.236	3.268	3.59	-0.413*
Ex. $\lambda_{3,1}$	1.198	1.231		8.689	3.913	2.68	-0.896**
Ex. $\lambda_{3,2}$	-0.234	1.198	7.131		4.392	1.518	-1.149***
Ex. $\lambda_{3,3}$	0.295	1.235	6.049	8.718		3.484	-0.347
Ex. λ_4	1.433	1.533	4.195	7.312	3.634		-0.508*

Industrial

Model	λ_1	λ_2	$\lambda_{3,1}$	$\lambda_{3,2}$	$\lambda_{3,3}$	λ_4	Avg. Log. Likelihood
Base Model	-0.021	1.52	4.721	8.726	3.998	4.127	-0.227
Ex. λ_1		1.515	4.7	8.717	3.988	4.133	-0.288
Ex. λ_2	2.743		2.954	7.888	4.147	4.611	-0.508*
Ex. $\lambda_{3,1}$	0.549	1.325		8.892	4.494	3.517	-0.421*
Ex. $\lambda_{3,2}$	0.953	1.217	7.194		4.093	1.689	-1.170***
Ex. $\lambda_{3,3}$	-0.236	1.661	6.129	8.71		3.194	-0.492*
Ex. λ_4	0.913	1.767	3.507	7.158	3.142		-0.356

Office

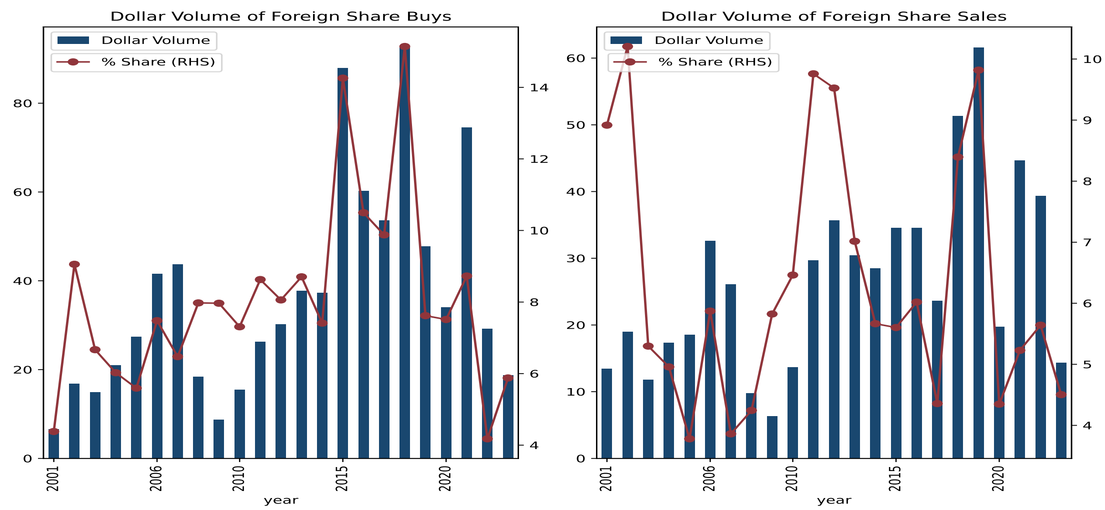
Model	λ_1	λ_2	$\lambda_{3,1}$	$\lambda_{3,2}$	$\lambda_{3,3}$	λ_4	Avg. Log. Likelihood
Base Model	0.346	1.297	4.691	8.435	3.897	3.991	-0.244
Ex. λ_1		1.366	4.775	8.418	3.903	4.084	-0.286
Ex. λ_2	2.827		3.903	7.868	3.564	4.149	-0.681*
Ex. $\lambda_{3,1}$	0.94	1.161		8.589	4.203	3.337	-0.414
Ex. $\lambda_{3,2}$	0.927	0.936	6.573		4.127	1.774	-1.254***
Ex. $\lambda_{3,3}$	0.285	1.246	5.643	8.495		3.335	-0.518*
Ex. λ_4	1.421	1.431	3.53	6.988	3.21		-0.391

Retail

Model	λ_1	λ_2	$\lambda_{3,1}$	$\lambda_{3,2}$	$\lambda_{3,3}$	λ_4	Avg. Log. Likelihood
Base Model	-0.752	1.94	6.137	9.863	4.57	4.904	-0.267
Ex. λ_1		1.916	6.146	10.029	4.777	4.886	-0.328
Ex. λ_2	1.13		4.357	8.514	4.647	5.692	-0.581*
Ex. $\lambda_{3,1}$	-0.304	1.722		9.981	5.131	4.104	-0.393
Ex. $\lambda_{3,2}$	0.119	1.405	7.838		4.278	1.798	-1.306***
Ex. $\lambda_{3,3}$	-0.931	1.977	7.56	9.706		4.226	-0.489*
Ex. λ_4	-0.022	2.25	4.544	7.778	3.855		-0.757**

Notes: Figure shows the point estimates of the meeting model for each of the four sectors. The last column reports the log-likelihood. The first row in each panel is the baseline model. The successive rows leave out one of the variables in the meeting model, each of which is a special case of the baseline model. We test the null hypothesis that the log-likelihood is the same in the special case and the baseline model using a likelihood ratio test for each year and report the significance using Fisher's combined probability test. Stars indicate whether the null hypothesis can be rejected at the 10% (*), 5% (**), or 1% level (***).

Figure A.5: Foreign Investor Share



Notes: Figures show foreign investor activity in US Commercial Real Estate Ecosystem during 2001-2023. The blue bars show the dollar value (in Billions, 2023 USD) traded, aggregated by property type while the red lines show the share of foreign investors among all investors. The left panel summarizes foreign investor buys while the right panel summarizes foreign investor sales.

Table A.11: RCA Markets to Aggregate Markets map

Our Market	RCA Markets
Atlanta	Atlanta
Austin	Austin
Baltimore	Baltimore
Birmingham	Birmingham (AL)
Boston	Boston
Broward - Palm Beach	Broward, Palm Beach Co
Charlotte	Charlotte
Chicago	Chicago
Cincinnati	Cincinnati
Cleveland	Cleveland
Columbus	Columbus
DC	DC
DC MD burbs	DC MD burbs
DC VA burbs	DC VA burbs
Dallas	Dallas
Denver	Denver
Detroit	Detroit
Hartford	Hartford
Houston	Houston
Indianapolis	Indianapolis
Inland Empire	Inland Empire
Jacksonville	Jacksonville
Kansas City	Kansas City
Las Vegas	Las Vegas
Los Angeles	Los Angeles
Manhattan	Manhattan
Memphis	Memphis
Miami/Dade Co	Miami/Dade Co
Minneapolis	Minneapolis
NYC Boroughs	NYC Boroughs
NYC Suburbs	Long Island, No NJ, Stamford, Westchester
Nashville	Nashville
Orange Co	Orange Co
Orlando	Orlando
Philadelphia	Philadelphia
Phoenix	Phoenix
Pittsburgh	Pittsburgh
Portland	Portland
Raleigh/Durham	Raleigh/Durham
Richmond/Norfolk	Norfolk, Richmond
Sacramento	Sacramento
Salt Lake City	Salt Lake City
San Antonio	San Antonio
San Diego	San Diego
San Francisco	San Francisco

Notes: This table provides the mapping we use from the raw RCA markets to the aggregated RCA markets we use.

Table A.11: RCA Markets to Aggregate Markets map: Continued

Our Market	RCA Markets
San Francisco Suburbs	East Bay, North Bay, SF Metro Other
San Jose	San Jose
Seattle	Seattle
St Louis	St Louis
Tampa	Tampa
Tertiary Mid-Atlantic	All Others - DE,MD,VA,WV; Ann Arbor, Albany, All Others - NJ, All Others - NY, All Others - PA, Allentown, Buffalo, Camden, NJ, Eastern PA, Harrisburg, Rochester, Syracuse, Trenton, NJ
Tertiary Midwest	Lexington, Louisville, Akron, All Others - IL, IN, OH, Dayton, Kankakee, IL, Toledo, All Others - Plains States, Des Moines, Omaha, Wichita, All Others - MI,WI, Flint, MI, Grand Rapids, Lansing, Madison, WI, Milwaukee, Monroe, MI, Racine, WI, All Others - MN.
Tertiary Northeast	All Others - CT, All Others - New England, New Haven, Portland, ME, Providence, Southern NH, Worcester
Tertiary Southeast	Polk Co, All Others - FL, Brevard Co, Daytona Beach, Florida Panhandle, Fort Myers, Gainesville, Martin/Saint Lucie, Sarasota, Tallahassee, All Others - GA,NC,SC,TN,KY, Charleston, Chattanooga, Columbia, Greensboro, Greenville, Knoxville, Myrtle Beach, Savannah, Winston-Salem, All Others - AR,LA, All Others - MS,AL, Baton Rouge, Huntsville, AL, Jackson, MS, Little Rock, New Orleans
Tertiary Southwest	Boulder, Greeley, CO, Albuquerque, All Others - Southwest, Tucson, All Others - OK,TX, Corpus Christi, El Paso, McAllen, Oklahoma City, Tulsa, Co Springs, Fort Collins, Provo
Tertiary West	Santa Barbara, Ventura Co, All Others - Northwest, Spokane, All Others - Southern California, Reno, Bakersfield, Fresno, Madera, CA, Modesto, Salinas, Vallejo-Fairfield, CA, All Others - Northern California, Napa, CA, Santa Rosa, CA, All Others - ID,MT,WY; Boise, Hawaii, Honolulu

Notes: This table provides the mapping we use from the raw RCA markets to the aggregated RCA markets we use.

References

Ang, Andrew, Chen, Bingxu, Goetzmann, William N. and Phalippou, Ludovic. (2018). 'Estimating Private Equity Returns from Limited Partner Cash Flows', *Journal of Finance* 73(4), 1751–1783.

Arora, Abhishek and Dell, Melissa. (2023). 'LinkTransformer: A Unified Package for Record Linkage with Transformer Language Models'.

Badarinza, Cristian, Balasubramaniam, Vimal and Ramadorai, Tarun. (2024), In Search of the Matching Function in the Housing Market, Technical report, SSRN.
URL: *https://ssrn.com/abstract=12345678*

Badarinza, Cristian and Ramadorai, Tarun. (2018). 'Home Away from Home? Foreign Demand and London House Prices', *Journal of Financial Economics* 130(3), 532–555.

Badarinza, Cristian, Ramadorai, Tarun and Shimizu, Chihiro. (2022). 'Gravity, Counterparties, and Foreign Investment', *Journal of Financial Economics* 145(2, pages = 132–152, month = August, doi = 10.1016/j.jfineco.2021.09.011, url = https://doi.org/10.1016/j.jfineco.2021.09.011,).

Cvijanović, Dragana, Milcheva, Stanimira and van de Minne, Alex. (2022). 'Preferences of Institutional Investors in Commercial Real Estate', *The Journal of Real Estate Finance and Economics* 65, 321–359.

Driessen, Joost, Lin, Tse-Chun and Phalippou, Ludovic. (2012). 'A new method to estimate risk and return of nontraded assets from cash flows: The case of private equity funds', *Journal of Financial and Quantitative Analysis* 47(3), 511–535.

Fox, Jeremy T. (2018). 'Estimating Matching Games with Transfers', *Quantitative Economics* 9(1), 1–38.
URL: *https://www.nber.org/papers/w14382*

Gabaix, Xavier, Koijen, Ralph S.J., Richmond, Robert and Yogo, Motohiro. (2024), Asset Embeddings, Technical report, SSRN Working Paper Series.
URL: *https://www.ssrn.com/abstract=4507511*

Ghent, Andra C. (2021). 'What's wrong with Pittsburgh? Delegated investors and liquidity concentration', *Journal of Financial Economics* 139(2), 337–358.
URL: *https://doi.org/10.1016/j.jfineco.2020.08.015*

Giacoletti, Marco. (2021). 'Idiosyncratic Risk in Housing Markets', *The Review of Financial Studies* 34(8), 3695–3741.
URL: *https://doi.org/10.1093/rfs/hhab033*

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

Griliches, Zvi. (1971), Introduction: Hedonic Price Indexes Revisited, *in* **Zvi Griliches**., ed., 'Price Indexes and Quality Change: Studies in New Methods of Measurement', Harvard University Press, Cambridge, MA, pp. 3–15.

Gupta, Arpit, Mittal, Vrinda and Van Nieuwerburgh, Stijn. (2022). 'Work from home and the office real estate apocalypse'.

Gupta, Arpit and Van Nieuwerburgh, Stijn. (2021). 'Valuing private equity investments strip by strip', *The Journal of Finance* 76(6), 3255–3307.

Hug, Nicolas. (2020), 'Surprise: A Python library for recommender systems'.

Kaplan, Steven N. and Schoar, Antoinette. (2005). 'Private equity performance: Returns, persistence, and capital flows', *Journal of Finance* 60(4), 1791–1823.

Kim, Min. (2024), Comparative R-Squared. SSRN Working Paper No. 3790200.

Koijen, Ralph S.J., Richmond, Robert J. and Yogo, Motohiro. (2024). 'Which Investors Matter for Equity Valuations and Expected Returns', *Review of Economic Studies* 91(4), 2387–2424.
URL: *https://doi.org/10.1093/restud/rdad083*

Koijen, Ralph S.J. and Yogo, Motohiro. (2019). 'A Demand System Approach to Asset Pricing', *Journal of Political Economy* 127(4), 1475–1515.
URL: *https://doi.org/10.1086/701414*

Korteweg, Arthur and Nagel, Stefan. (2016). 'Risk-Adjusting the Returns to Venture Capital', *Journal of Finance* 71(3), 1437–1470.

Korteweg, Arthur and Sørensen, Morten. (2010). 'Risk and Return Characteristics of Venture Capital-Backed Entrepreneurial Companies', *Review of Financial Studies* 23, 3738–3772.

Lancaster, Kevin. (1966). 'A New Approach to Consumer Theory', *Journal of Political Economy* 74, 132–157.

Ma, Zhuang and Collins, Michael. (2018). 'Noise Contrastive Estimation and Negative Sampling for Conditional Models: Consistency and Statistical Efficiency', *CoRR* abs/1809.01812.
URL: *http://arxiv.org/abs/1809.01812*

Mikolov, Tomás, Sutskever, Ilya, Chen, Kai, Corrado, Greg and Dean, Jeffrey. (2013). 'Distributed Representations of Words and Phrases and their Compositionality', *CoRR* abs/1310.4546.
URL: *http://arxiv.org/abs/1310.4546*

Mittal, Vrinda. (2024), Desperate Capital Breeds Productivity Loss: Evidence from Public Pension Investments in Private Equity. SSRN Working Paper No. 4283853.

Peng, Liang. (2016). 'The risk and return of commercial real estate: A property level analysis', *Real Estate Economics* 44(3), 555–583.

Plazzi, Alberto, Torous, Walter and Valkanov, Ross. (2010). 'Expected Returns and the Expected Growth in Rents of Commercial Real Estate', *Review of Financial Studies* 23, 3469–3519.

Plazzi, Alberto, Torous, Walter and Valkanov, Rossen. (2008). 'The Cross-Sectional Dispersion of Commercial Real Estate Returns and Rent Growth: Time Variation and Economic Fluctuations', *Real Estate Economics* 36(3), 403–439.

Rosen, Sherwin. (1974). 'Hedonic Prices and Implicit Markets: Production Differentiation in Pure Competition', *Journal of Political Economy* 82, 34–55.

Sagi, Jacob S. (2021). 'Asset-Level Risk and Return in Real Estate Investments', *Review of Financial Studies* 34(8), 3647–3686.
URL: *https://academic.oup.com/rfs/article/34/8/3647/5941474*

Van Nieuwerburgh, Stijn. (2019). 'Why Are REITs Currently So Expensive?', *Real Estate Economics* 47(1), 18–65.

Van Nieuwerburgh, Stijn, Stanton, Richard and de Bever, Leo. (2015), 'A Review of Real Estate and Infrastructure Investments by the Norwegian Government Pension Fund Global (GPFG)'.

Wallace, Nancy E. (1996). 'Hedonic-based price indexes for housing: Theory, estimation, and index construction', *Economic Review* pp. 34–48.

Witte, Ann D., Sumka, Howard J. and Erekson, Homer. (1979). 'An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets', *Econometrica* .