



Preferences of Institutional Investors in Commercial Real Estate

Dragana Cvijanović^{1,2} · Stanimira Milcheva³  · Alex van de Minne⁴

Accepted: 16 December 2020 / Published online: 15 May 2021
© The Author(s) 2021

Abstract

In this paper we analyze market segmentation by firm size in the commercial real estate transaction process. Using novel micro-level data, we look at the probability distribution of investors acquiring a specific bundle of real estate characteristics, distinguishing between investors based on the size of their real estate portfolio. We find evidence of market segmentation by investor size: institutional investors segment across property characteristics based on the size of their real estate portfolio. The probability that a large (small) seller will sell a property to a similar-sized buyer is higher, keeping all else equal. We explore potential drivers of this market segmentation and find that it is mainly driven by investor preferences. During the Global Financial Crisis (GFC), large investors were less likely to buy the ‘average’ property, as compared to the period before or after the crisis, indicating time-varying investor preferences.

Keywords Institutional investor · Preferences · Real estate investment · Market segmentation · Size

✉ Stanimira Milcheva
s.milcheva@ucl.ac.uk

Dragana Cvijanović
dc998@cornell.edu

Alex van de Minne
avdminne@uconn.edu

¹ SC Johnson College of Business, Cornell University, Ithaca, NY, USA

² Warwick Business School, University of Warwick, Coventry, UK

³ Bartlett Faculty, University College London, London, UK

⁴ School of Business, University of Connecticut, Storrs, CT, USA

Introduction

Around 90% of investible commercial real estate in the United States is owned and traded in private markets (Ling and Archer 2008; Fisher et al. 2009). Anecdotal evidence from real estate practitioners suggests that since early 2000s private real estate markets have experienced significant inflows of institutional capital. Institutional¹ investors play an important role not only in mainstream asset markets, but also in direct real estate. However, given their dominance, not enough is known about the interplay among institutional investors on commercial real estate markets. Given the significant heterogeneity among institutional investors in real estate and the illiquid nature of the asset class, we are provided with a natural laboratory to analyse the matching between buyers and sellers and the investment preferences of these market players.

Unlike the transaction process in residential real estate, a commercial transaction is larger in value and does not take place in an auction. Instead, the process is rather less transparent and not always the highest bidder would be chosen (see Chinloy et al. 2013). This can result in market segmentation in commercial real estate markets. Some preliminary evidence by Geltner and van de Minne (2017) and Cvijanović et al. (2020) suggests that the law of one price may be violated with different investors paying different prices for a comparable property.

Given the nature of the direct real estate asset class, market segmentation can be present when there is asymmetric information about the property or the actors; other drivers of segmentation in real estate markets can be associated with differences in bargaining power, economies of scale for large players, and style investing. Market segmentation can also be enhanced by the nature commercial real estate is transacted, in that brokers advertise certain types of properties to certain investor clienteles. Another way that segmentation can be propagated is through knowledge networks of investors who prefer to trade with each other only.

In this paper we explore if market segmentation exists by looking at whether the size of the investor plays a role in their preferences for commercial real estate. More specifically, we assess the probability that an investor of a certain size purchases a specific ‘bundle’ of real estate characteristics.

Using a rich micro-level data set on a universe of commercial real estate transactions enables us to observe the dynamics of buying and selling at the individual property level and allows us to try to answer the following questions: 1) Do institutional buyers and sellers segment based on the size of their real estate portfolio? What is the probability that a specific property is bought by a large or small investor? 2) What kind of properties do institutional investors choose to buy based on property’s size, quality, Net Operating Income (NOI) and other characteristics? 3) Is there a time- and geographical variation in the distribution of investors across properties?

Our main findings can be summarized as follows: (1) Institutional investors segment across property characteristics based on the size of their real estate portfolio.

¹We refer to institutional investors for all investors who are not retail investors, including developers, owners and operators.

Large investors tend to buy *large* properties, properties with higher NOI, properties of high quality, young properties, and properties located close to the Central Business District (CBD). The reverse is true for small investors. (2) The probability that a large seller will sell a property to a similar-sized buyer is higher than to other size quantiles. (3) Exploring this segmentation we conclude that it is most likely driven by investor preferences or different riskiness of real estate and investor styles, rather than i.e. financial constraints, based on evidence from variations in net operating income (NOI). This is validated by using bilateral buyer-seller matching pairs. (4) During the Global Financial Crisis (GFC), large investors were less likely to buy the ‘average’ property on the market as compared to the period before or after the crisis. Moreover, the size of the average investor declined across all states during the GFC. This finding implies that the presence of large institutional investors may exacerbate the real estate cycle leading to faster decline in property values during a downturn. Importantly, it also highlights time-variation in investor preferences.

The classic financial constraints literature highlights the importance of size as a proxy for such constraints. For example, small investors may be more likely to buy small properties as they may face binding financial constraints and/or operational limits, and as such tend to be priced out of the larger deals, thus perpetuating market segmentation. In addition, small investors can face barriers to entry as brokers advertise certain properties only to select investor clientele.

The opposite is true for large players. Large investors may be more likely to buy large, centrally located, “safe” properties. One explanation for this can be that large investors do not have the required in-house management expertise or capacity to manage a large portfolio of numerous small properties. Liquidity and liquidation constraints can also play a role, in that transaction costs associated with liquidation of a diversified portfolio can be high. Previous research by Ghent (2018) shows that delegated and direct investors have different preferences for liquidity and hence may have an effect on the overall liquidity of a certain market. Therefore, large investors may behave like delegated investors with size of the investor providing an important dimension for segmentation.

Market segmentation by investor size may also be associated with different risk attitudes and investor styles. The latter would mean that large investors may prefer to invest in the ‘core’ properties, while small investors may position themselves so that they invest in ‘opportunistic’ real estate. That may be related to small investors potentially having a well developed local network and not be as diversified as large players. Therefore, large investors may be seen as picking rather ‘safe’ (‘core’) investments. While those property classifications have been widely used to describe investment styles for real assets, the definition of those risk profiles is highly subjective. In this paper we shed some light on this debate in the context of investor characteristics.

A significant contribution of the paper is the methodology we use to account for market segmentation. We use ordered logistic models in which real estate investors are grouped into different size quantiles and assess what is the probability that a property is bought by a certain investor size quantile (ISQ). More specifically, we follow Donald et al. (2000) by estimating a quantile model in a survival framework. While this method has not been used in real estate research, its application is common in labor economics. The methodology we use provides a way to conduct a logit-style

regression with a large number of categories—the size quantiles. We also model the buyer and seller side of the transaction simultaneously and account for property-level net operating income (NOI).

A further contribution is the construction of investor total real estate portfolio values which we further break down into investor size quantiles (ISQs). We use data from Real Capital Analytics (RCA) from 2002 and 2018. The ISQs are constructed by taking the average sales price of the target properties and building a balanced panel of price estimates for every year and every property. We subsequently link the information about the investor (buyer and seller) of a property and add up the total real estate portfolio value of every investor for every year to get to the investor size.

The paper is organized as follows: we discuss the relevant literature in Section “[Literature Review](#)”. Section “[Methodology](#)” described the empirical methodology, while in Section “[Data and Descriptive Statistics](#)” we describe the data. In Section “[Results](#)” we discuss the results and in Section “[Concluding Remarks](#)” we conclude.

Literature Review

Our paper relates to several strands of research. One strand of literature focuses on segmentation. Ghent (2018) assesses what makes an asset institutional quality and shows that delegated investors—those acting on behalf of their clients—invest in cities with high turnover. She argues that the reason for this is that those investors value liquidity more in the sense that they are not necessarily buy-and-hold investors and hence prefer to invest in more liquid markets. This heterogeneity in preferences for liquidity leads to market segmentation by investor type. This has a feedback effect on liquidity with market segmentation affecting the liquidity on the markets.

Ling et al. (2018) study the transaction prices of commercial real estate by comparing how those differ based on the distance of the investor from the property in question. They find that distant investors face higher search costs and information asymmetries as compared to local investors.

Related to this is the literature on how market segmentation affects prices on real estate markets. Clayton et al. (2009) talk about real estate markets being highly segmented in which mispricing can be observed which they contribute to sentiment and the lack of sophisticated traders. Fisher et al. (2009) investigate how institutional capital flows affect returns in private real estate markets. They find that unlike what theory predicts, capital flows do affect returns in those markets due to the high segmentation. They also argue that in addition to capital flows creating price pressure in segmented real estate markets, they may also affect investor expectations. That would mean that an increase in capital flows into a sector or market could act as a signal that investors are revising upwards their expectations regarding future income streams which may lead to higher asset prices. Beracha et al. (2018) find that hotel properties are segmented by hotel class arguing that the notion of aggregate pricing models bunching properties by their type would bias estimators.

Sagi (2017) and Badarinza et al. (2018) use search models to explain returns on the real estate markets. Sagi (2017) focuses on the property-specific part of the returns

as a large number of investors hold small portfolios and may be affected by idiosyncratic risks. So, for small investors, understanding what drives idiosyncratic returns is important. He develops a model in which he allows for investors to vary in their valuations of the income stream of properties and shows that returns are not generated by a random walk process which may be explained by the illiquidity of the asset. Piazzesi et al. ([Forthcoming](#)) also use a search model to study segmented housing markets due to heterogeneous clients. In particular they analyse the search behavior of buyers and how integrated different housing segments are. They observe that there are two types of searchers—narrow ones and broader ones, that span over several segments—which may be interested in areas with different levels of inventory. They find that the broad searchers spread local shocks across segments and reduce the effects of local market activity.

Geltner and van de Minne (2017) run a quantile regression to estimate price indices for different price segments. They aim to analyse if the law of one price in real estate markets which may be violated as different investors may be willing to pay different prices for the same product. They find that different price point properties have different return-risk dynamics and momentum.

Another strand of research related to the one above is the literature on style investing in stock markets. There is extensive literature that shows that small stocks are priced differently to large stocks and investors price size in. Another way to look at it is, given that in direct real estate there are large information asymmetries, large institutional investors may decide to adopt style investing if they cannot, or are not willing to, make decisions based on the fundamentals of the building and the location. Barberis and Shleifer (2003) argue that investors may allocate funds not at the asset level but at the style level, meaning the assets are first grouped into categories—styles—and then a decision is made based on relative past performance rather than on absolute performance. Froot and Teo (2008) show that retail and institutional investors indeed allocate capital at the style level and the way they classify assets is by using size and value-growth dimensions. Value investors would buy cheap stocks which are associated with higher risk. Growth investors would buy more expensive stocks which are perceived as less risky. Similar to the stock market, institutional investors in real estate may also follow a specific style. Real estate investors can be classified into core or stabilized investors versus opportunistic investors. Core investors prefer low-risk low-return property, meaning a property with stable NOI for which they are prepared to accept low cap rate. Such investors may be large investors in terms of portfolio size. Opportunistic investors may be smaller firms that would buy cheaper properties which have high cap rates and low NOI. While we do not look at the effect of prices, as is not the focus of this paper, the literature that shows how different types of investors affect pricing is relevant to explain the rationale.

Methodology

Our main objective is to estimate the probability that a certain investor size purchases a specific bundle of real estate characteristics. The size is based on the dollar value of the current domestic real estate portfolio which is presented in the form of a full

panel of real estate portfolio values for every year for every company operating in the US. Investor size is a continuous variable in itself, and a hedonic analysis could be done following (Rosen 1974). However, we choose to subdivide the continuous variable into multiple investor size quantiles (ISQs). We do this for multiple reasons. First of all, we are interested in probability distributions of different ISQs purchasing certain properties rather than point estimates. In other words, we are not interested in one linear prediction, but in the entire probability density function (pdf). From a modelling point of view, this also makes more intuitive sense, given that we are talking about preference or “choice” models.

Finally, given the high variance in investor sizes, we also expect considerable non-linearities. Breaking up our investor sizes allows us to analyze such non-linearities.²

We compute the quantiles as follows. Firstly, we demean our full panel of investor sizes for every year. Next—and again for every period separately—we number the investors (from the smallest portfolio to the largest) from $q = 1, 2, \dots, Q$. We follow Donald et al. (2000) in that we group the first 5% and the largest 5%, and use 1% increments in between. This results in 92 quantile buckets.

Then we proceed to estimate an *ordered* multinomial logistic model, which is also referred to as a proportional odds model. Such models are given by

$$\text{logit}(\gamma_{ij}) = \log \frac{\gamma_{ij}}{1 - \gamma_{ij}} = \theta_j + X_i \beta. \quad (1)$$

where property i has probability γ to be bought by ISQ j . In this formulation θ_j is a constant representing the baseline value of the transformed cumulative probability for ISQ j . X is a matrix of covariates for property i , and β is the corresponding vector of parameter estimates. It is custom to choose $\beta_1 = 0$ for identification purposes. The estimates of β are interpreted as odds ratios, that is how much does a characteristic of a property change the probability of ISQ $j \neq 1$ purchasing property i over ISQ $j = 1$ (in case of $\beta_{j=1} = 0$).

Exponentiating (1), we find that the odds of $Y_{ij} \leq j$, in other words, the odds of a response in category j or below, are

$$\frac{\gamma_{ij}}{1 - \gamma_{ij}} = \lambda_j \exp\{X_i \beta\}, \quad (2)$$

where $\lambda_j = \exp\{\theta_j\}$. The λ_j may be interpreted as the baseline odds of a response in ISQ j or below when $X = 0$. The effect of the covariates X is to raise or lower the odds of a response in category j or below by the factor $\exp\{X' \beta\}$. Note that the effect is a proportionate change in the odds of $Y_i \leq j$ for all response categories j . If a certain combination of covariate values doubles the odds of being in category 1, it also doubles the odds of being in category 2 or below, or in category 3 or below. Hence the name proportional-odds.

Proportional-odds models can be estimated using maximum likelihood as well. Interestingly, note that the representation of the proportional-odds model in Eq. 2 is very similar to a Cox-proportional hazard model (Cox and Reid 1987, 1993). The

²Also, see Donald et al. (2000) and Machado and Mata (2005) for further reading on such non-parametric models.

difference, or rather why the proportional-odds and hazard model are the same in this specific case, is that there is no right censoring. In the sense that all properties will get sold, because we have a transaction dataset.

Finally, we also add two sets of Gaussian random effects to our hazard model for every individual property and/or investor. In survival literature, this is also referred to as *frailty* (Ripatti and Palmgren 2000). We prefer random effects over fixed effects, because (1) approximately half of our properties got sold only once in our sample (meaning we cannot identify the fixed effect), and (2) many (mostly smaller) investors only invest in 1 market (meaning the market and investor fixed effect are perfectly co-linear). As a result, we would have to omit a lot of data if we would use property and investor fixed effects. It is well known from indexing literature that only including properties that sold multiple times in the data introduces a bias toward “winners”, or “better” properties, see Abraham and Schauman (1991), Shiller (1993), Clapp and Giaccotto (1999), Clapham et al. (2006), and Van de Minne et al. (2019) for more details on this bias. These random effects “capture” all (non time-varying) unobserved heterogeneity of properties and investors. As shown in Francke and Van de Minne (2020b), as long as a few properties transact more than once, one can identify the random effect for all transactions. Such Gaussian random effects are given by

$$\delta_i^k \sim N(0, \sigma_{\delta^k}^2), \tag{3}$$

where $k = \{\text{property ID, investor ID}\}$. Given that we have tens of thousands of individual properties and investors in our data we run in some computational challenges. To overcome these issues we replace the semi-parametric Cox baseline with a Weibull distribution. Also, we do not allow for correlation between the random effects. To increase computation time even further, we estimate these frailty models using LaPlace Approximation, a fast Bayesian procedure (Rue et al. 2009). More details on initialization and specifications are omitted here to conserve space, but are available upon request.

In addition to the role of size, we also assess the role of the type of investor for robustness purposes. The investor type depends on the institutional setting in which the investor operates. Those types include public and delegated investors, among others. For the modelling of the investor type one might assume some form of ordering of the data since it is categorical in nature. However, we use the *unordered* responses, in the sense that we do not assume that one category of investors is “better” compared to another. As such we use a multinomial logistic model. Suppose that property i has probability γ to be bought by type of investor j . The multinomial logistic model is now given by

$$\gamma_{ij} = \frac{\exp\{X_i \beta_j\}}{1 + \sum_{j=2}^J \exp\{X_i \beta_j\}}. \tag{4}$$

The model can be estimated by maximum likelihood.

Data and Descriptive Statistics

Our main data provider is Real Capital Analytics Inc (RCA). RCA is a data and analytics company that focuses particularly on the investment market for commercial real estate in the United States and various international markets. RCA has been collecting, analyzing, and interpreting comprehensive commercial real estate transaction information within all investment-grade strata of United States commercial property asset markets starting in 2000. RCA has a capture rate of over 90% for transactions of commercial “investable” real estate.³ As a result, the data should be close to the entire population of commercial real estate properties. Furthermore, RCA employs hundreds of researchers across the globe to find the ultimate investor of a property as accurately as possible. For example, RCA knows when a firm bought a property on behalf of someone else, or if a firm is “simply” a subsidiary / vehicle for a different firm. This allows us to assign property ownership to the “correct” firm.

In the next subsection we will discuss how we construct the US real estate portfolio sizes of all investors in the RCA dataset. In Section “[Data and Descriptive Statistics of the Other Variables](#)”, we discuss the descriptive statistics of the remaining variables used in this research.

Measuring Investor Size

The main variable of interest is the value of investors’ real estate portfolio. This allows us to construct the investor size quantiles (ISQ) of both buyers and sellers as explained in Section “[Methodology](#)”. We use two datasets (RCA, and RCA CPPIs) to construct this value. The RCA CPPIs (or Real Capital Analytics Commercial Property Price Indexes in full) are asset price indexes based on a structural time series repeat sales model given in Van de Minne et al. (2019). The RCA CPPIs are location and sector (i.e. property types) specific. In total we observe 80 of such RCA CPPIs.

First, we compute the average transaction price and transaction date for every unique property in the data. Next, we take the index corresponding to the market (location and property type specific) of the property, and assess the value of the property in any given year between 2001 and 2018. This gives us a full panel of all (appraised) property values between 2001 and 2018.

We also know from the RCA data who purchased and sold the property at time of sale of all the individual properties. This allows to find who owned which property between what time period. For example, if investor A sold a property in 2005 to investor B, we know that investor A owned the property before 2005 and investor B owned said property starting in 2005. There are some complexities though. For one, in approximately 5.5% of the transactions we find multiple investors (mostly designated “joint venture”), but we do not know the shares of ownership. In that case we give equal share of ownership to all investors. In some cases we also find that—for example—investor A bought a property in 2001, but investor B sold it in 2011. In this case we split the ownership in the middle, and assume investor B bought the

³In practice, RCA only includes real estate that once in its history sold for more than \$2.5 million.

property in 2006. In less than 4% of the cases did we find a missing transaction in the middle.

Figure 1 gives the average real estate portfolio size between 2002 and 2018, together with the 10th and 90th percentile of the portfolio sizes (grey shaded area). Average portfolio sizes first went up, from \$70 million to \$120 million, subsequently crashed and then after 2010, bounced back up again. In 2018 the average US real estate portfolio sizes are larger compared to its previous (pre-GFC) peak, at over \$160 million. Note that the data is skewed to the right, because there are a few very large companies that increase the mean. For example, the median is closer to \$10-15 million during this period. This is also visible from Fig. 2 by looking at the lower 10th percentile.

Table 1 gives summary statistics of the portfolio sizes by investor type. The first column gives the investor type. RCA has over 20 categories of investor types. For readability, we group them in the remainder of this paper into 4 larger groups (which are given in bold in Table 1); delegated, public, non-investor and direct. This categorization was used in previous literature as well, see for example Ghent (2018). Delegated investors include pension funds, equity funds, investment managers and banks. Public investor types include the real estate investment trusts (REITs) and real estate operating companies (REOCs). REITs and REOCs are under more scrutiny and are more transparent compared to the other firm types. In addition, these public companies have long holding periods by statute (Mühlhofer 2019), i.e. they make their money by collecting rents, and not by buying and selling of real estate. The non-investor group includes investor types finance, corporate, government, and non-profit. Finally, we

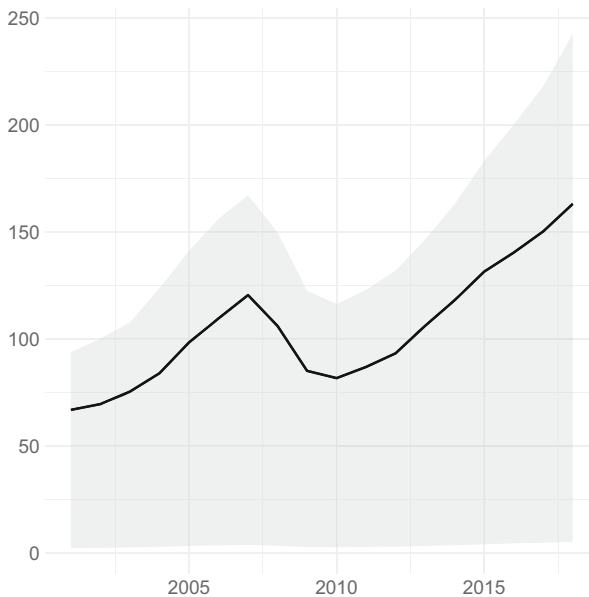


Fig. 1 Average value of Firm's US real estate portfolios in \$1M between 2002 and 2018. Grey shaded area in both plots represent the 10th and 90th percentile of said variable per year

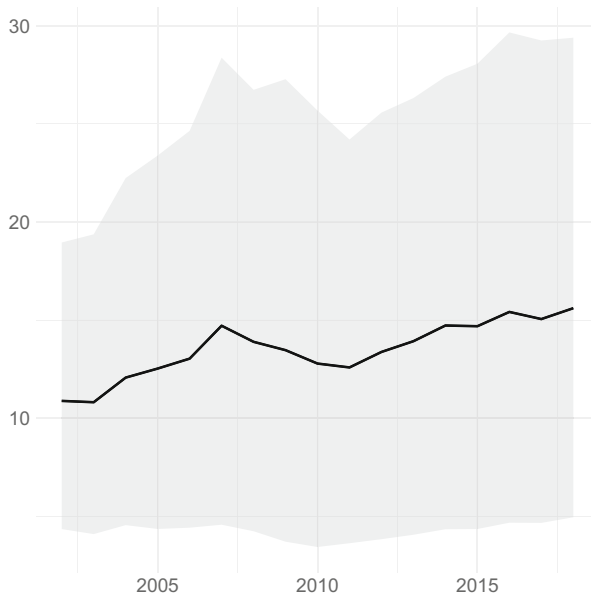


Fig. 2 Average Net-Operating-Income per square foot between 2002 and 2018. Grey shaded area in both plots represent the 10th and 90th percentile of said variable per year

have a group called “direct” which includes developer/owner/operators as well as all remaining investor types, like high net worth individuals and sovereign wealth funds (SWFs).

The second column of Table 1 gives how many firms of the corresponding investor type we identify in our dataset. The third column gives the average size of the firm’s US real estate portfolio per investor type, our main variable of interest. The fourth column provides the average value of the individual properties held by each investor type, and the final column gives how many properties they own on average (again in 2018). Values in different years are available upon request.

Note that the developer/owner/operator (DOO) investor type dominates the RCA dataset, and especially the “direct” group. Indeed, according to Table 1 the total value of all the commercial (investable) real estate is \$4.7T in 2018 based on the transactions in our sample. The DOO investors own 48% of that. A large proportion of those DOO investors adopt the Limited Liability Company (LLC) structure which is typical for private equity funds. Those specialized investors may include family offices and small developers that act as investors rather than speculative developers. Overall, we think that those investors adopt the mandate to invest in real estate for investment purposes and not for owner occupation. Therefore, we think that while DOO is a broad category, it actually encompasses investors that specialize in investing in direct real estate. Their assets would almost entirely consist of properties and their main focus would be real estate investment.

Our summary statistics show that, on average, SWFs have the biggest real estate portfolios. However, note that only 15 companies are designated SWFs in our data.

Table 1 Average investor size by investor type (RCA classification and our own in bold) in the year 2018

Investor type	Number of investors	Average value of portfolio	Average value of real estate	Number of properties
Delegated	1, 499	\$752, 540, 612	\$38, 535, 797	19.5
Pension Fund	91	\$1,563,425,522	\$ 63,316,298	24.7
Equity Fund	486	\$ 1,105,258,068	\$ 35,103,609	31.5
Investment Manager	528	\$ 740,159,903	\$ 40,393,223	18.3
Bank	394	\$ 146,768,542	\$ 28,221,965	5.2
Public	578	\$1, 448, 861, 132	\$28, 138, 898	51.5
Traded REIT	268	\$ 2,420,689,849	\$ 33,521,670	72.2
Non Traded REIT	132	\$ 928,560,340	\$ 17,178,692	54.1
REOC	178	\$ 371,499,382	\$ 20,203,755	18.4
Non Investor	2, 877	\$70, 204, 113	\$18, 080, 497	3.9
Finance	121	\$ 95,595,206	\$ 19,571,946	4.9
Corporate	1,979	\$ 80,656,489	\$ 18,419,016	4.4
Government	455	\$ 40,046,753	\$ 16,981,615	2.4
Non-Profit	322	\$ 39,036,490	\$ 14,946,195	2.6
Direct*	23, 793	\$105, 988, 429	\$21, 111, 971	5.0
Sovereign Wealth Fund	15	\$ 3,539,124,362	\$ 78,414,868	45.1
Insurance	125	\$ 797,077,974	\$ 58,608,675	13.6
Developer/Owner/Operator	22,406	\$ 100,303,746	\$ 19,992,934	5.0
High Net Worth	835	\$ 93,814,870	\$ 23,334,947	4.0
Educational	201	\$ 90,171,855	\$ 28,999,268	3.1
Religious	55	\$ 50,627,476	\$ 16,095,440	3.1

*The statistics of the investors in the “direct” group do not perfectly sum up to the group’s total. This is because we omitted smaller investor types from the table for readability, like endowments and CMBS investors

Their average real estate portfolio is worth over \$3B in 2018. They own 45 properties on average, which are worth almost \$80M each on average. This also means that on average SWFs own the most expensive real estate of all investor types. Non-profit organisations and governments are the smallest real estate owners on average. Their real estate is also the cheapest on average (together with religious buyer types) and have the least amount of real estate in general. Their portfolios are close to \$40M on average in 2018. REITs own on average the most properties, with 72 followed by the non traded REITs with an average of 54 properties per non-traded REIT. The DOO investors, even though large by summation, are relatively small per investor. The average property value they own is almost \$20M, and they “only” hold 5 properties on average in 2018, far less than for example (traded and non-traded) REITs and equity funds.

In order to get a feeling of how accurate our estimates of the investor sizes are, we will now take a detailed look at a selection of investors in the data. Table 2 gives the top 10 biggest real estate owners according to our estimates 2018. Unsurprisingly, number one is Blackstone with \$81 billion in (US) real estate. For comparison, according to their own reports, Blackstone had \$120B in assets under management (AuM) in real estate *world wide* in 2018. TIAA is the pension fund with the largest real estate exposure according to our estimates, with \$33B. We also find three REITs in our top 10, Simon, Vornado, and SL Green. Given that REITs focus their investments primarily on real estate, the firm's enterprise value should come close to the value of the underlying real estate (as valued by the stock market). As an extra robustness we therefore compare the (log) enterprise value of REITs taken from SNL Financial with our estimates of the (log) real estate portfolio value for REITs that we were able to match in Fig. 3. We match the REITs both, on firm and on an year level.

The overlap between our value estimates and the reported enterprise values is not perfect (Fig. 3), which might be caused by either stock market noise or inaccuracies in the RCA data, but overall there is a clear linear relationship between the two variables. It should be noted that our estimate of the real estate portfolio value is 10% less compared to the reported enterprise values of the REITs on average. This could indicate that REITs are overvalued (Van Nieuwerburgh 2019) or the other way around. However there are other explanations as well. First of all, REITs are mandated to invest *most* of its assets in real estate, but not all. More specifically, 75% of a REITs assets and income should come from real estate. Secondly, we know that private real estate returns typically lag public real estate returns. Reasons for this include informational inefficiencies and the slow (double-sided) search / negotiations in private real estate markets. As a result, previous literature found a lag of somewhere between 6 months to 1,5 years (Barkham and Geltner 1995; Francke and Van de Minne 2020a). Given that our analyzed time period has mostly been characterized by increasing prices, it *should* result in REITs having higher enterprise values compared to their contemporaneous private market valuations on average.

Table 2 Top 10 largest owners of US real estate in 2018

Company	Portfolio size	Type of investor
Blackstone	\$ 81,401,517,183	Equity Fund
Brookfield AM	\$ 79,811,188,929	Equity Fund
Simon Property Group	\$ 42,507,957,360	REIT
TIAA	\$ 32,794,867,587	Pension Fund
JP Morgan	\$ 32,740,438,261	Investment Manager
Vornado	\$ 31,789,241,472	REIT
Irvine Co	\$ 26,143,141,046	Developer/Owner/Operator
PGIM Real Estate	\$ 25,295,347,514	Insurance
Starwood Capital	\$ 23,926,530,059	Equity Fund
SL Green	\$ 23,312,834,837	REIT

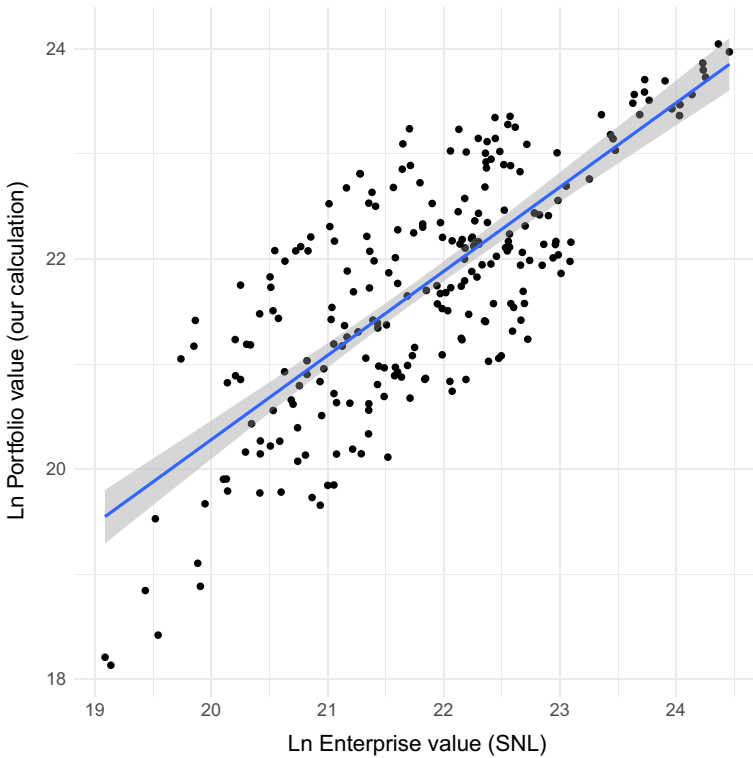


Fig. 3 Comparison of the (log) value of the real estate portfolios of REITs according to our calculations (y-axis) and REITs’ enterprise values taken from SNL (x-axis). Blue line represents a linear fit according to a linear regression model. Every dot represents a specific REIT in a specific year

Data and Descriptive Statistics of the Other Variables

Next we will discuss the transaction data, which is relevant for our analysis. Note that in Section “[Measuring Investor Size](#)” we appraised a full panel of the firm values. In contrast, the transaction data is a very unbalanced panel of property level transactions. Properties are only sold once every 7–8 years. Table 3 gives the summary statistics of our most important variables.

Note that the average investor size quantile (ISQ, both for buyers and sellers) is 46 (the mean is between 1 and 92). This indicates that larger investors transact more on average. The variables which we believe are important in order to predict preferences/segmentation are: (1) closeness to the Central Business District, (2) size of the property (in square feet), (3) the quality of the structure (measured by a “Q score”, see below), (4) age of the structure, (5) whether the property is green certified, (6) property type, (7) Net Operating Income of the property at time of sale, (8) market, and (9) year of sale.

Table 3 Main descriptive statistics of our data

	Mean	sd	Lower(10%)	Higher(90%)
Investor size quantile (buyer)	48.534	26.684	10.000	85.000
Investor size quantile (seller)	52.260	27.172	13.000	89.000
Distance to CBD (in km)	34.640	44.605	5.320	76.927
Square feet ($\times 1,000$)	197.623	227.007	25.920	409.624
Q score	1.518	0.276	1.130	1.890
Age (in years)	28.140	24.235	6.000	51.000
NOI per square foot	\$11.930	\$ 26.190	\$ 1.652	\$ 23.458
Selection of categorical variables				
Green	0.057			
Apartment	0.380			
Industrial	0.118			
Office	0.253			
Retail	0.249			

Investor size quantile = a quantile distribution from 1 (lowest) to 92 (highest) indicating the size of the domestic real estate portfolio of the investors in our data. The quantile score is relative within a year. *Q score* = a continuous indicator score between 1 (lowest) and 2 (highest) capturing “unobserved” quality of properties, essentially based on historic price per square foot compared to its direct peers. Exact method is proprietary to Real Capital Analytics. *NOI* = Net Operating Income of the property during the time of sale. Counts of sales in every year, per buyer type and per metro can be found in the [Appendix](#)

We show that the centroid of the *closest* Central Business District is approximately 35 km (or 22 miles) away on average. The CBD areas are provided to us by RCA. The average transacted property size is close to 200 thousand square feet, with an average age of 28 years. RCA also produces so-called “Q scores” for all properties. The exact methodology is proprietary to RCA, but it in essence is a quantile score between 1 and 2 based on the historic price per square foot as compared to its direct peers (which is sector and location specific). The average value is therefore unsurprisingly 1.5.

We get the NOI on the date of transaction from two sources. First, for about 35% of all transactions we observe the property’s NOI directly from the RCA transaction dataset. However, we also use mortgage data, also tracked by RCA. In many cases, whenever a property gets refinanced, an “underwritten NOI” is reported. Typically, there are a few years between the latest refinancing and the sale. We therefore change the NOI using “NOI indexes” from RCA. The NOI indexes are market (i.e. location and property type) specific. In total we observe approximately 300 of such NOI indexes in the US. This allows us to map an NOI at transaction for 48% of the entire dataset. NOI is important, because extant literature (Geltner et al. 2014, p. 554–556) shows idiosyncratic price movement specific to individual assets or granular market segments, largely reflects the condition of rental (space) markets, as proxied by the NOI. At the same time, asset-valuation risk reflects changes over time in the capital market that cause changes in the opportunity cost of capital. Hence, time variation in

the discount rate causes at least as much volatility in prices (see for example Geltner and Mei 1995).⁴ However, the opportunity cost of capital is highly correlated over time across space markets. Mostly because the risk-free rate, which drives the cost of capital to a large extent, is identical no matter where the investment is made. The other determinant of the opportunity cost of capital—the risk premium—can differ per market, or even property, but rarely changes over time (Geltner and Mei 1995; Geltner et al. 2014). In further analysis we therefore also use time and location dummies to account for such offsets.

We find that approximately 6% of all properties in our transaction dataset are green certified. The apartment property type was transacted the most with 38%, and “only” 12% are industrial properties. In the appendix we also provide the distributions over year of sale (Table 9), location (Table 10) and investor type (Table 11). Given the often found positive price-volume correlation in real estate (van Dijk et al. 2019; De Wit et al. 2013), it is not surprising to find that the smallest percentage of transactions was during the GFC, in 2008–2010 (Table 9). Transaction volume is now higher compared to the previous, pre-GFC peak in 2007. Real Capital Analytics defines 134 different metro areas, see Table 10. Los Angeles is by far the largest market volume-wise, with 9.5% of all transactions being in said metro area. New York is second with over 6% of all transactions. In Table 11 we split the investors by buyer and seller. The “direct” investor group is again (see Table 1) the largest, being responsible for more than half of all the transactions.⁵ The “non-investor” group is the smallest investor type in our data. Finally, we find that approximately 10% of all transactions are done by a foreign investor.

Since the variation of the buyer size is our main variable of interest, we will explore this variable in more detail. First of all, if indeed specific properties are only bought by specific sized investors, we should find more variation in investor size in the data as a whole, than within the individual properties. To confirm this, we first measure the standard deviation of the ISQ within each property ID (if they are repeated observations). We subsequently order the standard deviations per property from lowest to highest. Some percentiles based on this ordering can be found in the second column of Table 4. We also subtract the average standard deviation of the ISQ of the entire dataset (26.7, see Table 3) from these percentiles in the third column. For completeness, we redo this exercise with the within investor ID standard deviation of the ISQ variable. Both property ID and investor ID will also be added as random effects in the frailty models.

Table 4 confirms that properties tend to be sold to similar sized ISQ investors. Only at the 83rd percentile is the within property σ^{ISQ} larger compared to the data as a whole. Within the investor IDs the distribution is even less equal. Only at the 99nd percentile is the within investor σ^{ISQ} larger compared to the data as a whole. In other words, once an investor is designated large, it will remain relatively large, and

⁴This phenomenon is also widely accepted in the stock market (Shiller 1981).

⁵It is best practice to define volume as the amount of transactions (reported here) *times* the sales prices. Note that on average, the “direct” investors purchase cheaper properties (Table 1). If we multiply the average transaction price with the amount of properties sold, the share of direct investors is closer to 40%.

Table 4 Variation of Investor sizes within sub-groups

Percentiles	σ^{ISQ}		σ^{ISQ}	
	Within Property ID	Difference within Property ID and full dataset	Within Investor ID	Difference within Investor ID and full dataset
10%	2.121	-24.562	0.000	-26.684
25%	6.364	-20.320	2.121	-24.562
50%	14.119	-12.565	5.433	-21.251
75%	22.627	-4.056	9.195	-17.488
90%	31.820	5.136	14.009	-12.675
Unique obs	29,788		10,251	

ISQ = Investor Quantile Distribution; a quantile distribution from 1 (lowest) to 92 (highest) indicating the size of the domestic real estate portfolio of the investors in our data. The quantile score is relative within a year

vice versa. Finally, note the final row of Table 4. Here we give the unique amount of properties and investors. We have 44 thousand observations in total and 30 thousand unique properties. Thus, most transactions are single sales. On average every investor was involved with over 4 transactions in the data.

Results

Main Results

Here we discuss the results of our baseline hazard model which can be interpreted as an ordered multinomial logit model looking at the size of the investors. Our baseline is the yearly de-measured quantile distribution of size of the buyers' real estate portfolio (ISQ_{buyer}). As discussed in previous sections, this is a variable which we constructed using the RCA transaction data and is unique in its nature. The explanatory variables are object characteristics, like size of the property and NOI, but also seller types. We have four different specifications. In the first (I) model we do not include any variables on the sellers. The second (II) model includes the investor size quantile, and model (III) introduces the seller types. We also include a dummy on whether the seller is a foreign investor or not. The *ISQ* score enters in the model with a log transformation. In the fourth (IV) model we enter the *ISQ*-score non-barometrically, i.e. we take up a separate dummy for every individual *ISQ* ($q = 1, 2, \dots, 92$) score. The estimates are expected to be somewhat noisy, however we have enough degrees of freedom to make inferences from the results. All models have year and metro level fixed effects. The results are given in Table 5.

It should be noted that the interpretation of these results can be less than straightforward, other than the sign and significance. A negative parameter estimates, means

Table 5 Results of our main hazard models. The baseline is the investor size quantile of the buyer (ISQ_{buyer})

	(I)	(II)	(III)	(IV)
A. Main estimates				
ln ISQ _{seller}		-0.254***	-0.246***	Fig. 8
ln dist to CBD	-0.002	0.007	0.008	0.010*
ln structure size	-0.418***	-0.338***	-0.336***	-0.338***
Green	-0.058**	-0.054**	-0.054**	-0.041*
ln Q score	-0.973***	-0.847***	-0.872***	-0.872***
ln age	0.076***	0.085***	0.087***	0.089***
ln NOI per square foot	-0.233***	-0.200***	-0.197***	-0.198***
B. Property type fixed effects				
Apartment (<i>ref</i>)	0.000	0.000	0.000	0.000
Industrial	-0.437***	-0.398***	-0.386***	-0.394***
Office	-0.132***	-0.111***	-0.100***	-0.103***
Retail	-0.415***	-0.376***	-0.362***	-0.353***
C. Seller type fixed effects				
Delegated (<i>ref</i>)			0.000	0.000
Direct			0.034***	-0.049***
Non Investor			-0.048**	-0.115***
Public			-0.053***	-0.024
Foreign			-0.068***	-0.101***
D. Other fixed effects				
Year fixed effects	✓	✓	✓	✓
Metro fixed effects	✓	✓	✓	✓
E. Diagnostics				
loglik	-420,145	-419,206	-419,173	-418,822
AIC	840,607	838,730	838,672	838,150
Concordance	0.688	0.698	0.699	0.700
Wald test	11,163	14,131	14,119	14,278
N	43,928	43,928	43,928	43,928

*** = significant within the 1% level. ** = significant within the 5% level. * = significant within the 10% level

The Wald test indicates whether the final coefficients differ from the initial values. *ref* = reference category, CBD = closest Central Business District, as defined by Real Capital Analytics. NOI = Net-Operating-Income

that larger (ISQ) investors are more likely to purchase the property, and the other way around. For example, the parameter estimate on the (log) size of the structure is negative and statistically significant in all model specifications, meaning that the larger the property, the larger the buyer on average. In contrast, older properties are more likely to be bought by smaller investors, all things equal as can be seen by the

positive and significant effect of the log of the age of the property. Most parameters have the expected sign. Green properties, high Q score properties, and high NOI per square foot properties are more likely to be bought by largest investors. The distance to the closest CBD is mostly insignificant, or barely significant. This might seem surprising, but note that we control for metro fixed effects and property level NOI. Especially the latter is expected to be somewhat colinear with distance to CBD (i.e. properties further away from the CBD will have lower NOI per square foot). Apartment (industrial) properties are sold to the smallest (largest) investors, all things equal. Finally, even after controlling for the type of seller, we find that the size of the seller (ISQ_{seller}) has a significant impact on the buyer size. In all models we find that the larger sellers, tend to sell to larger buyers.

Again, even though the signs of the parameters are interesting in itself, it is difficult to get an intuition on the actual magnitude of the impact of said variables. We will therefore visualize the change in baseline (i.e. the probability distribution function (pdf)) for a specific set of linear combinations of covariates. More specifically, we can make a “representative” property (Diewert et al. 2011; de Haan and Diewert 2011). This “average” property can serve as an overall benchmark. It has the average characteristics of the sample of all properties in the RCA database. Its values can be found in Table 3. This property is a mix of 38% apartment, 25% office, 25% retail and 11% industrial, it is 200,000 square feet, 28 years old, has an NOI of \$12 thousand per square foot, and is 34 km away from the CBD. We subsequently run said covariates through our model estimates to get a pdf of buyers of that representative property. The pdf of buyers for Models (I) through (IV) is given in Fig. 4.⁶

Even though the models include very different covariates, on average the pdfs are indistinguishable from each other. Indeed, in all cases small investors have relative small probability of purchasing this “average” benchmark property, but the same goes for the very large investors to the right of the pdf. The investor quantile most likely to purchase this average property has an ISQ score around the 50 mark for all models.

In the remainder of this section we take this benchmark property, but change a few property characteristics to see how much the pdf changes/skews. For example, we can take a few existing properties within the RCA dataset, and view the estimated probability to be sold to any investor size. We also include an online Appendix to this paper where one can simulate how changes to the covariates impact on the distribution of investors based on the size. Table 6 gives three examples of actual transactions in the RCA dataset. All estimates are based on Model (I), as for now we are not yet interested in sellers’ characteristics.

The first property is an old office property in the suburbs of Los Angeles. The property is up for redevelopment, and has no NOI as of the transaction date. Its a relatively small property (51 thousand square feet) and has the lowest Q score possible. Our model overwhelmingly predicts that a small investor would be interested in this property. The actual buyer was of the lowest category, i.e. the $ISQ_{buyer} = 1$. The

⁶These estimates can be a bit “noisy”. We therefore smooth the line using a Kalman filter with a local linear trend (Francke and van de Minne 2017). After the smoothing we re-weight the points so they sum up to one again.

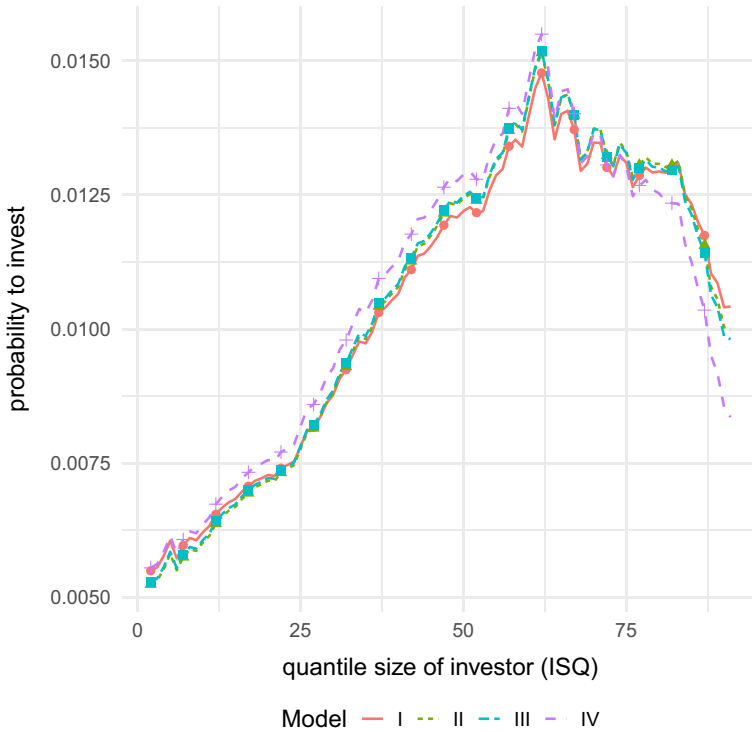


Fig. 4 Probability density function of buyers’ ISQ for the “benchmark” property for all our Models. This property has the average characteristics of all properties in the RCA dataset. The horizontal axis is the ISQ score of the buyers, and the vertical axis is the probability that a specific ISQ will purchase the property. These probabilities always add up to 1

middle panel is an apartment complex in Phoenix. The property has average scores throughout, and was bought (in 2006) by a mid-tier investor with an ISQ of 47. Our model also predicts that investors with a median portfolio size have the largest probability of purchasing this property. Finally, the third panel gives an industrial property, with a relatively low NOI per square foot. The property is younger than average, and has a Q score that is slightly above average with 1.71. However, our model predicts that only large investors are interested in this property. This is mostly caused by the fact that this is one of the largest properties in the dataset with 1.2M square feet. Above results can be explained in several ways. One the one hand, the findings are indicative of potential financing constraints as the small-sized buildings are predominantly bought by small investors and the large buildings by large investors. Small investors may face financial constraints to buy large and hence more expensive⁷ buildings. On the other hand, investors may prefer to buy certain buildings as

⁷Expensive means buildings with low cap rates and high transaction prices. While this paper does not look at prices as such, we use this term rather descriptively. Buildings with high NOI are also classified as expensive.

they choose to adopt different investment styles. For example, small investors may adopt a value-add or opportunistic strategy by buying properties with low Q score and low NOI, then refurbish those and benefit from the capital gains. Large investors instead may be interested in the so called ‘core’ buildings which tend to be large and more expensive. They may be more interested in stable cash flows instead.


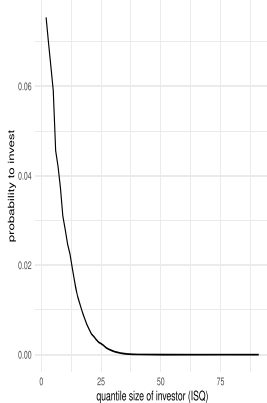

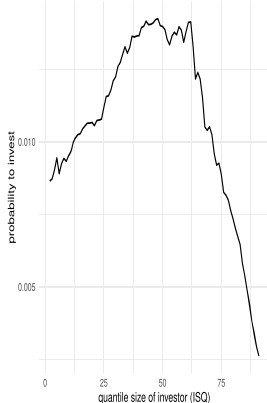

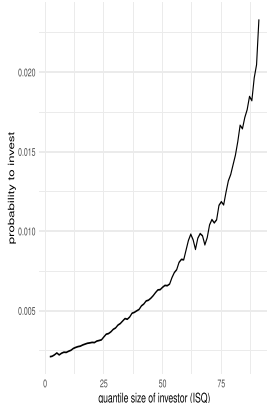
A useful outcome of above exercise is that we are able to pin down what a core, a value-add and an opportunistic building may look like if we want to keep working with the common terminology. The property characteristics associated with core buildings may vary depending on the sorting of investors across properties. Looking at which properties get bought by the top end of the pdf—the large players—can help us identify what core is for those investors. One example is the real estate displayed at the bottom panel in Table 6. In our example, a very large young industrial building with high quality and high overall NOI. A value-add property may look like the real estate displayed in the middle panel. And an opportunistic—at the top panel. The latter is a small old office building with very low NOI and the lowest possible quality. It is still located in a major city—Los Angeles—but it is away from the CBD.

In Fig. 5 we keep the average representative property, as given in Fig. 4, however we change two characteristics: (1) NOI per square foot, (2) and square footage itself. More specifically, we look how investors sort themselves across properties varying the combination of low/high NOI and small/large structures. The combination of low and small is designated to be at the 15th-percentile of investor size. That is associated with \$5.09 NOI per square foot, and 39 thousand square feet. That means that small investors would buy the small properties with low NOI per square foot all else equal. The high/large property is designated to be bought by the 75th-percentile ISQ. That is associated with an NOI per square foot of \$16.26 and 258 thousand square feet. From Fig. 5 we can clearly see that the largest “jump” comes from the building size. Even at a low NOI per square foot, larger investors still prefer larger properties. It does drop off a little for the very largest investors, although it should be noted that the total amount of NOI (NOI per square foot times the square footage) is twice as high for the combination of low NOI and large building (\$1.3M), compared to the high NOI and small building combination (\$0.6M). That means that large buildings have larger NOIs per square foot than smaller buildings.

Next we are interested in some of the time-varying effects associated with buyer preferences. For this we use our benchmark property again but we change two variables. The first is the NOI per square foot. This is a variable that is affected by the economic environment of that year. We use the means that are shown in Fig. 2. The second variable we change are the year dummies, to correspond with the year of interest. By doing this, we can see if the preference for the same property changes over time, by holding the other characteristics of said property fixed. We are specifically interested in the GFC period which spans from 2007 to 2009, for reasons we will show below. Because we are still not interested in the seller characteristics, we show the results of Model (I), although other models remain very similar and are available upon request. The results for a selection of years are given in Fig. 6.⁸

⁸Note again that we demean the full panel of investors in the first step in Section “Data and Descriptive Statistics” but we do not demean the ISQ scores in the transaction data. This is important because otherwise

Table 6 Three different properties and its probability to be sold to different investors sizes (ISQ_{buyer}). All estimates are based on Model (1)

Picture	Description	Hazard (pdf)
	<ul style="list-style-type: none"> • 100 E Nutwood St., Los Angeles • Office • year sold; 2011 • size; 51,841 square feet • Q score; 1.01 • NOI; \$ 0.0M • age; 59 years • distance to CBD; 13km • $ISQ_{buyer} = 1$ 	
	<ul style="list-style-type: none"> • 15050 N 59th Ave., Phoenix • Apartment • year sold; 2006 • size; 137,725 square feet • Q score; 1.44 • NOI; \$1.6M • age; 21 years • distance to CBD; 22km • $ISQ_{buyer} = 47$ 	
	<ul style="list-style-type: none"> • 1 True Temper Dr, Harrisburg PA • Industrial (logistics) • year sold; 2016 • size; 1,226,525 square feet • Q score; 1.71 • NOI; \$ 6.4M • 17 years • distance to CBD; 31km • $ISQ_{buyer} = 92$ 	

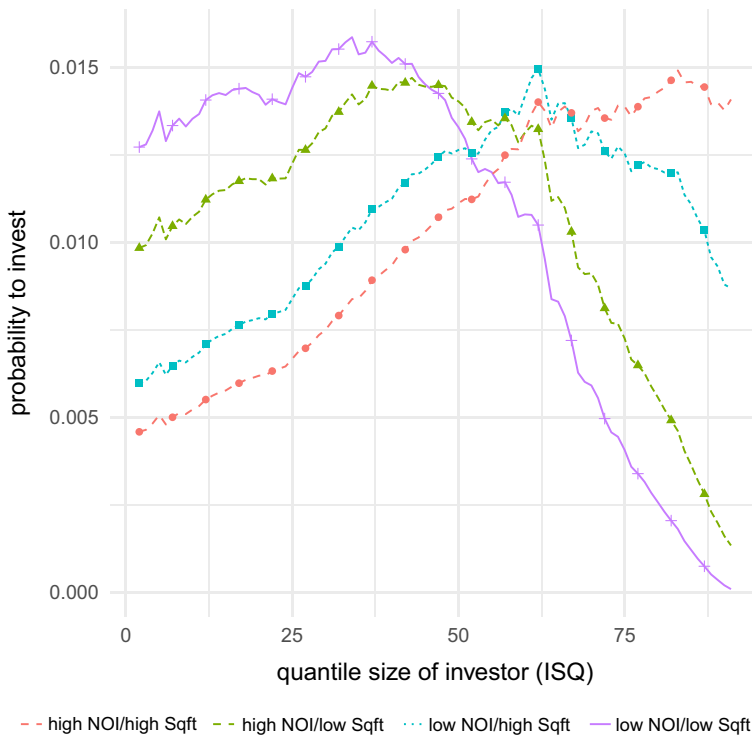


Fig. 5 Probability density function of buyers' ISQ for the “benchmark” property, using the average characteristics as found in Table 3. The only variable we change are the Net Operating Income (NOI) per square foot and the square footage itself (Sqft). Low (high) means its the variable at the 15% (75%) percentile as found in the data, ordered from low to high values. All estimates are based of Model (I)

What is interesting about Fig. 6 is that most of the changes of the distribution of buyer ISQs during the GFC is only affecting the higher ISQ investors. Note that the pdfs always sum to one. Thus, one cannot say that the probability of sale changed from one year to the next, as all results are to be interpreted within the confines of the given year. That being said, it is evident that the larger investors “pulled out” of the real estate market, compared to other investor sizes. Instead, the probability that the average property was bought by lower and mid-tier investors increased. The opposite picture is observed just prior to the crisis, with the largest investors being the most likely to purchase the average property. In 2009, the probability that a large investor would purchase the average property was at a historic low. (Since then, the probability has gone up again, but not as high as the immediate pre-GFC levels.)

Above observations are somewhat counter-intuitive as one would expect that the large investors have the least credit constraints and in a crisis period would be the

the mean of the ISQ scores per year would always be $(92/2 =) 46$. However, we might demean the full panel of investors, but if only smaller investors purchase properties, we will find a low ISQ score on average for that year. That is why we can track changes in investor sizes over the years.

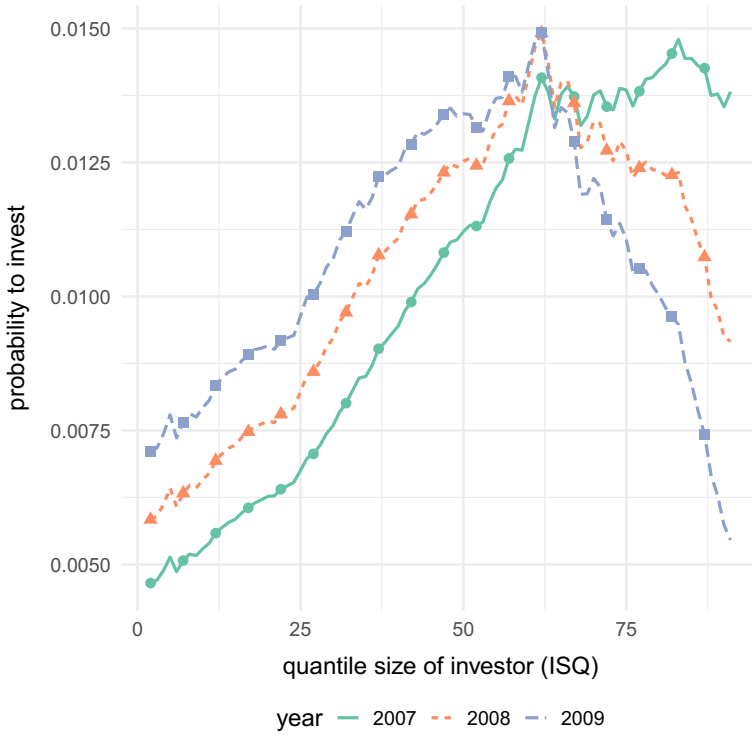


Fig. 6 Probability density function of buyers’ ISQ. Property is the “average” benchmark property, using the average characteristics as found in Table 3. The only variables that are changed are the NOI per square foot, see to Fig. 2, and the time dummies to the corresponding year. The results are for three years during the Great Financial Crisis, 2007 (green) 2008 (orange) and 2009 (grey). All estimates are based of Model (I)

predominant player on the markets. Clearly, they were the ones that actually have reduced the likelihood of buying the benchmark property. That finding suggests that they seem to be the most responsive ones to an economic downturn by deciding not to trade on the real estate market. An economic downturn may therefore be associated with the largest investors more strongly lowering trading activity than small and medium investors. This can be because large players wait for the real estate market to recover before they start buying or selling. They may be the least flexible to manoeuvre the GFC market and find good deals. Alternatively, those investors may struggle to justify a transaction during the GFC for their shareholders or lenders. This is because large investors (1) are subject to more stringent regulatory requirements, (2) may be listed on the stock exchange and subject to more scrutiny from shareholders, (3) or/and have strong dependence on issuing bonds and hence be wary of their credit ratings. Pension funds and insurance companies may have been affected by the introduction of new financial market regulations in the US and in Europe (i.e. Dott Frank, Basel III, Solvency II, etc.). REITs may face devaluation of their assets and as a result may find it harder to refinance and prefer not to opt for new stock issuance.

Independently from which channel dominates, the conclusion is that being a large investor and potentially having established good and cheap financing structures does not result in increased or even a continued trading activity in a downturn. This may be associated with a shock to liquidity for those large players. Smaller players may face financing constraints and illiquidity too but it seems they are more agile during a crisis and keep trading activity going. That can be linked to style investing—small investors look for opportunistic properties and the possibility for such deals increases in a downturn. The fact that the small players continue to trade in the GFC suggests against the presence of a financing constraint channel that explains the pdf of ISQs.

Above findings suggest that the potential effects of a change of the market composition of buyers and sellers may lead to overall lower prices in addition to any other downward pressure on prices during the GFC. That means that the presence of large investors and their diminishing likelihood to transact during a downturn and their higher likelihood to transact in a growing market can exaggerate the real estate cycle. The presence of small players may have the opposite effect on the real estate market and can help smooth the cycle.

This is also buttressed by looking at the expected buyer sizes per state for a selection of years, see Fig. 7. The expected buyer size is the sum product between the probabilities and the ISQ score of the pdfs. To construct Fig. 7 we first construct an

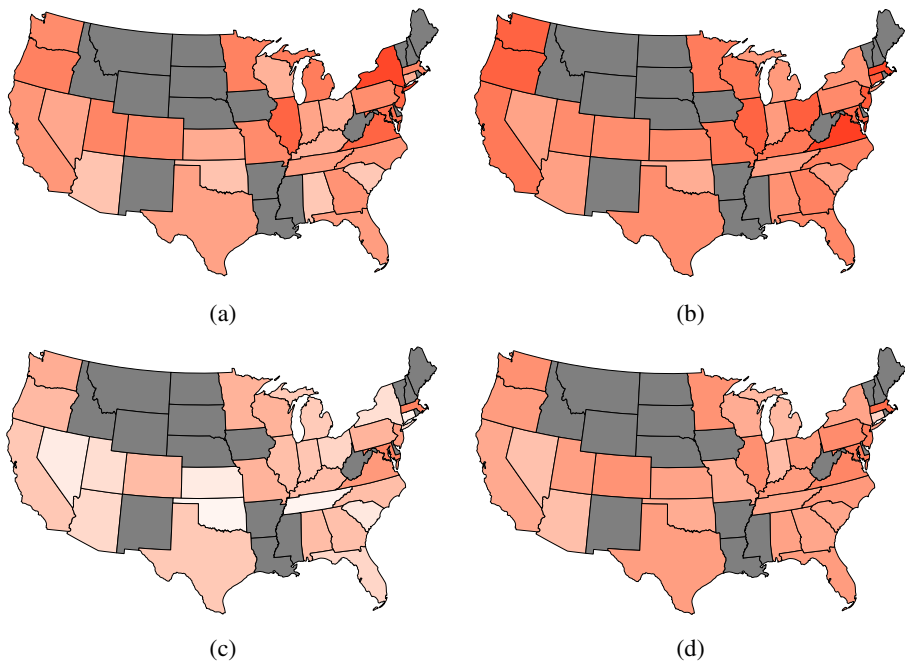


Fig. 7 Average expected size of the buyer based on Model (I) per state for a selection of years. The darker the color, the more probable a larger investor will be the buyer. Grey states include those states for which we have less than 10 observations per year. The representative property is the “average” benchmark property per state. We subsequently only change the average NOI per square foot and the time dummy

average benchmark property per state. The average structure size, property type distribution, etc., is thus different for every state. Next, we change the average NOI per square foot **per state** per year, and the year of sale dummy. These estimates are still based on Model (I). States with less than 10 observations per year are omitted, and are given in grey.

We find that in 2017, the distribution of expected buyer sizes across states is approximately similar to that in 2003, albeit a bit less concentrated in New York and Chicago (Illinois). In fact, the largest investors in 2017 can be found in DC and Massachusetts (dominated by Boston metro). Just pre-GFC in 2007, we find that the average size of investors increased. It did so quite in-discriminatory. In other words, larger investors become more active in almost every market. Two of the exceptions are the states of Michigan and Utah (dominated by Detroit and Salt Lake City respectively) where the average investor size actually went down. More interestingly, the state of New York (dominated by New York City) also saw decreases in average investor sizes between 2003 and 2017. States that saw its investor sizes increase the most between 2003–2007 were Wyoming, Ohio, Kentucky, and Alabama. These are states that are not typically known to attract large institutional investors. In the build-up to the GFC hence there seems to have been a shift of large players moving away from the more established commercial real estate markets into states with previously less institutional grade properties. That may have been due to a search for yield and better investment opportunities. At the end of the GFC, large institutional investors reduced their real estate investments in most states. In fact, none of the states had an average increase in institutional investor sizes. The largest drop in investor size between 2007–2009 was in Connecticut, followed by Ohio. Large institutional investors found their way back to real estate in 2017, although not equally in every state. States that only saw a small increase (none of the states had a decrease between 2009–2017) were Wyoming, Indiana, Missouri. In levels, the smallest investors can be found in South Carolina in 2017.

Next, we are interested in the size of the sellers' US real estate portfolios (ISQ_{seller}). In models (II) and (III) we include the log of the ISQ of the seller. In model (IV) we use a dummy for every of the 92 ISQ categories of the sellers instead. The estimates of the dummies are graphically given in Fig. 8. Note again in Table 5 that the estimates on the covariates hardly change between the models, so earlier results still hold for these models as well. Here we are mostly going to discuss the added variables on the sellers.

For models (II) and (III) we both find a negative effect of the log of the sellers' ISQ on the buyers' ISQ, meaning larger sellers sell to larger buyers, holding all things equal. Model (IV) gives a similar image. Even though the estimates can be a bit "noisy" at times, overall it is clear that the larger the sellers' ISQ, the larger the buyers' ISQ. If there would be no preference to trade with a specific size of investor, these dummies should be insignificantly different from zero. The results are very robust.

Next we compare the effect the sellers' size has on the probability to sell it to a buyer of a certain ISQ. For this we again use the benchmark property. The only

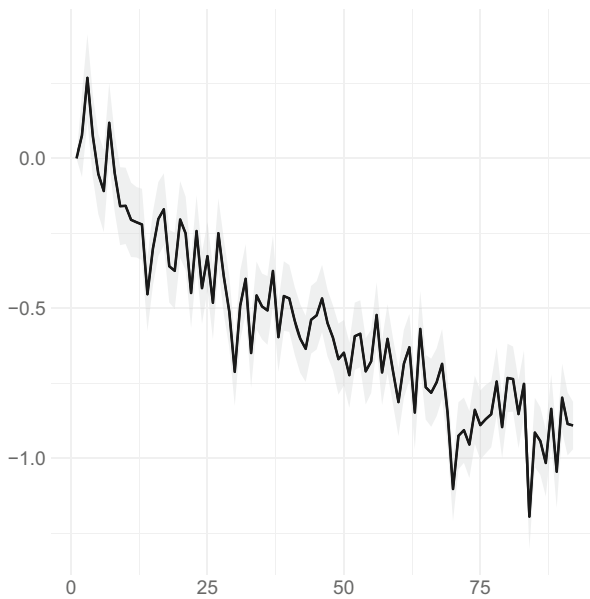
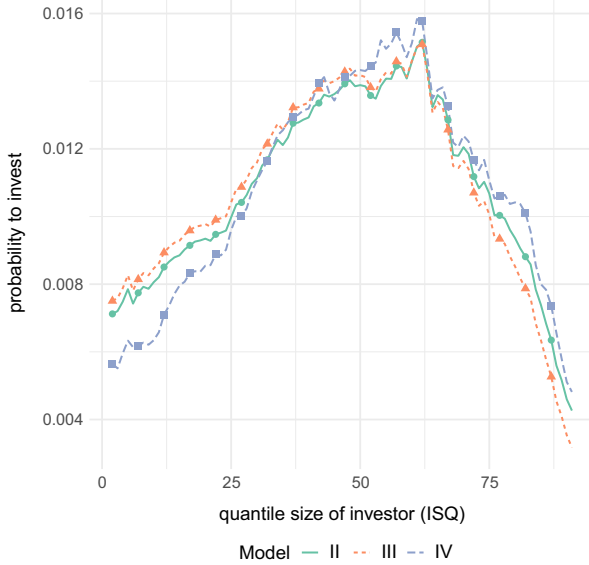


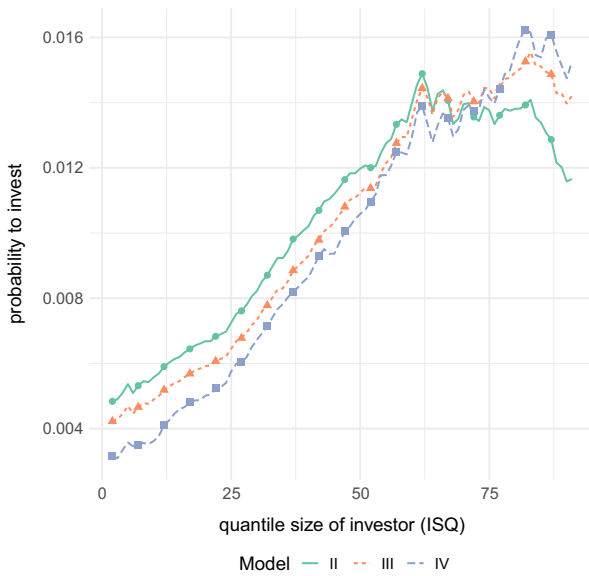
Fig. 8 Non parametric dummy estimates of ISQ seller quantiles, $ISQ_{seller=1}$ is the reference group and fixed at zero. Grey bars give the 95% confidence bounds. The estimates come from Model (IV)

parameters we change are the sellers' ISQ. More specifically, we are going to compare the buyers ISQ pdf when the sellers' ISQ is **15** versus **75**. We do this for all the models (II through IV) that include the ISQ_{seller} variable.

Even though the parametrization is very different between the models (two include a log transformed ISQ score, and in the other ISQ is entered non-parametrically), the results are very similar. The results are presented graphically in Fig. 9. In all cases we find that if the seller has an ISQ of 15, i.e. the seller is small, it is likely to sell to a smaller investor on average. The average ISQ_{buyer} is expected to be 44. This is compared to when the seller is large, i.e. they have an ISQ of 75, in which case the average expected ISQ_{buyer} is approximately 50 on average. Again note that it is the same property in both cases, and the same investor type. The main difference is—similar to our GFC analysis in Fig. 6—in the high ISQ buyers. The probability that a smaller investor will sell the *same* property to a large investor is relatively slim. Large investors overwhelmingly sell to other large investors, irrespective of the type of investor. These findings align with what has currently been observed in the industry where investors within the same size and background sell to each other. Above findings suggest that the real estate market is segmented and investors trade with each other in separate brackets depending on their size. That may be due to established networks within a given investor sub-market and familiarity with the players. Large sellers are mostly selling the average property to large buyers while small sellers can sell to a range of buyers but the likelihood to sell to the biggest buyers is very small. It is important to keep in mind that in each of those scenarios, we are dealing with the same property and not with larger or smaller properties depending on the size of



(a)



(b)

Fig. 9 Pdf of expected buyer ISQ. Property is the “average” benchmark property, using the average characteristics as found in Table 3. The only variable that is changed is the sellers’ ISQ

the seller. That may suggest that the large sellers mostly sell properties within their network and that the network effect is stronger for the very large investors.

Finally, we will shortly discuss the model diagnostics. The differences in the concordance score (Harrell et al. 1996) between models (I) through (IV) are negligible although the likelihood and AIC do improve after introducing sellers characteristics.⁹ The AIC prefers the third model, but the likelihood the fourth model. The Wald test shows that all parameter estimates are different from their starting value.

Robustness: Frailty and Prior Exposure

Next we will discuss two sets of auxiliary results. First we will discuss our hazard models with frailty for individual properties and investors. These results can be found in Table 7. In Model (V) we only include the property level random effects, in Model (VI) only the investor level random effect, and Model (VI) contains both random effects.

Only including the random effects for the properties does not alter the results erratically. It seems in general that there is not much unobserved heterogeneity with this set of covariates, evident from the low standard deviation on the random effects ($\sigma_{k=\text{property id}} = 0.006$, see Table 7). Only the green certification and the dummy for direct investor seller type become insignificant. Including the investor random effects does attenuate most of the results. For example, none of the seller type variables are significant anymore. Also the effect of Q-score reduces, but is still highly significant. The coefficients on property types also reduce considerably. This can be explained by the fact that most investors tend to focus on one sector, thus the investor random effect effectively controls implicitly for property types. The same argument can be used to explain why the age variable becomes insignificant; Investors tend to focus on specific vintages. Interestingly, the effect of NOI per square foot and the size of the structure remain very robust. The effect of the prior investor size (ISQ_{seller}) stays highly significant, but is attenuated as well. All of this is very encouraging, given the low amount of variation we found within investors' ISQ ([Data and Descriptive Statistics of the Other Variables](#)) and the fact that the standard deviation on the investors random effects is quite large ($\sigma_{k=\text{investor id}} = 3.2$ and 3.5 , see Table 7) for both models with the investor random effects. Finally, the fit—according to the marginal log likelihood—is best for the model with both random effects included.

Next, we decide to swap out our dependent variable—ISQ—with a variable for how much prior exposure the investor had in a specific market before the transaction. We—again - use the 80 market definitions provided to us by RCA (which are location and sector specific), and compute the dollar value the investor already had in said market prior to the transaction, as a percentage of the total value of the investors' real estate portfolio. As with the ISQ, we make quantiles of this variable by grouping investors who had less than 5% of their portfolio in a market and 95% or more in a market, and with 1% increments in between, resulting in—again—92 quantiles.

⁹The concordance test gives the probability that the prediction x goes in the same direction as the actual data y . The value will therefore be between 0 and 1, with 1 being the best fit possible.

Table 7 Results of our frailty hazard models. The baseline is the investor size quantile of the buyer (ISQ_{buyer})

	(V)	(VI)	(VII)
A. Main estimates			
ln ISQ _{seller}	-0.234***	-0.123***	-0.114***
ln dist to CBD	-0.001	-0.007	-0.007
ln structure size	-0.291***	-0.251***	-0.233***
Green	-0.019	0.033	0.031
ln Q score	-0.711***	-0.243***	-0.222***
ln age	0.062***	0.007	0.006
ln NOI per square foot	-0.165***	-0.190***	-0.180***
B. Property type fixed effects			
Apartment (<i>ref</i>)	0.000	0.000	0.000
Industrial	-0.301***	-0.074**	-0.067**
Office	-0.087***	-0.081***	-0.073***
Retail	-0.242***	-0.030	-0.029
C. Seller type fixed effects			
Delegated (<i>ref</i>)	0.000	0.000	0.000
Direct	0.019	-0.003	-0.003
Non Investor	-0.088***	0.008	0.008
Public	-0.035**	-0.003	-0.003
Foreign	-0.034*	-0.045**	-0.042*
D. Other fixed effects			
Year fixed effects	✓	✓	✓
Metro fixed effects	✓	✓	✓
E. Diagnostics			
m. loglik	-203,828	-177,566	-177,360
α	1.850	4.458	4.359
$\sigma_{k=\text{property id}}$	0.006	-	0.006
$\sigma_{k=\text{investor id}}$	-	3.188	3.463
N	43,928	43,928	43,928

*** = significant within the 1% level. ** = significant within the 5% level. * = significant within the 10% level

m. loglik is that marginal log likelihood. Parameter α gives the shape of the Weibull distribution of the baseline. *ref* = reference category, CBD = closest Central Business District, as defined by Real Capital Analytics. NOI = Net-Operating-Income

We call this variable PEX. This variable is of interest to us, as it highly relates to the ISQ variable. The downside is that it is heavily endogenous, and we therefore cannot simply add it to our main specifications. Indeed, only larger investors have the capacity to diversify. For this reasons we estimate a separate model and see if our findings are consistent. The results are given in Table 8.

Table 8 Results of our main hazard models. The baseline is the prior exposure (as a % of its entire portfolio) in the market where the transaction takes place of the buyer (PEX_{buyer})

	(I)	(II)	(III)	(IV)
A. Main estimates				
ln PEX _{seller}		-0.064***	-0.058***	✓
ln dist to CBD	0.034***	0.025***	0.025***	0.024***
ln structure size	0.222***	0.207***	0.205***	0.203***
green	0.052**	0.040*	0.041*	0.040*
ln Q score	0.765***	0.726***	0.733***	0.737***
ln age	-0.081***	-0.085***	-0.086***	-0.086***
ln NOI per square foot	0.007	0.011	0.010	0.008
B. Property type fixed effects				
Apartment (<i>ref</i>)	0.000	0.000	0.000	0.000
Industrial	0.271***	0.255***	0.247***	0.247***
Office	0.071***	0.057***	0.053***	0.052***
Retail	0.306***	0.286***	0.281***	0.280***
C. Seller type fixed effects				
Delegated (<i>ref</i>)			0.000	0.000
Direct			-0.030**	-0.028**
Non investor			0.072***	0.076***
Public			0.038**	0.038**
Foreign			-0.040**	-0.041**
D. Other fixed effects				
Year fixed effects	✓	✓	✓	✓
Metro fixed effects	✓	✓	✓	✓
E. Diagnostics				
loglik	-422,986	-422,768	-422,752	-422,703
AIC	846,288	845,854	845,830	845,912
Concordance	0.631	0.638	0.639	0.639
Wald test	5,298	5,722	5,757	5,826
N	43,928	43,928	43,928	43,928

*** = significant within the 1% level. ** = significant within the 5% level. * = significant within the 10% level

The Wald test indicates whether the final coefficients differ from the initial values. *ref* = reference category, CBD = closest Central Business District, as defined by Real Capital Analytics. NOI = Net-Operating-Income

Note that the results of our PEX model mirrors the findings of our main models in Table 5. The reason being is that larger investors tend to have less prior exposure on average per market because they are relatively diversified. Thus a positive estimate on ISQ, translates to a negative estimate in the PEX model.

From our results, we find that investors with little prior exposure in a market tend to purchase properties that are larger, green certified, high Q score, closer to the CBD and of low age. These properties are arguably more informationally efficient (Geltner and van de Minne 2017), and tend to be “safer”/core investments. This could also explain partly why larger investors tend to go with these larger properties. Investors with a lot of prior exposure also tend to purchase apartments, whereas industrial and retail is more likely to be bought by investors with less prior exposure to the market. As was the case with the ISQ variables, we also find that the sellers’ prior exposure (PEX_{seller}) explains the prior exposure of the buyer. “Local” investors (high PEX) tend to sell to other “local” investors, and the other way around as documented by the significant coefficient of PEX_{seller} in Models (II)–(IV). The only estimate that is inconsistent with previous findings is that of NOI per square foot. This estimate is insignificant for all models, whereas it had a large impact on ISQ.

Even though we will not show all the pdfs you can make from the results in Table 8, we do want to highlight one of the results. Figure 10 gives the pdf of the “average” representative property being bought by an investor with little prior exposure for different years. Little experience is defined as having less than 5% of its portfolio in the target market prior to the transaction. We find that on average, our representative property has a 13%–14% change of being bought by an investor with little experience in that market. This number is relatively stable for the different years. However, during the GFC we find a few basis points drop in this probability. This

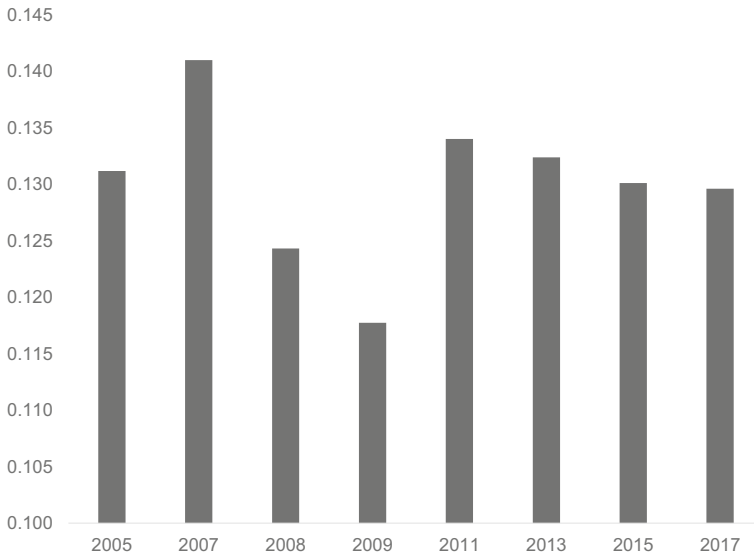


Fig. 10 Probability that the “average” representative property gets bought by an investor in a market that is “new” to this investor for a selection of years. Our definition of “new” is when the investor owned less than 5% of its portfolio (= the PEX variable) in the target market before the transaction. The results are based on the estimates of Model (I)

could mean that during the crisis investors were more afraid to invest in markets they did not know well.

Robustness Results by Investor Type

Finally, we present alternative estimations based on the multinomial logistic regression model and looking at investor types instead of size. Our dependent variable is categorical and consists of the four investor groups we constructed in Section “[Data and Descriptive Statistics](#)”. We use the same explanatory variables as in the baseline model (I). The actual estimation results, with model fit and significance can be found in the Appendix, see Table 12. The marginal effects of the estimated coefficients can be found in Table 13. Note that the rows always sum up to zero by construct.

Starting with the delegated investor, we find that they prefer large, high Q score, high NOI, green certified, young properties, close to the Central Business District (CBD). A negative coefficient for CBD means that the further away the property is from the closest CBD, the lower the probability a delegated investor will purchase it. In terms of the Q score, for every 1% increase in it, the probability that a delegated investor will purchase the property, increases by 0.15%. In other words, delegated investors are interested in “A-class” stabilized properties, essentially “cash-cows”. Delegated investors also prefer industrial properties the most as compared to the other investor types. Those industrial properties mostly consist of logistics centers and warehouses. Delegated investors are also the ones most likely to buy real estate from foreign sellers.

Public investors fit that profile to a large extent as well. Indeed, they also prefer large properties (positive coefficient on structure size, with 0.025), young properties and high Q scores. Compared to direct and the non-investor groups they also value high NOI per square foot more, but not as much as delegated investors. This effect is insignificant (Table 12). In contrast to delegated investors, public investors care less about proximity to CBDs. In general we find that public investors are spatially well diversified, as is to be expected. Relatively speaking, public investors buy less of buildings which have green certification. Also, they are most likely to buy retail properties.

Direct investors seem to have the complete opposite preferences compared to delegated or public investors. Compared to those investors, they buy smaller properties, with low Q scores, that are older, with relative low NOI per square foot. In other words, the direct investors are more opportunistic and seek to add value. They are most likely to purchase apartments. The non-investors are middle of the road in almost all categories. Compared to the other investor types they are not interested in particularly large or small, high or low Q score properties.

Finally, we will discuss the bottom panel of Table 13. Here we find the estimates for the seller type. The results indicate that—holding all things equal—direct investors are most likely to purchase from other direct investors; non-investors buy mostly from non-investors; public investors have the highest probability to buy their real estate from other public investors. Only delegated investors mostly buy from an

other category of investors different to theirs, namely the public firms. However, their second favorite seller are other delegated investors.

In summary, we find that in the absence of investor size (ISQ), investor type has been used to sort investors across properties. Larger, high quality properties are mostly sold to delegated and public firms. This is in line with our findings about large investors.

Furthermore, we find that investors tend to trade within their own investor type, holding constant for the property quality. This means that even if another type of investor had a property with the same characteristics—investors would choose to buy from the same category as theirs. This may be due to network effects—investors within the same category belong to the same networks and may feel more accustomed to do business with familiar players. We will not test the channel at hand in this paper but it remains for future research to identify the mechanisms behind this sorting between buyers and sellers.

Concluding Remarks

This paper uses novel micro-level data on commercial real estate transactions to assess market segmentation by firm size.

Considering the specifics of commercial real estate as an asset class such as low transparency, high capital intensity, indivisibility, high transaction costs, illiquidity, the paper provides evidence for market segmentation by investor size. Investor size is a latent variable which has not been directly observed previously due to the lack of available data. To estimate the size of an institutional investor we look at the properties they have purchased over a considerably long time period and compute the value of their real estate portfolios.

We then group investors into different size quantiles to assess the probability that a property is bought by a certain investor size. Following Donald et al. (2000) we estimate the quantile model in a survival framework. We find evidence of market segmentation by investor size. The main observation is that the probability that a large seller (small) will sell a property to a similar-sized buyer is higher. Moreover, large institutional investors tend to buy *larger* properties, properties with higher NOI, properties with higher quality, younger properties, and properties located closer to the CBD. The opposite is true for small investors. The observed segmentation can be attributed to investor preferences rather than financial constraints as evident by analysis of NOI. During the GFC, large investors were less likely to buy an average property, as compared to the period before or after the crisis. Moreover, the size of the average investor declined across all states during the GFC. This finding implies that the presence of large institutional investors may exacerbate a downturn in property values, as small investors would be rather interested in lower NOI properties. Further implications of our findings call for a more efficient matching of buyers and sellers by estate agents given that the market does not operate as an auction. This could potentially increase the efficiency of the real estate markets overall.

Appendix

Table 9 Year of sale distribution

Year of sale	% Count
2002	0.023
2003	0.030
2004	0.040
2005	0.062
2006	0.068
2007	0.070
2008	0.029
2009	0.015
2010	0.029
2011	0.051
2012	0.069
2013	0.076
2014	0.085
2015	0.098
2016	0.093
2017	0.082
2018	0.080

Table 10 Locational distribution

Metro	% Count	Metro	% Count	Metro	% Count	Metro	% Count
Akron	0.12%	Brevard Co	0.18%	Houston	3.12%	Pittsburgh	0.25%
Albany	0.06%	Buffalo	0.16%	Huntsville, AL	0.12%	Polk Co	0.17%
Albuquerque	0.22%	Charleston	0.22%	Indianapolis	0.60%	Portland	1.11%
Others - AR,LA	0.22%	Charlotte	1.67%	Jackson, MS	0.07%	Portland, ME	0.03%
Others - CT	0.01%	Chattanooga	0.11%	Jacksonville	0.98%	Providence	0.23%
Others - DE, MD, VA, WV	0.30%	Chicago	3.66%	Kansas City	0.62%	Provo	0.04%
Others - FL	0.17%	Cincinnati	0.56%	Knoxville	0.15%	Raleigh	1.50%
Others - GA, NC, SC, TN, KY	1.62%	Cleveland	0.37%	Los Angeles	9.53%	Reno	0.25%
Others - ID, MT, WY	0.05%	Co Springs	0.44%	Lansing	0.07%	Richmond	0.49%
Others - IL, IN, OH	0.64%	Columbia	0.31%	Las Vegas	1.72%	Rochester	0.09%
Others - MI, WI	0.27%	Columbus	0.78%	Lexington	0.08%	Sacramento	1.31%
Others - MN	0.08%	Corpus Christi	0.08%	Little Rock	0.10%	Salinas	0.07%

Table 10 (continued)

Metro	% Count	Metro	% Count	Metro	% Count	Metro	% Count
Others - MS,AL	0.43%	Dallas	4.36%	Louisville	0.25%	Salt Lake City	0.45%
Others - New England	0.20%	Dayton	0.14%	Madera, CA	0.00%	San Antonio	0.77%
Others - NJ	0.02%	Daytona Beach	0.17%	Madison, WI	0.08%	San Diego	2.17%
Others - Northern CA	0.23%	DC Metro	3.57%	Martin/St Lucie	0.08%	Santa Barbara	0.09%
Others - Northwest	0.33%	Denver	2.80%	McAllen	0.03%	Sarasota	0.27%
Others - NY	0.12%	Des Moines	0.09%	Memphis	0.66%	Savannah	0.08%
Others - OK,TX	0.55%	Detroit	0.63%	Miami/So Fla	2.79%	Seattle	2.50%
Others - PA	0.27%	Eastern PA	0.07%	Milwaukee	0.49%	San Francisco	5.14%
Others - Plains States	0.40%	El Paso	0.09%	Minneapolis	1.15%	Spokane	0.05%
Others - Southern CA	0.04%	FL Panhandle	0.27%	Modesto	0.25%	St Louis	0.63%
Others - Southwest	0.28%	Fort Collins	0.13%	Myrtle Beach	0.09%	SW Florida	0.05%
Allentown	0.16%	Fort Myers	0.25%	Naples	0.13%	Syracuse	0.05%
Ann Arbor	0.06%	Fresno	0.18%	Nashville	0.92%	Tallahassee	0.20%
Atlanta	5.02%	Gainesville	0.09%	New Haven	0.15%	Tampa	1.89%
Austin	1.18%	Grand Rapids	0.08%	New Orleans	0.11%	Toledo	0.10%
Bakersfield	0.11%	Greeley, CO	0.04%	Norfolk	0.36%	Trenton, NJ	0.06%
Baltimore	0.90%	Greensboro	0.40%	New York	6.24%	Tucson	0.54%
Baton Rouge	0.09%	Greenville	0.44%	Oklahoma City	0.28%	Tulsa	0.20%
Birmingham (AL)	0.36%	Harrisburg	0.15%	Omaha	0.17%	Wichita	0.05%
Boise	0.07%	Hartford	0.25%	Orlando	1.62%	Winston	0.18%
Boston Metro	1.76%	Hawaii	0.07%	Philadelphia	1.48%		
Boulder	0.19%	Honolulu	0.10%	Phoenix	5.12%		

Table 11 Buyer and seller type distribution

	Buyer	Seller
Delegated	0.223	0.275
Direct	0.611	0.532
Non investor	0.021	0.052
Public	0.145	0.141
Foreign investor	0.130	0.088

Table 12 Results from the Multinomial Logistic regression. Dependent variable are the different buyer types as found in bold in the column headers. The direct investor is the reference group

	Delegated	Non investor	Public
(Intercept)	-12.499***	-4.331***	-15.147***
ln dist to CBD	-0.019	-0.022	0.113***
ln structure size	0.846***	0.033	0.878***
green	0.213***	-0.288**	-0.219***
ln Q score	1.245***	-0.089	3.574***
ln age	-0.008	-0.017	-0.214***
ln NOI per square foot	0.442***	-0.164***	-0.021
Property type fixed effects			
Apartment (<i>ref</i>)	0.000	0.000	0.000
Industrial	0.802***	2.398***	1.814***
Office	0.444***	1.997***	1.740***
Retail	0.048	1.318***	2.271***
Seller type fixed effects			
Delegated (<i>ref</i>)	0.000	0.000	0.000
Direct	-0.409***	-0.006	-0.170***
Non Investor	-0.047	0.439***	0.382***
Public	0.313***	0.510***	0.216***
Foreign	0.063	0.068	-0.175***
Other fixed effects and diagnostics			
Year	✓	✓	✓
Metro	✓	✓	✓
Likelihood	-63,932		
N	43,928		

*** = significant within the 1% level, ** = significant within the 5% level, * = significant within the 10% level

ref = reference category. CBD = Central Business District, as defined by Real Capital Analytics. NOI = Net-Operating-Income at time of sale.

Table 13 Marginal effects of our Multinomial Logit Model. Dependent variable are the different buyer types as found in bold in the column headers. Parameter estimates can be found in the Appendix in Table 12

	Direct	Delegated	Non Investor	Public
ln dist to CBD (km)	-0.0011	-0.0077	0.0004	0.0084
ln structure size (in square foot)	-0.1173	0.0961	-0.0041	0.0253
Q score	-0.3302	0.1508	-0.0009	0.1803
ln age	0.0188	-0.0081	-0.0010	-0.0097
ln NOI per square foot green	-0.0435	0.0577	-0.0077	-0.0065
Property type fixed effects				
Apartment (<i>ref</i>)	0.0000	0.0000	0.0000	0.0000
Industrial	-0.2489	0.1218	0.0304	0.0967
Office	-0.1727	0.0604	0.0322	0.0801
Retail	-0.1372	-0.0161	0.0280	0.1254
Seller type fixed effects				
Delegated (<i>ref</i>)	0.0000	0.0000	0.0000	0.0000
Direct	0.0325	-0.0404	0.0042	0.0037
Non Investor	0.0048	-0.0187	0.0082	0.0057
Public	-0.0662	0.0502	0.0080	0.0080
Foreign	-0.0034	0.0140	0.0012	-0.0118
Year fixed effects	✓	✓	✓	✓
Metro fixed effects	✓	✓	✓	✓

ref = reference category. All investor explanatory variables refer to the seller type. CBD = Central Business District, as defined by Real Capital Analytics. NOI = Net-Operating-Income at time of sale

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Abraham, J.M., & Schauman, W.S. (1991). New evidence on home prices from Freddie Mac repeat sales. *Real Estate Economics*, 19, 333–352.

Badarizna, C., Ramadorai, T., Shimizu, C. (2018). Gravity, counterparties, and foreign investment. Working Paper.

Barberis, N., & Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, 68, 161–199.

Barkham, R., & Geltner, D.M. (1995). Price discovery in American and British property markets. *Real Estate Economics*, 23, 21–44.

- Beracha, E., Hardin, W. III., Skiba, H.M. (2018). Real estate market segmentation: hotels as example. *Journal of Real Estate Finance and Economics*, *56*, 252–273.
- Chinloy, P., Hardin, W. III., Wu, Z. (2013). Price, place, people and local experience. *Journal of Real Estate Research*, *35*, 477–505.
- Clapham, E., Englund, P., Quigley, J.M., Redfearn, C.L. (2006). Revisiting the past and settling the score: index revision for house price derivatives. *Real Estate Economics*, *34*, 275–302.
- Clapp, J.M., & Giaccotto, C. (1999). Revisions in repeat-sales price indexes: here today, gone tomorrow. *Real Estate Economics*, *27*, 79–104.
- Clayton, J., Ling, D.C., Naranjo, A. (2009). Commercial real estate valuation: fundamentals versus investor sentiment. *The Journal of Real Estate Finance and Economics*, *38*, 5–37.
- Cox, D.R., & Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, *55*, 1–39.
- Cox, D.R., & Reid, N. (1993). A note on the calculation of adjusted profile likelihood. *Journal of the Royal Statistical Society B*, *55*, 467–471.
- Cvijanović, D., Milcheva, S., van de Minne, A. (2020). Does investor size matter? Evidence from commercial real estate transactions. Technical report, SSRN Working Paper.
- de Haan, J., & Diewert, W.E. (2011). *Handbook on residential property price indexes*. Luxembourg: Eurostat.
- De Wit, E.R., Englund, P., Francke, M.K. (2013). Price and transaction volume in the Dutch housing market. *Regional Science and Urban Economics*, *43*, 220–241.
- Diewert, E., de Haan, J., Hendriks, R. (2011). The decomposition of a house price index into land and structures components: a hedonic regression approach. Technical report, UBC Department of Economics Discussion Paper.
- Donald, S.G., Green, D.A., Paarsch, H.J. (2000). Differences in wage distributions between Canada and the united states: an application of a flexible estimator of distribution functions in the presence of covariates. *The Review of Economic Studies*, *67*, 609–633.
- Fisher, J., Ling, D.C., Naranjo, A. (2009). Institutional capital flows and return dynamics in private commercial real estate markets. *Real Estate Economics*, *37*, 85–116.
- Francke, M.K., & van de Minne, A.M. (2017). The hierarchical repeat sales model for granular commercial real estate and residential price indices. *The Journal of Real Estate Finance and Economics*, 1–22.
- Francke, M.K., & Van de Minne, A.M. (2020a). Daily valuation of private commercial real estate: a new mixed frequency approach. Technical report, Working Paper.
- Francke, M.K., & Van de Minne, A.M. (2020b). Modelling unobserved heterogeneity in hedonic price models. *Real Estate Economics*.
- Froot, K., & Teo, M. (2008). Style investing and institutional investors. *Journal of Financial and Quantitative Analysis*, *43*, 883–906.
- Geltner, D., & Mei, J. (1995). The present value model with time-varying discount rates: implications for commercial property valuation and investment decisions. *The Journal of Real Estate Finance and Economics*, *11*, 119–135.
- Geltner, D.M., & van de Minne, A.M. (2017). Do different price points exhibit different investment risk and return in commercial real estate? *Journal of Portfolio Management*, *43*, 105–119.
- Geltner, D.M., Miller, N.G., Clayton, J., Eichholtz, P.M.A. (2014). *Commercial real estate, analysis & investments*, 3rd edn (OnCourse Learning).
- Ghent, A. (2018). What's wrong with Pittsburgh. SSRN Working Paper.
- Harrell, F.E. Jr., Lee, K.L., Mark, D.B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, *15*, 361–387.
- Ling, D.C., & Archer, W.R. (2008). *Real estate principles: a value approach*. McGraw-Hill/Irwin series in finance, insurance, and real estate (McGraw-Hill/Irwin).
- Ling, D.C., Naranjo, A., Petrova, M.T. (2018). Search costs, behavioral biases, and information intermediary effects. *Journal of Real Estate Finance and Economics*, *57*, 114–151.
- Machado, J.A.F., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, *20*, 445–465.
- Mühlhofer, T. (2019). They would if they could: assessing the bindingness of the property holding constraints for REITs. *Real Estate Economics*, *47*, 431–477.
- Piazzesi, M., Schnedeir, M., Stroebel, J. (Forthcoming). Segmented housing search. *American Economic Review*.

- Ripatti, S., & Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, *56*, 1016–1022.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *The Journal of Political Economy*, *82*, 34–55.
- Rue, H., Martino, S., Chopin, N. (2009). Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*, 319–392.
- Sagi, J. (2017). Asset-level risk and return in real estate investments. Working Paper.
- Shiller, R.J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, *71*, 421–436.
- Shiller, R.J. (1993). *Macro markets, creating institutions for managing society's largest economic risks*. Oxford: Oxford University Press.
- Van de Minne, A.M., Francke, M.K., Geltner, D.M., White, R. (2019). Using revisions as a measure of price index quality in repeat-sales models. *The Journal of Real Estate Finance and Economics*, 1–40.
- van Dijk, D., Geltner, D., van de Minne, A. (2019). Revisiting supply and demand indexes in real estate. *Journal of Real Estate Finance and Economics*.
- Van Nieuwerburgh, S. (2019). Why are REITs currently so expensive? *Real Estate Economics*, *47*, 18–65.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.