Information Asymmetries, Financial Constraints and Institutional Investment: Evidence from the Real Estate Market

Dragana Cvijanović[1], Stanimira Milcheva[b], Alex van de Minne [c],

COMMENTS WELCOME.

[a] *Warwick Business School, University of Warwick*
[b] *Department of Construction and Project Management, University College London*
[c] *School of Business, University of Connecticut*

Abstract

In this paper we analyze the underlying economic mechanisms that might be driving the observed patterns in commercial real estate prices, as an interplay between buyer and seller characteristics (in terms of their size, capital constraints, management skill and market knowledge), and the timing and geographical location of these transactions. By jointly modelling the institutional investors' decision to invest in a particular real estate market and the effect of such decision on real estate prices and holding period, we find that, controlling for property-, and time-varying location specific-, as well as investor characteristics: largest buyers (sellers) tend to pay (sell for) a price premium for the otherwise identical property at the time of purchase (sale), relative to smallest buyers (sellers). Keeping investor size and investor financing constraints constant, more informed sellers tend to sell at a premium, while more informed buyers tend to buy at a discount. Furthermore, more informed sellers (buyers) hold properties longer. These results point to a significant role of private valuations and investor market informedness in commercial real estate markets, when financing choice is taken into account.

Email addresses: `dragana.cvijanovic@wbs.ac.uk` (Dragana Cvijanović), `s.milcheva@ucl.ac.uk` (Stanimira Milcheva), `avdminne@uconn.edu` (Alex van de Minne)

In case of heterogeneous assets, such as real estate, not all buyers and sellers are equally well informed about a particular market. Evidence from the housing market suggests that the amount of information market participants have about local markets has important implications for subsequent real estate price dynamics (Kurlat and Stroebel, 2015). While the information literature in (commercial) real estate markets is in its infancy (Badarinza et al., 2019; Agarwal et al., 2018; Garmaise and Moskowitz, 2003), few studies have examined the interplay of information asymmetry and capital constraints on valuation and time to sale of heterogeneous assets, such as real estate. In this paper, we try to fill this gap. Given significant heterogeneity among institutional investors in real estate, we are provided with a natural laboratory to analyze the underlying economic mechanisms that might be driving the observed patterns in commercial real estate prices and associated holding period, as an interplay between the characteristics of buyers and sellers, in terms of their size, capital constraints, management skill and market informedness.

By jointly modelling the institutional investors' decision to invest in a particular real estate market (both in terms of property type and property location) and the effect of such decision on real estate prices using longitudinal and survival models, we find that, controlling for property characteristics, time-varying location specific characteristics (such as local demand) and buyer (seller) size, indebtedness, management skill and prior local market knowledge : (1) Buyers of the largest quartile (portfolio wise) pay approximately 5% more for the otherwise identical property, compared to a similar investor from the smallest quartile. There is no significant different between the investor size quartiles when the property is sold. (Although the very smallest investors do sell with a modest discount.) This price premia at the time of purchase and sale cannot be explained by investors' financing constraints, nor by the level of information they have about the particular market prior to transaction, suggesting that investor private valuations have a significant role in pricing of commercial real estate.

In the spirit of Garmaise and Moskowitz (2003) we find that (2) the relationship between seller Debt-to-Asset ratio (DTA) and sales prices is almost flat, but that buyers faced with binding financing constraints, as measured by high DTA ratios, tend to be more careful and purchase otherwise identical assets at a discount. In both cases, however, the probability of transaction taking place decreases in the level of financial constraints faced by the sellers (buyers), consistent with the evidence from the housing markets (Genesove and Mayer, 1997, 2001). (3) Keeping investor size and investor financing constraints constant, more informed sellers tend to sell at a premium, while more informed buyers tend to buy at a discount, suggesting that information asymmetries are a significant driver of commercial real estate pricing, even when financing choice is taken into account. Furthermore, more informed sellers (buyers) are more likely to sell (buy), providing support to the argument that reduction in information asymmetry has a positive effect on asset liquidity (Ghent, 2019; Sagi, 2017). Finally (4) we find that real estate firms with more management skill sell properties for a premium. Although there is no evidence that high management skill results in a discount when buyer a property. We also find that firms with high management skill hold properties longer compared to their lower management skill counterpart.

The observed evidence can be interpreted as a manifestation of different economic mechanisms: in the commercial real estate market, buying and selling does not happen as an auction with the highest bidder getting the property. Instead, the process is rather less transparent and not always the highest bidder would be chosen. This may be due to unobserved information about the property or the actors, such as asymmetric information and search costs, differences in bargaining power or capital constraints, managerial skill, or heterogeneous underlying investors preferences, as captured by their reservation prices. In this paper, we find evidence of economically significant effects of investor private valuations (unobserved preferences) and prior informedness on commercial real estate prices and probabilities of sale.

While institutional ownership of fixed income assets and equities has been extensively researched in the finance literature, little is yet known about the drivers and impact of institutional investment on commercial real estate markets. [1] Why would we expect there to be an impact of investor characteristics on pricing and holding period in the first place? Unlike publicly traded assets, whose prices should not be affected by capital flows or trading activity, assuming rational investors and costless arbitrage, real estate assets are heterogeneous and real estate markets are highly segmented. This market segmentation may create conditions in which capital flows (or changes in aggregate demand) affect real estate asset valuations, due to inelastic supply

[1] In this paper institutional investment is defined broadly, to include: private equity funds, endowments, sovereign wealth funds, real estate developers, pension funds, insurance companies, management companies, Real Estate Investment Trusts (REITs), real estate operating companies (REOCs), etc.

of real estate or due to changes in investor expectations (Fisher et al. (2009), Clayton et al. (2009)).

Our empirical tests are built on a large sample of commercial real estate transactions covering all metro areas in the U.S. and property types; and coupled with the 17 years of transaction data, we can empirically test both cross-sectional and temporal variation in the role of asymmetric information and financial constraints in commercial real estate markets. We construct a comprehensive data set of all commercial real estate transactions[2] in the U.S. between 2000 and 2018, available through Real Capital Analytics (RCA). For each transaction, we obtain detailed hedonic property characteristics, transaction price, ownership structure and data on any loans used to finance the transaction. Coupled with detailed data on 98 property-type-specific and metro-specific Commercial Property Indices (RCA CPPI) for the U.S., we are able to estimate the size of each investor's real estate portfolio, as well as both property- and investor-level Debt-to-Asset (DTA) ratios. To proxy for the amount of information an investor has about a particular market, we compute their relative prior exposure to each submarket by dividing the total number of real estate assets they own in a particular market, as a percentage of their overall portfolio.

The observed contemporaneous correlation between commercial real estate prices and institutional investor characteristics, falls short of providing evidence of a causal effect of institutional investor characteristics on commercial real estate values, both directly, and indirectly.

To address this issues, we identify the effect of seller and buyer characteristics on prices and probability of sale by following the *identical* property over time. Property level Net-Operating-Income (NOI) in combination with property random effects (and a comprehensive matrix of property-level covariates) captures most of the (partly unobservable) heterogeneity of transactions. Any change in sales prices, or holding period must therefore be caused by the seller and buyer characteristics. Note that this is essentially a difference-in-differences setup, taking care of any endogeneity. Also, we model the property holding period simultaneously with prices by allowing the property specific random effects to correlate, because it is well established in real estate literature that prices and liquidity/probability of sale co-move (Van Dijk et al., 2018). The holding period is modelled as a survival model,

[2]Larger than $2M

with the holding period as a baseline. All seller and buyer characteristics are subdivided into 40 quantiles each, allowing for a near non-parametric relationship with prices and the probability of sale. We subsequently structure the parameters (on the seller and buyer characteristics) to follow a second order random walk. Given the complexity of the joint longitudinal/survival model with many random effects, we estimate the system using Integrated Nested Laplace Approximation, developed by Rue et al. (2009).

Studies that use commercial real estate data are far and few between; and many of them focus either only in a handful of markets in the US (Garmaise and Moskowitz, 2003), or they are interested in the international dimension of the role of asymmetric information in institutional investment (Badarinza et al., 2019; Agarwal et al., 2018). The role of information asymmetries has received a lot more attention in the housing markets (Kurlat and Stroebel, 2015; Chinco and Mayer, 2015). At the same time, the role of capital constraints has been analyzed also in the context of housing markets (Genesove and Mayer, 1997, 2001). Our paper contributes to both of these strands of literature by examining investor behaviour in commercial real estate markets, and the joint role of asymmetric information and search costs, and liquidity constraints in pricing and probability of sale of commercial real estate assets.

The paper is also broadly related to a body of work that analyzes real estate markets using search and matching models. While a number of papers have used search and matching models to understand the housing market[3], the literature that looks at CRE markets in this context is only in its infancy. In one of the fist studies in this field, (Sagi, 2017) explains the returns on individual properties with a search model. While Badarinza et al. (2019) uses a search model to quantify how search frictions arising from differences in investor nationality affect cross-border capital flows, Ghent (2019) studies the effects of heterogeneity in the frequency of valuation shocks. In this paper, we look at the matching process between buyers and sellers with heterogenous liquidity constraints and varying levels of information about a particular real estate market.

The rest of the paper is organized as follows: in Section 1 we describe our data set. In Section 2 we discuss our empirical strategy, while Section 3 discusses our main findings. In Section 4 we conclude.

[3]See Han and Strange (2015) for a summary of early literature on housing search models.

1. Data and Descriptive Statistics

Our analysis uses a unique data set of commercial real estate transactions in the U.S., containing fine micro-level transaction details such as property type, location, and profiles of both buyers and sellers, from Real Capital Analytics (RCA), a commercial data analytics firm.[4] RCA collects the sales data for commercial properties and portfolios transacted at a minimum price of US$1 million and claim to have achieved a capture rate of over 90% for transactions of commercial "investable" real estate. As a result, the RCA data set is considered industry standard and is said to capture close to the entire population of commercial real estate properties.

Our initial dataset covers all commercial real estate transactions across the U.S. of at least US$2 million, between 2000 and 2018. For every transaction we observe the transaction price, Net-Operating-Income (NOI), the size of the property, location[5], property type (office, retail, apartment and industrial), construction year and year of sale. For each transaction, we also observe up to four buyers and up to four sellers. We also have information on the structure of the deal, whether it is a Joint Venture, or if the buyers/sellers are delegated investors. Furthermore, for each real estate asset, we observe the loan amount at transaction date, and all subsequent refinances. This data is often supplemented with information about the loan interest rate for most loan originations or refinances.

Using this data, we construct a full panel of property values, hedonic characteristics and buyer and seller characteristics over time. For each real estate asset, we first take the average sales price (and appraised values during refinances) and average transaction (and refinance) year. For example, say that a property was sold in 2000 for 10M and refinanced in 2010 with an appraised value of 20M. In this case the average property value is 15M and the average year is 2005. To obtain the real estate asset's value for other years,

[4] RCA is an independent real estate data analytics firm headquartered in New York, with offices in San Jose, London and Singapore. It collects data on commercial property transactions with a minimum size of US$2 million across 146 countries. The database contains, cumulatively, US$18 trillion in commercial property transactions linked to over 200,000 investors and lenders. Source: https://www.rcanalytics.com/.

[5] RCA has a definition for location which largely follows MSA defnitions. However, for the larger MSAs, RCA defines more granular "sub-markets", like Manhattan, Boroughs and Suburbs for New York. In total, RCA has 98 of such market definitions. We will use these market definitions throughout the paper.

we use the RCA Commercial Property Price Indexes (CPPI). The RCA CPPIs are price indexes based on the Bayesian repeat sales methodology laid out in Van de Minne et al. (2019), while using mostly the same underlying RCA transaction data described above. RCA produces 98 (monthly) location- and property-type-specific RCA CPPIs in the US. All indexes are non overlapping, spanning time period between December 2000 and 2018. [6]

We use the buyer and seller information to find out in what year which investor (or combination of investors) owned which real estate asset. For example, if investor A sold property X in 2005 to investor B, we assume that investor A was the owner of property X between 2000 and 2004, and investor B started owning said property in 2005 (until it was sold again, or we are at the end of our sample). If there are multiple owners of the asset, we divide the ownership equally over all partners.[7]

1.1. Investor Metrics

An important feature of our data set, is that it allows us to compute several (institutional) buyer and seller metrics, that can be used to capture buyer and seller size (in terms of their real estate portfolio), their real estate leverage ratios ((real estate) debt to total (real estate) assets) and their prior exposure to a particular real estate market.

First of all, we can easily compute the **value** of investors real estate portfolio in every single year.[8] We simply "sum" all the real estate values for every year for every investor.

To proxy for buyer(seller) liquidity constraints (Genesove and Mayer, 1997), we compute investor-level real estate debt to total (real estate) assets ratio (**Debt-to-Asset** or DTA) in each year by aggregating property-level debt to investor level and dividing by the size of the investor real estate portfolio, as defined above. For each property in our sample, we have data on the loan amount at the time of sale and refinancing.[9] [10] To calculate

[6]The RCA CPPI is based on the structural time series modelling developed by Van de Minne et al. (2019). The indexes capture all price movements created by macro economic and local factors, as well as market level depreciation.

[7]The relative investor shares in joint ventures are not available in RCA data.

[8]Note that we do not take into account other assets than real estate (since they are not available for most investors in our sample), or real estate assets outside of the US.

[9]If there are multiple owners, we divide the stated transaction level debt equally over all partners.

[10]On average, a property gets refinanced every three years and gets transacted every six

the debt amount between transactions (and refinancings) we use a 30-year annuity calculation in combination with the reported interest rate.[11] As such we can get a good proxy of remaining mortgage debt on each property in every year. [12]

To proxy for investors informativeness about a particular property market (Kurlat and Stroebel, 2015; Agarwal et al., 2018; Badarinza et al., 2019) we construct a measure of investors **prior exposure** to a local sub-market, by taking the total number of real estate assets the investor already owns in that market and dividing it by the total number of assets the investor owns across the U.S. For example, if investor A only has properties in market Z, this measure will get a value of 1. If investor B has 10 properties in 10 different markets, the value of prior exposure associated with every market will be 1/10, etc. This relative measure captures the prior knowledge or **market information** the buyer (or seller) have about each local sub-market. [13]

Our fourth investor metric is based on the amount of refinancing the owners are associated with, which we see as a proxy for **management skill**. More specifically, we first *count* the amount of times owners refinanced their properties whenever the interest rates are lower now, compared to when the mortgage was originated. Note that one property can be refinanced multiply times throughout its history by the same owner. Next, we divide this number by the total amount of properties held by the corresponding owner. Our proxy for management skill therefore mostly takes a value between 0 and 1. However larger than 1 is possible, in case the owners refinance the same property multiple times.

Finally, we only keep the properties with *accurate* transaction prices and non-missing values for the other explanatory variables, including the NOI. If we do not observe an accurate price for a transaction, we omit that entire property from our data, not just that transaction, to avoid generating a bias in our holding period data.[14] We also limit our sample to transac-

years in our sample.

[11]In cases where we do not observe loan interest rate, we simply take the average sample interest rate in that year.

[12]By definition, this means that we compute both property-level Loan-to-Value (LTV) ratios, and investor level LTVs, which we will call Debt-to-Asset ratios throughout.

[13]By scaling the number of real estate assets in each sub-market by total number of assets, we are also implicitly controlling for investor size.

[14]Suppose property A is sold in 2001, 2006 and 2018. If we then only drop the middle

tions after 2006, because by then we have more accurate estimates of the levels of debt. Our final sample contains around 15,000 observations. Some descriptive statistics of the this transaction data can be found in Table B.1.

[Place Table B.1 about here]

The average property transaction price is \$37M, and the average value of the investor real estate portfolio (*Value*) as calculated by us is \$550M. Note that this number is not for individual investors per se, but also for the joint ventures, in which case we sum the portfolio values. Thus, on average our institutional investors (and joint ventures of investors) own approximately ($\frac{\$550\text{M}}{\$37\text{M}} \approx$) 15 properties in any given year. The average holding period is 6 years, but most properties in our sample haven't been sold yet (for a second time). Variable *Censored* which is 0.12 suggests that 88% of the properties in our sample period were only sold once. The holding period is either the time between buy and sell (in which case the *Censored* dummy will get a value of 1), or the time between buy and the end of our sample, which is 2018 (in which case the *Censored* dummy will get a value of 0).

On average we find a Debt-to-Asset (*DTA*) of 55% between the buyers and sellers. As Figure A.2 shows, there have been relatively large swings in combined buyer and seller *DTA* ratios between 2006 and 2018. During the crisis, the DTA of the average buyer and seller was over 75%, whereas at the end of the sample, the DTA is as low as 50%.

[Place Figure A.2 about here]

Our measure of investor market informativeness, *PriorExposure* shows that on average buyers and sellers have approximately one-fifth of their assets *already* in the market they are about to transact in. Since there are a multitude of property-type- and location-specific local sub-markets (98 in total) in comparison the average properties held (18, see discussion above), if the choice to buy a real estate asset in a particular market was random, the average value of our informativeness measure *PriorExposure* would be close to zero. Given that the actual number is a lot higher, suggests that the locational and property-type choices of investors are not random and that

transaction, we would attribute a holding period of 17 years to the first sale, whereas in reality it was "only" 5 years. If we would leave out the final year, we would forget to "censor" the 2006 transaction in 2018, also biasing our data.

investors at least to a certain degree specialize and prefer markets with lower amount of information asymmetries (Kurlat and Stroebel, 2015; Badarinza et al., 2019; Agarwal et al., 2018).

Our proxy for management skill indicates that approximately 12% of all properties are refinanced in case of a lower interest rate environment. Finally we distinguish between four categories of buyers and sellers. This categorization was used in previous literature as well, see for example Ghent (2019). Delegated investors include pension funds, equity funds, investment managers and banks. Public investor types include the REITs and REOCs. REITs and REOCs are under more scrutiny and are more transparent compared to the other firm types. In addition, these public companies have 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 have to group called "direct" which includes all remaining investor types, like high net worth individuals and Sovereign Wealth Funds. The "direct" group is the largest of them all, and the "non investor" the smallest.

2. Methodology

2.1. Reservation Prices and Liquidity

Following Fisher et al. (2003, 2007) and Van Dijk et al. (2018) we assume that heterogeneous properties are traded among heterogeneous agents in a double-sided search market. Both buyers and sellers set their reservation prices based on property and buyer characteristics and the current market conditions. These can be represented by two hedonic/structural pricing equations, as in (Rosen, 1974):

$$P^B_{ijt} = \mu^B_t + X_{it}\beta^B + Q^B_{jt}\alpha^B + \epsilon^B_{ijt},$$
$$P^S_{ijt} = \mu^S_t + X_{it}\beta^S + Q^S_{jt}\alpha^S + \epsilon^S_{ijt},$$

where P^B_{ijt} (P^S_{ijt}) denotes the reservation price of buyer B (seller S respectively). The time varying constant μ captures changes in the macro-economic environment (which could be market specific), whereas vector X denotes the publicly observable characteristics of the real estate asset i and vector Q denotes mostly unobservable characteristics of agent j. Subscript t denotes the time of sale. Vectors β and α capture the corresponding parameter estimates. The residuals (ϵ) are assumed to be normally distributed with mean zero and variance σ_ϵ^2.

A sale with transaction price P_{it} only takes place if the reservation price of a buyer is higher (or equal) to that of the seller ($P_{ijt}^B \geq P_{ijt}^S$), and if the transaction occurs $P_{ijt}^B \geq P_{it} \geq P_{ijt}^S$. This implies that the current real estate owners (i.e. the sellers) can "drive" realized prices to a large extent. Sellers will be reluctant to sell, for as long as they cannot get a satisfactory price for the property (Genesove and Mayer, 2001, 1997). It is therefore unsurprising that we are interested in **holding period** as well.

In the housing market context, De Wit and Van der Klaauw (2013) show how sellers use list prices to signal their reservation price to potential buyers in the single family housing market in the Netherlands. Low list prices typically result in lower transaction prices, but also in lower time on the market, and vice versa. Using evidence from the Boston housing market in 1990s, Genesove and Mayer (1997, 2001) find that high levels of sellers' mortgage debt (measured by Loan-to-Value ratios) at the asset level increase sales prices, but also increase time on market. The existing literature in the field of real estate repeat sales models find mixed evidence on the relationship between holding periods and (real) asset returns. Typically, extant literature finds that properties with shorter holding periods have higher returns, most likely driven by unobserved variation in property characteristics or very desirable elements to a property.[15] For example, Chinco and Mayer (2015) show that out-of-town second-house buyers behaved like misinformed speculators during the mid 2000s, driving up house-price appreciation rates. [16] Whereas the existing studies are set in the housing context (Kurlat and Stroebel, 2015) and focus on either pricing (Genesove and Mayer, 2001) or the probability of sale (Genesove and Mayer, 1997), or estimate these models separately, we build a simultaneous model of prices and "liquidity" (measured by holding period in our case) that accounts for these correlations in unobserved heterogeneity.

[15]See Abraham and Schauman (1991); Shiller (1993); Clapp and Giaccotto (1999); Clapham et al. (2006) among others for a discussion on this topic.

[16]In our dataset we observe properties that where redeveloped. We omit these from our analysis, however, large capital expenditures are still unobservable to us. For more recent work on ways to capture capital expenditures, see Sagi (2017).

2.2. Our Methodology

The joint modelling of longitudinal response measurements (like real estate prices in our case) and the timing of events is well researched in the medical sciences. (See Henderson et al., 2000; Faucett and Thomas, 1996; Lin et al., 2002; Brown and Ibrahim, 2003; Wang and Taylor, 2001, for example.) A familiar example is that of HIV clinical trials, where covariates, including treatment assignment, demographic information, and physiological characteristics, are recorded at baseline, and measures of immunologic and virologic status, such as CD4 count and viral RNA copy number, are taken at subsequent clinic visits. Separately modelling such longitudinal and timing-to-events models are inappropriate, because the longitudinal variables is typically correlated with patients survival endpoint. Therefore, over the years many models to overcome such issues have been developed in said field.

We follow Henderson et al. (2000), who proposed a very flexible joint model that allows a very broad range of dependencies between the longitudinal responses and the survival endpoints. The key insight of the methodology is that one ought to connect the longitudinal and survival processes with a latent bivariate Gaussian process. The longitudinal and time-to-event data can then be assumed independent given the linking latent process and the covariates.

Assume there are m properties, that we follow over a time interval $[0, \omega)$. The i^{th} subject provides a set of (possibly partly missing) longitudinal measurements $y_{ij}, j = 1, \ldots, n_i$ at times $s_{ij} = j = 1, \ldots, n_i$, and a possible censored survival time t_i to a certain endpoint. The joint model is composed of two submodels, one for each type of data. The longitudinal data y_{ij} are modeled as;

$$
\begin{aligned}
y_{ij}|\eta_{ij}, \sigma_\epsilon^2 &\sim N(\eta_{ij}, \sigma_\epsilon^2), \\
\eta_{ij} &= \mu_i(s_{ij}) + \theta_{i,1},
\end{aligned}
\tag{1}
$$

where $\mu_i(s)$ is given by $X_{i,1}\beta_1 + Q_{i,1}^B\alpha_1^B + Q_{i,1}^S\alpha_1^S$, the (potentially time-varying) property and agent characteristics with corresponding parameter estimates. Note that parameter $\theta_{i,1}$ is the property random effect, and will capture all unobserved heterogeneity, similar to the research by Francke and Van de Minne (2019). The response variable y in our case are real estate transaction prices. (Where the price is affected by both buyer *and* sellers' characteristics, as well as property characteristics.)

We subsequently also need a survival model for the holding periods, since many properties will have been bought, but are not yet sold at the end of the sample, and are thus *censored*. More specifically, we model the survival data by an Accelerated Failure Time (AFT) model. We assume that the survival time (read holding period) of the i^{th} subject follows a Weibull distribution,

$$\begin{aligned} t_i|\lambda_i(t) &\sim \text{Weibull}(\phi, \lambda_i(t)), \\ \log(\lambda_i(t)) &= X_{i,2}\beta_2 + Q_{i,2}^B\alpha_2^B + Q_{i,2}^S\alpha_2^S + \theta_{i,2}, \end{aligned} \tag{2}$$

where $X_{i,2}\beta_2 + Q_{i,2}^B\alpha_2^B + Q_{i,2}^S\alpha_2^S$ represent the possible time-varying covariates of both the property and the sellers/buyers with corresponding parameter vectors. Note that the covariates do not need to be similar between the longitudinal and survival model. Similar to Eq. (1), we also include a random effect (called a "frailty" in survival literature), denoted by $\theta_{i,2}$.

We jointly model the longitudinal and survival processes via a latent zero-mean bivariate Gaussian process on $(\theta_{i,1}, \theta_{i,2})^T$, which is independent across different subject, as described in Henderson et al. (2000). For low-dimensional problems, such multivariate normal random-effects distributions might be best specified and understood in terms of conditional distributions, for example,

$$\begin{pmatrix} \theta_{i,1} \\ \theta_{i,2} \end{pmatrix} \sim \text{NVM}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\theta_1}^2 & \rho\sigma_{\theta_1}\sigma_{\theta_2} \\ \rho\sigma_{\theta_1}\sigma_{\theta_2} & \sigma_{\theta_2}^2 \end{pmatrix} \right)$$

is equivalent to,

$$\begin{aligned} \theta_{i,1} &\sim N(0, \sigma_{\theta_1}^2), \\ \theta_{i,2} &\sim N\Big(\frac{\sigma_{\theta_2}}{\sigma_{\theta_1}}\rho\theta_{i,1}, (1-\rho^2)\sigma_{\theta_2}^2\Big). \end{aligned} \tag{3}$$

We constraint parameter $-1 \leq \rho \leq 1$ and both variance parameters $(\sigma_{\theta_1}^2, \sigma_{\theta_2}^2)$ have to be larger than zero. Note that with a $\rho < 0$ the random effects of prices and liquidity (holding periods) are *positively* correlated, and vice versa. (A negative coefficient in the hazard model indicates shorter holding periods, i.e. more liquidity.) With $\rho = 0$, we get a model with uncorrelated unobserved heterogeneity, and we get $\theta_{i,2} \sim N(0, \sigma_{\theta_2}^2)$ for the survival model as well. Since there can also be a difference in "scale" in the random effects, this scale is captured by the fraction of the variance terms, or; $\frac{\sigma_{\theta_2}}{\sigma_{\theta_1}}$.

Such models can be estimated by Bayesian procedures. For example, Guo and Carlin (2004) develop a fully MCMC procedure and Rizopoulos (2011) use a specialized EM algorithm. However, given the relative large dataset we are using (see Section 1) we use Integrated Nested Laplace Approximation (INLA), developed by Rue et al. (2009).[17] INLA has been used in real estate application before, see Francke and Van de Minne (2019). Using INLA to model a longitudinal and timing-of-events study is not straightforward. However, Martino et al. (2011) show how INLA can be adapted and applied to such complex models.

In essence, INLA computes an approximation to the posterior marginal distribution of the hyper-parameters. Operationally, INLA proceeds by first exploring the marginal joint posterior for the hyper-parameters in order to locate the mode, a grid search is then performed and produces a set of "relevant" points together with a corresponding set of weights, to give the approximation of the distributions. Each marginal posterior can be obtained using interpolation based on the computed values and correcting for (probably) skewness, by using log-splines. For each hyper-parameter, the conditional posteriors are then evaluated on a grid of selected values for the prior and the marginal posteriors are obtained by numerical integration. In this paper we have a flat prior for all hyper-parameters.

2.3. Parametrization

Longitudinal price changes in real estate can be very local, since changes in local demand and supply can arise quickly (Francke and van de Minne, 2017; Van de Minne et al., 2019). DiPasquale and Wheaton (1994) show that if demand for real estate increases, rents go up as a function of prevailing cap rates and how well the construction industry responds to increases in demand. (Think of zoning restrictions and imperfect foresight; Saiz, 2010; Harter-Dreiman, 2004, among others.). It is therefore custom to interact (sub)location dummies with dummies for time of sale. However, given that in our study we analyze the entire commercial real estate market of the U.S., that would result in a big loss in degrees of freedom and computational efficiency. Furthermore, the estimated dummies will be affected by noise, implying that the (real estate) indexes will not be accurate or reliable (Geltner and Ling, 2006).

[17] Note that in the medical science datasets are rarely larger than 1,000 subjects, whereas we work with over 10K transactions.

In order to accurately control for time-variation in local market conditions, one can alternatively add variables to the model on the right-hand side that "capture" local price changes. For example Fisher et al. (2003, 2007) use the appraised values of the underlying real estate as an independent variable, using NCREIF data. A change in appraised values could indicate a change in demand/supply on a very local level. To capture changes in local demand, we use Net-Operating-Income (NOI).

Extant literature (Geltner et al., 2014, p. 554-556) shows idiosyncratic price movement specific to individual assets or granular market segments, largely reflect 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).[18]

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 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). A stylized example of this phenomena is given in Figure A.1.

[Place Figure A.1 about here]

Figure A.1 gives a few examples of market capitalization rates ("caprates") in the US, as given by Real Capital Analytics. Relationships depicted in the graph show that the correlations between caprates are very high, although we are talking about very different markets across the US. This motivates our basic longitudinal model below:

$$\eta_{ijt} = \ln \frac{P_{ijt}}{S_{it}} = \gamma_t + \ln \frac{N_{it}}{S_{it}} \beta^1 + d_{it}^{\text{Mkt}} \zeta + X_{it} \beta^2 + Q_{jt}^B \alpha^B + Q_{jt}^S \alpha^S + \theta_{i,1} + \epsilon_{ijt} \quad (4)$$

where P are transaction prices, S is the size in square footage, N is the property NOI at time of transaction. Time varying constant γ_t essentially captures the effect of the risk free rate on property prices, and is therefore

[18] This phenomenon is also widely accepted in the stock market (Shiller, 1981).

similar to all properties and sellers and buyers alike in a given year. As argued above, the property-level NOI captures a broad matrix of time-varying observed and unobserved characteristics in the very local (i.e. property level) space markets. We also estimate a parameter estimate on NOI (designated β^1), as there could be some non-linearities between the amount of NOI and prices. Subsequently, we have an offset of the "market" (a combination of location and property type), given by d_{it}^{Mkt}, with corresponding parameter vector ζ, and we have property specific offsets, capturing other unobserved heterogeneity in the capital market on a property level in θ.

Our matrix of property characteristics (X) includes the following variables. We first include the square footage of the property itself. It is well established in real estate literature that larger properties tend to sell for relatively less compared to smaller properties, *ceteris paribus* (Francke and Van de Minne, 2017). It is not known a priori if this will be fully captured by NOI, and we therefore include the (log) of square footage as well. Next we compute a variable based on construction year, by subtracting the construction year of the property from the first year of our sample (2006, see Section 1). This "age difference" *and* this "age difference" squared are both added to the regression model (Bokhari and Geltner, 2018; Geltner and Bokhari, 2015). Since the year of construction serves as an implicit control for asset depreciation, and given that we already control for NOI, our estimates will tell something about "caprate creep" (i.e. the depreciation of caprates). The effect of depreciation on caprates is not settled in existing literature. This is most likely caused by two opposing forces at play: (1) older properties generate less and less NOI, increasing the caprate (or decreasing the growth), however (2) older properties have a higher redevelopment potential, increasing future NOI and thus decreasing the caprate (or increasing the growth). [19]

Note that we use year of construction instead of age, as is more typical. The only reason for this, is that it is very difficult to enter age in the survival model, for every x years holding period, the property also aged x years. For the sake of consistency we therefore also use construction year in the longitudinal model, even though it is not necessary.

As explained in Section 1, our sellers/buyers characteristics (Q) consists

[19] See Clapp et al. (2012); Geltner et al. (2018) for a discussion on redevelopment option value.

of: (1) the firm's value, (2) firm's Debt-to-Assets, (3) firm's informdness, and (4) the management skill of the firm proxied by the amount of refinances. Given that we do not have a prior believe on the effect of these variables on prices and holding period, and because we want to allow for non-linearities in said variables, we want to enter the sellers/buyers characteristics non-parametrically. More specifically, we first "bucket" the seller/buyer characteristics in Z quantiles. In our case we subdivide the variables in 40 quantile buckets each increasing 2.5%, from $Z = \{q_0, q_{0.025}, q_{0.050}, \ldots, q_{0.975}, q_1\}$. Next we enter the 40 "dummies" in the regression model. Given the high amount of estimated parameters (40 per variables $\times 3 = 120$, we structure the parameters in a way that we keep the flexibility, but also have a more parsimonious model. We use a random walk model of order 2. Take the Gaussian vector of parameter estimates $\alpha = (\alpha_1, \ldots, \alpha_Q)$, with Q being all the quantile groups. We construct or second order random walk by assuming independent second order increments;

$$\begin{aligned} \Delta^2\alpha_q &= \alpha_q - 2\alpha_{q+1} + \alpha_{q+2} + \xi_q \\ \xi_q &\sim N(0, \sigma_\xi^2). \end{aligned} \tag{5}$$

Note that we estimate the set of parameters for each buyer and seller's characteristic. In order to identify to parameters, we put the constraint on α that the values must sum to 0, or: $\sum_{q=1}^{Q} \alpha_q = 0$.

We also include (non time varying) the investor type of buyer and seller in Eq. (4). As noted in Section 1, we define four types of investors; delegated, non investor, direct and delegated.

For the survival model, we use the same variables as in the longitudinal model. We also put the same structure on the parameters α. The only difference is that we effectively replace the calendar time effects with a baseline for holding period. It *is* possible to add calendar time effects to survival models, see for example De Wit et al. (2013). However, we feel that this extra added layer of complexity is outside the scope of this paper. Also note that we have multiple time-varying parameters, which, as explained earlier, should capture much of the very local calendar time effects.

3. Results

To summarize our methodology; We identify the effect of seller and buyer characteristics on prices and holding period by following the *same* property over time. Net-Operating-Income (NOI) in combination with property random effects (and some other covatiates) captures most of the (unobserved) heterogeneity of properties. Any change in sales prices, or probability of sale must therefore be caused by the seller and buyer characteristics. Note that this is essentially a difference-in-differences setup, taking care of any endogeneity. All of this allows us to see the effect if - for example the DTA - for sellers affect prices differently compared to buyers. Also, we model the holding period simultaneously with prices through the random effects term (and by using similar covariates), because it is well established in real estate literature that prices and liquidity/probability of sale co-move (Van Dijk et al., 2018).

For readability we essentially "split" our results into two groups. First of all, you can find our results for the "fixed parameters", including: NOI per square foot, the square footage itself, construction year and the fixed effects in Tables B.2. The estimates of the most important dummy variables can be found in Table B.3.

The second set of results show the results for the "random parameters", which include all four of our buyer and seller characteristics. We present these findings in Figures A.5 – A.8. The variance, or hyperparameters can be found in Table B.4. The estimated coefficients for hazard models with a Weibull distribution can be difficult to interpret. Therefore, we rescale all the parameters by $\tilde{\beta} = \frac{-\beta}{\phi}$, with ϕ being the estimated shape parameter of the Weibull distribution. This makes the interpretation more in line with "standard" cox proportional hazard (Cox and Reid, 1987, 1993) and probit/logit models, meaning that a positive estimate means a longer holding period of the real estate and vice verse. Although it should be noted that the interpretation of the results for the buyers' characteristics on holding period is still not so straightforward. Indeed, how could a buyer affect the holding period of a property not (yet) owned by this buyer? In fact, the interpretation is more about *preferences/matching*. More specifically, a positive estimate for a buyer (characteristic), indicates that this buyer is more likely to buy from a seller that is inclined to have longer holding periods, ceteris paribus. As a robustness we also ran the models without the buyer characteristics, see also Section 3.3.

We will first discuss the results of the fixed parameters (NOI, square footage, construction year and other fixed effects). Subsequently, we discuss the effects the seller and buyers have on transaction prices and probability to sell in Section 3.2. Finally, we shortly lay out all the robustness checks we ran in Section 3.3.

3.1. Results of Fixed Parameters

The results of the fixed parameters are as expected, see Tables B.2 – B.3. For example, we find that the size of the property does not impact the price per square foot. A property twice the size, but with a similar NOI per square foot, will not have a higher price per square foot. NOI per square foot on the other hand, does impact the price substantially. We roughly find that a doubling of the NOI per square foot, will result in a 70% increase in price per square foot, *ceteris paribus*.[20] One way to interpret these results is that cap rates are not different for properties of different sizes in the cross-section, and that price differences are thus mainly driven by the NOI. Square footage does impact the holding period. Larger properties that transact have lower holding periods, see negative coefficient on the corresponding parameter. The NOI per square foot also impact the holding period significantly. Investors that own a high NOI per square foot property, will tend to hold on to this property for a longer period (see positive coefficient in Table B.2).

[Place Table B.2 about here]

We also find similar results between buyers and sellers for the effect of construction year (and construction year squared) on prices per square foot and the holding period. Because the coefficients themselves are difficult to grasp, we plotted the effect of construction year on prices per square foot in Figures A.3a – A.3b for the effect of age on price per square foot and the holding period respectively. Properties built in the 1950s will have a 10% – 12% discount. However, properties older than that actually transact with a premium, *ceteris paribus*. Again, this relates to redevelopment option value, increasing property values through cap rates (Geltner et al., 2018). This does not mean that older properties are valued more per se, because we control for NOI. It is well known that depreciation mostly affects the NOI, see Bokhari

[20] This might be difficult to see from the estimates themselves at first sight. However, say we have property that is 100,000 square feet, with a \$ 100,000 NOI per year. The (log) predicted price per square foot taking the sellers results in Table B.2, ignoring the other variables is; $0.213 + 0.974 \times \ln 100,000\text{Sqft} + 0.966 \times \ln \frac{\$100,000\text{NOI}}{100,000\text{Sqft}} = 11.3$. If we only double the square feet, we will get an estimate log price per square foot of 11.3, or no difference. If we keep the size at 100,000 but we double the NOI we get an estimated (log) price per square foot of 12.0, a log difference of 0.7.

and Geltner (2018); Geltner and Bokhari (2015). We also find that properties of older vintages have higher holding periods, as shown in Figure A.3b.

[Place Figure A.3 about here]

[Place Figure A.4 about here]

The estimates of the year of sale dummies can be found in Figure A.4. During the Great Financial Crisis (GFC) property values decreased with 30% controlling for NOI. Given that we control for NOI, we cannot directly interpret this index as a asset value price index, as NOI is part of the real estate cycle. The index is also "constant liquidity" in the sense that we model "liquidity" (proxied by holding period) simultaneously (Van Dijk et al., 2018). Prices are now slightly over the previous peak. The city fixed effects are not given here to conserve space, but these are available upon request.

[Place Table B.3 about here]

From the other fixed effects in Table B.3, we conclude the following. All things equal, industrial properties trade with the lowest price per square foot (longitudinal model), but they also have the longest holding period (survival model). Apartments have both the shortest holding period, and the highest price per square foot. We do not find any significant effect of investor type on prices, other than a 4.5% discount when public investors sells a property. In contrast, we do find (large) significant effects of investor types on holding periods. (Other than the effects of the direct investors, which is insignificant for both buyers and sellers.) Non investors (see the sellers results) tend to hold on to their real estate the shortest, whereas public investors have the longest holding periods. The latter should come as no surprise, because most public real estate investors are virtually mandated to hold on to their properties, see Mühlhofer (2019). On the buyer-side, we find that the "non investors" and public firms tend to purchase properties from investors with long holding periods. The estimates of the MSA fixed effects are omitted to conserve space, but are available upon request.

3.2. The Effects of Seller and Buyer Characteristics on Prices and Holding Period

In this Section we will discuss the effects of buyer and seller characteristics on both prices and holding periods. Given that we estimate 92 parameters for all four buyer/seller characteristic (total asset value, Debt-to-Assets, informedness, skill) visualize the estimation results in Figures A.5 – A.8. We will also discuss the model fit in this Section, see Table B.4.

The effect of investors real estate portfolio sizes on prices and holding period. Figure A.5 gives the effect of the firms real estate portfolio size on prices and holding period, for both the buyers and sellers. First of all, the size of the portfolio hardly impact the sales price (Figure A.5a). The smallest investors do tend to sell for an approximate 3% discount, but this only counts for the smallest 10 quantiles or so. After that, there is no significant difference between the sales prices and the quantiles. On the buyer side, we do see a larger difference, see Figure A.5b. Generally speaking, the larger the investor, the larger the premium is. From Figure A.5c we find that the effect of investor size on holding period is U-shaped. "Medium" investors have the shortest holding period. The longer a property is held, the more likely it will be sold to a larger investor, see Figure A.5d.

[Place Figure A.5 about here]

The effect of debt to assets on prices and holding period. Figure A.6 gives the effect of the debt to asset ratios of firms on both prices and holding period. Note that the quantile distribution goes from a DTA of 0 (the first quantile) to a DTA of 1.3 (and higher, the 92^{nd} quantile), see x-axis in Figure. On the seller side, we find that firms with a high DTA sell properties for a discount (Figure A.6a), and hold the properties shorter (Figure A.6c). This findings are the not completely consistent with previous research in the single family housing market (Genesove and Mayer, 1997, 2001), who found that households with high Loan to Value (LTV) ratios only sold their properties for a higher price, *if* they would be able to sell it. (Resulting in higher time on market and more withdrawals.) However, our setting is very different from housing literature. Buyers that have a high DTA (i.e. "distressed" buyers) will only purchase a property is they can buy it for a good price, see Figure A.6a. Firms with a DTA of 1 or higher will purchase the same property with a 7% discount, compared to firms with a DTA of zero.

[Place Figure A.6 about here]

The effect of prior exposure on prices and holding period. Figure A.7 gives the effect prior exposure has on the price and holding period on real estate. Again, prior exposure ("PriorExposure") is measured as the number of properties a certain firm had in a predefined market (defined by our data provider) before the transaction, as a fraction of all the properties the firm owns. This number is therefore between 0 (no real estate in said market before the transaction) to 1 (all of the real estate the firm had before the transaction was in said market). The results are very clear. The more prior exposure a firm has in a market, the higher price it can bargain when selling (Figure A.7a), and lower price when buying (Figure A.7b). More specifically, if a firm had a prior exposure of 100% in a certain market, it will sell its real estate with a 2% *premium*, and buy real estate in said market with a 4% *discount*, compared to someone with 25% prior exposure. This is a clear indication that information asymmetry exists in commercial real estate, and that companies can "learn" about a market the more it invests in it. Firms with higher prior exposure also tend to hold on to its real estate longer, see Figure A.7c. Likewise, properties that are held for a long time, are more likely to sell to a firm with a lot of prior exposure in the corresponding market, see Figure A.7d.

[Place Figure A.7 about here]

The effect of management skill on prices and holding period. Our proxy for management skill (essentially the fraction of properties that are refinanced when interest rates go down) affect the price only modestly. We find a significant premium when selling of high skill management versus low skill management of around 4%, but only at the 10% level. When buying a property the results are less robust. Note from Figure A.8a that the effect of skill of the buyer on prices goes up and down, and that the differences are mostly insignificant from each other. We also find (Figure A.8c) that firms with a high skilled management also tends to increase the holding period. As such, the high skilled firms keep a property, but only sell it if they can bargain a higher price. We find the same for the buyer skill on prices (Figure A.8d. Properties with a long holding period a more likely to be sold to buyers with high management skill.

[Place Figure A.8 about here]

Model fit and hyperparameters. Finally, Table B.4 provides the variance parameters. As expected, we find that higher prices (measured by the random effects) results in shorter holding periods. See the positive coefficient on ρ. It should be noted that the effect is very small (although significant). A possible implication of this, is that most of the positive/price volume correlation found in literature actually goes through the income of real estate and less through the cap rates. Even though this is not the main topic of this research, it is worth investigating this more in future research. The variance on the random effects in the hazard model (σ_{θ_1}) is larger than those on the longitudinal model (σ_{θ_2}) by a factor of almost three. However, the models are so different from each other, that it is difficult to give any meaningful interpretation to this difference. The Weibull scale parameter is approximately 2 (ϕ) and the noise in the longitudinal model is 0.15 (σ_ϵ). The variance parameters on the second order random walk models are also given in Table B.4, although we will not discuss them for the sake of brevity. A more meaningful way of interpreting said hyperparameters is by looking at estimated trends in the different Figures A.5 – A.8.

[Place Table B.4 about here]

3.3. Robustness

As a robustness check we ran multiple models using only a subset of variables we have. We are interested in how "stable" the results are, and whether or not the shift when one of the other variables is omitted. For example, we ran separate models with only buyer and only seller characteristics. All results remained very similar to what was shown previously. It was also helpful to run the models without the buyer characteristics of the hazard model alone. For example, if the DTA estimate on sellers would change erratically, this could mean the matching between buyers and sellers would go through the DTA of the sellers. Unfortunately, we did not observe large shifts in any of the seller's characteristics estimates. Finally, we also ran ran four separate models, only including 1 of our 4 variables on buyer/seller characteristics. Also here the results remain robust. All of these results are available upon request.

4. Concluding Remarks

In this paper we analyze the underlying economic mechanisms that might be driving the observed patterns in commercial real estate prices, as an interplay between buyer and seller characteristics (in terms of their size, capital constraints, skill and market knowledge), and the timing and geographical location of these transactions. Given significant heterogeneity among institutional investors in real estate, we are provided with a natural laboratory to analyze the underlying economic mechanisms that might be driving the observed patterns in commercial real estate prices and associated probabilities of sale, as an interplay between the characteristics of buyers and sellers, in terms of their size, capital constraints, management skill and market informedness.

By jointly modelling the institutional investors' decision to invest in a particular real estate market (both in terms of property type and property location) and the effect of such decision on real estate prices using longitudinal and survival models, we find that, controlling for property characteristics, time-varying location specific characteristics (such as local demand) and buyer (seller) size, indebtedness and prior local market knowledge :

(1) Buyers of the largest quartile (portfolio wise) pay approximately 5% more for the otherwise identical property, compared to a similar investor from the smallest quartile. There is no significant different between the investor size quartiles when the property is sold. (Although the very smallest investors do sell with a modest discount.) This price premia at the time of purchase and sale cannot be explained by investors' financing constraints, nor by the level of information they have about the particular market prior to transaction, suggesting that investor private valuations have a significant role in pricing of commercial real estate.

In the spirit of Garmaise and Moskowitz (2003) we find that (2) the relationship between seller Debt-to-Asset ratio (DTA) and sales prices is almost flat, but that buyers faced with binding financing constraints, as measured by high DTA ratios, tend to be more careful and purchase otherwise identical assets at a discount. In both cases, however, the probability of transaction taking place decreases in the level of financial constraints faced by the sellers (buyers), consistent with the evidence from the housing markets (Genesove and Mayer, 1997, 2001). (3) Keeping investor size and investor financing constraints constant, more informed sellers tend to sell at a premium, while more informed buyers tend to buy at a discount, suggesting that information asymmetries are a significant driver of commercial real estate pricing, even when financing choice is taken into account. Furthermore, more informed sellers (buyers) are more likely to sell (buy), providing support to the argument that reduction in information asymmetry has a positive effect on asset liquidity (Ghent, 2019; Sagi, 2017). Finally (4) we find that real estate firms with more management skill sell properties for a premium. Although there is no evidence that high management skill results in a discount when buyer a property. We also find that firms with high management skill hold properties longer compared to their lower management skill counterpart.

Taken together, evidence shown in this paper points to economically significant effects of investor private valuations (unobserved preferences) and prior informedness on commercial real estate prices and holding period.

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

Agarwal, S., T. F. Sing, and L. Wang, 2018, Information asymmetries and learning in commercial real estate markets, *Available at SSRN: 3022705* .

Badarinza, C., T. Ramadorai, and C. Shimizu, 2019, Gravity, counterparties, and foreign investment, *Available at SSRN 3141255* .

Bokhari, S.J., and D.M. Geltner, 2018, Characteristics of depreciation in commercial and multi-family property: An investment perspective, *Real Estate Economics* 46, 745–782.

Brown, E. R., and J. G. Ibrahim, 2003, A bayesian semiparametric joint hierarchical model for longitudinal and survival data, *Biometrics* 59, 221–228.

Chinco, A., and C. Mayer, 2015, Misinformed speculators and mispricing in the housing market, *The Review of Financial Studies* 29, 486–522.

Clapham, E., P. Englund, J. M. Quigley, and C. L. Redfearn, 2006, Revisiting the past and settling the score: Index revision for house price derivatives, *Real Estate Economics* 34, 275–302.

Clapp, J. M., and C. Giaccotto, 1999, Revisions in repeat-sales price indexes: Here today, gone tomorrow, *Real Estate Economics* 27, 79–104.

Clapp, J.M., K.S. Bardos, and S.K. Wong, 2012, Empirical estimation of the option premium for residential redevelopment, *Regional Science and Urban Economics* 42, 240–256.

Clayton, J., D. C. Ling, and A. Naranjo, 2009, Commercial real estate valuation: fundamentals versus investor sentiment, *The Journal of Real Estate Finance and Economics* 38, 5–37.

Cox, D. R., and N. Reid, 1987, Parameter orthogonality and approximate conditional inference, *Journal of the Royal Statistical Society B* 55, 1–39.

Cox, D. R., and N. Reid, 1993, A note on the calculation of adjusted profile likelihood, *Journal of the Royal Statistical Society B* 55, 467–471.

De Wit, E. R., P. Englund, and M. K. Francke, 2013, Price and transaction volume in the Dutch housing market, *Regional Science and Urban Economics* 43, 220–241.

De Wit, E. R, and B. Van der Klaauw, 2013, Asymmetric information and list-price reductions in the housing market, *Regional Science and Urban Economics* 43, 507–520.

DiPasquale, D., and W. C. Wheaton, 1994, Housing market dynamics and the future of housing prices, *Journal of Urban Economics* 35, 1–27.

Faucett, C. L., and D. C. Thomas, 1996, Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach, *Statistics in medicine* 15, 1663–1685.

Fisher, J., D. Gatzlaff, D. M. Geltner, and D. Haurin, 2003, Controlling for the impact of variable liquidity in commercial real estate price indices, *Real Estate Economics* 31, 269–303.

Fisher, J., D. M. Geltner, and H. Pollakowski, 2007, A quarterly transactions-based index of institutional real estate investment performance and movements in supply and demand, *The Journal of Real Estate Finance and Economics* 34, 5–33.

Fisher, J., D. C. Ling, and A. Naranjo, 2009, Institutional capital flows and return dynamics in private commercial real estate markets, *Real Estate Economics* 37, 85–116.

Francke, M. K., and A. M. van de Minne, 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., and A. M. Van de Minne, 2017, Land, structure and depreciation, *Real Estate Economics* 45, 415–451.

Francke, M. K., and A. M. Van de Minne, 2019, Dealing with unobserved heterogeneity in hedonic price models, *Available at SSRN: 3249256* .

Garmaise, M. J., and T. J. Moskowitz, 2003, Confronting information asymmetries: Evidence from real estate markets, *The Review of Financial Studies* 17, 405–437.

Geltner, D., and J. Mei, 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., and S. Bokhari, 2015, Commercial buildings capital consumption in the United States, *Final Report, MIT Center for Real Estate, November* .

Geltner, D. M., A. Kumar, and A. van de Minne, 2018, Riskiness of real estate development: A perspective from urban economics & option value theory, *Real Estate Economics* online on 30 July 2018.

Geltner, D. M., and D. Ling, 2006, Considerations in the design and construction of investment real estate research indices, *Journal of Real Estate Research* 28, 411–444.

Geltner, D. M., N. G. Miller, J. Clayton, and P. M. A. Eichholtz, 2014, *Commercial Real Estate, Analysis & Investments*, third edition (OnCourse Learning).

Genesove, D., and C. Mayer, 1997, Equity and time to sale in the real estate market, *American Economic Review* 87, 255–269.

Genesove, D., and C. Mayer, 2001, Loss aversion and seller behavior: Evidence from the housing market, *The Quarterly Journal of Economics* 116, 1233–1260.

Ghent, A. C., 2019, What's wrong with pittsburgh? delegated investors and liquidity concentration .

Guo, X., and B. P. Carlin, 2004, Separate and joint modeling of longitudinal and event time data using standard computer packages, *The American Statistician* 58, 16–24.

Han, L., and W. C. Strange, 2015, The microstructure of housing markets: Search, bargaining, and brokerage, in *Handbook of regional and urban economics*, volume 5, 813–886 (Elsevier).

Harter-Dreiman, M., 2004, Drawing inferences about housing supply elasticity from house price responses to income shocks, *Journal of Urban Economics* 55, 316–337.

Henderson, R., P. Diggle, and A. Dobson, 2000, Joint modelling of longitudinal measurements and event time data, *Biostatistics* 1, 465–480.

Kurlat, P., and J. Stroebel, 2015, Testing for information asymmetries in real estate markets, *The Review of Financial Studies* 28, 2429–2461.

Lin, H., B. W. Turnbull, C. E. McCulloch, and E. H. Slate, 2002, Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer, *Journal of the American Statistical Association* 97, 53–65.

Martino, S., R. Akerkar, and H. Rue, 2011, Approximate Bayesian inference for survival models, *Scandinavian Journal of Statistics* 38, 514–528.

Mühlhofer, T., 2019, They would if they could: Assessing the bindingness of the property holding constraints for REITs, *Real Estate Economics* 47, 431–477.

Rizopoulos, D., 2011, Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data, *Biometrics* 67, 819–829.

Rosen, S., 1974, Hedonic prices and implicit markets: Product differentiation in pure competition, *The Journal of Political Economy* 82, 34–55.

Rue, H., S. Martino, and N. Chopin, 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 commercial real estate returns, Technical report, Working Paper, University of North Carolina-Chapel Hill.

Saiz, A., 2010, The geographic determinants of housing supply, *The Quarterly Journal of Economics* 125, 1253–1296.

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 University Press).

Van de Minne, A. M., M. K. Francke, D. M. Geltner, and R. White, 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., D. Geltner, and A. van de Minne, 2018, The dynamics of liquidity in commercial property markets:revisiting supply and demand indexes in real estate, *Journal of Real Estate Finance and Economics* .

Wang, Y., and J. M. G. Taylor, 2001, Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome, *Journal of the American Statistical Association* 96, 895–905.

Appendix A. Figures

Figure A.1: **Caprates for different markets in the US.** The capitalization rates can be downloaded from the website of Real Capital Analytics, for 100s of markets across the world.

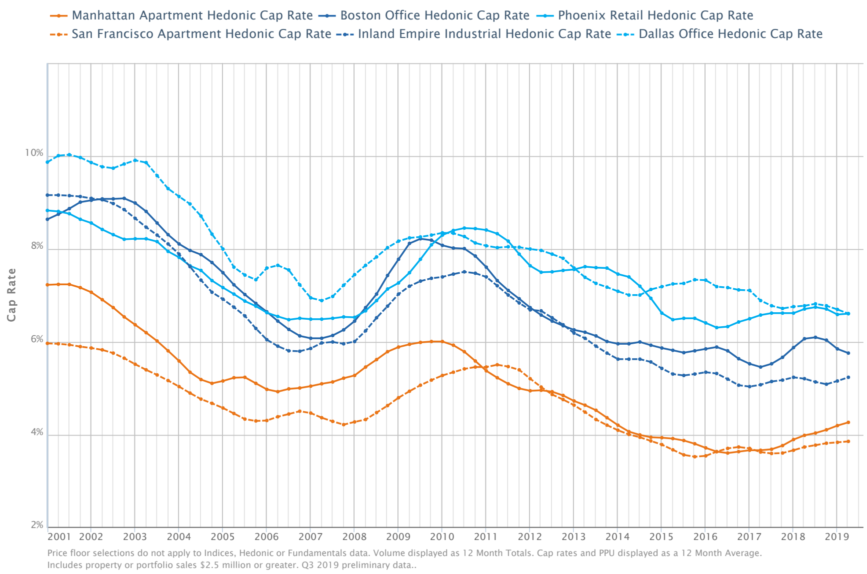


Figure A.2: **Debt-to-Asset (DTA) between 2006 and 2018.** These are the average of both the buyers *and* sellers aggregated. Numbers are based on transaction the dataset.

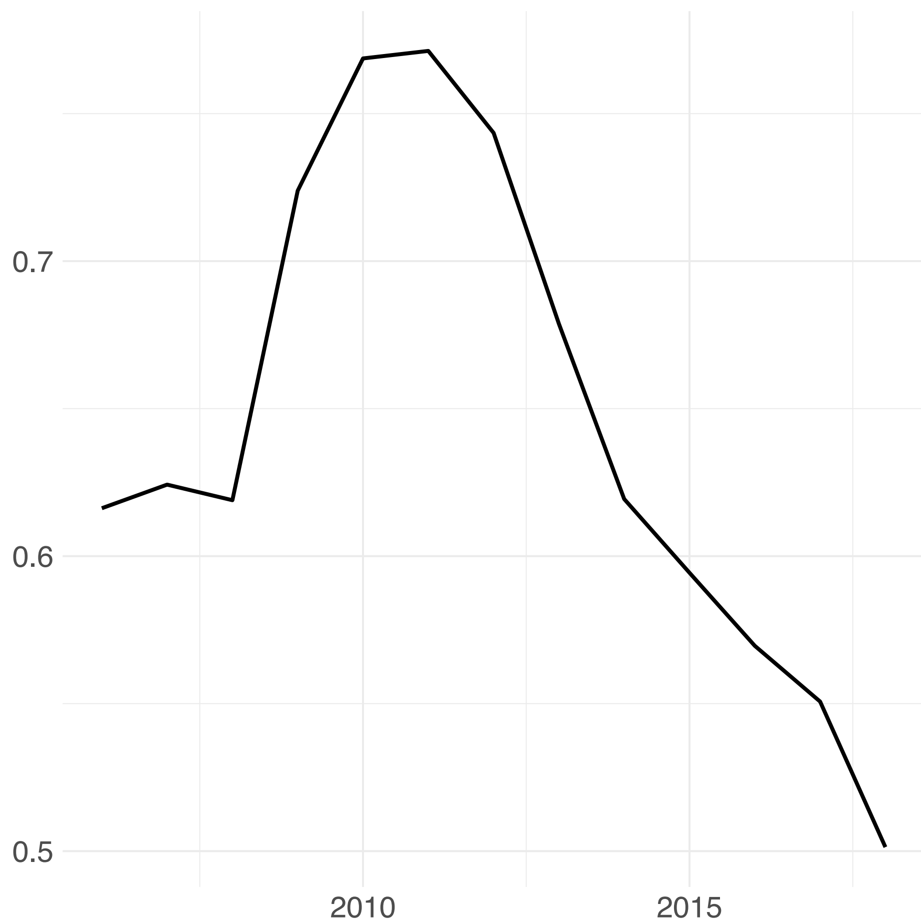
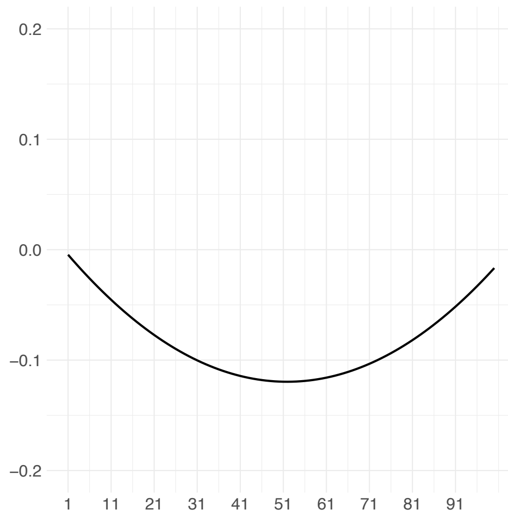


Figure A.3: **Effect of construction year.** y-axis gives the price per square foot discount or premium in log levels. The x axis gives the age of the properties in 2006. Shaded area give the 95% credible intervals. Note that the estimate of the first construction year (2006) is fixed to zero to avoid the dummy trap.

(a) Longitudinal model.



(b) Survival model.

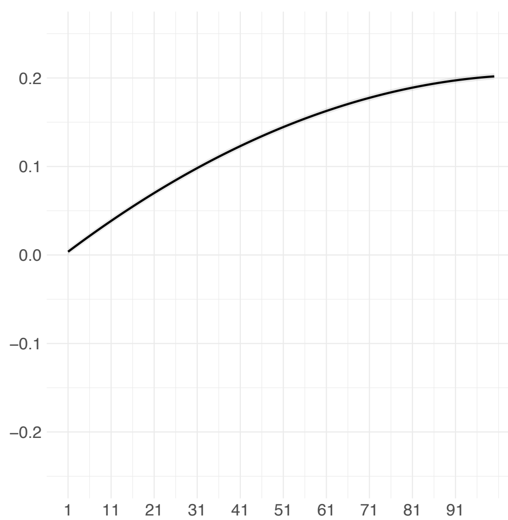


Figure A.4: **Year fixed effects of longitudinal pricing model.** y-axis gives the price per square foot discount or premium in log levels. Shaded area give the 95% credible intervals. Note that the estimate of the first year (2006) is fixed to zero to avoid the dummy trap.

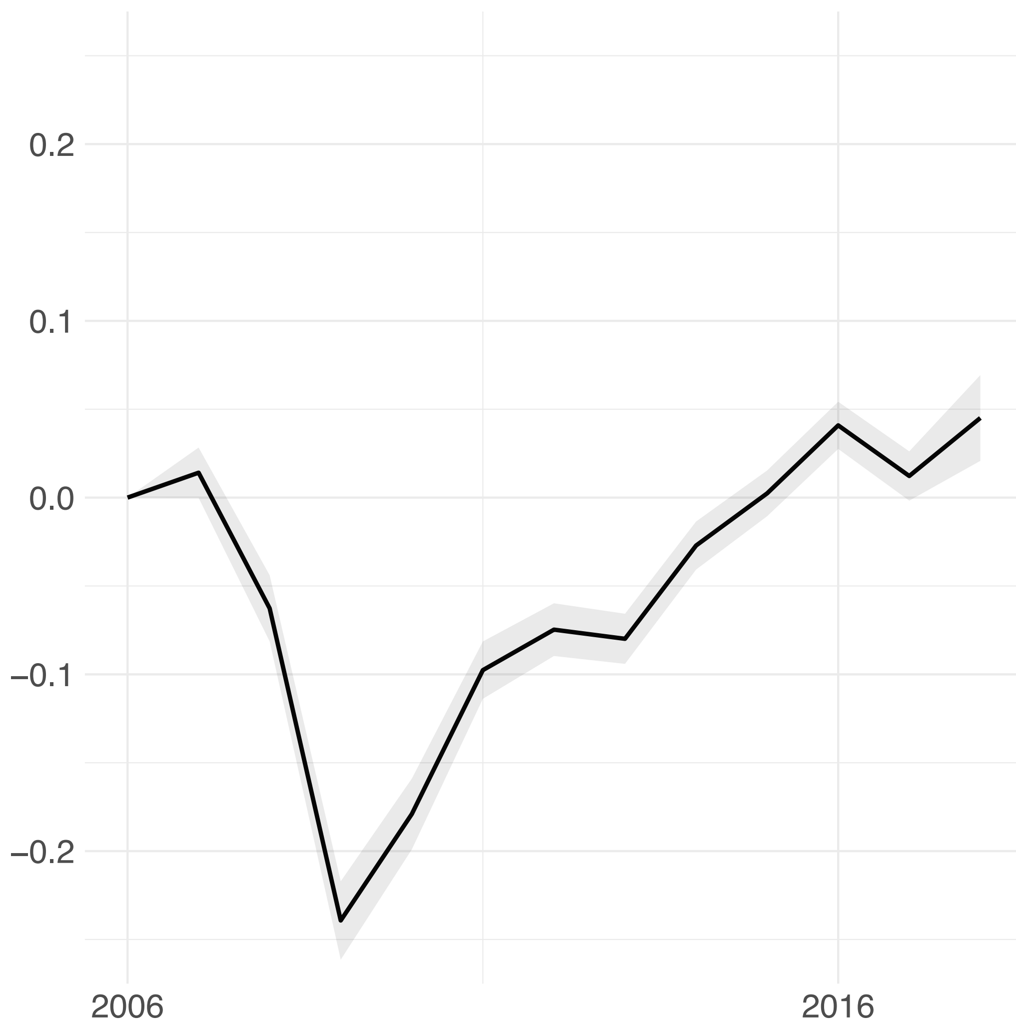
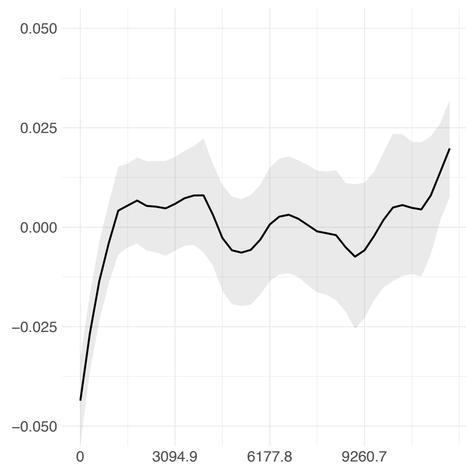
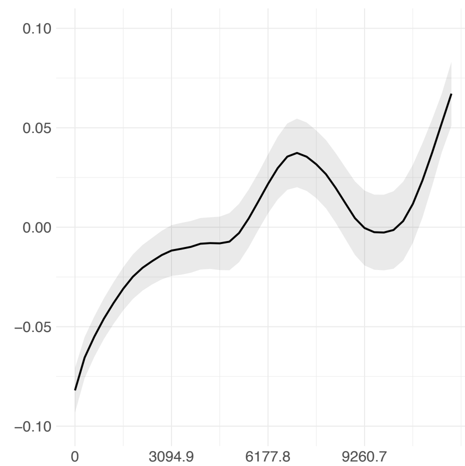


Figure A.5: **The effect of the value of the investor on prices and holding period.** Left column is for the sellers, whereas in the right column gives the results for the buyers. y-axis is the price per square foot premium or discount in log levels for the longitudinal model (top two panels), and log of the hazard ratio for the (i.e. probability to sell) for the survival model (bottom two panels). x-axis for the quantile distribution of the value of investors.

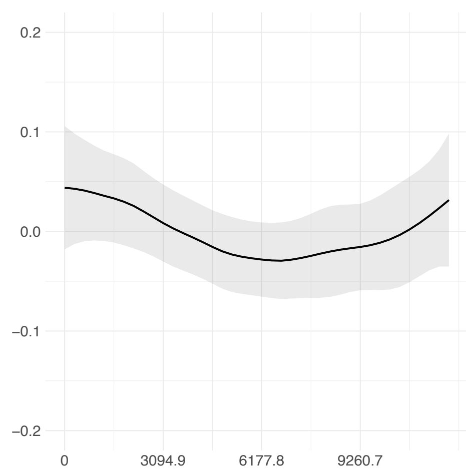
(a) Sellers' value on prices.



(b) Buyers' value on prices.



(c) Sellers' value on holding period.



(d) Buyers' value on holding period.

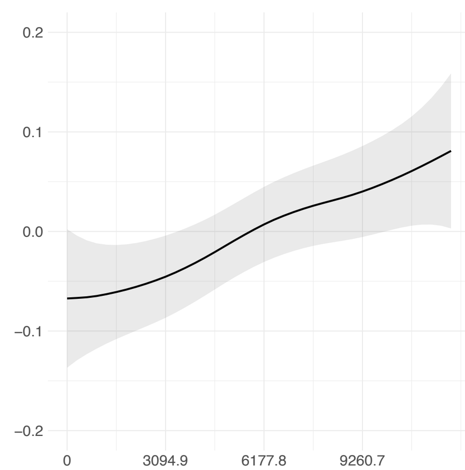
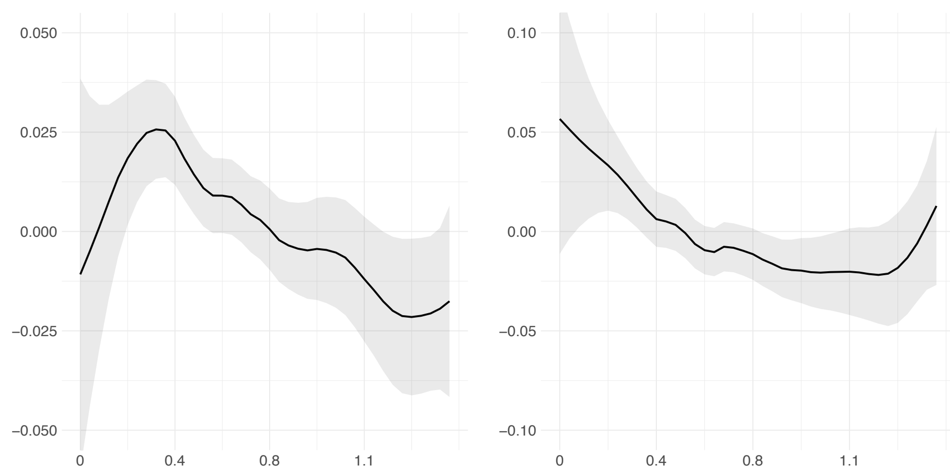


Figure A.6: **The effect of investors' Debt to Asset (DTA ratio on prices and holding period.** Left column is for the sellers, whereas in the right column gives the results for the buyers. y-axis is the price per square foot premium or discount in log levels for the longitudinal model (top two panels), and log of the hazard ratio for the (i.e. probability to sell) for the survival model (bottom two panels). x-axis for the quantile distribution of the value of investors.

(a) Sellers' DTA on prices. (b) Buyers' DTA on prices.



(c) Sellers' DTA on holding period. (d) Buyers' DTA on holding period.

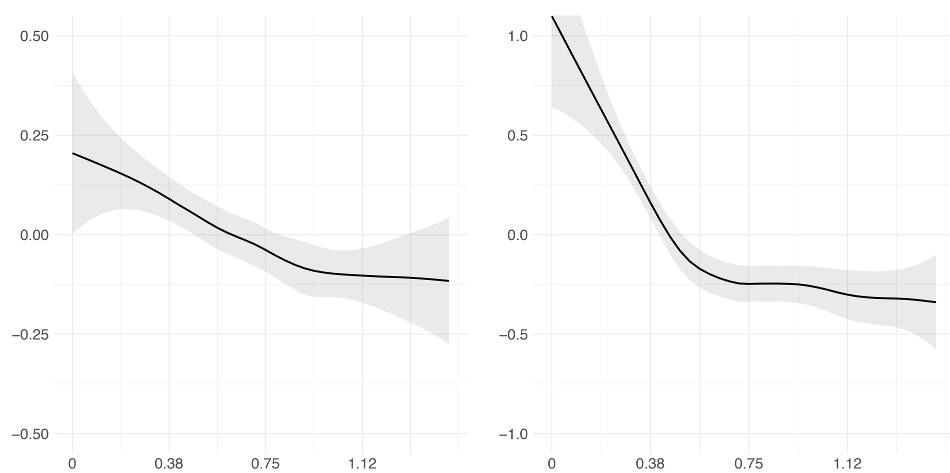
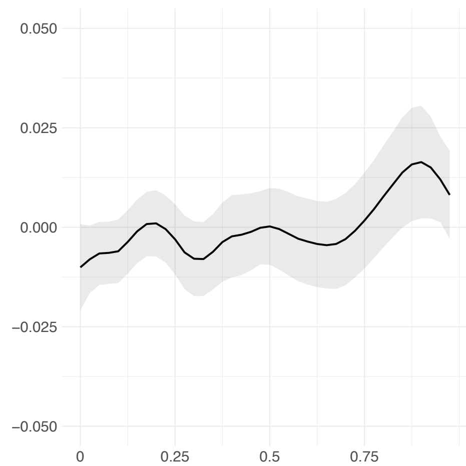
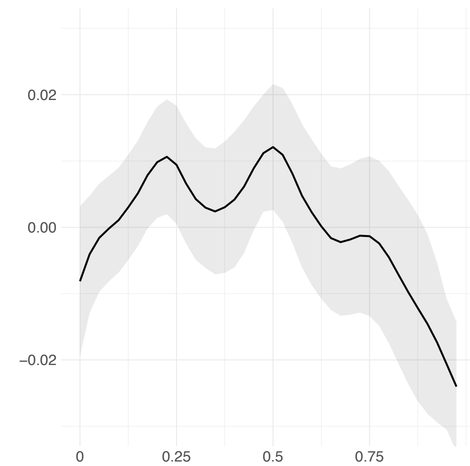


Figure A.7: **The effect of the investors' PriorExposure on prices and holding period.** Left column is for the sellers, whereas in the right column gives the results for the buyers. y-axis is the price per square foot premium or discount in log levels for the longitudinal model (top two panels), and log of the hazard ratio for the (i.e. probability to sell) for the survival model (bottom two panels). x-axis for the quantile distribution of the value of investors.

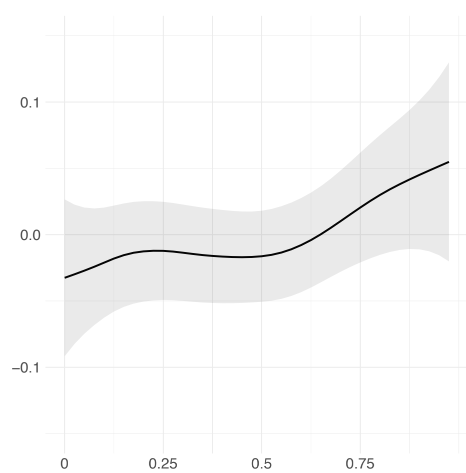
(a) Sellers' PriorExposure on prices.



(b) Buyers' PriorExposure on prices.



(c) Sellers' PriorExposure on holding period.



(d) Buyers' PriorExposure on holding period.

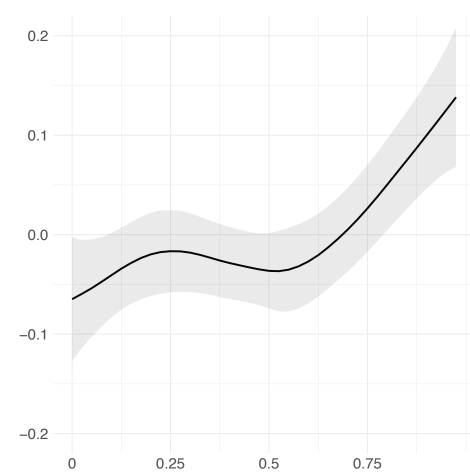
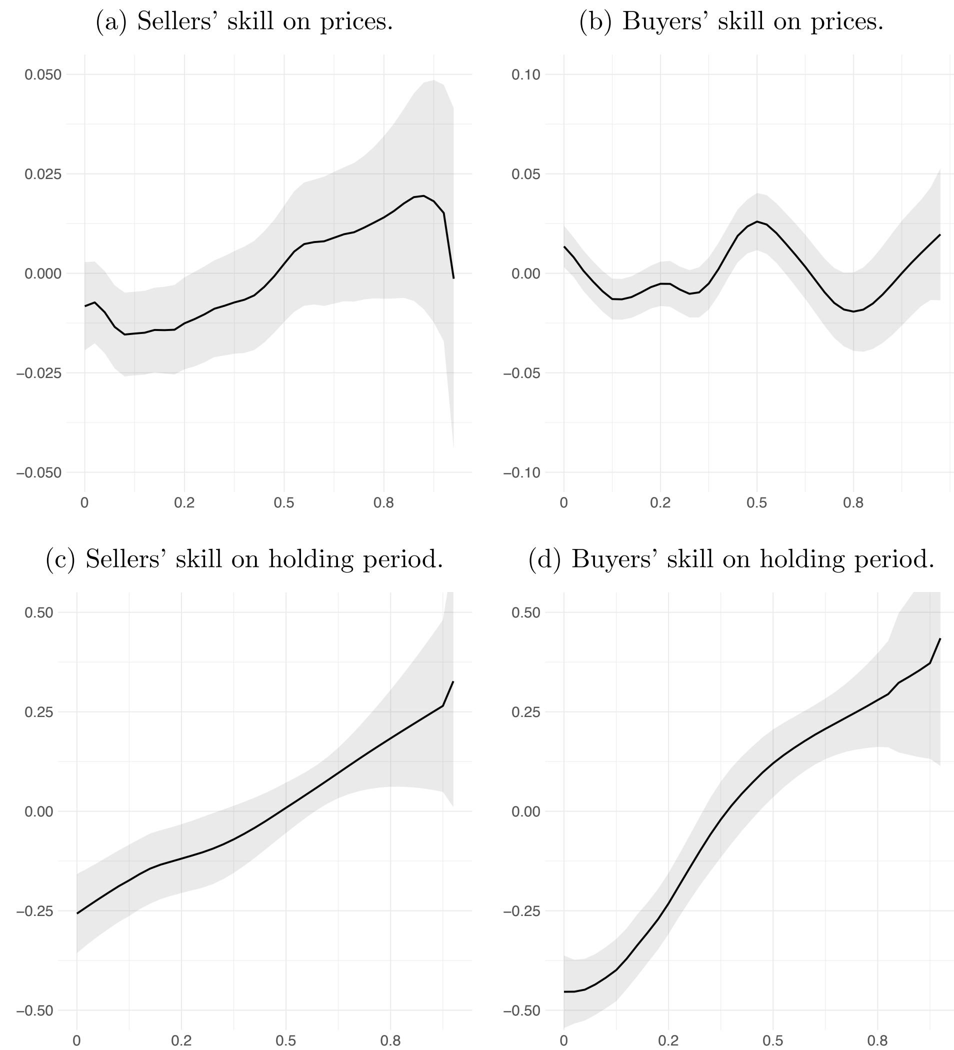


Figure A.8: **The effect of investor skill on prices and holding period.** Left column is for the sellers, whereas in the right column gives the results for the buyers. y-axis is the price per square foot premium or discount in log levels for the longitudinal model (top two panels), and log of the hazard ratio for the (i.e. probability to sell) for the survival model (bottom two panels). x-axis for the quantile distribution of the value of investors.



(a) Sellers' skill on prices.

(b) Buyers' skill on prices.

(c) Sellers' skill on holding period.

(d) Buyers' skill on holding period.

Appendix B. Tables

Table B.1: **Descriptive statistics.** *NOI* = property-level Net-Operating-Income. *Value* = value of investor real estate portfolio. *DTA* = total investor level debt divided by value of investor real estate portfolio. *PriorExposure* = the percentage of properties an investor has in a property-type- and location-specific market as a fraction of its total real estate portfolio. *Censored* is a dummy variable which equal 1, if the property is sold for the second time.

	Mean	**1st Quantile**	**3rd Quantile**
Price	\$37,206,087	\$7,750,000	\$37,800,000
$\frac{\text{NOI}}{\text{Square feet}}$	\$ 14.01	\$6.79	\$ 17.91
Square feet	186,982	55,500	248,160
Construction year	1988	1980	2002
Holding period (years)	6	3	8
Censored	0.121		
N	15,278		
	Seller characteristics		
Value	\$ 494,699,039	\$ 104,211,659	\$ 2,464,369,382
DTA	0.549	0.467	0.646
PriorExposure	0.213	0.075	0.500
Management	0.128	0.043	0.261
Delegated	3,264		
Direct	8,702		
Non Investor	627		
Public	2,703		
	Buyer characteristics		
Value	\$ 585,042,940	\$ 121,194,075	\$ 3,288,607,108
DTA	0.552	0.468	0.651
PriorExposure	0.200	0.067	0.500
Management	0.107	0.042	0.233
Delegated	2,727		
Direct	9,528		
Non Investor	264		
Public	2,759		

Table B.2: **Results of the Fixed Parameters.** We present the posteriors of the estimates with some statistics. The **mean** of the posteriors (i.e. the "estimate", the **sd** = standard deviation of the posterior, and the 2.5% and 97.5% credible intervals denoted **lower** and **higher** respectively. *Sqft* is square foot of structure, *NOI* is property-level Net-Operating-Income, and *Construction year* is the amount of years between the year built and the start of our sample (2006). The survival parameters are rescaled by $\tilde{\beta} = \frac{-\beta}{\phi}$, with ϕ being the Weibull scale parameter.

	mean	**sd**	**lower**	**higher**
Longitudinal				
(Intercept)	0.321	0.021	0.279	0.363
ln Sqft	0.942	0.006	0.930	0.954
$\ln \frac{\text{NOI}}{\text{Sqft}}$	0.955	0.005	0.946	0.964
Construction year	-0.005	0.000	-0.005	-0.004
(Construction year)^2	0.000	0.000	0.000	0.000
Year Fixed Effects		*Yes*		
City Fixed Effects		*Yes*		
Property Type Fixed Effects		*Yes*		
Buyer/Seller Type Fixed Effects		*Yes*		
Survival				
(Intercept)	3.065	0.138	2.803	3.336
ln Sqft	-0.109	0.042	-0.192	-0.027
$\ln \frac{\text{NOI}}{\text{Sqft}}$	0.070	0.034	0.003	0.137
Construction year	0.004	0.002	-0.001	0.008
(Construction year)^2	0.000	0.000	0.000	0.000
Baseline		*Holding period (in years)*		
City Fixed Effects		*Yes*		
Property Type Fixed Effects		*Yes*		
Buyer/Seller Type Fixed Effects		*Yes*		
N		15,278		

Table B.3: **Property type and buyers/seller characteristics results.** We present the posteriors of the estimates with some statistics. The **mean** of the posteriors (i.e. the "estimate", the **sd** = standard deviation of the posterior, and the 2.5% and 97.5% credible intervals denoted **lower** and **higher** respectively. The survival parameters are rescaled by $\tilde{\beta} = \frac{-\beta}{\phi}$, with ϕ being the Weibull scale parameter. *ref* is the reference category to avoid the dummy trap.

		mean	**sd**	**lower**	**higher**
	Longitudinal				
Property type:	Apartment (*ref*)	0.000	0.000	0.000	0.000
	Industrial	-0.168	0.008	-0.183	-0.152
	Office	-0.145	0.007	-0.159	-0.132
	Retail	-0.119	0.007	-0.132	-0.107
Buyers:	Delegated (*ref*)	0.000	0.000	0.000	0.000
	Direct	-0.005	0.006	-0.017	0.008
	Non Investor	0.021	0.017	-0.012	0.054
	Public	-0.013	0.007	-0.027	0.001
Sellers:	Delegated (*ref*)	0.000	0.000	0.000	0.000
	Direct	-0.005	0.006	-0.017	0.007
	Non Investor	0.003	0.012	-0.019	0.026
	Public	-0.045	0.007	-0.059	-0.032
	Survival				
Property type:	Apartment (*ref*)	0.000	0.000	0.000	0.000
	Industrial	0.775	0.069	0.646	0.910
	Office	0.254	0.044	0.167	0.341
	Retail	0.407	0.050	0.311	0.505
Buyers:	Delegated (*ref*)	0.000	0.000	0.000	0.000
	Direct	0.060	0.041	-0.021	0.140
	Non Investor	0.405	0.188	0.085	0.774
	Public	0.291	0.055	0.185	0.400
Sellers	Delegated (*ref*)	0.000	0.000	0.000	0.000
	Direct	0.061	0.037	-0.012	0.135
	Non Investor	-0.160	0.094	-0.333	0.025
	Public	0.261	0.048	0.168	0.355

Table B.4: **Results of the Hyperparameters.** We present the posteriors of the estimates with some statistics. The **mean** of the posteriors (i.e. the "estimate", the **sd** = standard deviation of the posterior, and the 2.5% and 97.5% credible intervals denoted **lower** and **higher** respectively. DTA is investor-level Debt-to-Asset ratio, VAL is the value of the investor real estate portfolio, and EXP the percentage of properties an investor has in a property-type- and location-specific market as a fraction of its total real estate portfolio.

		mean	**sd**	**lower**	**higher**
Measurement:	σ_ϵ (Noise)	0.148	0.003	0.143	0.153
	ϕ (Weibull scale)	2.001	0.050	1.897	2.093
	σ_{θ_1} (random effect - long)	0.111	0.004	0.104	0.118
	σ_{θ_2} (random effect - surv)	0.296	0.044	0.226	0.398
	ρ	0.033	0.005	0.043	0.023
		Longitudinal			
Sellers:	$\sigma_\xi^{\mathrm{DTA}}$	0.005	0.001	0.003	0.007
	$\sigma_\xi^{\mathrm{VAL}}$	0.006	0.002	0.003	0.010
	$\sigma_\xi^{\mathrm{MAN}}$	0.005	0.001	0.003	0.008
	$\sigma_\xi^{\mathrm{EXP}}$	0.004	0.001	0.003	0.007
Buyers:	$\sigma_\xi^{\mathrm{DTA}}$	0.005	0.002	0.003	0.009
	$\sigma_\xi^{\mathrm{VAL}}$	0.006	0.002	0.003	0.010
	$\sigma_\xi^{\mathrm{MAN}}$	0.005	0.002	0.003	0.009
	$\sigma_\xi^{\mathrm{EXP}}$	0.004	0.001	0.003	0.007
		Survival			
Sellers:	$\sigma_\xi^{\mathrm{DTA}}$	0.009	0.004	0.004	0.020
	$\sigma_\xi^{\mathrm{VAL}}$	0.007	0.003	0.003	0.015
	$\sigma_\xi^{\mathrm{MAN}}$	0.010	0.006	0.004	0.025
	$\sigma_\xi^{\mathrm{EXP}}$	0.007	0.003	0.003	0.013
Buyers:	$\sigma_\xi^{\mathrm{DTA}}$	0.032	0.013	0.015	0.064
	$\sigma_\xi^{\mathrm{VAL}}$	0.007	0.003	0.003	0.013
	$\sigma_\xi^{\mathrm{MAN}}$	0.018	0.012	0.004	0.049
	$\sigma_\xi^{\mathrm{EXP}}$	0.009	0.004	0.004	0.020