Non-risk Determinants of Investment Returns: Quality, Deal Size, and Returns of Commercial Real Estate

Liang Peng*

Smeal College of Business
The Pennsylvania State University
University Park, PA 16802
Phone: (814) 863 1046
Fax: (814) 865 6284
Email: liang.peng@psu.edu

April 2019

Abstract

For many important assets in the economy that have high values, are indivisible, opaque, costly to acquire, and cannot be shorted, individual deals' attributes may affect their investment returns for reasons that are related to risk. This paper empirically tests whether commercial real estate's investment returns are related to properties' quality, which is measured with net operating income per square foot, and their deal size, which is measured with acquisition prices. Analyzing 6,215 properties with a total value of $230 billion over the 1977 to 2017 period, I find that properties with higher quality tend to have higher returns, both before and after adjusting for risk, though such returns diminish over time. I also find that larger deals tend to have lower returns, both before and after adjusting for risk.

Key words: Commercial real estate, investment returns, quality, and deal size

JEL classification: G12, R33

---

* I thank Jeff Fisher, Andy McCulloch, and Bob White for numerous constructive suggestions and comments. I thank the Real Estate Research Institute for a research grant and thank NCREIF for providing commercial real estate data. All errors in this paper are my sole responsibility.

1. Introduction

Risk is a key determinant of investment returns and is almost the sole focus of the entire asset pricing literature. However, non-risk attributes may also affect investment returns of many important assets in the economy that have high values, are indivisible, opaque, costly to acquire, and cannot be shorted. Such assets include housing, commercial real estate, private equity, and venture capital, and they are significant components of the total wealth in the economy (see, e.g. Case and Shiller (1989), Cochrane (2005), Kaplan and Schoar (2005), Chetty and Szeidl (2007), Plazzi, Torous and Valkanov (2010), Franzoni, Nowak and Phalippou (2012), Piketty and Zucman (2014), Hochberg and Mühlhofer (2015), Peng (2016), Chetty, Sandor and Szeidl (2017), Korteweg and Soren (2017), and Sagi (2017)). The importance of housing alone has been manifested in the recent global financial crisis. For these assets, each deal requires significant amount of capital to acquire, has high due diligence costs, and constitutes a large portion of an investor's portfolio. Therefore, individual deals' characteristics, despite not necessarily related to risk, may affect asset prices and returns. However, the non-risk determinants of investment returns for these assets have been largely ignored in the literature.

This paper aims to fill the void by empirically testing whether commercial real estate's investment returns, before and after adjusting for risk, are related to individual properties' two notable attributes, their quality and deal size, at the time when they are acquired. While "quality" is a generic concept, in this paper, it refers to a property's ability to generate net income, which is specifically measured with net operating income (NOI) per square foot. I measure deal size with properties' acquisition prices.[1]

This paper analyzes quality because it is essentially a proxy for location, which is arguably the single most important determinant for real estate values. Space in more scarce and highly sought locations tends to generate higher rental income per square foot in the user/rental market, and thus higher price per square foot in the property market. However, the literature is virtually silent on whether quality should be related to

[1] Virtually all properties in the sample were sold as individual deals, not in bundles, so acquisition prices seem to be a reasonable measure for deal size.

properties' investment returns theoretically and whether it is indeed related to returns empirically, for reasons related to risk or not. This paper aims to provide direct empirical evidence on whether properties' quality is related to their investment returns, before and after adjusting for risk.

It is apparent that deal size may affect returns for risk-related reasons. For example, larger deals likely constitute larger portions of investors' portfolios and thus may expose investors to more liquidity risk or more non-systematic risk. As a result, investors may demand higher returns for investing in larger deals. This paper controls for properties' systematic and nonsystematic risk that is related to deal size, and tests whether deal size affects returns for non-risk-related reasons.

Deal size may positively or negatively affect commercial real estate investment returns for at least two reasons that have nothing to do with risk. First, due to the large amount of capital required to acquire real estate, financial/capital constraints, which are typically inconsequential for investing in stocks and bonds, become relevant. Fewer investors can afford larger deals; therefore, possible abnormal returns, if exist at all, are less likely traded away for larger deals due to fewer able buyers and thus less competition among them. As a result, larger deals may have higher returns. Evidence consistent with this notion has been found in the housing market. Built on the emerging literature of the "risk segmentation" of housing (see, e.g. Peng and Thibodeau (2013), and Hartman-Glaser and Mann (2016), Peng and Thibodeau (2017), and Peng and Zhang (2019)), Peng and Zhang (2019) uses micro-level data to find that more expensive houses have higher price appreciation rates, both before and after adjusting for risk, if they are held for longer than about two years.

Second, investors may also be willing to accept lower returns for investing in larger deals, as doing so may achieve economy of scale in due diligence costs. Since the real estate market is opaque, it is costly to conduct due diligence for each deal, regardless their size. For investors who need to deploy a large amount of capital, the saving in due diligence cost achieved by investing into a small number of larger deals might overweight losses due to their lower returns. Consequently, investors can be better off investing in larger deals with lower returns. This, once again, has nothing to do with risk. The economy of scale gained by investing in larger deals is similar in spirit to the phenomenon that stock investors with limited or scarce attention may process more market and sector-wide information than firm-specific information (Peng and Zhang (2019)). Note that homebuyers cannot achieve similar economy of scale as their primary goal is searching for one house, not deploying a fixed amount of capital.

It is an empirical question whether deal size is positively or negatively related to investment returns before and after adjusting for risk. What this paper focuses on is testing whether deal size has a net aggregate effect on individual properties' investment returns, not distinguishing different channels through which deal size may affect returns.

My analysis uses a high-quality proprietary database of commercial real estate maintained by the National Council of Real Estate Investment Fiduciaries (NCREIF). The dataset covers 36,718 properties with the total value at acquisition being about $1.1 trillion (in 2017:Q2 dollars) over a period of 40 years (from 1977:Q2 to 2017:Q2). It contains detailed information regarding individual properties' time invariant attributes, including their locations, square footage, investors/managers, types, etc. It also contains quarterly cash flow and valuation, which allows me to calculate properties' investment returns. The final clean sample used in this paper consists of 6,215 properties with a total value of about $230 billion.

I first test whether individual properties' modified internal rates of returns (MIRRs) during holding periods are related to their quality and deal size. With each property's type, investor/manager, location (the Core Business Statistical Area where it is located), and acquisition period being controlled, I find that MIRRs are higher for properties with higher quality and smaller deal size. This result is generally robust across properties that have been sold and those that had not been sold. This result is also generally robust across four main property types – apartment, industrial, office, and retail – and subsamples with different quality (high vs. low) and deal size (large vs. small).

I further use a holding-period total return model, which is an extension of the repeat sales regression that is widely used to construct house price indices (see, e.g. Bailey, Muth and Nourse (1963), Case and Shiller (1989), Goetzmann (1992), and Peng (2012)), to control for the average market-wide investment returns during each property's holding period, in addition to controlling for its type, investor/manager, and location, and allow both quality and deal size to affect both *per-investment* and *per-period* abnormal returns. I find that high quality properties have positive per-investment abnormal returns but negative per-period abnormal returns. In other words, they have positive abnormal returns that diminish over time. Regarding deal size, I find that larger deals tend to have negative per-investment abnormal returns, which is not affected by the duration of holding periods.

I then test whether risk-adjusted returns are related to quality and deal size. I measure a property's systematic risk with its factor loadings in conventional asset pricing factor models, and measure its non-systematic risk with two variables: an ex post measure that I call "peculiar risk", which equals the squared residual from estimating a factor model, and the property's acquisition cap rate (see, e.g. Peng (2019) for evidence that cap rate predicts future risk) as an ex ante "catch-all" risk measure that helps capture non-systematic risk.

A central econometric challenge is that I do not observe time series of returns for each property, so I am unable to use time series regressions to estimate each property's factor loadings. I overcome this challenge with an approach proposed by Peng (2019) in the framework of holding-period factor models. This approach allows individual properties' factor loadings as well as their risk-adjusted returns to be functions of their attributes. It also allows individual properties to have different exposure to non-systematic risk. This approach allows me to directly test whether properties' risk-adjusted returns are related to their quality and deal size, while allowing their systematic risk (i.e. factor loadings) as well as exposure to non-systematic risk to also be functions of these two variables.

Another challenge is that real estate might have unknown factors that differ from common stock and bond factors. To mitigate this problem, I use residuals from fitting the data to a model that contains stock and bond factors to construct a "real estate factor", which is essentially an index of real estate specific risk premium that cannot be explained by the stock and bond factors. I validate this real estate factor by showing that it has strong out-of-sample explanatory power, and include it as an additional factor in the holding period factor models I estimate.

I find that properties with higher quality have positive *per-investment* risk-adjusted returns but negative *per-period* risk-adjusted returns. In other words, they have positive risk-adjusted returns that diminish and may even become negative over time. I also find that larger deals have negative risk-adjusted returns that do not seem to be affected by the duration of their holding periods. These results are robust when I include the real estate factor as well as peculiar risk or cap rates in the models. These novel results constitute original contributions to the asset pricing literature by showing that assets' attributes, such as quality and deal size, may have significant impact on investment returns of many important assets in the economy that have high values, are indivisible, opaque, costly to acquire, and cannot be shorted, for reasons that are not related to risk.

The rest of this paper is organized as follows. Next section describes the data. The third section tests whether properties' investment returns, before adjusting for risk, are related to their quality and deal size. The fourth section tests whether properties' returns, after adjusting for both systematic and non-systematic risk, are related to quality and deal size. The last section concludes.

2. Data

2.1. Database and main variables

This paper uses the proprietary dataset of the National Council of Real Estate Investment Fiduciaries (NCREIF). NCREIF is a not-for-profit real estate industry association, which collects, processes, and disseminates information on the operation and transactions of commercial real estate. Its members are typically investment companies, pension funds, and life insurance companies.[2] The database contains information on property attributes, such as property type, street address, square footage, etc., as well as quarterly financial and accounting information for each property. Subsets or earlier releases of this dataset have been used in Pivo and Fisher (2011), Plazzi, Torous and Valkanov (2012), Peng (2016), Gang, Peng and Thibodeau (2017), and Sagi (2017). This paper uses the 2017:Q2 release of the database, which consists of 36,718 properties invested or managed by NCREIF members from the third quarter of 1977 to the second quarter of 2017.

A caveat of using this dataset is worth noting: properties appear in this dataset because members of NCREIF have invested in them. Therefore, they are a selected/conditional sample. Consequently, while the results found in this paper might reflect general economic relationships that apply to all commercial real estate, I am unable to rule out the possibility that they may also be driven by institutional investors' behavior. In this sense, readers should be cautious when trying to generalize this paper's findings.

This paper focuses on three features of each property. The first is its investment performance, which I measure with the Modified Internal Rate of Return (MIRR) as well as the corresponding total return during its holding period. The second is its "quality", which I measure with the net operating income per square foot when it was acquired. The third is its deal size, which I measure with the acquisition price. Other information used for each property includes the NCREIF member who invests in or manages the property, the CBSA where it is located, type (apartment, industrial, office, or retail), acquisition and holding periods, and its acquisition cap rate.

I calculate each property's modified IRR and total return during its holding period if it had been sold by the end of the sample period. For each property i, I calculate its quarterly Modified IRR, which is denoted by $MIRR_i$, using its quarterly cash flow series. The cash flows consists of the acquisition cost in the acquisition quarter,[3] the NOI plus

[2] Examples of NCREIF members are Blackrock, Citi group, TIAA, New York Life, Invesco, Heitman/JMB, and Cornerstone real estate advisers.

[3] We assume that all acquisitions and dispositions take place at the end of quarters. For a small number of properties, the database shows positive net operating income in the recorded acquisition quarters, possibly

proceeds from partial sales minus capital expenditures in subsequent quarters, and the net sale proceeds plus NOI minus capital expenditures in the disposition quarter. To obtain the two rates needed for the calculation of MIRRs – the financing rate and the reinvestment rate – I calculate quarterly equally weighted average total returns for each property type, using appraised values unless transaction prices are observed. I use return series of these indices as financing and reinvestment rates. Note that MIRRs are similar to IRRs but seem superior in measuring real estate returns, because the present value equations of commercial real estate investments often have multiple IRR solutions, mainly due to long holding periods and irregular cash flows. Results in this paper are robust when I use IRRs in all analyses.[4] After calculating MIRRs, I calculate the holding period total gross return, which is denoted by R_i, as

$$R_i = \left(1 + MIRR_i\right)^{sell_i - buy_i}, \tag{1}$$

where $sell_i$ is the quarter when the property is sold, and buy_i is the quarter when the property is acquired.

Using the same approach described above, I calculate quarterly MIRRs and total returns over a five-year holding period since acquisition for unsold properties and include these properties in my analyses. Including unsold properties helps mitigate a possible sample selection problem that sold properties can be selected samples if disposition decisions are related to returns (see, e.g. Gatzlaff and Haurin (1997), Gatzlaff and Haurin (1998), Fisher, Gartzlaff, Geltner and Haurin (2003), Goetzmann and Peng (2006), Korteweg and Sorensen (2010), and Sagi (2017)). Since these properties hadn't been sold, I use their appraised values (minus estimated selling cost) at the end of year 5 since acquisition as the net sale proceeds. In unreported robustness checks, I calculate and use three-year total returns too, which provide robust results.

---

because their acquisitions took place in the middle of those quarters. For these properties, we assume the acquisitions took place at the end of the previous quarters.

[4] When there are multiple solutions for total return IRRs for a property, I select the smallest one from all solutions that are higher than the capital appreciation IRR, which is unique for each property.

I use the net operating income (NOI) per spare foot in the first year after acquisition to measure each property's quality. It is important to note that NOI varies across time and is affected by inflation as well as real estate rental market conditions. Therefore, the same dollar amount in different time periods is not comparable. To allow NOI to be comparable across time, I construct a NOI growth index for each property type to reflect inflation and changing rental market conditions, and use these indices to inflate NOI in each quarter for each property to 2017:Q2 dollars. Specifically, for each period, I first calculate the NOI growth rate for each property. I then remove properties with growth rates lower than -80% or higher than 100%, which likely indicate data errors, as well as those with growth rates in the lowest and highest 5 percentiles of the remaining properties. I then calculate the across-property equally weighted average NOI growth rate as the NOI growth index if there are at least 60 properties available for this calculation. I use the NOI index to inflate each property's NOI in each quarter to 2017:Q2 dollars. I then calculate the sum of NOI in the first four quarters after acquisition (in 2017:Q2 dollars) and divide it by the gross square feet of the property. Note that I also apply the above algorithm to calculate the rent in the first year after acquisition per square foot, which I use in robustness checks and data cleaning.

I use the acquisition price to measure deal size. It is apparent that prices are also affected by inflation and property market conditions, so the same dollar amount in different time periods is not comparable either. I following the same algorithm used to calculate the NOI growth index to construct price appreciation index for each property type – this time using transaction prices and appraised prices when transaction prices are not available. I use these price appreciation indices to inflate all acquisition prices to 2017:Q2 dollars.

I calculate the acquisition cap rate for each property whenever the data permit. The cap rate of property i acquired at the end of quarter t, denoted by $C_{i,t}$, is defined as

$$C_{i,t} = \frac{\sum_{s=t+1}^{t+4} NOI_{i,s}}{P_{i,t}} \tag{2}$$

where $P_{i,t}$ is the acquisition price and $NOI_{i,s}$ is the quarterly net operating income. I am able to calculate acquisition cap rates for 18,543 properties but not for others due to missing prices or income. This paper uses cap rates as a proxy for investors' ex ante risk, as Peng (2019) shows that cap rates help predict individual properties' ex post investment returns and risk. To mitigate biases due to a mechanical relationship between investment returns and cap rates as the acquisition price is used in both calculations, I also calculate cap rates based on appraised values four quarters after the acquisition, which are highly correlated with cap rates and are used in our regressions.

2.2. Data cleaning and summary

I clean the data using the following procedure. I first remove 4,422 properties with missing inflated purchase price. An inflated purchase price may be missing either because the purchase price itself is missing or because the property price index used to inflate prices has missing values during the property's holding period. I then remove 1,896 properties with inflated purchase price lower than $2 million, which are too small deals compared with most properties in the sample. After that, I remove 14,279 properties with missing inflated NOI in the first year after acquisition, due to missing NOI or missing NOI index values, and 949 properties with negative inflated NOI in this year, which is likely due to data errors or atypical investments. There are 14,613 properties remaining in the sample.

I then apply more filtering rules. For a property to stay in the final sample, it must satisfy *all* of the following conditions: (1) it must belong to one of the four main property types: apartment, industrial, office, and retail; (2) each of the following variables must be within 3.5 standard deviation of the mean of the cross-property distribution of the variable: log NOI (in 2017:Q2 dollars) in the first year after acquisition, log *per-square-foot* NOI (in 2017:Q2 dollars) in the first year after acquisition, log *per-square-foot* inflated rent (in 2017:Q2 dollars) in the first year after acquisition, log *per-square-foot* purchase price (in

2017:Q2 dollars); and (3) the holding period total return of the property must be reasonable.[5] The final sample consists of 6,540 properties.

Table 1 reports basic statistics of the clean sample of 6,540 properties located in 227 different Core Business Statistical Areas (CBSAs), including 4,414 properties that had been sold by the end of the sample period and 2,126 properties that had not been sold. The properties consist of 1,465 apartment, 2,382 industrial, 1,601 office, and 1,092 retail properties. This table presents the minimum, 25 percentile, median, 75 percentile, maximum, mean, and standard deviation of properties' annualized MIRRs, first year NOI (2017:Q2 dollars) per square foot, acquisition price (2017:Q2 dollars), and the duration of properties that had been sold.

I visualize the main variables in Figures 1, 2, and 3, which respectively plot the histograms of the annualized MIRRs, the year 1 NOI per square foot (2017:Q2 dollars, log values), deal size (20l7:Q2 dollars, log values) of all the 6540 properties. Figure 4 plots the duration (log quarters) of the holding periods of sold properties.

Table 2 reports correlations between the main variables for the whole sample and each of the four main property types. For the whole sample, the correlation is 0.15 between the annualized MIRR and NOI per square foot (2017:Q2 dollars, log values), is -0.14 between the annualized MIRR and deal size (acquisition price in 2017:Q2 dollars, log values), and is -0.32 between the annualized MIRR and duration (log quarters). These numbers seem to suggest that investment returns are higher for properties with higher quality, smaller deal size, and shorter holding periods. The correlation is 0.10 between NOI per square foot and deal size, is 0.21 between NOI per square foot and duration, and is -0.09 between deal size and duration. Table 2 also shows that the correlations between main variables are similar across all four types of properties.

[5] The *quarterly* total return MIRR must be higher than -10% and lower than 40%. Further, a property's quarterly total return MIRR must be highly correlated with the same property's quarterly capital appreciation MIRR. Specifically, its residual from a linear regression of capital appreciation MIRRs against total return MIRRs needs to be within three standard deviations of the mean of residuals.

I further use Figure 5 to visualize the relationship between year 1 NOI per square foot and annualized MIRRs, which seems to confirm the positive correlation between them. Figure 6 plots annualized MIRRs against deal size, which appears to confirm their negative relationship. Figure 7 plots annualized MIRRs against duration (log quarters).

3. Investment returns, property quality, and deal size before adjusting for risk

This section tests whether properties' investment returns, before adjusted for risk, are related to property quality and deal size. To make it easier to interpret results, I normalize quality and deal size by subtracting their means from each value and then dividing by their respective standard deviations. As a result, 0 means that the value equals the mean, and 1 means the value is one standard deviation above the mean. We normalize deal size with its national mean and national standard deviation because investors are free to invest in different CBSAs so deal size seems comparable across CBSAs. An empirical question is should we normalize year-1 NOI per square foot with local (CBSA) or national means and standard deviations. It turns out that the two normalized variables are highly correlated, as Figure 8 illustrates, and results in this paper are robust regardless how we normalize quality. However, when we run horse races between locally and nationally normalized year-1 NOI per square foot by including both of them in all regressions in this paper, locally normalized values are always statistically significant but the nationally normalized ones become insignificant. Therefore, I use locally normalized year-1 NOI per square foot in all analyses below. For the local normalization to be less affected by small samples in thin markets, I only conduct the normalization for properties located in CBSAs that have at least 10 samples. As a result, the sample size becomes 6,215 in all analyses below.

The first model analyzes whether annualized MIRRs are significantly related to quality, deal size, and duration of their holding periods.

$$MIRR_i = \alpha + \beta_c Q_i + \beta_D D_i + \beta_P U_i + \sum_{k=1}^{K} \lambda_k X_{i,k} + \varepsilon_i \tag{3}$$

In equation (3), for property i, Q_i is quality (year-1 NOI per square foot, 2017:Q2 dollars, log values), D_i is deal size (acquisition price in 2017:Q2 dollars, log values), and U_i is duration of its holding period (log quarters). The model also includes a variety of control variables, $X_{i,k}$, including fixed effects of property types, CBSAs, investors/managers, and acquisition periods.

Table 3 reports regression results for the whole sample, the sold properties, and the properties that had not been sold. For the whole sample, MIRRs are significantly higher for properties with higher quality, smaller deal size, and shorter holding periods. Specifically, when year-1 NOI per square foot increases by 1 standard deviation, the annualized MIRR increases by about 0.009 and this increase is statistically significant at the 1% level. This relationship remains significant for sub-samples of sold and unsold properties. When the deal size increases by 1 standard deviation, the annualized MIRR decreases by 0.003, which is also statistically significant at the 1% level. This relationship is significant for unsold properties but becomes insignificant for sold properties. If the duration increases by 1 (log value), then the annualized MIRR decreases by about 0.0534, which is significant at the 1% level. The relationship remains significant for sold properties but cannot be tested for unsold properties, because all unsold properties have a 5-year artificial holding period, so there is no variation in their holding periods.

I then estimate (3) for each type separately, and report results in Table 4. Table 4 shows that, first, MIRRs are higher for higher quality for all four types. Coefficients of NOI per square foot are 0.0179, 0,0072, 0.0109, and 0.0166 respectively and all significant at the 1% level. Second, MIRRs are lower for larger deals. Coefficients of deal size are -0.0177, -0.0044, -0.0053, and 0.0001. The first three are significant at the 1% level but the last one, which is for retail properties, is insignificant. Third, MIRRs are higher for shorter duration. Coefficients of duration are -0.0549, -0.0528, -0.0574, and -0.0350 respectively and all significant at the 1% level. Overall, it seems reasonable to conclude that the results found for the whole sample – MIRRs are higher for higher quality, smaller deals, and shorter duration – are generally robust across property types.

I further investigate whether the results are robust in sub-samples with different quality level and deal size. This is to check whether outliers are driving the results. I first partition the full sample into high and low quality groups. The high (low) quality group consists of properties with quality being above (below) local means. I then partition the whole sample into large deals and small deals. The group of large (small) deals consists of properties with deal size being above (below) the national mean. I estimate the model in (3) for these four groups respectively and report the results in Panel A of Table 5. The first result is that the MIRR increases with quality in all sub-samples, and this relationship is significant at the 1% level for all sub-samples except the high-quality group, for which the coefficient has the same sign but insignificant. Second, the MIRR decreases with deal size in all four sub-samples, and the relationship is significant at the 1% level for all sub-samples except large deals, for which the coefficient has the same sign but insignificant. Third, the MIRR decreases with duration, which is statistically significant at the 1% level in all four sub-samples.

I then further split the whole sample into four mutually exclusive groups: high quality and large deals, high quality and small deals, low quality and large deals, and low quality and small deals. I estimate the model in (3) for these sub-samples and report results in Panel B of Table 5. The results are very similar: the MIRR increases with quality in all sub-samples except the group of high quality and small deals, decreases with deal size in all sub-samples except the group of high quality and large deals, decreases with duration in all sub-samples. The coefficients are generally statistically significant at the 1% level.

The model in (3) directly relates investment returns to quality, deal size, and duration. However, it only partially controls for time varying performance of the over all real estate market using acquisition period dummies. Now I use the following model to more carefully control for time varying market-wide performance by including dummies for each quarter within each property's holding period.

$$\log(R_i) = \alpha_Q Q_i + \rho_Q Q_i U_i + \alpha_D D_i + \rho_D D_i U_i + \sum_{t=buy_i+1}^{sell_i} M_t + \sum_{k=1}^{K} \lambda_k X_{i,k} + \varepsilon_i \tag{4}$$

In equation (4), for property i, R_i is gross return over its entire holding period; the coefficient of quality Q_i, α_Q, captures the per-investment abnormal return related to quality; the coefficient of the interaction term between quality Q_i and duration U_i, ρ_Q, captures the per-period abnormal return related to quality; the coefficient of deal size, α_D, captures the *per-investment* abnormal return related to deal size; the coefficient of the interaction term between deal size D_i and duration U_i, ρ_D, captures the *per-period* abnormal return related to deal size; M_t are coefficients of dummies for each quarter within the property's holding periods and capture the average real estate market performance; and $X_{i,k}$ are dummy variables for property types, investors/managers, and CBSAs where the property is located. Note that the model in (4) is essentially an extended version of the repeat sales regression (see, e.g. Bailey, Muth and Nourse (1963), Case and Shiller (1989), Goetzmann (1992), Peng (2012)), with the extension being that it includes not only holding period dummies but also other variables.

Table 6 reports results of estimating the model in (4) using the whole sample, sold properties, and unsold properties respectively, which are consistent with those in Table 3. The results suggest that properties with higher quality have significantly higher per-investment returns. When quality increases by one local standard deviation, the holding-period return (log gross return) increases by about 0.13 for the whole sample, 0.14 for sold properties, and about 0.05 for unsold properties. However, this abnormal return seems to diminish over time, as suggested by the coefficient of the interaction term between quality and duration. The per-period decrease of the abnormal return is about 0.003 for the whole sample and about 0.004 for sold properties, which are statistically significant but economically small compared with the per-investment positive abnormal returns. The results in Table 6 also suggest that larger deals have lower per-investment abnormal returns. When the deal size increase by one standard deviation, the holding-period return (log gross return) decreases by about 0.04 for the whole sample and about 0.05 for sold properties. The negative abnormal per-investment return does not seem to change over time: while the coefficient of the interaction term between deal size and duration is positive and significant at the 10% level for the whole sample, it is economically small (0.008). It is worth noting that the adjusted R2s in Table 6, which are 0.54, 0.57, and 0.37, are much higher than those in Table 3, which are 0.43, 0.50 and 0.37. This seems to suggest that including dummies for each quarter in properties' holding periods helps explain properties' holding period returns.

4. Risk-adjusted returns, property quality, and deal size

4.1. Research design

There are two possible reasons why properties with higher quality and smaller deals may have higher returns. First, they may have higher returns solely because they have higher risk. Second, they may have higher returns even after adjusting for risk. These two possibilities have very different implications for investors. I now analyze whether quality and deal size affect returns for reasons that are not related to risk by testing whether they are related to properties' risk-adjusted returns.

It is crucial in my tests to allow properties with different quality and deal size to have different systematic and non-systematic risk. I measure a property's systematic risk with its factor loadings in conventional asset pricing factor models, and measure its non-systematic risk with two variables: an ex post measure that equals the squared "residual" from fitting its holding period return to a factor model, and an ex ante "catch-all" risk measure, which is simply its acquisition cap rate (see, e.g. Peng (2019) for evidence that cap rates help predict future risk).

It is challenging to allow individual properties to have different systematic risk because I do not observe time series of returns for each property; as a result, I am unable to use time series regressions to estimate each property's factor loadings. I overcome this challenge with an approach proposed by Peng (2019) in the framework of holding-period factor models, which allows individual properties' factor loadings as well as their risk-adjusted returns to be functions of their attributes. This approach allows me to formally test whether properties' risk-adjusted returns are related to their quality and deal size, while allowing their systematic risk (i.e. factor loadings) to be also functions of their quality and deal size.

The same approach also makes it possible to allow individual properties to have different exposure to its non-systematic risk, or, in other words, different "loadings" of the two non-systematic risk measures. By allowing both the non-systematic risk measures themselves and their loadings to vary across properties with different quality and deal size, I am able to mitigate possible biases in my tests that are due to the either heterogeneous non-systematic risk or its heterogeneous effects on properties' returns.

Peng (2019)'s approach is built on a holding-period factor model. This model is used by Cochrane (2005) to estimate the beta of venture capital investments. Similar models are used by Korteweg and Sorensen (2010), Driessen, Lin and Phalippou (2012), and Franzoni, Nowak and Phalippou (2012) to estimate factor loadings for private equity, and by Peng (2016) to estimate factor loadings of private commercial real estate.

Consider a property i that was acquired in period buy_i and sold in period $sell_i$, I assume that the single-period return for this property in period t, $R_{i,t}$ (a gross return), is generated from the following log-linear factor model,

$$\log\left(R_{i,t}\right)-\log\left(T_t\right)=\alpha_i+\sum_{k=1}^{K}\beta_i^k F_{k,t}+\upsilon_{i,t} \tag{5}$$

where T_t is the risk-free interest rate (a gross return), $\left\{F_{k,t}\right\}_{k=1}^{K}$ are k factors, α_i and β_i^k are property i's risk adjusted return and the loading on factor $F_{k,t}$, and $\upsilon_{i,t}$ is an error term. Note that the model allows factor loadings and the alpha to vary across properties.

I then aggregate both sides of (5) across periods within the property's holding period, and have the following.

$$\sum_{t=buy_i+1}^{sell_i}\log\left(R_{i,t}\right)-\sum_{t=buy_i+1}^{sell_i}\log\left(T_t\right)$$
$$=\alpha_i\left(sell_i-buy_i\right)+\sum_{k=1}^{K}\left(\beta_i^k\sum_{t=buy_i+1}^{sell_i}F_{k,t}\right)+\sum_{t=buy_i+1}^{sell_i}\upsilon_{i,t} \tag{6}$$

Note that the duration of the holding period, U_i, is essentially

$$U_i=sell_i-buy_i, \tag{7}$$

and

$$\log\left(R_i\right)=\sum_{t=buy_i+1}^{sell_i}\log\left(R_{i,t}\right). \tag{8}$$

I further simplify the notation for the error term as follows.

$$\sum_{t=buy_i+1}^{sell_i}\upsilon_{i,s}=\varepsilon_i \tag{9}$$

The model becomes

$$\log\left(R_i\right)-\sum_{s=buy_i+1}^{sell_i}\log\left(T_t\right)=\alpha_i U_i+\sum_{k=1}^{K}\left(\beta_i^k\sum_{s=buy_i+1}^{sell_i}F_{k,t}\right)+\varepsilon_i. \tag{10}$$

Since real estate returns may have non-temporal components (see, e.g. Goetzmann and Spiegel (1995)), I add a non-temporal term z_i to the model.

$$\log\left(R_i\right)-\sum_{s=buy_i+1}^{sell_i}\log\left(T_t\right)=z_i+\alpha_i U_i+\sum_{k=1}^{K}\left(\beta_i^k\sum_{s=buy_i+1}^{sell_i}F_{k,t}\right)+\varepsilon_i \tag{11}$$

To test whether risk-adjusted returns are related to property quality and deal size, I let z_i to be a function of quality and deal size as follows.

$$z_i=z+\rho_Q Q_i+\rho_D D_i \tag{12}$$

I also let the per-period alpha to be a function of a variety of dummy variables $X_{i,n}$, including dummies for CBSAs, investors/managers, and property types, and quality and deal size as follows.

$$\alpha_i=\sum_{n=1}^{N}\eta_n X_{i,n}+\gamma_Q Q_i+\gamma_D D_i \tag{13}$$

I include dummies to allow properties located in different CBSAs, invested or managed by different NCREIF members, and belonging to different types to have their own base levels of alphas.

I also allow factor loadings to be functions of property quality and deal size.

$$\beta_i^k = \beta^k + \lambda_Q^k Q_i + \lambda_D^k D_i \tag{14}$$

The model becomes

$$\begin{aligned}
&\log\left(R_i\right) - \sum\nolimits_{s=buy_i+1}^{sell_i} \log\left(T_t\right) = z + \rho_Q Q_i + \rho_D D_i \\
&+\alpha \sum_{n=1}^{N} \eta_n X_{i,n} + \gamma_Q Q_i U_i + \gamma_D D_i U_i + \sum\nolimits_{k=1}^{K}\left(\beta^k \sum\nolimits_{s=buy_i+1}^{sell_i} F_{k,t}\right) \\
&+\sum\nolimits_{k=1}^{K}\left(\lambda_Q^k Q_i \sum\nolimits_{s=buy_i+1}^{sell_i} F_{k,t}\right) + \sum\nolimits_{k=1}^{K}\left(\lambda_D^k D_i \sum\nolimits_{s=buy_i+1}^{sell_i} F_{k,t}\right) + \varepsilon_i.
\end{aligned} \tag{15}$$

The null hypotheses I test are whether ρ_Q, ρ_D, γ_Q, and γ_D are zero. If ρ_Q and ρ_D are not zero, I conclude that risk-adjusted returns have non-temporal (per-investment) components that are related to quality and deal size. If γ_Q and γ_D are not zero, I conclude that per-period risk-adjusted returns are related to quality and deal size. Note that the model allows factor loadings to be functions of quality and deal size, which is reflected in the interaction terms between factors and quality and deal size in (15). As a result, the test results are not biased by heterogeneity in properties factor loadings. For example, the results are not biased by the possibility that larger deals may have larger loadings on the liquidity risk, which is already captured by the interaction term between the liquidity risk factor and deal size.

4.2. Variables

When estimating (15), I include stock market factors, bond market factors, and a "real estate factor" that I construct, which captures the common component of properties' risk premium that is not explained by stock and bond factors. The stock market factors consist of the union of the six factors in Fama and French (2018) and the five factors in Hou, Mo, Xue and Zhang (2018). The bond market factors include the term spread (the difference between the 10-year treasury annual yield and 1-year treasury annual yield) and the credit spread (the difference between the BAA corporate bond annual yield and AAA corporate bond annual yield) and their first order quarterly differences, which are shown by Peng (2016) to help explain real estate returns. Since (15) is a log-return model, all the stock and bond market factors are in log gross returns.

I create and include the “real estate” factor to mitigate the problem of missing factors. No matter how many factors I include in the model, it is always possible that some unknown factors are missing. The real estate factor is a “catch-all” variable that captures the average effect of all the missing factors on properties’ risk premium. In other words, it is an index for real estate specific risk premium that is orthogonal to the factors included.

I use a larger sample from the NCREIF database, 10,898 properties to be specific, to construct the real estate factor. I have a larger sample because the construction of the real estate factor requires fewer variables than what my main analyses in this paper do, and thus it is reasonable to impose fewer cleaning filters; as a result, more properties remain in the sample. For a property to be in this sample, it needs to satisfy the following three conditions. First, the acquisition price (2017:Q2 dollars) is higher than $2 million. Second, it belongs to the four main types (apartment, industrial, office, and retail). Third, the holding period total return of the property must be reasonable as I described earlier in this paper.

The first step in constructing the real estate factor is to estimate the model in (11), which is a simplified version of (15) that lets all properties have identical alphas and identical loadings for each factor, using the 10,898 properties.[6] I include in the model all the stock and bond market factors I use in this paper. Note that the Case and Shiller (1989) three-stage approach can be used if the variance of ε_i increases with the duration of each

[6] The results are robust when I allow alphas and loadings to vary across properties as functions of quality and deal size.

property's holding period. However, I find no evidence for such a relationship;[7] therefore, I estimate the model with OLS and obtain the residual $\hat{\varepsilon}_i$ for each property, which measures the component of its risk premium what is not explained by the stock and bond market factors.

The second step works on the residuals $\hat{\varepsilon}_i$ from the first step. According to equation (9), each residual from the first step can be considered as the sum of the latent per-period residual $\upsilon_{i,t}$ across the holding period. I further assume that each single-period residual $\upsilon_{i,t}$ contains a common component I_t, which is the index for the real estate market-wide risk premium, and an error $e_{i,t}$.

$$\upsilon_{i,t} = I_t + e_{i,t} \tag{16}$$

Note that the real estate index I_t essentially captures the common risk-premium component of all properties that is orthogonal to the included stock and bond factors.

Combining (9) and (16) leads to the following model, which is essentially the repeat sales regression.

$$\hat{\varepsilon}_i = \sum\nolimits_{t=buy_i+1}^{sell_i} I_t + \sum\nolimits_{t=buy_i+1}^{sell_i} e_{i,t} \tag{17}$$

Note that the above model essentially regresses each property's holding-period residual against dummies for each quarter of the property's holding period, and the real estate index I_t is the coefficients of the quarter dummies. Since I find no evidence that the variance of the error term in (17) increases with the holding period duration, I estimate (17) with OLS. Figure 9 plots the time series of the real estate factor $\hat{I}_t$.

I then validate the real estate factor by showing that it has out-of-sample explanatory power for properties' risk premium. To do so, I randomly split the whole sample of

[7] Following Case and Shiller (1989), I first obtain residuals from OLS regression of (9), and then regress squared residuals against the duration of the holding period. The result shows that squared residuals, which are proxies for variance, are not increasing with duration.

properties into two groups with equal number of properties, say groups A and B. I then use residuals of properties in group A, which are obtained in the first step of constructing the real estate index, to estimate the index, which is denoted by I_t^A, and then test whether I_t^A explains return residuals of properties in group B, ε_i^B, using the following regression.

$$\varepsilon_i^B = \lambda \sum_{t=buy_i+1}^{sell_i} I_t^A + e_i^B \tag{18}$$

A significant and positive λ would indicate that the real estate index has out-of-sample explanatory power for property risk premium.

I conduct 1,000 rounds of the out-of-sample test, randomly splitting the sample each time, and plot the histogram of λ from the 1,000 rounds in Figure 10. It is apparent that λ is positive and significantly different from 0, which is also confirmed by a formal t-test. This is strong evidence that the real estate index helps capture the common components of properties risk premium that are orthogonal to the stock and bond market factors.

Another variable I construct pertains to non-systematic risk of individual properties, which I call "peculiar risk". I construct it using the following two-step approach. First, I estimate the model in (15) that includes all stock and bond factors and the real estate factor I just constructed and obtain residuals for each property. Second, I calculate the squared values of the residuals and call them "peculiar risk", denoted by K_i. The peculiar risk measures the deviation of individual properties' risk premium during its holding period from what can be explained by all the factors. Figure 11 plots the histogram of this risk measure, which seems to be consistent with a Chi-squared distribution.

Figure 12 plots the histogram of individual properties' acquisition cap rates, which I use as a proxy for investors' ex ante risk measure to help capture non-systematic risk of each property. The cap rates used in regressions discussed in next section are calculated using appraised values at the end of the first year after acquisition, which are very similar to cap rates plot in the figure.

4.3. Empirical results

Table 7 report results of estimating (15) by including stock market factors only (specification I), bond market factors only (specification II), and both stock and bond factors (specification III). The first result is that properties with higher quality have positive *per-investment* abnormal returns. The coefficient of quality is 0.0453 when only stock market factors are included, 0.0613 when only bond market factors are included, and 0.0439 when both types of factors are included. All three coefficients are statistically significant at the 1% level. The second result is that properties with higher quality appear to have negative *per-period* abnormal returns. The coefficient of the interaction term between quality and duration is insignificant when including stock factors only but significant when including bond factors or both stock and bond factors. Putting together the above results, it seems that while properties with higher quality might have higher risk-adjusted returns, the returns diminish over time.

The third result in Table 7 is the larger deals have negative *per-investment* abnormal returns. The coefficient of deal size is -0.0398 when including stock factors only, -0.0386 when including bond factors only, and -0.0406 when including both types of factors. All three coefficients are statistically significant at the 1% level. The fourth result is that larger deals appear to have positive *per-period* abnormal returns. The coefficient of the interaction term between deal size and duration is 0.041 when including stock factors only, 0.0046 when including bond factors only, and 0.0026 when including both types of factors. The first two coefficients are statistically significant at the 1% level but the last one is insignificant. Overall, when both stock and bond factors are included, larger deals seem to have lower risk-adjusted returns, which do not appear to change over time.

It is also worth noting that Table 7 suggests that real estate has significant loadings on most stock and bond factors. For example, the loading on the market risk premium is 0.1156 and statistically significant in specification III, which is consistent with Peng (2016). While this paper does not focus on interpreting these loadings or discussing their implications for investments, the results expand the existing literature on property-level risk and return characteristics of commercial real estate.

I then further add to (15) the real estate factor, each property's peculiar risk, and each property's acquisition cap rate that is calculated using appraised value one year after the acquisition to mitigate a mechanical relationship between cap rates and investment returns due to the fact that acquisition prices enter both calculations. Table 8 reports the results of three specifications. The first specification adds the real estate factor as well as its interaction terms with quality and deal size. A few things are worth noting. First, higher quality properties still have positive and significant *per-investment* abnormal returns and negative and weakly significant (10%) *per-period* abnormal returns. Second, larger deals still have lower and significant *per-investment* abnormal returns and insignificant *per-period* abnormal returns. Third, the real estate factor has a positive and significant coefficient: 0.5408, which corroborates earlier finding that it helps provide additional explanatory power for individual properties' risk premium. Fourth, larger deals seem to have smaller loadings on the real estate factor, as the coefficient of the interaction term between deal size and the real estate factor is -0.0089 and statistically significant. The implication is that larger deals are less sensitive to the real estate market average performance. Finally, property quality does not appear to matter for the loading on real estate factors.

The second specification further adds peculiar risk as well as its interaction terms with quality and deal size. Results regarding abnormal returns related to quality and deal size remain robust: properties with higher quality have positive per-investment abnormal returns and negative per-period abnormal returns, and larger deals have negative per-investment abnormal returns and insignificant per-period abnormal returns. The new result is that peculiar risk has a positive loading, which is consistent with the notion that part of the risk premium is compensating for this measure of non-systematic risk. Furthermore, the interaction term between quality and peculiar risk has a significantly negative coefficient, which seems to show that properties with higher quality are less exposed to peculiar risk. The interaction term between deal size and peculiar risk has an insignificant coefficient, which appears to indicate that deal size does not affect properties' exposure to this risk.

The third specification uses each property's acquisition cap rate as a proxy for ex ante risk, which helps capture non-systematic risk of individual properties. First, results regarding abnormal returns related to quality and deal size are still robust: properties with higher quality have positive per-investment abnormal returns and negative per-period abnormal returns, and larger deals have negative per-investment abnormal returns and insignificant per-period abnormal returns. Second, the cap rate has a positive loading, which seems to indicate that it contains information regarding risk beyond what the factors capture and such risk is being compensated. Third, the interaction term between quality and cap rate has a positive and statistically significant coefficient: 0.0035. This implies that properties with higher quality have more exposure to the risk captured by cap rates. Third, the interaction term between deal size and cap rate has a negative and statistically significant coefficient: -0.0152. This appears to suggest that larger deals have less exposure to risk captured by cap rates.

Overall, Tables 7 and 8 seem to provide very robust evidence that properties with higher quality have higher risk-adjusted returns, which, however, seem to diminish over time and might become negative for investments with long holding periods. Furthermore, larger deals have lower risk-adjusted returns, which do not appear to change with the holding period duration. These two tables also indicate that real estate has a unique risk premium component that is common for all properties, which is substantiated by the significant coefficient of the real estate factor. Another finding is that non-systematic risk appears to be also priced, which is not surprising as individual properties can be significant components of investors' portfolios so non-systematic risk matters.

5. Conclusions

Many important assets in the economy, such as housing, commercial real estate, private equity, and venture capital, are very different from the most-studied assets such as stocks and bonds in the sense that they have high values, are indivisible, opaque, costly to acquire, and cannot be shorted. The asset pricing literature, which virtually treats risk as the sole determinant of returns, seems inadequate for these assets, because individual assets' attributes may affect their returns for reasons that have nothing to do with risk.

This paper uses a high-quality proprietary dataset to analyze whether individual properties' quality and deal size are related to their returns, both before and after adjusting for risk. The dataset contains very rich property-level information, which allows me to control for many features that might be related to properties' risk and returns, including their types, investors/managers, CBSAs, and holding periods. Results indicate that properties' investment returns, before adjusting for risk, are positively related to their quality and negatively related to their deal size. More specifically, properties with higher quality tend to have higher returns, which, however, appear to diminish over time, and larger deals tend to have lower returns, which do not seem to vary across time. These results remain after adjusting for both systematic and non-systematic risk.

The above results are original and constitute novel contributions to the literature, as they highlight the importance of non-risk attributes in determining investment returns of many assets in the economy. Deal-level attributes that could have been treated as "frictions" in the asset pricing literature, such as due diligence costs for acquisition, may turn out to be return determinants for these assets. Given the significant roles many of these assets play in the economy, the literature seems to be able to benefit from more future research that further explores what other attributes may affect investment returns and why.

References

Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse, 1963, A regression method for real estate price index construction, *Journal of the American Statistical Association* 58, 933-942.

Case, Karl E., and Robert Shiller, 1989, The efficiency of the market for single family homes, *American Economic Review* 79, 125-137.

Chetty, R., and A. Szeidl, 2007, Consumption commitments and risk preferences, *Quarterly Journal of Economics* 122, 831-877.

Chetty, Raj, Laszlo Sandor, and Adam Szeidl, 2017, The effect of housing on portfolio choice, *Journal of Finance* 72, 1171-1212.

Cochrane, John, 2005, The risk and return of venture capital, *Journal of Financial Economics* 75, 3-52.

Driessen, Joost, Tse-Chun Lin, and Ludovic Phalippou, 2012, A new method to estimate risk and return of non-traded assets from cash fows: The case of private equity funds, *Journal of Financial and Quantitative Analysis* 47, 511-535.

Fama, Eugene, and Kenneth R. French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234–252.

Fisher, Jeffrey, Dean Gartzlaff, David Geltner, and Donald Haurin, 2003, Controlling for the impact of variable liquidity in commercial real estate price indices, *Real Estate Economics* 31, 269-303.

Franzoni, Francesco, Eric Nowak, and Ludovic Phalippou, 2012, Private equity performance and liquidity risk, *Journal of Finance* 67, 2341–2373.

Gang, Jianhua, Liang Peng, and Thomas G. Thibodeau, 2017, Risk and returns of income producing properties: Core versus non-core, *Real Estate Economics* DOI: 10.1111/1540-6229.12208.

Gatzlaff, Dean H, and Donald R. Haurin, 1997, Sample selection bias and repeat-sales index estimates, *Journal of Real Estate Finance and Economics* 14, 33-50.

Gatzlaff, Dean H, and Donald R. Haurin, 1998, Sample selection and biases in local house value indices, *Journal of Urban Economics* 43, 199-222.

Goetzmann, William N., 1992, The accuracy of real estate indices: Repeat sale estimators, *Journal of Real Estate Finance and Economics* 5, 5-53.

Goetzmann, William N., and Liang Peng, 2006, Estimating house price indices in the presence of seller reservation prices, *Review of Economics and Statistics* 88, 100-112.

Goetzmann, William N., and Matthew Spiegel, 1995, Non-temporal components of real estate returns, *Review of Economics and Statistics* 77, 199-206.

Hartman-Glaser, Barney, and William Mann, 2016, Collateral constraints, wealth effects, and volatility: Evidence from real estate markets, *UCLA working paper*.

Hochberg, Yael, and Tobias Mühlhofer, 2015, Capital-market competitiveness and managerial investment decisions: Evidence from commercial real estate, *Working paper*.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2018, Q5, *Working paper*.

Kaplan, Steven, and Antoinette Schoar, 2005, Private equity performance: Returns, persistence, and capital flows, *Jurnal of Finance* 60, 1791–1823.

Korteweg, Arthur, and Morten Soren, 2017, Skill and luck in private equity performance, *Journal of Financial Economics* 123, 535-562.

Korteweg, Arthur, and Morten Sorensen, 2010, Risk and return characteristics of venture capital-backed entrepreneurial companies, *Review of Financial Studies* 23, 3738–3772.

Peng, Liang, 2012, Repeat sales regression on heterogeneous properties, *Journal of Real Estate Finance and Economics* 45, 804-827.

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

Peng, Liang, 2019, Is risk of real estate predictable?, *Penn State University Working Paper*.

Peng, Liang, and Thomas Thibodeau, 2013, Risk segmentation of american homes: Evidence from denver, *Real Estate Economics* 41, 569-599

Peng, Liang, and Thomas G. Thibodeau, 2017, Idiosyncratic risk of house prices: Evidence from 26 million home sales, *Real Estate Economics* 45, 340-375.

Peng, Liang, and Lei Zhang, 2019, House prices and systematic risk: Evidence from micro data, *Real Estate Economics* Forthcoming.

Peng, Liang, and Lei Zhang, 2019, Is housing investment a "rich man's game"?, *Penn State University Working Paper*.

Piketty, Thomas, and Gabriel Zucman, 2014, Capital is back: Wealth-to-income ratios in rich countries, 1700–2010, *Quarterly Journal of Economics* 129, 1255-1310.

Pivo, Gary, and Jeffrey D. Fisher, 2011, The walkability premium in commercial real estate investments, *Real Estate Economics* 39, 185-219.

Plazzi, Alberto, Walter Torous, and Rossen Valkanov, 2010, Expected returns and the expected growth in rents of commercial real estate, *Review of Financial Studies* 23, 3469-3519.

Plazzi, Alberto, Walter Torous, and Rossen Valkanov, 2012, Exploiting property characteristics in commercial real estate portfolio allocation, *Journal of Portfolio Management* 35, 39-50.

Sagi, Jacob S., 2017, Asset-level risk and return in real estate investments, *Manuscript, University of North Carolina-Chapel Hill*.

Table 1. Data summary

This table reports the number of properties (all, sold, and unsold), the number of metro areas where the properties are located, and summary statistics of annualized total return modified IRR (1 means 100%, actual for sold properties and estimated using appraised values for unsold properties), NOI in the first year after acquisition per square foot (2017:Q2 dollars), and the acquisition price (2017:Q2 million dollars), duration of holding period (quarters) for all properties and each of the four main property type.

	All	Apartment	Industrial	Office	Retail
All properties	6,540	1,465	2,382	1,601	1,092
Sold properties	4,414	1,084	1,442	1,167	721
Not-sold properties	2,126	381	940	434	371
Metro areas	227	118	116	103	173
Annualized Modified IRR (1=100%)					
Minimum	-0.094	-0.075	-0.082	-0.094	-0.073
25%	0.025	0.013	0.029	0.018	0.038
Median	0.078	0.076	0.081	0.066	0.095
75%	0.126	0.121	0.123	0.117	0.147
Maximum	0.398	0.393	0.395	0.398	0.396
Mean	0.084	0.080	0.085	0.076	0.099
Standard Deviation	0.085	0.089	0.082	0.086	0.084
First year NOI per square foot ($)					
Minimum	0.27	1.40	0.27	1.40	0.61
25%	9.55	9.41	6.71	24.99	15.76
Median	16.48	13.01	10.18	38.98	22.12
75%	31.02	17.53	17.84	62.91	31.52
Maximum	686.77	208.99	686.77	654.67	291.42
Mean	26.79	15.05	15.71	54.59	25.98
Standard Deviation	33.63	11.33	20.99	50.17	19.09
Acquisition price ($ million)					
Minimum	2.00	2.70	2.00	2.04	2.22
25%	9.87	23.35	5.73	12.01	12.50
Median	21.40	35.28	10.74	26.83	24.50
75%	41.84	54.64	20.94	56.73	45.67
Maximum	1,430.54	295.99	687.39	1,430.54	880.99
Mean	36.72	44.83	17.75	52.23	44.47
Standard Deviation	58.07	34.99	27.36	85.11	70.87
Duration for sold properties (quarters)					
Minimum	4	4	4	4	4
25%	17	18	17	17	16
Median	28	26	29	28	27
75%	38	34	40	38	39
Maximum	124	92	124	104	112
Mean	28.79	27.24	29.69	29.11	28.83
Standard Deviation	15.16	13.03	16.07	15.52	15.53

Table 2. Correlation between variables

This table reports correlation between pairs of the following variables: annualized modified IRR (log gross returns), the first year NOI per square foot (2017 Q2 dollars, log values), purchase price (2017:Q2 dollars, log values), and duration (log quarters, for sold properties only) for the whole sample and the four property types.

	NOI per SF	Acquisition price	Duration
		Full sample	
Annualized MIRR	0.15	-0.14	-0.32
NOI per SF		0.10	0.21
Acquisition price			-0.09
		Apartment	
Annualized MIRR	0.22	-0.16	-0.31
NOI per SF		0.28	0.10
Acquisition price			-0.01
		Industrial	
Annualized MIRR	0.22	-0.13	-0.31
NOI per SF		-0.16	0.32
Acquisition price			-0.13
		Office	
Annualized MIRR	0.21	-0.16	-0.35
NOI per SF		0.09	0.30
Acquisition price			-0.14
		Retail	
Annualized MIRR	0.19	-0.18	-0.35
NOI per SF		0.13	0.16
Acquisition price			0.04

Table 3. MIRRs, quality, deal size, and duration

This table reports results of regressions of properties' holding period annualized MIRRs (log gross returns) against their quality, which is measured with year-1 NOI per square foot (2017:Q2 dollars, log values) normalized by local (CBSA) means and standard deviations, deal size, which is acquisition price (2017:Q2 dollars, log values) normalized by whole-sample mean and standard deviation), and duration (log quarters), dummies of property types, CBSAs where properties are located, investors/managers, and acquisition periods. White's heteroscedasticity-consistent standard deviations are reported in parentheses. ***, **, and * indicate significant levels of 1%, 5%, and 10% respectively.

	Whole sample	Sold	Not-sold
Quality	0.0086***	0.0092***	0.0096***
	(0.0013)	(0.0017)	(0.0019)
Deal size	-0.0031***	-0.0005	-0.0033*
	(0.0011)	(0.0014)	(0.0018)
Duration	-0.0534***	-0.0618***	NA
	(0.0021)	(0.0022)	
Property type dummy	Yes	Yes	Yes
CBSA dummy	Yes	Yes	Yes
Investor dummy	Yes	Yes	Yes
Acquisition period dummy	Yes	Yes	Yes
Sample size	6,215	4,192	2,126
Adjusted R2	0.43	0.50	0.37

Table 4. MIRRs across property types

This table reports results of regressions of properties' holding period annualized MIRRs (log gross returns) against their quality, which is measured with year-1 NOI per square foot (2017:Q2 dollars, log values) normalized by local (CBSA) means and standard deviations, deal size, which is acquisition price (2017:Q2 dollars, log values) normalized by whole-sample mean and standard deviation), and duration (log quarters), dummies of property types, CBSAs where properties are located, investors/managers, and acquisition periods for each of the four property types. White's heteroscedasticity-consistent standard deviations are reported in parentheses. ***, **, and * indicate significant levels of 1%, 5%, and 10% respectively.

	Apartment	Industrial	Office	Retail
Quality	0.0179***	0.0072***	0.0109***	0.0116***
	(0.0034)	(0.0021)	(0.0028)	(0.0033)
Deal size	-0.0117***	-0.0044**	-0.0053**	0.0001
	(0.0036)	(0.0017)	(0.0023)	(0.0029)
Duration	-0.0549***	-0.0528***	-0.0574***	-0.0350***
	(0.0050)	(0.0034)	(0.0039)	(0.0058)
CBSA dummy	Yes	Yes	Yes	Yes
Investor dummy	Yes	Yes	Yes	Yes
Acquisition period dummy	Yes	Yes	Yes	Yes
Sample size	1,394	2,315	1,553	953
Adjusted R2	0.51	0.47	0.43	0.52

Table 5. MIRRs across subsamples

This table reports results of regressions of properties' holding period annualized MIRRs (log gross returns) against their quality, which is measured with year-1 NOI per square foot (2017:Q2 dollars, log values) normalized by local (CBSA) means and standard deviations, deal size, which is acquisition price (2017:Q2 dollars, log values) normalized by whole-sample mean and standard deviation), and duration (log quarters), dummies of property types, CBSAs where properties are located, investors/managers, and acquisition periods for sub-samples. The four sub-samples in Panel A are high quality (year-1 NOI per square foot is above local average), low quality (year-1 NOI per square foot is below local average), large deals (deal size is above its national average), and small deals (deal size is below its national average). The four mutually exclusive sub-samples in Panel B are high quality and large deals, high quality and small deals, low quality and large deals, and low quality and small deals. White's heteroscedasticity-consistent standard deviations are reported in parentheses. ***, **, and * indicate significant levels of 1%, 5%, and 10% respectively.

Panel A

	High quality	Low quality	Large deals	Small deals
Quality	0.0012	0.0145***	0.0093***	0.0084***
	(0.0022)	(0.0021)	(0.0018)	(0.0019)
Deal size	-0.0054**	-0.0034**	-0.0021	-0.0062**
	(0.0016)	(0.0016)	(0.0022)	(0.0024)
Duration	-0.0566***	-0.0517***	-0.0494***	-0.0561***
	(0.0028)	(0.0030)	(0.0030)	(0.0030)
Property type dummy	Yes	Yes	Yes	Yes
CBSA dummy	Yes	Yes	Yes	Yes
Investor dummy	Yes	Yes	Yes	Yes
Acquisition period dummy	Yes	Yes	Yes	Yes
Sample size	2,968	3,247	3,270	2,945
Adjusted R2	0.42	0.46	0.42	0.46

Panel B

	High quality & Large deals	High quality & Small deals	Low quality & Large deals	Low quality & Small deals
Quality	0.0078*	-0.0048	0.0164***	0.0136***
	(0.0034)	(0.0033)	(0.0032)	(0.0030)
Deal size	-0.0025	-0.0067*	-0.0108***	-0.0072**
	(0.0030)	(0.0035)	(0.0034)	(0.0034)
Duration	-0.0487***	-0.0629***	-0.0453***	-0.0506***
	(0.0040)	(0.0042)	(0.0045)	(0.0043)
Property type dummy	Yes	Yes	Yes	Yes
CBSA dummy	Yes	Yes	Yes	Yes
Investor dummy	Yes	Yes	Yes	Yes
Acquisition period dummy	Yes	Yes	Yes	Yes
Sample size	1,626	1,342	1,644	1,603
Adjusted R2	0.39	0.45	0.46	0.50

Table 6. Holding-period returns

This table reports results of regressions of properties' holding period total returns (log gross returns) against their quality, which is measured with year-1 NOI per square foot (2017:Q2 dollars, log values) normalized by local (CBSA) means and standard deviations, deal size, which is acquisition price (2017:Q2 dollars, log values) normalized by whole-sample mean and standard deviation), and dummies of property types, CBSA where properties are located, investors/managers, and quarters within holding periods. White's heteroscedasticity-consistent standard deviations are reported in parentheses. ***, **, and * indicate significant levels of 1%, 5%, and 10% respectively.

	Whole sample	Sold	Not-sold
Quality	0.1332***	0.1422***	0.0481***
	(0.0112)	(0.0128)	(0.0094)
Quality * duration	-0.0032***	-0.0036***	NA
	(0.0004)	(0.0004)	
Deal size	-0.0438***	-0.0489***	-0.0150*
	(0.0114)	(0.0133)	(0.0088)
Deal size * duration	0.0008*	0.0007	NA
	(0.0004)	(0.0005)	
Property type dummy	Yes	Yes	Yes
CBSA dummy	Yes	Yes	Yes
Investor dummy	Yes	Yes	Yes
Investment period dummies	Yes	Yes	Yes
Sample size	6,215	4,192	2,126
Adjusted R2	0.54	0.57	0.37

Table 7. Risk-adjusting returns and stock and bond factors

This table reports results of estimating holding-period factor models in (15) that (1) allow properties' risk-adjusted returns to vary across CBSAs where they are located and their property types and investors/managers, and (2) allow properties' factor loadings to be functions of quality and deal size. White's heteroscedasticity-consistent standard deviations are reported in parentheses. ***, **, and * indicate significant levels of 1%, 5%, and 10% respectively.

	I	II	III
Dummies	Yes	Yes	Yes
Dummies * duration	Yes	Yes	Yes
Factors * quality	Yes	Yes	Yes
Factors * deal size	Yes	Yes	Yes
Quality	0.0453***	0.0613***	0.0439***
	(0.0115)	(0.0113)	(0.0116)
Quality * duration	0.0009	-0.0104***	-0.0083**
	(0.0013)	(0.0012)	(0.0033)
Deal size	-0.0398***	-0.0386***	-0.0406***
	(0.0117)	(0.0110)	(0.0116)
Deal size * duration	0.0041***	0.0046***	0.0026
	(0.0013)	(0.0013)	(0.0036)
Mkt.Rf	0.1606***		0.1156**
	(0.0359)		(0.0476)
SMB	-1.0768***		-0.9258***
	(0.1537)		(0.2080)
HML	0.2385***		0.0823
	(0.0399)		(0.0517)
RMW	0.1067		0.0775
	(0.0744)		(0.0808)
CMA	-0.5653***		-0.4484***
	(0.1477)		(0.1537)
LIQ	0.2278***		0.2255***
	(0.0242)		(0.0346)
MOM	0.2021***		0.0331
	(0.0433)		(0.0513)
Q.ME	0.9286***		0.7722***
	(0.1563)		(0.1917)
Q.IA	0.5830***		0.6782***
	(0.1533)		(0.1599)
Q.ROE	-0.3649***		-0.2353**
	(0.1030)		(0.1095)
Credit spread		-10.0080***	-5.7362***
		(0.5469)	(1.2235)
Term spread		-1.0553***	0.2760
		(0.2252)	(0.4689)
Change in credit spread		0.2184	13.1050**
		(4.1538)	(6.3187)
Change in term spread		-2.7016**	-5.4406***
		(1.2964)	(1.8567)
Sample size	6,215	6,215	6,215
Adjusted R2	0.32	0.29	0.33

Table 8. Risk-adjusting returns, stock, bond, and real estate factors
This table reports results of estimating holding-period factor models in (15) that (1) allow properties' risk-adjusted returns to vary across CBSAs where they are located and their property types and investors/managers, and (2) allow properties' factor loadings to be functions of quality and deal size. White's heteroscedasticity-consistent standard deviations are reported in parentheses. ***, **, and * indicate significant levels of 1%, 5%, and 10% respectively.

	I	II	III
Dummies	Yes	Yes	Yes
Dummies * duration	Yes	Yes	Yes
Factors * quality	Yes	Yes	Yes
Factors * deal size	Yes	Yes	Yes
Quality	0.0459***	0.0498***	0.0517***
	(0.0105)	(0.0115)	(0.0125)
Quality * duration	-0.0062*	-0.0067**	-0.0084**
	(0.0032)	(0.0032)	(0.0036)
Deal size	-0.0419***	-0.0435***	-0.0328***
	(0.0114)	(0.0113)	(0.0126)
Deal size * duration	0.0018	0.0022	-0.0047
	(0.0036)	(0.0035)	(0.0041)
Real estate factor	0.5408***	0.5420***	0.5963***
	(0.0365)	(0.0362)	(0.0421)
Quality * real estate factor	-0.0560	-0.0570	-0.0767*
	(0.0382)	(0.0376)	(0.0426)
Deal size * real estate factor	-0.0889**	-0.0867**	-0.0916**
	(0.0369)	(0.0369)	(0.0437)
Peculiar risk		0.2345***	
		(0.0630)	
Quality * Peculiar risk		-0.1593***	
		(0.0579)	
Deal size * Peculiar risk		0.0934	
		(0.0628)	
Cap rate			0.0099***
			(0.0028)
Quality * Cap rate			0.0035***
			(0.0008)
Deal size * Cap rate			-0.0152***
			(0.0048)
Sample size	6,215	6,215	6,215
Adjusted R2	0.35	0.36	0.38

Figure 1. Histogram of log annualized MIRRs

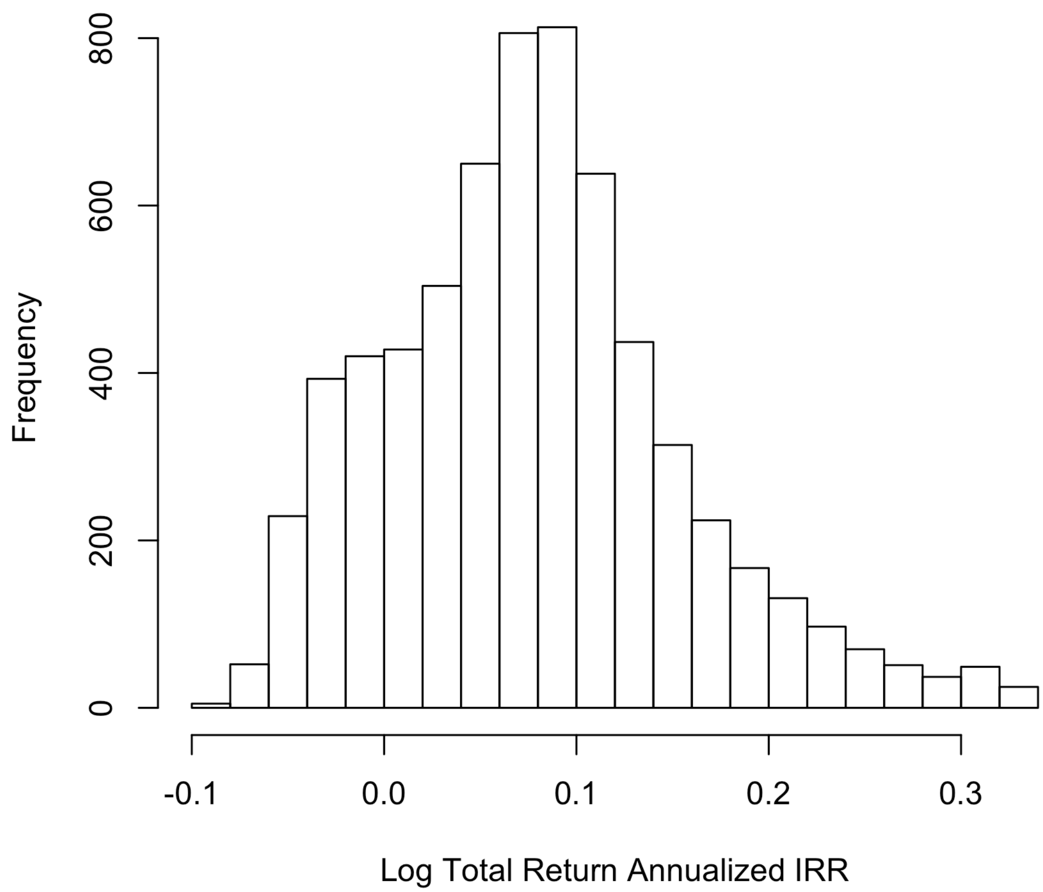

Figure 2. Histogram of first year NOI per square foot (2017:Q2 dollars, log values)

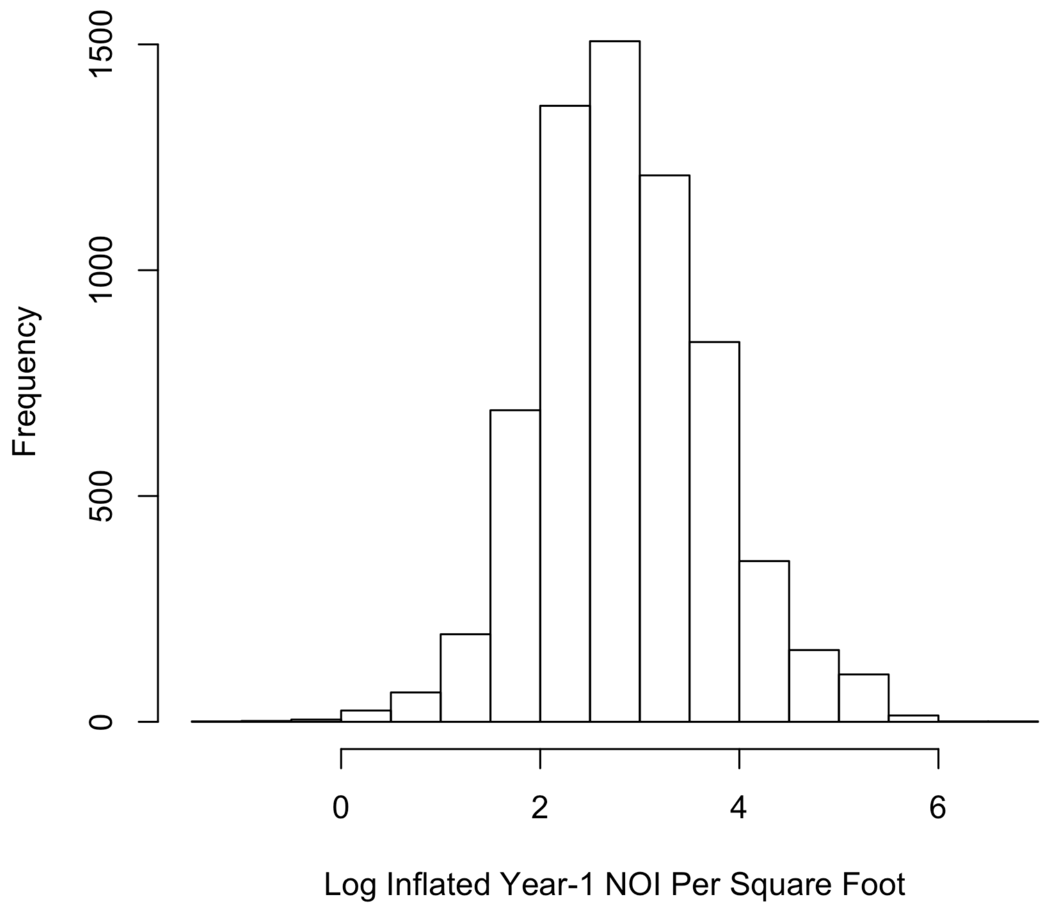

Figure 3. Histogram of purchase prices (2017:Q2 dollars, log values)

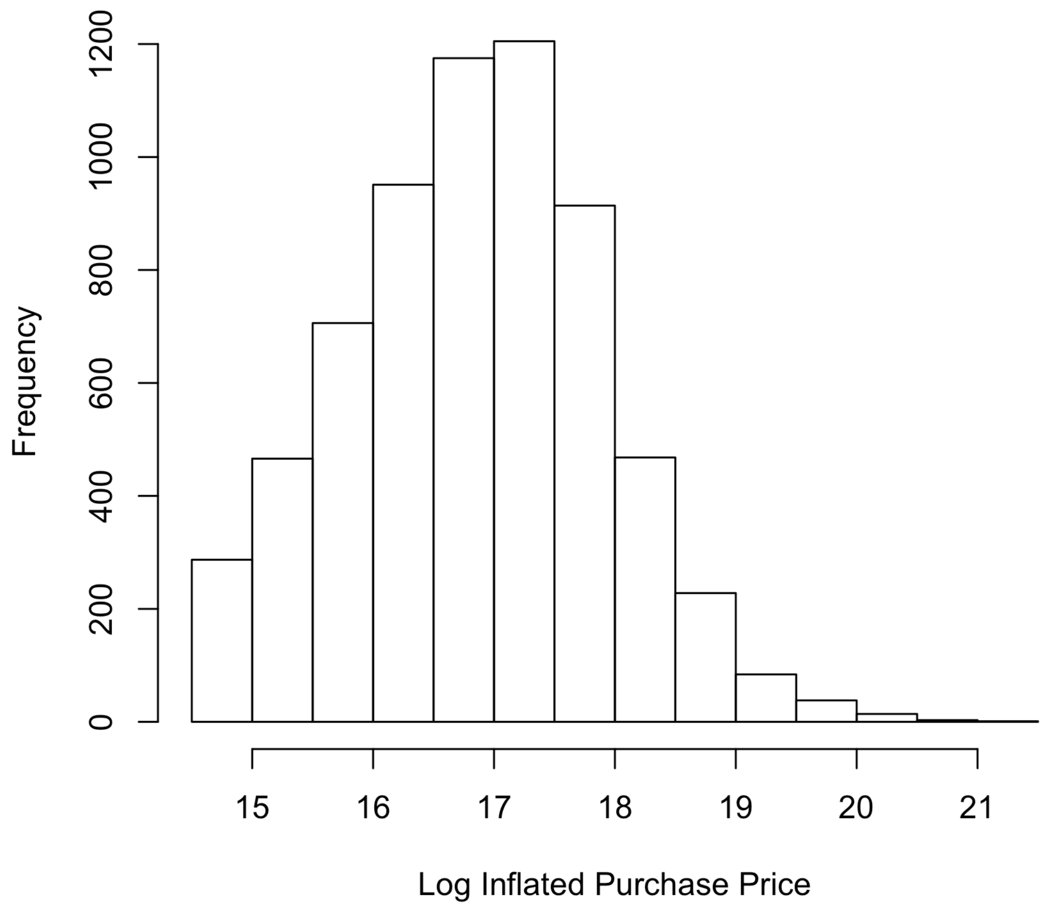

Figure 4. Histogram of duration (log quarters) of sold properties

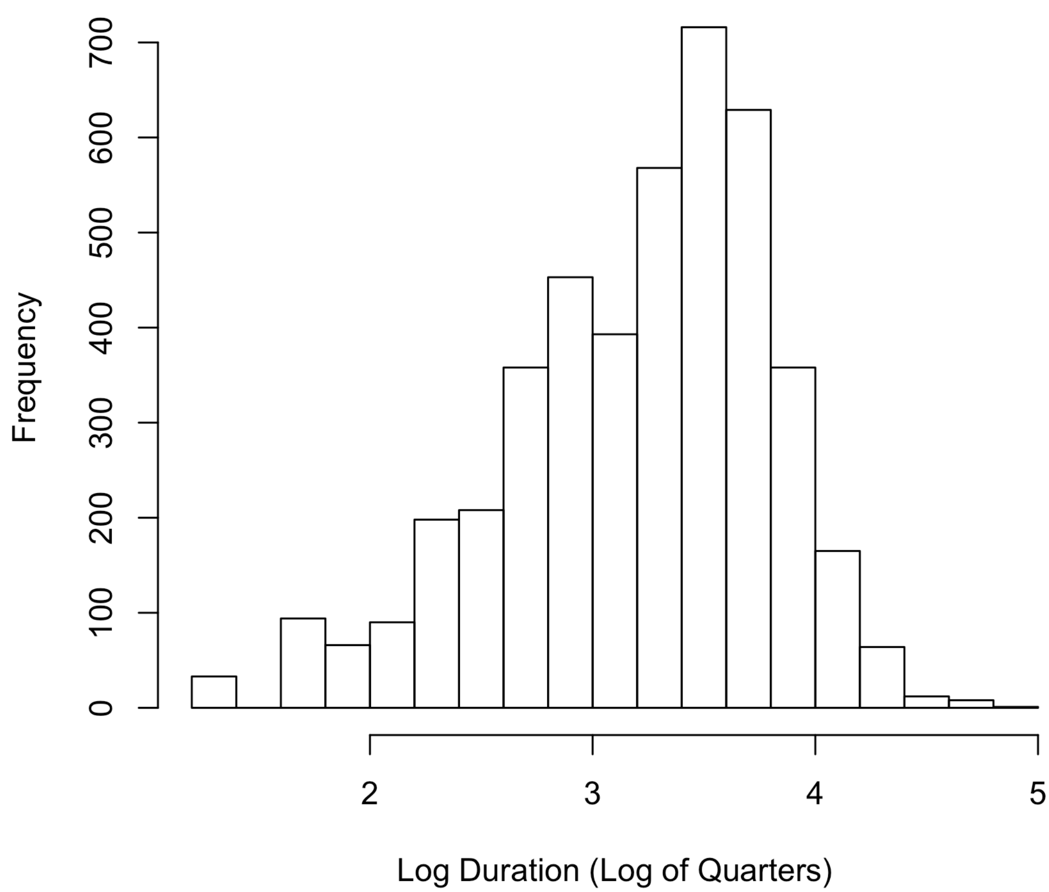

Figure 5. Year-1 NOI per square foot and and annualized MIRRs

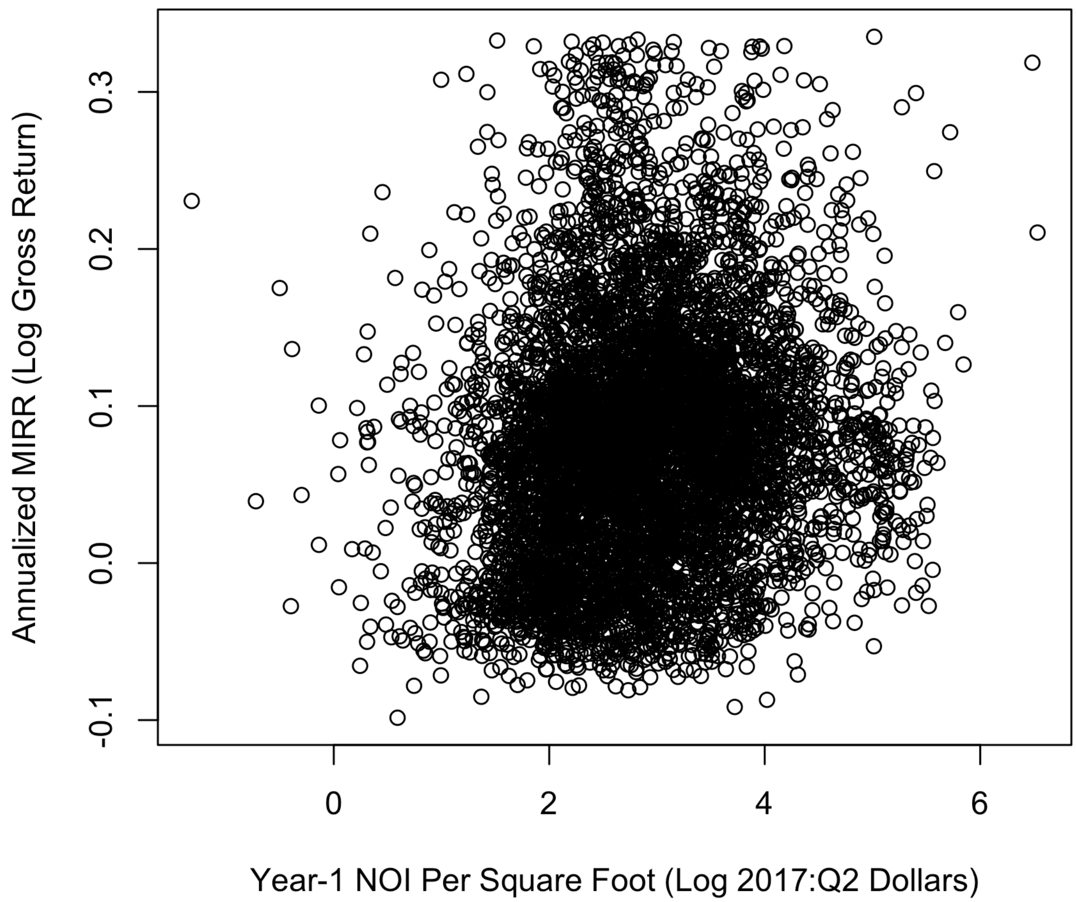

Figure 6. Acquisition price and annualized MIRRs

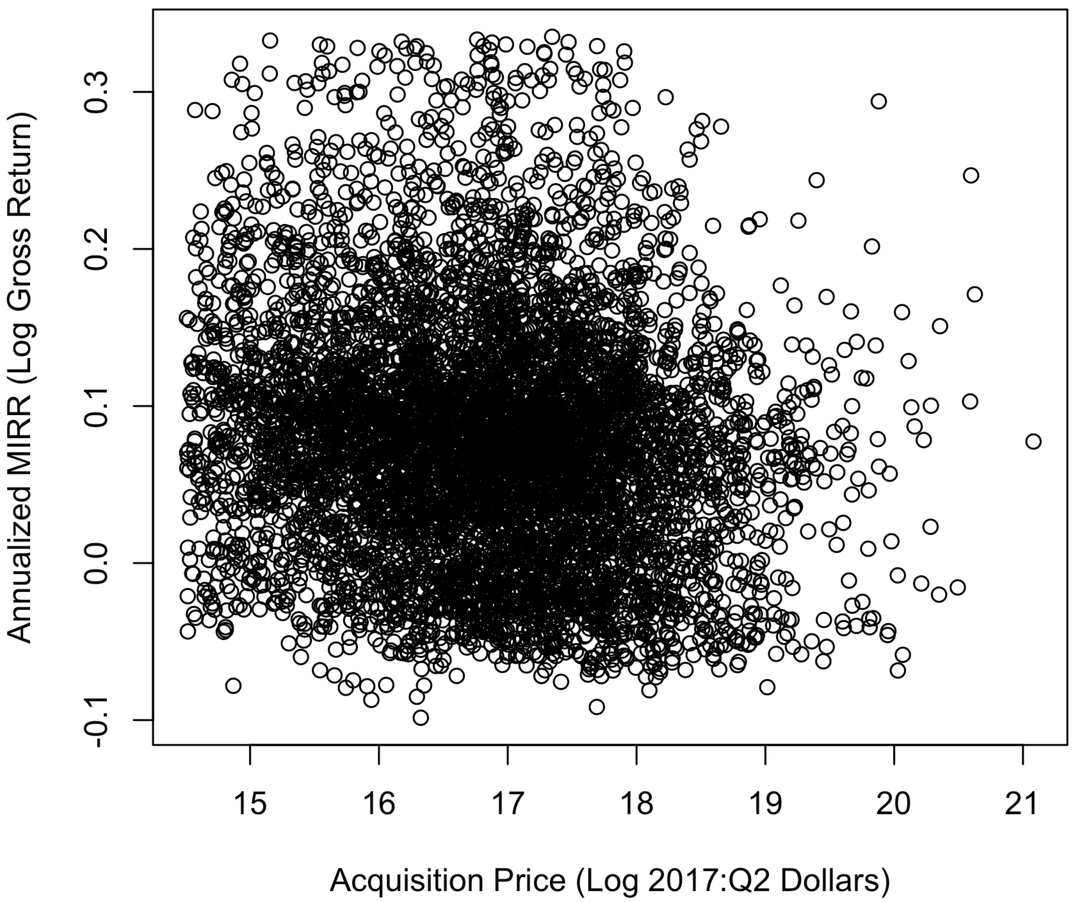

Figure 7. Duration and annualized MIRRs

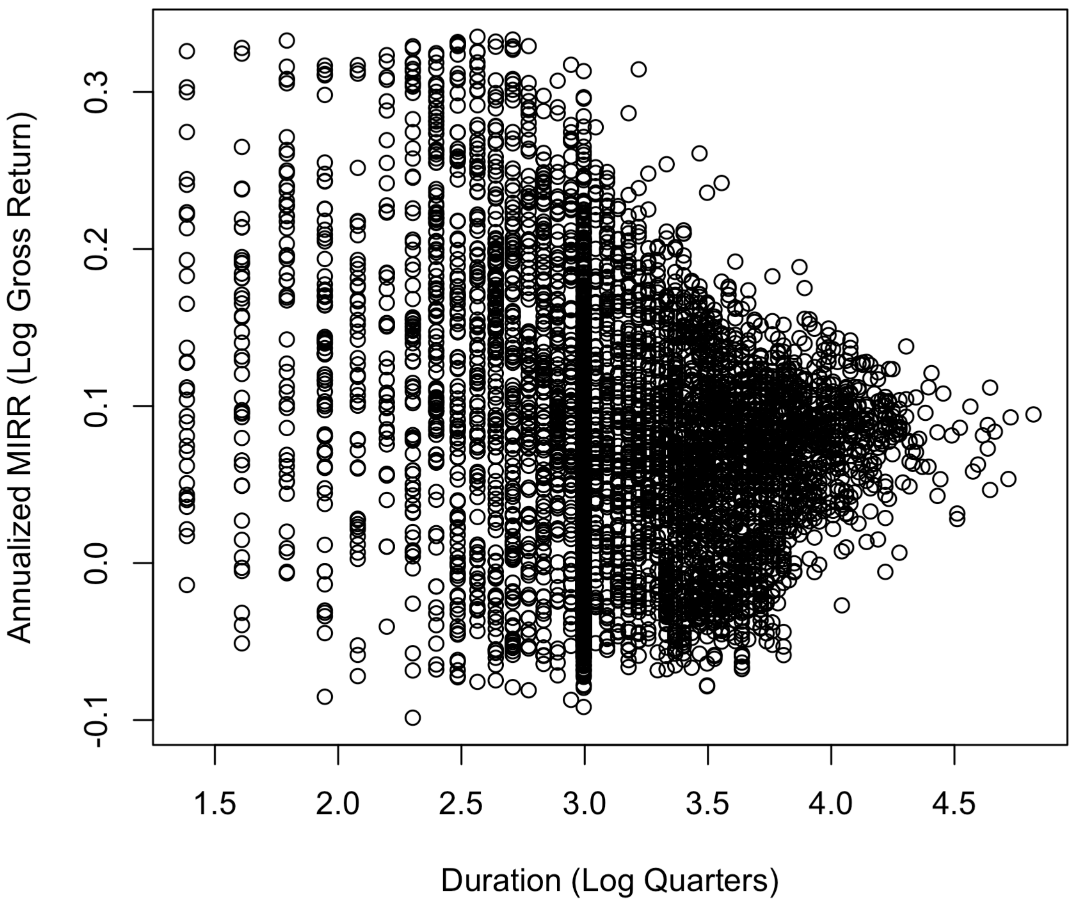

Figure 8. Nationally- and locally-normalized quality

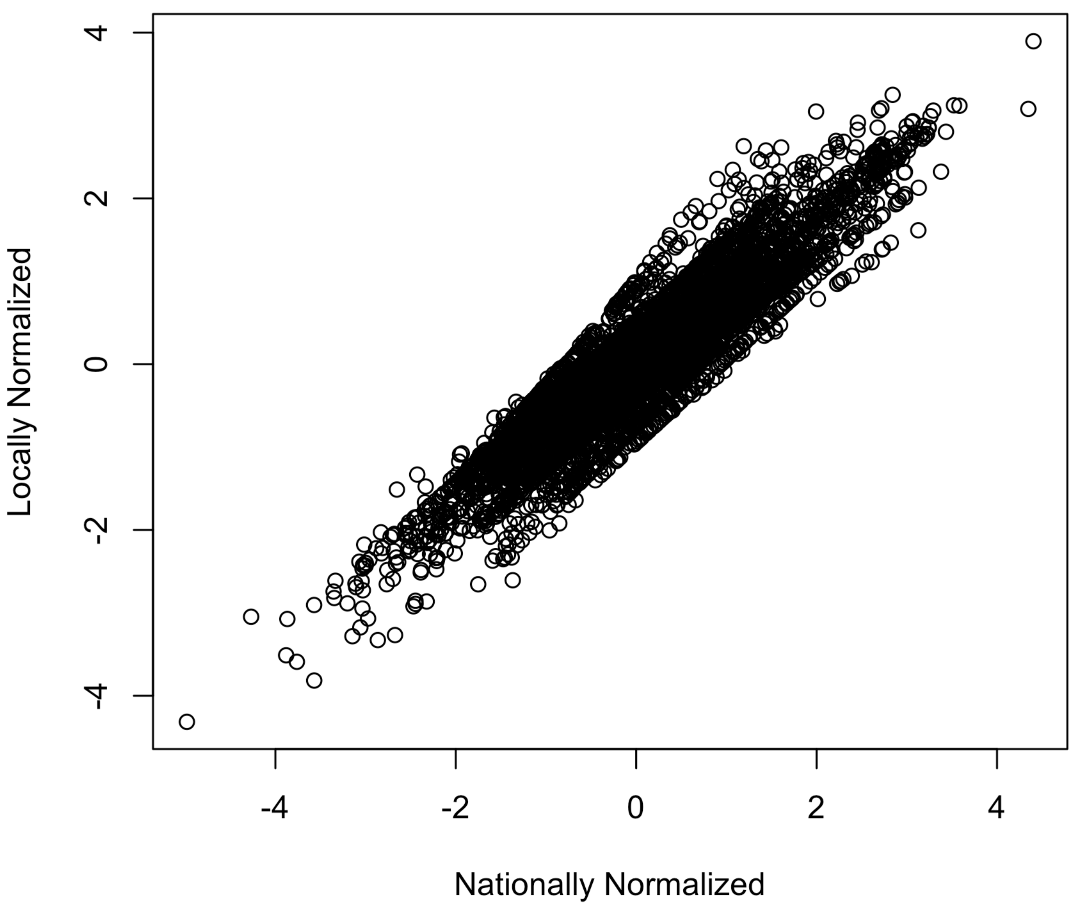

Figure 9. The real estate factor (real estate risk premium index)

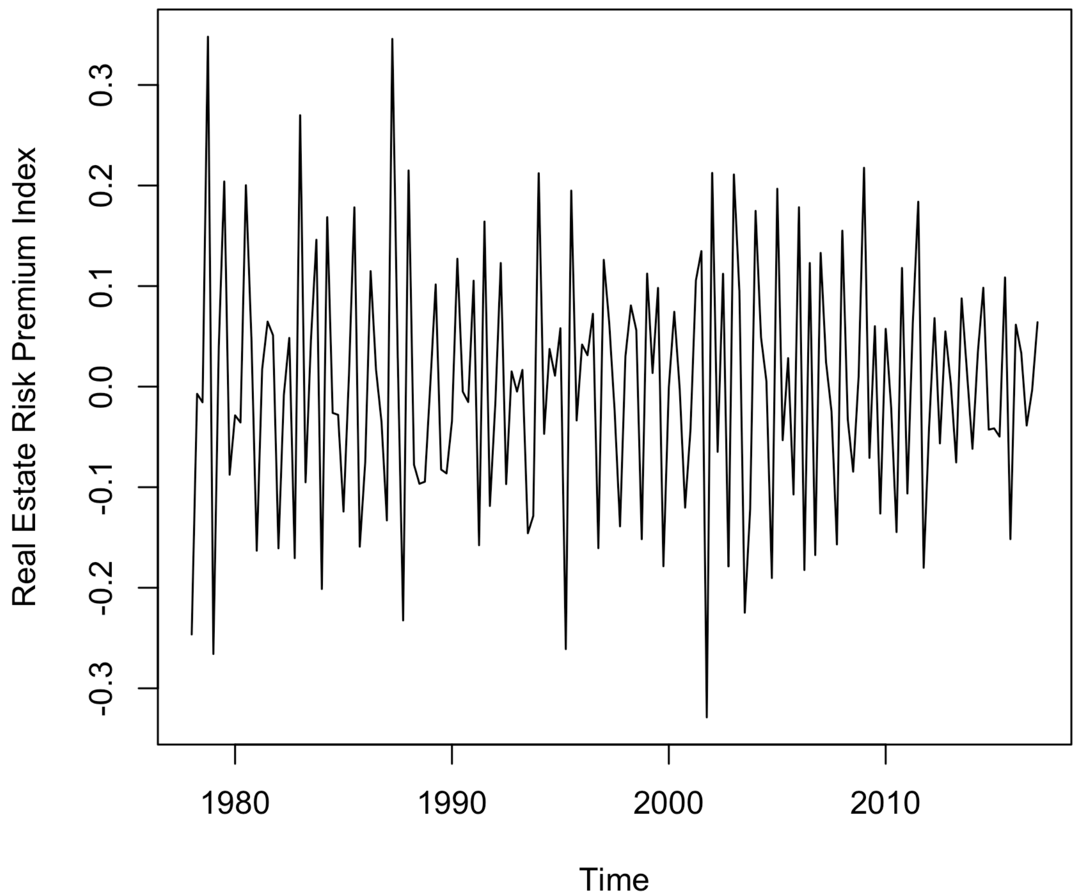

Figure 10. Out-of-sample explanatory power of the real estate factor (real estate risk premium index): histogram of coefficients

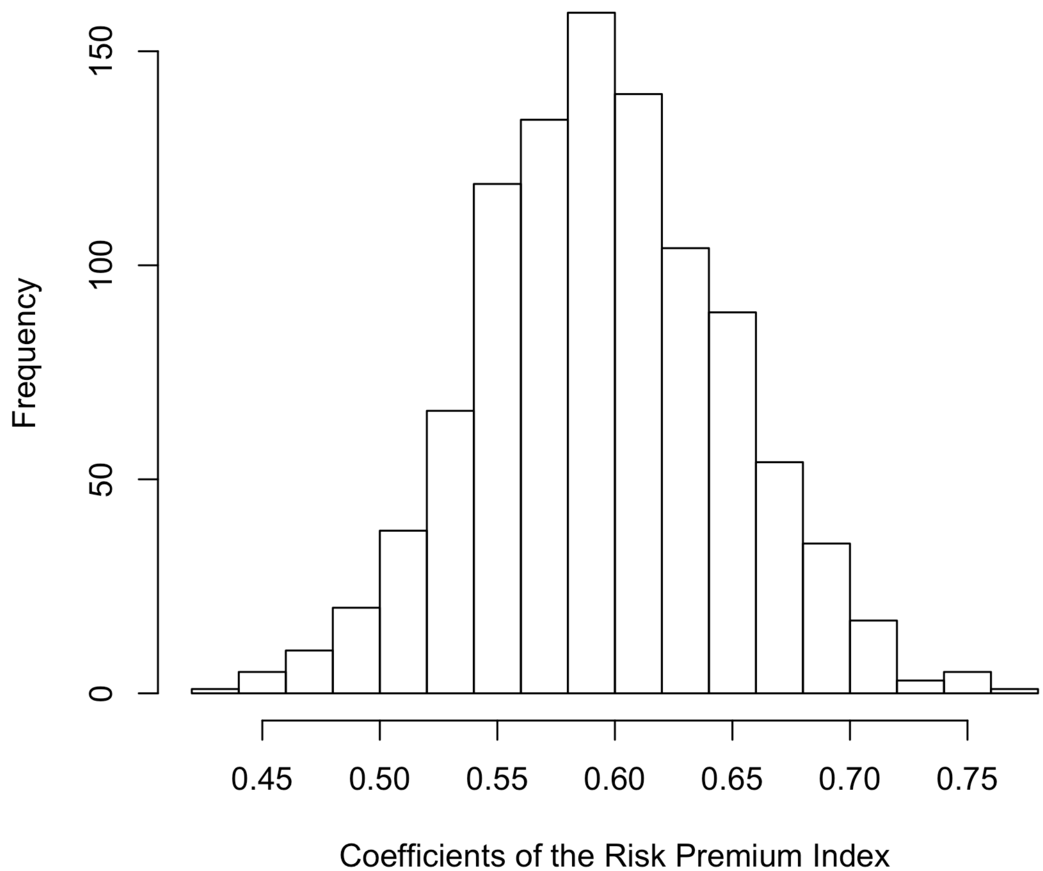

Figure 11. Histogram of peculiar risk

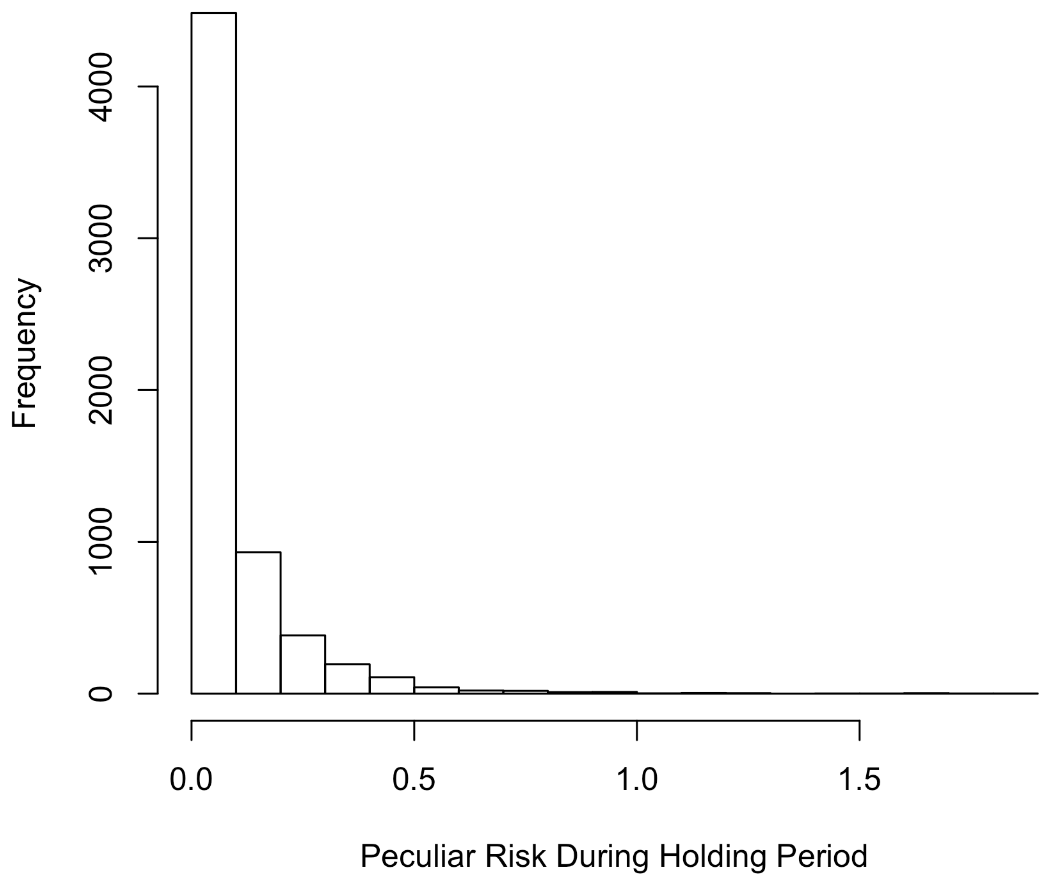

Figure 12. Histogram of acquisition cap rates