Market Timing and Investment Selection: Evidence from Real Estate Investors*

Yael V. Hochberg
Jones Graduate School of Business
Rice University & NBER

Tobias Mühlhofer
School of Business Administration
University of Miami

August 31, 2016

Abstract

We examine commercial real estate fund managers' abilities to generate abnormal profits through selection of outperforming property sub-market segments or through the timing of entry into and exit from sub-markets. The vast majority of portfolio managers exhibit little market timing ability, with the exception of non-NYSE REITs after the financial crisis. A substantial fraction of managers seem able to successfully select property sub-markets. Selection performance exhibits significant persistence. Managers that are active in more liquid markets tend to exhibit better timing performance, while managers exhibiting better selection ability appear to be active in less liquid markets.

Keywords: Active Portfolio Management, Alternative Asset Classes, Commercial Real Estate, Market Timing, Investment Selection.
JEL Codes: G11, G23, R33.

*We thank Jim Clayton, Jeff Fisher, Shaun Bond, Jay Hartzell, Sheridan Titman, Charles Trzcinka, Russ Wermers, Barney Hartman-Glaser and seminar and conference participants at the Western Finance Association Annual Meetings, the University of California at Berkeley, the University of Texas, Indiana University, the Securities and Exchange Commission, the University of Cincinnati, Syracuse University, Cornerstone Research and the Real Estate Research Institute Annual Conference for helpful discussions and suggestions. We thank NCREIF for provision of data on private property holdings. Both authors gratefully acknowledge funding from the Real Estate Research Institute. Hochberg additionally acknowledges funding from the Zell Center for Risk Research at the Kellogg School of Management. Address correspondence to hochberg@rice.edu (Hochberg), tobias.muhlhofer@gmail.com (Mühlhofer). Portions of this research were conducted while Mühlhofer was a visiting faculty member at the University of Texas at Austin.

Active money management plays a significant role in the financial services industry. Whether the substantial fees charged by professional portfolio management companies are justified by their ability to produce abnormal returns has been the subject of a large research literature, beginning with Jensen (1968). While much of the literature has focused on mutual funds, more recently, researchers have begun to examine similar questions for alternative asset classes. In this paper, we examine the ability of portfolio managers to generate abnormal profits in a large alternative asset market: Commercial Real Estate.

The commercial real estate (CRE) market represents a significant portion of the investment universe, rivaling the size of the publicly traded equities market. The total value of the asset class was estimated at $11 trillion as of the end of 2009 (Florance, Miller, Peng and Spivey (2010)).[1] In contrast to the mutual fund setting, where markets are thought to be relatively efficient and evidence on managerial value-added is mixed at best, the CRE market represents a relatively inefficient, illiquid and opaque market with greater scope for informational advantages that may lead to abnormal profits.

In this paper, we examine the ability of CRE portfolio managers to generate abnormal profits through two specific forms of value-added ability: the selection of outperforming property sub-market segments, and the timing of entry into, and exit from, sub-markets. The primary investment decisions made by a CRE portfolio manager are typically the choice of geographic and property-type sub-markets in which they invest. Thus, portfolio managers' ability to generate abnormal profits is largely viewed by practitioners as their ability to select a city (CBSA) and property type in which properties will outperform the broader commercial real estate market.[2] Evidence in the academic literature, however, suggests market timing may also play a role in portfolio manager performance. Existing

[1]For comparison, Wilshire estimates the total market capitalization of US publicly traded equities at $12 trillion at the same point in time.

[2]Evidence for this statement can be found, for example, in the annual reports and 10-K filings of REITs. As an example, consider Simon Property Group (currently the largest REIT) and its 10-K for the year 2010. Simon, in its portfolio description (p.13), characterizes its investment choices primarily by subtype and location. In its property table (pp.14–32), once again the primary attributes for the firm's investment properties are size and location (CBSA). The discussion of the company's development pipeline (starting p. 81) also characterizes investment choices exclusively by city. Similarly, Camden, a large apartment REIT, lists on page 6 of its 2010 annual report, some highlights of its portfolio. In this listing, investment choices are only defined by city and state (i.e. CBSA). Property type is superfluous, as Camden only invests in apartment complexes. Starting page 10 of its 10-K filing for the same year, in the "Property Table", all investment choices are characterized primarily by size and location.

studies of the real estate market suggest that property markets display evidence of predictability (see e.g. Plazzi, Torous and Valkanov (2010), Liu and Mei (1992, 1994), Barkham and Geltner (1995), Case and Shiller (1990), Case and Quigley (1991)). Furthermore, studies employing simulated technical trading strategies suggest that market timing profits can be made in the real property market (e.g. Geltner and Mei (1995) and Mühlhofer (2015)).

We utilize a complete dataset of property holdings by REITs invested in by major institutions, distinguishing between REITs traded on the NYSE, who tend to hold institutional grade properties, and those traded on AMEX or NASDAQ, who often do not. We augment this data with a dataset of property holdings of portfolio managers of private entities, such as commingled real-estate funds. [3] Our data represent approximately half of the equity invested into the CRE asset class as a whole, with the remaining half encompassing, to a large extent, sub-institutional grade properties and owner-occupiers who do not engage in delegated portfolio management.

To assess portfolio managers' abilities to generate abnormal profits, we calculate measures similar in nature to those of Daniel, Grinblatt, Titman and Wermers (1997), utilizing sub-market returns where DGTW employ individual investment returns and using a broader style benchmark return where DGTW employ a characteristic benchmark.[4] Our timing measure captures a portfolio manager's tendency to take advantage of mean reversion and tilt the composition of their broader portfolio towards certain sub-markets when these outperform and away from these sub-markets when they underperform. Our selection measure captures a portfolio manager's ability to consistently select sub-market categories (or 'styles') which outperform broader benchmarks. Our analysis is thus similar in spirit to the style-based analysis of Fung and Hsieh (2002, 2004), who perform such an analysis for hedge funds. Fung and Hsieh (2002, 2004) argue that investment classification along style-based lines (here, sub-markets) is both better-suited and economically warranted for alternative asset classes relative to analysis using the common asset pricing factors, as is frequently done for mutual funds. Assuming that benchmarks capture the risk characteristics of the market, no further

[3] In contrast to the mixed nature of properties held by so-called institutional grade REITs, properties held by co-mingled real estate funds (CREFs) are all institutional grade.

[4] While we do not employ individual property returns, prior work has shown that such returns are largely similar to the returns of the sub-market in which they are located. See e.g. Crane and Hartzell (2007).

risk-adjustment of returns should be necessary. However, as is the case in the setting of Fung and Hsieh (2002, 2004), we cannot unequivocally rule out risk as an explanation for observed return differentials.

Consistent with the common practitioner view, we find that the vast majority of REITs (both NYSE- and non-NYSE-traded) and private portfolio managers exhibit little ability to successfully time their investments. Throughout the sample period, both NYSE REITs and private managers (both of whom hold primarily institutional grade properties) exhibit similar patterns of negligible timing ability. Examining performance before and after the financial crisis of 2007-2008, however, reveals different patterns for NYSE REITs and private managers than those observed for NASDAQ and AMEX traded REIT managers, who hold smaller, non-institutional grade properties and properties in smaller markets. These latter types of properties fell precipitously in value during the financial crisis of 2007-2008, and to a much greater extent than the fall for larger, institutional grade properties. Furthermore, the recovery of these smaller markets lagged the recovery in more institutional-grade markets. Consistent with the hypothesis that non-NYSE traded REIT managers were able to take advantage of this lagged recovery in those markets, we find that AMEX and NASDAQ traded REIT managers exhibit positive timing measures in the post-crisis period, suggesting that managers of funds holding these smaller properties were able to take advantage of the large magnitude of mean reversion and increased predictability in prices for these properties post-crisis.

In contrast, a substantial fraction of all types of managers exhibit an ability to select outperforming sub-markets. Sub-market selection ability is strongly persistent, especially in performance rankings of managers. Moreover, sub-market selection ability and timing ability appear to be negatively correlated.

Despite differences in taxation and trading restrictions for private and REIT portfolio managers,[5] we observe little difference between the overall timing and selection abilities of private and REIT

[5] While REITs do face some holding period restrictions, we do not believe these should truly restrain REIT managers from rebalancing portfolios to time markets. Institutional features allow most REIT managers to circumvent such restrictions. For example the so-called "*dealer rule*" states that REITs have to hold each property they purchase for four (or after 2008, two) years, and can only sell 10% of their portfolio in a given tax year. However, Mühlhofer (2014, 2015) show that Umbrella Partnership REITs (UPREITs) are not significantly affected by this constraint, and that UPREITs constitute the majority of REITs in our sample period.

portfolio managers beyond the differences in timing measures between NYSE and NASDAQ/AMEX REITs during the post-crisis period. Overall, our findings suggest that in a market such as the CRE market, in which transactions are often costly and time-consuming,[6] both REIT and private fund managers appear to generate value primarily through asset subclass selection.[7]

If the major barrier to market timing in the CRE setting is a lack of liquidity, we would expect to observe an association between the liquidity of the sub-markets in which a manager invests and his success in timing- or selection-performance. A timing strategy requires relatively quick transactions in order to move in- and out of the market when necessary and should therefore be more feasible in a sub-market segment with higher liquidity. Conversely, selection strategies, which are more passive, do not require this type of liquidity, and so managers who engage in this type of strategy may choose to invest in less liquid markets and avoid paying liquidity premia.

When we regress timing and selection performance on measures of market liquidity, we find widespread evidence consistent with the idea that market liquidity is associated with relatively better timing performance, and market illiquidity with better selection performance. This further supports the hypothesis that the poor timing performance exhibited by managers is partly attributable to the property market's overall lack of liquidity. These relationships are present even when restricting the sample period to pre-crisis years, suggesting that the pattern we observe is not simply the result of prices of properties in less liquid markets falling further than in more liquid markets during the financial crisis.

Our work contributes to the large literature exploring the generation of abnormal returns by portfolio managers. While much of this literature has focused on mutual fund managers,[8] our work contributes to an emerging literature exploring manager skill in alternative asset markets (see e.g.

[6]Commercial property practitioners regard three to twelve months to complete a property sale as reasonable.

[7]Since real estate is fixed in location, practitioners argue that managers are essentially betting on the economic base of a given location. In gateway cities (e.g. San Francisco) where supply is constrained, there is little new construction, and demand for properties is high, abnormal returns should occur in the short run. Thus, for example, to the extent that REIT and private managers invest in gateway cities, they should experience abnormal returns. Hess and Liang (2000) observe that while both REITs and private managers share a preference for large MSAs, private (REIT) investors have a greater concentration of investments than REIT investors in western (Midwest and South) cities.

[8]Abnormal profits (or the lack thereof) for mutual funds in the stock market have been studied extensively in the literature (see e.g. Jensen (1968, 1969), Brown and Goetzmann (1995), Gruber (1996), Carhart (1997)). Managers' ability to select individual investments versus benchmarks and to time the market have been studied by Daniel et al. (1997) as well as e.g. Wermers (2003), Kacperczyk, Sialm and Zheng (2005).

Kaplan and Schoar (2005) in private equity and venture capital markets and Bond and Mitchell (2010) in real estate). In contrast to existing studies of other alternative asset markets, which are often limited by data availability[9], the CRE data we employ in this study allows us to conduct a detailed analysis of manager returns in such markets. While contemporaneous studies such as Bond and Mitchell (2010) also use private real estate fund data to assess outperformance, their work focuses on performance measures such as overall *alpha*, whereas our analysis assesses the timing and sub-market selection components that may drive such *alpha*.

The remainder of this paper is structured as follows. Section 1 describes the data we employ for our analysis. Section 2 details our primary empirical analysis. Section 3 discusses and concludes.

1 Data

Our analysis is focused on the institutional grade CRE market, a segment valued at $3-4 trillion in 2009.[10] The data for our analysis are obtained from three primary data sources.

We obtain property transaction data for REIT portfolio managers from SNL Financial, which aggregates data from 10-K and 10-Q reports of all publicly traded REITs that are considered *institutional-grade*. The SNL Financial DataSource dataset provides comprehensive coverage of corporate, market, and financial data on institutional-grade publicly traded REITs and selected privately held REITs, and REOCs (Real Estate Operating Companies). While the resulting dataset consists of "institutional grade" REITs, i.e. REITs held by major institutions, this does not ensure that these REITs actually hold institutional grade properties. Rather, larger REITs, typically traded on the NYSE, tend to hold institutional grade properties, while smaller REITs, typically traded on NASDAQ or AMEX, hold smaller, often non-institutional grade properties. Smaller properties experienced larger declines during the financial crisis, and their recovery lagged that of larger properties in the post-crisis period. We therefore split our REIT sample by exchange, separately analyzing

[9]Data limitations in VC and PE include such difficulties as being able to observe only venture-capital financed firms that went public, having to rely on voluntarily reported investment returns, or by being forced to use other indirect public-market related measures to infer information about the more inefficient private market.

[10]Institutional grade in this setting refers to being held by institutions, though the properties within the individual portfolios may themselves not necessarily be investment grade.

NYSE-traded REITs and NASDAQ- and AMEX-traded REITs.

The SNL data contains accounting variables for each firm, as well as a listing of properties held in each firm's portfolio. For each property, SNL lists a variety of property characteristics, as well as property transaction data. By aggregating across these properties on a firm-by-firm basis in a particular time period, we can compute a REIT's fractional exposure to particular sets of characteristics such as property type and geographic segment. The SNL REIT sample runs from Q2 1995 through Q4 2013.

We obtain property transactions data for private real estate portfolio managers from the National Council of Real Estate Investment Fiduciaries (NCREIF), which collects transaction-level data for private entities (primarily pension funds). While membership in NCREIF (and thus reporting of transactions to NCREIF) is voluntary, inclusion in NCREIF's database is considered desirable and prestigious on the part of private managers. NCREIF's stated policy is to only report data on high-grade institutional-quality CRE. Inclusion of one's property transactions in NCREIF's database and indices is viewed as confirming a level of quality on the included investor. As a result, most eligible managers choose to become members of NCREIF, and thus subject themselves to quarterly reporting of transactions. NCREIF membership constitutes a long-term contract and commitment, and once included, it is not possible for an investor to report performance only in certain quarters and not in others; the investor is contractually obligated to report all transactions going forward. Data reported by NCREIF members to NCREIF is protected by a strict non-disclosure agreement.[11] Thus, manipulating performance numbers is viewed as ineffective. This arrangement gives us the opportunity to examine trades in a large private asset market in some detail. The NCREIF sample runs from Q1 1978 through Q2 2014. We note that all our analysis is robust within sub-periods and to excluding the portions of the NCREIF sample prior to 1995:Q2 and after 2013:Q4 when the REIT sample begins and ends.[12]

[11] As academic researchers, we are given access to NCREIF's raw data under the same non-disclosure agreement.

[12] Directly evaluating the portfolio of property holdings of REITs and private investors allows a more micro-level of analysis than the oft used analysis of mutual funds of REITs enables (see e.g. Kallberg, Liu and Trzcinka (2000), Hartzell, Mühlhofer and Titman (2010)), in that it allows us to paint an exhaustive picture of the actual investment decisions in the privately traded CRE market. As we are able to observe the individual property characteristics, we can readily calculate the portfolio weights of each geographic and property-type sub-market in the portfolio, and select benchmarks that are appropriate for computing timing and selection ability measures. Knowing the timing of individual

For our sub-market segment returns and benchmarks, we use two data sources. First, we employ real estate market returns at various levels of aggregation obtained from the National Property Index (NPI) series (also compiled by NCREIF). Second, we utilize the Commercial Property Price Indices (CPPI) from Real Capital Analytics (RCA).

The NCREIF NPI is considered the de-facto standard performance index for investible US CRE.[13] Index series are available on a national level, as well as disaggregated by region, division, state, CBSA, property type, property sub-type, and by interactions of the property type and geographic sub-categories. The NPI consists of three indices: appreciation, income and total return. As would be the case in analysis of stock market returns, we employ the total return indices for our analysis. NCREIF only reports NPI series that constitute 'market' returns, and do not report series that are dominated by a few properties or investors (see methodology descriptions at http://www.ncreif.org). We take the further conservative approach of discarding all markets in which the NPI series is built from the trades and returns of less than 10 properties.[14] The NCREIF data sorts properties into five major type categories: Apartment, Hotel, Industrial, Office and Retail, with all but hotel broken down into two to eight further subtypes (e.g. Apartment: Garden, Apartment: High-rise and Apartment: Low-rise). Additionally, properties are classified as belonging to one of four Regions: East, Midwest, South and West, which in turn are broken down first into two Divisions each (e.g. East: Mideast and East: Northeast) and then further by State and CBSA (not detailed for brevity).

While the NCREIF NPI index returns are commonly used as a benchmark in the CRE setting, the price appreciation portion of the NPI series is based on appraised values where transaction prices are not available. Biases associated with the use of property appraisal data are well-documented. Due to these concerns, we repeat our analysis using the RCA CPPI. This set of indices is based on a repeat-sales methodology and should therefore only contain transaction-based data. A sub-

property transactions also allows us to more accurately compute portfolio weights across time.

[13]NCREIF gathers data on the property investments of private institutions for the purpose of constructing the NPI.

[14]Technically, the performance of the NPI indices is based on the trading performance of all NCREIF members combined. Thus, upon first glance these indices may not seem to constitute a passive benchmark. However, NCREIF's universe covers such a large portion of privately-held institutional-grade commercial real estate that when aggregated, this trading performance essentially shows the entire market's transactions. It is therefore reasonable to view these benchmarks as passive in like manner to a stock index, which ultimately also reflects the aggregate trading behavior of the market. This view of the NPI indices as passive benchmarks is also generally held by market practitioners.

stantial set of properties is required to make such an index possible, and therefore the lowest level of aggregation for which the RCA indices have nationwide coverage is Region/Type. The data for RCA's indices begins at the end of 2000, and consists solely of the price appreciation series. As we wish to examine overall investment performance–appreciation plus income[15]–we add the NCREIF income-return component constructed for that particular sub-market to the RCA-based benchmarks. NCREIF's income return component comes directly from a property's proforma, and involves no appraisal data. The geographic regions used by RCA are defined slightly differently from the NCREIF regions. To compare across benchmark sets, we re-compute NCREIF indices that match the RCA regions, by taking weighted averages of NCREIF's state-level series for each state in an RCA region. Utilizing the full geographic coverage of the data allows us a level of de-aggregation that stops at the region $\times$ property-type level, as compared to the lower levels of aggregation obtainable using the NCREIF data. As discussed in the robustness section of the paper, subject to the caveats of shorter time series and higher levels of aggregation, utilizing the RCA data provides us with similar results to those obtained using NCREIF data.

Our data allow for many levels of disaggregation at both the geographical and property type levels, as well as the interaction thereof. While this creates many degrees of freedom for analysis, it allows us a very complete view of how value may be generated by managers through market timing or sub-market selection. NCREIF subdivides property types into subtype for all types, while SNL does not provide certain subtypes for some property types. Wherever SNL data does not provide a property subtype, we employ the property type benchmark (i.e. one level higher of aggregation) in place of a subtype benchmark. Table 1 describes the breakdown of property types and sub-types as well as geographical regions and division for non-NYSE-traded REITs, NYSE-traded REITs and private manager samples. Our data can additionally be disaggregated to the state and Core-Based Statistical Area (CBSA) level. [16] The table details the number of unique properties transacted in by REITs and private portfolio managers for each property type and subtype, as well as each Region and Division. While benchmarks are thus available at a variety of levels of aggregation,

[15]We do not distinguish whether a manager creates value by investing in property that generates superior rental income, versus investing in property whose market price increases generate superior returns

[16]For brevity, we do not detail State and CBSA-level breakdowns in the table.

in much of the analysis, we see substantially similar patterns regardless of the level of benchmark aggregation. We thus report results for a limited number of levels of benchmark aggregation, focusing on the Divisional level of geographic aggregation (which constitutes the *middle* of the geographic size classifications), for most of our univariate analysis.[17]

Table 2 presents summary statistics for the sample, broken out by type of manager (non-NYSE REIT, NYSE REIT, private). The table presents time-series statistics of quarterly holdings. Our sample consists of 166 non-NYSE REIT portfolio managers[18] transacting in 4,159 properties, 136 NYSE-traded REIT portfolio managers transacting in 13,025 properties, and 124 private managers transacting in 22,322 properties over the course of the time period in question. Private managers hold the largest portfolios by square feet (55.6 million), approximately twice the size of portfolios held by NYSE-traded REITs. Non-NYSE traded REITs hold the smallest portfolios by square feet, at roughly 17 million square feet per portfolio, on average. The individual properties held by private managers are also slightly larger, on average, at 272,000 square feet per property, while non-NYSE REITs and NYSE REITs hold somewhat smaller properties, at 193,000 and 189,000 square feet per property. REIT managers of both types hold portfolios that are more concentrated by both geographical segment and property type than do private managers. Of the three groups, NYSE REITs exhibit the longest property holding periods, at over 7 years, compared to the non-NYSE and private managers, who both typically hold properties for a little over 5 years.

While our underlying datasets allow us to calculate sub-market exposures by square footage, we do not observe market valuations of individual properties for REIT holdings, and therefore cannot calculate exposures by property value for these managers. For the private portfolio manager sample, we do observe quarterly estimates of market value for each property held. We note that all the results reported hereafter for our private manager sample are robust to the use of value weights rather than square footage weights.

[17] Results for other levels of aggregation are available from the authors upon request.

[18] Throughout this study we use the term "manager" to refer to entire organizations (REITs or private funds) in charge of portfolio management. Our data does not allow us to see the turnover of actual management teams within organizations.

2 Empirical Analysis

Our goal is to identify managerial ability to generate abnormal profits through sub-market selection and timing. We thus begin by classifying the holdings and trades of portfolio managers by geographic location, property class, and size (square footage). We use this classification to characterize each portfolio manager's holdings as part of a sub-market investment category (i.e. style). We then construct measures to evaluate the manager's ability to generate value within- or across categories over time. Our approach is thus similar in spirit to the asset-based style factors analysis of Fung and Hsieh (2002, 2004), who show that an investment classification and performance assessment along such lines is better suited and more economically warranted for alternative investments than an analysis using common asset pricing factors.[19] As in Fung and Hsieh (2002, 2004), we match on characteristics. As our setting is real estate markets, our characteristics are property type and location, which are analogous to the use of styles in hedge funds. Assuming that benchmarks capture the risk characteristics of the market, no further risk adjustment of returns should be necessary. However, as is the case in Fung and Hsieh (2002, 2004), we cannot unequivocally rule out risk as an explanation for differences in returns.

Data for privately traded CRE suffers from well-known issues and biases related to the market's low number of trades and to the subsequent necessity to make some use of appraisals in its analysis. As is widely documented in the real estate literature (see e.g. Geltner (1991), Clayton, Geltner and Hamilton (2001)), the resulting smoothing problems primarily lead to errors estimating second moments and co-moments of returns, while leaving longer-term first moments largely unaffected in multiple period analysis.[20] The lack of reliable second moments makes parametric, regression-based analysis of investment performance (along the lines of Carhart (1997), Treynor and Mazuy (1966) or Henriksson and Merton (1981)) problematic. Instead, we resort to a holdings-based, non-parametric approach, which consists of constructing weighted sums of outperformance across a

[19] Further discussion of the appropriateness of various factor types for hedge funds abound in the hedge fund literature. See, for example, Titman and Tiu (2011), for a discussion of the appropriateness of different types of factor model for this asset class, as well as a review of the literature that treats this question.

[20] Stale pricing issues have also been examined in other asset classes, such as the literature examining bond fund performance (for example Chen, Ferson and Peters (2010)).

manager's portfolio at each point in time.[21] This outperformance can occur across time, within a style choice (*Market Timing*), or cross-sectionally between the performance of a manager's actual investment choice and the performance of a larger encompassing style portfolio (*Market Selection*). Our analysis is both robust to desmoothing of the NPI indices to account for possible appraisal bias or smoothing, and to the use of the alternate RCA indices, which do not suffer from appraisal bias concerns.

2.1 Market Timing

To measure managers' ability to time entry and exit from a style, we compute a *Market Timing* measure, defined as follows:

$$MT_t = \sum_{j=1}^{N}(w_{j,t-1}R_{s_j,t} - w_{j,t-5}R_{s_j,t-4}), \tag{1}$$

where $w_{j,\tau}$ is a fund's fractional exposure to property j at the end of quarter τ, and $R_{s_j,\tau}$ is the return to a passive portfolio that mirrors the broad investment style s_j to which property j belongs.[22]

The measure compares the actual style-choice induced return component from a particular sub-market exposure in a particular quarter to the style-choice induced return component that was earned by the portfolio manager's exposure to this sub-market a year earlier. The measure will be positive for any time period in which a fund's weighted return derived from exposure to a particular style exceeds the weighted returns from that fund's exposure to this style a year earlier. If the manager increases portfolio exposure to a style in an upturn and decreases exposure in a downturn, such positive timing ability will be captured by the MT measure.[23] We calculate the measure using

[21] Non-parametric, holdings-based procedures of performance evaluation are also used in other literatures on managerial value-added, such as the mutual fund literature, as an alternative to return-based factor models. See, for example Daniel et al. (1997) whose methodology resembles ours in terms of mechanics. Other studies, such as Jiang, Yao and Yu (2007) also argue in favor of using non-parametric, holdings-based techniques in such a setting.

[22] Employing non-parametric techniques of the nature of Equation 1 has an additional advantage in our setting. As noted before, commercial property returns data is likely to suffer from appraisal smoothing which overstates autocorrelation in returns. However, under the assumption that the price series and the one-year lagged price series are cointegrated, this concern will be mitigated in performance measures that utilize lagged differencing of the returns series, as the procedure should largely remove such overstated persistence.

[23] Methodologically, this approach resembles the measures constructed by Daniel, Grinblatt, Titman and Wermers

annualized quarterly return series.

We compute the fractional exposure weights for Equation 1 using the fraction of the manager's total square footage under management in a particular quarter that is constituted by properties in a particular sub-market (i.e. the sum of individual property square footages held by the portfolio manager in that sub-market divided by total portfolio square footage). For our sub-market returns, we employ NCREIF's National Property Index (NPI) total return indices at the lowest levels of aggregation. For example, if a manager owned office buildings in Chicago's Central Business District (CBD) in the first quarter of 2006, the relevant return to the passive portfolio that mirrors the most narrowly-defined associated sub-market would be the total return to NCREIF's Chicago CBD Office sub-index for the quarter. Summing up weights across all properties managed by this manager in the particular quarter and sub-market yields the manager's total fractional exposure to this style, which is multiplied by the return to the relevant NCREIF sub-index to generate the weighted return for that style in that quarter. Thus, if Chicago CBD Office property represented 25% of the manager's total property holdings in that quarter, we then multiply the 25% fractional exposure by the return of the Chicago CBD Office sub-market in that quarter. We similarly construct the fractional exposure and sub-market return for the prior year and compute the weighted return the same way.

Following the example in the prior paragraph, and using this approach, a high positive MT value would be generated by the manager's decision to increase portfolio exposure to Chicago CBD Office ahead of a rise in the market, and/or decrease exposure ahead of a slump in the Chicago CBD Office market. This would then be considered positive timing ability with respect to the Chicago CBD Office market overall. We proceed analogously for all other styles (property geography and class sub-markets) to which the manager's portfolio has exposure. The sum across all properties and thereby all sub-markets yields the manager's MT_t measure for that quarter. We repeat this procedure for each quarter the manager appears in our dataset, and then compute time-series statistics by manager.

As our dataset contains benchmarks for various levels of aggregation at the property geogra-

(1997) to measure characteristic timing in the mutual fund market. Here, as in DGTW, we employ a weighted sum of outperformance across time which is constructed for a portfolio in each time period. However, economically, our approach differs markedly from that of Daniel et al. (1997), as their study measures performance over investment *characteristics*, defined along the lines of asset pricing factors, while we measure outperformance of particular sub-market (i.e. *style*) choices over time.

phy/class level, it affords us the ability to examine timing ability at a variety of levels of specialization. Continuing the prior example, while a Chicago CBD office property could, economically, be a bet on the Chicago Office market, it could also be considered as part of a more general bet on the overall Chicagoland Commercial Property Market, a bet on the Midwest Office Market, Midwest Commercial Property Market, or a bet on the nationwide Office market, etc. We thus can construct our MT measures using multiple levels of aggregation in our style benchmarks. At the geography level, we use property portfolio index returns for the CBSA level, the state level, the divisional level, the regional level and at the whole national level. We additionally then interact each geographical level of specialization with property type and sub-type to indicate property class. This gives us 15 variants of the MT measure, each constructed using portfolio weights at different levels of aggregation, for each quarter of the sample. We then additionally compute a time-series average MT for each manager at each level of aggregation. To enable calculation of reliable t-statistics, we include all observations where the manager is present in the time series data for at least 12 quarters.[24] [25]

As a representative benchmark, we report all analysis results with measures computed at the Divisional level of geographic aggregation. Throughout this paper, we find that varying the benchmark aggregation level does not typically qualitatively alter results. Where we find significant variation across aggregation levels, we report multiple levels.[26]

Table 3, Panel A, presents distributional statistics of the time-series average MTs from the cross-section of the three types of portfolio managers. To obtain the statistics in the table, we first compute the MT_t measure described in Equation (1) for each level of aggregation for every manager in every quarter. For each manager, within each level of aggregation, we then compute a time-series average

[24]In similar exercises using weekly data, Hartzell et al. (2010) restrict analysis to observations where a manager is present in the time series data for at least 24 observations. Our results are robust to imposing this more strict requirement. If we eliminate the restriction completely, we obtain similar means, medians and standard deviations for the sample.

[25]Note that measuring MT against the National-level benchmark (which is an average performance of the entire commercial property market) would measure the value generated by a manager's moving funds into and out of commercial property as a whole. Because we do not have data on non-real-estate holdings for the entities we examine (in as far as this is even meaningful), we are unable to assess performance along this dimension. Therefore, we do not calculate measures related to *market timing* with respect to the National benchmark.

[26]This overall picture is consistent with the findings of Boudry, Coulson, Kallberg and Liu (2013), who find that return-benchmarking performance is largely invariant to the number of properties in the benchmark portfolio, once this reaches a level of approximately 20.

over the quarters for which the manager is active in the sample, to obtain a single average MT statistic per manager. The table displays the mean, standard deviation, minimum, first quartile, median, third quartile, and maximum of these measures across managers. As stated before, due to invariance of estimates to specification of benchmark aggregation level, we report measures computed at the Divisional level. For ease of interpretation, the distributional statistics are illustrated in the first panel of Figure 1. The dark line at the middle of each box represents the median of the cross-section of MTs. The boxes represent the inter-quartile spread of the distribution, while the whiskers demarcate 1.5 times the interquartile range from the edge of box. The circles represent outliers that do not fall within the whiskers.

As is apparent both from the table and the figure, neither of the two types of REIT managers nor private portfolio managers appear to exhibit particular skill at timing. Across the board for all three types of managers, a significant fraction of managers (over half) exhibit negative MT point estimates. Non-NYSE REIT managers have significantly more dispersion in timing measures than do their NYSE-traded REIT and private counterparts. 31% of non-NYSE REIT managers have positive point estimates for timing ability, in contrast to 12% for NYSE-traded REIT managers and 13% for private managers. In this respect, NYSE REITs look similar to private managers, likely due to the similarities in the properties in which they transact. This is true regardless of whether we employ a pure geographic benchmark or a geography-type interaction benchmark.

To better assess whether some managers are able to successfully and significantly time the market versus the style benchmarks, we further compute distributional statistics for the t-statistic testing the hypothesis that a manager has zero timing ability against the two-sided alternative (see e.g. Hartzell et al. (2010)). These distributional statistics are reported in Table 3, Panel B. The t-statistics are calculated for the hypothesis that the time-series mean MT is equal to zero for each manager (such t-statistics are often refereed to as an "Information Ratio" in the mutual fund literature). For ease of interpretation, these distributional statistics are illustrated in the second panel of Figure 1.

If fund returns are purely random in the cross-section (with the mean manager generating zero outperformance), pure statistical chance would on average cause 2.5% of funds to appear to have statistically positive outperformance at the 5% significance level (assuming a two-tailed test). To distinguish true outperformance from such "False Discoveries", we follow the procedure suggested by Storey (2002) and adapted to the purpose of cross-sectional performance studies by Barras, Scaillet and Wermers (2010), and compute a *False Discovery Ratio*.[27] This measure hinges upon the recognition that the cross-section of p-values associated with a hypothesis test of zero outperformance should, if fund returns are entirely random, show a uniform distribution from zero to one. Identifying the existence of true outperformance is then accomplished by comparing the fraction of managers that show an apparent statistical outperformance to the fraction of managers that should show such outperformance if p-values were uniformly distributed. The *False Discovery Ratio* (FDR^+) is calculated as the fraction of apparent outperformance that we should see purely by chance for the specific empirical distributions we encounter in the data. When the fraction of t-statistics that are above the 5% significance level ("Fraction Sig.") is in excess of FDR^+, this then indicates the existence of managerial outperformance in excess of what could be expected by chance. We report the positive *False-Discovery Ratio* (FDR^+) in Table 3, Panel B.

Looking at the distributional statistics in the table (and as illustrated by the figure), it is apparent that up to and beyond the third quartile, the timing abilities of all types of managers are statistically indistinguishable from zero.[28] Some managers in the upper quartile of each institution type appear to possess significantly positive timing measures. Comparing the fraction of managers in each category for whom timing measures are significantly positive to the FDR^+, however, it is only for non-NYSE REITs that this fraction exceeds the number that should appear positively significant by mere statistical chance. Altogether, the observed patterns suggest that while some managers in the

[27] A second common technique applied in the recent literature examining cross-sectional distributions of outperformance is the bootstrap-analysis proposed by Kosowski, Timmermann, Wermers and White (2006). Such a technique, however, is not feasible in our setting, as it is based upon the use of a parametric factor-model analysis to generate the original cross-sectional distribution of outperformance, and therefore requires reliable second moments. A similar concern would apply to a procedure such as Fama and French (2010). An additional impediment to this alternative approach in our setting is that the number of portfolios in our study is small compared to the samples in mutual-fund studies, reducing the effectiveness of a bootstrap.

[28] Below the median for private managers, and below the first quartile for both NYSE and non-NYSE REITs, timing abilities become significantly negative. Studies using parametric approaches in the mutual fund literature, such as Ferson and Schadt (1996), argue that negative timing ability does not make economic sense, because an investor could become a good "timer" by simply trading in the opposite direction of such portfolio managers. Ferson and Schadt (1996) argue that the evidence of negative timing can come from using inaccurate benchmarks, such as ignoring the effect of public information. However, the arguments in this literature are less applicable in the CRE setting, as properties cannot be shorted, and holdings and trades in manager portfolios cannot be readily observed.

upper quartile of non-NYSE-traded REITs appear to have significant timing ability, managers of NYSE REITs and private portfolio managers do not appear to exhibit any evidence of significant timing ability, even in the upper quartile.

Of course, our sample period covers the financial crisis of 2007-2008. While the crisis was driven primarily by housing, CRE markets were also severely affected during this time. CRE markets fell sharply during the crisis, reverting post-crisis. Importantly, smaller markets, typically transacted in by non-NYSE-traded REITs, fell further than larger markets during the crisis, and were slower to recover afterwards. Could the differences between the distribution of timing ability seen for non-NYSE REITs versus NYSE REITs and private managers be attributable to non-NYSE REIT managers' being able to take advantage of the slower recovery in these lagging markets?

To explore this idea, we split our sample into the period before the financial crisis (up to the end of 2007), and during- and after the crisis (from the beginning of 2008 onward).[29] Figure 2 shows differences in the distributions of MT measures and MT-measure t-statistics respectively, before and after the crisis, for the three categories of managers.[30]

While both private managers and NYSE REIT managers exhibit little evidence of positive timing ability both before and after the crisis, for non-NYSE-traded REITs, Figure 2 shows a dramatic upward shift in timing performance after the financial crisis. Comparing across manager types, the distributions of timing measures before the crisis look very similar for all three types of managers, with negative medians and positive performance only occurring well inside the top quartile. After the crisis, private managers and NYSE REITs exhibit only a slight upward shift in the median and an overall shrinkage in the total performance dispersion. Here, positive timing results are still confined to well within the top quartile. In contrast, non-NYSE-traded REITs experience a large shift upwards in mean MT measures, with all but the bottom quartile of managers exhibiting positive timing point estimates. In further (untabulated) analysis we find that, in the pre-crisis period, none of the three groups of managers exhibited a fraction of significant timing performance that exceeds the

[29] In results untabulated for brevity, we test alternative time-period cuts and find that the results are not sensitive to moving the cutoff by several quarters in either direction.

[30] For brevity, we do not include tables; however these are available upon request. Further, in this figure we only include the geographic benchmark; the results for the interacted benchmarks, available upon request, are qualitatively unchanged.

FDR^+. Post-crisis, however, the fraction of portfolios of non-NYSE REITs exhibiting significantly positive performance exceeds the FDR^+ by several percentage points. For private and NYSE REIT managers, no such contrast is visible.

The patterns we observe post-crisis are consistent with the hypothesis that small (i.e. non-NYSE) REITs may have achieved improved timing performance by moving into smaller markets with less institutional presence, which fell further during the crisis and tended to lag larger markets in the recovery from the crisis. This lag may have made the recovery in these markets easier to predict and therefore to time. The limited dispersion for NYSE REITs and private managers, on the other hand, is consistent with the hypothesis that these types of managers largely pursued buy-and-hold strategies after the crisis. This idea is also consistent with a *flight to quality* by large managers in the wake of the crisis, a common view held by practitioners. The overall pattern in the data, however, suggest that outside of such one-time crisis periods, CRE portfolio managers of all stripes exhibit limited evidence of an ability to earn excess returns by timing the markets.

Our timing measure, MT, compares current sub-market returns for each sub-market to returns in the prior year in order to capture the extent to which managers are increasing weights in sub-markets that improve in performance and exit markets which subsequently see a drop in performance. As a robustness check, we also consider an alternate measure of timing ability to ensure we account for the lengthy nature of transactions in the CRE market. MT_{2yr} compares the most recent sub-market return and weight to sub-market returns and weights lagged by two years (rather than one year). In untabulated analysis, we find that this alternate measure yields timing estimates that are lower than those reported in Table 3 by 5 to 10 basis points. The fraction of managers achieving positive timing performance under this alternative measure is also lower, by 1 to 2 percentage points. When we examine the t-statistics for zero-outperformance, we find slightly higher fractions of portfolios that yield statistically significantly positive outperformance (by about two percentage points). Thus, overall, we find that making this alteration yields results that are qualitatively very similar to those generated under our original measure.

It is important to note that poor timing ability need not be solely attributable to managers' inherent (lack of) abilities to read markets and time their movements. Rather, the absence of positive timing returns may be a result of the microstructure of the CRE market, in which property transactions may involve extremely prolonged transaction times. Compared to other asset markets, the commercial property market suffers from slow execution of transactions and very high transaction costs. These frictions may make it difficult to execute a timing strategy, and thus may be an important driver behind the timing results we observe.

2.2 Market Selection

Our second measure captures a portfolio manager's ability to add value cross-sectionally by making particularly good sub-market investment choices within a more broadly defined style category, and holding those sub-market investments for some time. To measure managers' ability to add value along this dimension, we use a *Market Selection* measure, defined as follows:

$$MS_t = \sum_{j=1}^{N} w_{j,t-1}(R_{s_j,t} - R_{Q_j,t}) \tag{2}$$

In this equation, R_{s_j} is the return to the geography and property class sub-market portfolio j, while R_{Q_j} is the return to a more broadly defined style portfolio to which the sub-market belongs. The weight w_j is defined as property square-footage over total portfolio square footage for the properties held by the manager in period t in sub-market j.[31] This measure will be positive if a manager invests in geographic and property class sub-markets that outperform more broadly defined markets within a similar "style."[32]

Returning to the example from the previous section, we would attribute the return from all holdings of Chicago CBD office buildings to the most narrowly defined geographic and class style of

[31] Analogous to the case of our market timing measure, the methodological approach used to construct this sub-market selection measure resembles the characteristic selectivity measure of Daniel et al. (1997), in that here too we employ a measure that constitutes the weighted sum of outperformance versus a benchmark. As before our approach differs economically from that of Daniel et al. (1997), however, in that we compare narrower style choices to more broadly defined ones, as opposed to comparing specific asset choices to asset-pricing characteristics-based benchmarks.

[32] As with the MT measure, the construction of MS should, to a large extent, alleviate concerns arising from smoothed property portfolio returns, as here, too, we are differencing return series.

Chicago CBD Office, and thus R_{s_j} for this sub-market will be the NCREIF NPI return for Chicago CBD Office. The return on this sub-market would then be compared to the return of a larger market of which Chicago CBD Office is a subset, for example, the overall Chicago Office market, which becomes R_{Q_j}. The manager's specific choice of Chicago CBD Office within the overall set of possible Chicago Office properties would have proved to be successful, if the Chicago CBD Office market outperforms the overall Chicago Office market, and this would generate a positive excess return for that property. The weighted average of all portfolio exposures at a particular time constitute overall *Market Selection* performance for that manager in that quarter.

As is the case for our timing measure, given the wide variety of benchmark aggregations available to us, we can construct MS by using the returns to ever larger markets for R_{Q_j}, while preserving R_{s_j} at the smallest possible style definition. Thus, after the first run described above, we would compute a version of MS that compares the performance of the Chicago CBD Office market with the Chicagoland Office market, then the Illinois Office market, then the overall Illinois market, and so forth, once again using all levels of aggregation or specialization, geographically and by class.[33] [34] As in the case of timing, the patterns we uncover are largely invariant to level of geographic aggregation of the benchmark. We therefore tabulate and graph the Divisional level and its interactions with property class as representative.

Table 4, Panel A, presents the distributional statistics of the time-series average *Market Selection* measure for the cross-sections of non-NYSE REIT managers, NYSE REIT managers, and private portfolio managers. As in Table 3 we first compute the MS_t measure described in Equation (2) for

[33]Note that the NPI series disaggregated by CBSA and property subtype simultaneously contains a large amount of missing values, due to insufficient portfolio size. We overcome this limitation by defining R_{s_j} as the return to the CBSA/Type portfolio to which property j belongs. This becomes our most narrowly defined style and we construct MS to compare this style choice to larger encompassing styles.

[34]An alternative to our approach, which evaluates managers' abilities to select sub-markets, would be to evaluate the outperformance of managers' selections of individual properties within sub-markets (akin to Daniel et al. (1997)). The lack of reliable price or return series for the individual properties held by the managers in our data, however, prevents us from conducting such analysis. Given that these properties are not traded while they are held in a manager's portfolio, they are also not priced. For REITs, this limitation is absolute, as these firms also do not disclose appraisals of the properties in their portfolio. While NCREIF members do disclose periodic appraisals, using these would prevent us from conducting a clean comparison between private and REIT managers. However, as argued previously, the choice of CBSA and property type constitutes the defining characteristics of a commercial property investment and so a portfolio that reflects these choices should capture the concept of narrowly defined choice of style, which is the goal of our analysis. Furthermore, Crane and Hartzell (2007) compare the use of a CBSA/Type portfolio from the NPI to using direct returns for a particular property where available, and find that the returns derived through the index have a correlation of more than 0.96 with the actual property returns.

every manager in every quarter. We then compute a time-series average for each manager over the quarters for which the manager is active in the sample, to obtain a single average MS statistic per manager. We illustrate these statistics in Figure 3.

For all three types of managers, both the mean and median MS measure is very close to zero. In contrast to what we observed for the MT (timing ability) measure, where very few managers exhibited positive point estimates, here approximately half the managers in our sample have negative *Market Selection* measures, and the other half have positive *Market Selection* measures. Furthermore, it is apparent from the table and figure that sub-market selection ability shows more dispersion with respect to the geographic-only benchmark than with respect to benchmarks disaggregated by both geography and property class. Similar patterns can be seen by examining the median and the inter-quartile spread; the first panel of Figure 3 confirms this pattern. This is true whether we employ the benchmarks at the regional, divisional, or state level, and is consistent across all three manager types. This pattern suggests that selection of the right property class within geographic subdivisions is an important part of managerial value added, and that this component of ability contains a high degree of heterogeneity.

As we did for the MT measure, we next compute t-tests for the hypothesis that a manager has zero selection ability against the two-sided alternative. The distributions of the resulting t-statistics are reported in Table 4, Panel B, and the second panel of Figure 3. All statistics shown (FDR^+) are analogous to Table 3.

From the figure and table, it is apparent that the entire range of manager MS measures that fall between the first and the third quartile are statistically indistinguishable from zero at the 5% significance level. That said, a t-statistic of around $+1.96$ is still well within the range covered by the whiskers (the same can be said about a value of -1.96). Put differently, one does not have to look for extreme positive outliers to find significant outperformance (or underperformance) along the selection dimension. Comparing the fraction of managers that shows significant outperformance to the respective FDR^+, we see that for all three types of managers, this fraction well exceeds the amount of outperformance that would be expected to be observed by mere chance. While a substantial fraction of all managers appear to have positive selection ability, this fraction is considerably larger for private managers, at approximately 12% of managers, about double the percentage observed for both types of REIT managers (approximately 6%).

Overall, these results suggest considerable heterogeneity among managers in sub-market selection ability. An appreciable fraction of managers of all types appear to have significant ability to create value through property class selection. For timing, in contrast, most managers exhibit negative performance, with only the extreme positive outliers showing significantly positive value-added.

As with timing, we also compare performance before and after the crisis for MS. The results are presented in Figure 4. While non-NYSE REITS exhibit significantly different timing measures pre- and post- crisis, for selection, generally speaking, the differences are less stark. Overall, for all three types of managers, selection ability post-crisis appears to be slightly lower than pre-crisis. That said, for all three types of managers, a significant fraction of the upper quartile exhibit significantly positive selection ability both pre- and post-crisis. For all three groups of managers, the fraction of portfolios with significantly positive selection exceeds the respective FDR^{+}, both before and after the crisis, suggesting that a fraction of all three types of managers exhibit significant positive selection ability in both time periods.

2.3 Correlation between Timing and Selection Ability

Next, we compute correlations between timing and selection ability. We calculate correlation coefficients over the entire panel dataset (by manager and quarter). Specifically, for manager m, and MT or MS performance over benchmark i at time t, we calculate the correlation of all $MT_{m,i,t}$ with $MS_{m,i,t}$, together with t-statistics for the hypothesis that the true correlation is zero, against the two-sided alternative. Manager MS and MT at the quarterly level are significantly negatively correlated, with coefficients generally around $-.1$. In other words, managers who exhibit the ability to generate returns through market selection are still performing somewhat poorly on average at timing their property trades, or may even be selecting *at the expense* of timing performance. This negative correlation would also be consistent with managers choosing to invest in long-term selection ability, rather than attempts to time the market, when operating in sub-markets where properties are more difficult to trade (or trading times more protracted). We explore this liquidity hypothesis in more detail in Section 2.7.[35]

2.4 Time-Series Persistence of Selection Ability

If some managers display selection ability, a natural question to ask is whether such abilities are persistent. To answer this question, we examine time-series persistence of the *Selection* measures by manager, examining rank persistence and *permanence* of managers in the top decile or quartile. We then examine the forward returns to investing in the portfolios of managers who ranked in the top decile on selection ability in the past.[36]

We begin by examining the persistence of manager relative rankings in selection ability over time. For each year we construct a percentile rank for each manager, based on his or her realized performance with respect to *Market Selection* over the past year. We then compute autocorrelations of manager percentile ranks over one-, two-, and three-year horizons. Table 5 presents these estimates.

The Table shows that, across the board, for all three types of managers, autocorrelations in rank are significantly positive and economically large at both the one- and two-year horizons (24% to 41% at one year and 9% to 20% at two years, depending on manager type and benchmark specification). At a three-year horizon, we do not find evidence of rank persistence. The bottom part of the table presents the fraction of managers in the top quartile (decile) in year t that remain in the top quartile - 75*plus* (decile - 95*plus*) in years $t+1$, $t+2$, $t+3$. If there is no rank persistence, and managers randomly appear in top quartile (decile), we would expect this fraction to be 0.25 (0.1) assuming all managers survive. For all three types of managers, there is widespread evidence of top quartile and decile permanence at both the one and two year horizons. Between 37% and 51% of top quartile managers remain in the top quartile in the following year, and between 24% and 39% of top decile

[35]Henriksson (1984) argues (in a parametric setting, for mutual funds) that negative correlation between timing and selection may be a sign of mis-specified benchmarks. However, other studies such as Jagannathan and Korajczyk (1986) refute this argument and show that such an outcome is possible even with correctly specified benchmarks.

[36]We separately compute the same measures for persistence of the timing ability measures. However, since timing performance is generally low, these results do not carry much salience and so we do not report them in the paper. The results, however, are available from the authors upon request.

managers remain in the top decile in the following year, depending on manager type. Similar patterns are present at the two year horizon, though these patterns weaken at the three year horizon. Thus, in this dimension, past positive selection performance may give some degree of confidence of such performance in the future, making asset allocation to outperforming managers feasible.

In this vein, Table 6 illustrates the returns to a trading strategy that allocates capital in year $t+1$ to all portfolios ranked in the top decile of MS for year t. For each year, ending at time t, we rank managers according to annualized MS performance over the previous year, with respect to each benchmark level i and within their entity type p (i.e. REIT versus NYSE REIT versus private). We then simulate (for each i and p) the returns to investing equal amounts of money at time t into each portfolio that ranked in the top decile in terms of MS performance in the previous year for the same benchmark and entity type, and holding those investments until the end of year $t+1$. The table presents time-series averages of equal-weighted cross-sectional mean MS returns for each year, obtained per manager over the year, with respect to benchmark i.

The table shows that investing in year $t+1$ in the portfolios of private managers who ranked in the top decile on selection ability in year t returns a statistically significant *positive* return at all benchmark levels, with positive returns ranging from 0.53% per year to 2.63% per year depending on the level of benchmark aggregation. For NYSE REITs, forward investment returns to this strategy are significantly positive at nearly every benchmark level, ranging from 0.41% per year to 2.7% per year depending on benchmark aggregation level. Returns are generally higher when employing a geographic-only benchmark, which is consistent with our earlier findings. More generally, across our various analyses, more value seems to be created by selecting property type than by just selecting geography. This basic pattern is also apparent for non-NYSE managers, who exhibit significantly positive forward investment returns only when measuring against a geographic-only benchmark.

2.5 Appraisal-Effects on NPI-Based Measures

As the NPI indices are constructed partially using appraisal data (which, as mentioned earlier, has well-known limitations), we conduct several tests to ensure that our results are not driven by potential smoothing effects present in appraisal data.

First, in untabulated results, we employ the desmoothing methodology developed in Cho, Kawaguchi and Shilling (2003) as adapted for panel datasets by Mühlhofer (2015) to construct desmoothed return series. We then rerun our analyses using these desmoothed indices. We obtain similar distributions for the desmoothed MT and MS measures, and the persistence results remain largely unchanged. After desmoothing, the one-year forward returns from investment in the top decile of managers are amplified: the returns from investment in the top decile of managers by MS are higher than before desmoothing, and the returns from investment in the top decile of managers by MT are more negative. This suggests that the empirical patterns uncovered by our analysis are not driven by smoothing present in appraisal data.

As a further robustness test, we re-run parts of our analysis using the transaction-based Commercial Property Price Indices (CPPI) from Real Capital Analytics (RCA) discussed in Section 1 as benchmarks instead of NPI. Since RCA does not have coverage at the same level of disaggregation that NCREIF does, we have to slightly alter our analysis. For MT we construct measures in the same way, but can only compute benchmarks for National/Type, Regional, and Regional/Type. For MS we construct measures in such a way as to measure the specific investment return, $R_{s_j,t}$ in Equation 2 (page 18) as the return to the Region/Type market in which the property is located. We therefore are only able to assess the impact which the choice of Region/Type (i.e., for example, Midwest Office) makes in comparison to a larger benchmark aggregation (i.e., for example Midwest Commercial Property, National-Office, and National Commercial Property). We thus find it most intuitive to report tests for statistical differences in these modified measures constructed with RCA benchmarks and the same modified measures constructed with NCREIF benchmarks.[37]

Table 7 shows the mean obtained with RCA, the mean obtained with NCREIF's NPI, and a t-statistic testing the hypothesis that the two are the same, for each entity type, measure type and benchmark level. We observe little to no statistical difference for REIT manager selection ability, except for a very small reduction in selection performance for non-NYSE REITs at the National

[37]Similar tabulations for the MT measure are available from the authors upon request.

level.

In further untablulated analysis, we also test whether the top-decile forward investment returns (Table 6) differ statistically between the two different data series used for constructing benchmarks. Here, we find no statistical difference between the results generated with the two data series. This test suffers from low power, however, as we measure annual returns and RCA's data limits allows us only to begin testing in 2001.

Overall, these tests suggest that our results are not an artifact of appraisal smoothing present in NCREIF's NPI data. The much larger scope of the NCREIF's data compared to RCA's warrants the choice of the NCREIF benchmarks to conduct our primary analysis.

2.6 REIT Stock Returns

Since REITs are publicly traded and have traded prices, it should be the case that the stock market recognizes and rewards "good" REIT managers and penalizes "bad" REIT managers. To test this idea, we map the performance of the portfolio in direct real estate markets to stock market performance of the REITs. In other words, we ask whether the stock market recognizes when REIT managers outperform (on either timing or selection) and rewards this through higher returns. We thus conduct an exercise akin to the one described in Table 6, but on REIT stock returns. More specifically, in year $t-1$, we rank managers according to their selection performance with respect to a particular benchmark. We simulate investing into an equal-weighted portfolio of all managers ranked in the top decile according to that benchmark and holding that portfolio for year t. We then estimate a factor model regression for each such portfolio with excess returns to the portfolio on the left-hand side and the set of four REIT factors of Hartzell et al. (2010) on the right.[38] [39] We then ask whether this strategy generates positive *alpha*. As we do throughout our analysis, we separately analyze REITs traded on the NYSE and REITs traded on the other exchanges.

[38]Hartzell et al. (2010) construct a set of four factors to benchmark REIT portfolio returns, which are methodologically akin to the Carhart (1997) four-factor model, but for which all factors are constructed using REITs (which are excluded in the construction of the common asset-pricing factors). Hartzell et al. (2010) show that these factors are better suited to benchmark REIT returns than the common asset pricing factors.

[39]Given that our portfolios are equal-weighted, it is especially important to use a four factor model in order to assure that the strategy we propose does not just load on size-, book-to-market, or momentum risk.

We leave the estimates untabulated for brevity (available upon request). For NYSE REITs we observe significant *alphas* when portfolios are constructed based on MS measured for the smaller, geographic-only benchmarks, consistent with the evidence on selection presented earlier in this study. For non-NYSE REITs, while we find consistently positive point estimates for *alpha*, none of these are statistically significant. This difference could possibly be attributed to closer analyst following of NYSE REITs, which might help these stocks reflect information on outperformance more unambiguously and with less noise.

2.7 Liquidity and Timing- and Selection-Performance

In our final set of analysis we test the hypothesis that underlying market liquidity may affect the ability to successfully time the market. A market timing strategy requires transactions that are frequent (and relatively fast). Conversely, a market-selection strategy, which is more passive, would not be as dependent on underlying market liquidity. Thus, we would expect to see that managers that are trading in more liquid markets are more likely to be those pursuing timing strategies and/or exhibiting better results at timing. In contrast, managers who are trading in less liquid markets may be more likely to pursue selection-oriented strategies and may perform relatively better at selection. In fact, this separation may become endogenous, with managers whose talents lie in timing properties choosing to trade in more liquid markets, which in equilibrium likely would require a premium for such liquidity, and managers better at selection avoiding such markets and therefore such premia. The relatively low overall liquidity of commercial property markets could then account for why timing performance is worse overall than selection performance.

To empirically test this hypothesis, we begin by constructing a liquidity measure for each property sub-market. Our measure is based on the illiquidity measure of Amihud (2002), as modified for commercial property markets by Mühlhofer (2015). Specifically, we define illiquidity for sub-market s in quarter t as:

$$illiq_{s,t} = \frac{|R_{s,t}|}{frac.volume_{s,t}} \tag{3}$$

Here, $R_{s,t}$ is the total return to sub-market s in quarter t, and $frac.volume_{s,t}$ is defined as square footage bought into the NCREIF portfolio plus square footage sold out of the NCREIF portfolio, divided by total sub-market square footage. The measure assesses price impact of trading. If the value is high, there are large price changes caused by relatively small trades, which implies an illiquid market; if the measure is low, the opposite is the case. For each manager i in each quarter t, we then compute a *manager illiquidity* measure defined as:

$$mgr.illiq_{i,t} = \sum_{s=1}^{S} w_{i,s,t} illiq_{s,t} \tag{4}$$

which is the weighted average illiquidity across manager i's portfolio in quarter t, weighted by the manager's fractional exposure to each sub-market.

Our illiquidity measure by construction cannot be negative, and its realizations are log-normally distributed. Thus, taking the logarithm of the measure makes the variable normally distributed and gives it better statistical properties. We thus use the log of the illiquidity measure as a primary explanatory variable in the following panel-regression models:

$$MT_{i,t} = \alpha + \beta_1 private_i + \beta_2 log(mgr.illiq_{i,t}) + \vec{\beta_3}' \vec{controls}_{i,t} + \epsilon_{i,t} \tag{5}$$

$$MS_{i,t} = \alpha + \beta_1 private_i + \beta_2 log(mgr.illiq_{i,t}) + \vec{\beta_3}' \vec{controls}_{i,t} + \epsilon_{i,t} \tag{6}$$

The dependent variables in the regressions are the current MT measure and the current MS measure, respectively, for each manager in each quarter. In addition to our illiquidity measure, we also include two indicator variables. These are *private*, which is equal to one for a manager of a private portfolio, and zero otherwise, and *nyse*, defined analogously for managers of REITs traded on the NYSE, leaving the omitted category as non-NYSE REITs. We also include a set of portfolio characteristics as control variables. The set of controls consists of $log(size_{i,t})$, which is the natural logarithm of manager i's average portfolio size in square feet between quarter $t-3$ and t, and $geo.spec_{i,t}$ and $type.spec_{i,t}$, which measure the level of geographic and property type specialization of a manager's portfolio, respectively, and which are computed as Hirschman-Herfindahl Indices

$$H_{i,t} = \sum_{s=1}^{N} w_{s,t}^2 \tag{7}$$

In line with the other independent variables, $geo.spec_{i,t}$ and $type.spec_{i,t}$ are defined as moving averages from quarter $t-3$ to quarter t of the respective series of $H_{i,t}$. The final control variable is average property size for each manager at time t. This controls for the possibility that managers dealing in larger properties may need to be active in more liquid markets in order to raise the likelihood of a timely exit from an investment. With the exception of the *private* and *nyse* dummies, all independent variables are taken as one-year moving averages or trailing sums of each of the time series employed, to allow for slow changes in property markets and portfolio characteristics.

Market liquidity should have the most significant impact on a disaggregated, local level.[40] We therefore focus these tests on the smallest level of benchmark aggregation at which we have consistent data, namely the State level, and its interactions with property class.

It is apparent in the data that lower-liquidity secondary markets took longer to recover than higher-liquidity primary markets in the wake of the financial crisis, which made the former easier to time. It is this collinear phenomenon and not inherently the markets' lack of liquidity which determines timing performance in this time period. In addition, there is a general consensus among practitioners of a *flight to quality* during the financial crisis, which led to larger institutional investors' selling out of smaller (and less liquid) markets, thus causing these markets to fall further than larger markets. This can again be seen as a deviation from the usual dynamics of this market. To rule out the possibility of the recovery period biasing our conclusions, we conduct our estimation solely for the pre-crisis period from 1995 to 2007.

We estimate two versions of each panel regression set. The first includes time fixed effects, and clusters robust heteroskedasticity-consistent standard errors by manager. As Petersen (2009) suggests that a panel setting with group effects may lead to inefficient OLS estimates, we also estimate each

[40] For example, if trying to sell an office property specifically in San Diego, the necessary market liquidity will be much more difficult to find, than if trying to sell an office property anywhere on the West Coast.

set of panel regressions with random effects, using feasible GLS.

Table 8 shows the results for these tests. For brevity, we do not report the coefficients on the constant term or control variables, though these are included in all models. The statistics reported below each set of Random-Effects coefficients are R^2_{TOT} (a pseudo R-square for the entire model), R^2_{FE} (a pseudo R-square for the fixed-effects portion only, excluding the variation explained by the random effects), χ^2_{MOD} (a model χ^2 statistic, testing the joint significance of all coefficients), and χ^2_{BP} (the result of the Breusch-Pagan test, of the null hypothesis that no random effects exist in the data). Note that χ^2_{BP} strongly rejects for all specifications, suggesting that a random-effects model is likely to be appropriate in this setting.

The negative and significant coefficients on *log.mgr.illiq* evident in Table 8, Panel A overwhelmingly suggest that managers that are active in less illiquid (i.e. more liquid) markets are also those that exhibit better market timing performance. These results hold broadly across all three benchmark levels, whether geographic only or geography-class interactions. We also observe a consistently negative and significant coefficient on the *private* dummy in our models. Holding other portfolio characteristics constant, private portfolio managers thus appear to obtain timing profits that, on average, are 26 to 40 basis points per quarter lower than those achieved by REIT managers. A possible explanation for this finding, which is explored in more detail in Hochberg and Mühlhofer (2014), is that the governance and compensation structures for private and REIT managers differ significantly. This seems to be associated with varying degrees of capital-market competitiveness, which better incentivizes performance in REIT managers. For NYSE managers the point estimates are negative, but the coefficients are not significant. When we compare the estimates from the Fixed-Effects model with the estimates obtained from the Random-Effects models, we find coefficients of approximately equal value in both specifications, though the random-effects models exhibit slightly higher statistical significance levels for the variables of interest. Thus, our results appear to be fairly robust to different model specifications. The R^2 values we obtain are fairly high (.2 to .5 range).[41]

Table 8, Panel B, reports the estimates from our panel regressions on market selection perfor-

[41] Note that in the Random-Effects model the R^2_{FE} is low, and so a large part of timing performance seems to be associated with individual manager characteristics or time-periods (such as large market turnarounds which offer the possibility to time successfully).

mance. We find positive coefficient estimates for *log.mgr.illiq* at the geographic-only benchmark level, suggesting that managers that invest in less liquid markets also have a tendency to select better sub-markets. The observed pattern is consistent with previous results on selection, which were stronger and showed wider dispersion at geographic-only levels, where part of the manager's selection task consists of choosing the property class to invest into. The *private* and *nyse* dummies are primarily statistically insignificant. The values for R^2 for the selection models are generally much lower than for timing, suggesting a greater prevalence of idiosyncratic effects along the lines of unobserved manager skill or characteristics.[42]

Overall, our analysis suggests that managers that tend to invest in more liquid markets have a tendency to exhibit relatively better timing, while managers that tend to invest in less liquid markets have a tendency to exhibit relatively better selection. This choice is likely endogenous, in that managers with naturally better timing abilities should tend to invest in markets that allow them to capitalize on these.[43] However, since the effective timing profits that are realized tend to be quite low (as shown in the univariate statistics), the better investment choice for investors considering investing with a CRE portfolio manager may be to be to invest with managers that exhibit good selection performance.

3 Conclusion

The ability of active money management to generate abnormal returns that justify their fees has long been a subject of academic debate. In this paper, we examine the ability of portfolio managers to generate abnormal profits in the CRE market through the timing of entry and exit into real estate sub-markets and through selection of outperforming sub-markets versus broader style-matched benchmarks. Like many alternative asset classes, the CRE market, which is characterized by trans-

[42]A natural follow-on question is how individual manager characteristics–such as experience, background, education, etc.– relate to the manager's ability to either time property purchases and sales or to select individual property subclasses of investment versus style-matched similar investments. Unfortunately, the limitations of the data available to us prevents us from exploring cross-sectional relationships of this nature. We therefore leave this question to future research.

[43]The documented negative correlations between timing and selection are also consistent with the argument presented here.

actions that primarily occur in relatively illiquid and opaque private asset markets, may provide greater opportunities for managerial skill and informational advantages to add value and to lead to abnormal profits.

Our analysis suggests that both large NYSE-traded REITs and private portfolio managers exhibit little or negative ability to successfully time the market, on average. A substantial proportion of managers, both private and REIT, however, exhibit some level of positive selection ability. Selection ability appears to be somewhat persistent, allowing for the possibility of efficient portfolio allocation to portfolio managers with selection ability. Managers that are active in more liquid markets tend to exhibit better market timing performance, while managers exhibiting better selection ability appear to be active in less liquid markets. Our findings are consistent with the commonly held practitioner notion that in commercial property markets, where transactions are costly and may involve prolonged transacting periods, managers may primarily create value for their investors through selection of geographic and property-type sub-markets versus a broader passive portfolio, rather than through precise timing of sub-market entry and exit.

To the best of our knowledge, our study is the first to examine issues of market timing and sub-market selection ability in the CRE market. As such, it potentially may provide insights for investors, portfolio managers and academics into how managers may earn abnormal profits in these markets. Overall, our results suggest that in the CRE market, managerial value added, where it exists, appears to come primarily in the form of investment selection and investors should search for managerial value-added along this dimension.

References

Amihud, Y.: 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* **5**, 31–56.

Barkham, R. and Geltner, D.: 1995, Price discovery in american and british property markets, *Real Estate Economics* **23**(1), 21–44.

Barras, L., Scaillet, O. and Wermers, R.: 2010, False discoveries in mutual fund performance: Measuring luck in estimated alphas, *The Journal of Finance* **65**(1), 179–216.

Bond, S. A. and Mitchell, P.: 2010, Alpha and persistence in real estate fund performance, *Journal of Real Estate Finance and Economics* **41**, 53–79.

Boudry, W. I., Coulson, N. E., Kallberg, J. G. and Liu, C. H.: 2013, On indexing commercial real estate properties and portfolios, *The Journal of Real Estate Finance and Economics* **47**(4), 617–639.

Brown, S. J. and Goetzmann, W. N.: 1995, Performance persistence, *Journal of Finance* **50**, 679–698.

Carhart, M. M.: 1997, On persistence in mutual fund performance, *Journal of Finance* **52**, 57–82.

Case, B. and Quigley, J.: 1991, Dynamics of real-estate prices, *Review of Economics and Statistics* **73**(1), 50–58.

Case, K. and Shiller, R.: 1990, Forecasting prices and excess returns in the housing market, *AREUEA Journal* **18**(3), 253–273.

Chen, Y., Ferson, W. and Peters, H.: 2010, Measuring the timing ability and performance of bond mutual funds, *Journal of Financial Economics* **98**(1), 72–89.

Cho, H., Kawaguchi, Y. and Shilling, J. D.: 2003, Unsmoothing commercial property returns: A revision to fisher-geltner-webb's unsmoothing procedure, *Journal of Real Estate Finance and Economics* **27**(3), 393–405.

Clayton, J., Geltner, D. and Hamilton, S.: 2001, Smoothing in commercial property valuations: Evidence from individual appraisals, *Real Estate Economics* **29**(3).

Crane, A. D. and Hartzell, J. C.: 2007, Is there a disposition effect in corporate investment decisions? evidence from real estate investment trusts, *SSRN eLibrary, http://ssrn.com/paper=1031010* .

Daniel, K., Grinblatt, M., Titman, S. and Wermers, R.: 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* **52**(3), 1035–1058.

Fama, E. and French, K.: 2010, Luck versus skill in the cross-section of mutual fund returns, *The Journal of Finance* **65**(5), 1915–1947.

Ferson, W. and Schadt, R.: 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* pp. 425–461.

Florance, A., Miller, N., Peng, R. and Spivey, J.: 2010, Slicing, dicing, and scoping the size of the us commercial real estate market, *Journal of Real Estate Portfolio Management* **16**(2), 101–118.

Fung, W. and Hsieh, D. A.: 2002, Asset-based style factors for hedge funds, *Financial Analysts Journal* **58**(5), pp. 16–27.
URL: *http://www.jstor.org/stable/4480414*

Fung, W. and Hsieh, D. A.: 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* **60**(5), pp. 65–80.
URL: *http://www.jstor.org/stable/4480604*

Geltner, D.: 1991, Smoothing in appraisal-based returns, *Journal of Real Estate Finance and Economics* **4**(3), 327–345.

Geltner, D. and Mei, J.: 1995, The present value model with time-varying discount rates: Implications for commercial property valuation and investment decisions, *Journal of Real Estate Finance and Economics* **11**(2), 119–135.

Gruber, M. J.: 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* **51**, 783–810.

Hartzell, J., Mühlhofer, T. and Titman, S.: 2010, Alternative benchmarks for evaluating mutual fund performance, *Real Estate Economics* **38**(1), 121–154.

Henriksson, R.: 1984, Market timing and mutual fund performance: An empirical investigation, *Journal of Business* pp. 73–96.

Henriksson, R. and Merton, R.: 1981, On market timing and investment performance. ii. statistical procedures for evaluating forecasting skills, *Journal of Business* pp. 513–533.

Hess, R. and Liang, Y.: 2000, Regional investment preferences: Ncreif investors vs. reits. Prudential Real Estate Investors Research Notes.

Hochberg, Y. V. and Mühlhofer, T.: 2014, Capital-market competitiveness and managerial investment decisions:
evidence from commercial real estate. Working paper, available at: http://tobias.muhlhofer.com.

Jagannathan, R. and Korajczyk, R. A.: 1986, Assessing the market timing performance of managed portfolios, *The Journal of Business* **59**(2), 217–35.
URL: *http://ideas.repec.org/a/ucp/jnlbus/v59y1986i2p217-35.html*

Jensen, M.: 1968, The performance of mutual funds in the period 1945-1964, *Journal of Finance* **23**, 389–416.

Jensen, M.: 1969, Risk, the pricing of capital assets, and the evaluation of investment portfolios, *Journal of Business* **42**, 167–247.

Jiang, G., Yao, T. and Yu, T.: 2007, Do mutual funds time the market? evidence from portfolio holdings, *Journal of Financial Economics* **86**(3), 724–758.

Kacperczyk, M., Sialm, C. and Zheng, L.: 2005, On the industry concentration of actively managed equity mutual funds, *The Journal of Finance* **60**(4), 1983–2011.

Kallberg, J. G., Liu, C. L. and Trzcinka, C.: 2000, The value added from investment managers: An examination of funds of reits, *Journal of Financial and Quantitative Analysis* **35**, 387–408.

Kaplan, S. N. and Schoar, A.: 2005, Private equity performance: Returns persistence and capital, *Journal of Finance* **60**, 1791–1823.

Kosowski, R., Timmermann, A., Wermers, R. and White, H.: 2006, Can mutual fund stars really pick stocks? new evidence from a bootstrap analysis, *The Journal of finance* **61**(6), 2551–2595.

Liu, C. H. and Mei, J.: 1992, The predictability of returns on equity reits and their comovement with other assets, *Journal of Real Estate Finance and Economics* **5**(4), 401–418.

Liu, C. H. and Mei, J.: 1994, An analysis of real estate risk using the present value model, *Journal of Real Estate Finance and Economics* **8**, 5–20.

Mühlhofer, T.: 2014, Why do reit returns poorly reflect property returns? unrealizable appreciation gains due to trading constraints as the solution to the short-term disparity, *Real Estate Economics* **41**, 814–857.

Mühlhofer, T.: 2015, They would if they could: Assessing the bindingness of the property holding constraint for reits., *Real Estate Economics* **Forthcoming**.

Petersen, M. A.: 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of financial studies* **22**(1), 435–480.

Plazzi, A., Torous, W. and Valkanov, R.: 2010, Expected returns and expected growth in rents of commercial real estate, *Review of Financial Studies* **23**(9), 3469.

R Development Core Team: 2008, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL: *http://www.R-project.org*

Storey, J.: 2002, A direct approach to false discovery rates, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(3), 479–498.

Titman, S. and Tiu, C.: 2011, Do the best hedge funds hedge?, *Review of Financial Studies* **24**(1), 123–168.

Treynor, J. and Mazuy, K.: 1966, Can mutual funds outguess the market, *Harvard Business Review* **44**(4), 131–136.

Wermers, R.: 2003, Is money really 'smart'? new evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working Paper, University of Maryland.

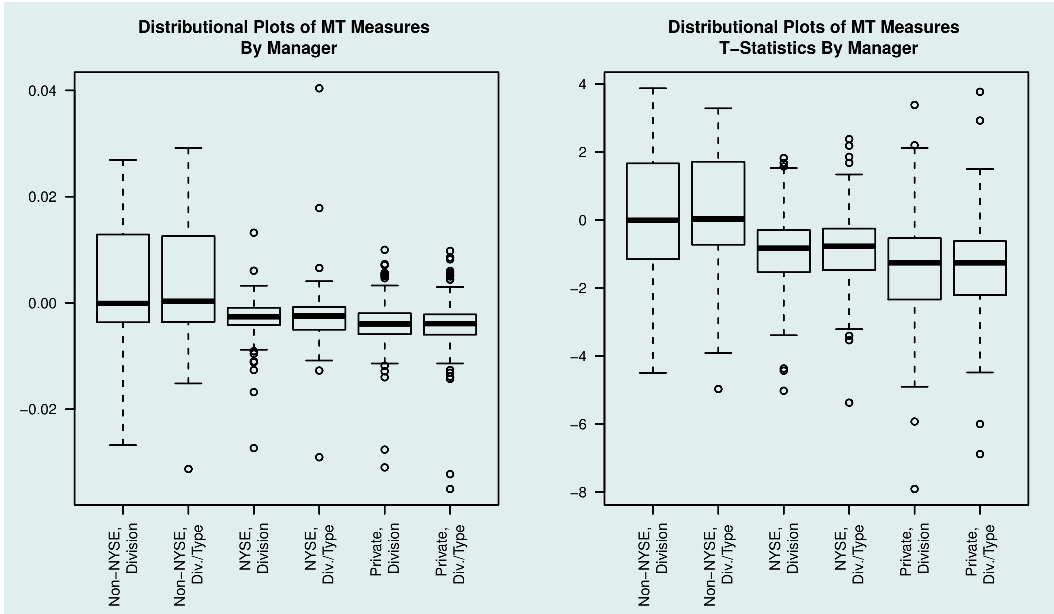

Figure 1: Timing Measures.

This figure shows box-and-whisker diagrams comparing the cross-sectional distributions of MT measures obtained by managers with respect to the Divisional and Divisional/Type benchmarks. The data presented consists of Non-NYSE REIT managers, NYSE-REIT managers, and private managers. The left figure shows cross-sectional distributions of managers' time-series means. The right figures shows cross-sectional distributions of the t-statistic that a manager's lifetime MT-measure is different from zero.

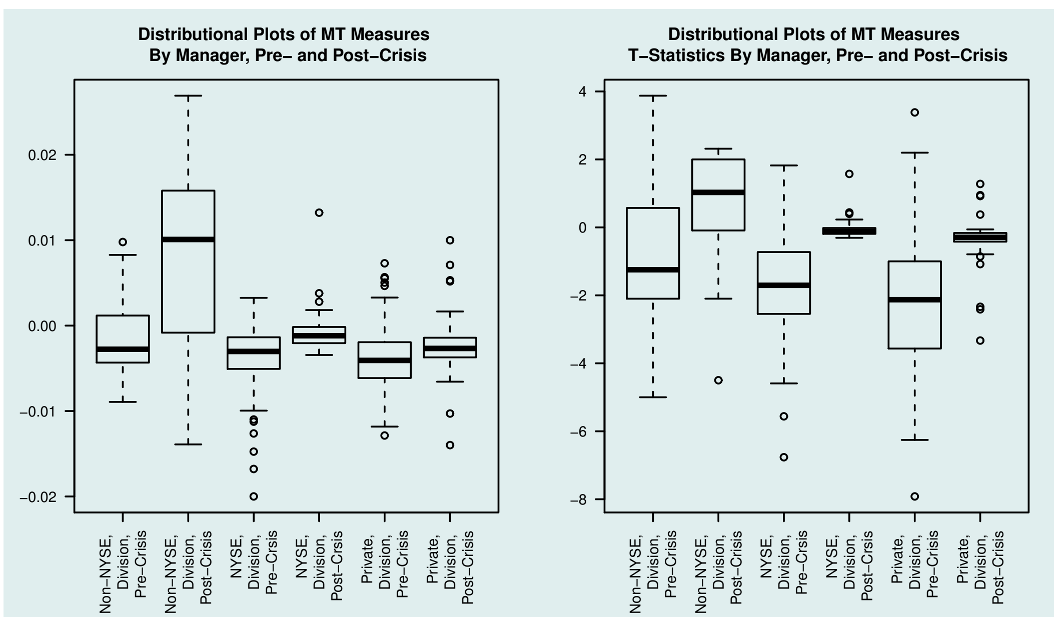

Figure 2: Timing Measures, Pre- and Post Crisis

This figure shows box-and-whisker diagrams comparing the cross-sectional distributions of MT measures obtained by managers with respect to the Divisional benchmark, and compares these before and after the financial crisis of 2008. The data presented consists of Non-NYSE REIT managers, NYSE-REIT managers, and private managers. The left figure shows cross-sectional distributions of managers' time-series means. The right figures shows cross-sectional distributions of the t-statistic that a manager's lifetime MT-measure is different from zero.

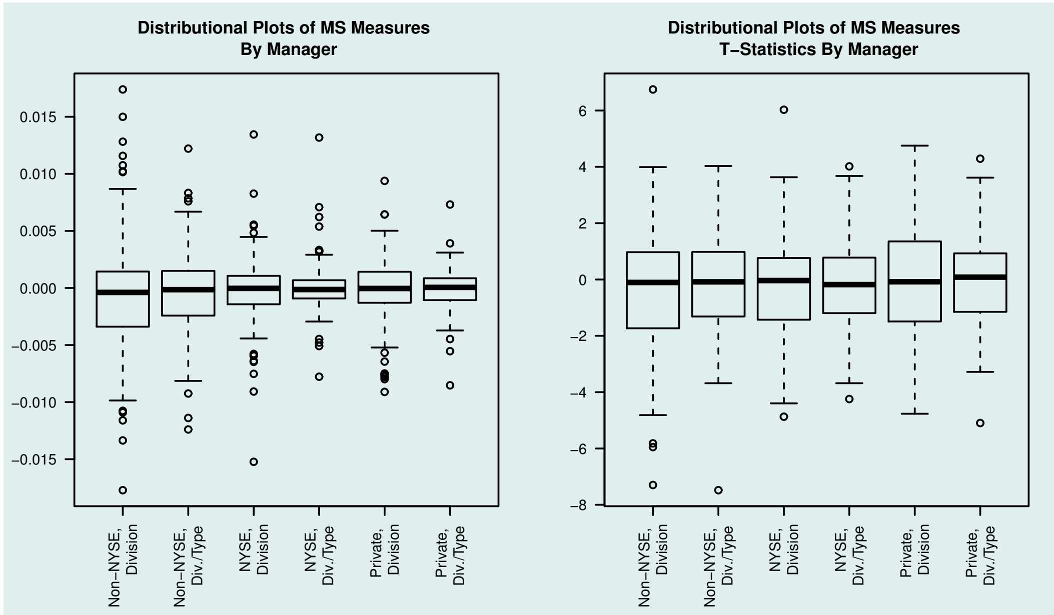

Figure 3: Selection Measures

This figure shows box-and-whisker diagrams comparing the cross-sectional distributions of MS measures obtained by managers with respect to the Divisional and Divisional/Type benchmarks. The data presented consists of Non-NYSE REIT managers, NYSE-REIT managers, and private managers. The left figure shows cross-sectional distributions of managers' time-series means. The right figures shows cross-sectional distributions of the t-statistic that a manager's lifetime MT-measure is different from zero.

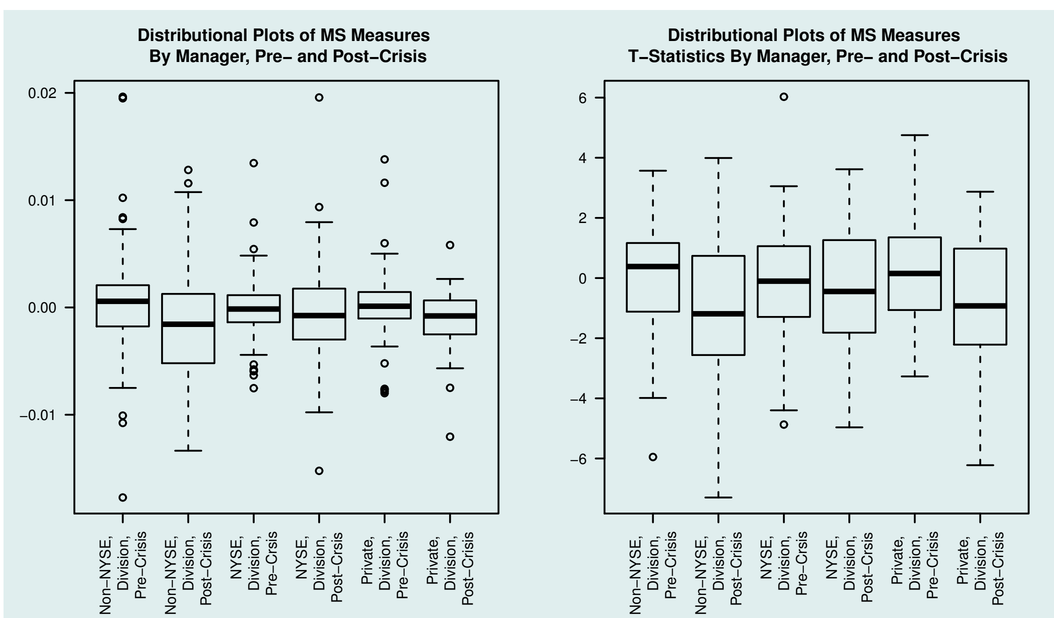

Figure 4: Selection Measures Pre- and Post Crisis

This figure shows box-and-whisker diagrams comparing the cross-sectional distributions of MS measures obtained by managers with respect to the Divisional benchmark, and compares these before and after the financial crisis of 2008. The data presented consists of Non-NYSE REIT managers, NYSE-REIT managers, and private managers. The left figure shows cross-sectional distributions of managers' time-series means. The right figures shows cross-sectional distributions of the t-statistic that a manager's lifetime MT-measure is different from zero.

Table 1: Subdivisions by Geography and Property Type

This table presents the numbers of properties held by private investors, REITs that are traded on the NYSE, and REITs that are not traded on the NYSE (i.e. that are traded on Amex or Nasdaq). Property counts are organized by NCREIF Type, Subtype, Region, and Division. NCREIF also offers organizations by state and CBSA, which we do not present here.

		Non-NYSE REITs	**NYSE REITs**	**Private**
Type and Subtype				
Apartment		173	2426	4574
	Garden			2889
	High-rise			1119
	Low-rise			522
Hotel		19	9	449
Industrial		524	3492	8218
	Warehouse	221	1068	6556
	R&D	11	142	790
	Flex Space			941
	Manufacturing	30	124	53
	Showroom			31
	Other	5	76	159
Office		916	4080	5539
	CBD	785	3834	1222
	Suburban	131	255	4373
Retail		2535	3048	3701
	Community			1206
	Theme/Festival			13
	Fashion/Specialty			109
	Neighborhood			1394
	Outlet	38	51	11
	Power Center	47	65	319
	Regional	78	183	383
	Super Regional			236
	Single Tenant	134	480	347
Regions and Divisions				
East		1138	3598	5083
	Mideast	518	1691	2493
	Northeast	620	1907	2591
Midwest		753	2419	3411
	East North Central	498	1844	2505
	West North Central	255	575	907
South		1464	3770	6452
	Southeast	771	2281	3451
	Southwest	693	1489	3005
West		804	3238	7406
	Mountain	282	877	1759
	Pacific	522	2361	5647

Table 2: Summary Statistics

This table presents summary statistics for the sets of properties held by both private investors and Real Estate Investment Trusts.

	Mean	**Std. Dev.**	**1st Quartile**	**Median**	**3rd Quartile**
Non-NYSE REITs					
Property Sizes (1000 Sq. ft.)	193.61	260.82	52.81	111.55	226.71
Portfolio Sizes (1000 Sq. ft.)	17,064	20,804	2,706	7,860	23,892
Portfolio Presence in Number of CBSAs	9.25	12.86	1	4	11
Number of Property Types in Portfolios	1.67	0.89	1	1	2
Number of Property Subtypes in Portfolios	2.2	1.62	1	2	3
Property Holding Periods (years)	5.51	3.67	3.37	4.49	6.82
Manager Number of Years in Sample	5.56	4.16	2.75	4.75	6.69
Manager HHI, by Region	0.72	0.26	0.47	0.74	1
Manager HHI, by Division	0.66	0.3	0.39	0.68	1
Manager HHI, by State	0.61	0.32	0.32	0.57	1
Manager HHI, by CBSA	0.57	0.33	0.29	0.51	1
Manager HHI, by Type	0.89	0.17	0.83	1	1
Manager HHI, by Subtype	0.84	0.2	0.73	0.99	1
Number of Properties: 4,159		Number of Managers: 166			
NYSE REITs					
Property Sizes (1000 Sq. ft.)	189.22	234.92	56.58	119.5	237.13
Portfolio Sizes (1000 Sq. ft.)	24,804	27,135	7,889	17,372	28,496
Portfolio Presence in Number of CBSAs	16.75	15.4	4.25	13	23.5
Number of Property Types in Portfolios	2.06	1.08	1	2	3
Number of Property Subtypes in Portfolios	3.55	2.2	2	3	5
Property Holding Periods (years)	7.22	5.57	3.27	5.97	9.75
Manager Number of Years in Sample	12.5	6.14	7.5	11.75	19
Manager HHI, by Region	0.7	0.24	0.5	0.69	0.98
Manager HHI, by Division	0.6	0.27	0.38	0.55	0.82
Manager HHI, by State	0.51	0.29	0.27	0.43	0.66
Manager HHI, by CBSA	0.42	0.3	0.19	0.33	0.6
Manager HHI, by Type	0.87	0.19	0.75	0.97	1
Manager HHI, by Subtype	0.76	0.22	0.57	0.8	0.98
Number of Properties: 13,025		Number of Managers: 126			
Private Portfolios					
Property Sizes (1000 Sq. ft.)	271.85	332.5	99.63	181.98	323.68
Portfolio Sizes (1000 Sq. ft.)	51,583	48,714	14,249	36,174	75,466
Portfolio Presence in Number of CBSAs	32.67	30.14	10.75	24	43
Number of Property Types in Portfolios	3.37	1.28	3	4	4
Number of Property Subtypes in Portfolios	8.03	5.03	4	7	11
Property Holding Periods (years)	5.11	3.93	2.25	4	7.5
Manager Number of Years in Sample	11.98	9.24	4.5	10.38	17.31
Manager HHI, by Region	0.51	0.23	0.34	0.41	0.59
Manager HHI, by Division	0.4	0.24	0.23	0.3	0.49
Manager HHI, by State	0.33	0.24	0.15	0.25	0.39
Manager HHI, by CBSA	0.25	0.22	0.1	0.17	0.35
Manager HHI, by Type	0.62	0.23	0.44	0.55	0.78
Manager HHI, by Subtype	0.48	0.21	0.33	0.43	0.58
Number of Properties: 22,322		Number of Managers: 144			

Table 3: Market Timing Measures

This table presents distributional characteristics across managers, for quarterly Market Timing (MT) measures. The measures are computed relative to style benchmarks at the National, Regional, Divisional, State, and CBSA level. At each geographic level of disaggregation, we use the benchmark for all property types combined, as well as benchmarks separated and matched by type and subtype. To save space, we only show results for measures computed at the Divisional level, as well as its property-class interactions. Panel A shows the cross-sectional distribution of time-series means by manager. Panel B shows the cross-sectional distribution of t-statistics testing the hypothesis $H_o : MT = 0$ against the two-sided alternative, over a manager's time series of quarterly MT measure observations. We include only managers for whom we can compute a MT measure over 12 quarters or more. *Fraction Sig.* is the fraction of portfolios that show a significantly positive t-statistic, at the 5% level (two-tailed test). FDR^+ is the positive-side False-Discovery Ratio, or the fraction of portfolios with significantly positive performance that would be observed if returns were purely determined by luck.

Panel A: Means

	Non-NYSE REITs			**NYSE REITs**			**NCREIF Private**		
	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype
Mean	0.0032	0.0033	0.0033	−0.0031	−0.0026	−0.0030	−0.0038	−0.0040	−0.0039
StdDev	0.0102	0.0102	0.0107	0.0043	0.0062	0.0059	0.0053	0.0058	0.0055
Min.	−0.0268	−0.0313	−0.0305	−0.0273	−0.0290	−0.0293	−0.0310	−0.0350	−0.0308
1st Qu.	−0.0036	−0.0036	−0.0038	−0.0041	−0.0049	−0.0052	−0.0059	−0.0060	−0.0060
Median	0	0.0003	0	−0.0026	−0.0025	−0.0027	−0.0040	−0.0039	−0.0042
3rd Qu.	0.0129	0.0126	0.0116	−0.0010	−0.0008	−0.0013	−0.0019	−0.0022	−0.0018
Max.	0.0269	0.0292	0.0317	0.0132	0.0404	0.0404	0.0100	0.0098	0.0091
Fraction > 0	0.3133	0.3253	0.3133	0.1190	0.1429	0.1190	0.1319	0.1250	0.1389

Panel B: T-Statistics

	Non-NYSE REITs			**NYSE REITs**			**NCREIF Private**		
	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype
Mean	−0.0271	0.1426	0.1261	−0.9666	−0.8770	−0.8867	−1.3632	−1.3704	−1.3087
Min.	−4.4984	−4.9755	−4.9929	−5.0263	−5.3771	−5.1698	−7.9175	−6.8903	−6.4165
1st Qu.	−1.1547	−0.7276	−0.7373	−1.5228	−1.4740	−1.3479	−2.3405	−2.2105	−2.1355
Median	−0.0092	0.0277	−0.0057	−0.8303	−0.7730	−0.7969	−1.2610	−1.2612	−1.1968
3rd Qu.	1.6644	1.7153	1.5581	−0.3093	−0.2578	−0.3429	−0.5356	−0.6242	−0.5020
Max.	3.8741	3.2836	3.1273	1.8225	2.3794	2.5283	3.3828	3.7695	3.1956
Fraction Sig.	0.0422	0.0964	0.0783	0	0.0159	0.0159	0.0208	0.0139	0.0139
FDR^+	0.0297	0.0331	0.0339	0.0203	0.0228	0.0218	0.0182	0.0174	0.0195

Table 4: Market Selection Measures

This table presents distributional characteristics across managers, for quarterly Market Selection (MS) measures. The measures are computed relative to style benchmarks at the National, Regional, Divisional, State, and CBSA level. At each geographic level of disaggregation, we use the benchmark for all property types combined, as well as benchmarks separated and matched by type and subtype. To save space, we only show results for measures computed at the Divisional level, as well as its property-class interactions. Panel A shows the cross-sectional distribution of time-series means by manager. Panel B shows the cross-sectional distribution of t-statistics testing the hypothesis $H_o : MS = 0$ against the two-sided alternative, over a manager's time series of quarterly MS measure observations. We include only managers for whom we can compute a MS measure over 12 quarters or more. *Fraction Sig.* is the fraction of portfolios that show a significantly positive t-statistic, at the 5% level (two-tailed test). FDR^+ is the positive-side False-Discovery Ratio, or the fraction of portfolios with significantly positive performance that would be observed if returns were purely determined by luck.

Panel A: Means

	Non-NYSE REITs			**NYSE REITs**			**NCREIF Private**		
	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype
Mean	−0.0004	−0.0003	−0.0009	−0.0002	0	−0.0005	−0.0002	−0.0001	0
StdDev	0.0055	0.0037	0.0041	0.0033	0.0023	0.0026	0.0030	0.0018	0.0018
Min.	−0.0177	−0.0124	−0.0148	−0.0152	−0.0078	−0.0109	−0.0091	−0.0085	−0.0084
1st Qu.	−0.0034	−0.0024	−0.0027	−0.0014	−0.0009	−0.0014	−0.0013	−0.0011	−0.0006
Median	−0.0004	−0.0001	−0.0002	0	−0.0001	−0.0004	0	0	0
3rd Qu.	0.0014	0.0015	0.0015	0.0010	0.0007	0.0004	0.0014	0.0008	0.0006
Max.	0.0174	0.0122	0.0079	0.0135	0.0132	0.0122	0.0094	0.0073	0.0072
Fraction > 0	0.3434	0.3313	0.3193	0.4286	0.4365	0.3333	0.3819	0.4028	0.3958

Panel B: T-Statistics

	Non-NYSE REITs			**NYSE REITs**			**NCREIF Private**		
	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype	Divisional	Div./Type	Div./Subtype
Mean	−0.3808	−0.1866	−0.4096	−0.1328	−0.1810	−0.4396	−0.1737	−0.0399	0.0962
Min.	−7.2973	−7.4810	−7.4810	−4.8725	−4.2446	−4.2515	−4.7676	−5.0960	−5.1639
1st Qu.	−1.7069	−1.3070	−1.5427	−1.4138	−1.1796	−1.3502	−1.4851	−1.1352	−0.7257
Median	−0.1053	−0.0839	−0.1814	−0.0371	−0.1813	−0.4243	−0.0806	0.0835	0.0734
3rd Qu.	0.9580	0.9695	0.7696	0.7556	0.7740	0.4152	1.3524	0.9257	1.0844
Max.	6.7475	4.0294	3.6118	6.0273	4.0167	4.0528	4.7520	4.2894	3.8652
Fraction Sig.	0.0602	0.0783	0.0663	0.0635	0.0635	0.0556	0.1458	0.1181	0.1181
FDR^+	0.0297	0.0331	0.0339	0.0203	0.0228	0.0218	0.0182	0.0174	0.0195

Table 5: Rank Persistence in Selection Measures

This table shows the persistence of relative rankings of managers, according to their MS Measure. Each year, we rank managers according to their realized Market Selection performance. $\rho_{t,\tau}$ shows the correlation between a manager's rank at time τ and at time t. $75plus$ and $90plus$ show permanence above the 75th and 90th percentile. Specifically, $75plus_{\tau,t}$ shows the fraction of all managers whose performance was at or above the 75th percentile at time τ, whose performance is still at or above the 75th percentile at time t. $90plus$ shows the analogous statistic for the 90th percentile.

Market Selection

	Non-NYSE REITs			**NYSE REITs**			**NCREIF Private**		
	Divisional	Divisional/Type	Divisional/Subtype	Divisional	Divisional/Type	Divisional/Subtype	Divisional	Divisional/Type	Divisional/Subtype
$\rho_{t,t-1}$	0.4129**	0.3357**	0.2742**	0.3946**	0.2598**	0.2335**	0.3714**	0.2807**	0.2849**
$\rho_{t,t-2}$	0.1837**	0.1179**	0.1649**	0.1791**	0.0857*	0.1136**	0.2002**	0.1417**	0.1317**
$\rho_{t,t-3}$	0.0214	0.0153	0.0054	0.0036	−0.0337	0.0357	0.0278	0.0596	0.0158
$75plus_{t-1,t}$	0.44	0.44	0.37	0.51	0.4	0.37	0.48	0.39	0.4
$75plus_{t-2,t}$	0.34	0.32	0.27	0.33	0.31	0.32	0.39	0.34	0.31
$75plus_{t-3,t}$	0.23	0.25	0.2	0.23	0.28	0.31	0.25	0.27	0.22
$90plus_{t-1,t}$	0.29	0.28	0.24	0.39	0.29	0.35	0.38	0.3	0.28
$90plus_{t-2,t}$	0.19	0.13	0.14	0.21	0.22	0.14	0.17	0.19	0.18
$90plus_{t-3,t}$	0.15	0.09	0.06	0.19	0.21	0.16	0.05	0.12	0.12

*:significance level $\leq$ 5%. **: significance level $\leq$ 1%.

Table 6: Top-Decile Forward Investment Returns

This table shows Market Selection-based returns with respect to a particular benchmark level, obtained by investing during year $t+1$ into all portfolios ranked in the top decile for year t, with respect to the same measure and benchmark level. The figures shown are time-series averages of equal-weighted cross-sectional mean returns for each year, and t-statistics testing the hypothesis that the time-series average of cross-sectional means is equal to zero, against the two-sided alternative.

	Non-NYSE REITs		**NYSE REITs**		**NCREIF Private**	
MS	Mean	t-statistic	Mean	t-statistic	Mean	t-statistic
National	0.0271	3.91**	0.0228	3.74**	0.0103	3.13**
National/Type	0.0172	2.27*	0.0161	3.83**	0.0063	2.37*
National/Subtype	0.0096	1.28	0.0098	1.94	0.0053	1.95
Regional	0.0273	4.06**	0.0262	3.77**	0.0172	5.02**
Regional/Type	0.0107	1.60	0.0075	1.90	0.0094	4.99**
Regional/Subtype	0.0073	1.06	0.0085	1.46	0.0105	4.83**
Divisional	0.0272	3.52**	0.0271	3.89**	0.0174	5.27**
Divisional/Type	0.0067	1.02	0.0081	1.93	0.0083	3.90**
Divisional/Subtype	0.0045	0.60	0.0128	2.67*	0.0070	3.00**
State	0.0246	3.19**	0.0220	3.00*	0.0160	4.75**
State/Type	0.0042	0.59	0.0064	1.82	0.0094	3.84**
State/Subtype	0.0055	0.65	0.0041	0.96	0.0078	3.23**
CBSA	0.0203	3.20**	0.0179	2.81*	0.0079	2.73**

*: significance level $\leq 5\%$. **: significance level $\leq 1\%$.

Table 7: Differences Between RCA-Benchmark and NPI-Benchmark Measures

This table shows means for Market Selection measures computed with RCA's transaction-based benchmarks ($Mean_{RCA}$), and means computed with the same methodology, but using NCREIF NPI benchmarks ($Mean_{NPI}$),for both public and private managers. We then show a t-statistic for the null hypothesis that the two means are identical.

Market Selection

	Non-NYSE-REITs		**NYSE-REITs**		**NCREIF Private**	
	National	National/Type	National	National/Type	National	National/Type
$Mean_{RCA}$:	−0.0018	−0.0007	−0.0007	−0.0003	−0.0013	−0.0003
$Mean_{NPI}$:	0	−0.0006	0.0002	−0.0005	−0.0010	−0.0007
T-test:	−2.52*	−0.13	−1.68	0.41	−0.47	0.82
	Regional	Regional/Type	Regional	Regional/Type	Regional	Regional/Type
$Mean_{RCA}$:	−0.0004		0.0003		0	
$Mean_{NPI}$:	0.0001		0.0006		−0.0006	
T-test:	−0.82		−0.51		1.41	

*: significance level $\leq 5\%$. **: significance level $\leq 1\%$.

Table 8: Regression Results, Liquidity

Dependent variable: Quarterly MT or MS measure by manager, for state-level benchmarks. This table presents results from panel regressions, by time and manager, for both public and private portfolios. The independent variables are the log of our manager-illiquidity measure, which indicates the liquidity of the markets in which the manager's portfolio is invested into, as well as a dummy *nyse*, which is one for portfolios of NYSE REITs and zero for public portfolios, and a dummy *private* defined analogously for private portfolios. The controls consist of the log of the manager's portfolio size, average property size, and manager portfolio specialization (by both geography and type). All independent variables except *private* and *nyse* are specified as one-year moving averages. For each specification, we show a model with date fixed effects and standard errors clustered by manager, and one with random effects. For the random effects, we show a pseudo-R-squared for the total variation captured by the model (R^2_{TOT}) and one for the variation captured by the fixed portion of the model (R^2_{FE}). The two subsequent statistics are a Chi-squared test of joint significance for all variables (χ^2_{MOD}), as well as a Breusch-Pagan test of the null hypothesis that no random effects exist (χ^2_{BP}).

Panel A: Market Timing

	State		**State/Type**		**State/Subtype**	
	Date Fixed Effects, Clustered by Manager					
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
log.mgr.illiq	−0.0007	−1.84	−0.0012	−1.76	−0.0017	−2.07*
nyse	−0.0006	−1.01	−0.0010	−1.41	−0.0004	−0.49
private	−0.0027	−3.89**	−0.0034	−4.1**	−0.0038	−4.05**
$\overline{R^2}$	0.435		0.271		0.209	
F	114.843		55.281		36.786	
Intercept and Controls	Yes		Yes		Yes	
	Random Effects					
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
log.mgr.illiq	−0.0009	−2.91**	−0.0016	−3.55**	−0.0018	−2.97**
nyse	−0.0010	−1.36	−0.0013	−1.7	−0.0006	−0.69
private	−0.0026	−3.15**	−0.0033	−3.81**	−0.0038	−3.93**
R^2_{TOT}	0.5087		0.3231		0.2469	
R^2_{FE}	0.0018		−0.0011		−0.0021	
χ^2_{MOD}	36.684**		53.252**		60.565**	
χ^2_{BP}	1868538**		739573**		403207**	
Intercept and Controls	Yes		Yes		Yes	

Panel B: Market Selection

	State		**State/Type**		**State/Subtype**	
	Date Fixed Effects, Clustered by Manager					
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
log.mgr.illiq	0.0029	4.49**	−0.0001	−0.17	0.0001	0.22
nyse	−0.0001	−0.11	0.0002	0.52	0.0004	0.81
private	−0.0003	−0.54	0	−0.02	0.0003	0.68
$\overline{R^2}$	0.02		0.013		0.018	
F	3.756		2.752		3.173	
Intercept and Controls	Yes		Yes		Yes	
	Random Effects					
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
log.mgr.illiq	0.0023	6.37**	−0.0002	−0.73	0	−0.02
nyse	−0.0006	−1.05	0.0001	0.2	0.0002	0.61
private	−0.0012	−1.74	−0.0002	−0.58	−0.0000	−0.07
R^2_{TOT}	0.1094		0.0498		0.0636	
R^2_{FE}	0.007		0.0025		0.0057	
χ^2_{MOD}	51.18**		10.256		15.191	
χ^2_{BP}	88202**		30326**		93525**	
Intercept and Controls	Yes		Yes		Yes	

*: significance level ≤ 5%. **: significance level ≤ 1%.