Conditional Geographical Clustering on REIT Performance, Efficiency and Shareholder Value

Daniel Huerta
Florida Gulf Coast University
Lutgert College of Business
Department of Economics and Finance
Lucas Institute for Real Estate Development & Finance
10501 FGCU Blvd. S.
Lutgert Hall LH-3361
Fort Myers, FL 33965
Phone: (239) 590-7315
E-mail: dhuerta@fgcu.edu

Christopher Mothorpe
College of Charleston
School of Business
Department of Economics
5 Liberty Street
Charleston, SC 29401
Phone: (843) 953-7273
E-mail: mothorpeca@cofc.edu

Working Paper

Conditional Geographical Clustering on REIT Performance, Efficiency and Shareholder Value

Project Description and Background

Conditional geographical clustering is the strategy of grouping real estate properties within a contiguous region to exploit economies of scale through spatial proximity. We expect significant benefits from this strategy as a result of gains in local market expertise and cost reductions associated with improved operational performance from the efficient management of a portion of a Real Estate Investment Trust (REIT) property portfolio. This strategy differs from both geographical diversification and agglomeration strategies. Geographical diversification is the strategy of acquiring properties in distinct geographical markets as to take advantage of the diversification effect of the differing economic conditions in the multiple markets. However, managing a property portfolio that is geographically disperse may pose challenges such as lack of expertise in the multiple markets, difficulty in property monitoring, lower management efficiency, and higher agency costs.[1] Prior literature finds REIT geographical diversification either destroys firm value or has little to no benefit. Ambrose, Ehrlich, Hughes, and Wachter (2000), Capozza and Sequin (1998, 1999), Gyourko and Nelling (1996), and Demirci, Eichholtz, and Yonder (2020) find either no, or limited, evidence of economic benefits. Whereas, Campbell, Petrova, and Sirmans (2003), Cici, Corgel, and Gibson (2011), Cronqvist, Hogfeldt, and Nilsson (2001), and Hartzell, Sun, and Titman (2014) present results that indicate discounts in value for geographically diversified REITs. More recently, Feng, Pattanapanchai, Price and Sirmans (2019) find geographical diversification benefits arise for REITs with high levels of institutional ownership and which invest in core property types. Agglomeration, on the other hand, refers to the strategy of locating properties near concentrations of economic activity such as in areas of fast economic growth or areas where similar properties owned by other firms are located. Prior literature explains agglomeration economies benefits firm productivity and provides positive externalities (Henderson 1986; Henderson 2003; Rosenthal and Stranges 2008; Melo et al., 2009; Greenstone et al. 2010;

[1] Feng et al. (2021) posit REIT geographical diversification may increase agency costs as managers employ capital into new markets where they lack expertise and professional relationships with local vendors.

Koster et al. 2014) which may explain the concentration of REIT properties in certain U.S. markets. However, agglomeration generally refers to the location of properties neighboring other properties that are not owned by the REIT.

In this paper, we examine the impact of conditional geographical clustering on REIT operations and firm value. Specifically, we test whether a strategy of property clustering translates into improved efficiency and performance that may impact REIT firm value and stock returns. That is, we explore channels through which conditioned geographical clustering contributes to REIT shareholder wealth. Such channels include operational efficiency, operational performance, and credit risk.

We contribute to the literature by measuring the optimal REIT cluster size (in terms of number of property units) and distance (in terms of amplitude of radii) by property-type specialization. This analysis provides REIT managers with indications of if property clustering is an effective strategy for all REIT specializations. Moreover, for those property-types for which clustering matters, our results provide guidance on the optimal proportion of the portfolio that should be clustered and the size of the cluster that will provide most benefit. The analysis by property-type specialization is of particular importance since each property sector has unique characteristics, distinct demand and supply drivers, and responds to economic factors in different ways. Each REIT asset class signifies a distinct business line with different economic sensitivities and which calls for a particular investment strategy that corresponds to the idiosyncrasies of the property type. Prior literature highlights the importance of property-type specialization segmentation in REIT studies finding, for example, that specialized REITs show varying degree of business cycle exposure, tend to have distinct levels of correlation with the economy, show markedly dissimilar capital structures, varying risk-return characteristics and deviations from net asset value, and are prone to different pricing anomalies (Wheaton, 1999; Reddy and Cho, 2018; Van Nieuwerburgh, 2019; Huerta et al., 2020).

Data and Sample Description

We obtain data for all publicly traded equity REITs from S&P Global Market Intelligence (formerly SNL) for the sample period 1993-2019. Our sample consists of 3,441 REIT-year observations for 310 REITs. For each REIT, we obtain a detailed description of their property portfolio for each sample year allowing us to observe property acquisitions, dispositions, values, and location, among other variables. Our dataset includes roughly 660,000 property-year observations of approximately 85,000 properties. We obtain REIT stock prices from the Center for Research in Security Prices (CRSP) and market factors from Kenneth French's website.[2]

Conditional Geographical Clustering

We identify property clusters using the density based spatial clustering of application with noise (DBSCAN) learning algorithm.[3] The DBSCAN algorithm classifies each property as a core point, a border point, or a noise point based on the minimum number of properties (*M)* in a cluster and the distance radius (*R*).[4] A core point is a property with at least *M* other properties within distance *R* of itself while a border point has at least one core point within distance *R*. Noise points are those not classified as a core or border point. A cluster is then defined as every core point within distance *R* of any other core point in the cluster as well as any border point within distance *R* of at least one core point in the cluster. There are two required inputs for the DBSCAN algorithm: (1) a minimum cluster size; and (2) a distance radius. We chose a minimum cluster size of two arguing efficiency gains from conditioned geographical clustering begin once a second property is located near the first property in a given location. We examine six different distance radii to determine the optimal cluster size per property type: (1) 5 miles; (2) 12.5 miles; (3) 25 miles; (4) 50 miles; (5) 75 miles; and (6) 100 miles.

---

[2] Kenneth French's data can be accessed at: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Last access on April 4, 2026.

[3] Available through ERSI's ArcGIS Pro package.

[4] For a complete discussion on the DBSCAN algorithm see Ester et al. (1996).

We run the DBSCAN algorithm for each REIT-year observation in our sample for the six distance radii and use the resulting information to calculate a *Cluster Average* variable, representing the degree of clustering for a REIT in a particular year. Equation (1) estimates *Cluster Average* for REIT *i* in year *t*, defined as the weighted average of the cluster size with weights determined by the size of the property measured in square feet relative to aggregate square footage of all properties in the REIT portfolio.[5]

$$Cluster\ Average_{it} = \sum_{j=1}^{N_{it}} j * \frac{w_{jit}}{S_{it}} \qquad (1)$$

where j represents the size of the cluster, N_{it} represents the largest cluster size for REIT *i* in year *t*, w_{jit} is the square footage of all properties within a cluster of size *j* for REIT *i* in year *t* and S_{it} represents the square footage of all properties within REIT *i* in year *t*.[6] If the value of *Cluster Average* is one, then the REIT consists of all "noise" properties (i.e., a portfolio of only free-standing properties that cannot be considered part of a cluster). Larger values of *Cluster Average* represent higher degrees of clustering.

Summary Statistics

Panel A of Table 1 provides descriptive statistics for our dependent variables. The average *Tobin's Q* is 1.19 and the mean *Firm Q* is 1.32 with standard deviations of .37 and .41, respectively. The average raw operational efficiency ratios are .67 for *OER1* and .45 for *OER2*, which mirror the averages reported by Beracha, Feng and Hardin (2019a).

Panel B displays summary statistics for our clustering measure (*Cluster Average*) defined by the various distance radii. The mean values are interpreted as the average degree of clustering,

[5] REIT property sizes are provided by S&P Global Market Intelligence depending on REIT property focus. We employ a unit multiplier (converter) provided by S&P Global Market Intelligence that converts the multiple size metrics to square feet.
[6] To calculate the missing property sizes, we replace the missing values with the average of the property size by primary and secondary property types.

defined by the weighted average of all clusters within a REIT-year (i.e., for a radius of 50 miles, the average REIT-year has 33.69 properties in a cluster). The minimum value for each distance radii is 1, indicating at least one REIT-year observation has an average of 1 property in every cluster. The degree of clustering and the large dispersion for the *Cluster Average* variable is significantly influenced by the distance radius employed. For example, at a distance radius of 25 miles the average degree of clustering is 19.61 with a range between 1 and 225.86 while at a distance radius of 75 miles the average degree of clustering is 48.77 with a range between 1 and 3,738.70. We investigate the optimal cluster size in section 5.5.

Figure 3 further illustrates the distribution of *Cluster Average* variable for a distance radius of 50 miles, and, at this distance, the DBSCAN algorithm identifies 109,726 clusters. Panel A of Figure 3 shows the distribution of all clusters by cluster size. Of the 109,726 clusters, 51,880 (47%) have 1 property, 36,735 (33%) contains between 2 and 5 properties, 9,172 (8%) contain between 6 and 10 properties, 10,041 (9%) contain between 11 and 49 properties, and 8,098 (2%) have greater than 49 properties. Panel B of Figure 3 shows the average number of clusters by size across all observations. The average REIT-year has 16 clusters with 1 property, 6.3 clusters with 2 properties, 3.8 clusters with 3 properties, 2.7 clusters with 4 properties, 2.2 clusters with 5 properties, 1.9 clusters with 6 properties, and 1.25 clusters with 7 or more properties.

Panel C of Table 1 presents summary statistics for our control variables. *Agglomeration* has a mean value of 170.92, indicating the average REIT has properties located near approximately 171 other REIT properties in a given year. *Agglomeration* has a range between 1 and 1,246 indicating there is at least one REIT where all its properties are isolated (more than 2 miles from any other REIT property) and at least one REIT owns properties concentrated in areas of extremely dense economic activity. *Self-advised* and *Self-managed* show average values of .89 and .80, respectively, suggesting most REITs in our sample make internal management decisions as well as self-manage their operations. The averages of the geographic HHIs range from -0.46 (*Region HHI*) to -0.27 (*MSA HHI*), and, as expected, the level of geographic diversification increases (larger values in absolute terms) as the geographical classification becomes more specific (e.g. from Regions to MSAs). The average *Property HHI* is -0.73 indicating REITs tend to concentrate on a particular property type. The average ratio of total debt to total assets (*Leverage*) is .50, the average *Firm age* is 2.30, and the average *REI growth* is 19.03. We winsorize *Tobin's Q*, *Firm Q*, *OER1*, *OER2*, *Leverage*, *Size*, *Firm age*, and *REI Growth* at the 1% level to mitigate the influence of outliers.

Empirical Strategy: REIT Operating Performance, Efficiency, and Value

To test for the impact of conditional geographical clustering on REIT operating performance and efficiency, we employ ordinary least squares (OLS) regressions with the following functional form:

$$\begin{aligned} Y_{it} = {} & \beta_0 + \beta_1 Cluster\ Average_{it} + \beta_2 \Delta Y_{it} + \beta_3 Size_{it} + \beta_4 Age_{it} + \beta_5 REI\ Growth_{it} + \beta_6 Price\ to\ NAV_{it} \\ & + \beta_7 Debt\ to\ NAV_{it} + \beta_8 Self\ Advised_{it} + \beta_9 Geographical\ Diversification_{it} \\ & + \beta_{10} Agglomeration_{it} + \beta_j Property\ Focus_{jt} + \varepsilon_t \end{aligned} \tag{2}$$

Where Y_{it} is, alternatively, return on assets (*ROA=FFO/TA*), return on equity (*ROE=FFO/TE*), operational efficiency (OER), credit risk, Tobin's Q, and Firm Q. ΔY_{it} is the change in performance *i* ($Y_t - Y_{t-1}$) to control for performance persistence. *Cluster Average*$_{it}$ is as defined in the previous section. *Size*$_{it}$ is the natural logarithm of market capitalization. *Age*$_{it}$ is the natural logarithm of the age of the REIT squared. *REI Growth*$_{it}$ is the percentage growth in real estate investments. *Price to NAV*$_{it}$ is the ratio of market price of equity to net asset value. *Debt to NAV*$_{it}$ is total debt as a percent of net asset value. *Self Advised*$_{it}$ is a dummy variable that takes the value of one if the REIT is self-advised and zero if the REIT is externally advised. *Geographical Diversification*$_{it}$ is a Herfindhal index measuring geographical concentration based on three orderings: by Metropolitan Statistical Areas (MSAs)[7], State, and National Council of Real Estate Investment Fiduciaries (NCREIF) regions. *Agglomeration*$_{it}$, a proxy for concentration of economic activity, is computed as the agglomerations of REIT-owned properties calculated using the DBSCAN algorithm with a minimum cluster size of 2 and a cluster radius of 2 miles. *Property Focus*$_{it}$ is a vector of binary variables that indicate whether a REIT primarily operates or invest in a specific property type or if it is property-type diversified. Property focus classifications include *Health Care*, *Hotel*, *Industrial/Office*, *Multifamily*, *Retail*, and *Self-Storage*. All models include year dummies to control for potential year fixed effects.

Empirical Strategy: REIT Stock Price Performance

We further examine the impact of conditional geogrphical diversification on the REIT stock return generating process. We test whether *Cluster Average* help explain the cross-sectional stock return of REITs. Specifically, we regress annual excess REIT stock return using the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model and the Fama and French (2015) five-factor model while including *Cluster Average*. Model specifications are as follows:

$$R_{REIT,it} - R_{f,t} = \alpha_i + \beta_1 MKT_{RF_t} + \beta_2\, SMB_t + \beta_3\, HML_t + \beta_4\, Cluster\ Average_{it} + \varepsilon_{it} \quad (3)$$

$$R_{REIT,t} - R_{f,t} = \alpha_i + \beta_1 MKT_{RF_t} + \beta_2\, SMB_t + \beta_3\, HML_t + \beta_4\, MOM_t + \beta_5\, Cluster\ Average_{it} + \varepsilon_{it} \quad (4)$$

$$\begin{aligned} R_{REIT,t} - R_{f,t} = \alpha_i + \beta_1 MKT_{RF_t} + \beta_2\, SMB_t + \beta_3\, HML_t + \beta_4\, RMW_t + \beta_5\, CMA_t \\ + \beta_6\, Cluster\ Average_{it} + \varepsilon_{it} \end{aligned} \quad (5)$$

where $R_{REIT,it}$ is the annual stock return of REIT i at time t, R_f is the risk-free rate at time t, and MKT_{RFt} is the value-weighted market return minus the risk-free rate at year t. SMB_t (Small minus Big), HML_t (High

[7] We employ the US Census Bureau's 2010 MSA definitions. When employing the MSA-level classification, if a property is located outside of a formally identified MSA, we place properties in their respective state.

minus Low), MOM_t (Momentum), RMW_t (Profitability) and CMA_t (Investment) are the year *t* Fama-French-Carhart return to zero investment factor-mimicking portfolios calculated to capture size, book-to-market, momentum, profitability and investment effects, respectively. The variable of interest, *Cluster Average*, determines whether geographical conditional clustering is related to REIT stock returns after controlling for market risk.

References

Ambrose, B. W., Ehrlich, S. R., Hughes, W. T., & Wachter, S. M. (2000). REIT economies of scale: Fact or fiction? *Journal of Real Estate Finance and Economics*, 20(2), 211-224.

Campbell, R.D., Petrova, M., & Sirmans, C.F. (2003). Wealth effects of diversification and financial deal structuring: Evidence from REIT property portfolio acquisitions. *Real Estate Economics*, 31(3), 347-366.

Capozza, D. R., & Seguin, P. J. (1998). Managerial style and firm value. *Real Estate Economics*, 26(1), 131–150.

Capozza, D. R., & Seguin, P. J. (1999). Focus, transparency and value: The REIT evidence. *Real Estate Economics*, 27(4), 587–619.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of finance*, 52(1), 57-82.

Cici, G., Corgel, J., & Gibson, S. (2011). Can fund managers select outperforming REITs? Examining fund holdings and trades. *Real Estate Economics*, 39(3), 455–486.

Cronqvist, H., Hogfeldt, P., & Nilsson, M. (2001). Why agency costs explain diversification discounts. *Real Estate Economics*, 29(1), 85–126.

Demirci, I., Eichholtz, P., & Yonder, E. (2020). Corporate diversification and the cost of debt: Evidence from REIT bank loans and mortgages. *Journal of Real Estate Finance and Economics*, 61, 316-368.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd* (Vol. 96, No. 34, pp. 226-231).

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1), 1-22.

Feng, Z., Pattanapanchai, M., Price, S. M., & Sirmans, C.F. (2021). Geographic diversification in real estate investment trusts. *Real Estate Economics*, 49 (1), 267-286.

Greenstone, M., Hornbeck, R., & Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 188, 536-598.

Gyourko, J., & Nelling, E. (1996). Systematic risk and diversification in the equity REIT market. *Real Estate Economics*, 24(4), 493–515.

Hartzell, J. C., Sun, L., & Titman, S. (2014). Institutional investors as monitors of corporate diversification decisions: Evidence from real estate investment trusts. *Journal of Corporate Finance*, 25, 61-72.

Henderson, J. V. (1986). Efficiency of resource usage a city size. *Journal of Urban Economics*, 19, 47-70.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics*, 53, 1-28.

Huerta-Sanchez, D., Jafarinejad, M., Kim, D., & Soyeh, K. W. (2020). Disentangling bubbles in equity REITs. *Quarterly Review of Economics and Finance, 76*, 357-367.

Koster, H. R. A., Van Ommeren, J. & Rietveld, P. (2014). Agglomeration economies and productivity: a structural estimate approach using commercial rents. *Economica*, 81, 63-85.

Reddy, W., & Cho, H. (2018). Emerging sector REITs. *The Routledge REITs Research Handbook* (pp. 56-76). Routledge.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and source of agglomeration economies. *Handbook of Regional and Urban Economics*, 4, 2119-2171.

Van Nieuwerburgh, S. (2019). Why are REITs currently so expensive? *Real Estate Economics, 47*(1), 18-65.

Wheaton, W. C. (1999). Real estate "cycles": some fundamentals. *Real estate economics, 27*(2), 209-230.

Table 1. Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Panel A: Dependent variables				
Tobin's Q	1.19	0.37	0.53	3.14
Firm Q	1.32	0.41	0.62	3.47
OER 1	0.67	0.28	0.17	2.54
OER 2	0.45	0.29	0.05	2.30
Panel B: Cluster variables				
Clustering, 5 miles	6.27	9.54	1.00	76.98
Clustering, 12.5 miles	14.34	24.03	1.00	189.65
Clustering, 25 miles	19.61	31.30	1.00	225.86
Clustering, 50 miles	33.69	137.04	1.00	3453.36
Clustering 75 miles	48.77	185.57	1.00	3738.70
Clustering, 100 miles	63.31	207.98	1.00	4003.29
Panel C: Control variables				
Agglomeration	170.92	191.70	1.00	1246.32
Self-advised	0.89	0.31	0.00	1.00
Self-managed	0.80	0.40	0.00	1.00
Region HHI	-0.46	0.28	-1.00	-0.13
State HHI	-0.33	0.26	-1.00	-0.04
MSA HHI	-0.27	0.26	-1.00	-0.01
Property HHI	-0.73	0.23	-1.00	-0.19
Leverage	0.50	0.17	0.00	0.98
Size	13.57	1.83	8.54	17.20
Firm age	2.30	1.01	0.00	3.99
REI growth	19.03	42.51	-40.81	301.59

This table reports summary statistics of variables used in this paper. There are 3,441 observations (*Firm Q* only has 3,433 observations). *Tobin's Q* is the ratio of the market value of equity plus the book value of debt to the book value of assets. *Firm Q* is the ratio of the implied market capitalization plus total assets minus the book value of equity to total assets. *OER1* is operational efficiency measured as the ratio of total expenses minus real estate depreciation and amortization to total revenue. *OER2* is operational efficiency measured as the ratio of total expenses minus real estate depreciation and amortization minus rental operating expenses to total revenue minus expense reimbursements. *Cluster Average* is a clustering score that measures the degree of conditioned geographical clustering. *Agglomeration* is a continuous variable representing the degree to which a REIT's properties are located near concentrations of economic activity. *Self-advised* is a binary variable indicating if the company makes acquisition and management decisions internally. *Self-managed* is a binary variable specifying if a REIT manages the day-to-day operations of its own properties. *Region HHI*, *State HHI*, and *MSA HHI* are geographical diversification Herfindahl indices by National Council of Real Estate Investment Fiduciaries (NCREIF) regions, state, and Metropolitan Statistical Areas (MSAs), respectively. *Property HHI* is property-type Herfindahl index. *Leverage* is the ratio of REIT total debt to total assets. *Size* is the natural logarithm of total market capitalization. *Firm age* is the natural logarithm of one plus the smaller of either the number of years since a REIT's initial public offering or the number of years since the firm adopted REIT status. *REI Growth* is the growth rate of real estate investment as defined by S&P Global Market Intelligence. The following variables are winsorized at the 1% level: *Tobin's Q*, *Firm Q*, *OER1*, *OER2*, *Leverage*, *Size*, *Firm age* and i.