Rental Price Adjustment, Volatility and Clustering:

Evidence from U.S. Office Space Market

Tian, Baiyue

Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Science

Columbia University

May 2013
Acknowledgments

I would like to express my sincere gratitude to my advisor Prof. David A. King and my thesis reader Prof. Jesse M. Keenan. I would like to thank you for your continuous support of my research, as well as offering me invaluable opportunity to participate in your research projects as research assistant at Columbia Center for Urban Real Estate during the past two years. I could not have imagined having a better advisor and mentor for my study at Columbia!

My sincere thanks also go to Dr. Youguo Liang, Dr. Tim Wang, and Mr. Robert D. Naso, for discussing with me on my thesis topic, sharing with me their insightful thoughts on regional economics and real estate, and advising me on my career.

I also want to thank all my friends and classmates at Columbia. I cannot name all of you in this brief paragraph, but I will remember all the happy times we enjoyed together on this fabulous campus and all the sleepless nights we were working together before deadlines.

Finally, I want to express my heartfelt appreciation to my entire family. Without your love and support, I would not have had the courage to get through all those difficult times studying abroad. Words are not enough to express what this has meant to me! I dedicate this thesis, however imperfect, to my family!
Abstract

Since last decades, researchers and practitioners became more and more interested in strategically allocating real estate assets in regions with common features instead of viewing each individual market separately. They conducted researches to develop their own classification systems which can be used in this strategic asset allocation process. Following a similar logic in grouping markets with common features together, I started a series of researches in developing a classification system on office space markets under three market conditions - rising markets, turbulent markets and recovering markets, and comparing the stability or the potential structural changes in the classification system under different market conditions. This study as the first research of a series of three researches intends to identify and estimate U.S. office space market's clustering dynamic between 2007 and 2012, which is developed as a classification system under turbulent market condition. The commonalities in this research are defined and measured by minimum average of all distances between the pair of observations from the pair markets based on three metrics – average effective rent, standard deviation of effective rent, and rental price elasticity. The outcomes are then tested through splitting dataset to validate the stability of this classification system. The deliverables from this research are a framework, which can be applied to all similar researches focused on market segments, as well as a classification system containing seven groups of markets, which sets up the foundation for future researches.

Key Words: Office Space Market, Rental Price Adjustment, Volatility, Cluster Analysis
Chapter 1: Background ................................................................. 1
Chapter 2: Study Purpose ............................................................... 3
Chapter 3: Literature Review ........................................................ 4
  1) Classification System in Real Estate Research ............................. 4
  2) Clustering Analysis in Real Estate Research ............................... 5
  3) Market Rent Determinant Model ............................................. 6
Chapter 4: Methodology ................................................................. 7
  1) Inputs used in Hierarchical Cluster Analysis ............................... 8
  2) Hierarchical Cluster Analysis .................................................. 10
  3) Validation of Hierarchical Classifications by Splitting Dataset ...... 11
Chapter 5: Data .................................................................................. 12
Chapter 6: Statistical Procedures and Results .................................. 13
  1) Rental Price Elasticity of Vacancy Rate ..................................... 13
  2) Average Linkage Cluster Analysis .......................................... 17
  3) Validation by Data Splitting ..................................................... 21
  4) Observations from Cluster Analysis Outcome ......................... 25
Chapter 7: Conclusions and Suggestions for Further Research .......... 30
  1) Conclusions ........................................................................... 30
  2) Suggestions for Further Research ............................................ 30
Reference ......................................................................................... 31
List of Figures

Figure 1 Dendrogram for Average Linkage Cluster Analysis ........................................ 19
Figure 2 Dendrogram for Average Linkage Cluster Analysis – Training Group ............ 23
Figure 3 Dendrogram for Average Linkage Cluster Analysis – Testing Group ............ 23
Figure 5 Average Effective Rent .................................................................................... 28
Figure 6 Standard Deviation ......................................................................................... 29
Figure 7 Rental Price Elasticity of Vacancy Rate* ....................................................... 29
List of Tables

Table 1 Rental Price Elasticity of Vacancy Rate.............................................................. 15
Table 2 Average Rent, Standard Deviation, and Elasticity (2007 Q1 -2012 Q3) ........ 17
Table 3 Average Linkage Cluster Analysis* ................................................................. 20
Table 4 Summary Statistics: Number of Markets, Mean, Standard Deviation........... 24
List of Equations

Equation 1 \( A_i = \frac{1}{n} \times \sum_{j=1}^{n} x_j \) ..........................................................8

Equation 2 \( s_i = \sqrt{\frac{\sum (x_j - \bar{x})^2}{n-1}} \) .........................................................9

Equation 3 \( \ln(y_i) = \alpha_i + \beta_i \ln(x_i) \) ..........................................................9

Equation 4 \( \epsilon_i = \frac{d \ln(y_i)}{d \ln(x_i)} = \beta_i \) ..............................................................10

Equation 5 \( D(X,Y) = \frac{1}{N_X \times N_Y} \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} d(x_i, y_j) \) .................................................11
Rental Price Adjustment, Volatility and Clustering:

Evidence from U.S. Office Space Market

Chapter 1: Background

Since last decades, researchers and practitioners became more and more interested in strategically allocated real estate assets in regions with common features instead of viewing each individual markets separately. They conducted researches from different perspectives to develop such kind of classification systems to capture the common features within individual markets and then group them together, and demonstrate the advantage of their classification system in asset diversification. This research also focuses on developing classification systems to capture commonalities within markets. To uncover and illustrate the stability or potential structural changes in the classification systems under different business cycles, I started a series of research in developing classification systems on office space markets under three business cycles, defined as rising markets (2002-2006), turbulent markets (2007-2012) and recovering markets (2013-2018), and comparing the stability or the potential structural changes in the classification systems under these three different business cycles. Although to define a specific turning point of each cycle is almost impossible, we can still define a stage within a whole cycle from observing historical performance. Also, even though we are not able to capture the accurate turning point, since we are more interested in the dynamic in a period rather than a single point, it is still acceptable.

Market rents are seen to adjust in response to the local as well as the national
economic conditions (Hekman 1985). Researchers have demonstrated that there are commonalities existing among groups of markets, and based on those commonalities, different markets can be divided and grouped into a same cluster as long as they share a common dominant feature. (Deng, Fisher 2003, Goetzmann, 1995). Furthermore, the common market dynamic revealed and demonstrated in each group can be ascribed to the common characteristics of their local economy shared by these markets within a same group. Researchers have developed various classification systems to research on the homogeneous economic variables’ influence on real estate market performance under each classification criteria.

This research is the first one of the three researches to be conducted. I focus on 2007-2012 in this paper, developing a classification system under the turbulence in the market. Since the crisis in 2008, market rent of office space market plunged dramatically after previous years’ outstanding performance in most of the cities in the US. This provides us a good opportunity to illustrate the dynamics in office space market, with a particular focus on post-crisis era to uncover what are the reactions in different markets to this turbulence in the economy and the potential commonalities showed in their behavior under this turbulent market condition. Although, neo-classical economic theory states that it is the interaction between supply and demand side factors that determine the market rental price, considering real estate space market’s unique characteristic that the supply is inelastic during the short run, I assume a constant supply in this research as most urban economics and real estate researchers did in their researches on real estate or urban economics (Mueller, 1992).
Under this assumption, I identify the three metrics which is used in the cluster analysis in this paper from demand side.

In this study, a classification system based on three metrics: average effective rent, standard deviation of effective rent and rental price elasticity of vacancy rate is developed and under this framework I firstly tested the validity of this system and then discussed the common features observed in each group identified by this classification system. Although as aforementioned, researches also pointed out there are potential linkage between the local economy and the commonality in real estate markets’ performance, to test such kind of linkage is not the focus of this paper.

Chapter 2: Study Purpose

Firstly, the methodologies in constructing and validating the classification system employed in this research provides a framework for this kind of analysis on market segments from macro level by observing and comparing inter-market relations which can be performed before focusing on one specific market.

Secondly, the outcomes of the series of this research can help urban policy makers or real estate professionals to deepen their knowledge of local real estate market dynamic and the relationships of this market dynamic to other major office markets in the United States. They can view and monitor the market dynamic patterns of major “families” of markets and by investigating the inter-market commonalities and distinctions at a higher level can form a big picture and avoid deeply analysis on one specific market while losing the whole forest. This is more important in terms of strategic analysis and asset allocation.
Finally, combing the findings in the three researches to be conducted around this topic as aforementioned, it can provide evidence of stability or potential structural changes in market dynamic under different market conditions, thus providing information which is useful in risk management from real estate portfolio’s perspective.

Chapter 3: Literature Review

1) Classification System in Real Estate Research

Researchers are interested in identifying commonalities existing in cities with common economic features or same geographic locations and developed various classification systems to support their findings. Over the past thirty years, there is a shift from geographic classification to a pure local economic classification.

Researchers developed and improved classification system with an initial idea to diversify risks by allocating properties into areas with distinctive characteristics. The conventional approach is to diversify real estate portfolio by property types and/or geographic regions. Miles and McCue (1982) concluded that a strategy that diversified real estate portfolio by property types showed higher risk-adjusted cash yields than a four-region geographic strategy. In a later study, by using property-specific data from a large real estate commingled fund and tested correlation among returns in two different diversification strategy, Miles and McCue (1984) found the same conclusion as previous study. However, another group of researchers, such as Hartzell (1986) provided a contrary view that states geographic region diversification strategy worked better than property–type diversification strategy.
Grissom (1987) showed evidence that support the importance of geographic diversification within industrial real estate. Recent years, more and more researchers as well as professionals shifted from purely geographic diversification to an economic diversification strategy which sheds the geography altogether and diversify along purely economic lines. Mueller (1992) compared three strategies, including the NCREIF four geographic regions, the Solomon Brothers eight regions (a combination of economics and geography) and a purely economic grouping of the 316 Metropolitan Statistics Areas (MSA) in the United States. He found that the purely economic diversification strategy increases risk-adjusted returns compared to the other two aforementioned strategies. This new trend emphasizes the common features in local economic structure and their impacts on real estate performance.

2) Clustering Analysis in Real Estate Research

Prudential Real Estate Investors’ research team started to focus on developing economic location classification system since 1992 and had been consistently contributing to developing and updating a classification system by grouping Metropolitan Statistics Areas (MSA) with similar local economic features into same cluster. The latest version was updated by Fiorilla, Liang and Lumban-Tobing (2010). In their research, by applying multivariate cluster analysis on five economic features proposed, the thirty five top U.S. metro areas are divided into seven clusters: Capital Metro, New York Corridor, Tech Centers, Southern Growth, Lifestyle Centers, and Southern California. Goetzmann and Wachter (1995) also conducted research on this topic. In their research, K-means clustering algorithm is applied to effective rents and
vacancy data for twenty-one metropolitan U.S. office markets, respectively to identify a few major “families” of cities as they defined it, such as an oil and gas group and an industrial Northeast group, and so forth. Deng, Fisher, Sanders, and Smith (2003) applied clustering algorithm to net operating income (NOI) of commercial properties to identify clusters of properties with similar NOI performance and related those clusters to the local economic influences. Clustering techniques have become powerful tools for researchers who are interested in investigating similarities in real estate market performance and relating them to their local economic characteristics.

3) Market Rent Determinant Model

Since this research focuses on rental price’s dynamic patterns, there is a necessity to review the various models developed by previous researchers to describe or estimate market rent determinant or rental price adjustments.

Some of the researchers focus on supply and demand drivers to explain the market rent dynamic. Rosen (1984) provides a theoretical view of the supply and demand of office space. Shilling, Sirmans and Corgel(1987) and Wheaton and Torto, (1988, 1994) modified the initial model proposed by Rosen( 1984). The common premise of these researches is that the rent change is triggered by the un-equilibrium in demand and supply in office space market and vacancy rate is employed as a measure in the model. In addition to vacancy rate, there are many other demand side drivers used by researchers, such as unemployment rate, interest rate, GDP and so forth. The only supply side driver being tested is office stock (D’Arcy, 1996, Tsolacos, 1998), however, the result shows that supply side drivers do not have significant
influence on market rent.

Another group of researchers are also interested in interpreting market rent dynamic through investigating the property specific physical characteristics (Hough and Kratz 1983, Vandell and Lane 1989 and Doiron, Shilling and Sirmans 1992). Researchers also attempted to develop rent determinant model focusing on spatial allocation. Colwell and Sirmans (1978) and Colwell and Munneke (1997)’ researches demonstrated the relationship between property’s distance to the city center and rent variability. Archer and Smith (1994), studied the significant role of downtown office properties in local economies through investigating the viability of downtown office properties in growing metropolitan communities.

Although the aforementioned models come from different perspectives, they show their usefulness in explaining and sometimes forecasting the price dynamic in real estate markets. To tell which of them is better or more accurate is not the purpose of this paper. As I am more interested in identifying the office rental prices’ dynamic and developing a classification system based on those metrics that can be used to measure the rental prices’ dynamic in the twenty major office space markets. I follow the first group of researchers’ thread to develop inputs for the clustering model by building a multivariate regression model between rental price and demand side factors.

Chapter 4: Methodology

This paper employs two-step procedure to firstly identify metrics that represent the inter-markets’ commonalities and then develop a classification system based on
those commonalities. The first step is to identify and calculate the inputs for the log-linear model. The second step is to develop a classification system, using a multivariate regression technique known as cluster analysis, which identifies homogeneous groupings of major office markets based on the common characteristics within each market.

1) **Inputs used in Hierarchical Cluster Analysis**

Commonality is a broad and abstract concept. Researchers have derived various classification systems from different inputs used in their models to describe and represent those specific common features they are interested in (Goetzmann and Wachter 1995; Fiorilla, Liang and Lumban-Tobing 2010; Anderson and Shain 2001; Mueller 1993). Due to the special focus on a turbulent period in the U.S.’s real estate markets, this paper weighs more on the reaction of office space markets to the market conditions from 2007 to 2012. Consequently, the three specific metrics employed by this research are: average effective rent, standard deviation of the effective rent and rental price elasticity of vacancy rate.

Average effective rent is calculated as arithmetic mean from the quarterly effective rent data provided by Reis database on twenty major office space markets in the U.S. from Q1 2007 to Q3 2013. Effective rents are asking rents net of any rental concessions, expressed over the life of the lease term. Mathematically, the average effective rent can be calculated as:

\[
A_t = \frac{1}{n} \times \sum_{j=1}^{n} x_j \quad (1)
\]

Where,
RENTAL PRICE ADJUSTMENT, VOLATILITY AND CLUSTERING

\( A_i \) is the average (or arithmetic mean) effective rent in market \( i \);

\( n \) is the number of quarters in the dataset;

\( x_j \) is the actual value of effective rent in quarter \( j \).

Standard deviation of effective rent measures the dispersion of actual effective rent to average effective rent. The larger this dispersion or variability of actual effective rent to its average level is, the higher the standard deviation. The smaller this dispersion or variability is, the lower the standard deviation. We can use this to measure the expected risk brought by the effective rent’s movement. The standard deviation of effective rent is calculated as:

\[
 s_i = \sqrt{\frac{\sum(x_j - \bar{x})^2}{n-1}} \tag{2}
\]

Where,

\( s_i \) is the standard deviation of effective rent in market \( i \);

\( x_j \) is the actual effective rent in quarter \( j \)

\( N \) is the number of quarters in the dataset.

The rental price elasticity of vacancy rate reflects how sensitive the rental price of a market is to the recovery of office space market as the vacancy rate declines. To estimate rental price elasticity of vacancy rate, a log-linear model is applied to the quarterly effective rent data and quarterly vacancy rate data for twenty major Office space markets in the United States. As both the dependent variable quarterly effective rent and the independent variable quarterly vacancy rate cannot be zero, a log-linear model is specified as:

\[
 \ln(y_i) = \alpha_i + \beta_i \ln(x_i) \tag{3}
\]
RENTAL PRICE ADJUSTMENT, VOLATILITY AND CLUSTERING

Where,

\( \ln(y_i) \) is the natural log value of quarterly effective rent for market \( i \);

\( \ln(x_i) \) is the natural log value of quarterly vacancy rate for market \( i \).

Then, after several steps’ algebraic manipulation, the log-linear model implies that the rental price elasticity \( \epsilon_i \) of vacancy rate for market \( i \) can be defined as following:

\[
\epsilon_i = \frac{d \ln(y_i)}{d \ln(x_i)} = \beta_i \quad (4)
\]

2) Hierarchical Cluster Analysis

The second step is to apply hierarchical cluster analysis technique, using the three inputs- average effective rent, rent volatility and rental price elasticity of vacancy rate, which we identified and computed in the first step. The Hierarchical clustering is a data analysis technique to build a binary tree of the data that successively merges similar groups of points. This algorithm initially places each data point into its own singleton group, and then iteratively merges the two closest groups until all the data are merged into a single cluster. The aforementioned algorithm then results in a sequence of groupings. The user chooses a “natural” clustering from this sequence.

Average Linkage Algorithm is one of the widely used techniques within Hierarchical Clustering. This algorithm examines the structure presents in a pairwise distance matrix to construct a dendrogram. In each step, the nearest two clusters are merged and formed into a higher-level cluster. This algorithm specifically defines the distance between two clusters as the average of all distances between the pairs of observations from the two clusters. Mathematically, Average Linkage Algorithm is
described by the following expression:

\[ D(X,Y) = \frac{1}{N_X \times N_Y} \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} d(x_i, y_j) \]  

Where,

\( d(x, y) \) is the distance between clusters, \( x \in X \) and \( y \in Y \);

\( X \) and \( Y \) are two sets of observations (clusters);

\( N_X \) and \( N_Y \) are the numbers of observations in cluster \( X \) and cluster \( Y \) respectively.

3) Validation of Hierarchical Classifications by Splitting Dataset

By applying the aforementioned numerical classification technique to the dataset, we are not only interested in the structure identified in the sample, but interested in the structure of the statistical population from which the sample is derived. It is possible that among the clusters gained by the classification there are some, which are representative only for the sample and not for the whole statistical population, thus these clusters can be called “artificial”. (Botta-Dukat, 2008) To avoid the interpretation of artificial clusters problem widely associated with numerical classification techniques, researchers have devised and introduced various validation techniques. This paper uses one of those methods called validation by data splitting, which was initially proposed by Botta-Dukat in 2008.

Data splitting is widely used when there are no well-specified hypotheses before analysis starts. Under this method, data are randomly divided into two parts, the same analysis technique is applied to each part, respectively. A cluster will be regarded to be valid, if there is one and only one cluster in the other dataset with similar characteristics. If there is no such cluster in the other parts, it means that the cluster is
characteristic only for the sample. If there is more than one such cluster in the other datasets, it indicates arbitrary divisions. This procedure will be applied to each level of the dendrogram generated by hierarchical classifications, starting from the highest level and continues on lower levels, until number of valid clusters start to decrease as the emergency of arbitrary divisions. The dendrogram should be interpreted on that level then.

However, we should be very cautious when applying this method, as we will never know the real structure of data in real world, we cannot conclude that the result is accurate. Accuracy can be measured only if the real structure of the statistical population is known (Botta-Dukat, 2008). Actually, unable to evaluate the accuracy of the data structure is a disadvantage existing in all the validation procedures. But the result from this validation method is still useful in practice as it at least demonstrate the stability of the structure identified. Fortunately, stability and accuracy are related, i.e. stability is necessary for accuracy, but it is not true vice versa (Botta-Dukat, 2008).

Chapter 5: Data
To identify the inputs for cluster models, we use quarterly effective rent and quarterly vacancy rate from twenty major office space market, provided by Reis database. Most of these markets are defined by Metropolitan Statistics Areas (MSA) or combined Metropolitan Statistics Areas (CMSA). The time period of this dataset ranges from 2007 Q1 to 2012 Q3, capturing the turbulence of the market during and after the crisis.
The first two inputs, average effective rent and the standard deviation of the effective rent are calculated directly from the quarterly effective rent data for each market. The third input discussed in this paper is rental price elasticity of vacancy, and it is calculated by applying log linear model to the dependent variable quarterly effective rent and independent variable quarterly vacancy rate for each market.

Chapter 6: Statistical Procedures and Results

1) Rental Price Elasticity of Vacancy Rate

Table 1 listed the results from log-linear regressions on 20 major markets. Most of the results show statistically significance at confidence level of 95% or above. The only two exceptions are New Jersey and DC. For DC, the P-value 0.051 is very close to statistical significance at confidence level of 95%. For New Jersey, the P-value 0.15, however this does not mean we have to exclude New Jersey market from our dataset, as P-value only shows statistically significance at given confidence level. If we widen our confidence level, we can accept this result at a confidence level of 85%, which is still reasonable. Based on aforementioned reasons, we still include these two markets in the following analysis.

Table 2 presents a summary of the results from the first step. Different markets show different dynamics in the period from 2007 Q1 to 2012 Q3. Some of the markets stand out of all the rest and show unique dynamic patterns. For example, New York the ranks top among all the twenty markets both in terms of average effective rent and standard deviation of the effective rent. The high rent generation ability is associated with high volatility. The rental price elasticity of vacancy rate ranks top three among
all the markets, which means the rental price is more sensitive to the recovery in market demand compared to those markets show less elasticity. We also recognize that there are several markets show similar performance in all the three metrics, such as Charlotte and Denver, Houston and Philadelphia and so forth. In the second step, we will further investigate the commonalities and the distinctions among the twenty major markets, by applying Average linkage Cluster Analysis to the metrics identified and calculated in first step.
## Table 1 Rental Price Elasticity of Vacancy Rate

<table>
<thead>
<tr>
<th></th>
<th>New York</th>
<th>New Jersey</th>
<th>DC</th>
<th>Baltimore</th>
<th>Philadelphia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacancy Rate</td>
<td>-0.400***</td>
<td>-0.0853</td>
<td>-0.0738</td>
<td>-0.140***</td>
<td>-0.181***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.150)</td>
<td>(0.051)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.800***</td>
<td>3.123***</td>
<td>3.907***</td>
<td>3.324***</td>
<td>3.487***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.804</td>
<td>0.101</td>
<td>0.177</td>
<td>0.435</td>
<td>0.535</td>
</tr>
</tbody>
</table>

P-values in parentheses, * P<0.05, ** p<0.01, *** p<0.001

Source: Reis Database

---

## Table 1 Rental Price Elasticity of Vacancy Rate (Continued)

<table>
<thead>
<tr>
<th></th>
<th>Long Island</th>
<th>Boston</th>
<th>Oakland-E Bay</th>
<th>Seattle</th>
<th>Atlanta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacancy Rate</td>
<td>-0.299***</td>
<td>-0.314***</td>
<td>-0.194**</td>
<td>-0.283***</td>
<td>-0.201***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.006)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.872***</td>
<td>4.252***</td>
<td>3.568***</td>
<td>3.951***</td>
<td>3.432***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.491</td>
<td>0.326</td>
<td>0.420</td>
<td>0.767</td>
<td>0.852</td>
</tr>
</tbody>
</table>

P-values in parentheses, * P<0.05, ** p<0.01, *** p<0.001

Source: Reis Database
Table 1 Rental Price Elasticity of Vacancy Rate (Continued)

<table>
<thead>
<tr>
<th></th>
<th>Dallas</th>
<th>Charlotte</th>
<th>Denver</th>
<th>Houston</th>
<th>Chicago</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacancy Rate</td>
<td>-0.440*</td>
<td>-0.118**</td>
<td>-0.348***</td>
<td>-0.177**</td>
<td>-0.293***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.003)</td>
<td>(0.000)</td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.133***</td>
<td>3.201***</td>
<td>3.855***</td>
<td>3.499***</td>
<td>3.911***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.265</td>
<td>0.372</td>
<td>0.619</td>
<td>0.386</td>
<td>0.609</td>
</tr>
</tbody>
</table>

P-values in parentheses, * P<0.05, ** p<0.01, *** p<0.001
Source: Reis Database

Table 1 Rental Price Elasticity of Vacancy Rate (Continued)

<table>
<thead>
<tr>
<th></th>
<th>Detroit</th>
<th>Minneapolis</th>
<th>Phoenix</th>
<th>Los Angeles</th>
<th>Orange County</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacancy Rate</td>
<td>-0.194*</td>
<td>-0.238***</td>
<td>-0.346***</td>
<td>-0.356***</td>
<td>-0.466***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.228***</td>
<td>3.521***</td>
<td>3.991***</td>
<td>4.233***</td>
<td>4.416***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.205</td>
<td>0.765</td>
<td>0.847</td>
<td>0.767</td>
<td>0.740</td>
</tr>
</tbody>
</table>

P-values in parentheses, * P<0.05, ** p<0.01, *** p<0.001
Source: Reis Database
### Table 2 Average Rent, Standard Deviation, and Elasticity (2007 Q1 -2012 Q3)

<table>
<thead>
<tr>
<th>Market</th>
<th>Average Rent ($ PSF)</th>
<th>Standard Deviation</th>
<th>Elasticity ($i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>49.30913</td>
<td>5.065811</td>
<td>-0.40024</td>
</tr>
<tr>
<td>New Jersey</td>
<td>19.51174</td>
<td>0.509266</td>
<td>-0.08528</td>
</tr>
<tr>
<td>DC</td>
<td>41.84696</td>
<td>0.891493</td>
<td>-0.07382</td>
</tr>
<tr>
<td>Baltimore</td>
<td>18.87043</td>
<td>0.428162</td>
<td>-0.14026</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>19.8413</td>
<td>0.278115</td>
<td>-0.18057</td>
</tr>
<tr>
<td>Long Island</td>
<td>22.7487</td>
<td>0.829246</td>
<td>-0.29884</td>
</tr>
<tr>
<td>Boston</td>
<td>30.57522</td>
<td>1.805023</td>
<td>-0.31447</td>
</tr>
<tr>
<td>Oakland-East Bay</td>
<td>21.21783</td>
<td>1.217486</td>
<td>-0.19435</td>
</tr>
<tr>
<td>Seattle</td>
<td>24.61087</td>
<td>1.471304</td>
<td>-0.28311</td>
</tr>
<tr>
<td>Atlanta</td>
<td>17.16826</td>
<td>0.44108</td>
<td>-0.20132</td>
</tr>
<tr>
<td>Dallas</td>
<td>15.59565</td>
<td>0.691072</td>
<td>-0.44045</td>
</tr>
<tr>
<td>Charlotte</td>
<td>17.84435</td>
<td>0.422835</td>
<td>-0.11761</td>
</tr>
<tr>
<td>Denver</td>
<td>17.0587</td>
<td>0.771205</td>
<td>-0.3478</td>
</tr>
<tr>
<td>Houston</td>
<td>19.96174</td>
<td>0.890978</td>
<td>-0.1769</td>
</tr>
<tr>
<td>Chicago</td>
<td>21.42652</td>
<td>0.536357</td>
<td>-0.29343</td>
</tr>
<tr>
<td>Detroit</td>
<td>14.80043</td>
<td>0.978856</td>
<td>-0.19416</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>17.05652</td>
<td>0.51285</td>
<td>-0.23827</td>
</tr>
<tr>
<td>Phoenix</td>
<td>18.71043</td>
<td>1.51141</td>
<td>-0.34587</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>27.08739</td>
<td>1.369261</td>
<td>-0.35594</td>
</tr>
<tr>
<td>Orange County</td>
<td>22.40957</td>
<td>2.742393</td>
<td>-0.46644</td>
</tr>
</tbody>
</table>

Source: Reis Database

2) **Average Linkage Cluster Analysis**

The dendrogram in Figure 1 shows the results from Average linkage Hierarchical Algorithm. From observing the visual result shows in Figure 1, we recognize that the markets with similar features are listed close to each other, from the bottom to the top. The dendrogram is helpful in presenting the relationships visually; however, it is far less enough for research purpose. To continue analysis, we have to decide the number of groups.
Unfortunately, there is no technique for us to identify an optimal number of groups objectively, and this is the drawback existing in all kinds of cluster analysis techniques, no matter it is Non-hierarchical or Hierarchical. This is the art part of cluster analysis. We assign different group number based on our observation from Figure 1 to generate outcomes and compare those outcomes resulting from various group numbers. By observing the dendrogram in Figure 1, we initially assume the potential group number may be five, six, seven, eight, or nine. Assigning seven as the group number shows the most representative results. When we choose five or six, Boston, New York, and DC are always standing alone as one group each, while the rest markets are grouped together to form a large group, which weaken the interpretation of those rest markets as they are merged together. When we tested eight or nine groups, more individual markets were separated, which weakened the interpretation ability of the outcome. The result of seven groups is shown in Table 3.
Figure 1. Dendrogram for Average Linkage Cluster Analysis
Table 3 Average Linkage Cluster Analysis*

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
<th>Group 6</th>
<th>Group 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Island</td>
<td>Houston</td>
<td>Dallas</td>
<td>Seattle</td>
<td>Boston</td>
<td>DC</td>
<td>New York</td>
</tr>
<tr>
<td>Chicago</td>
<td>Baltimore</td>
<td>Detroit</td>
<td>Los Angeles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orange County</td>
<td>Philadelphia</td>
<td>Minneapolis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oakland-East Bay</td>
<td>New Jersey</td>
<td>Denver</td>
<td>Charlotte</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phoenix</td>
<td>Atlanta</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Outcome from cluster analysis
3) **Validation by Data Splitting**

Even though we have tried different potential groups and conducted careful analysis to finalize the group structure in previous section, there is still no assurance of having arrived at a meaningful and useful set of clusters. Some test or set of tests must be applied to determine whether the solution differs significantly from a random solution (Punj, Stewart, 1983).

I employ the Data Splitting validation technique in this section to test the stability of the outcome in previous cluster analysis. To conducted Data Splitting validation, I first excluded three groups consists of single cities. The essence of Data Splitting is to identify similar structure as training group shows in testing group. For groups consists of single market, there is no reference market in another group. Through reasoning analysis, we can tell that the solution is still stable even if we include those three markets in validation analysis, although it is not straight forward without reference group. We can simplify this complicated process by referring to the descriptive statistics of the three measures we used in cluster analysis. As will be shown in Table 5 in the following discussion, the descriptive statistics of those three markets disclose their unique market dynamic patterns, and distinguish each of them from the rest markets. Moreover, in prior cluster analysis, no matter by how many groups I differentiate the markets’ performance, as long as the number is above 3, those three markets are separated from the rest. This evidence also shows the stability. Consequently, we can safely exclude New York, D.C. and Boston in our validation analysis in this section.
After excluded the aforementioned three groups, I randomly split the rest markets into two parts. The first part is named as Training Group; the second is named as Testing Group. Then, I apply Average Linkage Algorithm to two groups respectively. The results are shown in Figure 2 and Figure 3. In previous section, I discussed the validation rule: A cluster will be regarded to be valid if there is one and only one cluster in the other dataset with similar characteristics (Botta-Dukat, 2008). And the similar characteristics here are visually shown as they are grouped together at a same level within the dendrogram. This rule is applied to each level of the dendrograms in Figure 2 and Figure 3. I compare the two dendrograms by each level from top to bottom, until arriving at a level that violates our validation rule. At the level of four groups, for each group of the training part, there is one and only one similar group in testing parts. At higher levels, such as the level of three groups, group1 and group 2 as identified in previous cluster analysis merges in both parts. At the level of 5 groups, group 2 as identified in prior cluster analysis is further divided into two groups in both Training part and Testing part, however, such a pattern violates the validation rule, as there will be one group being similar to more than one group in the other dataset. Thus, the level of four-group is accepted by this validation procedure, which supports our cluster analysis.
Figure 2 Dendrogram for Average Linkage Cluster Analysis – Training Group

Figure 3 Dendrogram for Average Linkage Cluster Analysis – Testing Group
### Table 4 Summary Statistics: Number of Markets, Mean, Standard Deviation

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
<th>Group 6</th>
<th>Group 7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Effective Rent</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>21.95</td>
<td>19.38</td>
<td>19.38</td>
<td>25.85</td>
<td>30.58</td>
<td>41.85</td>
<td>49.31</td>
<td>22.88</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.56</td>
<td>0.56</td>
<td>1.75</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>8.74</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1.33</td>
<td>0.72</td>
<td>0.64</td>
<td>1.42</td>
<td>1.81</td>
<td>0.89</td>
<td>5.07</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.50</td>
<td>0.21</td>
<td>0.07</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>1.09</td>
</tr>
<tr>
<td>Elasticity*</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>-0.31</td>
<td>-0.19</td>
<td>-0.26</td>
<td>-0.32</td>
<td>-0.31</td>
<td>-0.74</td>
<td>-0.40</td>
<td>-0.26</td>
</tr>
<tr>
<td></td>
<td>0.11</td>
<td>0.10</td>
<td>0.12</td>
<td>0.05</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.11</td>
</tr>
</tbody>
</table>

*The minus sign in elasticity only shows the negative relationship between the rental price and vacancy rate.*
4) Observations from Cluster Analysis Outcome

Table 5 shows the summary statistics for each group. The first line shows the number of markets belonging to this group and the bottom line shows the standard deviation of each observation of the variable from the markets within the group. We are more interested in the middle line in each row in Table 5 as the figure in that line shows the average level of each metric for each group. By observing and summarizing the average value of all the three metrics used in the clustering analysis, we can obtain the features as described by those metrics for each group. I assign a number to each group, for example, New York belongs to Group 7. However, the number by itself has no meaning as they are just an assigned number to distinguish between each group. Recall from the literature review chapter, I reviewed past researches on this topic. In their researches, researchers always assigned a specific name to define the nature of each group. I will not assign a name to each group, because as stated in the background and study purpose chapter, this research is focused on the market dynamic and the potential structural changes or stability of such market dynamic, while to further link this market dynamic to the underlying industry or the local economy is beyond the scope of this research. I will further my research in the future to explore and illustrate the drivers of the market dynamic patterns from the perspective of local demand side or local economy structure. But before any further empirical analysis is conducted on this potential linkage, I will not assign a specific name to each group and continue to use those numbers to simply distinguish between each group. In the following, I will discuss the observations from the cluster analysis.
and the summary of descriptive statistics achieved from each group.

Group 1 consists of four markets: Long Island, Chicago, Orange County and Oakland-East Bay. The average effective rent of this group is $21.95 per square foot, which is close to the total average effective rent of the total seven groups. The standard deviation is 1.33, ranking in the mid of all the seven groups. The elasticity of this group is 0.31, which is above the total average elasticity of 0.26.

Group 2 consists of five markets: Houston, Baltimore, Philadelphia, New Jersey, and Phoenix. The average effective rent of this group is $19.38 per square foot, the same as that of Group 3, and this is the lowest average effective rent among all the seven groups. When it comes to the standard deviation, it is 0.72, ranking the second lowest. The rental price elasticity of this group is 0.19, also ranking the lowest among all the seven groups, which means, the rental price of this group is the least sensitive to the recovery in market demand.

Six markets belong to Group 3. They are Dallas, Detroit, Minneapolis, Denver, Charlotte, and Atlanta. The market dynamic of this group is very similar to that of Group 2. The average effective rent is $19.38 per square foot, the same as that of Group 2 and the lowest among all the seven groups. The standard deviation is lower than Group 2, showing relatively low market volatility even under the turmoil in the whole country’s economy. While the rental price elasticity is 0.26, higher than that of Group 2 and equal to the total average rental price elasticity.

Seattle and Los Angeles form Group 4. The average effective rent of this group is $25.85 per square foot, above the total average effective rent and ranking the fourth
among all the seven groups. The standard deviation of this group is 1.42, ranking the third highest among all the seven groups, reflecting a relatively high volatility in this group. The rental price elasticity is 0.32, also ranking the third highest of all the seven groups.

From Group 5 to Group 7, each group is formed by a single market. The unique characteristics in those three markets make them outstanding all of the rest and cannot be grouped with any other together.

Group 5 is formed by Boston market. The average effective rent of Boston is $30.58 per square foot, ranking the third highest among all the seven groups. The standard deviation is 1.81, which ranks the second highest. While its rental price elasticity is 0.31 and it is similar to Group 1 and Group 4.

Group 6 is formed by Washington DC. The average effective rent of this market is $41.85, ranking the second highest among all the seven groups and almost twice that of total average effective rent. The standard deviation is only 0.89, which is below the total average standard deviation. This market is relatively less volatile during this period, compared to the rest groups with relatively high average effective rent, such as Group 5, and Group 7. While, in terms of the rental price elasticity, the 0.74 elasticity is the highest among all the seven groups, which reflect the sensitivity of this market’s rental price to its recovery in the market’s demand.

Group 7 is formed by New York. The $49.31 per square foot average effective rent of this market is the highest among all the seven groups. At mean time, the standard deviation of this market is also the highest among all the seven groups. The
5.07 standard deviation is far beyond the total average of 1.17 and this number is almost 2.5 times that of the second highest market - Boston’s 1.81. This market is extremely volatile during this time period. The rental price elasticity of this market is 0.40, and this is the second highest among all the seven groups, well above the total average. The lowest is 0.19, reflected in Group 2 (Houston, Baltimore, Philadelphia, New Jersey, and Phoenix).

**Figure 4 Average Effective Rent**

![Average Effective Rent Chart](chart.png)

Source: Reis Database
### Figure 5 Standard Deviation

<table>
<thead>
<tr>
<th>Group</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 7</td>
<td>5.07</td>
</tr>
<tr>
<td>Group 6</td>
<td>0.89</td>
</tr>
<tr>
<td>Group 5</td>
<td>1.81</td>
</tr>
<tr>
<td>Group 4</td>
<td>1.42</td>
</tr>
<tr>
<td>Group 3</td>
<td>0.64</td>
</tr>
<tr>
<td>Group 2</td>
<td>0.72</td>
</tr>
<tr>
<td>Group 1</td>
<td>1.33</td>
</tr>
</tbody>
</table>

Source: Reis Database

### Figure 6 Rental Price Elasticity of Vacancy Rate*

<table>
<thead>
<tr>
<th>Group</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 7</td>
<td>-0.4</td>
</tr>
<tr>
<td>Group 6</td>
<td>-0.74</td>
</tr>
<tr>
<td>Group 5</td>
<td>-0.31</td>
</tr>
<tr>
<td>Group 4</td>
<td>-0.32</td>
</tr>
<tr>
<td>Group 3</td>
<td>-0.26</td>
</tr>
<tr>
<td>Group 2</td>
<td>-0.19</td>
</tr>
<tr>
<td>Group 1</td>
<td>-0.31</td>
</tr>
</tbody>
</table>

Source: Reis Database

*The minus sign in this table shows the negative relationship between the rental price and vacancy rate. In analysis, the minus sign is always ignored. We only care about absolute value.
Chapter 7: Conclusions and Suggestions for Further Research

1) Conclusions

Based on the above analysis, I identified seven groups according to the commonalities existing among the characteristics of the markets within the same group. I also realize that there are three markets behaving uniquely and cannot be grouped together with any other markets, they are Boston, D.C. and New York.

Among all the seven groups, Group 7 (New York) shows the highest average effective rent of $49.31 per square foot, while the lowest average effective rent during this period is $19.38, which are reflected in both Group 2 (Houston, Baltimore, Philadelphia, New Jersey, Phoenix) and Group 3 (Dallas, Detroit, Minneapolis, Denver, Charlotte, and Atlanta). When it comes to the standard deviation, which measures the volatility of rental price, the most volatile market is Group 7 (New York). The standard deviation of this market is 5.07. The lowest standard deviation is 0.64 which is reflected in Group 3 (Dallas, Detroit, Minneapolis, Denver, Charlotte, and Atlanta). In terms of rental price elasticity, the highest is Group 6 (DC)’s -0.74.

2) Suggestions for Further Research

Firstly, as aforementioned, I am going to extend the time range of my dataset, so that I can apply the same cluster algorithm to the other two time periods, and test whether there is a structural change in the market dynamic under different market conditions. Secondly, I am intended to discuss the common features in local economic structure in the markets within each group, further investigate the potential linkage between the local economy and the office space market dynamics.
RENTAL PRICE ADJUSTMENT, VOLATILITY AND CLUSTERING

Reference


RENTAL PRICE ADJUSTMENT, VOLATILITY AND CLUSTERING


RENTAL PRICE ADJUSTMENT, VOLATILITY AND CLUSTERING


Voith, R. and T. Crone, *National Vacancy Rates and the Persistence of Shocks in the U.S.*