Theoretical and Empirical Researches in Urban Management
Volume 9 Issue 1 / February 2014

THE EFFECTS OF CLUSTERING ON OFFICE RENTS: EMPIRICAL EVIDENCE FROM THE RENTAL OFFICE MARKET IN HO CHI MINH CITY

Du T. HUYNH
Fulbright Economics Teaching Program and Harvard Graduate School of Design
232/6 Vo Thi Sau, District 3, Ho Chi Minh City, Vietnam
dhuynh@gsd.harvard.edu

Abstract
This article investigates the positive effects on rents by office clustering in Ho Chi Minh City. The density of buildings in the CBD increases labor productivity and creates added value for the whole economy, stimulating economic growth and development. Applying similar methodologies and approaches used to identify the clustering effects in Houston and Amsterdam, I separated the rent effects of location density. After controlling main characteristics of buildings including: the location, age, size, quality and operational status, I find a strong positive effect of being located in dense office areas. For every time the size of the office district doubles, the rents tenants are willing to pay increases by 4.3 to 7.1 percent. This finding extends the literature of the clustering effects from the developed world to the developing one. In other words, the results support the argument that real estate market behavior is remarkably similar from place to place.
Keywords: Clustering effects, agglomeration economies, Ho Chi Minh City, rental office market.

1. INTRODUCTION

The classical theory of office location argues that office rents are based on face-to-face contact possibilities which are a decreasing function of distance to the central business district (CBD)(Heilbrun 1987). Even with the exponential change of technology, face-to-face contact is still an important factor in determining office rent (Bollinger et al. 1998). Proximity to other office buildings creates positive effects on office rents.Finding empirical evidence of the clustering effect, especially in the developed world, has been a significant interest of researchers (Archer & Smith 2003; Jennen & Brounen 2009; Shilton & Stanley 1999).

Approaching the issue from economies of scale, a synonym of agglomeration, Archer & Smith (2003) explain office clustering using demand economies of scale and demand for marketing image. The authors have found empirical evidence of clustering effects on the office market in Houston, Texas. Jennen & Brounen (2009) extended the existing literature by testing the rent effects of office clusters empirically outside the United States.They also found empirical evidence of clustering effects on office rents in Amsterdam. However, both works by Archer & Smith (2003) and Jennen & Brounen (2009) are still limited to the developed world. To extend the coverage of the literature, following similar methodologies and approaches, I demonstrate that the effects of clustering on office rents also exist in the thin and emerging market of HCMC. In other words, results of this article support the argument by Malpezzi (1999) that real estate market behavior is remarkably similar from place to place.

The remainder of this article is organized as follows: The second section is a literature review of agglomeration economies. The third section explains the ways to measure clusters. The fourth section briefly surveys the office market in HCMC and explains rationales for choosing HCMC. The data is described in the fifth section. The sixth and seventh sections present the empirical results and robustness check. The eighth section compares the results in HCMC to those in Houston and Amsterdam. The conclusion will be in the final sections.

2. AGGLOMERATION ECONOMIES

Agglomeration economies are the benefits resulting from businesses located in close proximity to one another in cities and industrial clusters (Glaeser 2010). It is widely observed that skyscrapers (primarily occupied by business offices) are concentrated in most cities' central business districts (CBD) around the world, from New York to HCMC. High rise buildings are highly clustered together (Shilton & Stanley 1999). This phenomenon is due to advantages of proximity (Rosenthal & Strange 2004). There are efficiency gains and cost savings that result from proximity to customers, suppliers, workers, and even competitors (Yankow 2006), and the benefits of the agglomeration economies come from transportation cost savings and increasing returns (Duranton & Puga 2004).

There are different ways to explain sources of agglomeration, but they are basically based on the argument by Marshal (1890) that knowledge spillovers, input sharing, and labor market pooling are the three main sources (Duranton & Puga 2004; Rosenthal & Strange 2004). Information exchange (Jaffe et al. 1993), industrial linkages (Henderson et al. 1995), and labor market search (Kim 1989) create agglomeration economies (Timothy & Wheaton 2001). Proximity facilitates information exchange to create knowledge spillovers, which makes industries locate close together, and pooling of industries become larger and denser. The clustering enables firms to recruit appropriate employees and enables employees to find work which commensurate with their abilities and competence. This means that the cost of finding jobs and filling vacancies is less and productivity increases.

Approaching slightly differently, Duranton (2004) argues that sharing, matching, and learning are three main mechanisms or micro-foundations of agglomeration economies. Sharing mechanisms deal with sharing indivisible facilities, sharing the gains from the wider variety of input suppliers that can be sustained by a larger final-goods industry, sharing the gains from the narrower specialization that can be sustained with larger production, and sharing risks; matching mechanisms by which agglomeration improves either the expected quality of matches or the probability of matching, and alleviates hold-up problems; and learning mechanisms are based on the generation, the diffusion, and the accumulation of knowledge. Sharing, matching and learning are not economies of scale within firms but across firms, and thus are positive externalities.

Higher productivity is the main benefit of agglomeration (Yankow 2006) while congestion, pollution and crime are the three major factors that cause its costs (Kahn 2010). In order to measure the effects of agglomeration, costs and benefits should be estimated simultaneously and comprehensively. Unfortunately, good estimates has not yet found (Combes et al. 2010). Therefore, the effort as so far is to measure major components of agglomeration and Glaeser (2010) points out three components which absorb benefits of agglomeration:

Urban economists infer urban success from high local wages, robust real estate prices, and growth in the number of people within an area. If a place is doing well, then employers should be willing to pay more for workers in that area, people should be willing to pay more for access to that place, and more people should move to that area.

Basically, the aggregated benefits of agglomeration should be the total value added from the increase of real estate prices and the increase of wages of a whole region. The value added from wage increase is consisted of two components: wage increase for each individual and the numbers of jobs. The higher wage in urban is due to higher productivity (Yankow 2006). However, according to spatial equilibrium theories, the difference between the wage in the CBD and the wage in the periphery is exactly the same as the transportation costs borne by individual workers. If there were a difference between wage increase and transportation costs, there would be a shift of jobs from one to another location. Therefore, the magnitude of the wage difference depends on the transportation costs and all extra profit from agglomeration is left over for landowners (DiPasquale & Wheaton 1993). The wage increase for individuals from agglomeration may not create added value for the whole economy because it is offset by individual transportation costs. Moreover, wage increase may be even lower than the cost of the whole economy because individual transportation cost, when congestion exists, is usually less than the cost of the whole society due to externalities (Gomez-Ibanez and Small 1999).

As a result, the increase of real estate values is the most obvious sign of agglomeration economies and there are different approaches to quantifythis increase. Measuring the increase of rents from clustering effects is one of them.

3. DETERMINING CLUSTERS

Measuring cluster sizes is complicated. Fundamentally, the larger the cluster, the larger the expected positive impacts and effects from the surroundings will have on a building, however. These effects attenuate with distance (Duranton & Puga 2004). Nearby objects are more related than distant objects (Tobler 1970). If agents are in close proximity to one another then there is more potential for business interaction (Rosenthal & Strange 2004). Interpersonal contact is a prominent factor in creating clustering effects. Conceptually, a precise measurement of cluster sizes should be the quantity of interaction among agents or people in a cluster. This quantity depends on the number of people and the frequencies of interaction among them. Unfortunately, it is not usually possible to determine these two factors, especially the frequency of interaction. Therefore, two alternative proximities have typically been applied and are measured by either total employment or total physical areas.

Regarding indentifying clusters based on employment, there are two major methods by Giuliano & Small (1993), Giuliano et al. (2005), McMillen (2001), or McMillen & McDonald (2004). However, due to the lack of employment data in the case of HCMC, these methods were not reviewed for this study. For the purpose of this article, the focus is on defining clusters throughout physical areas in which the total floor area of buildings is most commonly used as cluster size (Gat 1998; Archer & Smith 2003; Jennen & Brounen 2009). Even using this method, there are several ways to determine clusters.

The simplest way to determine the existence of a cluster is through direct. Archer & Smith (2003) applied this method to identify the clusters in Houston. The size of a cluster is the aggregate floor area of all buildings within visible clusters. These authors created 76 clusters representing 647 buildings; and there are two types of clusters. The eleven largest clusters were coded as dummy variables and the other variable is sizes of clusters measured by the total floor area that were generated into a natural logarithm in regressions. The regression results support the argument on clustering effect hypothesized in the work of Archer & Smith (2003). However, there are several problems that arise applying this methodology. First, it assumed clustering effects are uniform across all buildings in a cluster. It is the same for higher density buildings at the center and buildings on the periphery of a cluster. Clustering effect attenuating with distance was not considered. Second, the method of identifying the 76 clusters,

Theoretical and Empirical Researches in Urban Management
Volume 9 Issue 1 / February 2014

Theoretical and Empirical Researches in Urban Management
Volume 9 Issue 1 / February 2014

as Jennen & Brounen (2009) pointed out, is subjective, depending entirely on the authors experience and adjustment.

Taking advantage of geographic information system (GIS) software, Jennen & Brounen (2009) were able to address the possibility that the effect of a cluster declined with distance – distance decay clusters. Initially, the authors approached the problem by indentifying simple clusters using simple threshold measure, which was more sophisticated than those of Archer & Smith (2003). Simple clusters were determined by aggregating the floor area of all buildings which a certain distacne of the subject building (threshold), 0.5 km2, for example. The size of the cluster includes the floor area of the building chosen for its cluster and was calculated by applying the following formula:

$$Cluster_k = \sum_{i=1}^{n} w_{ik} size_i$$

$size_i$: total floor area of building i

w_{ik}=1 for $d_{ik} \leq$ threshold

w_{ik}=0 for $d_{ik} \geq$ threshold

d_{ik} is the distance between $building_k$ and its nearby $building_i$

This definition of the "simple cluster" is basically similar to that by Archer & Smith (2003), and ignores the possibility that clustering effects attenuate with distance. The alternative is to include measure of distance decay to reflect the possibility that the farther from a building, the less impact the other buildings have on a cluster. The distance decay clusters are determined as following formula:

$$Cluster_k = \sum_{i=1}^{n} exp(-\lambda * d_{ik}) size_i ,$$

with λ as the distance decay exponent and other parameters being the same as the above formula.

The level of distance decay is affected by the choice of λ (lamda). A higher level of λ means faster distance decay and steeper decreasing weight with an increase in distance. The left part of Figure 1 shows that at cluster k, the floor area of the building k contributes totally to the cluster; one square meter of a building 500 m away from building k will contribute 1 m^2 in the simple cluster, 0.6m^2 in the distance decay cluster with λ = 0.001; 0.1m^2 with λ = 0.005; and almost nothing with λ = 0.01.

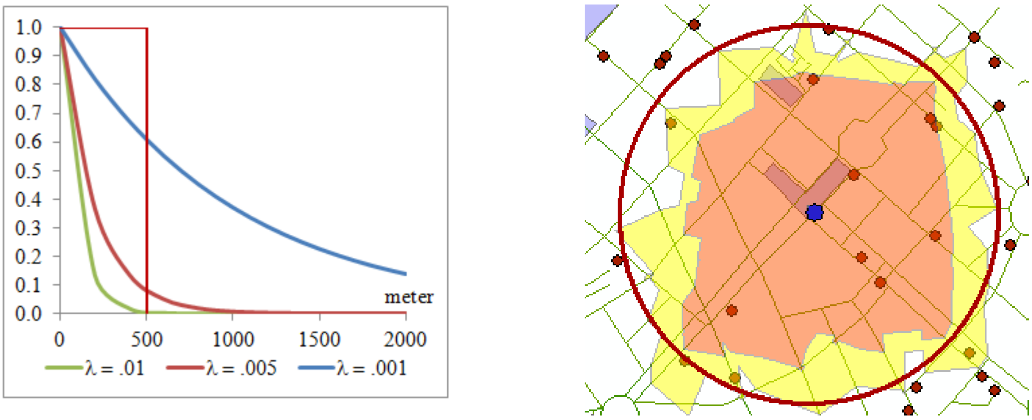

FIGURE 1 - ILLUSTRATION OF CLUSTERS
Source: Author's

One flaw in Jennen & Brounen (2009)'s approach to determining clusters is the assumption that distances between buildings are straight lines, while the actual distances depend on networks of roads, potential barriers and transportation conditions. There is no distinction between a building separated by a river and a building located on a large road, in relation to a defined building. To overcome this issue, this article uses actual distances based on real networks of roads as opposed to straight lines. The right part of Figure 1 shows clear differences between simply assumed and actual distances. There are 13 buildings within a circle buffer of a 400 m radius, while there are only nine buildings in the orange buffer with actual distance (400 m) from the center building. There are 13 buildings in the yellow buffer with actual distance (500 m) from the center building, but one building is different from the circle buffer.

While using actual distances is an improvement, it does not recognize that the traffic conditions are not the same everywhere and for all types of roads regardless of their size and conditions. Actual distance should be measured through travel time because time is a major proportion of transportation cost (Timothy & Wheaton 2001). Unfortunately, such a measurement is not available in HCMC yet.

Regarding choosing thresholds of simple clusters and λ of distance decay clusters: there is no clear cutoff points or obsticals. To eliminate the effect of choices made about the level of distance decay, following the same method by Jennen & Brounen (2009), this work tests the sensibility of the results to different assumptions about decay functions.

4. THE RENTAL OFFICE MARKET IN HO CHI MINH CITY

Conceptually, to evaluate economic impact of density or clustering, two critical sets of data are needed: added values created by clusters and cluster sizes. For the first set, data of added values, unfortunately, are not available in most cases. To test agglomeration economies realistically, at least one of the three signs are generally found, including: high local wages, robust real estate prices, and population growth, as listed by Glaeser (2010). In this article, real estate prices or rents are tested. For the second set of data, when using total floor areas to measure clusters, floor area of all buildings in a certain distance or area – regardless whether they are rental or direct ownership – should be included. Unfortunately, data on the size and rents of buildings that are occupied by other owners are not ready available for HCMC. Therefore, in this constraint, we examine the agglomeration economies in HCMC by quantitatively analyzing its rental office market with two concerns.

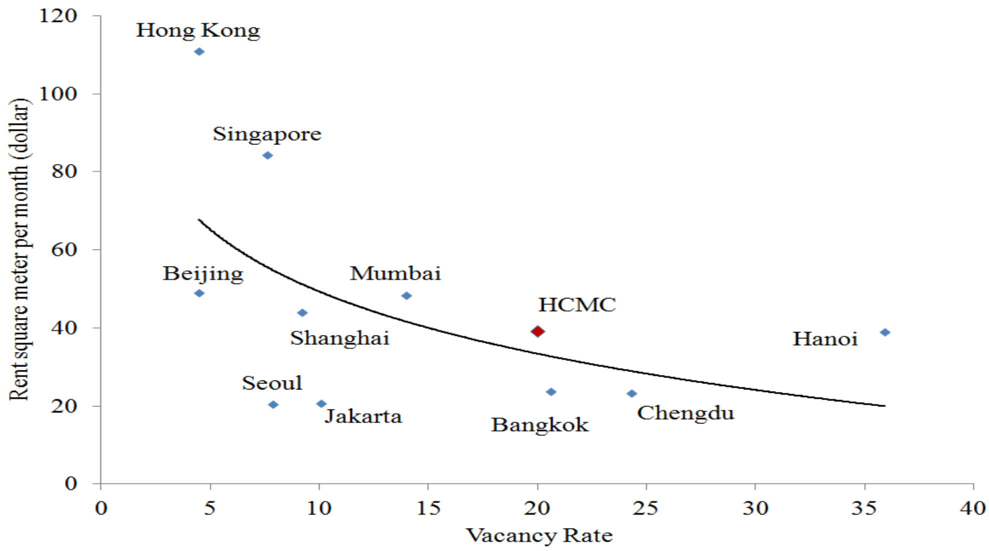

FIGURE 2 - RENT AND VACANCY RATES IN SELECTED CITIES IN ASIA
Source: Author's calculations from Colliers International (2011)

Notes: The horizontal axis represents the vacancy rate in percentage and the vertical axis represents the rent per square meter per month in dollars. This figure is to show the negative relationship between the vacancy and the rent

The first concern is that the office rental market in HCMC is in a highly unstable period. Right after Vietnam joined the World Trade Organization (WTO) in early 2007, the rental rate of grade-A buildings was about USD35/m^2 (CBRE Vietnam 2011), which was much higher than those of neighboring cities such as Bangkok and Kuala Lumpur, even thought total GDP or GDP per capita of these cities are much higher than those of HCMC. The price reached a peak at USD70/m^2 in the second quarter of 2008 with the vacancy rate of almost zero. Since then it has continuously declined to its early 2007 level, with around 20 percent vacancy at the end of 2011 (CBRE Vietnam 2011; Colliers International 2011; Savills Vietnam 2011) due to a rapid increase of supply and contraction of demand, which in turn was due to the global economic crisis and internal issues of Vietnam's economy. Despite the weak economy and high vacancy rates, the rental rates in HCMC at the end of 2011 were still over 50 percent higher than those of its neighboring cities (Fig. 2). HCMC's rental rates are indeed high by world standards, but which may not effect whether they are much higher for buildings in big clusters than for buildings in smaller clusters. In other words, what this study is less concerned with is not how HCMC rents compare with those in other cities, but rather with how they vary within the HCMC market.

The second concern is that the rental office market in HCMC is very small. There is considerable disagreement about the total stock of rental offices. The highest number is from the statistics produced by Colliers International (2011). According to the Second Quarter Report of this consultant company, the forecasted office space for rent area in HCMC by end of 2011 was 2.1 million m^2. Otherwise, in the April Report of 2011,CBRE Vietnam (2011) reported that there were 266 buildings with the total floor area of only 1.7 million m^2. The lowest number is from Savills Vietnam (2011). Its second quarter report in 2011 provided detail statistics on 171 buildings with a total rentable floor area of only 1.2 million m^2.

Even with the highest number, the stock of office for rent in HCMC is the lowest among 20 selected cities in Asia recorded by Colliers International (2011). In comparison to these cities, HCMC is basically in the lowest corner (Fig. 3). Moreover, the rental office space in HCMC is so tiny in comparison to 80,000 enterprises at the end of 2009 (PSO-HCMC 2011). The reason is that most businesses in HCMC are of small and medium size, and their offices are usually in small row houses on streets while big companies usually have their own office buildings that are close to office buildings for rent. If their buildings are included, office buildings in HCMC are denser, especially in the CBD.

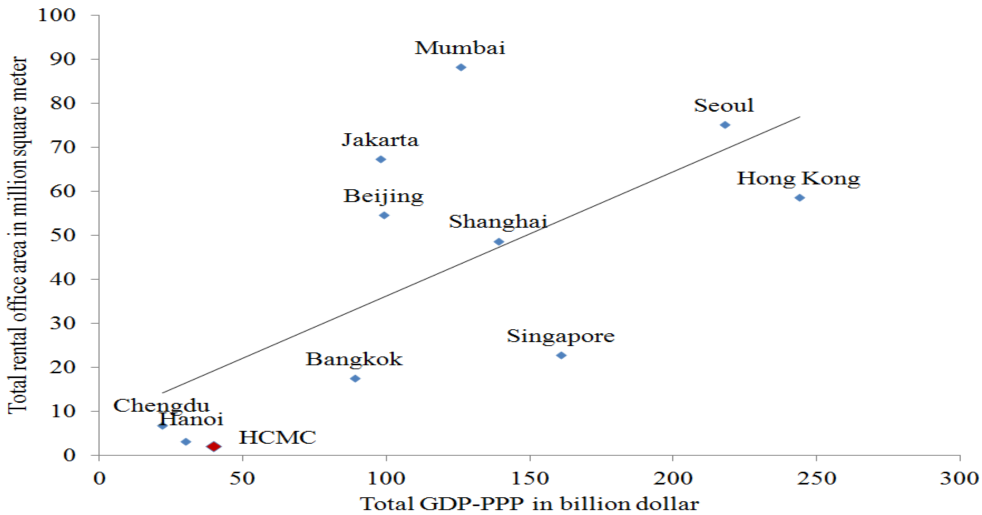

FIGURE 3 - RENTAL OFFICE STOCKS AND GDP OF SELECTED CITIES IN ASIA IN 2010
Source: Colliers International (2011) for office data, Habitat (2011b) for GDP data

Notes: The horizontal axis represents the total GDP-PPP (purchasing power parity) in billion dollars and the vertical axis represents the rental office are in million square meters. This figure is to show the positive relationship between the size of economies of cities and their rental office markets.

In addition to the two concerns previously mentioned, the transportation and urban pattern in HCMC are unique. Row houses, motorcycles, and sidewalk businesses (nhà phố, xe máy và kinh doanh vỉ a hè) are three typical characteristics of HCMC in particular, and Vietnamese cities in general (Nguyen, 2011). Motorcycle is the chosen mode of transportation by a vast number of residents and as such is one of the most unusual characteristics in HCMC in particular and Vietnam in general. Motorbikes carry over 80 percent of the passenger trips (DOT-HCMC 2011). Most streets are full of motorcycles from dawn to midnight every day. In general, the transportation mode in HCMC is very different from many cities around the world, especially developed cities.

The data set reflects the nature of the developing world where markets are growing but thin and data are incomplete. If the expected result hold true in highly unstable and thin markets, however, then the same should be more obvious in stable and matured markets. As this study will show, quantitative results reflecting the nature of the rental office market in HCMC are very similar to those in more developed cities, more precisely Huston and Amsterdam. Basic laws of agglomeration economies or clustering effects still work well in HCMC as will be demonstrated in the following sections.

5. THE DATA AND MAIN VARIABLES OF REGRESSION MODELS

For the quantitative analysis of this article, a data set of 171 buildings in HCMC is used and shown in Figure 4.

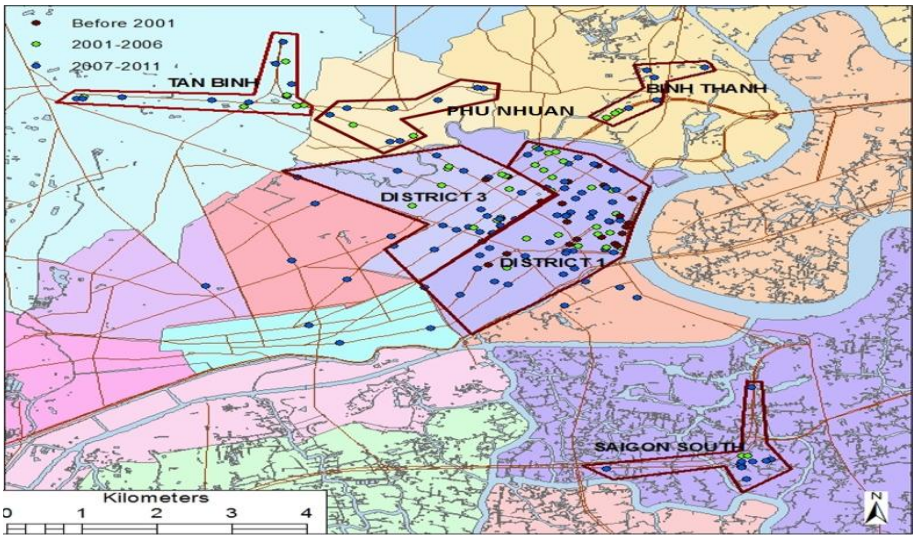

FIGURE 4 - THE DISTRIBUTION OF RENTAL OFFICE BUILDINGS IN HCMC
Source: Author's compilation in Arcmap 10 from various sources

Notes: The figure shows the distribution of rental office buildings in HCMC. There are six main groups or clusters. The two largest clustersare in the central business district (district 1& district 3). The clusters in Phu Nhuan and Tan Binh are along the arterial road connecting the CBD and the city's airport. The cluster in Binh Thanh is along the arterial road connecting the CBD and the national road #1. The cluster in Saigon South is in a newly built town.

The data is the combination of three sources: 1) names, addresses, rentable areas, rents, and vacancies were collected by Savills, a global real estate services provider listed on the London Stock Exchange in the second quarter, 2011; 2) ages (years since completion); and 3) heights (number of stories) of buildings and collected by the author mainly from websites: http://muabannhadat.com.vn/ and http://www.diaoconline.vn/, the two most visited online, commercial real estate companies in Vietnam. For buildings of uncertain age and number of stories, information was verified using the 1080 service, Vietnam's most popular information provider via telephone and internet. Definitions of variables are in Table 1 and the statistics summary of the sample is in Table 2.

TABLE 1: VARIABLES AND DEFINITIONS.

Variable	Definition
Dependent variable	
Rent	Rent in dollars per square meter per month
Cluster measurement (1000 m^2)	
Cluster	Total rentable area of buildings within 0.5 km^2
Cluster 001	Cluster at λ = 0.001
Cluster 002	Cluster at λ = 0.002
Cluster 005	Cluster at λ = 0.005
Spatial and location attributes	
Distance to CBD	Distance from each building to the central business district (km)
Physical characteristics of buildings	
Building size	Total rentable area of a building (m^2)
Age	Age of buildings
Height	Number of stores
Grade A	Dummy variable equals 1 if buildings is graded A
Grade B	Dummy variable equals 1 if buildings is graded B
Market operation	
Vacancy	Vacancy rate (%)

TABLE 2 - STATISTICS SUMMARY.

Variable	Observations	Mean	Std. Dev.	Min	Max
Independent variables					
Rent (USD)	173	23.94	8.45	10	55
Cluster variables (1000 m^2)					
Cluster	171	48.4	42.0	0.8	185.7
Cluster 001	171	155.2	83.9	14.9	304.6
Cluster 002	171	64.2	42.1	3.4	167.9
Cluster 005	171	20.1	15.9	0.9	80.4
Spatial and location attributes					
Distance to CBD (km)	171	2.7	2.0	0.2	8
Building characteristics					
Building size (m^2)	171	6,757	8,139	800	62,846
Building age (year)	171	5.8	4.8	1.0	40.0
Grade A	171	0.0	0.2	0.0	1.0
Grade B	171	0.2	0.4	0.0	1.0
Building height (storey)	171	12.8	6.7	3	68
Market operation variable					
Vacancy (%)[a]	171	11.6	18.1	0.0	92.0

The mean (unweighted avearge) of vacancy here is only 11.6% because of the dominant of grade-C buildings (small size). The weighted vacancy was 18% as calculation of Savills (2011).

Dependent Variable

The dependent variable is the log of the asking rent per square meter per month in US dollars or USD, similar to the dependent variable used by Jennen & Brounen (2009) while Archer & Smith (2003) used rent per square foot as the dependent variable. In Vietnam, all prices have to be quoted in the domestic currency, dong or VND by law. However, prices of goods or services related to foreigners (imported goods or partners, mainly are foreigners), are usually quoted in USD, even when transactions are in VND. The rents range between USD10 for grade-C buildings and USD55 for the grade-A buildings. The mean is USD24, skewed to the low end because of the domination of grade-C buildings, which account for over 70 percent of the total number of buildings.

Cluster Measurement and Spatial Attributes

Two key variables in this article are the cluster size and the distance to the CBD.[1] The former is used to demonstrate clustering effects or agglomeration economies while the latter is used to analyze the role of the transportation system.

The coefficient of the cluster size variable is expected to be positive. The larger the cluster, the larger the positive impact on rent rates. Four alternative sets of clusters have been created. The first set is 171 simple clusters calculated by aggregating floor areas of all buildings within 500 m of each buidling. The second to fourth clusters are distance decay clusters with λ = 0.001, 0.002, and 0.005, respectively. All distances have been generated through the real road network in HCMC by using ArcMap 10 –a professional GIS software.

There are two other variables including the distance to the CBD from clusters and visible clusters (clusters observed directly as described in the literature review). Archer & Smith (2003) included both, but the coefficient of the distance to the downtown is positive, contrary to conventional expectations. The authors explained this unusual result, but we speculate that it is because of the multicollinearity issue when the square of distance to downtown was also included and its coefficient is negative. The absolute value of the coefficient of the square of distance to downtown is much higher than that of the distance to the downtown. Jennen & Brounen (2009) ignored the distance to the CBD, and used visible

[1]Proximity to a main road also refects better accessibilty. However, in this article, the near main road dummy variable is not included as it was in the housing chapter because almost all buildings are located on main roads. Itwas tested in regressions, but it is not statistically significant and the R^2 is not improved.

clusters. No explanations for the omission were found, but multicollinearity is a likely reason. For the purpose of this article, the variable measuring the distance to the CBD from the cluster is included while a variable representing for visible clusters is omitted because of two reasons. First, one of the main focuses is the transportation system for travel to the CBD, so that the rent gradient from the CBD needs to be measured. Second, putting these two variables (the distance to the CBD and visible clusters) into the same regression causes a multicollinearity problem.[2] This issue is confirmed in the test of Variance Inflation Factors (VIF) in Table 5. The coefficients of the distance to the CBD variable should be negatively significant, e.g., the further away from the CBD, the lower the rents (Alonso 1964; Mills 1967; Muth 1969).

Building Characteristics

Several variables measuring the physical characteristics of buildings are included. The total floor area or building size is expected to create positive effects on rents (Archer & Smith 2003; Jennen & Brounen 2009). Historic buildings may create a positive impact on real estate prices (Coulson & Leichenko 2001; Jennen & Brounen 2009), although age in general causes negative impact on real estate prices. In this case, the coefficient of age is expected to be negative. Finally, a higher grade building means higher amenity and construction quality. Therefore, it is expected the coefficient of building grades to be positive.

Regarding age and grades of buildings, all buildings built before 2001, the year Vietnam and the United States signed the Bilateral Trade Agreement (BTA), are located in District 1 and District 3. Buildings built from 2001-2006, the time between BTA and when Vietnam joined the World Trade Organization (WTO), are located mainly in the downtown. After Vietnam's joining WTO, more buildings have been built but spread out beyond the CBD area.

In comparison to grade-C buildings, the coefficients of grade-A buildings and grade-B buildings are expected to be positive and the magnitude of the coefficient of grade-A bigger than that of grade B. There are 7 grade-A buildings, all of which are in the CBD; 40 grade-B buildings, most of which are in the CBD. The remaining 126 buildings are grade-C buildings and are located throughout the city.[3] Fourth, height, an indicator of density, is expected to be negative in the hedonic model for housing markets (DiPasquale & Wheaton 1993), but for office buildings, the situation is more complicated.

[2] The regression with visible cluster was tested and presented in Appendix 1. The result is consistent with the chosen regression model as presented in Table 3 below.
[3] Building grades are based on Savills' report. However, the grading is commonly used by most participants including CBRE.

Theoretical and Empirical Researches in Urban Management
Volume 9 Issue 1 / February 2014

Higher density can reduce amenities, but taller buildings can mean higher quality and better design, which could create a positive impact. Therefore, the impact of height of office buildings is difficult to predict. It may cause positive, negative, or have no impact on rent rates.

Market Operation

Theoretically, there is a natural vacancy rate at which market equilibrium is established. When vacancy rate is above the natural rate, rents will decrease (Grenadier 1995). There is no defined natural vacancy rate, but in common perception, it should be below 10 percent, at least, or even 5 percent. The vacancy rate in HCMC market is so high (18% of geometric mean) that is considered above the market equilibrium and vacancy means supply excess. In any event, the higher the vacancy of the building, the lower the price, except new or newly renovated buildings are being put into operation. Thus, the coefficient of vacancy is expected to be negative.

Empirical results

In this article, the log-log model with some modifications is used, which allows elasticities to be easily estimated and is similar to the model used by Jennen & Brounen (2009) for the purpose of comparison. Based on the availability of data, the hedonic regression is as follows:

$$\begin{aligned} Ln(Rent) = {} & \beta 0+\beta_1 ln(\text{Cluster measure})+\beta_2 ln(\text{Building size})+ \beta_3 ln(\text{Building age}) + \\ & +\beta_4 ln(\text{Building height}) +\beta_5 \text{Grade A}+ \beta_6 \text{Grade B} +\beta_7 ln(\text{Vacancy}) + \qquad (1) \\ & +\beta_8 \text{Distance to CBD} +\varepsilon \end{aligned}$$

The regression results are in Table 3. In the base model, most variables are statistically significant. The adjusted R-square is high at 0.701. The signs of coefficients of variables are exactly as expected, except for building height which is not statistically significant. The regressions are reasonable after adding the cluster variable. The coefficients of the cluster variable in all four regressions are statistically significant at the 1 percent level. The results of the distance decayed cluster with λ =0.002 and the simple cluster are close and they are the best (both R^2 and t-stat are high). The 0.72 adjusted R^2 means that this model accounted for nearly three-quarters of the office for rent market in HCMC. Therefore, these regressions will be chosen to explain related issues in this article.

The regression results confirm the clustering effects or agglomeration economies in HCMC. As the cluster size doubled, the rent would increase from 4.7-7.1 percent, and if the building size doubled, the rent per square meter would increase 5.3 percent. The implication of this is that buildings in the CBD tend to be bigger and they will be denser in the future.

TABLE 3 - REGRESSION RESULTS

Log of rent	Base Model	Simple cluster	Distance Decay Cluster		
			λ =.002	λ =.005	λ =.001
Log of Cluster size		0.047***	0.071***	0.062***	0.12***
		(3.7)	(3.57)	(3.34)	(3.37)
Log of building size	0.074***	0.053**	0.053**	0.039	0.062***
	(3.28)	(2.38)	(2.36)	(1.59)	(2.81)
Log of building age	-0.041**	-0.054***	-0.05**	-0.053***	-0.046**
	(-2.02)	(-2.76)	(-2.53)	(-2.66)	(-2.34)
Log of building height	0.005	0.014	0.017	0.012	0.014
	(0.13)	(0.36)	(0.44)	(0.32)	(0.37)
Grade A	0.629***	0.61***	0.611***	0.623***	0.61***
	(7.65)	(7.68)	(7.68)	(7.8)	(7.63)
Grade B	0.233***	0.21***	0.204***	0.213***	0.203***
	(5.57)	(5.15)	(4.96)	(5.2)	(4.88)
Log of vacancy	-0.017*	-0.015	-0.013	-0.015	-0.012
	(-1.74)	(-1.52)	(-1.32)	(-1.51)	(-1.19)
Distance to CBD	-0.08***	-0.07***	-0.061***	-0.072***	-0.045***
	(-10.98)	(-9.45)	(-6.93)	(-9.58)	(-3.62)
Constant	2.718***	2.389***	2.06***	2.396***	1.294***
	(15.43)	(12.47)	(8.22)	(12.22)	(2.84)
Adjusted R-square	0.701	0.7227	0.7212	0.7187	0.7190

Notes: Coefficients are standardized betas; T-stats in parentheses; *** at 1% significant level; ** at 5% significant level; and * at 10% significant level.

The coefficient of the distance to the CBD is negative as expected. The further from the CBD the building, the lower the rent will be. Specifically, each kilometer further from the CBD, the rent decreases 6-7 percent. This number has significant policy implications for HCMC.

Most of the variables measuring the physical characteristics and market operation follow the standard expectation. As buildings age their rents decrease. Similar to the office market in Houston, Texas in the United States, vacancy of the buildings in HCMC has no impact on rent. However, this result may be misleading because vacancy of individual buildings is less important in determining rents than the vacancy in the office market as a whole. Average rent of grade-A buildings is about two-thirds higher than that of grade-C buildings; and average rent of grade-B buildings is about a fifth higher than that of grade-C buildings.

6. ROBUSTNESS CHECK

The empirical results confirm the clustering effects on the office market in HCMC. However, weakness in explanation of adding more variables and multicollinearity may seriously affect results. To eliminate

Theoretical and Empirical Researches in Urban Management
Volume 9 Issue 1 / February 2014

errors that may seriously affect the results, the goodness-of-fit test, correlation, variance inflation factor (VIF) test, and cutoff points will be conducted in this section.

First, there are cases in which models with additional variables appear better than models with fewer variables, but in reality the additional variables do not improve the explanatory power of the regression. The goodness-of-fit test is used to determine whether there is a significant improvement of models with the cluster variable and the base model without the cluster variable.

Or more precisely, the model with the cluster variable explains the office market in HCMC better than the base model. The goodness-of-fit test is based on following formula:

$$f = \frac{(R_1^2 - R_0^2)/J}{(1-R_1^2)/(N-K)}$$, with

R^2: Adjusted R square of models

J: different in freedom degrees between base model and model with the cluster variable

K: number of variables

N: number of observations

TABLE 4 - JOINT TEST OF THE SIGNIFICANCE OF GOODNESS-OF-FIT REGRESSION MODEL

	Base model	Simple cluster	Cluster with λ =.002
N	171	171	171
R^2	0.7011	0.7227	0.7212
K		8	8
J		1	1
Critical value			
1%		6.64	6.64
5%		3.84	3.84
F statistics		12.70	11.75

Notes: The goodness-of-fit test is used to determine whether there is a significant improvement of models with the cluster variable and the base model without the cluster variable.

Or more precisely, the model with the cluster variable explains the office market in HCMC better than the base model. The goodness-of-fit test is based on following formula:

$$f = \frac{(R_1^2 - R_0^2)/J}{(1-R_1^2)/(N-K)}$$, with

R^2: Adjusted R square of models

J: different in freedom degrees between base model and model with the cluster variable

K: number of variables

N: number of observations

The F statistics results show that models with a cluster variable are better than the base model, which means that the clustering effects exist.

Second, as Sheppard's (1999) has shown multicollinearity can be a major concern when using a hedonic model. It seems natural to expect multicollinearity to pose a problem for the estimation of hedonic prices.

The multicollinearity effect, in which some independent variables are highly correlated with each other, creates biased coefficient magnitudes and indicators.

Therefore, it is necessary to test multicollinearity between variables. In this article, two types of multicollinearity tests were pursued: the test of variance inflation factors (VIF).

TABLE 5 - VARIANCE INFLATION FACTORS (VIF)

Rejected model			Chosen model		
Variable	**VIF**	**1/VIF**	**Variable**	**VIF**	**1/VIF**
Distance to the CBD	10.50	0.10	Log of building size	2.68	0.37
District 1*	6.51	0.15	Log of cluster with λ =.002	2.06	0.48
Saigon South*	4.57	0.22	Distance to the CBD	1.92	0.52
Log of Cluster = λ 0.002	4.49	0.22	Grade B	1.92	0.52
Tan Binh*	4.30	0.23	Log of Height	1.66	0.60
District 3*	3.59	0.28	Grade A	1.57	0.66
Log of Building size	2.94	0.34	Log of Age	1.31	0.76
Grade B	2.07	0.48	Log of Vacancy	1.16	0.86
PhuNhuan*	1.75	0.57			
Log of Height	1.74	0.58			
Grad A	1.62	0.62			
BinhThanh*	1.60	0.63			
Log of Age	1.36	0.73			
Log of Vacancy	1.32	0.76			
Mean VIF	**3.46**		**Mean VIF**	**1.79**	

*Visible clusters

Notes: This table shows the results of testing the variance inflation factor (VIF) which quantifies the severity of multicollinearity in an ordinary least squares regression analysis. The results of the rejected model to the left of the table, the VIF of the distance to the CBD is above the threshold of 10, as argued by O'Brien (2007). The problem is that a major proportion of buildings are located in Discrict 1 (40%) of the CBD. Consequently, there is a significant correlation among variables. In the chosen model, the results of the variance inflation factor test are consistent with our previous expectation. All VIFs are far below the threshold of 10, which means that multicollinearity is not a significant concern.

Two VIF tests were run. In the regressions with both sets of variables: distance to the CBD and visible variables, there is indeed a multicollinearity problem. The VIF of the distance to the CBD is above the threshold of 10, as argued by O'Brien (2007). The problem is that a major proportion of buildings are located in Discrict 1 (40%) of the CBD. Consequently, there is a significant correlation among variables.

Theoretical and Empirical Researches in Urban Management
Volume 9 Issue 1 / February 2014

There are two options to minimize the multicollinearity problem: 1) increase the sample size, and/or 2) eliminate variables. In this case, it is not possible to increase the sample size. The second option was chosen and as previously analyzed, for the purpose of this article, visible variables have been eliminated and the distance to the CBD has been used.

In the chosen model, the results of the variance inflation factor test are consistent with our previous expectation. All VIFs are far below the threshold of 10, which means that multicollinearity is not a significant concern.

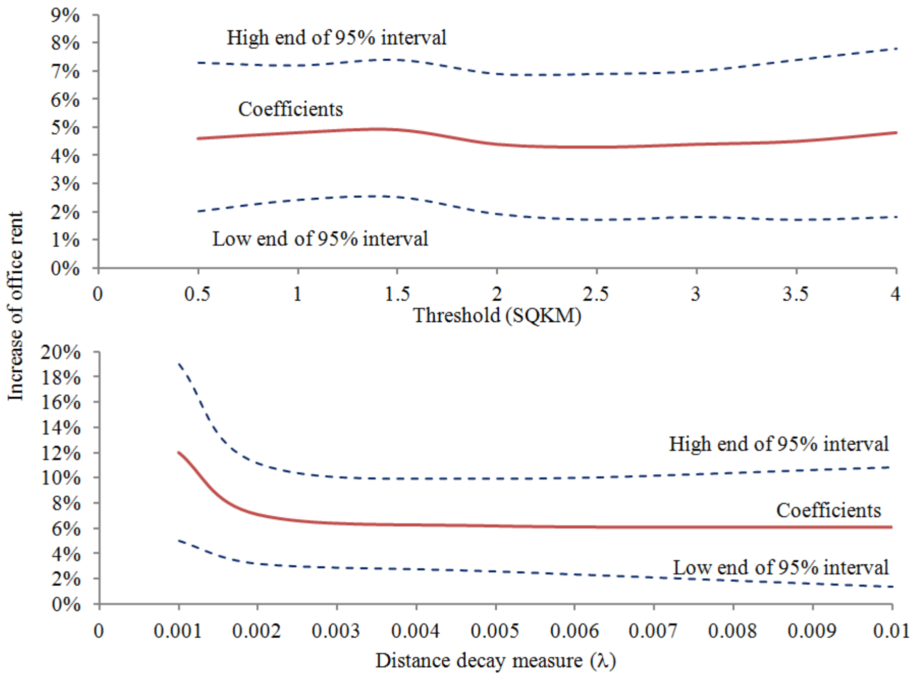

FIGURE 5 - REGRESSION COEFFICIENTS OF CLUSTER VARIABLES FOR DIFFERENT CUT-OFF POINTS.
Source: Author's work

Note: This figure is to test different cutoff points. For the simple cluster, the coefficients of regressions of cluster sizes from 0.5km^2 to 4km^2 are smooth and vary from 4.3-4.9 percent. All coefficients are statistically significant, but there are no statistical differences among them. For the distance decay cluster, however there is a difference. The coefficients of regressions of cluster sizes with λ from 0.001 to 0.002 vary drastically (from 7.1-12% when λ is doubled). However, those with λ from 0.002 to 0.01 are smooth. They vary from 6.1 to 7.1 percent. All coefficients are statistically significant, but there are no statistical differences among them.

Third, regarding determining cutoff points of thresholds in the simple clusters and λ of distance decay measure, there is no common rule or chosen points. Following previous studies, especially Jennen & Brounen (2009), the different cutoff points presented in Figure 5 were tested. For the simple cluster, the coefficients of regressions of cluster sizes from 0.5km^2 to 4km^2 are smooth and vary from 4.3-4.9 percent. All coefficients are statistically significant, but there are no statistical differences among them. For the distance decay cluster, however there is a difference. The coefficients of regressions of cluster sizes with λ from 0.001 to 0.002 vary drastically (from 7.1-12% when λ is doubled). However, those with λ from 0.002 to 0.01 are smooth. They vary from 6.1 to 7.1 percent. All coefficients are statistically significant, but there are no statistical differences among them.

Two conclusions can be withdrawn from the cutoff analysis. There are clustering effects or agglomeration economies with a wide range of cluster sizes. There are high probabilities that the cluster effects vary between 4.3 percent (the lowest point in the stable range of simple clusters) and 7.1 percent (the highest point in the stable range of the distance decay cluster). This range will be used to estimate values of agglomeration economies in the existing CBD.

7. COMPARISON TO AMSTERDAM AND HOUSTON

Two earlier studies of the effect of clustering on office rent in Amsterdam and Houston show similar results, especially the signs of estimated coefficients related to cluster effect (Tab. 7).

TABLE 6 - MEANS OF MAIN STATISTICS INDICATORS IN THREE ARTICLES.

Items	HCMC	Amsterdam	Houston
Observations	171	1,465	647
Rent (dollar/m^2)	23.94	178	165.55
Building size (m^2)	6,757	6,995	9,927
Simple cluster (1000m^2)	49	107	
Distance decayed cluster (1000m^2)	64	284	
Age	5.77		17
Adjusted R-squared	0.725	0.545	0.557

Source: HCMC's is author's results; data of Amsterdam is from Jennen & Brounen (2009); and data of Houston is from Archer & Smith (2003)

Note: This table presents main statistics indicators of the rental office markets in three cities including: Ho Chi Minh City, Amsterdam and Houston.

TABLE 7 - MAIN EMPIRICAL RESULTS OF THREE CITIES.

HCMC		AMSTERDAM		HOUSTON	
Dependent variable					
Log of rent (dollar/m^2)		Log of rent (dollar /m^2)		Rent (dollar /square foot)	
Independent variables					
Log of cluster	0.071***	Log of cluster	0.078*	Log of cluster	0.389**
Log of building size	0.053**	Log of building size	0.105*	Log of building size	1.350***
Log of building age	-0.05**	1971-90/1991-98[a]	-0.179*	Log of building age	-0.589***
Log of vacancy	-0.016			Vacancy	-0.769
Log of building height	-0.001			Log of building height	1.654***
Distance to CBD	-0.061***	Log of distance to highway	0.032*	Distance to downtown	41.6***

Source: HCMC's is author's results; data of Amsterdam is from Jennen & Brounen (2009); and data of Houston is from Archer & Smith (2003)

Notes: *** at 1% significant level; ** at 5% significant level; and * at 10% significant level.

There are six time periods with the earliest being the years 1991-1998, but only the period of 1971-1990 is used to make a preliminary comparison.

Note: This table presents main empirical results of the rental office markets in three cities including: Ho Chi Minh City, Amsterdam and Houston.

The most obvious evidence is that the coefficients of clusters of all three cities are statistically significant. More specifically, the magnitutes of coefficients of clusters between HCMC and Amsterdam are close. With distance decay clusters, when the cluster size changes 1 percent, the rent in HCMC and Amsterdam will change 0.071 and 0.078 percent, respectively. With the simple clusters, HCMC and Amsterdam are 0.047 and 0.057 percent, repectively. It is difficult to compare HCMC and Amsterdam with the Houston case because the linear-log equation was used in Houston instead of log-log equation. The signs of the coefficients of cluster variables are as expected, even it it is difficult to compare trends of magnitudes of coefficients. One can conclude that the clustering effects exist in all three cities inspite of many differences due to stages of development, political institutions and their impact, and other issues effecting clustering.

8. CONCLUSIONS

Looking around the world, social life and culture in cities are so different from one place to another. However, the nature and behavior of human beings are not very different. Similar to Houston and Amsterdam, the empirical results in this article illustrate that there is a positive clustering effect on office rents in HCMC. Some would argue that the situation in developing countries is too different to replicate or apply research or theories from the developed world. However, all fundamentals working in the mature markets also work surprisingly well in a thin market during a highly volatile period. This article extends existing literature by examining the influence of clustering outside the developed world. In other words, the findings of this article support the argument that real estate market behavior is remarkably similar from place to place (Malpezzi 1999).

Findings in this article give several implications for HCMC. That office rents in places where office buildings are denser tend to be higher than office rents where buildings stand alone or are less dense means that buildings in the CBD of HCMC tend to be bigger and they will be denser in the future. In perspectives of private investors or developers, it is better to build office buildings in the CBD and in places where office buildings are dense. From the perspectives of the municipal government or policymakers, this will cause an increase in transportation demand. This should make the concerns of the city's government focus on congestion in the downtown of the city more serious.

ACKNOWLEDGMENT

This article has been developed from a major chapter of my Doctor of Design Dissertation at Harvard Graduate School of Design. I am particularly grateful to José A. Gómez-Ibáñez, my primary advisor. I am fortunate to also have Richard B. Peiser and Annette M. Kim as my advisors. Comments of William Wheaton on the quantitative model and methodology have been especially helpful. I thank conference participants at the Fulbright Economics Teaching Program on December 21, 2012.

REFERENCES

Alonso, W. 1964. *Location and Land Use*, Cambridge: Harvard University Press.
Archer, W.R. & Smith, M.T. 2003. Explaining Location Patterns of Suburban Offices. *Real Estate Economics*, 31(2), pp.139–164.
Bollinger, C.R., Ihlanfeldt, K.R. & Bowes, D.R. 1998. Spatial Variation in Office Rents within the Atlanta Region. *Urban Studies*, 35(7), pp.1097–1118.
CBRE Vietnam. 2011. *The Second Quarter Report*,
Colliers International. 2011. *Asia Pacific Office Market Overview: 2Q-2011*,
Combes, P.-P. et al. 2010. Estimating agglomeration economies with history , geology , and worker effects. In E. L. Glaeser, ed. *Agglomeration Economics*. Chicago and London: The University of Chicago Press, pp. 15–65.
Coulson, N.E. & Leichenko, R.M. 2001. The Internal and External Impact of Historical Designation on Property Values. *Real Estate Finance*, pp.113–124.
DiPasquale, D. & Wheaton, W.C. 1993. *Urban Economics and Real Estate Markets*, Englewood Cliffs, NJ: Prentice Hall.
DOT-HCMC. 2011. Transportation Data. Available at: In Vietnamese.
Duranton, G. & Puga, D. 2004. Micro-foundations of Urban Agglomeration Economies. In V. Henderson & J. F. Thisse, eds. *Handbook of Regional and Urban Economics, Volume 4*. Amsterdam: North-Holland, pp. 2063–2117.
Gat, D. 1998. Urban Focal Points and Design Quality Influence Rents: The Case of the Tel Aviv Office Market. *Journal of Real Estate Research 16*, 16, pp.229–247.
Giuliano, G. et al. 2005. Not All Sprawl: Evolution of Employment Concentration in Los Angels, 1980-2000. Available at: http://www.usc.edu/schools/price/lusk/research/pdf/wp_2005-1002.pdf.
Giuliano, G. & Small, K. 1993. Is the Journey to Work Explained by Urban Structure? *Urban Studies*, 30(9), pp.1485–1500.
Glaeser, E.L. 2010. *Agglomeration Economics*, The University of Chicago Press.
Gomez-Ibanez, J.A. & Small, K.A. 1999. Urban Transportation. In P. Cheshire & E. Mills, eds. *Handbook of Regional and Urban Economics*. Amsterdam: North-Holland, pp. 1937–1999.
Grenadier, S.R. 1995. Local and National Determinants of Office Vacancies. *Journal of Urban Economics*, 37(1), pp.57–71.
Heilbrun, J. 1987. *Urban Economics and Public Policy*, New York: St Martin's Press.
Henderson, V., Kuncoro, A. & Turner, M. 1995. Industrial Development in Cities. *Journal of Political Economy*, 103(5), pp.1067–1090.
Jaffe, A.B., Trajtenberg, M. & Henderson, R. 1993. Geographic Localization of Knowledge Spillover as Evidence by Patent Citations. *Quarterly Journal of Economics*, 108(3), pp.577–598.
Jennen, M.G.J. & Brounen, D. 2009. The Effect of Clustering on Office Rents: Evidence from the Amsterdam Market. *Real Estate Economics*, 37(2), pp.185–208.

Kahn, M.E. 2010. New Evidence on Trends in the Cost of Urban Agglomeration. In E. L. Glaeser, ed. *Agglomeration Economics*. Chicago and London: The University of Chicago Press, pp. 339–353.

Kim, S. 1989. Labor Specialization and the Extent of the Market. *Journal of Political Economy*, 97(3), pp.692–705.

Malpezzi, S. 1999. Economic Analysis of Housing Markets in Developing and Transition Economies. In E. S. Mills & P. Cheshire, eds. *Handbook of Regional and Urban Economics*. Amsterdam: North-Holland, pp. 1791–1864.

Marshal, A. 1890. *Principles of Economics*, London: Macmillan.

McMillen, D.P. 2001. Nonparametric Employment Subcenter Identification. *Journal of Urban Economics*, 50(3), pp.448–473.

McMillen, D.P. & McDonald, J. 2004. Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983-1999. *Real Estate Economics*, 32(3), pp.463–486.

Mills, E.S. 1967. An Aggregative Model of Resource Allocation in a Metropolitan Area. *American Economic Review*, 57(2), pp.197–210.

Muth, R.F. 1969. *Cities and Housing*, University of Chicago Press.

O'brien, R.M. 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality Quantity*, 41(5), pp.673–690.

PSO-HCMC. 2011. *Statistical Yearbook, 2010*,

Rosenthal, S. & Strange, W. 2004. Evidence on the Nature and Sources of Agglomeration Economies. In V Henderson & J. Thisse, eds. *Handbook of Regional and Urban Economics: Cities and Geography*. Amsterdam: North-Holland, pp. 2119–2171.

Savills Vietnam. 2011. *The Second Quarter Report*,

Sheppard, S. 1999. Hedonic Analysis of Housing Markets. In P. Cheshire & E. Mills, eds. *Handbook of Regional and Urban Economics*. Amsterdam: North-Holland, pp. 1595–1635.

Shilton, L. & Stanley, C. 1999. Spatial Patterns of Headquarters. *Cities*, 17(3), pp.341–364.

Timothy, D. & Wheaton, W.C., 2001. Intra-Urban Wage Variation, Employment Location, and Commuting Times. *Journal of Urban Economics*, 50(2), pp.338–366.

Tobler, W.R.. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46(2), pp.234–240.

Yankow, J.J. 2006. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics*, 60(2), pp.139–161.

Appendix 1: Regression Results of Alternative Models

	With Distance to CBD			With Visible Clusters		
	Base model	Simple cluster	DDC* λ =.002	Base model	Simple cluster	DDC λ =.002
Log of Cluster size		0.047***	0.071***		0.044***	0.070***
		(3.7)	*(3.57)*		*(3.1)*	*(3.01)*
Log of building size	0.074***	0.053**	0.053**	0.063***	0.051**	0.053**
	(3.28)	*(2.38)*	*(2.36)*	*(2.81)*	*(2.29)*	*(2.38)*
Log of building age	-0.041**	-0.054***	-0.05**	-0.034*	-0.047**	-0.044**
	(-2.02)	*(-2.76)*	*(-2.53)*	*(-1.73)*	*(-2.41)*	*(-2.26)*
Log of building height	0.005	0.014	0.017	0.005	0.005	0.007
	(0.13)	*(0.36)*	*(0.44)*	*(0.12)*	*(0.12)*	*(0.18)*
Grade A	0.629***	0.61***	0.611***	0.658***	0.639***	0.633***
	(7.65)	*(7.68)*	*(7.68)*	*(8.27)*	*(8.22)*	*(8.11)*
Grade B	0.233***	0.21***	0.204***	0.200***	0.190***	0.186***
	(5.57)	*(5.15)*	*(4.96)*	*(4.7)*	*(4.58)*	*(4.45)*
Log of vacancy	-0.017*	-0.015	-0.013	-0.014	-0.012	-0.012
	(-1.74)	*(-1.52)*	*(-1.32)*	*(-1.36)*	*(-1.25)*	*(-1.18)*
Distance to CBD	-0.08***	-0.07***	-0.061***			
	(-10.98)	*(-9.45)*	*(-6.93)*			
District 1				0.345***	0.270***	0.239***
				(6.68)	*(4.86)*	*(3.91)*
District 3				0.261***	0.201***	0.182***
				(4.77)	*(3.55)*	*(3.06)*
BinhThanh				-0.041	-0.015	0.015
				(-0.57)	*(-0.21)*	*(0.21)*
Tan Binh				-0.120*	-0.163**	-0.157**
				(-1.85)	*(-2.52)*	*(-2.44)*
Saigon South				0.045	0.029	-0.016
				(0.68)	*(0.41)*	*(-0.23)*
PhuNhuan				0.055	-0.037	0.048
				(0.75)	*(-0.52)*	*(0.68)*
Constant	2.718***	2.389***	2.06***	2.374***	2.110***	1.791***
	(15.43)	*(12.47)*	*(8.22)*	*(12.7)*	*(10.49)*	*(6.72)*
Adjusted R-square	0.701	0.723	0.721	0.722	0.736	0.735
F-statistics		12.7***	11.75***		8.84***	8.12***

*DDC: Distance decay cluster

Notes: Coefficients are standardized betas; T-stats in parentheses; *** at 1% significant level; ** at 5% significant level; and * at 10% significant level.