

Use of Geographically Weighted Regression (GWR) Method to Estimate the Effects of Location Attributes on the Residential Property Values

Mohd Faris Dziauddin and Zulkefli Idris

Received: 08 04 2016 / Accepted: 24 08 2016 / Published online: 30 06 2017
© 2017 Faculty of Geography UGM and The Indonesian Geographers Association

Abstract This study estimates the effect of locational attributes on residential property values in Kuala Lumpur, Malaysia. Geographically weighted regression (GWR) enables the use of the local parameter rather than the global parameter to be estimated, with the results presented in map form. The results of this study reveal that residential property values are mainly determined by the property's physical (structural) attributes, but proximity to locational attributes also contributes marginally. The use of GWR in this study is considered a better approach than other methods to examine the effect of locational attributes on residential property values. GWR has the capability to produce meaningful results in which different locational attributes have differential spatial effects across a geographical area on residential property values. This method has the ability to determine the factors on which premiums depend, and in turn it can assist the government in taxation matters.

Keywords: Geographically weighted regression; Kuala Lumpur ; Location attributes; Residential property values

Abstrak Studi ini memperkirakan pengaruh atribut lokasi terhadap nilai properti residensial di Kuala Lumpur, Malaysia. Regresi berbobot secara geografis (GWR) diperkirakan memungkinkan penggunaan parameter lokal daripada parameter global, dengan hasil disajikan dalam bentuk peta. Hasil penelitian ini mengungkapkan bahwa nilai properti residensial terutama ditentukan oleh atribut fisik (struktural) properti, namun kedekatan dengan atribut lokasional juga berkontribusi sedikit. Penggunaan GWR dalam penelitian ini dianggap sebagai pendekatan yang lebih baik daripada metode lain untuk menguji pengaruh atribut lokasi terhadap nilai properti residensial. GWR memiliki kemampuan untuk menghasilkan hasil yang berarti dimana atribut lokasi yang berbeda memiliki efek spasial yang berbeda di wilayah geografis mengenai nilai properti residensial. Metode ini memiliki kemampuan untuk menentukan faktor-faktor dimana premi bergantung, dan pada gilirannya dapat membantu pemerintah dalam hal perpajakan.

Kata kunci: Atribut lokasi; Kuala Lumpur ; Nilai properti perumahan; Regresi tertimbang secara geografis.

1. Introduction

Location is acknowledged as an important attribute of immovable objects such as land and improvements to land. In the last five decades or so, many studies have been conducted to investigate the effect of location on land values and subsequently residential property values. In the standard practice of property valuation, property location is considered the most important factor in determining property value because the physical structure of the property depreciates over time. The land underneath the physical structure is what increases in value, and this is why the most common mantra we usually hear from real estate practitioners is 'location, location and location'. The theory of property value formation examines location as a composite effect of a set of locational

attributes [Kauko, 2003]. In other words, attributes such as the central business district (CBD), parks, shopping malls, schools, rail transit systems, highway, etc. may contribute positively towards the formation of a residential property value, particularly in urban areas. It has to be noted that, the benefits of locational attributes are realised mainly in the form of externalities, hence they are collectively shared by a large number of people and houses [Kauko, 2003; Orford, 1999]. Externalities can be positive (beneficial) and negative (harmful). For example, rail transit systems can generate both positive (improve accessibility) and negative externality (traffic congestion and noise pollution) effects. Since an assessment of the relationship between residential property values and a locational externality effect will benefit the government, housing developers and home buyers, it is important to evaluate their relationship in monetary terms. The most commonly used method to estimate the value formation of a residential property in monetary terms is hedonic pricing model (HPM). HPM is a well-established technique used for analysing a market for a single commodity with many attributes, in particular

Mohd Faris Dziauddin

Department of Geography & Environment Faculty of Human Sciences, Sultan Idris Education University, 35900 Tanjung Malim, Perak, Malaysia.

Zulkefli Idris

UMW Technology Sdn. Bhd.16-1, Level 16, Menara Prestige No 1, Jalan Pinang, 50450 Kuala Lumpur, Malaysia.
Correspondent E-mail: faris@fsk.upsi.edu.my

that of housing. HPM is based on the idea that residential properties are not homogeneous and can differ in respect to a variety of attributes. These various attributes will determine the residential property value. However, spatial issues within the housing market, such as spatial heterogeneity and spatial autocorrelation, have not been given adequate attention in HPM. Therefore, this study dissects the spatial distribution of residential property value formation for proximity to locational attributes. Knowledge about the spatial distribution of value formation is not only important to understand the effects of locational attributes on residential property values, but it also has important implications 'from a policy perspective as it underpins equity considerations of tax implementation' [Mulley, 2014]. Given that ensuring a higher degree of fairness and equity in a property tax implementation is likely to depend on such a spatial distribution, then it must be included in property valuation models. As Lockwood and Rossini [2012] argue, a professional valuer is obligated to specify and calibrate the best models to produce the accuracy required in accounting for location in the valuation process. Having said all the above, this paper seeks to contribute to the existing literature by focusing on the effects of locational attributes in a developing country setting and by considering the spatial distributional effects of residential property value formation. To map these effects, a method known as geographically (or locally) weighted regression (GWR) is used. This method enables the local parameters rather than the global parameters to be estimated, and thus it provides a way of accommodating the local geography of residential property value–locational attribute relationship.

The remainder of this paper is organised as follows. Section 2 reviews the related literature on the effect of locational attributes on residential property values. Section 3 discusses the Kuala Lumpur housing market. Section 4 presents the estimation methods in which the use of the GWR technique is justified and explained. Section 5 discusses the analysis and empirical findings. Section 6 concludes.

The theoretical framework of the relationship between location and land values developed by Alonso [1964], Muth [1969] and Mills [1972] indicated a distance decay relationship between land rent and distance from the central business district (CBD). In other words, land value decreases as the distance from the CBD increases. Note that the Alonso–Muth–Mills theoretical framework is a theory that underpins hedonic house price models, particularly the trade-off model of residential location and the distance decay relationship between land rent and distance from the CBD. According to Orford [1999], the motivation behind the early hedonic house price research was to provide empirical evidence of a negative land rent gradient as verification of the trade-off model. Note here that the distance from the CBD is not the only

factor influencing land values, but a variety of other locational, neighbourhood and environmental quality attributes on the location are also crucial to the formation of land values and subsequently property values [Freeman, 1979]. Ample evidence from the literature shows that the quality of location measured by locational, neighbourhood and environmental quality contributes significantly to the value formation of properties such as residential properties.

Thus far, the effect of locational attributes on residential property values has been estimated using HPM, although this method is found to be less sensitive to spatial variation, as the method assumes that the effect of locational attributes on residential property values is fixed across a geographical area (Fotheringham et al., 2002). Only a few studies have utilised spatial econometrics in property models (see, e.g., Orford, 1999; Du and Mulley, 2006; Bitter et al., 2007; Yu et al., 2007; Mulley, 2014; Dziauddin et al., 2015). For instance, in a study on residential property values in Milwaukee in the United States, Yu et al. [2007] used spatial regression and GWR to investigate the spatial variation of residential property values and found that the two methods produced more accurate and meaningful predictive results than the HPM. In addition, many studies have been conducted more on developed regions such as North America, Europe and Australia than on developing countries. Therefore, this study aims to contribute to this gap by using residential property markets in Kuala Lumpur, Malaysia, as a case study. Chau et al. [2003] provided major reviews of over 85 international studies on the effect of locational attributes such as air pollution, noise, view, zoning and neighbourhood. Reviews of previous work indicate that the location of a residential property is considered in terms of fixed and relative locational attributes. Fixed locational attributes capture the location of a residential property with respect to the whole urban area and is related to some form of distance and accessibility measures. Conversely, relative locational attributes are measures that reflect the externalities of the local neighbourhood and that are unique to an individual property such as the presence of amenities. Recently, the measurement of the effect of locational attributes on residential property values has been improved by considering the visibility of the amenity from the property, although in earlier studies distance and accessibility from locational attributes to the property were commonly used to measure their effect on value. Previous reviews have suggested that the effect of various types of locational attributes on residential property values is complex and has subtle variations in the estimation. The first study on the effect of locational attributes or specific environmental amenities on residential property values is that by Ridker and Henning [1967]. Using HPM in an attempt to estimate the effect of air pollution on property values in St. Louis, Missouri in the United States, they found a significant

negative relationship between air sulphation level and property values. Since then, many studies have been conducted to estimate the effect of locational attributes on residential property values. The empirical results indicate that most studies find the availability of public transport [Mulley, 2014; Dziauddin et al. 2015; Hess and Almeida, 2007], local schools quality [Gibbons and Machin, 2003; Mitchell, 2000], woodlands/urban forests [Powe et al. 1997; Tyrvaainen, 1997], green spaces [Vandergrift and Lahr, 2011; Jim and Chen, 2006], stadiums/sport facilities [Ahlfeldt and Kavetsos, 2014; Ahlfeldt and Maennig, 2008] and historically significant buildings [Lazrak et al. 2014] to have a positive and statistically significant effect on residential property values. For instance, Jim and Chen [2006] found that the view of green spaces and proximity to water bodies increased residential property values in Guangzhou and Shenzhen, China, by 7.1% and 13.2%, respectively. Other studies have also indicated that locational attributes such as mountains [Jim and Chen, 2009], lakes [Lansford and Jones, 1995] and coasts [Conroy and Millosch, 2011] have a positive and statistically significant effect on residential property values. A study by Conroy and Millosch [2011] in San Diego, California in the United States found that residential properties located within 500 feet from the coast sold 101.6% more residential properties than those located further away. However, a few studies found negative or no significant effect from locational attributes on residential property values due to environmental pollution and nuisance effects, including visual, aesthetic, noise, safety and traffic. For example, Simons et al. [2015] studied the effects on residential property values of refinery air pollution in Houston, Texas. They used hedonic pricing models to analyse sales of 3,946 homes from 2006 to 2011 based on proximity to the refinery and air pollution (sulphur dioxide – SO₂). Their analysis suggests residential properties within two miles radius from the refinery losses of 6-8% of the values. There have also been several studies on visual impacts on residential property values. For example, visual effects from wind farms are valued negatively by home buyers. A recent investigation by Gibbons [2014] in England, the United Kingdom, indicated that the price decreased by 5%–6% for residential properties located within 2 kilometres from wind farms (where wind farms are visible), by less than 2% for residential properties 2–4 kilometres away and by less than 1% for residential properties 14 kilometres away. Similar results were also observed for a busy road. A study conducted by Lake et al. [2000] in Glasgow, Scotland, based on over 3,500 property sales suggested that residential property values decreased by 0.2% for every 1 dB increase in noise. A recent study by Anderson et al. [2010] in Lerum, Sweden, suggested that an increase in noise pollution of 1 dB reduced the residential property values by over 2%.

Kuala Lumpur is not only the capital city of Malaysia but also the major financial and commercial centre and

home to hundreds of multinational companies. From 982,920 people in 1980, the population of Kuala Lumpur increased to 1.58 million people in 2010 (Department of Statistics Malaysia, 2010; 1980), with approximately 47% of the population in the 25–54 year old bracket and covering a total land area of 243 square kilometres. These characteristics have made Kuala Lumpur a densely populated city in Malaysia. Although Kuala Lumpur is host to 1.58 million permanent residents, the city's total population swells to more than 2.5 million people during the day as workers predominantly from the surrounding areas travel to the city for work. Among the population in Kuala Lumpur, 41% comes from the Malay ethnic group, 39% from the Chinese ethnic group, 9% from the Indian ethnic group, 1.6% from other ethnic groups and the rest are Non-Malaysian citizens [Department of Statistics Malaysia, 2010]. The gross domestic product per capita at purchasing power parity (PPP) of Kuala Lumpur increased to MYR 55,951 (USD 18,218) in 2010 from MYR 7,497 (USD 2,998) in 1980 [Department of Statistics Malaysia, 2010; 1980]. The labour force predominately consists of technicians, trade workers, labourers and clerical and administrative workers accounting for approximately 69%; professionals and managers only account for 18% of the workforce [Department of Statistics Malaysia, 2010].

The rapid growth of the population, employment, economic activities and services has resulted in a demand boom for housing. Kuala Lumpur's housing market is generally varied across geographical areas, with the highest priced housing found in the central and western areas (see Figure 1). The central and western parts of the city have larger and more expensive housing units than the southern and northern parts, although new development is occurring at the urban fringe in all directions. Note that the central and western areas of the city are home to major landmarks such as the Petronas Twin Towers; Malaysia's most famous department stores such as Suria KLCC Mall, Pavillion, Lot 10 and Star Hill and residences of Malaysia's royal families, ministers, middle and upper class families and expatriates (located in city centre, Bangsar, Dutamas, Mont Kiara, KL Sentral, Bukit Damansara, Taman Tun Dr. Ismail and Bukit Tunku).

With the exception of a slump during the Asian financial crisis in the late 1990s, the Kuala Lumpur housing market is generally vibrant. In recent years, residential property price appreciation has accelerated with strong price increases. For instance, in 2013, the residential property price appreciation increased to 16%. Specifically, Kuala Lumpur bungalows were in top demand as it surged by 22%, followed by high-rise buildings that increased by 17% and link residential properties that appreciated by 15%, which reflects the lack of new supply of residential properties in the capital [Department of Valuation and Property Services, 2013].

2. The Methods

Building on the traditional HPM, a spatial econometric method known as GWR is used to calibrate local regression parameters by weighting the distance between one data point and another through the data coordinates.

Residential properties are valued by their physical (structural), locational, neighbourhood and environmental attributes. Using the standard HPM, the contribution of these attributes to residential property values can be presented as follows:

$$P_i = f(S, L, N) + \epsilon_i \dots \dots \dots (1)$$

where P_i is the sale price of residential property i ; S is the vector of structural variables represented by floor size, number of bedrooms, type of house, etc.; L is the vector of locational variable effects measured as proximity to the CBD; N is the vector of the neighbourhood variables captured by the percentage of ethnic group; and ϵ_i is the vector of random error terms.

Since the first emergence of HPM in econometric literature, it has been considered to be the best method to estimate the effect of locational attributes on residential property values. Studies using HPM are too many to summarise in this section, but Malpezzi [2002], Chin and Chou [2003] and Sirmans and Macpherson [2003] provide a broad literature overview on the application of the method. HPM estimates the effect of locational attributes on residential property values by controlling the number of attributes in a multiple regression model with the dependent variable of property price [Mulley, 2014]. The housing market is an inherently dynamic and stochastic entity, in which the relationship between residential properties and locational attributes is the result of unobservable variation in the location across properties coupled with the heterogeneity of the housing markets, particularly in large housing markets [Tse, 2002]. Bitter et al. [2007] argued that changes in household preferences for physical, locational and neighbourhood attributes ‘may result in spatial mismatches between supply and demand as the housing stock available within a particular geographic area may not match current demand’. They asserted that ‘greater competition for those housing attributes that are in high demand, yet locally scarce, should result in higher marginal prices’. Therefore, supply and demand imbalances are expected to result in spatial heterogeneity in large housing markets. Based on this nature, producing an HPM is unlikely to sufficiently capture the spatial context and variation in which residential properties are located. The reason is that one of the limitations of the HPM is its insensitivity to deal with the existence of spatial effects in the residential property value–locational attribute relationship. Anselin [1988] found that the presence of spatial effects on the validity of traditional statistical

methodshas long been recognised as a potential problem. In addition, Fotheringham et al. [2002] acknowledged that the relationship between residential property values and locational attributes across geographical areas in a stationary fashion may not be representative of the situation in any particular part of the study area and may hide some interesting and important local differences in the determinants of residential property values. Therefore, as an alternative, the spatial econometric method of GWR developed by Fotheringham et al. [2002] is used to calibrate local regression parameters by weighting the distance between one data point and another through the data coordinates. This method enables the local parameters rather than the global parameters to be estimated, and thus it provides a way of accommodating the local geography of the residential property–locational attribute relationship. Although this study focuses primarily on the GWR method, it also recognises a few studies that have applied other spatial econometric methods, such as spatial expansion method and multilevel modelling, to examine the residential property price–locational attribute relationship [e.g. Jones and Bullen, 1993; Orford, 1999]. GWR is a model in which local variations in the parameter estimates are included using the coordinates of the variable. Thus, by including longitude and latitude coordinates (u_i, v_i) to Eq. (1) above, the general form of the HPM can be mathematically expressed at location i in space as follows [Crespo and Grêt-Regamey 2013]:

$$P_i(u_i, v_i) = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \epsilon_i, \quad i = 1, \dots, n, \quad (2)$$

- where
- P_i = response variable at point i ,
- u_i, v_i = spatial coordinates of point i ,
- $\beta_0(u_i, v_i)$ = location-specific intercept term parameter,
- $\beta_k(u_i, v_i)$ = k th location-specific parameter,
- p = number of unknown local parameters to be estimated (excluding the intercept term),
- x_{ik} = k th explanatory variable associated with β_k ,
- ϵ_i = random component assumed to be independently and identically distributed,
- n = number of observations.

According to Eq. 2, the location-specific parameters $\beta_k(u_i, v_i)$ are estimated using weighted least squares and can be expressed as follows [Crespo and Grêt-Regamey, 2013]:

$$\beta(u_i, v_i) = [XTWiX]^{-1}XTWi, \quad i = 1, \dots, n, \dots (3)$$

- where
- $\beta(u_i, v_i)$ = ($p \times 1$) vector parameter estimate at location i ,
- X = ($n \times p$) matrix of the observed explanatory variables,
- W_i = distance decay ($n \times n$) matrix
- p = ($n \times 1$) vector of the observed response variables.

Note that p and i are as defined in Eq. 2. Location i is also denoted as the regression point, the point at

which the parameters are estimated. As expressed in the equation above, the weighting of an observation is conducted through a distance decay matrix (W_i), so that observations located near the point in space are weighted more than the observations located further away. Through this geographically weighted calibration, continuous and smooth surfaces of local parameter estimates can be mapped across the geographical area.

Data Acquisition

Residential property data The residential property price transaction data used in this study were obtained from the sale of landed residential properties in the Kuala Lumpur market in 2010. The geographical area of the study is shown in Fig. 1. In total, 4,393 units of selling prices together with their structural attributes were collected. The data were collected from the Department of Valuation and Services, Malaysia

(Kuala Lumpur branch). Ownership status (freehold/ leasehold) of the residential properties, floor size, lot size and number of bedrooms were the structural attributes of the residential properties obtained from the data provider and used for the analysis.

Proximity to Locational Attributes Data To measure the distance for a given observation to the locational attributes, a geographical information system (GIS) was used to position each observation accurately on a local map using geographical coordinates (latitude and longitude) obtained from Google Maps. GIS and spatial analysis were integrated into this study, and the integration was particularly useful because the proximity from observations to the locational attributes was measured accurately using network distance. The distance in metres was measured along the street network using a GIS programme called Multiple Origins to Multiple Destinations obtained

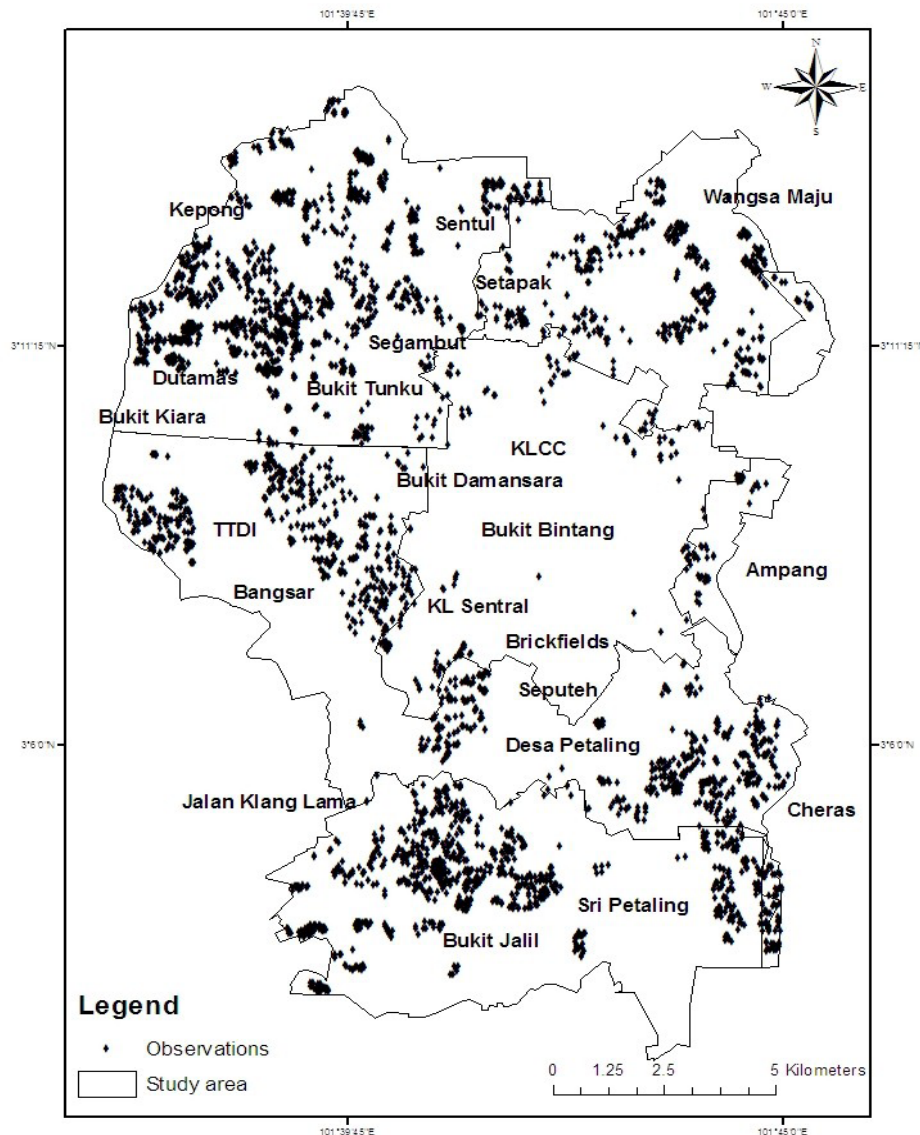


Figure 1. Map of study area and spatial distribution of observations

from the Environmental Systems Research Institute support centre. The network–distance measurement using this programme requires three layers of spatial data: points of origin (observations), points of destinations (locational attributes) and road network data. This measurement allows the shortest route from each observation to the locational attributes to be calculated. Moreover, the Multiple Origins to Multiple Destinations programme enables more than one destination to be selected at any one time. Thus, proximity to the locational attributes can be calculated simultaneously for each observation. Table 1 presents a list of explanatory variables considered for inclusion in the HPM together with their descriptive statistics.

In all regression-based analysis, some explanatory variables are usually multicollinear. To handle this problem, the correlations among the explanatory variables used for the inclusion in the final models were detected by using Pearson’s correlation coefficient and variance inflation factors (VIFs). Following Orford [1999], a Pearson’s correlation coefficient above 0.8 and a VIF above 10 indicate harmful collinearity. This rule was applied in this study.

3. Results and Discussion

The results of the HPM and GWR models are presented in two stages: the first part shows the results from the HPM and the second part shows the results from the GWR model.

The HPM Estimation

According to the seminal works in the area of hedonic pricing of Rosen [1974], Freeman [1979b], Halvorsen and Pollawski [1981] and Palmquist [1984], four common functional forms are used to measure the relationship between the dependent and the explanatory variables in HPM: linear, semi-log, double-log and Box-Cox transformations. Unfortunately, no studies have been conducted on how to choose a particular functional form in HPM except for the work of Cropper et al. [1988], which extensively reviewed the choice of functional form for HPM by previous researchers. This study concluded that linear, semi-log, double-log and Box-Cox linear perform best with quadratic forms, with the quadratic Box-Cox faring relatively poorly. After performing and comparing these functional forms for the data in this study, we found that the semi-log functional form performed the best, and thus we used it estimate the relationship being studied. The semi-log regression equation is regressed on a set of determinants as follows:

$$\ln P_i = \beta_0 + \beta_1 \text{FREEHOLD}_i + \beta_2 \text{FLOORSIZE}_i + \beta_3 \text{LOTSIZE}_i + \beta_4 \text{BEDROOMS}_i + \beta_5 \text{TYPE_TERRACED}_i + \beta_6 \text{TYPE_TOWNHOUSE}_i + \beta_7 \text{TYPE_SEMIDETACHED}_i + \beta_8 \text{TYPE_BUNGALOW}_i + \beta_9 \text{PRIMARY_SCHOOL}_i + \beta_{10} \text{SECONDARY_SCHOOL}_i + \beta_{11} \text{SHOPMALL}_i + \beta_{12} \text{FOREST}_i + \beta_{13} \text{RAIL_STATION}_i + \beta_{14} \text{MALAY}_i + \beta_{15} \text{CHINESE}_i + \epsilon_i \dots (4)$$

Table 1. A list of explanatory variables and descriptive statistics of the model’s variables

Variable	Hedonic price model (HPM)			GWR model		
	Coefficients (β)	Standardised coefficients (β)	t-ratio	Lower quartile	Median	Upper quartile
Intercept	11.647	0.159	183.944	11.080	11.985	12.328
FREEHOLD	0.294	0.542	18.733	0.084	0.155	0.232
FLOORSIZE	0.005	0.086	49.578	0.004	0.005	0.005
LOTSIZE	1.860E-4	0.108	9.593	4.800E-4	4.030E-4	0.001
BEDROOMS	0.190	-0.043	11.497	0.076	0.132	0.173
TYPE_TERRACED	-0.084	-0.062	-2.841	0.038	0.290	0.391
TYPE_TOWNHOUSE	-0.223	0.124	-5.645	-0.210	0.024	0.179
TYPE_SEMIDETACHED	0.357	0.096	10.454	0.338	0.571	0.635
TYPE_BUNGALOW	0.442	0.031	9.869	0.182	0.361	0.497
PRIMARY_SCHOOL	2.832E-5	0.070	2.783	2.700E-5	1.400E-5	8.500E-5
SECONDARY_SCHOOL	5.244E-5	-0.112	6.881	9.000E-6	3.900E-5	1.070E-4
SHOPMALL	-8.112E-5	0.052	-11.385	-8.000E-5	-5.600E-5	-4.000E-6
FOREST	-2.973E-5	-0.018	-5.934	-2.200E-5	-4.000E-6	8.000E-6
RAIL_STATION	-0.036	0.117	-2.132	-0.073	-0.019	0.030
MALAY	0.041	0.067	12.629	0.006	0.033	0.070
CHINESE	0.025		7.416	-0.037	-0.002	0.025

Notes: Goodness of fit: Adjusted R2=0.73 (HPM); adjusted R2=0.85 (GWR). AICc=5663.63 (HPM); AICc=3185.55 (GWR)

where i is the subscript denoting each property, P_i is the price of property i in Malaysian Ringgit (MYR), \ln is the natural logarithm; FREEHOLD is the dummy variable that indicates the status of ownership of the property, FLOORSIZE is the floor area of the property measured in square metre, LOTSIZE is the lot size of the property measured in square metre, BEDROOMS is the dummy variable that indicates a property with four or more bedrooms, TYP xxx is the set of dummy variables that illustrates the type of house, and it is further described as follows:

TYPE_TERRACED is 1 if the property is cluster, 0 otherwise;
 TYPE_TOWNHOUSE is 1 if the property is town house, 0 otherwise;
 TYPE_SEMIDETACHED is 1 if the property is semi-detached, 0 otherwise;
 TYPE_BUNGALOW is 1 if the property is bungalow, 0 otherwise.

PRIMARY_SCHOOL, SECONDARY_SCHOOL, SHOPMALL and FOREST are the network–distance variables measured in metres from the property to primary schools, secondary schools, shopping malls and forests, respectively. RAIL_STATION is the dummy variable that indicates properties within a 2 km radius to the nearest station. The 2 km radius is chosen

in accordance with the theory of the rail transit system–residential property value relationship. MALAY and CHINESE are the percentages of residents from the Malay and Chinese ethnic groups. $\beta_0, \dots, \beta_{15}$ denotes a set of parameters to be estimated on the explanatory variables (including the intercept term), and ϵ_i is the standard error of the estimation assumed to be independently and identically distributed.

The parameter estimates for HPM together with the GWR model are shown in Table 2. In general, HPM explains approximately 73% of the variation in the dependent variable, consistent with other published empirical results. In the final model, all of the explanatory variables that influenced residential property values were significant at the 1% level.

Note that explanatory variables were incorporated in the final model based on significant coefficient values, and they alleviated the potential issues of multicollinearity. As the model is a semi-log functional form, the interpretation of the parameter estimates relates to their proportional (or their percentage when multiplied by 100) effect on price. Among the structural attributes, property size, which is measured by floor size (FLOORSIZE), is the most significant contribution to the value formation of a residential property. Specifically, for every square metre increase in floor size, the expected selling price of a residential

Table 2. Estimation results of the global (HPM) and local (GWR) models (n=4393)

Variable	Hedonic price model (HPM)			GWR model		
	Coefficients (β)	Standardised coefficients (β)	t-ratio	Lower quartile	Median	Upper quartile
Intercept	11.647		183.944	11.080	11.985	12.328
FREEHOLD	0.294	0.159	18.733	0.084	0.155	0.232
FLOORSIZE	0.005	0.542	49.578	0.004	0.005	0.005
LOTSIZE	1.860E-4	0.086	9.593	4.800E-4	4.030E-4	0.001
BEDROOMS	0.190	0.108	11.497	0.076	0.132	0.173
TYPE_TERRACED	-0.084	-0.043	-2.841	0.038	0.290	0.391
TYPE_TOWNHOUSE	-0.223	-0.062	-5.645	-0.210	0.024	0.179
TYPE_SEMIDETACHED	0.357	0.124	10.454	0.338	0.571	0.635
TYPE_BUNGALOW	0.442	0.096	9.869	0.182	0.361	0.497
PRIMARY_SCHOOL	2.832E-5	0.031	2.783	2.700E-5	1.400E-5	8.500E-5
SECONDARY_SCHOOL	5.244E-5	0.070	6.881	9.000E-6	3.900E-5	1.070E-4
SHOPMALL	-8.112E-5	-0.112	-11.385	-8.000E-5	-5.600E-5	-4.000E-6
FOREST	-2.973E-5	0.052	-5.934	-2.200E-5	-4.000E-6	8.000E-6
RAIL_STATION	-0.036	-0.018	-2.132	-0.073	-0.019	0.030
MALAY	0.041	0.117	12.629	0.006	0.033	0.070
CHINESE	0.025	0.067	7.416	-0.037	-0.002	0.025

Notes: Goodness of fit: Adjusted R²=0.73 (HPM); adjusted R²=0.85 (GWR). AICc=5663.63 (HPM); AICc=3185.55 (GWR)

property increases by 0.5% (at the mean value, this equates to a premium of MYR 4331 or USD 1379, with the FOREX rate at MYR 3.14 or USD 1.00 in December 2010). Among the locational attributes, distance to a local shopping mall is the most statistically significant. The model suggests that for every metre away from a local shopping mall, residential property values are likely to decrease at approximately 0.008% (at the mean value, this equates to MYR 70 or USD 22), thus signifying a strong appreciation of house buyers for this locational attribute category. The distance to rail transit stations (RAIL_STATION) is the least statistically significant locational attribute, and it has a counter-intuitive sign. Residential properties are estimated to be sold at discounts of up to 3.6%, which is equal to MYR 31,181 (USD 9930) at the mean value, in a radius of 2 km to the nearest station. Among neighbourhood variables, the percentage of residents from the Malay ethnic group contributes most significantly to the value formation of a residential property. For every 1% increase in ethnic Malay (MALAY), residential property values are likely to increase by about 4.1%, which is equal to MYR 35,512 (USD 11,309). In all regression-based analyses, the independent

variables possess various units of measurements, and thus their relative importance to the dependent variable is difficult to assess. To help overcome this difficulty, the computer-intensive LMG method developed by Lindeman et al. [1980] can be used to decompose R2 [Grömping, 2006]. The LMG method can be written in the following form [Christensen, 1992]:

$$LMG(x_k) = \frac{1}{p} \sum_{j=0}^{p-1} \left(\sum_{S \subseteq \{x_1, \dots, x_p\}; n(S)=j} \frac{seqR^2(\{x_k\} | S_k(r))}{\binom{p-1}{j}} \right)$$

where n(S) is the number of variables included in the set of regressors S. Therefore, LMG is defined as the average contributions in models of different sizes. The order of the regressors in any model is a permutation of available regressors x_1, \dots, x_p , which is denoted by the tuple of indices r .

When the unit of measurement for independent variables is standardised, the relative importance of these variables on the dependent variable can be determined. The results from the LMG method indicate that structural variables are more influential than locational and neighbourhood variables in estimating residential property values. Specifically, the structural attributes

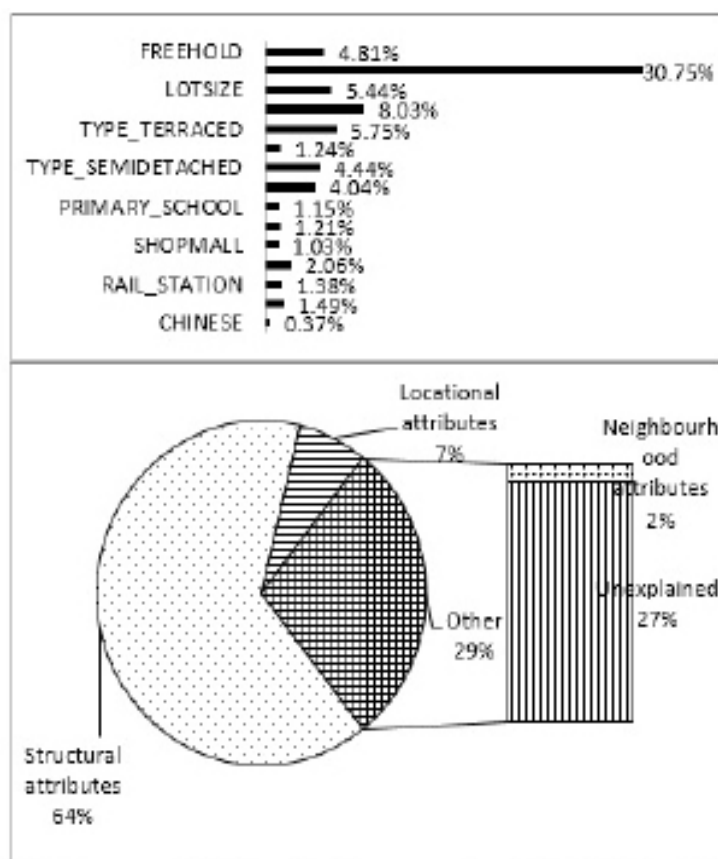


Figure 2. Shares of relative importance (LMG) for residential property values, OLS-Model (Note: The R2 with significant variables is 73%)

explain 64% of the variation in the residential property value, and the locational and neighbourhood attributes explain only 9%. Among the structural attributes, floor size (FLOORSIZE) and a property with four or more bedrooms (BEDROOMS) are the most influential variables in the value formation of a residential property. For example, floor size and a property with four or more bedrooms increase the property value by approximately 31% and 8%, respectively. Among the locational attributes, the existence of a forest within the vicinity is the most influential variable in the value formation of a residential property, as proximity to a forest increases the property value by approximately 2%. This finding contradicts the previous result before the standardisation was made, i.e. proximity to a shopping mall is found to be the most influential variable.

GWR Local Model Estimation

Table 2 presents the summary of the parameter estimates obtained from the HPM in a single value and the parameter estimates obtained from the GWR in multiple values (lower quartile, median quartiles and upper quartile). Clearly, the values of the parameter estimates are not homogeneous but vary for all of the explanatory variables. The summary results of the GWR in Table 2 also indicate that the GWR performs better than the HPM, as the adjusted R² improves from 73% to 85% in explaining the variation in the dependent variable and has a lower Akaike Information Criterion (from 5663.63 down to 3185.55). More importantly, a key advantage of GWR is the ability to visualise and explore the spatial variability of explanatory variables in the model, in which the parameter estimates of explanatory variables may vary significantly over a geographical area. For instance, some explanatory variables may exhibit counter-intuitive signs in the HPM but later reveal statistically significant local parameter estimates with an anticipated sign in the GWR modelling. Fig. 3a-e maps the local regression results for the locational attributes variables. Although all the local parameter estimates can be mapped, only these results are mapped because of this study's focus on the effect of locational attributes on residential property values and the space limitation. The best interpretation comes from the maps of the local parameter estimates alongside the maps of the local t-ratio, as they exhibit the local significance that accounts for the local varying estimate errors [Crespo and Grêt-Regamey, 2013]. Figs. 3a-e show the spatial variation over a geographical area of a premium on the residential property values provided by the locational attributes. In these five figures, the local parameter estimates are shown as different colour points that indicate the positive and negative premiums added to residential property values. With the exception of proximity to rail transit stations (RAIL_STATION), the negative coefficient values (red colour points) for the PRIMARY_SCHOOL, SECONDARY_SCHOOL, SHOPPING_MALL and FOREST

variables should be interpreted as positive premiums, whereas positive coefficient values (blue colour points) should be interpreted as negative premiums. While assessing the residential property value in relation to the existence of a primary school (PRIMARY_SCHOOL), the global model estimates that the residential property value increases by 0.002% for every metre away from a primary school, and this single value is applied to the entire area. An examination of Fig. 3a shows that both the t-ratio and the parameter estimates exhibit a spatial variation across a geographical area that cannot be seen in a global model. For instance, in the South West area, a positive effect is observed where it contributes to 0.008%–0.02%. At the mean value, this value is equal to a MYR 69–MYR 173 or USD 29–USD 55 reduction for every metre away from a primary school. Bangsar, the southwestern area of Kuala Lumpur, is an affluent area dominated by expatriates and high-income dwellers. Typically, the high-performing schools are considered as amenities sought by the community, as this segment of society does not simply send their children to any school. Moreover, the locals and expatriates who make up the affluent segment of this area place top priority in the safety of their children, and thus they prefer and appreciate primary schools that are near their homes. Although not as significant as the southwestern area, the southern and south-eastern areas also yield a positive result at 0.0006%–0.008%, which is equal to MYR 5–MYR 69 (USD 1.66–USD 29). As these two areas are populated by households with young schoolchildren, the primary school is considered an essential amenity and presents an opportunity for the residential property value to be positively affected. The northwest area of Kepong is not affected. For Malaysian parents, primary school education is delivered in a standard manner regardless of which school their children attend. Children are sent to any primary school for fundamental education. Another area of concern is Kepong, which is populated by Chinese residents who send their children to Chinese schools within the area or alternatively to the better established ones that are usually located in the city centre.

For SECONDARY_SCHOOL, the global model estimates that the residential property value increases by 0.005% for every metre away from a secondary school. Once again, an examination of Fig. 3b shows that the secondary school effect on residential property value varies by location, as indicated by the t-ratio and the parameter estimates. As shown by the parameter estimates in Fig. 3b, some residential properties located in the southern and north-western areas are statistically significant in relation to the existence of secondary schools, with positive residential property premiums of up to 0.001%–0.008% for every metre away from a secondary school. At the mean value, this value is equal to MYR 8–MYR 69 or USD 3–USD 22. The southern (Bukit Jalil) and north-western (Kepong) areas manifest their appreciation in the presence of secondary schools

within the localities. The only National Sports School, which is attended by young sporting talents with high potential, is located in Bukit Jalil. Although no direct connection exists between the appreciation of house prices and this educational institution, the school is an integral part of a multi-million national sports complex that also houses the main national stadium, Stadium Bukit Jalil. It is a bustling complex where many major sporting and international events are held regularly. Overall, the features of the complex contribute to the increase in property value. Interestingly, Kepong house prices have recorded no appreciation in primary schools as discussed previously. However, the opposite result is observed in the secondary school when buyers put more value in having a secondary school closer to their houses. Malaysian parents are generally more concerned with secondary school as that is when all major examinations take place. Additionally, having a secondary school closer to home minimises the time for teenagers to travel from their houses to school. This situation in turn minimises the risk of the children being involved in time-wasting activities away from the house, or worse, getting involved in juvenile crimes

Globally, a residential property value is estimated to decrease by 0.008% for every metre away from the nearest local shopping mall. However, by mapping the t-ratio and the parameter estimates from a local model, we can observe that this variable appears to exhibit a significant spatial variation (see Figure. 3c). Proximity to the nearest local shopping mall (SHOPMALL) shows a clustering of significantly negative values for properties located in the western, north-western, northern and northeast areas, which have statistically significant premiums of 0.003%–0.01% for every metre away from a local shopping mall. At the mean value, this value is equal to a premium of MYR 26–MYR 87 or USD 8–USD 28. The northwest section has the most significant effect in the appreciation of house prices in terms of a shopping mall. A large shopping mall called One Utama is located in Bandar Utama (i.e. Major Urban). The area of Bandar Utama now has key shopping outlets such as IKEA and TESCO, which directly contribute to the increased property value for these areas. Aside from these areas, other housing areas in the north-eastern part of Kuala Lumpur, i.e. Wangsa Maju, have a positive effect, although it is not as high as that in the Bandar Utama or Damansara areas. Therefore, the residents of the mentioned areas look for a local mall that is closer to their homes to minimise travel time.

Another influential locational attribute is proximity to forest areas (FOREST). In the global model, the residential property value is estimated to decrease by 0.003% for every metre away from the nearest forest. Again, an examination of Figure. 3d shows that the forest effect on residential property value varies across a geographical area, as indicated by the t-ratio and the parameter estimates. The map shows significantly negative values for properties located in the centre,

north and northwest, where statistically significant premiums of 0.002%–0.008% are recorded for every metre away from the forest. At the mean value, this value is equal to a premium of MYR 17–MYR 69 or USD 5–USD 22. The central area of Kuala Lumpur or Segambut has the largest effect in terms of forest as a parameter. This hilly area has been known for its greenery since its establishment, and the green surroundings are still currently intact. Affluent house buyers who have realised that the forest will be kept naturally are encouraged to own houses in this area. Moreover, the area only has a few new housing projects, and thus it is one of the few green lungs in Kuala Lumpur. Cheras (southeast area) has many forest areas, but its appreciation is not significant. Cheras is known as a less affluent area in Kuala Lumpur compared with other areas. Thus, the demographic view of homeowners or homebuyers here is that of the working class that tends to buy apartments in the area and has less appreciation for a green forest.

The remaining locational attribute is properties within a 2 km radius to the nearest rail transit station. In the global model, residential properties within a 2 km radius to the nearest rail transit station are sold at discounts of up to 3.6%. However, an examination of Figure. 3e suggests a positive relationship between the existence of a rail transit system and residential property values in the north-eastern and eastern areas, where statistically significant premiums of 1%–15.7% are recorded for properties within a 2 km radius to the nearest rail station. At the mean value, this value is equal to a premium of MYR 8,600–MYR 136,000 or USD 2700–USD 43,000. Housing areas along the train lines are generally positively affected by their closeness to train stations. The three areas observed to have significant effects are Cheras, Ampang and Wangsa Maju. These areas are served by two LRT lines, namely, STAR and Kelana Jaya, the lines that predominantly connect the major housing areas to the city centre of Kuala Lumpur. The north-western or Kepong area is the next area to be positively affected. This area is only served by a commuter line. Unlike the LRT, the commuter line connects only the secondary housing areas outside Kuala Lumpur, and thus it is not as efficient as the LRT lines. However, as it is the only train mode that is available in Kepong, assuming that the property values around the Kepong area increase due to the availability of the only train service into Kuala Lumpur is fair.

3. Conclusion

This study presents several important observations on the effects of locational attributes on the residential property value. Generally, the estimation results appear to suggest that residential property values are mainly determined by the property's physical (structural) attributes, consistent with other reported research estimating the residential property value uplift. More importantly, we find strong evidence of a spatial distribution of residential property value

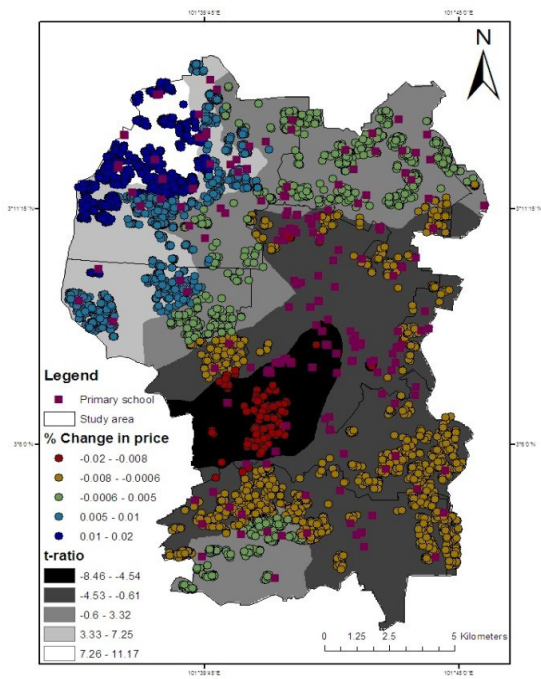


Figure 3a. Map of the local model estimates of the proximity to primary school

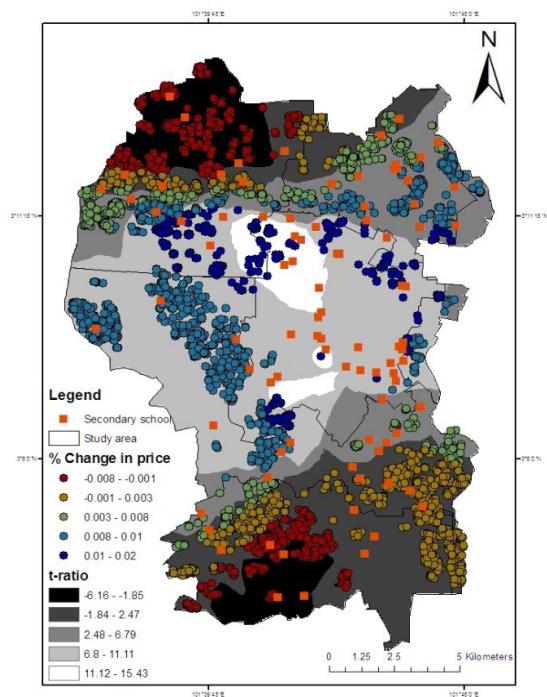


Figure 3b. Map of the local model estimates of the proximity to secondary school

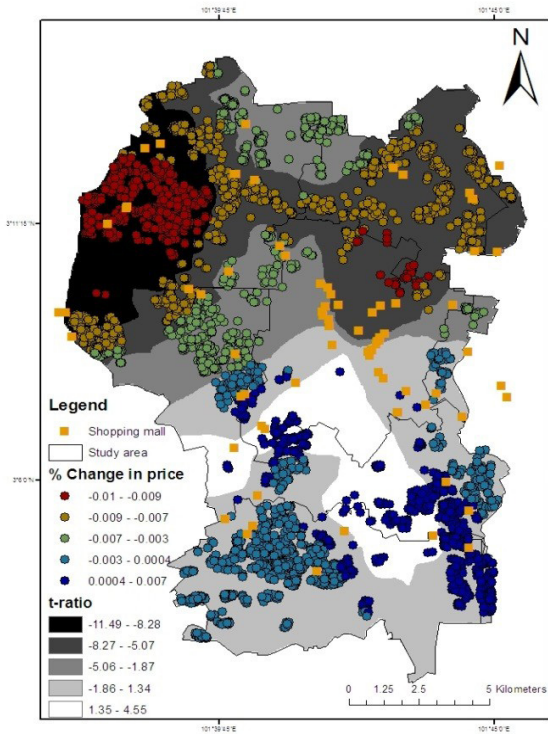


Figure 3c. Map of the local model estimates of the proximity to shopping mall

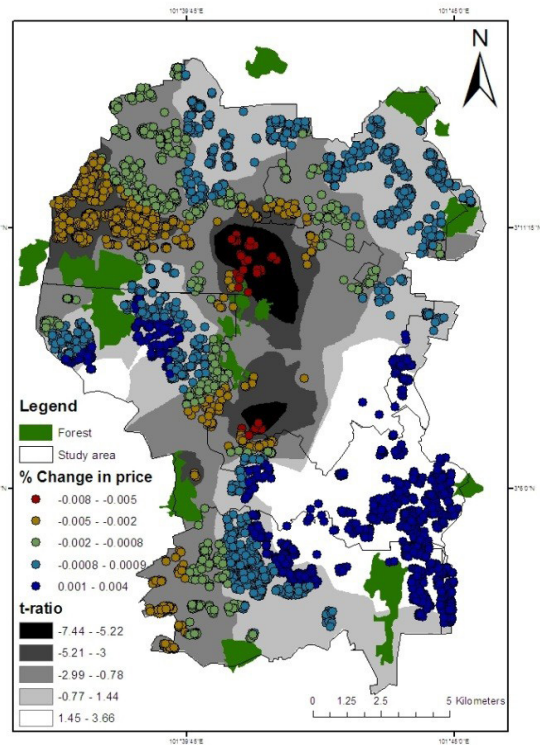


Figure 3d. Map of the local model estimates of the proximity to forest

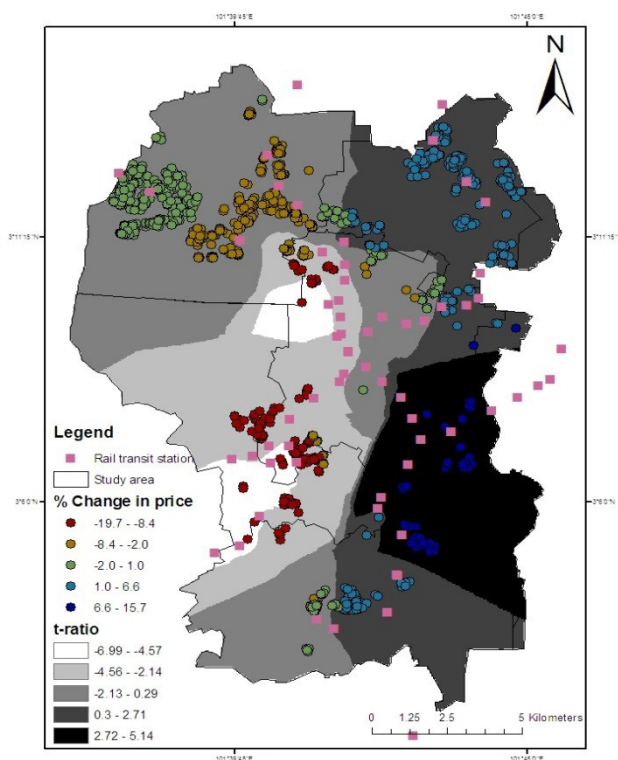


Figure 3e. Map of the local model estimates of the proximity to rail transit station

formation in relation to proximity to locational attributes. The use of the GWR specification of the hedonic model enables this study to successfully show distinct variations of premiums across the study area.

The study shows that different locational attributes have differential spatial effects on residential property premiums across the study area. The results indicate that the positive externalities generated by a shopping mall and a forest, for instance, have a greater effect than the negative externalities on residential property premiums for the majority of residential properties. On the contrary, the positive externalities generated by the proximity to primary and secondary schools and a rail transit station have a more marginal effect than the negative externalities on the residential property premiums for the majority of residential properties.

Subsequently, the spatial variation in property value can be used as a policy determinant in relation to the tax structure for specific households, i.e. houses that benefit more from an amenity (a higher tax is imposed on these homes than on those that are not highly affected). This fair taxation valuation generates more income for the local councils. Therefore, the study on the outcome of the effect of locational attributes on residential property values is highly useful and relevant. However, the entire process is complex. The estimation of this nature is not a trivial exercise. It requires a clear theoretical understanding of how location interacts with

residential property value, a data-rich environment and a method that enables these data to be modelled, and it takes into account the spatial context in which residential properties can be found. As we strive for a more credible interpretation of GWR results, future studies are suggested to use the qualitative approach, such as a focus group of affected areas, to reveal all qualitative explanations on the subjects within the area.

References

- Ahlfeldt G. M. and Kavetsos G. (2014). Form or Function? The Impact of New Sports Stadia on Property Prices in London, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177 (1), 169-190.
- Ahlfeldt, G. M. and Maennig, W. (2008). Impact of Sports Arenas on Land Values: Evidence from Berlin, *The Annals of Regional Science*, 44 (2), 205-227.
- Alonso, W. (1964). *Location and Land Use: Towards a General Theory of Land Rent*. Cambridge, MA: Harvard University Press.
- Anderson, H., Jonsson, L. and Ögren, M. (2010). Property Prices and Exposure to Multiple Noise Sources: Hedonic Regression with Road and Railway Noise, *Environmental and Resource Economics*, 45 (1), 73-89.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic.
- Bitter, C., Mulligan, G., Dall'erba, S. (2007). Incorporating Spatial Variation in Housing Attribute Prices: A Comparison of Geographically Weighted Regression and the Spatial Expansion Method, *Journal of Geographical Systems*, 9, 7-27.

- Chau, K.W., Yiu, C.Y., Wong, S.K. and Lai, L.W.C. (2003b). Hedonic Price Modelling of Environmental Attributes: A Review of the Literature and A Hong Kong Case Study, in L.W.C. Lai and F.T. Lorne (eds.) *Understanding and Implementing Sustainable Development*, New York: Nova Science, 87-110.
- Chin, T.L. and Chau, K.W. (2003). A Critical Review of Literature on the Hedonic Price Model, *International Journal for Housing Science and Its Applications*, 27 (2), 145-165.
- Conroy, S. J. and Milosch, J. L. (2011). An Estimation of the Coastal Premium for Residential Housing Prices in San Diego County, *The Journal of Real Estate Finance and Economics*, 42, 211-228.
- Crespo, R. and Grêt-Regamey, A. (2013). Local Hedonic House-Price Modelling for Urban Planners: Advantages of Using Local Regression Techniques, *Environment and Planning B: Planning and Design*, 40 (4), 664 - 682.
- Cropper, M.L., Deck, L.B. and McConnell, K.E. (1988). On the Choice of Functional Form for Hedonic Price Functions, *The Review of Economics and Statistics*, 70 (4), 668-675.
- Department of Statistics Malaysia (1980). *The Malaysia Population and Housing Census*. Kuala Lumpur: Department of Statistics Malaysia.
- Department of Statistics Malaysia (2010). *The Malaysia Population and Housing Census*. Kuala Lumpur: Department of Statistics Malaysia.
- Department of Valuation and Property Services (2013). *Property Market Report*. Putrajaya.
- Du, H. and Mulley, C. (2006). Relationship between Transport Accessibility and Land Value: Local Model Approach with Geographically Weighted Regression, *Transportation Research Record*, 1977 (1), 197-205.
- Dziauddin, M. F., Powe, N. and Alvanides, S. (2015). Estimating the Effects of Light Rail Transit (LRT) System on Residential Property Values Using Geographically Weighted Regression (GWR), *Applied Spatial Analysis and Policy*, 8, 1-25.
- Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. England: John Wiley & Sons Ltd.
- Freeman, A.M.III (1979a). The Hedonic Price Approach to Measuring Demand for Neighbourhood Characteristics, In D. Segal (ed.), *The Economics of Neighbourhood*. New York: Academic Press, 193-217.
- Freeman, A.M.III (1979b). Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues, *Scandinavian Journal of Economics*, 81, 154-173.
- Gibbons, S. (2014). *Gone with the Wind: Valuing the Visual Impact of the Wind Turbines Through House Prices*. Serc paper discussion: London School of Economics and Political Sciences & Spatial Economics Research Centre.
- Gibbons, S. and Machin, S. (2003). Valuing English Primary Schools, *Journal of Urban Economics*, 53, 197-219.
- Grömping, U. (2006). Relative Importance for Linear Regression in R: The Package Relaimpo, *Journal of Statistical Software*, 17, 1-27.
- Halvorsen, R. and Pollakowski, H. (1981). Choice of Functional Form for Hedonic Price Equations, *Journal of Urban Economics*, 10, 37-49.
- Hess, D.B. and Almeida, T.M. (2007). Impact of Proximity to Light Rail Rapid Transit on Station-Area Property Values in Buffalo, New York, *Urban Studies*, 44 (5-6), 1041-1068.
- Jim, C.Y. and Chen, W.Y. (2006b). Recreation-Amenity Use and Contingent Valuation of Urban Green Spaces in Guangzhou, China, *Landscape and Urban Planning*, 75 (25), 81-96.
- Jim, C.Y. and Chen, W.Y. (2009). Value of Scenic Views: Hedonic Assessment of Private Housing in Hong Kong, *Journal of Landscape and Urban Planning*, 91, 1-9.
- Jones, K. and Bullen, N. (1993). A Multilevel Analysis of the Variations in Domestic Property Prices: Southern England 1980-1987, *Urban Studies*, 30 (8), 1409- 1426.
- Lake, I. R., Bateman, I. J., Day, B. H. and Lovett, A. A. (2000). Assigning a Monetary Value to Noise Reduction Benefits: An Example From The UK, [online], Available at <http://www.iccr-international.org/trans-talk/docs/ws2-lake.pdf>
- Lansford, N.H., and Jones, L.L. (1995). Recreational and Aesthetic Value of Water Using Hedonic Price Analysis, *Journal of Agricultural and Resource Economics*, 20 (2), 341-355.
- Lazrak, F., Nijkamp, P., Rietveld, P. and Rouwendal, J. (2014). The Market Value of Cultural Heritage in Urban Areas: An Application of Spatial Hedonic Pricing, *Journal of Geographical Systems*, 16, 89-114.
- Lindeman, R. H., Merenda, P. F. and Gold, R. Z. (1980). *Introduction to Bivariate and Multivariate Analysis*. Scott Foresman, Glenview, IL.
- Malpezzi, S. (2002). Hedonic Pricing Models and House Price Indexes: A Select Review. In Kenneth Gibb and Anthony O'Sullivan (eds.), *Housing Economics and Public Policy: Essays in Honour of Duncan MacLennan*. Oxford: Blackwell Publishing, 67-89.
- Mills, E. S. (1972). *Urban Economics*. Illinois: Scott, Foresman and Company.
- Mitchell, D. (2000). *Cultural Geography: A Critical Introduction*. Oxford: Blackwell.
- Mulley, C. (2014). Accessibility and Residential Land Value Uplift: Identifying Spatial Variations in the Accessibility Impacts of A Bus Transitway, *Urban Studies*, 51 (8), 1707-24.
- Muth, R. F. (1969). *Cities and housing: The Spatial Pattern of Urban Residential Land Use*. Chicago: The University of Chicago Press.
- Orford, S. (1999). *Valuing The Built Environment: GIS and House Price Analysis*. England: Ashgate Publishing Ltd.
- Palmquist, R.B. (1984). Estimating the Demand for the Characteristics of Housing, *Review of Economics and Statistics*, 66, 394-404.
- Powe, N.A., Garrod, G.D., Brunsdon, C.F. and Willis, K.G. (1997). Using A Geographic Information System to Estimate an Hedonic Price Model of The Benefits of Woodland Access, *Forestry*, 70 (2), 139-149.
- Ridker, R.G. and Henning, J.A. (1968). The Determination of Residential Property Values with Special Reference to Air Pollution, *Review of Economics and Statistics*, 49, 246-257.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competitions, *Journal of Political Economy*, 72, 34-55.
- Simons, R.A., Seo, Y. and Rosenfeld, P. (2015). Modelling the Effects of Refinery Emissions on Residential Property Values, 37 (3), 321-342.
- Sirmans, G.S., and Macpherson, D. (2003). The State of Affordable Housing, *Journal of Real Estate Literature*, 11 (2), 131-156.
- Tse, R.Y.C. (2002). Estimating Neighbourhood Effects in House Prices: Towards A New Hedonic Model Approach, *Urban Studies*, 39 (7), 1165-1180.
- Tyrvaïnen, L. (1997). The Amenity Value of the Urban Forest: An Application of the Hedonic Pricing Method, *Landscape and Urban Planning*, 37 (3), 211-222.
- Vandegrift, D. and Lahr, M. (2011). Open

Space, House Prices, and the Tax Base, The
Annals of Regional Science, 46 (1), 83-100.