



ISSN 1732-4254 quarterly

BULLETIN OF GEOGRAPHY. SOCIO-ECONOMIC SERIES

journal homepages:

<http://www.bulletinofgeography.umk.pl/>
<http://wydawnictwoumk.pl/czasopisma/index.php/BGSS/index>
<http://www.degruyter.com/view/j/bog>

DE
G

Spatial heterogeneity and spatial bias analyses in hedonic price models: some practical considerations

Haniza Khalid^{CDFMR}

International Islamic University Malaysia, Faculty of Economics and Management Sciences, P.O. Box 10, 50728 Kuala Lumpur, Malaysia; e-mail: hanizamv@iium.edu.my

How to Cite:

Khalid, H., 2015: Spatial heterogeneity and spatial bias analyses in hedonic price models: some practical considerations. In: Szymańska, D. and Chodkowska-Miszczuk, J. editors, *Bulletin of Geography. Socio-economic Series*, No. 28, Toruń: Nicolaus Copernicus University, pp. 113–129. DOI: <http://dx.doi.org/10.1515/bog-2015-0019>

Abstract. A great number of contemporary studies are incorporating explicit consideration of spatial effects in the estimation of hedonic price functions. At the most basic level, interactive spatial regime models are employed to detect the presence of spatial heterogeneity in datasets. A full-scale spatial analysis would include determination and adjustments for spatial lag and spatial error dependences. However, there is still plenty of room for future research to help unravel the numerous modelling and practical issues associated with a comprehensive spatial examination, such as the specification of the spatial dependence structure or functional ‘neighbourhoods’. Another important issue relates to the use of spatial multipliers to filter spatial bias particularly in models which use log-transformed variables. Estimation of a hedonic price function using Malaysian dataset of agricultural land sale values indicates spatial disaggregation and spatial dependence. However, diagnostic tests and actual estimation of spatial models do not always provide unambiguous conclusions while predicted errors do not vary all that much from those generated by simpler models. Despite the conceptual appeal of spatial analyses, the inefficiency attributable to spatial biases may not be large enough to cause critical errors in policy decisions.

Article details:

Received: 23 June 2014
Revised: 25 February 2015
Accepted: 10 March 2015

Key words:

hedonic price analysis,
agricultural land prices,
spatial heterogeneity,
spatial dependence.

© 2015 Nicolaus Copernicus University. All rights reserved.

Contents:

1. Introduction	114
2. Methodological Background	115
3. Application: Malaysia Agricultural Land Price	116
3.1. Data	116
3.2. Model Specification	117
3.3. Data Analysis and Results	117

4. Discussion.....	122
4.1. How Big Should a 'Neighbourhood' Be?.....	122
4.2. To Row-standardise or Not to Row-standardise?.....	123
4.3. Determining Implicit Prices.....	123
4.4. Model Selection and Policy Application.....	125
5. Conclusion.....	125
Appendix A.....	126
References.....	127

1. Introduction

Empirical models estimating determinants of price are generally *ad hoc* in nature; they do not least because markets for different goods are affected by different contexts and factors. For any given good, data and measurement constraints often yield different model specifications in different studies, and hence, different results. The model selection is therefore extremely critical to unveil a 'final' model which best describes market realities and available data. For a good as heterogeneous as land, price studies usually employ the hedonic modelling approach to account for the different combination of attributes shown by each parcel of land. Spatially-adjusted models test for correlations between observed data due to geographical proximity and similarity. As with other modelling biases, failure to account for spatial correlations will lead to less than accurate prediction of values. Naturally, a greater number of price studies are now using varying forms of spatial analyses such as Brunstad et al. (1995), Huang et al. (2006), Kim et al. (2003), Bell and Bockstael (2000), Madison (2007), Benirschka and Binkley (1994), Patton and McErlean (2003), Cotteleer et al. (2008), and Kuethe (2012). Economic models of urban land development have incorporated greater spatial complexity, focusing on spatial simulation models with spatial endogenous feedbacks and multiple sources of spatial heterogeneity (Chen et al., 2011). Despite the conceptual appeal of spatial analysis to researchers and research-users, its execution can be quite involved, whereas the results do not always lend themselves easily to policy-making applications. Muller and Loomis (2008) also cautioned that the gap between coefficients corrected and uncorrected for spatial dependence may not always be econom-

ically significant, i.e. the inefficiency attributable to spatial influences may not be large enough to cause critical errors in policy decisions. These concerns together with other factors such as lack of access to GIS data in certain economies may continue to impede the progress of the spatial econometrics application in price studies.

The objective of this paper is to describe some of these concerns applying data from the Malaysian farmland market. Firstly, a spatial regime model is developed to capture the effect of regional location on land price. The underlying assumption is that land characteristics may have different shadow prices depending on the region the land parcel is in, simply because land markets in different regions may be driven by different topographical, economic, and institutional factors. Sample heterogeneity is considered a major source of heteroscedastic errors in a model. Secondly, if it is suspected that the price of a parcel is partly explained by prices of nearby and similar parcels, then spatial dependence models are necessary to correct for this effect and other spatial attributes not captured by the model. When using the spatial model results, several issues must be sufficiently addressed (i) how to determine the accuracy of spatial weights employed; (ii) how to interpret the results to help draw implicit values of the attributes studied particularly when more flexible functional forms are not readily implemented in models adjusting for spatial dependence (Kim et al., 2003); (iii) how model selection is best conducted; and subsequently (iv) to what extent are the differences in the models significant in policy formulations.

The paper is organised as follows. Section 2 briefly presents theoretical underpinnings of both the hedonic price model and its spatial-adjust-

ed variations: spatial regime and spatial dependence models. Section 3 describes the dataset, model specification, and results. Section 4 delves into detail several computational and application considerations. Section 5 summarises and concludes.

2. Methodological Background

Hedonic Pricing Model (HPM). The principle underlying HPM is that a good's overall value is simply an aggregation of the implicit value of its attributes (Rosen, 1974) with the following regression form,

$$p_i = \alpha + \beta_1 x_{i1} + \dots + \beta_m x_{mi} + \varepsilon_i, \text{ for } i = 1, \dots, I, \quad (\text{a})$$

where p_i is price of item i , $X = (x_1, x_2, \dots, x_m)$ is a vector of the $k = 1, \dots, m$ characteristics of land and β are unknown parameters to be estimated and ε is the error term presumed to have a multivariate normal distribution, $N(0, \sigma^2 I)$. Semi or double-log functional forms are often applied to accommodate non-linearity in the data.

Spatial Heterogeneity. In cross-sectional data, heteroscedasticity can imply that at least one type of group-effect was overlooked. For farmland, market segmentation persists if: (i) buyers and sellers are neither able nor interested to participate in more than one local market, and (ii) markets differ in their supply and demand structures (see Freeman, 1979). Consequently, each market would bear different shadow prices for a given attribute (Goodman, Thibodeau, 1998). In contrast, market disaggregation is usually absent if suppliers and demanders actively interact across geographic markets to arbi-

trage price differences in different locations (Palmquist, 1989) for instance: (i) if the crop cultivated is traded in national and international markets, giving rise to an integrated land market throughout all regions; (ii) all regions have similar land and agricultural policies; and (iii) market agents do not display particular regional preferences.

Delineation of sub-markets can be evidenced by statistically significant shifts in the model intercept, functional form or slopes. Anselin (1988: 129) provides a brief summary and commentary of alternative procedures to account for spatial variation including switching regressions and spatial adaptive filtering process. One powerful approach to analyse spatial heterogeneity is to apply an interaction model whereby the intercept and all slopes are allowed to vary across sub-markets (Anselin, 1988; Patton, McErlean, 2003). Model "a" can be rewritten as a spatial regime model:

$$p_i = \sum_{r=1}^s \alpha_r + \sum_{k=1}^m \sum_{r=1}^s \beta_{kr} x_{kri} + \varepsilon_i, \quad \text{for } i=1, \dots, I, \quad (\text{b})$$

where $r=1, 2, \dots, s$ submarkets. The mean or covariance and variance structures differ and in consequence there are clustered error variances, denoted as $\text{Var}[\varepsilon_i] = \sigma_r^2$.

Spatial Dependence. The interdependence among parcels of land due to their relative geographic locations from each other can be formally stated as $\text{cov}(y_i | x_i, y_j | x_j) \neq 0$ where y_i and y_j are observations on a random variable at locations i and j (see Fulcher, 2004). For each observation i , there is a number of j neighbours potentially able to influence i 's outcome. This leads to non-zero covariance between observations even after controlling for dif-

ferences in attributes locations (see Kim et al., 2003; Anselin, Bera, 1998).

Spatial error dependence, in particular, refers to the existence of patterns in the regression error terms caused by: (i) one or more omitted variable in the equation; and the variable exhibits a spatial pattern; or (ii) the aggregation bias from the use of variables measured at different spatial scales, e.g., district-level climate index versus parcel-level price

data. **Spatial lag dependence** occurs when there is interdependence of the dependent or independent variable across observations as a result of their relative locations to each other (Bell, Bockstael, 2000). The price of parcel i is partly determined by prices of parcel j which falls under a certain 'neighbourhood' definition. The hedonic method assumes buyers and sellers have perfect information regarding parcel attributes and are able to objectively value its attributes. However, as Elad et al. (1994) correctly point out, although land exists everywhere, the markets for land are often localised with only a relatively small percentage of land changing hands each year. The thin volume of trading contributes to inefficient market information. When information is imperfect and costly, parcels located within the same neighbourhood are usually assumed to share similar characteristics. Hence, a parcel is often priced according to the local average price of the class of land it 'belongs' to. Subsequently this tendency perpetuates "circularity of price-setting" within the land market (Taff, 1999; Patton, McErlean, 2003).

Spatial Weights. Spatially-adjusted models require at the outset determination of parameter values identifying the collection of observations potentially influential to a given parcel i . This definition of a 'neighbourhood' can be formally expressed in a spatial weight matrix, \mathbf{W} . Standard choices for \mathbf{W} are row normalized contiguity matrices, nearest neighbour matrices, or weights inversely proportional to distance or to the square of the distance; the further the neighbour, the smaller weight it carries. The fact that the weights matrix is arbitrarily chosen is unsatisfactory, as different matrices may lead to different results. Therefore, some authors test several matrices, in order to assess the robustness of the conclusions. But, in most cases, only one matrix is used. In distance-based spatial weight matrices, elements w_{ij} can be either the absolute or the inverse distance between the i^{th} and the j^{th} observations; provided parcel j falls within a pre-specified distance from parcel i . Therefore, each \mathbf{W} is a full matrix with zero elements only on the diagonal; which makes for computationally intensive estimations. For this reason, distance-based matrices are usually employed for smaller datasets. In binary spatial weight matrix, the elements in \mathbf{W} equals one for i, j pairs considered neighbours and zero otherwise.

The m -order nearest neighbours matrix comprises elements $w_{ij} = 1$; otherwise $w_{ij} = 0$ if j is one of the m nearest neighbours to i . The extent of the neighbourhood is controlled through m . The resulting matrix is sparse and therefore its calculations require much less computer memory and storage space. Another benefit is that there will be no 'islands' or observations without neighbours (Anselin, Bera, 1998). (Note: The other common spatial weight matrix is the contiguity matrix which basically only allows contiguous neighbours to affect each other and hence, is usually applied when observational unit is aggregated (known boundaries). Its w_{ij} elements are positive if i th and j th observations share a common boundary). In row-standardised spatial matrices, the normalized weight matrix, $\tilde{\mathbf{w}}$, is structured as $\tilde{w}_{ij} = \left(1 / \sum_{j=1}^N |w_{ij}|\right) w_{ij}$ whereby the total sum of weights is fixed at unity. Row-standardisation is popular in the literature because it allows easy comparison between models and data as well as facilitates the maximum likelihood (ML) estimation of spatial models (see Cotteleer et al., 2008).

3. Application: Malaysia Agricultural Land Price

3.1. Data

The agricultural land price data comes from 2,222 land sales transactions for a period of seven years, from 2001 to 2007, and from four states in the west coast of Peninsular Malaysia: Selangor, Perak, Negri Sembilan, and Malacca, which were selected on the basis of their relatively higher degree of non-agricultural investment and population growth compared to the rest of the country. The principle data source is the annual Property Market Reports (PMR) published by the National Property Information Centre, Malaysia. Pre-excluded from the report are non-competitive transfers such as land leases to government agencies, land transfers involving nominal or zero compensation, and transactions between related business parties.

Distances are calculated using spatial points representing the parcels rather than centroids of post-code areas or districts, where the parcels are located. All calculations follow the Euclidian distance defi-

dition, i.e. straight-line distance between the parcel and its neighbours and **nearest** city in kilometres, follows $z = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ where x_1 and x_2 are longitudes and y_1 and y_2 are latitudes of the two points. Demographic data are derived from the 1991 and 2000 Population and Housing Census of Malaysia.

3.2. Model Specification

A summary of variable definition and descriptive statistics is provided in Table 1. The dependent variable is price per hectare in ringgit, *rprice*, in respect of year 2000 prices. Parcels located within the Integrated Agricultural Development Area or agricultural Group Settlement Act schemes (GSA) are restricted in terms of use as well as ownership. Malay Reserve Land (MRL) enactment bars certain areas of land from being sold to non-Malays.

The two types of land restrictions are represented in the model by *gsa* and *mrl* dummies respectively.

Road frontage, *rdfnt*, is hypothesised to give positive value to parcel price, irrespective of parcel's potential use. The proximity of a parcel to the nearest major city (Kuala Lumpur, Ipoh, Malacca City and Seremban), *distown*, is expected to be positive for price. District population growth, *popgro*, and population density, *popden*, indicate levels of urbanisation pressure in the area. Given the very wide range of its values, the dependent variable, *rprice*, is log-transformed to minimise problems related to heteroscedastic errors. Other log-transformed variables encompass *distown* and *popden*. The regression function is estimated firstly using the OLS model, followed by a spatial regime model and then Spatial Error Correction (SEC), Spatial Autoregression (SAR), Spatial Durbin (SD), and General Spatial (GS) models, as described in Appendix A.

Table 1. Data Description and Summary Statistics: Full Sample (n=2,222)

Variable	Description	Mean	Std Devia-Tion	Min	Max
<i>rprice</i>	Sale Value per hectare (in RM) in 2000 prices	106.028	146.490	4.753	1,254.197
<i>rdfront</i>	1=Parcel with Road Frontage; 0=otherwise	0.202	0.402	0	1
<i>distown</i>	Euclidian distance to nearest town (in km)	40.54	24.32	1.81	126.62
<i>popden</i>	District's population density based on 2000 Census	228.78	303.61	13.09	2,516.08
<i>popgro</i>	Annualised district population growth based on 1991 & 2000 Census (in %)	1.96	2.66	-0.41	13.47
<i>gsa</i>	<i>gsa</i> =1 if located in Group Settlement Schemes	0.22	0.42	0	1
<i>mrl</i>	<i>mrl</i> =1 if located in Malay Reserve Land areas	0.22	0.41	0	1

Source: Author's own work

3.3. Data Analysis and Results

Spatial Heterogeneity. In order to test spatial heterogeneity in the Malaysian data, the dataset is divided into two regions: (i) Central which includes Selangor, Malacca, and Negri Sembilan, three small but highly industrialised and densely populated states; and (ii) Perak which has a vast land stock but lower population density. Table 2 compares the regression results from the basic and the spatial re-

gime model. Overall, the latter's specification does very little to improve the model's explanatory power, increasing R^2 by only 0.025 points, although the joint hypothesis that the two regions are the equal yields F-statistics_{8,2206} = 18.11. All the explanatory variables are significant in both groups except *ldistown* and *year7*. More importantly, the spatial regime model is unable to reject heteroscedasticity ($\chi^2 = 59.62$). In summary, the effect of attributes on price obviously differs across regions. Nevertheless, spatial heterogeneity is apparently not the main source of group-effects in the data.

Table 2. Partial Elasticities from Basic and Spatial Regime Model Estimation

Variable	Basic Model		Spatial Regime Model			
	Coefficient	SE	Perak	SE	Central	SE
Central					1.33***	0.390
Rdfnt	0.84***	0.041	1.01***	0.062	0.62***	0.052
Gsa	-0.37***	0.032	-0.36***	0.042	-0.39***	0.044
Mrl	-0.13***	0.033	-0.14**	0.054	-0.11*	0.044
popgro	0.12***	0.008	0.26***	0.033	0.10***	0.009
lpopden	0.21***	0.019	0.18***	0.024	0.14***	0.041
ldistown	-0.14***	0.029	0.04	0.044	-0.14**	0.052
year7	-0.18***	0.036	-0.15***	0.045	-0.08	0.059
constant	10.23***	0.169	9.46***	0.210		
N	2,222			2,222		
R ²	0.5158			0.5422		
Adj. R ²	0.5143			0.5391		
Breusch-Pagan χ^2	58.24			59.62 (<i>p-value</i> = 0.000)		
Jacques-Bera χ^2	60.67			110.9 (<i>p-value</i> = 0.000)		
AIC	4,646.9			4,538.4		
SIC	4,692.6			4,629.7		

Explanation: Dependent variable is log of real price per hectare. Robust standard errors in parentheses (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$)

Source: Author's own work

Spatial Weight Matrix Specification. To avoid very large spatial matrices, the parcels are divided according to their best use-potentials: *development*, *rubber*, *oil palm* and *paddy*. Another category, *vacant*, comprises parcels which at the time of sale were uncultivated or under-utilised for various structural and institutional reasons. The dataset is not segmented further, e.g., by year or region, to avoid serious imbalances in matrix sizes. Furthermore, the effect of time is already sufficiently accounted for in the model through *year7* dummy.

To ensure robustness of the results, three types of spatial weight matrices are employed. The inverse distance-squared decay function is written as **W1**: $W_{ij} = 1/d_{ij}^2$; if $d_{ij}^2 < 11.1$ kilometers and 0 if otherwise; row-standardised. The second type is simply the un-standardised version of the first, i.e.

W2: $W_{ij} = 1/d_{ij}^2$; if $d_{ij}^2 < 11$ kilometers and 0 if otherwise. The third is a nearest neighbour weight matrix, namely **W3**: $W_{ij} = 1$ if five nearest neighbours, 0 if otherwise.

Spatial Autocorrelation Tests. The Moran and Spatial Lagrange Multiplier tests procedures are conducted using all three spatial weight matrices for each of the five land groups. Moran's I statistics is significant in all but paddy land category (Table 3). **W2** appears incompetent to describe spatial bias and therefore is not pursued further. Based on the robust LM_lag and LM_error tests, both **W1** and **W3** yield the same conclusions i.e. there is significant **spatial lag** bias in the oil palm, rubber and vacant land group; no spatial dependence is detected in paddy group while in development land group, the tests were not conclusive.

Table 3. Results from Tests of Spatial Dependencies

Sub-sample	Tests	W1: inverse-distance ² row standardised		W2: inverse-distance ² unstandardised		W3: five nearest neighbours	
		Statistic	p-value	Statistic	p-value	Statistic	p-value
Development (n=506)	Moran's_I	7.58	0.0000	0.75	0.4532	11.05	0.0000
	LM_Error	53.61	0.0000	0.48	0.4895	109.83	0.0000
	Robust_LM_Error	0.57	0.4520	0.41	0.5226	2.06	0.1510
	LM_Lag	55.34	0.0000	0.76	0.3844	114.87	0.0000
	Robust_LM_Lag	2.30	0.1295	0.69	0.4069	7.10	0.0077
Oil Palm (n=462)	Moran's_I	7.46	0.0000	1.4	0.1610	12.06	0.0000
	LM_Error	51.81	0.0000	1.97	0.1609	132.76	0.0000
	Robust_LM_Error	1.25	0.2645	2.3	0.1293	3.93	0.0475
	LM_Lag	67.08	0.0000	2.75	0.0975	141.28	0.0000
	Robust_LM_Lag	16.51	0.0000	3.08	0.0792	12.45	0.0004
Paddy (n=94)	Moran's_I	0.49	0.6238	0.04	0.9716	0.73	0.4671
	LM_Error	0.00	0.9823	0.00	0.9772	0.01	0.9345
	Robust_LM_Error	1.46	0.2273	0.00	0.9845	2.35	0.1251
	LM_Lag	0.18	0.6750	0.04	0.8447	0.97	0.3246
	Robust_LM_Lag	1.63	0.2013	0.04	0.8455	3.32	0.0686
Rubber (n=623)	Moran's_I	10.36	0.0000	0.76	0.4477	13.82	0.0000
	LM_Error	101.49	0.0000	0.55	0.4579	176.52	0.0000
	Robust_LM_Error	1.49	0.2220	0.46	0.4991	1.98	0.1596
	LM_Lag	129.02	0.0000	1.22	0.2690	197.61	0.0000
	Robust_LM_Lag	29.02	0.0000	1.13	0.2883	23.07	0.0000
Vacant (n=537)	Moran's_I	7.75	0.0000	0.31	0.7546	10.63	0.0000
	LM_Error	55.44	0.0000	0.09	0.7611	102.29	0.0000
	Robust_LM_Error	0.01	0.9124	0.16	0.686	3.91	0.0480
	LM_Lag	65.38	0.0000	1.77	0.1828	105.38	0.0000
	Robust_LM_Lag	9.95	0.0016	1.85	0.1743	7.01	0.0081

Source: Author's own work

Basic and Spatial Model Regressions. Since both **W1** and **W3** gave similar test conclusions, the latter is dropped from further discussions in the paper. Using **W1**, although the diagnostic tests suggest a spatial lag process for 3 out of 5 land categories, we estimate each group using the standard OLS model and all four spatial models. For brevity, Tables 4a and 4b only show selected parameter estimates, i.e. the lagged explanatory and dependent variables along with several model performance indicators. Based on the AIC and BIC model selection criteria, the best-performing models are OLS

for paddy, SEC for developable land, SAR for oil palm land. Results were less conclusive for the remaining two categories of land. For rubber land, the AIC points in favour of the SD while BIC supports the GS model. In the vacant land group, the AIC supports the GS model while BIC supports the SAR model. Both spatial processes are statistically significant when tested in their respective models; the Wald and likelihood-ratio test on λ are statistically significant in the SEC model, but so is ρ in the SAR model. From the GS regressions, ρ is not significant in the developable land category but significant in

others. The spatial error coefficient λ is significant in all sub-samples although it posted negative values in the rubber and vacant land categories. In all SD regressions, ρ is always significant. Wald tests on coefficient of lagged independent variables are in many cases significant, even though the individual parameter variables are not. In SD models, the number of regressors is doubled through the addi-

tion of spatially lagged dependent and independent variables. Since the R^2 or squared correlation computation does not correct for large number of regressors, it automatically becomes less reliable as a model performance criteria. Despite earlier LM tests which indicated spatial lag process in almost all land categories, individual regression have shown contrastingly mixed results.

Table 4a. Results of Spatial Durbin Models

Variables	Paddy	Developable	Oil Palm	Rubber	Vacant
wx_rdfnt	0.00 (0.143)	-0.16* (0.076)	-0.09 (0.112)	-0.05 (0.089)	-0.21 (0.140)
wx_gsa	-0.13 (0.115)	-13.20 (9.515)	0.01 (0.090)	0.09 (0.063)	-0.09 (0.138)
wx_mrl	0.33** (0.116)	-0.04 (0.080)	0.04 (0.144)	-0.09 (0.070)	0.13 (0.089)
wx_popgro	-0.03 (0.301)	0.02 (0.034)	0.12 (0.076)	0.10* (0.042)	-0.01 (0.042)
wx_lpopden	-0.23 (0.245)	0.02 (0.092)	-0.24* (0.112)	-0.04 (0.077)	-0.03 (0.069)
wx_ldistown	0.02 (0.168)	0.39* (0.193)	-0.48* (0.192)	-0.43* (0.171)	-0.39* (0.160)
wx_year7	-0.02 (0.134)	0.02 (0.091)	-0.14 (0.089)	-0.10 (0.085)	-0.20* (0.100)
Constant	11.60*** (1.345)	7.29*** (0.636)	7.58*** (0.766)	6.31*** (0.608)	7.40*** (0.623)
Rho	-0.00 (0.124)	0.34*** (0.050)	0.34*** (0.060)	0.40*** (0.046)	0.32*** (0.048)
R ² /Squared correlation	0.547	0.347	0.412	0.477	0.454
Log likelihood	-18.41	-369.74	-294.67	-346.12	-402.50
AIC	70.817	773.48	623.35	726.24	839.00
SIC	114.053	845.32	693.65	801.63	911.86

Explanation: Dependent variable is log real price per hectare. Robust standard errors in parentheses (** p<0.001, * p<0.01, * p<0.05)

Source: Author's own work

Table 4b. Results of OLS, SEC, SAR and GS models for the different land-uses

Variables	Paddy			Developable			Oil Palm			Rubber			Vacant							
	OLS	ML-SEC	ML-SAR	ML-SEC	ML-SAR	ML-GS OLS	ML-SEC	ML-SAR	ML-SEC	ML-SAR	ML-GS OLS	ML-SEC	ML-SAR	ML-SEC	ML-SAR	ML-GS OLS	ML-SEC	ML-SAR		
Constant	11.0** (0.738)	11.1*** (0.459)	11.6*** (1.497)	13.4*** (2.283)	11.1*** (0.273)	11.3*** (0.358)	7.34*** (0.618)	12.7*** (2.169)	10.8*** (0.304)	10.5*** (0.386)	6.52*** (0.701)	13.6*** (1.338)	10.0*** (0.327)	5.44*** (0.505)	3.92*** (0.554)	10.4*** (0.285)	10.3*** (0.324)	6.60*** (0.549)	4.46*** (0.685)	
Rho	-	-	-0.05 (0.110)	-0.23 (0.212)	-	0.34*** (0.048)	-0.11 (0.173)	-	-	0.38** (0.055)	-0.32* (0.130)	-	-	0.45*** (0.040)	0.60*** (0.052)	-	-	0.36*** (0.045)	0.57*** (0.063)	
lambda	-	0.00 (0.132)	-	0.22 (0.224)	-	0.36*** (0.046)	0.45** (0.144)	-	0.41*** (0.053)	-	0.67*** (0.089)	-	0.46*** (0.042)	-	-0.27** (0.090)	-	0.37*** (0.049)	-	-0.32** (0.099)	
R ² /Squared Correlation	0.493	0.493	0.495	0.502	0.321	0.319	0.324	0.315	0.356	0.354	0.380	0.325	0.416	0.413	0.441	0.446	0.411	0.409	0.421	0.424
Breusch-Pagan χ^2	11.51	-	-	-	0.38	-	-	1.25	-	-	-	-	2.68	-	-	0.80	-	-	-	-
Log likelihood	-23.74	-23.74	-23.65	-23.26	-402.2	-375.9	-376.1	-375.8	-331.3	-306.1	-302.4	-305.3	-413.1	-364.3	-355.9	-352.6	-438.3	-413.4	-410.5	-407.4
AIC	63.488	67.487	67.302	68.527	820.46	771.92	772.15	773.64	678.69	632.27	624.83	632.67	842.28	748.75	731.87	727.36	892.71	846.92	841.01	836.76
SIC	83.834	92.920	92.735	96.504	854.27	814.18	814.42	820.13	711.77	673.63	666.19	678.16	877.76	793.09	776.21	776.15	927.00	889.78	883.87	883.91

Explanation: Dependent variable is log of real price per hectare. Robust standard errors in parentheses (***) p<0.001, ** p<0.01, * p<0.05)

Source: Author's own work

In-sample prediction errors using numerical criteria where estimated regression model is imposed on approximately 20% of the sample to obtain predicted values of the dependent variable, *rprice*, this time in their natural scale. The resulting pairs of predicted and actual prices are used to generate prediction errors for the four land categories (paddy is excluded) estimated according to three regression models each, OLS, ML-Spatial Error, and ML-Spatial Lag, as shown in Table 5. The MSE and MAE values are all under 0.5, while the AER is less than 5%. The ML-Spatial Lag model produces the lowest

outcome in all cases except MAE and AER for vacant parcels in which the OLS model appeared superior. Overall, the reduction in prediction errors is not really substantial, i.e. the initial hedonic specification is fairly competent in predicting the dependent variable. Gao et al. (2006) arrived at the same conclusion in his cross-validation exercise of OLS, spatial dependency and geographically weighted regression models. We are also of the opinion that it is not very likely that an out-sample prediction endeavour would bring about different conclusions regarding model selection either.

Table 5. Comparison of Models using Numerical Criteria

Mean Squared Error				
Model	Developable	Oil palm	Rubber	Vacant
OLS	0.3032	0.2256	0.2361	0.3289
ML-Spatial Error	0.3024	0.2251	0.2554	0.3323
ML-Spatial Lag	0.2953	0.2215	0.2293	0.3232
Mean Absolute Error				
OLS	0.4356	0.3898	0.3888	0.4639
ML-Spatial Error	0.4348	0.3924	0.4081	0.4669
ML-Spatial Lag	0.4266	0.3887	0.3807	0.4677
Average Error Rate				
OLS	0.0357	0.0375	0.0393	0.0464
ML-Spatial Error	0.0356	0.0378	0.0414	0.0468
ML-Spatial Lag	0.0349	0.0374	0.0385	0.0467

Source: Author's own work

4. Discussion

4.1. How Big Should a 'Neighbourhood' Be?

Estimated coefficients are often more sensitive to spatial weight definitions than the estimation method itself (Bell, Bockstael, 2000). For instance, Breustedt and Habermann (2011) found that marginal incidence amounts of EU agricultural subsidies ranged from €0.38 up to €0.45 per additional euro depending on the spatial weight matrices chosen. Hence, in determining actual boundaries of a 'neighbourhood' via the choice of spatial weights, one must try to ensure the following:

- the structure of spatial dependence must accurately capture the potential influence between

observations; not based on *ad hoc* descriptions of spatial patterns, e.g., administrative boundaries, zipcodes, voting constituencies, and so forth (Anselin, 1988).

- the size and density of the 'neighbourhood' matters, e.g., distance-based weights matrix is less suitable for rural areas because of the smaller number of neighbours per unit of land compared to urban areas (Goldsmith, 2004 in Wang, Ready, 2005).

Cotteleer et al. (2008) employed Bayesian Model Averaging in combination with the Markov Chain Monte Carlo Model Composition technique. However, they acknowledged that their method was time-consuming whilst the regression resulted in lower bounds on estimated means and t-statistics.

Fernandez-Vasquez and Rodriguez-Valez (2007) estimated individual elements of their spatial weight matrix by using Maximum Entropy econometrics. They argued that if spatial weight matrices were row standardised, each one of their rows could be approached as probabilities distributions. However, the technique still relies on **arbitrary** specification of priors on the values of the spatial parameters.

4.2. To Row-standardise or Not to Row-standardise?

Row-standardising ensures total impact of neighbours across observations always sums to one despite varying number and/or density of neighbours for each observation. Wang and Ready (2005) noted that if distance-based weights are row-standardised, absolute distance between neighbours for each row would be re-scaled, causing actual spatial relation-

ships between observations to be distorted. Say that two observations A and B have the same number of neighbours (see Table 6’s first two rows). A’s neighbourhood is relatively sparser; whereas all of B’s neighbours are located nearby at approximately the same distance. Based on a distance-decay weighting principle, A’s first neighbour is weighted 0.4, while all of B’s are weighted only 0.2 despite being located at the exact same distance from their respective base observations. On the other hand, A’s third neighbour, who is located twice as far, is given 0.2. This ‘distance effect’ means that remote neighbours of one observation can be weighted equally as nearer neighbours of another observation. The resulting spatial weights matrix is no longer symmetrical, thus, computation of test statistics becomes complicated. Similar types of distortions can be found where further neighbours of observations with few neighbours have higher weights than closer neighbours of observation with many more neighbours.

Table 6. Example of ‘distance effect’ and ‘number effect’ due to row-standardisation (spatial weights provided in brackets)

Observation	Neighbour 1	Neighbour 2	Neighbour 3	Neighbour 4	Neighbour 5	Total Weights
A	2 km (0.4)	4 km (0.2)	4 km (0.2)	8 km (0.1)	8 km (0.1)	(1.0)
B	2 km (0.2)	2 km (0.2)	2 km (0.2)	2 km (0.2)	2 km (0.2)	(1.0)
C	2 km (0.5)	2 km (0.5)	2 km (0.5)	-	-	(1.5)
D	2 km (0.5)	2 km (0.5)	2 km (0.5)	2 km (0.5)	2 km (0.5)	(2.5)

Source: Author’s own work

On the other hand, **not** row-standardising the spatial weight matrix implies that units with more neighbours will attract higher price-premium than those with fewer neighbours, *ceteris paribus*. Compare the total effects of neighbours on C who has only two neighbours and D who has five (see Tab. 6’s last two rows). All neighbours are located at the same distance away from the respective observations. The total neighbourhood effects are 1.5 for C and 2.5 for D. This unintentional result is called the ‘number effect’. In a **spatial lag model**, the number effect is relatively far more damaging than the distance effect, because total spill-over of prices could multiply as the number of neighbour increases. The number effect is not as serious in a **spatial er-**

ror model because the magnitude of errors cannot be affected by the number of neighbours.

4.3. Determining Implicit Prices

Through the hedonic model coefficients and subsequently the derived partial price elasticities, it is possible to estimate the marginal or implicit value of a good’s attribute. In a spatial error model, the marginal implicit price is the same as the standard linear model simply because the adjustments do not yield different parameter estimates, only smaller or larger standard errors.

However, in the presence of spatial lag dependence, partial elasticity and predicted implicit prices must be calculated differently. Take a spatial lag

model which follows a semi-log specification. Partial differentiation of the spatial lag function with respect to attribute x_k is expressed as

$$\frac{\partial \ln P}{\partial x_k} = \begin{bmatrix} \partial \ln P_1 / \partial x_{1k}, \partial \ln P_1 / \partial x_{2k}, \dots, \partial \ln P_1 / \partial x_{nk} \\ \partial \ln P_2 / \partial x_{1k}, \partial \ln P_2 / \partial x_{2k}, \dots, \partial \ln P_2 / \partial x_{nk} \\ \partial \ln P_n / \partial x_{1k}, \partial \ln P_n / \partial x_{2k}, \dots, \partial \ln P_n / \partial x_{nk} \end{bmatrix} \tag{j}$$

The price of a given parcel where $i=1$, which is P_i (first row of the matrix) is directly influenced by marginal changes in attribute x_k in location 1 as well as changes in x_k which occurred in other locations, i.e. at neighbouring parcels or $x_{2k}, x_{3k}, \dots, x_{nk}$. In other words, coefficient estimate in an OLS model in the presence of spatial lag effect tends to over-value the impact of the regressor x_k on price as a result of indirect influences attributable to x_k coming from neighbouring units which are not accounted for. Hence, it can be argued that even if parameter estimates from spatial models differ very slight-

ly from OLS estimates, the difference in marginal effects can be far more substantial. The effect of a unit change in x_k induced at every parcel location in P_i is called the spatial multiplier; its value given by the sum of each row of the inverse matrix of row standardised spatial weight matrix or $1/(1-\rho)$; introduced into the matrix as $\mathbf{A} = [\mathbf{I} - \rho\mathbf{W}]^{-1}$.

If the dependent variable is in the log, the partial derivative Jacobian matrix to show elasticity of price in the semi-log model with respect to x_k can be written as

$$\frac{\partial \ln P}{\partial x_k} = \begin{bmatrix} \left(\beta_k \cdot a_{11} \cdot \frac{P_1}{\partial x_{1k}} \right), \left(\beta_k \cdot a_{12} \cdot \frac{P_1}{\partial x_{2k}} \right), \left(\beta_k \cdot a_{1n} \cdot \frac{P_1}{\partial x_{nk}} \right) \\ \left(\beta_k \cdot a_{21} \cdot \frac{P_2}{\partial x_{1k}} \right), \left(\beta_k \cdot a_{22} \cdot \frac{P_2}{\partial x_{2k}} \right), \left(\beta_k \cdot a_{2n} \cdot \frac{P_2}{\partial x_{nk}} \right) \\ \left(\beta_k \cdot a_{n1} \cdot \frac{P_n}{\partial x_{1k}} \right), \left(\beta_k \cdot a_{n2} \cdot \frac{P_n}{\partial x_{2k}} \right), \left(\beta_k \cdot a_{nn} \cdot \frac{P_n}{\partial x_{nk}} \right) \end{bmatrix} = \beta_k A \left(\frac{P}{x_k} \right) \tag{k}$$

The matrix's diagonal elements represent the direct or own effects of x_k on price, while its off-diagonal elements represent the indirect/cross-effect coming from x_k changes in neighbouring units. Partial differentiation with respect to a log-transformed x_k variable requires that a specific value of x_k is introduced into the matrix. Usually the sample means that \bar{x}_k is sufficient, but in the presence of spatial lag, the mean values of x_k are equally relevant in other locations. Yet since these values are not the same in each location, i.e. $\bar{x}_{1k} \neq \bar{x}_{2k} \neq \bar{x}_{3k} \dots \neq \bar{x}_{nk}$, the task of computing the implicit prices becomes more tedious and arbitrary than usual. The extra computational burden may be avoided unless the empirical exercise aims to seek improved point estimates. OLS partial elasticities and predicted implicit prices can be sufficiently useful as an upper-bound guide to policy assessment or market analysis. Indeed various studies concluded that the marginal effects of land attributes on its price from both OLS and a spatial lag model estimation to be almost identical (Kim et al., 2003; Patton, McErlean,

2003; Mueller, Loomis, 2008; Nelson, 2010). They found relatively small differences between implicit prices from OLS and spatially-corrected estimates. This suggests that OLS in hedonic models may still give reasonable estimates even in the presence of the spatial dependence, particularly if only a small number of coefficients are affected and the degree of bias is relatively small.

4.4. Model Selection and Policy Application

In our examination of the Malaysian farmland data, the LM-tests, model regression outputs and predictive cross-validation exercises failed to help us decisively identify the nature and extent of spatial bias in the data. Recommended spatial multiplier values obtained were on the high side, hence, substantive inferences based on them could be markedly different from those based on OLS. However, parameter estimates in all of the spatial models were almost always relatively smaller than their OLS counterparts;

confirming our suspicion that OLS coefficients tend to overstate the impact of regressors on the dependent variable (implicit prices). OLS coefficients are not found to be remarkably different in the developable land group or the paddy group, where neither types of spatial bias were found. To adjust those, one needs to employ the computed spatial lag multiplier ($1/(1-\rho)$) to deflate the partial elasticity values, i.e. 1.613, 1.818, and 1.563 for rubber, oil palm, and vacant land categories respectively. However, since the multipliers are quite large, any downward corrections would yield very small estimate values, much smaller than the values suggested by the spatial lag model itself, or in any other spatial model for that matter. To illustrate that tendency it can be stated that if OLS coefficient for dummy *year7* in the vacant land regression is 0.20 then the spatial-lag-adjusted coefficient would be $0.20/1.563 = 0.13$, whereas the spatial lag model estimate is 0.16. This apparent lack of convergence could prevent unreserved and meaningful valuations regarding the true degree of the spatial lag bias.

Assuming that data constraints prevent the researcher from adopting spatial models, the problem of spatial bias can be minimised by introducing additional spatial information or variables into the model. For instance, Diniz-Filho et al. (2003) found that just by adding climatic variables into their ecological diversity function, spatial structure in the original data appeared to be sufficiently explained. They also argued if residuals were auto-correlated at small distance classes, observations which were only short distances apart might not provide independent data points for testing long distance spatial effects. More research needs to be conducted in order to identify the types of spatial information which can be proven to generally reduce spatial correlations in price models. This is a useful contribution to future empirical work. Intuitively, the higher the amount of spatial information provided, the smaller the risk of omitted variable bias, thus, the smaller the degree of spatial autocorrelation.

5. Conclusions

Spatial heterogeneity and spatial biases in hedonic studies are indisputably valid issues to be addressed

in hedonic price studies. This paper highlights some of the practical considerations when applying and using spatial analysis outcomes in empirical work, particularly relating to spatial weights and model selection. Spatial autocorrelation tests and its coefficients are very sensitive to spatial weight matrix specifications; such that different weights yield different conclusions on the same dataset. Future research is needed to facilitate the process of filtering out spatial lag bias particularly in log-log models or spatio-temporal models (Note: in spatio-temporal models, sales are ordered according to time. In addition to specifying the 'neighbourhood' extend, the researcher also needs to arbitrarily specify how long the influence of one sale prevails over another sale). Obviously, more theoretical and empirical studies are needed to help draw standardised principles for its implementation and interpretation across different functional forms, and indeed across different goods and spatial contexts. Whilst the growth of spatial analyses literature is greatly encouraging, each study should endeavour to meet at least two conditions in order for it to lead to improved recommendations regarding the inclusion of spatial heterogeneity and dependencies adjustments in policy interventions. Firstly, the researcher must ensure that the spatial interactions is accurately described by the spatial structure (spatial weights) proposed. Secondly, the presence of spatial autocorrelation must be proven to cause statistical inefficiency as well as overestimation of land values when the influence of spatial autocorrelation was left uncorrected. In other words, spatial models must outperform the OLS estimation in the presence of spatial autocorrelation. If the two conditions are not fulfilled satisfactorily or if data and time constraints prevent a full-scale spatial analysis, then it would remain reasonable to use standard model results, at least as upper bounds limits for coefficient estimates.

Appendix A

I. Spatial Models

Various adjustments to model the two types of spatial dependences are described as follows.

i) Spatial Error Correction Model (SEC)

The basic hedonic function can be extended to include a spatially autoregressive process in the error term

$$y = X\beta + u \quad \text{where} \quad u = \lambda Wu + \varepsilon \quad \text{and} \quad \varepsilon \sim N(0, \sigma^2 I) \quad (c)$$

where y is a $(n \times 1)$ vector of dependent variables, X a $(n \times k)$ matrix of explanatory variables, β is a $(k \times 1)$ vector of parameters, λ is the spatial scalar autocorrelation coefficient, W is the $(n \times n)$ spatial weight matrix, u is the vector of spatially correlated error terms, and ε is the vector of uncorrelated error term. The spatial autoregression coefficient, λ ,

indicates the correlation between parcel i 's error and a composite of the errors of its neighbours.

ii) Spatial Autoregressive Model (SAR)

This spatial autoregressive process can be formalised and added as an extension of the basic hedonic model to obtain

$$y = X\beta + \rho Wy + \varepsilon \quad (d)$$

where ρ is a scalar autoregressive parameter and ε is as usual, distributed according to $\varepsilon \sim N(0, \sigma^2 I)$. Technically, the spatial lagged dependent variable, Wy , is an endogenous variable which is always correlated with the error term, ε_i , as well as the error terms at all j locations. The model must be estimated by either Maximum Likelihood (ML) or instrumental variables (IV) techniques. The former involves maximizing the related log likelihood function with respect to the scalar autoregressive parameter, ρ . However, ML procedures are often challenging when the sample size is large. Moreover, they call for explicit distributional assumptions. Alternatively, the IV procedure which is computationally simple

and less restrictive regarding the distribution of the disturbances can be applied for cross-sectional spatial autoregressive models (Kelejian, Prucha, 1998).

ii) Spatial Durbin Model (SD)

If there are reasons to suspect that an observation is also affected by the explanatory variables of neighbouring observations, then the spatial Durbin or spatial common factor model is more appropriate (Anselin, 1988). In this model, a set of spatially-lagged explanatory variables is added to the right-hand side

$$y = \rho Wy + X\beta_1 + \rho Wx + X\beta_2 + \varepsilon \quad (e)$$

iv) General Spatial Model (GS)

The general spatial model basically incorporates the spatial error term into the spatially lag dependent model and therefore is considered to be a higher-order

model. A different weight matrix may be specified for each of the spatial dependence processes if it is believed that a different set of neighbours produces different type of influences. The general spatial model can be written as

$$y = \rho W_1 y + X\beta + u \quad \text{where} \quad u = \lambda W_2 u + \varepsilon \quad \text{and} \quad \varepsilon \sim N(0, \sigma^2 I)$$

Combining the two spatial processes in one expression yields,

$$\mathbf{y} = \rho \mathbf{W}_1 \mathbf{y} + \lambda \mathbf{W}_2 \mathbf{y} - \rho \lambda \mathbf{W}_1 \mathbf{W}_2 \mathbf{y} + \mathbf{X} \boldsymbol{\beta} - \lambda \mathbf{W}_2 \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (\text{f})$$

II. Model Selection Criteria

Generally, if a spatial model does not outperform the standard linear model, then the simple linear model is considered sufficiently robust to represent the data. Since spatial models are not estimable

by OLS method (see Anselin, 1988: 243), alternative measures of fit are obtained from information-based criteria, e.g., the Aikake Information Criterion (AIC) and the Schwartz Information Criterion (SIC). Gao et al. (2006) suggest comparing prediction powers of the models using in-sample or out-sample observations. Three useful numerical cross-validation criteria are:

$$\text{Mean of squares of prediction errors: } \frac{1}{n} \sum (y_i - \hat{y}_i)^2 \quad (\text{g})$$

$$\text{Mean of absolute errors: } \frac{1}{n} \sum |y_i - \hat{y}_i| \quad (\text{h})$$

$$\text{Average error rate: } \frac{1}{n} \sum \frac{|y_i - \hat{y}_i|^2}{y_i} \quad (\text{i})$$

where y_i and \hat{y}_i denote the observed and predicted dependent variables, respectively. This means the smaller the values given by these criteria, the better the model's performance.

References

- Anselin, L.**, 2003: Spatial Externalities, Spatial Multipliers and Spatial Econometrics. In: *International Regional Science Review*, Vol. 26, No. 2, pp. 153-166. DOI: <http://dx.doi.org/10.1177/0160017602250972>
- Anselin, L.**, 1988: *Spatial Econometrics: Method and Models*, Dordrecht: Kluwer Academic Publishers.
- Anselin, L.**, 1990: Some Robust Approaches to Testing and Estimation in Spatial Econometrics. In: *Regional Science Urban Economics*, Vol. 20, pp. 141–163.
- Anselin, L. and Bera, A.K.**, 1998: Spatial dependence in Linear Regression Models with an introduction to Spatial Econometrics. In: Ullah, A. and Giles, D. editors, *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, pp. 237-289.
- Breustedt, G. and Habermann, H.**, 2011: The Incidence of EU Per-Hectare Payments on Farmland Rental Rates: A Spatial Econometric Analysis of German Farm-Level Data. In: *Journal of Agricultural Economics*, Vol. 62, pp. 225-243. DOI: <http://dx.doi.org/10.1111/j.1477-9552.2010.00286.x>
- Bell, K.P. and Bockstael, N.E.**, 2000: Applying the generalized-moments estimation approach to spatial problems involving micro-level data. In: *Review of Economics and Statistics*, Vol. 82 (1), 72-82.
- Bell, K.P. and Irwin, E.G.**, 2002: Spatially Explicit Micro-level Modelling of Land Use Change at the Rural-urban Interface. In: *Agricultural Economics*, Vol. 27, pp. 217–232.
- Benirschka, M. and Binkley, J.M.**, 1994: Land Price Volatility in a Geographically Dispersed Market. In: *American Journal of Agricultural Economics*, Vol. 76, No. 2, pp. 185-195.
- Bockstael, N.E.**, 1996: Modeling Economics and Ecology: The Importance of a Spatial Perspective. In: *American Journal of Agricultural Economics*, Vol. 78, pp. 1168–80.
- Bowen, W.M., Mikelbank, A. and Prestegaard, D.M.**, 2001: Theoretical and Empirical Considerations Regarding Space in Hedonic Price Model Applications. In: *Growth and Change*, Vol. 32, No. 4, pp. 466-490.
- Brown, J. and Rosen, H.**, 1982: On Estimation of Structural Hedonic Price Models. In: *Econometrica*, Vol. 50, pp. 765-788.

- Brunstad, R.J., Gaasland, I. and Vårdal, E.,** 1995: Agriculture as a Provider of Public Goods: a Case Study for Norway. In: *Agricultural Economics*, Vol. 13, pp. 39-49.
- Chen, Y., Irwin, E. and Jayaprakash, C.,** 2011: Incorporating spatial complexity into economic models of land markets and land use change. In: *Agricultural and Resource Economics Review*, Vol. 40, No. 3, pp. 321-340.
- Cotteleer, G., Stobbe, T. and Van Kooten, G.C.,** 2008: A Spatial Bayesian Hedonic Pricing Model of Farmland Values. In: *12th European Association of Agricultural Economists*, Belgium: Ghent, August 2008.
- Diniz-Filho, J.A.F., Bini, L.M. and Hawkins, B.A.,** 2003: Spatial Autocorrelation and Red Herrings in Geographical Ecology. In: *Global Ecology and Biogeography*, Vol. 12, pp. 53-64.
- Dubin, R.,** 1988: Estimation of Regression Ccoefficients in the Presence of Spatially Autocorrelated Error Terms. In: *Review of Economics and Statistics*, Vol. 70, pp. 466–474.
- Elad, R.L., Clifton, I.D. and Epperson, J.E.,** 1994: Hedonic Estimation Applied to the Farmland Market in Georgia. In: *Journal of American Agriculture and Applied Economics*, Vol. 26, pp. 361-5.
- Fernandez-Vazquez, E. and Rodriguez-Valez, J.,** 2007: Taking off some hoods: Estimating spatial models with a non-arbitrary W matrix. In: *Meeting of the Spatial Econometric Association*, Cambridge: 11-14 July 2007.
- Freeman, A.M., III,** 1979: Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues. In: *Scandinavian Journal of Economics*, Vol. 81, No. 2, pp. 154-73.
- Fulcher, C.,** 2003: Spatial Aggregation and Prediction in the Hedonic Model. In: PhD. dissertation, North Carolina State University.
- Gao, X., Yasushi, A. and Chung, Ch-J.F.,** 2006: An Empirical Evaluation of Spatial Regression Models. In: *Computers and Geosciences*, Vol. 32, pp. 1040–1051. DOI: <http://dx.doi.org/10.1016/j.cageo.2006.02.010>
- Goodman, A.C. and Thibodeau, U.T.G.,** 1998: Housing Market Segmentation. In: *Journal of Housing Economics*, Vol. 7, pp. 121-143.
- Huang, H., Miller, G., Sherrick, B. and Gomez, M.,** 2006: Factors Influencing Illinois Farmland Values. In: *American Journal of Agricultural Economics*, Vol. 88, No. 2, pp. 458-470.
- Kelejian, H.H. and Prucha, I.R.,** 1998: A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. In: *Journal of Real Estate Finance and Economics*, Vol. 17 (1998), pp. 99–121
- Kim, C.W., Phipps, T.T. and Anselin, L.,** 2003: Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach. In: *Journal Environmental Economics and Management*, Vol. 45, No. 1, pp. 24–39.
- Kueth, T.H.,** 2012: Spatial fragmentation and the value of residential housing. In: *Land Economics*, Vol. 88, No. 1, pp. 16-27.
- Maddison, D.,** 2000: A Hedonic Analysis of Agricultural Land Prices in England and Wales. In: *European Review of Agricultural Economics*, Vol. 27, No. 4, pp. 519-532.
- Mueller, J.M. and Loomis, B.,** 2008: Spatial Dependence in Hedonic Property Models: Do Different Corrections For Spatial Dependence Result in Economically Significant Differences in Estimated Implicit Prices? In: *Journal of Agricultural and Resource Economics*, Vol. 33, No. 2, pp. 212-231.
- National Institute of Valuation, Valuation and Property Services Division. Property Market Report. Kuala Lumpur: Ministry of Finance Malaysia.
- Nelson, J.P.,** 2010: Valuing rural recreation amenities: hedonic prices for vacation rental houses at Deep Creek Lake, Maryland. In: *Agricultural and Resource Economics Review*, Vol. 39, No. 3, pp. 485-504.
- Palmquist, R.B.,** 1989: Land as a Differentiated Factor of Production: A Hedonic Model and Its Implications for Welfare Measurement. In: *Land Economics*, Vol. 6, No. 5, pp. 23-8.
- Patton, M. and Mcerlean, S.,** 2003: Spatial Effects within the Agricultural Land Market in Northern Ireland. In: *Journal of Agricultural Economics*, Vol. 54, pp. 35–54. DOI: <http://dx.doi.org/10.1111/j.1477-9552.2003.tb00047.x>
- Rosen, S.,** 1974: Hedonic Prices And Implicit Markets: Product Differentiation in Pure Competition. In: *Journal of Political Economy*, Vol. 82, pp. 32–55.
- Taff, S.J., Tiffany, D.G. and Weisberg, S.,** 1996: Measured Effects of Feedlots on Residential Property Values in Minnesota: A Report to the Legislature. In: *Staff Paper P96-12*, Saint Paul: Department of Applied Economics, University of Minnesota.
- Wang, L. and Ready, R.C.,** 2005: Spatial Econometric Approaches to Estimating Hedonic Property Value Models. In: *American Agricultural Economics Association Annual Meetings*, Providence, RI, July, 2005.
- Wilhelmsson, M.,** 2004. A Method to Derive Housing Sub-Markets and Reduce Spatial Dependency. In: *Property Management*, Vol. 22, No. 4, pp. 276-288.