Does it matter how far you live from the city centre? A study of apartment values in Vienna

Shanaka Herath

Research Institute for Spatial and Real Estate Economics
Vienna University of Economics and Business
Nordbergstrasse 15
A-1090 Vienna, Austria
Tel.: +43-1-31336-5764
shanaka.herath@wu.ac.at

The explanations provided in the literature for community-level variations in constant-quality house prices vary from urban economics (e.g., distance from the Central Business District) to local public economics (e.g., property tax rate) to urban-amenities theory (e.g., crime rate). This paper examines the impact of distance from the city centre on values of residential apartment units in Vienna. Data on individual owner assessments of apartment value along with other hedonic determinants of housing value is obtained from a popular online database of apartment units available for sale. The three candidates for the location variable include the district in which the apartment units are located, a categorical variable classifying location into three groups -the centre, neighbouring districts and outskirts- in addition to estimates of distance from a predefined point in the city centre to the apartment units. The control variables of the analysis include number of rooms, living area, number of the floor, number of bathrooms and toilets, availability of balcony / terrace / elevator / basement and type of flooring in addition to the information if the unit is furnished. The first part of the article examines the possible influence of distance and other location variables on the price of a constant-quality apartment. The results demonstrate a negative relation between distance from the city centre and apartment price. The tests employed to empirically explore if there is spatial autocorrelation in the residual series suggest use of the spatial autoregressive error model (SEM) as appropriate. The estimation results of the spatial model are very similar to those of the base model, suggesting that the results produced by the base model are not an outcome of any misspecification of the model.

Keywords: house prices, hedonic model, centralization

1. Introduction

The hedonic price method (HPM) is a popular tool commonly used to estimate demand or marginal price of different attributes of various commodities. Some applications of

HPM in environmental economics to estimate the implicit value of environmental amenities are good examples. The attractiveness of HPM comes from the fact that it has an intuitive appeal to the consumer theory of Lancaster and, from a practical point of view, the results are straightforward and easy to interpret. The semi-logarithmic form of the hedonic function, for example, has the advantage that the coefficient estimates are proportions of the price, i.e., marginal prices, which are directly attributable to the respective characteristics within a composite commodity (Herath and Maier, 2010). HPM is also widely used in real estate and housing economics research and spatial hedonic models of house prices appeared with the widespread use of spatial econometrics in the recent past.

The purpose of this study is to use a spatial hedonic model to examine the impact of distance from the city centre on values of residential apartment units in Vienna. The explicit research question this study attempts to address is whether constant-quality apartment values are influenced by the distance from the city centre. The study indirectly examines whether structural characteristics of apartment units have any influence on prices. Use of a spatial hedonic model is justified by the fact that there is overwhelming evidence in the literature that prices of housing units which are closer to each other are spatially autocorrelated. Similarity of structural attributes of units, the same environmental amenities in the surrounding area and similar income levels of units close to each other are seen as possible causes of this correlation.

The methodology with regard to study of cities evolved from the monocentric model to polycentric models. Early empirical studies using the monocentric model reported that the distance variable in house price models produced the expected negative coefficient, indicating that constant-quality house prices go down in value when houses are located further away from the city centre. However, Bender and Hwang (1985), Dubin (1992), Dubin and Sung (1987) and Olmo (1995), among many others, show that subsequent studies produced contradictory results. These studies point to the polycentric agglomeration of cities as a possible reason for this controversy.

However, the radial design of the public transport system, and the fact that majority of jobs are concentrated in the centre suggest that Vienna is necessarily a monocentric city. On the one hand, the public transport system, which consists of trains, trams and buses, is very efficient and comprehensively covers the whole city and, this well-developed public transport network is planned in such a way that the routes are by and large directed towards the centre from the outskirts. On the other hand, the first district is the largest employment locality in Vienna. This is partially due to tourism, as well as the presence of many corporate headquarters occasioned by the central location of the 1st district. The empirical part of the present article provides insights into the validity of the monocentric model for the city of Vienna and adds to the literature indicating that the prominence of the city centre still stands out in many cities as the focal economic centre. The originality of this study lies in the fact that this is the first paper to examine explicitly a version of the monocentric model in the context of Vienna.

In a related previous paper, Brunauer et al (2009) estimate hedonic price equations for rents in Vienna. The analysis employed attempts to explain unobserved district-specific heterogeneity with location-specific intercepts, with the postal code serving as a location variable. Multiplicative scaling factors are introduced in order to allow for spatial variation in the nonlinear price gradients. The authors conclude that there is substantial spatial variation in price gradients within Vienna, reflecting the existence of submarkets related to districts. The main point of departure of the present article is that it considers Vienna as a single housing market. The analysis employed here not necessarily disregards the polycentric argument, rather it suggests although Vienna has service centres such as supermarkets, recreation centres, employment centres and educational centres spread throughout the city (multi-centric), it demonstrates a monocentric state of affairs at the aggregate city level.

The remaining paper is organized as follows: Section 2 provides a brief literature review primarily on the monocentric model; Section 3 describes the data sources and variables; Section 4 discusses the methodology employed in the study which consists of the base model, spatial autocorrelation tests and the spatial hedonic model; Section 5 presents empirical results; and Section 6 concludes with a summary of main findings.

2. Brief literature review on the monocentric model

The literature on variations in constant-quality house prices provides approaches to estimate demand and prices of houses based on, among other things, urban economics, local public economics and urban-amenities theory. In urban economics, for example, it is established that distance from the CBD and accessibility influence house prices

(Alonso, 1964; Mills, 1967; Muth, 1969). On the other hand, public services and property tax rates have an influence on house prices based on local public economic theories (Hamilton, 1976; Yinger, 1982). Additionally, urban-amenities theory postulates that crime rates influence house prices (Jackson, 1979; Rosen, 1979).

Turning to urban economics, a number of alternative approaches were established over the years to capture the effects of location on value of real estate. Thünen (1826) included a distance variable in a study of agricultural land use and the urban monocentric model of Alonso (1964), which was popular after the mid-1980s, applied this to urban regions. The initial monocentric model was subsequently revised and generalised by Mills (1967, 1972a, 1972b) and Muth (1969) and was known as AMM model to symbolize contributions of Alonso, Mills and Muth to this theory. In its present form, the monocentric model suggests the following: house prices are influenced by distance from the CBD, transportation costs, household income, metro-area population size and agricultural rental rates. This implies that distance from the CBD should be included in any real estate price model, and the inclusion of distance from the CBD has numerous implications on all real estate valuation models. First and foremost among them is that firms and households are willing to bid more for land that is closer to the CBD because transport costs, in terms of out-of-pocket expenses or travel times to the CBD, will be lower. Over the last few decades, the urban monocentric model has been empirically tested by many scholars. Ball (1973) and Richardson (1988) provide literature surveys on this topic.

The urban monocentric model is an approach that takes into account the absolute location of points in space. Haynes and Fotheringham (1984) state that all entities on the face of the earth can be identified in absolute terms by referring to their respective longitude and latitude coordinates. It is then possible, by referring to these coordinates, to compare absolute positions of entities. For example, distance from one point on the surface to another can be specified in this way. Most scholarly work on urban monocentric models includes either distance to the city centre, travel time or travel cost in the model specification to capture these price dynamics generated by location in space. The monocentric model is, therefore, consistent with the idea of absolute location. Alternative models such as gravity models are, in contrast, based on relative location of points or regions in space.

3. Data sources and variables

The dataset employed in this study was provided by a popular online database (ERESNET GmbH) of apartments available for rent and sale in Vienna. The dataset covers the period from 11 December 2009 to 25 March 2010 and includes individual owner assessments of prices (offered prices) in addition to the other hedonic determinants of housing value. By using data collected over a short period of time, i.e., a little over 3 months, the analysis gains the advantage of studying a constant sample of properties over the cycle. In other words, it is safe to assume that the temporal dynamics do not have a significant impact on apartment values within this short period of time.

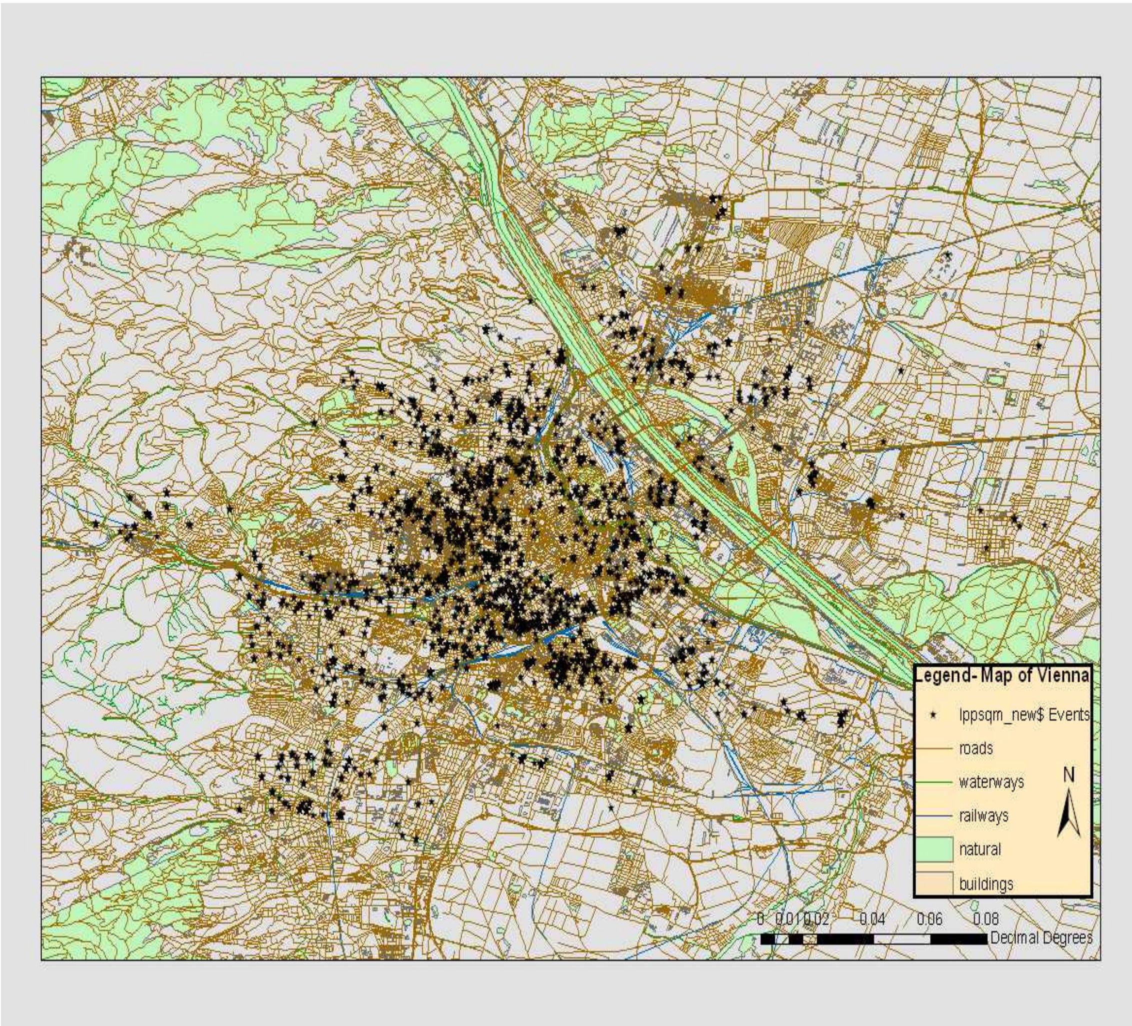

Fig.1 Location of the apartment units within the city of Vienna. Source: Author's own compilations.

Information regarding the location of apartment units is primarily required in order to perform the spatial analysis; therefore, the observations without address or other information about spatial location of the unit were removed from the dataset. Some data cleaning was required particularly to deal with the problem of outliers. The outliers detected, mainly in the price series, were carefully dealt with through consultation with the data provider. This process led to a dataset of 7028 observations of apartment units within the Vienna city limit suitable for the analysis (see Fig. 1).

The unit of analysis of the study (the dependent variable) is owner appraisal values of apartment units. There is a prolonged discussion in the literature whether the appraisal value is an accurate approximation of the actual price of housing units, although in the absence of transaction prices, a number of studies make use of assessed values to carry out analyses. The dataset also contains the following additional structural variables: number of rooms, living area (in square meters), number of the floor, condition of the apartment, number of bathrooms, number of toilets, availability of a balcony, terrace, elevator or basement, type of flooring and information on whether the unit is furnished. Appropriate use of these data in the analysis requires defining categorical variables as listed in Table 1. Most studies examine the effect of these characteristics on sales price, but the present study chooses to use price per square metre to homogenize the response variable.

Table 1 Categorical variables

Scale type	Variable	Categories
ordinal	Floor	Ground floor-3rd floor (0-3)
		4th floor and above (4)
		Top floor (5)
	Condition	Best-Bad (1-3)
	Furnished	Unfurnished-Furnished (0-2)
	Location	CBD-Outward (0-2)
binary	Balcony	Yes/No (1/0)
	Terrace	Yes/No (1/0)
	Elevator	Yes/No (1/0)
	Basement	Yes/No (1/0)
	Parquet flooring	Yes/No (1/0)

The dataset also contains the location variables postal address and the district (also the postal code- PLZ) in which the apartment units are located. Vienna is comprised of 23 districts (Bezirke), and they are numbered approximately in a clockwise pattern. The 1st district is the city centre - what is considered the CBD in this study. The 1st district used to be the entire city until the mid 19th century, and even today it is considered the centre of the expanded city. As an alternative location variable, estimates of distance from a predefined point in the CBD to apartment units are obtained based on longitude and latitude data calculated using Google Maps[1]. The historically important Stephansdom church is situated right in the middle of the 1st district, and therefore distance from the Stephansdom church to each apartment unit was obtained.

Fig. 2 Districts of Vienna. Source: Author’s own design.

However, the preferred way of incorporating location into this analysis is to define a dummy variable to capture apartments in the CBD into one category, in closer districts into a second category, and in farther ones into a third category. This mechanism assigns the dummy “0” to the apartment units in the 1st district (inner-ring or CBD), the dummy “1” to the apartment units within the “outer-ring” districts (neighbouring districts of the CBD) and the dummy “2” to the apartment units in the outskirts. Figure 2 demonstrates this classification.

[1] The mechanism involves providing the address in the graphical interface of Google Maps to track the exact location. It is then possible to calculate longitude and latitude data, look at the distribution by district, and to check for any possible outliers individually.

4. Methodology

The general form of the hedonic price function with regard to house prices takes the form

$$P_i = f(S_i, L_i, N_i)$$

where p_i is the log of the sales price of house i, S_i is a vector of structural housing characteristics, L_i is a vector of location variables, and N_i is the neighbourhood characteristics.

The initial regression model is estimated by means of OLS, and it takes the form

$$y_i = X_i \beta + \varepsilon_i$$

where y is a (N×1) stochastic variable, X is a (N×k) matrix of non-stochastic variables, and ε is an (N×1) error vector that is IID[2] (0, σ^2). In the case of the hedonic price function shown above, the log of the sales price is the stochastic variable (y) and all the hedonic variables such as structural housing characteristics, location variables, and neighbourhood characteristics are incorporated into the matrix of non-stochastic variables (X).

However, spatial econometrics literature illustrates that the relationship depicted above poses a problem when i observations represent regions or points in space. When regions or points are close to each other, the values observed at one location tend to depend on observations made at other locations. This idea is consistent with Tobler's first law of geography, which states that everything is related to everything else, but close things are more related than things that are far apart (Tobler, 1970). Similarly, LeSage and Pace (2009) define this spatial dependence as a situation where values observed at one location or region, say observation i, depend on the values of neighbouring observations at nearby locations. The presence of spatial dependence indicates that the observations are spatially autocorrelated, and if the spatial effects are not taken into account within the model, the traditional assumption of IID errors is violated as a consequence of this dependence.

[2] The classical assumption of independent and identically distributed (i.i.d.) variable

Wilhelmsson (2002) shows that spatial autocorrelation can arise from the following sources in the context of house prices: the price is affected by the price of neighbouring houses; relevant spatially correlated variables have been omitted; or the functional form is misspecified or suffers from measurement error.

There are several tests to detect spatial autocorrelation. One of the oldest test statistics among them, Moran's I coefficient, is still widely used to test for any spatial correlation in linear regression models. More recently, Lagrange Multiplier (LM) tests were developed within the framework of maximum likelihood theory. The LM tests became predominant in the recent past because they provide the possibility to test jointly the hypothesis of no spatial dependence due to an omitted spatial lag or due to spatially autoregressive errors.

The LM tests check for two types of spatial autocorrelation that occur due to the dependence in spatial lag ρ and spatial error λ

$$\begin{aligned} y &= X\beta + \rho W y + u, \\ u &= \lambda W u + e \end{aligned}$$

where e is an error term which is IID, and W is a spatial weight matrix. This spatial weight matrix is constructed to specify and standardize neighbours so that each row adds up to unity, creating a row-stochastic matrix for estimation of the parameter ρ, the spatial autoregression coefficient. Neighbours in these analyses can be defined either as k number of nearest neighbours or based on distance. If neighbours are based on distance, observations that fall within a lower distance bound and an upper distance bound (usually measured in kilometres) are included. λ is the error correlation coefficient.

Tests for a missing spatially lagged dependent variable test that $\rho = 0$, while tests for spatial autocorrelation of the error u test whether $\lambda = 0$. Therefore, the results of the LM tests point to the most appropriate spatial model from the spatial autoregressive error model and the spatial autoregressive lag model. See Anselin (1988) for detailed LM test equations.

These LM tests, however, have been improved by Bera and Yoon (1993) in order to be able to conduct specification testing with locally misspecified alternatives. The experimental simulation results of Anselin and Florax (1995) and Anselin et al (1996) show that these robust LM tests have more power to identify the most appropriate spatial model. The two variants of the robust LM tests are (1) a test for spatial error autocorrelation in the presence of a spatially lagged dependent variable, and (2) a test for endogenous spatial lag dependence in the presence of spatial error autocorrelation. The robust tests are similar to their classical counterparts, although they are extended with a correction factor to account for the local misspecification. Bera and Yoon (1993) provide detailed test equations of the robust tests.

There are two alternative spatial models to choose from once the source of spatial autocorrelation is determined. Spatial correlation among the dependent variables is defined as a spatial lag situation which is specified by the spatial autoregressive lag model (SAR)

$$\begin{aligned} &y = \rho W\, y + X\beta + \varepsilon, \\ &\varepsilon \sim N(0, \sigma^2 I_n), \end{aligned}$$

where y is a vector of dependent variables, X is a matrix of independent variables, ρ is the spatial autoregression coefficient and W is a spatial weight matrix.

When spatial dependence exists in the error term, a spatial autoregressive error model (SEM) is employed. The SEM model takes the form

$$\begin{aligned} &y = X\beta + u, \\ &u = \lambda W\, u + \varepsilon, \\ &\varepsilon \sim N(0, \sigma^2 I_n), \end{aligned}$$

where y is a vector of dependent variables, X is a matrix of independent variables, and W is a spatial weight matrix accounting for correlation in the error terms across space, and λ is the error correlation coefficient.

This article makes use of the specification search strategy proposed by Florax et al. (2003, p.562), incorporating all of the above-mentioned spatial econometric considerations. This hybrid specification strategy is based on the combined use of the classical and robust tests for spatial dependence. The steps involved in the approach are as follows:

1. Estimate the initial model by means of OLS.
2. Test the hypothesis of no spatial dependence due to an omitted spatial lag or due to spatially autoregressive errors using LM diagnostics (LMlag and LMerror tests).
3. If both tests are not significant, retain initial estimates from step 1 as final. Otherwise continue to step 4.
4. If both tests are significant, estimate the specification pointed to by the more significant of the two robust versions of the LM diagnostics. If RLMlag > RLMerror then estimate SAR model using MLLAG[3]. If RLMerror > RLMlag then estimate SEM model using MLERROR[4]. Otherwise continue to step 5.
5. If LMlag is significant but LMerror is not, estimate SAR using MLLAG. Otherwise continue to step 6.
6. Estimate SEM using MLERROR.

5. Empirical results

5. 1. Descriptive statistics

Summary statistics for the price per square metre, structural attributes of apartments, and location characteristics are included in Table 2. The average apartment in the sample is an approximately 87 square metre apartment with two to three rooms, located about 4.4 kilometres away from the CBD, and costs roughly 2581 euros per square metre. Most of the apartment units in the sample are located in the 12th district. The lowest price per square metre of apartment is 180 euros, although it can go up as high as 36,000 euros. Most apartments are on the 1st floor, have one toilet and are in good condition. The majority of the buildings in which these apartments are provide an elevator. The apartment unit that is farthest from the CBD is located 13.4 kilometres away from the centre of the city. The bulk of apartment units is located in the outskirts (outside the outer ring).

[3] Maximum likelihood estimator for the model including a spatially lagged dependent variable
[4] Maximum likelihood estimator for the spatially autoregressive error model

Table 2 Description of variables

Variable	Description	Scale type	Mean	Median	Min	Max
ppsqm	price per square metre of apartment (EUR)	ratio	2581.2	2551	180.3	36000
PLZ	district in which the apartment is located	nominal		1120	1010	1230
roomscount	number of rooms in the apartment	nominal	2.59	2	0	8
livingarea	size of the living area (square meters)	ratio	86.56	74.23	18	982
floor	number of floor the apartment is located on	ordinal		1	0	5
cond	condition of the apartment	ordinal		1	1	3
toiletcount	number of toilets in the apartment	nominal		1	0	5
bathroomcount	number of bathrooms in the apartment	nominal		0	0	4
balc	availability of a balcony	binary		0	0	1
terra	availability of a terrace	binary		0	0	1
elev	availability of an elevator (in the building)	binary		1	0	1
basem	availability of a basement (in the building)	binary		0	0	1
furn	if the apartment is furnished	ordinal		0	0	2
flooring	type of flooring	binary		0	0	1
distancecbd	distance from the centre of the city	ratio	4.3941	3.6633	0.044	13.4358
location	If the apartment is located inside the CBD	ordinal	1.625	2	0	2

5. 2. Estimation of the base model

The starting point of the empirical analysis is to estimate the base model. As mentioned in Section 3, this study estimates three regression analyses incorporating the three possible distance variables, namely the district in which the apartment units are located, distance from a predefined point in the centre of the city to each apartment unit, and a location dummy variable with three categories based on proximity to the CBD. Results for all three models based on a semi-log specification are shown in Table 3. The first model includes the location variable district (PLZ), the second model the three-level dummy variable, and the third model the distance from the CBD as follows:

$$\begin{aligned}\log(ppsqm)_i = {} & \alpha_0 + \alpha_1(PLZ_i) + \alpha_2(roomscount_i) + \alpha_3(floor_i) + \alpha_4(cond_i) + \alpha_5(toiletcount_i) + \\ & \alpha_6(bathroomcount_i) + \alpha_7(balc_i) + \alpha_8(terra_i) + \alpha_9(elev_i) + \alpha_{10}(basem_i) + \\ & \alpha_{11}(furn_i) + \alpha_{12}(flooring_i) + \varepsilon_i\end{aligned}$$

$$\log(ppsqm)_i = \alpha_0 + \alpha_1(location_i) + \alpha_2(roomscount_i) + \alpha_3(floor_i) + \alpha_4(cond_i) + \alpha_5(toiletcount_i) + \alpha_6(bathroomcount_i) + \alpha_7(balc_i) + \alpha_8(terra_i) + \alpha_9(elev_i) + \alpha_{10}(basem_i) + \alpha_{11}(furn_i) + \alpha_{12}(flooring_i) + \varepsilon_i$$

$$\log(ppsqm)_i = \alpha_0 + \alpha_1(distance_i) + \alpha_2(roomscount_i) + \alpha_3(floor_i) + \alpha_4(cond_i) + \alpha_5(toiletcount_i) + \alpha_6(bathroomcount_i) + \alpha_7(balc_i) + \alpha_8(terra_i) + \alpha_9(elev_i) + \alpha_{10}(basem_i) + \alpha_{11}(furn_i) + \alpha_{12}(flooring_i) + \varepsilon_i$$

Table 3 Regression results on the impact of hedonic characteristics on log of price per square metre

Dependent variable: log of price per square metre of apartment			
Explanatory variables	*Model 1*	*Model 2*	*Model 3*
Constant	3.8402*** (0.0197)	3.8685*** (0.0238)	3.5379*** (0.0209)
PLZ 1020	-0.2894*** (0.0127)		
PLZ 1030	-0.3530*** (0.0139)		
PLZ 1040	-0.3002*** (0.0163)		
PLZ 1050	-0.3974*** (0.0147)		
PLZ 1060	-0.3121*** (0.0145)		
PLZ 1070	-0.2831*** (0.0143)		
PLZ 1080	-0.2627*** (0.0170)		
PLZ 1090	-0.2950*** (0.0165)		
PLZ 1100	-0.5224*** (0.0139)		
PLZ 1110	-0.5276*** (0.0197)		
PLZ 1120	-0.4139*** (0.0134)		
PLZ 1130	-0.3051*** (0.0138)		
PLZ 1140	-0.3832*** (0.0141)		
PLZ 1150	-0.4453*** (0.0138)		
PLZ 1160	-0.4331*** (0.0142)		
PLZ 1170	-0.3871*** (0.0157)		
PLZ 1180	-0.3075*** (0.0146)		
PLZ 1190	-0.2823*** (0.0135)		

PLZ 1200	-0.4206*** (0.0176)		
PLZ 1210	-0.4652*** (0.0134)		
PLZ 1220	-0.4392*** (0.0171)		
PLZ 1230	-0.3660*** (0.0142)		
Location_outside inner ring		-0.3095*** (0.0150)	
Location_outside outer ring		-0.3991*** (0.0150)	
Distance_cbd			-0.0099*** (0.0009)
Rooms count 1	-0.0248 (0.0132)	-0.0430** (0.0158)	-0.0357* (0.0178)
Rooms count 1.5	0.0826 (0.0433)	0.0740 (0.0525)	0.0782 (0.0591)
Rooms count 2	0.0053 (0.0110)	-0.0126 (0.0132)	-0.0145 (0.0149)
Rooms count 2.5	-0.0631 (0.0391)	-0.0688 (0.0473)	-0.0957 (0.0534)
Rooms count 3	0.0192 (0.0109)	0.0092 (0.0131)	0.0077 (0.0148)
Rooms count 3.5	-0.0039 (0.0360)	-0.0421 (0.0435)	-0.0254 (0.0490)
Rooms count 4	0.0351** (0.0111)	0.0216 (0.0133)	0.0204 (0.0150)
Rooms count 4.5	0.0041 (0.0846)	-0.0091 (0.1026)	0.0018 (0.1156)
Rooms count 5	0.0342** (0.0125)	0.0288 (0.0151)	0.0349* (0.0170)
Rooms count 5.5	-0.0335 (0.0850)	0.0522 (0.1029)	0.0334 (0.1161)
Rooms count 6	0.0355 (0.0213)	0.0278 (0.0258)	0.0283 (0.0291)
Rooms count 7	0.1367*** (0.0306)	0.1062** (0.0370)	0.1153** (0.0417)
Rooms count 8	0.1364* (0.0607)	0.2401** (0.0735)	0.2155** (0.0828)
Floor 1	-0.0071 (0.0056)	-0.0146* (0.0067)	-0.0124 (0.0076)
Floor 2	-0.0054 (0.0061)	-0.0170* (0.0074)	-0.0106 (0.0083)
Floor 3	0.0021 (0.0065)	-0.0162* (0.0078)	-0.0106 (0.0087)
Floor 4	0.0189** (0.0067)	-0.0131 (0.0078)	-0.0013 (0.0088)
Floor 5	0.0318*** (0.0058)	0.0137* (0.0068)	0.0290*** (0.0077)
Condition_moderate	-0.1019*** (0.0059)	-0.1142*** (0.0071)	-0.1244*** (0.0080)
Condition_bad	-0.1559*** (0.0082)	-0.1680*** (0.0099)	-0.1739*** (0.0112)

	-0.0951***	-0.0932***	-0.0817***
Toilet count 1	(0.0115)	(0.0137)	(0.0155)
	-0.0886***	-0.0620***	-0.0470**
Toilet count 2	(0.0117)	(0.0141)	(0.0158)
	-0.0666***	-0.0058	-0.0044
Toilet count 3	(0.0162)	(0.0192)	(0.0217)
	0.0289	0.0276	0.0472
Toilet count 4	(0.0444)	(0.0537)	(0.0606)
	-0.0115	-0.0452	0.2684***
Toilet count 5	(0.0539)	(0.0653)	(0.0717)
	0.0009	0.0064	0.0049
Bathroom count 1	(0.0051)	(0.0061)	(0.0068)
	0.0286***	0.0369***	0.0458***
Bathroom count 2	(0.0069)	(0.0083)	(0.0093)
	0.0563*	0.0760**	0.0827**
Bathroom count 3	(0.0231)	(0.0277)	(0.0312)
	0.0232	0.0582	0.0609
Bathroom count 4	(0.0503)	(0.0610)	(0.0688)
	0.0403***	0.0277***	0.0044
Balcony	(0.0042)	(0.0049)	(0.0053)
	0.0635***	0.0628***	0.0365***
Terrace	(0.0046)	(0.0055)	(0.0061)
	0.0562***	0.0551***	0.0644***
Elevator	(0.0046)	(0.0054)	(0.0062)
	-0.0024	-0.0035	-0.0011
Basement	(0.0050)	(0.0059)	(0.0066)
	-0.0153**	-0.0310***	-0.0521***
Furnished_partially	(0.0056)	(0.0067)	(0.0075)
	-0.0515***	-0.0841***	-0.1033***
Furnished_fully	(0.0133)	(0.0161)	(0.0181)
	0.0105**	0.0068	0.0185***
Flooring_parquet	(0.0035)	(0.0039)	(0.0044)
R-squared	0.694	0.544	0.421
Adjusted R-squared	0.688	0.539	0.414
F-statistic	129.1	104.5	65.3
Prob(F-statistic)	0.000	0.000	0.000
Number of observations	3365	3365	3365

Source: Author's calculations.
Notes: Coefficient standard errors are given in parentheses. *, ** and *** stand for statistically different from zero at the 5%, 1% and 0.1% significance levels.

Structural coefficient estimates are of the expected sign and generally consistent across the three models. Presence of balcony, terrace, elevator, parquet flooring and the level of apartment in the top floor all contribute positively to apartment price, while moderate condition of the apartment lowers price and bad condition further deteriorates price. Fully or partially furnished apartments generate a lower price, an indication that empty apartments are preferred particularly when it comes to purchase. In general, there is a positive relation between the level (floor) of the apartment and price, indicating that apartments in the higher level are preferred. Presence of a second or a third bathroom increases price while existence of a second or a third toilet decreases price.

Particularly interesting in the context of this article is the impact of location on price. The reference variable in these estimations is location within the 1st district. All the coefficients in *Model 1* district variables 2nd to 23rd (PLZ 1020 to PLZ 1230) are negative and highly significant, indicating that location outside the 1st district, i.e., the CBD, decreases apartment price. In *Model 2*, both coefficients of the variables *location outside inner-ring* and *location outside outer ring* are negative and highly significant. The relatively large coefficient of the latter indicates that a location of apartments farther away from the CBD has a larger negative impact on price. The negative and highly significant distance variable in *Model 3* also demonstrates this negative relation between distance from the city centre and price of an apartment. All the three models point out that residents are willing to bid a higher price for an apartment unit located in the city centre, and the price of a constant-quality apartment unit decreases with an increase in distance from the city centre. These results are consistent with the negative rent gradient of the monocentric model.

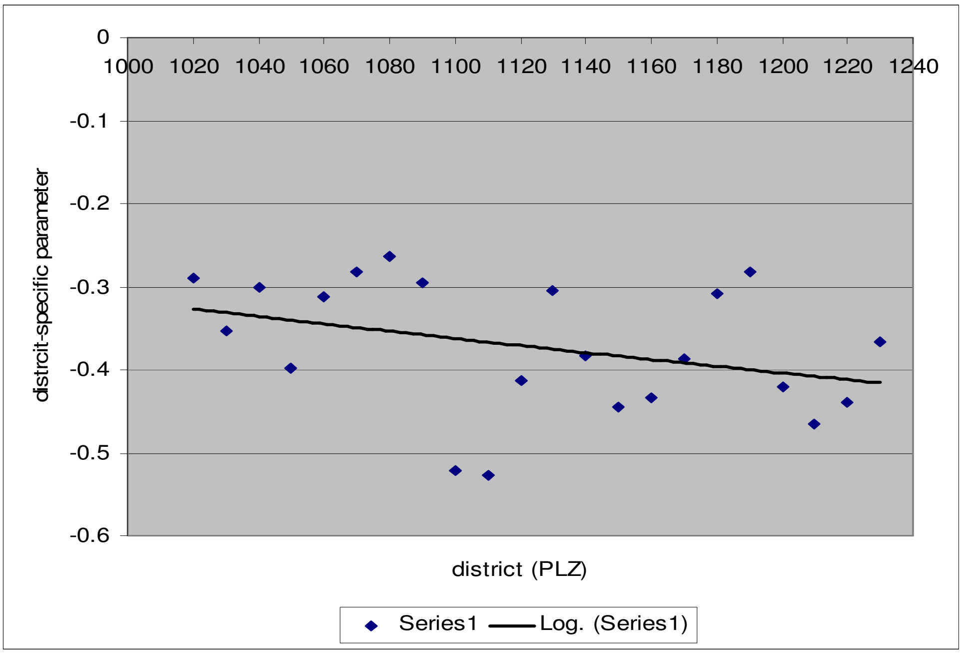

Fig. 3 Districts and district-specific regression parameters. Source: Author’s own compilations.

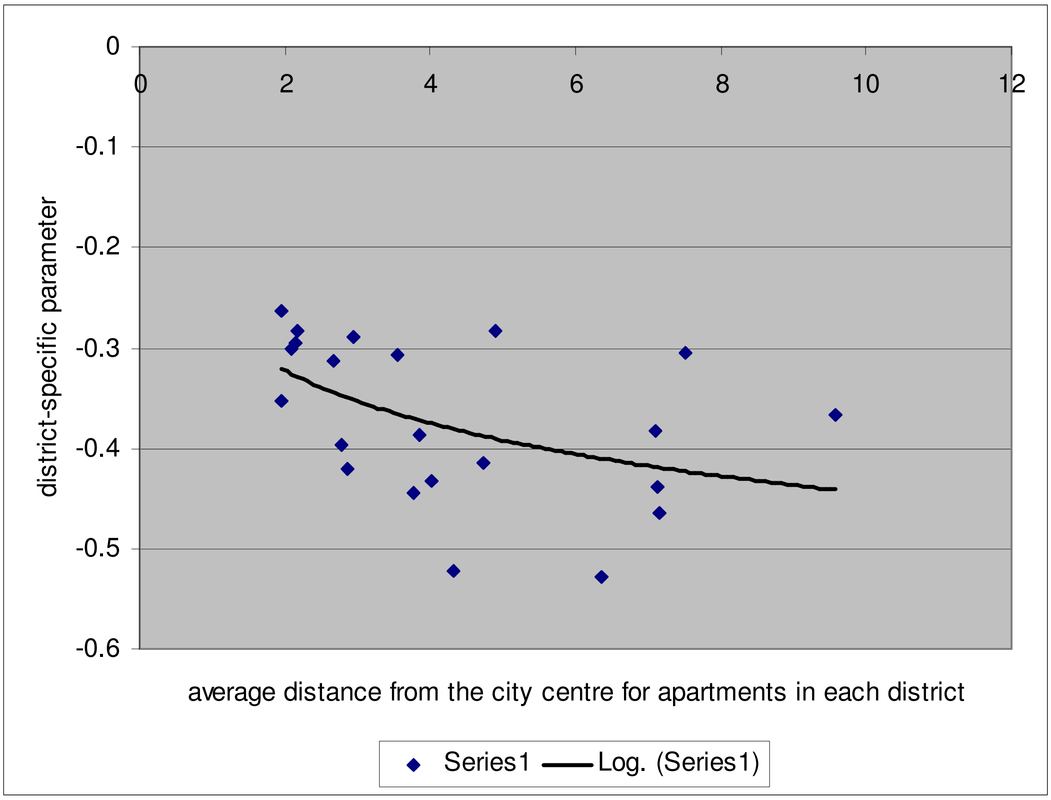

Fig. 3 District-specific regression parameters and average distance from the city centre for apartments in each district. Source: Author's own compilations.

Figure 3 and Figure 4 further elaborate this negative relation between distance from the city centre and apartment price. In Figure 3, district specific parameters are plotted against relevant districts showing that districts in the outskirts have a larger negative impact on price per square metre compared to the neighbouring districts to the CBD. In addition, Figure 4 shows the clear negative relation between district specific parameters and average distance from the city centre for apartments in each district. Note that both these trend lines are very similar to the negative rent gradient portrayed in the classical monocentric model.

5. 3. Spatial analysis

LeSage and Pace (2009, page 3) suppose that a spatial dependence pattern is explained by the model variables distance and density. However, Espey et al. (2007), among many others, maintain that even after accounting for spatial characteristics explicitly, spatial dependence may still exist, resulting in inefficient coefficient estimates. The hedonic house price models estimated above include distance from the city centre or similar proxy variables to capture spatial effects, although it is sensible to estimate a spatial hedonic model to further confirm the results produced by the base model.

In line with the specified methodology, the Moran's I test is a regular first step when testing for spatial effects. This test requires that a spatial weight matrix is constructed based on a reasonable definition of neighbours. Neighbours in this analysis are the apartment units defined by Great Circle distance of less than or equal to 100 metres (0.1 kilometres) from each apartment unit. This mechanism leaves out 364 housing units from tests of spatial autocorrelation due to the fact that these apartment units are without neighbours. The study uses row standardised style (W) spatial weight matrix following the common practise in spatial econometrics. The results are shown in Table 4.

Table 4 Moran's I test for spatial autocorrelation in residuals

	Definition of neighbours -> d1=0, d2=0.1		
	Model 1	*Model 2*	*Model 3*
Moran's I	0.374 (0.000)	0.511 (0.000)	0.582 (0.000)

Notes: The style of the weight matrix is W (row standardised). P-values follow in parentheses.

Moran's I is similar but not equivalent to a correlation coefficient, and its value ranges from −1 (perfect dispersion) to +1 (perfect correlation). The neighbours, based on an appropriate criterion, are specified using a weight matrix within the test equation. The zero value indicates a random spatial pattern, i.e., absence of autocorrelation. More details of the Moran's I test are provided in Anselin (1988). All test statistics presented above provide the same conclusions: the null hypothesis of no spatial effects has to be rejected.

Table 5 LM diagnostics for spatial dependence

	Definition of neighbours -> d1=0, d2=0.1		
	Type of weight matrix -> row standardized (W)		
	Model 1	*Model 2*	*Model 3*
LMerror	1343.81 (2.2e-16)	2511.12 (2.2e-16)	3261.24 (2.2e-16)
LMlag	76.56 (2.2e-16)	82.81 (2.2e-16)	86.44 (2.2e-16)
RLMerror	1313.51 (2.2e-16)	2462.74 (2.2e-16)	3201.26 (2.2e-16)
RLMlag	46.26 (1.034e-11)	34.43 (4.422e-09)	26.46 (2.694e-07)

The Moran's I test is suggestive of spatial effects. Thus, the next step is to apply the LM tests. The LM tests are based on least squares residuals and, make use of a spatial weight matrix which serves the task of incorporating the influence of the price of the nearest neighbouring houses on the price of any given house. Following the tradition, the study uses a row standardized weight matrix for the tests. The results of the LM tests for the three specifications of the spatial hedonic model are shown in Table 5. LM lag (LMlag) and LM error (LMerror) tests examine the hypothesis of no spatial dependence due to an omitted spatial lag and due to spatially autoregressive errors. Both tests are significant, which necessitates the application of robust LM tests. Both robust LM tests are also significant, although the RLMerror statistic is more significant, indicating the presence of spatial correlation in the error term in all the three specifications. Therefore, the specification strategy of Florax et al. (2003) point to the SEM model as the most appropriate spatial model. The results of the three versions of the SEM model estimated using MLERROR are shown in table 6.

Table 6 Spatial regression results on the impact of hedonic characteristics on log of price per square metre

Dependent variable: log of price per square metre of apartment			
Explanatory variables	*Model 1*	*Model 2*	*Model 3*
Constant	3.8074*** (0.0242)	3.8065*** (0.0288)	3.4724*** (0.0171)
PLZ 1020	-0.2950*** (0.0203)		
PLZ 1030	-0.3453*** (0.0219)		
PLZ 1040	-0.3098*** (0.0254)		
PLZ 1050	-0.3908*** (0.0232)		
PLZ 1060	-0.2983*** (0.0230)		
PLZ 1070	-0.2831*** (0.0224)		
PLZ 1080	-0.2488*** (0.0265)		
PLZ 1090	-0.3190*** (0.0247)		
PLZ 1100	-0.5271*** (0.0221)		
PLZ 1110	-0.5248*** (0.0281)		
PLZ 1120	-0.4186*** (0.0216)		
PLZ 1130	-0.2915*** (0.0218)		

PLZ 1140	-0.3905*** (0.0221)		
PLZ 1150	-0.4557*** (0.0219)		
PLZ 1160	-0.4258*** (0.0225)		
PLZ 1170	-0.4197*** (0.0247)		
PLZ 1180	-0.3090*** (0.0229)		
PLZ 1190	-0.2781*** (0.0213)		
PLZ 1200	-0.4396*** (0.0251)		
PLZ 1210	-0.4622*** (0.0215)		
PLZ 1220	-0.4117*** (0.0252)		
PLZ 1230	-0.3774*** (0.0221)		
Location_outside inner ring		-0.3185*** (0.0249)	
Location_outside outer ring		-0.4072*** (0.0247)	
Distance_cbd			-0.0090*** (0.0014)
Rooms count 1	0.0031 (0.0121)	0.0046 (0.0127)	0.0076 (0.0131)
Rooms count 1.5	0.0492 (0.0347)	0.0418 (0.0357)	0.0442 (0.0367)
Rooms count 2	0.0183 (0.0101)	0.0196 (0.0106)	0.0226* (0.0110)
Rooms count 2.5	-0.0347 (0.0320)	-0.0343 (0.0331)	-0.0396 (0.0340)
Rooms count 3	0.0257* (0.0102)	0.0291** (0.0107)	0.0311** (0.0110)
Rooms count 3.5	0.0249 (0.0309)	0.0149 (0.0324)	0.0159 (0.0335)
Rooms count 4	0.0392*** (0.0103)	0.0424*** (0.0108)	0.0437*** (0.0112)
Rooms count 4.5	0.0568 (0.0688)	0.0610 (0.0715)	0.0643 (0.0737)
Rooms count 5	0.0405*** (0.0114)	0.0425*** (0.0119)	0.0459*** (0.0123)
Rooms count 5.5	0.0686 (0.0628)	0.0877 (0.0629)	0.0944 (0.0640)
Rooms count 6	0.0060 (0.0185)	0.0037 (0.0193)	0.0082 (0.0199)
Rooms count 7	0.1075 (0.0275)	0.1003*** (0.0288)	0.1033*** (0.0298)
Rooms count 8	0.0536 (0.0517)	0.0584 (0.0544)	0.0476 (0.0564)
Floor 1	-0.0074 (0.0051)	-0.0113* (0.0054)	-0.0099 (0.0055)

Floor 2	-0.0011 (0.0056)	-0.0049 (0.0059)	-0.0024 (0.0061)
Floor 3	0.0079 (0.0058)	0.0031 (0.0062)	0.0073 (0.0064)
Floor 4	0.0253*** (0.0062)	0.0162* (0.0065)	0.0215** (0.0067)
Floor 5	0.0501*** (0.0054)	0.0447*** (0.0057)	0.0490*** (0.0059)
Condition_moderate	-0.0884*** (0.0054)	-0.0929*** (0.0056)	-0.0922*** (0.0058)
Condition_bad	-0.1344*** (0.0075)	-0.1391*** (0.0078)	-0.1402*** (0.0080)
Toilet count 1	-0.0768*** (0.0112)	-0.0667*** (0.0118)	-0.0701*** (0.0123)
Toilet count 2	-0.0706*** (0.0115)	-0.0535*** (0.0122)	-0.0565*** (0.0127)
Toilet count 3	-0.0576*** (0.0156)	-0.0374* (0.0166)	-0.0451** (0.0172)
Toilet count 4	0.0849* (0.0364)	0.1242** (0.0378)	0.1239** (0.0390)
Toilet count 5	0.0493 (0.0870)	0.0470 (0.1241)	0.3841** (0.1447)
Bathroom count 1	-0.0013 (0.0046)	-0.0006 (0.0049)	-0.0040 (0.0050)
Bathroom count 2	0.0200** (0.0063)	0.0211** (0.0066)	0.0176* (0.0068)
Bathroom count 3	0.0134 (0.0210)	0.0250 (0.0221)	0.0189 (0.0229)
Bathroom count 4	0.0105 (0.0405)	0.0190 (0.0422)	0.0133 (0.0436)
Balcony	0.0387*** (0.0040)	0.0387*** (0.0042)	0.0350*** (0.0043)
Terrace	0.0556*** (0.0044)	0.0566*** (0.0047)	0.0527*** (0.0048)
Elevator	0.0464*** (0.0047)	0.0408*** (0.0050)	0.0440*** (0.0053)
Basement	-0.0043 (0.0051)	-0.0032 (0.0055)	-0.0016 (0.0058)
Furnished_partially	-0.0081 (0.0057)	-0.0098 (0.0061)	-0.0153* (0.0063)
Furnished_fully	-0.0215 (0.0122)	-0.0244 (0.0127)	-0.0328* (0.0131)
Flooring_parquet	0.0175*** (0.0036)	0.0143*** (0.0039)	0.0197*** (0.0040)
λ	0.563	0.687	0.729
LR test value	853.75***	1650.9***	2132.5***
Log likelihood	4037.46	3767.74	3604.58
AIC	-7952.9	-7453.5	-7129.2
Number of observations	3365	3365	3365

Source: Author's calculations.

Notes: Coefficient standard errors are given in parentheses. *, ** and *** stand for statistically different from zero at the 5%, 1% and 0.1% significance levels.

The spatial hedonic model is perceived as an improved model since it captures the spatial effects that are present in the base model. The results obtained for all the three specifications of the spatial model are very similar to those of the corresponding basic models. The structural characteristics top floor, presence of balcony, terrace, elevator and parquet floor have a positive influence on price, while moderate and bad condition of an apartment affect price negatively. All three specifications of the spatial model provide evidence supporting these insightful results.

More importantly, the estimated coefficients of location variables employed in the spatial regression analysis empirically verify the monocentric structure of Vienna. The coefficients of the location variable PLZ (district) in *Model 1* are negative for all the districts away from the CBD, indicating that prices decrease moving away from the centre. Similarly, the categorical location variable employed in *Model 2* to capture the effects of location within the neighbouring districts to the CBD is significant and shows the expected negative sign. The categorical variable representing location in the outskirts is also significant and even has a large impact on price, which is consistent with the idea that constant-quality apartments located further away from the city centre are relatively inexpensive. The coefficient of the location variable in *Model 3*, i.e., distance from the CBD, provides additional support to substantiate the claim of a negative rent gradient for the apartment market in Vienna.

The calculation of impact measures is needed in order to interpret the regression coefficients correctly in the context of spatial lag and spatial Durbin models[5] because of the spillovers between the terms in these data-generation processes. However, this step is not required with regard to the present analysis since the spatial error model is employed.

6. Conclusion

There are different approaches to study house price dynamics. The monocentric theory in urban economics is a classical yet empirically useful model among them. Early studies using this model as the theoretical foundation produced evidence confirming a negative rent gradient. Subsequently, the multicentric model became predominant with

[5] Spatial Durbin model adds average-neighbour values of the independent variables to the spatial lag specification.

the emergence of the idea that cities evolve with polycentric developments. This study, using a version of the monocentric model, provides insights into the relation between spatial proximity to the city centre and the price of apartments. The unique dataset used in the analysis was provided by a popular online database of apartment units available for sale in Vienna.

The present paper is an application of the spatial hedonic house price model. Hedonic price theory, also known as hedonic regression, estimates the implicit price of utility-bearing characteristics of composite commodities. Early hedonic house price models, for instance, primarily included structural attributes of houses in the model specification although spatial hedonic models surfaced subsequently with location variables such as distance and travel time from the city centre and, travel cost. Nonetheless, it is shown in spatial econometrics that when the observations are regions or points in space, the conventional assumption of independent observations and errors in linear regression models may be violated. This requires that hedonic models are tested for spatial autocorrelation, and, if detected, spatial effects are incorporated into the model specification. Classical tests such as Moran's I, and new innovations such as LM tests, are useful for testing spatial autocorrelation. This paper makes use of a combination of these classical and new techniques. The methodology applied involves estimating the base model, testing spatial autocorrelation, specification search, and, based on the results of the econometric tests, estimating the spatial model. The paper borrows from the specification search strategy proposed by Florax et al. (2003) in choosing the most appropriate spatial model in the context of data used in the article.

The estimated base model suggests that structural characteristics such as top floor, availability of a balcony, terrace, elevator and parquet flooring have a positive impact on price. Moderate or bad condition of the apartment unit negatively affect price. The coefficients of the location variable in all three specifications are negative, providing proof for the negative rent gradient suggested by the monocentric model. Nevertheless, substantial spatial autocorrelation was detected in the residual series. The LM and RLM test results point to the spatial error model (SEM) as the most appropriate specification. The estimated spatial error model also provides overwhelming support for the negative rent gradient, confirming findings of the base model. The intrinsic preference of residents to live close to the centre primarily caused by the radial transport system and prominence of the central district as the economic centre is seen as a possible explanation in the context of Vienna. This paper adds to the literature by indicating that distance from the city centre matters when it comes to residential location.

References

Alonso, W. (1964), Location and Land Use, Harvard University Press, Cambridge, Mass.

Anselin, L. (1988), Spatial Econometrics: Methods and Models, Kluwer, Dordrecht.

Anselin, L. and Florax, R.J.G.M. (1995), "Small sample properties of tests for spatial dependence in regression models: some further results", in Anselin, L. and Florax, R.J.G.M. (Eds.), New Directions in Spatial Econometrics, Springer, Berlin, pp. 21-74.

Anselin, L., Bera, A., Florax, R.J.G.M. and Yoon, M. (1996), "Simple diagnostic tests for spatial dependence", Regional Science and Urban Economics, Vol. 26, pp. 77-104.

Ball, M. (1973), "Recent empirical work on the determinants of relative house prices", Urban Studies, Vol. 10, pp. 213-233.

Bender, B. and Hwang, H. (1985), "Hedonic house price indices and secondary employment centers", Journal of Urban Economics, Vol. 17, pp. 90-107.

Bera, A. and Yoon, M. (1993), "Specification testing with locally misspecified alternatives", Econometric Theory, Vol. 9, pp. 649-658.

Brunauer, W.A., Lang, S., Wechselberger, P. and Bienert, S. (2009), "Additive hedonic regression models with spatial scaling factors: an application for rents in Vienna", Journal of Real Estate Finance and Economics, Vol. 41 No. 4, pp. 390-411.

Dubin, R.A. (1992), "Spatial autocorrelation and neighborhood quality", Regional Science and Urban Economics, Vol. 22, pp. 433-452.

Dubin, R.A. and Sung, C.H. (1987), "Spatial variation in the price of housing: rent gradients in non-monocentric cities", Urban Studies, Vol. 24, pp. 193-204.

Espey, M., Fakhruddin, F., Gering, L.R. and Lin, H. (2007), "Living on the edge: residential property values in the urban-rural interface", Journal of Agricultural and Applied Economics, Vol. 39 No. 3, pp. 689-699.

Florax, R.J.G.M., Folmer, H. and Rey, S.J. (2003), "Specification searches in spatial econometrics: the relevance of Hendry's methodology", Regional Science and Urban Economics, Vol. 33, pp. 557-579.

Hamilton, B. (1976), "The effects of property taxes and local public spending on property values: a theoretical comment", Journal of Political Economy, Vol. 84, pp. 647-650.

Haynes, K.E. and Fotheringham, A.S. (1984), Gravity and Spatial Interaction Models, Sage Publications, California.

Herath, S. and Maier, G. (2010), "The hedonic price method in real estate and housing market research: a review of the literature", SRE-Discussion 2010/03, Institute for the Environment and Regional Development, Vienna University of Economics and Business, Vienna.

Jackson, J.R. (1979), "Intra-urban variation in the price of housing", Journal of Urban Economics, Vol. 6, pp. 464-479.

LeSage, J.P. and Pace, R.K. (2009), Introduction to Spatial Econometrics, Taylor & Francis, Boca Raton.

Mills, E.S. (1967), "Transportation and patterns of urban development", American Economic Review, pp. 197-210.

Mills, E.S. (1972a), Studies in the Structure of the Urban Economy, Johns Hopkins University Press, Baltimore, MD.

Mills, E.S. (1972b), Urban Economics, Scott Foresman, Glenview, Illinois.

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

Olmo, J.C. (1995), "Spatial estimation of housing prices and locational rents", Urban Studies, Vol. 32 No. 8, pp. 1331-1344.

Richardson, H. (1988), "Monocentric vs. policentric models: the future of urban economics in regional science, Annals of Regional Science, Vol. 22, pp. 1-12.

Rosen, S. (1979), "Wage based indexes of urban quality of life", in Mieszkowski, P. and Straszheim, M. (Eds.), Current Issues in Urban Economics, Johns Hopkins Univ. Press, Baltimore, MD, pp. 74-104.

Thünen, V. (1826), The Isolated State, Volumes I.

Tobler, W. (1970), "A computer movie simulating urban growth in the Detroit region", Economic Geography, Vol. 46 No. 2, pp. 234-240.

Wilhelmsson, M. (2002), "Spatial models in real estate economics", Housing, Theory and Society, Vol. 19, pp. 92-101.

Yinger, J. (1982), "Capitalization and the theory of local public finance", Journal of Political Economy, Vol. 90, pp. 917-943.