TEXTO PARA DISCUSSÃO Nº 101

HEDONIC PRICE MODELS WITH SPATIAL EFFECTS: AN APPLICATION TO THE HOUSING MARKET OF BELO HORIZONTE, BRAZIL

Paulo Brígido Rocha Macedo

Junho de 1996

Ficha catalográfica

330.115 M141h 1996	MACEDO, Paulo Brígido Rocha. Hedonic price models with spatial effects: an application to the housing market of Belo Horizonte. Belo Horizonte. Belo Horizonte : FACE/CEDEPLAR, 1996. 24p. (Texto para discussão; 101) 1. Economia espacial. 2. Modelos de preços hedônicos. 3. Mercado Imobiliário - Belo Horizonte (MG). I. Título. II. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. III. Série.

Versão preliminar não sujeita a revisão.

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL

HEDONIC PRICE MODELS WITH SPATIAL EFFECTS: AN APPLICATION TO THE HOUSING MARKET OF BELO HORIZONTE, BRAZIL

Paulo Brígido Rocha Macedo
Department of Economics, and Center for Regional Development and Planning (CEDEPLAR), Federal University of Minas Gerais, Brazil

CEDEPLAR/FACE/UFMG
BELO HORIZONTE
1996

SUMÁRIO

I INTRODUCTION

Hedonic price models (HPM), first developed in the 1960s, have frequently been applied to the study of housing markets. In the early 1990s, some studies incorporated spatial econometrics in HPM housing models to address methodological issues ignored in the standard approach.

This paper uses both standard and spatial autoregressive hedonic price models to analyze sample data from the housing market of Belo Horizonte, Brazil. Among the spatial econometric tools used are diagnostic tests for the detection of spatial dependence and heterogeneity, which provide the means to identify adjacency effects in the determination of housing prices. The paper also tests a number of alternative functional forms for both the standard and the spatial HPM models, using the Box-Cox transformation of the variables analyzed.

The results show that spillover (adjacency) effects are an important source of price variation in the residential apartment market in Belo Horizonte. The empirical findings support the need to incorporate spatial effects in studies of housing price determination. The importance of incorporating spatial factors makes the use of special spatial econometrics techniques essential to the analysis of economic performance in housing markets.

II LITERATURE REVIEW

II.1 Hedonic Price Models

According to Griliches (1971), "the hedonic characteristics approach to the construction of price indexes is based on the empirical hypothesis that the multitude of varieties (or models) of a particular commodity can be comprehended in terms of a much smaller number of basic attributes." Early examples of this approach include the empirical analysis of automobile prices by Griliches (1961) and Dhrymes (1967), the study of the real estate construction market by Bailey, Muth and Nourse (1963), and the analysis of technological change in the computer mainframe industry by Chow (1967). In the 1970s empirical applications, as well as theoretical developments, firmly established the hedonic price models (HPM) approach in the literature. A major contribution is that of Rosen (1974) who develops a theoretical framework "based on the hedonic hypothesis that goods are valued for their utility-bearing characteristics." Accordingly, "hedonic prices are defined as the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them."

Empirical analyses based on the hedonic approach must address two the following questions first proposed by Griliches (1971):

1) What are the relevant characteristics?
2) What is the form of the relationship between prices and characteristics?

With respect to the first question, the early HPM studies on automobile prices used three car characteristics: size, power, and accessories; Chow's (1967) analysis of the mainframe computer industry had two characteristics: memory capacity and speed of the instruction cycle. Urban housing markets, however, present a much larger number of potentially relevant characteristics. Butler (1982) notes that "data on many of these characteristics are either unavailable or of exceedingly poor quality. Even without data constraints, the intrinsic clustering of characteristics combinations into a relatively small number of configurations leads to considerable multicollinearity in estimates employing a generous selection of the relevant variables". Echoing the warning of Griliches (1971) against "the use of variables which are not direct characteristics of the commodity but an outcome of the market experiment," Butler comments that "this is the case of a number of studies on urban housing markets in which income and other demander characteristics were intended as proxies for neighborhood quality." He analyzes the specification bias cost of employing a simple model (four unit-specific variables) rather than a more extensive model which adds a list of demographic variables. Comparing the two hedonic indexes estimated over the same data base of a single metropolitan area at a given point in time, the empirical findings indicate little practical impact on the specification bias of the more restricted model.

With regard to Griliches' second question, the specification of the functional form in the price-characteristics relationship, a number of HPM studies in the literature use linear, semilog (dependent variable, price, being logarithmic), or doublelog functional forms. These provide straight interpretations of the estimated coefficients, respectively: the implicit marginal characteristic prices prevailing at a given market equilibrium; the percentage change in the commodity price for a unit change in any of its characteristics; the percentage change in the commodity price for a percentage change in any of its characteristics (elasticity). Economic theory has not yet developed criteria for the choice of functional forms, so most researchers view the choice as an empirical question to be decided by the best data fit. Many HPM applications to urban housing markets have used the Box-Cox tranformation-of-variables procedure to allow for the possibility that the best functional form could be non-linear.

II.2 Hedonic Price Models with Spatial Effects

The field of Regional Science and Urban Economics addresses issues related to human spatial behavior in cities, regions, and major geopolitical areas. Standard econometric techniques can be used in the statistical analysis of spatial interaction models and the calibration of regional econometric models, but there are specific aspects of spatial data that are beyond the reach of these techniques.[1] Anselin (1988) calls such aspects "spatial effects," whose principal types are: "spatial dependence," also called spatial autocorrelation or association, and "spatial heterogeneity." Spatial econometrics models takes these spatial effects explicitly into account.

[1] See Anselin (1988) and Anselin (1992).

With regard to the first effect, spatial dependence, Cliff and Ord (1973) state that if the presence of a phenomenon in one area (district, city, region) changes the likelihood of its presence in neighboring areas, the phenomenon is said to exhibit spatial autocorrelation. In other words, "Everything is related to everything else, but near things are more related than distant things."[2] Spatial dependence may result from the arbitrariness of borderlines between spatial units of observation such as districts, cities, states; the presence of spatial externalities such as shared neighborhood characteristics which affect housing prices; and/or spillover effects such as the impact of the price of one housing unit on the price of its adjacent neighbors.

The second spatial effect, spatial heterogeneity, refers to spatial variability (structural instability) in the parameters or even in the functional form. For example, a cross-sectional data set with very different spatial units, such as rich areas in a southern region and poor areas in the north, may exhibit spatial heterogeneity effects. Although standard econometric techniques such as switching regressions can cope with many of the problems of spatial heterogeneity, there are instances (such as error terms presenting spatial dependence) in which acknowledgement of the underlying spatial structure can improve the efficiency of the estimation procedures.

Only recently, in the early 1990s, have spatial econometrics techniques been used to study hedonic price models, specifically taking into account spatial effects. Can (1990, 1992) specifies a model in which the price of a housing unit in any location depends not only on its structural and neighborhood characteristics, as in the traditional HPM approach, but on the prices of adjacent units. (The model allows for checking of the strength of the price interdependence.) This approach closely resembles actual characteristics of an urban housing market, where realtors appraise housing units by their relevant individual characteristics and also by the price history of neighboring units.

III THEORETICAL CONSIDERATIONS

This work examines the observed marginal hedonic prices of a sample of apartment units in the city of Belo Horizonte for a number of alternative functional form specifications of a Hedonic Price Model (HPM). As Rosen (1974) points out, hedonic prices connect equilibrium reservation prices (minimum price for any package of characteristics), and characteristics, but require more information to identify the underlying supply and demand functions. Follain and Jimenez (1985) meet these information requirements by assuming a generalized constant elasticity-of-substitution utility function for the representative agent (no heterogeneous preferences allowed), whose arguments are a vector of housing characteristics and a composite of all other goods. Follain and Jimenez (FJ) use the Box-Cox technique to search for the best functional form for the hedonic function. Their equation (FJ,4), reproduced below, represents the price-characteristics relationship:

$$\frac{p^{\lambda}-1}{\lambda} = \beta_0 + \sum_1^m \beta_j Z_j + u$$

[2] Tobler (1979), "first law of geography," quoted in Anselin (1988).

where P is the market price of the housing unit, β is a vector of m regression coefficients, Z is a vector of m housing characteristics, u is a vector of error terms, and λ is a parameter used to transform P to do Box-Cox analysis.

After estimating the parameters of the Box-Cox functional form, FJ express $P_i = \partial P/\partial Z_i$, the unobserved marginal price of the ith characteristic, as:

$$\hat{P}_i = \hat{\beta}_i P^{(1-\hat{\lambda})},$$

derived from their estimated equation (FJ,4) through algebraic manipulation. It is worth noting that $\lambda=1$ yields the linear relationship where the estimated P_i is equal to β_i, the only instance in which it does not vary with the observation.

IV DESCRIPTION OF THE MODEL

This paper uses the hedonic price relationship in its reduced form, with a general Box-Cox model in two versions: one for the standard HPM approach and another which allows for spatial effects. The model for the standard approach is as follows:

$$\frac{P^{\lambda}-1}{\lambda} = \beta_0 + \sum_1^m \beta_i\left(\frac{Z_i^{\theta}-1}{\theta}\right) + u \qquad (1)$$

where P is the price of the housing unit, β is a vector of m regression coefficients, Z is a vector of m housing characteristics, u is a vector of error terms, and λ and θ are parameters used to transform P and Z, respectively, to do Box-Cox analysis.

The expression for P_i is derived analogously as that of FJ:

$$\hat{P} = [(1 + \hat{\lambda}\hat{\beta}_0) + \hat{\lambda}\sum_1^m \hat{\beta}_j \frac{(Z_j^{\hat{\theta}} - 1)}{\hat{\theta}}]^{1/\hat{\lambda}}$$

$$\frac{\partial P}{\partial Z_i} = \hat{P}_i = \hat{\beta}_i Z_i^{(\hat{\theta}-1)}[(1 + \hat{\lambda}\hat{\beta}_0) + \hat{\lambda}\sum_1^m \hat{\beta}_j \frac{(Z_j^{\hat{\theta}} - 1)}{\hat{\theta}}]^{1/\hat{\lambda}}$$

$$\hat{P}_i = \hat{\beta}_i Z_i^{(\hat{\theta}-1)} \hat{P}^{(1-\hat{\lambda})} \qquad (2)$$

For any λ, $\theta=1$ yields the hedonic equation linear in the characteristics analyzed by FJ. $\lambda=1$ and $\theta=1$ yield the linear hedonic equation whose coefficents are the implicit marginal characteristic price. For any other pair of values of λ and θ, the estimated value of the unobserved marginal price of the ith characteristic varies by observation, but a mean value may be computed by averaging the P_is corresponding to each observation.

To analyze spatial effects in the hedonic price-characteristic relationship, the following spatial lag model[3] is considered:

---

[3] See Anselin (1988) for a full discussion.

$$\frac{P^{\lambda}-1}{\lambda} = \beta_0 + \rho\lambda W(\frac{P^{\lambda}-1}{\lambda}) + \sum_1^m \beta_i \frac{(Z_i^{\theta} - 1)}{\theta} + \varepsilon \qquad (3)$$

where P is a vector of N observations of the dependent variable, price of a housing unit, β is a vector of m regression coefficients, ρ is the coefficient of the spatially lagged dependent variable, W is an NxN spatial weights matrix, Z is a vector of m housing characteristics, ε is a vector of N error terms, and λ and θ are parameters used to transform P and Z, respectively, to do Box-Cox analysis. The spatial weights matrix W plays a role similar to the time-lag operator in a time-series modeling context. It is built in such a way that each row and matching column correspond to an observation pair, ij, and its entry value signals when these observations i and j are considered to be neighbors. The relevant set of neighbors for each observation can be defined as either those that share a border, i.e. simple contiguity (for areal units), or those that are within a critical distance (for point data), as in the urban housing market).

The rejection of the hypothesis, H_0: ρ=0, implies the existence of adjacency (spillover) effects in the housing market, i.e. the price of one housing unit affects the prices of neighboring units.

If there are no spillover effects but there is spatial dependence in the regression error terms (spatial error model), the OLS estimates remain unbiased but the t- and F- statistics for tests of significance will be biased and the statistical interpretation of the regression model will be incorrect. The spatial error model is specified as follows:

$$P = \beta_0 + \sum_1^m \beta_i Z_i + \varepsilon \qquad (4)$$

$$\varepsilon = \delta W \varepsilon + \nu$$

Where P, β_i, Z_i, W and ε have the same meaning as in equation (3), v is a vector of N independent and normally distributed error terms, and δ is the residual spatial autoregression coefficient. The rejection of the hypothesis δ=0 implies the presence of spatial dependence in the residuals.

V THE DATA

The database analyzed has price and characteristic information for a sample of Belo Horizonte residential apartments lying within a spatial region of approximately 16 square kilometers. The apartments were included in a market survey of residential prices conducted by the Instituto de Pesquisas Econômicas e Administrativas (IPEAD) of the Universidade Federal de Minas Gerais in October 1995. The apartments' characteristics include variables such as apartment area (square meters), age, availability of garage space, local topography, and the level of public services such as piped water, electricity, and garbage collection. Topography is fairly homogeneous for the region studied, with a uniform index assigned to all apartments by city tax assessors, and this characteristic does not affect their relative market value according to realtors. The region is also well-provided with city services, and there is a homogeneous overall index of their availability. For this study, therefore, the sources of price variation, are the area of the housing unit in square meters, its age, and the availability of a garage space.

To build the spatial weights matrix, the geographic information necessary was derived from a city map of scale 1cm = 25,000cm. A binary contiguity weight matrix is constructed which defines a residential "neighbor" as a unit within a distance of 1.5 km.

VI EMPIRICAL RESULTS

The following section presents four specifications of the standard hedonic price model (HPM): semilog, doublelog, linear, and a nonlinear form estimated by means of the Box-Cox transformation procedure, which are commonly used in urban housing market studies. This section also considers the same four specifications when spatial effects are incorporated, as considered by Can (1990, 1992).

The Box-Cox procedure is implemented by iterated ordinary least squares, following Spitzer (1982). The systematic grid search checks the residual sum of squares (RSS) for different pairs of the parameters (λ,θ) within the range [-1.25,1.25] in increments of 0.25, by using the statistical package LIMDEP[4]. The best fit in this search is given by the pair of parameters $\lambda=0.25$ (dependent variable) and $\theta=0$ (independent variables).

Table I presents the summary statistics of the analyzed variables. The transformed price variables for $\lambda=0$ and $\lambda=0.25$ are shown with corresponding mnemonic names P0 and P025, and the transformed variables area ($\theta=0$) and age ($\theta=0$) with names A0 and I0, respectively.

[4] Software by William H. Greene.

TABLE I

SUMMARY STATISTICS

VARIABLE	MEAN	VARIANCE	ST.DEV.	SKEWNESS	KURTOSIS
P	72330.19	2.450735E+09	49504.9	1.860869	6.188621
A	132.4768	4071.908	63.8115	1.412476	5.311006
I	18.03774	89.47027	9.458873	-0.2272928	1.93313
P0	11.00997	0.329855	0.57433	0.4979435	2.871522
A0	4.784981	0.1964776	0.4432579	0.2700276	2.59841
I0	2.643905	0.7244319	0.8511357	-1.395625	4.093176
P-025	3.742347	0.001282378	0.0358103	0.1530793	2.61091

VARIABLE	MIN	MAX	RANGE
P	21500	245000	223500
A	50.87	368	317.13
I	1	34	33
P0	9.975808	12.40901	2.433205
A0	3.92927	5.90808	1.97881
I0	0	3.52636	3.52636
P-025	3.66967	3.82021	0.15054

P= Price in Brazilian Reais (RS)
A=Area of apartment in square meters
I = Age
P0=Transformed price variable (λ=0, i.e., lnP)
A0=Transformed area variable (θ=0, i.e, lnA)
I0= Transformed age variable (θ=0, i.e., lnI).
P-025=Transformed price variable (λ=-0.25)

Tables II - IV present the estimates, for both the standard hedonic price model and the hedonic regression with spatial effects, of the four functional forms considered in this work: semilog, doublelog, linear, and Box-Cox best fit. All estimations were derived with the software SPACESTAT[5], which provides regression diagnostics for multicollinearity, normality, heteroskedasticity, and spatial dependence in the case of the spatial effects model.

The spatial weights matrix W is based on distance. It defines the neighbors of a given housing unit as the set of other units located within a ray of 1.5 kilometers. The presence of spillover (adjacency) effects on housing prices leads to the non-rejction of the null hypothesis of $\rho=0$ in equation (3).

The four specifications analyze the transformed variables: price, P (parameter λ), area of the housing unit, A (parameter θ), age I (parameter), and the indicator of garage availability, G (a binary 0-1 variable in all specifications).

After Box-cox best fit ($\lambda=-0.25$; $\theta=0$), the best statistical results are obtained by the doublelog ($\lambda=0$; $\theta=0$) and semilog ($\lambda=0$; $\theta=1$) specifications, in that order. For all three of these specifications, the explanatory variables Area and Garage are statistically significant. Surprisingly, Age is not statistically significant, although it has the expected negative sign.

Spatial effects are strong in the three best specifications: Box-Cox best fit, semilog, and doublelog. For the first two, the tests for spatial lag dependence (adjacency effects) are highly significant (0.012742 and 0.002481, respectively), whereas the hypothesis of spatial error dependence is rejected (0.062875 and 0.32559, respectively) at the level of 5%. The picture is more complex for the doublelog specification, where both tests are highly significant: the null hypothesis of no spatial dependence is rejected with a probability of 0.034408 and the null hypothesis of no spatial error dependence is also rejected, with a probability of 0.018644.

5 Anselin (1990).

TABLE II

Semi Log
Ordinary Least Squares Estimation

DEPENDENCY VARIABLE		LNP	OBS 53	VARS 4	DF 49
R2	0.7821	R2-adj	0.7687		
LIK	-5.43897	AIC	18.8779 SC	26.7591	
RSS	3.81011	F-test	58.6105	Prob	3.10369e-16
SIG-SQ	0.0777574	(0.278850)	SIG-SQ(ML)	0.0718889	(0.268121

VARIABLE	COEFF	S.D.	T-Value	Prob
CONSTANT	9.94553	0.184192	53.995351	0.000000
A	0.00716456	0.000683761	10.478173	0.000000
I	0.00236667	0.00506867	-0.466922	0.642625
G	0.220356	0.0980662	2.247011	0.029177

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER		10.596769	
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	PROB
Kiefer-Salmon	2	0.194589	0.907289
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Breusch-Pagan test	3	10.623864	0.013944
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	PROB
White	8	17.104759	0.029037

Spatial Lag Model
Maximum Likelihood Estimation

DATA SET	IPTU	SPATIAL WEIGHTS MATRIX		WGH15	
DEPENDENT VARIABLE		LNP	OBS 53	VARS 5	DF48
R2	0.8139	Sq. Corr.	0.8189		
LIK	-0.861880	AIC	11.7238	SC	21.5752
SIG-SQ	0.0597292	(0.244396)			

VARIABLE	COEFF	S.D.	z-value	Prob
W_LNP	0.351693	0.115428	3.046850	0.002313
CONSTANT	6.08671	1.27201	4.785125	0.000002
A	0.00672707	0.000616892	10.904782	0.000000
I	-0.000712603	0.00447558	-0.159220	0.873495
G	0.2223	0.0859522	2.586320	0.009701

REGRESSION DIAGNOSTICS					
DIAGNOSTICS FOR HETEROSKEDASTICITY					
RANDOM COEFFICIENTS					
TEST	DF	VALUE	PROB		
Breusch-Pagan test	3	12.490019	0.005880		
Spatial B-P test	3	12.490032	0.005880		
DIAGNOSTICS FOR SPATIAL DEPENDENCE					
SPATIAL LAG DEPENDENCE FOR WEIGHTS MATRIX			WGH15 (row-standardized weights)		
TEST	DF	VALUE	PROB		
Likelihood Ratio Test	1	9.154185	0.002481		
LAGRANGE MULTIPLIER TEST ON SPATIAL ERROR DEPENDENCE					
WEIGHT	STAND	ZERO	DF	VALUE	PROB
WGH15	yes	no	1	0.966356	0.325591

TABLE III

Double Log
Ordinary Least Squares Estimation

DEPENDENT VARIABLE		LNP	OBS 53	VARS 4	DF 49
R2	0.8170	R2-adj	0.8058		
LIK	-0.801560	AIC	9.60312	SC	17.4843
RSS	3.19844	F-test	72.9429	Prob	4.35431e-18
SIG-SQ	0.0652742	(0.255488)	SIG-SQ(ML)	0.0603479	(0.245658)

VARIABLE	COEFF	S.D.	T-value	Prob
CONSTANT	5.9988	0.487105	12.315223	0.000000
LNA	1.05133	0.0892084	11.785152	0.000000
LNI	-0.0537355	0.0481798	-1.115312	0.270157
G	0.171038	0.0866143	1.974704	0.053951

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER		35.143450	
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	PROB
Kiefer-Salmon			
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Breusch-Pagan test	3	5.346875	0.148090
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	PROB
White	8	11.260179	0.187388

Spatial Lag Model
Maximum Likelihood Estimation

DEPENDENT VARIABLE	LNP	OBS 53	VARS 5	DF 48
R2	0.8328	0.8308	Sq. Corr	
LIK	1.43559	AIC	7.12881	SC 16.9803
SIG-SQ	0.0551611	(0.234864)		

VARIABLE	COEFF	S.D.	z-value	Prob
W_LNP	0.238328	0.107165	2.223943	0.026152
CONSTANT	3.58963	1.15914	3.096813	0.001956
LNA	0.996258	0.086449	11.524226	0.000000
LNI	-0.0401412	0.0447736	-0.896536	0.369967
G	0.176181	0.0796664	2.211481	0.027003

REGRESSION DIAGNOSTICS					
DIAGNOSTICS FOR HETEROSKEDASTICITY					
RANDOM COEFFICIENTS					
TEST		DF	VALUE	PROB	
Breusch-Pagan test		3	6.121818	0.105832	
Spatial B-P test		3	6.121818	0.105832	
DIAGNOSTICS FOR SPATIAL DEPENDENCE					
SPATIAL LAG DEPENDENCE FOR WEIGHTS MATRIX				WGH15 (row-standardized weights)	
TEST		DF	VALUE	PROB	
Likelihood Ratio Test		1	4.474309	0.034408	
LAGRANGE MULTIPLIER TEST ON SPATIAL ERROR DEPENDENCE					
WEIGHT	STAND	ZERO	DF	VALUE	PROB
WGH15	yes	no	1	5.534638	0.018644

TABLE IV

Linear Model
Ordinary Least Squares Estimation

DEPENDENT VARIABLE		P	OBS 53	VARS 4	DF 49
R2	0.7535	R2-adj	0.7384		
LIK	-611.017	AIC	1230.03	SC	1237.91
RSS	3.20210e+10	F-test	49.9208	Prob	6.24470e-15
SIG-SQ	6.53489e+08	(25563.4)	SIG-SQ(ML)	6.04169e+08	(24579.9)

VARIABLE	COEFF	S.D.	t-value	Prob
CONSTANT	-13425.6	16885.7	-0.795085	0.430399
A	635.232	62.6834	10.133975	0.000000
I	-239.985	464.667	-0.516466	0.607852
G	8272.33	8990.16	0.920153	0.362001

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER		10.596769	
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	PROB
Kiefer-Salmon	2	27.599150	0.000001
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Koenker-Bassett test	3	15.750672	0.001276
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	PROB
White	8	30.949319	0.000143

Spatial Lag Model
Maximum Likelihood Estimation

DATA SET	IPTU	SPATIAL WEIGHTS MATRIX		WGH15	
DEPENDENT VARIABLE	P	OBS 53	VARS 5	DF 48	
R2	0.7619	Sq. Corr.	0.7643		
LIK	-609.955	AIC	1229.91	SC	1239.76
SIG-SQ	5.77699e+08	(24035.4)			

VARIABLE	COEFF	S.D.	z-value	Prob
W_P	0.222975	0.142779	1.561680	0.118363
CONSTANT	-28863.3	18574.7	-1.553906	0.120207
A	617.948	60.3423	10.240718	0.000000
I	-185.45	439.945	-0.421530	0.673368
G	8086.41	8452.78	0.956657	0.338741

REGRESSION DIAGNOSTICS				
DIAGNOSTICS FOR HETEROSKEDASTICITY				
RANDOM COEFFICIENTS				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	3	43.102225	0.000000	
Spatial B-P test	3	43.102232	0.000000	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
SPATIAL LAG DEPENDENCE FOR WEIGHTS MATRIX		WGH15 (row-standardized weights)		
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	2.123592	0.145046	
LAGRANGE MULTIPLIER TEST ON SPATIAL ERROR DEPENDENCE				
WEIGHT	STAND ZERO	DF	VALUE	PROB
WGH15	yes no	1	5.342924	0.020807

TABLE V

Box-Cox Best Fit
Ordinary Least Squares Estimation

DEPENDENT VARIABLE		P-025	OBS 53	VARS 4	DF 49
R2	0.8232	R2-adj	0.8124		
LIK	147.179	AIC	-286.358	SC	-278.476
RSS	0.0120162	F-test	76.0516	Prob	1.88918e-18
SIG-SQ	0.000245228	(0.0156597)	SIG-SQ(ML)	0.000226720	(0.0150572)

VARIABLE	COEFF	S.D.	t-value	Prob
Constant	3.42792	0.0298563	114.813874	0.000000
A0	0.0654656	0.00546788	11.972758	0.000000
I0	-0.00290209	0.0029531	-0.982728	0.330569
G	0.0123463	0.00530889	2.325587	0.024224

REGRESSION DIAGNOSTICS			
MULTICOLLINEARITY CONDITION NUMBER		35.143426	
TEST ON NORMALITY OF ERRORS			
TEST	DF	VALUE	PROB
Kiefer-Salmon	2	0.342127	0.842768
DIAGNOSTICS FOR HETEROSKEDASTICITY			
RANDOM COEFFICIENTS			
TEST	DF	VALUE	PROB
Breusch-Pagan Test	3	1.632683	0.652002
SPECIFICATION ROBUST TEST			
TEST	DF	VALUE	PROB
White	8	4.765208	0.782351

Spatial Lag Model
Maximum Likelihood Estimation

DATA SET	IPTU	SPATIAL WEIGHTS MATRIX		WGH15			
DEPENDENT VARIABLE		P-025	OBS 53	VARS 5		DF 48	
R2	0.8415	Sq. Corr.	0.8438				
LIK	150.281	AIC	-290.562	SC	-280.711		
SIG-SQ	0.000200313	(0.0141532)					

VARIABLE	COEFF	S.D.	z-value	Prob
W_P-025	0.26404	0.102247	2.582389	0.009812
CONSTANT	2.4557	0.376838	6.516589	0.000000
A0	0.0614051	0.00521678	11.770680	0.000000
I0	-0.00191039	0.00269862	-0.707914	0.478998
G	0.0127601	0.00480118	2.657710	0.007867

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	3	1.905855	0.592175
Spatial B-P test	3	1.905855	0.592175

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHTS MATRIX WGH15 (row-standardized weights)

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	6.204556	0.012742

LAGRANGE MULTIPLIER TEST ON SPATIAL ERROR DEPENDENCE

WEIGHT	STAND	ZERO	DF	VALUE	PROB
WGH15	yes	no	1	3.459878	0.062875

Table VI presents a comparison of estimates of the mean partial derivatives (implicit marginal characteristic prices) derived from equation (2) with the corresponding values of the parameters λ and θ. The indicator variable G remains a binary 0-1 variable in all specifications, but the pair of parameters (λ,θ) are assigned values of (0;1), (0;0) and (-0.25;0) for the semilog, doublelog and Box-Cox best fit forms, respectively. In all cases, the partial derivatives with respect to the i-th characteristic vary by observation: sample averages of these derivatives are computed. Figures reported include the characteristics Area and Garage which have statistically significant estimates in the specifications analyzed.

The Box-Cox best fit specification yields an intermediary value for the mean partial derivative with respect to both characteristics Area and Garage, as compared to the doublelog and semilog forms, although it is closer to the latter for the variable Area. Among all specifications, the doublelog has by far the lowest coefficient of variation (ratio standard deviation/mean).

With respect to the availability of Garage (binary variable G), the figures computed show that the average sample values in the doublelog and Box-Cox best fit specifications with spatial effects are R$12.242 and R$15.031, respectively.

TABLE VI

Mean Partial Derivatives of the i-th Characteristic

VARIABLE	MEAN	VARIANCE	ST DEV	SKEWNESS	KURTOSIS
ADPISL	507.4805	143178.8	378.3898	3.346942	16.66854
ADPISLW	475.3914	110541.3	332.4775	2.870906	13.09932
ADPIDL	534.6828	3692.138	60.76296	-0.1555044	2.140095
ADPIDLW	507.7826	4371.829	66.11981	-0.2354811	2.035638
ADPIBF	529.0229	17932.61	133.9127	1.21987	4.53277
ADPIBFW	496.9555	15274.51	123.5901	0.6806048	3.073966
GDPISL	15608.27	1.354408E+08	11637.9	3.346942	16.66854
GDPISLW	15745.97	1.378411e+08	11740.57	3.346942	16.66854
GDPIDL	11884.18	4.519976E+07	6723.077	1.455062	5.177226
GDPIDLW	12241.53	4.795888E+07	6925.235	1.455062	5.177226
GDPIBF	14649.45	1.406193E+08	11858.3	2.33694	9.317868
GDPIBFW	15030.88	1.32674E+08	11518.42	2.034834	7.470509

ADPISL = Mean Partial Derivative w.r.t Area, Semilog
ADPISLW = Mean Partial Derivative w.r.t Area, Semilog Spatial Effects
ADPIDL = Mean Partial Derivative w.r.t Area, Double Log
ADPIDLW = Mean Partial Derivative w.r.t Area, Double Log Spatial Effects
ADPIBF = Mean Partial Derivative w.r.t Area, Box Cox Best Fit
ADPIBFW = Mean Partial Derivative w.r.t Area, Box Cox Best Fit, Spatial Effects

GDPISL = Mean Partial Derivative w.r.t Garage, Semilog
GDPISLW = Mean Partial Derivative w.r.t Garage, Semilog Spatial Effects
GDPIDL = Mean Partial Derivative w.r.t Garage, Double Log
GDPIDLW = Mean Partial Derivative w.r.t Garage, Double Log Spatial Effects
GDPIBF = Mean Partial Derivative w.r.t Garage, Box Cox Best Fit
GDPIBFW = Mean Partial Derivative w.r.t Garage, Box Cox Best Fit, Spatial Effects

VII CONCLUSION

This study examines sources of price variation in a sample of residential apartments of Belo Horizonte, Brazil, using both standard and spatial autoregressive hedonic price models (HPM). The Box-Cox transformation procedure is used to choose the best non-linear functional form. For the sake of comparison, other specifications are also analyzed, including the linear, semilog and doublelog forms.

The results show that spillover (adjacency) effects are an important source of price variation in this housing market. For two of the specifications analyzed (including the best fit non-linear functional form), the hypothesis of no spatial dependence (spillover) is strongly rejected and, at the same time, the hypothesis of no spatial heterogeneity (spatial error dependence) is not rejected. A third specification has both hypotheses strongly rejected, meaning that both spatial effects are present and further analysis is required to determine which one is relevant.

These empirical findings support the need to incorporate spatial effects in studies of housing price determination. The importance of incorporating spatial factors makes the use of special spatial econometrics techniques essential to the analysis of economic performance in housing markets.

BIBLIOGRAPHY

Anselin, Luc (1988). *Spatial Econometrics: Methods and Models* Dordrecht: Kluwer Academic.

Anselin, Luc (1989). "What is Special About Spatial Data? Alternative Perspectives on Spatial Data Analysis," UCLA-Santa Barbara Geography Department, National Center for Geographic Information & Analysis (NCGIA), Technical Paper 89-4 (June).

Anselin, Luc (1990). *SpaceStat, A Program for the Statistical Analysis of Spatial Data*. Manual of the Preliminary Beta Release.

Anselin, Luc (1992). "Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences," UCLA-Santa Barbara Geography Department, National Center for Geographic Information & Analysis (NCGIA), Technical Report 92-10 (Dec.).

Bailey, M.J., R.F. Muth and H.O. Nourse (1963). "A Regression Method for Real Estate Price Index Construction," *Journal of the American Statistical Association*, 58:933-942.

Butler, Richard V. (1982). "The Specification of Hedonic Indexes for Urban Housing," *Land Economics*, Vol.58, No.1(Feb.):96-108.

Can, Ayse (1992). "Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban Economics*, 22:453-474.

Chow, Gregory C. (1967). "Technological Change and the Demand for Computers," *American Economic Review*, 57: 1117-1130.

Cliff, A.D. and J.K. Ord (1973). *Spatial Autocorrelation* (London: Pion).

Follain, James R. and Emmanuel Jimenez (1984). "Estimating the Demand for Housing Characteristics," *Regional Science and Urban Economics*, 15:77-107.

Follain, James R. and Emmanuel Jimenez (1985). "The Demand for Housing Characteristics in Developing Countries," *Urban Studies*, 22:421-432.

Greene, William H. (1992). LIMDEP USER'S MANUAL REFERENCE GUIDE.

Griliches, Zvi, ed. (1971), *Price Indexes and Quality Change*, Cambridge: Harvard University Press.

Megbolugbe, Isaac F. (1989), "A Hedonic Index Model: The Housing Market of Jos, Nigeria," *Urban Studies*, 26(5), Oct.:486-494.

Rosen, Sherwin (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82(Jan./Feb.):34-55.

Spitzer, John J. (1981). "A Primer on Box-Cox Estimation," *The Review of Economics and Statistics*, 64:307-313.