ISSN 0102-9924

FACULDADE DE CIÊNCIAS ECONÔMICAS DA UFRGS

Análise Econômica

CAPITAL HUMANO NOS MUNICÍPIOS PARANAENSES
LUCIANO NAKABASHI E EVÂNIO FELIPE

CÂMBIO, INFLAÇÃO, JUROS E RESERVAS NA TRANSIÇÃO DE REGIMES CAMBIAIS: UMA INVESTIGAÇÃO ECONOMÉTRICA PARA O BRASIL
FLÁVIO VILELA VIEIRA E CARLOS DE ALMEIDA CARDOSO

INTEGRAÇÃO ECONÔMICA E POLÍTICAS DE DESENVOLVIMENTO: EXPERIÊNCIAS E PERSPECTIVAS PARA A AMÉRICA LATINA
RICARDO DATHEIN

MERCADO IMOBILIÁRIO E A IMPORTÂNCIA DAS CARACTERÍSTICAS LOCAIS: UMA ANÁLISE QUANTÍLICO-ESPACIAL DE PREÇOS HEDÔNICOS EM BELO HORIZONTE
BERNARDO FURTADO

ANÁLISE DA DINÂMICA DA PRODUTIVIDADE DO TRABALHO ENTRE SETORES E ESTADOS BRASILEIROS NA DÉCADA DE 90
ADELAR FOCHEZATTO E VALTER JOSÉ STULP

TEORIA ECONÔMICA DO SUICÍDIO: ESTUDO EMPÍRICO PARA O BRASIL
CLAUDIO DJISSEY SHIKIDA, ARI FRANCISCO ARAUJO JR E RAFAEL ALMEIDA VILHENA GAZZI

ÍNDICE RELATIVO DE DESENVOLVIMENTO ECONÔMICO E SOCIAL DOS MUNICÍPIOS DA REGIÃO SUDOESTE PARANAENSE
CARMEM OZANA MELO

DESEMPENHO COMERCIAL DA INDÚSTRIA DE MÁQUINAS E IMPLEMENTOS AGRÍCOLAS NO BRASIL
CARLOS ALBERTO CINQUETTI

COMPETITIVIDADE DAS EXPORTAÇÕES MUNDIAIS DE PLANTAS VIVAS E PRODUTOS DE FLORICULTURA
ETEVALDO ALMEIDA, PATRICIA SALES LIMA, LUCIA MARIA SILVA, RUBEN DARIO MAYORGA E FRANCISCO DE LIMA

UMA AVALIAÇÃO DA EFICÁCIA DO FNE, NO PERÍODO 1995-2000
ALEXANDRE MANOEL ANGELO DA SILVA E GUILHERME RESENDE

ANO 25
Nº 47
Setembro, 2007

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. José Carlos Ferraz Hennemann
FACULDADE DE CIÊNCIAS ECONÔMICAS
Diretor: Prof. Gentil Corazza

CENTRO DE ESTUDOS E PESQUISAS ECONÔMICAS
Diretor: Prof. Lovois de Andrade Miguel
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS
Chefe: Prof. Eduardo Ernesto Filippi
DEPARTAMENTO DE CIÊNCIAS CONTÁBEIS E ATUARIAIS
Chefe: Prof. Ceno Odilo Kops
CURSO DE PÓS-GRADUAÇÃO EM ECONOMIA
Coordenador: Prof. Marcelo S. Portugal
CURSO DE PÓS-GRADUAÇÃO EM DESENVOLVIMENTO RURAL
Coordenador: Prof. Paulo Dabdab Waquil
CONSELHO EDITORIAL: André Moreira Cunha (UFRGS), Carlos G. A. Mielitz Netto (UFRGS), Carlos Henrique Horn (UFRGS), Eduardo A. Maldonado Filho (UFRGS), Eleutério F. S. Prado (USP), Eugênio Lagemann (UFRGS), Fernando Cardim de Carvalho (UFRJ), Fernando Ferrari Filho (UFRGS), Fernando de Holanda Barbosa (FGV/RJ), Flávio Augusto Ziegelman (UFRGS), Flávio Vasconcellos Comim (UFRGS), Gentil Corazza (UFRGS), Giácomo Balbinotto Netto (UFRGS), Gustavo Franco (PUC/RJ), Hélio Henkin (UFRGS), Jan A. Kregel (University of Missouri at Kansas City), João Rogério Sanson (UFSC), Joaquim Pinto de Andrade (UnB), Júlio César Oliveira (UFRGS), Luiz Estrella Faria (UFRGS), Luis Paulo Ferreira Nogueról (UFRGS), Marcelo S. Portugal (UFRGS), Maria Alice Lahorgue (UFRGS), Octávio Augusto Camargo Conceição (UFRGS), Paul Davidson (Journal of Port Keynesian Economics), Paulo D. Waquil (UFRGS), Pedro C. D. Fonseca (UFRGS), Philip Arestis (University of Cambridge), Ricardo Dathein (UFRGS), Ronald Otto Hillbrecht (UFRGS). Sabino da Silva Porto Jr. (UFRGS), Sérgio M. M. Monteiro (UFRGS), Stefano Florissi (UFRGS) e Werner Baer (University of Illinois at Urbana - Champaign).

COMISSÃO EDITORIAL: Eduardo Augusto Maldonado Filho, Fernando Ferrari Filho, Hélio Henkin, Marcelo Savino Portugal, Paulo Dabdab Waquil. e Sérgio Marley Modesto Monteiro.
EDITOR: Sérgio Marley Modesto Monteiro
EDITOR ADJUNTO: Hélio Henkin
SECRETÁRIO: Emerson Douglas Neves
REVISÃO DE TEXTOS: Vanete Ricacheski
EDITORAÇÃO: Núcleo de Editoração e Criação da Gráfica da UFRGS – Janaína Horn
FUNDADOR: Prof. Antônio Carlos Santos Rosa
Os materiais publicados na revista Análise Econômica são da exclusiva responsabilidade dos autores. É permitida a reprodução total ou parcial dos trabalhos, desde que seja citada a fonte. Aceita-se permuta com revistas congêneres. Aceitam-se, também, livros para divulgação, elaboração de resenhas e recensões. Toda correspondência, material para publicação (vide normas na terceira capa), assinaturas e permutas devem ser dirigidos ao seguinte destinatário:

Prof. Sérgio Marley Modesto Monteiro
REVISTA ANÁLISE ECONÔMICA - Av. João Pessoa, 52
CEP 90040-000 PORTO ALEGRE - RS, BRASIL
Telefones: (051) 3308 3513 /3308 4164
Fax: (051) 3308 3990
Email: rae@vortex.ufrgs.br

Assinatura revista Análise Econômica: R$50,00
A assinatura anual dá direito a 2 números da revista.

Análise Econômica
Ano 25, n° 47, setembro, 2007 - Porto Alegre
Faculdade de Ciências Econômicas, UFRGS, 2000

Periodicidade semestral, março e setembro.
ISSN 0102-9924

1. Teoria Econômica - Desenvolvimento Regional - Economia Agrícola - Pesquisa Teórica e Aplicada - Periódicos. I. Brasil.
Faculdade de Ciências Econômicas,
Universidade Federal do Rio Grande do Sul.

CDD 330.05

CDU 33 (81) (05)

Mercado imobiliário e a importância das características locais: uma análise quantílico-espacial de preços hedônicos em Belo Horizonte

Bernardo Furtado*

Resumo: A motivação principal desse artigo é identificar quão relevante é a localização de determinado imóvel na composição do seu valor de mercado. Isso, no intuito de compreender melhor como as relações econômicas se dão espacialmente no tecido urbano. Para tal utiliza-se de uma miríade de conceitos – oriundos de várias ciências, da geografia, à arquitetura, ao urbanismo, à economia –, e metodologias, da estatística, à econometria e ao geoprocessamento. A proposta segue a literatura de preços hedônicos, mas propõe que a síntese da percepção das amenidades urbanas seja especificada pelo elemento do bairro (Lynch, 1997). Dois modelos são apresentados, os testes e resultados comentados, sendo que o de melhor ajuste é o modelo de erro auto-regressivo espacial com especificação de bairros classificados por um índice de renda. A análise quantílica acrescenta informações relevantes no entendimento do modelo.

Palavras-Chave: mercado imobiliário, econometria espacial, análise quantílica

Jel Classification: C21, R2, L85

1 Introdução

A motivação principal desse artigo é identificar quão relevante é a localização de determinado imóvel na composição do seu valor de mercado, no intuito de compreender melhor como as relações econômicas ocorrem espacialmente no tecido urbano. Para tal utiliza-se de todo um conjunto de conceitos oriundos de várias ciências, da geografia, à arquitetura, ao urbanismo, à economia, e metodologias, da estatística, à econometria e ao geoprocessamento.

Hipoteticamente, um imóvel com as mesmas características específicas, localizado porém em bairro diferente, teria qual alteração no seu preço final? Nesse sentido, o objetivo é identificar padrões e caracte-

* Doutorando em Economia (CEDEPLAR/UFMG), professor do Centro Universitário UNA. E-mail: furtado@cedeplar.ufmg.br

Artigo recebido em maio de 2007. Aceito em agosto de 2007

rísticas dos preços dos imóveis no município de Belo Horizonte. Dessa forma, segue a abordagem de Neto (2002) e Hermann e Haddad (2005), cuja pergunta fundamental é se a localização do imóvel é característica fundamental no estabelecimento do patamar do preço, ou se o são as características intrínsecas à unidade habitacional em si. Tradicionalmente na literatura "...é dada maior atenção econométrica às características habitacionais do que às características localizacionais..." (HERMANN e HADDAD, 2005, p. 13). Busca-se complementar esse quadro e detalhar a participação de ambos.

A análise hedônica de preços de imóveis na economia busca valorar concomitantemente especificidades do imóvel e amenidades urbanas de sua localização.[1] A análise desses fatores é geralmente feita em cidades consideradas como monocêntricas (ou duocêntricas) na qual é a distância em relação ao *business district center* o fator relevante. Ademais, medidas como distância até a rodovia, disponibilidade de serviços, presença de áreas verdes próximas ou índices de criminalidade (RONDON e ANDRADE, 2003) são sempre consideradas no desenho do modelo. Neste artigo segue-se abordagem de certo modo distinta. Norteia-se pela proposta do arquiteto Kevin Lynch (1997) e da discussão de Paisagem Urbana, segundo a qual a percepção do cidadão se circunscreve espacialmente em uma entidade reconhecida. Ou seja, o agente relaciona qualidades positivas e negativas (atributos) a determinadas porções espaciais que ele e seus interlocutores conseguem determinar com alguma precisão no espaço urbano.

Com efeito, a comercialização de imóveis (e a atribuição de seus preços[2]) sempre indica a unidade espacial do bairro como referência de localização. Nesse sentido, a medida de "proximidade ao centro", por exemplo, está embutida na sua localização e não é medida explicitamente.[3] Se o atributo for relevante, o bairro, como um todo, é valorizado (ou não). De outro modo, o argumento que se defende é que o bairro é visto como referência para uma série de variáveis cujos pesos são tratados em conjunto e não separadamente.

A própria natureza complexa do urbano e da cidade (MONTEMÓR, 2006) leva, na sua análise, a problemas de multicolinearidade no processo de estimação que são devidos em especial à não-neutralida-

[1] Veja, para o caso brasileiro, Hermann e Haddad (2005) e Macedo (1996) entre outros.

[2] Veja, por exemplo, a pesquisa de preços da Fundação IPEAD da UFMG, www.ipead.face.ufmg.br.

[3] Do ponto de vista urbanístico, até a determinação de um "centro" em uma cidade como Belo Horizonte se torna complicada dada sua dispersão espacial. Toda a região centro-sul da cidade poderia ser considerada como "centro". Medir a distância até esse centro expandido é inviável, assim como a distância interna à região. Ademais, utilizar o centro tradicional (a Praça 7 de setembro) como referência é incorrer em erro gratuito já que não é relevante para o mercado de imóveis residenciais.

de do espaço. Nesse sentido pressupõe-se que o espaço físico, rígido, condiciona as questões socioeconômicas, mas também é condicionado por elas (SOJA, 1993; BRUECKNER, THISSE *et al.*, 1999).[4] A análise multivariada de componentes principais contribui para essa questão, uma vez que, como se verá, possibilitou a construção de um índice de valores de renda domiciliar que resume satisfatoriamente a multiplicidade de fatores de cada vizinhança.

A natureza do urbano impôs ainda que a análise fosse quantílico-espacial. Ou seja, houve ganhos de especificação na comparação do modelo simples com o modelo espacial. O quantílico, por sua vez, ressaltou a valoração diferente da importância da localização para imóveis com níveis de preços diferentes.

O artigo conta ainda com a discussão de preços hedônicos e a percepção da paisagem urbana, como justificativa para a escolha do bairro como unidade de análise; a descrição do modelo simples, espacial e sua matriz de pesos e o modelo quantílico. Algumas limitações do modelo proposto são discutidas e, a seguir, o estudo de caso de Belo Horizonte é apresentado com a descrição da base de dados, bem como a análise dos resultados e as considerações finais.

2 Referencial Teórico

2.1 Preços hedônicos e percepção da paisagem urbana: bairros como unidade de análise

O mercado de terras urbanas tem a peculiaridade de cada localização ser sempre diferenciada em relação às demais o que leva à monopolização do espaço (SINGER, 1982). Como bem heterogêneo, o modelo mais adequado para tratá-lo é o de preços hedônicos. De acordo com Hermann e Haddad:

> [O modelo de preços hedônicos] considera um bem heterogêneo como um pacote fechado de atributos, e estima o preço marginal de cada atributo a partir da análise do valor observado do bem heterogêneo e de suas respectivas quantidades de atributos (2005, p. 239).

Sheppard, por sua vez, (1999, p. 1599) apresenta duas justificativas para utilizar as funções de preços hedônicos no mercado imobiliário: a) na construção de índices de preços que considerem mudanças na qualidade das mercadorias consumidas;[5] e b) como atributo informacional na construção de preços para mercadorias heterogêneas.

[4] Sobre o assunto, veja também Harvey (1973); Santos (1992) e Soja (1996).

[5] Nesse caso, imóvel construído em localidade diferente assume qualidade diversa.

No mercado de imóveis, embora os conceitos não sejam formalizados, estão presentes na medida em que, empiricamente, a diferença de preços em imóveis similares, mas com localizações diferentes, é de grande magnitude.

A disciplina que estuda a Paisagem Urbana entre os arquitetos e urbanistas foi em grande parte formulada por Kevin Lynch e Aldo Rossi (2001)[6] e especifica que a percepção da paisagem urbana, feita individual e coletivamente, contribui com a construção da imagem do lugar. Nesse sentido, elementos[7] físicos e imaginários ou simbólicos (tais como rodovias, linhas de trens ou rios e praças, esquinas ou monumentos) segmentam e nomeiam o espaço em unidades identificáveis pela população.

O bairro assume localmente a função dessa unidade espacial homogênea sobre o qual atributos são associados. De acordo com essa lógica de construção da imagem percebida, uma amenidade urbana,[8] tal como presença de áreas verdes, associa-se a certa unidade espacial e não é linear no espaço. Dito de outra forma, um imóvel localizado mais próximo dessa amenidade pode ter sua imagem mais ou menos modificada, já que se localiza em outra unidade (bairro) na construção perceptiva da população.

Note que, dada a heterogeneidade do urbano,[9] as unidades espaciais variam de tamanho conforme seu grau interno de homogeneidade varia. Em termos práticos, uma ocupação irregular, do tipo favela, nunca se classifica na mesma unidade que o bairro adjacente, embora o desvalorize.

2.2 O modelo proposto e as conseqüentes necessidades da modelagem

A escolha da forma funcional seguida será a abordagem conhecida como "geral para específico", ou seja, parte-se de um modelo com todas as variáveis contidas na amostra e identificam-se aquelas relevantes, mais significativas, para o modelo em questão. Ademais, segundo orientação de especificação de modelos espaciais de Anselin (2005, p. 199), parte-se da análise tradicional, utiliza-se a estatística I de Moran e o teste do Multiplicador de Lagrange (ML) para identificar se há necessidade de modelar espacialmente.

[6] Para outras abordagens, veja também Cullen (1961) e o as obras do geógrafo, Yi-Fu. Tuan (1980; 1983).

[7] Os cinco elementos analíticos propostos por Lynch (1997) são: vias, limites, marcos, bairros e nós.

[8] Para derivação de modelo formal de amenidades urbanas na Nova Economia Urbana, veja Brueckner, Thisse e Zenou (1999).

[9] Essa heterogeneidade é facilmente comprovada pela análise da diversidade de renda, explicita nos resultados do censo demográfico realizado pelo IBGE (2000). Para análise visual do caso de Belo Horizonte veja Furtado (2006).

A busca do modelo teve como premissas a escolha de atributos que sintetizassem características várias. Ou seja, ao invés de listar um número expressivo de variáveis, optou-se pela escolha de indicativos-síntese que fossem representativos. Buscou-se, com isso, minimizar problemas de multicolinearidade, ao mesmo tempo em que possibilitou acesso a banco de dados mais amplos. Os resultados indicam que essa escolha foi razoável.

Com base no descrito acima, o modelo pressupôs que o preço dos imóveis, a variável dependente, deveria ser explicado a) pela sua área, já que é natural que quanto maior o imóvel, mais caro seja; b) pela sua idade, sendo, nesse caso, negativo o sinal esperado do coeficiente; c) pelo imposto predial e territorial urbano – IPTU – que serve de *proxy* para as características próprias do imóvel, tais como qualidade do acabamento; d) pelo valor do condomínio vigente, que busca refletir as qualidades e serviços do prédio como um todo. Ou seja, taxas de condomínio mais elevadas indicam a presença de mais funcionários permanentes, incluindo porteiros de tempo integral, elevadores, manutenção de piscinas, cuidados com jardins, por exemplo;[10] e) o número de vagas, já que apartamentos idênticos, no mesmo prédio, podem ter número de vagas alocadas diferentemente.

O controle das características locais foi especificado de duas formas. O primeiro por um índice calculado por uma instituição de pesquisa que tem como base da sua constituição a informação de rendas domiciliares fornecidas pelo IBGE (censo 2000). Esse índice é subdivido em quatro categorias: popular, médio, luxo e alto luxo e é utilizado rotineiramente como referência para o "preço do metro quadrado" que é o principal balizamento das imobiliárias da cidade.

Os resultados iniciais indicaram que essa especificação não foi suficiente para a explicação completa dos imóveis, o que levou à construção de outro índice mais específico e que será discutido na próxima seção.

2.3 Análise Multivariada – componentes principais

A análise de componentes principais – ACP – é uma técnica de estatística que tem sido usada largamente em análise e compressão de dados. Segundo Mingoti:[11]

> ...seu objetivo principal é o de explicar a estrutura de variância e covariância de um vetor aleatório (...) através da construção de combinações lineares

[10] De fato, uma regressão com o valor da taxa de condomínio como variável dependente e piscina, salão, quadra e classificação IPEAD como indpendentes,, mostrou-se com bom ajuste e R2 ajustado de 0.6175.

[11] Veja ainda Johnson e Wichern (1998).

das variáveis originais. Estas combinações lineares são chamadas de componentes principais e são não correlacionadas entre si (2005, p. 59).

Seja $X = (X_1, X_2, ..., X_p)$ um vetor aleatório; $\sum_{pxp}$ sua matriz de covariâncias com $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$ os respectivos auto-valores da matriz e e_1, e_2, ... e_p os auto-vetores normalizados. A j-ésima componente principal da matriz $\sum_{pxp}$ pode ser definida como:

Equação 1 – Definição de componente principal

$$Y_j = e'_j X = e'_{j1} X_1 + e'_{j2} X_2 + ... + e'_{jp} X_p \qquad \text{(MINGOTI, 2005, p. 60-61)}$$

A forma de construção das componentes implica que a primeira é sempre mais representativa em termos de variância total e, assim, sucessivamente. Os valores numéricos das componentes principais, denominadas de *escores*, podem ser utilizados em análises de regressão (MINGOTI, 2005, p. 60).

Na efetivação do procedimento, realizado automaticamente pelo *software* ArcGIS e inúmeros outros pacotes estatísticos, há a padronização dos dados, o cálculo da matriz de correlação para, em seguida, realizarem-se os cálculos dos auto-vetores e auto-valores que permitirão a construção dos valores individuais para cada observação, de acordo com a fórmula da equação 1.

Uma das vantagens da utilização de ACP é a não-necessidade de assumir a normalidade da distribuição. Por outro lado, alguns estudos indicam que mudanças na escala das variáveis podem alterar o resultado (NAIK e KHATTREE, 1996). A padronização dos dados contribui para minimizar esse problema.

A análise de componentes principais foi utilizada no contexto deste artigo para se construir um índice de relevância da vizinhança baseada na renda dos domicílios fornecida pelos resultados para setores do censo (IBGE) 2000.

Especialmente, na análise do mercado imobiliário a ACP foi aplicada por Hermann e Haddad (2005) no Brasil e por Can e Megbolugebe (1997) nos Estados Unidos.

2.4 Análise espacial e modelos de econometria espacial

O conceito essencial na compreensão da análise espacial é o de dependência espacial. Esse conceito explica o que é comumente chamado de primeira Lei da Geografia de Tobler, segundo a qual: "todas as coisas são parecidas, mas coisas mais próximas se parecem mais que coisas mais distantes..." (CÂMARA, 2004, p. 11). A autocorrelação espacial, por sua vez, é a expressão que mede essa dependência espacial. De

fato, Anselin (2005) sugere primeiramente a implementação do modelo tradicional para, em seguida, avaliar a presença de autocorrelação espacial e, conseqüentemente, a necessidade de um modelo espacial. O primeiro indicativo dessa presença é o índice I de Moran:

Equação 2 – Índice I de Moran

$$I = \frac{\sum_{j=1}^{n} \sum_{i=1}^{n} w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}, \text{ na qual} \qquad (1)$$

w_{ij} é a matriz de pesos que traz a consideração espacial para a fórmula, n é o número de variáveis, x_i é a variável de interesse. São várias as possíveis matrizes de pesos a serem construídas e serão discutidos no item seguinte. Embora a estatística I de Moran seja adequada para indicar a presença de autocorrelação espacial, ela não explicita de que forma a presença ocorre (ANSELIN, 2005, p. 197).

O modelo espacial deve ser utilizado quando há sugestões de ordem teórica que indiquem haver influência espacial entre as observações. No caso aqui analisado, por exemplo, deve-se incluir a possibilidade do caso espacial se há indicação que o preço, qualidade, tamanho ou oferta de serviços dos imóveis próximos influenciem os valores do imóvel analisado. De acordo com o modelo simples de identificação do processo espacial proposto por Anselin (2005, p.199), caso não haja relação espacial, pode-se retornar ao modelo tradicional com maior segurança do seu ajuste.

Dentre os modelos econométricos espaciais, os mais citados na literatura são o de defasagem espacial (*spatial lag*), o qual prevê a inclusão de um termo de influência dos vizinhos no processo regressivo:

Equação 3 – Modelo de defasagem espacial

$$Y = \rho Wy + X\beta + \varepsilon, \text{ onde} \qquad (2)$$

Y é a variável dependente; Wy é o vetor de defasagens espaciais, derivado da escolha da matriz de pesos espaciais;[12] ρ é o coeficiente autoregressivo espacial; X é a matriz de variáveis explicativas independentes com seus coeficientes β e ε é o vetor de erros. Outro modelo é o de erro auto-regressivo espacial (*error lag*)[13] que é capaz de captar um "...efeito

[12] Não é objetivo deste artigo discutir a teoria subjacente. Para maiores detalhes consulte, dentre outros, Anselin (1988).

[13] O modelo de erro espacial auto-regressivo foi proposto inicialmente por Whittle (1954).

espacial não-modelado..." (ALMEIDA, 2004, p. 61). Nesse caso, há algum efeito espacial implícito no problema, mas que não está captado pelo coeficiente. Seu efeito é então expresso pelo erro:

Equação 4 – Modelo auto-regressivo espacial

$$Y = X\beta + u, \text{ e} \tag{3}$$

$$u = \lambda Wu + \varepsilon, \text{ onde} \tag{4}$$

λ é o parâmetro do erro auto-regressivo espacial.[14]

Os testes recomendados para identificar o modelo de autocorrelação espacial são: a) o *Lagrange Multiplier Lag*, para identificar se é o caso de modelo de defasagem espacial, b) o *Lagrange Multiplier Error* para identificar se é o caso de utilização do modelo de erro espacial. Há ainda o LM-SARMA que indicaria a hipótese de modelo mais complexo no qual há autocorrelação espacial tanto nos coeficientes quanto nos erros.

Vale ressaltar ainda que modelos de Mínimos Quadrados Ordinários – MQO – não são adequados para análises espaciais já que a literatura (ALMEIDA, 2004) indica que quando há autocorrelação espacial os MQO apresentam resultados inconsistentes. A alternativa sugerida é a de se utilizar o estimador de Máxima Verossimilhança – MV – como sendo o parâmetro estimado que gerou, com a maior probabilidade, a amostra observada. A sugestão de utilização do estimador de MV encontra-se em Anselin (2005) e segue desenvolvimento metodológico de Anselin (1988), Anselin e Bera (1998) e Smirnov e Anselin (2001).[15]

2.5 Matrizes de peso

A matriz de pesos é parte essencial da análise espacial e sua função é representar como se dá a influência dos vizinhos em cada observação e como essa influência decai no espaço. Vários autores insistem que matrizes diferentes podem alterar significativamente os resultados de modo que é prudente fazer a escolha da matriz cuidadosamente e testar as várias possibilidades.

[14] Outros modelos espaciais incluem: modelo com erro de média móvel espacial, modelo regressivo cruzado espacial, modelo de Durbin espacial, modelo misto com defasagem e erro de média móvel de primeira ordem (ALMEIDA, 2004, cap. 4).

[15] De acordo com Anselin, "*However, due to the simultaneity implied by the spatial nature of the dependence, these procedures [estimated general least square] are not applicable in the spatial case and a full maximum likelihood estimation must be carried out*" (1992, p. 214). De todo modo, o resultado estimado pelo método de momentos generalizados em dois passos (GM *two-step*) apresentou resultados bem próximos (da ordem de 0,01), com exceção do índice ACP, cuja diferença foi maior (0,1).

As matrizes mais simples são as de contigüidade espacial que indicam, por meio de uma relação binária, o fato de ser ou não vizinho. Dentre essas há dois tipos: matriz *queen* que considera como vizinhos aqueles que têm fronteiras ou apenas vértices em comum com a observação analisada e a matriz do tipo rook que considera somente aqueles com fronteiras comuns. Em ambos os casos é possível considerar as matrizes de maior ordem de contigüidade na qual seriam incluídos os vizinhos dos vizinhos e assim por diante. A análise teórica subjacente e o fenômeno analisado contribuem para indicar a escolha mais adequada.

Outras possibilidades para a construção da matriz de pesos são as baseadas em distância espacial, nas quais pode se explicitar um limite *ad hoc* para a influência. Há ainda as chamadas *k-nearest neighbours*, na qual também se determina *ad hoc* o número de vizinhos que influenciam cada observação e garante-se que, independente da distância ou da contigüidade, todas as observações terão o mesmo número de vizinhos.

2.6 Análise quantílica

A análise quantílica[16] é baseada em pontos derivados a intervalos regulares da função distribuição acumulada[17] de uma variável estocástica. Os intervalos que denotam os subconjuntos de análise da amostra podem ser de qualquer ordem (por exemplo: 100 para os centis, cinco para quintis). Nesse estudo, a análise quantílica segue a proposta teórica de Buchinsky (1997) segundo a qual ele reforça a necessidade de verificação se os padrões da regressão (feitos, por exemplo, por MQO) se repetem nos quantis. Caso contrário, a análise quantílica adiciona informações essenciais à interpretação do fenômeno. Ademais, a regressão quantílica não é tão sensível à presença de *outliers* (como em MQO) e permite examinar as mudanças "*at different points of the distribution*". O autor resume a relevância da análise quantílica assim: "clearly, it is not enough to investigate changes in the mean when the entire shape of the distribution changes dramatically" (*op.cit.*, p. 90).[18]

Wooldridge recomenda a regressão quantílica quando há heterogeneidade na população:

> *Median regression is a special case of quantile regression, where we model quantiles in the distribution of y given x. For example, in addition to the median, we can estimate how the first and third quartiles in the distribution of y given x change with x. Except for the median (which leads to Least Absolute*

[16] Para maiores detalhes veja também Buchinsky (1997) e os artigos originais de Koenker & Bassett (1978).

[17] $F(x) = P(X \leq x)$, no qual F(x) é a probabilidade de que a variável X apresente valor inferior ou igual a x. Desse modo, a probabilidade de que X se situe num intervalo (a, b) é F(b) − F(a) se a ≤ b.

[18] Essa assertiva será confirmada pela análise dos dados.

Deviations), the objective function that identifies a conditional quantile is asymmetric about zero. See, for example, Koenker and Bassett (1978) and Manski (1988, Section 4.2.4). Buchinsky (1994) applies quantile regression methods to examine factors affecting the distribution of wages in the United States over time (2002, p. 367).

2.7 Limitações dos modelos

Como já explicitado, esse artigo é exploratório e pretende apenas avançar no conhecimento da dinâmica espacial do mercado imobiliário de Belo Horizonte. Quanto às limitações dessa proposta, vale destacar que as informações coletadas para o preço do imóvel são provenientes de corretoras de imóveis e, portanto, não passam de estimativas de preço de mercado que, provavelmente, são diferentes dos preços pelos quais os imóveis são vendidos. Esse fato, já detalhado na literatura, não causa problemas significativos já que se a variável dependente é medida com erro, temos:

Equação 5 – Modelo com variável medida com erro

$$Y^* = \alpha + \beta X + \varepsilon, \text{ onde } Y^* \text{ é a variável medida com erro} \quad (5)$$

Equação 6 – Adequação da variável

$$Y = Y^* + u, \text{ e } Y \text{ é a variável correta.} \quad (6)$$

Nesse caso, procede-se da seguinte forma:

Equação 7 – Ajuste do modelo

$$Y - u = \alpha + X + \varepsilon \quad (7)$$

Equação 8 – Modelo com incorporação do erro

$$Y = \alpha + X + (\varepsilon + u) \quad (8)$$

Equação 9 – Novo modelo com ajuste

$$Y = \alpha + X + w, \quad (9)$$

os erros de medida são incorporados aos erros e volta-se ao modelo original, no qual os estimadores de MQO são não-viesados, consistentes e eficientes.

Outra limitação importante, entretanto, é a correta indicação do nome do bairro dada ao imóvel que se publica para a venda. De fato, seria tentador anunciar um imóvel em bairro mais valorizado para que alcance maior valor no mercado. Contra este fato, acredita-se que a indicação errada sistemática contribua para minar a credibilidade do corretor frente ao cliente. Contribuem também para mitigar esse problema o grande número de bairros considerados, a possibilidade de bairros próximos terem valoração próxima e, finalmente, o tamanho da amostra do modelo principal (510 observações).

A análise espacial, por sua vez, sofre a limitação de que os imóveis não estão localizados nas coordenadas exatas, já que não se obteve o endereço de cada um, mas apenas o bairro no qual está localizado. Há, dessa forma, uma simplificação da localização e os imóveis são georeferenciados de acordo com os centróides do bairro em que se encontram. Seria como se todos os imóveis de determinado bairro estivessem concentrados, para efeito da análise espacial, no centro geográfico do bairro. Tal limitação está em sintonia com a proposta teórica que sustenta que os atributos espaciais são percebidos na estrutura do bairro e não na análise micro-urbana. Dessa forma, a limitação não compromete os resultados exploratórios do estudo.[19]

3 Especificação dos modelos para o caso de Belo Horizonte

3.1 Base de dados geral

A composição da base de dados foi feita por meio de coleta de informações disponibilizadas publicamente pelas corretoras de imóveis de Belo Horizonte. A coleta dos dados foi feita por estagiários[20] no *site* da empresa Rede NetImóveis, uma rede de associação de imobiliárias que compartilham suas carteiras de imóveis. A busca no *site*[21] permite a seleção por cidade, bairro e tipo (casa, apartamento de 2 quartos, 3 quartos, 4 quartos, cobertura). Uma vez selecionados os imóveis, suas características foram automaticamente computadas em planilhas. Ao todo, foram coletados dados sobre 3196 imóveis à venda em outubro de

[19] Outro estudo, com a informação de endereço de cada imóvel já está em andamento. Nesse caso, as características completas dos imóveis foram fornecidas pela empresa Rede NetImóveis.

[20] Cristiane Nobre Prudente fez a coleta de apartamentos de 2 quartos, no período de 6 a 14 de outubro de 2005. Leila Luiza Efigênio fez a coleta de apartamentos de 3 quartos, em 05 de outubro de 2005. Vladimir Augusto fez a coleta para apartamentos de 4 quartos entre 12 e 18 de outubro de 2005. O autor fez a coleta para casas e coberturas, entre 20 e 29 de outubro de 2005.

[21] <http://www.netimoveis.com.br>.

2005 em 156 bairros distintos. De todo modo, os dados não cobrem todos os 267 bairros da cidade. O viés implícito é referente à área de atuação das imobiliárias parceiras da Rede NetImóveis, bem como a formalidade da negociação. Ou seja, aquelas operações oferecidas diretamente pelos vendedores sem a intermediação de corretores membros da Rede não foram computadas. Embora esse fato limite a generalização da análise proposta, grande parte da cidade foi coberta e garante resultados exploratórios com forte indicação de comportamento.

As informações referentes a cada imóvel (tabela 1) foram: a) bairro; b) classificação de bairros por renda de acordo com a tabela de janeiro de 2005 da Fundação Ipead da UFMG;[22] c) preço (em reais); d) área (em metros quadrados); e) tipo (casa, apartamento, cobertura); f) número de quartos; g) número de banhos; h) número de vagas de garagem; i) número de piscinas; j) existência de quadras esportivas; k) existência de jardins e varanda; l) idade do imóvel, m) valores do Imposto Predial e Territorial Urbano (IPTU), (mensais, em reais) e n) valores para a taxa de condomínio em reais.

Como se verá, há observações que não contêm todas as informações. O modelo principal contou com apenas 510 observações, especialmente pela falta de informações para a variável IPTU, considerada fundamental. O detalhamento do modelo principal é feito a seguir.

Em especial, em relação a casas e coberturas só foram coletados dados referentes ao preço e à área. Em relação à informação sobre idade do imóvel, houve dificuldades no momento da coleta em discernir se a informação de "0" anos indicava lançamento de imóvel, ou informação inexistente. Por essa razão, e no intuito de garantir a qualidade dos dados, foram consideradas nas análises com essa variável somente aquelas observações (802) com idade superior ou igual a 1 ano.

Tabela 1 – Número de observações para variáveis coletadas

Dado	Número de observações
Área	3119
Bairro	3196
Banho	1875
Classificação IPEAD	3152
Idade	802
IPTU	541
Jardins	1923

[22] Essa é a primeira variável que busca representar a vizinhança. No segundo modelo, será substituída pelo índice da análise multivariada.

Número de quartos	1923
Número de vagas	1900
Piscina	1923
Preço	3195
Preço por metro quadrado	3079
Quadra	1922
Salão de festas	1923
Tipo	3196
Valor do condomínio	816
Varanda	1579

Fonte: elaboração do autor

Pode-se perceber claramente que há viés na **veiculação de imóveis para venda** já que 71% dos dados da amostra referem-se a bairros classificados como alto (3) e luxo (4) pelo Ipead. Disso resulta uma média para as observações na classificação de bairros de 2,92, enquanto a média geral, para todos os 267 bairros é de 1,65.[23] Isso gerou média geral de preços (3196 observações) relativamente elevada de R$ 245.537,00, com área de 183,98 m^2. Conseqüentemente, o preço médio do metro quadrado para a amostra é de R$ 1.447,00. Dentre os imóveis pesquisados, 73% têm uma e duas vagas de garagem e a média de idade, para aqueles imóveis com 1 ou mais anos, é de 12,75. Se desconsiderado o problema de distinção entre lançamento e informação desconhecida, a média de idade dos imóveis abaixaria para cerca de 6 anos. O preço do IPTU mensal médio foi de R$85,00 e do condomínio de R$ 305,00.

3.2 Caracterização da Amostra

Na construção dos modelos foram utilizadas somente as variáveis referentes ao a) bairro; b) classificação IPEAD (modelo 1) e índice ACP (modelo 2), c) área construída, d) valores do IPTU e do condomínio e e) idade. Os imóveis que continham informações de todas essas variáveis somaram 510 observações (tabela 2). Estão situados em 71 bairros diferentes; têm média de preço de 184 mil reais e área de cerca de 121 metros quadrados, o que leva a um preço médio do metro quadrado de R$ 1522,00. A idade média do imóvel da amostra principal é de 14 anos, ele está situado em bairro classificado como de renda alta pelo IPEAD e paga cerca de R$ 300,00 a taxa de condomínio mensal e R$ 90,00 de IPTU.

[23] Desvio-padrão de 0,92 para a amostra de 156 bairros e 0,99 para os 267 bairros.

Tabela 2 – Caracterização da amostra

	Média	Desvio-padrão	Máximo	Mínimo
Classificação bairros IPEAD	3.14	0.96	4	1
Preço (Reais)	184025	161139	1200000	30000
Área (m^2)	120.88	57.69	450	40
Idade em anos	14.31	10.05	50	1
Valor IPTU mensal	88.10	73.76	400	12.76
Valor condomínio mensal	299.08	242.66	1500	25
Vagas	1.75	0.95	6	0

Fonte: elaboração do autor

Dentre os 16 bairros com maior preço médio do metro quadrado, 14 estão classificados pelo IPEAD como do tipo 4. A distribuição do preço médio do metro quadrado acompanha a distribuição de renda do IBGE e a classificação de bairros do IPEAD.

3.3 Construção do índice de renda (ACP) e geoprocessamento – modelo 2

O índice construído com base na análise de componentes principais utilizou dez variáveis de renda domiciliar do censo demográfico do IBGE (informações por setores censitários,[24] disponibilizados pelo programa *estatcart*),[25] a saber: v0145 (sem rendimento); v0136 a v0144 (que correspondem aos extratos de: 0 a 0,5 salários mínimos; 0,5 a 1; 1 a 2; 2 a 3; 3 a 5; 5 a 10; 10 a 15; 15 a 20 e mais de 20). Na combinação linear resultante – a componente principal – as três variáveis correspondentes aos grupos com maiores rendas (acima de 10 s.m.) foram positivamente adicionadas com os seguintes coeficientes: 0.287; 0.3178 e 0.2502; os domicílios com renda entre 5 e 10 s.m. foram multiplicados por 0.0209 e as outras classes remanescentes receberam coeficientes negativos da ordem de 0.4. O programa ArcGIS apresenta os resultados no formato *raster*.[26]

Posteriormente foi necessário utilizar a camada que contém os polígonos de bairros fornecidos pela Prefeitura de Belo Horizonte para

[24] O IBGE divide o município de Belo Horizonte em 2562 setores censitários.

[25] *Estatcart* é um programa desenvolvido pelo IBGE que facilita a manipulação dos dados do censo demográfico. Lançado em 2002 está disponível para venda, em CD-ROM para municípios com maiores de 25 mil habitantes e fornece o agregado dos resultados do universo por setores censitários. Veja ainda Brasil (2003).

[26] Formato *raster* é um formato matricial georeferenciado, sendo que a cada célula corresponde um valor. Para maiores detalhes, veja Druck e outros (2004) ou livros-texto sobre geoprocessamento e análise espacial, tais como: Lillesand, Thomas M. e Ralph W. Kiefer. Remote Sensing and Image Processing. John Wiley and Sons. 1987 e Richards, John A. Remote Sensing Digital Image Analysis: An Introduction. Berlin: Springer–Verlag. 1986..

tirar a média dos valores contidos no *raster*. Ou seja, utiliza-se uma técnica de análise espacial para agregar valores ao nível da unidade espacial desejada, nesse caso o bairro. Na prática, transformou-se a informação inicialmente fornecida como total por setores censitários para média por bairro.

Note que, dessa forma, na prática, procedeu-se a uma transformação e síntese de variáveis espaciais nas quais informações por setores censitários foram sintetizadas e o resultado apresentado por bairros. Construiu-se, portanto, uma hierarquia de renda de todos os 286 bairros de Belo Horizonte.[27] Como, nesse caso, cada bairro tem seu valor específico, a representatividade da vizinhança é muito mais precisa que a simples divisão em 4 tipos fornecida pelo IPEAD.

Na análise de regressão do modelo com o índice, o valor referente a cada bairro foi adicionado para a observação equivalente. No modelo com 510 observações, temos as seguintes estatísticas básicas (tabela 3):

Tabela 3 – Dados básicos da variável índice ACP

Variável	Obs.	Média	Desvio-padrão	Mínimo	Máximo	Valores únicos
Índice ACP	510	0.59634	0.1737317	0.24401	0.91642	67

Fonte: elaboração do autor.

3.4 Divisão espacial: Bairros

A Prefeitura de Belo Horizonte utiliza no seu planejamento municipal as subdivisões espaciais em Regionais (9) e Unidades de Planejamento (81). A Fundação IBGE, por sua vez, subdivide o município em 2561 setores censitários (2000). Embora reconhecida, a divisão por bairros não é comumente utilizada pela prefeitura na disponibilização e análise de dados. No contexto desse artigo, supõe-se que a influência espacial no valor dos imóveis se dá em escala maior que a unidade de planejamento. Considerando-se, porém, que não foi possível acesso à informação por setores censitários (ou por endereços), optou-se por utilizar a divisão de bairros da Fundação IPEAD em um primeiro momento. Segundo essa divisão, tradicional no mercado imobiliário, há 267 bairros classificados em uma escala de renda de 1 a 4, popular, médio, alto e luxo[28] que tem por base a classificação de renda dos setores censitários do IBGE de 2000. No segundo momento, optou-se pela análise utilizando-se a variável índice ACP.

[27] Essa tabela está disponível e pode ser solicitada ao autor.

[28] Veja a referência para todos os bairros no *site* da Fundação IPEAD: www.ipead.face.ufmg.br.

3.5 Modelo 1 – com tipificação de bairros do IPEAD

Na análise de mínimos quadrados ordinários (MQO) tradicional, bem como na análise espacial, a forma funcional utilizada foi semi-logarítmica, formalmente:

Equação 10 – Modelo 1

$$lnPreço = \alpha + \beta_1 lnArea + \beta_2 lnIPTU + \beta_3 lnCond + \beta_4 lnIdade + \beta_5 IPEAD + \beta_6 Vagas + \varepsilon, \text{ onde:} \quad (10)$$

lnPreço é o logaritmo natural (ln) do preço em reais; lnArea é o ln da área em metros quadrados; lnIPTU é o ln da valor mensal do IPTU, em reais; lnCond é o ln do valor da taxa de condomínio paga pelo morador do imóvel; lnIdade é o ln da idade em anos do imóvel; IPEAD é o valor (de 1 a 4) da classificação do bairro no qual está situado o imóvel e vagas é o número (escalar) de vagas por apartamento.

3.6 Modelo 2 – com índice ACP

Para o modelo 2, a composição é a mesma da fórmula 10, no entanto, o item IPEAD é substituído pelo índice ACP, descrito anteriormente (tabela 3).

4 Testes e resultados

A seguir, apresentamos alguns testes básicos que garantem a qualidade do modelo. O modelo semi-logarítmico, também sugerido por Neto (2002) e Hermann e Haddad (2005) é adequado por trazer os resíduos para perto da normalidade. Ademais, o teste de Shapiro-Wilk indica a normalidade, rejeitando a hipótese nula. No intuito de testar a presença de *outliers*, além da inspeção visual da tabela dos dados, foi utilizado o teste de *inter-quartile range* que assume a simetria da distribuição. O resultado do teste não identificou nenhum *severe outlier*. O teste de Breusch-Pagan/Cook-Weisbert indicou que a variância do resíduo é homogênea e que não há heterocedasticidade.

No intuito de checar a multicolinearidade dos dados, procedeu-se ao teste de *variance inflation factor* – VIF, para o qual a regra empírica utilizada (*rule of thumb*) indica que números maiores que 10 denotam alta multicolinearidade. Os resultados mostraram que não há problema de multicolinearidade para a regressão proposta (MQO) (maior valor para a variável "lncond" 3.10 e média de 2.31). Com todas as variáveis coletadas, foi gerada a tabela de correlação (Anexo I) que demonstra que apenas as variáveis IPTU e condomínio estão mais correlacionados

com o preço (0.7880 e 0.7951, respectivamente). Entende-se que isso não compromete o modelo, apesar de já ser esperado.

Os testes de especificação realizados também indicaram que o modelo apresenta bom ajustamento e especificação. O teste de Ramsey, por exemplo, aceitou a hipótese que o modelo não contém variáveis omitidas.

Neste artigo, a matriz escolhida foi a de distância espacial dado que a configuração espacial da amostra está organizada da seguinte maneira: há numero variável de observações por bairro; todas as observações estão localizadas no seu centro geográfico; há bairros mais próximos entre si e outros mais distantes. Dessa forma, a matriz que melhor captou a influência dos vizinhos, dada a distribuição espacial muito pouco homogênea, foi a matriz de distância espacial com distância mínima para que nenhuma observação ficasse sem vizinhos.

Dada a não-contiguidade dos dados, as matrizes do tipo *queen* e *rook* não são indicadas. Matrizes com número igual de vizinhos (*k-nearest neighbors*) não contêm relação de contigüidade simétrica e, portanto, não são adequadas em modelos de estimação por defasagem espacial (ANSELIN, 2005, p. 220).[29]

Os resultados da análise tradicional (tabela 4) confirmam que, embora a localização seja importante, são determinantes as características próprias do imóvel para a constituição do preço. Todos os coeficientes são significativos e têm o sinal esperado. De acordo com o modelo proposto, o tamanho do imóvel (ln área) é o fator mais significativo na composição do preço. Um aumento de 1% na área do imóvel, levaria a aumento de 0,67 % no preço. Os serviços ofertados no prédio de apartamentos, com conseqüente aumento do valor mensal para o condomínio é o segundo fator a impactar o preço, indicando que aumento de 1% no valor da taxa de condomínio encarece o imóvel em cerca de 0,27 %. Em relação à idade do imóvel, como seria de se esperar, o aumento de, digamos, 10% na idade do imóvel, resultaria em redução do preço em cerca de 1%. Uma vaga a mais no apartamento, eleva o preço do imóvel em 14%, em média.

Em relação à variável de controle espacial, a classificação do IPEAD para bairros, o indicativo é que a mudança de um padrão de classificação, por exemplo, de médio para alto, elevaria o preço do imóvel em 11%.

[29] Ainda assim, testes revelaram valores próximos, independentemente da matriz utilizada.

Tabela 4 – Resultado modelo 1, análise MQO

Estatísticas: R^2 ajustado = 0.8896; F(6, 503) = 684,79; 510 obs.				
Variáveis	Coeficientes	Desvio-padrão	T de student prob.	>ITI
LN Área	0,6719	0,0437394	15,36	0,000
LN IPTU	0,1450	0,0254278	5,7	0,000
LN Cond	0,2728	0,0248425	10,98	0,000
LN Idade	-0,0912	0,118035	-7,73	0,000
IPEAD	0,1126	0,0149485	7,53	0,000
Vagas	0,1385	0,0170757	8,11	0,000

Fonte: elaboração do autor, *software* STATA, laboratório do CEDEPLAR/UFMG

4.1 Modelo 1 – diagnóstico espacial

Antes de validar os resultados apresentados é necessário que se testem modelos espaciais.

A estatística I de Moran indicou a presença de autocorrelação espacial, como era esperado (figura 1), e o resultado dos testes indicou a utilização de um modelo de erro auto-regressivo espacial (veja equações 3 e 4).

O valor esperado da estatística I de Moran é igual a – [1/(n-1)]. Nesse caso, como temos n = 510, se não houvesse correlação espacial dos dados, o valor seria de – 0,0020. O resultado de 0,3754, portanto, indica forte correlação, com as observações concentradas no segundo (alto-alto) e quarto (baixo-baixo) quadrantes. Ademais, a análise de pseudo-significância, por meio do processo de permutações (9999), mostrou que a estatística é altamente significante (p = 0.0001), com média de – 0,0021 e desvio-padrão de 0,0099. O método de "envelopes"[30] também garantiu a validade do resultado.

[30] Veja Anselin (2005).

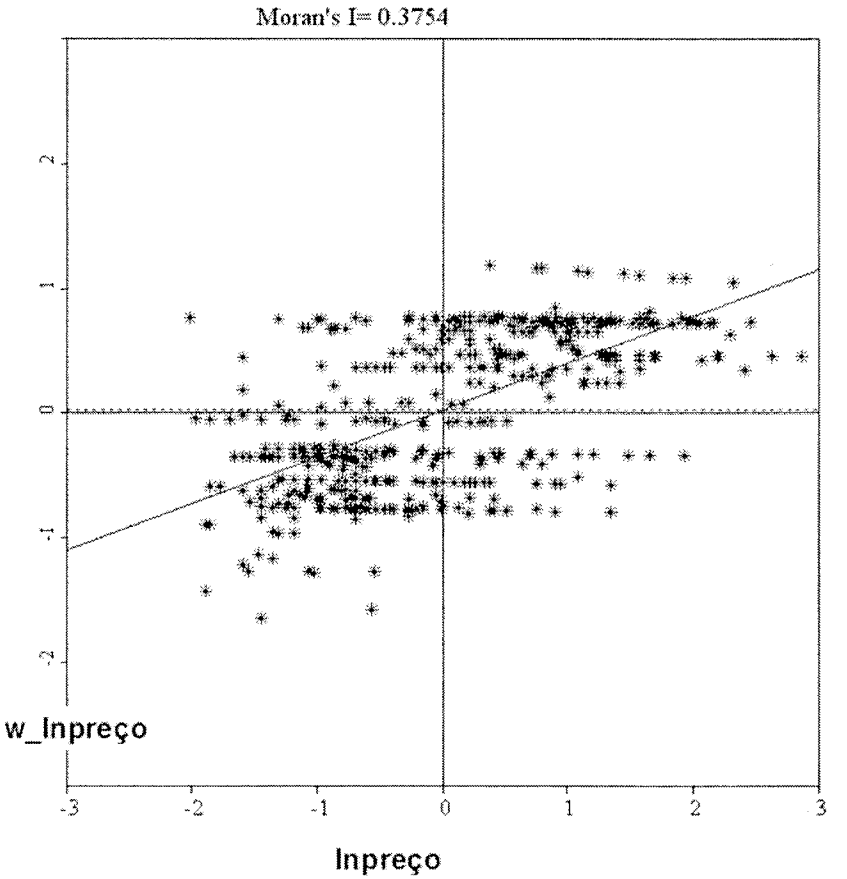

Figura 1 – Estimador I de Moran para a variável lnpreço

Fonte: elaboração do autor, *software* GeoDa, laboratório do CEDEPLAR/UFMG

Os resultados da análise de dependência espacial e dos testes (tabela 5) explicitam que o modelo de erro auto-regressivo espacial é indicado, o que enfraquece os resultados da análise tradicional (MQO/tabela 4).

Os resultados para *log-likelihood*, *Akaike Information Criterion*[31] e *Schwarz Criterion* servem para comparação entre modelos espaciais. Para o *log-likelihood* quanto maior o escalar (na reta real) melhor o ajuste. Para os *criteria* vale o oposto, ou seja, quanto menor o valor, melhor.

[31] AIC = -2L + 2K, onde L é Log-likelihood e K é o número de parâmetros do modelo. Para mais detalhes, veja Anselin, 1988.

Tabela 5 – Resultados de testes modelo 1 e diagnóstico para dependência espacial

TESTES	VALORES	Prob.
Log-likelihood	-19,8565	
Akaike Information Criterion	53,7129	
Schwarz Criterion	83,3538	
Multicollinearity Condition Number	55,50054	
Teste Jarque-Bera	6,47195	0,0393219
Teste Breusch-Pagan	19,19366	0,0038488
Koenker-Basset	15,49534	0,0167349
Teste White	173,9153	0,0000000
DIAGNÓSTICO DE DEPENDÊNCIA ESPACIAL		
Lagrange Multiplier (lag)	13,6356271	0,0002219
Robust LM (lag)	6,0663148	0,0137784
Lagrange Multiplier (error)	30,2931197	0,0000000
Robust LM (error)	22,7238075	0,0000019
Lagrange Multiplier (SARMA)	36,3594345	0,0000000

Fonte: elaboração do autor, *software* GeoDa, laboratório do CEDEPLAR/UFMG

Outros indicativos fornecidos pelos resultados são o *multicollinearity condition number*, cuja *rule of thumb* indica que números maiores que 30 são problemáticos e cujo resultado no modelo foi de 55,63. O teste Jarque-Bera confirma a normalidade dos erros e os testes Breusch-Pagan e Koenker-Basset rejeitam a presença de heterocedasticidade. Embora o teste de White (tabelas 5 e 7) sugira a presença de heterocedasticidade, Anselin (2005, p. 177) lembra que os testes contra heterocedasticidade são muito sensíveis à presença de dependência espacial e que para amostras grandes os testes Breusch-Pagan e Koenker-Basset são adequados.[32]

O diagnóstico da dependência espacial aponta ainda que "... *in the rare instance that both would be highly significant, go with the model with the largest value for the test statistic* ..." (ANSELIN, 2005, p. 218).[33]

A tabela 6 contém os resultados do modelo de erro auto-regressivo espacial. Os valores dos coeficientes, que se mantiveram pouco vari-

[32] De todo modo, para o modelo 2, foi estimado o modelo de erro heterocedástico por meio de uma estimação FGLS. Os resultados foram próximos dos apresentados na tabela 8.

[33] A título de checagem, foi realizada a regressão de defasagem espacial. No entanto as estatísticas para o *log-likelihood* e os *criterion* apresentaram valores bem inferiores aos do modelo de erro espacial.

áveis ao longo de todo o exercício, variam com pequena magnitude. A elasticidade preço-condomínio reduz sua influência (de 0,27 para 0,247) e a pequena elevação para o valor do coeficiente de vagas. Isso significa que, levando em conta a questão espacial, o valor do condomínio perde importância em relação às suas características próprias e às do prédio. Ou dito de outra forma, efeitos provocados pela vizinhança estavam sendo capturados no valor do coeficiente do condomínio. De fato, o modelo espacial apresenta melhor ajuste que o de MQO, como indicam os valores de *log-likelihood* (de -19,85 para -11,96) e os de AIC (de 53,71 para 37,92) e SC (de 83,35 para 67, 56). Todavia, ressalte-se que a presença do *lambda* significativo e a validade do modelo indicam que a especificação não está razoável e que há características espaciais importantes que não estão sendo modeladas. Nesse sentido, a classificação do IPEAD por si só não explica a diferença de preços entre os bairros, havendo distinção entre bairros mais detalhada do que sugere a proposta.

Tabela 6 – Resultados modelo 1 com Erro Espacial

TESTES	VALORES	Prob.
Log-likelihood	-11,963751	
Akaike Information Criterion	37,9275	
Schwarz Criterion	67,568376	
Teste Breusch-Pagan	24,1121	0,000498
DIAGNÓSTICO DE DEPENDÊNCIA ESPACIAL		
Log-likelihood Ratio Test	15,78545	0,0000709

Estatísticas: R^2 ajustado = 0.895085; σ^2 = 0,06881; 510 obs.				
Variáveis	Coeficientes	Desvio-padrão	Valor de Z	Prob.
LN Área	0,6696382	0,04242995	15,78221	0,0000000
LN IPTU	0,1439533	0,02478086	5,809051	0,0000000
LN Cond	0,2476281	0,02483174	9,972242	0,0000000
LN Idade	-0,09782459	0,01155754	-8,464135	0,0000000
IPEAD	0,1184579	0,01871408	6,329881	0,0000000
Vagas	0,12417038	0,01703637	8,317725	0,0000000
Lambda	0,4896931	0,1178465	4,155349	0,0000325

Fonte: elaboração do autor, *software* GeoDa, laboratório do CEDEPLAR/UFMG

Anselin (2005) recomenda ainda que se verifique a ordem das estatísticas de W, que é o quadrado do *z-value* do *lambda*; que deve

ser maior que a estatística calculada para o *likelihood ratio test* que, por sua vez, deve ser superior ao *lagrange multiplier*. Nesse exercício, essa ordem é violada, pois: $W = (4,155349)^2 = 17,2669 > LR = 15,78545 < LM = 30,29$.[34]

De fato, as considerações finais desse trabalho ressaltam que um modelo de análise do mercado imobiliário deveria contar, além das variáveis utilizadas, com maiores informações sobre as amenidades urbanas presentes em cada bairro.

O passo seguinte é exatamente a de reunir informações, no nível do bairro, para que se possa detalhar melhor o modelo.

4.2 Modelo 2 – análise clássica e diagnóstico espacial

Um segundo modelo foi implementado no sentido de capturar melhor as informações referentes aos bairros. Como já dito, substitui-se o índice produzido pelo IPEAD (que tinha apenas quatro categorias), pelo índice ACP, segundo o qual cada bairro tem seu valor específico.

Conforme se nota (tabela 7), esse modelo apresenta melhor ajuste que o anterior,[35] com todos os coeficientes significativos e ligeiramente diferentes. Nota-se que na medida em que se detalha a configuração espacial – nesse caso com um índice mais fidedigno – os coeficientes apresentam seus valores reduzidos. A variável que expressa o número de vagas, entretanto, apresenta aumento no coeficiente.

Mais uma vez, o diagnóstico indica a necessidade de implementar o modelo de erro auto-regressivo espacial, o que é feito em seguida.

Tabela 7 – Resultados MQO modelo 2 e diagnóstico espacial

Estatísticas: R^2 ajustado = 0.900637; F(7, 503) = 769.938; Prob F = 0; 510 observações				
Variáveis	Coeficientes	Desvio-padrão	Valor de t	Prob.
LN Área	0,669	0,041387	16,16218	0,0000000
LN IPTU	0,133	0,024149	5,488307	0,0000001
LN Cond	0,236	0,024070	9,804033	0,0000000
LN Idade	-0,094	0,011202	-8,364138	0,0000000
Índice ACP	0,882	0,80971	10,89738	0,0000000
Vagas	0,140	0,016166	8,63495	0,0000000

[34] Embora essa ordem também seja violada no modelo 2, os outros parâmetros de ajuste indicam que a especificação é melhor. De todo modo, ao longo de todos os modelos testados os valores dos parâmetros permaneceram muito próximos com variações da ordem de 0.01, o que equivale a cerca de um ponto percentual de variação.

[35] Expressa pelos índices: *log-likelihood*, AIC e SC.

Testes	Valores	Prob.
Log-likelihood	6,93388	
Akaike Information Criterion	0,132235	
Schwarz Criterion	29,7731	
Multicollinearity Condition Number	55,39	
Teste Jarque-Bera	8,713438	0,0128204
Teste Breusch-Pagan	14,51	0,0244301
Koenker-Basset	11,33058	0,0786813
Teste White	147,1872	0,0000000
DIAGNÓSTICO DE DEPENDÊNCIA ESPACIAL		
Lagrange Multiplier (lag)	0,1405153	0,7077693
Robust LM (lag)	2,4064662	01208349
Lagrange Multiplier (error)	21,9410448	0,0000028
Robust LM (error)	24,2069957	0,0000009
Lagrange Multiplier (SARMA)	24,347511	0,0000052

Fonte: elaboração do autor, *software* GeoDa, laboratório do CEDEPLAR/UFMG

4.3 Modelo 2 – análise espacial

O modelo de erro auto-regressivo espacial, com índice ACP, apresenta o melhor ajuste dentre os testados. Há mais uma rodada de pequenas reduções nos valores dos coeficientes e pequeno incremento no coeficiente do número de vagas. O valor para o *lambda*, embora significativo, tem seu valor de z reduzido. Isso é interessante porque indica que a especificação do modelo está captando melhor as relações espaciais.[36]

Tabela 8 – Resultados erro auto-regressivo espacial – modelo 2

Testes	Valores	Prob.
Log-likelihood	13,139767	
Akaike Information Criterion	-12,2795	
Schwarz Criterion	17,36134	
Teste Breusch-Pagan	20,24231	0,0001294
DIAGNÓSTICO DE DEPENDÊNCIA ESPACIAL		
Log-likelihood Ratio Test	12,41177	0,0004266

[36] O modelo SARMA completo também foi testado com estimação pelo método dos momentos generalizados (*two-step* e *iterated*). Mais uma vez os resultados são bastante próximos (com diferenças da ordem de 0.005), porém **Wy** não foi significativo a 5%. As maiores diferenças nos valores dos coeficientes ocorreram no índice ACP (-0.04); condomínio (+0.03) e vagas (-0.015).

Estatísticas: R^2 = 0.904840; σ^2 = 0,055220; 510 obs.				
Variáveis	Coeficientes	Desvio-padrão	Valor de Z	Prob.
LN Área	0,66793	0,04035	16,55151	0,0000000
LN IPTU	0,13369	0,02363	5,657413	0,0000000
LN Cond	0,22244	0,02389	9,311671	0,0000000
LN Idade	-0,09897	0,01101	-8,991497	0,0000000
Índice ACP	0,97956	0,09882	9,912638	0,0000000
Vagas	0,14497	0,01622	8,939214	0,0000005
Lambda	0,46669	0,12187	3,82951	0,0001284

Fonte: elaboração do autor, *software* GeoDa, laboratório do CEDEPLAR/UFMG

Numericamente, a título de exemplo, pode-se dizer que, em média, um aumento de 1% na área do imóvel (apartamentos), levaria a um aumento de cerca de 0,67 % no preço. Outro fator relevante são os serviços ofertados no prédio de apartamentos o que é aproximadamente medido pelo valor mensal da taxa de condomínio. Há indicação de que um aumento de 1% no valor dessa taxa encarece o imóvel em cerca de 0,22 %. Em relação à idade do imóvel, como seria de se esperar, um aumento de 10% na idade do imóvel, resultaria em redução do preço em cerca de 1%. Uma vaga de garagem a mais no apartamento, por sua vez, eleva o preço do imóvel em 14,5%, em média. Uma piora no padrão construtivo de 1%, medido aproximadamente pelo IPTU, levaria à redução de 0,13%. Um apartamento com as mesmas características, mas em bairro diferente, digamos, do centro (índice ACP 0,456219) para o Buritis (0,567284), elevaria o preço do imóvel em cerca de 11%.

Tome-se um imóvel da amostra: um apartamento no Buritis, de 3 quartos, com 2 vagas de garagem, 6 anos de idade e 110 m^2 de área (tabela 9). Se esse imóvel estivesse localizado no bairro Centro, seu preço cairia de 115.000 para 102.350. Por outro lado, se ele fosse 10 m^2 maior, com 120 m^2 de área, seu valor subiria para 121.983. Dessa forma, o comprador pode dispor de uma base comparativa. Ainda, segundo o exemplo, uma vaga a mais nesse imóvel vale mais (131.672) do que 10 m^2 a mais de área.

Tabela 9 – Comparação numérica ilustrativa

Preço (R$)	**115000**	**Se tivesse as mesmas características, mas**	**Novos valores**
Bairro	Buritis (0,567284)	estivesse localizado no bairro Centro (0,456219)	102350
Área (m^2)	110	tivesse 120 m^2	121983
No. Vagas	2	tivesse 3 vagas	131672

Idade (anos)	6	tivesse 7 anos	113103
Valor IPTU (R$)	73	tivesse um IPTU de 64	113105
Condomínio (R$)	120	tivesse um condomínio de 132	117558

Fonte: elaboração do autor.

4.4 Modelo 2 – análise quantílica

O modelo dois com a análise quantílica é somente exploratório, já que não é modelado espacialmente.[37] Entretanto contribui de forma relevante no entendimento do comportamento das variáveis analisadas na medida em que explicita diferenças significativas de valores dos coeficientes para quantis distintos.

Os resultados apresentados (tabela 9) são significativos[38] e consistentes com a análise teórica. Note-se que a importância do tamanho do imóvel decresce na medida em que seu preço sobe (figura 2). A presença de vagas e a idade do imóvel são menos importantes para aqueles com preços intermediários. O fator de destaque na análise quantílica é o comportamento da variável índice ACP que apresenta aumento do coeficiente na medida em que aumenta o valor do imóvel. Imóveis mais caros pagam proporcionalmente mais (em relação aos mais baratos) por localizações mais privilegiadas.[39]

Tabela 10 – Modelo 2 – análise quantílica

lnpreço	**(q.1)**	**(q.25)**	**(q.5)**	**(q.75)**	**(q.9)**
lncond	0,305 (6,68)**	0,228 (7,09)**	0,213 (8,21)**	0,259 (6,88)**	0,252 (6,00)**
lnarea	0,674 (10,19)**	0,637 (11,02)**	0,636 (14,33)**	0,634 (11,28)**	0,593 (9,97)**
lniptu	0,079 -1,71	0,153 (4,37)**	0,197 (7,55)**	0,157 (3,89)**	0,09 (1,99)*
lnidade	0,105 (4,95)**	0,06 (3,75)**	-0,071 (5,92)**	-0,097 (5,85)**	0,095 (6,36)**
vagas	0,169 (4,69)**	0,154 (6,30)**	0,12 (6,91)**	0,122 (5,53)**	0,164 (7,51)**
Índice ACP	0,638 (5,00)**	0,827 (7,93)**	1,008 (10,98)**	1,022 (7,57)**	1,104 (7,18)**

[37] A possibilidade de se rodar simultaneamente um modelo quantílico-espacial ainda está sendo investigada.

[38] À exceção dos valores do IPTU para o quantil de 10%.

[39] Note que no decil superior (q9) 1.104 o valor é muito superior ao do decil inferior (q1) 0.638. Aliás, o parâmetro é sempre crescente.

(*continuação*)

Constant	6,011 (27,23)**	6,345 (33,80)**	6,34 (43,23)**	6,39 (34,76)**	7,013 (37,65)**

Fonte: elaboração do autor, *software* Stata, laboratório do CEDEPLAR/UFMG
* Significante a 5%.
** Sugnificante a 1%.
Valor absoluto da estatística t entre parênteses.

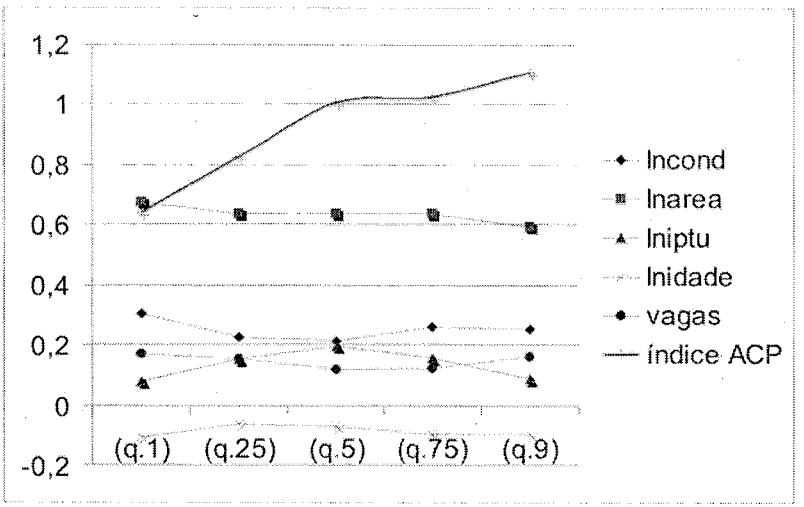

Figura 2 – evolução dos coeficientes da análise quantílica

Fonte: elaboração do autor, com base nos dados da tabela 9.

5 Considerações Finais

A discussão empírica nacional sobre o mercado imobiliário é extremamente relevante tendo-se em vista o peso da habitação nos orçamentos familiares. Entretanto, as estimativas correntemente utilizadas por corretores e imobiliárias são altamente discricionárias. É do interesse dos atores da cidade que as relações entre preços e usufruto estejam mais bem delineadas. Simultaneamente, a disponibilidade de dados georeferenciados que caracterizam a cidade permite detalhamento em qualidade e escalam antes impensáveis. Isso indica que a influência de porções pequenas da cidade, ao nível da edificação, podem agora ser mensuráveis e incluídas no modelo.

Ressalte-se ainda que embora esse artigo apresente boas indicações de variáveis relevantes e parâmetros estimados, o modelo, por ser bastante complexo, ainda se encontra incompleto. Avanços podem ser

feitos com melhorias na disponibilidade dos dados das prefeituras em relação a transações efetivamente realizadas. A análise temporal que indicasse mudanças de comportamento ao longo do tempo poderia delinear com clareza a influência de mudanças no padrão urbano, tais como construção de equipamentos públicos ou privados, abertura de lojas ou padrão de vizinhança.

Enfim, o artigo cumpre sua função ao "identificar padrões e características dos preços dos imóveis no município de Belo Horizonte". Apresenta indicativos de que há dependência espacial e que a classificação espacial por tipo de bairro não é suficiente para explicar a diferenciação espacial. Ademais, explicita a importância primeira do tamanho do imóvel, do seu tipo, do número de vagas e de banho no seu preço, permite a comparação entre bairros e mantém, ao longo do exercício, coeficientes praticamente constantes.

6 Referências

1. ALMEIDA, E. S. Curso de econometria espacial aplicada. Piracicaba: ESALQP-USP. 2004

2. ANSELIN, L. Spatial econometrics: methods and models. Dordrecht: Kluwer Academic. 1988. 284 p.

3. ANSELIN, L. SpaceStat tutorial: a workbook for using SpaceStat in the analysis of spatial data. University of Illinois. Urbana-Champaign, p.263. 1992

4. ANSELIN, L. Exploring spatial data with GeoDaTM: a workbook. Urbana-Champaign, IL: Spatial Analysis Laboratory, Department of Geography, University of Illinois. 2005. 244 p.

5. ANSELIN, L. e BERA, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: A. Ullah e D. E. A. Giles (Ed.). Handbook of Applied Economic Statistics. New York: Marcel Dekker, 1998, p.237-289.

6. BRASIL. Censo demográfico 2000: agregado por setores censitários dos resultados do universo. Documentação do arquivo. M. D. Planejamento e I. B. D. G. E. Estatística: 157 p. 2003.

7. BRUECKNER, J. K., THISSE, J.-F., et al. Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. European Economic Review, v.43, n.1, p.91-107. 1999.

8. BUCHINSKY, M. Recenty advances in quantile regression models. Journal of Human Resources, v.33, n.1, p.88-126. 1997.

9. CÂMARA, G. Análise espacial e geoprocessamento. In: S. S. Druck, M. S. Carvalho, et al (Ed.). Análise espacial de dados geográficos. Brasília: Embrapa, 2004, p.209.

10. CAN, A. e MEGBOLUGBE, I. Spatial Dependence and House Price Index Construction. Journal of Real Estate Finance and Economics, v.14, p.203-222. 1997.

11. CULLEN, G. Townscape. London. 1961. 315 p.

12. DRUCK, S., CARVALHO, M. S., et al. Análise espacial de dados geográficos. Brasília: Embrapa. 2004. 209 p.

13. FURTADO, B. A. Metrópole fractal: um modelo com autômatos celulares para análise do espaço urbano. XXXIV Encontro Nacional de Economia. Salvador, BA: ANPEC - Associação Nacional dos Centros de Pós-graduação em Economia 2006. 20 p.

14. HARVEY, D. Social justice and city. Baltimore: Johns Hopkins University. 1973. 336 p.

15. HERMANN, B. M. e HADDAD, E. A. Mercado Imobiliário e Amenidades Urbanas: A View Through the Window. Estudos Econômicos v.35, n.2, Abril-junho, p.237-269. 2005.

16. JOHNSON, R. e WICHERN, D. Applied multivariate statistical analysis: Prentice Hall. 1998. 799 p.

17. KOENKER, R. e BASSETT, G. Regression Quantiles. Econometrica, v.46, n.1, January, p.33-50. 1978.

18. LYNCH, K. e CAMARGO, J. L. A imagem da cidade. São Paulo: Martins Fontes. 1997. 227 p.

19. MACEDO, P. B. R. Hedonic price model with spatial effects: an application to the housing market of Belo Horizonte, Brazil. Textos para discussão. Belo Horizonte: CEDEPLAR/FACE/UFMG: 22 p. 1996.

20. MINGOTI, S. A. Análise de dados através de métodos de estatística multivariada. Belo Horizonte: Editora UFMG'. 2005. 297 p.

21. MONTE-MÓR, R. L. D. M. O que é o urbano no mundo contemporâneo. Textos para discussão. 281: 14 p. 2006.

22. NAIK, D. N. e KHATTREE, R. Revisiting olympic track records: some practical considerations in the principal component analysis. The American Statistician, v.50, n.2, p.140-144. 1996.

23. NETO, E. Estimação do preço hedônico: uma aplicação para o mercado imobiliário da cidade do Rio de Janeiro. (dissertação de mestrado). EGFV, Rio de Janeiro, 2002.

24. RONDON, V. V. e ANDRADE, M. V. Impactos da criminalidade no valor dos aluguéis em Belo Horizonte. Textos para discussão. Belo Horizonte: CEDEPLAR/UFMG 2003.

25. ROSSI, A. A arquitetura da cidade. São Paulo: Martins Fontes. 2001. 309 p.

26. SANTOS, M. Espaço e método. São Paulo: Nobel. 1992. 88 p.

27. SHEPPARD, S. Hedonic analysis of housing markets. In: P. Cheshire e E. S. Mills (Ed.). Handbook of regional and urban economics. Amsterdam: North-Holland, v.3, 1999, p.1595-1636.

28. SINGER, P. O Uso do Solo na Economia Capitalista. In: E. Maricato (Ed.). A produção capitalista da casa e da cidade no Brasil industrial. São Paulo: Alfa-ômega, 1982, p.21-36.

29. SMIRNOV, O. e ANSELIN, L. Fast maximum likelihood estimation of very large spatial autoregressive models: A characteristic polynomial approach. Computational Statistics and Data Analysis, v.35, p.301-319. 2001.

30. SOJA, E. W. Geografias pós-modernas: a reafirmação do espaço na teoria social critica. Rio de Janeiro: Jorge Zahar. 1993. 324 p.

31. SOJA, E. W. Third space: expanding the geographical imagination. Oxford: Blackwell Publishers. 1996. 354 p.

32. TUAN, Y. Topofilia: um estudo de percepção, atitudes e valores do meio ambiente. São Paulo. 1980. 288 p.

33. TUAN, Y. Espaço e lugar: a perspectiva da experiência. São Paulo: DIFEL. 1983. 250 p.

34. WHITTLE, P. On stationary processes in the plane. Biometrika, v.41, p.434-449. 1954.

35. WOOLDRIDGE, J. M. Econometric analysis of cross section and panel data. Cambridge, Mass.: MIT Press. 2002. 752 p.