



UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

MARIA TERESA COELHO MORETTINI

**SEGMENTAÇÃO DO MERCADO HABITACIONAL DO
MUNICÍPIO DE SÃO PAULO**

Florianópolis, 2009

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

MARIA TERESA COELHO MORETTINI

**SEGMENTAÇÃO DO MERCADO HABITACIONAL DO
MUNICÍPIO DE SÃO PAULO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Santa Catarina como requisito para obtenção do título de Mestre em Engenharia Civil

Área de Concentração: Construção Civil

Orientador: Prof. Dr. Antonio Edésio Jungles

Florianópolis, 2009.

**SEGMENTAÇÃO DO MERCADO HABITACIONAL DO
MUNICÍPIO DE SÃO PAULO**

MARIA TERESA COELHO MORETTINI

Dissertação julgada adequada para a obtenção do Título de MESTRE em Construção Civil e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Civil - PPGEC da Universidade Federal de Santa Catarina - UFSC.

COMISSÃO EXAMINADORA:

AGRADECIMENTOS

Agradeço a Deus pela clareza e paz que me possibilitaram realizar este estudo e por ter colocado pessoas tão especiais em meu caminho:

- Meus três amores, Mauricio e Renato, meus “filhos queridos” e Gerson, meu “bem” - minha linda família - razão da minha vida.

- Meu orientador Edésio, sempre pronto para atender minhas solicitações, o que possibilitou que a barreira da distância fosse vencida; os membros da banca, que souberam entender minha urgência; Heineck, banca da qualificação e professor de várias disciplinas que cursei, que me incentivou e promoveu em mim o gosto pela pesquisa; Haddad, que me possibilitou participar de suas aulas na FEA-USP; Prof. Norberto, que com intervenções tão apropriadas, possibilitou que seu pudesse fazer os ajustes para a conclusão deste trabalho; as meninas que estiveram no apoio, Mari e Dayana; todos foram além do que eu pudesse esperar.

- Meus pais, irmãos e irmã, avós, tios e tias - a grande família - plena de modelos de superação e realização, especialmente os que deixaram tantas saudades: meu pai, que me proporcionou intensas alegrias, meu avô, que me ensinou a encontrar felicidade nas coisas mais singelas da vida e minha avó, pelos momentos de ternura que ficam para sempre.

RESUMO

Na literatura sobre o mercado habitacional há um consenso em se adotar como hipótese de trabalho a existência de submercados. Com isso, o mercado imobiliário urbano deveria ser retratado como um conjunto de submercados distintos e inter-relacionados. Entender como as áreas metropolitanas são repartidas em submercados habitacionais é importante por várias razões, entre as quais: aumentar a precisão dos modelos estatísticos que são utilizados para estimar o preço das residências; melhorar a habilidade de financiadores e investidores em precificar os imóveis reduzindo o risco associado ao processo e fornecer informações para o comprador para encontrar uma residência.

Com o objetivo de se verificar a ocorrência de dependência espacial, utilizando-se dados dos lançamentos imobiliários do município de São Paulo dos últimos dez anos (3.590 condomínios verticais), foi aplicado o programa GeoDa, que permite a realização de testes estatísticos a partir de dados georreferenciados.

Quando existe dependência espacial, deve-se aplicar modelos espaciais hedônicos para análise do mercado, que consideram não apenas as características físicas das residências, mas também atributos da localidade e aspectos da vizinhança. O GeoDa possui diversos testes estatísticos para verificar a dependência espacial dos dados e também está apto a efetuar a modelagem de Regressão Clássica e Regressão Espacial, além da identificação do modelo espacial que melhor se adapta aos dados.

Foi constatada a ocorrência de autocorrelação espacial e verificou-se que o modelo de erro espacial é o que melhor se ajusta ao conjunto de dados. Pelas análises efetuadas conclui-se que o mercado habitacional no município de São Paulo é segmentado.

ABSTRACT

On the housing market literature there is a consensus to adopt the existence of sub-markets and the recognition that the urban housing market should be portrayed as a set of distinct and inter-related sub-markets. It is important to understand how the metropolitan areas are divided into housing sub-markets for several reasons, including: increasing the accuracy of statistical models that are used to estimate the price of houses, improving the ability of lenders and investors in define the risk financing; provide information to the buyer to find a residence.

Therefore, this study aims to determine the housing market segmentation in São Paulo using property data from the last ten years (3,590 vertical condominiums). The program GeoDa was used, since it allows doing statistical tests from georeferenced data.

The price of a house is likely to be related to the price of adjacent properties. Hedonic price models cannot perfectly capture the effects of location. The aim of this work is to analyse how best to take into account spatial dependence to our understanding of the urban housing price determination process. The use of spatial models to estimate hedonic functions provides a way to define housing submarkets through statistical modeling.

The results indicates the occurrence of autocorrelation and the segmentation of São Paulo housing market.

SUMÁRIO

1 - INTRODUÇÃO.....	23
1.1 Motivação para a pesquisa.....	24
1.2 Justificativa da pesquisa.....	25
1.3 Objetivos.....	26
1.3.1 Objetivo geral.....	26
1.3.2 Objetivos específicos.....	26
1.4 Premissas.....	26
1.5 Delimitações.....	27
1.6 Estrutura do trabalho.....	28
2 - REVISÃO BIBLIOGRÁFICA.....	29
2.1 Marketing.....	29
2.1.1 O mercado.....	30
2.1.2 Segmentação do mercado como estratégia de marketing.....	32
2.1.3 Sistemas de informação.....	35
2.2 Mercado habitacional.....	36
2.2.1 Produto - características da habitação.....	36
2.2.2 Eficiência do mercado.....	38
2.2.3 Informações sobre o imóvel.....	39
2.2.4 Estabilidade do mercado – mudança dos preços (“ <i>filtering</i> ”).....	40
2.2.5 Equilíbrio do mercado.....	41
2.3 Estudos sobre o mercado habitacional - natureza dos submercados.....	43
2.3.1 Primeiros Estudos – conceito de substituição.....	44
2.3.2 Modelos de equilíbrio urbano.....	45
2.3.3 Heterogeneidade do produto – demanda por atributos....	47
2.3.4 Características da oferta e da demanda.....	48
2.3.5 Dinâmica do mercado.....	51

3 - ECONOMETRIA ESPACIAL.....	53
3.1 Autocorrelação espacial.....	55
3.2 Heterogeneidade espacial	57
3.3 Matriz de pesos espaciais – W –.....	59
3.4 Estudos sobre heterogeneidade do mercado habitacional – segmentação	60
4 - MÉTODOS PARA A SEGMENTAÇÃO DO MERCADO HABITACIONAL.....	65
4.1 Identificação dos submercados habitacionais.....	65
4.1.1 Submercados espaciais.....	67
4.1.2 Submercados estruturais.....	69
4.1.3 Submercados espaciais e estruturais em conjunto.....	69
4.2 Métodos para segmentar o mercado habitacional.....	70
4.2.1 Submercados pré-definidos ou definidos <i>a priori</i>	72
4.2.2 Análise dos componentes principais - PCA (Principal Component Analysis).....	73
4.2.3 Análise de <i>cluster</i>	74
4.2.4 Modelos hedônicos.....	75
4.3 Modelos de Regressão com Efeitos Espaciais.....	76
4.3.1 Modelo Geral com Efeitos Espaciais Globais.....	77
4.3.2 Modelo Espacial Autoregressivo de Primeira Ordem.....	77
4.3.3 Modelo de defasagem espacial – (<i>Spatial Lag Model</i>) ...	78
4.3.4 Modelo de erro espacial	78
4.3.5 Modelos com Efeitos Espaciais Locais.....	78
4.3.6 Modelagem de forma discreta.....	79
4.4 GeoDa.....	80
4.4.1 Testes de autocorrelação espacial	80
4.4.1.1 Índice <i>I</i> de Moran Global – Índice de Associação Espacial Global.....	81
4.4.1.2 LISA (Local Indicator of Spatial Analysis) – Índice de Autocorrelação Local.....	83
4.4.2 Análise gráfica de dependência espacial.....	86
4.4.2.1 Diagrama de espalhamento de Moran (Moran Scatterplot Map).....	86
4.4.2.2 LISA map e Moran map	88

4.4.3 Análise exploratória	88
5 - METODOLOGIA PARA ANÁLISE DE DADOS.....	91
5.1 Etapas para verificação da dependência espacial e ajuste do modelo de regressão espacial.....	92
5.2 Verificação das suposições do modelo, para constatação de necessidade de ações corretivas – validade do modelo.....	94
5.3 Delineamento da pesquisa	95
5.3.1 Variáveis.....	97
5.3.2 Dados - Coleta e preparação.....	98
5.3.3 Testes estatísticos	99
5.4 Dados do setor imobiliário.....	99
6 - RESULTADOS – ANÁLISE DE DADOS	103
6.1 Pontos plotados.....	103
6.2 Histograma de conectividade.....	104
6.3 Matriz de pesos espaciais baseada na distância entre pontos..	106
6.4 I de Moran para a variável Preço	106
6.5 Teste de significância do Índice de Moran	107
6.6 Autocorrelação espacial local – LISA	108
6.7 Gráfico de Moran.....	109
6.8 Regressão: verificação dos pressupostos e escolha de variáveis a serem incluídas no modelo	111
6.9 Dependência espacial.....	113
6.10 Análise dos resíduos para a verificação de clusters e segmentação.....	115
6.11 Mapa LISA – Clusters e significância	116
7 - CONCLUSÃO.....	119
REFERÊNCIAS BIBLIOGRÁFICAS	123
ANEXOS	133

LISTA DE FIGURAS

Figura 1 - Distribuição de referência para o I de Moran constatando a não-aleatoriedade. Fonte: ANSELIN, 2003.	85
Figura 2 - Distribuição de referência para o I de Moran constatando a aleatoriedade. Fonte: ANSELIN, 2003.	85
Figura 3 - Gráfico de espalhamento de Moran.	87
Figura 4 - Fluxograma: processo de decisão para escolha do modelo de regressão. Fonte: ANSELIN, 2005.	92
Figura 5 - Mapa do município de São Paulo contendo lançamentos de imóveis	102
Figura 6 - Pontos plotados correspondentes a lançamentos verticais. .	103
Figura 7 - Histograma de conectividade.....	104
Figura 8 - Pontos com 1 a 104 vizinhos (em amarelo).....	105
Figura 9 - Randomização dos dados.....	108
Figura 10 - Preço em relação ao Preço dos vizinhos (alto-alto).	110
Figura 11 - Preço em relação ao Preço dos vizinhos (baixo-baixo)	111
Figura 12 - Mapa LISA clusters.	117
Figura 13 - Mapa LISA de significância	118
Figura 14 - Mapa geográfico do município de São Paulo.	135
Figura 15 - Mapa da região metropolitana e distritos do município de São Paulo – taxa de crescimento anual da população.	136
Figura 16 - Mapa do município de São Paulo – Regiões, subprefeituras e distritos.....	138

LISTA DE TABELAS

Tabela 1 - Autocorrelação espacial de acordo com a matriz de pesos para ajuste da distância.....	106
Tabela 2 - Normalidade dos resíduos	112
Tabela 3 - Modelo de regressão tradicional: Mínimos Quadrados Ordinários.....	113
Tabela 4 - Dependência espacial.....	114
Tabela 5 - Modelo de regressão espacial (<i>Spatial error</i>).	114
Tabela 6 - Maiores aglomerados urbanos do mundo.....	135
Tabela 7 - Dados da região metropolitana e do município de São Paulo.....	136
Tabela 8 - Domicílios segundo tipologias residenciais. Município de São Paulo – (Censo 2000).	140

LISTA DE GRÁFICOS

Gráfico 1 - Preço do lançamento versus média de preços dos lançamentos vizinhos localizados dentro de um raio de 150m.	107
Gráfico 2 - Preço em relação ao Preço dos vizinhos (alto-alto).	109
Gráfico 3 - Preço em relação ao Preço dos vizinhos (baixo-baixo). ...	110
Gráfico 4 - Resíduos do modelo.....	112
Gráfico 5 - Resíduo do preço de um lançamento versus média do resíduo do preço dos lançamentos vizinhos dentro de um raio de 1000m.....	116
Gráfico 6 - Número de apartamentos lançados e média de área útil: 1992 a 2007 – Município de São Paulo – Fonte: Embraesp.....	139
Gráfico 7 - Domicílios segundo tipologias residenciais. Município de São Paulo – (Censo 2000).	141

1 - INTRODUÇÃO

O desenvolvimento de métodos de análise de cenários e de tendências de mercado aplicados à construção civil é uma prioridade na busca de inovação para o setor.

O mercado imobiliário representa um dos segmentos mais importantes na economia nacional. O investidor deste mercado busca atingir benefício financeiro com a venda de unidades habitacionais e tem como base o objetivo de obter a satisfação dos clientes.

As análises de viabilidade econômicas em empreendimentos habitacionais estão sempre condicionadas a expectativa de aceitação do produto.

O processo de decisão de investimentos para lançamento de novas construções, na etapa de análise de viabilidade, requer uma avaliação cuidadosa de oportunidades no ambiente externo à empresa, além da avaliação interna, pois os riscos do setor são altos.

Mercados habitacionais são fenômenos complexos, que consistem de um balanço extremamente heterogêneo, com produtos diversificados, muitos produtores e com consumidores com diversos gostos e orçamentos. Em grandes áreas metropolitanas é provavelmente demasiado simplista para muitos propósitos se pensar em termos de um mercado unitário, pois pode haver importantes discontinuidades em toda região e residências diversificadas (BOURASSA et al., 1997).

Informações sobre a oferta concorrente além do comportamento da demanda, em termos de preferência e necessidades são fundamentais para a análise do mercado no processo de concepção do empreendimento.

Na tentativa de se encontrar nichos de mercado com necessidades não atendidas, se refinam técnicas de segmentação de mercado e da subdivisão dos segmentos.

A tarefa da divisão de um grande mercado em submercados suscita numerosas questões teóricas e metodológicas (PALM, 1978). Um problema é a definição do submercado.

É importante entender o mercado e segmentá-lo, antes de decidir o que produzir, para se obter um produto rentável e que atenda as necessidades do público alvo.

Entender como as áreas metropolitanas são repartidas em submercados habitacionais é importante por várias razões. Primeiro, apontar propriedades dos submercados habitacionais aumenta precisão dos modelos estatísticos que são utilizados para estimar o preço das residências. Segundo, identificar as fronteiras do submercado dentro de áreas metropolitanas permite a pesquisadores encontrarem modelos melhores da variação espacial e temporal em seus preços. Terceiro, uma determinação correta da propriedade dos submercados melhora a habilidade de financiadores e investidores de precificar o risco associado ao financiamento para o mutuário. Finalmente, fornecer informações sobre a delimitação dos submercados para o comprador reduz seu custo para encontrar a residência (GOODMAN e THIBODEAU, 2007).

Através de métodos estatísticos e com base na análise e na interpretação de dados oferecidos pelas instituições ligadas ao setor da construção civil, pode-se segmentar o mercado habitacional.

A análise econômica do mercado habitacional no Reino Unido e nos Estados Unidos tem sido extensa nas últimas décadas e, com isso, tem-se uma ampla literatura disponível. O primeiro foco de atenção foi para as causas e efeitos da dispersão geográfica. Uma das abordagens teóricas dominantes toma como ponto de partida a relação entre acessibilidade e preços, conhecida como teoria dos “trade-off”. Esta tem recebido críticas especialmente dirigidas ao realismo de suas premissas, pois assume um mecanismo de equilíbrio estático como base da teoria e a suposição de que o mercado imobiliário urbano pode ser visto como um todo, no qual um único conjunto de processos gera determinantes comuns dos preços das residências. Ao contrário do último pressuposto, pode-se argumentar que o mercado imobiliário deve ser tratado como uma série de submercados.

A habitação se diferencia dos outros bens principalmente por dois fatores: a durabilidade e a imobilidade. Como consequência, a dimensão espacial desempenha um fator importante no mercado habitacional.

A maioria das análises é orientada para esclarecer a relação entre o preço e os atributos das residências com uma interpretação do significado espacial.

1.1 Motivação para a pesquisa

A produção do mercado imobiliário quando bem ajustada às reais necessidades dos usuários, promove ganhos não só para os

empreendedores, mas também para a sociedade que passa a ser mais bem atendida quando aos seus anseios e preferências são considerados.

O produto habitação exige elevados investimentos e, por ter características que o distinguem de outros bens do mercado, como a imobilidade e a heterogeneidade, necessita de estudo teórico diferenciado.

Adotando uma metodologia para a avaliação de dados oferecidos por empresas de pesquisa ligadas ao mercado imobiliário, o empreendedor pode melhorar suas condições para planejar os investimentos.

1.2 Justificativa da pesquisa

O empreendedor do mercado habitacional precisa conhecer alternativas que maximizem o retorno de seus investimentos. O objetivo da análise do mercado é de reduzir riscos e incertezas inerentes à atividade empresarial e auxiliar o empreendedor a conceber um projeto rentável e que atenda as necessidades do público alvo (FERNANDEZ, ILHA e OLIVEIRA, 2003).

Muitas vezes, a decisão de investimento é tomada pelo empresário de forma intuitiva, de acordo com a sua experiência e percepção das condições do mercado, sem ter como base uma análise criteriosa, fundamentada em dados (GONZALEZ, 2000).

Pode-se conseguir uma grande quantidade de informações sobre o mercado imobiliário, através de dados divulgados por associações ligadas ao setor. Utilizando-se métodos estatísticos, pode-se buscar interpretá-los para se obter indicações sobre o comportamento do mercado habitacional.

No ambiente estudado, as incertezas são muitas e os riscos são altos. É importante melhorar o sistema de avaliação das informações para que o investidor possa tomar as decisões em um patamar menor de risco.

Há aproximadamente cinco décadas, uma importante contribuição da economia de mercado habitacional é o reconhecimento de que o mercado imobiliário urbano deveria ser retratado como um conjunto de submercados distintos, mas inter-relacionados, através dos quais os preços hedônicos são significativamente diferentes. Na literatura há um consenso de que submercados habitacionais deveriam ser adotados como hipóteses de trabalho. Suas aplicações são espalhadas através de avaliação de propriedade e previsão de preço (ADAIR et al.,

apud TU, 2007), análises do mercado habitacional local (MACLENNAN e TU, 1996), planejamento urbano (JONES, 2002), assim como submercados baseados em análise de risco de pagamento de hipoteca.

1.3 Objetivos

1.3.1 Objetivo geral

Verificar a segmentação do mercado habitacional de São Paulo, para auxiliar o processo de decisão dos empreendedores da construção habitacional na escolha do produto a ser ofertado.

1.3.2 Objetivos específicos

- Verificar a autocorrelação espacial, ou seja, a relação do preço de um imóvel com os preços dos imóveis próximos a ele;
- Verificar quais as características físicas e locacionais do imóvel são relevantes em relação ao preço, através da análise das variáveis: área útil, área de condomínio (área total do apartamento menos área útil), distância à estação de metrô mais próxima, distância à escola particular mais próxima, distância ao centro da cidade (distância até a Praça da Sé) e data do lançamento do condomínio;
- Avaliar a importância da localização e verificar a ocorrência de clusters, ou seja, a existência de diferentes padrões locais na área pertencente ao município de São Paulo, para se constatar se este mercado habitacional está segmentado;
- Verificar a localização dos submercados habitacionais;
- Verificar a adequabilidade do uso do programa GeoDa para análises do Mercado Imobiliário.

1.4 Premissas

- As características da oferta, quanto a variáveis e atributos não testados, são consideradas homogêneas;

- Diferentes níveis de oferta não influenciam o resultado;
- Fatores macro econômicos são considerados estáveis para o escopo da análise e não interferem no resultado;
- Diferentes formas de pagamento dos imóveis não alteram o preço. Condições de financiamento da habitação são consideradas estáveis no período, não afetando significativamente as condições de acesso à compra do imóvel. A flutuação do dólar no período não interfere no resultado;
- Municípios limítrofes são considerados homogêneos e não causam diferenciais ao longo do perímetro do município;
- Não existem diferenças substanciais no sistema urbano por todo o território do município de São Paulo que possam interferir no resultado, especialmente referente à malha urbana e, assim, as distâncias têm representatividade e equivalem à acessibilidade.

1.5 Delimitações

- O estudo utiliza dados secundários e as variáveis testadas são apenas as divulgadas nas pesquisas;
- Os dados divulgados por entidades setoriais não abrangem todos os lançamentos imobiliários;
- Não são considerados fatores macroeconômicos e aspectos demográficos das regiões estudadas;
- A organicidade da demanda não é levada em consideração.

Para Tu (2007), modelos de submercado devem ser usados em análise de mercado, somente se os submercados estiverem cuidadosamente definidos, refletindo diferenças verdadeiras em preços hedônicos.

Watkins (1998) conclui que existem imprecisões consideráveis na conceitualização de submercados imobiliários, que limitam as aplicações de conceitos de submercado, causadas por critérios de demarcação ou de definições, por métodos estatísticos inadequados, assim como pela seleção das unidades de análise.

Palm (1978) considera que a maior fraqueza na tentativa de definir submercado empiricamente é a falha de levar em conta os

caminhos pelos quais o imóvel foi adquirido e usar as informações de agentes imobiliários, mas ela demonstrou que submercados definidos por este tipo informação são melhores que aqueles definidos por critérios econômicos.

1.6 Estrutura do trabalho

A dissertação de mestrado encontra-se organizada em cinco capítulos:

No primeiro capítulo tem-se a introdução na qual serão abordadas questões pertinentes ao tema estudado, motivação e justificativa, seguida dos objetivos (gerais e específicos), das premissas e das delimitações do estudo e da estruturação da proposta.

O segundo capítulo apresenta uma revisão dos conceitos sobre marketing, mercado habitacional, submercados, sistema de informação e fontes de pesquisas do setor imobiliário, incluindo também estudos relevantes sobre segmentação do mercado imobiliário. O terceiro, revisão conceitual sobre autocorrelação espacial.

No quarto capítulo apresenta-se o procedimento metodológico aplicado ao estudo de forma a se obter os objetivos estabelecidos no primeiro capítulo. São também abordados neste capítulo os métodos utilizados por autores dos estudos que foram relacionados nos capítulos anteriores delineamento da pesquisa.

O quinto capítulo refere-se aos resultados encontrados a partir da investigação realizada e o sexto capítulo às conclusões extraídas das análises realizadas.

No final, são apresentadas as referências bibliográficas.

2 - REVISÃO BIBLIOGRÁFICA

2.1 Marketing

O marketing é definido como a arte de descobrir oportunidades, desenvolvê-las e lucrar com elas (KOTLER, 1999). O objetivo das empresas deve ser o crescimento lucrativo. Assim, o profissional responsável pelo marketing deve avaliar necessidades do consumidor, medir extensão e intensidade destas.

Além das necessidades, deve-se buscar entender os desejos e as demandas do mercado. As necessidades são requisitos humanos básicos como alimento, abrigo, instrução e diversão, que se tornam desejos quando direcionadas a objetos específicos que possam satisfazê-las. O marketing, da mesma forma que certos fatores sociais, pode influenciar os desejos, mas não cria necessidades, pois elas existem antes deles. (KOTLER e KELLER, 2006).

Na década de 1990 ocorreu uma transição do marketing tradicional para o marketing de relacionamento, o que exige um conhecimento mais profundo e detalhado dos determinantes do comportamento do cliente e de seu processo decisório. O termo cliente, mais abrangente que consumidor, passou a ser utilizado referindo-se às pessoas que assumem diferentes papéis no processo de compra, com participação direta ou indireta, como o especificador, o influenciador, o comprador ou decisor, o pagante e o usuário ou aquele que consome o produto (BRETZKE, 2003).

Kotler e Keller (2006) consideram que as decisões de compra baseiam-se em como os consumidores percebem o preço. O processo de compra segue uma seqüência de etapas: reconhecimento do problema, busca de informações, avaliação de alternativas, decisão de compra e comportamento pós-compra. Para planejar e implantar estratégias de preço com eficácia é preciso entender o comportamento do comprador, assim como as influências que ele recebe.

O cliente desempenha papel forte e decisivo na formação de preços e a exploração correta dos atributos do produto que constituirão a percepção de valor do consumidor, deve ser sustentada pelo preço sugerido. Cada produto é formado por um conjunto de atributos que o caracterizam e que possuem diferentes graus de importância para o

consumidor, portanto também é relevante na formação dos preços a análise de atributos do produto.

O critério mais utilizado para classificar as situações de compra é aquele que considera a percepção de risco e a experiência anterior com compras do mesmo tipo. Relacionado com o envolvimento está o nível de planejamento da compra. A compra de valor ou compra planejada associa o montante investido com o risco da decisão. Na compra de imóveis residenciais, além do alto valor a ser investido no bem, existem fatores como o comprometimento da renda disponível, tipo de construção e localização do imóvel, que representam decisões importantes e associam um grande custo psicológico à compra, gerando uma percepção de risco muito maior do que aquela associada ao valor do investimento (BRETZKE, 2003).

Para uma correta interpretação das preferências do cliente, de acordo com Oliveira et al. (1998), devem ser avaliadas variáveis econômicas, como renda mensal e valor patrimonial familiar, em conjunto com a variável ciclo de vida. O ciclo de vida de uma família compreende as formas como a família está estruturada ao longo da vida de uma pessoa e o comportamento de compra pode estar associado a cada ciclo. Estes autores verificam que a evolução entre estágios do ciclo de vida familiar reflete na preferência por determinados atributos de um produto, como, na escolha da residência, o número de quartos.

De acordo com Brandstetter e Heineck (2004), o estudo sobre o comportamento do cliente em relação à habitação requer uma adaptação da teoria do marketing devido ao custo elevado do produto, compras não frequentes, alto envolvimento, classe e marca desconhecidas e significativo tempo gasto em reflexão e procura do produto. Para estes autores, devem ser considerados diferentes conceitos tratados em estudos sobre a interação entre o homem e a habitação, provenientes da literatura da arquitetura, e sobre o comportamento do consumidor em relação à habitação, discutidos nos textos da geografia sócio-econômica.

2.1.1 O mercado

O mercado corresponde ao conjunto de pessoas com necessidades que podem ser supridas por produtos, e que dispõe de renda para adquiri-los. O mercado-alvo é aquele para o qual a empresa vai direcionar seus investimentos e focar sua estratégia de marketing. Ele é selecionado com base na análise de atratividade, considerando os resultados de receitas e lucros a serem alcançados pela empresa, com a

análise das vantagens competitivas em relação à concorrência e da demanda potencial (LIMEIRA, 2003).

De acordo com Kotler (1999), uma das decisões mais importantes que uma empresa precisa tomar é se vai tratar o mercado de maneira homogênea e até que ponto. No marketing de massa, a empresa oferece um produto ou serviço padrão para todo o mercado. Kotler e Keller (2006) citam que os mercados podem ser separados em quatro níveis: segmentos, nichos, áreas locais e mercado individual.

O marketing de segmentos é praticado por empresas que projetam seus produtos para um ou mais segmentos específicos e não para todo o mercado. Segmento significa um grande grupo de consumidores, que possuem as mesmas preferências. Ao decidir adotar o marketing de segmentos, uma empresa pode decompor o mercado em segmentos cada vez menores (KOTLER e KELLER, 2006).

A empresa pode procurar por um nicho atraente dentro de um segmento e definir seu mercado alvo, visto que, um nicho envolve clientes com necessidades distintas e assim, tem potencial para gerar maiores lucros pela especialização. Nicho de mercado corresponde a um grupo definido mais estritamente que procura por um *mix* de benefícios distintos. Enquanto os segmentos atraem vários concorrentes, os nichos são pequenos e atraem uma ou duas empresas (KOTLER e KELLER, 2006).

O marketing local, ainda de acordo com estes autores, considera programas de marketing idealizados de acordo com os desejos e as necessidades de grupos de clientes locais, de determinadas áreas comerciais ou bairros, e pode ser direcionado conforme as características demográficas. Em um nível mais específico está o marketing para clientes, praticado por empresas que se concentram em cada cliente individual, adaptando a ele seus produtos ou serviços e sua comunicação.

O marketing de mercado-alvo eficaz exige a segmentação do mercado, a definição de um ou mais segmentos-alvo (mercado-alvo), e o posicionamento de mercado. Posicionamento consiste em fazer com que os clientes percebam quais são os principais benefícios ou o que diferencia o produto da empresa.

Quanto à determinação do preço, as empresas geralmente não determinam um preço único, mas têm uma estrutura de preços que reflete variações geográficas na demanda e no custo, exigências de segmento de mercado, oportunidade de compra além de outros fatores, e adotam certas estratégias de adequação de preços como: preço com

desconto, preço promocional, preço geográfico e preço discriminatório (KOTLER e KELLER, 2006).

Os preços geográficos são preços diferentes para clientes em diferentes locais; os discriminatórios são preços diferentes para diferentes grupos de clientes, para um mesmo produto. Os autores acima citados indicam que, para que a discriminação de preços funcione, uma condição essencial é que o mercado seja segmentável, e os segmentos devem mostrar diferentes níveis de demanda.

Smith, em uma publicação de 1956, que, segundo Kotler (1999), foi um artigo revolucionário, comparou a empresa que oferecia uma variedade de produtos (diferenciação por produto) com a empresa que projetava produtos para segmentos específicos do mercado (segmentação de mercado), e considerou que a segmentação é condição indispensável para o crescimento da empresa.

2.1.2 Segmentação do mercado como estratégia de marketing

O processo de segmentação é fundamental para se desenvolver qualquer plano ou estratégia de marketing e consiste em dividir grupos de consumidores potenciais, com necessidades e características semelhantes, já que, provavelmente, eles apresentarão comportamento de compra semelhante (DIAS, 2003).

Na segmentação do mercado, um produto é diferenciado pelo produtor para que seja atingido um subgrupo de consumidores. Quando o mercado é subdividido em classes distintas de consumo homogêneo, ou seja, com consumidores com mesmas preferências, obtém-se submercados (KOTLER e KELLER, 2006).

De acordo com Smith (1956), segmento significa um grupo de clientes com necessidades e anseios comuns entre si e um tanto diferentes dos clientes de outros segmentos. Desta forma, pode-se considerar que o mercado seja heterogêneo, formado por mercados homogêneos menores.

Deve-se ter o cuidado de assegurar que os segmentos reflitam exatamente os consumidores. Porém, raras vezes, se identificam segmentos totalmente homogêneos, em que todos os clientes potenciais são semelhantes. Na segmentação de mercado tenta-se isolar as características que distinguem um certo grupo de consumidores do mercado e entender quais as características do grupo, como idade, localização geográfica, poder aquisitivo e padrões de compra, que afetam seu comportamento de compras (BOONE e KURTZ, 2009).

As principais bases para segmentação de mercado são: segmentação geográfica, segmentação demográfica, psicográfica e comportamental, além de aspectos relacionados com o produto (KOTLER e KELLER, 2006).

Na segmentação geográfica ou espacial os segmentos são identificados com base em critérios geográficos, como localização e densidade populacional. De acordo com Weiss (apud DIAS, 2003), o lugar em que se vive determina a maneira como se vive e conhecendo-se onde e como as pessoas vivem fica mais fácil de atender às suas necessidades.

A localização geográfica não garante que todos os consumidores de uma determinada localidade tomarão as mesmas decisões de compra, mas essa abordagem de segmentação ajuda a identificar alguns padrões gerais. A segmentação geográfica faz distinções úteis quando existem necessidades ou preferências regionais, mas não deve ser o único método de segmentação, porque distinções entre consumidores também existem dentro de uma determinada localização geográfica (BOONE e KURTZ, 2009).

De acordo com Kotler e Keller (2006), as necessidades e desejos dos consumidores, assim como as preferências por marcas e produtos, estão frequentemente associados às variáveis demográficas. A segmentação demográfica diz respeito a grupos de variáveis básicas como faixa etária, sexo, tamanho das famílias, ciclo de vida, renda, classe social e outras como religião e raça. Alguns autores (DIAS, 2003; FREITAS, 2000) consideram como sócio-econômica a segmentação ligada à renda, escolaridade, ocupação profissional e classe social.

Doling (apud Freitas, 2002) associa a passagem pelos estágios do ciclo de vida familiar ao acúmulo de patrimônio, relacionando a compra de uma residência de maior valor à venda de uma moradia anterior.

Para Fernandez e Hochheim (2000), o estudo sobre atributos da localização na escolha da habitação e sua correlação por cada estágio do ciclo de vida familiar, fornece bases para a elaboração de uma efetiva estratégia de segmentação.

Já que o mercado pode ser definido como pessoas com poder de compra, Boone e Kurtz (2009) consideram que os padrões de despesa e a renda sejam bases para segmentar o mercado consumidor. Os autores citam que os profissionais de marketing, geralmente, têm como alvo áreas geográficas, pois estas estão associadas à renda de seus moradores. Estes profissionais podem, também, considerar a idade ou tipo de família para determinar o poder de compra de um grupo de consumidores.

Ainda, para os autores acima, de acordo com as “leis de Engel” - publicadas mais de um século atrás, e que ainda são válidas - que se baseiam no estudo do impacto das mudanças de renda familiar sobre o comportamento de gasto dos consumidores, a porcentagem gasta com moradia e serviços domésticos permanece constante mesmo quando acontece mudança da renda.

Para Bretzke (2003), a situação econômica, ou seja, a renda disponível, a economia ou a facilidade de financiamento, considerando o crédito e as taxas de juros, determinam não só o que comprar, mas quando comprar. Em uma análise da segmentação do mercado habitacional em termos de tipologias habitacionais, Freitas (2000) faz uma associação dos atributos físicos do imóvel com as características sócio-econômicas das famílias que os demandam.

A segmentação por renda é uma prática bastante utilizada em várias categorias de produtos. Schnare e Struyk (1976) examinaram grupos de acordo com a faixa de renda, alta ou baixa, e Munro (1986, apud WATKINS, 1998) além da faixa de renda também examinou subgrupos de compradores de acordo com o grau de dificuldade experimentado durante o processo de escolha das residências.

As características geográficas e demográficas são as bases primárias para dividir os consumidores em segmentos homogêneos de mercado. Porém, com o reconhecimento da necessidade de retratar de maneira mais completa e natural os consumidores, e ter uma visão mais exata do comportamento de compra destes, muitas empresas investem na pesquisa psicográfica, e assim, podem identificar o que os consumidores de vários segmentos demográficos e geográficos querem e necessitam (BOONE e KURTZ, 2009).

Na segmentação psicográfica, os grupos homogêneos são formados com base no modo como as pessoas pensam, ou seja, a personalidade, os valores e o auto-conceito, e na forma como elas levam suas vidas, ou melhor, estilo de vida, atitudes, atividades, interesses e opiniões (DIAS, 2003).

A segmentação comportamental se refere a grupos de consumidores reunidos de acordo com seus conhecimentos, uso e resposta ao produto. Variáveis comportamentais, como índice de utilização e atitudes em relação ao produto, podem ser pontos de partida para definir segmentos de mercado (KOTLER e KELLER, 2002). A identificação de segmentos através das respostas do consumidor em relação ao produto é realizada com estudo de padrões de consumo e benefícios procurados.

Segmentar pela busca de benefícios é uma abordagem que enfoca os atributos que as pessoas buscam e os benefícios que esperam receber de um produto ou serviço (BOONE e KURTZ, 2009).

Smith (1956) considera que a segmentação é uma estratégia relativa à oferta de mercado, com um ajuste mais racional e preciso do produto, que representa um esforço para entender as necessidades do consumidor.

2.1.3 Sistemas de informação

O sistema de informações de marketing tem como objetivo coletar, analisar e distribuir informações relevantes sobre o mercado e os indicadores de desempenho a fim de detectar problemas, identificar oportunidades e ajudar o processo de decisão empresarial. Os dados têm de ser relevantes para o processo de decisão (PARENTE, 2003).

O sistema de inteligência de marketing é o conjunto de procedimentos e fontes para se obter informações e envolve pesquisas para se coletar dados, regular e sistematicamente, sobre o que está acontecendo no mercado. O sistema deve ser capaz de monitorar o desenvolvimento da concorrência e acompanhar indicadores de mercado que afetam o comportamento do consumidor.

Os dados são classificados como primários e secundários. Os dados primários são aqueles que não estão disponíveis e precisam ser coletados para uma finalidade específica ou para um projeto específico de pesquisa. Os dados secundários são os existentes, que já foram coletados para outra finalidade e podem ser encontrados em algum lugar (KOTLER e KELLER, 2006). Eles são um ponto de partida já que se encontram disponíveis e geralmente não tem custo ou tem um custo menor que os dados primários. Podem ser utilizados para ajudar a mapear oportunidades.

Há dois tipos de dados secundários: os internos à empresa e os externos. Os dados externos, como é o caso dos fornecidos por pesquisas feitas por órgãos públicos ou entidades particulares como os órgãos setoriais, bem como o último censo, nas pesquisas exploratórias, como é o caso de pesquisas para lançamentos de novos produtos, são muito apropriados e constituem a principal fonte para o sistema de inteligência de marketing (PARENTE, 2003).

Os sistemas de informação geográfica (SIGs) são sistemas informatizados que montam, armazenam, manipulam e exibem dados por sua localização. O resultado é um mapa geográfico revestido com

dados digitais sobre consumidores em uma área particular (BOONE e KURTZ, 2009).

A partir de uma ampla base de dados, através do *data mining*, ou seja, da utilização de técnicas estatísticas e matemáticas, como análise de agrupamento, detecção de interação automática e modelagem, pode-se extrair informações úteis sobre pessoas, tendências e segmentos (KOTLER e KELLER, 2006).

2.2 Mercado habitacional

O mercado habitacional é incomum em termos de complexidade e dimensões. A complexidade deriva da natureza heterogênea do produto, pois cada residência é única, imóvel e tem preço alto (GALSTER, 1996).

Além de produtos que não são homogêneos, no mercado imobiliário são violados outros pressupostos clássicos de um mercado eficiente, pois as negociações envolvem poucos compradores e vendedores e estes não têm pleno conhecimento do mercado. A ineficiência relativa deste mercado tem implicações na forma como ele opera e a consequência é que fatores não internos ao mercado irão afetar o preço de cada propriedade vendida (EVANS, 1995).

2.2.1 Produto - características da habitação

Cada definição de habitação ou residência satisfaz o propósito de uma pesquisa ou de um mercado habitacional particular. A definição mais comumente utilizada deriva do conceito de que esta é uma mercadoria complexa com múltiplas características (LANCASTER, 1966).

Fisher e Fisher (1954) consideram que as duas características principais das unidades imobiliárias, a durabilidade e a imobilidade, têm implicações importantes com respeito ao comportamento do mercado imobiliário. A imobilidade implica que os serviços devem ser consumidos no próprio local. Poucas unidades podem ser adaptadas facilmente para demandas que mudam por diferentes tipos de serviços no local em que elas estão. A durabilidade das residências é de algumas décadas e eventuais alterações são caras, pois envolvem muito trabalho e aplicação de materiais duráveis e, como não costumam gerar retorno, geralmente não são executadas. Assim, o estoque muda lentamente com a adição de novas unidades, exceto para alguns locais particulares.

Devido às duas características mais importantes da habitação, sua imobilidade espacial e sua durabilidade, a dimensão espacial desempenha um papel importante no mercado e a explicação disso constitui a finalidade predominante da teoria habitacional urbana local. Observações em qualquer conjunto de habitações de uma zona urbana revelam que, em cada local existe semelhanças, como o caso de famílias do mesmo grupo sócio-econômico e com rendimentos similares que tendem a se agrupar em áreas bem definidas. Uma explicação de influências espaciais no mercado habitacional urbano é que este é separado em uma série de submercados quase independentes. (BALL e KIRWAN, 1977).

Quigley (1985) especifica que residências podem ser diferenciadas em termos de três dimensões distintas, cada qual refletindo os maiores atributos destas. Primeiro, as residências são espacialmente fixadas, e assim diferenças dos atributos da localização ou da característica da vizinhança fazem com que cada habitação seja única; isto constitui a dimensão espacial ou locacional de uma residência. Duas residências do mesmo tipo, tamanho e idade, mas localizadas em vizinhanças diferentes, terão valores diferentes e atrairão compradores diferentes. Em segundo lugar, as residências são duráveis e assim a dimensão temporal também não pode ser ignorada. Terceiro, as residências podem ser diferenciadas por tamanho, tipo, ano de construção e outras características que satisfazem demandas dos compradores o que constitui a dimensão estrutural ou física das residências.

Para Tu (1997), a residência é definida como um produto com múltiplas características classificadas em espacial ou não espacial; assim, considera que características da vizinhança em relação à aparência física das residências do entorno e as atributos do local constituem a diferenciação espacial e as outras características como o tipo de residência ou ano de construção compõem a dimensão não espacial. Cada característica ampla inclui um grupo de componentes constituintes, sendo que alguns têm contribuição significativa nas diferenças entre residências e contribuem com diferença significativa no preço e outros componentes, apesar de afetar o preço da residência, não têm influência na escolha do consumidor.

As residências podem ser classificadas em grupos de produto em termos de componentes de cada característica em alguma dimensão (MACLENNAN et al., 1987). Os componentes de cada característica das residências podem ser identificados de várias formas, como o tamanho da residência que é medido pela área ou pelo número de

quartos. Várias técnicas de agrupamento estão disponíveis e não há um grupo padronizado ou uma técnica padrão utilizada; depende do propósito do agrupamento.

Em um estudo de Jones (1979), o autor considera que há uma extensa variação na ocorrência e na qualidade dos atributos das residências, mas que há menos variação na extensão dos tipos de serviços fornecidos por esses atributos, e os atributos individuais são categorizados em grandes grupos, com base no tipo de serviço que eles fornecem. Os atributos específicos escolhidos são selecionados com base nas respostas sobre por que as pessoas mudam, e os selecionados e suas utilidades são descritos por categoria, como: quanto à acomodação, considera o tamanho da residência, se a casa é isolada ou geminada, condições sobre o estado de conservação e presença de banheiro interno; quanto aos atributos do local, o ambiente físico percebido como limpo ou não, e o ambiente social em termos de nível de barulho e qualidade das escolas; e quanto à localização, a acessibilidade ao trabalho, escolas, compras, transporte público, amigos.

Schnare e Struyk (1976) consideram duas características que distinguem habitação de outros serviços fornecidos em áreas urbanas, ou seja, a durabilidade do estoque e a espécie dos serviços produzidos por atributos estruturais da unidade residencial e atributos da vizinhança onde a unidade está localizada. Disto resulta que uma oferta de certos tipos de habitação pode ser basicamente fixa por períodos de tempo relativamente longos.

2.2.2 Eficiência do mercado

O conceito de um mercado eficiente, utilizado na microeconomia, deriva do modelo de um mercado perfeitamente competitivo, onde se parte do princípio que há compradores e vendedores suficientes e que cada participante não pode, individualmente, determinar os preços de mercado. Preços refletem opiniões consensuais sobre os valores dos bens ou produtos. Evans (1995) demonstra que os preços dos imóveis não são totalmente determinados pelo mercado e nele, e conclui que o mercado imobiliário deve ser considerado como um mercado ineficiente.

As negociações no mercado habitacional são poucas, isto pela baixa frequência na qual uma propriedade é vendida e existem, relativamente, poucos participantes em qualquer venda. Faltam informações completas e outro sintoma da ineficiência deste mercado, se

relaciona à capitalização da nova informação e ao movimento dos preços no tempo, pois, enquanto o mercado de produtos usa informação diária dos preços, o mercado imobiliário usa informações anuais.

Mas, para Evans (1995), o problema crucial é a heterogeneidade das residências, pois poucas propriedades são tão semelhantes que possam ser consideradas como substitutas exatas. O preço não é como o de um simples produto, já que cada característica incorporada em uma propriedade possui uma série de preços, sendo cada preço unitário de cada característica variável de acordo com as diferentes localizações. As condições podem ser redefinidas para cada segmento de mercado, onde os produtos são homogêneos, mas, um mercado eficiente requer também que se tenha um grande número de compradores e de vendedores além de ampla informação, e os segmentos não são grandes o suficiente para que isto ocorra.

2.2.3 Informações sobre o imóvel

Apesar da natureza multidimensional da oferta de residências, da natureza indeterminada dos preços de mercado e da dificuldade para decodificar preços de imóveis, a suposição adotada em estudos empíricos e teóricos para análise do mercado habitacional de que este mercado é eficiente, implica em participantes deste mercado perfeitamente informados e direcionados (EVANS, 1995).

Watkins (1996) pesquisou o nível de informação para consumidores do mercado habitacional ao comparar novos entrantes - compradores da primeira residência ou vindos de outro local - com os que já participavam previamente do mercado, ao verificar se os preços pagos por uma hipotética residência padrão eram significativamente diferentes. O autor verificou que todos os compradores, sendo as compras repetidas ou não, entram no mercado com informação imperfeita, já que a informação disponível para estes é imprecisa pela heterogeneidade do mercado habitacional, e obsoleta por causa da natureza dinâmica do mercado. De acordo com o mesmo autor, as famílias encontram considerável dificuldade para coletar informações atualizadas sobre preços e vacância de residências.

De acordo com Tu e MacLennan (1996, apud Watkins, 1996), em média, as famílias no Reino Unido compram uma nova residência a cada 8,5 anos, e o envolvimento é pouco freqüente quando comparado com a compra de outros bens dado o tempo que decorre entre transações e as informações sobre vacâncias ficam obsoletas um tanto rapidamente. A

natureza dinâmica do mercado habitacional faz com que e características locais e da vizinhança, que têm considerável influência no desejo dos consumidores para escolha das residências, se alterem através do tempo.

Para Watkins (1996), todas as famílias devem entrar em um processo de escolha potencialmente prolongado, caro e que consome tempo para obter informações relevantes. O mesmo autor considera que, pela complexa estrutura do mercado habitacional caracterizada por submercados espaciais e estruturais, as famílias focam sua procura e obtenção de informação em segmentos do mercado e que, pode ser que as informações dos corretores de imóveis e o diferente grau com que estes se especializam em algum segmentos, faça com que haja desvantagem de informação para compradores de algum tipo de imóvel ou particular localização.

Além disto, é improvável que a família precise uma residência que seja um substituto próximo da que ela reside, de mesmo tamanho ou similar, com mesmo estilo ou idade que a residência existente, já que a maioria se muda por razões ligadas ao ciclo de vida, estilo de vida ou transferência de trabalho (WATKINS, 1996).

A teoria da demanda de Lancaster diz que, do ponto de vista das pessoas, os bens que elas compram é um conjunto de características, cada um com seu valor, e a decisão de comprar um “conjunto” ao invés de outros, não depende apenas do desejo por características constituintes da residência, mas da habilidade relativa de cada um em avaliar os imóveis em oferta.

2.2.4 Estabilidade do mercado – mudança dos preços (“*filtering*”)

Fisher e Fisher (1954) desenvolveram modelos de “*filtering*”, um conceito amplamente utilizado em estudos do mercado habitacional, que considera a mudança de preços da propriedade ou de valor do aluguel através do tempo para uma certa unidade imobiliária de determinada área. A unidade pode ser uma residência ou um grupo de residências, a metragem quadrada de escritório ou comercial, ou qualquer utilidade em estudo. A área estudada geralmente se estende pela vizinhança ou bairro, cidade ou a área metropolitana. Supõe-se que certas unidades ocupam uma posição, expressa por seus preços ou valores de aluguel, com uma distribuição de valor similar a outras unidades dentro de um local em um certo momento; se após decorrer um período de tempo, a posição relativa das mesmas unidades tiver mudado, ocorreu *filtering*.

O fenômeno do *filtering* acontece geralmente reduzindo o valor do imóvel, com as unidades ocupando uma posição inferior devido à taxa de deterioração dos imóveis ou à mudança nos desejos dos usuários, mas, mudanças externas em relação ao transporte ou infraestrutura do entorno podem tornar uma região mais atrativa. No caso da habitação, o surgimento de novas construções, particularmente se apresentarem inovações em termos de design ou lay-out, pode acelerar o processo de *filtering* do estoque existente (FISHER e FISHER, 1954).

De acordo com Galster (1996), o fenômeno central na determinação dos submercados é a dinâmica da mudança dos preços e a movimentação associada à qualidade das moradias. Os autores elaboram um modelo com uma estrutura fechada dentro da qual os componentes do *filtering* – que são as mudanças do valor das residências, os grupos de bairros com residências de qualidade semelhante e possíveis mudanças desta qualidade - podem ser compreendidos e correlacionados. Esta abordagem considera que estes fatores possuem uma relação direta com as mudanças na avaliação do mercado, onde grupos de residências e a mudança da qualidade destas são livremente associados com *filtering*. Neste modelo, o equilíbrio da avaliação dos submercados é estabelecido através da interação das condições de oferta e demanda no médio prazo e possíveis mudanças quando algum determinante de uma dessas funções é alterado.

Quando a avaliação do valor de mercado de uma determinada região tem queda, em resposta, por exemplo, a novas construções, algum local que previamente ocupava uma posição inferior no submercado pode se movimentar para cima em qualidade. Alguns locais que originalmente poderiam ter sido vistos como uma ótima escolha em submercados de qualidade superior podem, então, perder seu valor. Além disto, as mudanças com o tempo da oferta e da demanda em um submercado, levam a uma complexidade devido a um sucessivo e desigual grau de incrementos do nível de qualidade dos submercados, afetando também os submercados próximos. O ajustamento complexo interno ao submercado irá mostrar sempre simultaneamente a elevação e a queda no valor de mercado, na qualidade da residência e na qualidade da vizinhança (ROTHENBERG et al., 1991).

2.2.5 Equilíbrio do mercado

Estudos sobre o comportamento do consumidor buscam entender o equilíbrio entre oferta e demanda do mercado habitacional. Watkins (1998) verifica que submercados são gerados por processos complexos

entre oferta e demanda, que refletem influências tanto estruturais como espaciais.

Whitehead e Odling-Smee (1975) argumentam que o conceito convencional de equilíbrio de mercado é particularmente inapropriado no mercado habitacional urbano, porque fatores como custo de transação e custo de informação são importantes para este mercado. O objetivo fundamental da pesquisa destes autores é o de desenvolver modelos microeconômicos de comportamento, que combinados revelam as relações do mercado. Para estes autores, juntar micro efeitos à análise macro com a introdução de fatores não coletivos como distribuição da renda, idade da residência ou distribuição da localização é complexo, mas pode ser útil ao se trabalhar com o mercado econômico da habitação.

Rothenberg et al. (1991) demonstram que submercados são estáveis em uma série de cidades dos Estados Unidos, quando estudam a dinâmica da mudança dos preços e a movimentação associada à qualidade das moradias.

Mas, segundo Bajic (1985), em geral, os preços implícitos dos atributos podem ou não refletir um equilíbrio de curto prazo do mercado. Rápidas mudanças nos preços das residências mostram que o mercado não está em equilíbrio de curto prazo, nem próximo, já que existe a possibilidade de se encontrar grandes diferenças nos preços individuais dos atributos.

Tu (1997), no estudo das relações entre dois submercados habitacionais, verifica que no longo prazo, teoricamente existem cinco fatores importantes que podem influenciar preços nos submercados, que são:

- Novos proprietários ocupantes das residências em estoque na área. Como as novas residências têm preços mais altos que as mais antigas, novos estoques deslocam para cima o preço da residência no submercado. Novos estoques de residências também melhoram o ambiente físico da vizinhança, a qual pode atrair mais famílias para a região. Isto irá também aumentar o preço;
- Renovação das residências. A reforma e melhoria das residências antigas melhoram a qualidade destas, o que valoriza o ambiente local, e assim, provoca aumento dos preços das residências;
- A elevação e queda da economia. Não apenas a taxa de emprego e o nível de renda local, que podem atrair as famílias

imigrantes, mas também mudanças nos atributos do local afetam os preços das residências;

- O desenvolvimento de sistema de transporte local. Isto altera a acessibilidade ao centro de negócio ou centro da cidade, o que irá influenciar as preferências das famílias por residências naquela área, e, portanto, afetam o preço;
- Mudança na estrutura demográfica local e na preferência dos compradores. O autor comenta que houve aumento do número de pessoas solteiras em Glasgow e que isto aumentou a demanda por residências pequenas, o que provocou um aumento no preço deste setor.

Entretanto, não se tem evidências empíricas que, ao nível de submercados, como e em que extensão esses fatores influenciam no preço das residências.

2.3 Estudos sobre o mercado habitacional - natureza dos submercados

As abordagens teóricas para explicar a segmentação do mercado podem ser classificadas em teoria econômica neoclássica ou comportamental-cultural ou ainda de uma forma intermediária entre esse dois extremos. Na teoria econômica neoclássica, uma área urbana pode ser segmentada se as preferências e/ou a renda das famílias for diferente com relação ao espaço e acessibilidade, se não existirem outros fatores substanciais em relação ao valor das residências. Além disto, a abordagem sobre o uso da terra e preferências por determinados ambientes podem ser vistos como uma explicação para a ocorrência de segmentação, e a segmentação dentro das áreas urbanas pode ser baseada em mais fatores como o tipo de construção ou atributos que dominam as residências.

Os modelos culturais e comportamentais têm origem em fatores sócio-culturais e enfatizam a diversidade de composição das famílias. De acordo com a perspectiva construcionista a segregação é construída socialmente, assim ela não existe, mas deve ser elaborada. As teorias sobre o modo de vida e prática de segregação são resultados de escolhas sócio-culturais dos consumidores. O que elas têm em comum é que surgiram como uma desaprovação da racionalidade do modelo neoclássico e começaram a ser difundidas a partir dos anos 1980.

A teoria do desequilíbrio local, diz que diferentes consumidores têm preferências diferentes em relação à moradia e face a grande diversidade de alternativas de residências não se pode considerar o mercado como homogêneo. Este desequilíbrio é ocasionado por fatores espaciais e setoriais e pela diversificação dos dois lados: do estoque e da demanda.

2.3.1 Primeiros Estudos – conceito de substituição

A estrutura conceitual para análise do mercado habitacional foi desenvolvida por economistas dos Estados Unidos nos anos 1950 e 60, e tem início com o estudo de Rapkin et al. (1953, apud FISHER e FISHER, 1954) que postularam que, em um amplo sentido, cada unidade residencial dentro de um mercado habitacional pode ser considerada uma substituta de qualquer outra unidade, mas que esta visão só pode ser mantida para uma análise geral, pois as unidades residenciais são diferenciadas pela localização, tamanho, layout, ano de construção, preço e outras características, e nenhuma residência é uma substituta perfeita de outra.

Substitutos, de acordo com Bourassa et al. (2007), são pares de bens nos quais um aumento no preço de um deles, leva a um aumento na demanda do outro. No caso da habitação que é um produto heterogêneo, considera-se que os pares de bens com características similares são como substitutos próximos.

Para Fischer e Fisher (1954), existe um grau de substituição muito próximo para certos tipos ou grupos de unidades residenciais, quando competem entre si como alternativa para os compradores do mercado habitacional. Consideram que, no mercado imobiliário duas unidades estão no mesmo mercado se são substitutas, em algum grau, entre si. Assim, apesar de nenhuma unidade ser uma perfeita substituta de outra, quando o grau de substituição é próximo, podem ser incluídas no mesmo mercado, e, se cada unidade dentro de um mercado local pode ser considerada como substituta de cada outra, pode-se dizer que todas as habitações formam um único mercado. Entretanto, esses autores também consideram que esse ponto de vista pode ser mantido apenas em uma análise geral e com certa dificuldade, pois, salvo por uma grande subdivisão homogênea, as unidades habitacionais não são mercadorias padronizadas.

Grigsby (1963, apud JONES et al., 2001) introduz o conceito de submercados habitacionais em termos de substituição, ao considerar que um submercado é compreendido por um grupo de residências que

representam substitutos relativamente próximos, mas não substituíveis por outro grupo de residências, e não necessariamente localizadas na mesma região. Galster (1996) cita que o mesmo autor analisa não apenas o comportamento das ligações entre várias características da residência, mas também a dinâmica da mudança do preço e da qualidade destas para elucidar o fenômeno do *filtering*.

Para Rothenberg et al. (1991), também, no conceito central de submercados o bem habitação tem substituição relativamente próxima. Considera que o padrão de substituição está ligado ao preço, custo pretendido, desejável proximidade a utilidades, trabalho, e amigos, qualidade ambiental, segregação étnica, e demanda por atributos individuais das residências, como o número de quartos e o tamanho do terreno.

Watkins (1998) cita que substituição requer consumidores indiferentes entre as características da habitação - físicas, atributos do local e da vizinhança - mas, que nestes primeiros estudos, as características físicas, também citadas na literatura como características estruturais, foram consideradas mais importantes para a formação de submercados do que as demais.

Estes primeiros estudos introduziram o conceito de submercado em termos de substituição, considerando que duas unidades habitacionais estão em um mesmo submercado se competem entre si como alternativa para os consumidores deste mercado.

2.3.2 Modelos de equilíbrio urbano

Com a contribuição de Alonso (1965), estudos empíricos e teóricos do mercado habitacional passam a utilizar modelos de equilíbrio da nova economia urbana com pesquisas desenvolvidas a partir da observação de que a moradia e a acessibilidade ao emprego eram itens ligados.

O autor acima desenvolve a teoria das trocas (*trade-off*), que tem como ponto de partida a relação entre acessibilidade à residência e preços. Utiliza um modelo simples de equilíbrio espacial, no qual considera que os consumidores, ao se mudarem para locais mais distantes de seus centros de trabalho aumentam seus custos de locomoção, que, em contrapartida, são compensados pelo menor custo dos terrenos.

Alonso (1965) considera, em seu estudo, que a cidade é construída com um plano homogêneo, sem traços característicos, que o

sistema de transporte é também homogêneo através de toda a cidade e que existe apenas um centro, que é o centro de negócios, onde todos os empregos estão concentrados. Conclui que, para um padrão de equilíbrio locacional, é necessário que a renda seja maior com o aumento da distância ao trabalho.

Wheaton (1977) revisa e atualiza o estudo de Alonso (1965), para buscar o determinante do padrão americano renda-localização, pois nas áreas urbanas das cidades americanas, as famílias de renda média e alta vivem em subúrbios mais longe do centro, e a mobilidade deles para fora deixa as cidades segregadas ao domínio da pobreza, onde os mais pobres pagam por serviços que se encontram deteriorados.

Os resultados deste estudo de Wheaton sugerem que, apesar de um aumento da renda tender a gerar a locação nos subúrbios, no longo prazo e com a competitividade do mercado, as diferenças dos preços de terrenos da oferta acabam sendo pequenas e estatisticamente insignificantes para justificar esse padrão, assim, a suburbanização das classes média e alta nas cidades americanas é uma resposta do mercado a exterioridades e incentivos fiscais da fragmentação municipal.

Ball e Kirwan (1977), que também assumem que as características espaciais são mais importantes que as estruturais na formação dos submercados, citam que pode surgir segmentação do mercado se houver restrições, para compradores ou vendedores, com referência a determinadas áreas geográficas; restrições para os compradores, neste sentido, significa que estes não podem ocupar determinadas regiões ou também a falta de recursos ou de financiamento para a compra do imóvel. No estudo destes autores sobre a abordagem das trocas nas transações e a emergência de submercados, eles consideram questionável colocar demasiada ênfase na acessibilidade para explicar a distribuição espacial das habitações, pois, apareceu clara evidência da importância tanto de aspectos históricos como de outros derivados de políticas de planejamento sobre as características da distribuição espacial da oferta de habitação.

A partir da década de 1980, o modelo de Alonso passou a ser questionado por suas premissas, ou seja, a hipótese de um mecanismo de equilíbrio estático que constitui a base da teoria, e também pela suposição de que um mercado habitacional urbano pode ser visto como um todo, no qual um conjunto único de processos gera determinantes comuns a todas as habitações do mercado.

2.3.3 Heterogeneidade do produto – demanda por atributos

Lancaster (1966), com foco na demanda, desenvolve uma linha da teoria microeconômica na qual a utilidade é gerada não pelo bem em si, mas pelas características dos bens. A essência da nova abordagem é resumida em três itens, cada um representando uma quebra com a teoria tradicional: o bem, em si, não traz utilidade ao consumidor, ele possui características e essas características promovem a utilidade; em geral, um bem possui mais de uma característica e várias características são compartilhadas por mais de um bem; bens em combinação podem possuir características diferentes das pertencentes aos bens separadamente. A aplicabilidade na habitação é direta devido à heterogeneidade do produto.

Para Maclennan et al. (1987), a partir da definição de Lancaster, pode-se concluir que o estoque habitacional é dividido em grupos distintos de produtos. A decisão de comprar uma residência ao invés de outras depende não só do desejo pelas características constituintes destas, mas da condição relativa de cada um e do estoque.

De acordo com Malpezzi (2002), a residência é, neste sentido, entendida como o “lar”, e as características devem gerar conforto e bem estar. Assim, os atributos do produto relativos à localização, ambiente da vizinhança e estrutura da residência, em conjunto ou isoladamente, conferem ao produto uma combinação de características, que são entendidas objetivamente ou subjetivamente pela demanda.

Jones (1979), que segue Lancaster (1966), cita que a habitação é como um vetor de características e atributos e que, quando as famílias se mudam, elas buscam satisfazer seus desejos e necessidades através dos benefícios derivados dos vários atributos da nova residência.

A teoria de Lancaster (1966) foi criticada pela ausência de um embasamento teórico, porém, mais tarde, foi formalizada por Rosen (1974), que com enfoque nas características dos produtos, enfatiza a forma como ocorre a interação dos consumidores com uma estrutura de características da oferta. Em seu modelo de pura competição entre bens heterogêneos, Rosen sugere que existem mercados implícitos para cada atributo da residência, e que o preço de equilíbrio de longo-prazo é estabelecido por cada atributo.

Com os trabalhos de Lancaster e Rosen tem início uma linha de estudos econômicos através de equações de regressão do preço – equações hedônicas - que pressupõe que o preço das moradias é determinado pelo valor de cada característica, física e espacial da unidade habitacional. A aceitação econômica da relação entre preços

observados e características dos produtos torna evidente que preços diferentes entre os bens são reconhecidos como uma diferente equalização para um conjunto de alternativas que eles contém.

De acordo com Pasha e Butt (1996), Rosen (1974) não propõe nenhum método particular para estimação dos preços, mas delinea uma estratégia de modelagem, deixando a escolha do procedimento de estimação para ser determinado conforme a aplicação em cada caso particular.

2.3.4 Características da oferta e da demanda

Para Straszheim (1975, apud GOODMAN e THIBODEAU, 2003), o preço das residências pode variar consideravelmente dentro da área metropolitana devido à heterogeneidade das residências do estoque, e assim, a região metropolitana deve ser segmentada em submercados menores devido a fatores relacionados às características da oferta de residências. De acordo o mesmo autor (1975, apud MACLENNAN, 1977) certas características individuais dos consumidores, como idade, sexo ou ocupação, ou ainda a existência de discriminação racial, fatores que podem influenciar a parte financeira de acesso à habitação, também conduzem a uma segmentação complexa do mercado.

Outros autores como Munro (1986), Schnare e Struyk (1976) e Palm (1978) também entendem que ocorre segmentação do mercado habitacional por certas características das famílias, como o nível de renda, e consideram que os compradores competem por uma parte particular das residências do estoque.

Schnare e Struyk (1976) argumentam que a segmentação do mercado habitacional ocorre quando a demanda das famílias por uma determinada estrutura física da residência ou relacionada à vizinhança é extremamente inflexível, ou seja, altamente inelástica e a preferência é compartilhada por um número de famílias relativamente grande. Consideram que, em alguns casos, demanda altamente inelástica para certos bairros ou tipos de grupos, junto com ofertas inelásticas, podem segmentar o mercado em um número de setores independentes, produzindo grandes diferenciais nos preços dos atributos individuais das residências e conseqüentemente no preço total por unidade de serviço da habitação. Os autores consideram que, com a pronunciada segmentação do mercado, a análise tradicional, em particular dos preços, que considera o mercado habitacional único, se torna inapropriada.

A linha de pesquisa de Schnare e Struyk (1976), que consideram que submercados têm origem justamente pela competição insuficiente

no mercado local, segue um ponto de vista que significa rejeição da situação ideal assumida pela teoria neoclássica, que expressa a noção de que o mercado se ajusta para remover o desequilíbrio entre oferta e demanda, e que esse ajuste ocorre rapidamente. Nos testes, sem preço, a segmentação é observada empiricamente com base em seis fatores: própria ou alugada; tipo de residência, número de quartos; e características qualitativas, recurso do financiamento, idade do estoque e localização.

Para Bajic (1985), as principais razões para a segmentação do mercado são o alto custo do produto e a fixação locacional das unidades residenciais. A interação destes dois fatores do bem habitação com a alta inelasticidade da demanda por certas características estruturais e da vizinhança pode criar vários submercados independentes. Em seu estudo, o autor inclui informações socioeconômicas das famílias, como a idade, escolaridade, renda normal e número de crianças e adultos nas residências, além de dados relativos ao sistema de transporte da região e características físicas das residências.

Goodman e Kawai (1982) examinam a demanda por habitação por uma ótica específica, com as variáveis, renda e preço, baseadas em dados individuais das famílias. Estes autores dividem a renda em permanente e transitória, sendo que na transitória consideram fatores humanos e outros, que denominam não humanos, de acumulação. Constatam uma redução do efeito da renda permanente na demanda por habitação quando as pessoas já possuem bens anteriormente, especialmente quando a família já tem uma residência. Assim, concluem que a separação, para cada classe de renda, em componentes permanentes e transitórios melhora substancialmente o poder de prever a estimação da demanda por habitação. Concluem também, que a demanda não é função apenas da renda e do preço da residência, mas também de diversas características das famílias, como o número de membros, presença ou não de crianças em escola, raça e ainda de outras variáveis que são relativas ao bem estar.

Quigley (1985) apresenta uma análise empírica da escolha habitacional baseada nas famílias individuais e nas características das residências. O autor considera a escolha das famílias em três situações: a escolha da residência dentro de uma região de uma determinada cidade, a escolha da vizinhança dentro de uma cidade ou a escolha marginal de serviços e atributos locais de qualquer cidade. São incluídas medidas de qualidade das residências, referentes ao tamanho, como número de banheiros e de quartos por pessoa, e também o valor que a família gasta com outros bens que não a habitação, o que é considerado através do

cálculo da renda mensal menos o custo com pagamento de aluguel. Os resultados sugerem que a escolha da habitação pode ser mais sensível à variação da acessibilidade ao trabalho do que é indicado em modelos mais restritivos de escolha das famílias.

Harsman e Quigley (1995) comparam o nível de segregação por raça e por tipo de família em duas cidades distintas, mas que possuem mesmo nível de renda e mesma composição demográfica, e pequena variação na provisão de serviços públicos. Consideram que a classificação da população por tipo de família é correta (os tipos incluem famílias formadas por casais com ou sem filhos, adultos solteiros morando sozinhos e por sexo, famílias com parentes solteiros e por sexo, e não familiares morando juntos sendo dois ou mais adultos). Como resultado verificam que a segregação por classe de renda é levemente menos pronunciada que a por tipo de família e que, por alguma razão, as forças para que haja segregação por tipo de família são maiores em uma cidade do que na outra.

Goodman (1981) nota que no processo de procura para compra ou locação são examinadas as residências de uma determinada área, de acordo com o preço, segregação racial ou proximidade desejável a amigos ou ao local de trabalho, e assim, características do estoque também podem introduzir uma segmentação.

De acordo com Watkins (1998), a diferenciação do produto ou a existência de grupos ou subgrupos da demanda não são suficientes para que exista um submercado mas, o submercado irá existir somente onde a interação entre uma oferta segmentada e uma demanda diferenciada conduzem a variações significantes no preço depois da padronização por complexidade do produto. Para o autor, ao se caracterizar sempre a estrutura dos submercados de acordo com as variações da oferta, o lado da demanda será ignorado, ou seja, falha-se em reconhecer que diferentes tipos de compradores provavelmente saíram de diferentes classes de características que dependem de suas necessidades, preferências e recursos.

Segmentação, de acordo com Bourassa et al. (1997), significa diferentes tipos de bens sendo vendidos mercados totalmente distintos, consumidores e produtores com preferências diferentes e preços diversificados, e, também para Kauko (2004), a segmentação é identificável com base no estoque, na demanda e nos preços.

2.3.5 Dinâmica do mercado

Jones et al. (1998) ressaltam a necessidade de que se entenda melhor a estrutura do mercado habitacional, como ele opera e sua dinâmica, com o propósito de estabelecer se existem preços significativamente diferentes para residências padronizadas nos diferentes submercados e verificar se existe dinâmica destes, já que a teoria sugere que mudanças no padrão relativo de preço entre submercados podem revelar que partes do mercado estão sujeitas a uma pressão maior da demanda. Consideram que a mudança das famílias entre os submercados ocorre principalmente porque elas desejam uma residência de um padrão superior ou inferior àquela onde residem e não conseguem encontrar o imóvel no mesmo submercado.

Os autores acima verificam que, no período do estudo, o aumento de 46% no preço das residências no submercado do oeste da cidade de Glasgow, na Escócia, comparado com apenas 8% no sudoeste pode refletir excesso de demanda para o oeste. O resultado tem várias implicações, pois a evidência suporta um esquema de definição e identificação raramente utilizado, que sugere que os submercados são gerados por um processo complexo em que a dinâmica da oferta e da demanda estão ligadas ao processo de escolha da residência e no desenho urbano. Visto que a limitação do estoque é determinada, ao menos em parte, por políticas de planejamento público na cidade onde se deu o estudo, consideram, ainda, que os profissionais ligados ao planejamento urbano poderiam monitorar o mercado habitacional.

Os mesmos autores consideram que os testes estatísticos que costumam ser adotados tradicionalmente para análise do mercado habitacional se mostram incompletos e insuficientes para a determinação de submercados e assim, entendem que os padrões de mobilidade das famílias devem ser incluídos nas análises. Em outro estudo Jones et al. (2004), examinam novamente a área urbana de Glasgow e verificam que é possível demonstrar que os submercados habitacionais tendem a ser independentes entre si e que a inclusão da mobilidade das famílias forneceu uma justificativa adicional para a existência de submercados.

Outro fator relacionado à dinâmica do mercado foi verificado por Gabriel (1984) no estudo do desenvolvimento das cidades de Israel, visto que, mesmo com a tentativa de integrar os imigrantes de regiões distintas como África, Ásia e Europa, padrões de desenvolvimento indicam que ocorreu a estratificação dos bairros de acordo com a origem da população, data de imigração e condições locais de habitação.

O mesmo autor cita que as áreas habitacionais metropolitanas não são frequentemente caracterizadas por um grau de substituição com convergência dos preços dos atributos para uma constante, como em um modelo de equilíbrio. Com respeito a isso, a demanda das famílias por uma vizinhança particular ou por características estruturais da habitação pode ser relativamente inelástica. Em conjunto com restrições na oferta por atributos da habitação, emerge a segmentação do mercado.

Gabriel segue estudo de Schnare e Struyk, porém a conclusão é diferente, ou seja, de que o mercado é segmentado, mas no caso do desenvolvimento de Israel, a formação e a definição das vizinhanças e a divisão em regiões refletem características sociais e econômicas de grupos de imigrantes. O resultado obtido enfatiza a importância da segmentação do mercado habitacional em um contexto substancialmente diferente dos estudos anteriores.

Galster (1997) utiliza a estrutura dos submercados para analisar a dinâmica do mercado, com modelos que fornecem a disponibilidade de habitações pode-se comparar a efetividade de políticas de iniciativas do lado da demanda versus iniciativas da construção, do lado da oferta, e no centro desta estrutura está a idéia de que o mercado se ajusta nos dois caminhos. Verifica que há um espaço para um ajustamento independente da oferta e da demanda interiormente ao submercado e que existe uma substituição entre submercados quando as famílias mudam de residência. O autor argumenta que políticas direcionadas a submercados particulares podem iniciar uma cadeia de ajustamento tanto internamente como entre estes submercados.

3 - ECONOMETRIA ESPACIAL

Anselin (1988) define econometria espacial como um conjunto de técnicas que tratam das peculiaridades causadas pelo posicionamento espacial - os efeitos espaciais - e promovem uma maneira de se encontrar especificação apropriada na análise estatística de modelos regionais. Os efeitos espaciais são de dois tipos, a saber, autocorrelação espacial e heterogeneidade espacial.

A econometria espacial é relevante no estudo do mercado habitacional porque leva em conta explicitamente, nos modelos, a influência da localização espacial. A análise de regressões é a técnica mais freqüentemente utilizada desde os anos 1970, mas a influência dos efeitos espaciais só foi adicionada aos estudos nos anos 1990 e, de acordo com Anselin (2002), têm recebido mais atenção nos últimos anos.

Na modelagem hedônica, uma técnica matemática de regressão, formalizada por Rosen (1974) para aplicação na análise do mercado habitacional, considera-se que uma residência é formada por um conjunto de atributos, e cada um contribui para a formação do preço da propriedade. Os atributos, individualmente, não têm preços explícitos, mas preços implícitos atribuídos pelo mercado, que podem ser considerados como sendo os preços hedônicos dos atributos. Considera-se que o preço explícito de uma propriedade particular é a soma dos preços hedônicos dos vários atributos contidos no conjunto.

As regressões envolvem três tipos de atributos: estruturais, relativos às características físicas da residência; espaciais, de dois tipos, um relativo à localização ou acessibilidade, e outro relativo à vizinhança, ou seja, às características das propriedades vizinhas. A inclusão de variáveis representando vários atributos pode ocasionar multicolinearidade, autocorrelação espacial e heterocedasticidade. Esses modelos podem ser utilizados para identificação da influência de certos fatores no preço das propriedades, porém, de acordo com Anselin (1988), modelos hedônicos de preços de residências que não levam em consideração a autocorrelação espacial podem conter pouca especificação e levar a conclusões erradas. Estudos hedônicos que consideram aspectos espaciais estão no campo da econometria espacial que é um subcampo da estatística espacial.

Os modelos espaciais são especificados por regressões lineares com interdependência espacial, o que inclui a observação da relação entre cada unidade com aquelas que são suas vizinhas. Isto se baseia na primeira lei da geografia enunciada por Tobler em 1970, que diz que tudo tem relação com tudo mais, porém, o que está próximo está mais relacionado. Anselin e Bera (1998) citam que a dependência espacial é a regra e não a exceção.

O desenvolvimento da estatística espacial ocorreu pela necessidade de quantificação da dependência espacial presente em um conjunto de geodados, que são dados que estão vinculados à sua localização geográfica. Segundo Anselin (1992), a característica da análise estatística de dados espaciais é ter seu foco em identificar padrões espaciais de lugares e valores, verificar associação espacial entre eles e a variação sistemática do fenômeno por localização. Ele ainda ressalta que estes padrões espaciais, conhecidos como efeitos espaciais, afetam a validade dos métodos estatísticos tradicionais.

No modelo clássico de regressão linear o pressuposto é que os erros devem ser homocedásticos e não autocorrelacionados. Assim, o modelo de regressão linear tradicional tem a limitação de não ter a capacidade de controlar efeitos espaciais, já que o pressuposto seria violado (Almeida e Haddad, 2004).

O padrão da interação no espaço econômico tem uma natureza multidirecional o que produz efeitos espaciais: a heterogeneidade espacial e a autocorrelação espacial. A heterogeneidade diz respeito a aspectos da estrutura socioeconômica no espaço, ocasionando diferentes respostas dependendo do lugar. Na autocorrelação espacial o valor de uma variável de interesse em uma certa localidade, depende do valor dessa variável nas localidades vizinhas.

Em um estudo para residências alugadas, com dois dormitórios, na cidade de São Paulo, Hermann e Haddad (2005) consideram que um imóvel seja um bem heterogêneo composto de várias características físicas da residência e atributos de sua localidade, ou seja, distância e acessibilidade ao centro urbano, além de características naturais como áreas verdes e outras, que consideram amenidades urbanas, tais como trânsito, poluição e segurança. O resultado obtido pelos autores aponta maior importância para os itens ligados à localização distância e acessibilidade ao centro no valor do aluguel.

Além dos estudos sobre o mercado habitacional, a econometria espacial tem sido aplicada em outras áreas do conhecimento, com estudos ligados à agricultura, onde aspectos físicos do local com o tipo de solo e clima, que podem facilitar o cultivo de certos produtos, e o

vento que pode ser responsável pelo espalhamento de sementes nas localizadas contíguas, evidenciam a ocorrência de autocorrelação espacial. Estudos sobre criminalidade, pobreza, mortalidade infantil e difusão de tecnologia também têm utilizado a econometria espacial para explicar a similaridade de fenômenos em locais próximos. Ainda alguns estudos ligados à proliferação de doenças, que utilizam ao invés da distância, a conectividade entre pontos, podem ilustrar a amplitude de aplicação da análise exploratória de dados espaciais (AEDE).

3.1 Autocorrelação espacial

Na estatística clássica, o conceito de correlação diz respeito à relação de duas variáveis, ou seja, mede a concentração de uma variável com relação à concentração da outra, de modo que, quando as concentrações das duas aumentam simultaneamente, a correlação é positiva, e quando as duas variações nas concentrações têm sentidos contrários é negativa. Na estatística espacial, os padrões de associação são diferentes da convencional, pois a correlação se refere ao conceito de vizinhança e mensura a semelhança entre o valor de uma variável em uma área e o valor da mesma variável em áreas consideradas vizinhas.

Autocorrelação espacial, ou dependência espacial, significa que o valor de uma variável, como o preço de um imóvel, em uma certa localidade depende do preço dos imóveis nas localidades vizinhas. A autocorrelação pode estar associada, à baixa especificação dos modelos estatísticos, pela ausência de variáveis importantes ou ao uso de funções inadequadas que não capturam os efeitos das características no preço.

No mercado habitacional, a autocorrelação espacial ocorre especialmente por causa das características das residências. Pode-se verificar que residências próximas tendem a ter semelhanças físicas, como a área ou a idade da construção, além de atributos do local em que se situam também semelhantes, e assim, as famílias que residem em uma mesma vizinhança utilizam as mesmas estruturas urbanas. De acordo com Gillen et al. (2001), estas famílias seguem padrão de trocas parecido, sugerindo também que os padrões de acessibilidade são similares.

Na análise de dependência espacial, busca-se verificar a associação que depende, especialmente, da distância, mas que pode ser representada também pela conectividade, como o tempo ou custo de

deslocamento entre os pontos, ou seja, o termo “vizinho” está baseado no padrão espacial adotado.

Na maioria dos estudos do mercado habitacional sobre a autocorrelação espacial, assume-se que esta é isotrópica, isto é, que é função apenas da distância. Gillen et al. (2001) verificam que a autocorrelação dos preços das residências, bem como a dos resíduos de equação hedônica de preços é anisotrópica, pois sofre influência não só da distância, mas também da direção que as separa, e que a autocorrelação é maior em direção ao centro da cidade ou centro de negócios e em direção a grandes vias de transporte. Porém, de acordo com Tu (2007), se isto for considerado, a análise torna-se complexa, pois demanda que se conheça como a direção está condicionada, e ao se assumir que o processo é isotrópico e estacionário, o significado da média e variância de cada resíduo distribuído é constante para todos os locais e a correlação entre os resíduos é função apenas da distância que separa as propriedades.

É importante que se detecte a ocorrência de autocorrelação espacial nos modelos, pois isto pode implicar que o modelo de regressão especificado tenha estrutura autoregressiva, com dependência não linear entre as variáveis dependentes e independentes ou que uma ou mais variáveis regressoras importantes estejam faltando. De acordo com Anselin e Bera (1998), a presença de dependência espacial em dados georreferenciados deve ser utilizada ao se procurar uma melhor adequação de modelos, mas também quando o objetivo é a interpretação de como a forma de interação espacial está relacionada ao processo econômico e social que existe por trás.

É possível se destacar duas fontes de dependência espacial: em primeiro lugar, erros de medida para observações em lugares contíguos podem causar dependência no espaço; em segundo lugar, a existência de uma variedade de modelos de interação social introduz dependência entre os agentes econômicos num sistema, como em processos de difusão de tecnologia ou de imitação (ANSELIN, 1988).

A autocorrelação do tipo erro espacial aparece quando existe correlação espacial no erro das variáveis explicativas ou variáveis omitidas, como é o caso de variáveis ligadas a poluição ou barulho, que estão associadas a todas observações de uma vizinhança, mas são difíceis de se medir ou de se saber o quanto podem influenciar nos preços das residências. Este tipo de autocorrelação também pode ser ocasionado por erro de medida das variáveis, que podem estar vinculadas a divisões artificiais da unidade geográfica que não

coincidem com a verdadeira dimensão do fenômeno observado, ou ainda, a pouca especificação da função utilizada.

A autocorrelação espacial positiva indica similaridade, pois ocorre quando valores altos - ou baixos - de uma variável tendem a se agrupar no espaço, enquanto a negativa indica que, em determinado local, os vizinhos apresentam valores diferentes. Verifica-se uma autocorrelação espacial positiva, quando valores de certa variável - como a variável resíduo do preço de modelos hedônicos - tendem a se agrupar aos similares das observações adjacentes, formando *clusters*.

3.2 Heterogeneidade espacial

Além da dependência espacial, verifica-se um segundo tipo de efeito espacial, a heterogeneidade, que se refere à unicidade de cada lugar, ou seja, à diferenciação espacial ou regional. Cada localização é possuidora de características únicas, assim, a heterogeneidade está associada a uma ausência de estabilidade do comportamento das variáveis sobre o espaço, e os parâmetros que caracterizam as unidades variam com a localização.

Heterogeneidade espacial se refere à variação na relação através do espaço ou a uma variação sistemática de comportamento de um certo processo no espaço. Isto geralmente provoca heterocedasticidade nos termos do erro, o que viola um pressuposto do modelo clássico de regressão (CAN, 1990).

Os parâmetros das equações hedônicas de preços de residências são estimados usualmente pelo método dos mínimos quadrados ordinários e neste procedimento assume-se que os resíduos são independentes e uniformemente distribuídos, com média zero, variância constante e covariância zero. Quando os resíduos estão espacialmente correlacionados, o pressuposto é violado, já que a covariância não é zero e a estimação deixa de ser eficiente e assim, os parâmetros devem ser estimados por um modelo que leve em consideração a autocorrelação espacial.

De acordo com Gillen et al. (2001), os resíduos dos preços podem estar espacialmente autocorrelacionados por três motivos: primeiro, porque as externalidades em uma região influenciam o valor de mercado das residências vizinhas de maneira similar. Thibodeau (1990, apud GILLEN et al., 2001) demonstrou que a construção de grandes edifícios de escritórios reduz o valor de mercado das residências que se situam nas suas imediações, em cerca de quinze

porcento. Segundo, que outras informações sobre estrutura e características do local não são facilmente obtidas e também ficam excluídas nos estudos, como a qualidade do ensino ou taxa de criminalidade em uma região que são difíceis de se observar. Além disso, as informações socioeconômicas e demográficas são obtidas pelo censo que só é realizado a cada dez anos. Quando as equações hedônicas são utilizadas nos modelos de preços das residências, o resíduo conterá informações destas características que não foram observadas. Em terceiro lugar, mesmo em uma situação ideal onde todas as características estejam disponibilizadas, é difícil se ajustar um modelo especificamente correto, como o caso de se especificar corretamente o quanto a qualidade das escolas influencia o preço das residências.

Informações sobre os determinantes das externalidades da vizinhança no mercado residencial são difíceis de se obter, conseqüentemente estas variáveis ficam omitidas nos estudos empíricos sobre o preço das habitações. A baixa especificação dos modelos pode contribuir para a autocorrelação espacial dos resíduos dos preços das equações hedônicas das habitações. Vários pesquisadores (Goodman e Thibodeau, 2003; Bourassa et al., 2007; Long et al., 2007; Bourassa et al., 2008) desenvolveram modelos de preços hedônicos para habitação que examinam a autocorrelação espacial dos preços e dos resíduos dos preços das residências.

A heterogeneidade espacial diz respeito a aspectos da estrutura socioeconômica do espaço geográfico (ANSELIN, 1988). A instabilidade estrutural no espaço pode ser observada na forma de coeficientes das variáveis do modelo que variam conforme a localização, de variância não constante dos erros (heterocedasticidade) ou da existência de formas funcionais diferentes para determinados subconjuntos de dados. Nesse caso, a conseqüência prática é a inadequação de se ajustar um mesmo modelo teórico para todo o conjunto de dados. Se a localização das observações for o fator determinante da instabilidade, a estrutura espacial, ou seja, a heterogeneidade, deve ser considerada explicitamente, e as ferramentas econométricas tradicionais não funcionam (PERUZZO TRIVELLONI e HOCHHEIM, 2006).

Recentemente, em estudos referentes a autocorrelação espacial, está em foco, especialmente nos Estados Unidos, o desenvolvimento de ferramentas de análise que utilizam as coordenadas de localização, com o objetivo de melhorar a acuidade e realismo dos modelos.

3.3 Matriz de pesos espaciais – W –

Em amostragens transversais, onde não se considera a variação dos dados em relação ao tempo, fenômenos de heterogeneidade espacial e autocorrelação espacial podem ser considerados equivalentes e não há informação suficiente para se estimar a matriz de covariâncias diretamente dos dados. É necessário que se imponha uma estrutura para essa matriz, e se crie um conceito de operador espacial de defasagem, com definição de uma matriz de pesos espaciais (W) para definir a estrutura de covariância dos dados de vizinhança de cada observação. O operador espacial calcula a média ponderada da variável na vizinhança para cada observação por meio de pesos que representam a covariâncias entre dois elementos (PERUZZO TRIVELLONI e HOCHHEIM, 2006). Assim, para uma certa variável y , a matriz de pesos espaciais W_y é a média ponderada do valor da variável y da vizinhança.

Vizinhança, na autocorrelação espacial ou associação espacial pode ser entendida como uma tendência de que o valor de uma variável associada a uma determinada localização assemelha-se mais ao valor de amostras vizinhas suas do que ao restante das localizações, em um conjunto amostral. Nesse sentido, considera-se que, em um conjunto de entes geográficos, ou seja, pontos, linhas ou polígonos, as variáveis com valores similares tendem a estar próximas e pode ser formado um aglomerado ou cluster.

De acordo com Gillen (2001), a autocorrelação espacial dos preços, assim como a dos resíduos dos preços, é função da distância e da direção que separa as residências. Mas a inclusão da direção além da distância nos modelos, torna a análise complexa e assim, nos estudos do mercado habitacional, não se considera usualmente esta dimensão e busca-se verificar a associação que depende, especialmente, da distância (TU, 2007). Na análise de dependência espacial, a associação pode ser representada, também, pela conectividade, como tempo ou custo de deslocamento entre os pontos, ou seja, o termo “vizinho” está baseado no padrão espacial adotado. De acordo com Peruzzo Trivelloni e Hochheim (2006), a matriz de pesos espaciais W influencia fortemente os resultados dos modelos.

A defasagem espacial ocorre quando a dependência é originada pela interação espacial da variável dependente, como no caso em que o preço de uma unidade residencial que influi na determinação do preço dos vizinhos.

3.4 Estudos sobre heterogeneidade do mercado habitacional – segmentação

A heterogeneidade na determinação dos preços das residências através da identificação de submercados tem sido abordada considerando essa como um caso especial da segmentação. O maior problema metodológico, neste caso, é como delinear os submercados. Na literatura, isto tem sido abordado de várias formas, buscado relacionar este delineamento com os limites políticos, como dados do censo e código postal das regiões, ou então pela determinação das áreas definidas pelos agentes do mercado. Contudo, diversos autores têm abordado em seus estudos técnicas estatísticas de delimitar os submercados, através dos modelos que consideram o agrupamento das residências de acordo com suas características e preço (ANSELIN, 2008).

O foco do estudo de Goodman (1981) é verificar qual deve ser o melhor esquema para se dividir uma área metropolitana. São experimentadas várias configurações de submercados para comparar coeficientes de preços hedônicos com os obtidos em testes de equilíbrio de longo prazo. A análise da covariância dos preços hedônicos residuais é combinada com o comportamento dos preços das residências de uma determinada área geográfica, para discutir as variações de preços dentro de uma área metropolitana.

De acordo com Goodman (1981), possíveis consumidores e locadores examinam as habitações apenas em uma região limitada, por causa do custo da procura, por discriminação racial ou pelo desejo de proximidade de amigos ou do local de trabalho, e este comportamento tende a reduzir a pressão por uma uniformidade do preço da habitação no mercado como um todo.

Tu (1997) apresenta uma estrutura simples, em dois níveis, para submercados habitacionais. No primeiro nível os submercados são ligados à vizinhança e se distinguem por seus componentes espaciais, e irão existir se houver persistência significativa de disparidades hedônicas nos preços da habitação, em relação aos seus componentes “chave” espaciais e os da vizinhança. No segundo nível, as residências pertencentes a cada submercado espacial são agrupadas de acordo com suas características não espaciais, como o tamanho. Se, dentro dos submercados espaciais, os grupos de produtos tiverem diferenças significantes em seus preços em relação a características “chave” eles pertencem a submercados distintos estruturalmente. As características

“chave” são as que trazem contribuição significativa para as diferenças entre as residências. As residências dentro de cada submercado devem ser semelhantes em termos de suas características “chave”.

Informações sobre o preço adicionadas ao endereço de cada propriedade, podem possibilitar o uso de um sistema de informações geográficas para calcular várias medidas de proximidade, incluindo medidas que podem ser utilizadas em modelos de dependência espacial (BOURASSA et al., 1997).

Goodman e Thibodeau (1998) definem submercados habitacionais como zonas geográficas em que o preço da habitação, por unidade de serviço, é constante, e consideram que a segmentação pode ser atribuída a características individuais das residências, a atributos do local ou a características da vizinhança. No estudo destes autores, o principal objetivo é identificar as áreas geográficas onde o preço das habitações é constante. Constatam que as famílias têm preferência explícita com relação aos atributos de certos locais, como o ensino público da região, e assim, o mercado habitacional é segmentado com relação aos serviços de cada região.

A teoria locacional da habitação, que tem base na teoria de Alonso (1965) propõe que o consumidor permuta o preço da residência por custos com transporte quando escolhe onde morar. Para reduzir os gastos com transporte, as propriedades são desenvolvidas próximas a grandes avenidas que levam ao centro de negócios. Este padrão de desenvolvimento habitacional pode resultar em forte autocorrelação espacial nos preços das residências e nas equações hedônicas de preços dos resíduos ao longo das principais vias de ligação e na direção do centro de negócios.

Bourassa e Hoesli (1999) consideram que as três principais dimensões em submercados habitacionais estão relacionadas às características físicas das residências, características da vizinhança (socio-econômicas e outras) e acessibilidade ao centro da região metropolitana.

Tu (1997) comenta que as forças que direcionam cada submercado podem ser diferentes, pois a estrutura de um submercado ou mercado local e sua evolução estão relacionadas à mudança da população local. As mudanças no tamanho da família, na estrutura familiar e no estilo de vida irão moldar as preferências dos compradores e irão criar uma nova demanda para os diferentes tipos de habitação. Nesta pesquisa, o autor investiga a ligação entre o mercado habitacional local que as famílias escolhem para morar e outras regiões

urbanas. Discute a descoberta de um sistema de equilíbrio das residências no mercado habitacional local, e as razões determinantes do equilíbrio estável de curto prazo, assim como os padrões de evolução do equilíbrio de longo prazo, para explorar algumas possibilidades de ligação entre o desenvolvimento dos submercados e do mercado habitacional global.

O mesmo autor ressalta que o desenvolvimento urbano, como mudança dos locais onde existe oportunidade de emprego e regeneração de áreas urbanas, orienta padrões de migração intra-urbana e remodela a distribuição das residências e pode ainda influenciar o tamanho da cidade. O desenvolvimento de um sistema de transportes local, em geral, também está relacionado com a evolução do sistema residencial local, porém não se sabe em que grau esta relação acontece. Conclui que o estudo da estrutura dos submercados locais e suas propriedades em comparação com o desenvolvimento do mercado habitacional global pode conduzir a um novo entendimento do complexo sistema urbano.

Tu (2000), em um trabalho sobre o mercado habitacional da Austrália chega a duas conclusões importantes: primeiro, que existe um certo número de relações sólidas entre os preços reais nacionais da habitação e alguns fatores econômicos e que, quando não há recessão, o salário, as taxas nominais para financiamento hipotecário, a taxa de desemprego e atividades da construção habitacional são as principais forças motrizes do mercado habitacional nacional. O autor conclui que o mercado habitacional australiano é altamente segmentado, ou seja, compreende uma série de submercados habitacionais regionais, que são impulsionados essencialmente por fatores econômicos internos. Três importantes implicações podem ser derivadas das conclusões acima: que a segmentação do mercado habitacional australiano em um nível subnacional ocorre porque o mercado da habitação e dinâmicas do desempenho do preço da habitação são determinados pelo desempenho econômico e há disparidade no desempenho econômico regional; que a política habitacional nacional pode não ter o mesmo impacto sobre as diferentes capitais, e que devem ser consideradas as disparidades regionais devido aos diferentes padrões de dinâmica regional do mercado habitacional e que o modelo nacional de preços para residências pode não representar modelos do preço regional da habitação.

Goodman e Thibodeau (2003) realizam uma análise do submercado baseada em dados de setores censitários e código postal

das residências, e concluem que os submercados obtidos com utilização dos códigos postais foram bem caracterizados e os resultados reduziram a variância, o que os autores consideram um benefício importante deste estudo.

O estudo de Bourassa et al. (2003) tem como objetivo testar se as previsões de valor hedônico são comprovadamente perfeitas quando um mercado é subdividido em submercados. Os autores comparam submercados de determinadas localizações divididos por faixa de valor com outros divididos por mesma faixa de valor mas não contíguos. Argumentam sobre a importância do tamanho da amostra e outras variáveis estatísticas da localização utilizadas. Utilizam dados sobre operações realizadas da cidade de Auckland, Nova Zelândia e confirmam a importância da localização e a importância da segmentação do mercado na determinação do valor do imóvel.

Maclennan (1977) analisa a hipótese implícita feita em estudos hedônicos de preços do mercado habitacional e conclui que a técnica pode ser aplicada apenas em cidades com extensa análise estática quando os dados da amostra são capazes de desagregação do significado.

Goodman e Thibodeau (2007), em um novo estudo, avaliam uma questão importante relacionada à construção de submercados: se áreas geográficas precisam ser espacialmente adjacentes para serem consideradas no mesmo submercado ou não. Consideram que o consumidor, quando toma a decisão de comprar uma moradia, não necessariamente limita sua escolha a áreas espacialmente concentradas e pode procurar similaridade de preços em vizinhanças localizadas através da área metropolitana. Examinam duas alternativas para delinear submercados: uma que combina regiões adjacentes em áreas com transações com dados suficientes para estimar parâmetros de equações de preços hedônicos e a segunda que permite descontinuidade espacial em submercados.

Tu et al. (2007) deixam que os dados definam a segmentação urbana residencial, substituindo as delimitações convencionais administrativas, ou qualquer outra delimitação pré-definida usada na literatura de submercados imobiliários. O estudo é experimental e representa uma das primeiras tentativas de investigar a segmentação no mercado habitacional através autocorrelação espacial do preço. Os dados de transações de imóveis são utilizados com aplicação de uma função hedônica convencional. Os resultados são comparados com os obtidos em estudos anteriores sobre a estrutura dos segmentos do

mercado habitacional de Singapura. O teste utilizado mostra que esta estrutura de submercados pode melhorar a precisão da previsão de preços em 17,5% em relação aos definidos em estudos anteriores que utilizavam limites pré-definidos realizados na mesma área.

4 - MÉTODOS PARA A SEGMENTAÇÃO DO MERCADO HABITACIONAL

4.1 Identificação dos submercados habitacionais

Watkins (1998) discute que através da literatura sobre o mercado habitacional há uma imprecisão considerável na conceitualização de submercado. As várias definições, os critérios de demarcação, as diferentes técnicas estatísticas empregadas e ainda o fato de que os estudos utilizam dados de regiões diferentes e diversos períodos de tempo têm impedido o desenvolvimento de uma abordagem analítica coerente. Com dados do mercado habitacional de Glasgow são comparadas várias definições alternativas e esquemas de identificação. Fica evidente a importância dos submercados e que estes possuem, simultaneamente, dimensões espaciais e estruturais.

De acordo com Goodman et al. (2007) é importante que se entenda como as regiões metropolitanas são repartidas em submercados habitacionais por várias razões. Primeiro, conhecendo as propriedades dos submercados pode-se aplicar modelos estatísticos para melhorar a estimação do preço das residências. Segundo, a identificação das fronteiras dos submercados dentro das áreas metropolitanas melhora a modelagem espacial e temporal de variações dos preços. Terceiro, um melhor conhecimento dos submercados aumenta a habilidade em se calcular o risco associado ao investimento em imóveis. Finalmente, informações sobre as delimitações dos submercados podem reduzir o custo da procura de imóveis para os compradores.

O conceito original de submercado considera que a segmentação do mercado ocorre quando o mercado de um certo bem é subdividido em grupos homogêneos de consumidores (Kotler, 1999). A segmentação é uma forma utilizada pelo produtor para atingir determinados grupos de consumidores através da diferenciação de um produto (SMITH, 1956).

Mas, para que um submercado exista, não é suficiente a diferenciação do produto ou a existência de diferentes grupos de consumidores. Deve haver uma interação entre oferta segmentada e conduta diferenciada da demanda a variações significantes do preço para produtos padronizados. Não se pode ignorar as ligações que ocorrem

entre as características dos consumidores e dos produtos, porém, freqüentemente, se conceitualiza a estrutura do mercado habitacional de acordo com a oferta, ou seja, pelas características dos produtos (WATKINS, 1998).

Para Fisher e Fisher (1954), o mercado imobiliário pode ser dividido de acordo com o uso, tipo de posse, localização ou características físicas. Quanto ao tipo de uso, o mercado habitacional inclui apenas as unidades do estoque que são usadas preferencialmente para propósitos residenciais, e estas não pertencem ao mercado comercial ou ao de escritórios, pois não são substitutas das unidades destes outros mercados; quanto ao tipo de posse, o imóvel pode ser ocupado pelo proprietário ou alugado. Vários serviços devem ser consumidos no próprio lugar, assim, o mercado local fica confinado a certas áreas restritas: em relação à localização considera-se o estudo de certas regiões, como os bairros ou toda uma região metropolitana. O mercado habitacional pode também ser descrito pelas características físicas das unidades em estoque incluindo o número de quartos, área, tipo de serviços disponíveis ou idade da construção .

Na definição de Rapkin et al. (1953, apud FISHER e FISHER 1954), a posição espacial exerce influência apenas referente aos serviços contidos naquela localização. Eles explicam que o mercado habitacional é caracterizado por uma demanda segmentada e diferentes estoques. Segmentação do estoque e da demanda são conseqüências da heterogeneidade do estoque, da imobilidade espacial e da durabilidade. A existência de imobilidade espacial implica que os compradores irão selecionar as residências com base em uma ordem de serviços disponíveis, públicos e privados, em uma determinada região, e, pelas diferenças na qualidade dos serviços, o mercado compreende uma agregação de submercados que não estão em competição entre si.

O mercado habitacional pode estar dividido em submercados devido a fatores relacionados tanto à oferta quanto à demanda. Os submercados podem ser definidos pelo tipo de estrutura das residências (como casa isolada, casa em condomínio, apartamento), por características estruturais da unidade residencial (como número de dormitórios e tamanho) e do condomínio quando for o caso, ou por características do local onde se situa (educação pública, segurança e facilidade de transporte, entre outras). Alternativamente o mercado pode ser segmentado por características da demanda como renda familiar e raça (GOODMAN e THIBODEAU, 1998).

Em diversos estudos, especialistas enfatizam o produto, considerando a importância das dimensões espaciais e estruturais, separadamente ou em conjunto na determinação das dimensões dos submercados.

4.1.1 Submercados espaciais

Bajic (1985) segmentou o setor de residências ocupadas por proprietários através de linhas estruturais, assim como Rothenberg et al. (1991). Dale-Johnson (1982) também construiu segmentos de mercado compreendendo residências com características físicas similares.

Em vários estudos empíricos sobre a existência de submercados assume-se que características espaciais são mais importantes que características estruturais. Straszheim (1975, apud GOODMAN e THIBODEAU, 2003) argumenta que o preço das residências pode variar consideravelmente dentro da área metropolitana notando que a característica fundamental do mercado habitacional urbano é a variação dos preços, que são diferentes de acordo com a localização do imóvel.

Ball e Kirwan (1977) notaram que submercados emergem se houver restrições para compradores e vendedores de entrarem no mercado de determinadas áreas geográficas. Especialmente, eles consideram as restrições na demanda pela falta de financiamento ou pela baixa disponibilidade de residências em uma determinada área. Entretanto, argumentam que características históricas dos estoques também podem introduzir uma segmentação espacial.

Palm (1978) analisa os preços de residências na área da baía de São Francisco, nos Estados Unidos, e verifica que ocorre uma diferente relação de preços por bairros e quando se considera o mercado como um todo. Pelos padrões correlacionados aos modelos do mercado global, diferente dos de submercados, conclui que o modelo de segmentação deve ser usado nos estudos do mercado habitacional, porém, ressalta que fronteiras dos submercados devem ser cuidadosamente definidas, pois a diferença pode ser discreta entre os grupos.

No mesmo estudo, o resultado da análise leva a conclusão que a fronteira, ou seja, a delimitação dos submercados estudados, exerce maior influência do que qualquer outra relação. Assim, o mais importante é delimitar corretamente os mercados para as análises feitas dentro das áreas nas quais é provável que as residências sejam trocadas. Na região analisada, notou ainda que, apenas em áreas restritas, certos

atributos da vizinhança exercem efeitos positivos ou negativos na tendência dos preços.

De acordo com Watkins (1998), existe indicação de que características espaciais são importantes devido aos atributos relativos à acessibilidade da localização e às características da vizinhança.

Munro (1986), em um estudo do mercado de Glasgow, na Escócia, verifica que uma barreira física pode ser determinante na ocorrência da segmentação de uma área urbana, com a constatação da divisão do mercado habitacional em submercados distintos ao norte ou ao sul do Rio Clyde.

Goodman (1978, 1981) fornece suporte empírico adicional para mercados habitacionais segmentados geograficamente. Ele comparou os coeficientes hedônicos de características estruturais e dos bairros em cinco áreas metropolitanas. Como verificou que os coeficientes não foram constantes, concluiu que o mercado metropolitano é segmentado.

Michaels e Smith (1990) concluem que a identificação de submercados distintos deve ser importante para a caracterização de quanto o atributo da localização de uma residência influencia equilíbrio de preços. Gabriel (1984) também encontra suporte, nos resultados de seu estudo, para segmentação baseada em diferenças na vizinhança.

De acordo com Watkins (1998), não se sabe como os limites geográficos são delineados. Hancock (1991, apud Watkins 1998) delimitou uma área e testou a validade dos dados associados a variáveis com equações de regressão. Estudos de outros autores buscam a dimensão dos submercados empiricamente.

Goodman e Thibodeau (2007) avaliam uma questão importante relacionada à construção de submercados: se áreas geográficas precisam ser espacialmente adjacentes para serem consideradas no mesmo submercado. Consideram que o consumidor não necessariamente limita sua escolha a áreas espacialmente concentradas e pode procurar similaridade de preços em vizinhanças localizadas através de toda a área metropolitana, quando estão tomando a decisão de comprar uma moradia. Examinam duas alternativas para delinear submercados: uma que combina regiões adjacentes e a segunda que permite descontinuidade espacial em submercados, delineando-os pelo tamanho da residência e preço médio do metro quadrado. Concluem que ambas definições são importantes e a aplicação de uma ou outra depende do objetivo de cada estudo.

4.1.2 Submercados estruturais

É discutível que os submercados não estejam vinculados espacialmente. Embora Straszheim (1975) tenha encontrado que certos preços sejam diferentes em algumas partes da cidade, Rothemberg et al. (1991) não admitem variações nos preços dos atributos dentro de uma região metropolitana.

De acordo com Watkins (1998) alguns autores consideram que o tamanho e a renda das famílias necessariamente restringem o tipo de residência que elas procuram. Assim, o tamanho da residência, o número de quartos e outras características físicas podem ser os fatores determinantes para a segmentação do mercado habitacional.

Dale-Johnson (1982) testou treze variáveis com base em residências de mesmas características e além de obter uma confirmação adicional da hipótese de segmentação do mercado habitacional, constatou a importância da dimensão estrutural das residências.

Bajic (1985) utilizou dados de imóveis dos quais tinha informações sobre características sócio-econômicas dos proprietários como idade, ocupação profissional, educação e número de crianças e adultos nas residências. Coletou dados sobre acesso, como transporte e vias. O autor verificou a existência de três submercados com condições de oferta e demanda diferentes em cada um de acordo com características estruturais como a área do lote e a área do imóvel.

4.1.3 Submercados espaciais e estruturais em conjunto

Maclennan e Tu (1996, apud WATKINS, 1998) explicam que ambos os fatores, espaciais e estruturais, separadamente e interagindo, podem gerar submercados.

A segmentação do mercado pode ser atribuída a diferenças das características estruturais, atributos do local em que se situa, ou a ambos. Imóveis que foram construídos em épocas diferentes podem ter características que não são facilmente modificadas ou reproduzidas e as amenidades de cada localização também dificilmente são duplicadas. Para Goodman e Thibodeau (1998), os dois fatores estão ligados ao processo de desenvolvimento de cada região, e para certos grupos da demanda, os imóveis não são compreendidos como substitutos no processo de escolha da residência.

Jones et al. (2001) concluem que, apesar de submercado definidos espacialmente, bem como os que o são estruturalmente

promoverem a evidência sobre a existência de submercados, as dimensões dos submercados são determinadas pelas duas dimensões do produto, simultaneamente.

Malpezzi (2002) sugere que, algumas das variáveis locais e estruturais das residências podem ser relevantes de acordo com o propósito do estudo, que são: número de quartos, área, idade da unidade, padrão do acabamento, presença de garagens, distância ao centro e acesso a escolas, shoppings e além de outros atributos da localidade e características da vizinhança.

A estrutura proposta por Adair et al. (1996, apud WATKINS, 1998) é que o mercado é primeiramente segmentado espacialmente e em segundo lugar os submercados são diferenciados estruturalmente, dentro de cada região. Nesta conceituação, os submercados são denominados ‘nested’, com a subdivisão estrutural inserida na espacial.

4.2 Métodos para segmentar o mercado habitacional

Existe uma ampla literatura que suporta teoricamente a existência de submercados habitacionais, porém pouco consenso sobre como deve ser o modelo (WATKINS, 1998).

Os atributos do produto devem ser similares dentro de um submercado e diferentes em relação aos demais submercados. Os padrões de substituição podem estar ligados ao preço, custo da procura, desejável proximidade a serviços essenciais, como os serviços públicos, qualidade do ambiente, segregação étnica ou à demanda pelos atributos de um tipo de residências.

Diferentes equações devem ser estimadas para cada submercado. Nos primeiros estudos os pesquisadores determinavam um grupo de residências com características em comum, conforme o objeto do estudo, para então, testar se os grupos podiam ser considerados como submercados habitacionais. Em estudos mais recentes os agrupamentos passam a ser identificados através de técnicas estatísticas. Diversos estudos sobre o mercado habitacional têm como objetivo a comparação dos submercados definidos *a priori*, ou seja, os previamente classificados com os definidos por técnicas de agrupamento, como a análise de *cluster*. Assim, para se testar a segmentação espacial, a base é uma determinada localidade; para a segmentação estrutural, um determinado tipo de residência.

Uma opção, também muito utilizada nos estudos dos submercados habitacionais envolve a redução do número de variáveis para que fiquem apenas as que tenham significância estatística.

Para que um grupo de residências possa ser considerado um submercado residencial, esse deve ser significativamente distinto dos demais grupos e do mercado global em estudo. Um procedimento amplamente utilizado para esta verificação é a aplicação de equações de regressão, que associam o preço ou valor da habitação aos seus atributos, as equações hedônicas. Nestas equações os coeficientes das variáveis representam o peso que cada atributo possui que significam o quanto cada característica contribui para a formação do preço da residência, pois, como já foi amplamente discutido neste estudo, a melhor forma de se tratar o bem “habitação” é através de suas características, visto sua natureza heterogênea.

Nos estudos mais recentes, os pesquisadores, na tentativa de melhorar os modelos, procuram entender o significado da parcela não explicada destes, alguns procuram entender padrões de migração urbana e outros a influência das interligações que existem entre as residências que se situam em um determinado local.

Na última década surge a estatística espacial, e a introdução de extensões dos modelos tradicionais permite que se possa captar padrões globais de uma área geográfica bem como padrões locais, que distinguem como uma determinada região.

Nem todos os estudos tiveram sucesso na determinação das dimensões dos submercados, mas exceto o de Schnare e Struyk (1976) todos comprovaram a existência de segmentação no setor habitacional.

Kauko (2004) relaciona alguns modelos para determinação de submercados:

- Modelos hedônicos: Rothenberg et al. (1991); Watkins (1998); MacLennan e Tu (1996);
- Repetitividade de vendas em combinação com cointegração: Jones et al. (2003);
- Preços hedônicos hierárquicos: Goodman e Thibodeau (1998, 2003);
- Projeção de clusters, às vezes combinado com análise de regressão: Bourassa et al. (1997, 1999, 2003);
- Rede neural Kauko et al. (2002);
- Modelo de mobilidade familiar: Jones (2002).

Abaixo estão apresentados os métodos para determinação de submercados habitacionais mais utilizados nos trabalhos que foram reunidos neste estudo. Outros métodos de segmentação adequados a vários propósitos foram revisados por Freitas (2000).

4.2.1 Submercados pré-definidos ou definidos *a priori*

Nos primeiros estudos hedônicos, os submercados são pré-definidos e delineados com base em critérios da localização e do ambiente (BOURASSA et al., 1997).

Nos critérios para definir os limites dos submercados eram consideradas a ocorrência de certas características estruturais nas residências e também de características espaciais, ou seja, barreiras físicas, como os rios; fronteiras geográficas ou administrativas, como as divisas de bairros (STRASZHEIM, 1975), e fatores locacionais, ligados à distância e acessibilidade ao centro de negócios da cidade (ALONSO, 1965; WHEATON, 1977).

Schnare e Struyk (1976) dividem Boston, nos Estados Unidos, de acordo com a localização dentro ou fora da cidade, renda e número de quartos e testam a hipótese de que os mercados habitacionais urbanos são segmentados em submercados definidos espacialmente ou estruturalmente no sentido de diferenças significativas nos preços por unidade de serviços habitacionais.

Alguns pesquisadores passaram a utilizar métodos estatísticos para formar fronteiras geográficas e definiram a região com estas limitações geográficas como as menores unidades de análise. Goodman e Thibodeau (1998) argumentam que se a menor unidade de análise for imposta ao invés de modelada, não se pode confiar que os submercados resultantes sejam identificados em um caminho ótimo.

Maclennan e Tu (1996 apud TU, 1997) consideram que estudos econômicos da habitação tem a tarefa de revelar, ao invés de assumir *a priori* a estrutura do mercado. Mas, para estabelecer empiricamente esta estrutura, deve-se utilizar conceitos consistentes, assim como o de grupos de produto ou submercados.

Bourassa e Hoesli (1999) verificam se a estrutura de submercados habitacionais pré-definidos difere da dos encontrados através de análises estatísticas (PCA e análise de cluster), em Auckland, na Nova Zelândia, com o objetivo de avaliar o impacto de se impor restrições para a construção de submercados. Utilizam três critérios diferentes para a construção dos submercados pré-definidos: por tipo de propriedade

(isolada ou agrupada), valor da residência e área geográfica. Os submercados classificados de acordo com o valor enfatizam todas as dimensões do mercado habitacional, visto que esse tipo de classificação reflete os vários atributos das residências. O resultado é que o mercado habitacional da região pesquisada está relacionado, principalmente, com as características físicas das residências. Assim, a classificação melhor entre os tipos de submercados pré-definidos, é por tipo de propriedade, mas concluem que ao se pré-definir os submercados estes ficam direcionados e não refletem as dimensões da segmentação habitacional no que diz respeito à região metropolitana estudada.

Bourassa et al. (1997) comparam submercados pré-definidos com outros construídos por técnica estatística de agrupamento (cluster) para analisar o mercado habitacional em duas cidades, Sydney e Melbourne, na Austrália. As equações calculadas para as classificações dos submercados definidos *a priori* são comparadas com a do mercado global de cada uma das cidades e o mesmo é feito para os submercados definidos pela análise de agrupamento. Apenas um dos submercados de Melbourne, que foi definido pela técnica de agrupamento, é significativamente melhor do que todos os outros, considerando-se todos, construídos pelos dois métodos. Nos outros casos, a análise estatística também produz submercados que são melhores, mas a melhora não é significativa na comparação destes submercados com os construídos pela classificação *a priori*. Segundo os autores, isto pode ter ocorrido, em particular neste estudo, como consequência da qualidade dos dados utilizados, pois a variável dependente utilizada foi uma estimativa do valor de mercado feita pelos proprietários das residências e não o preço, visto que este não estava disponível. Para eles, dados mais precisos podem permitir a constatação de que os submercados construídos devem ser claramente superiores aos submercados definidos *a priori*.

4.2.2 Análise dos componentes principais - PCA (Principal Component Analysis)

PCA ou análise de fatores é um procedimento que deriva um pequeno número de combinações lineares para reduzir o número de variáveis originais enquanto retém a quantidade substancial de informações. O uso do PCA é importante porque identifica as dimensões básicas que distinguem os submercados habitacionais.

Este teste de experimentos fatoriais é utilizado para se comprovar, com uma certa confiabilidade, a significância ou a influência de características de produtos ou processos. Para se comparar duas populações, por meio de seus parâmetros, deve-se verificar sua forma de distribuição. Especificada a forma, a igualdade dos parâmetros que identificam a curva implica em igualdade das populações. No caso de populações com distribuição normal, antes de se processar o teste de comparação de médias, deve-se fazer o teste de comparação das variâncias. Deve-se verificar se as variâncias podem ser consideradas iguais, para que ambas as populações tenham a mesma dispersão.

Se as médias diferem significativamente, pode-se afirmar que existe preferência por determinada tipologia em relação ao parâmetro analisado. Com a análise da variância, pode-se testar a hipótese da influência de um ou mais fatores no resultado de um experimento e a existência de inter-relações entre variáveis. As operações processadas para análise da variância podem ser agrupadas em uma única tabela, conhecida como ANOVA (Analysis of Variance).

Maclennan et al. (1987) utilizam PCA em uma região com estoque de residências semelhantes e características socioeconômicas também semelhantes.

4.2.3 Análise de *cluster*

A técnica da análise de cluster é similar à análise fatorial, mas trata os objetos enquanto a análise fatorial se preocupa com as variáveis.

Esta análise tem sido utilizada em pesquisas de marketing para segmentação do mercado, para a compreensão do comportamento do comprador e para identificação das oportunidades para um novo produto. É uma técnica usada para classificar objetos ou produtos em grupos relativamente homogêneos chamados clusters ou conglomerados. Os objetos em cada cluster tendem a ser semelhantes entre si, mas diferentes de objetos de outros clusters (Malhotra, 2006).

Os clusters são sugeridos pelos dados e não definidos a priori. A análise de clusters examina a interdependência entre todo o conjunto de variáveis, pois não faz distinção entre variáveis dependentes e independentes.

Conforme Malhotra (2006), para a aplicação da análise de cluster, inicialmente é necessário definir o problema de aglomeração e as variáveis a serem tratadas estatisticamente. Escolhe-se, então, uma medida de distância dos conglomerados. Após, define-se o processo de

aglomeração que dependerá das variáveis em estudo e do problema em foco. Neste ponto, a intuição do pesquisador deve ser utilizada para a escolha do melhor processo e definição do número de conglomerados da próxima etapa. Os conglomerados resultantes devem ser interpretados em termos das variáveis usadas para constituirlos e de outras variáveis adicionais importantes. Finalmente, o pesquisador precisa avaliar a validade do processo de aglomeração.

Maclennan e Tu (1996) produziram uma matriz de submercados com quatro setores da cidade e cinco grupos de tipologias de produto. Utilizaram PCA e análise de clusters.

4.2.4 Modelos hedônicos

Tu et al. (2007) modelaram dados de negociação de residências usando função hedônica convencional. As correlações entre os resíduos hedônicos são usadas como identificadores para incluir unidades residenciais em *clusters*. Os testes padrão para identificação de submercados, equações hedônicas, são aplicados para cada zona para examinar a segmentação do mercado.

São dois tipos de variáveis envolvidas nos modelos hedônicos de preços de residências, as dependentes e as independentes. A variável dependente é uma medida do valor da habitação representada pelo preço de venda, preço desejado, valor do aluguel, ou até pela estimação do valor da residência pelo proprietário. Nos estudos hedônicos, o uso de diferentes medidas para o valor do imóvel é evidente. O valor do aluguel é normalmente usado em estudos que envolvem propriedades alugadas, enquanto os demais são utilizados em estudos que envolvem residências ocupadas pelos proprietários.

O segundo tipo de variável na análise de regressão são as variáveis independentes, que incluem todos os fatores significantes que afetam a medida do preço. São duas as dimensões mais importantes dos fatores: estrutural e locacional ou espacial. Enquanto as características estruturais são conceitualizadas diretamente, fatores locacionais podem ser de certa forma confusos. Isto por causa de diferentes termos em diferentes estudos representando fatores locacionais. Em alguns estudos são consideradas que são três as dimensões, ou seja, características estruturais das propriedades, acessibilidade e qualidade da vizinhança. Outras variáveis com base no censo ou elementos sobre as características da região também são utilizadas, como a qualidade das escolas públicas ou ainda qualidade do meio ambiente, do ar ou da água.

Atributos estruturais envolvem o tipo da residência, se é isolada ou em condomínios, se é apartamento e ainda a área do terreno, a área útil, além de características como o número de quartos, de banheiros ou de garagens, qualidade ou estado de conservação e há quanto tempo está construída. Acessibilidade inclui medidas baseadas em distâncias ou custo envolvido para o deslocamento ao centro de negócios, escolas, transporte públicos entre outros. Qualidade da vizinhança inclui características socioeconômicas e demográficas como a renda média, o nível de educação, composição étnica, qualidade da escola mais próxima e taxa de criminalidade e, ainda, amenidades do local, que são os atributos agradáveis do ambiente, como a proximidade de parques. Alguns atributos da acessibilidade e atributos da vizinhança são difíceis de se identificar e quantificar (TU et al., 2004), pois em alguns casos, apesar da conveniência de uma localidade conter um atributo relativo à acessibilidade, como uma estação de trem, a proximidade desta pode conter desvantagens como ruídos, poluição ou grande movimentação de pessoas, o que faz com que residir no entorno desta estação não promova vantagem para algumas famílias.

4.3 Modelos de Regressão com Efeitos Espaciais

Os modelos de regressão espacial também necessitam dos três principais pressupostos do modelo de regressão convencional, ou seja, os erros são normais com média zero; os erros têm variância constante (homocedasticidade); os erros são não correlacionados. Porém, ao se incorporar na sua estrutura o fator “espaço”, na maioria das vezes, são eliminados os problemas de autocorrelação e heterocedasticidade mencionados anteriormente, isso porque esse último ocorre, dentre outros motivos, devido à ausência de variáveis, sejam observáveis ou não observáveis como é o caso do “espaço”.

Os modelos de regressão espacial, por terem o fator “espaço” em sua estrutura, corrigem os valores dos coeficientes angulares β , além de se ajustarem melhor aos dados. Isto se deve ao fato da contribuição do fator “espaço” estar distribuída, em um modelo de regressão convencional, entre as variáveis do modelo e assim, o valor do β estimado na amostra está super ou subestimando o verdadeiro β populacional.

4.3.1 Modelo Geral com Efeitos Espaciais Globais

A dependência espacial pode ser incorporada aos modelos clássicos de regressão pela adição de uma variável independente correspondente à variável dependente defasada (\mathbf{W}_Y) ou por uma estrutura espacialmente defasada no erro da regressão (\mathbf{W}_ε). Segundo Anselin (1988), a forma geral de um modelo espacial auto-regressivo global está representado na Equação (1). Um processo auto-regressivo é aquele em que uma variável é explicada por si própria, como acontece nos modelos de séries temporais e, no caso dos modelos espaciais, a variável é explicada também pelos seus vizinhos.

$$\mathbf{Y} = \rho \mathbf{W}_Y + \beta \mathbf{X} + \varepsilon$$

$$\varepsilon = \lambda \mathbf{W}_\varepsilon + \mathbf{u}$$

Então:

$$\mathbf{Y} = \rho \mathbf{W}_Y + \beta \mathbf{X} + \lambda \mathbf{W}_\varepsilon + \mathbf{u} \quad (1)$$

Onde: \mathbf{Y} é a variável dependente; ρ e λ são coeficientes de autocorrelação espacial; \mathbf{X} é a matriz das variáveis independentes dos dados; \mathbf{W}_Y e \mathbf{W}_ε são matrizes de pesos espaciais; β é o vetor dos coeficientes; ε é o vetor de resíduos e \mathbf{u} é o vetor dos resíduos não correlacionados ou erro aleatório.

Este modelo considera a dependência espacial na variável dependente \mathbf{Y} e no erro aleatório ε , e não necessariamente as matrizes \mathbf{W}_Y e \mathbf{W}_ε precisam ser diferentes. Os outros modelos são derivados deste, como serão apresentados a seguir.

4.3.2 Modelo Espacial Autoregressivo de Primeira Ordem

Esse modelo é utilizado quando se deseja explicar a variável dependente \mathbf{Y} a partir de seus vizinhos. A dependência espacial está na própria variável \mathbf{Y} . Fazendo $\mathbf{W}_\varepsilon = 0$ e $\mathbf{X} = 0$ chega-se a Equação (2)

$$\mathbf{Y} = \rho \mathbf{W}_Y + \mathbf{u} \quad (2)$$

onde todos os parâmetros são como na Equação (1).

4.3.3 Modelo de defasagem espacial – (*Spatial Lag Model*)

É utilizado quando se deseja explicar a variável dependente Y a partir dela mesma e de outras variáveis explicativas. É obtido da Equação (1), fazendo $W_{\varepsilon} = 0$.

$$Y = \rho W_Y + \beta X + u \quad (3)$$

onde todos os parâmetros são como na Equação (1).

Se o parâmetro espacial ρ na Equação (3) for zero, então o modelo resultante é exatamente igual a um modelo de regressão convencional. Quando o valor de ρ estiver próximo de zero (baixa dependência espacial), pouca informação será agregada a β , enquanto que se estiver próximo de +1 ou -1 (alta dependência espacial), um valor significativo será agregado a β . Assim, pode-se considerar que a regressão espacial corrige os parâmetros do modelo, quando comparada à regressão convencional (Câmara et al., 2002).

4.3.4 Modelo de erro espacial

Neste modelo a estrutura espacial está no erro aleatório. É obtido a partir da Equação (1) fazendo-se $W_Y = 0$

$$Y = \beta X + \lambda W_{\varepsilon} + u \quad (4)$$

onde todos os parâmetros são como na Equação (1).

Este modelo é apropriado quando o resíduo resultante de um modelo de regressão convencional possui dependência espacial, constatada pelo I de Moran, ou seja o erro é função do erro dos vizinhos. Outros fatores que corroboram a utilização desse modelo são os erros de medida, ausência de variáveis explicativas ou variáveis não-observáveis, além da heterocedasticidade (CÂMARA et al., 2002).

4.3.5 Modelos com Efeitos Espaciais Locais

Os modelos espaciais globais partem do pressuposto que o processo espacial analisado é estacionário, ou seja, possui um único padrão de associação espacial. Câmara et al. (2002) argumentam que a estacionariedade nem sempre ocorre, pois na prática, para conjuntos de dados censitários de médio e grande porte, a natureza dos processos espaciais é tal que diversos padrões de associação espacial podem estar

presentes. Esta hipótese, que pode ser verificada, por exemplo, pelos indicadores locais de autocorrelação espacial, está na origem aos modelos cujos parâmetros variam no espaço.

Dessa forma foram desenvolvidos os modelos com efeitos espaciais locais, a fim de se ter um refinamento maior do modelo e um melhor entendimento do processo espacial, que são utilizados quando o processo é não-estacionário, e por isso, há a necessidade de que os coeficientes da regressão reflitam essa heterogeneidade espacial (CÂMARA et al., 2002).

Existem duas formas de se trabalhar com efeitos locais: modelar a variação espacial de forma discreta, dividindo o espaço em sub-regiões estacionárias denominadas regimes espaciais, ou modelar a tendência espacial de forma contínua, com parâmetros variando no espaço.

4.3.6 Modelagem de forma discreta

Na modelagem de forma discreta a idéia é identificar primeiramente os regimes espaciais, com a ajuda do diagrama de espalhamento de Moran e/ou dos indicadores locais. Dessa forma é possível dividir os dados em subregiões, criando indicadores como por exemplo na Equação (5), e assim estimar os parâmetros do modelo conjuntamente.

$$Y_1 = X_1 \beta_1 + \varepsilon_1 \quad (5)$$

$$Y_2 = X_2 \beta_2 + \varepsilon_2 \quad (6)$$

onde:

- Y_1 e Y_2 são variáveis dependentes;
- X_1 e X_2 são variáveis independentes;
- β_1 e β_2 são parâmetros do modelo;
- ε_1 e ε_2 são erros aleatórios.

Câmara et al. (2002) comentam sobre a utilização prática desses modelos: “Na prática, para os dados sócio-econômicos típicos de cidades brasileiras, o modelo de regimes espaciais tende a apresentar resultado melhor que os modelos de regressão simples ou de regressão espacial com efeitos globais. Isto ocorre em função das fortes

desigualdades sociais no Brasil, que ocasionam descontinuidades abruptas nos fenômenos estudados, como no caso do recorte entre favelas e áreas ricas, como é freqüente em nossas grandes cidades”.

4.4 GeoDa

O GeoDa é um programa que realiza análises estatísticas a partir de dados georreferenciados, ou seja, dados que possuem coordenadas geográficas, podendo ser traduzidos em pontos de um mapa. Cada ponto está interligado às informações relacionadas a ele através de uma tabela que contém os dados.

Com o GeoDa efetua-se a modelagem de Regressão Clássica e de Regressão Espacial: defasagem espacial (*Spatial Lag*) e erro espacial (*Spatial Error*).

Para verificar a autocorrelação espacial, o GeoDa utiliza seis testes estatísticos: o principal é o *I* de Moran, e os demais são referentes ao multiplicador de Lagrange que são: LM-Lag; Robust LM-Lag, LM – Error, Robust LM – Error, LM-SARMA. Os testes com multiplicador de Lagrange devem ser aplicados em seqüência, assim, os testes Robust LM-Lag e Robust LM – Error só devem ser considerados se os testes LM-Lag e LM – Error forem significativos.

O GeoDa possui três testes de heterocedastidade (não-constância da variância dos erros) que são: Breusch-Pagan, Koenker-Bassett e White. Os dois primeiros são semelhantes e podem ser aplicados sobre coeficientes aleatórios, cuja forma funcional no GeoDa é de função dos quadrados das variáveis. A diferença é que no segundo os resíduos são não “studentizados”, ou seja, o teste pode ser feito na presença da não normalidade dos erros. O teste White não assume forma funcional específica para a heterocedasticidade, por isso ele pode ser usado como teste mais geral, porém existe uma restrição: que a amostra seja grande.

4.4.1 Testes de autocorrelação espacial

O “coeficiente de correlação de Pearson”, índice r , mede a correlação entre duas variáveis x e y . O valor de r varia entre -1 e $+1$, sendo que $+1$ indica correlação positiva máxima, -1 correlação negativa máxima e quando o índice r é igual a zero não ocorre correlação.

Os testes de autocorrelação espacial têm o mesmo objetivo do coeficiente de correlação de Pearson, apenas introduzindo em sua

estrutura a informação referente ao fator “espaço”. Segundo Câmara et al. (2002), eles medem o nível de interdependência geográfica entre os valores da variável, bem como a natureza e a força deste relacionamento.

Existem duas formas para medir a dependência espacial: através de indicadores globais que capturam tal dependência em apenas um fator, e através de indicadores locais que verificam a dependência espacial em cada área estudada.

4.4.1.1 Índice *I* de Moran Global – Índice de Associação Espacial Global

Este índice é o mais difundido para verificação de dependência espacial e mede a autocorrelação espacial a partir do produto dos desvios em relação à média. Ele representa uma medida global da autocorrelação espacial, pois indica o grau de associação espacial presente no conjunto de dados. É uma medida importante, porque sintetiza inúmeras possibilidades de padrões de associação local entre as amostras georreferenciadas. São nestes padrões locais que em grande parte reside a informação relevante para análises geográficas. Especificamente em análises intra-urbanas é fundamental que se investigue as configurações locais de associação espacial pois o foco da análise é justamente a observação da organização territorial interna, ou seja, importa saber como determinadas características geográficas se distribuem espacialmente, se há concentrações ou tendências de determinadas características que possam revelar os elementos territoriais estruturais.

O índice *I* de Moran Global é semelhante ao coeficiente de correlação convencional, já que possui um termo que é produto dos desvios em relação à média global, ponderada pela proximidade geográfica entre áreas vizinhas, ou seja, é um coeficiente de correlação entre valores da mesma variável em locais vizinhos.

Neste índice a hipótese nula é de independência espacial. Como o coeficiente de correlação de Pearson, o *I* de Moran varia de -1 a +1, onde valores próximos de zero indicam ausência de autocorrelação espacial, valores positivos (entre 0 e +1) indicam autocorrelação direta e valores negativos (entre 0 e -1) autocorrelação inversa.

O *I* de Moran Global para *N* observações é dado por:

$$I = \frac{N \sum_{i=1}^N \sum_{j=1, i \neq j}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{w_0 \sum_{i=1}^n (x_i - \bar{x})^2} \quad (7)$$

onde:

- x_1, x_2, \dots, x_N são as observações de N locais;
- N representa o número de observações;
- w_{ij} são os elementos da matriz de contigüidade binária normalizada (W) representando a proximidade de diferentes localizações, ou seja, se i e j são locais vizinhos ($i \neq j$), w_{ij} terá o valor unitário, caso contrário terá o valor nulo (essa é a hipótese nula do teste);
- x_i e x_j são os valores da variável analisada e

$$w_0 = \sum_{i=1}^N \sum_{j=1}^N w(i,j).$$

Deve ser realizado um teste para se estabelecer a validade estatística do Índice de Moran, sendo que o mais utilizado é o de pseudo-significância. Neste caso são geradas diferentes permutações dos valores de atributos associados às regiões e cada permutação produz um novo arranjo espacial, onde os valores estão redistribuídos entre as áreas. Pode-se construir uma distribuição empírica de I e, se o valor do índice I medido originalmente corresponder a um “extremo” da distribuição simulada, significa que este arranjo, corresponde à situação observada, não é aleatório e corresponde a uma situação específica.

Os gráficos de médias e medianas segundo linhas e colunas permitem explorar simultaneamente a presença de tendência (não-estacionariedade de primeira ordem), e não estacionariedade de segunda ordem, onde a variância e a covariância entre vizinhos não se mantêm constante.

4.4.1.2 LISA (Local Indicator of Spatial Analysis) – Índice de Autocorrelação Local

Os indicadores globais de autocorrelação espacial fornecem apenas um valor como medida de associação espacial para todo o conjunto de dados, contudo, ao se lidar com grande número de áreas é possível que diferentes regimes de associação espacial ocorram e que surjam pontos extremos de autocorrelação espacial, onde a dependência pode ser ou não mais percebida. Dessa forma, é necessário utilizar indicadores de associação espacial que possam estar associados às diferentes localizações de uma variável distribuída espacialmente, e isto é possível através de um indicador local de associação espacial (Local Indicator of Spatial Association – LISA).

São utilizados quando se deseja analisar a contribuição espacial de cada área individualmente (ANSELIN, 1995). Servem também para verificar a hipótese de estacionariedade do processo e a identificação de agrupamentos típicos e pontos atípicos, ou *outliers*. Um LISA, segundo Anselin (1995) é qualquer estatística que satisfaça as seguintes condições: o LISA para cada observação fornece uma indicação de aglomerações significativas ao redor daquela observação; a soma dos LISAs para todas as observações é proporcional ao indicador global de associação espacial.

Por meio deste procedimento, desenvolvido por Luc Anselin a partir do Índice de Moran, é possível se avaliar a hipótese de aleatoriedade espacial comparando os valores do indicador de cada região com os indicadores das regiões vizinhas.

Este indicador mensura a autocorrelação espacial de uma localização específica com os seus vizinhos. Da mesma forma que o *I* de Moran Global, o *I* de Moran Local (LISA) significativamente positivo indica que os valores do local em questão e seus vizinhos são similares, ou seja, quando há autocorrelação positiva, há padrões de similaridade espacial. Já o *I* de Moran Local significativamente negativo indica que o valor da localização em análise é desigual em relação a seus vizinhos, ou seja, há padrões de dissimilaridade espacial. O *I* de Moran Local pode ser calculado através da equação a seguir, para uma localização *i*, e quando os valores de I_i são diferentes de zero indica que a unidade *i* está espacialmente associada aos seus vizinhos.

$$I_i = \frac{x_i - \bar{x}}{S_i^2} \sum_{j=1}^N w_{ij} (x_j - \bar{x}) \text{ onde } S_i^2 = \frac{\sum_{j=1, j \neq i}^N x_j^2}{N-1} - \bar{x}^2 \quad (8)$$

O LISA produz valor específico para cada objeto ou área, o que possibilita a identificação de *clusters* (atributos semelhantes), *outliers* (discrepantes) e a presença de mais de um regime espacial.

De acordo com Anselin (1995), existe uma proporcionalidade direta entre o valor da correlação global e os valores das autocorrelações locais. Ele demonstra que os LISAs permitem a decomposição dos indicadores globais em contribuições individuais, indicando porções territoriais de não estacionariedade e identificando aglomerados (*clusters*) significativos de valores semelhantes em torno de determinadas localizações. Dentre os LISAs mais difundidos estão o Índice Local de Moran (I_i) e os G_i e G_i^* de Geary, sendo que os últimos são utilizados quando existe uma pequena quantidade de vizinhos.

No índice local de Moran a autocorrelação espacial é calculada a partir do produto dos desvios em relação à média como uma medida de covariância, dessa forma, valores significativamente altos indicam altas probabilidades de que haja locais de associação espacial tanto em regiões com altos valores associados como com baixos valores associados. Por outro lado, baixos valores apontam para um padrão que pode ser entendido como locais de comportamento mais errático da variável observada entre um ponto e seus vizinhos (QUEIROZ, 2003).

Se o valor do índice de Moran local for positivo ($I_i > 0$) indica *clusters* de valores similares (alto-alto ou baixo-baixo), e se for negativo ($I_i < 0$), *clusters* de valores distintos (alto-baixo ou baixo-alto). Em resumo, alto-alto e baixo-baixo indicam um padrão de similaridade entre uma unidade e seus vizinhos.

A significância do Índice de Moran Local (Lisa) é realizada da mesma forma que para o Índice de Moran Global (I), avaliada sob a hipótese de normalidade; para cada área é calculado o índice local e em seguida faz-se a permutação aleatória dos valores das demais áreas (ANSELIN, 1995). Para isso, no GeoDa, são geradas até 999 permutações que constituirão a distribuição empírica do índice. Caso o valor calculado – representado na figura 1 pela faixa amarela - esteja na cauda dessa distribuição, constituirá um evento significativo, se não,

indica que os dados estão distribuídos aleatoriamente (figura 2). A hipótese nula destes testes é a não existência de dependência espacial.

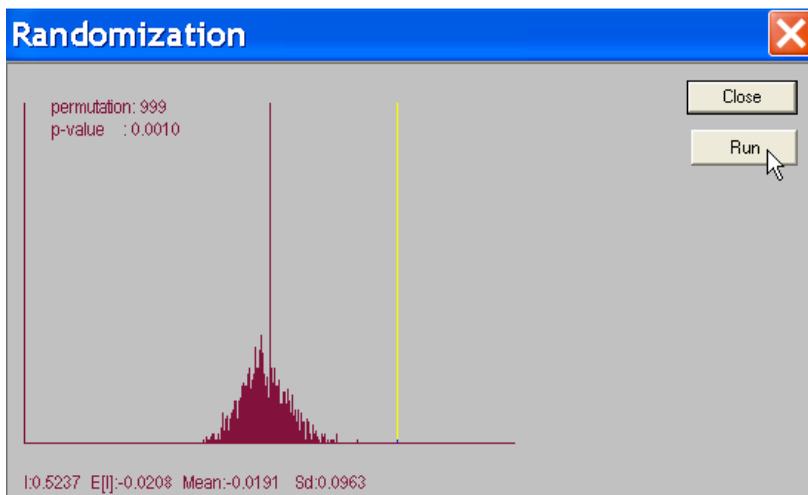


Figura 1 - Distribuição de referência para o I de Moran constatando a não-aleatoriedade. Fonte: ANSELIN, 2003.

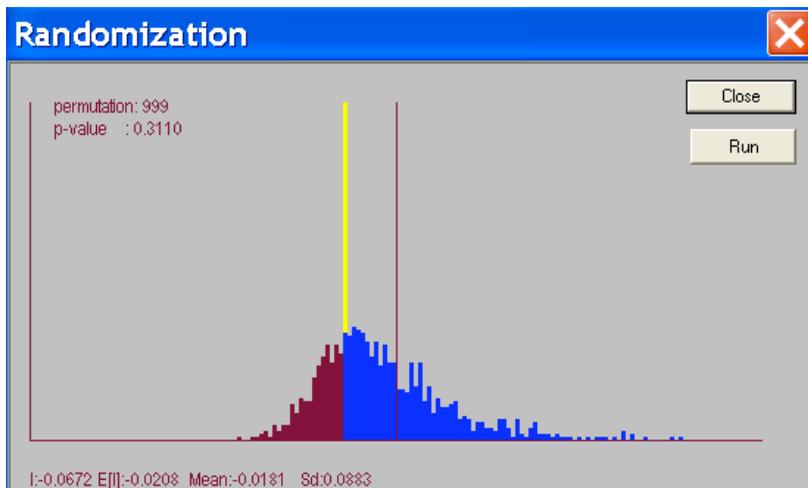


Figura 2 - Distribuição de referência para o I de Moran constatando a aleatoriedade. Fonte: ANSELIN, 2003.

Uma vez determinada a significância estatística deste índice é útil elaborar um mapa indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados. Estas regiões podem ser vistas como bolsões de homogeneidade, no caso, regiões de concentração de valores elevados dos atributos e regiões com valores reduzidos dos atributos. Essas áreas possuem dinâmica espacial própria e merecem análise detalhada.

O Índice de Moran Local (LISA) pode ser visualizado por meio de mapas (LISA cluster map). Tais mapas permitem a identificação de conglomerados de áreas que podem ser caracterizadas por áreas com altas proporções do indicador utilizado e áreas vizinhas com proporções também relativamente altas (quadrante alto-alto) ou conglomerados de áreas com baixas proporções do indicador e vizinhos também com proporções relativamente baixas (quadrantes baixo-baixo). Na geração do LISA MAP, os valores do índice local de Moran são classificados em quatro grupos: não significantes (valor zero na legenda), com significância de 95% (classe 1), 99% (classe 2) e 99,9% (classe 3).

A capacidade de detecção de agrupamentos significativos de amostras de valores próximos em torno de cada localização permite identificar os pontos onde tais características predominam, apontando assim, potenciais territórios desenhados a partir destas características.

4.4.2 Análise gráfica de dependência espacial

4.4.2.1 Diagrama de espalhamento de Moran (Moran Scatterplot Map)

Uma forma adicional para a visualização da dependência espacial é o Diagrama de Espalhamento de Moran que indica os diferentes regimes espaciais presentes nos dados, ou seja, o relacionamento entre os valores de cada variável e o valor da média de seus vizinhos.

O I de Moran (coeficiente de regressão linear) corresponde a tangente α da reta de regressão de z (valores normalizados) e wz (média dos vizinhos, também normalizadas), conforme a figura 3. Construído com base nos valores normalizados permite analisar o comportamento da variabilidade espacial, ou seja, pode-se comparar os valores normalizados de determinado atributo em uma área com a média de seus vizinhos, com um gráfico bidimensional dividido em quatro quadrantes.

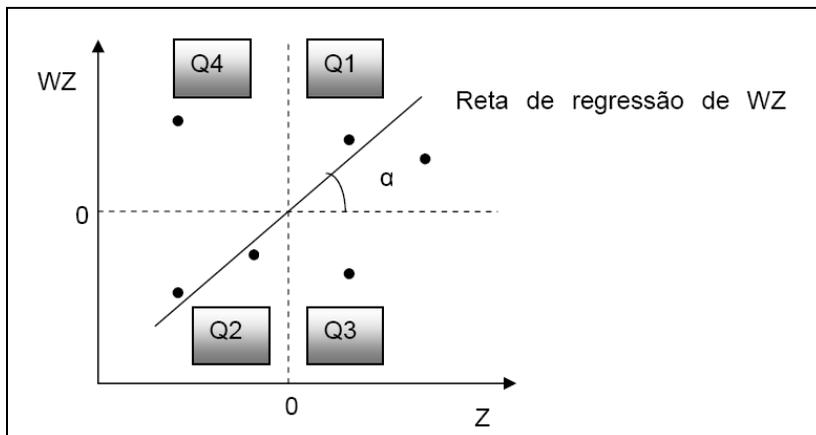


Figura 3 - Gráfico de espalhamento de Moran.

Os valores encontrados nos quadrantes Q1 (desvios positivos e médias locais positivas) e Q2 (desvios negativos e médias locais negativas) correspondem a pontos de associação espacial positiva, cuja localidade possui vizinhos com valores semelhantes, ou seja, clusters espaciais com valores similares. No sentido de uma localização, verifica-se a existência de vizinhos com valores semelhantes.

Nos quadrantes Q3 (desvios positivos e médias locais negativas) e Q4 (desvios negativos e médias locais positivas) estão os pontos de associação espacial negativa onde a região apresenta valores distintos das localidades vizinhas, isto é, existem *clusters* espaciais com valores diferentes entre a localidade e sua vizinhança. Os pontos nesses dois últimos quadrantes marcam regiões de transição entre regimes espaciais distintos e podem ser considerados como extremos, tanto por estarem distantes da reta de regressão, quanto por indicarem regiões que não seguem o mesmo processo de dependência espacial das demais observações (ANSELIN, 1995). No sentido de localização, verifica-se a existência de vizinhos com valores distintos.

Para se identificar valores extremos (*outliers*) pode-se verificar os pontos no diagrama de Moran que são extremos em relação à tendência central refletida pela inclinação da reta de regressão. A outra maneira consiste em localizar os pontos cujos valores estão acima de dois desvios padrões da média. Estes, então, podem ser considerados bolsões de não-estacionaridade. A presença de valores extremos pode ainda significar problemas com a especificação da matriz de proximidade ou

com a escala espacial de observação dos dados. Eles também podem indicar regiões de transição entre regimes espaciais distintos, os quais geralmente pertencem aos quadrantes Q3 e Q4 (QUEIROZ, 2003).

A dependência espacial apresenta maior intensidade à medida que mais pontos se aglomeram no primeiro e no terceiro quadrante, enquanto que a existência de pontos no segundo e quarto quadrante descaracteriza esse fato. Esses últimos sugerem que o fator “espaço” não influencia diretamente na valoração da variável y , caracterizando-se assim, como eventos aleatórios.

Os indicadores globais fornecem um único valor como medida da associação espacial para o conjunto de dados. Quando a área de estudo está muito subdividida, é muito provável que ocorram diferentes regimes de associação espacial e que apareçam máximos locais de autocorrelação espacial, onde a dependência espacial é ainda mais pronunciada.

4.4.2.2 LISA map e Moran map

O Diagrama de Espalhamento de Moran também pode ser visto através de um mapa (quando existir a malha geográfica georreferenciada) *LISA Cluster Map*. Os quatro quadrantes são representados por cores diferentes para facilitar a identificação.

O *LISA Significance Map* é utilizado para localizar, no mapa, os pontos estatisticamente significantes, calculados pelos índices locais. Geralmente usam-se níveis de significância de 99,9%; 99% e 95%. Os indicadores locais servem para indicar a significância estatística da concentração espacial indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados.

Constatada a autocorrelação espacial com as ferramentas descritas anteriormente, o próximo passo é utilizar modelos de regressão que incorporem os efeitos espaciais. Estes modelos podem ser globais ou locais. Os modelos globais utilizam um único parâmetro para capturar a estrutura de correlação espacial, enquanto que os locais utilizam vários parâmetros que variam continuamente no espaço.

4.4.3 Análise exploratória

As técnicas de análise exploratória aplicadas a dados espaciais são essenciais ao desenvolvimento das etapas da modelagem estatística espacial, em geral sensível ao tipo de distribuição, à presença de valores

extremos e à ausência de estacionariedade. As técnicas empregadas são, em geral, adaptações das ferramentas usuais. Além disso, a não estacionariedade do processo espacial na região de estudo também deve ser investigada, nos seus vários aspectos: variação na média (primeira ordem), na variância e na covariância espacial. A forma mais simples e intuitiva de análise exploratória é a visualização de valores extremos nos mapas (CÂMARA et al., 2002).

Um aspecto fundamental da análise exploratória espacial é a caracterização da dependência espacial, mostrando como os valores estão correlacionados no espaço. Neste contexto, as funções utilizadas para estimar quanto o valor observado de um atributo numa região é dependente dos valores desta mesma variável nas localizações vizinhas são a autocorrelação espacial e o variograma. A dependência espacial pode ser medida de diferentes formas. O índice de Moran mede a autocorrelação espacial a partir do produto dos desvios em relação à média, porém há outras medidas que apesar de similares, medem a dependência espacial a partir de outras operações, ou seja, diferença simples como no índice de Geary ou soma simples como nas estatísticas G (ANSELIN, 1992).

5 - METODOLOGIA PARA ANÁLISE DE DADOS

A primeira etapa da análise refere-se à verificação da necessidade ou não de se utilizar um modelo de regressão espacial para a modelagem dos dados. Caso não seja constatada a existência de dependência espacial, um modelo de regressão convencional apresentará os mesmos resultados de se utilizar um modelo de regressão espacial.

A estratégia básica para se encontrar o modelo espacial mais adequado proposta por Anselin (2005), de acordo com fluxograma abaixo, é a seguinte: estimar o modelo por MQO (Mínimos Quadrados Ordinários); diagnosticar dependência espacial de acordo com o teste global (I de Moran) e testes específicos (Multiplicador de Lagrange - LM); em caso de detecção de dependência espacial, reestimar o modelo (aquele com menor p-valor nos testes LM): modelo de erro auto-regressivo espacial ou modelo de defasagem espacial.

Uma vez que os fenômenos da heterogeneidade espacial e da autocorrelação espacial, muitas vezes, não são distinguidos facilmente na prática, costuma-se adotar uma dessas duas abordagens: modela-se inicialmente a heterogeneidade espacial e testa-se para a autocorrelação espacial nos resíduos; ou modela-se inicialmente a autocorrelação espacial e testa-se para a heterogeneidade espacial. Sempre se procura corrigir estes dois efeitos espaciais no modelo final a fim de que se possa obter estimativas não-viesadas e precisas (ALMEIDA e HADDAD, 2004).

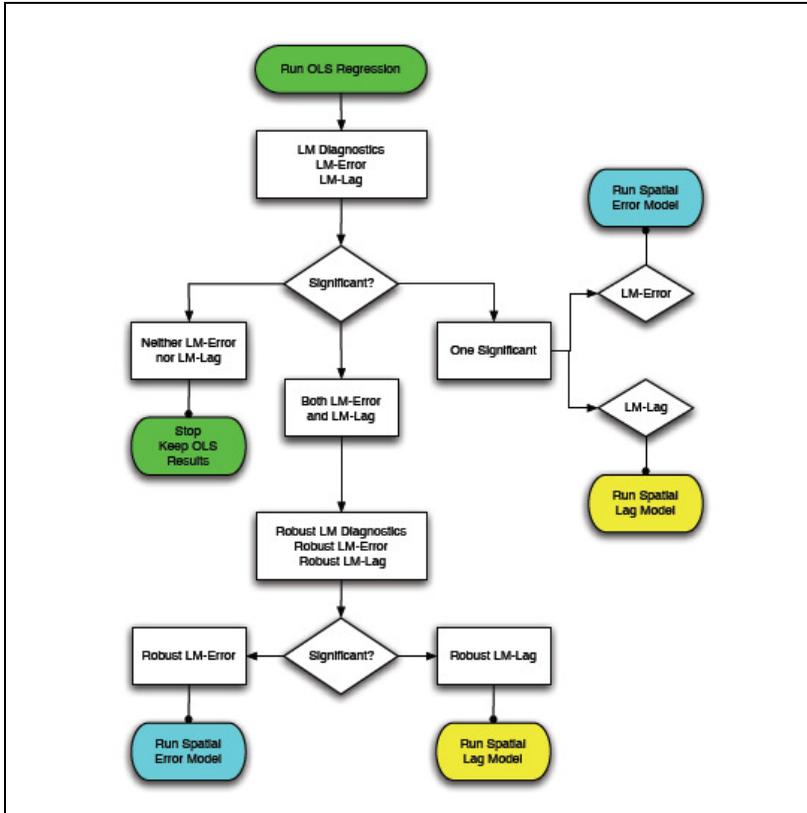


Figura 4 - Fluxograma: processo de decisão para escolha do modelo de regressão. Fonte: ANSELIN, 2005.

5.1 Etapas para verificação da dependência espacial e ajuste do modelo de regressão espacial

- Definição da matriz W : com a definição do objeto do estudo e a partir dos dados, deve-se definir as matrizes de proximidade espacial. É importante definir mais de uma matriz W a fim de refinar e melhor caracterizar a dependência espacial, que pode estar presente tanto na variável dependente quanto no erro aleatório, ou em ambos.

- Teste I de Moran, Moran Scatterplot e Moran Map : para se verificar a dependência espacial pode-se gerar um mapa onde o valor da variável dependente é diferenciado por cores, para se observar a existência de tendência espacial e quantificar a dependência através do *I* de Moran ou *G* de Geary. Ou, ainda, podem ser usadas também as formas gráficas Moran Scatterplot e Moran Map para auxiliar na constatação da dependência espacial.

Se não for constatada a dependência espacial pode-se aplicar o Modelo de Regressão Tradicional (Mínimos Quadrados Ordinários), mas se for constatada a dependência espacial deve-se fazer as verificações para se encontrar o modelo espacial mais adequado.

Depois de verificada a existência de dependência espacial, o passo seguinte é a construção do modelo. O melhor tipo de modelo a ser utilizado (global ou local) depende primeiramente de uma análise a respeito da estacionariedade, onde os indicadores locais podem ser utilizados para tal verificação.

- Estacionariedade, análise de correlações entre variáveis e verificação de relação linear: depois de definido o melhor modelo a ser utilizado, deve ser construída uma matriz de correlações a fim de identificar as variáveis explicativas mais relevantes e as possíveis correlações entre as mesmas. Alta correlação entre as variáveis independentes pode levar a um problema denominado multicolinearidade, o que pode invalidar o modelo.

Outro ponto que deve ser analisado é a relação linear entre as variáveis, o que pode ser facilmente visualizado pelos gráficos de dispersão. Na falta de linearidade, medidas corretivas devem ser aplicadas para garantir tal pressuposto. Também é fundamental verificar os pressupostos básicos da análise de regressão, pois são eles que garantem a validade do modelo. Recapitulando, os pressupostos são normalidade dos resíduos, erros com variância constante (homocedasticidade) e erros não correlacionados.

- Normalidade dos resíduos: os testes, nos modelos de regressão linear, só são válidos quando os erros seguem distribuição normal.

- **Heterocedasticidade:** é uma das hipóteses do modelo de regressão, isto é, a de que a variância teórica do termo de distúrbio aleatório condicional em relação as variáveis independentes seja constante. Quando a variância teórica (não observável) do distúrbio aleatório muda ao longo de diferentes segmentos do intervalo de tempo considerado ou em função de variáveis independentes temos o caso de heterocedasticidade. Neste caso os estimadores de mínimos quadrados deixam de ser estimadores lineares não tendenciosos ótimos e perdem sua eficiência assintótica. Além disso, todos os testes de hipóteses baseados em estatística t, F e qui-quadrado deixam de ser válidos. Hipótese nula: Não há heterocedasticidade.

5.2 Verificação das suposições do modelo, para constatação de necessidade de ações corretivas – validade do modelo

As estatísticas R^2 e $R^2_{ajustado}$ indicam a **qualidade do ajuste**. Quanto mais próximo os valores de R^2 e $R^2_{ajustado}$ estiverem de 1, mais ajustado está o modelo; enquanto que mais próximo de 0, pior é esse ajuste. O valor do $R^2_{ajustado}$ é mais aconselhável do que R^2 por levar em consideração a quantidade de parâmetros do modelo, pois à medida que novas variáveis são inseridas, mais próximo de 1 tende a ser o valor de R^2 (quando existem poucas variáveis essas estatísticas tendem a ser muito próximas). Os modelos sem o intercepto, ou seja, com a reta passando pela origem apresentam maiores valores de R^2 e $R^2_{ajustado}$, no entanto essas medidas perdem o sentido por poderem assumir valores negativos.

A estatística-F verifica se todos os parâmetros do modelo são não significativos, ou seja, verifica se os parâmetros são iguais a zero (conjuntamente). Em outras palavras, ao se rejeitar a hipótese nula pode-se dizer que o modelo existe porque pelo menos um parâmetro é diferente de zero. Por outro lado, a estatística-t indica a significância de cada variável individualmente. Ambas estatísticas são utilizadas para verificar a influência ou não de cada variável independente na variável dependente em estudo.

O critério **Akaike** é uma estatística frequentemente utilizada para a escolha da especificação ótima de uma equação de regressão. O Critério de **Schwarz**, também é uma estatística semelhante ao critério de Akaike, apropriada para se comparar a adequabilidade dos dados aos modelos. Na comparação de modelos, para ambos os critérios, um valor

menor do índice indica melhor ajuste. Ainda para a escolha da melhor especificação, pode-se utilizar o valor do **Log likelihood**, sendo que valores maiores para este índice indicam melhor ajuste.

- O modelo final de Regressão Espacial - *Spatial lag* ou *Spatial error* – é definido, com utilização da matriz de pesos espaciais mais apropriada. Sempre se busca corrigir estes dois efeitos espaciais no modelo final a fim de que se possa obter estimativas não viesadas.

5.3 Delineamento da pesquisa

O uso de econometria para pesquisa empírica envolve uma etapa de especificação que pode vir da experiência pessoal do pesquisador ou ser baseada em princípios teóricos. Depois de especificada e estimada a regressão o passo seguinte consiste em avaliar a qualidade da especificação escolhida sob diversos aspectos.

A partir do resultado dos diversos testes pode parecer interessante rever a especificação inicial, e fazer todos os testes novamente. Em algum momento encontra-se a especificação que resista bem a todos os testes e pareça fazer sentido do ponto de vista da teoria e da experiência do pesquisador. Neste ponto, o processo de pesquisa empírica atingiu o seu objetivo: uma representação empírica exata da relação matemática entre determinadas variáveis.

Os testes para ajuste da equação de regressão partem da definição da “hipótese nula” a ser testada. O resultado do teste é composto pelos valores amostrais de uma ou mais estatísticas do teste e de probabilidades a ela associadas (p-valores).

É necessário verificar, antes de analisar a regressão, se as premissas para que os testes estatísticos sejam eficientes estão presentes, ou seja, se a distribuição é normal, se os resíduos são independentes entre si, com valor esperado igual a zero e variância constante. Isto pode ser verificado através da análise dos resíduos.

A análise gráfica dos resíduos revela se a presunção de normalidade da distribuição de resíduos se confirma. Além disto, pode também indicar se a variância dos resíduos é realmente constante, ou seja, se a dispersão dos dados em torno da reta de regressão é uniforme e ainda se há ou não uma variável não identificada que deve ser incluída no modelo. Revela ainda se a ordem em que os dados foram coletados, como o tempo de observação, tem algum efeito sobre os dados, ou se a

ordem deve ser incorporada como variável no modelo e também se a presunção de que os resíduos não são correlacionados está satisfeita.

Portanto, na checagem das premissas através de gráficos, com a plotagem dos resíduos (eixo y: resíduos), quando os dados tendem às premissas, o gráfico deve mostrar uma faixa horizontal centrada em torno do zero, sem mostrar uma tendência positiva ou negativa.

Quanto ao teste da adequação de modelos, se o gráfico dos resíduos mostrar uma tendência sistemática positiva ou negativa significa que uma outra função (não linear) deve ser escolhida.

Para teste da existência de variáveis esquecidas, deve-se verificar se os resíduos não estão aleatoriamente distribuídos em torno do zero. Além disto, se o gráfico dos resíduos demonstra um padrão quando plotado contra determinada variável, esta variável deve ser incluída no modelo como uma variável independente.

- Normalidade dos resíduos:

Como os resíduos devem apresentar distribuição normal, compara-se a distribuição dos resíduos com a curva normal. Se a distribuição dos resíduos não for normal, as estimativas não serão eficientes e com maior erro padrão. Assim, deve-se incluir novas variáveis no modelo ou formular corretamente a relação funcional. As possíveis causas são: omissão de variáveis explicativas importantes ou formulação matemática (forma funcional) incorreta.

- Heterocedasticidade:

Para o teste da heterocedasticidade, deve-se verificar a igualdade da variância dos resíduos. A variância dos resíduos é indicada pela largura da dispersão dos resíduos, quando o valor de uma variável independente aumenta. Se essa largura aumenta ou diminui quando o valor da variável aumenta, a variância não é constante. Então, os resíduos devem ser aleatórios, sem padrão, próximos da reta zero.

Quando existe heterocedasticidade, o método dos mínimos quadrados não pode ser usado para estimar a regressão, devendo então ser usado um método mais complexo chamado mínimos quadrados geral.

Através da homocedasticidade, ou seja, resíduos com a mesma variância para cada observação, pode-se avaliar o conteúdo informacional dos resíduos. Para identificação da homocedasticidade analisa-se a evolução da dispersão dos resíduos e examina-se a

distribuição dos resíduos para cada observação. Neste caso, a solução é mudar a forma funcional através de transformação das variáveis e estimar a regressão via mínimos quadrados ponderados. Pode ocorrer no caso de diferenças entre os dados da amostra, modelo de aprendizagem, discricionariedade no uso da renda ou erro de especificação.

O modelo pressupõe ausência de correlação, ou seja, que a correlação entre os resíduos é zero, o que significa que o efeito de uma observação é nulo sobre a outra. Se os resíduos forem correlacionados as estimativas não são eficientes e têm maior erro padrão. Possíveis causas da ocorrência de correlação são: viés de especificação, ou seja, falta de variáveis ou forma funcional incorreta ou problemas com o manuseio de dados (interpolação/ extrapolação). Assim, se ocorrer, a solução é formular corretamente a relação funcional.

A multicolinearidade ocorre quando se tem duas ou mais variáveis independentes do modelo explicando o mesmo fenômeno, com variáveis contendo informações similares. Se duas ou mais variáveis independentes forem altamente correlacionadas ocorre dificuldade na separação dos efeitos de cada uma das variáveis e a multicolinearidade tende a distorcer coeficientes estimados. Como consequência os erros padrão são maiores, o modelo tem menor eficiência com estimativas mais imprecisas, os estimadores ficam sensíveis a pequenas variações dos dados e há dificuldade na separação dos efeitos de cada uma das variáveis.

5.3.1 Variáveis

Na especificação de modelos econométricos é importante que se determine quais as variáveis que devem ser incluídas, pois na prática é raro conhecermos as variáveis corretas.

Nos modelos hedônicos para análise do mercado habitacional, as regressões envolvem três tipos de variáveis independentes, ou seja, as físicas ou próprias do imóvel, e as variáveis relativas à localização, que podem ser classificadas como de acessibilidade e da vizinhança. A inclusão de variáveis representando várias características pode ocasionar multicolinearidade, autocorrelação espacial e heterocedasticidade. A autocorrelação espacial pode causar imperfeições na aplicação de modelos hedônicos utilizados na análise do mercado habitacional.

Quando as coordenadas da localização são incluídas junto com um grupo de variáveis independentes numa equação hedônica de unidades residenciais, na avaliação do impacto da localização no preço

do imóvel, o modelo produz uma melhora nas previsões (Bourassa et al., 2007).

As variáveis explicativas utilizadas nos modelos hedônicos de preços das habitações referentes a características físicas da habitação podem se referir ao tamanho do imóvel, como área e número de dormitórios, aos padrões de construção, à idade da construção (que significa há quanto tempo foi construída), e, ainda, à características do condomínio, como a área proporcional de condomínio, o número de andares e de apartamentos por andar, número de elevadores e equipamentos de lazer da área comum.

Variáveis de acessibilidade se referem a distâncias, ao centro da cidade ou a um determinado pólo de valorização ou desvalorização. Estas podem se mostrar significativas em um modelo de regressão por vários motivos, relacionados à composição da amostra ou área de influência real que um pólo possa ter em relação ao conjunto de imóveis da região (PERUZZO TRIVELLONI E HOCHHEIM, 2006).

Com relação às variáveis de qualidade da vizinhança, podem ser considerados fatores econômicos da região, sociais e ambientais. Essas variáveis podem conter erros pois são extraídas de zonas geográficas que podem ter sido delineadas arbitrariamente, ou podem não estar em conformidade com os objetivos do estudo. Além disto, as informações podem não ser abrangentes ou estar desatualizadas.

5.3.2 Dados - Coleta e preparação

Inicialmente foram coletados dados sobre os lançamentos imobiliários correspondentes a edifícios de apartamentos da cidade de São Paulo, vinculados ao endereço de cada empreendimento.

Para cada um dos lançamentos, de acordo com o endereço do condomínio, foram estabelecidas as coordenadas geográficas (latitude e longitude) utilizando-se recursos do Google Earth. Estas coordenadas, por se referirem a distâncias angulares, foram transformadas em distâncias lineares através da aplicação de fator multiplicativo de conversão correspondente a 111 km para a latitude e a 102 km para a longitude, que foram adotados de acordo com o posicionamento geográfico geral da cidade de São Paulo (32°S e 46°W - Fonte: Instituto Geográfico e Cartográfico – IGC).

Os dados georreferenciados foram transformados para o formato shapefile, que é o formato compatível para os testes do programa GeoDa, a partir de um arquivo XLS do Excel.

No caso de condomínios com dois ou mais edifícios, visto que o programa GeoDa não aceita duas entradas de dados iguais em uma determinada coordenada, os dados foram revisados e as coordenadas foram ajustadas, considerando-se que a distância entre edifícios localizados em um mesmo endereço é de 10 metros.

5.3.3 Testes estatísticos

Para obtenção do modelo de regressão linear e para a verificação dos pressupostos deste foram aplicados testes estatísticos do software TS-Sisreg. Foram testadas algumas combinações de variáveis para se verificar quais deveriam ser incluídas nos modelos e a forma destas para linearização dos parâmetros, o que possibilitou que fosse gerado o modelo melhor ajustado aos dados.

Com o objetivo de se detectar a ocorrência de autocorrelação espacial e verificar a segmentação do mercado habitacional de São Paulo foram aplicados os testes do programa GeoDA 0.9.5-i (Anselin, 2004). No caso de dependência espacial significativa, como a análise de regressão padrão deixa de ser apropriada, um dos modelos alternativos que incorporam dependência espacial, modelo de erro ou de defasagem espacial, deve ser aplicado para se obter a modelagem mais adequada.

Quatro testes foram considerados para avaliação da dependência espacial: Multiplicador de Lagrange (LM) e Multiplicador de Lagrange Robusto (RLM) para dependência do tipo defasagem (lag) e erro (error). Foram realizados, além de testes para os preços, outros para os resíduos dos preços, para se verificar a formação de clusters significantes que correspondem a submercados habitacionais.

5.4 Dados do setor imobiliário

Os dados que compõe este estudo foram selecionados a partir de uma base de dados da Embraesp (Empresa Brasileira de Estudos de Patrimônio) com lançamentos imobiliários do município de São Paulo, que incluem unidades residenciais em condomínios verticais, horizontais e também unidades comerciais, de dezembro de 1998 até novembro de 2008, e foram reunidos a partir de uma divulgação on line no site do Secovi de São Paulo, sendo que estavam disponibilizados para acesso nos últimos meses de 2008, quando se deu esta pesquisa.

São utilizados apenas os dados referentes às unidades verticais, ou seja, aos edifícios de apartamentos. De acordo com a Prefeitura de

São Paulo, apartamento é um domicílio particular localizado em edifício de um ou mais pavimentos, com mais de um domicílio, servidos por espaços comuns (hall de entrada, escadas, corredores, portaria ou outras dependências) e, ainda, o domicílio que se localiza em prédio de dois ou mais andares em que as demais unidades são não-residenciais e, também, aqueles localizados em edifícios de dois ou mais pavimentos com entradas independentes.

Considera-se, neste estudo, que decorre um prazo médio de 30 meses para execução dos edifícios, assim, as unidades residenciais em lançamentos que ocorreram a partir de dezembro de 1998 passam a integrar o estoque habitacional após o ano de 2000. Com isso, os dados do último censo (2000) divulgados pela Prefeitura de São Paulo sobre o estoque habitacional (Anexo) compõem este estudo, como uma contextualização e podem ser considerados como os valores do estoque inicial de unidades habitacionais do município.

Os dados reunidos correspondem a 3.590 lançamentos imobiliários, ou seja, condomínios verticais, com um total de 237.255 apartamentos situados em edifícios estritamente residenciais.

Alguns empreendimentos reúnem mais de um lançamento imobiliário. Assim, em um mesmo endereço pode-se verificar a ocorrência de até 9 lançamentos, que neste estudo foram tratados individualmente por possuírem características diferentes. Os 3.590 lançamentos ou condomínios verticais estão distribuído por endereços conforme segue: 1710 endereços com 1 lançamento, 664 endereços com 2 lançamentos, 112 endereços com 3 lançamentos, 37 endereços com 4 lançamentos, 3 endereços com 5 lançamentos, 5 endereços com 6 lançamentos, 2 endereços com 7 lançamentos, 1 endereços com 9 lançamentos.

Nesta base de dados os lançamentos ou condomínios têm até 1152 apartamentos, sendo que estes, na maioria dos casos, estão distribuídos em 1 bloco ou edifício de apartamentos, porém verifica-se a ocorrência de até 18 blocos em um mesmo condomínio, que é o caso do condomínio que tem 1152 apartamentos de 2 dormitórios e 54,81 m² de área útil por unidade.

O preço atribuído ao lançamento é o preço de venda de um dos apartamentos do condomínio; considera-se que é igual para todos os apartamentos de um mesmo condomínio. Apesar de existir a influência de fatores como o número do andares, andar em que o apartamento se situa e posicionamento relativo da unidade no andar, considera-se que os preços são os mesmos para todas as unidades. Cada preço está associado

com variáveis indicando o endereço completo e características físicas da residência.

Entre as variáveis disponibilizadas pela base de dados consultada, algumas foram selecionadas para serem testadas, referentes às características físicas do apartamento: área útil, número de dormitórios; características do condomínio: área total menos área útil, número de unidades por andar; e características da localização: distância à Praça da Sé, distância à escola particular mais próxima e distância à estação de metro mais próxima.

Os lançamentos ocorreram ao longo de dez anos. Para Schnare e Struyk (1976), uma oferta de certos tipos de habitação pode ser basicamente fixa por períodos de tempo relativamente longos. Apesar disso, neste estudo, a data do lançamento (mês) é considerada, pela possibilidade de sua interferência no preço.

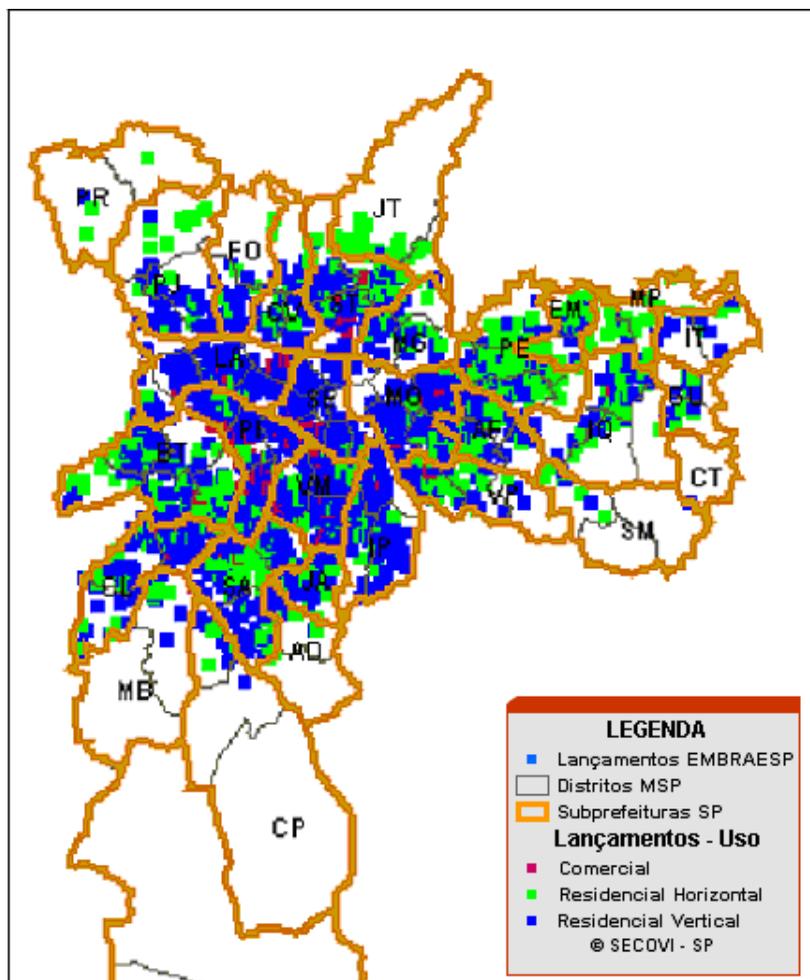


Figura 5 - Mapa do município de São Paulo contendo lançamentos de imóveis

6 - RESULTADOS – ANÁLISE DE DADOS

6.1 Pontos plotados

Os pontos plotados de acordo com a figura 6 abaixo correspondem à localização geográfica dos empreendimentos que se referem apenas aos lançamentos residenciais verticais na cidade de São Paulo, nos últimos dez anos. Cada ponto está relacionado a algumas características do apartamento, ou seja, preço (em R\$), área útil, área de condomínio, além de variáveis como distância à estação do metrô mais próxima e distância à escola particular mais próxima, distância à praça da Sé e data (correspondente ao mês do lançamento). Com estas informações pode-se fazer o teste estatístico do I de Moran e a verificação da ocorrência de autocorrelação espacial entre os imóveis. Pode-se também, à partir destes dados, identificar o modelo de regressão mais apropriado, com a determinação das variáveis significantes e forma destas, para se proceder a verificação da segmentação do mercado.

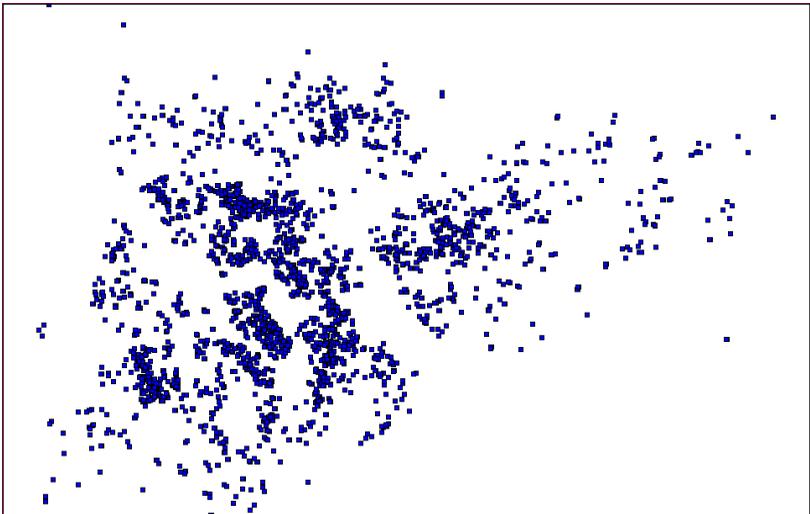


Figura 6 - Pontos plotados correspondentes a lançamentos verticais.

Inicia-se a análise pela verificação da distância mínima na qual todos os imóveis tenham pelo menos um vizinho, considerando-se toda a área do município. O Geoda fornece a distância aproximada de 4,5 km como este valor mínimo. Com isto, todos os condomínios têm pelo menos um vizinho a uma distância de 4,5 km. Isto significa que se for especificado um valor menor que este, aparecem pontos desconectados, também denominados “ilhas”, ou seja, pontos isolados, que não possuem vizinhos.

6.2 Histograma de conectividade

O histograma de conectividade descreve a distribuição das localizações em relação ao número de vizinhos (figura 7). Este gráfico permite que sejam detectadas situações estranhas na distribuição que podem afetar a autocorrelação espacial e a especificação da regressão, como a ocorrência de ilhas ou também uma distribuição bimodal e com algumas localizações com muitos vizinhos e outras com poucos.

Neste histograma, o valor que aparece na parte superior de cada barra representa o número de observações, neste caso, o número de empreendimentos, que possuem uma determinada quantidade de vizinhos (ou o intervalo) indicada na legenda correspondente à cor daquela barra.

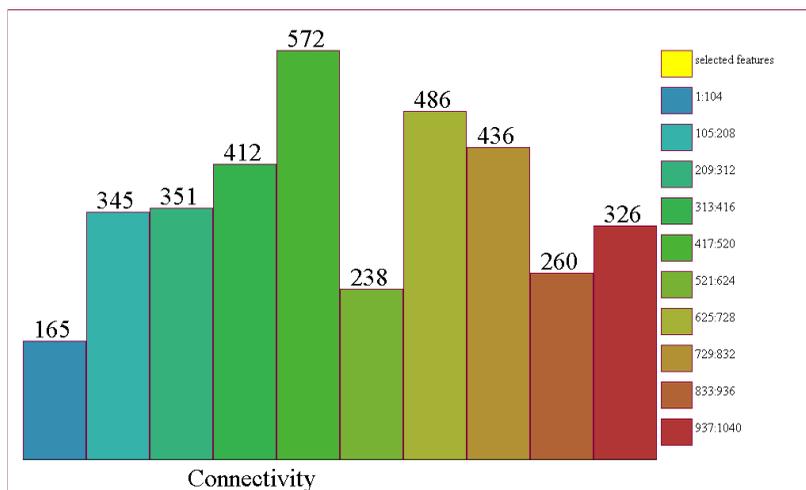


Figura 7 - Histograma de conectividade.

Como se pode verificar na figura acima, existem 326 empreendimentos que possuem cerca de mil empreendimentos (entre 937 e 1040) em sua vizinhança, quando se considera que a vizinhança abrange os imóveis localizados em um raio de 4,5 km; ao passo que, existem 165 empreendimentos que possuem poucos vizinhos, entre 1 e 104, localizados na sua vizinhança, considerando-se também os mesmos 4,5 km de distância.

A localização geográfica dos empreendimentos com um determinado número de vizinhos (ou intervalo) correspondente a cada barra do histograma pode ser visualizada, conforme se verifica na figura 8. Estão destacados (marcados em amarelo) os empreendimentos que possuem poucos vizinhos, ou seja, menos de 104 vizinhos. Estes se localizam nas regiões periféricas do município.

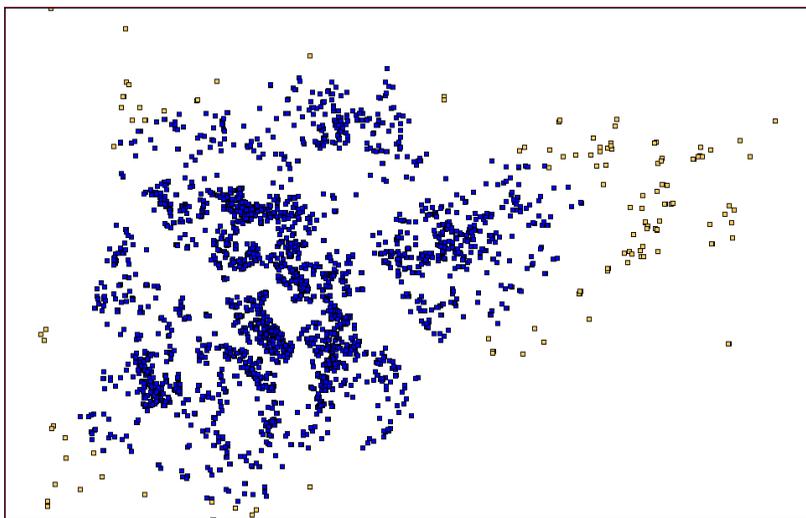


Figura 8 - Pontos com 1 a 104 vizinhos (em amarelo).

Como o Índice Global de Moran é uma medida da autocorrelação espacial para toda a área, deve-se identificar a distância mais apropriada para a matriz de pesos espaciais que define a vizinhança, adequada a configuração geral da área. Isto pode ser obtido calculando-se o índice de Moran para distâncias menores do que aquela que promove a configuração “sem ilhas”.

6.3 Matriz de pesos espaciais baseada na distância entre pontos

Para o ajuste da distância que melhor representa o padrão geral de vizinhança do município de São Paulo foram feitos testes no GeoDa. A matriz de pesos baseada na distância considera que cada par de vizinhos tem um número de identificação da origem, um do destino e a distância entre os dois pontos.

Baumont et al. (2002) propõem que o ajuste da distância deve ser realizado de modo que se encontre a matriz de pesos que melhor represente o padrão geral de vizinhança, que, para eles, é aquela que corresponde ao maior valor do Índice de Moran. Seguindo o procedimento proposto pelos autores, foram testadas algumas matrizes e os resultados obtidos estão apresentados abaixo (tabela 1).

Tabela 1 - Autocorrelação espacial de acordo com a matriz de pesos para ajuste da distância

PREÇO		P-VALUE= 0,0010		
Distância	I de Moran	E(I)	Mean	Sd
4,527 km	0,1094	-0,0003	-0,0002	0,0013
3 km	0,1498	-0,0003	-0,0003	0,0020
2 km	0,2031	-0,0003	-0,0003	0,0039
1 km	0,3295	-0,0003	-0,0001	0,0053
0,50 km	0,4646	-0,0003	0,0002	0,0089
0,20 km	0,5351	-0,0003	0,0005	0,0033
0,15 km	0,5389	-0,0003	-0,0001	0,0146
0,10 km	0,5130	-0,0003	-0,0006	0,0156

Verifica-se que a distância limite mais apropriada para o cálculo da matriz de pesos espaciais, utilizando-se a variável Preço, é 0,15 km, com um Índice de Moran Global de 0,5389.

6.4 I de Moran para a variável Preço

O Índice de Moran é formalmente equivalente ao coeficiente de regressão e a matriz de pesos espaciais é que possibilita seu cálculo. Conforme se determinou acima, a distância de 0,15 km é a melhor

ajustada para o cálculo do I de Moran. É verificada a presença de autocorrelação espacial global, com um grau de associação entre os valores observados e a média ponderada dos valores da vizinhança de 0,5389, o que indica que o padrão de distribuição espacial apresentado pelos dados não é aleatório.

O valor de 0,5389 do índice de Moran, por ser positivo, indica a existência de autocorrelação positiva, que é uma associação de similaridade entre os valores da variável estudada com as que estão na localização espacial da mesma. Desta forma, como se pode verificar no gráfico 1, empreendimentos com valores altos tendem a estar localizados na vizinhança do que também possuem valores altos e, os de valores baixos estão cercados dos que têm valores baixos.

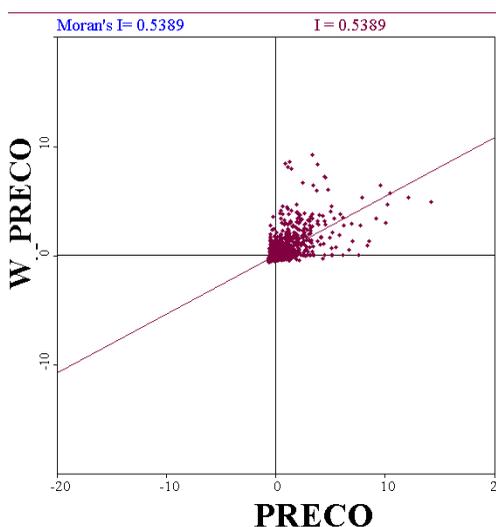


Gráfico 1 - Preço do lançamento versus média de preços dos lançamentos vizinhos localizados dentro de um raio de 150m.

6.5 Teste de significância do Índice de Moran

Para se examinar a significância estatística de um valor do Índice de Moran, a forma mais usual é através de um teste de pseudo-significância onde são geradas permutações com os valores redistribuídos pela área de estudo e o I de Moran é recalculado com os valores da variável recombinados.

Na randomização todos os dados são distribuídos aleatoriamente pela área estudada, o que é realizado 999 vezes (figura 9). O valor do I de Moran da distribuição correspondente à situação observada (representada pela marca amarela na figura) se situa em um dos extremos da distribuição simulada, ou seja, das distribuições criadas aleatoriamente, o que significa que a distribuição observada tem uma situação específica e não provém de uma situação aleatória.

O valor do I de Moran indica a existência de dependência espacial, e para 999 permutações apresentou valor esperado $E(I)$ igual a $-0,0003$ e uma probabilidade de significância de $0,0010$ (P-value), ou seja rejeita-se a hipótese nula, de independência espacial, ao nível de significância de 10%. Com este teste do GeoDa pode-se descartar a aleatoriedade da distribuição e se confirmar a autocorrelação espacial.

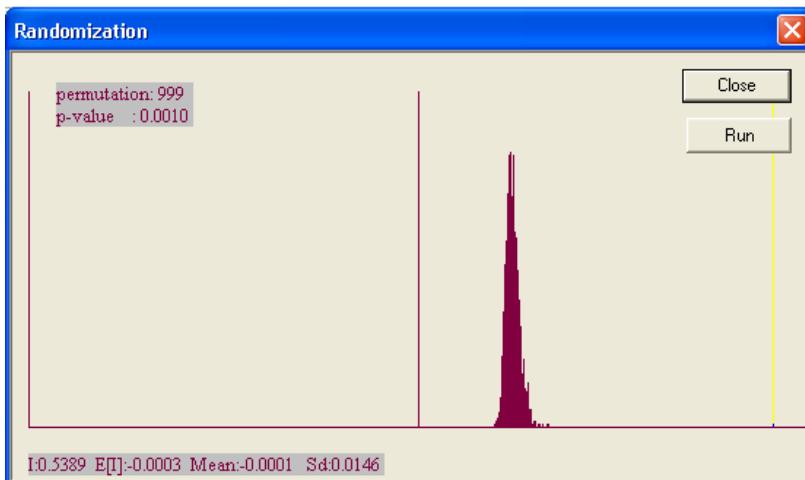


Figura 9 - Randomização dos dados.

6.6 Autocorrelação espacial local – LISA

O I de Moran é uma medida global e não revela a estrutura regional de autocorrelação espacial. O Índice Local de Associação Espacial permite a identificação de agrupamentos (clusters) com valores dos atributos semelhantes, objetos discrepantes (outliers) e de mais de um regime espacial.

Como ocorre no indicador global, a existência de autocorrelação positiva resulta numa relação espacial do tipo alto-alto (high-high). Quando não há autocorrelação espacial, os valores não estão associados espacialmente a seus vizinhos.

Existem duas ferramentas para se verificar essa estrutura que são o diagrama de dispersão de Moran e os indicadores locais de associação espacial – LISA.

6.7 Gráfico de Moran

Pode-se visualizar melhor a associação espacial entre o valor da variável estudada (Preço) com a média dos valores de seus vizinhos (W_Preço) através do gráfico de espalhamento de Moran (Anselin, 1996). No eixo das abscissas do gráfico de Moran entra o valor padronizado de uma variável e no das ordenadas, o valor padronizado dos vizinhos destas unidades. Caso não haja qualquer autocorrelação espacial, as observações ficam bem distribuídas pelos quatro quadrantes.

No gráfico 2, do Preço em relação ao Preço dos vizinhos (W_Preço), verifica-se no primeiro quadrante as observações com Preços acima da média com vizinhos também acima da média (high-high).

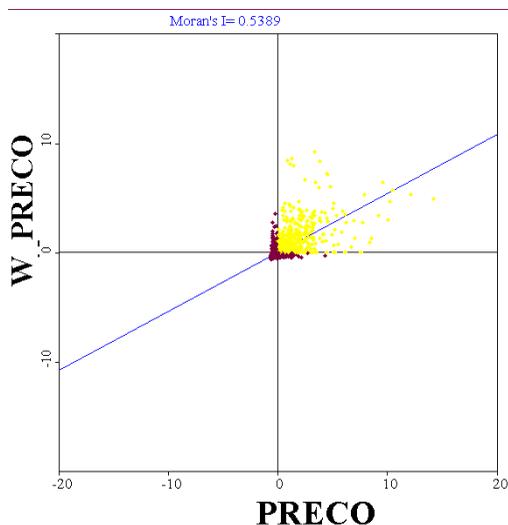


Gráfico 2 - Preço em relação ao Preço dos vizinhos (alto-alto).

Na figura abaixo estão indicados (em amarelo) os imóveis cujos preços foram marcados no primeiro quadrante do gráfico 2.

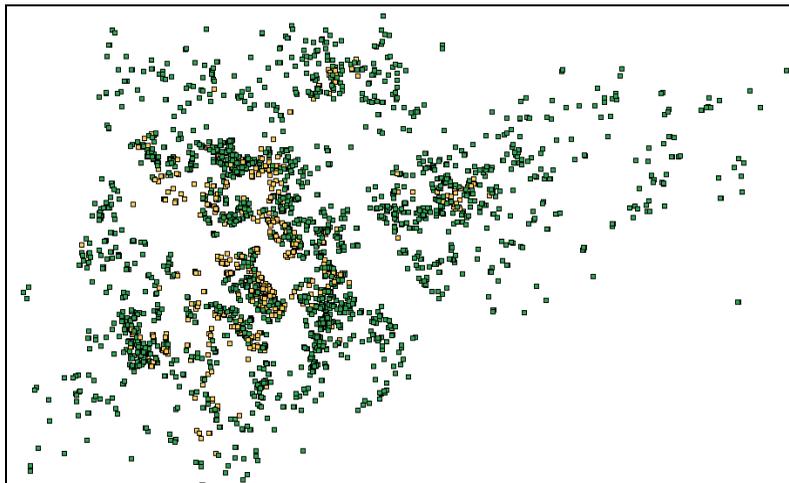


Figura 10 - Preço em relação ao Preço dos vizinhos (alto-alto).

As observações com preços abaixo da média que possuem vizinhos também estão abaixo ocupam o terceiro quadrante (low-low) do gráfico, conforme se pode ver no gráfico 3.

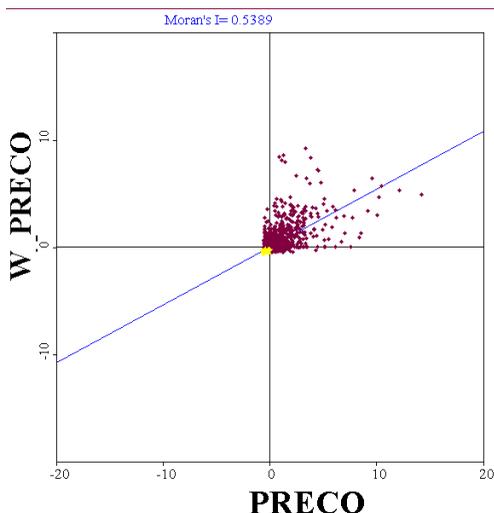


Gráfico 3 - Preço em relação ao Preço dos vizinhos (baixo-baixo).

Na figura abaixo estão indicados (em amarelo) os imóveis cujos preços foram marcados no terceiro quadrante do gráfico 3.

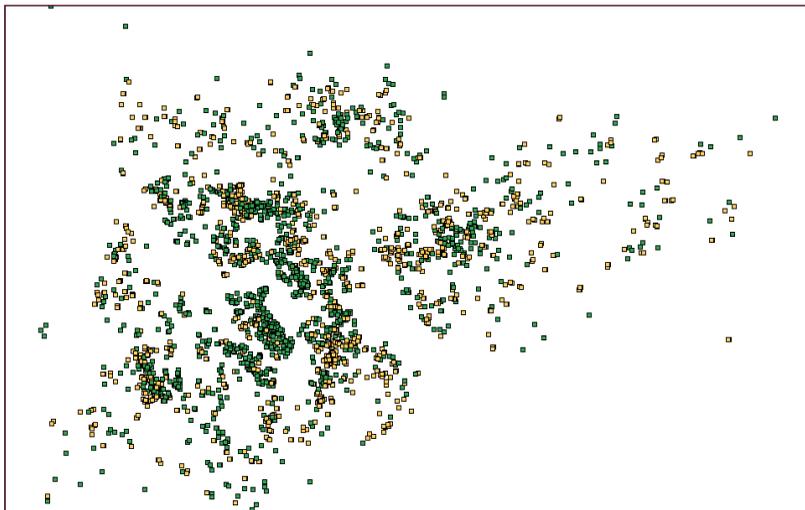


Figura 11 - Preço em relação ao Preço dos vizinhos (baixo-baixo)

Pode-se verificar a existência de padrões de similaridade, sendo que os do tipo alto-alto (FIGURA 10) estão localizados nas regiões mais próximas ao centro do município, enquanto que os padrões baixo-baixo estão espalhados por todo o território (FIGURA 11).

6.8 Regressão: verificação dos pressupostos e escolha de variáveis a serem incluídas no modelo

Foi feita a plotagem dos resíduos contra as variáveis, com o objetivo de se verificar a existência de padrões de comportamento e avaliação das variáveis a ser incluídas no modelo, além de testes para checagem das premissas do modelo. A normalidade do resíduo foi constatada, conforme tabela 2.

Tabela 2 - Normalidade dos resíduos

NORMALIDADE DOS RESÍDUOS (%)		
-1 a +1	-1,64 a +1,64	-1,96 a +1,96
70	88	93
Padrão	68	90
	95	

Foram eliminados os outliers acima de três desvios e o número de dados foi reduzido para 3529. Este procedimento para o tratamento dos dados pode ser adotado devido ao tamanho da amostra.

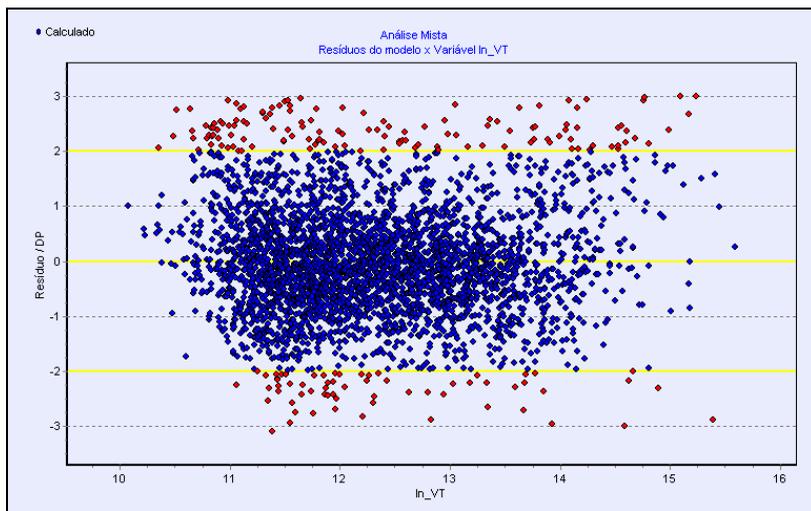


Gráfico 4 - Resíduos do modelo

As variáveis incluídas no modelo foram: área útil, área proporcional de condomínio (correspondente à área total menos área útil), distância à Praça da Sé e data (que relaciona os meses em que os empreendimentos foram lançados). Foram testadas, também, outras variáveis, porém não representaram melhora do modelo ou se mostraram incoerentes e, assim, não foram incluídas, como o caso da distância ao metrô, distância à escola particular mais próxima e o número de unidades por andar.

Ainda com referência ao ajuste da forma funcional, algumas variáveis foram transformadas através da função logarítmica. Foi também necessária a modificação de uma variável para evitar a ocorrência de multicolinearidade, ou seja, a área de condomínio foi utilizada na forma de índice: área de condomínio/área útil.

O modelo de regressão tradicional apresentou os resultados de acordo com a tabela 3.

Tabela 3 - Modelo de regressão tradicional: Mínimos Quadrados Ordinários

REGRESSION				
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION				
Data set	: São Paulo			
Dependent Variable	: ln(PREÇO R\$)	Number of Observations	: 3529	
Mean dependent var	: 12.3199	Number of Variables	: 5	
S.D. dependent var	: 0.975774	Degrees of Freedom	: 3524	
R-squared	: 0.950820	F-statistic	: 17032.9	
Adjusted R-squared	: 0.950764	Prob(F-statistic)	: 0	
Sum squared residual	: 165.248	Log likelihood	: 394.27	
Sigma-square	: 0.0468921	Akaike info criterion	: -778.54	
S.E. of regression	: 0.216546	Schwarz criterion	: -747.696	
Sigma-square ML	: 0.0468257			
S.E of regression ML	: 0.216392			
Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANTE	5.980424	0.03043444	196.5018	0.0000000
LN(ÁREA ÚTIL)	1.378944	0.006123853	225.1759	0.0000000
DISTÂNCIA À SE	-0.03414735	0.0009285055	-36.77668	0.0000000
LN(ÁREA COND/UT)	0.2017768	0.01164798	17.32291	0.0000000
DATA LANÇAMENTO	0.006393586	0.0001062571	60.17092	0.0000000

O R^2 ajustado do modelo é igual a 0.950820, o que indica um ajuste bom.

A seguir, foi feita uma verificação para se detectar a ocorrência de dependência espacial.

6.9 Dependência espacial

Para o diagnóstico de dependência espacial, de acordo com o procedimento indicado por Anselin (1988), deve-se inicialmente testar o modelo tradicional dos Mínimos Quadrados Ordinários. Constatando-se sua ocorrência, deve-se verificar, pelas estatísticas de Lagrange, qual o modelo espacial mais adequado.

Foram calculadas diversas matrizes de pesos espaciais e a que apresentou o melhor resultado foi a composta por vizinhos localizados dentro de um raio de 1 km (W1000).

Tabela 4 - Dependência espacial.

DIAGNOSTICS FOR SPATIAL DEPENDENCE			
FOR WEIGHT MATRIX : W1000.GWT (row-standardized weights)			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.395741	71.6719317	0.0000000
Lagrange Multiplier (lag)	1	415.6427138	0.0000000
Robust LM (lag)	1	208.8708881	0.0000000
Lagrange Multiplier (error)	1	5045.1027056	0.0000000
Robust LM (error)	1	4838.3308798	0.0000000
Lagrange Multiplier (SARMA)	2	5253.9735936	0.0000000

Verifica-se, na tabela 4, pelo valor do *Robust LM (error)* de 4838.3308798, com probabilidade associada de 0,0000000, a ocorrência de heterogeneidade espacial, o que indica que o modelo de erro espacial é o mais adequado aos dados.

Segue a aplicação do teste de regressão do modelo de erro espacial, com matriz de pesos espaciais W1000.

Tabela 5 - Modelo de regressão espacial (*Spatial error*).

REGRESSION				
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION				
Data set	: ln(PREÇO R\$)	Spatial Weight	: W1000.GWT	
Dependent Variable	: LN_VT	Number of Observations:	: 3529	
Mean dependent var	: 12.319864	Number of Variables	: 5	
S.D. dependent var	: 0.975774	Degree of Freedom	: 3524	
Lag coeff. (Lambda)	: 0.865423			
R-squared	: 0.974551	R-squared (BUSE)	: -	
Sq. Correlation	: -	Log likelihood	: 1415.995122	
Sigma-square	: 0.024231	Akaike info criterion	: -2821.99	
S.E of regression	: 0.155662	Schwarz criterion	: -2791.146395	

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANTE	6.637256	0.0431345	153.8735	0.0000000
LN (ÁREA ÚTIL)	1.23082	0.005479824	224.6094	0.0000000
DISTÂNCIA À SE	-0.0390706	0.002944636	-13.2684	0.0000000
LN (ÁREA COND/UT)	0.1652627	0.008969761	18.42443	0.0000000
DATA LANÇAMENTO	0.006812857	7.865922e-005	86.61232	0.0000000
LAMBDA	0.865423	0.01081361	80.03088	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE				
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : W1000.GWT				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	2043.45	0.0000000	

O R^2 no modelo *error* é 0.974551, ou seja, superior ao valor obtido na aplicação do modelo de regressão tradicional. Porém, em modelos de regressão espacial este índice não é a medida mais apropriada para que se verifique o ajuste. De acordo com Anselin (2004), o valor obtido corresponde a um pseudo R^2 que não pode ser comparado diretamente com o R^2 de modelos clássicos. Nestes casos, é mais conveniente a avaliação por outros critérios como o Critério Akaike, Critério Schwarz e Log likelihood. .

O Critério Akaike passou de -778.54 do modelo clássico para -2821.99 no modelo espacial e o Critério Schwarz passou de -747.696 do modelo clássico para -2791.146395 no modelo espacial; ambos diminuíram em relação ao modelo clássico, o que indica uma melhora no ajuste da especificação do modelo espacial. O *Log likelihood* cresceu de 394.27 do modelo clássico para 1415.995122 no modelo de erro espacial, confirmando a melhora do ajuste pelo acréscimo da variável defasagem espacial do resíduo como variável independente.

Os coeficientes das variáveis independentes permanecem altamente significantes ($P < 0,0000000$), e como o λ também é altamente significativo, pode-se considerar que o poder explicativo do modelo está ligado a valores da vizinhança.

No modelo de erro espacial (*spatial error*) tem-se um número mais limitado de diagnósticos que no modelo clássico. O teste Likelihood Ratio para a autocorrelação espacial do resíduo compara com o modelo nulo, ou seja, a regressão clássica com o modelo espacial. Pelo valor da probabilidade do Likelihood Ratio (0.0000000), pode-se ter a confirmação da grande significância do coeficiente espacial autorregressivo.

6.10 Análise dos resíduos para a verificação de clusters e segmentação

Os parâmetros das equações hedônicas dos preços das residências são geralmente estimados com a utilização dos Mínimos Quadrados Ordinários. Neste procedimento assume-se que os resíduos são independentes e identicamente distribuídos com média zero, variância constante e covariância zero. Quando os resíduos são autocorrelacionados, a suposição de covariância zero é violada e os Mínimos Quadrados Ordinários se tornam ineficientes para estimar os parâmetros. Neste caso, passa a ser mais adequado que estes sejam obtidos com a modelagem que considera a autocorrelação espacial.

O gráfico abaixo mostra a autocorrelação dos resíduos.

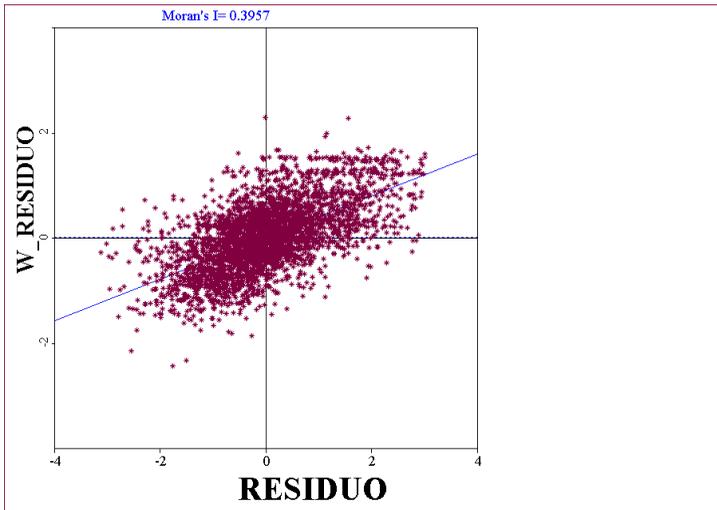


Gráfico 5 - Resíduo do preço de um lançamento versus média do resíduo do preço dos lançamentos vizinhos dentro de um raio de 1000m.

O I de Moran para o resíduo do preço é igual a 0.395741 (gráfico 4), sendo que o valor é significativo ao nível de significância de 10%.

Como os resíduos dos preços estão autocorrelacionados e possuem significância estatística, o que mostra a segmentação do mercado habitacional do município de São Paulo.

6.11 Mapa LISA – Clusters e significância

O mapa LISA permite que se verifique a ocorrência dos padrões de cada localidade. Os clusters espaciais verificados no mapa LISA, mostram como ocorre a segmentação do mercado habitacional do Município de São Paulo.

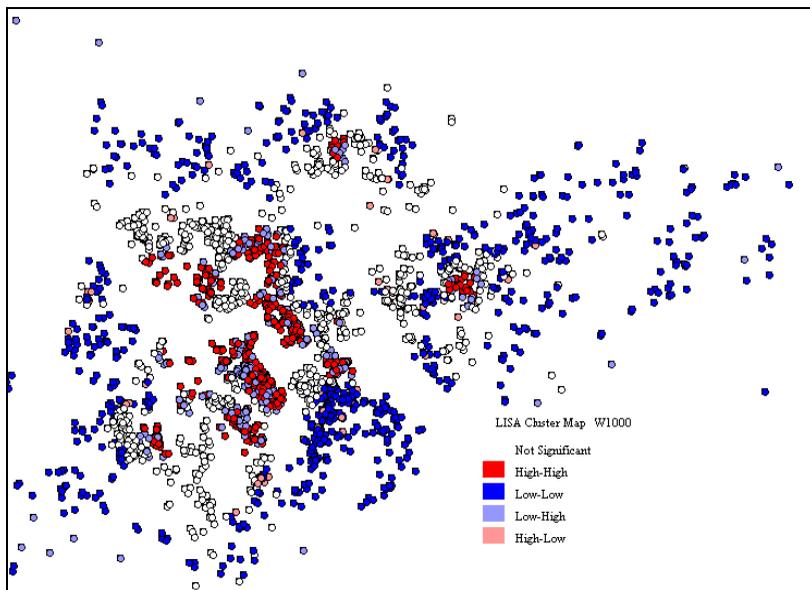


Figura 12 - Mapa LISA clusters.

Pode-se visualizar, na figura acima, os clusters onde prevalecem apartamentos com preços altos, ou seja, acima da média (destacados em vermelho), localizados principalmente na região centro-sul do município. Ocorrem também clusters ou grupamentos de imóveis com preços baixos (destacados em azul escuro), sendo que estes se localizam especialmente na região periférica, mas, na zona norte da cidade destaca-se um cluster de preços altos, e na zona leste um outro.

Os pontos que estão em branco na figura acima, correspondem aos dados situados em locais onde não se verifica um padrão específico. No mapa, todos os pontos coloridos correspondem a dados significantes, e o nível de significância destes pode ser observado na figura 13.

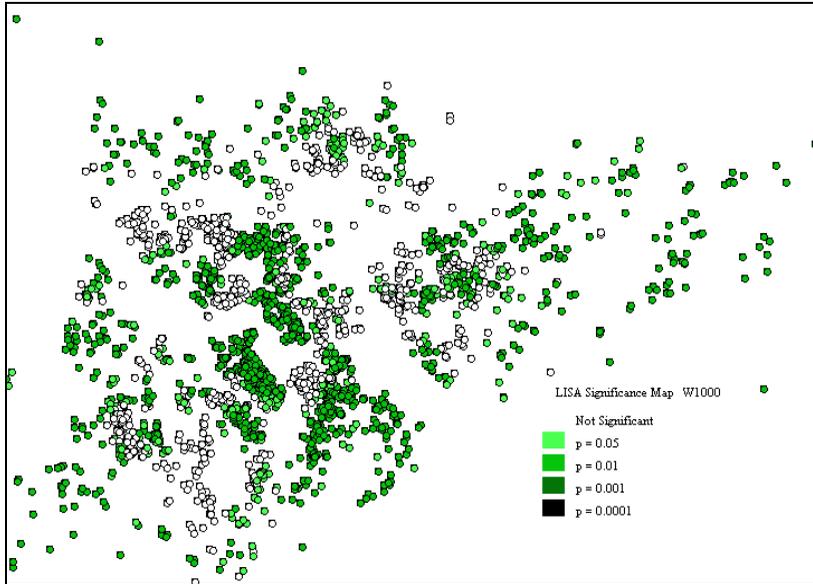


Figura 13 - Mapa LISA de significância

A figura acima mostra a significância estatística que se verifica tanto no caso dos clusters de imóveis com preços altos como nos de preços baixos, o que valida a distribuição espacial da figura 12.

7 - CONCLUSÃO

Nesta pesquisa propôs-se analisar os lançamentos de empreendimentos do município de São Paulo, de forma a se obter um modelo para identificar a relação existente entre o preço do empreendimento e sua localização. Foi feita uma ampla revisão de textos internacionais sobre segmentação do mercado habitacional e sobre econometria espacial. Foram realizados testes que permitiram que se verificasse a autocorrelação espacial dos edifícios estudados, a definição das variáveis mais relevantes no preço do imóvel e a constatação da segmentação do mercado habitacional.

Pela análise dos dados dos preços dos empreendimentos referentes aos lançamentos de edifícios de apartamentos do município de São Paulo nos últimos dez anos, pode-se concluir que existe autocorrelação espacial global. O valor positivo do índice global de Moran indica que a autocorrelação espacial é positiva, ou seja, o valor do preço de um apartamento tende a ser semelhante aos valores dos seus vizinhos. A validade estatística do índice, verificada pelo software GeoDa, que disponibiliza o teste de pseudosignificância (técnica de permutação), pode ser verificada pois o valor do índice global de Moran encontrado para os dados corresponde ao extremo da distribuição simulada, tratando-se assim, de evento não aleatório e com significância estatística.

A autocorrelação espacial dos preços verificada demonstra que o preço de um apartamento em uma região é dependente dos preços dos apartamentos localizados na vizinhança, em uma área abrangida por um raio de 150m. Esta distância foi determinada através da matriz de pesos espaciais para qual os dados mostravam o maior valor para a estatística de Moran referente à variável dependente do estudo, ou seja, o preço dos apartamentos. Isto significa que se a média do preço dos imóveis localizados na vizinhança tem influência significativa sobre o preço daquele imóvel, os modelos de regressão devem incorporar efeitos espaciais.

A consequência prática é que, na estimação do valor de mercado de um lançamento imobiliário, pode-se ajustar o preço e as características do imóvel de acordo com dados da oferta da localidade. Distâncias maiores mostraram autocorrelação menor, o que significa

menor padrão de similaridade das unidades. Distâncias menores, também mostraram autocorrelação menor, o que pode significar perda de informações por se deixar de incluir todos os imóveis relevantes, que são os situados na área delimitada através do cálculo da melhor distância.

Para a modelagem da equação dos preços foram utilizadas variáveis para representar atributos dos apartamentos, ou seja, a área útil (correspondente ao tamanho); para se considerar o condomínio, a área proporcional de condomínio; referente à localização espacial, a distância à Praça da Sé; e também a data em que ocorreu o lançamento. Obteve-se uma boa explicação do preço, com um R^2 igual a 0.950820 (R^2 ajustado 0.950764).

Algumas variáveis testadas foram desconsideradas por motivos diversificados: o número de dormitórios porque ocorreu multicolinearidade (duas variáveis explicando o tamanho do apartamento); a distância à escola mais próxima porque não melhorou o modelo, possivelmente a causa é que os dados incluem desde pré-escola até faculdade, sem considerar as escolas públicas; a distância à estação de metrô mais próxima, porque além de considerar estações previstas junto com as operantes, os problemas ligados ao entorno da estação podem ter contribuído para que a variável não melhorasse o modelo.

Verificou-se que a área útil dos apartamentos é uma variável importante na determinação do preço dos lançamentos. É provável que nesta variável estejam contidas informações sobre o padrão dos imóveis no modelo encontrado.

O modelo espacial melhorou ainda mais a explicação dos preços em relação ao modelo tradicional dos Mínimos Quadrados Ordinários, o que se pode verificar pela comparação entre valores da probabilidade log, do Critério Akaike e do Critério Schwarz e o R^2 encontrado para o modelo de erro foi 0.974551. O modelo espacial que melhor se ajustou foi o do tipo *spatial error*, que sugere a ocorrência de heterogeneidade espacial, e a matriz de pesos espaciais mais adequada foi a correspondente à distância de 1 km.

A heterogeneidade se refere à unicidade de cada lugar, ou seja, à diferenciação espacial ou regional. Cada localização é possuidora de características únicas, assim, a heterogeneidade está associada a uma ausência de estabilidade do comportamento das variáveis sobre o espaço, e os parâmetros que caracterizam as unidades variam com a localização. A heterogeneidade espacial diz respeito a aspectos da estrutura socioeconômica do espaço geográfico (Anselin, 1988). Diferentes consumidores têm preferências diferentes em relação à

moradia e face a grande diversidade de alternativas de residências não se pode considerar o mercado como homogêneo, observando-se a ocorrência de submercados.

Porém, existe a possibilidade de os dados conterem erros de medida, ocasionados por dados agrupados de forma não condizente com o propósito da investigação. De acordo com Gillen et al. (2001), os resíduos dos preços podem estar espacialmente autocorrelacionados porque as externalidades em uma região influenciam o valor de mercado das residências vizinhas da mesma maneira. Mas, é possível que isto ocorra pela dificuldade em se obter informações sobre a estrutura e características de um local, e, conseqüentemente, algumas variáveis ficam omitidas nos estudos empíricos sobre o preço das habitações. Além disto, as informações socioeconômicas e demográficas, que são obtidas pelo censo, realizado a cada dez anos, podem não estar atualizadas o suficiente, para determinados propósitos. Quando equações hedônicas são utilizadas nos modelos de preços das residências, o resíduo conterá informações destas características que não foram observadas. Mesmo em uma situação ideal onde todas as características estejam disponibilizadas, é difícil se ajustar um modelo especificamente correto, pela quantificação da influência de certas variáveis (que é uma dificuldade que ocorre em algumas situações como quando se quer determinar corretamente o quanto a qualidade das escolas influencia o preço das residências).

Não se pode deixar de citar, também, a existência de outros fenômenos que influem na distribuição das habitações no território municipal, como é o caso, no município de São Paulo, da influência de fatores históricos, mencionado por Haddad (1987).

Apesar das dificuldades e limitações, Malinvaud (apud Gujarati, 2004) cita que: “A arte do econometrista consiste em achar um conjunto de hipóteses que sejam tanto suficientemente específicas quanto realistas, para lhe permitir tirar o máximo proveito possível dos dados à sua disposição”.

Desta forma, através do LISA, que é um teste para a verificação da autocorrelação espacial local e que identifica localidades com de padrões distintos do conjunto, obteve-se a indicação de *clusters* espaciais significativos, de valores similares ao redor das observações, o que pode ser verificado no mapa de clusters. Além disto, pelo mapa LISA de significância, pode-se constatar que as mesmas observações destacadas são significantes.

Através dos resultados obtidos ao longo do presente trabalho pode-se concluir que o mercado habitacional do município de São Paulo

é segmentado. Esta informação contribui para o melhor entendimento do mercado habitacional da cidade, fornecendo subsídios para os diversos agentes do setor. Portanto, o objetivo da pesquisa foi atingido e, apesar da grande quantidade de dados para serem analisados, com a aplicação do GeoDa, os resultados puderam ser validados com êxito.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, Eduardo S.; HADDAD, Eduardo A. MEECA: Um modelo Econométrico Espacial para projeção consistente de culturas agropecuárias. Revista de Economia e Sociologia Rural. Vol 42 (3). 2004.

ALONSO, William. Location and Land Use: Towards a General Theory of Land Rent. Harvard University Press. Massachusetts. 1965.

ANSELIN, L. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht, Netherlands. 1988.

ANSELIN, L. Local Indicators of spatial Association – LISA. Geographical analysis Vol 27, No. 2. 93-115. 1995.

ANSELIN, L. The Moran Scatterplot as an ESDA tool to assess local instability in spatial association. Spatial Analytical perspectives in GIS. Taylor and Francis, London. P. 111-125. 1996.

ANSELIN, Luc. Under the hood. Issues in the specification and interpretation of spatial regression models. Agricultural Economics. Vol 27, 247-267. 2002

ANSELIN, Luc.. GeoDa™ 0.9 User's Guide. Spatial Analysis Laboratory. Department of Agricultural and Consumer Economics. University of Illinois, 2003

ANSELIN, Luc. Exploring Spatial Data with GeoDa: A Workbook. Center for Spatially Integrated Social Science . 2004. Disponível para downloading no site <https://www.geoda.uiuc.edu/pdf/geodaworkboob.pdf>

ANSELIN, Luc. Exploring Spatial Data with GeoDa: A Workbook. Spatial Analysis Laboratory. Department of Geography. University of Illinois, Urbana. Revised Version, 2005.

ANSELIN, L., Bera, A., Florax, R., and Yoon, M. Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*. 1996.

ANSELIN, L ; Bera, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A.; Giles, D. (Ed.) *Handbook of applied economic statistics*. Giles: Marcel Dekker, 1998

ANSELIN, Luc; LOZANO-GRACIA, Nancy. *Spatial Hedonic Models*. Working Paper 2008-02. GeoDa Center for Geospatial Analysis and Computation. Arizona State University. 2008

BAJIC, Vladimir. Housing-Market Segmentation and Demand for Housing Attributes: Some Empirical Finding. *AUREA- Journal (Journal of the American Real Estate & Urban Economic Association)*. Vol. 13, N^o 1, 58-75. 1985.

BALL, M.; KIRWAN, R. Accessibility and supply constraints in the urban housing market, *Urban Studies*. Vol. 14, 11-32. 1977.

BAUMONT, Catherine.; ERTUR, Cem; LE GALLO, Julie. *The European regional convergence process 1980-1995: Do Spatial Regimes and Spatial Dependence Matter?* University of Burgundy, França. 2002. Disponível em: <<http://129.3.20.41/eps/em/papers/0207/0207002.pdf>> Acesso em julho de 2009.

BOONE, Louis E.; KURTZ, David L. *Marketing Contemporâneo*. Tradução da 12^a. Edição norte-americana. Cengage Learning, São Paulo – SP. 2009.

BOURASSA, Steven C.; CANTONI, Eva; HOESLI, Martin. Spatial dependence, housing submarkets, and house price prediction. *Journal of Housing Economics*. Vol. 12, 12-28. 2007.

BOURASSA, Steven C.; HAMELINK, Foort; HOESLI, Martin; MACGREGOR, Bryan D. *Defining residential submarkets: evidence from Sydney and Melbourne*. Paper presented at the RICS “Cutting Edge” Conference. Dublin. 1997.

BOURASSA, Steven C.; HAMELINK, Foort; HOESLI, Martin; MCGREGOR, Bryan D. Defining housing submarkets. *Journal of Housing Economics*. Vol. 8, 160-183. 1999.

BOURASSA, Steven C.; HOESLI, Martin. *The Structure of Housing Submarkets in a Metropolitan Region*. 1999.

BOURASSA, Steven C.; HOESLI, Martin ; PENG, Vincent S. Do housing submarkets really matter? *Journal of Housing Economics*. Vol.12, 12-18. 2003.

BRANDSTETTER, Maria Carolina Gomes de Oliveira; HEINECK, Luiz Fernando M. Discussão sobre os principais aspectos do comportamento do consumidor do mercado imobiliário. I Conferência Latino-Americana de Construção Sustentável- X ENTAC - Encontro Nacional de Tecnologia do Ambiente Construído. São Paulo. 2004.

BRETZKE, Miriam. Comportamento do cliente. Capítulo 3. *Gestão de Marketing / Coordenação Sérgio Roberto Dias*. Saraiva. São Paulo. 2003.

CÂMARA, Gilberto; CARVALHO, Marília Sá; CRUZ, Oswaldo Gonçalves; CORREA, Virginia. Análise espacial de áreas. Instituto Nacional de Pesquisas Espaciais. Divisão de Processamento de Imagens, São José dos Campos, SP. 2002. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/>>, Acesso em: julho 2009.

CAN, A. The measurement of neighborhood dynamics in urban house prices. *Economic Geography*, Vol. 66, 254-272. 1990

DALE-JOHNSON, David. An alternative approach to housing market segmentation using hedonic price data. *Journal of Urban Economics*. Vol. 11, 311-332. 1982.

DIAS, Sérgio Roberto. Análise do mercado. Capítulo 2. *Gestão de Marketing / Coordenação Sérgio Roberto Dias*. Saraiva. São Paulo. 2003.

EVANS, Alan W. The property market: ninety per cent efficient? *Urban Studies*. Vol. 32, N° 1, 5-29. 1995.

FERNANDEZ, João Alberto da Costa Ganzo; HOCHHEIM, Norberto. A variável localização e suas implicações mercadológicas. Artigo Técnico. ENTAC. 8°. Salvador, BA. 2000.

FERNANDEZ, João Alberto da Costa Ganzo; ILHA, João Carlos Godoy; OLIVEIRA, Roberto de. Análise de mercado para empreendimentos habitacionais multifamiliares. III Workshop Brasileiro de Gestão de Processos de Projeto na Construção de Edifícios. Belo Horizonte, MG. 2003.

FISHER, Ernest M.; FISHER, Robert M. Urban Real Estate. Henry Holt and Company. New York. 1954.

FREITAS, Ana Augusta Ferreira de. Segmentação do mercado imobiliário utilizando dados de preferência declarada. Tese. Universidade Federal de Santa Catarina. Departamento de Engenharia de Produção de Sistemas. 2000. Florianópolis.

GABRIEL, Stuart. A. A note on housing market segmentation in an Israeli development town. Urban Studies. Vol. 21, 189-194. 1984.

GALSTER, George C. William Grigsby and the Analysis of Housing Sub-markets and Filtering. Urban Studies Vol. 33, N° 10, 1797-1805. 1996.

GALSTER, George C. Comparing demand-side and supply-side housing policies: sub-market and spatial perspectives. Housing Studies. Vol. 12, 549-560. 1997.

GILLEN, Kevin; THIBODEAU, Thomas; WACHTER, Susan. Anisotropic Autocorrelation in House Prices. Journal of Real Estate Finance and Economics. 23:1, 5 – 30. 2001.

GONZALEZ, Marco Aurélio Stumpf; FORMOSO, Carlos Torres. Análise de viabilidade econômico-financeira de construções residenciais. Encontro Nacional de Tecnologia do Ambiente Construído (ENTAC). Salvador, BA. 2000.

GOODMAN, Allen C. Hedonic prices, price indices and housing markets. Journal of Urban Economics. Vol. 5, 471-484. 1978.

GOODMAN, Allen C. Housing submarkets within urban areas definitions and evidence. *Journal of Regional Science*. Vol. 21, N° 2, 175-185. 1981.

GOODMAN, Allen C.; KAWAI, Masahiro. Permanent income, hedonic prices, and demand for housing: new evidence. *Journal of Urban Economics*. Vol. 12, 214-237. 1982.

GOODMAN, Allen C.; THIBODEAU, Thomas G. Housing market segmentation. *Journal of Housing Economics*. Vol.7, 121-143. 1998.

GOODMAN, Allen C.; THIBODEAU, Thomas G. Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics*. Vol. 12, 181-201. 2003.

GOODMAN, Allen C.; THIBODEAU, Thomas G. The spatial proximity of metropolitan area housing submarkets. *Real Estate Economics*. Vol.35 , 209-232. 2007.

GUJARATI, Damodar N. *Econometria básica*. Terceira edição. Pearson Education do Brasil. São Paulo. 2000

HADDAD, Eduardo. EAE 503 - Economia regional e urbana. Disciplina ministrada na Faculdade de Economia e Administração da Universidade de São Paulo. Notas de aula. 2009. Disponível em: <http://www.econ.fea.usp.br/nereus/eae503.htm>. Acesso em julho de 2009.

HADDAD, Emílio. Sobre o estudo da divisão de cidade em zonas homogêneas: aplicação para o município de São Paulo. Tese. Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo. 1987.

HARSMAN, Bjorn; QUINGLEY, John M. The spatial segregation of ethnic and demographic groups: comparative evidence from Stockholm and San Francisco. *Journal of Urban Economics*. Vol. 37, 1-16. 1995.

HERMANN, Bruno M.; HADDAD, Eduardo A. Mercado Imobiliário e Amenidades Urbanas: A View Through the Window . *Estudos Econômicos*. Vol. 35 – N.2. São Paulo. 2005.

JONES, Colin. Housing: The Element of Choice. Urban Studies. Vol. 16, 197-204. 1979.

JONES, Colin. The definition of housing market areas and strategic planning. Urban Studies. Vol. 13. Nº 3, 549-564. 2002.

JONES, Colin; LEISHMAN, Chris; WATKINS, Craig. Housing market processes, urban housing submarkets and planning policy. Paper presented at RICS Cutting Edge Conference. Oxford. 2001.

JONES, Colin; LEISHMAN, Chris; WATKINS, Craig. Intra-Urban migration and housing submarkets: theory and evidence. Housing Studies. Vol. 19, Nº 2, 269-283. 2004.

KAUKO, Tom. A comparative perspective on urban spatial housing market structure: some more evidence of local submarkets based on neural network classification of Amstendam. Urban Studies. Vol. 41, Nº 13, 2555-2579. 2004.

KAUKO, Tom; HAKFOORT, Jacco; HOOIMEIJER, Pieter. Capturing housing market segmentation: an alternative approach based on neural network modelling. Housing Studies. Vol.17, Nº 6, 875-894. 2002.

KOTLER, Philip. Marketing para o séc XXI. Como criar, conquistar e dominar mercados. 4ª edição. Editora Futura. 1999

KOTLER, Philip; KELLER, Kevin Lane. Administração de marketing. 12a edição. Pearson Prentice Hall. São Paulo. 2006.

LANCASTER, Kelvin J. A new approach to consumer theory. Journal of Political Economy, Vol. 74, 132-157. 1966.

LIMEIRA, Tânia Maria Vidigal. Fundamentos de Marketing. Capítulo 1. Gestão de Marketing / Coordenação Sérgio Roberto Dias. Saraiva. 2003. São Paulo.

MACLENNAN, Duncan. Some thoughts on the nature and purpose of house price studies. Urban Studies. Vol. 14, 59-71. 1977.

MACLENNAN, D; MUNRO, M; WOOD, G. Housing choice and the structure of urban housing markets. *Between State and Market Housing in the Post-industrial era*. 1987.

MACLENNAN, D; TU, Y. Economic perspectives on the structure of local housing markets. *Housing Studies*. N.11. 1996.

MALHOTRA, Naresh K. *Pesquisa de Marketing: uma orientação aplicada*. Bookman. Tradução Laura Bocco. 4a. edição. 2006. Porto Alegre.

MALPEZZI, Stephen. Hedonic Pricing Models: A Selective and Applied Review. *Housing Economics: Essays in Honor of Duncan MacLennan*. Edited by Kenneth Gibb and Anthony O'Sullivan. 2002.

MICHAELS, R.; SMITH, V. K. Market segmentation and valuing amenities with hedonic models: the case of Hazardous Waste Sites. *Journal of Urban Economics*. Vol. 28, 223-242. 1990.

MUNRO, M. Testing for segmentation in the private housing market in Glasgow. Centre for Housing Research. Discussion paper n.8. University of Glasgow. 1986.

OLIVEIRA, Maria Carolina Gomes de; FREITAS, Ana Augusta F.; HEINECK, Luiz Fernando M. Explicação da formação de preferências habitacionais utilizando o conceito de ciclo de vida. VII Encontro Nacional de Tecnologia do Ambiente Construído (ENTAC). Florianópolis. 1998.

PALM, Risa. Spatial segmentation of the urban housing market. *Economic Geography*. Vol.54, 210-221. 1978.

PARENTE, Juracy Gomes. O Sistema de informação de marketing e a pesquisa de marketing. Capítulo 14. *Gestão de Marketing / Coordenação Sérgio Roberto Dias*. Saraiva. 2003. São Paulo.

PASHA, Hafiz A.; BUTT, Mohammad S. Demand for housing attributes in developing countries: a study of Pakistan. *Urban Studies*. Vol. 33, N° 7, 1141-1154. 1996.

PERUZZO TRIVELLONI, Carlos A.; HOCHHEIM, Norberto. O valor de localização dos imóveis: determinação por métodos de análise especial. XIII COBREAP. Fortaleza-CE. 2006.

QUIGLEY, John M. Consumer Choice of Dwelling Neighborhood and Public Services. *Regional Science and Urban Economics*. Vol.15, 41-63. 1985.

ROSEN, Sherwin. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*. Vol. 82, 34-55. 1974.

ROTHENBERG, Jerome; GALSTER, George C.; BUTLER, Richard V.; PITKIN, John R. The maze of urban housing markets. University of Chicago Press, IL. 1991. Partes do livro disponibilizadas na Internet em dezembro de 2008.

SCHNARE, Ann B.; STRUYK, Raymond J. Segmentation in urban housing markets. *Journal of Urban Economics*. Vol. 3, 146-166. 1976.

SMITH, Wendell R. Product differentiation and market segmentation as alternative marketing strategies. *The journal of marketing*. Vol.21. N° 1. 1956.

STRASZHEIM, Mahlon R. An econometric Analysis of the market. *Urban and Regional Studies*. N° 2. New York: National Bureau of Economic Research.. 1975.

TU, Yong. The local housing sub-market structure and its properties. *Urban Studies*. Vol. 34, N° 2, 337-353. 1997.

TU, Yong; SUN, Hua; YU, Shi Ming. Spatial autocorrelations and urban housing market segmentation. *The Journal of Real Estate Finance and Economics*. Vol. 34, N° 3. 2007.

WATKINS, Craig. Do new market entrants pay more for housing? Aberdeen Papers in Land Economy. Department of Land Economy. University of Aberdeen. Escócia. 1996.

WATKINS, Craig. The definition and identification of housing submarkets. Aberdeen Papers in Land Economy. University of Aberdeen. Department of Land Economy. Escócia. 1998.

WHEATON, William C. Income and urban residence analysis of consumer demand for location. The American Economic Review. Vol. 69, N^o 4, 620-631. 1977.

WHITEHEAD, C.M.E.; ODING-SMEE, J.C. Long-run equilibrium in urban housing – a note. Urban studies. Vol. 12, 315-318. 1975.

ANEXOS

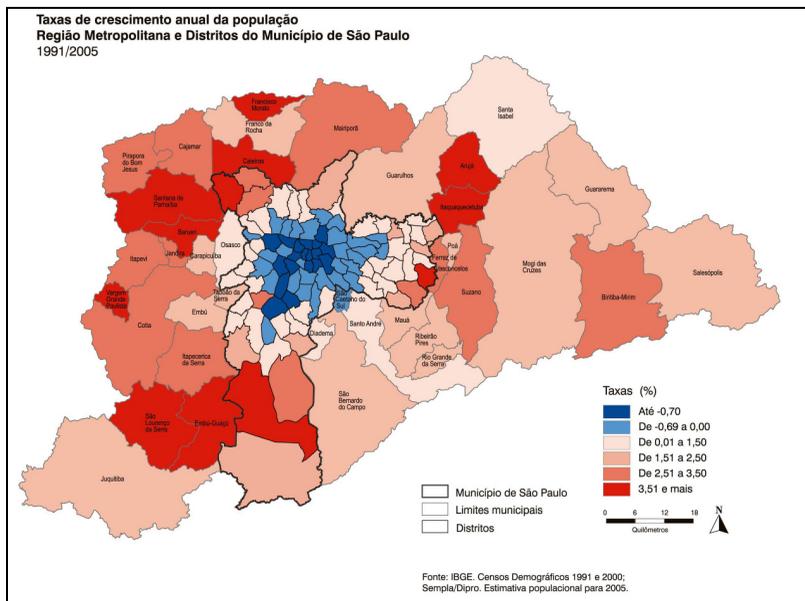


Figura 15 - Mapa da região metropolitana e distritos do município de São Paulo – taxa de crescimento anual da população.

O município de São Paulo aproxima-se atualmente da cifra dos 11 milhões de habitantes sendo a maior cidade do Brasil e da América Latina.

Tabela 7 - Dados da região metropolitana e do município de São Paulo.

Área Total, População Residente e Densidade Demográfica
Região Metropolitana e Município de São Paulo 1991 e 2000

Unidade Territorial	Área (km ²)	População		Densidade (pop/km ²)	
		1991	2000	1991	2000
Região Metropolitana de São Paulo	8.051	15.444.941	17.834.664	1.918	2.215
Município de São Paulo	1.509	9.646.185	10.434.252	6.392	6.915

Fonte: IBGE, Censos Demográficos

Elaboração: Sempla/Dipro – Prefeitura do município de São Paulo

A população, distribuída em uma área de 1.509 km², apresenta algumas especificidades em termos demográficos e urbanos, dentre as quais merece destaque a questão da ocupação territorial.

O Censo Demográfico 2000 (IBGE) mostra, em primeiro lugar, que o município de São Paulo concentra 94,7% dos seus domicílios em área urbana e o restante 5,3 %, em área rural, na periferia da cidade. Além da questão do urbano e rural na cidade, outros pontos interessantes a se destacar são o processo de ocupação do município e a discussão sobre o número de favelas existentes.

As alterações na dinâmica de ocupação urbana são marcantes nas últimas décadas. Até os anos 1980, todas as regiões da cidade apresentavam crescimento positivo, com exceção de um núcleo muito central formado pelos distritos de Belém, Bom Retiro, Brás e Pari. A partir de então, esta tendência ao crescimento negativo espalhou-se para áreas adjacentes como o início da zona leste e partes das regiões norte, oeste e sul. Nos anos 1990, cerca de 60% dos distritos registravam redução populacional. Apenas os distritos situados nas áreas mais extremas da cidade, a maior parte limítrofes com outros municípios da Região Metropolitana, continuavam apresentando crescimento, o que indica que o processo de *periferização* ultrapassou os limites de São Paulo.

Regiões, Subprefeituras e Distritos
Município de São Paulo

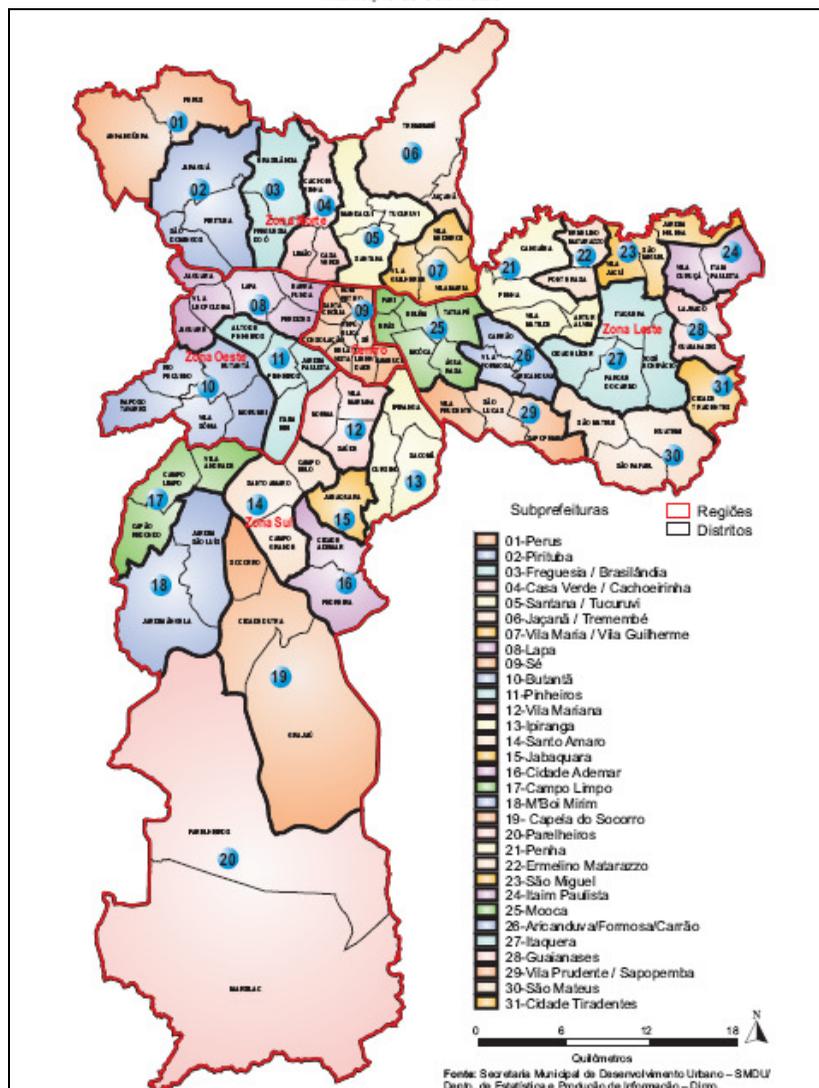


Figura 16 - Mapa do município de São Paulo – Regiões, subprefeituras e distritos.

O processo de verticalização mais significativo segundo os índices de tipologia apartamento do IBGE, se dá pela ordem, nas Subprefeituras Sé, Vila Mariana, Pinheiros, Lapa, Mooca, Santana/Tucuruvi e Ipiranga. De outro lado, a maior incidência de casas (domicílios residenciais horizontais) se dá nas Subprefeituras de Parelheiros, cerca de 98% do total, Capela do Socorro e Perus, 93% cada, São Mateus 92% e Jaçanã/Tremembé 91%.

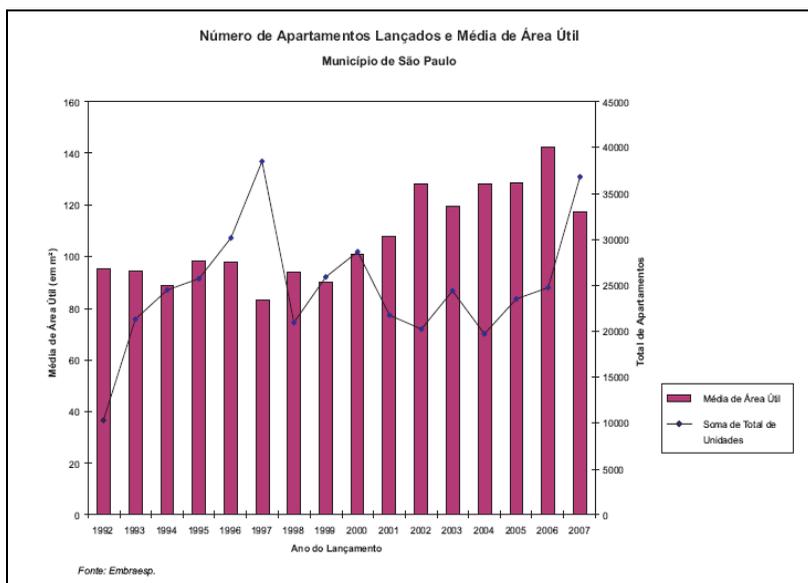


Gráfico 6 - Número de apartamentos lançados e média de área útil: 1992 a 2007 – Município de São Paulo – Fonte: Embrasesp

Tabela 8 - Domicílios segundo tipologias residenciais. Município de São Paulo – (Censo 2000).

Domicílios - segundo tipologias residenciais – 2000
Município de São Paulo, Subprefeituras e Distritos Municipais

Unidades Territoriais	Total domicílios MSP (*)	Casa	Apartamento	Cômodo
MSP	2.985.977	2.193.866	751.401	40.710
Aricanduva/Formosa/Carrão	76.867	66.584	9.774	509
Butantã	106.548	76.897	28.323	1.328
Campo Limpo	137.737	109.708	25.694	2.335
Capela do Socorro	148.737	137.879	8.564	2.294
Casa Verde/Cachoeirinha	87.509	74.198	12.380	931
Cidade Ademar	99.377	90.424	6.744	2.209
Cidade Tiradentes	49.677	23.161	26.082	434
Ermelino Matarazzo	56.881	51.679	4.423	779
Freguesia/Brasilândia	107.296	91.516	14.070	1.710
Guaianases	66.151	56.955	8.648	548
Ipiranga	127.094	92.465	32.950	1.679
Itaim Paulista	93.627	80.528	12.077	1.022
Itaquera	130.074	102.319	26.378	1.377
Jabaquara	62.416	45.591	15.946	879
Jaçanã/Tremembé	68.290	62.210	5.060	1.020
Lapa	87.703	44.043	43.121	539
M'Boi Mirim	130.958	116.558	8.882	5.518
Mooça	97.227	58.907	36.236	2.084
Parelheiros	28.074	27.523	106	445
Penha	137.264	110.586	25.815	863
Perus	28.109	26.020	1.848	241
Pinheiros	99.227	25.154	73.851	222
Pirituba	108.867	88.763	19.472	632
Santana/Tucuruvi	95.848	61.753	33.566	529
Santo Amaro	67.644	41.009	26.128	507
São Mateus	100.920	92.845	7.124	951
São Miguel	100.511	88.978	9.977	1.556
Sé	140.193	17.862	118.350	3.981
Vila Maria/Vila Guilherme	87.523	75.133	10.693	1.697
Vila Mariana	109.470	33.296	75.611	563
Vila Prudente/Sapopemba	148.158	123.322	23.508	1.328

Fonte: IBGE - Censo 2000 - Elaboração: Sempla/Dipro

(*)Total de domicílios MSP : Base Universo do Censo 2000. Dado levantado apenas para os domicílios particulares permanentes

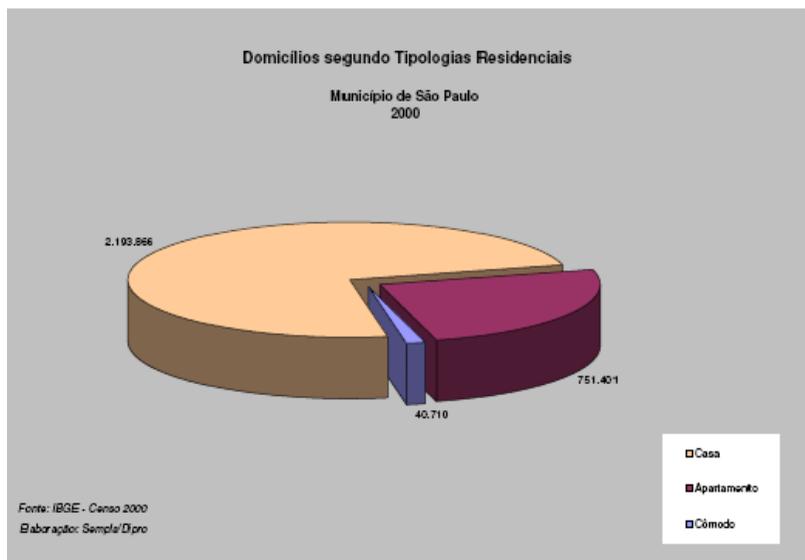


Gráfico 7 - Domicílios segundo tipologias residenciais. Município de São Paulo – (Censo 2000).