

## 1. Introdução

Tradicionalmente, na literatura econômica, o bem-estar da sociedade está associado à sua capacidade de gerar os bens necessários e desejados ao consumo dos indivíduos. Quanto maior o consumo de bens, satisfeitas as hipóteses de racionalidade, maior o nível de bem-estar atingido. A noção de consumo está relacionada, de maneira geral, à transformação de bens materiais. Entretanto, quando introduzimos a dimensão espacial, observamos que uma parcela da satisfação dos indivíduos depende das características específicas de cada localização. Os agentes econômicos possuem um endereço, residem e trabalham majoritariamente em cidades. Portanto, observamos que a presença ou a ausência de amenidades urbanas interfere diretamente em seu bem-estar. Neste caso, podemos extrapolar a noção de consumo, incorporando um conjunto específico de amenidades urbanas à cesta dos agentes. Assim, entendemos que as escolhas dos indivíduos dependem, além da preferência por um conjunto de bens materiais, de uma parcela de bens não-materiais que refletem a qualidade de vida dos centros urbanos.

O estudo de valoração dessas amenidades revela-se importante para a compreensão do impacto do meio urbano sobre os indivíduos. Qualquer alteração no espaço físico das cidades pode ter conseqüências sobre o preço das moradias e sobre o bem-estar dos agentes. A escolha residencial das famílias e a migração intra-urbana dependem da localização de amenidades. Assim, além das variáveis usuais de análise, como produção, emprego, entre outras, a presença de amenidades urbanas motiva estudos para a compreensão da satisfação dos indivíduos. Políticas públicas ou ações privadas com potencial impacto sobre o meio urbano seriam mais bem avaliadas sob essa perspectiva.

Entendemos por amenidades urbanas um conjunto de características específicas de uma localidade com contribuição positiva ou negativa para a satisfação dos indivíduos.<sup>1</sup> As amenidades não estão restritas a características naturais, como áreas verdes, praias, clima, etc. Também estão incluídos na definição os bens (ou males) gerados pelo próprio homem, tais como trânsito, poluição, oferta de entretenimento, segurança, etc.

A valoração de amenidades urbanas não é algo trivial. Enquanto na maioria dos mercados podemos observar as quantidades transacionadas e os preços dos produtos, no caso das amenidades isto não acontece.<sup>2</sup> Embora possamos supor que exista uma oferta, uma demanda e um preço de equilíbrio para as diferentes amenidades, não podemos coletar esses dados diretamente do mercado. Não existe explicitamente um mercado de compra e venda de trânsito, ou um mercado de compra e venda de poluição. Dadas essas dificuldades, devemos, pois, recorrer a algum modelo alternativo que seja capaz de revelar essas informações.

O modelo mais utilizado na literatura para esse fim é o de preços hedônicos. Este considera um bem heterogêneo como um pacote fechado de atributos e estima o preço marginal de cada atributo a partir da análise do valor observado do bem heterogêneo e de suas respectivas quantidades de atributos.<sup>3</sup> Podemos considerar um imóvel como um bem heterogêneo, com valor observado, composto de um pacote de características, no qual se incluem as amenidades urbanas de sua localidade. Supomos que as famílias, quando fazem escolhas residenciais, não estão preocupadas apenas com o que existe da janela para dentro do imóvel, mas também com o que existe da janela para fora. Ou seja, a vizinhança interfere na qualidade do imóvel e, portanto, afeta seu preço de locação. Analisando o mercado privado de aluguéis, podemos inferir sobre o preço das amenidades.

---

<sup>1</sup> Ver Bartik *et al.* (1987).

<sup>2</sup> Ver Sheppard (1999).

<sup>3</sup> Ver Rosen (1974).

O modelo de preços hedônicos para habitação vem sendo largamente utilizado para medir o valor marginal das características intrínsecas, ou estruturais do imóvel, e também para estimar variáveis sócio-ambientais correlacionadas. Dentre os estudos recentes podemos destacar: Bowes e Ihlanfeldt (2001), que avaliam o impacto das estações de trem sobre o valor dos imóveis para Atlanta, Estados Unidos; Espey e Lopez (2000), que pesquisam o efeito do barulho dos aeroportos em Reno-Sparks, Estados Unidos; Benson *et al.* (1998), que avaliam o impacto da vista sobre o preço dos imóveis em Bellingham, Canadá; Macedo (1998), que estima o preço de variáveis estruturais em Belo Horizonte, Brasil; e Andersson (1997), que analisa a qualidade dos condomínios em Singapura, Malásia. Para a cidade de São Paulo, especificamente, há pelo menos dois estudos que aplicam o modelo hedônico: Biderman (2001) avalia a demanda por imóveis novos e Oliveira (1997) estima o preço negativo da poluição do ar.

O objetivo deste trabalho é procurar, na medida do possível, esgotar todo o efeito vizinhança sobre o valor dos aluguéis. O artigo está desenvolvido da seguinte maneira: na seção 2, fazemos um breve resumo do modelo de preços hedônicos e sua aplicação para o mercado de aluguéis; na seção 3, apresentamos um conjunto de problemas recorrentes na estimação de preços hedônicos e explicitamos o método escolhido; na seção 4, fazemos algumas considerações sobre a configuração urbana assumida para a cidade de São Paulo; na seção 5, apresentamos os dados utilizados na pesquisa e desenvolvemos um critério para seleção de variáveis ambientais; na seção 6, apresentamos os resultados obtidos; e, na seção 7, tecemos algumas conclusões sobre o trabalho.

## **2. Considerações Teóricas Iniciais**

A maioria dos modelos econômicos trata de interações entre os agentes que se dão via mercado. Produtores e consumidores podem tomar suas decisões otimizadoras, baseados em informações sobre preços e suas próprias capacidades de produção e consumo. Entretanto, para estudarmos o impacto de amenidades urbanas sobre a moradia, precisamos extrapolar essa hipótese. Consumidores não fazem escolhas diretas sobre a quantidade de poluição ou de segurança que desejam adquirir. Um novo edifício erguido em uma quadra horizontal toma a privacidade dos antigos moradores, que, por sua vez, não pagaram ou obtiveram nada por isso. Os agentes econômicos importam-se com essas características locais, que, contudo, não são vendidas no mercado.

Quando estudamos as externalidades geradas pelas amenidades urbanas, enfrentamos a limitação de desconhecer o seu verdadeiro valor. Ou seja, não existe propriamente um mercado para esse tipo de “produto” e, portanto, não podemos observar o seu preço. Contudo, podemos assumir que existe implicitamente uma oferta e uma demanda por essas características e, assim, tentar inferir seus respectivos preços de equilíbrio. Este é o roteiro traçado pelo modelo de preços hedônicos. O preço marginal dos atributos é estimado, fazendo regressão o preço do bem heterogêneo sobre as quantidades de características associadas.

Entretanto, a implementação do modelo para o mercado de imóveis esbarra em mais um problema. Sabemos exatamente quais são as amenidades e as magnitudes consumidas pelos agentes? Provavelmente não. Para Sheppard (1999) a pesquisa sobre preços hedônicos assemelha-se a uma situação na qual um pesquisador, impossibilitado de entrar num supermercado, deve inferir sobre o preço dos produtos, apenas com base na informação do valor total da conta de cada família, de suas respectivas características, como gostos e renda, e em uma visão turva das quantidades de produtos que circulam dentro dos carrinhos.

Isso é essencialmente o que acontece quando empreendemos uma análise hedônica para o mercado de imóveis (ou aluguéis). O preço do bem final geralmente é conhecido com clareza. Entretanto, informações sobre as quantidades de características consumidas são bastante imprecisas. Embora

possamos aceitar que as características observadas da janela para fora dos imóveis afetam seu preço de locação, não sabemos quais dessas amenidades estão sendo consumidas e em que quantidades exatas. Como veremos, a solução frequentemente utilizada para esse problema é assumir *ad hoc* o pacote de atributos relevantes e designar suas quantidades de acordo com a unidade territorial escolhida.

Como observamos, a análise de preços hedônicos é caracterizada pela estimação de uma “função preço”, com o preço do aluguel de um lado e o pacote de características de outro. A apresentação isolada dessa função mascara o conteúdo microeconômico do modelo, elegantemente desenvolvido em Rosen (1974) para o caso de mercados competitivos, provendo a formalização necessária.

Na próxima seção, trataremos das dificuldades enfrentadas para estimar a função de preços hedônicos. Relataremos as diversas opções adotadas e explicitaremos o método de estimação escolhido.

### 3. Método de Estimação

#### 3.1. Problemas na Estimação de Preços Hedônicos

A estimação da equação hedônica apresenta um grande conjunto de dificuldades. Primeiro, a teoria não determina uma forma funcional nem as variáveis relevantes para a estimação. Segundo, a regressão esbarra frequentemente no problema de multicolineariedade. E, terceiro, mais recentemente alguns autores vêm alertando para a presença de autocorrelação espacial dos resíduos. Faremos uma rápida revisão desses problemas, a partir da extensa literatura já produzida sobre o assunto, determinando o que nos parece ser mais apropriado para a execução do presente trabalho.

Tradicionalmente os estudos adotam uma forma linear ou log-linear para a função de preços hedônicos. Entretanto, a partir da década de 1980, começaram a surgir alguns trabalhos que adotam a transformação Box-Cox. Como a teoria não determina uma forma funcional específica, a vantagem da transformação Box-Cox é permitir que os próprios dados revelem a melhor forma funcional.

O método Box-Cox tem a vantagem de gerar resíduos homocedásticos e simétricos. Entretanto, como o parâmetro de transformação é dependente das variáveis explicativas escolhidas, o método revela-se incapaz de atestar consistentemente a significância dos coeficientes estimados.<sup>4</sup> Como estamos interessados justamente em revelar as amenidades urbanas que causam impacto sobre o preço dos aluguéis, ou seja, como desconhecemos as variáveis relevantes, parece-nos mais conveniente adotar um método de estimação que teste adequadamente a significância de cada uma delas.<sup>5</sup>

Ainda sobre o problema de especificação, precisamos fazer algumas observações sobre as variáveis explicativas que devem entrar na função de preços hedônicos. Embora o modelo hedônico não determine *a priori* as variáveis relevantes, muitos estudos ignoram alguns conceitos clássicos da economia urbana. As contribuições de Alonso (1964), Mills (1972), Fujita e Ogawa (1982), entre outros, evidenciam o papel da localização dos imóveis para a formação de um gradiente de preços.

---

<sup>4</sup> Ver Andersson (1997).

<sup>5</sup> O método Box-Cox ainda tem a desvantagem de gerar resultados não muito intuitivos. Como comenta Bowen *et al.* (2001), o coeficiente estimado expressa a variação no preço do aluguel transformado dada uma variação marginal no atributo transformado. Se adotamos uma forma linear, o preço hedônico marginal passa a ser uma constante. Embora essa alternativa esbarre na crítica de que a função de preços hedônicos seja não-linear, essa simplificação será útil para dedicarmos maior atenção a outros problemas estatísticos. Na verdade, experimentamos outras formas pré-especificadas, mas a versão linear foi sempre superior.

Segundo estas, à medida que o imóvel se afasta do centro de negócios, o preço da moradia deve compensar o aumento nos custos de transporte. Voltaremos a esta questão de forma mais aprofundada na seção 4. Por enquanto, é relevante apenas destacar a necessidade de incluir um vetor que expresse o efeito acessibilidade.

A teoria econômica sugere também a inclusão da oferta de bens públicos e a qualidade da vizinhança na análise sobre o preço de moradia. Esse é justamente o foco do presente trabalho, e talvez um dos tópicos mais problemáticos na literatura sobre o modelo hedônico para habitação. Muitos estudos vêm apresentando resultados insignificantes ou mesmo invertidos para os coeficientes das amenidades urbanas. O problema principal reside na dificuldade de mensurar exatamente aquilo que está sendo consumido pelos moradores. Andersson (1997) apresenta uma pesquisa com os fatores em ordem de importância que determinam a escolha residencial dos agentes. Em sua maioria, os principais atributos eleitos são de natureza intangível, como o conceito de apazibilidade, por exemplo, o que dificulta a escolha de uma *proxy* para seus efeitos. Na tentativa de preencher esses fatores, podemos selecionar uma variável bastante correlacionada com alguma outra, o que leva ao problema de multicolinearidade. Ou uma *proxy* pode igualmente estar captando efeitos distintos, muitos vezes ambíguos, o que impacta na significância do coeficiente.

O problema de medida ainda está relacionado à exata fronteira de impacto das amenidades. Geralmente os dados são fornecidos para alguma fronteira administrativa – município, distrito, setor censitário ou bairro –, o que não condiz necessariamente com o alcance das externalidades provocadas. A falta de informação sobre a qualidade da vizinhança é um dos motivos apontados para a presença de autocorrelação espacial.<sup>6</sup> O aluguel de um imóvel pode estar correlacionado com o aluguel de seus vizinhos, porque, na verdade, eles comungam as mesmas amenidades urbanas.

Como observamos, os estudos empíricos sobre preços hedônicos para habitação enfrentam regularmente o problema de multicolinearidade. As consequências, nesse caso, são bastante indesejadas: torna-se muito difícil identificar separadamente o efeito das variáveis envolvidas; os parâmetros estimados tendem a apresentar baixa significância; e qualquer variação no número de observações ou no quadro das variáveis explicativas pode alterar sensivelmente os coeficientes estimados. A solução para esse caso envolve a inclusão de mais informação no banco de dados – que nem sempre é disponível – ou a adoção de técnicas para correção da não-ortogonalidade das variáveis explicativas. Seguindo os passos de Kain e Quinley (1970) e Can (1990), apresentaremos uma especificação alternativa que inclui a técnica de análise fatorial.

Na literatura mais recente, alguns autores começaram a atentar para a presença de autocorrelação espacial nos resíduos.<sup>7</sup> Dada a natureza geográfica da pesquisa, os imóveis podem apresentar algum grau de dependência espacial. Isso pode ocorrer pela falta de discriminação das variáveis ambientais, como exposto anteriormente, ou pela presença de efeitos *spillover* entre imóveis. A quebra na hipótese de independência entre observações provoca a estimação de parâmetros viesados e a determinação de níveis de significância enganosos. Portanto, é necessária a adoção de testes para averiguação de uma possível dependência espacial entre as observações. Caso o resultado dos testes seja positivo, deve-se reespecificar o modelo.

### 3.2. Especificação e Teste

Feitas as considerações necessárias sobre a forma funcional e as variáveis relevantes, a equação a ser estimada será especificada da seguinte maneira:

$$p(S, A, E) = \alpha_0 + \alpha_1 S + \alpha_2 A + \alpha_3 E + u, \quad (1)$$

<sup>6</sup> Ver Can (1990) e Andersson (1997).

<sup>7</sup> Ver Can (1990) e Dubin (1992).

onde  $S$  representa uma matriz de características estruturais do imóvel,  $A$  uma matriz que designa acessibilidade e  $E$  uma matriz de variáveis ambientais.

Além da especificação tradicional, vamos estimar uma forma alternativa, incluindo análise fatorial. O método busca solucionar o problema de multicolinearidade, criando fatores que são transformações lineares das variáveis independentes, de maneira que sejam ortogonais ou quase ortogonais entre si. Utilizaremos a análise fatorial apenas para as variáveis ambientais, mantendo o vetor original de variáveis estruturais e de acessibilidade.

A primeira etapa resume-se à obtenção de uma solução linear para um conjunto de variáveis em termos de fatores hipotéticos, como segue:

$$z_i = \sum_j a_{ij} F_j + d_i U_i, \quad (2)$$

onde  $z_i$  é a versão normalizada do atributo  $i$ ,  $F_j$  é o fator comum  $j$ ,  $a_{ij}$  é a carga fatorial,  $U_i$  é o fator singular  $i$  e  $d_i$  é o parâmetro correspondente.

O número de fatores comuns pode ser reduzido em relação ao número de variáveis iniciais. Dependendo da quantidade de restrições em relação ao número de incógnitas, as cargas fatoriais não são unicamente determinadas. Pelo contrário, existem infinitas formas de construção de fatores ortogonais ou quase-ortogonais. Este problema de indeterminação permite a seleção arbitrária de alguma solução mais favorável à interpretação dos fatores. A partir de uma extração inicial, podemos rotar os fatores para selecionar a solução “preferida”.<sup>8</sup> Adotamos o método de rotação *varimax*, que aumenta a carga fatorial de cada variável para um fator específico. Com isso, podemos melhor definir a característica de cada fator. Ou seja, como abandonamos as variáveis iniciais, desejamos criar fatores que apresentem alguma interpretação econômica razoável. A rotação *varimax* facilita essa construção, pois maximiza a variância de uma variável inicial sobre o mesmo fator.

Na segunda etapa, os fatores devem ser descritos em função das variáveis. Como o número de fatores é inferior ao de variáveis, a matriz de cargas fatoriais é singular. O procedimento mais adequado neste caso é a simples regressão pelo método de mínimos quadrados. A regressão linear é expressa da seguinte forma:

$$F_j = \sum_i \beta_{ji} z_i + \varepsilon_j \quad (3)$$

Conhecidos os fatores, podemos então estimar a função de preços hedônicos, como segue:

$$P = \alpha_0 + \alpha_1 S + \alpha_2 A + \sum_{j=3}^J \alpha_j F_j(z_1, z_2, \dots, z_I) + u \quad (4)$$

O modelo pode, ainda, ser reespecificado de acordo com a presença de dependência espacial. Existem duas maneiras para incorporar esse efeito à especificação tradicional. A primeira é o modelo com autocorrelação nos resíduos:

$$p = \alpha Z + u, \quad (5)$$

---

<sup>8</sup> Ver Harman (1967).

com

$$u = \gamma Wu + e,$$

onde  $\gamma$  é um coeficiente auto-regressivo espacial,  $W$  uma matriz de pesos espaciais e  $e$  um termo de erro padrão.

A segunda forma adiciona uma variável dependente espacialmente defasada na especificação tradicional:

$$p = \rho Wp + \alpha Z + u \quad (6)$$

A segunda alternativa apresenta um significado econômico mais explícito. Estamos assumindo que o valor de um aluguel é influenciado pelos aluguéis vizinhos. De fato, no mercado imobiliário é comum valorar aluguéis baseando-se na precificação dos imóveis próximos. A presença de dependência espacial ainda pode ser explicada por variáveis ambientais omitidas. Neste caso, imóveis próximos comungam a mesma qualidade de vizinhança, e o efeito das variáveis ambientais omitidas é captado no coeficiente da variável dependente espacialmente defasada. Mostraremos na seção 6 que, à medida que incorporamos variáveis de conteúdo geográfico, a significância dos coeficientes de autocorrelação espacial tende a diminuir.

A potencial ocorrência de dependência espacial pode ser avaliada utilizando-se o teste de Multiplicadores de Lagrange (LM). O teste LM é baseado na estimação por máxima verossimilhança. Contudo, o teste não necessita da estimação do modelo alternativo. O modelo restrito sem um parâmetro de dependência espacial é assumido como hipótese nula. Assim, podemos proceder ao teste LM pelo método de mínimos quadrados.<sup>9</sup>

A hipótese de dependência espacial é incorporada às equações (5) e (6) através da matriz de pesos generalizados  $W$ . A matriz de dimensão  $n \times n$  avalia o grau de vizinhança entre as observações. Diferentemente dos modelos de séries de tempo, a construção de um operador defasado para fenômenos é um tanto arbitraria.<sup>10</sup> Como trabalhamos com unidades espaciais na forma de pontos, adotaremos medidas diretas de distância entre as observações, da seguinte forma:

$$W1: w_{ij} = d_{ij}^{-1}$$

$$W2: w_{ij} = d_{ij}^{-2}$$

onde  $d_{ij}^{-1}$  é o inverso da distância entre o domicílio  $i$  e o domicílio  $j$  e  $d_{ij}^{-2}$  é o inverso do quadrado da distância.

Na seção 6, apresentaremos os resultados das estimações para a função de preços hedônicos, utilizando as variáveis ambientais originais e a construção dos fatores, equações (1) e (4). O trabalho ainda envolve três versões para cada caso, dependendo do modelo de acessibilidade assumido, que discutiremos a seguir.

#### 4. Acessibilidade

Como já havíamos mencionado, um estudo empírico sobre o modelo hedônico para habitação não pode estar desprovido de um embasamento teórico em economia urbana. Um dos motivos é o papel

---

<sup>9</sup> Ver Anselin (1988).

<sup>10</sup> *Idem*.

determinante da localização na formação de preços de imóveis e aluguéis. Para selecionar o vetor que capte o efeito da acessibilidade na função hedônica, precisamos primeiro considerar as diferentes configurações urbanas que mais se aproximam do nosso objeto de estudo.

#### 4.1. Configuração Urbana Monocêntrica

A economia urbana tradicional baseia-se num modelo de cidade circular com um único centro de negócios responsável por todo o emprego local. O modelo monocêntrico foi desenvolvido nos trabalhos pioneiros de Alonso (1964), Mills (1972) e Wheaton (1974), com sua apresentação sendo consolidada em Brueckner (1987). Nesse modelo, o valor do aluguel e a densidade estrutural da cidade são funções negativas da distância ao centro de negócios. A região central é determinada *ad hoc* e representa um pico no gradiente de preços de aluguel.

Em geral, os estudos empíricos sobre o modelo hedônico para habitação assumem a forma urbana monocêntrica, incorporando uma variável de distância linear ou logarítmica de cada domicílio ao centro de negócios  $A_i = d_i$  ou  $A_i = \ln(d_i)$ . É evidente que nem todas as cidades apresentam essa configuração. Muito pelo contrário, especialmente nas grandes metrópoles observamos formatos multicêntricos. As cidades, em geral, têm uma origem monocêntrica, mas, com a expansão territorial, os custos de transporte induzem o surgimento de novos centros. É possível, ainda, que as cidades não apresentem crescimento radial, pois características topográficas podem limitar essa expansão. Um outro aspecto bastante relevante é que o pico do gradiente de preços nem sempre se localiza nos centros de negócios. A presença de amenidades ou “desamenidades” urbanas pode modificar essa estrutura. Um exemplo clássico é a deterioração dos centros históricos. Nesse caso, o gradiente de preços, ao invés de apresentar um formato monotonicamente decrescente, apresentaria, a partir da origem, um movimento ascendente, para depois assumir uma inclinação negativa.

A cidade de São Paulo apresenta algumas vantagens para a adoção de um formato monocêntrico. Primeiro, encontramos um centro histórico bem definido, tradicionalmente *locus* das atividades de comércio e serviços, e podemos assumir a Praça da Sé como marco fundamental. Depois, a cidade não sofreu nenhuma restrição geográfica que impedisse sua configuração bastante circular. Entretanto, observando a distribuição dos postos de trabalho por distrito no município, percebemos que a incidência de emprego não se limita ao antigo centro de negócios.<sup>11</sup>

Encontramos ainda um outro *cluster* de empregos a sudoeste, na região das avenidas Berrini e Faria Lima. Este ponto pode ser adotado como um novo centro de negócios, o que indicaria um formato duocêntrico para a cidade. Entretanto, o formato com dois centros ainda é uma boa simplificação, visto que os empregos não estão restritos a essas áreas.

Como, na verdade, estamos interessados no fluxo de pessoas e no impacto dos custos de transporte, uma melhor *proxy* de avaliação é a razão empregos sobre população economicamente ativa. Dessa forma, podemos separar os distritos eminentemente importadores de mão-de-obra dos demais. A Tabela 1 resume as informações para os dez principais distritos importadores de mão-de-obra.

---

<sup>11</sup> Observamos uma aglomeração em torno do centro histórico, na Praça da Sé, mas com uma proeminente expansão na direção oeste. A avenida Paulista, por se localizar muito próxima do centro histórico, pode, para efeito de cálculos de distância, ser considerada parte deste núcleo. O único fato indesejado para a análise de emprego por distrito é que a avenida se localiza na fronteira dos distritos Bela Vista e Jardim Paulista.

**Tabela 1:** Emprego Total por Distrito - São Paulo

<i>Classif.</i>	<i>Distritos</i>	<i>Emprego total</i>	<i>Total acumulado</i>	<i>Total/PEA</i>
1	Sé	91,520	3.70%	6.1993
2	Barra Funda	50,127	5.72%	5.5796
3	República	103,116	9.89%	2.8311
4	Santo Amaro	119,757	14.73%	2.7275
5	Bela Vista	119,820	19.57%	2.5365
6	Vila Leopoldina	47,755	21.50%	2.4997
7	Pinheiros	94,555	25.32%	2.0798
8	Pari	20,120	26.14%	1.9758
9	Brás	33,907	27.51%	1.9177
10	Itaim Bibi	110,842	31.99%	1.8813

Fonte: Sempla (2002).

Observamos que os seis distritos que incorporam os dois centros assinalados, Sé, República e Bela Vista, para o centro histórico, e Santo Amaro, Pinheiros e Itaim Bibi, para o centro novo, encontram-se na lista acima. O grau de aglomeração ainda é baixo quando comparado ao índice de empregos total do município. Os dez principais distritos importadores de mão-de-obra representam apenas 32% do total de empregos na cidade.<sup>12</sup>

Embora a região central ainda represente um papel importante como *locus* de emprego no Município de São Paulo, a opção tradicional pelo modelo monocêntrico parece um pouco limitada. É evidente a participação do centro de serviços, complexo Berrini-Faria Lima, na atividade econômica da cidade, como já se havia discutido em Biderman (2001). Portanto, devemos avançar um pouco mais nessa questão e analisar modelos alternativos para a seleção do vetor de acessibilidade.

#### 4.2. Configuração Urbana Duocêntrica

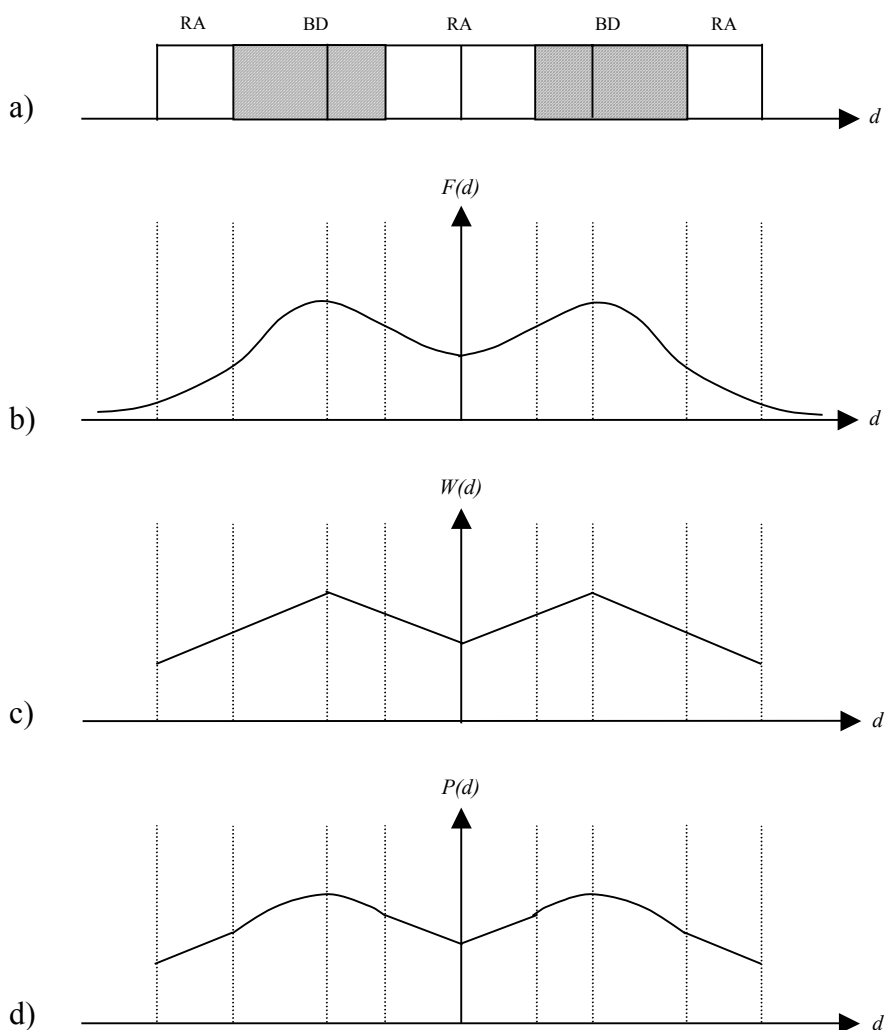
Muitos estudos teóricos já foram desenvolvidos com o objetivo de modelar áreas metropolitanas não-monocêntricas. No trabalho pioneiro de Fujita e Ogawa (1982), a configuração urbana passa a ser determinada endogenamente pelo modelo. Firms e famílias competem entre si na decisão de localização e enfrentam *trade-off's* por estarem mais ou menos próximas dos respectivos centros. Para as firmas, os ganhos de aglomeração são contrapostos aos crescentes custos de produção e, no caso das famílias, o custo de transporte é compensado pela queda no valor dos aluguéis. Este modelo é um tanto mais complexo, mas, de qualquer forma, reproduz o gradiente de preços negativamente inclinado com origem no centro (ou centros) de negócios.

No Gráfico 1, observamos o caso de uma cidade duocêntrica. Na parte (a), é apresentada a configuração urbana, onde *BD*, *business district*, representa a região na qual se localizam exclusivamente as firmas e *RA*, *residential area*, a região de localização exclusiva das famílias. Na parte (b), são ilustrados os ganhos de aglomeração das firmas através da função  $F(d)$ . Podemos observar que à medida que nos aproximamos dos centros os ganhos são maiores. O mesmo efeito é notado para os salários  $W(d)$  e para os aluguéis  $P(d)$ , partes (c) e (d).

<sup>12</sup> É importante salientar que os dados para emprego fornecidos pela SEMPLA (2002) apresentam um viés de não-aglomeração, isso porque não contabilizam os empregos no setor de administração pública, concentrados no centro histórico, e desconsideram o mercado informal. Embora não existam números precisos sobre a distribuição espacial do setor informal, este parece também estar concentrado no centro histórico.



**Gráfico 1: Modelo Duocêntrico**



A incorporação de um vetor de acessibilidade na equação hedônica, neste caso, ainda é pouco discutida na literatura. Plaut (1998) adota simplificada duas medidas de distância para os respectivos centros:  $\alpha A_i = \alpha^1 d_i^1 + \alpha^2 d_i^2$ , onde  $\alpha^n$  é o coeficiente da distância ao centro  $n$  e  $d^n$  é a distância ao centro  $n$ . Entretanto, esta interpretação não parece coerente com o modelo de Fujita e Ogawa (1982), já que a distância de ambos os centros afetaria o preço de todas as moradias. Segundo o modelo, as famílias trabalham exclusivamente nos centros mais próximos. O valor de salários e aluguéis decresce na mesma proporção, independentemente da direção. Uma medida de acessibilidade para o caso duocêntrico deveria retratar apenas a distância para o centro mais próximo, como  $A = \min(d_i^1, d_i^2)$ . Considerando que pode haver diferença na altura dos interceptos das regiões centrais, devemos distinguir no entorno de que vizinhança localiza-se a observação, incluindo uma *dummy* relativa à proximidade a um dos dois centros.

Temos agora um quadro completo das estimações a serem realizadas. Procuraremos extrair os preços hedônicos marginais das amenidades urbanas através das variáveis ambientais originais ou de fatores. Para cada caso serão apresentadas três versões de acessibilidade: o caso monocêntrico, o caso duocêntrico com duas distâncias e o caso duocêntrico com distância mínima. Antes de procedermos às estimações, resta ainda apresentar o banco de dados.

## 5. O Banco de Dados

### 5.1 Amostra de Domicílios

Os dados necessários para a estimação das equações (1) e (4) foram, em parte, obtidos diretamente de fontes primárias e, em parte, gerados através da utilização de um sistema geográfico de informação (SIG). Os dados agrupados no SIG apresentam três formas distintas de representatividade espacial: pontos, distância entre pontos e polígonos. As informações sobre imóveis estão expressas na forma de pontos no espaço, sendo que cada ponto representa um domicílio ao qual está associado um valor de aluguel e um conjunto de características estruturais. As informações sobre acessibilidade e um grupo de amenidades estão expressas na forma de distância entre pontos. Isto é feito calculando-se a distância em linha reta entre uma observação e um outro ponto de referência (e.g. a distância de cada domicílio ao centro de negócios). Um último tipo de informação é aquela representativa de um território específico, um polígono – no presente estudo, um distrito correspondente no município de São Paulo. Quando um domicílio é georeferenciado, automaticamente assume o valor do polígono que o circunda. Portanto, para grande parte das amenidades, assumimos o distrito como a vizinhança relevante para a caracterização do imóvel. Como já discutimos, essa decisão arbitrária embute um problema de erro de medida, pois os distritos não são homogêneos o suficiente para retratarmos com o mesmo valor domicílios estabelecidos em lugares diferentes do mesmo distrito.

As informações sobre imóveis foram extraídas da “Pesquisa de Orçamentos Familiares” (POF) produzida pela Fundação Instituto de Pesquisas Econômicas – Fipe (1999). São ao todo 497 observações, cada uma com o valor do aluguel pago pela família e diversas características estruturais do imóvel. Esta amostra foi georeferenciada de acordo com os endereços presentes nos questionários da POF.

A adoção do valor do aluguel, em vez do valor de compra e venda do imóvel, parece-nos mais conveniente. O valor do imóvel expressa uma expectativa futura, pois seu cálculo é baseado no valor presente de um número infinito de aluguéis. Portanto, alterações espaciais futuras, ou sua simples expectativa, podem impactar o valor venal do imóvel. Como nossa análise limita-se às características espaciais em um dado período de tempo, o valor do aluguel não capta esse distúrbio.

Na Tabela 2, observamos algumas estatísticas descritivas sobre as características estruturais dos imóveis e as variáveis de distância calculadas entre cada domicílio e os centros de negócios eleitos. Os dados da Fipe (1999) não esgotam todo o escopo de variáveis estruturais. Particularmente faltam duas informações bastante usuais nos estudos hedônicos: a idade do imóvel e a qualidade do acabamento. Estas duas variáveis de controle apresentam elevada significância e sinal esperado negativo para a primeira e positivo para a segunda na maioria das pesquisas. Sua ausência, além de diminuir a capacidade de explicação do modelo, acarretará um certo grau de viés sobre os parâmetros estimados. Outra variável bastante usual nos trabalhos com modelos hedônicos para habitação é o tamanho em metro quadrado do imóvel. Neste estudo, o número de cômodos funciona como uma aproximação do tamanho do imóvel, embora sob a forte hipótese implícita de que todos os cômodos têm o mesmo padrão.

**Tabela 2:** Variável Independente e Variáveis de Controle

<i>Variável</i>	<i>Descrição</i>	<i>Média</i>	<i>Dv-padrão</i>
ALUG	aluguel em R\$	353.20	211.55
SALA	nº de salas	0.7042	0.5110
GARAG	nº de garagens	0.3581	0.5164
DORM_AG	nº de dormitórios + 0.5*dorm. de empregada	1.5755	0.8223
COZ_AG	nº de cozinhas + 0.5*(copas + lavanderias)	1.5453	0.5701
BAN_AG	nº de banheiros + 0.5*(lavabos + banh. de empregada)	1.1932	0.5981
OUT_DEP	nº de áreas externas + quintais + outros depósitos	0.6338	0.5451
DIST_SE	distância à Praça da Sé, no centro histórico, em km	11.81	6.19
DIST_BER	distância ao cruzamento da Av. das Nações Unidas e Av. Bandeirantes em km	14.20	7.18
DIST_CBD	menor distância entre o centro histórico e o centro novo em km	10.452	5.924
DUM_SE	dummy: menor distância à Praça da Sé	0.71	0.45

Fonte: elaborado a partir de dados do LUME (2002) e FIPE (1999).

## 5.2. Seleção de Variáveis Ambientais

A seleção das variáveis ambientais costuma ser bastante arbitrária. Alguns trabalhos vinculam a escolha delas diretamente ao seu objeto de análise. Outros simplesmente as ignoram e limitam-se a estudar o comportamento das variáveis estruturais. Este trabalho, ao contrário, não tem o foco nas variáveis estruturais nem em uma amenidade específica. Temos, então, o problema de decidir o espectro de variáveis ambientais a ser analisado.

Ainda que a seleção seja *ad hoc*, partimos para um critério um pouco mais intuitivo. Regredindo o valor do aluguel sobre as variáveis de controle da Tabela 2, obtemos um resíduo que incorporaria todas as externalidades urbanas e, ainda, algumas variáveis de controle não consideradas. Se aceitarmos este resíduo como *proxy* dos efeitos ambientais, ainda que estejamos conscientes de que uma parte de sua variância depende de fenômenos não ambientais, poderemos gerar curvas de nível de externalidades urbanas, de negativas a positivas, sobre a cidade.

Iniciamos regredindo o aluguel apenas sobre as variáveis estruturais (Tabela 3). Com isso, obtemos um resíduo que basicamente representa o valor da terra.<sup>13</sup> Excetuando-se alguns trechos, aparentemente viesados sob a influência de pontos isolados, o comportamento das curvas é bastante intuitivo.<sup>14</sup> Na região leste, norte e extremo oeste da cidade, prevalecem externalidades negativas, enquanto na região centro-oeste, considerada a área nobre da cidade, dominam as externalidades positivas. Os picos de gradientes de preços estão muito próximos dos pontos escolhidos como centro de negócios. Isto reforça o papel da proximidade aos locais de trabalho na formação de preços de aluguel. O pico ao redor do complexo Berrini-Faria Lima apresenta-se um pouco mais elevado se comparado à Praça da Sé. Essa diferença pode ser resultado das amenidades, já que o centro novo cresceu próximo a regiões de maior poder aquisitivo, enquanto o centro velho, como já falamos, sofre com a degradação provocada pelo excesso de densidade estrutural, trânsito e criminalidade. Voltaremos a este assunto quando incluirmos as variáveis ambientais.

<sup>13</sup> Ver Dubin (1992).

<sup>14</sup> Todos os mapas gerados encontram-se disponibilizados pelos autores.

**Tabela 3:** Regressão Linear sobre Variáveis de Controle

	<i>Estruturais</i>		<i>Monocêntrico</i>		<i>Duocêntrico.a</i>		<i>Duocêntrico.b</i>	
	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>
C	5.4269	0.8000	110.467	0.0000	141.730	0.0000	127.296	0.0000
SALA	98.0056	0.0000	86.3086	0.0000	88.1702	0.0000	89.6973	0.0000
GARAG	35.7380	0.0132	33.1132	0.0136	32.5380	0.0137	33.8935	0.0107
DORM_AG	85.1302	0.0000	79.5253	0.0000	75.1090	0.0000	75.5536	0.0000
COZ_AG	31.9309	0.0086	34.6248	0.0022	32.3206	0.0037	31.7707	0.0046
BAN_AG	88.8489	0.0000	86.1656	0.0000	85.5305	0.0000	85.3293	0.0000
OUT_DEP	-37.1034	0.0012	-18.2659	0.0915	-16.0606	0.1314	-18.3653	0.0859
DIST_SE			-8.4636	0.0000	-4.3816	0.0012		
DIST_BER					-4.9783	0.0000		
DIST_CBD							-9.3011	0.0000
DUM_SE							-14.7637	0.2467
R <sup>2</sup> adj	0.5786		0.6353		0.6477		0.6437	

Na segunda regressão adotamos o modelo monocêntrico, incorporando uma variável de distância ao centro velho (Tabela 3). Quando expurgamos o efeito acessibilidade, notamos uma menor diferenciação centro-periferia tanto no sentido leste quanto no sentido sudeste e norte. A região central agora apresenta resíduos num patamar negativo. A região sudoeste continua fortemente influenciada pelas amenidades positivas, enquanto as regiões ao extremo leste e sul passam a revelar sinais positivos. Isso talvez seja um efeito indesejado da introdução de uma variável de distância linear ao centro. A periferia talvez esteja sendo compensada excessivamente.

No caso duocêntrico com duas distâncias e com distância mínima (Tabela 3), ao expurgarmos o efeito distância ao complexo Berrini-Faria Lima, a região no entorno e no próprio centro novo permanece com valores bastante positivos. Isso confirma que o pico de gradiente no centro novo deve-se, em parte, à boa qualidade da zona residencial próxima.

O vetor de acessibilidade com duas distâncias acentua o efeito de compensação na vizinhança. Isso é um pouco amenizado quando utilizamos apenas a distância mínima. Por outro lado, regiões relativamente mais próximas ao centro novo, como oeste, sudoeste e sudeste, passam a revelar resíduos mais negativos. No caso das duas últimas regiões, a pesquisa confirma resultados esperados, entretanto, com relação à zona oeste, acreditamos que algumas observações podem não estar correspondendo à média dos distritos. Para amostras maiores, distorções locais como esta seriam compensadas pelas demais observações e poderíamos ter maior confiança nos resultados apresentados. Mais que isso, poderíamos assumir resultados inesperados como reveladores. Entretanto, a amostra da POF é representativa apenas do município como um todo. Assim, preferimos encarar com reservas resultados para os distritos.

Os resultados obtidos através dos mapas de resíduos, em geral, ratificam os fenômenos esperados e servem de indicadores para a seleção de variáveis ambientais. Para cada região procuramos encontrar a explicação adequada para o tipo de ocorrência. Na região extremo leste, norte e sul, acreditamos que parte do efeito negativo seja explicado pela violência e pelas condições precárias de moradia, como a presença de favelas; na região centro-oeste, alguns prováveis motivos de valorização devem ser, além da menor ocorrência de violência e de condições precárias de moradia, a concentração de áreas verdes. Para o centro histórico da cidade gostaríamos de testar o efeito saturação, entretanto não encontramos uma *proxy* satisfatória. O problema do trânsito poderia ser uma indicação, mas não havia uma medida que pudesse ser estendida para toda a cidade. A verticalidade das construções ou a densidade estrutural, por definição, seria a melhor variável para captar o uso excessivo do solo. Entretanto, enfrentaríamos o problema de endogeneidade, já que o

preço da terra condiciona o tipo de construção, pois zonas mais valorizadas tendem a ser mais verticalizadas.<sup>15</sup>

Além disso, incorporamos um conjunto de variáveis geralmente presentes nos estudos de preços hedônicos para habitação: uso do solo para fins comerciais e industriais, transporte público e poluição atmosférica. Resolvemos ainda incluir o zoneamento urbano. A política de zoneamento constantemente é alvo dos *rent-seekers*, terminologia econômica para os caçadores de renda, aqueles que conseguem auferir renda por mecanismo de pressão aos órgãos públicos. Na Tabela 4, estão listadas as variáveis ambientais selecionadas.

**Tabela 4:** Variáveis Ambientais

<i>Variável</i>	<i>Descrição</i>	<i>Média</i>	<i>Dv-padrão</i>
DIST_METR	distância à estação de metrô mais próxima em km	3.84	3.27
DIST_TREM	distância à estação de trem mais próxima em km	3.00	1.81
TCONST_C	área construída comercial sobre área construída total (*100) por distrito	17.17	10.01
TCONST_I	área construída industrial sobre área construída total (*100) por distrito	5.58	4.78
PART_IN	partículas inaláveis em $\mu\text{g}/\text{m}^3$ de acordo com estação medidora mais próxima	47.08	5.40
COEF_ARB	área arbórea sobre área urbana (*100) por distrito	6.21	5.25
CRIME	taxa de homicídios por 100.000 habitantes	54.87	25.81
POP_FAV	população favelada sobre pop. total (*100) por distrito	6.39	7.17
Z_ER	zona estritamente residencial	0.0121	0.1093
Z_RB	zona residencial de densidade demográfica baixa	0.7465	0.4355
Z_RM	zona residencial de densidade demográfica média	0.1408	0.3482
Z_M	zona mista	0.0805	0.2723
Z_I	zona industrial	0.0060	0.0775
Z_E	zonas especiais	0.0141	0.1180

Fonte: elaborado a partir de dados do LUME (2002), SEADE (2000), CETESB (1999) e SVMA (1994).

Todas as variáveis ambientais têm relação direta com o espaço e com a qualidade do bairro. Entretanto, em muitos estudos de preços hedônicos para habitação, os autores confundem qualidade da vizinhança com o nível sócio-econômico das pessoas que moram nessa mesma vizinhança. Assim, é comum a inclusão de variáveis sócio-econômicas, como renda, educação e saúde. Se o bairro é bom, então o preço médio dos aluguéis será elevado e, assim, apenas os indivíduos de maior poder aquisitivo poderão escolher esse local para moradia. Portanto, preço do aluguel explica renda, ou seja, novamente há um problema de endogeneidade. Ademais, como esperamos que as variáveis sócio-econômicas, como renda, estejam altamente correlacionadas com a presença de amenidades, sua inclusão torna difícil a estimação dos parâmetros ambientais.

Acreditamos que, de fato, algum efeito sócio-econômico influencia a escolha residencial das famílias. O “status social” de um bairro é um bom exemplo disso. Estratégias de marketing, além de evidenciarem a qualidade ambiental dos empreendimentos, por causa de sua localização, costumam

<sup>15</sup> Ver Brueckner (1987).

destacar o fato de eles se situarem em regiões emergentes. De qualquer maneira, considerando o que foi exposto, optamos por incluir apenas as variáveis estritamente ambientais ao modelo.

## 6. Resultados

Inicialmente, aplicamos o modelo para estimar o preço hedônico marginal das variáveis ambientais originais, de acordo com a equação (1). Foram testadas as três opções de vetores de acessibilidade: monocêntrico; duocêntrico(a), que inclui duas variáveis de distância; e duocêntrico(b), que considera apenas a distância mínima e uma *dummy* demonstrando a que centro se refere. Em seguida, selecionamos as variáveis ambientais mais significativas e construímos fatores, resumindo essa informação. Esses fatores, então, foram utilizados numa segunda rodada de estimações, como exposto na equação (4), novamente para os três casos de acessibilidade. Por último, investigamos a presença de dependência espacial no modelo. O resultados são apresentados nas Tabelas 5-7.

**Tabela 5:** Regressão Linear sobre Variáveis Ambientais

	<i>Monocêntrico</i>		<i>Duocêntrico.a</i>		<i>Duocêntrico.b</i>	
	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>
C	114.676	0.0492	156.866	0.0082	181.074	0.0046
SALA	71.0099	0.0000	71.0431	0.0000	71.4649	0.0000
GARAG	32.4128	0.0105	33.0117	0.0085	33.5320	0.0077
DORM_AG	68.5297	0.0000	67.4612	0.0000	68.2021	0.0000
COZ_AG	39.1733	0.0003	39.2764	0.0002	38.2536	0.0004
BAN_AG	77.8238	0.0000	76.5457	0.0000	77.3382	0.0000
OUT_DEP	-14.7701	0.1689	-12.9842	0.2226	-14.6740	0.1693
DIST_SE	-7.1298	0.0000	-0.0073	0.9978		
DIST_BER			-6.9341	0.0014		
DIST_CBD					-7.2613	0.0000
DUM_SE					-33.7772	0.1095
DIST_MET	2.4165	0.2957	0.4349	0.8543	2.0125	0.3755
DIST_TRE	-6.2551	0.0720	-5.5220	0.1093	-5.4355	0.1055
TCONST_C	0.6449	0.4328	0.3171	0.6991	0.2034	0.8071
TCONST_I	-1.4735	0.2350	-1.8638	0.1293	-1.6918	0.1677
PART_IN	0.8027	0.4515	1.2501	0.2407	0.7128	0.4987
COEF_ARB	4.4152	0.0001	0.1455	0.9335	0.5912	0.7222
CRIME	-0.9983	0.0068	-0.9665	0.0082	-1.0670	0.0030
POP_FAV	0.1100	0.9295	-1.3242	0.3122	-1.3503	0.3052
Z_ER	354.010	0.0000	343.342	0.0000	338.855	0.0000
Z_RM	1.2518	0.9443	0.8265	0.9628	2.0784	0.9066
Z_M	2.8441	0.9080	4.9350	0.8395	4.8928	0.8408
Z_I	-11.4153	0.8708	1.0192	0.9883	2.8386	0.9676
Z_E	-4.8121	0.9193	-6.4959	0.8901	-10.4534	0.8244
R <sup>2</sup> adj	0.6816		0.6878		0.6858	

A primeira rodada de regressões revela que as variáveis estruturais são para todos os casos bastante significativas. Como optamos pela função linear, os coeficientes revelam preços constantes. Assim, cada sala aumenta o valor do aluguel em aproximadamente R\$ 71,00, cada garagem em R\$ 33,00, cada dormitório em R\$ 68,00, e assim sucessivamente. Lembramos que os valores refletem o mercado em 1998, antes da ruptura cambial e da retomada da inflação, que impactaram os preços dos aluguéis nos últimos anos.

Considerando a configuração monocêntrica, observamos que o modelo se encaixa de acordo com o retratado na teoria. A distância em relação à região central apresenta coeficiente negativo e significativo, formando, assim, o gradiente de preços declinante em direção à periferia. Sob essa hipótese, o modelo revela apenas algumas variáveis ambientais significativas: distância às estações de trem (significativa a 10%), coeficiente arbóreo, crime e zona estritamente residencial.

Em grande parte dos estudos hedônicos para habitação, o impacto dos transportes públicos é pouco significativo. De acordo com Bowes e Ihlanfeldt (2001), esse resultado decepcionante revela, na verdade, o efeito ambíguo do fenômeno. As estações de trem e metrô podem estar contribuindo, simultaneamente, de forma positiva e negativa para a vizinhança. Uma estação beneficia a redondeza pelo aspecto de acessibilidade e fomento de atividades de comércio varejista. Por outro lado, aumenta a incidência de criminalidade e provoca na região adjacente mais distúrbios, como a poluição sonora. Biderman (2001) encontra resultados para a região metropolitana de São Paulo que destacam esse efeito diferenciado do metrô sobre as regiões adjacentes e as regiões mais afastadas. Os efeitos negativos tendem a prevalecer num raio próximo à estação, e os efeitos positivos passam a dominar a partir dessa fronteira. Neste trabalho, os resultados obtidos para estações de trem destacam seu benefício. Com relação às estações de metrô, acreditamos que a baixa significância e o sinal positivo para a distância podem estar revelando a ausência de informações ambientais adequadas para a região central. Como as linhas de metrô concentram-se no centro da cidade, é possível que a variável esteja captando outros efeitos não considerados sobre a região.

As taxas de construção comercial e industrial, embora os coeficientes apareçam com sinal esperado, não apresentam significância. No caso do comércio, a provável elevada correlação com distância à região central pode estar elevando a variância do parâmetro estimado. No caso da indústria, existe uma forte autocorrelação com a linha férrea. Podemos insistir nessas amenidades quando da construção dos fatores.

No caso da poluição, de fato a regressão não revela nenhuma significância possível. A principal razão para isso reside no fato de que existe pouca variância intra-urbana de partículas inaláveis. Portanto, a variável é praticamente uma constante.

A melhor estimativa de qualidade ambiental parece ser o coeficiente arbóreo. O coeficiente é positivo – cada ponto percentual a mais de área verde sobre área urbana aumenta o aluguel aproximadamente em R\$ 4,50 – e é significativo a 1%. A região mais nobre da cidade, a região centro-oeste, apresenta um dos mais elevados coeficientes arbóreos, acompanhada pela região sudeste e pela região sul, onde é definida a Área de Proteção aos Mananciais.

A criminalidade também aparece como um elemento decisivo na formação de preços para o mercado de aluguéis no Município de São Paulo. Os índices de violência na cidade afetam principalmente a periferia: regiões norte, leste e sudeste. Concomitantemente, a incidência de favelas ocorre preponderantemente nas mesmas regiões, mas, neste caso, a variável não apresenta significância. A elevada correlação entre a presença de favelas e criminalidade deve estar prejudicando a estimação dos coeficientes.

Quanto às *dummies* de zoneamento, o único resultado significativo foi a circunscrição de zonas estritamente residenciais (ZER). Esta parece ser a melhor *proxy* para o que chamamos de apazibilidade. As ZER são pequenas glebas de padrão de construção horizontal inseridas nos distritos que resguardam a tranquilidade dos moradores. Entretanto, acreditamos que o seu valor, R\$ 354,00 aproximadamente, dado os R\$ 353,20 de média dos aluguéis para toda a amostra, esteja superestimado. As ZER abrigam geralmente as residências de famílias mais abastadas, devendo estar, portanto, captando o efeito de variáveis omitidas, como, por exemplo, a qualidade de acabamento das construções.

A inclusão de uma variável de distância ao centro novo, no modelo duocêntrico(a), altera pouco o quadro das estimativas, e o ajuste do modelo tem uma melhora desprezível: o  $R^2$  ajustados passa de 0.6816 para 0.6878. Entretanto valem algumas observações. A distância ao centro histórico perde sua significância, e todo efeito acessibilidade passa a ser dirigido ao centro novo. A taxa de construção industrial passa a ser quase significativa a 10%, mas, por outro lado, o coeficiente arbóreo perde sua importância. Como havíamos discutido, os pontos de referência eleitos para distâncias podem estar representando picos de gradientes de preços que não necessariamente expressam acessibilidade. No caso, o centro novo, que se desenvolveu no interior do *cluster* de regiões privilegiadas da cidade, pode estar expressando, paralelamente ao efeito acessibilidade, a qualidade ambiental de seu entorno e, assim, extraindo resultados que deveriam ser captados pelo vetor ambiental.

Para o modelo duocêntrico(b), com uma variável de distância, os resultados permanecem estáveis. As variáveis estruturais e ambientais mantêm aproximados os valores dos coeficientes. O ajuste do modelo também pouco se altera, caindo para 0.6858. O gradiente de preços, representado pela variável DIST\_CBD, é negativo e significativo, mas a *dummy* para a região central não se revela importante.

Nossa segunda opção foi regressir o valor do aluguel sobre fatores que condensariam as informações presentes nas variáveis ambientais. A vantagem deste procedimento é escapar da multicolinearidade que aumenta a variância dos parâmetros e reduz sua significância. A desvantagem é que os fatores nem sempre apresentam algum significado econômico intuitivo. Assim, corremos o risco de estar estimando o preço de algo pouco compreensível. Para facilitar esta tarefa, primeiro optamos pela rotação *varimax*, que maximiza a variância de cada variável num mesmo fator e por transformar todas as variáveis com conotação negativa – as “desamenidades” – no seu inverso. Feitas algumas combinações, conseguimos construir fatores relativamente compreensíveis, como mostra a Tabela 6.

**Tabela 6:** Matriz de Cargas Fatoriais Rotacionadas

	<i>Fatores</i>		
	1	2	3
MET_INV	0.458	-0.058	0.103
TRE_INV	0.021	0.006	-0.056
IND_INV	0.853	0.033	-0.088
TCONST_C	0.758	-0.026	-0.069
COEF_ARB	-0.177	0.978	-0.114
CRI_INV	0.367	0.213	0.905
Z_ER	0.011	0.119	0.023

A terminação INV indica o inverso de uma variável. Assim, temos o inverso da distância às estações de metrô, o inverso às estações de trem, o inverso da taxa de construção industrial e o inverso do crime. Taxa de construção industrial, coeficiente arbóreo e zona estritamente residencial permanecem como estavam.

O inverso de criminalidade aparece razoavelmente representado em todos os fatores, entretanto sua incidência é preponderante, e praticamente exclusiva, no Fator 3. Assim, “batizamos” o terceiro fator de “segurança”. O Fator 1 concentra as variâncias do inverso da distância ao metrô, do inverso da presença de atividade industrial e da presença de atividade comercial. Portanto, é um fator associado à proximidade da região mais urbanizada, à presença de infraestrutura. Chamamos esse fator, então, de “infra-estrutura”. O segundo fator também é bastante evidente. Praticamente toda a



variância é motivada pela presença de áreas verdes, restando um pouco para ZER. Denominamo-lo, então, de “aprazibilidade”.

Rodando a regressão com os fatores para os três casos de acessibilidade, chegamos aos resultados expostos na Tabela 7. Observamos, primeiramente, que os preços marginais das variáveis estruturais permanecem estáveis. É interessante notar que a ausência de amenidades primárias eleva a significância da constante, que agora possui um valor inferior, caindo de R\$ 150,00 para R\$ 100,00 na média. Como dissemos, a exígua diferença intra-urbana de poluentes torna PART\_IN praticamente uma segunda constante, perturbando a variância do coeficiente da constante original.

**Tabela 7:** Regressão Linear sobre Fatores

	<i>Monocêntrico</i>		<i>Duocêntrico.a</i>		<i>Duocêntrico.b</i>	
	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>	<i>coef.</i>	<i>prob.</i>
C	85.0424	0.0007	102.448	0.0001	108.244	0.0001
SALA	78.2854	0.0000	79.8526	0.0000	80.1177	0.0000
GARAG	33.4091	0.0010	32.7800	0.0109	34.3425	0.0077
DORM_AG	75.6848	0.0000	73.6109	0.0000	74.1721	0.0000
COZ_AG	35.4688	0.0013	35.4914	0.0012	34.6267	0.0015
BAN_AG	82.3986	0.0000	82.4077	0.0000	82.0941	0.0000
OUT_DEP	-2.7021	0.8050	-0.9902	0.9275	-1.5819	0.8844
DIST_SE	-5.8936	0.0000	-1.0157	0.6236		
DIST_BER			-5.1934	0.0046		
DIST_CBD					-6.0609	0.0000
DUM_SE					-39.0400	0.0496
F1	18.3058	0.0182	22.3779	0.0043	23.5992	0.0023
F2	26.3504	0.0000	7.1443	0.4157	3.9966	0.6443
F3	26.3834	0.0000	30.8029	0.0000	33.9533	0.0000
R <sup>2</sup> adj	0.6611		0.6660		0.6659	

Para os três casos de acessibilidade, os resultados com fatores acompanham os apresentados na Tabela 5. No modelo monocêntrico, a distância ao centro histórico é negativa e significativa a 1%. Todos os fatores são altamente significativos e, como esperávamos, com valores positivos. “Infra-estrutura” possui um preço marginal de R\$ 18,30, “aprazibilidade”, de R\$ 26,35 e “segurança”, de R\$ 26,38.

Para o modelo duocêntrico(a), novamente a distância à Praça da Sé perde significância, enquanto a distância ao Complexo Berrini-Faria Lima mantém-se significativo a 1%, com sinal negativo. Neste caso, o fator “aprazibilidade”, composto essencialmente da variância de coeficiente arbóreo, perde significância, repetindo o caso das variáveis ambientais originais. Os outros fatores permanecem relevantes.

As alterações são quase imperceptíveis para o modelo duocêntrico(b). Apenas os fatores “infra-estrutura” e “segurança” revelam-se significativos – R\$ 23,60 para o primeiro e R\$ 33,95 para o segundo. A variável de acessibilidade implica no gradiente de preços decrescente, mas agora a *dummy* relativa ao centro histórico é significativa a 5%. O intercepto do centro histórico é R\$ 39,00 menor. É provável que a perda parcial de informações ambientais tenha sido transferida para o vetor de acessibilidade, configurando, deste modo, a degradação do centro histórico.

Concluídas as estimações, passamos para os testes de dependência espacial. Utilizamos o teste LM, que exhibe isoladamente a necessidade de incorporação de uma variável dependente defasada ( $LM_{lag}$ ) ou a reespecificação do resíduo ( $LM_e$ ). Para os dois casos experimentamos a matriz de defasagens

especiais  $W$ , com o inverso das distâncias e o inverso do quadrado das distâncias. Os resultados para todas as equações estimadas estão na Tabela 8.

**Tabela 8:** Análise de Dependência Espacial

Regressões	$W: d^1$		$W: d^2$	
	$LM_e$ ( <i>prob</i> )	$LM_{lag}$ ( <i>prob</i> )	$LM_e$ ( <i>prob</i> )	$LM_{lag}$ ( <i>prob</i> )
(1) $p = \alpha_0 + \alpha_1 S$	0.0000	0.0000	0.1700	0.3473
(2) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Mono})$	0.0073	0.4499	0.3410	0.7453
(3) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Duo.a})$	0.0668	0.1768	0.3542	0.6112
(4) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Duo.b})$	0.0332	0.0287	0.3774	0.5860
(5) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Mono}) + \alpha_3 E$	0.2332	0.7779	0.3444	0.9700
(6) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Duo.a}) + \alpha_3 E$	0.2985	0.8699	0.3334	0.9966
(7) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Duo.b}) + \alpha_3 E$	0.0509	0.5770	0.4050	0.9462
(8) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Mono}) + \alpha_3 F(E)$	0.5203	0.8338	0.5402	0.6262
(9) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Duo.a}) + \alpha_3 F(E)$	0.5780	0.6765	0.5098	0.5655
(10) $p = \alpha_0 + \alpha_1 S + \alpha_2 A(\text{Duo.b}) + \alpha_3 F(E)$	0.5047	0.3804	0.5622	0.5346

As regressões para as variáveis ambientais expostas na Tabela 5 estão dispostas nas linhas 5-7, e as regressões com fatores apresentados na Tabela 7 estão dispostas nas linhas 8-10. Em todos os casos, tanto para  $W: d^1$  como para  $W: d^2$ , a reespecificação para um modelo hedônico espacial não se revela necessária. Portanto, os preços marginais estimados anteriormente sustentam-se.

Entretanto, é interessante observar que a dependência espacial do modelo vai regredindo à medida que incluímos informações espacializadas, conforme havíamos adiantado. Por exemplo, o teste de autocorrelação nos resíduos para a matriz do inverso das distâncias  $W: d^1$  vai perdendo significância a cada linha de regressões. Na regressão 1, composta apenas por variáveis explicativas estruturais, o *p-value*, que é de 0.0000, vai para 0.0073, com a introdução da distância ao centro histórico, na regressão 2. Com a inclusão de mais uma variável de distância, na regressão 3, o *p-value* já é de 0.0668. Quando introduzimos, então, o vetor de variáveis ambientais, o teste já não é mais significativo: 0,2332 na regressão 5 e 0,2985 na regressão 6. Portanto, o problema de dependência espacial pode ser minimizado desde que disponibilizemos informações geográficas relevantes para o modelo.

## 7. Considerações Finais

Para estimar os preços implícitos das amenidades urbanas para a cidade de São Paulo, enfrentamos o dilema de aceitar a baixa significância dos parâmetros na regressão sobre o vetor de variáveis ambientais iniciais, ou construir fatores que corrigiriam a não-ortogonalidade das variáveis e, portanto, melhorariam a sua significância, mas gerariam resultados de menor intuição econômica.

No primeiro caso, foi possível revelar que a presença de estações de trem, áreas verdes e zonas estritamente residenciais contribuem para a elevação do aluguel, enquanto a criminalidade reduz o seu valor. A baixa significância das estações de metrô e da atividade comercial pode ser duplamente explicada pela forte multicolinearidade das variáveis que compõem o vetor e pelo próprio efeito ambíguo dos fenômenos, que contribuem ao mesmo tempo positiva e negativamente para a valorização dos imóveis. Por outro lado, a poluição definitivamente não parece influir na formação de preços. O mesmo não se pode concluir da presença de favelas, dado que existe uma forte correlação entre esta e criminalidade.

O modelo com fatores, além de repetir qualitativamente os resultados apresentados anteriormente, revela a implicação da oferta de infra-estrutura urbana sobre a moradia. Neste caso, podemos concluir que, de fato, regiões mais centralizadas, servidas de comércio, serviços e transporte público, apresentam valores proporcionalmente mais elevados.

A construção de curvas de nível baseadas nos resíduos da regressão sobre variáveis de controle ajudou na seleção das variáveis explicativas. Este método pode contribuir para pesquisas com modelos hedônicos para habitação, bem como para um conjunto de outros estudos espaciais. Sua utilização deve ser acompanhada por um conjunto de cartas temáticas sobre a região de estudo para que se possa extrair alguma informação relevante sobre o fenômeno.

Entretanto, nem sempre as variáveis desejadas estão disponíveis. Algumas variáveis de controle estão ausentes, como a idade do imóvel e a qualidade de acabamento. Entre as variáveis ambientais, não encontramos uma *proxy* adequada para saturação urbana, como trânsito e densidade estrutural. Existe ainda o problema da escala selecionada. No caso, a maioria das informações está disposta por distrito, o que provavelmente não revela a verdadeira percepção de vizinhança de cada observação.

Quanto aos problemas econométricos levantados: sobre o problema de especificação, resolvemos por simplicidade assumir um modelo linear; o efeito de multicolinearidade foi resolvido com a construção de fatores; e a possibilidade de dependência espacial foi testada. Apenas omitimos a análise de heterocedasticidade espacial. Entretanto, seus efeitos restringem-se à confiabilidade do nível de significância dos parâmetros, sem impor nenhum viés sobre as estimativas.

## 8. Referências

- ALONSO, W. (1964). *Locational and Land Use*. Cambridge: Harvard University Press.
- ANDERSSON, D. E. (1997). *Hedonic Prices and Center Accessibility: Conceptual Foundations and an Empirical Hedonic Study of the Market for Condominium Housing in Singapore*. Taipei: National Taiwan University.
- ANSELIN, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht, Kluwer Academic Publishers.
- BARTIK, T.; SMITH, K. (1987). "Urban Amenities and Public Policy". *Handbook of Regional and Urban Economics*, v. 2: 1207-1254.
- BENSON, E. D.; HANSEN, J. L., SCHWARTZ, A. L. Jr, SMERSH, G. T. (1998). "Pricing Residential amenities: The Value of a View". *Journal of Real State Finance and Economics*, 16:1, 55-73.
- BIDERMAN, C. (2001). *Forças de Atração e Expulsão na Grande São Paulo*. São Paulo: Fundação Getúlio Vargas.
- BOWEN, W. M.; MIKELBANK, B. A.; PRESTEGAARD, D. M. (2001). "Theoretical and Empirical Considerations Regarding Space in Hedonic Housing Price Model Applications". *Growth and Change*. v. 32, n. 4 (Fall 2001): 466-490.
- BOWES, D. R.; IHLANFELDT, K. R. (2001). "Identifying the Impacts of Rail Transit Stations on Residential Property Values". *Journal of Urban Economics*, 50: 1-25.
- BRUECKNER, J. (1987). "The Structure of Urban Equilibria: a Unified Treatment of the Muth-Mills Model". *Handbook of Regional and Urban Economics*. v. 2: 821-845.
- CAN, A. (1990). "The Measurement of Neighborhood Dynamics in Urban House Prices". *Economic Geography*, 66, 254-272.
- CHESHIRE, P.; SHEPPARD, S. (1998). "Estimating the Demand for Housing, Land and Neighborhood Characteristics". *Oxford Bulletin of Economics and Statistics*, 60 (3): 357-382.
- DUBIN, R. A. (1992). "Spatial Autocorrelation and Neighborhood Quality". *Regional Science and Urban Economics*, 22, 433-452.
- ESPEY, M. e LOPEZ, H. (2000). "The Impact of Airport Noise and Proximity on Residential Property Values". *Growth and Change*. v. 31 (Summer 2000): 408-419.
- FIPE (1999). *Pesquisa de Orçamentos Familiares*. São Paulo, Universidade de São Paulo
- FUJITA, M. e OGAWA, H. (1982). "Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configuration". *Regional Science and Urban Economics*, 12, 161-196.

- HALVORSEN, R. e POLLAKOWSKI, H. O. (1981). "Choice of Functional Form for Hedonic Price Equations". *Journal of Urban Economics*, 10:37-49.
- HARMAN, H. (1967). *Modern Factor Analysis*. Chicago: The University of Chicago Press.
- KAIN, J. F. e QUINLEY, J. M. (1970). "Measuring the Value of Housing Quality". *Journal of the American Statistical Association*, 65:532-548.
- MACEDO, P. B. R. (1998). "Hedonic Price Models with Spatial Effects: an application to the house market of Belo Horizonte, Brazil". *Revista Brasileira de Economia*, 52(1), 63-81.
- MILLS, E. (1972). *Studies in the Structure of the Urban Economy*. Baltimore: Johns Hopkins University Press.
- OLIVEIRA, R. G. de (1997). *Dois Estudos sobre a Poluição do Ar na Cidade de São Paulo*. São Paulo: Universidade de São Paulo.
- PLAUT, P. O.; PLAUT, S. E. (1998). "Endogenous Identification of Multiple Housing Price Centers in Metropolitan Areas". *Journal of Housing Economics*, 7, 193-217.
- ROSEN, S. (1974). "Hedonic Price and Implicit Markets: Product Differentiation in Pure Competition". *Journal of Political Economics*, 82: 34-55.
- SHEPPARD, S. (1999). "Hedonic Analysis of Housing Markets". *Handbook of Regional and Urban Economics*. v. 3, 1.595-1.635.
- WHEATON, W.C. (1974). "Comparative Static Analysis of Urban spatial Structure". *Journal of Economic Theory*, 9:223-237.