



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CURSO DE PÓS-GRADUAÇÃO EM ECONOMIA**  
**MESTRADO PROFISSIONAL EM ECONOMIA**

**LORENA CANDIDO FERNANDES DE ALMEIDA**

**MODELO DE DEMANDA HABITACIONAL PARA A CIDADE DE SÃO PAULO**

**FORTALEZA**

**2021**

LORENA CANDIDO FERNANDES DE ALMEIDA

MODELO DE DEMANDA HABITACIONAL PARA A CIDADE DE SÃO PAULO

Dissertação apresentada ao Mestrado Profissional em Economia Empresarial da Universidade Federal do Ceará. Como requisito parcial à obtenção do título de Mestre. Área de concentração: Economia de empresas .

Orientador: Prof.º Dr. Sérgio Aquino de Souza

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

D1m DE ALMEIDA, LORENA CANDIDO FERNANDES.  
Modelo de demanda habitacional para a cidade de São Paulo / LORENA CANDIDO FERNANDES DE ALMEIDA. – 2021.  
26 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Economia, Administração, Atuária e Contabilidade, Mestrado Profissional em Economia de Empresas, Fortaleza, 2021.  
Orientação: Prof. Dr. Sérgio Aquino de Souza.

1. Modelos de preços hedônicos. 2. Demanda por habitação. 3. Preços hedônicos. I. Título.

CDD 330

---

LORENA CANDIDO FERNANDES DE ALMEIDA

MODELO DE DEMANDA HABITACIONAL PARA A CIDADE DE SÃO PAULO

Dissertação apresentada ao Mestrado Profissional em Economia Empresarial da Universidade Federal do Ceará. Como requisito parcial à obtenção do título de Mestre. Área de concentração: Economia de empresas.

Aprovada em: 29/06/2021.

BANCA EXAMINADORA

---

Prof. Dr. Sérgio Aquino de Souza (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Andrei Gomes Samonassi (Membro)  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Eveline Barbosa Silva Carvalho (Membro)  
Universidade Federal do Ceará (UFC)

## **AGRADECIMENTOS**

A Deus, pelo dom da vida e por sua infinita graça e misericórdia, que tem transbordado em minha vida.

A minha querida mãe Fátima, seu exemplo foi minha principal inspiração. Sua força e coragem me impulsionou. Você foi a mãe possível de ser.

A Tércio, meu companheiro incansável e admirador eterno, que cuidou incansavelmente dos nossos filhos enquanto eu estudava. Seria muito mais difícil sem você.

Aos meus filhos Heitor, Enzo e Olívia, com quem eu compartilho o melhor desta vida. Gratidão pela paciência e espera neste tempo de entrega a este projeto.

Ao meu irmão Igor, por me amar sem reservas. E por me emprestar sua casa para me refugiar e escrever esta dissertação.

Ao professor doutor Sérgio Aquino, pelo incentivo e confiança depositados neste desafio. Sem o seu apoio eu não teria concluído com sucesso e no prazo estabelecido.

Aos membros da banca, Andrei e Eveline, pela gentileza em colaborar com esta dissertação.

## RESUMO

Este trabalho consiste em apresentar a análise dos principais atributos na formação de preços de imóveis residenciais para venda na cidade de São Paulo, SP, em 2019. Selecionamos o modelo de preços hedônicos que tem sido apresentado na literatura como o de melhor acurácia para previsão no mercado imobiliário. Estimou-se o modelo a partir de uma regressão linear múltipla, que permitiu associar o preço dos imóveis às características estruturais e locacionais dos bens. Os resultados sugerem que de acordo com a base de dados se trata de um mercado produto diferenciado, em que as variáveis de lazer e de localização do imóvel foram relevantes para explicar a variabilidade do preço, e refletiram a realidade do mercado imobiliário da região analisada.

**Palavras-Chaves:** modelos de preços hedônicos; demanda por habitação; preços hedônicos

## **ABSTRACT**

This work consists in presenting the analysis of the main attributes in the formation of prices of residential properties for sale in the city of São Paulo, SP, in 2019. We selected the hedonic pricing model which has been presented in the literature as the best accurate one for forecasting in the real estate market. The model was estimated from a multiple linear regression, which allowed associating the house prices with the structural and locational characteristics of such goods. The results suggest that according to the database is a differentiated product market, in which leisure and property location variables were relevant to explain the price variability, and they reflected the reality of the real estate market in the analyzed region.

**Keywords:** hedonic pricing models; demand for housing; hedonic prices

## LISTA DE GRÁFICOS

GRÁFICO 1 - RESÍDUOS.....	22
---------------------------	----

## LISTA DE TABELAS

TABELA 1 - ESTATÍSTICAS DESCRITIVAS POR CARACTERÍSTICA DAS VARIÁVEIS DE CONTROLE .	19
TABELA 2 - AJUSTAMENTO DO MODELO DE DEMANDA HABITACIONAL PARA A CIDADE DE SÃO PAULO .....	21
TABELA 3 - MATRIZ DE CORRELAÇÃO .....	21
TABELA 4 - COMPARATIVO ENTRE O MODELO DE ESTIMADORES MQO E NEW-WEST .....	23

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>10</b>
<b>2</b>	<b>REFERENCIAL TEORICO .....</b>	<b>12</b>
2.1	MODELO DE PREÇOS HEDÔNICO .....	13
<b>3</b>	<b>DESCRIÇÃO DOS DADOS E ESTIMAÇÃO DO MODELO .....</b>	<b>16</b>
3.1	SELEÇÃO DAS VARIÁVEIS .....	17
<b>4</b>	<b>RESULTADOS EMPIRICOS .....</b>	<b>19</b>
<b>5</b>	<b>CONCLUSÃO E CONSIDERAÇÕES FINAIS .....</b>	<b>25</b>

## 1 INTRODUÇÃO

O mercado imobiliário já representa parcela significativa do Produto Interno Bruto (PIB), como também tem sido responsável pelas transformações estruturais das grandes metrópoles, nas últimas décadas. Mudanças estruturais implementadas entre os anos de 2004 e 2014, como a avanço de normas e leis que estabeleceram o Estatuto da Cidade (Lei 10.257/2001), e a criação do Ministério das Cidades em 2003, além das políticas de expansão ao crédito imobiliário e de subsídios habitacionais, impulsionaram profundamente os investimentos no setor.

Tais avanços vem proporcionando ligeira redução do déficit habitacional, mas que segundo Fundação João Pinheiro (1995), ainda constitui um dos principais problemas sociais do país. A estimativa de déficit para o Brasil, em 2019, foi de 5,876 milhões de domicílios, responsável por 8% do estoque total de domicílios adequados nesse ano. Em 2015, o déficit habitacional brasileiro era de 6,187 milhões de moradias, o que representava 9,3% do estoque de moradia adequadas nesse ano.

O direito a habitação é caracterizado como necessidade básica e fundamental na promoção da estabilidade e segurança familiar. O consumo habitacional é, portanto, inerente ao ser humano e pode ser agregado em dois grandes grupos: o primeiro são aqueles que adquirem o bem para atender sua necessidade imediata por moradia, e o segundo que adquirem para atender aos anseios de compor uma cesta de bens de investimento.

Neste sentido, para compreender os diferentes aspectos da demanda por habitação, a literatura divide os modelos teóricos em dois pilares: o macroeconômico, que utiliza um sistema de formação de preço mais complexo e robusto; e o microeconômico, que se atem às características do comportamento individual do consumidor que contribuem para a formação de preço de habitação.

Tais atributos devem ser elaborados a partir das características extrínsecas e intrínsecas do ativo e não somente por características físicas do bem, como tem-se observado na maioria das avaliações imobiliárias. Ratificado pela Síntese de Indicadores Sociais (2018), do IBGE, que contextualiza que, para a análise das condições habitacionais: [...] a área construída do domicílio, o número médio de pessoas por domicílio e por dormitório, o acesso a água e ao esgotamento sanitário, o estado de conservação e a condição de ocupação são, em linhas gerais, indicadores mais eficazes do nível de bem-estar das moradias do que propriamente o tipo de material utilizado na construção do domicílio”.

Um dos primeiros autores que contribuiu para a abordagem metodológica apresentando a curva de oferta e demanda em que os preços são funções das características do bem e não do bem em si, foi Rosen em 1974. Tais preços foram chamados de preços implícitos ou preços hedônicos. O método de preços hedônicos foi desenvolvido, inicialmente, por Court (1939) que estabelece a importância relativa dessas variáveis implícitas na formação dos preços, mensurando a disposição marginal a pagar, de um consumidor, dado características intangíveis de um bem, como uma vista para o mar, a distância até o shopping mais próximo e a disponibilidade de serviços nas proximidades.

Particularmente, tal metodologia de análise microeconômica do mercado imobiliário, permite estimar as variáveis que impactam na formação de preços que acrescentados determinadas características sociodemográficas da população, possibilita prever a função demanda por moradia. E a partir desta pode-se compreender a relação entre o consumo individual das famílias por habitação e a sua renda ou no próprio preço da habitação, ou seja, conhecer a elasticidade preço-demanda da habitação. Assim contribuir para ampliar as ferramentas que possibilitem estabelecimento de políticas públicas que contribuam para o bem-estar da coletividade e dos indivíduos.

O objeto de estudo aqui escolhido é a Região Metropolitana de São Paulo (RMSP), que é a 8ª cidade mais populosa do planeta, com cerca de 21 milhões de habitantes, sendo a 10ª maior aglomeração urbana do mundo, chegando a abrigar cerca de 5,9% de toda a população brasileira (11,3 milhões de habitantes). É considerado o principal centro financeiro, corporativo e mercantil da América do Sul. O Produto Interno Bruto (PIB), em 2010, da RMSP foi da ordem de R\$ 701 bilhões, cerca de 56% do PIB do Estado, e cerca de 20% de toda a arrecadação nacional; considerado o maior polo de riqueza do Brasil.

Neste cenário, a presente pesquisa pretende oferecer uma contribuição objetiva ao tema, analisando o mercado imobiliário residencial da RMSP, no Estado de São Paulo no ano de 2021. Que além desta introdução o trabalho apresentará uma segunda seção que trata do referencial teórico de suporte às teorias de formação de preços e dos fatores que influenciam a demanda por imóveis residenciais. A terceira que consiste na metodologia adotada no trabalho, com informações sobre as variáveis utilizadas na análise econométrica. E finalmente, na quarta e quintas seções apresentaremos, respectivamente, os resultados empíricos das análises e as considerações finais e perspectivas para o mercado futuro.

## 2 REFERENCIAL TEÓRICO

O setor imobiliário tem desempenhado papel importantíssimo na economia dos países industrializados; sua relação com os principais indicadores econômicos tem se dado particularmente pela geração de riqueza de um país ou região; pelo seu efeito multiplicador nos demais setores da economia; pela criação de novos postos de trabalho; e pela vinculação entre o crescimento econômico e a atividade do setor. (TABALES, 2007)

O mercado de habitação residencial, pensado para além da produção habitacional, na atualidade, tem ocupado papel fundamental de produtor do espaço urbano, transformando as cidades em espaços de materialização dos interesses individuais e coletivos da sociedade onde a racionalidade das classes sociais se manifesta no decorrer do processo de produção do espaço urbano. (MELO, 2020)

No Brasil, o mercado imobiliário passou por profundas transformações, principalmente entre os anos de 2004 e 2014, que impulsionou o aumento da produção de imóveis residenciais. A compreensão do funcionamento do mercado, a partir da análise dos preços das habitações, tem sido fundamental para a elaboração e implementação de políticas habitacionais, urbanas e de expansão do crédito e aumento de investimentos para redução do déficit habitacional. (DANTAS; MAGALHÃES; VERGOLINO, 2010)

O bem habitação distingue-se dos demais mercados da economia, por apresentar características peculiares e formar um mercado heterogêneo, cíclico, dominado por grandes agentes econômicos, e porque sua demanda é afetada diretamente pelas alterações políticas, econômicas e climáticas.

As unidades habitacionais guardam relação, não só com o espaço urbano onde está inserida, como também existe relação espacial de dependência entre os imóveis próximos, que por sua vez exibem características estruturais semelhantes; já que, muitas vezes, foram projetados e construídos ao mesmo tempo. E, portanto, como aponta Heslin *et al.* (2009), as análises empíricas têm demonstrado que há forte dependência espacial entre os resíduos do preço dessas residências.

Neste sentido, a literatura tem abordado diversas técnicas de modelagem dos preços de imóveis; elas vão desde modelos tradicionais com abordagens convencionais que relacionam os fatores de decisão de compra do imóvel com sua localização territorial, devido a facilidade de acesso a centros de negócios e serviços, assim como à externalidades ambientais e culturais; passando por modelos dinâmicos e pela metodologia de preços hedônicos, que apesar de suas

limitações tem sido o mais utilizado e será o foco da presente pesquisa; e finalmente, os mais recentes estudos de redes neurais no campo da Inteligência Artificial.

Em termos gerais, os modelos de preços hedônicos vêm ocupando a vanguarda nas pesquisas sobre métodos preditivos de preços das habitações, por melhor se adaptar à construção de índices de preços, análises de bolhas imobiliárias e desempenho de carteiras de investimento; como também tem sido utilizado para avaliar o impacto das externalidades sobre o valor dos imóveis. Porém, um limitador observado na comparação entre os diversos métodos utilizados, é que a natureza dos dados obtidos guarda forte relação com os resultados alcançados. Diversos autores estabelecem comparação entre diversas técnicas como forma de melhorar a previsibilidade das estimativas.

O modelo hedônico tem sido escolhido por sua capacidade preditiva do preço de mercado do imóvel a partir de um conjunto de atributos individuais cujos preços implícitos contribuem para formação do preço de cada unidade habitacional. A forma funcional comumente utilizada é a *semi-logarítmica*, a partir da técnica dos Mínimos Quadrados Ordinários (MQO), por se adaptar bem aos dados e porque os coeficientes estimados podem ser interpretados como sendo a contribuição relativa do bem que é diretamente atribuída as respectivas características deste bem. (SELIM, 2009)

## 2.1 Modelo de preços hedônico

A contribuição preliminar, para os modelos teóricos de estudos de preço no mercado habitacional, foi oferecida por Lancaster (1966), que propôs amplificar a abordagem da Teoria do Consumidor, onde tradicionalmente o consumidor exerce preferências, maximizando a utilidade; expandindo tal medida de análise de consumo de um bem para suas características intrínsecas e não somente a partir do bem em si.

Referida abordagem propõe que o consumo de um bem pode ser expresso de forma linear, assumindo relação entre o nível de atividade  $k$ , o número de bens consumidos  $y$  e o  $j$ -ésimo bem consumido  $x_j$ , como segue:

$$x_j = \sum_k a_{jk} y_k \quad (1)$$

Supondo ainda que  $z_i$  é a quantidade da  $i$ -ésima característica adquirida em uma atividade de consumo, assim:

$$z_i = \sum_k b_{ik} y_k \quad (2)$$

Nota-se que a relação entre os vetores  $x$  e  $z$ , que são respectivamente, relação de bens de consumo com as preferências dos consumidores, é medida pelo vetor  $y$ , que é exatamente a atividade de consumo.

Como simplificação do modelo, Lancaster sugere que a escolha do consumidor é dada pelo “pacote” de atributos de um bem  $(x_1, \dots, x_n)$  definido por  $Z = B \cdot x$ , onde  $Z = [Z_i]$  é o vetor de características,  $X = [X_i]$  o vetor de bens, e  $B = [b_{ij}]$  a matriz dos coeficientes que relacionam bens as características. A função utilidade (estritamente concava), considerando um ambiente de equilíbrio competitivo entre os agentes, é definida no espaço amostral do vetor  $z_i$ , a saber:  $U(x, z_1, z_2, \dots, z_n)$ , em que  $x$  representa o valor disponível para a compra de outros bens consumidos. Portanto, a maximização da utilidade dependerá exatamente da escolha ótima de  $x$  e  $(z_1, z_2, \dots, z_n)$  que satisfaça a restrição orçamentária  $y = x + p(z)$ , formalmente definida:

$$\text{Max } U(x, z), \text{ sujeito a } y = x + p(z) \quad (3)$$

Rosen (1974) foi o responsável por ampliar a abordagem de preços hedônicos ao contexto de mercado, apresentando as equações de oferta e demanda em que os preços são funções das características; tornando-o em instrumento de avaliação mais utilizado para modelar o mercado imobiliário urbano.

Griliches (1971) corroborando com os demais, citado por Fávero (2003, p. 10), define que preços hedônicos se referem à construção de índices baseada em modelos econométricos, em que o preço de um determinado bem é explicado pelas características pertencentes ao mesmo e, portanto, o que representa utilidade ao consumidor não é o bem em si, mas seus respectivos atributos.

Tais atributos podem ser um conjunto de características específicas de uma região influenciando na percepção positiva ou negativa para a satisfação dos indivíduos. Podem fazer parte desta cesta, não só os recursos naturais, como praia, área verde, clima etc.; mas também questões estruturais provocadas pelo próprio homem, tais como trânsito, poluição, segurança, oferta de serviços e entretenimentos etc.

Segundo Selim (2009) a teoria dos preços hedônicos foi formulada como um problema, na economia, de equilíbrio espacial, onde todo o conjunto de preços implícitos

orientam a tomada de decisão tanto de consumidores como de produtores no espaço de suas características.

De acordo com Dantas (2010): pela equação de preços hedônicos, podem ser obtidos os preços implícitos ou hedônicos de cada característica individual da habitação, que refletem a utilidade que cada uma gera para o consumidor, como também o preço total que ele está disposto a pagar pela cesta de serviços de habitação com que se defronta no mercado.

No mercado imobiliário se transaciona bens imóveis com atributos diferentes (bens heterogêneos), onde o preço de equilíbrio reflete a quantidade de atributos deste bem; e quanto melhor o atributo, maior o preço que o consumidor atribuirá a ele. Desta forma, a análise da utilidade do bem no mercado habitacional se torna mais realista, ampliando as possibilidades de um modelo mais assertivo.

Nas últimas três décadas de aplicação desses modelos de precificação de residências, a principal lição aprendida, segundo Cheshire & Sheppard (2004), citado por Alves *at al.* (2011), é a de que o valor do imóvel varia sistemática e substancialmente com a sua localização. Por isso, características da localização são tão importantes quanto aos atributos físicos da casa, como quantidade de quartos, lavatórios e vagas de garagem etc., na determinação de seu preço de mercado.

É sabido que informações sobre preço e quantidade das externalidades não são conhecidas no mercado, parecendo um limitador para a implementação do modelo. Contudo, para Hermann (2003), podemos assumir que existe implicitamente uma oferta e uma demanda por essas características e, dessa forma, tentar inferir seus respectivos preços de equilíbrio. Roteiro este, traçado pelo modelo de preços hedônicos, onde o preço marginal dos atributos é estimado, fazendo regredir o preço do bem heterogêneo sobre as quantidades de características associadas.

A evolução das técnicas de modelagem do mercado habitacional tem se dado a partir da abordagem econométrica dos MQO, que quando utilizada com parcimônia e comparativamente com outros métodos, tem apresentado resultados significativos para explicar os movimentos dos agentes neste setor. Como a questão espacial tem sido apresentada como relevante para explicação dos modelos, e como os dados associados a posição geográfica são caracterizados pela heterogeneidade espacial ou dependência, os resultados obtidos podem apresentar problemas de tendenciosidade. A solução apresentada nas pesquisas tem sido proceder com a análise de modelos estimados pela metodologia denominada Econometria Espacial como e mais recentemente a análise de redes neurais tem sido apresentada como alternativa que melhor prediz os preços de imóveis, segundo Selim (2009).

### 3 DESCRIÇÃO DOS DADOS E ESTIMAÇÃO DO MODELO

A escolha por um determinado bem, segundo a teoria clássica do consumidor, é baseada nas suas características intrínsecas. Logo, a função utilidade será atribuída conforme o grau de satisfação que o consumidor obtém a partir de uma combinação de bens, em um período em que ele será consumido. Segundo Henderson e Quandt (1988, p.8), citado por Fávero (2003): “normalmente, o consumidor obtém utilidade da variedade e da diversificação de bens que consome e, desta forma, a função utilidade não deve ser definida para um período tão curto que não seja possível satisfazer seu desejo de variedade”.

O modelo de preços hedônicos tem sido amplamente utilizado para estabelecer o comportamento da demanda por bens diferenciados como automóveis, vinhos e computadores; a sua forma funcional *linear* tem sido adotada com maior frequência na literatura, em detrimento da forma *semi-logatitímica* e *logaritímica*. Porém várias pesquisas têm se preocupado em ampliar os limites da abordagem *linear* estabelecendo comparativos entre as diversas técnicas de modelagem.

Os preços implícitos, econometricamente, são estimados a partir de análises de regressão múltiplas em função de suas características, dado pela equação simplificada:

$$Y = \beta_0 + \sum_{j=1}^n \beta_j X_j + \varepsilon \quad (4)$$

Em que  $Y$  é a variável dependente e  $X_j$ ,  $j = 0, 1, \dots, n$ , são as variáveis independentes,  $\beta$ s são os coeficientes que mede a variação esperada na variável dependente dado variação de uma unidade em  $X_j$  e  $\varepsilon$  é um termo de erro não explicado pelos regressores da equação. De acordo com o método dos Mínimos Quadrados Ordinários (MQO), para encontrar os estimadores  $\beta$ s faz-se necessário minimizar a soma dos quadrados dos erros, conforme segue:

$$\sum_{i=1}^n \varepsilon_i^2 = [\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3 \cdots \varepsilon_n] \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} = \varepsilon' \varepsilon \quad (5)$$

Observando que:

$$Y = X\beta + \varepsilon \Rightarrow Y - X\beta = \varepsilon' = (Y - X\beta') \quad (6)$$

Sendo assim, a soma dos quadrados dos erros é determinada por:

$$L = \varepsilon' \varepsilon = (Y - X\beta')(Y - X\beta) \quad (7)$$

A solução para o problema de minimização da soma dos quadrados dos erros é dada por:  $\frac{\partial L}{\partial \beta} = 0$ , ou de outra forma:

$$\beta = (X'X)^{-1}X'Y \quad (8)$$

A análise de regressão múltipla tem sido a metodologia utilizada para estimar os preços implícitos dos produtos que por sua vez, são regredidos em função das suas características. Esses modelos sugerem métodos que podem identificar índices de interesse para cada uma das características abordadas. A estrutura de demanda será explicitada a partir da diferenciação observada por cada agente econômico, a partir da comparação do “pacote” de atributos dados seus respectivos preços no mercado.

### 3.1 Seleção das Variáveis

Para determinar as variáveis de análise do produto imobiliário, a literatura supõe que a pesquisa pode ser feita em duas etapas: define-se a demanda por habitação, inicialmente, com base nos resultados obtidos através da estimação de preços hedônicos, e nos atributos sociodemográficos da amostra. A partir da equação de preços hedônicos identifica-se a relação entre o preço implícito de cada característica individual do imóvel, que maximiza a utilidade para cada consumidor, assim como a propensão a pagar por determinada cesta de serviços com que o consumidor se depara no mercado.

Posteriormente, no segundo estágio é possível definir as elasticidades-preço da habitação, permitindo esclarecer o comportamento do consumo individual de habitação quando da alteração de preço da unidade residencial. Após análise das interações e das interdependências entre as variáveis, selecionou-se aquelas que obtiveram maior significância global para o modelo.

Considera-se nesta pesquisa, como preditores microeconômicos, os dados referentes aos atributos estruturais dos imóveis que serão chamados de características intrínsecos (CI), a saber: número de quartos, banheiros, suítes, área total etc. Além das variáveis

físicas do apartamento, será também analisado o impacto de itens de lazer, como a existência ou não de piscina, elevador e/ou mobília. Será utilizado a variável de controle extrínseca de acessibilidade que relaciona a distância dos imóveis a grandes centros comerciais. Em vista disso, o conjunto de variáveis analisadas permitirá apreender a dinâmica do mercado imobiliário para venda na capital paulistana relacionando seus aspectos físicos e locacionais dos bens de acordo com a equação especificada abaixo:

$$\ln PR_i = \alpha + \sum_{j=1}^n \beta_j Z_i + \varepsilon \quad (10)$$

Em que  $PR_i$  refere-se ao preço de cada imóvel e o subscrito “i” indica cada unidade imobiliária.  $Z_i$  representa um vetor contemplando características intrínsecas do imóvel: número de quartos, número de suítes (quarto com banheiro), vagas de garagem, tamanho do imóvel, existência de elevador, mobílias e piscina. Além da variável de acessibilidade, que abrange a distância da unidade habitacional ao Museu de Arte de São Paulo Assis Chateaubriand (MASP), não somente por sua importante localização, como também por ser considerado uma das instituições culturais brasileiras mais importante. Ademais podemos incluir a variável dummy que assume valor 1 se o imóvel  $i$  for novo, e zero caso contrário (usado).  $\varepsilon_i$  Indica o termo do erro aleatório.

A base de dados utilizada apresenta estrutura básica de dados em corte transversal, formada quase que exclusivamente por imóveis usados, a venda na capital paulista, distribuídos em 96 bairros *clusterizados* nas quatro principais zonas da cidade: norte, sul, centro e leste. A base é bastante heterogênea, apresenta imóveis com preços e tamanhos distintos, sendo 72,5% destes localizados nas zonas centro e leste, e alta variabilidade nas áreas de lazer, mesmo nas unidades similares.

#### 4 RESULTADOS EMPIRICOS

O modelo de preços hedônicos apresentado nesta seção, ora sugerido e aplicado na literatura por diversos autores já apresentados nas discussões anteriores, foi especificado por meio da forma funcional *semi-logaritimica*, ajustando-se especificamente para o caso analisado.

Foi utilizado uma amostra de um conjunto de dados coletada pelo site kaggle, composta por 6412 apartamentos para venda na cidade de São Paulo, distribuídos em 96 bairros da cidade, anunciados em vários sites classificados de imóveis, no mês de abril de 2019. Entre os dados coletados, além do preço do imóvel, tamanho e valor de condomínio, foi coletado também atributos físicos dos apartamentos – área privativa, número de quartos, banheiros, vagas de estacionamento, quantidade de suítes; e ainda, variáveis *dummys* para os itens existência de elevador, piscina, mobília e se imóvel é novo ou usado.

A Tabela 1 apresenta a descrição e fonte dos dados utilizados no modelo, assim como as principais estatísticas descritivas destas variáveis, tais como: média, mediana, máximo e mínimo (medidas de posição) e desvio-padrão (medidas de dispersão). Observa-se que pelos valores mínimos e máximos da variável preço, condomínio e área, que a base de dados é heterogênea: o preço dos imóveis variam de mínimo de R\$ 442mil a R\$ 3,8milhões; o valor da taxa de condomínio tem média de R\$ 537,34; e a área com variabilidade entre mínimo de 30m<sup>2</sup> e máximo de 620m<sup>2</sup> de área construída.

Tabela 1 - Estatísticas descritivas por característica das variáveis de controle

	<b>Média</b>	<b>Mediana</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Desvio Padrão</b>	<b>Descrição</b>
PREÇO	608624.1	380000.0	42000.0	10000000	740451.5	Preço do imóvel (R\$) a venda em SP
CONDOMÍNIO	537.349	400	0.0000	8920.0	624.632	Valor Condomínio (R\$)
ÁREA	78.611	62	30	620	50.872	Área construída (m <sup>2</sup> )
QUARTOS	2.322	2	1	6	0.714	Nº Dormitórios do imóvel
BANHEIROS	2.037	2	1	7	0.917	Nº de banheiros do imóvel
SUÍTES	0.931	1	0	6	0.772	Nº de suítes do imóvel
VAGA GARAGEM	1.326	1	0	7	0.751	Nº de vagas de garagem do imóvel
ELEVADOR	0.415	0	0	1	0.492	Dummy para existência de elevador
MOBÍLIA	0.117	0	0	1	0.321	Dummy para existência de mobília no imóvel
PISCINA	0.539	1	0	1	0.498	Dummy para existência de piscina
NOVO	0.032	0	0	1	0.176	Dummy para imóvel novo

DISTMASP	0.019	0	0	1	0.137	Dummy para imóvel localizado cerca de 1km do MASP
----------	-------	---	---	---	-------	---

Fonte: dados da pesquisa

Vale apontar que a variável sobre a localização do imóvel ao MASP, foi calculada a partir da distância euclidiana entre cada apartamento e o ponto de referência. E definido como variável *dummy* 1 para as unidades localizadas até 1km de distância do museu.

De acordo com os resultados do modelo estimado na Tabela 2, observa-se que os sinais apresentados das variáveis explicativas legitimam o comportamento do mercado de que os preços dos imóveis estão diretamente relacionados as suas características estruturais e locais. Percebe-se que os compradores de apartamentos atribuem maior importância aos aspectos estruturais do empreendimento em si, como existência de elevador e piscina, do que propriamente as suas características físicas; as variáveis condomínio e área do imóvel tem efeito marginal positivo, mas pouco expressivo, para a amostra coletada. Já a variável binária da distância do imóvel ao MASP captura o efeito positivo sobre o preço dos imóveis mais próximos desta região.

A maioria dos coeficientes apresentaram-se estatisticamente significantes ao nível de 1%, apresentando um alto poder explicativo, com coeficiente de determinação de 0,73, e rejeitando completamente, através do teste F, a hipótese de que o conjunto de variáveis adotadas no modelo não é importante para explicar a variabilidade observado no preço. No entanto, variáveis como *quartos*, *elevador* e *imóvel novo*, apresentam insignificância, podendo indicar um problema de multicolinearidade entre as variáveis, já que não faz sentido na prática tais observações.

As variáveis *condomínio* e *área* do imóvel, apesar de serem estatisticamente significantes, tem baixíssima influência sobre a decisão de compra do consumidor. Enquanto alterações na ordem de uma unidade de *vaga de garagem* devem elevar os preços de venda em torno de 15,7%. Demonstrando que na base de dados analisada temos um mercado com produto diferenciado. A análise da elasticidade-preço confirma tal afirmação, já que o fato de existir ou não *piscina* ou *mobília*, podem elevar o preço do imóvel, em média 21,5% e 6,7%, respectivamente.

Segundo Haddad e Hermann (2003), modelos que consideram fatores de acessibilidade revelam a influência da oferta de infraestrutura urbana sobre a moradia, tendendo a valorizar esses locais. Para o modelo em questão, esse critério pode ser observado a partir da análise da variável DISTMASP, que conforme estimativa, um aumento em distância de 1 metro mais próximo do centro leva a um aumento do preço do imóvel em SP em torno de 41%,

comprovando sua influência no contexto urbano da cidade, não apenas como um centro de negócios, como também principal centro de lazer e cultura para os paulistanos.

Tabela 2 - Ajustamento do modelo de demanda habitacional para a cidade de São Paulo

Variável	Coefficiente	Desvio Padrão	Probabilidade
C	11.80505	0.019334	0.0000
CONDOMÍNIO	0.000176	1.04E-05	0.0000
ÁREA	0.007588	0.000178	0.0000
QUARTOS	0.006920	0.009190	0.4515
BANHEIROS	0.077641	0.012589	0.0000
SUÍTES	-0.037216	0.014246	0.0090
VAGA DE GARAGEM	0.157201	0.010465	0.0000
ELEVADOR	0.018666	0.010498	0.0754
MOBÍLIA	0.067359	0.014976	0.0000
PISCINA	0.215327	0.010466	0.0000
NOVO	-0.042944	0.027812	0.1226
DUMMYMASP	0.414176	0.035253	0.0000
$R^2$	0.735015		
$R^2$ Ajustado	0.734559		
F	1613.844		
Prob(F)	0.000000		
AIC	0.901840		

Fonte: dados da pesquisa

Para esclarecer o fato da multicolinearidade apresentamos a matriz de correlação e como se vê, várias das correlações são bastantes altas. Mas ainda assim, essa ainda pode ser uma condição suficiente, mas não necessária para tal afirmação. Omitir as variáveis com elevada colinearidade pode cair em um viés de especificação, portanto, para tentar corrigir o efeito em questão iremos optar por transformar os dados, inserindo termos quadráticos para analisar novamente as estatísticas e compará-las com a primeira regressão.

Tabela 3 - Matriz de correlação das variáveis explicativas

	Preço	Condomínio	Área	Quartos	Banheiros	Suítes
Preço	1.000000					
Condomínio	0.614456	1.000000				
Área	0.825660	0.662869	1.000000			
Quartos	0.490255	0.408126	0.658609	1.000000		

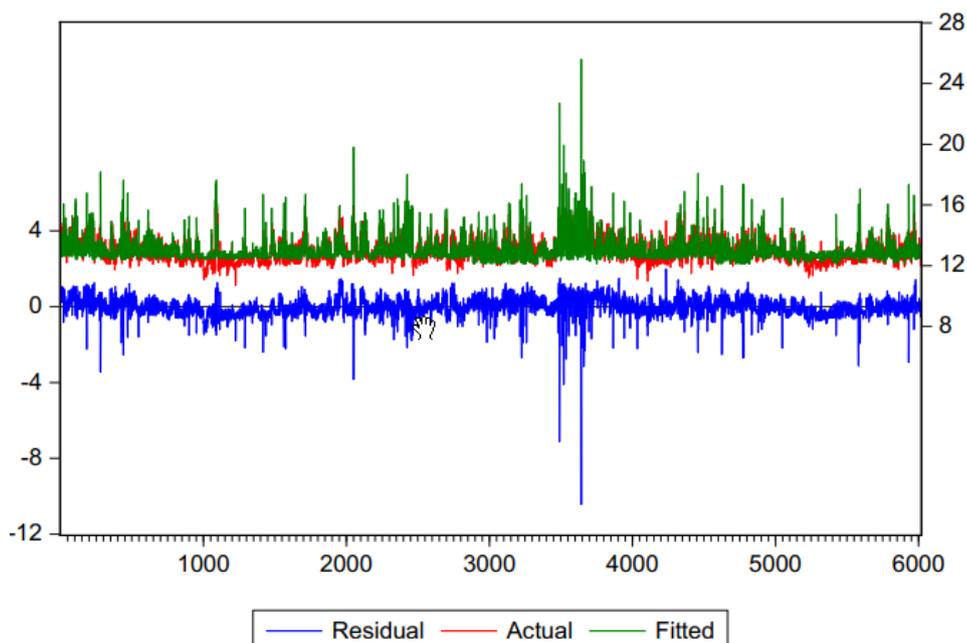
Banheiros	0.636782	0.472210	0.673858	0.552663	1.000000	
Suítes	0.623177	0.434446	0.632128	0.503019	0.896459	1.000000

Fonte: dados da pesquisa

Ao executar a transformação quadrática, observa-se aumento do poder explicativo da variável *quartos* e dos itens de lazer (*vaga de garagem, mobília, piscina*) na formação dos preços dos imóveis residenciais. Todos os coeficientes apresentaram-se estatisticamente significantes ao nível de 1%, apesar da relativa redução do grau de explicação do modelo para 0,61. As variáveis *suítes* e *banheiros*, indicaram relação inversa a variável dependente; podendo ser consideradas redundantes dentro do modelo, portanto, retirando-as da regressão sem prejuízo dado  $R^2$  permanecer inalterado após a mudança. Assim optou-se por retirar a variável *novo* que também se apresenta redundante e desnecessária, já que a base se refere apenas a apartamentos novos.

Para testar a hipótese de heterocedasticidade, aplicou-se o teste residual de White, confirmando a forte dispersão dos dados no modelo original (Gráfico1); resultado esperado para amostra de dados do tipo corte transversal. Segundo Gujarati (2006), em grandes amostras, a medida corretiva através do método de Newey-West, que nada mais é do que uma extensão do método de erros-padrão consistentes de White para heterocedasticidade, tem sido capaz de corrigir os efeitos, gerando estimadores eficientes.

Gráfico 1 – Relação entre os termos de erro da equação



Na Tabela 4 são apresentados os resultados comparativos entre os dois modelos, onde o (\*) denota erros-padrão de Newey-West. Comparando as regressões, verificamos que os valores para os coeficientes estimados e o valor de  $R^2$  permaneceram inalterados. Enquanto os erros-padrão do segundo modelo são maiores que os obtidos por meio dos MQO, e assim, as razões do teste  $t$  no modelo Newey-West são muito menores que as razões do primeiro modelo. Mostrando que o modelo MQO de fato subestimou os verdadeiros erros-padrão.

Dentre as variáveis quantitativas, observa-se que o preditor *vaga de garagem* manteve relação positiva com a variável dependente, porém mais que dobrou a relação, de 15,7% para 34,9%. O preditor número de dormitórios, que na Tabela 2 era estatisticamente não significativo ao nível de 5%, passou a ser representativo para o modelo da Tabela 4, impactando em 16,9% a variação do preço, dado variação de uma unidade de quarto. Enquanto a variável binária *piscina* se manteve, praticamente, constante na relação com a elasticidade-preço da habitação.

Tabela 4 - Comparativo entre o modelo de estimadores MQO e Newey

	Variável	Coefficiente	Desvio Padrão	Probabilidade
MQO	C	11.87463	0.021630	0.0000
	ÁREA^2	1.51E-05	4.79-E07	0.0000
	QUARTOS	0.169632	0.010292	0.0000
	VAGA DE GARAGEM	0.349171	0.011323	0.0000
	PISCINA	0.207274	0.012322	0.0000
	DISTMASP	0.663030	0.041302	0.0000
	NEWY-WEST	C	11.87463	0.045452*
ÁREA^2		1.51E-05	2.13E-06*	0.0000
QUARTOS		0.169632	0.017513*	0.0000
VAGA DE GARAGEM		0.349171	0.022238*	0.0000
PISCINA		0.207274	0.017690*	0.0000
DISTMASP		0.663030	0.052039*	0.0000
$R^2$		0.622394		

Fonte: dados da pesquisa

Por fim, a variável binária distância do imóvel ao MASP, que foi incluída com intuito de capturar diferenças entre os efeitos dos apartamentos mais próximo do centro de negócios de São Paulo, mantido constante as demais variáveis, apresentou elevada significância. Tais resultados podem ser observados também na relação conjunta com as

variáveis *quartos* e *vaga de garagem*, que apontam para segmento de produtos altamente diferenciados.

A significância apontada pode se dar pelo fato de a amostra concentrar-se principalmente na região centro da capital paulista; corroborando para a hipótese de um mercado produto diferenciado. Como também, pode refletir um problema de especificação do modelo, que segundo Tabales (2007) modelos lineares para estimar preços de habitação, pode não corresponder com a realidade já que se deve levar em conta que os atributos do citado imóvel estabelecem relação conjunta, de forma simultânea e não de forma aditiva. Formando o que os autores chamam de correlação espacial.

## 5 CONCLUSÃO E CONSIDERAÇÕES FINAIS

Considerando a natureza dos dados em corte transversal, formada por grande amostra de dados com 6412 observações, bem como alto grau de diferenciação e a relação de dependência espacial observados, o objetivo geral da pesquisa em estimar um modelo confiável, que atendesse aos aspectos de não tendenciosidade, eficiência e consistência foram atendidos.

Os resultados revelam que os atributos implícitos e as variáveis binárias referente a itens de lazer e distância ao centro de negócios da cidade, se apresentaram correlacionadas positivamente com a variável dependente, sendo responsáveis por provocarem aumento substancial no preço final de venda. Sugerindo que o consumidor para este mercado de produto diferenciado atribui mais importância a características de lazer e comodidades, do que propriamente as características físicas dos imóveis.

A partir do modelo de preços hedônicos estimado, foi possível verificar a importância dos preços marginais na formação dos preços para implantação de novos empreendimentos. Por outro lado, percebeu-se que a variável condomínio revelou-se estatisticamente não significantes; e que a retirada de variáveis redundantes no modelo não influenciou no seu poder de explicação.

Finalmente, os resultados encontrados devem permitir um entendimento da dinâmica do mercado imobiliário residencial para venda na grande São Paulo no período analisado. Muito embora a maioria das variáveis consideradas no modelo ter relação com a variabilidade dos preços de venda, notou-se limitações no que concerne as medidas corretivas para as premissas dos estimadores MQO para dados de corte transversal; direcionando para uma análise de correlação na econometria espacial, ou para análise de modelos de redes neurais, que já consta em boa parte dos estudos recentes fora do Brasil.

Sugere-se, para futuras pesquisas, formular um comparativo de técnicas para dados em corte transversal, com inclusão de uma matriz de pesos que capture a relação dos diferentes tipos de unidades de um mesmo empreendimento, para inferir, inclusive, relação espacial com imóveis vizinhos. Pode-se incluir atributos econômico ou financeiros que seccionem a demanda de consumo para habitação e a demanda de consumo para investimento, que compõem o enorme e complexo mercado habitacional desta cidade. Por fim, deve-se explorar o comparativo de estimativas de MQO com as técnicas da Inteligência Artificial para a casos brasileiros, a exemplo do que já ocorre com as pesquisas internacionais.

## REFERÊNCIA BIBLIOGRAFICA

- ALVES, D., YOSHINO, J., PEREDA, P., AMREIN, C. Modelagem dos preços de imóveis residenciais paulistanos. **Revista Brasileira de Finanças, Rio de Janeiro**, RJ, v.9, n.2, p. 167-187, jun. 2011.
- ARRAES, R. A.; SOUSA FILHO, E. Externalidades e formação de preços no mercado imobiliário urbano brasileiro: um estudo de caso. **Revista de Economia Aplicada**, Ribeirão Preto, SP, v. 12, n. 2, p. 289-319, abr./jun. 2008.
- DANTAS, R.; MAGALHÃES, A.; VERGOLINO, J.R. Um modelo espacial de demanda habitacional para a cidade de Recife. *Estudos Econômicos*, v.40, p.891-916, out-dez.2010.
- FAVERO, L.P. Modelos de preços hedônicos aplicados a imóveis residenciais em lançamento no município de São Paulo. Dissertação (Mestrado em Administração) – FEA/USP, São Paulo, 2003.
- FAVERO, L.P.; LUPPE, M.R. Modelos de preços hedônicos aplicados ao mercado imobiliário: uma revisão da literatura. **Revista de Economia e Administração**, v.5, n.4, 424-448p, out./dez. 2006.
- MELO, F. A valorização do capital e a produção do espaço urbano: a produção imobiliária habitacional do segmento econômico em Manaus-AM. Dissertação (Mestrado em Geografia) – UFAM, Manaus, 2020.
- FUNDAÇÃO JOÃO PINHEIRO. **Déficit habitacional no Brasil 2016-2019**. Belo Horizonte, 2020.
- GUJARATI, D. N. **Econometria básica**. 4. ed. São Paulo: Makron Books, 2006.
- HADDAD, E. A.; HERMANN, B. M. **Muito além do jardim**: mercado imobiliário e amenidades urbanas. São Paulo: USP, Dep. Economia, 2003. (Texto para Discussão, 04).
- HERMANN, B. M. **Estimando o preço implícito de amenidades urbanas: evidências para o município de São Paulo**. Dissertação (Mestrado em Economia) – USP, São Paulo, 2003. 74 p.
- HOESLIN, M.; CANTONI, E.; BOURASSA, S.C. Predicting house prices with spatial dependence: A comparison of alternative methods. **The Journal of Real Estate Research**. April 2014.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE), **Síntese de Indicadores Sociais: Uma análise das condições de vida da população brasileira 2018**. Rio De Janeiro, 2018.
- LANCASTER, K. A new approach to consumer's theory. **Journal of Political Economy**, Chicago, v. 74, n. 2, p. 132-57, 1966.

PINTO, V.H.; FERNANDES, R.A. Análise de preços hedônicos no mercado imobiliário residencial de Conselheiro Lafaiete, MG. *Interações*, Campo Grande, MS, v.20, n.2, p. 627-643, abr./jun. 2019.

ROSEN, S. Hedonic prices and implicit markets: product differentiations in pure competition. *journal of political Economy*, v. 82, p. 34-55, 1974.

SELIM, H. 2009. Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. ScienceDirect. Disponível em: [www.sciencedirect.com](http://www.sciencedirect.com). Acesso em: 15/06/2021.

TABALES, J. Mercados Inmobiliarios: Modelización de los precios. Tese (Doutorado em Organização de Empresas) – Universidad Cordoba, Cordoba, 2007.