

**FACULDADE DE ECONOMIA**  
Universidade do Algarve

Campus de Gambelas, 8000-117 Faro, Portugal  
Tel.: +351 289 800 915; Fax: +351 289 815 937



---

**UNIVERSIDADE DO ALGARVE**

**FACULDADE DE ECONOMIA**

**A Formação dos Preços dos Imóveis para Habitação**

VANDA MARIA FILIPE GUERREIRO

**Mestrado em Finanças Empresariais**

**2014**



**FACULDADE DE ECONOMIA**  
Universidade do Algarve

Campus de Gambelas, 8000-117 Faro, Portugal  
Tel.: +351 289 800 915; Fax: +351 289 815 937



---

**UNIVERSIDADE DO ALGARVE**  
**FACULDADE DE ECONOMIA**

**A Formação dos Preços dos Imóveis para Habitação**

VANDA MARIA FILIPE GUERREIRO

**Mestrado em Finanças Empresariais**  
**Dissertação orientada por:**  
**Professor Doutor Paulo Manuel Marques Rodrigues**  
**Licenciado António Daniel Correia dos Santos**

**2014**

# **Declaração de autoria de trabalho**

## **A Formação dos Preços dos Imóveis para Habitação**

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Assinatura do candidato

---

Copyright © 2014 Vanda Maria Filipe Guerreiro

A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

## ÍNDICE GERAL

	Página
1. Principais objetivos do projeto .....	6
2. Revisão de Literatura.....	9
a) Variáveis que contribuem para a formação dos preços no mercado imobiliário .....	10
b) Aplicações do modelo de preços hedónicos.....	18
3. Metodologias para analisar a formação dos preços no mercado imobiliário ...	20
4. Análise Empírica .....	23
a) Tratamento e análise dos dados.....	23
b) Ensaios Iniciais.....	26
c) Colinearidade .....	27
d) Componentes principais .....	31
e) Modelo Hedónico .....	36
f) Variáveis de interação .....	40
g) Preço médio e modelo por tipologia .....	41
h) Testes à especificação do modelo .....	45
i) Heterocedasticidade .....	46
j) Testes aos parâmetros do Modelo 10 .....	50
k) Significância económica dos estimadores.....	51
5. Conclusão .....	52
6. Desenvolvimentos futuros .....	55
7. Bibliografia de base:.....	57
8. Anexos.....	61
I. Anexo A: Teste F ao Modelo 1 .....	61
II. Anexo B: Teste F ao Modelo 3 .....	62
III. Anexo C: Modelo por Tipologia .....	63

## ÍNDICE DE FIGURAS

<b>Figura 4.1</b> .....	<b>38</b>
<b>Figura 4.2</b> .....	<b>39</b>
<b>Figura 4.3</b> .....	<b>44</b>
<b>Figura 4.4</b> .....	<b>46</b>
<b>Figura 4.5</b> .....	<b>51</b>

## ÍNDICE DE TABELAS

<b>Tabela 3.1</b> .....	<b>20</b>
<b>Tabela 4.1</b> .....	<b>27</b>
<b>Tabela 4.2</b> .....	<b>28</b>
<b>Tabela 4.3</b> .....	<b>29</b>
<b>Tabela 4.4</b> .....	<b>29</b>
<b>Tabela 4.5</b> .....	<b>29</b>
<b>Tabela 4.6</b> .....	<b>30</b>
<b>Tabela 4.7</b> .....	<b>31</b>
<b>Tabela 4.8</b> .....	<b>31</b>
<b>Tabela 4.9</b> .....	<b>32</b>
<b>Tabela 4.10</b> .....	<b>33</b>
<b>Tabela 4.11</b> .....	<b>33</b>
<b>Tabela 4.12</b> .....	<b>34</b>
<b>Tabela 4.13</b> .....	<b>34</b>
<b>Tabela 4.14</b> .....	<b>34</b>
<b>Tabela 4.15</b> .....	<b>35</b>
<b>Tabela 4.16</b> .....	<b>35</b>
<b>Tabela 4.17</b> .....	<b>36</b>
<b>Tabela 4.18</b> .....	<b>36</b>
<b>Tabela 4.19</b> .....	<b>38</b>
<b>Tabela 4.20</b> .....	<b>39</b>
<b>Tabela 4.21</b> .....	<b>40</b>
<b>Tabela 4.22</b> .....	<b>41</b>
<b>Tabela 4.23</b> .....	<b>42</b>
<b>Tabela 4.24</b> .....	<b>43</b>
<b>Tabela 4.25</b> .....	<b>43</b>
<b>Tabela 4.26</b> .....	<b>44</b>
<b>Tabela 4.27</b> .....	<b>45</b>
<b>Tabela 4.28</b> .....	<b>48</b>
<b>Tabela 4.29</b> .....	<b>49</b>

<b>Tabela 4.30</b> .....	<b>50</b>
<b>Tabela 5.1</b> .....	<b>54</b>
<b>Tabela 8.1</b> .....	<b>62</b>
<b>Tabela 8.2</b> .....	<b>63</b>
<b>Tabela 8.3</b> .....	<b>63</b>
<b>Tabela 8.4</b> .....	<b>64</b>
<b>Tabela 8.5</b> .....	<b>64</b>
<b>Tabela 8.6</b> .....	<b>65</b>
<b>Tabela 8.7</b> .....	<b>65</b>
<b>Tabela 8.8</b> .....	<b>66</b>

### LISTA DE ABREVIATURAS

IPC	Índice de Preços no Consumidor
OOHPI	Owner Occupied House Price Index
HICP	Harmonised Index of Consumer Prices
SPAR	Sales Price Appraisal Ratio
OLS	Ordinary Least Squares
FWLS	Feasible Weighted Least Squares
AIC	Critério de Informação de Akaike
BIC	Critério de Informação Bayesiano ou Critério de Informação de Schwarz
HQC	Critério de Informação de Hannan- Quinn
IMI	Imposto Municipal sobre Imóveis
IRS	Imposto sobre o Rendimento das Pessoas Singulares
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics

## **AGRADECIMENTOS**

Este estudo, mais do que o resultado do meu trabalho, esforço e dedicação, é fruto de um conjunto de esforços que o tornaram possível e sem os quais teria sido muito mais difícil realizar este percurso. Este é o espaço onde manifesto a minha gratidão aos meus orientadores Professor Paulo Rodrigues e Dr. Daniel Santos pela forma como me orientaram, pela sua disponibilidade, e pelo entusiasmo e motivação. Devo de igual modo agradecer ainda, a generosidade do colega Dr. António Machado Lopes, pela partilha dos seus conhecimentos técnicos que contribuíram decisivamente para a qualidade deste trabalho.

Agradeço igualmente, ao INE pela disponibilidade dos dados.

Em especial, quero agradecer à minha irmã e às duas sobrinhas ela que me deu, pelo seu amor, que me dá a força para nunca desistir; à minha prima e às duas priminhas que ela me deu, por aquele carinho tão especial; e principalmente à minha mãe, pela motivação, afeto, carinho, compreensão e pelo tempo da sua vida que me concedeu, mas sobretudo por tudo o que tenho, tudo o que fiz e tudo o que sou.

## RESUMO

O presente estudo destina-se a determinar um modelo para formação do preço no imobiliário.

O conhecimento do preço marginal que o consumidor está disposto a pagar por um determinado serviço de habitação, também chamado de preço implícito ou hedônico, pode ser de grande interesse quer para o empreendedor quer para o decisor político, na medida em que pode ajudar na análise de custo-benefício, como também para elaboração e implantação de políticas habitacionais e urbanas.

*“... no longo prazo, são as preferências do consumidor que determinam a configuração das cidades.”* [Dantas et al. \(2007\)](#).

Após uma breve introdução ao tema, ao longo do capítulo 1, apresenta-se no capítulo 2 um levantamento e análise dos estudos, e principalmente dos artigos científicos, mais relevantes para esta dissertação. É analisada por um lado literatura respeitante às variáveis que influenciam o valor dos imóveis e por outro, estudos realizados com recurso a modelos hedônicos. Os modelos hedônicos são, entre outros, brevemente descritos no capítulo 3. No capítulo 4 é descrito todo o processo de estimação do um modelo linear pelo método dos mínimos quadrados ordinários, realizado a partir de um conjunto de dados com referência nacional sobre o valor dos apartamentos em 2009. São abordadas, entre outras, as questões relativas à colinearidade, testes de especificação e de significância e heteroscedasticidade. São experimentadas diversas metodologias, entre as quais componentes principais, modelos hedônicos, combinação não linear de variáveis, e partição da base de dados. Do estudo decorrem algumas conclusões que são explanadas no capítulo 5. Por fim são expostas breves sugestões para futuros desenvolvimentos da matéria em apreço que por se encontrarem além do âmbito do presente estudo não foram aprofundadas.

**Termos chave:** Preços; imóveis; modelos hedônicos; Mínimos Quadrados Ordinários.

## ABSTRACT

### The price determinants of residential property

This study is intended to determine a model for price formation in the real estate.

The knowledge of the marginal price which consumers are willing to pay for a particular housing service, also called hedonic or implicit price, can be of great interest for both, entrepreneurs and policymakers, to the extent that it can help in the cost- benefit analysis, but also in the design and implementation of housing and urban policies .

*"... In the long run, are the consumer preferences that determine the configuration of cities."* Dantas , Magalhães et al . (2007).

After a brief introduction on the topic, in Chapter 1, it is presented in Chapter 2 a survey and analysis of the studies, and especially of the most relevant scientific papers. It analyses on one hand the literature concerning variables that influence the value of dwellings and on the other hand empirical studies using the hedonic models. The hedonic models are, among others, briefly presented along Chapter 3. In Chapter 4 is described the process of estimating a linear model by the Ordinary Least method, is conducted from a national data set with reference to the value of the apartments in 2009. Among others, issues related to collinearity, model specification tests, tests of significance, and heteroscedasticity are addressed. Various methodologies including the principal components, hedonic models, non-linear combination of variables, and the database and partition are experienced. This study derives some conclusions that are explained in Chapter 5. Finally, are exposed some brief suggestions for future developments of the matter under consideration that because they are beyond the scope of this study were not investigated.

**Keywords:** Prices; homes; hedonic models, Ordinary Least Squares.

## 1. Principais objetivos do projeto

Em Setembro de 2000 o Comité de Programa Estatístico da EU decidiu criar um índice de preços à habitação própria OOHPI (*Owner Occupied Housing Price Index*) tendo por base a abordagem, *net acquisitions*. O projeto iniciou em 2001 com 5 países. Atualmente, já na sexta fase são 30 os países europeus envolvidos. O artigo de [Makaroidis et al. \(2006\)](#) e o mais recente de [Marola et al. \(2012\)](#) descrevem as ações e os objetivos alcançados pelo EUROSTAT no cálculo de um OOHPI compatível com o HICP (*Harmonised Index of Consumer Prices*) bem como o tratamento que o OOHPI tem conhecido no âmbito do cálculo do CPI (*Consumer Price Index*). Os objetivos do HICP são medir a convergência dos preços entre os estados membros e a variação dos preços. Considerando que as despesas com a habitação são uma parcela importante do consumo das famílias, e que em alguns países o arrendamento (incluído no HICP) tem maior peso do que a aquisição de habitação própria (não incluído no HICP), e noutros verifica-se o inverso, então os objetivos do HICP estão comprometidos a menos que o OOHPI seja incluído no HICP. Das ações do EUROSTAT resulta um Manual Técnico do OOHPI ([EUROSTAT \(2010\)](#)). Paralelamente existe um outro projeto liderado pelo EUROSTAT destinado à elaboração de um Manual para a produção de índices de preços residenciais com vista à normalização internacional de metodologias [Balk et al. \(2011\)](#)

De acordo com [Balk et al. \(2011\)](#) a utilidade de medir os preços residenciais assenta nos seguintes aspetos:

- *“ Um indicador macro económico para a inflação.*
- *Um instrumento para a política monetária e definição de objetivos.*
- *Uma medida da riqueza.*
- *Um indicador de estabilidade ou solidez financeira para medir a exposição ao risco.*
- *Um deflator das contas nacionais.*

- *Um elemento na escolha individual de um cidadão de adquirir um imóvel residencial.*
- *Um input para outros índices de preços, particularmente o Índice de Preços ao Consumidor, para uso, entre outros, como argumento nos acordos salariais e indexações. “*

Este estudo visa colmatar a lacuna da falta de informação acerca das preferências dos consumidores num sector que em 2008 já em crise contribuiu 7,29%<sup>1</sup> para o VAB nacional. Segundo dados do Banco de Portugal a construção é o terceiro principal sector ao nível do número de pessoas ao serviço (13%) e volume de negócios (10%), e o segundo em número de empresas (14%).

O mercado imobiliário além do rendimento das atividades diretas, das empresas da chamada “fileira da construção” e do imobiliário, também tem efeitos noutras atividades a jusante, nomeadamente lojas de mobiliário e decoração, advogados, notários, contabilistas, gestão de condomínio, limpeza, manutenção de jardins e piscinas, serviços de internet e televisão, venda de eletrodomésticos, entre muitos outros.

Apesar da sua forte expressão, o mercado imobiliário em Portugal não é fonte de inspiração académica. Não existem artigos científicos publicados dedicados ao mercado imobiliário português. As teses de mestrado ou doutoramento (não publicadas) existentes acerca do tema são essencialmente destinadas a encontrar modelos de valorização do imóvel para fins de financiamento bancário ou fiscais, que se baseiam muitas vezes em variáveis objetivas, essencialmente de estrutura. [Reichert \(1990\)](#) confirmou que as variáveis de estrutura têm capacidade de explicar mais de 90% da formação do preço. Mas levanta uma questão: O que explica que casas com características semelhantes apresentem preços díspares? Claramente as características de localização tem relevo, mas quais? E em que medida? Outros estudos académicos existentes

---

<sup>1</sup> Fonte: INE

dedicam-se mais a explorar metodologias de cálculo ou segmentos muito específicos do mercado imobiliário.

A desvalorização dos imóveis teve em grande medida na base da crise que hoje enfrentamos. Numa economia sustentada por um consumo alavancado pela concessão de crédito, que como descrito por [Santos et al. \(2007\)](#) assentou na quebra das taxas de juro, e que tinha como colateral imóveis em permanente capitalização, é fundamental conhecer os *drivers* dessa capitalização. Conhecer as determinantes do preço poderá permitir conhecer quais as variáveis que podem assegurar que o imóvel mantém o seu valor numa situação de crise. Estes elementos poderão ainda ser úteis para a gestão e antecipação de potenciais crises no sistema financeiro.

O imóvel para habitação é o maior investimento que uma pessoa faz na vida e tem um elevado peso no orçamento familiar, é por isso um fator a ter em consideração em muitas políticas de coesão social e regional [Bourassa et al. \(2006\)](#). Por exemplo, através da medição da taxa de esforço (relação entre o preço do imóvel e o salário médio) na compra da habitação e sua comparação entre regiões ou países. Se dois imóveis, com as mesmas características de estrutura e de envolvente exigem diferentes taxas de esforço, esta divergência deve ser corrigida através de medidas de política económica.

Finalmente os outros beneficiários dos resultados deste estudo serão os *developers* e os agentes imobiliários que terão uma ferramenta para tornar o seu produto mais competitivo. Considerando que a atual tendência de decréscimo da procura se irá alterar para os próximos anos, saberão como adequar as características do produto oferecido à procura existente, estando melhor preparados para enfrentar os novos desafios.

## 2. Revisão de Literatura

Um imóvel pode ser visto como um bem composto que proporciona diferentes níveis de serviços consoante as suas características.

A análise de [Jansen et al. \(2011\)](#) identifica diferentes funções de uma habitação. Além de ser um abrigo, uma habitação tem diversas funcionalidades socioeconómicas, enquadra um conjunto de atividades diárias externas e internas nomeadamente trabalho, lazer, compras, relacionamento com os vizinhos e amigos. É igualmente visto como um bem de consumo durável e muitas vezes de investimento.

Ainda de acordo com os citados autores, uma habitação é um tipo de bem com algumas especificidades que o diferenciam em termos de análise: é muito caro; é espacialmente imóvel; tem uma elevada durabilidade; é muito heterogéneo e tem a possibilidade de alterar as suas características.

Se fosse possível ter o preço de cada uma das características, o preço do imóvel seria o resultado do somatório de todos esses preços. Este é o conceito de preço implícito, definido no trabalho pioneiro de [Rosen \(1974\)](#), onde *“Os preços observáveis e as quantidades específicas das características associadas a cada bem definem um conjunto de preços implícitos ou hedónicos”*.

Os preços hedónicos têm vindo a ser utilizados desde então para o estudo dos preços de diversos bens onde as combinações possíveis de características, e a possibilidade de alteração são de tal forma amplas que os efeitos resultantes das diferenças na qualidade têm de ser expurgados a fim de se analisar a evolução pura do preço.

### **a) Variáveis que contribuem para a formação dos preços no mercado imobiliário**

Existem inúmeros estudos acerca da influência das mais diversas variáveis que podem contribuir para a formação do preço da habitação. [Fischel \(2001\)](#) chega mesmo a afirmar que “*Tudo parece ser capitalizável*”.

A literatura refere as variáveis de estrutura e de localização como as mais determinantes. Um modelo que inclua a área do imóvel como variável explicativa corre um sério risco de qualquer uma das outras variáveis não ser significativa. Foi o que sucedeu a [Reichert \(1990\)](#) o que o levou a incluir a área na variável dependente. Passou a estudar o preço por metro quadrado. Esta reformulação já lhe permitiu determinar a influência das outras variáveis. Utilizando dados para os EUA de 1975 a 1987 analisou a influência da variação da população, do rendimento, das taxas de juro e da taxa de emprego nos preços dos imóveis. Concluiu que destas o rendimento é que teria um maior efeito com uma elasticidade de 3,78, ou seja, o aumento de 1 por cento no rendimento levaria a um acréscimo de 3,78 por cento nos preços da habitação. Destas variáveis, aquela que registaria menor influência positiva é a taxa de emprego com uma elasticidade de 0,91 por cento. Como seria de esperar a taxa de juro exerce uma influência negativa sobre os preços e de acordo com este estudo 1 por cento de aumento da taxa de juro levaria a uma queda nos preços de 0,19 pontos percentuais. Este autor desenvolveu com os mesmos dados uma análise regional onde concluiu que a influência das variáveis varia de acordo com a região. Ou seja não existe um mercado nacional da habitação mas sim regional, por exemplo as alterações na taxa de juro tem maior influência na região da Nova Inglaterra. [Gerkman \(2010\)](#) apresenta um estudo bastante complexo com base em econometria espacial, onde as variáveis são agrupadas em três conjuntos: localização, estrutura e envolvente. A localização é medida através de atributos como, vista mar, distância ao centro de negócios, proximidade a transportes públicos; as variáveis de estrutura são as habituais, nomeadamente, idade, área, número de divisões; os atributos da

envolvente, utilizados são a taxa de crime, as características socioeconómicas dos vizinhos, a poluição e o nível de ruído.

As características da vizinhança em termos de bairro são também um fator com grande influência na capitalização imobiliária. Gerkman parte também do princípio que o preço de uma habitação depende do valor, praticado no mercado, dos imóveis vizinhos com as mesmas características, e da relação entre a oferta e a procura num determinado bairro. Tal como descrevem [Dantas et al. \(2007\)](#) “*cada edifício funciona como um micro polo de influência sobre os seus vizinhos*”. Por exemplo, o preço pelo qual foi vendido um apartamento num determinado prédio vai influenciar o preço de venda de outros apartamentos nesse mesmo prédio ou bairro. Acrescentam-se à regressão hedónica tradicional variáveis de desfasamento espacial a fim de determinar a autocorrelação espacial nos dados [Bruch et al. \(2011\)](#) verificaram a influência da estabilidade da vizinhança, do *status* socioeconómico, nível de rendimento e etnia, como variáveis exploratórias no modelo de determinação do preço da habitação.

Os agentes imobiliários referem-se frequentemente à abordagem 3L para a determinação do preço das casas. De acordo com esta abordagem, também referida por Gerkman e desenvolvida por [Greene \(2012\)](#) existem três variáveis que influenciam o preço das casas: localização, localização e localização. Estes três níveis de localização respeitam à área metropolitana, à cidade e à rua. Estes autores concluíram que estes níveis de localização têm o mesmo grau de influência já que ao retirar o conjunto de variáveis respeitante a cada nível o desvio padrão aumenta 2,7%, 2,3% e 2,2%, respetivamente.

[Oates \(1969\)](#) considera a capitalização dos efeitos dos orçamentos públicos locais nos valores do imobiliário de uma comunidade, tanto no lado da receita como na despesa, i.e., considera os efeitos do imposto local sobre a propriedade e o valor da despesa publica local em escolas por aluno. Conclui que se a receita gerada por um aumento de impostos for aplicada na melhoria das escolas o valor

dos imóveis se manterá. O estudo de [Oates \(1969\)](#) veio a ser complementado e discutido por [Pollakowski \(1973\)](#), já que o próprio tinha deixado um alerta para o nível de confiança dos seus resultados. As dúvidas colocam-se na utilização do investimento público por aluno como uma boa *proxy* de serviço público, bem como a relação desta variável com outras variáveis relativas as políticas públicas locais que não foram consideradas na equação. A correlação das variáveis com o termo de erro (endogeneidade) é outro problema que não foi devidamente ultrapassado.

A qualidade da escola é uma variável frequentemente utilizada como *proxy* do investimento público em condições de bem-estar. [Haurin et al. \(1996\)](#) analisaram o efeito desta variável para imóveis de qualidade-constante. O resultado revelou que a capitalização positiva da proximidade a uma escola de qualidade era igual para todas as habitações na área de influência dessa escola. Vários estudos, nomeadamente de [Downes et al. \(2002\)](#), [Reback \(2005\)](#) e [Nguyen-Hoang et al. \(2011\)](#), vieram confirmar que a qualidade da escola<sup>2</sup> avaliada pelos proprietários não é medida em termos de montante do investimento público mas em termos de resultados dos alunos.

Anteriormente referido, [Fischel \(2001\)](#) acrescenta outra perspetiva, considerando que as políticas públicas locais (política fiscal, política de educação, política ambiental entre outras) afetam mais os preços das casas do que as políticas centrais, e que os eleitores são sensíveis a isso, decorrendo que as políticas locais funcionam melhor. Em Portugal, no que respeita ao lado da receita pública, tanto o imposto sobre a propriedade (IMI) como outras taxas, impostos e preços dos serviços públicos (derrama, IRS, água, recolha de resíduos, etc.) que podem ser fixados pelos municípios tem sido até ao momento, poucos e de pequeno valor. Além disso, os próprios municípios tendem a harmonizar e fixar as mesmas taxas e preços, funcionando como um oligopólio. Assim, não se prevê

---

<sup>2</sup> O ranking nacional de escolas elaborado pelo Ministério da Educação, em Portugal pode servir para a medição dos impactos desta variável na escolha do local para adquirir habitação.

que em Portugal, as taxas e impostos municipais tenham influenciado o preço da habitação. Presentemente, os municípios estão a ser pressionados para encontrar formas de autofinanciamento, pelo que é previsível o aumento destas taxas de forma diferenciada de acordo com as necessidades de cada município. Assim a política fiscal pública local pode vir a ser um fator discriminante na escolha do município para adquirir habitação.

Os autores [Ribeiro \(2013\)](#) analisaram para o Japão a reação dos proprietários perante eventuais soluções públicas para o risco de terramoto e concluíram que o programa de apoio à remoção de barreiras dos espaços e edifícios tem uma influência positiva significativa nas preferências dos proprietários. Já os testes sísmicos ou programas de revitalização levadas a cabo pelas câmaras não tem impacto. No entanto, no caso de ações despoletadas ao nível das áreas metropolitanas o seu impacto é relevante.

De facto, as opções de investimento público são passíveis de influenciar a escolha do local para habitação. É certo que o comprador não analisa o orçamento do município aquando da compra, mas tal como [Bickerdike \(1902\)](#) lembrou "*a iluminação e a limpeza das ruas é visível a olho nu*". Mais recentemente [Stadelmann \(2010\)](#) encontra para o Cantão de Zurique uma forte influência na capitalização do investimento público em cultura, saúde e bem-estar social. É razoável considerar que, conforme uma sociedade se torne mais esclarecida e que os indicadores de nível de escolaridade e de rendimento aumentam, a valorização de bens culturais e ambientais venham a registar um peso crescente nas opções dessa sociedade, nomeadamente na formação dos preços da habitação.

Neste sentido a capitalização dos imóveis pode medir a propensão para pagar por certos bens públicos, materiais ou imateriais, sendo o contrário também uma realidade. Externalidades públicas negativas, nomeadamente de carácter ambiental, desvalorizam os imóveis que se encontram no seu perímetro.

[Reichert \(1997\)](#) verificou que os terrenos urbanos em Uniontown, Ohio sofreram uma desvalorização entre 5% a 15% até uma distância de 2.057,40 metros de um aterro de resíduos tóxicos. Os proprietários foram indenizados pelas autoridades públicas pelas perdas agregadas. [Kiel et al. \(1995\)](#) alertam para o facto de um rumor acerca da potencial contaminação de um aterro próximo, ser suficiente para afetar negativamente os preços e que esses efeitos se prolongam no tempo.

[Farber \(1998\)](#) analisou 25 estudos que se debruçam acerca do estudo de locais de depósito de lixo tóxico, dos quais 19 concluíram que existe uma importante influencia nos preços e uma forte tendência para o seu aumento à medida que o imóvel se afasta da lixeira.

[Tobin et al. \(1988\)](#) comprovaram que a ocorrência de uma cheia registada recentemente reduz o valor dos imóveis. Mais recentemente [Deaton et al. \(2004\)](#) usaram preços hedónicos para concluir que as pessoas associam uma serie de desamenidades a zonas industriais semelhantes a zonas de lixeiras, estimando 2 modelos verificam que se for incluído o fator industrial o efeito da lixeira é menor e sem significância.

Pelo contrário, se a envolvente ambiental for favorável o imóvel poderá registar uma forte capitalização. [Luttik \(2000\)](#) comprovou que dois imóveis de estrutura similar podem ter preços muito diferentes consoante a agradabilidade da envolvente onde se localizam. Uma casa localizada numa zona próxima de um espaço verde, com boa qualidade da água e do ar regista um maior valor acrescentado do que uma casa com localização menos favorável. O prémio pode atingir os 29% sobre o preço da localização menos favorável. Também [Hanley et al. \(1993\)](#) avaliam o valor da qualidade ambiental refletida no preço dos imóveis.

Devido a uma maior consciência ambiental, ou por motivos de poupança financeira em climatização, impermeabilização, e estética, ou por obrigação

legislativa, os aspetos da eficiência energética estão cada vez mais no centro das atenções, gerando algum valor acrescentado para o imobiliário. Os telhados verdes têm vindo a ganhar alguma popularidade, mas até que ponto estão os consumidores dispostos a pagar por isso? [Ichihara et al. \(2010\)](#) analisaram o efeito nas rendas deste tipo de valência para Nova Iorque e concluíram que existe um prémio de 16%.

A distância para o emprego ou para a escola são fatores a considerar na escolha de habitação. Os efeitos dos índices de congestionamento de tráfego automóvel no percurso para o emprego, por tipo de profissão, foram analisados por [Franklin et al. \(2002\)](#) que concluíram que este índice está positivamente correlacionado com os preços da habitação para os trabalhadores do comércio e universitários, e negativamente correlacionado para os trabalhadores industriais e estudantes até ao 12º ano de ensino.

A medição do custo de vida e a convergência inter-regional com vista à coesão é uma premissa fundamental das políticas regionais, no entanto esta medição e comparabilidade está comprometida pela falta de um índice de preços à habitação. A aquisição de habitação representa o maior peso no gasto mensal de uma família e o maior investimento que um indivíduo faz na vida. As diferenças inter-regionais entre os preços da habitação, em princípio, levariam a uma negociação dos salários a fim de nivelar o poder de compra entre as regiões, à correção dos indicadores económicos regionais, e à correta medição das taxas de pobreza regionais. Num estudo de [Paredes \(2011\)](#) para duas regiões Chilenas, são identificados imóveis em cada uma das regiões que tendo as mesmas características registam preços de mercado diferentes. Estas diferenças devem ser usadas para desenvolver políticas de convergência de rendimentos inter-regional.

O preço da habitação pode ser visto como um fator de alavancagem do consumo, tal como demonstra [Iacoviello \(2004\)](#). A capacidade de endividamento de um indivíduo está estreitamente ligada com o valor do seu imóvel. Em

circunstâncias de capitalização imobiliária o empréstimo que é feito para a compra da habitação é por vezes superior ao preço que é pago pela mesma, sendo o adicional aplicado em consumo. Igualmente anos mais tarde quando parte do empréstimo já foi amortizado e/ou o imóvel valorizou, poderá ser escriturada nova hipoteca que permite ao proprietário aceder mais uma vez ao crédito que se traduzirá em aumento do consumo. O modelo desenvolvido por este autor permite estimar que um aumento do preço do imobiliário irá provocar um aumento de curto prazo no consumo entre 1,33% e 1,94%, por permitir aumentar o endividamento. Ainda confirmando a influência no consumo da valorização imobiliária e da facilidade de acesso ao crédito decorrente da mesma, [Girouard et al. \(2001\)](#) analisaram os países da OCDE. O valor estimado para a elasticidade do consumo à riqueza imobiliária bruta entre 4% a 6% em 1998 para os países do G7. Este efeito pode também ser comprovado pelo lado da poupança. Um estudo de [Engelhardt \(1996\)](#) verificou que os proprietários que obtiveram mais-valias na venda de um imóvel tendem a diluir essas mais-valias na redução do seu padrão de poupança. A abordagem de [Campbell et al. \(2007\)](#) acrescenta que apenas a percepção do aumento de riqueza decorrente de uma valorização do património imobiliário de um indivíduo, é suficiente para estar na base de um aumento do consumo.

Este efeito de alavancagem do consumo provocado pela facilidade de acesso ao crédito tendo como colateral um imóvel em continua valorização beneficiando também de uma diminuição dos constrangimentos regulamentares veio a potenciar o efeito das políticas monetárias sobre o consumo de acordo com o estudo de [Aoki et al. \(2002\)](#) para o Banco de Inglaterra. O Fundo Monetário Internacional reconheceu a importância do efeito gerado pelo significativo aumento do crédito à habitação sobre o aumento do consumo, revelado num documento produzido por [Ludwig et al. \(2001\)](#).

Os ciclos do consumo e da valorização imobiliária têm-se vindo a acompanhar ao longo dos últimos 25 anos, em Inglaterra, de acordo com o estudo de [Attanasio et al. \(2005\)](#) para o Banco de Inglaterra. No entanto, estes concluem

que a relação entre estas duas variáveis se deve ao facto de serem influenciadas pelas mesmas causas, e não de serem dependentes uma da outra. Estas conclusões remetem-nos para a influência das variáveis macroeconómicas, tais como taxas de juro, PIB *per capita*, financeiras, de acesso ao crédito e demográficas. No contexto da formação de preços, [Égert et al. \(2007\)](#) desenvolveram um estudo focalizado na Europa Central e de Leste, que revela igualmente os efeitos da integração na CEE. A influencia de longo prazo (30 anos) da atividade económica sobre os preços dos imóveis em 15 países foi observada por [Adams et al. \(2010\)](#). O resultado empírico foi o de um acréscimo de 0,6% no preço dos imóveis em resposta a um aumento de 1% na atividade económica. Esta variável “atividade económica” é composta por um *mix* das variáveis: emissão de moeda, emprego, consumo, produção industrial e PIB. Alguns estudos encontram a influência das variáveis macroeconómicas no preço da habitação, outros pelo contrário, conseguem comprovar a influência do preço da habitação nas variáveis macroeconómicas, por exemplo o [Elbourne \(2008\)](#). Este autor conclui que 14,3% do decréscimo no consumo pode ser explicado pela queda dos preços na habitação na sequência de um choque nas taxas de juro.

A compra de imóveis destinados à habitação pode ser efetuada por dois grupos com diferentes objetivos e registando determinantes de preferência diferentes: consumidores e investidores. A este respeito [Arraes et al. \(2008\)](#) fizeram um levantamento bastante completo das externalidades que influenciam essa escolha numa cidade brasileira, Fortaleza, Estado do Ceara. Nesse estudo, os autores também consideram edifícios comerciais, o que estará fora do âmbito deste estudo. Os resultados revelam no caso dos apartamentos que “A *concentração de unidades, quando excessiva, reduz o bem-estar por externalidades negativas, atuando negativamente no preço do imóvel.*”, e que o lazer é um fator importante pois permite o bem-estar social, a melhoria da qualidade da vizinhança e das suas relações. A distância ao centro não é relevante, tal como não o é para o caso dos consumidores finais, a existência de áreas comerciais no bairro. Tal como esperado a variável segurança tem um papel preponderante na formação dos preços. No entanto, ao contrário do que seria

eventualmente espectável a proximidade de hospitais e escolas atua negativamente. Aspectos como o tráfego e a poluição sonora estão na origem desta tendência. O risco medido pelo custo de oportunidade do capital e a existência de vários imóveis vagos no mesmo bairro, contribuem negativamente para o preço, sendo considerados pelos autores [Arraes et al. \(2008\)](#) como características financeiras que afetam principalmente o grupo dos investidores.

*“Para muitas pessoas, a sua casa ideal seria uma vivenda espaçosa com a parte da frente próxima de serviços urbanos e as traseiras localizadas numa área verde e calma, como por exemplo um parque público.”* [Jansen et al. \(2011\)](#)

Como este tipo de casas nem sempre estão disponíveis no mercado ou tem um valor superior aquele que um indivíduo está disposto a pagar, na realidade a escolha de uma casa é feita com base numa hierarquia de satisfação de necessidades (perto do trabalho ou da escola, etc.) e num conjunto de restrições (orçamento, disponibilidade no mercado, etc.).

Existe assim uma grande diferença entre o comportamento efetivo que se traduz em preferências reveladas (*revealed preferences*) e as pretensões originais preferências relatadas (*stated preferences*).

As preferências reveladas estão no centro da análise neoclássica e é com base nelas que o presente estudo será desenvolvido.

## **b) Aplicações do modelo de preços hedónicos**

Os autores [Reis et al. \(2006\)](#) estudaram o índice de preços para os automóveis de passageiros em Portugal, no período 1997-2001. Consideram os resultados obtidos robustos, e concluem que o IPC calculado pelo INE está sobreavaliado em cerca de 0.15%. Recomendam por isso que esta análise seja

realizada para outro tipo de bens. Muito embora afirmem que a complexidade deste tipo de estudos torna o seu uso pelos institutos de estatística impraticável. Contrariamente um artigo de [Moulton \(2001\)](#) faz um apanhado da evolução do uso de preços hedónicos nas estatísticas produzidas por quatro institutos oficiais de estatística nos EUA desde 1968: *Bureau of Economic Analysis* (BEA); *Bureau of Labor Statistics* (BLS); *Bureau of the Census*; *Federal Reserve Board*. Estes institutos produzem já um considerável número de índices de preços com base no modelo hedónico, tendo a sua utilização registado um acréscimo mais acentuado desde os anos 80. As implicações no IPC são sempre consideráveis atingindo em alguns casos variações de 6 pontos percentuais. A título de exemplo são produzidos índices de preços hedónicos para casas unifamiliares em construção; preço no produtor para computadores e equipamento periférico; estruturas residenciais multifamiliares; centrais telefónicas digitais; *software*; rendas e equivalente de rendas dos proprietários; vestuário; computadores; microprocessadores; LAN *routers* e *switches*; televisões; equipamento áudio; *camcorders*; colégios; livros; máquinas de lavar e secar roupa; leitores de DVD; fornos micro-ondas; frigoríficos.

[West \(2011\)](#) analisa as componentes que caracterizam o comportamento do preço no carvão térmico marítimo que é importado para a Índia com origem na Indonésia, África do Sul e Austrália. Concluiu que a aparente arbitrariedade nos preços desaparece quando introduzida a questão da qualidade, medida em termos de unidades de energia produzidas, e que não é o custo do transporte que dará vantagem a algum mercado de origem.

Um produto que tem conhecido diversos estudos de aplicação do modelo de preço hedónico é o vinho. Este modelo permite identificar o preço que o consumidor está disposto a pagar e as qualidades mais valorizadas pelos consumidores ([Lima \(2006\)](#); [Gustafson et al. \(2011\)](#); [Lecocq et al. \(2006\)](#); [Nerlove \(1995\)](#)).

### 3. Metodologias para analisar a formação dos preços no mercado imobiliário

Existem três abordagens mais experimentadas para estudar os preços no mercado imobiliário anulando os efeitos da qualidade de acordo com o artigo de [Wooldridge \(2013\)](#): modelo hedônico, método das vendas repetidas e o SPAR (*Sales Appraisal Ratio*). Na impossibilidade de obter dados acerca das características dos imóveis que permitam a utilização do modelo hedônico o autor compara os resultados dos outros dois métodos, para a Holanda. Concluí que a diferença no longo prazo da tendência registada não é negligenciável e que no curto prazo o SPAR é menos volátil.

Igualmente [Bourassa et al. \(2006\)](#) comparou as três abordagens “... *as vantagens e os relativamente limitados inconvenientes do método SPAR torna-o o candidato ideal para ser usado pelas agências governamentais no desenvolvimento dos índices de preços da habitação.*”

O [Jansen et al. \(2011\)](#) numa exaustiva publicação focalizada em métodos para medir as preferências na escolha de um imóvel, identificou 9 métodos e técnicas de análise, que sistematizou na seguinte tabela:

**Tabela 3.1**

Nove métodos e técnicas analíticas destinadas a medir as preferências e escolha na habitação

<b>Métodos e técnicas analíticas</b>	<b>Objetivo</b>
Método de investigação tradicional da procura de habitação	Obter uma visão exata da atual e futura procura de habitação, em termos quantitativos e qualitativos.
Método do plano de decisão em rede	Revelar o processo de escolha pessoal baseado num cabaz de características de uma habitação, o que é essencial, aquelas que podem ser compensadas e as que são consideradas irrelevantes.

<b>Métodos e técnicas analíticas</b>	<b>Objetivo</b>
Método das estruturas relevantes	Para aferir acerca das preferências das pessoas e o porquê dessas escolhas.
Método da utilidade multi-atributo	Fazer uma escolha racional entre as alternativas disponíveis, baseado no perfil da habitação que capitaliza maior utilidade.
Método da análise conjunta	Estimar a função de utilidade que pode ser usada para prever a utilidade global dos perfis residenciais e depois comparar alternativas residenciais em termos das preferências das pessoas.
Método das imagens residenciais	Examinar holisticamente preferências para novas alternativas.
Método do estilo de vida	Construir e localizar as habitações de acordo com as preferências grupos com estilos de vida comuns.
Análise económica neo clássica	Hierarquizar e aferir acerca das preferências por alternativas.
Análise longitudinal	Perante uma questão específica avaliar como as características e circunstâncias num determinado ponto no tempo influenciam a resposta a essa questão.

O [Goodman et al. \(1998\)](#) utiliza os modelos hierárquicos para a área metropolitana de Dallas, recorrendo à qualidade do ensino como variável de segmentação. Mais tarde em [Goodman et al. \(2003\)](#) utilizando mais outros dois fatores de segmentação confirmaram que a utilização de segmentos melhora os resultados da função hedónica.

Tal como já referido anteriormente nos trabalhos de [Gerkman \(2010\)](#); [Goodman et al. \(2003\)](#); [Greene \(2012\)](#); [Dantas et al. \(2007\)](#); [Bruch et al. \(2011\)](#); [Jansen et al. \(2011\)](#); a interdependência espacial das variáveis e as questões de vizinhança são relevantes na avaliação do mercado imobiliário, pelo que as técnicas de econometria espacial começaram a ser aplicadas. [Goodman et al.](#)

(2003) estimam preços hedónicos com desagregação espacial e concluem que a capacidade de previsão do modelo hedónico é melhorada.

Os autores [Bourassa et al. \(2008\)](#) desagregam espacialmente aplicando o modelo hedónico estimado pelo OLS à análise das componentes principais e *cluster* ou usando *dummies* para esses segmentos. Concluiu que a conjugação do modelo hedónico com os métodos de desagregação espacial permitem atingir resultados mais precisos.

[Ramalho \(2010\)](#) identificam quatro métodos para estimar índices de preços hedónicos: método do índice de preços imputados, onde os preços das características variam ao longo do tempo, método das variáveis *dummy*, neste método os preços das características mantêm-se constantes só o termo constante é que muda, método *re-pricing* e o método das características do preço. Também os autores [Silver et al. \(2004\)](#) aplicam os dois primeiros métodos referidos por [Ramalho \(2010\)](#) estimando o índice de preços imputados através de uma média de *Törnqvist*. [Haan \(2004\)](#) demonstra que a imputação de *Törnqvist* pode resultar de um caso particular imputado do método das variáveis *dummy*.

[Dorsey et al. \(2010\)](#) encontrou profundas diferenças entre os resultados obtidos pelo modelo hedónico e pelo método das vendas repetidas. Este último apresenta algumas desvantagens:

- Perdem-se muitas observações;
- Não considera as alterações nas características dos imóveis ao longo do tempo;
- Toda a série tem de ser recalculada cada vez que se inclui um novo período.

Por sua vez o modelo hedónico pode incluir os fatores de correlação espacial, as preferências e as condições de mercado.

[Elliott et al. \(2001\)](#), num estudo realizado para a Direção Geral do Ambiente da Comissão Europeia, concluem que o método dos preços hedónicos é o mais adequado no que respeita a avaliar a contribuição dos fatores ambientais para a formação de preços no imobiliário.

#### **4. Análise Empírica**

##### **a) Tratamento e análise dos dados**

Este estudo pretende encontrar as variáveis que mais contribuem para a formação do preço dos imóveis pela estimação de um modelo econométrico de regressão linear múltipla.

A estimação empírica foi produzida com o uso de uma base de dados de cobertura nacional do registo de apartamentos no Imposto Municipal sobre Imóveis (IMI) respeitante ao ano 2009 contendo 50532 observações. Das variáveis disponíveis foram selecionadas as seguintes:

- Código de freguesia – nomenclatura que codifica as freguesias
- Idade ( IMI\_IDADEP ) – Idade do Imóvel
- Tipologia ( TIPOLO ) – número de assoalhadas
- Valor global do ato ( VAL\_GLOB\_ACTO ) – Valor Total da Aquisição
- Área ( AREA\_D ) – Área Bruta Privativa da Fração
- Preço médio ( Preço\_medio ) – Valor global do ato / Área

À base de dados inicial acrescentou-se as seguintes variáveis obtidas nos CENSOS 2011:

- População ( POP\_2011 ) – população residente em Portugal por freguesia
- Edifícios ( EDIFICIOS ) – número de edifícios em Portugal por freguesia
- Ensino superior ( ENS\_SUPERIOR ) – número de indivíduos com o ensino superior em Portugal por freguesia

O modelo base a estudar é o seguinte:

$$\text{VAL\_GLOB\_ACTO}_i = \beta_0 + \beta_1 \text{IMI\_IDADEP}_i + \beta_2 \text{TIPOL}_i + \beta_3 \text{AREA\_D}_i + \beta_4 \text{POP\_2011}_i + \beta_5 \text{EDIFICIOS}_i + \beta_6 \text{ENS\_SUPERIOR}_i + u_i$$

Onde  $u_i$  é uma variável aleatória, independente e identicamente distribuída (i.i.d.) de média zero e variância constante.

A relevância da localização, enfatizada por [Gerkman \(2010\)](#), é no presente estudo aferida pelas variáveis População ( POP\_2011 ) e Edifícios ( EDIFICIOS ). Estas variáveis são medidas de concentração demográfica, que representam *proxis* de um conjunto de fatores de localização. Considerando que a administração pública e o setor empresarial privado estabelecem o seu mapa de implantação espacial ponderando critérios de ordem demográfica, compreende-se que, quanto maior o número de edifícios e de residentes na freguesia, maior a probabilidade da proximidade a escolas, hospitais, centros comerciais e outros serviços públicos e privados.

A variável “Código de freguesia” foi a chave primária das duas bases de dados utilizadas, ou seja, destinou-se a associar as observações da base de dados

dos CENSOS 2011 à base de dados do IMI, mais concretamente às variáveis “POP\_2011” e “EDIFICIOS”.

A etapa que se seguiu à compilação dos dados numa única base foi a limpeza “económica” da base de dados original, eliminando observações com base nos seguintes critérios:

- Indicação de um uso diferente do residencial;
- *Missing values*;
- Observações com o código de freguesia 141003. Este código existe na base de dados, mas não se encontra nem no CENSUS 2011 nem nos códigos de freguesias das finanças;
- Tipologias superiores a 6 divisões.

A limpeza anterior foi depois complementada com uma eliminação de *outliers* recorrendo à função do *software* **Gretl** “Estatísticas da avaliação da Predição” que assinala um resíduo que excede 2,5 vezes o erro padrão. Esta função foi aplicada ao modelo e as observações assinaladas foram consideradas *outliers*. Este processo foi efetuado 11 vezes até todos os *outliers* terem sido eliminados.

Após os procedimentos de adaptação e limpeza da base de dados o número de observações fixou-se em 33816.

A variável dependente no presente caso de estudo é em termos económicos o preço do imóvel e que de acordo com as variáveis disponíveis pode ser representado pelo valor global do ato ou pelo preço médio. Muito embora [Reichert \(1990\)](#) tenha optado no seu estudo por estimar o preço médio, entendeu-se que se deveria testar ambas as hipóteses de modelo, com o valor global do ato e com o preço médio, como variáveis dependentes. A pertinência de estudar as duas

hipóteses de modelo prende-se com a relevância de confirmar empiricamente a teoria estudada no que respeita às variáveis área e localização. Se alguns autores, como [Reichert \(1990\)](#), concluem que a área é o principal fator na valorização do imóvel, outros autores, como [Gerkman \(2010\)](#), confirmam a valorização dos fatores de localização.

Os modelos foram ensaiados com os logaritmos das variáveis explicativas já que o fato de estas se encontrarem em diferentes dimensões ou seja, escalas muito dísparas resultou em *t-ratio* anormalmente elevados. Igualmente na literatura descrita as análises empíricas resultam da estimação de elasticidades. Após a transformação das variáveis, os *t-ratio* registam ainda valores acima do habitual neste indicador em resultado da fraca variabilidade das variáveis censitárias.

Os resultados apresentados foram obtidos através dos *softwares* Gretl e SPSS, este último essencialmente para o cálculo das Componentes Principais.

Uma vez identificados problemas de heteroscedasticidade todos os modelos foram estimados utilizando a correção de *White* (erros padrão robustos).

## **b) Ensaio Iniciais**

A função explicativa da formação dos preços nos apartamentos foi estimada através de uma regressão múltipla pelo método dos mínimos quadrados ordinários, com 6 variáveis, e 33816 observações da base de dados agora expurgada de *outliers*, da qual resultou o **Modelo 1 (Tabela 4.1)**.

**Tabela 4.1**  
**Modelo 1**

OLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	8.9378	0.0471	189.6156	<0.00001	***
l_POP_2011	-0.1247	0.0057	-22.0283	<0.00001	***
l_EDIFICIOS	-0.0580	0.0034	-17.4717	<0.00001	***
l_ENS_SUPERIOR	0.2463	0.0040	62.2838	<0.00001	***
l_AREA_D	0.5673	0.0100	56.4823	<0.00001	***
l_IMI_IDADEP	-0.1501	0.0014	-106.9463	<0.00001	***
TIPOLO	0.0060	0.0034	2.0245	0.04292	**

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3318.6970	S.E. of regression	0.3133
R-squared	0.5864	Adjusted R-squared	0.5863
F(6, 33809)	6225.0210	P-value(F)	0.0000
Log-likelihood	-8733.2380	Akaike criterion	17480.4800
Schwarz criterion	17539.4800	Hannan-Quinn	17499.3000

### c) Colinearidade

A não colinearidade entre as variáveis independentes é uma das condições de *Gauss-Markov*. Duas variáveis independentes correlacionadas implica que a matriz de observações dos estimadores (X) tem duas colunas lineares e que por isso a sua matriz de variâncias/covariâncias não pode ser invertida, inviabilizando assim a utilização de OLS dado que:

$$\hat{\beta} = (X'X)^{-1} \cdot X'Y$$

O teste à tolerância da colinearidade reflete a possibilidade de existência de colinearidade entre uma variável e as restantes. No presente caso, conclui-se que não existe um problema de multicolinearidade, embora a variável “l\_POP\_2011” tenha um valor muito próximo do limite.

Variance Inflation Factors	
Minimum possible value = 1.0	
Values > 10.0 may indicate a collinearity problema	
l_POP_2011	9.879
l_EDIFICIOS	2.257
l_ENS_SUPERIOR	7.235
l_AREA_D	3.148
l_IMI_IDADEP	1.333
TIPOLO	2.762
VIF(j) = 1/(1 - R(j)^2), where R(j) is the multiple correlation coefficient between variable j and the other independent variables	
Properties of matrix X'X:	
1-norm = 11377207	

Existem neste modelo variáveis correlacionadas como se pode verificar na matriz de correlações apresentada entre as duas variáveis indicadoras da dimensão de um apartamento: “TIPOLO” e “l\_AREA\_D” (**Tabela 4.2**).

**Tabela 4.2**  
**Correlation coefficients**

5% critical value (two-tailed) = 0.0107 for n = 33816

l_AREA_D	TIPOLO	
1.0000	0.7643	l_AREA_D
	1.0000	TIPOLO

Igualmente entre as variáveis de localização “l\_EDIFICIOS” e “l\_POP\_2011” verifica-se alguma correlação (**Tabela 4.3**).

**Tabela 4.3**  
**Correlation coefficients**

5% critical value (two-tailed) = 0.0107 for n = 33816

I_POP_2011	I_EDIFICIOS	
1.0000	0.7120	I_POP_2011
	1.0000	I_EDIFICIOS

Face a esta situação realizaram-se alguns testes de colinearidade, nomeadamente uma regressão entre a variável independente e uma das variáveis com forte correlação e comparou-se o resultado do *t-ratio* resultante desta regressão e esse valor no caso do modelo original.

**Tabela 4.4**

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	10.9073	0.0268	407.4210	<0.00001	***
I_POP_2011	0.0669	0.0028	23.8729	<0.00001	***

**Tabela 4.5**

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	10.9097	0.0092	1180.8409	<0.00001	***
TIPOLO	0.1934	0.0027	71.1834	<0.00001	***

O *t-ratio* associado ao coeficiente da variável “I\_POP\_2011” inicial foi de -21.8228 e isoladamente foi de 23.8729. A diferença de grandeza é considerável e quando conjugada com a mudança de sinal revela um comportamento diferente dessa variável independente quando se encontra em conjunto com as restantes e de forma individual. No caso da variável “TIPOLO” o sinal mantém-se, mas a diferença de grandeza é elevada, passou de 2.2368 para 71,1834. Estes resultados indiciam a existência de colinearidade a qual irá ser tratada.

No âmbito do teste F, foi experimentada uma outra transformação do modelo que permitiria eliminar o problema resultante da colinearidade. Esta proposta incide na eliminação da variável “TIPOLO” já que o *t-ratio* se encontra próximo da não rejeição da hipótese nula para o caso de um teste *t* bilateral com nível de significância de 0,01.

Test for omission of variables -  
 Null hypothesis: parameters are zero for the variables  
 TIPOLO  
 Test statistic:  $F(1, 33809) = 5.00328$   
 with p-value =  $P(F(1, 33809) > 5.00328) = 0.0253059$

Esta transformação melhora o resultado do Critério de Informação Bayesiano ou de Schwarz (BIC) à especificação do modelo. O resultado do teste F é superior ao valor crítico e o *p-value* associado assume um valor muito pequeno, indicando a rejeição da hipótese nula. Conclui-se que a variável “TIPOLO” embora tenha uma fraca significância estatística quando analisada de forma isolada, no conjunto do modelo revela-se estatisticamente significativa. O modelo na hipótese de não rejeição da hipótese nula seria o seguinte (**Tabela 4.6**):

**Tabela 4.6**  
**Modelo 2**

OLS, using observations 1-33816  
 Dependent variable: l\_VAL\_GLOB\_ACTO  
 Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value	
const	8.8883	0.0388	229.1437	<0.00001	***
l_POP_2011	-0.1237	0.0057	-21.8986	<0.00001	***
l_EDIFICIOS	-0.0596	0.0034	-17.6283	<0.00001	***
l_ENS_SUPERIOR	0.2459	0.0040	62.3109	<0.00001	***
l_AREA_D	0.5825	0.0061	94.7315	<0.00001	***
l_IMI_IDADEP	-0.1491	0.0013	-114.9369	<0.00001	***

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3319.1880	S.E. of regression	0.3133
R-squared	0.5864	Adjusted R-squared	0.5863
F(5, 33810)	9585.3810	P-value(F)	0.0000
Log-likelihood	-8735.7400	Akaike criterion	17483.4800
Schwarz criterion	17534.0500	Hannan-Quinn	17499.6200

Nesta fase existe um problema de colinearidade que não pode ser corrigido com a eliminação de uma das variáveis envolvidas. Nenhuma das variáveis pode ser eliminada porque ambas são importantes para a especificação do modelo (teste

F) ainda que o *t-ratio* da variável “TIPOLO” esteja perto da região de não rejeição da hipótese nula.

No caso da colinearidade entre as variáveis “l\_POP\_2011” e a “l\_EDIFICIOS” não se colocou a hipótese da eliminação já que o *t-ratio* indica a significância estatística de ambas.

#### d) Componentes principais

Uma segunda proposta para a resolução do problema da colinearidade é o recurso a componentes principais. Esta transformação é construída para que a PC1 (Componente Principal 1) explique o máximo de variabilidade dos dados e a PC2 (Componente Principal 2) contenha a máxima variância condicionada ao facto de ser ortogonal ou seja não haver correlação entre as duas componentes.

##### I. Componentes principais - “TIPOLO” e “l\_AREA\_D”

Estimaram-se os valores próprios (*Eigenvalue*) (**Tabela 4.7**) e os vetores próprios (*Eigenvectors*) (**Tabela 4.8**) das componentes principais formadas pelas variáveis “TIPOLO” e “l\_AREA\_D”.

**Tabela 4.7**  
**Eigenanalysis of the Correlation Matrix**

Component	Eigenvalue	Proportion	Cumulative
1	1.7643	0.8822	0.8822
2	0.2357	0.1178	1.0000

**Tabela 4.8**  
**Eigenvectors (component loadings)**

	PC1	PC2
TIPOLO	0.939	-0.343
l_AREA_D	0.939	0.343

$$PC1\_AREA\_TIPOLO = (l\_AREA\_D + TIPOLO) * 0.939 \quad (1)$$

$$PC2\_AREA\_TIPOLO = (l\_AREA\_D - TIPOLO) * 0.343 \quad (2)$$

Conclui-se que a correlação entre as componentes (**Tabela 4.9**) é significativamente diferente de zero (próxima de -0.86) pelo que o problema da colinearidade ainda se mantém. As componentes estão correlacionadas entre si.

**Tabela 4.9**  
**Correlation coefficients**

5% critical value (two-tailed) = 0.0107 for n = 33816

PC1_AREA_TIPOLO	PC2_AREA_TIPOLO	
1.0000	-0.8587	PC1_AREA_TIPOLO
	1.0000	PC2_AREA_TIPOLO

A correlação entre as variáveis pode ser resultado da grande diferença de escala das variáveis pelo que se procedeu a sua normalização, utilizando a seguinte formulação:

$$N\_l\_AREA\_D = \frac{l\_AREA\_D - \overline{l\_AREA\_D}}{se(l\_AREA\_D)} \quad (3)$$

$$N\_TIPOLO = \frac{TIPOLO - \overline{TIPOLO}}{se(TIPOLO)} \quad (4)$$

Após a normalização das variáveis as componentes principais já apresentam o valor zero para a sua correlação (**Tabela 4.10**). Os valores e vetores próprios mantêm-se.

**Tabela 4.10**  
**Correlation coefficients**

5% critical value (two-tailed) = 0.0107 for n = 33816

PC1_N_AREA_TIPOLO	PC2_N_AREA_TIPOLO	
1.0000	0.0000	PC1_N_AREA_TIPOLO
	1.0000	PC2_N_AREA_TIPOLO

Procedeu-se à estimação do modelo com as componentes principais das variáveis normalizadas que resultou no seguinte (**Tabela 4.11**):

**Tabela 4.11**  
**Modelo 3**

OLS, using observations 1-33816

Dependent variable: l\_VAL\_GLOB\_ACTO

Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value	
const	11.5532	0.0230	501.8592	<0.00001	***
l_POP_2011	-0.1247	0.0057	-22.0283	<0.00001	***
l_EDIFICIOS	-0.0590	0.0034	-17.4717	<0.00001	***
l_ENS_SUPERIOR	0.2463	0.0040	62.2838	<0.00001	***
l_IMI_IDADEP	-0.1501	0.0014	-106.9463	<0.00001	***
PC1_N_AREA_TIPOLO	0.1110	0.0012	94.7129	<0.00001	***
PC2_N_AREA_TIPOLO	0.2854	0.0092	30.829	<0.00001	***

Mean dependent var	11.5433	S.D. dependent var	0.4871
Sum squared resid	3318.6970	S.E. of regression	0.3133
R-squared	0.5864	Adjusted R-squared	0.5863
F(6, 33809)	6225.0210	P-value(F)	0.0000
Log-likelihood	-8733.2380	Akaike criterion	17480.4800
Schwarz criterion	17539.4800	Hannan-Quinn	17499.3000

Neste modelo os coeficientes das componentes apresentam o sinal expectável em termos económicos e os correspondentes *t-ratio* são significativos.

Apesar da significância estatística das componentes evidenciada pelos *t-ratio* realizou-se igualmente o teste F sobre as componentes sendo o resultado apresentado no Anexo B.

## II. Componentes principais - “1\_EDIFICIOS” e “1\_POP\_2011”

Executaram-se as operações anteriores também para o par de variáveis “1\_EDIFICIOS” e “1\_POP\_2011”, nomeadamente no que diz respeito à normalização das variáveis.

**Tabela 4.12**  
**Eigenanalysis of the Correlation Matrix**

Component	Eigenvalue	Proportion	Cumulative
1	1.7120	0.8560	0.8560
2	0.2880	0.1440	1.0000

**Tabela 4.13**  
**Eigenvectors (component loadings)**

	PC1	PC2
1_POP_2011	0.925	-0.379
1_EDIFICIOS	0.925	0.379

O modelo que resulta destas componentes principais (**Tabela 4.14**) apresenta estimadores significantes.

**Tabela 4.14**  
**Modelo 4**

OLS, using observations 1-33816  
Dependent variable: 1\_VAL\_GLOB\_ACTO  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	7.2897	0.0480	151.7329	<0.00001	***
1_ENS_SUPERIOR	0.2463	0.0040	62.2838	<0.00001	***
1_IMI_IDADEP	-0.1501	0.0014	-106.9463	<0.00001	***
1_AREA_D	0.5673	0.0100	56.4823	<0.00001	***
TIPOLO	0.0070	0.0034	2.0245	0.04292	**
PC1_N_1_POP_EDIF	-0.0876	0.0025	-35.7074	<0.00001	***
PC2_N_1_POP_EDIF	0.0944	0.0092	10.2150	<0.00001	***

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3318.6970	S.E. of regression	0.3133
R-squared	0.5864	Adjusted R-squared	0.5863

F(6, 33809)	6225.0210	P-value(F)	0.0000
Log-likelihood	-8733.2380	Akaike criterion	17480.4800
Schwarz criterion	17539.4800	Hannan-Quinn	17499.3000

Considerando a matriz de correlação (**Tabela 4.15**) e a parte da variação total das duas variáveis que é explicada por cada componente (**Tabela 4.16**) optou-se por não considerar a componente 2 uma vez que a sua capacidade explicativa e correlação com as variáveis que a compõem é reduzida.

**Tabela 4.15**  
**Correlation coefficients**

5% critical value (two-tailed) = 0.0107 for n = 33816

I_POP_2011	I_EDIFICIOS	PC1_N_POP_EDIFICIOS	PC2_N_POP_EDIFICIOS	
1.0000	0.7120	0.9252	-0.3795	I_POP_2011
	1.0000	0.9252	0.3795	I_EDIFICIOS
		1.0000	-0.0000	PC1_N_POP_EDIFICIOS
			1.0000	PC2_N_POP_EDIFICIOS

**Tabela 4.16**  
**Capacidade explicativa por componente**

PC1_N_I_POP_EDIF	$1.712/2 = 0.856$
PC2_N_I_POP_EDIF	$0.288/2 = 0.144$

Embora o resultado dos testes de significância estatística e a significância económica sejam favoráveis à escolha do **Modelo 3**, considerou-se que seria indispensável a este estudo a formulação de um modelo hedónico que também permita ultrapassar o problema da colinearidade.

O **Modelo 4** apresenta-se como uma boa solução para o problema da colinearidade entre as variáveis “I\_POP\_2011” e “I\_EDIFICIOS”.

e) **Modelo Hedónico**

Outra técnica possível e eventualmente a mais indicada para tratar a variável discreta e qualitativa “TIPOLO”, e em coerência metodológica com a literatura descrita, é proceder a uma estimação do modelo hedónico. Este modelo implica a criação de variáveis *dummy* (binárias) definidas da seguinte forma (**Tabela 4.17**):

**Tabela 4.17**  
**Variáveis *Dummy***

	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>	<b>D5</b>	<b>D6</b>
<b>Tipologia</b>	T1=1	T2=1	T3=1	T4=1	T5=1	T6=1
<b>Outras tipologias</b>	0	0	0	0	0	0

Estimou-se o **Modelo 5** com todas as variáveis e os resultados obtidos foram os melhores até ao momento (**Tabela 4.18**).

**Tabela 4.18**  
**Modelo 5**

OLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
l_ENS_SUPERIOR	0.2114	0.0025	84.0764	<0.00001	***
l_AREA_D	-0.1508	0.0013	-112.1598	<0.00001	***
l_IMI_IDADEP	0.5900	0.0085	69.6701	<0.00001	***
PC1_N_I_POP_EDIF	-0.0680	0.0017	-39.4740	<0.00001	***
D1	7.5699	0.0399	189.9795	<0.00001	***
D2	7.4883	0.0407	183.8681	<0.00001	***
D3	7.4585	0.0432	172.7353	<0.00001	***
D4	7.4518	0.0453	164.4571	<0.00001	***
D5	7.5723	0.0480	157.7914	<0.00001	***
D6	7.6741	0.0537	142.8212	<0.00001	***

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3283.5530	S.E. of regression	0.3117
R-squared	0.9993	Adjusted R-squared	0.9993
F(10, 33806)	4644010	P-value(F)	0.0000
Log-likelihood	-8553.2350	Akaike criterion	17126.4700
Schwarz criterion	17210.7600	Hannan-Quinn	17153.3600

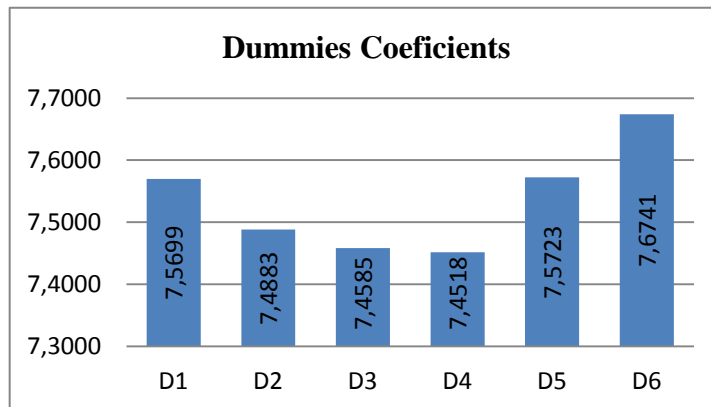
Uma especificidade deste modelo é a não existência de constante. Esta foi a resposta escolhida para evitar a armadilha da variável *dummy*. Outra hipótese seria considerar uma das variáveis *dummy* como “grupo de controle”, ou seja não a considerar na regressão e assim a sua influência seria resultado da situação na qual os coeficientes de todas as outras variáveis são zero. A vantagem de não incluir a constante é incluir todas as variáveis e evitar o problema da escolha do “grupo de controle”, sendo mais imediato e intuitivo verificar a sua contribuição económica para o modelo bem como a verificação do seu *t-ratio*. A desvantagem de não incluir a constante é que o cálculo para a soma do quadrado dos resíduos (SQR) não centra os  $y_i$  na sua média, consequentemente,  $SQR_D > SQR$ , e assim o  $R_D^2 > R^2$ . A bondade do ajustamento terá sempre um valor muito superior no modelo hedónico e por isso devemos continuar a considerar o  $R^2$  anterior.

O estimador da variável *dummy* é um indicador do preço implícito de cada assoalhada adicionada. Um apartamento com duas assoalhadas tem um valor 8% inferior a um com apenas uma assoalhada<sup>3</sup>. Quando aumentamos uma assoalhada a um apartamento com duas assoalhadas, o seu preço diminui em 3%. O aumento de três assoalhadas para quatro assoalhadas também regista uma diminuição do preço em 0.7%. Quando se aumenta uma assoalhada a um apartamento com 4 assoalhadas acontece que o preço aumenta e o mesmo acontece na tipologia seguinte (**Figura 4.1**).

---

<sup>3</sup>  $(7.56992 - 7.48832) = 8,16\%$

**Figura 4.1**  
**Coefficientes das variáveis *Dummy***



A relação positiva que se espera entre preço e aumento do número de assoalhadas nem sempre se verifica. Esta situação decorre da existência de casas com uma tipologia inferior que correspondem a áreas maiores e vice-versa.

A área está mais correlacionada com o preço do que a tipologia como se pode verificar nesta matriz de correlações (**Tabela 4.19**).

**Tabela 4.19**  
**Correlation coefficients**

5% critical value (two-tailed) = 0.0107 for n = 33816

I_VAL_GLOB_ACTO	TIPOLO	AREA_D	
1.0000	0.3610	0.5823	I_VAL_GLOB_ACTO
	1.0000	0.7354	TIPOLO
		1.0000	AREA_D

Fazendo uma análise aos quartis por tipologia verifica-se uma intrusão dos quartis das áreas correspondentes a uma tipologia nas tipologias seguintes. Por exemplo na **Tabela 4.20** pode-se verificar que, existem apartamentos da tipologia 1 com áreas pertencentes à tipologia 3. A área mais elevada das tipologias 1 e 2 encontram-se no 4ºquartil das tipologias 3,4 e 5. O mesmo acontece para a tipologia 3 onde se verifica que a área mais elevada encontra-se no 4ºquartil das

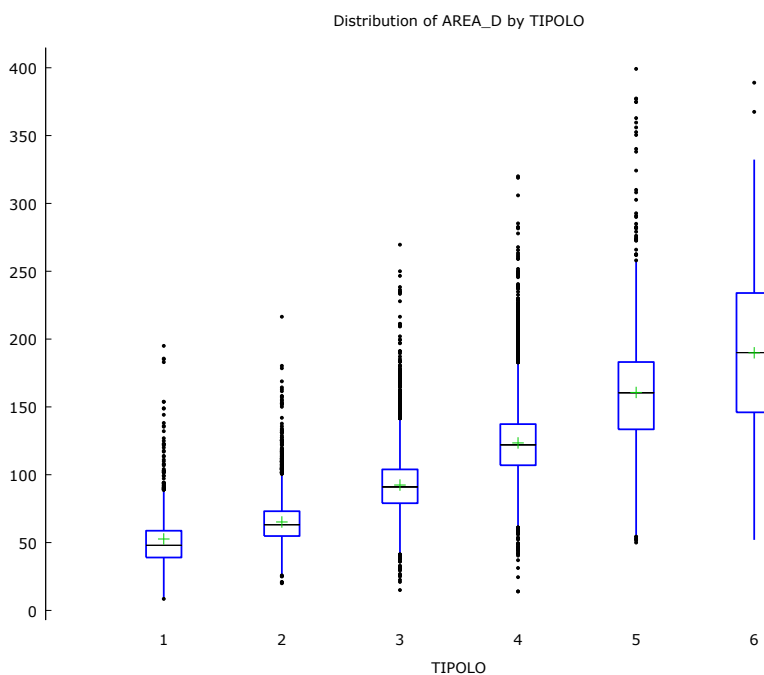
tipologias 4, 5 e 6. Isto explica o decréscimo de preços em certas tipologias quando o número de assoalhadas aumenta.

**Tabela 4.20**  
**Análise de quartis**

Quartis da área	Número de divisões					
	Tip. 1	Tip. 2	Tip. 3	Tip. 4	Tip. 5	Tip. 6
1º quartil	39,00	54,83	79,00	107,00	133,57	146,14
2º quartil	48,00	63,12	91,00	122,00	160,38	190,00
3º quartil	58,70	73,12	103,93	137,33	183,01	232,49
4º quartil	195,00	216,48	269,60	320,00	399,17	389,00
Media	52,65	65,17	92,41	123,45	160,65	189,89
Nº Observações	973,00	5130,00	13611,00	11953,00	1950,00	199,00
Peso	2,9%	15,2%	40,3%	35,3%	5,8%	0,6%
2º quartil - 1º quartil	9,00	8,29	12,00	15,00	26,81	43,87
3º quartil - 2º quartil	10,70	10,00	12,93	15,33	22,63	42,49
4º quartil - 3º quartil	136,30	143,37	165,67	182,67	216,16	156,51

Como se verifica no gráfico seguinte (**Figura 4.2**) existem sempre observações em determinada tipologia com a mesma área que observações na tipologia seguinte.

**Figura 4.2**



**f) Variáveis de interação**

Considerando a hipótese de que em diferentes épocas a tendência no que respeita à dimensão das assoalhadas se terá alterado, testou-se a hipótese da interação entre as variáveis área e tipologia resultar da idade do imóvel.

Construiu-se um modelo com variáveis de interação onde se incluiu as *dummies* definidas anteriormente para a variável “TIPOLO” e a combinação de cada *dummy* com a variável “1\_IMI\_IDADEP” (Tabela 4.21).

**Tabela 4.21**  
**Modelo 6**

OLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO

	Coefficient	Std. Error	t-ratio	p-value	
l_ENS_SUPERIOR	0.2120	0.0025	84.2774	<0.00001	***
l_AREA_D	0.5850	0.0085	68.8852	<0.00001	***
PC1_N_POP_EDIFICIOS	-0.0682	0.0017	-39.7546	<0.00001	***
D1	7.5667	0.0408	185.2907	<0.00001	***
D2	7.4932	0.0409	183.3706	<0.00001	***
D3	7.4777	0.0433	172.8844	<0.00001	***
D4	7.4712	0.0454	164.3996	<0.00001	***
D5	7.6241	0.0487	156.5384	<0.00001	***
D6	7.8250	0.0594	131.6408	<0.00001	***
D1_IDADE	-0.1377	0.0068	-20.2963	<0.00001	***
D2_IDADE	-0.1432	0.0030	-47.7406	<0.00001	***
D3_IDADE	-0.1512	0.0020	-77.4767	<0.00001	***
D4_IDADE	-0.1504	0.0021	-70.4494	<0.00001	***
D5_IDADE	-0.1740	0.0049	-35.8142	<0.00001	***
D6_IDADE	-0.2215	0.0133	-16.6843	<0.00001	***

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3277.5350	S.E. of regression	0.3114
R-squared	0.5915	Adjusted R-squared	0.5914
F(14, 33801)	3228975	P-value(F)	0.0000
Log-likelihood	-8522.2190	Akaike criterion	17074.4400
Schwarz criterion	17200.8700	Hannan-Quinn	17114.7700

Com este novo modelo podemos aferir acerca da influência da idade por cada tipologia uma vez que este resulta num modelo por cada tipologia. Por exemplo, para a tipologia 2 o modelo é o seguinte:

$$\begin{aligned} \text{I VAL GLOB ACTO} = & 7.49315 + 0.211975 \text{ I\_ENS\_SUPERIOR} \\ & + 0.584973 \text{ I\_AREA\_D} - 0.089647 \text{ PC1\_N\_1\_POP\_EDIF} \\ & - 0.143216 \text{ D2\_IDADE} \end{aligned}$$

Os resultados são interessantes, por permitirem a movimentação da reta do ajustamento quer na origem quer no declive, a robustez dos estimadores que apresentam *t-ratios* muito elevados para as novas variáveis combinadas e *p-values* pouco diferentes de zero indica-nos que esta formulação deve ser adotada.

#### g) Preço médio e modelo por tipologia

Experimentou-se uma análise, à semelhança do estudo de [Reichert \(1990\)](#), replicando o modelo anterior embora nesta experimentação a variável dependente considerada seja o preço médio por m<sup>2</sup> e não o valor global do imóvel.

**Tabela 4.22**  
**Modelo 7**

OLS, using observations 1-33816  
Dependent variable: l\_Pre\_\_o\_m\_\_dio  
Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value	
I_ENS_SUPERIOR	0.2010	0.0028	74.3971	<0.00001	***
PC1_N_POP_EDIFICIOS	-0.0643	0.0018	-35.0256	<0.00001	***
D1	5.9391	0.0276	214.9257	<0.00001	***
D2	5.7459	0.0227	253.4464	<0.00001	***
D3	5.5778	0.0213	261.7617	<0.00001	***
D4	5.4553	0.0212	257.7364	<0.00001	***
D5	5.4889	0.0236	232.2834	<0.00001	***
D6	5.5918	0.0464	120.5830	<0.00001	***
D1_IDADE	-0.1207	0.0097	-12.4153	<0.00001	***
D2_IDADE	-0.1171	0.0034	-34.3328	<0.00001	***

D3_IDADE	-0.1207	0.0019	-64.1539	<0.00001	***
D4_IDADE	-0.1175	0.0020	-57.4591	<0.00001	***
D5_IDADE	-0.1310	0.0052	-25.1401	<0.00001	***
D6_IDADE	-0.1680	0.0179	-9.4004	<0.00001	***

Mean dependent var	6.9737	S.D. dependent var	0.4085
Sum squared resid	3509.1410	S.E. of regression	0.3222
R-squared	0.3780	Adjusted R-squared	0.3778
F(13, 33802)	1190785	P-value(F)	0.0000
Log-likelihood	-9676.6910	Akaike criterion	19381.3800
Schwarz criterion	19499.3800	Hannan-Quinn	19419.0300

O modelo seguinte é semelhante ao **Modelo 1** utilizando o preço médio por m<sup>2</sup> como variável dependente e a variável “TIPOLO” como variável explicativa (**Tabela 4.23**).

**Tabela 4.23**  
**Modelo 8 - Global**

OLS, using observations 1-33816  
Dependent variable: l\_Preço\_medio

Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value	
const	7.230	0.0250	288.8173	<0.00001	***
l_POP_2011	-0.0964	0.0058	-16.5745	<0.00001	***
l_EDIFICIOS	-0.0669	0.0035	-19.1449	<0.00001	***
l_ENS_SUPERIOR	0.2322	0.0040	57.4436	<0.00001	***
TIPOLO	-0.1195	0.0022	-54.2598	<0.00001	***
l_IMI_IDADEP	-0.1186	0.0013	-91.6279	<0.00001	***

Mean dependent var	6.9737	S.D. dependent var	0.4085
Sum squared resid	3573.9720	S.E. of regression	0.3251
R-squared	0.3665	Adjusted R-squared	0.3664
F(5, 33810)	3841.4920	P-value(F)	0.0000
Log-likelihood	-9986.2150	Akaike criterion	19984.4300
Schwarz criterion	20035.0000	Hannan-Quinn	20000.5600

Paralelamente experimentou-se separar a base de dados original por tipologia e estimou-se um modelo para cada tipologia, apresentados no **Anexo C**.

Na **Tabela 4.24** apresenta-se um quadro comparativo dos coeficientes para os modelos por tipologia e para o modelo que inclui todas as tipologias (**Tabela 4.23**). Verifica-se que a constante assume um valor muito elevado e que a significância económica dos coeficientes é reduzida. Apesar de algumas variações no valor pode-se considerar que as variáveis influenciam o preço de forma similar em todas as tipologias exceto no caso da influência da variável “EDIFICIOS” que para os apartamentos com 1 assoalhada tem um sinal positivo e nas restantes tem um sinal negativo. Depreende-se que os indivíduos que procuram apartamentos pequenos têm preferência por zonas com maior concentração.

**Tabela 4.24**  
**Análise aos Coeficientes**

<i>Variáveis</i>	<i>Coefficientes</i>						
	<i>Global</i>	<i>Tip. 1</i>	<i>Tip. 2</i>	<i>Tip. 3</i>	<i>Tip. 4</i>	<i>Tip. 5</i>	<i>Tip. 6</i>
const	7.230	7.2170	7.3166	6.9896	6.4128	6.5184	5.8100
l_POP_2011	-0.0964	-0.1899	-0.1572	-0.1188	0.0024	-0.1024	0.1001
l_EDIFICIOS	-0.0669	0.1003	-0.0126	-0.0611	-0.1125	-0.1019	-0.1340
l_ENS_SUPERIOR	0.2322	0.1722	0.2115	0.2359	0.1983	0.2813	0.1601
l_IMI_IDADEP	-0.1195	-0.1202	-0.1099	-0.1176	-0.1227	0.0131	-0.0220
TIPOLO	-0.1186						

Os *t-ratios* apresentam uma grande variabilidade o que é um indicador de que a opção pelo modelo não linear que permite uma maior adaptação do modelo para cada tipologia pode conferir alguma consistência aos resultados.

**Tabela 4.25**  
**Análise aos t-ratio**

<i>Variáveis</i>	<i>t-ratio</i>						
	<i>Global</i>	<i>Tip. 1</i>	<i>Tip. 2</i>	<i>Tip. 3</i>	<i>Tip. 4</i>	<i>Tip. 5</i>	<i>Tip. 6</i>
const	288.8173	45.1729	110.6885	198.6157	179.9467	54.0991	11.9490
POP_2011	-16.5745	-4.1718	-9.4338	-13.5292	0.2601	-3.3103	0.8070
EDIFICIOS	-19.1449	4.0066	-1.3779	-11.8427	-20.1126	-5.7797	-1.9591
ENS_SUPERIOR	57.4436	5.6759	18.6104	38.3008	31.3881	14.0392	2.2270
IMI_IDADEP	-91.6279	-13.1521	-31.9463	-62.1497	-61.3312	1.9033	-0.7305
TIPOLO	-54.2598						

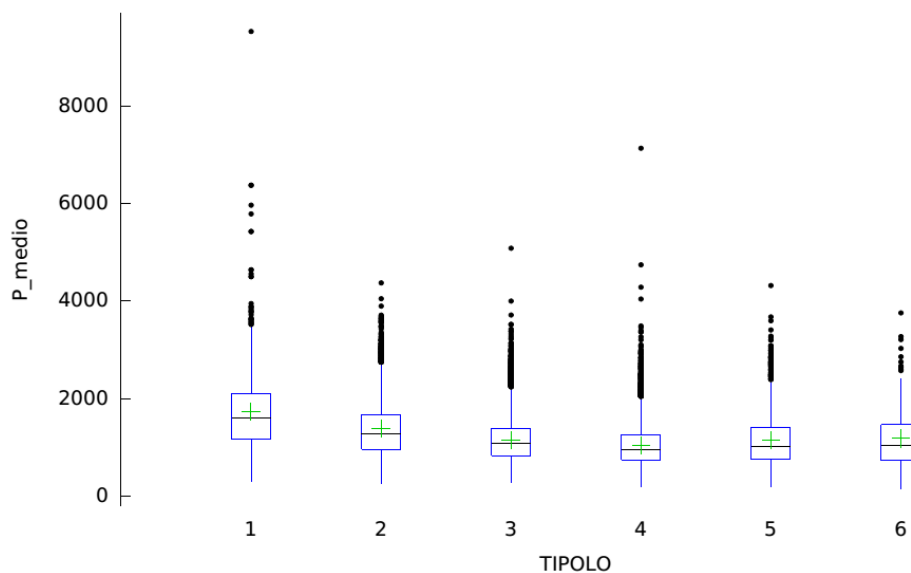
**Tabela 4.26**  
**Análise aos testes de especificação do modelo**

Testes	Global	Tip. 1	Tip. 2	Tip. 3	Tip. 4	Tip. 5	Tip. 6
F	4028.65	65.23	395.63	1661.36	1920.49	54.10	11.95
Max Verossimilhança	-9803.31	-508.85	-2007.31	-3323.05	-2613.62	-3.31	0.81
AIC	19669.19	1052.10	4057.33	6693.69	5274.18	-5.78	-1.10
BIC	19618.62	1027.70	4024.62	6656.10	5237.23	14.04	2.23
HQC	19634.75	1036.99	4036.07	6668.63	5249.63	1.903	-0.73

Na **Figura 4.3** podemos confirmar os resultados obtidos na estimação hedónica. Enquanto poderia ser esperado que o preço do imóvel aumentasse com o número de assoalhadas podemos verificar que isso nem sempre acontece já que a área é um fator mais indicativo da dimensão enquanto o número de assoalhadas é apenas uma característica cuja quantidade nem sempre é valorizada de forma positiva. Neste sentido o preço médio por metro quadrado diminuiu entre as tipologias 1 e 4.

**Figura 4.3**

Distribution of P\_medio by TIPOLO



## h) Testes à especificação do modelo

Testar a especificação do modelo a fim de garantir que se obteve o modelo mais parcimonioso, isto é, encontrar o modelo que envolva o mínimo de parâmetros possíveis a serem estimados e seja explicativo do comportamento da variável dependente, implica um conjunto testes. Analisou-se a função da máxima verossimilhança, o Critério de Informação de Akaike (AIC), o Critério Bayesiano de Schwarz (BIC) e o Critério de Informação de Hannan-Quinn (HQC).

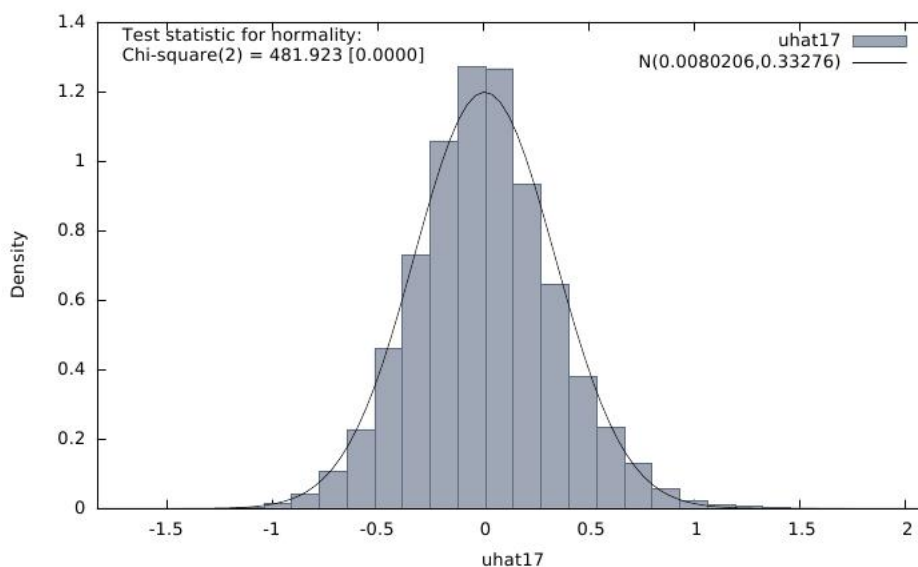
Estes testes (Critérios) permitem avaliar, por comparação, qual dos modelos tem a melhor especificação.

**Tabela 4.27**  
**Análise comparativa dos modelos**

	<b>Modelo 1</b>	<b>Modelo 5</b>	<b>Modelo 6</b>	<b>Modelo 7</b>
<b>F</b>	7989.598	4644010	3496.616	1580.301
<b>Max Verossimilhança</b>	-8733.238	-8553.235	-8522.219	-9676.691
<b>AIC</b>	17539.48	17210.76	17200.87	19499.38
<b>BIC</b>	17480.48	17126.47	17074.44	19381.38
<b>HQC</b>	17499.30	17153.36	17114.77	19419.03

Para o **Modelo 6** com as componentes principais e variáveis não lineares os valores para os testes de especificação do modelo são mais favoráveis, quando comparado com os restantes modelos, como se pode verificar no quadro acima onde são apresentadas as medidas de especificação do modelo para alguns dos vários modelos estimados. Verifica-se igualmente que para o **Modelo 6** com as componentes principais ambos, os *t-ratio* e os coeficientes, são mais significativos comparativamente ao **Modelo 1** ainda com a vantagem de se ter eliminado a colinearidade. Igualmente sem colinearidade o modelo hedónico (**Modelo 5**) melhora ligeiramente todas as estatísticas.

**Figura 4.4**  
**Análise à normalidade dos resíduos**



Confirmando-se que o modelo mais bem especificado é o **Modelo 6**, é agora importante identificar-se e aplicar-se diferentes metodologias de ajustamento aos problemas de heteroscedasticidade.

#### **i) Heterocedasticidade**

Quando a variância do erro, i.e. dos fatores não observados, não é constante o modelo sofre de heterocedasticidade. Esta característica embora não afete a consistência dos estimadores nem a bondade do ajustamento,  $R^2$ , resulta na falta de consistência dos testes t, F e intervalos de confiança. Considerando que estes testes são fundamentais na avaliação de um modelo econométrico os testes à homoscedasticidade são indispensáveis. Vamos testar o modelo hedónico que se apresenta como sendo o melhor modelo.

A hipótese nula de homoscedasticidade é:

$$H_0 : E(u^2 | x_1, x_2, \dots, x_k) = E(u^2) \neq \sigma^2$$

Utilizaram-se os testes de *Breusch-Pagan* e *White* para testar a hipótese nula de homoscedasticidade. O *pvalue* indica que a hipótese nula de homoscedasticidade é rejeitada. Assim, concluímos que o modelo é heterocedástico.

### **Breusch-Pagan test for heteroskedasticity**

Null hypothesis: heteroskedasticity not present

Test statistic: LM = 737.336203,

with p-value = P(Chi-square(14) > 737.336203) = 0.000000

### **White's test for heteroskedasticity**

Test statistic: TR<sup>2</sup> = 2382.490456,

with p-value = P(Chi-square(59) > 2382.490456) = 0.000000

Uma das soluções possíveis para o problema da heteroscedasticidade é a estimação de um Modelo FWLS (*Feasible Weighted Least Squares*)<sup>4</sup>. A construção de um modelo deste tipo implica a estimação de um fator de ponderação que resulta das seguintes operações:

1. Obter os resíduos da regressão;
2. Logaritmizar o quadrado dos resíduos;
3. Estimar a regressão com o vetor anteriormente calculado
4. Como variável independente;
5. Calcular o inverso da exponencial dos valores estimados da regressão anterior.

---

<sup>4</sup> Pag. 287, Wooldridge, J. M., Ed. (2013). Introductory econometrics - A modern approach. Mason, Cengage Learning.

O vetor obtido é utilizado como ponderador na regressão pelo método dos mínimos quadrados ponderados.

**Tabela 4.28**  
**Modelo 9**

WLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO  
Variable used as weight: h\_2

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
l_ENS_SUPERIOR	0.2126	0.0025	84.7993	<0.00001	***
l_AREA_D	0.5862	0.0086	67.8228	<0.00001	***
PC1_N_I_POP_EDIF	-0.0892	0.0022	-39.8017	<0.00001	***
D1	7.5603	0.0421	179.5760	<0.00001	***
D2	7.4830	0.0413	181.3828	<0.00001	***
D3	7.4675	0.0436	171.3742	<0.00001	***
D4	7.4616	0.0459	162.7278	<0.00001	***
D5	7.6112	0.0495	153.7737	<0.00001	***
D6	7.8087	0.0694	112.4689	<0.00001	***
D1_IDADE	-0.1372	0.0091	-15.0663	<0.00001	***
D2_IDADE	-0.1431	0.0035	-41.0161	<0.00001	***
D3_IDADE	-0.1517	0.0018	-82.2027	<0.00001	***
D4_IDADE	-0.1508	0.0021	-72.5473	<0.00001	***
D5_IDADE	-0.1746	0.0056	-31.4118	<0.00001	***
D6_IDADE	-0.2215	0.0184	-12.0181	<0.00001	***

Statistics based on the weighted data:

Sum squared resid	125102.3000	S.E. of regression	1.9238
R-squared	0.5748	Adjusted R-squared	0.5746
F(15, 33801)	3045.9570	P-value(F)	0.0000
Log-likelihood	-70101.8400	Akaike criterion	140233.7000
Schwarz criterion	140360.1000	Hannan-Quinn	140274.0000

Statistics based on the original data:

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3277.6470	S.E. of regression	0.3114

Não existem diferenças significativas entre os valores estimados no **Modelo 6** e o **Modelo 9** o que parece indicar que a forma de heteroscedasticidade está bem especificada.

Outra solução pode ser o cálculo de erros padrão robustos e os correspondentes *t-ratio* robustos. Estes erros padrão são validos haja ou não heteroscedasticidade. Optou-se pela estimação do modelo com a correção de White, pelo método do HCCME (*Heteroscedasticity-Corrected Covariance Matrices*) - HC1.

**Tabela 4.29**  
**Modelo 10**

OLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
l_ENS_SUPERIOR	0.2120	0.0028	76.5458	<0.00001	***
l_AREA_D	0.5850	0.0101	57.7246	<0.00001	***
PC1_N_I_POP_EDIF	-0.0897	0.0024	-37.9791	<0.00001	***
D1	7.5667	0.0488	155.0807	<0.00001	***
D2	7.4932	0.0489	153.2965	<0.00001	***
D3	7.4777	0.0518	144.4315	<0.00001	***
D4	7.4712	0.0542	137.8225	<0.00001	***
D5	7.6241	0.0575	132.5208	<0.00001	***
D6	7.8250	0.0755	103.6383	<0.00001	***
D1_IDADE	-0.1377	0.0086	-15.9656	<0.00001	***
D2_IDADE	-0.1432	0.0033	-43.8036	<0.00001	***
D3_IDADE	-0.1512	0.0019	-78.3521	<0.00001	***
D4_IDADE	-0.1504	0.0021	-70.7988	<0.00001	***
D5_IDADE	-0.1740	0.0054	-32.0637	<0.00001	***
D6_IDADE	-0.2215	0.0186	-11.8795	<0.00001	***

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	3277.5350	S.E. of regression	0.3114
R-squared	0.5916	Adjusted R-squared	0.5914
F(14, 33801)	3228975	P-value(F)	0.0000
Log-likelihood	-8522.2190	Akaike criterion	17074.4400
Schwarz criterion	17200.8700	Hannan-Quinn	17114.7700

Verifica-se que se mantem a significância estatística para todos os parâmetros e que os erros padrão robustos são muito semelhantes aos do **Modelo 6**.

### j) Testes aos parâmetros do Modelo 10

Todos os *t-ratio* estão na região de rejeição da hipótese nula para o caso de um teste bilateral com nível de significância de 0,01. Conclui-se que todos os estimadores são estatisticamente significantes, sendo o mais significativo o estimador da variável *dummy* “D1” onde a tipologia é 1 e em contraposição o menos significativo é o estimador da variável de interação “D6\_IDADE”.

Os *p-value* são todos quase nulos pelo que a probabilidade de se cometer um erro de 1ª espécie ou seja rejeitar a hipótese nula quando esta é verdadeira para qualquer nível de significância é nulo.

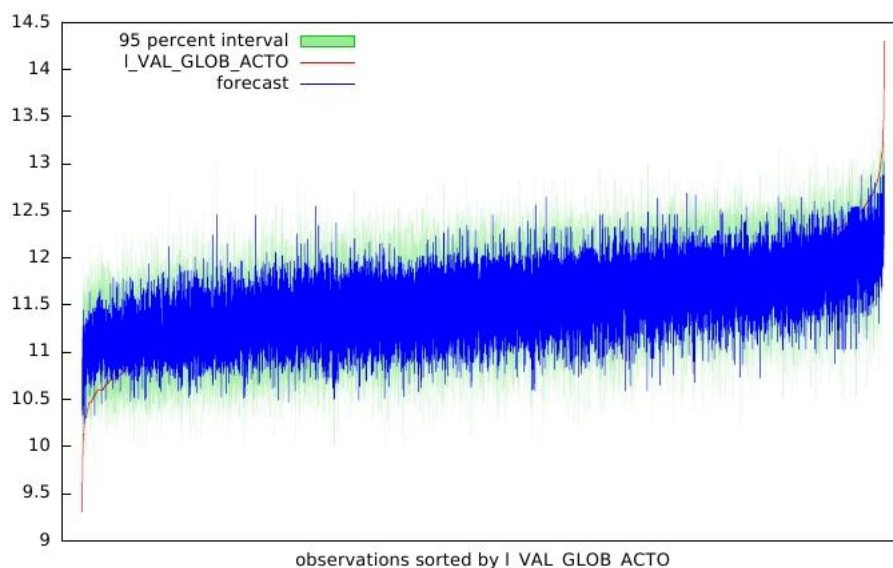
Igualmente se confirma que em 95% dos casos os Intervalos de Confiança (IC) construídos desta maneira contêm os respectivos estimadores.

**Tabela 4.30**  
**t(33801, 0.025) = 1.960**

Variable	Coefficient	95 confidence interval
l_ENS_SUPERIOR	0.211975	(0.206547, 0.217402)
l_AREA_D	0.584973	(0.565111, 0.604836)
PC1_N_I_POP_EDIF	-0.0896470	(-0.0942735, -0.0850205)
D1	7.56672	(7.47108, 7.66235)
D2	7.49315	(7.39734, 7.58896)
D3	7.47769	(7.37622, 7.57917)
D4	7.47116	(7.36491, 7.57741)
D5	7.62411	(7.51135, 7.73688)
D6	7.82500	(7.67702, 7.97299)
D1_IDADE	-0.137667	(-0.154568, -0.120766)
D2_IDADE	-0.143216	(-0.149624, -0.136808)
D3_IDADE	-0.151193	(-0.154975, -0.147410)
D4_IDADE	-0.150355	(-0.154518, -0.146193)
D5_IDADE	-0.173951	(-0.184585, -0.163318)
D6_IDADE	-0.221544	(-0.258097, -0.184991)

Os valores observados (linha vermelha) e os valores estimados (mancha azul) encontram-se dentro do intervalo de confiança a 95% (mancha verde).

**Figura 4.5**  
**Análise preditiva**



### **k) Significância económica dos estimadores**

Conjuntamente os estimadores explicam 59,16% da variabilidade da variável dependente de acordo com a estatística  $R^2$ .

O sinal negativo do coeficiente associado à variável “PC1\_N\_1\_POP\_EDIF” indica-nos que o aumento da componente que associa a população na freguesia e a concentração imobiliária, mantendo todas as outras variáveis constantes, provoca uma diminuição no preço dos apartamentos. Considerando que esta componente resulta da soma ponderada de duas variáveis logaritimizadas o seu coeficiente representa a elasticidade compósita da variável dependente face às variações conjuntas das variáveis que compõem a componente. A um acréscimo de 1% nesta componente corresponde uma diminuição média de 0,09% no preço dos

apartamentos, mantendo as restantes variáveis o mesmo valor. A variável “ENS\_SUPERIOR”, também é considerada no modelo na forma logarítmica número de indivíduos com ensino superior na freguesia, que pretende ser uma *proxi* do *status* social da localização contribui positivamente para a valorização imobiliária. Mantendo inalteradas todas as variáveis, um acréscimo de 1% no número de indivíduos residentes na freguesia com ensino superior resultam numa variação no preço médio de 0,21%.

A variável “l\_AREA\_D” apresenta um coeficiente com o sinal positivo tal como seria expectável, indicando que quanto maior a área do apartamento maior o seu preço. Igualmente esta variável foi logaritmizada e o seu coeficiente representa uma elasticidade. A um aumento de 1% na área de um apartamento corresponde um acréscimo de preço de 0.59%. O elevado peso económico revelado por este coeficiente confirma a teoria estudada.

Os coeficientes das variáveis de interação resultado do cruzamento entre a tipologia e o logaritmo da idade do imóvel representam a elasticidade da variável idade para cada uma das tipologias. Esta elasticidade é sempre crescente com o aumento de uma assoalhada exceto quando há um acréscimo de 3 para 4 assoalhadas. Estes coeficientes são negativos para todos os casos indicando que quanto mais antigo for o apartamento menor será o seu preço. Em termos gerais corresponde ao revelado pela revisão bibliográfica e por isso uma tendência económica aceite.

## **5. Conclusão**

A área é a variável com maior poder explicativo para a formação do preço, mas não é a única. Tal como [Reichert \(1990\)](#) também o presente estudo o comprovou: outras variáveis também contribuem para a formação do preço dos imóveis nomeadamente aquelas que representam fatores de localização e de *status* social.

[Bourassa et al. \(2006\)](#) comparou as três abordagens e concluiu que “... *as vantagens e os relativamente limitados inconvenientes do método SPAR torna-o o candidato ideal para ser usado pelas agências governamentais no desenvolvimento dos índices de preços da habitação.*” Estas avaliações estão dependentes das condições no mercado de financiamento e nesse sentido o valor determinado por estes processos não é um bom indicador dos preços de mercado. O objeto de análise quando se estuda a evolução de um índice de preços ou quando se pretende determinar as características que o comprador mais valoriza são os preços efetivamente praticados no mercado.

O método dos preços hedônicos é, desde o trabalho de [Rosen \(1974\)](#), uma metodologia recorrentemente utilizada e reconhecida como uma das mais indicadas para analisar a evolução de preços de produtos onde as diversas características podem exercer um efeito de qualidade que influencia o preço. O modelo hedônico apresenta-se como sendo o de melhor especificação tendo em conta os resultados da função da máxima verossimilhança e dos testes AIC, BIC e HQC. Com a introdução de *dummies* elimina-se o problema da colinearidade entre “TIPOLO” e “AREA\_D”, mantendo ambas as variáveis no modelo. Conclui-se que com o modelo obtido com a regressão OLS contém heteroscedasticidade. O problema da heteroscedasticidade é ultrapassado pela regressão com erros padrão robustos.

A tipologia está associada a uma medida de dimensão do apartamento, pelo que apresenta uma elevada correlação com a área. No entanto, esta correlação não é perfeita e muitas vezes estas variáveis não evoluem no mesmo sentido. Assim, ambas têm significância conjunta elevada e devem manter-se na regressão. A tipologia é uma característica que contribui de forma significativa para a formação do preço, ainda que nem sempre possa ser considerada como medida do tamanho do imóvel.

O número de residentes (POP\_2011) e o número de edifícios (EDIFICIOS), são variáveis utilizadas neste estudo como *proxis* de localização que se pressupõe indicarem a proximidade a serviços públicos e privados, comércio, etc. Conclui-se que existe uma preferência por zonas com uma menor concentração populacional. As pessoas não querem ter um hospital ou uma escola no outro lado da rua. A proximidade a centros urbanos é desejada, mas tem uma “distância ótima” ou seja é desejada até um certo nível.

[Arraes et al. \(2008\)](#) “A concentração de unidades, quando excessiva, reduz o bem-estar por externalidades negativas, atuando negativamente no preço do imóvel.”.

O modelo recomendado pelo presente estudo é o seguinte:

$$\begin{aligned} \widehat{L\_VAL\_GL\ddot{O}B\_ACTO} = & 0.211975 \text{ I\_ENS\_SUPERIOR} + 0.584973 \text{ I\_AREA\_D} - 0.0896470 \text{ PC1\_N\_1\_POP\_EDIF} \\ & (0.0027693) \qquad (0.010134) \qquad (0.0023604) \\ & + 7.56672 \text{ D1} + 7.49315 \text{ D2} + 7.47769 \text{ D3} + 7.47116 \text{ D4} \\ & (0.048792) \qquad (0.048880) \qquad (0.051773) \qquad (0.054209) \\ & + 7.62411 \text{ D5} + 7.82500 \text{ D6} - 0.137667 \text{ D1\_IDADE} - 0.143216 \text{ D2\_IDADE} \\ & (0.057531) \qquad (0.075503) \qquad (0.0086227) \qquad (0.0032695) \\ & - 0.151193 \text{ D3\_IDADE} - 0.150355 \text{ D4\_IDADE} - 0.173951 \text{ D5\_IDADE} - 0.221544 \text{ D6\_ID} \\ & (0.0019297) \qquad (0.0021237) \qquad (0.0054252) \qquad (0.018649) \end{aligned}$$

$$T = 33816 \quad \bar{R}^2 = 0.5914 \quad F(14, 33801) = 3.2290e+006 \quad \hat{\sigma} = 0.31139$$

(standard errors in parentheses)

**Tabela 5.1**  
**Descrição das variáveis do Modelo 10**

VARIÁVEL	DESCRIÇÃO
I_ENS_SUPERIOR	Logaritmo do número de indivíduos com o ensino superior em Portugal por freguesia
I_AREA_D	Logaritmo da Área Bruta Privativa da Fração
PC1_N_1_POP_EDIF	Componente principal 1 resultante da análise às componentes principais do conjunto de variáveis normalizadas e logaritmizadas “POP_2011” e “EDIFICIOS”
D1	Variável <i>dummy</i> para TIPOLOGIA=1 e restantes = 0

D2	Variável <i>dummy</i> para TIPOLOGIA=2 e restantes = 0
D3	Variável <i>dummy</i> para TIPOLOGIA=3 e restantes = 0
D4	Variável <i>dummy</i> para TIPOLOGIA=4 e restantes = 0
D5	Variável <i>dummy</i> para TIPOLOGIA=5 e restantes = 0
D6	Variável <i>dummy</i> para TIPOLOGIA=6 e restantes = 0
D1_IDADE	Interação da <i>dummy</i> D1 com a variável IMI_IDADEP (Idade do Imóvel)
D2_IDADE	Interação da <i>dummy</i> D2 com a variável IMI_IDADEP (Idade do Imóvel)
D3_IDADE	Interação da <i>dummy</i> D3 com a variável IMI_IDADEP (Idade do Imóvel)
D4_IDADE	Interação da <i>dummy</i> D4 com a variável IMI_IDADEP (Idade do Imóvel)
D5_IDADE	Interação da <i>dummy</i> D5 com a variável IMI_IDADEP (Idade do Imóvel)
D6_IDADE	Interação da <i>dummy</i> D6 com a variável IMI_IDADEP (Idade do Imóvel)

## 6. Desenvolvimentos futuros

Para fins de política económica é importante saber o quanto as pessoas valorizam os bens ambientais, a cultura, a educação, a proximidade a centros comerciais ou os investimentos públicos e em que áreas. As variáveis determinantes da capitalização dos imóveis podem medir a propensão para pagar por certos bens públicos e são muitas vezes o reflexo de fatores que se prendem com a competitividade intermunicipal. Por exemplo, uma cidade mais atrativa terá indicadores de capitalização imobiliária superiores a uma cidade menos atrativa. Quais são os fatores determinantes da atratividade das cidades?

A valorização dos imóveis pode medir a propensão para pagar pela proximidade a bens públicos, materiais ou imateriais, sendo o contrário também

uma realidade. O presente estudo conclui que existe uma “distância ótima” que seria interessante calcular ainda que esta possa variar consoante o bem em causa. Por exemplo, a “distância ótima” a um hospital poderá ser maior do que a um espaço de lazer.

Este estudo foi realizado para apartamentos em 2009, mas admite-se que as conclusões poderiam ser diferentes uma vez alargada a amostra a outras tipologias como por exemplo moradias, ou a outros anos.

O modelo poderá vir a ser melhorado com a incorporação de mais variáveis de localização e de caracterização do imóvel.

## 7. Bibliografia de base:

- Adams, Z. and R. Füss (2010). "Macroeconomic determinants of international housing markets." Journal of Housing Economics **19**(1): 38-50.
- Aoki, K., et al. (2002). House Prices, Consumption, and Monetary Policy: A Financial Accelerator Approach, Bank of England: 38.
- Arraes, R. A. and E. d. Sousa Filho (2008). "Externalidades e formação de preços no mercado imobiliário urbano brasileiro: um estudo de caso." Economia Aplicada **12**: 289-319.
- Attanasio, O., et al. (2005). Consumption, House Prices and Expectations. Leicester, Bank of England.
- Balk, B. M., et al. (2011). Handbook on Residential Property Price Indices. EUROSTAT.
- Bickerdike, C. F. (1902). "Taxes on Land and property." Economic Journal **XII**(48): 12.
- Bourassa, S. C., et al. (2008). Predicting House Prices with Spatial Dependence: Impacts of Alternative Submarket Definitions. Research Paper Series. Geneva, Swiss Finance Institute 26.
- Bourassa, S. C., et al. (2006). "A simple alternative house price index method." Journal of Housing Economics **15**(1): 80-97.
- Bruch, E. E. and R. D. Mare (2011). Methodological Issues in the Analysis of Residential Preferences & Mobility Population Studies Center Universidade de Michigan.
- Campbell, J. Y. and J. F. Cocco (2007). "How Do House Prices Affect Consumption " Journal of Monetary Economics **3**(54): 29.
- Dantas, R. A., et al. (2007). "Avaliação de imóveis: a importância dos vizinhos no caso de Recife." Economia Aplicada **11**: 231-251.
- Deaton, B. J. and J. P. Hoehn (2004). "Hedonic analysis of hazardous waste sites in the presence of other urban disamenities." Environmental Science & Policy **7**(6): 499-508.
- Dorsey, R. E., et al. (2010). "Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle." Journal of Housing Economics **19**(2): 75-93.
- Downes, Thomas A. and Jeffrey E. Zabel (2002). "The impact of school characteristics on house prices: Chicago 1987-1991." Journal of Urban Economics **52**(1): 1-25.
- Égert, B. and D. Mihaljek (2007). Determinants of House Prices in Central and Eastern Europe. William Davidson Institute, University of Michigan.

- Elbourne, A. (2008). "The UK housing market and the monetary policy transmission mechanism: An SVAR approach." Journal of Housing Economics **17**(1): 65-87.
- Elliott, M. and Partners (2001). Study on the Valuation and Restoration of Biodiversity Damage for the Purpose of Environmental Liability, EUROPEAN COMMISSION  
DIRECTORATE-GENERAL ENVIRONMENT.
- Engelhardt, G. V. (1996). "House prices and home owner saving behavior." Regional Science and Urban Economics **26**(3-4): 313-336.
- EUROSTAT (2010). Technical manual on owner-occupied housing.
- Farber, S. (1998). "Undesirable facilities and property values: a summary of empirical studies." Ecological Economics **24**(1): 1-14.
- Fischel, W. A. (2001). The Homevoter hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land Use Policies.
- Fischel, W. A., Ed. (2001). The Homevoter hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land Use Policies.
- Franklin, J. P. and P. Waddell (2002). A Hedonic Regression of Home Prices in King County, Washington, using Activity-Specific Accessibility Measures TRB 2003 Annual Meeting.
- Gerkman, L. (2010). Topics in Spatial Econometrics - With Applications to House Prices. Economics and Society. Helsingfors, Publications of the Hanken School of Economics: 33.
- Gerkman, L. (2010). Topics in Spatial Econometrics - With Applications to House Prices. . Publications of the Hanken School of Economics. Helsingfors, Hanken School of Economics: 33.
- Girouard, N. and S. Blöndal (2001). House Prices and Economic Activity. O. E. D. W. Ppares.
- Goodman, A. C. and T. G. Thibodeau (1998). "Housing Market Segmentation." Journal of Housing Economics **7**(2): 121-143.
- Goodman, A. C. and T. G. Thibodeau (2003). "Housing market segmentation and hedonic prediction accuracy." Journal of Housing Economics **12**(3): 181-201.
- Greene, W. H., Ed. (2012). Econometric Analysis.
- Gustafson, C. R., et al. (2011). Consumer Characteristics, Identification, and Hedonic Valuation of Wine Attributes: Exploiting Data from a Field Experiment. . RMI-CWE, University of California Davis.
- Haan, J. d. (2004). Hedonic Regression: The Time Dummy Index As a Special Case of the Imputation Törnqvist Index Voorburg Statistics Netherlands 13.
- Hanley, N. and C. L. Spash (1993). Cost-Benefit Analysis and the Environment.

- Haurin, D. R. and D. Brasington (1996). "School Quality and Real House Prices: Inter- and Intrametropolitan Effects." Journal of Housing Economics **5**(4): 351-368.
- Iacoviello, M. (2004). Consumption, House Prices and Collateral Constraints: a Structural Econometric Analysis. Department of Economics. Boston, Boston College.
- Ichihara, K. and J. P. Cohen (2010). "New York City property values: what is the impact of green roofs on rental pricing?" Springer-Verlag.
- Jansen, S. J. T., et al., Eds. (2011). The Measurement and Analysis of Housing Preference and Choice. Science+Business Media Dordrecht Heidelberg London New York, Springer
- Kiel, K. A. and K. T. McClain (1995). "House Prices during Siting Decision Stages: the Case of an Incinerator from Rumor through Operation." Journal of Environmental Economics and Management **28**: 241-255.
- Lecocq, S. and M. Visser (2006). "What Determines Wine Prices: Objective vs. Sensory Characteristics." Journal of Wine Economics **1**(1): 42-56.
- Lima, T. (2006). "Price and Quality in the California Wine Industry: An Empirical Investigation." Journal of Wine Economics **1**(2).
- Ludwig, A. and T. Slok (2001). The Impact of Stock Prices and House Prices on Consumption in OECD Countries, International Monetary Fund: 37.
- Luttik, J. (2000). "The value of trees, water and open space as reflected by house prices in the Netherlands." Landscape and Urban Planning **48**(3-4): 161-167.
- Makaronidis, A. and K. Hayes (2006). Owner-Occupied Housing for the Harmonized Indices of Consumer Prices OECD-IMF WORKSHOP Paris, Price Statistics, Eurostat, Luxembourg2.
- Marola, B., et al. (2012). Working Towards a Comparable House Price Inflation Measure in Europe. Geneva, UNECE.
- Moulton, B. R. (2001). The Expanding Role of Hedonic Methods in the Official Statistics of the United States. Washington DC, Bureau of Economic Analysis.
- Nerlove, M. (1995). "Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers." European Economic Review(39).
- Nguyen-Hoang, P. and J. Yinger (2011). "The capitalization of school quality into house values: A review." Journal of Housing Economics **20**(1): 30-48.
- Oates, W. E. (1969). "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and The Tiebout Hypothesis." The Journal of Political Economy **77**(6).
- Paredes, D. J. C. (2011). "Methodology to compute regional housing price index." Springer-Verlag 2009 **46**.

- Pollakowski, H. O. (1973). "The Effects of Property Taxes and Local Public Spending on Property Values: A Comment and Further Results." Journal of Political Economy **81**(4): 994-1003.
- Ramalho, E. A. a. J. J. S. R. (2010). On the Econometrics of Hedonic Housing Price Indexes: A Survey. Évora, Universidade de Évora: 61.
- Reback, R. (2005). "House prices and the provision of local public services: capitalization under school choice programs." Journal of Urban Economics **57**(2): 275-301.
- Reichert, A. K. (1990). "The Impact of Interest Rates, Income, and Employment upon Regional Housing Prices." The Journal of Real Estate Finance and Economics **3**(4).
- Reichert, A. K. (1990). "The impact of interest rates, income, and employment upon regional housing prices." The Journal of Real Estate Finance and Economics **3**(4): 373-391.
- Reichert, A. K. (1997). Impact of a Toxic Waste Superfund Site on Property Values. Ohio, Cleveland State University.
- Reis, H. J. and J. M. C. Santos Silva (2006). "Hedonic prices indexes for new passenger cars in Portugal (1997-2001)." Economic Modelling **23**(6): 890-908.
- Ribeiro, C. S. (2013). Apontamentos de econometria. Lisboa, ISEG.
- Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." Journal of Political Economy **82**: 34-55.
- Santos, D. and A. M. Lopes (2007). "A crise do imobiliário nos EUA." Confidencial Imobiliário Nov-07: 4.
- Silver, M. and S. Heravi (2004). The Difference Between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes IMF Working Paper International Monetary Fund: 20.
- Stadelmann, D. (2010). "Which factors capitalize into house prices? A Bayesian averaging approach." Journal of Housing Economics **19**(3): 180-204.
- Tobin, G. A. and B. E. Montz (1988). "Catastrophic flooding and the response of the real estate market." The Social Science Journal **25**(2): 167-177.
- West, J. (2011). The Effect of Quality Differentials on Integration of the Seaborne Thermal Coal Market. Nathan,, Griffith University.
- Wooldridge, J. M., Ed. (2013). Introductory econometrics - A modern approach. Mason, Cengage Learning.

## 8. Anexos

### I. Anexo A: Teste F ao Modelo 1

Realizou-se o teste F a fim de confirmar a relevância estatística das variáveis explicativas do **Modelo 1** para as variáveis cujo estimador apresenta um menor *t-ratio*: Tipologia; Edifícios.

Test for omission of variables -  
Null hypothesis: parameters are zero for the variables  
  1\_EDIFICIOS  
  TIPOLO  
Test statistic:  $F(2, 33809) = 163.302$   
with p-value =  $P(F(2, 33809) > 163.302) = 2.62617e-071$

O teste F para a hipótese de rejeitar conjuntamente as variáveis “TIPOLO” e “1\_EDIFICIOS” foi largamente superior ao valor crítico (3,84) apesar da omissão das variáveis não ter melhorado nenhuma das estatísticas de estimação. Ainda o *p-value* indica que a probabilidade de rejeitar a hipótese nula (parâmetros iguais a zero para as variáveis selecionadas) quando esta é verdadeira é muito próxima de zero. Rejeitou-se a hipótese nula, ou seja conclui-se que conjuntamente estas variáveis têm significância estatística. O modelo na hipótese de não rejeição da hipótese nula seria o seguinte (**Tabela 8.1**):

**Tabela 8.1**  
**Modelo 11**

OLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	8.8049	0.0341635	257.7282	<0.00001	***
l_POP_2011	-0.180613	0.00475312	-37.9987	<0.00001	***
l_ENS_SUPERIOR	0.26432	0.00376448	70.2143	<0.00001	***
l_AREA_D	0.585166	0.00515489	113.5169	<0.00001	***
l_IMI_IDADEP	-0.145269	0.00127046	-114.3441	<0.00001	***

Mean dependent var	11.54336		S.D. dependent var	0.487133
Sum squared resid	3350.756		S.E. of regression	0.314806
R-squared	0.582422		Adjusted R-squared	0.582372
F(4, 33811)	11789.56		P-value(F)	0.000000
Log-likelihood	-8895.790		Akaike criterion	17801.58
Schwarz criterion	17843.72		Hannan-Quinn	17815.03

## II. Anexo B: Teste F ao Modelo 3

O resultado do teste F à significância conjunta das componentes principais para o **Modelo 3** foi o seguinte:

Test for omission of variables -  
Null hypothesis: parameters are zero for the variables  
PC1\_N  
PC2\_N  
Test statistic:  $F(2, 33809) = 6441.18$   
with p-value =  $P(F(2, 33809) > 6441.18) = 0$

Deste teste resulta que as componentes têm conjuntamente significância estatística. Como se pode verificar pelo modelo estimado sem ambas as componentes (**Tabela 8.2**):

**Tabela 8.2**  
**Modelo 12**

OLS, using observations 1-33816  
Dependent variable: l\_VAL\_GLOB\_ACTO  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	11.8657	0.0260	456.4774	<0.00001	***
l_POP_2011	-0.1599	0.0067	-23.9263	<0.00001	***
l_EDIFICIOS	-0.0707	0.0039	-18.1324	<0.00001	***
l_ENS_SUPERIOR	0.2707	0.0046	59.3734	<0.00001	***
l_IMI_IDADEP	-0.1986	0.0014	-140.0939	<0.00001	***

Mean dependent var	11.5434	S.D. dependent var	0.4871
Sum squared resid	4583.2310	S.E. of regression	0.3682
R-squared	0.4288	Adjusted R-squared	0.4288
F(4, 33811)	6346.2090	P-value(F)	0.0000
Log-likelihood	-14191.6800	Akaike criterion	28393.3600
Schwarz criterion	28435.5100	Hannan-Quinn	28406.8100

### III. Anexo C: Modelo por Tipologia

**Tabela 8.3**  
**Tipologia 1**

OLS, using observations 1-973  
Dependent variable: l\_Preço\_medio  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	7.21709	0.159766	45.1729	<0.00001	***
l_POP_2011	-0.189995	0.0455431	-4.1718	0.00003	***
l_EDIFICIOS	0.100394	0.0250571	4.0066	0.00007	***
l_ENS_SUPERIOR	0.172292	0.030355	5.6759	<0.00001	***
l_IMI_IDADEP	-0.120298	0.0091467	-13.1521	<0.00001	***

Mean dependent var	7.353718	S.D. dependent var	0.460190
Sum squared resid	162.1383	S.E. of regression	0.409266
R-squared	0.212327	Adjusted R-squared	0.209072
F(4, 968)	65.23420	P-value(F)	7.04e-49
Log-likelihood	-508.8511	Akaike criterion	1027.702
Schwarz criterion	1052.104	Hannan-Quinn	1036.989

**Tabela 8.4**  
**Tipologia 2**

OLS, using observations 1-5130  
Dependent variable: Preço\_medio  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	7.31668	0.0661016	110.6885	<0.00001	***
l_POP_2011	-0.157268	0.0166706	-9.4338	<0.00001	***
l_EDIFICIOS	-0.0126344	0.00916946	-1.3779	0.16830	
l_ENS_SUPERIOR	0.211573	0.0113685	18.6104	<0.00001	***
l_IMI_IDADEP	-0.109978	0.0034426	-31.9463	<0.00001	***

Mean dependent var	7.147884		S.D. dependent var	0.409425
Sum squared resid	656.9190		S.E. of regression	0.358021
R-squared	0.235933		Adjusted R-squared	0.235337
F(4, 5125)	395.6324		P-value(F)	2.0e-297
Log-likelihood	-2007.310		Akaike criterion	4024.619
Schwarz criterion	4057.333		Hannan-Quinn	4036.070

**Tabela 8.5**  
**Tipologia 3**

OLS, using observations 1-13611  
Dependent variable: Preço\_medio  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	6.98961	0.0351916	198.6157	<0.00001	***
l_POP_2011	-0.118841	0.00878406	-13.5292	<0.00001	***
l_EDIFICIOS	-0.061063	0.00515617	-11.8427	<0.00001	***
l_ENS_SUPERIOR	0.235881	0.00615863	38.3008	<0.00001	***
l_IMI_IDADEP	-0.117594	0.00189211	-62.1497	<0.00001	***

Mean dependent var	6.973868		S.D. dependent var	0.376853
Sum squared resid	1298.604		S.E. of regression	0.308939
R-squared	0.328148		Adjusted R-squared	0.327950
F(4, 13606)	1661.368		P-value(F)	0.000000
Log-likelihood	-3323.049		Akaike criterion	6656.098
Schwarz criterion	6693.691		Hannan-Quinn	6668.630

**Tabela 8.6**  
**Tipologia 4**

OLS, using observations 1-11953

Dependent variable: Preço\_medio

Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	6.4128	0.0356372	179.9467	<0.00001	***
l_POP_2011	0.00244314	0.00939185	0.2601	0.79476	
l_EDIFICIOS	-0.112522	0.00559462	-20.1126	<0.00001	***
l_ENS_SUPERIOR	0.198285	0.00631719	31.3881	<0.00001	***
l_IMI_IDADEP	-0.12271	0.00200078	-61.3312	<0.00001	***

Mean dependent var	6.872882		S.D. dependent var	0.385971
Sum squared resid	1083.743		S.E. of regression	0.301173
R-squared	0.391338		Adjusted R-squared	0.391134
F(4, 11948)	1920.486		P-value(F)	0.000000
Log-likelihood	-2613.617		Akaike criterion	5237.234
Schwarz criterion	5274.178		Hannan-Quinn	5249.629

**Tabela 8.7**  
**Tipologia 5**

OLS, using observations 1-1950

Dependent variable: Preço\_medio

Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	6.51843	0.120491	54.0991	<0.00001	***
l_POP_2011	-0.102385	0.0309289	-3.3103	0.00095	***
l_EDIFICIOS	-0.101846	0.0176214	-5.7797	<0.00001	***
l_ENS_SUPERIOR	0.281266	0.0200343	14.0392	<0.00001	***
l_IMI_IDADEP	0.0131312	0.00689931	1.9033	0.05715	*

Mean dependent var	6.946022		S.D. dependent var	0.446168
Sum squared resid	289.7272		S.E. of regression	0.385953
R-squared	0.253241		Adjusted R-squared	0.251706
F(4, 1945)	164.8975		P-value(F)	1.2e-121
Log-likelihood	-907.9514		Akaike criterion	1825.903
Schwarz criterion	1853.781		Hannan-Quinn	1836.152

**Tabela 8.8**  
**Tipologia 6**

OLS, using observations 1-199  
Dependent variable: Preço\_medio  
Heteroskedasticity-robust standard errors, variant HC1

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	5.80994	0.486228	11.9490	<0.00001	***
I_POP_2011	0.100101	0.124038	0.8070	0.42064	
I_EDIFICIOS	-0.134043	0.0684198	-1.9591	0.05153	*
I_ENS_SUPERIOR	0.160132	0.0719054	2.2270	0.02710	**
I_IMI_IDADEP	-0.0219824	0.0300927	-0.7305	0.46597	

Mean dependent var	6.939174		S.D. dependent var	0.529529
Sum squared resid	45.06204		S.E. of regression	0.481953
R-squared	0.188354		Adjusted R-squared	0.171620
F(4, 194)	11.25515		P-value(F)	3.11e-08
Log-likelihood	-134.5849		Akaike criterion	279.1699
Schwarz criterion	295.6364		Hannan-Quinn	285.8343