

Avaliação Hedónica da ocupação urbana residencial: uma análise empírica aplicada a um centro urbano

João Lourenço Marques

Eduardo Anselmo de Castro

Centro de Estudos em Governação e Políticas Públicas (CEGOPP); Secção Autónoma de Ciências Sociais,
Jurídicas e Políticas; Universidade de Aveiro; Tel.: 234 370200; E-mail: j.marques@csjp.ua.pt

Resumo

O trabalho que se propõe apresentar é uma análise exploratória que tem como objectivo central estudar os padrões de localização residencial em áreas urbanas. Para tal aplica-se uma metodologia de preços hedónicos que explique o valor patrimonial das habitações em Portugal. Este estudo combina métodos de análise de dados convencionais com tecnologias de informação geográfica. Uma das maiores dificuldades inerentes à metodologia aplicada é o acesso a uma base de informação que permita avaliar não só os preços das habitações como também as suas características específicas. O presente artigo incide na área urbana e peri-urbana do concelho de Aveiro no qual fazem parte da amostra 175 imóveis que estão para transacção. Para a referida área urbana são utilizados dois tipos de informação: i) Informação específica aos imóveis para transacção em 2007, nomeadamente, elementos físicos e de conforto da habitação (tipologia, áreas, idade, garagem, etc., etc.) e ii) Informações de localização habitacional geradas a partir da utilização de um sistema de informações geográfica (SIG). A combinação desta informação, permite compreender quais, e em que medida, os atributos da habitação e da sua envolvente explicam o valor da habitação. Tendo consciência de que o processo hedónico representa preferência de consumo com elevado nível de subjectividade, esta análise, não deixa de ser determinante na actividade de planeamento urbano, como instrumento de apoio à decisão que ajuda a avaliar a importância dos vários elementos que compõem a estrutura e a dinâmica urbana.

1. Introdução

Relevância do estudo

O estudo da habitação assume-se como um tópico de análise muito importante nas economias modernas, quer à micro quer à macro escala. Numa análise micro, ao nível do indivíduo e das famílias, a habitação afigura-se como o mais valioso recurso da maioria da população. A parte dos rendimentos gastos na habitação representa uma fracção muito grande de despesa total. Estimativas do Banco de Portugal, das despesas em habitação, dão conta de um acréscimo, entre 1990 e 1999, de 14 a 47 por cento do rendimento mensal das famílias. Numa análise macro, num contexto mais alargado, o mercado da habitação tem sido considerado um factor essencial da economia nacional. De acordo com dados do Instituto Nacional de Estatística (INE), uma média de 10 por cento de Produto Interno Bruto em Portugal é gasto na construção de novas habitações. Por outro lado o estudo da habitação tem um papel importante no ordenamento do território dado que a relação entre a procura e a oferta estabelece regras e padrões de localização que determinam irreversivelmente, por vezes de forma prejudicial, a organização do espaço.

O sobrepovoamento (quantidade) e a falta de alojamentos adequados (qualidade) são fenómenos amplamente observados e importantes do problema urbano. Dados do INE indicam que as duas últimas décadas foram períodos de grande intensidade de crescimento do parque habitacional em Portugal, especialmente no número de alojamentos para uso sazonal. Estes obtiveram um aumento, entre 1991 em 2001, de 9,1% para 18,4%, no total de alojamentos. Ainda dados do INE, referem que o rácio alojamentos¹/famílias em Portugal atinge o segundo valor mais elevado da Europa (1,38), imediatamente a seguir à Espanha.

Apesar dos números indicarem uma dinâmica de construção acentuada, perspectivam-se mudanças no perfil de produção habitacional em Portugal nos próximos anos, resultado de um declínio recente na construção de novas habitações, onde a o desafio passará pela preocupação crescente ao nível do conforto da habitação. (INE; 2003)

Neste sentido é desejável que os decisores políticos, os planeadores do território, os economistas, os geógrafos e os arquitectos compreendam a estrutura da procura do mercado

¹ O total de alojamentos engloba a residência sazonal e a residência habitual.

da habitação de modo a intervirem sobre ele de forma compreensível e racional. O melhor entendimento do mercado da ocupação residencial no território passará pelo conhecimento de 3 vertentes:

- de como o **mercado** responde à procura de casa e aloca os recursos (financeiros, materiais e humanos) no sector da construção;
- de como os **agentes do planeamento** estimulam ou constroem a actividade do mercado imobiliário;
- de como os **proprietários** estão dispostos a vender as habitações;
- de como as **famílias** reagem à oferta da habitação determinando a procura de habitação residencial.

Como se pode observar, o sector da habitação nas suas múltiplas incidências tem impactos consideráveis que importam analisar: por um lado ao nível do território e da organização do espaço vivencial; por outro lado a economia - habitação como factor de desenvolvimento económico e social; e por fim as famílias ao nível da sua qualidade e bem-estar e pelo investimento que acarreta no orçamento familiar.

Melhorar a oferta da habitação, quer qualitativa quer quantitativamente, implica levar em linha de conta cada um dos aspectos enumerados, principalmente quando há indícios de fortes incertezas e especulação no mercado da habitação.

O artigo apresentado visa dar um contributo na compreensão do mercado da habitação, nomeadamente, ao nível dos factores que determinam a procura de casa, e que se reflecte na vontade de o futuro proprietário pagar um conjunto de atributos implícitos no preço da habitação. O objectivo é então avaliar a contribuição de cada atributo residencial ou de localização na valorização do preço de uma habitação.

Combinando os resultados analíticos com a visualização espacial desses resultados num sistema de informação geográfica é possível obter informação relevante para se perceberem as dinâmicas e tendências do mercado imobiliário.

Refira-se que os métodos apresentados não devem ser considerados como uma ferramenta de avaliação do mercado imobiliário mas apenas um caminho para melhor compreender as dinâmicas urbanas com respeito à formação de preços na habitação.

2. Método de preços hedónicos

2.1. Revisão bibliográfica

A análise hedónica de preços tem os seus princípios na teoria do consumo de Lancaster, no qual refere que a utilidade de um bem deriva das suas propriedade ou das suas características. (Lancaster; 1966) Posteriormente esta teoria foi estendida ao mercado residencial por Rosen tornando os preços hedónicos uma ferramenta de análise usada para a avaliação do mercado imobiliário e a análise urbana. (Rosen; 1974)

Estes modelos consideram que existe um conjunto de características habitacionais, quer físicas quer locacionais que podem explicar o valor do preço da habitação. Neste sentido o preço da habitação pode ser estimado por regressão a partir desse conjunto de atributos. A função pode ser expressa pela seguinte equação:

$$P(h) = F(A_H; A_E) + \varepsilon$$

onde:

$P(h)$ – preço da habitação

A_H – atributos físicos da habitação

A_E – atributos físicos da envolvente (amenidades ou desamenidades)

ε – valor residual / valor especulativo

A regressão expressa a melhor estimativa da variável dependente $P(h)$, dada pelas variáveis independentes (A_H , A_E). Contudo, poderá existir uma parte da variação que não é explicada e que pode ser por dois motivos:

- não foram consideradas todas as variáveis explicativas no modelo;
- uma parte pouco racional é importante na capacidade explicativa do modelo.

Existe um alargado debate acerca da melhor regressão a utilizar num estudo hedónico de preços, no entanto a bibliografia não determina uma forma funcional específica, e a regressão difere consoante o conjunto de dados e da realidade que se está a estudar. O modelo escolhido será aquele, que para o caso concreto, obtiver o maior valor de R^2 .

2.2. Limitações e premissas básicas:

Os factores determinantes considerados nos modelos explicativos do preço da habitação são particularmente diferentes de área para área urbana, dificultando a selecção dos atributos mais significativos para o modelo. Note-se que estes modelos têm a sua origem e aplicação predominante em países anglo-saxónicos, onde a realidade urbana é significativamente diferente da Portuguesa.

Não só os objectos de análise são diversos como também a aplicação de modelos hedónicos de preços é feita em diversas áreas/domínios dificultando também aqui a selecção dos atributos mais significativos para o modelo.

Existem estudos que analisam e caracterizam a habitação numa perspectiva integrada de modo a avaliar o preço da habitação em vários contextos espaciais. (A. S. Adair 1996; Bourassa, Hamelink et al. 1999; Gregory 2000; Jennifer and Radha 2005; Gang-Zhi, Seow Eng et al. 2006; Harald 2006; Sirmans, Lynn et al. 2006). Estes estudos mostram que existe uma parte substancial do preço da habitação que fica por explicar.

Outros procuram avaliar a importância das amenidades urbanas na escolha residencial, como sejam: espaços verdes; baixas densidades de ocupação; ruído, qualidade do ar, acessibilidade, eficiência energética etc. (Blomquist and Worley 1981; Dinan and Miranowski 1989; Donnelly 1991; So, Tse et al. 1997; Tyrvaïnen 1997; Earl, Julia et al. 1998; Luttik 2000; Bowes and Ihlanfeldt 2001; Smith, Poulos et al. 2002; Brasington and Hite 2005; Anderson and West 2006)

Outros ainda, procuram estudar o valor da habitação em contextos socio-económicos diferenciados e de segmentação racial das populações. (Largey 2002; Steve 2003; Wakefield and Inman 2003; Myers 2004; Wang and Li 2006; Hui, Chau et al. 2007)

No que respeita às restrições metodológicas dos modelos de preços hedónicos destacam-se três: a multicolinearidade, heterodasticidade e autocorrelação:

- i) A multicolinearidade é um problema comum em regressões, e acontece quando variáveis independentes têm fortes correlações entre si, o que leva, por um lado, ao difícil entendimento do contributo de cada variável no modelo, não se sabendo quais as

causas e quais os feitos, e por outro lado, leva a diminuição acentuada da eficiência do modelo. Um indício da existência de multicolinearidade é quando o R^2 é bastante alto, mas nenhum dos coeficientes da regressão é estatisticamente significativo, quando usado o teste t de Student². A Análise Factorial e mais concretamente o método das Componentes Principais permite contornar este problema na medida em que os factores resultantes desta análise são por natureza independentes entre si.

- ii) De acordo com os pressupostos da regressão os resíduos devem distribuir-se aleatoriamente em torno de zero mas pode acontecer que estes tenham uma mancha de largura não uniforme levando a um desajustamento do modelo. A heterodasticidade (dasticidade ou *scedastic* vem do Grego e que significa variância) é precisamente a violação da premissa básica dos métodos de estimativas dos mínimos quadrados, variância não constante.
- iii) O preço das habitações têm por vezes uma autocorrelação espacial, isto é, casas localizadas numa área próxima têm preços semelhantes³. Este facto não é uma total limitação do modelo, pois permite, caso tal aconteça, analisar os efeitos das externalidades territoriais. Com base em metodologias específicas, existem estudos que procuraram isolar tal efeito (Sabyasachi and Thomas 1998; Yong, Hua et al. 2007), no entanto, este ponto não é desenvolvido neste trabalho.

3. Modelo empírico de preços hedónicos

3.1. Recolha de dados e selecção das variáveis

Para a realização de um modelo de preços hedónicos é necessária informação de base respeitante às transacções imobiliárias residenciais por um período de tempo determinado e num espaço territorial bem delimitado. Na impossibilidade de se obter informação de imóveis já transaccionados, utilizou-se informação respeitante a imóveis que estão para

² Um teste para o coeficiente de regressão, com a hipótese nula H_0 , de que o coeficiente de regressão das variáveis independentes da população a que esta amostra pertence, é igual a zero.

³ i) a unidades de vizinhança têm características estruturais semelhantes (materiais de construção; áreas; anos de construção; garagens); ii) partilham serviços sociais, escolas, centros de saúde, etc.; iii) partilham distâncias a centros administrativos, económicos comerciais e semelhantes.

venda. Analisados os resultados verificou-se que não existem fortes enviesamentos nos dados recolhidos.

Foram seleccionadas para este estudo 6 das 14 freguesias do concelho de Aveiro, por se considerar serem as freguesias com maior dinamismo de crescimento habitacional (ver figura 2). A amostragem resultou de uma análise por quotas, relativas ao peso dos efectivos populacionais por freguesia, segundo os dados apresentados no quadro 1:

Quadro 1.- Distribuição da amostra por freguesias

Freguesia	População		Habitações	
Aradas	7628	15%	22	13%
Esgueira	12262	24%	48	27%
Glória	9917	19%	27	15%
Santa Joana	8652	17%	13	7%
São Bernardo	4079	8%	9	5%
Vera Cruz	8652	17%	56	32%
Total	51190	100%	175	100%



Figura 1.- Localização das freguesias estudadas do concelho de Aveiro

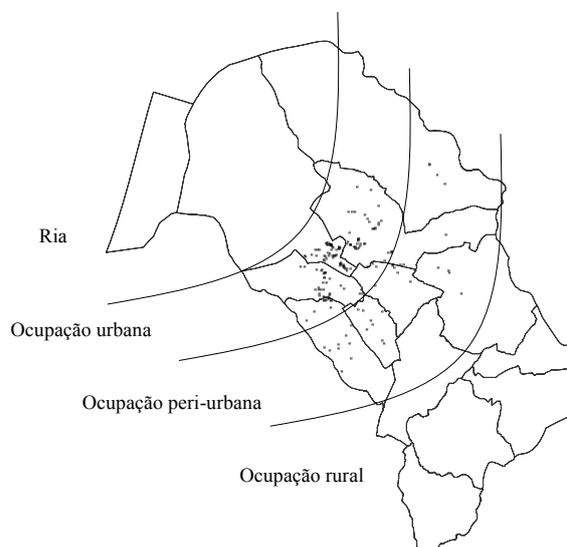
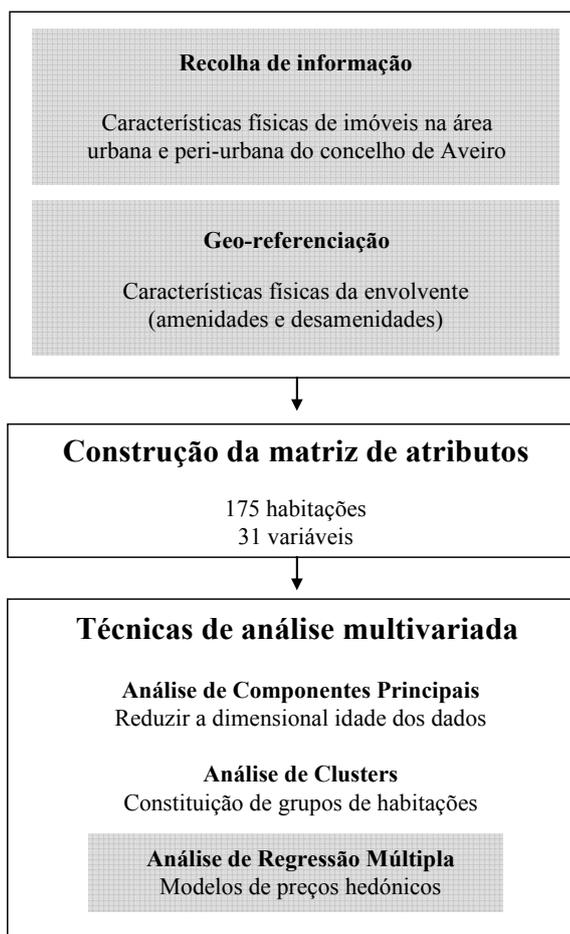


Figura 2.- Localização das habitações nas freguesias estudadas do concelho de Aveiro

O Modelo de preços hedónicos para os imóveis da área urbana e peri-urbana do concelho de Aveiro foi construído segundo a metodologia apresentada na figura seguinte:

Figura 3.- Esquema da metodologia adoptada



A recolha de informação é relativa a 175 imóveis a transaccionar na cidade de Aveiro. Para estes imóveis foram associadas um conjunto de características físicas da própria habitação bem como outros atributos de localização, construídos a partir de uma prévia geo-referenciação e utilizando o nome da rua como critério de localização. A posição de cada imóvel permitiu assim construir factores de localização que se consideram poder influir no valor do preço, nomeadamente a proximidade/distância a determinados equipamentos e serviços; e proximidade/distância algumas amenidades urbanas (*para uma análise mais*

detalhada dos atributos físicos da habitação e dos atributos físicos da envolvente veja-se o quadro 2).

Como facilmente se pode perceber, existem dois tipos de informação: por um lado, dados estatísticos que advêm da fonte primária (atributos físicos dos imóveis e elementos de qualidade e conforto) da agência imobiliária; e por outro lado, atributos construídos segundo critérios pré-estabelecidos de localização residencial (com recurso à base georeferenciada da informação primária).

A literatura existente sugere um conjunto de variáveis a serem incluídas nos modelos de preços hedónicos e que se podem classificar em 3 categorias básicas:

- Atributos estruturais
- Características de localização
- Características de vizinhança

Neste estudo não foram consideradas quaisquer variáveis referentes às características da vizinhança, pelo facto de que a informação estatística disponível dos censos 2001, de caracterização socio-económica de cada lugar ter uma desactualização considerável. Acresce também que em países onde a estrutura étnica e racial da população é mais evidente, a utilização deste variáveis é considerada imprescindível.

Outras variáveis independentes consideradas importantes não foram introduzidas no modelo. Se considerarmos que existem variáveis que ficaram fora do modelo e que são factores importantes na decisão final do consumidor, o modelo poderá ter um certo nível de enviesamento. Contudo, também é verdade que nenhum modelo pode captar todos os aspectos comportamentais do consumidor.

A recolha de dados resultou na selecção das seguintes variáveis:

Quadro 2.- Análise descritiva das variáveis

	Tipo de variável*	Casos válidos	Casos em falta	Média	Desvio Padrão	Mínimo	Máximo
<i>PHC – atributos físicos da habitação</i>							
Preco	E	175	0	139378,9	61136,9	39400	385000
TipoHabitacao	N	173	2	0,1	0,3	0	1
Tipologia	O	173	2	2,3	0,8	1	5
Duplex	N	170	5	0,2	0,4	0	1
AreaTotal	E	115	60	113,5	43,3	33,2	250,3
AreaCozinha	E	147	28	12,4	3,8	5,5	24,9
AreaQuartos	E	155	20	12,7	2,7	8,3	28,5
AreaSala	E	136	39	23,2	7,1	8,95	49,5
NumWC	E	149	26	1,7	0,8	1	6
NumPISOS	E	175	0	3,4	2,1	1	12
Piso	E	130	45	2,8	2,0	1	11
Varanda	N	175	0	0,2	0,4	0	1
Terraco	N	175	0	0,1	0,3	0	1
Garagem	N	175	0	0,6	0,5	0	1
NumElevadores	E	93	82	1,3	0,5	0	2
NumFrentes	E	154	21	2,0	0,7	1	4
NumAnos	E	92	83	10,8	10,9	0	67
TVCabo	N	175	0	0,3	0,4	0	1
Gas	N	175	0	0,4	0,5	0	1
Estado	N	174	1	0,1	0,3	0	1
<i>PAC – atributos físicos da envolvente (amenidades ou desamenidades)</i>							
Ruido	O	130	45	1,6	0,7	1	3
Bairro Social	N	175	0	0,0	0,2	0	1
Dist_Saúde	E	175	0	2423	946	1500	6514
Dist_Infantários	E	175	0	2627	707	1833	5840
Dist_Escolas	E	175	0	2043	1134	989	5809
Dist_Ens_Sup_	E	175	0	1978	1000	900	5938
Dist_Mercados	E	175	0	2114	1310	769	6620
Dist_Hipermercados	E	175	0	3567	568	2682	5715
Dist_Centro_Comercial	E	175	0	1904	1398	175	6260
Dist_Desporto	E	175	0	3662	607	2889	6679
Dist_Hotelaria	E	175	0	2412	1149	1178	6045
Dist_Cultura_e_Lazer	E	175	0	2054	1308	655	6491
Dist_Segurança	E	175	0	2042	1076	1073	6064
Dist_Centros_sociais_e	E	175	0	3898	620	2560	6084
Dist_Serviços_da_Adm	E	175	0	1858	1110	769	5637
Dist_Centro	E	175	0	2013	1419	261	6403

*E-Variável Escalar; O-Variável Ordinal; N-Variável Nominal

Redução da dimensionalidade dos dados

Neste ponto procura-se perceber o nível de associação das variáveis de modo construírem-se novas variáveis que traduzam a estrutura essencial dos dados iniciais. Para o efeito, utiliza-se uma análise factorial.

Utilizando os critérios de validação da matriz de atributos associados ao método das componentes principais (KMO; matriz anti-imagem, matriz de correlações, quadro das comunalidades) foram excluídas as seguintes variáveis: Tipo de habitação; Duplex; Varanda; Terraço; Garagem; Numero de anos; Estado; Ruído; Área quartos; Bairro Social.

As variáveis indicadas foram excluídas essencialmente por duas razões: por um lado, pela falta de dados na maioria das habitações e por outro lado pela pouca correlação que algumas dessas variáveis apresentavam com o conjunto das outras variáveis, o que obrigavam á retenções de factores isoladas para as explicarem.

Como resultado da análise factorial e utilizando o método de rotação *varimax*, obtiveram-se 5 factores/componentes que explicam aproximadamente 78% da variação inicial dos dados

Quadro 3.- Total Variância explicada por componente

Componentes	Valor próprio			Rotação Varimax		
	Total	% da Variância	% Cumulativa	Total	% da Variância	% Cumulativa
1	11,6	46,5	46,5	11,2	44,7	44,7
2	2,8	11,4	57,9	2,8	11,2	55,9
3	2,2	9,0	66,8	2,4	9,5	65,4
4	1,8	7,3	74,1	1,7	6,8	72,2
5	1,2	4,8	79,0	1,7	6,8	79,0

O quadro 4 representa os *loadings* de cada variável e de cada factor, expressando as correlações mais elevadas (acima de 0,5). É com base nestes valores que se interpretam e classificam cada factor.

Quadro 4.- Matriz dos loadings após rotação varimax

	1	2	3	4	5
Dist_Cultura_e_Lazer	0,99				
Dist_Hotelaria	0,99				
Dist_Centro_Comercial	0,99				
Dist_Escolas	0,99				
Dist_Segurança	0,99				
Dist_Centro	0,98				
Dist_Mercados	0,98				
Dist_Serviços_da_Administração	0,96				
Dist_Ens_Sup_	0,94				
Dist_Saúde	0,91				
Dist_Hipermercados	0,82				
Dist_Infantários	0,78		0,58		
NumWC		0,77			
AreaTotal		0,76			
Tipologia		0,68			
AreaSala		0,68			
NumFrentes		0,58			
AreaCozinha		0,54			
Dist_Centros_sociais_e_paroquiais			0,97		
Dist_Desporto	0,57		0,79		
Gas				0,82	
TVCabo				0,82	
Piso					0,77
NumPISOS					0,73
NumElevadores					0,65

Factor 1

O primeiro factor, responsável pela explicação da máxima variância dos dados iniciais (quase metade, 44,7%), está associado a variáveis de distâncias, com níveis de correlação bastante elevados, superiores a 0,8.

Analisando os resultados da figura 4, onde estão representados os *scores* padronizados (com média zero e desvio padrão igual a 1)⁴ do factor 1, agrupados por 4 classes, é evidente a representação radial das habitações com um epicentro amarrado ao centro da cidade de Aveiro. As habitações localizadas nesse epicentro com valores inferiores a 1,5 desvios padrões abaixo da média (distâncias menores) são aquelas com maiores acessibilidades aos equipamentos, comércio e serviços associados a este factor. À medida que as habitações se

⁴ Os valores positivos indicam quantos desvios padrões acima cada habitação tem relativamente ao valor médio, enquanto que os valores negativos indicam o número de desvios padrões abaixo que cada habitação está relativamente ao valor médio.

afastam desse centro, a distância média aos referidos equipamentos, comércios e serviços vai aumentando, o que leva à perda de centralidade.

Tal interpretação leva a caracterizar este factor como um indicador de acessibilidade aos equipamentos terciários (distância 1) explícitos nas respectivas variáveis (cultura e lazer, hotelaria; centros comerciais; mercados; serviços de administração; escolas de ensino superior; outras escolas; infantários; equipamentos de saúde; hipermercados).

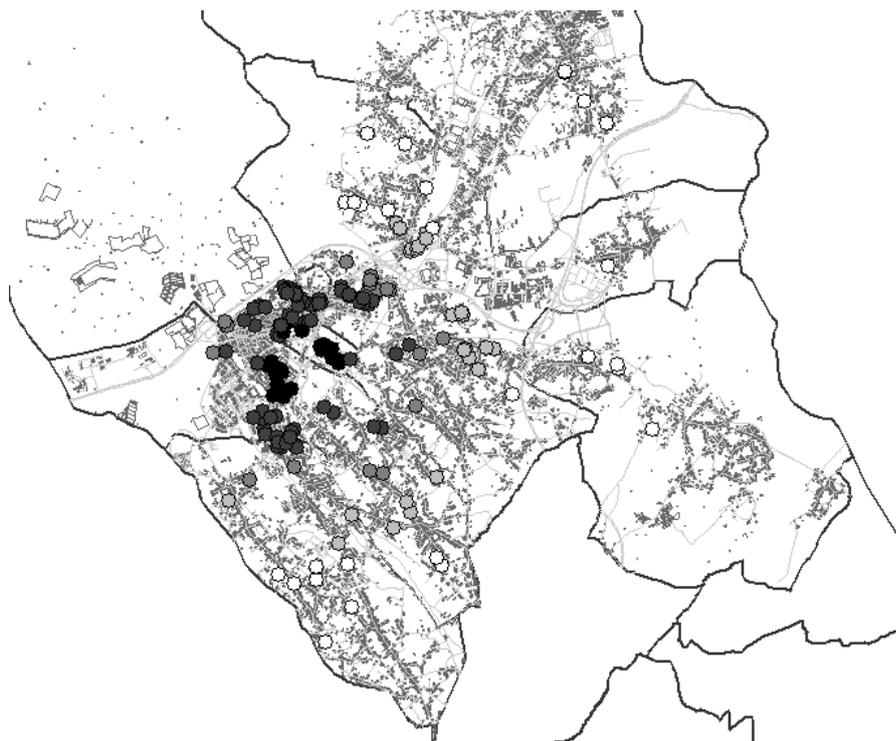


Figura 4.- Acessibilidade aos equipamentos terciários - *Distância 1 (factor 1)*

Factor 2

O segundo factor, com capacidade explicativa de 11% está associado a variáveis que caracterizam o aspecto dimensional da habitação. O tamanho da habitação está bem evidenciado nas variáveis números de casas de banho; área total e tipologia da habitação; área da sala; número de frentes e área da cozinha.

Numa análise geo-referenciada deste factor, e adoptando também 4 classes de intervalo dos valores dos *scores*, é possível observar alguma dispersão dos padrões dimensionais da habitação. Ao contrário do que sucedeu no anterior factor, em que havia uma distribuição

padrão dos imóveis, em virtude das variáveis em análise serem de cariz espacial, neste não é evidente uma padronização espacial.



Figura 5.- Dimensão habitacional (factor 2)

Factor 3

O factor 3 tem uma capacidade explicativa de 9,5% da variância inicial dos dados. Tal como aconteceu no factor 1, neste factor estão associadas variáveis que expressam distância a equipamentos desportivos, sociais e paroquiais (lares, centros sociais e paróquias, campos e pavilhões desportivos, piscinas etc.), no entanto distinguem-se dos anteriores por terem uma localização mais dispersa (distância 2). São equipamentos que não têm uma localização amarrada ao centro da cidade e todas as freguesias acabam por ter um.

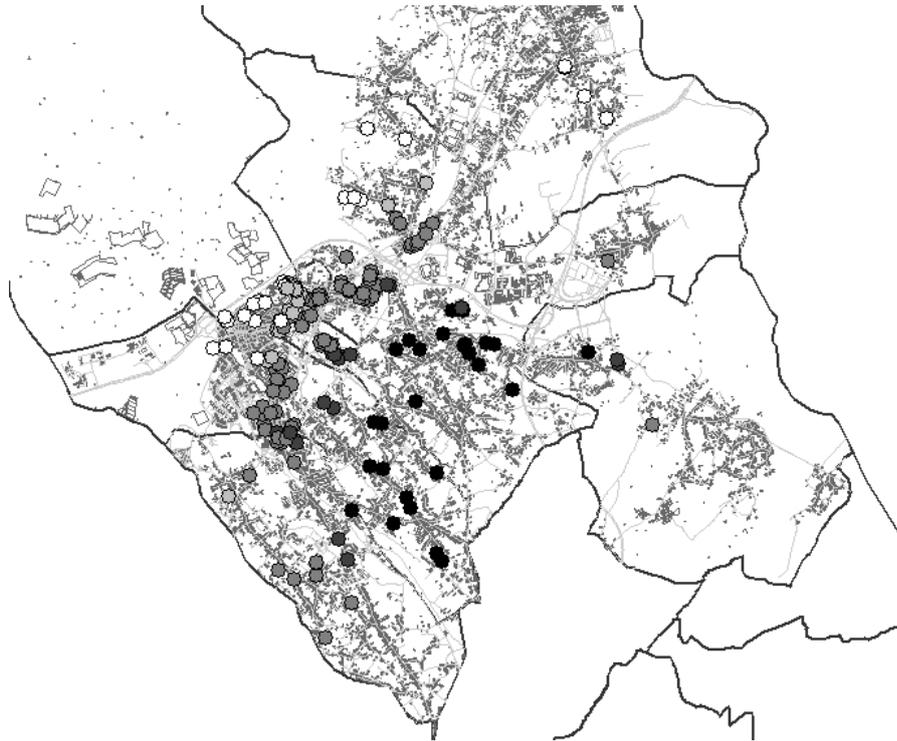


Figura 6.- Acessibilidade a equipamentos desportivos, sociais e paroquiais - *Distância 2*
(*factor3*)

Factor 4

O quarto factor agrega apenas duas variáveis, o acesso a televisão por cabo e gás canalizado. Com uma capacidade explicativa de 6,8%, este factor pretende evidenciar as habitações que têm ou não acesso a estes dois tipos de serviços.

Ainda que pareça abusivo optou-se por classificar este factor como um indicador de qualidade e conforto. No conjunto das variáveis iniciais não foram consideradas características objectivas de qualidade e conforto por falta de informação de base. Variáveis que traduzissem a qualidade dos materiais de construção e/ou tipo de equipamentos disponíveis na habitação ajudariam a reforçar melhor este conceito.

Também neste factor não existe uma padronização das localizações destes imóveis.

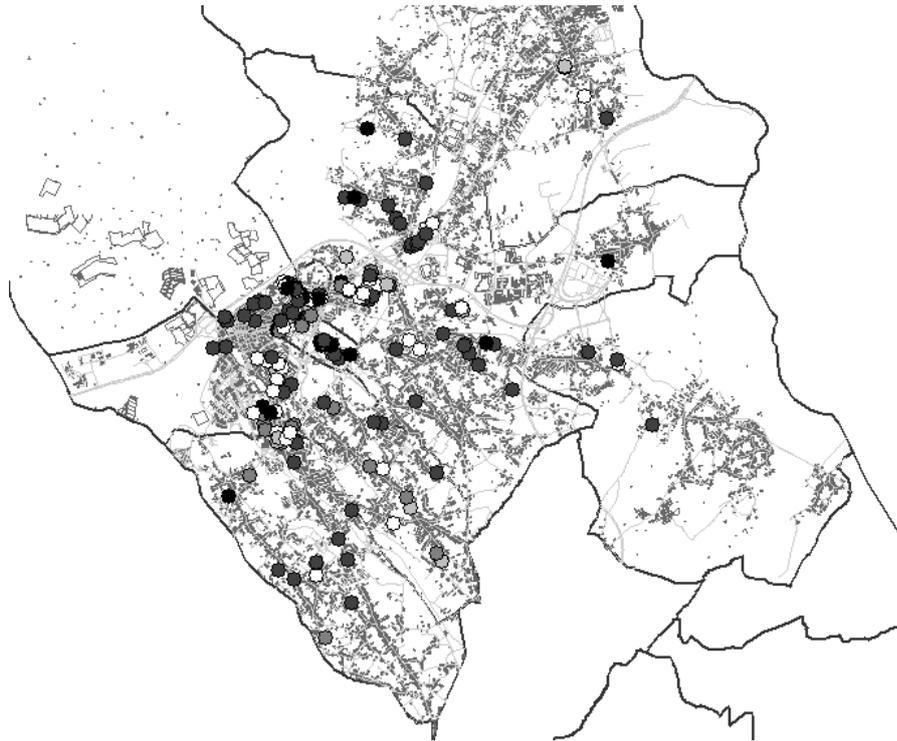


Figura 7.- Acesso a infra-estruturas / Qualidade e conforto (factor 4)

Factor 5

Por último, o quinto factor retido nesta análise, com uma explicação de 7,1%, tem associado a si variáveis que retratam as características dimensionais do edificado, volumetria do edificado. Quer as variáveis piso ou número de pisos, quer a variável que traduz o número de elevadores, identificam-se com a volumetria do edifício. Em termos da sua expressão no território e numa análise mais geral, não se encontram padrões de localização específicos.



Figura 8.- Volumetria (factor 5)

Definição de agrupamentos de imóveis

Neste ponto analisam-se os resultados de uma análise de *clusters*, que tem como objectivo se perceber a organização espacial dos imóveis, no que respeita às diferentes características estudadas: acessibilidades, dimensão habitacional, indicador de qualidade e conforto, volumetria do edificado).

Para a determinação de agrupamentos homogéneos nos dados analisados, foi realizada uma análise de *clusters* hierárquica, usando como variáveis caracterizadoras dos imóveis, os factores que resultaram da análise factorial (ponto anterior). Aplicaram-se vários critérios de agregação (*nearest neighbor*, *furthest neighbor*, *between groups*, *within groups*, *centroid groups* e *wards*) recaindo a escolha no método *furthestneighbor*. Usando uma distância euclidiana, como medida de aproximação, definiram-se 9 agrupamentos. A figura 9 mostra o resultado da análise.

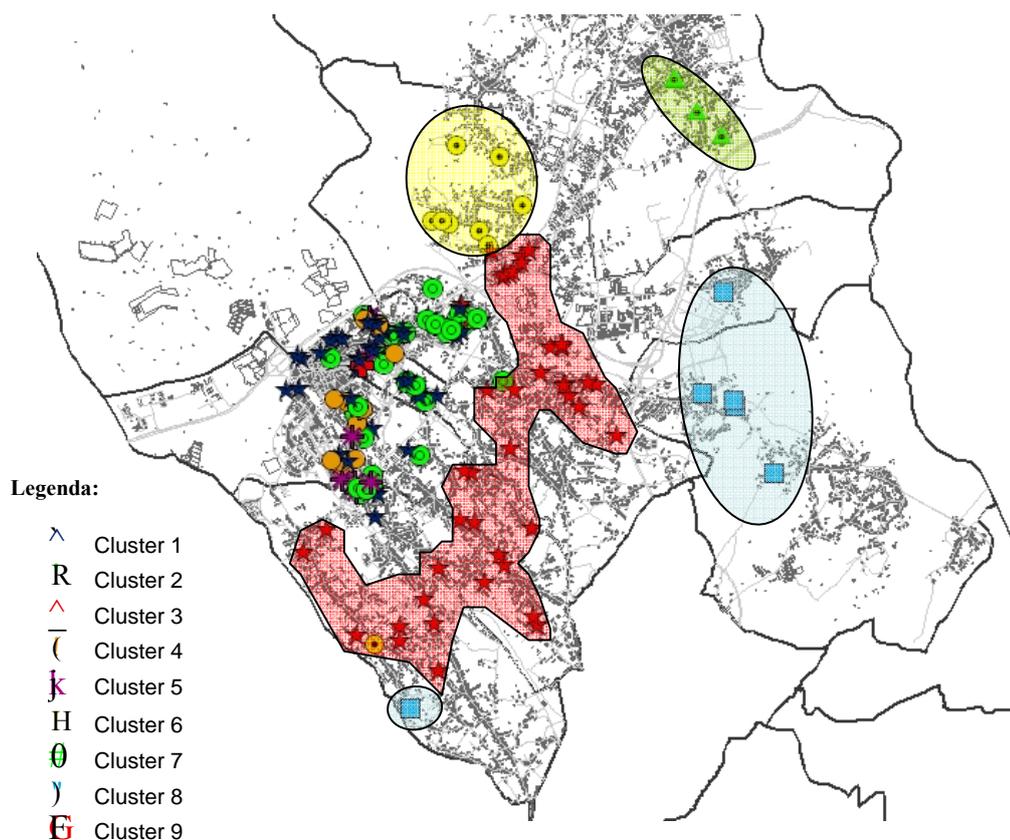


Figura 8.- Volumetria (factor 5)

Quadro 5.- Classificação dos clusters

	F1 Distância 1	F2 Dimensão	F3 Distância 2	F4 Qual. e conforto	F5 Volumetria
C1	↓			↓	
C2	↓	↗		↗	↓
C3	↑	↘	↓	↘	↘
C4	↓	↓	↑	↑	↘
C5	↓	↘		↑	↗
C6	↑	↑	↑	↓	
C7	↑	↘	↑	↗	↘
C8	↑		↓	↘	↗
C9	↘	↑	↗	↑	

Cluster 1 – Habitações com grandes acessibilidades (valores das distâncias baixos) e baixos níveis de qualidade e conforto (não são servidas por gás canalizado e televisão por cabo)

Cluster 2 – Habitações com boas acessibilidades, com áreas grandes e volumetria dos edifícios baixa. Distinguem-se do cluster anterior, por serem habitações cobertas por serviços gás canalizado e televisão por cabo.

Cluster 3 – Habitações com baixos níveis de centralidade ao CBD da cidade (Central Business District), mas com altos níveis de centralidade a equipamentos de desporto e centros sociais (lares 3ª idade) localizados pelo concelho. São habitações que apresentam dimensões médias e volumetrias do seu edificado relativamente baixas e nos indicadores de qualidade e conforto considerados apresentam também valores baixos.

Cluster 4 – Habitações com fortes níveis de centralidade, com áreas e volumetrias do seu edificado de reduzidas dimensões e com indicadores de qualidade e conforto acima da média.

Cluster 5 – Tal como no cluster 4, estas habitações caracterizam-se por terem uma localização privilegiada relativamente ao centro e cobertura de televisão por cabo e gás canalizado, mas distinguem deste por terem volumetrias dos edifícios mais elevadas.

Cluster 6 – Este cluster é constituído por habitações com boas áreas sem acesso a gás e televisão por cabo, mas afastadas quer do centro da cidade quer da segunda centralidade considerada no modelo (equipamentos sociais e desportivos).

Clusters 7 – Em termos de centralidade este cluster é idêntico ao anterior (cluster 6) distinguindo-se deste, ao nível das dimensões da habitação e do acesso à televisão por cabo e gás canalizado. São habitações que apresentam volumetrias, do seu edificado, relativamente baixas.

Cluster 8 – No que respeita à centralidade este cluster tem características semelhantes ao cluster 3, isto é, pouco centrais relativamente CBD da cidade, mas com altos níveis de centralidade a equipamentos de desporto e centros sociais. Distingue-se deste cluster relativamente ao indicador de qualidade e conforto e quanto à volumetria (valores baixos no primeiro e altos no segundo).

Cluster 9 – Este cluster agrega apenas uma habitação. É um outlier no contexto da amostra deste estudo pois apresenta valores elevados na sua dimensão.

Construção da equação funcional (modelo básico, extensões)

Neste ponto procura estimar-se um modelo estatístico para prever o comportamento da variável dependente (preço da habitação) a partir de um conjunto de variáveis explicativas (atributos físicos da habitação e atributos de localização).

Serão apresentados 3 modelos:

- Modelo L1 – variável dependente é o preço da habitação medido em euros e as variáveis independentes são os factores provenientes da Análise de Factorial (bom descriptor das características gerais da habitação, mas não do preço)
- Modelo L2 – variável dependente é o preço da habitação medido em euros e as variáveis independentes são as variáveis iniciais que deram origem a esses factores (bom descriptor das características específicas da habitação, e também do preço. Este modelo é uma especificação do Modelo L1, em virtude do detalhe das características habitacionais consideradas).
- Modelo L3) – variável dependente é o preço da habitação medido em euros por m² e as variáveis independentes são as variáveis iniciais (explicação unitária do preço por m²).

Em ambos os modelos, o objectivo é determinar a relação entre todas as variáveis independentes com o valor do preço da habitação (medido em euros ou em euros por m²) e perceber assim a relação de cada um dos atributos na valorização de uma habitação.

O método utilizado na regressão múltipla é o método de estimação *stepwise* (através do software *SPSS - Statistical Package for Social Sciences*). Em alternativa a considerar todas as variáveis do estudo no modelo, excluem-se aquelas que não contribuem significativamente para o aumento do valor dos coeficientes de correlação, ajudando também a perceber a multicolinearidade entre as variáveis independentes.

Modelo L1: Preço = f(factores)

Quadro 6.- Modelo sumário

R	R ²	Signific. F	Durbin-Watson
0,72	0,52	0,00	2,148

Predictors: (Constant), Factor 1; Factor 2

Quadro 7.- Coeficientes da regressão

Modelo	Não padronizados (B)	Padronizados (Beta)	t	Signif.
(Constant)	139378,91		43,34	**
Dimensão habitacional				
Factor 2	43364,11	0,71	13,45	**
Distância 1				
Factor 1	-8263,71	-0,14	-2,56	*

** Significante; ** Altamente significante*

Variáveis excluídas do modelo: Factor 3; Factor 4; Factor 5

A análise de regressão mostra que existe uma razoável correlação entre cada uma das variáveis independentes e a variável dependente. Validado também pelo teste F, levando a considerar o modelo estimado bastante significativo. De entre os 5 factores, considerados na análise factorial, apenas o factor 1 (Acessibilidade aos equipamentos terciários) e o factor 2 (dimensão habitacional) são estatisticamente significantes, os outros factores foram excluídos deste modelo.

Uma vez que o objectivo deste estudo não é estimar o valor do preço da habitação em função dos seus atributos, mas sim perceber a relação funcional da capacidade explicativa das variáveis do modelo, a função é expressa em termos de coeficientes beta.

A função para um modelo explicativo de 52% é:

$$\text{Preço(€)} = 139378,9 + 0,71 F2 - 0,14 F1$$

Em que:

F1- Dimensão da habitação

F2 - Distância/Acessibilidade a serviços, comércio e equipamentos

Os coeficientes Beta, por serem medidas estandardizadas, permitem comparar directamente os coeficientes das variáveis e o seu poder explicativo. Estes indicam que o preço da habitação é directamente proporcional à dimensão da habitação e inversamente proporcional à distância. Relativamente aos outros factores, estes não são significativos na explicação do preço da habitação.

Como foi referido anteriormente, os modelos de regressão devem considerar algumas premissas:

i) Distribuição normal dos resíduos

Para a análise da normalidade dos erros é usado o teste de Kolmogorov-Smirnov que testa a hipótese nula de os erros terem uma distribuição normal. Sendo o p-value 0,001 rejeita-se a hipótese de que a variável segue uma distribuição normal para os habituais nível de significância de ($\alpha=0,01$ e $\alpha=0,05$)

A figura 9 mostra a existência de valores atípicos (distanciados 1,5 e 3 vezes a distância interquartis) para os resíduos. O quadro 9 permite fazer um diagnóstico dos erros. Os valores dos resíduos estandardizados são obtidos pela divisão de cada desvio pelo seu desvio-padrão, têm média zero e desvio-padrão 1. Podemos verificar que em 3 habitações o erro associado é superior a 3 desvios-padrão, o que evidencia a presença de *outliers*. Ao excluírem-se os 3 valores atípicos mais elevados (a habitação 43, 104 e 135) a capacidade explicativa do modelo aumenta 2 % e a normalidade dos resíduos é assegurada.

Figura 9.- Diagrama de caixas

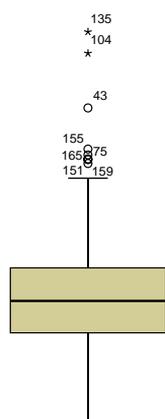
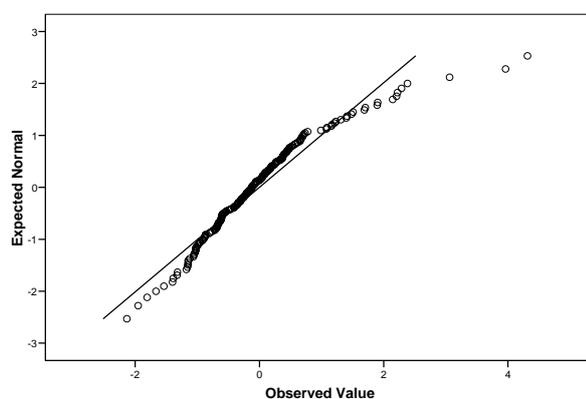


Figura 10.- Desvios à normalidade dos erros



Quadro 8.- Teste da Normalidade dos resíduos

Kolmogorov-Smirnov Test	Resíduos
N	175
Kolmogorov-Smirnov	0,095
Sig.	0,001

a Test distribution is Normal

ii) Erros independentes

A autocorrelação pode ser verificada pela estatística de Durbin-Watson (quadro 6), onde a hipótese nula é a existência de autocorrelação entre os resíduos. Assim de uma forma mais ou menos empírica, rejeita-se H_0 se $d \sim 2,0 (+0,2)$. Uma forma mais exacta, proposta por Durbin e Watson é comparar, através de uma tabela de valores críticos, o valor d com um limite superior (d_U) e um limite inferior (d_L) para testar a H_0 de não existir autocorrelação entre resíduos. (Maroco; 2003) Como se pode verificar no quadro 6, o teste Durbin-Watson tem um valor de 2,148 o que leva à rejeição da hipótese de existir autocorrelação entre os resíduos, isto é, assumem-se erros independentes.

Modelo L2: Preço = f(variáveis)

Quadro 9.- Modelo sumário (para 12 variáveis)

R	R ²	Signific. F	Durbin-Watson
0,893	0,797	0,000	2,477

Predictors: (Constant), AreaSala, AreaTotal_2, Estado, NumAnos, TipoHabitacao, Dist_Centro, Dist_Escolas, Tipologia, Garagem, AreaQuartos, Ruido, Dist_Hotelaria

Quadro 10.- Coeficientes da regressão (para 12 variáveis)

Variáveis independentes	Não padronizados (B)	Padronizados (Beta)	t	Signif.
(Constant)	14439,2		-2,69	*
AreaSala	2151,81	0,22	4,90	**
AreaTotal	292,56	0,21	3,81	**
Estado	46753,18	0,24	6,45	**
NumAnos	-1821,42	-0,24	-5,67	**
TipoHabitacao	58765,66	0,33	7,38	**
Dist_Centro	6,01	0,14	0,30	-
Dist_Escolas	95,34	1,77	3,91	**
Tipologia	21610,82	0,29	5,20	**
Garagem	7587,38	0,12	3,00	**
AreaQuartos	2521,15	0,11	2,81	*
Ruído	9913,16	0,10	2,57	*
Dist_Hotelaria	-109,82	-2,06	-2,43	*

*Significante; ** Altamente signficante

Variáveis excluídas do modelo: Duplex; AreaCozinha; NumWC; NumPISOS; Piso; Varanda; Terraço; NumElevadores; NumFrentes; TVCabo; Gás; Habitação a custos controlados; Dist_Saúde; Dist_Infantários; Dist_Ens_Sup_; Dist_Mercados; Dist_Hipermercados; Dist_Centro_Comercial; Dist_Desporto; Dist_Cultura_e_Lazer; Dist_Segurança; Dist_Centros_sociais_e_paroquiais; Dist_Serviços_da_Administração;

O Modelo L2 relaciona o preço com as características das habitações (quer físicas quer locais). No total das 31 variáveis inicialmente consideradas no modelo, retiveram-se 12 com uma capacidade explicativa de 78% da variação do preço das habitações. Tal como já tinha sido indiciado no modelo anterior (Modelo L1) as variáveis que mais explicam o valor da habitação são aquelas relacionadas com as suas características dimensionais e das acessibilidades. Numa análise mais detalhada verifica-se que a localização privilegiada para a valorização de uma habitação tende a não ser mais o centro da cidade, mas sim outras centralidades que despontam no concelho. É ainda significativo o facto da dimensão da habitação ser considerada factor determinante para a explicação do preço, arredando os factores locais para segundo plano.

i) Distribuição normal dos resíduos

Também para esta análise foi testada a normalidade dos erros usando o teste de Kolmogorov-Smirnov. Neste caso sendo o *p-value* 0,026 rejeita-se a hipótese de que a variável segue uma distribuição normal para um nível de significância de 5%. Neste caso verifica-se a existência de um *outliers*, distanciado a mais de 3 vezes da distância inter-

quartis. Não considerando este valor crítico na análise, a capacidade explicativa do modelo passaria a ser de aproximadamente 80% e os critérios de normalidade dos resíduos são assegurados.

Figura 11.- Diagrama de caixas

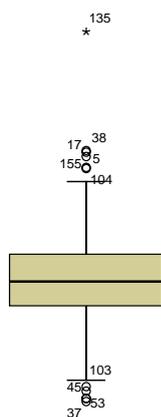
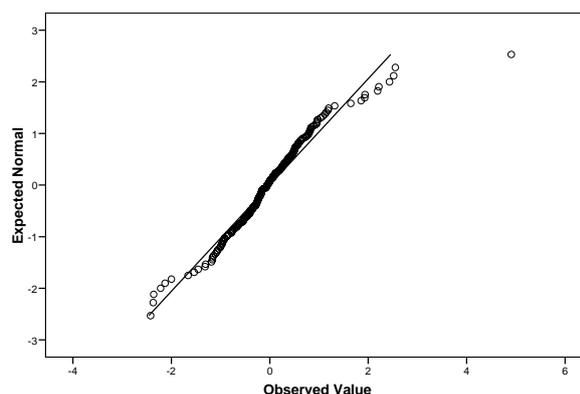


Figura 12.- Desvios à normalidade dos erros



Quadro 11.- Teste da Normalidade dos resíduos

Kolmogorov-Smirnov Test		Resíduos
N		175
Kolmogorov-Smirnov		0,073
Sig.		0,026

a Test distribution is Normal

ii) Erros independentes

Da mesma forma do modelo anterior, e consultando o quando 6, o teste Durbin-Watson tem um valor de 2,364 o que leva à rejeição da hipótese de existir autocorrelação entre os resíduos.

Modelo L3: $Preço(m^2) = f(\text{variáveis})$

Quadro 8.- Modelo sumário (10 variáveis)

R	R ²	Significância F	Durbin-Watson
0,828	0,686	0,009	2,461

Predictors: (Constant), AreaTotal, AreaSala, Estado, TipoHabitação, NumAnos, Dist_Centro, Dist_Escolas, Tipologia, Ruído, Dist_Hotelaria

Quadro 12.- Coeficientes da regressão para o modelo 10

Variáveis independentes	Não padronizados (B)	Padronizados (Beta)	t	Signif.
(Constant)	1824,9		3,091	*
AreaTotal	-7,209	-0,698	-10,611	**
AreaSala	18,368	0,257	4,981	**
Estado	380,451	0,272	5,837	**
TipoHabitacao	618,411	0,470	8,648	**
NumAnos	-21,011	-0,372	-7,686	**
Dist_Centro	-0,039	-0,125	-0,217	-
Dist_Escolas	0,822	2,090	3,746	**
Tipologia	128,458	0,238	3,476	*
Ruido	87,172	0,115	2,512	*
Dist_Hotelaria	-0,861	-2,219	-2,121	*

*Significante; ** Altamente significativa

Variáveis excluídas do modelo: AreaCozinha; AreaQuartos; NumWC; NumPISOS; Piso; Varanda; Terraço; Garagem; NumElevadores; NumFrentes; TVCabo; Gás; Habitação a custos controlados; Dist_Saúde; Dist_Infantários; Dist_Ens_Sup_; Dist_Mercados; Dist_Hipermercados; Dist_Centro_Comercial; Dist_Desporto; Dist_Cultura_e_Lazer; Dist_Segurança; Dist_Centros_sociais_e_paroquiais; Dist_Serviços_da_Administração

Este modelo não difere substancialmente do Modelo L2 no que respeita às variáveis do modelo. A função apresentada, tem uma capacidade explicativa de 69%, menos 10% que a do Modelo L1. O facto do Modelo L2 ter um coeficiente de determinação maior justifica-se em virtude de a variável preço, medida em valor absoluto, ter associada uma componente dimensional que está retratada também nas variáveis independentes. Já no Modelo L3, usando como variável dependente, o preço por m², a componente dimensional é retirada.

A análise apresentada pode não ir ao encontro dos resultados esperados, dado que as características locacionais não são consideradas determinantes para a valorização do preço. Reflectindo sobre a natureza da informação, no sentido de perceber a lógica dos resultados, importa dizer que as habitações analisadas correspondem maioritariamente ao mercado de usados (11% do total). E neste sentido poderá pensar-se que numa análise comparativa, entre habitações usadas nas freguesias mais centrais/urbanas e as mesmas habitações nas freguesias mais peri-urbanas, a lógica seja diferente, isto é, o estado de conservação dos imóveis de localização mais central pode ser substancialmente mais precária (mesmo devolutas) do que aquelas localizadas na periferia, desvalorizando assim o seu valor. De qualquer forma não deixa de ser relevante o peso secundário da variável distância ao

centro, levando a pensar-se que a valorização dos preços habitacionais se vai construindo baseada em novas centralidades da cidade.

Conclusões

Este estudo foi elaborado numa perspectiva exploratória no qual se pretendeu aplicar a um centro urbano a metodologia de preços hedónicos. Mesmo com a dificuldade inerente à recolha de informação foi possível construir uma matriz de atributos representativa do mercado imobiliário (essencialmente para habitações usadas). Algumas variáveis não foram consideradas no modelo em virtude da falta de dados. Os próprios preços das habitações podem não traduzir correctamente a vontade de cada indivíduo ou família pagar por aquele imóvel e pelas suas características, uma vez que os preços apresentados não correspondem ao valor transaccionado mas sim ao valor pedido pelo proprietário, depois de normalizado pela agência imobiliária.

Os modelos apresentados são um exemplo da panóplia de modelos que se poderiam apresentar. Existem bastantes desenvolvimentos de modelos aplicados aos preços hedónicos, no entanto optou-se, pela aplicação de modelos simples (lineares) pelo carácter exploratório desta reflexão.

Agradecimentos

Os autores deste trabalho querem expressar o seu reconhecimento à agência imobiliária, ERA Aveiro, que facultou a informação dos imóveis analisados neste estudo e também aos alunos da disciplina de Diagnóstico Socio-económicos, do curso de Planeamento Regional e Urbano, pelo trabalho de georeferenciação do mesmos. A estes últimos, um agradecimento especial pela disponibilidade demonstrada na reflexão desta temática.

Bibliografia

- A. S. Adair, J. N. B. W. S. M. (1996). "Hedonic modelling, housing submarkets and residential valuation." *Journal of Property Research* **13**(1): 67-83.
- ANDERSON, S. T. AND S. E. WEST (2006). "Open space, residential property values, and spatial context." *Regional Science and Urban Economics* **36**(6): 773-789.
- BLOMQUIST, G. AND L. WORLEY (1981). "Hedonic prices, demands for urban housing amenities, and benefit estimates." *Journal of Urban Economics* **9**(2): 212-221.
- BOURASSA, S. C., F. HAMELINK, ET AL. (1999). "Defining Housing Submarkets." *Journal of Housing Economics* **8**(2): 160-183.
- BOWES, D. R. AND K. R. IHLANFELDT (2001). "Identifying the Impacts of Rail Transit Stations on Residential Property Values." *Journal of Urban Economics* **50**(1): 1-25.
- BRASINGTON, D. M. AND D. HITE (2005). "Demand for environmental quality: a spatial hedonic analysis." *Regional Science and Urban Economics* **35**(1): 57-82.
- DINAN, T. M. AND J. A. MIRANOWSKI (1989). "Estimating the implicit price of energy efficiency improvements in the residential housing market: A hedonic approach." *Journal of Urban Economics* **25**(1): 52-67.
- DONNELLY, W. A. (1991). "A survey in applied environmetrics: The hedonic valuing of environmental amenities." *Environment International* **17**(6): 547-558.
- EARL, D. B., L. H. JULIA, ET AL. (1998). "Pricing Residential Amenities: The Value of a View." *The Journal of Real Estate Finance and Economics* **16**(1): 55-73.
- GANG-ZHI, F., O. SEOW ENG, ET AL. (2006). "Determinants of House Price: A Decision Tree Approach." *Urban Studies* **43**(12): 2301-2316.
- GREGORY, J. C. (2000). "Pricing size effects in housing markets." *Journal of Property Research* **17**(3): 203-219.
- HARALD, N. (2006). "Pricing Location: A Case Study of the Munich Office Market." *Journal of Property Research* **23**(2): 93-107.
- HUI, E. C. M., C. K. CHAU, ET AL. (2007). "Measuring the neighboring and environmental effects on residential property value: Using spatial weighting matrix." *Building and Environment* **42**(6): 2333-2343.
- JENNIFER, L. AND B. RADHA (2005). "Prediction Performance of a Hedonic Pricing Model for Housing." *The Appraisal Journal* **73**(2): 198.
- LARGEY, A. G. (2002). *Social issues in urban economics*. United States -- California, University of California, Irvine.
- 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.
- MAROCO, J. (2003) "Análise estatística com utilização do SPSS", Edições sílabo.
- MYERS, C. K. (2004). "Discrimination and neighborhood effects: understanding racial differentials in US housing prices." *Journal of Urban Economics* **56**(2): 279-302.
- PESTANA, H., "Análise de dados para ciências sociais", Edições Sílabo.
- SABYASACHI, B. AND G. T. THOMAS (1998). "Analysis of Spatial Autocorrelation in House Prices." *The Journal of Real Estate Finance and Economics* **17**(1): 61-85.
- SIRMANS, G. S., M. LYNN, ET AL. (2006). "The Value of Housing Characteristics: A Meta Analysis." *The Journal of Real Estate Finance and Economics* **33**(3): 215-240.

SMITH, V. K., C. POULOS, ET AL. (2002). "Treating open space as an urban amenity." *Resource and Energy Economics* **24**(1-2): 107-129.

SO, H. M., R. Y. C. TSE, ET AL. (1997). "Estimating the influence of transport on house prices: evidence from Hong Kong." *Journal of Property Valuation & Investment* **15**(1): 40.

STEVE, G. (2003). "Paying for Good Neighbours: Estimating the Value of an Implied Educated Community." *Urban Studies* **40**(4): 809-833.

TYRVAINEN, L. (1997). "The amenity value of the urban forest: an application of the hedonic pricing method." *Landscape and Urban Planning* **37**(3-4): 211-222.

WAKEFIELD, K. L. AND J. J. INMAN (2003). "Situational price sensitivity: the role of consumption occasion, social context and income." *Journal of Retailing* **79**(4): 199-212.

WANG, D. AND S.-M. LI (2006). "Socio-economic differentials and stated housing preferences in Guangzhou, China." *Habitat International* **30**(2): 305-326.

YONG, T., S. HUA, ET AL. (2007). "Spatial Autocorrelations and Urban Housing Market Segmentation." *The Journal of Real Estate Finance and Economics*.