



Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

**MESTRADO**  
**GESTÃO E AVALIAÇÃO IMOBILIÁRIA**

**TRABALHO FINAL DE MESTRADO**  
**DISSERTAÇÃO**

**DETERMINANTES DO NÚMERO DE VISITANTES DE UM CENTRO  
COMERCIAL NA ÁREA METROPOLITANA DE LISBOA**

**ANDREIA SOFIA MORENO FONSECA DE ALMEIDA**

**MAIO - 2013**



Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

# **MESTRADO EM GESTÃO E AVALIAÇÃO IMOBILIÁRIA**

## **TRABALHO FINAL DE MESTRADO DISSERTAÇÃO**

**DETERMINANTES DO NÚMERO DE VISITANTES DE UM CENTRO  
COMERCIAL NA ÁREA METROPOLITANA DE LISBOA**

**ANDREIA SOFIA MORENO FONSECA DE ALMEIDA**

**ORIENTAÇÃO:**

**PROF. DRA. SOFIA MARGARIDA FERNANDES FRANCO**

**MAIO - 2013**

## AGRADECIMENTOS

Foram diversas as pessoas que me ajudaram, direta ou indiretamente, a realizar este trabalho final de mestrado, e às quais estou profundamente agradecida.

À minha orientadora, a Prof. Dra. Sofia Franco, pela disponibilidade manifestada em orientar este trabalho, e pelo precioso e incansável apoio ao longo da elaboração do mesmo.

Aos meus professores e colegas da Pós-Graduação, e posterior complemento de Mestrado, em Gestão e Avaliação Imobiliária, pela partilha de conhecimento e espírito de entreajuda e amizade sempre presentes.

À Cushman & Wakefield, por todo o apoio que recebi a diversos níveis. Particular agradecimento à minha chefe de departamento Marta Costa, pelo incentivo que sempre me deu.

A todas as pessoas que me facultaram os dados que serviram de base ao desenvolvimento do meu trabalho, particularmente à Sandra Campos e Fernando Ferreira pela ajuda na obtenção dos mesmos.

Ao Cristiano, pelo apoio, carinho e compreensão ao longo de todos estes meses.

Finalmente, à minha mãe Nair Moreno, por seres quem és, por estares sempre presente, por querereres o melhor para mim.

A todos muito obrigada por me terem ajudado nesta etapa da minha formação académica.

**RESUMO**

Um dos principais desafios para os proprietários de centros comerciais é como atrair visitantes aos seus centros. Esta tese examina os determinantes do número de visitantes de um centro comercial na Área Metropolitana de Lisboa através de um modelo de *spatial lag*. Esta metodologia tem sido raramente utilizada em estudos sobre a atratividade dos centros comerciais e demonstra-se que é melhor face a uma estimativa com um modelo de estimação OLS. A análise também mostra que a presença de uma âncora alimentar e a proximidade a uma estação de metro são determinantes importantes na atração de visitantes a um centro comercial. No geral, os resultados também indicam que, no caso da Área Metropolitana de Lisboa, os centros comerciais competem uns com os outros, mas há também efeitos positivos de aglomeração.

**PALAVRAS CHAVE**

Centro comercial; retalho; modelo espacial; *spatial lag*; *spatial error*; e autocorrelação espacial.

**ABSTRACT**

One of the main challenges for mall owners is how to attract shoppers to patronize their malls. This thesis examines the determinants of the number of visitors to a shopping center in the Lisbon's Metropolitan Area using a spatial lag model. This methodology has been rarely used in studies exploring shopping center attractiveness and is shown to improve upon standard OLS estimation of cross-sectional data. The analysis also shows that housing a supermarket and proximity to a metro station are important determinants of shopping center patronage. Overall, the results also indicate that in the case of the Lisbon Metro Area, shopping centers do compete with each other, but there are also positive agglomeration effects.

**KEYWORDS**

Shopping centre; retail; spatial model; spatial lag; spatial error; and spatial autocorrelation.

## ÍNDICE

<b>1. INTRODUÇÃO</b> .....	5
<b>2. REVISÃO DA LITERATURA</b> .....	10
2.1. Literatura Teórica .....	10
2.2. Literatura Empírica .....	12
2.2.1. Análise das Rendas como Variável Endógena .....	12
2.2.2. Análise das Vendas como Variável Endógena .....	16
2.2.3. Análise do Número de Visitantes como Variável Endógena .....	17
<b>3. DADOS</b> .....	19
3.1. Variável endógena .....	19
3.2. Variáveis explicativas .....	19
3.2.1. Características do centro comercial .....	19
3.2.2. Característica demográfica da área de influência .....	20
3.2.3. Características da área envolvente.....	21
<b>4. MODELO EMPÍRICO</b> .....	22
4.1. Modelo Base .....	23
4.2. Efeitos Espaciais e Econometria Espacial .....	24
4.2.1. Spatial Lag.....	25
4.2.2. Spatial Error.....	26
4.2.3. Matrizes de Distância .....	27
4.2.4. Análise pela Econometria Espacial .....	28
<b>5. ANÁLISE DOS RESULTADOS</b> .....	30
5.1. Análise geral do modelo .....	31
5.2. Análise agregada do modelo base.....	32
5.3. Análise do modelo espacial com spatial lag .....	33
<b>6. CONCLUSÕES</b> .....	37
REFERÊNCIAS BIBLIOGRÁFICAS.....	39
ANEXO 1.....	41
ANEXO 2.....	46

## 1. INTRODUÇÃO

De acordo com a Associação Portuguesa de Centros Comerciais (APCC) (2011), o mercado de retalho português tem sido marcado pela cada vez maior predominância do formato dos centros comerciais, sendo hoje a sua importância estratégica e económica indiscutível<sup>1</sup>. Sendo um local onde se concentram um número significativo de lojas, com gestão integrada e horário de funcionamento alargado, verificou-se uma alteração gradual nos hábitos de compra da população em direção a este formato. Segundo esta fonte, os centros representados pela APCC (90% dos existentes em 2011) recebem aproximadamente 700 milhões de visitas anuais, refletindo cerca de 2 milhões de visitas diárias em média.

Em Portugal, os centros comerciais surgiram nos anos 70, com um conceito ainda pouco estruturado e que passava por unidades de pequena dimensão, em meios urbanos, com comércio de especialidade e sem uma gestão integrada. Como a APCC (2011) sugere, durante os anos 80 dá-se a grande revolução dos mesmos, passando a haver uma preocupação com o *mix* de retalhistas<sup>2</sup>, ou seja com a variedade da oferta existente; uma gestão integrada do espaço; e a centralização da propriedade (por oposição à venda unitária das lojas), o que permitia uma agilização nas tomadas de decisões. O principal exemplo desta evolução é o Amoreiras Shopping Centre, aberto ao público desde 1985 e promovido pela Mundicenter, ainda hoje um centro comercial de referência em Portugal.

---

<sup>1</sup> Segundo a APCC (2011), “os quase 3 milhões de metros quadrados de área bruta locável (ABL) representados pela APCC refletem um volume total de negócios próximo dos 10 mil milhões de euros/ano (5,8% do PIB), sendo uma indústria que tem no ativo mais de 100 mil postos de trabalho diretos”.

<sup>2</sup> Conjunto de lojistas que ocupam um centro comercial (Brown, 1992).

Durante os anos 90 diversificam-se os tipos de centros comerciais, até então de pequena dimensão (tipologias apresentadas na tabela I do anexo 1), começando a surgir os centros de média e grande dimensão, para atrair mais visitantes pelo aumento da oferta de lojistas, em localizações mais periféricas, junto a zonas maioritariamente residenciais e com bons acessos rodoviários, onde o investimento era mais vantajoso pelo menor valor do terreno<sup>3</sup>.

Dado o seu peso em termos do número de habitantes, a Área Metropolitana de Lisboa (AML) concentra o maior número de centros comerciais do país, 43 num total nacional de 120, o que corresponde a cerca de 1,2 milhões de m<sup>2</sup> de ABL<sup>4</sup>.

Analisando mais especificamente a década entre 2001 e 2011<sup>5</sup>, segundo a C&W o novo ABL a entrar no mercado correspondeu a 47% da oferta total existente, espelhando o grande incremento recente em termos de oferta existente e a maturação do mercado de centros comerciais. Conforme resumido na tabela II do anexo 1, durante este período o ABL por 1.000 habitantes na AML, um indicador usado para medir o volume de oferta existente, aumentou 61%. Passou dos 256 m<sup>2</sup>/1.000 habitantes em 2001 para os 413 m<sup>2</sup>/1.000 habitantes em 2011<sup>6</sup>, um valor bastante acima da média dos 27 países da União Europeia que, em Julho de 2011 se

---

<sup>3</sup> Este é o caso do CascaiShopping, inaugurado em 2001 pela Sonae Sierra, o primeiro grande centro regional. Na referida década, em 1997, abre o Centro Colombo também promovido pela Sonae Sierra e que pela sua dimensão, cerca de 120.000 m<sup>2</sup> de ABL, marca outro momento fulcral do setor. Em 2009 abriu o Dolce Vita Tejo, o maior centro comercial da Península Ibérica, com 122.000 m<sup>2</sup> de ABL e desenvolvido pela Chamartín. Este centro espelha a diversificação em termos de oferta para a qual os promotores começaram a evoluir, concentrando valências adicionais como uma grande praça central, onde se podem efetuar espetáculos e eventos, e uma área considerável de lazer, da qual se destaca a Kidzania.

<sup>4</sup> Dados de Maio de 2012 da empresa Cushman & Wakefield (C&W).

<sup>5</sup> Ver figura 1 do anexo 1.

<sup>6</sup> Fonte: ABL Total - Cushman & Wakefield (Maio de 2012) e População Total – INE – Censos 2001 e 2011 (Fevereiro de 2012).



encontrava nos 240 m<sup>2</sup>/1.000 habitantes<sup>7</sup>. Assim, verifica-se que há neste momento uma grande concorrência em termos de oferta na AML.

O que se pretende analisar nesta tese são quais os determinantes do número de visitantes de um centro comercial na Área Metropolitana de Lisboa. Para tal, usou-se uma base de dados em painel, entre 2006 e 2011, referente a 22 centros comerciais localizado na AML. A variável endógena é o número de visitantes anuais por centro comercial e as variáveis exógenas, de carácter espacial e não espacial, caracterizam o centro comercial, a demografia na sua área de influência e a zona envolvente do mesmo. Seguidamente foi usado um modelo de regressão linear múltipla e, havendo problemas de correlação espacial, o modelo usado teve uma componente de ajustamento espacial por *error lags*. Finalmente foi efetuada uma estimação destes modelos e subsequente análise da relevância estatística das variáveis usadas.

Que seja do nosso conhecimento em Portugal foi apenas publicado um estudo sobre o setor dos centros comerciais, um *paper* de Damian e Pinto (2011). Os autores analisaram o impacto das lojas âncora<sup>8</sup> na performance de 35 centros comerciais da Sonae Sierra, entre 2005 e 2007, através da análise dos determinantes das rendas, vendas e número de visitantes. Relativamente a esta última variável, igual à estudada nesta tese, procuraram explicá-la através das variáveis ABL total, idade e número de lojas do centro comercial, número e ABL total das âncoras, lugares de estacionamento e estacionamento pago. Com recurso a um modelo de regressão linear *stepwise* verificaram que, em cada ano, o número de visitantes tem como variáveis explicativas o número de lojas total e número de lojas âncora. O estudo é relevante ao evidenciar a importância da oferta

---

<sup>7</sup> Cushman & Wakefield – *Shopping Centre Development Report* (Setembro de 2011).

<sup>8</sup> Loja que pela reputação da sua marca, atrai consumidores ao centro comercial onde se encontra, aumentando o número de visitantes, vendas e lucros das lojas nas imediações (Konishi e Sandfort, 2003).

quantitativa dos centros comerciais e o impacto que as lojas âncoras têm em atrair visitantes. No entanto, foca-se somente nos centros geridos pela mesma entidade, de qualidade reconhecida, podendo tal facto, à partida, enviesar os resultados. Na presente tese, analisamos os centros comerciais na maior área metropolitana de Portugal e incluímos mais variáveis do que aquelas usadas pelos autores, incluindo a introdução de variáveis espaciais, como forma de contribuir em termos de fundamentação empírica para este setor tão importante a nível nacional.

Em termos internacionais, são diversos os estudos (Sirmans e Guidry, 1993; Gatzlaff, Sirmans e Diskin, 1994; Carter e Haloupek, 2000; Des Rosiers, Thériault e Ménétrier, 2005)<sup>9</sup> que se debruçam sobre a capacidade de atração dos centros comerciais. Embora estes analisem maioritariamente as rendas ou vendas<sup>10</sup> enquanto variáveis exógenas, parte das variáveis usadas nestes estudos, especialmente as não financeiras, procuram captar a capacidade de atração de um centro comercial, sendo este o mesmo tipo de análise em que se baseia o estudo dos determinantes dos visitantes de um centro comercial.

Em linha com a literatura teórica, estes estudos incluem variáveis espaciais e não-espaciais na sua análise, demonstrando ambas capacidade explicativa conforme os modelos criados. Adicionalmente, a correção da autocorrelação espacial tem vindo a ganhar relevância ao melhorar os modelos desenvolvidos, tendo também sido testada a sua existência no modelo econométrico usado, por forma a contribuir mais para a compreensão do que leva os consumidores a deslocarem-se a um centro comercial.

---

<sup>9</sup> Discutidos em mais detalhe na revisão da literatura.

<sup>10</sup> Dado haver menor confidencialidade nos países estudados face ao verificado em Portugal.

Para analisar os determinantes do número de visitantes de um centro comercial na AML, foi inicialmente usado um modelo de regressão linear, considerado o modelo base, com a variável endógena a estar representada em logaritmo e sendo explicada por quinze variáveis exógenas. Uma vez que, havendo presença de efeitos espaciais no modelo base, o critério de não correlação entre os erros é violado e o modelo estimado enviesado e inconsistente, foi testada a presença de dependência espacial neste modelo através dos testes de Moran  $I$  e LM, os quais evidenciam a necessidade de usar um modelo de regressão linear espacial, nomeadamente um modelo de *spatial lag*. Tendo sido testadas diversas matrizes, por resultar no modelo com o menor AIC, o modelo de regressão com dependência espacialmente desfasada final usou como matriz a distância inversa ao quadrado em minutos do centróide das freguesias dos centros comerciais.

Comparando os resultados dos modelos base e *spatial lag*, verifica-se que a capacidade de explicação do modelo é maior no segundo ao apresentar um AIC menor, com as variáveis a terem os sinais esperados e dez a serem significativas até 10%, às quais acresce  $\rho$ , significativo a 5%. Para ser possível comparar os coeficientes estimados nos dois modelos usou-se a metodologia sugerida por LeSage e Pace (2009), separando no modelo de *spatial lag* entre efeitos diretos e indiretos. Com os primeiros a serem diretamente comparáveis com os coeficientes estimados no modelo base, verificou-se que este estava a sobrestimar o impacto das variáveis explicativas.

No capítulo seguinte é efetuada uma revisão da literatura relevante para este trabalho, sendo primeiro apresentados os principais conceitos e modelos teóricos e de seguida analisados os estudos empíricos comparáveis. No terceiro capítulo são apresentadas e analisadas as variáveis usadas. No capítulo seguinte é descrita a metodologia usada nesta tese, incluindo os modelos de

regressão usados e testes aplicados. No quinto capítulo são analisados e comparados os resultados das regressões finais e no sexto capítulo são apresentadas as principais conclusões deste trabalho.

## 2. REVISÃO DA LITERATURA

Esta revisão da literatura é composta por duas partes. Na primeira, são descritos os conceitos e modelos teóricos para a presente tese, e na segunda, é efetuada uma análise dos estudos empíricos considerados relevantes para este trabalho.

### 2.1. *Literatura Teórica*

Ao longo dos últimos anos, o estudo sobre a temática dos centros comerciais tem vindo a alargar o foco sobre variáveis espaciais (mercado de oferta e procura, localização e edifício) à inclusão de variáveis não espaciais (imagem e *mix* de retalhistas<sup>11</sup>).

Relativamente às variáveis espaciais, a sua análise teve início com a teoria do lugar central, abordada por Reilly (1931) e melhorada por Huff (1964, 1966) e Christaller (1966), que estabelecia que a capacidade de atracção de um polo de retalho face a outros varia positivamente com o seu tamanho e negativamente com a sua distância aos consumidores.

Com a proliferação do uso automóvel e a maior mobilidade dos consumidores, torna-se mais fácil para estes efetuarem diversas tarefas apenas numa viagem, crescendo a preferência por centros comerciais mais longe e com maior número de retalhistas (Ertekin, Dokmeci, Unlukara e Ozus, 2008). O *mix* de retalhistas torna-se mais desafiante para os gestores dos centros, ao

---

<sup>11</sup> Conjunto de lojistas que ocupam um centro comercial (Brown, 1992).

procurarem tirar proveito da aglomeração de diversas lojas num só local. Por um lado, os custos de deslocação total do consumidor reduzem e por outro, os lojistas beneficiam de uma maior presença de consumidores, aumentando a capacidade de atração dos centros maiores (Eaton e Lipsey, 1982; Mulligan, 1983; Eppli e Benjamin, 1994).

Os estudos teóricos começam a analisar o impacto das variáveis não espaciais nos centros comerciais, entre as quais as externalidades positivas que são geradas, quer pela presença de retalhistas homogéneos (pela compra por comparação), quer por retalhistas heterogéneos (pelas compras diversificadas) num centro comercial.

O benefício da aglomeração de retalhistas semelhantes dá-se por efeito da comparação do preço de um produto entre lojas, previamente à sua compra. Introduzido por Hotelling (1929), consiste num modelo de concorrência espacial em que o consumidor escolhe entre duas lojas baseado no custo do produto e da deslocação. Uma vez atingido o preço de equilíbrio pela maximização do lucro das lojas, a concorrência efetua-se pela localização cada vez mais central destas, para minimizar os custos de deslocação dos clientes mais distantes. O autor conclui que ambas as lojas maximizam o seu lucro caso se localizem juntas no centro de determinado mercado, pois será neste caso que o preço é igual.

Brueckner (1993) alarga a análise das externalidades por efeito de aglomeração de lojistas num centro comercial aos retalhistas heterogéneos, criando um modelo em que as vendas de determinada loja dependem da área por si ocupada e também da área ocupada pelos restantes lojistas. Conclui que as vendas de uma loja aumentam à medida que a área ocupada pelas restantes é maior, pois torna o centro mais atrativo no geral. Contudo, ressalva que no caso de

lojistas homogêneos, poderão haver casos em que são originadas externalidades negativas entre lojas semelhantes por efeito do aumento da concorrência<sup>12</sup>, cabendo à gestão do centro fazer um balanço entre os ganhos e perdas dos lojistas no total.

Esta relação entre tamanho das lojas e maior capacidade de atração do centro comercial é mais forte no caso da presença de lojas âncora, pelo seu impacto ao nível da imagem<sup>13</sup> do mesmo, abordado por Brueckner (1993) e Eppli e Shilling (1995). Estes últimos criaram um modelo de análise do efeito cruzado entre a preferência dos consumidores por loja âncora e lojas satélite<sup>14</sup>. O modelo considera que a maior atração dos centros são as âncoras, dado o número de consumidores que originam. Concluem que, quanto maior for o efeito positivo na procura de lojas satélite pela presença de lojas âncora, maiores serão os centros, assim como a área alocada às âncoras. Este aumento origina um maior número de consumidores, crescendo assim o lucro do promotor e as oportunidades de desenvolvimento de centros comerciais.

## ***2.2. Literatura Empírica***

Os estudos empíricos apresentados de seguida incluem na sua análise variáveis explicativas espaciais e não espaciais. Em termos de metodologia, apenas o de Carter e Haloupek (2000) usa um modelo de regressão espacial.

### ***2.2.1. Análise das Rendas como Variável Endógena***

Os estudos empíricos abaixo analisados usam como variável endógena as rendas praticadas. No entanto, conforme mencionado por Des Rosiers, Thériault e Ménétrier (2005), uma maior

---

<sup>12</sup> Mais verificado no caso de lojas que vendem bens iguais, como sejam uma tabacaria ou uma farmácia.

<sup>13</sup> A percepção que o consumidor tem dos atributos de uma loja ou centro comercial através de elementos como a qualidade, serviço e conveniência dos artigos para venda (Mejia e Benjamin, 2002).

<sup>14</sup> Lojas não-âncora.

natureza confidencial das mesmas (igualmente verificada em Portugal) tem limitado o número de estudos comparativamente com outros sectores imobiliários.

Sirmans e Guidry (1993) basearam-se em dados de 1989 a 1991, de 55 centros comerciais em Baton Rouge, Luisiana (EUA), para analisarem os determinantes das variações da renda por pé quadrado ( $\text{ft}^2$ ) nos centros. Agruparam as variáveis exógenas enquanto *proxies* de quatro categorias: DRAW (poder de atracção do consumidor) - tamanho, idade e presença ou não de âncora nacional; DESIGN (design do edifício) - *dummies* sobre se o centro é fechado, resulta da aglomeração de edifícios independentes, é em formato em L ou em U; LOC (localização) - *dummies* para distinguir a localização e volume de tráfego nas imediações e MKT (mercado) - área desocupada, variáveis de tendência do tempo, população e rendimento *per capita*.

O modelo de regressão linear original revela existência de heterocedasticidade, sendo usado um modelo de mínimos quadrados ponderados. Para MKT usaram-se dois conjuntos de *proxies*: Modelo 1 com *dummies* que identificam o ano dos dados e Modelo 2 com população e rendimento *per capita*. Ambos conseguem explicar 85% das variações nas rendas por  $\text{ft}^2$  e as variáveis apresentam os sinais esperados. Tudo o resto constante, as rendas por  $\text{ft}^2$  dos centros sobem \$2,25 a \$2,43 com âncoras nacionais e \$0,5 por cada 10.000  $\text{ft}^2$  adicionais e diminuem \$0,1 por cada ano de idade ou 10.000  $\text{ft}^2$  vagos adicionais.

Este modelo tem sido muito referido e aplicado na análise das rendas dos centros comerciais, usando variáveis chave na explicação das rendas nos centros. As variáveis de DESIGN abordam um conjunto de *layout* de centros mais específico dos EUA e sem aplicação em Portugal, e o tipo

de tráfego parece desadequado, sendo o número de visitantes mais indicado. A estrutura deste modelo será aproveitada para o presente trabalho.

Gatzlaff, Sirmans e Diskin (1994) examinaram qual o impacto que a perda de uma âncora teria nas rendas de um centro comercial, usando dados de 1990 de 36 centros próximos uns dos outros na Florida e Geórgia (EUA). Baseados na metodologia aplicada por Sirmans e Guidry (1993), o modelo de regressão linear considerou como variável dependente a renda média anual por ft<sup>2</sup> dos retalhistas não âncora do centro, e como variáveis independentes as categorias: DRAW (replicaram as mesmas variáveis e adicionaram *dummy* da perda de loja âncora), LOCATION (localização por região) e MARKET (área desocupada, adicionada do poder de compra efetivo).

Usam um modelo de mínimos quadrados ponderados, cujos resultados demonstram que as rendas do centro diminuem com a sua idade e aumentam com o poder de compra da área de influência. A perda de uma âncora resulta numa diminuição média de rendas de 26,14%. Estimam um segundo modelo de variável instrumental ponderada, assumindo que a perda de uma âncora afeta primeiro a área vaga do centro (resultados demonstram que esta aumenta em 33% da área total), tendo depois um impacto negativo nas rendas, como forma dos gestores do centro manterem os retalhistas presentes (resultados apresentam uma quebra de 27.3% das mesmas). Desta forma, este estudo confirma a robustez do anterior assim como a importância das âncoras para os gestores dos centros e restantes retalhistas.

Carter e Haloupek (2000) usaram informação de 1991 e 1992, de 689 lojas satélite, localizadas em 8 centros comerciais nos EUA para analisar a variação da renda total das lojas satélite de um centro. Usaram como variáveis exógenas: *dummy* relativa à loja fazer parte de uma cadeia



regional/nacional; tamanho ( $\text{ft}^2$  e  $\text{ft}^2$  ao quadrado); duração do contrato; *dummy* sobre loja vender produtos comparáveis; *dummy* de localização de cada centro e medidas de acessibilidade da loja (distâncias ao centro do centro comercial e à loja semelhante, vaga e saída mais próximas).

A análise visual do *scatter plot* dos resíduos do modelo de mínimos quadrados ponderados identifica a existência de autocorrelação espacial, com uma não dispersão aleatória dos mesmos, e resíduos positivos elevados perto do centro físico dos centros comerciais e negativos elevados na periferia. Efetuaram-se ajustamentos ponderados às observações localizadas até 200 pés do centro físico do centro comercial para corrigirem esta autocorrelação, afetando os resultados da regressão, com um maior  $R^2$  e  $R^2$  ajustado e uma soma dos erros quadrados menor em 4%.

Este estudo demonstra que o recurso a modelos de regressão espacial para corrigir a existência de autocorrelação espacial permite melhorar o modelo, assim como retificar a violação da hipótese de distribuição idêntica e independente dos resíduos deste. Os autores usaram ajustamentos ponderados das observações mas, havendo autocorrelação espacial neste trabalho serão usados modelos com componente de ajustamento espacial por *error lags* e/ou *spatial lags*.

Des Rosiers, Thériault e Ménétrier (2005) analisaram dados financeiros (1998-2000) e de arrendamentos (1976-2000) de 8 centros comerciais no Quebeque (Canadá). A renda base unitária por  $\text{ft}^2$  era a variável endógena e as exógenas *proxys* de: Tamanho ( $\text{ft}^2$  do centro); *Mix* (categoria de artigos vendidos e do centro), Potencial e Atração. Para os dois últimos criaram índices baseados num inquérito a consumidores (onde concluíam que apenas 69% das visitas se

originavam a partir de casa<sup>15</sup>): Índice de Potencial Económico com base no número de visitantes e seu rendimento anual e Índice de Atracção do Centro adaptado do modelo de Reilly (1931).

São estimadas diversas versões do modelo, concluindo-se que os coeficientes de regressão deste em centros suprarregionais têm maior magnitude do que em centros regionais, corroborando as externalidades positivas da aglomeração de retalhistas em centros maiores. Parte dos fatores do *mix* de retalhistas apresentam poder explicativo, assim como o Índice de Potencial Económico.

Este estudo introduz variáveis que serão usadas neste trabalho, nomeadamente o *mix* de retalhistas. Por forma a captar parte dos visitantes a centros que não se deslocam a partir de casa (31% no inquérito referido) serão ainda usados os dados de trabalhadores na área de influência.

Damian e Pinto (2011) estudaram o impacto de lojas âncora na performance de 35 centros comerciais da Sonae Sierra em Portugal e Espanha, com base em dados de 2005-2007. A quarta hipótese que testam é se as lojas que beneficiam de externalidades positivas pela presença de lojas âncora no centro pagam um prémio por tal: a renda paga pelas lojas satélite (variável endógena) é função do número, venda total, renda fixa e renda variável das lojas âncora. O modelo final é explicado pela venda total das lojas âncora, comprovando em parte a hipótese, validando a importância das variáveis não espaciais na análise dos centros comerciais.

### ***2.2.2. Análise das Vendas como Variável Endógena***

Outra variável usada na análise em centros comerciais são as vendas, dado serem a variável considerada como medidor do nível de performance dos centros, portanto críticas quando

---

<sup>15</sup> As restantes tiveram origem do local de trabalho, escola ou outros locais de atividade, incluindo outras lojas.

analisados projetos no setor de retalho (Mejia e Benjamin, 2002; Okoruwa, Nourse e Terza, 1994). Tal como as rendas, esta variável tem natureza confidencial generalizada em Portugal.

A primeira hipótese testada no estudo de Damian e Pinto (2011) é se o número de âncoras e a área que ocupam influencia as vendas totais do centro comercial, a variável endógena. Esta será explicada pelo número e área das âncoras, lugares de estacionamento e estacionamento pago (S/N). Usando um modelo de regressão *stepwise*, as variáveis exógenas finais são número de âncoras e lugares de estacionamento, corroborando em parte esta hipótese. A terceira hipótese do estudo é que as lojas âncoras têm impacto nas vendas totais por visitante (variável endógena), explicada pelo ABL total, idade e número de lojas do centro, número e ABL total das âncoras, lugares de estacionamento e estacionamento pago (S/N). O modelo de regressão *stepwise* demonstra que, em 2007 e 2006, as variáveis significativas são os lugares de estacionamento e idade do centro e para 2005, é o número de âncoras, só sendo para este ano suportada a hipótese.

### ***2.2.3. Análise do Número de Visitantes como Variável Endógena***

Uma terceira variável explicada nos estudos é o número de visitantes do centro comercial. Embora esta não origine diretamente rendimento para lojistas e proprietários, é determinante para um funcionamento saudável do centro, com um estudo efetuado por Nicholls, Li, Kranendonk e Roslow (2002) a revelar que cerca de 55% dos visitantes inquiridos num centro em Miami, Florida (E.U.A.) em 1999 fizeram compras durante a sua visita a este. Em Portugal, estes valores são obtidos pelos principais proprietários dos centros comerciais e, embora a sua disponibilização a terceiros não seja prática generalizada, será esta a variável usada no presente estudo.

A segunda hipótese do estudo de Damian e Pinto (2011) é a de que as âncoras providenciam aos centros comerciais rendimento e atraem consumidores. Esta segunda parte é analisada através de um modelo de regressão linear *stepwise* para cada ano, em que o número de visitantes anuais (variável endógena) é explicado pelas mesmas variáveis usadas na terceira hipótese já apresentada. Para os três anos (2005 a 2007), o modelo tem como variáveis explicativas o número de lojas total e número de lojas âncora, suportando a segunda hipótese.

Da literatura empírica analisada, verifica-se que a maioria se foca na componente financeira dos centros comerciais. No entanto, a capacidade de um centro gerar melhores resultados financeiros depende, em parte, das mesmas variáveis que determinam a procura deste pelos visitantes, ou seja, da sua capacidade de atração. Quantos mais consumidores se atraírem, mais compras serão efetuadas, aumentando as vendas do mesmo e subseqüentemente as rendas cobradas. Assim, são usadas nesta tese variáveis (espaciais e não espaciais) usadas para explicar os determinantes das rendas e vendas, que se consideram ter também impacto na capacidade de gerar maior número de visitantes num centro comercial.

Por outro lado, a autocorrelação espacial é ainda pouco abordada, assim como é reduzido o uso de sistemas de informação geográfica. Ambos são usados neste trabalho, como forma de procurar trazer uma mais-valia aos estudos empíricos sobre este segmento de retalho.

### 3. DADOS

Os dados usados nesta tese procuram captar as características associadas aos principais centros comerciais na Área Metropolitana de Lisboa. São valores anuais, entre 2006 e 2011, de 22 centros comerciais, cuja localização geográfica se encontra na figura 2 do anexo 1. Obtidos entre Janeiro e Junho de 2012, junto de diversas fontes, foram agregadas no total 113 observações<sup>16</sup>.

#### 3.1. *Variável endógena*

A variável endógena, referente ao número de visitantes anual por centro comercial, foi obtida junto da APCC (através dos dados publicados nos Anuários dos Centros Comerciais) e dos promotores e/ou gestores dos respetivos centros comerciais.

#### 3.2. *Variáveis explicativas*

##### 3.2.1. *Características do centro comercial*

Como forma de caracterizar cada centro comercial, agregou-se para cada centro e anualmente, as variáveis: número de lojas satélite e *mix* comercial<sup>17</sup> das lojas âncora, dados também obtidos junto da APCC, promotores e/ou gestores dos centros comerciais, assim como através da C&W (pelas suas bases de dados internas). Para o *mix* das lojas âncora, usou-se o ABL total ocupado pela âncora alimentar (supermercado ou hipermercado) e pelas restantes lojas âncora.

Para efeitos da caracterização demográfica e da envolvente foi gerada uma área de influência para cada centro comercial, medida através do tempo (em minutos) percorrido de automóvel,

---

<sup>16</sup> Dado que nem todos os centros comerciais estavam abertos desde 2006 ou não dispunham de informação para todos os anos analisados.

<sup>17</sup> Conjunto de lojistas que ocupam um centro comercial (Brown, 1992).

tendo para tal sido usado o *software* informático *Insite Europe*, utilizado pela C&W sob licença da CACI<sup>18</sup>. O *Insite Europe* assume percursos de carro, realizados fora das horas de ponta, em todas as direções possíveis a partir de uma determinada localização. Consoante a duração máxima dos vários percursos de carro, são posteriormente identificadas as freguesias que ficam dentro<sup>19</sup> da área de influência escolhida.

Cada área de influência pode ter mais do que uma isócrona, ou seja, ter subdivisões em termos da distância máxima percorrida. Para esta tese foi criada uma área de influência de 30 minutos para cada centro comercial, com isócronas de 0-10, 10-15 e 15-30 minutos.

Para uma análise do impacto da melhoria das infraestruturas rodoviárias nas áreas de influência, seria ideal gerar cada uma com base nas estradas existentes em cada ano analisado. Tal não foi possível pois o *software* usado tem vindo a ter versões atualizadas que implicam a perda de acesso aos mapas de estradas anteriores. No entanto, dos 63,2 km de itinerários complementares ou principais concluídos entre 2006 e 2011 na AML<sup>20</sup> (tabela III do anexo 1), acreditamos que o maior impacto destes foi na redução do tempo de deslocação em horas de ponta, não prejudicando desta forma os resultados obtidos.

### 3.2.2. *Característica demográfica da área de influência*

Quanto à caracterização demográfica, calculou-se a população residente nas freguesias da isócrona de 0-30 minutos. Para tal, foram usados os dados da população residente por freguesia

---

<sup>18</sup> Empresa britânica dedicada à conceção e comercialização de *software* informático, especializada em sistemas de informação geográfica.

<sup>19</sup> A freguesia é incluída caso o centróide, ou seja, o ponto geográfico central dos limites da freguesia, se situe dentro da área de influência gerada.

<sup>20</sup> Com base na informação publicada pelo Instituto de Infraestruturas Rodoviárias.

publicados nos Censos de 2001 e de 2011 do Instituto Nacional de Estatística (INE). Para a estimativa do valor de cada variável por freguesia para os anos de 2006 a 2010, foi calculada a taxa média de crescimento geométrico anual<sup>21</sup> entre 2001 e 2011 e aplicada em termos de capitalização anual aos valores de 2001.

### 3.2.3. *Características da área envolvente*

Para a caracterização da área envolvente, obtiveram-se dados dos trabalhadores nas imediações do centro comercial e que são potencialmente visitantes deste, tendo para tal sido usada a remuneração média mensal na freguesia do centro comercial, através dos dados anuais (2006 a 2010) por freguesia, obtidos junto do Ministério do Trabalho e da Segurança Social (MTSS). Os valores de 2011 foram estimados aplicando-se aos dados de 2010 a taxa de crescimento média anual por freguesia entre 2006 e 2010.

Adicionalmente, foi agregada a distância à paragem de metro mais próxima, medida em minutos percorridos a pé recorrendo ao Google Maps<sup>22</sup>, tendo-se seguidamente gerado uma variável *dummy* igual a 1 caso houvesse uma estação a menos de 10 minutos a pé.

Finalmente, por forma a captar as acessibilidades na envolvente de cada centro comercial obteve-se o número de centros comerciais com ABL inferior e igual ou maior a 0-10 minutos; e totais a 10-15 minutos, calculados recorrendo ao *Insite Europe*.

As tabelas IV e V do anexo 1 apresentam, respetivamente, a definição das variáveis e a descrição estatística da base de dados usada neste trabalho. Em média, o número de visitantes

---

<sup>21</sup>  $10^{\sqrt[10]{\text{Variável}_{2011}/\text{Variável}_{2001}} - 1}$

<sup>22</sup> [www.maps.google.pt](http://www.maps.google.pt)

anuais de um centro comercial na AML é de cerca de 8,87 milhões, tendo no entanto uma variação dispersa, entre um mínimo de 1,27 milhões e um máximo de 26,9 milhões. A base de dados é constituída por dados em painel com efeitos fixos, ou seja, cuja adaptação a um modelo de regressão linear se efetuou com a introdução de *dummies* do ano a que as variáveis dizem respeito. Em média os centros comerciais têm 129 lojas satélite e, embora pelo menos numa observação de um centro não hajam lojas âncora, a âncora alimentar ocupava em média 10.911 m<sup>2</sup> e as restantes lojas âncora 9.725 m<sup>2</sup> de ABL.

A população residente nas freguesias até 30 minutos de carro de cada centro comercial varia entre 464 mil e 2.190 mil habitantes, sendo em média de 1.880 mil. A remuneração dos trabalhadores na freguesia do centro é de 1.344€ e 39% dos centros comerciais têm metro até 10 minutos percorridos a pé. Os centros comerciais costumam ter 2,5 centros comerciais com ABL inferior e 2 com ABL igual ou maior até 10 minutos de carro, aumentando para 6,5 na distância entre 10 e 15 minutos de carro. A variação destas três variáveis é entre nenhum concorrente e 9, 8 e 13, respetivamente, centros comerciais.

#### 4. MODELO EMPÍRICO

De seguida, são apresentados os dois tipos de modelos empíricos que foram usados para analisar os determinantes do número de visitantes de um centro comercial na Área Metropolitana de Lisboa (AML). O modelo base corresponde a um modelo de regressão linear e o segundo, por forma a captar a influência da localização das variáveis entre si, corresponde a um modelo de regressão linear espacial.



#### 4.1. Modelo Base

O modelo base, de regressão linear através do método dos mínimos quadrados, usado para explicar o número de visitantes de um centro comercial na AML tem esta estrutura:

$$(1) \quad \begin{aligned} \ln visit_{it} = & \beta_0 + \beta_1 ljsat_i + \beta_2 lnablalim_i + \beta_3 lnablljancsalim_i + \beta_4 lnpop30_i \\ & + \beta_5 lnremuntrabfreg_i + \beta_6 metro10min_i + \beta_7 cc10abABL_i \\ & + \beta_8 cc10ABL_i + \beta_9 cc1015_i + \sum_{t=2006}^{2011} \beta_t dumano_{it} + \varepsilon_{it} \end{aligned}$$

onde  $\ln visit$  corresponde ao logaritmo natural do número de visitantes do centro comercial  $i$  no ano  $t$ , com  $i = 1, \dots, 22$  e  $t = 1, \dots, 6$ . As variáveis estão representadas por vetores, conforme as características apresentadas no capítulo anterior, nomeadamente: centro comercial ( $ljsat$ ,  $lnablalim$  e  $lnablljancsalim$ ), demografia ( $lnpop30$ ) e envolvente ( $lnremuntrabfreg$ ,  $metro10min$ ,  $cc10abABL$ ,  $cc10ABL$  e  $cc1015$ ). A definição das variáveis de (1) encontram-se na tabela IV do anexo 1.

Devido à maior capacidade explicativa em termos de  $R^2$  ajustado, assim como pelo facto de resultar num modelo mais robusto, optou-se por um modelo em que a variável endógena  $visit$  está em logaritmo e a maioria das variáveis explicativas em valores lineares (modelo log-lin)<sup>23</sup>. Assim, os  $\beta_i$  estimados correspondem à taxa de variação aproximada do número de visitantes de um centro comercial na AML quando a variável  $X_i$  varia em uma unidade.

---

<sup>23</sup> Em comparação com um modelo com variáveis endógena e explicativas em valores lineares (modelo lin-lin) e variáveis endógena e explicativas com valores em logaritmo (modelo log-log).

#### ***4.2. Efeitos Espaciais e Econometria Espacial***

Os dados espaciais são caracterizados pela presença de autocorrelação e heterogeneidade espaciais. Referidas como efeitos espaciais, estas características são importantes na análise econométrica dos dados, pois podem invalidar os resultados metodológicos dos modelos de regressão não espaciais. Tal levou à necessidade de desenvolver um conjunto de técnicas que lidam com as peculiaridades causadas pelo espaço na análise estatística dos modelos, sendo estas consideradas como o domínio da econometria espacial (Anselin, 1988).

A dependência ou autocorrelação espacial deriva da existência de uma relação de dependência entre as observações no espaço. Dois tipos de efeitos podem levar a que tal aconteça, o primeiro devido a erros de medição das observações em unidades espaciais contiguas. Uma vez que os dados são obtidos de forma espacial agregada, podem ocorrer efeitos de *spillover* entre as fronteiras das observações. O segundo, resulta da existência de uma variedade de fenómenos de interação nos quais o espaço é um dos elementos estruturantes na explicação do comportamento humano, ou seja, o que é observado num local é em parte influenciado pelo que acontece nos restantes locais (Anselin, 1988).

A heterogeneidade espacial resulta da não uniformidade no espaço do comportamento de determinadas variáveis, significando que a forma funcional e parâmetros estimados de um dado modelo variam consoante a sua localização, não sendo portanto homogêneos.

Consoante a relação espacial encontrada nos dados, são geralmente estimados modelos que levam em conta os efeitos espaciais, nomeadamente os modelos com dependência espacialmente desfasada ou com dependência espacial dos erros.

No caso do processo de escolha de deslocação de um consumidor a um centro comercial na Área Metropolitana de Lisboa, a atratividade de determinado centro pode não depender exclusivamente de si, mas também do tipo de capacidade dos centros comerciais vizinhos em atraírem visitantes. Numa análise inicial, o impacto espacial de variáveis pode ser incluído num modelo de regressão linear com a inclusão do número de centros concorrentes nas isócronas 0-10 e 10-15 minutos. No entanto, a influência dos centros comerciais vizinhos pode ser mais intrínseca e originar dependência das observações entre um ou mais centros próximos uns dos outros, originando autocorrelação espacial.

#### 4.2.1. *Spatial Lag*

A primeira forma de incorporar os efeitos espaciais num modelo econométrico passa pelos valores da variável endógena das localizações vizinhas serem incluídos como uma variável explicativa adicional, criando-se o modelo de regressão com dependência espacialmente desfasada (*spatial lag*):

$$(2) \quad Y = \rho WY + X\beta + \varepsilon$$

onde:  $Y$  é o vetor  $n \times 1$  das observações da variável endógena número de visitantes;  $X$  a matriz  $n \times k$  das observações das variáveis explicativas;  $\beta$  o vetor  $n \times 1$  dos coeficientes a estimar;  $\varepsilon$  o vetor dos erros com distribuição normal, homocedásticos e não correlacionados;  $\rho$  o coeficiente espacial autorregressivo; e  $W$  a matriz  $n \times n$  de distâncias. Esta última consiste na interpretação da relação espacial assumida entre as observações, em que aquelas que não interagem têm um valor de 0 e as que interagem têm um valor diferente de 0.  $Wy$  é assim a variável dependente espacialmente desfasada, ou seja, a variável endógena ponderada em termos de distância às

observações vizinhas. Desta forma, o coeficiente  $\rho$  mede a forma como as observações vizinhas afetam a variável endógena e, caso não exista dependência espacial,  $\rho = 0$ .

Resolvendo (2) em ordem a  $Y$  tem-se:

$$(3) \quad Y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon$$

onde:  $I$  é a matriz de identidade  $n \times n$  e as duas expressões inversas são consideradas os multiplicadores espaciais. Desta forma,  $\beta_\kappa$  corresponde, não só ao efeito estimado da variação da variável  $x_\kappa$ , mas também ao impacto das variações desta mesma variável nas restantes localizações.

#### 4.2.2. *Spatial Error*

Caso a autocorrelação espacial esteja presente nos resíduos, incluem-se, como variável explicativa adicional, os valores dos resíduos estimados das localizações vizinhas, criando-se um modelo de regressão com dependência espacial dos erros (*spatial error*):

$$(4) \quad Y = X\beta + \varepsilon$$

$$(5) \quad \varepsilon = \lambda W\varepsilon + \mu$$

onde:  $\lambda$  é o coeficiente espacial autorregressivo;  $W$  a matriz de distâncias; e  $\mu$  o vetor dos erros que satisfaz as assunções clássicas de distribuição idêntica e independente. Neste caso,  $W\varepsilon$  consiste numa variável dos erros ponderada pelas observações vizinhas e o coeficiente  $\lambda$  capta a forma como o termo de erro das observações vizinhas influencia a variável endógena e as variáveis explicativas. Na ausência de correlação espacial entre os erros,  $\lambda = 0$ .

Resolvendo (4) e (5) em ordem a  $Y$  tem-se:

$$(6) \quad Y = X\beta + (I - \lambda W)^{-1}\mu$$

pelo que neste caso  $\beta_\kappa$  corresponde ao efeito estimado da variação da variável  $x_\kappa$ , sendo portanto a sua interpretação semelhante à dos coeficientes estimados do modelo base.

#### 4.2.3. Matrizes de Distância

Para se estimarem os modelos espaciais apresentados e efetuarem os testes de seguida apresentados, é necessário criarem-se diversas matrizes de distância. Estas servem para estabelecer quais as observações no espaço que têm influência sobre determinada observação, resultando da interpretação que seja feita, para cada matriz, dos conceitos de vizinhança e vizinhos. Não há atualmente um critério definido para a escolha entre o tipo de matrizes de distância a usar nos modelos de regressão, pelo que a solução passa pela criação de diversas matrizes, optando-se pelas que resultam em modelos mais robustos.

Cada matriz  $W$ , é composta por elementos  $w_{ij}$ , que expressam a interação espacial entre a variável  $i$  e a variável  $j$ . Estas matrizes podem ser binárias (tomando o valor de 1 se as variáveis  $i$  e  $j$  são consideradas vizinhas e 0 no caso contrário) ou não binárias (em que a relação geográfica entre as variáveis vizinhas é definida através de uma função da distância entre as duas). No caso desta tese as variáveis correspondem ou à localização de cada centro comercial ou ao centróide<sup>24</sup> da freguesia onde se localiza cada centro comercial, sendo de seguida descrito o tipo de matrizes criadas neste trabalho.

---

<sup>24</sup> Descrito no capítulo anterior.

Relativamente à vizinhança, consideram-se as definições em termos de contiguidade (partilha de fronteira) e distância (quilómetros e minutos). Para considerar a existência de vizinhança pela contiguidade<sup>25</sup>, usou-se o critério de matriz *rook* em que caso houvesse partilha de um ou mais lados da fronteira entre duas freguesias estas eram consideradas vizinhas, sendo neste caso usada uma matriz binária. Adicionalmente foi gerada uma matriz *row-standardized*, em que o valor das células de cada linha (*row*) é dividido pela sua soma, dando-se maior peso às observações com menos vizinhos e vice-versa.

Quanto à definição em termos de distância<sup>26</sup> ( $d$ ), foi calculada a distância em quilómetros e minutos percorridos de carro e os quilómetros em linha reta entre cada variável. Os valores obtidos de cada uma das matrizes resultantes foram ainda transformados no seu valor inverso ( $1/d$ ) e inverso ao quadrado ( $1/d^2$ ), dando assim maior peso aos vizinhos mais próximos e vice-versa. Foram igualmente criadas matrizes onde os valores acima do valor mínimo para cada localização ter pelo menos um vizinho eram 0, por forma a limitar a dependência entre as variáveis.

#### ***4.2.4. Análise pela Econometria Espacial***

A análise da presença de dependência espacial no número de visitantes de um centro comercial da AML começa com o teste da presença de efeitos espaciais no respetivo modelo de regressão linear através do método dos mínimos quadrados. Havendo presença destes efeitos, o critério de não correlação entre os erros é violado e o modelo estimado enviesado e inconsistente (Anselin, 1988). Frequentemente é usado um modelo de regressão linear com componentes

---

<sup>25</sup> Recorreu-se ao GeoDa.

<sup>26</sup> Recorreu-se ao InSite.

espaciais através do método da máxima verosimilhança, o qual será também usado neste trabalho para o modelo espacial.

No caso desta tese, ao se focar na Área Metropolitana de Lisboa, a questão do impacto da localização dos centros comerciais entre si é relevante, dada a já referida oferta elevada existente<sup>27</sup>, o que pode levar a que a capacidade de atração de alguns seja influenciada pela proximidade dos outros. Adicionalmente, a distribuição geográfica destes, com centros comerciais de dimensão semelhante localizados próximos entre si, poderá também causar interdependências espaciais que enviesam os resultados num modelo não espacial.

A expectativa *à priori* é de que o modelo final a usar seja o de *spatial lag*, pois do ponto de vista da teoria económica este é o mais apropriado a usar, ao permitir capturar os efeitos *spillover* entre os centros comerciais, ou seja, o impacto no número de visitantes dos centros concorrentes causado pela interdependência espacial existente relativamente a um centro, que tende a ser mais forte consoante a sua proximidade.

Uma discussão mais teórica sobre os testes de Moran I e do Multiplicador de Lagrange (LM), e sua aplicação ao modelo base, encontram-se no anexo 2. Conclui-se que o modelo a usar é de *spatial lag*, tendo como matriz a distância inversa ao quadrado em minutos do centróide das freguesias dos centros comerciais. Os resultados deste modelo demonstram que o parâmetro  $\rho$  é positivo (0,058) e significativo (a 5%), indicando que há autocorrelação espacial no número de visitantes de centros comerciais na AML para centros comerciais localizados em freguesias próximas umas das outras.

---

<sup>27</sup> Tabela II do anexo 1 demonstra que existe uma grande oferta face à média europeia (UE-27).

## 5. ANÁLISE DOS RESULTADOS

TABELA I  
REGRESSÕES COM LN<sub>VISIT</sub> COMO VARIÁVEL DEPENDENTE

Variáveis	Modelo Base	Modelo de <i>spatial lag</i> (1/d <sup>2</sup> minutos da freguesia)
a1	0.094 (0.89)	0.118 (1.23)
a2	0.114 (1.10)	0.139 (1.47)
a3	0.157 (1.55)	0.176* (1.90)
a4	0.057 (0.56)	0.076 (0.82)
a5	0.042 (0.42)	0.050 (0.55)
ljsat	0.003*** (4.32)	0.002*** (3.84)
lnablalim	0.037** (2.51)	0.046*** (3.26)
lnabljancsalim	0.113*** (3.49)	0.098*** (3.25)
lnpop30	0.317** (2.19)	0.323** (2.45)
lnremuntrabfreg	0.300 (1.51)	0.333* (1.84)
metro10min	0.266** (2.43)	0.236** (2.34)
cc10abABL	0.019 (1.00)	0.001 (0.06)
cc10ABL	-0.124*** (-5.82)	-0.162*** (-6.35)
cc1015	-0.034*** (-2.90)	-0.057*** (-3.90)
constante	7.686*** (3.10)	7.400*** (3.28)
$\rho$		0.058** (2.30)
Nº. de observações	113	113
R <sup>2</sup>	0.851	



R <sup>2</sup> Ajustado	0.829	
L	-19,21	-16,62
AIC	68,42	67,24
RSS	9,295	

Nota: \*\*\*  $p$ -value <0,01; \*\*  $p$ -value <0,05; e \*  $p$ -value <0,1.  $T$ -values em parêntesis.  $\rho$  corresponde ao parâmetro estimado no modelo de *spatial lag*.  $Z$ -values em parêntesis neste caso. L corresponde ao valor do *log-likelihood* e RSS à soma residual dos quadrados.

Para o modelo base, a hipótese nula de não significância conjunta dos coeficientes é rejeitada a um nível de 1% e a hipótese nula de existência de homocedasticidade através do teste White não é rejeitada a um nível de 1%. De acordo com o indicador *Variance Inflation Factor* (VIF), não há indícios de problemas de estimação dos  $\beta_i$  devido a multicolinearidade, encontrando-se os resultados na tabela VI do anexo 1.

### 5.1. Análise geral do modelo

Na tabela I são apresentados os resultados das regressões estimadas, modelo base e modelo de *spatial lag*, tendo como variável dependente o número de visitantes anual de um centro comercial na AML, expressa na forma funcional de logaritmo.

Quanto ao modelo base, este consegue explicar cerca de 83% da variação do número de visitantes anuais de um centro comercial na AML. O resultado de quase todas as variáveis tem os sinais expectáveis e oito são significativas até pelo menos 10%, incluindo a constante.

No entanto, ao se incorporarem os efeitos espaciais através do modelo de regressão com dependência espacialmente desfasada, a capacidade de explicação do modelo aumenta, ao apresentar um AIC<sup>28</sup> menor que o modelo base. Neste modelo as variáveis mantêm os sinais, aumentando o número de variáveis significativas até 10% para dez, incluindo a constante, às quais acresce  $\rho$ , significativa a 5%.

Conforme apresentado no capítulo anterior, os coeficientes estimados pelos dois modelos não são comparáveis pois enquanto o modelo base apenas captura os efeitos diretos das variáveis, no

<sup>28</sup> Comparação pelo R<sup>2</sup> ou R<sup>2</sup> ajustado não é aplicável aos modelos espaciais.

modelo de *spatial lag* são obtidos de forma agregada os efeitos diretos e indiretos, correspondendo estes últimos ao impacto que as variáveis têm sobre as localizações vizinhas. Para que a comparação seja possível, será usada a metodologia sugerida por LeSage e Pace (2009), com a separação no modelo de *spatial lag* entre os efeitos diretos e indiretos, sendo os primeiros diretamente comparáveis com os coeficientes estimados no modelo base. No anexo 2 é explicada como é feita esta separação.

### ***5.2. Análise agregada do modelo base***

Analisando os resultados do modelo base, verificamos que embora as *dummies* dos anos 2006 a 2008 não sejam significativas, estas refletem a esperada diminuição do número de visitantes, com as variáveis dos anos mais recentes a apresentarem coeficientes inferiores.

Relativamente às variáveis usadas para caracterizar os centros comerciais, as variáveis que dizem respeito à distribuição do centro comercial (número de lojas satélite) e seu *mix* comercial (ABL ocupada pela âncora alimentar e pelas restantes lojas âncora) são positivas e significativas a pelo menos 5%.

A variável de descrição da demografia, população nas freguesias até 30 minutos, percorridos de carro, do centro comercial é positiva e significativa a 5%.

O modelo não consegue capturar todas as variáveis de caracterização da área envolvente, nomeadamente o impacto da remuneração dos trabalhadores na freguesia e do número de centros comerciais concorrentes com ABL inferior até 10 minutos de carro que não são significativas, embora a primeira tenha o sinal esperado. Os restantes coeficientes são significativos a pelo menos 5%, com a existência de metro até 10 minutos a pé com capacidade de atração positiva, e

o número de centros concorrentes com ABL igual ou maior até 10 minutos de carro e totais entre 10 e 15 minutos o impacto negativo esperado.

### 5.3. Análise do modelo espacial com *spatial lag*

A tabela II resulta da aplicação da metodologia apresentada no anexo 2, sendo os efeitos diretos médios do modelo de *spatial lag* comparáveis com os coeficientes do modelo base.

TABELA II  
COEFICIENTES DAS REGRESSÕES COM LNVISIT COMO VARIÁVEL DEPENDENTE

Variável	Modelo Base	Modelo de <i>spatial lag</i> (1/d2 minutos da freguesia)		
	Coeficiente	Efeito Direto Médio	Efeito Indireto Médio	Efeito Total Médio
a1	0,09353	0,11801	0,00409	0,12210
a2	0,11421	0,13928	0,00483	0,14410
a3	0,15704	0,17567*	0,00609	0,18176
a4	0,05680	0,07596	0,00263	0,07860
a5	0,04169	0,05008	0,00174	0,05182
ljsat	0,00261***	0,00221***	0,00008	0,00229
lnablalim	0,03745**	0,04590***	0,00159	0,04749
lnabljancsalim	0,11297***	0,09820***	0,00340	0,10161
lnpop30	0,31737**	0,32344**	0,01121	0,33465
lnremuntrabfreg	0,30006	0,33343*	0,01155	0,34499
metro10min	0,26643**	0,23591**	0,00818	0,24409
cc10abABL	0,01905	0,00111	0,00004	0,00115
cc10ABL	-0,12355***	-0,16185***	-0,00561	-0,16746
cc1015	-0,03447***	-0,05743***	-0,00199	-0,05942
constante	7,68567***	7,40005***	0,25644	7,65650
$\rho$		0,05834**	0,00202	0,06037

Nota: \*\*\* *p-value* <0,01; \*\* *p-value* <0,05; e \* *p-value* <0,1.

Em relação ao modelo de *spatial lag*, verificamos um já referido maior número de variáveis significativas face ao modelo base. Uma vez que o modelo base apenas captura os efeitos diretos, serão os efeitos diretos médios do modelo de *spatial lag* a ser usados na comparação dos

coeficientes dos dois modelos. Efetuando esta comparação verifica-se que na maioria dos coeficientes o modelo base estava a sobrestimar os efeitos das variáveis explicativas na variação do número de visitantes anuais.

Neste modelo, a maioria das *dummies* dos anos continua a não ser significativa, à exceção da *dummy* do ano de 2008 significativa a 10%. Contudo, todas são positivas e tendem a decrescer, refletindo a atual conjuntura económica que tem vindo a levar à diminuição do número de visitantes dos centros comerciais. Por exemplo, em 2008 o número de visitantes de um centro comercial era aproximadamente maior em 18% face a 2011, *ceteris paribus*.

Em relação às características do centro comercial, o número de lojas satélite tem um efeito positivo e estatisticamente significativo no número de visitantes. Por cada loja adicional deste tipo, o número de visitantes de um centro comercial aumenta diretamente cerca de 0,2%. Em relação ao efeito indireto, cada loja satélite adicional está associada ao aumento em média de 0,01% nos restantes centros.

Verifica-se que as variáveis do *mix* comercial das lojas âncora são também relevantes na capacidade de atração de visitantes, o que reflete a importância de uma boa distribuição das lojas num centro comercial, especialmente quando estas são âncora. Relativamente ao modelo base, o peso da âncora alimentar aumenta, enquanto das restantes âncoras diminui, o que é esperado, pois muitas das deslocações aos centros são unicamente motivadas pelas compras habituais de bens para casa em supermercados ou hipermercados. No caso da âncora alimentar, o aumento de 1% do ABL da âncora alimentar, aumenta em média o número de visitantes de um centro comercial em 0,05% de forma direta.

Quanto à característica demográfica da área de influência, representada pela população residente nas freguesias até 30 minutos de carro do centro comercial, esta é positiva e significativa a 5%, revelando os benefícios positivos nos visitantes que um centro comercial obtém por estar localizado perto de uma elevada concentração populacional. Neste caso, a variação de 1% da população residente até 30 minutos de um centro origina um acréscimo direto em média de 0,3% no número de visitantes. Já o efeito médio de *spillover* traduz o aumento de 0,01% no número de visitantes dos restantes centros, com a variação de um ponto percentual da densidade populacional de um centro comercial.

Analisando as características da área envolvente, o impacto direto da remuneração dos trabalhadores na freguesia do centro passa a ser relevante face ao modelo base a 10% e positivo. Este resultado revela que pessoas com remunerações em média mais elevadas, ou seja, com maior poder de compra, terão maior tendência para visitar um centro comercial.

Conforme seria expectável, a existência de metro até 10 minutos a pé é significativa a 5% e positiva, demonstrando que a procura por centros comerciais aumenta caso exista uma estação de acesso a este meio de transporte até essa distância.

Quanto aos efeitos da concorrência, os resultados demonstram na quase totalidade o impacto negativo que a presença de centros comerciais concorrentes até 15 minutos de carro têm ao nível da capacidade de atração de visitantes. À exceção da variável do número de centros concorrentes com ABL inferior até 10 minutos, as restantes continuam significativas a 1%, apresentando um sinal negativo esperado. Acresce que se consegue captar a concorrência de uma melhor forma face ao modelo base, dado que este último estava a subestimar o impacto negativo das mesmas.

Por exemplo, por cada número adicional de centros comerciais até 10 minutos com ABL igual ou superior, o número de visitantes de um centro comercial diminui diretamente cerca de 16%, um impacto maior que a redução de cerca de 12% estimada no modelo base. Esta variável tem o maior impacto absoluto, confirmando que estes centros constituem a principal concorrência, ao terem uma maior oferta, seguidos do número total de centros comerciais entre 10 e 15 minutos que, estando mais afastados, têm um menor impacto.

Contudo, analisando o coeficiente  $\rho$ , que mede o grau de dependência espacial, este é positivo (5,8%) e confirma a existência da mesma ao ser significativo a 5%. Este resultado é interessante uma vez que, embora o conjunto das variáveis de concorrência demonstre um impacto negativo do número de centros concorrentes, há uma evidência que o número de visitantes anuais de um centro comercial é influenciado positivamente pelo número de visitantes dos centros vizinhos. Este resultado indica que, apesar de um efeito contrário dos centros comerciais mais próximos, no global os centros da amostra têm um impacto positivo entre si na capacidade de atrair visitantes, usando como medida da sua distribuição a distância inversa ao quadrado em minutos entre o centróide das suas freguesias.

Desta forma, o modelo de *spatial lag*, através do coeficiente  $\rho$ , consegue captar algo que o modelo base não consegue, ou seja, que o efeito líquido dos *spillovers* que a localização dos centros comerciais na AML relativamente uns aos outros tem no número de visitantes é positivo. Ou seja, apesar do impacto negativo da localização de centros concorrentes na proximidade, especialmente os centros maiores e que à partida terão maior capacidade de atração, há um impacto positivo e maior fruto das sinergias criadas pela localização destes. Estas poderão dever-se ao incremento dos visitantes, devido à proximidade a um centro considerado bom e que gera

uma maior afluência à envolvente, pelo fato de os centros mais perto serem de pior qualidade e captarem menos visitantes ou pelo benefício de centros diretamente concorrentes estarem a uma maior distância dos visitantes.

## 6. CONCLUSÕES

Neste trabalho analisaram-se os determinantes do número de visitantes de um centro comercial na Área Metropolitana de Lisboa entre 2006 e 2011. Com recurso a um total de 113 observações, foram agregadas variáveis exógenas, de carácter espacial e não espacial, por forma a caracterizar cada centro comercial, a demografia na sua área de influência e a zona envolvente do mesmo. Foi testada a existência de dependência espacial no modelo de regressão linear inicialmente estimado, nomeadamente com os testes de Moran *I* e LM, concluindo-se sobre a necessidade de usar um modelo de *spatial lag*. O não recurso a um modelo deste género iria comprometer os resultados obtidos uma vez que o critério base do modelo de regressão linear de não correlação entre os erros seria violado e, desta forma, os coeficientes estimados seriam enviesados e inconsistentes.

Adicionalmente, o modelo de *spatial lag* revelou-se melhor a explicar o que influencia a deslocação de uma pessoa a um centro comercial na AML, nomeadamente ao apresentar um AIC menor e a quase totalidade das variáveis a serem significativas até 10% e a terem o sinal esperado, além do coeficiente espacial autorregressivo, significativo a 5%. Comparando os efeitos diretos do modelo com dependência espacialmente desfasada com os coeficientes estimados do modelo OLS, verificou-se também que o modelo base estava a sobrestimar o impacto das variáveis explicativas.

Desta análise conclui-se que o número de lojas satélite e o *mix* comercial das lojas âncora têm um impacto positivo na capacidade de atrair visitantes. Verifica-se também que a população residente, remuneração dos trabalhadores e existência de metro perto são fatores importantes na capacidade de atração de visitantes a um centro comercial.

Relativamente ao impacto da existência de concorrência, demonstra-se que, apesar do impacto imediato negativo da presença de centros concorrentes, no global os centros da amostra têm um impacto positivo entre si na capacidade de atração de visitantes.

Seria interessante alargar este trabalho à escala de Portugal Continental para averiguar se os resultados verificados com esta amostra se mantêm, ou até voltar a efetuar o estudo uma vez ultrapassada a atual crise económica, para perceber se ocorreram alterações naquilo que motiva uma pessoa a deslocar-se a um centro comercial. Este tipo de estudos são importantes para a compreensão de um setor do imobiliário tão importante a nível nacional como o são os centros comerciais.



**REFERÊNCIAS BIBLIOGRÁFICAS**

- Anselin, Luc, "Spatial econometrics: methods and models", Dordrecht: Kluwer Academic Publishers, 1988
- Associação Portuguesa de Centros Comerciais, "'Shopping – centros comerciais em revista" – 20 anos a acompanhar o sector", Shopping – Centros Comerciais em Revista, 2011, 81: 12-21
- Brown, Stephen, "Tenant mix, tenant placement and shopper behavior in a planned shopping centre", The Service Industries Journal, 1992, 12(3): 384-403
- Brueckner, Jan K., "Inter-store externalities and space allocation in shopping centers", Journal of Real Estate Finance and Economics, 1993, 7: 5-16
- Carter, Charles C. e Haloupek, William J., "Spatial autocorrelation in a retail context", International Real Estate Review, 2000, 3(1): 34-48
- Christaller, W., "Central places in southern Germany", Englewood Cliffs, NJ: Prentice Hall, Inc., 1966
- Damian, Diana S.; Curto, José D.; Pinto, José C., "The impact of anchor stores on the performance of shopping centres: the case of Sonae Sierra", International Journal of Retail & Distribution Management, 2011, 39(6): 456-475
- Des Rosiers, François; Thériault, Marius; Ménétrier, Laurent, "Spatial versus non-spatial determinants of shopping center rents: modeling location and neighborhood-related factors", The Journal of Real Estate Research, 2005, 27(3): 293-319
- Eaton, B. Curtis e Lipsey, Richard G., "An economic theory of central places", The Economic Journal, 1982, 92(365): 56-72
- Eppli, Mark e Shilling, James D., "Large-scale shopping center development opportunities", Land Economics, 1995, 71(1): 35-41
- Eppli, Mark J. e Benjamin, John D., "The evolution of shopping center research: a review and analysis", The Journal of Real Estate Research, 1994, 9(1): 5-32
- Ertekin, Ozhan; Dokmeci, Vedia; Unlukara, Tuba; Ozus, Evren, "Spatial distribution of shopping malls and analysis of their trade areas in Istanbul", European Planning Studies, 2008, 16(1): 143-155
- Gatzlaff, Dean H., Sirmans, G. Stacy e Diskin, Barry A., "The effect of anchor tenant loss on shopping center rents", The Journal of Real Estate Research, 1994, 9(1): 99-110
- Hotelling, H., "Stability in competition", Economic Journal, 1929, 39: 41-57

Huff, David L., "A programmed solution for approximating an optimum retail location", *Land Economics*, 42(3): 293-303

Huff, David L., "Defining and estimating a trading area", *Journal of Marketing*, 1964, 28(July): 34-38

Instituto de Infraestruturas Rodoviárias, "Relatório de Monitorização da Rede de IP's e IC's", 2009

Instituto de Infraestruturas Rodoviárias, "Relatório de Monitorização da Rede Rodoviária Nacional", 2010

Konishi, Hideo e Sandfort, Michael T., "Anchor stores", *Journal of Urban Economics*, 2003, 53: 413-435

LeSage, James P. e Pace, R. Kelley, "Introduction to spatial econometrics", CRC Press, 2009

Mejia, Luis C. e Benjamin, John D., "What do we know about the determinants of shopping center sales? Spatial vs. non-spatial factors", *Journal of Real Estate Literature*, 2002, 10(1): 3-26

Mulligan, Gordon F., "Consumer demand and multipurpose shopping behavior", *Geographical Analysis*, 1984, 15: 76-81

Nicholls, J.A.F., Li, Fuan, Kranendonk, Carl J. e Roslow, Sydney, "The seven year itch? Mall shoppers across time", *Journal of Consumer Marketing*, 2002, 19(2): 149-165

Okoruwa, A.A., Nourse, H.O. e Terza, J.V., "Estimating Sales for retail centers: an application of the poisson gravity model", *Journal of Real Estate Research*, 1994, 9(1): 111-123

Reilly, W.J., "The Law of Retail Gravitation", Nova Iorque: Knickerbocker Press, 1931

Sirmans, C.F. e Guidry, Krisandra A., "The determinants of shopping center rents", *The Journal of Real Estate Research*, 1993, 8(1): 107-115

## ANEXO 1

## TABELA I

## TIPOS DE CENTROS COMERCIAIS

Tipo	Área Bruta Locável (m <sup>2</sup> )
Muito Grande	80.000 m <sup>2</sup> e superior
Grande	40.000 - 79.999 m <sup>2</sup>
Médio	20.000 - 39.999 m <sup>2</sup>
Pequeno	5.000 - 19.999 m <sup>2</sup>

Fonte: International Council of Shopping Centres

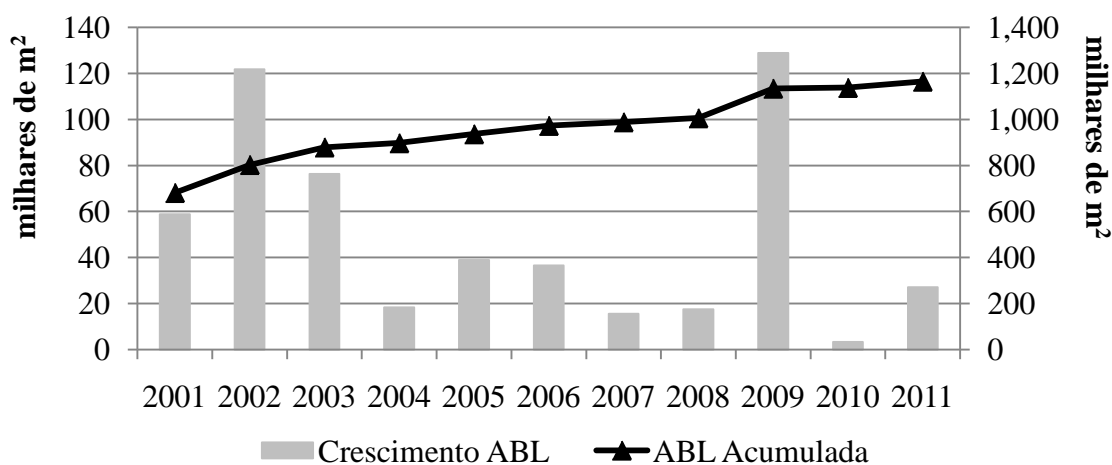


Figura 1 - Centros Comerciais na Área Metropolitana de Lisboa – 2001/2011  
Fonte: Cushman & Wakefield (Maio de 2012)

## TABELA II

## INDICADORES DE CENTROS COMERCIAIS - ÁREA METROPOLITANA DE LISBOA E PORTUGAL

Indicador	Portugal	AML
Centros Comerciais (n°.)	120	43
ABL Atual (milhões de m <sup>2</sup> )	3,0	1,2
Indicador	2001	2011
ABL (m <sup>2</sup> ) / 1.000 habitantes		
AML	256	413
UE-27		240

Fonte: Cushman & Wakefield e INE (Maio de 2012)

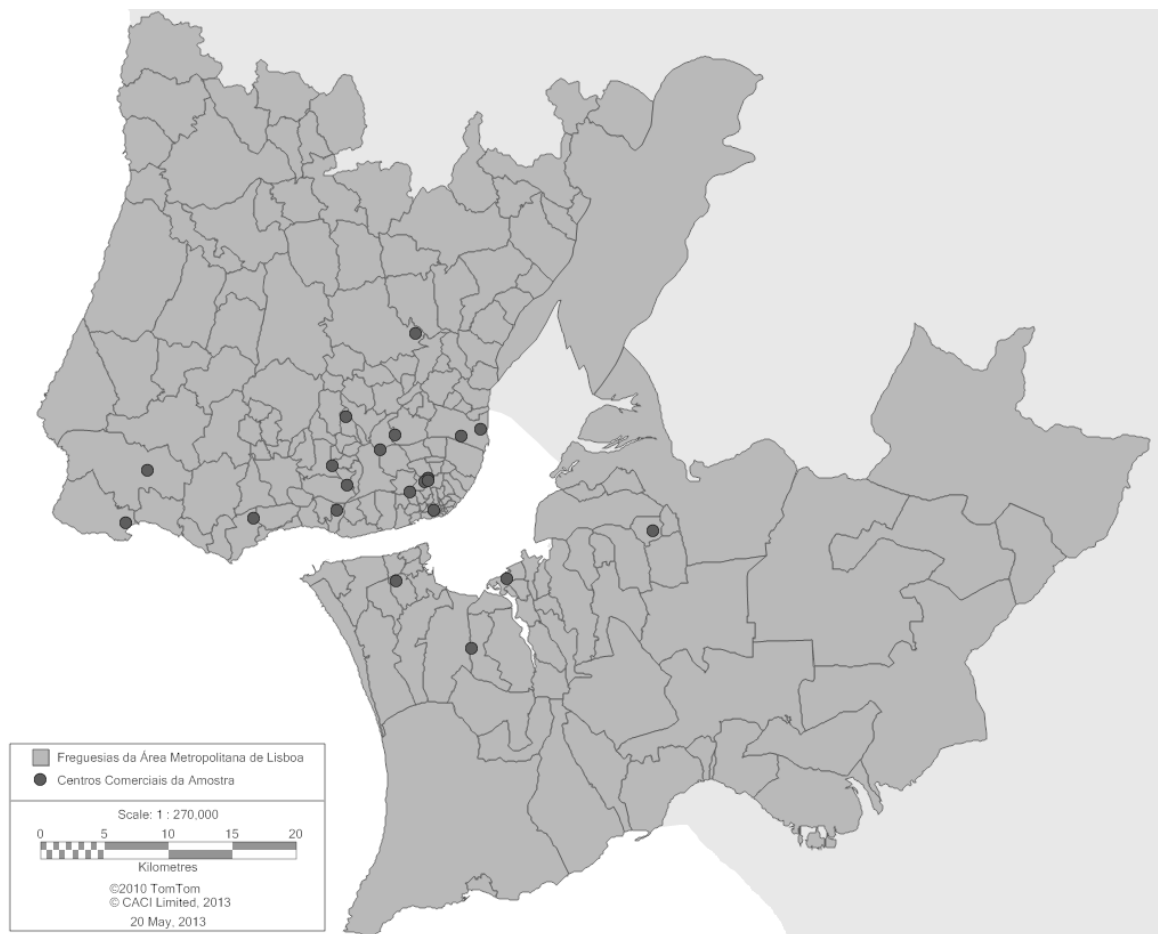


Figura 2 – Localização dos Centros Comerciais da Amostra  
 Fonte: InSite (Maio de 2013)

TABELA III

ITINERÁRIOS COMPLEMENTARES OU PRINCIPAIS CONCLUÍDOS ENTRE 2006 E 2011 NA AML

Itinerário	Lanço	Extensão (km)	Abertura
IP7	CRIL/Eixo NS – Av. Padre Cruz	4	2007
IC2	Carregado - Alenquer	9,3	2010
IC2	Arruda dos Vinhos – Carregado (A1/A10)	11	2006
IC11	Carregado (IP1) - Benavente	16	2007
IC17	Buraca - Pontinha	3,6	2011
IC16	Lourel – CREL (Belas)	10,9	2009
IC30	A5 (Alcabideche) - Ranholas	8,4	2009
<b>TOTAL</b>		<b>63,2</b>	

Fonte: Instituto de Infraestruturas Rodoviárias

TABELA IV  
DEFINIÇÃO DAS VARIÁVEIS

<b>Variável</b>	<b>Descrição</b>
visit	Número de visitantes do centro comercial.
ljsat	Número de lojas satélite.
ablalim	Área bruta locável (ABL) total ocupada pela loja âncora alimentar (supermercado ou hipermercado).
ablljancsalim	ABL total ocupada pelas lojas âncora, excluindo âncora alimentar.
pop30	População residente na isócrona de 0-30 minutos do centro comercial.
remuntrabfreg	Remuneração média (em euros) na freguesia do centro comercial.
metro10min	Variável <i>dummy</i> que toma o valor 1 caso exista metro até 10 minutos percorridos a pé do centro comercial.
cc10abABL	Número de centros comerciais com ABL inferior na isócrona de 0-10 minutos do centro comercial.
cc10ABL	Número de centros comerciais com ABL igual ou maior na isócrona de 0-10 minutos do centro comercial.
cc1015	Número de centros comerciais na isócrona de 10-15 minutos do centro
dumano	Variável <i>dummy</i> indicando o ano (2006-2011) a que se referem os dados . O ano de 2011 é excluído para evitar perfeita multicolineariedade.

TABELA V  
VARIÁVEIS E DESCRIÇÃO ESTATÍSTICA

Variável	Descrição	Média	Desvio Padrão	Mínimo	Máximo
visit	Nº. de visitantes anuais	8.875.759	6.346.247	1.268.497	26.900.000
a1	= 1 se variáveis de 2006	0,16	0,37	0	1
a2	= 1 se variáveis de 2007	0,16	0,37	0	1
a3	= 1 se variáveis de 2008	0,17	0,38	0	1
a4	= 1 se variáveis de 2009	0,17	0,38	0	1
a5	= 1 se variáveis de 2010	0,18	0,38	0	1
a6	= 1 se variáveis de 2011	0,17	0,38	0	1
ljsat	Nº. de lojas satélite	129,10	87,47	32	384
ablalim	ABL total ocupada pela loja âncora alimentar	10.911,43	9.610,99	0,00	29.489
ablljancsalim	ABL total ocupada pelas restantes lojas âncora	9.724,99	9.812,26	0,00	43.418,28
pop30	População residente até 30 minutos <sup>a</sup>	1.879.520	355.691,20	464.475,90	2.190.051
remuntrabfreg	Remuneração (€) dos trabalhadores na freguesia	1.344,16	350,83	782,89	1.854,92
metro10min	= 1 se existe metro até 10 minutos <sup>b</sup>	0,39	0,49	0	1
cc10abABL	Nº. de centros comerciais c/ ABL inferior até 10 minutos <sup>a</sup>	2,53	2,39	0	9
cc10ABL	Nº. de centros comerciais c/ ABL igual/maior até 10 minutos <sup>a</sup>	2,00	2,34	0	8
cc1015	Nº. de centros comerciais entre 10 a 15 minutos <sup>a</sup>	6,50	4,14	0	13

<sup>a</sup> Percorridos de carro

<sup>b</sup> Percorridos a pé

TABELA VI  
TESTES DE COLINEARIEDADE

<b>Variável</b>	<b>Tolerância</b>	<b>VIF</b>
lnablalim	0,238	4,20
lnremuntrabfreg	0,284	3,52
metro10min	0,294	3,40
ljsat	0,303	3,30
cc10ABL	0,342	2,92
cc1015	0,349	2,86
cc10abABL	0,410	2,44
lnabljanccsalim	0,434	2,30
lnpop30	0,535	1,87
a1	0,574	1,74
a5	0,579	1,73
a2	0,583	1,72
a3	0,586	1,71
a4	0,589	1,70
Média		2,53

## ANEXO 2

### 1. TESTES DOS EFEITOS ESPACIAIS

#### 1.1. Teste de Moran $I$

Para testar a existência de autocorrelação espacial, foi inicialmente calculado<sup>29</sup> o teste de Moran  $I$  global aos valores da variável dependente, o logaritmo do número de visitantes anuais dos centros comerciais na AML. Este teste é dado por:

$$(7) \quad I = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} Z_i Z_j}{S_0 m_2}$$

onde:  $N$  é o número de observações;  $w_{ij}$  corresponde aos elementos da matriz de distâncias  $W$  que expressam a interação espacial entre a variável  $i$  e a variável  $j$ ;  $Z_i = Y_i - \bar{Y}$ , com  $Y_i$  a representar o valor da variável  $Y$  na localização  $i$  e  $\bar{Y}$  a média dessa variável;  $S_0 = \sum_i \sum_j w_{ij}$ ; e  $m_2 = \sum_i Z_i^2 / N$ .

Com a hipótese nula de que não existe autocorrelação espacial, o valor esperado de  $I$  é  $E(I) = -1/(N - 1)$ , o que neste caso equivale a -0,01. Caso  $I$  seja maior do que o seu valor esperado, então a distribuição global da variável  $Y$  pode ser caracterizada por autocorrelação espacial positiva, o que significa que os valores tomados por  $Y$  em cada localização  $i$  tendem a ser semelhantes aos valores tomados por  $Y$  nas localizações consideradas como vizinhas. Por oposição, se  $I$  for menor do que o seu valor esperado, a distribuição global da variável  $Y$  pode ser caracterizada por autocorrelação espacial

---

<sup>29</sup> Com recurso ao programa de econometria informático Stata.



negativa, o que significa que os valores tomados por  $Y$  em cada localização  $i$  tendem a ser diferentes dos valores tomados por  $Y$  nas localizações consideradas vizinhas.

Posteriormente, e após ser efetuada a regressão linear do modelo base foi efetuado o teste de Moran  $I$  aos resíduos. Formalmente, esta estatística  $I$  é dada por:

$$(8) \quad I = [N/S_0]\{[\varepsilon'W\varepsilon]/\varepsilon'\varepsilon\}$$

onde:  $N$  é o número de observações;  $S_0 = \sum_i \sum_j w_{ij}$ ;  $\varepsilon$  o vetor dos resíduos da regressão linear; e  $W$  a matriz de distâncias. A hipótese nula é de que não existe autocorrelação espacial.

TABELA I

TESTES DE MORAN I AO MODELO BASE

	<b>Matriz Distância (1/d<sup>2</sup>) em Minutos do Centro Comercial</b>	<b>Matriz Distância (1/d) em Quilómetros da Freguesia</b>	<b>Matriz Distância (1/d<sup>2</sup>) em Minutos da Freguesia</b>
Moran $I$ Global			
$I$	1,42	-0,10	0,03
$E(I)$	-0,01	-0,01	-0,01
$z_I$	24,21 (0,00)	-8,44 (0,00)	2,37 (0,01)
Moran $I$ aos resíduos			
$z_I$	-2,34 (1,98)	-0,40 (0,69)	-14,84 (2,00)

Nota:  $z_I = (I - E(I))/sd(I)$  é a estatística de Moran  $I$  estimada, seguindo uma distribuição assintoticamente normal, cuja hipótese nula é a não existência de autocorrelação espacial.  $p$ -value em parêntesis.

Os resultados dos testes de Moran  $I$  são apresentados na tabela I, sendo as respetivas matrizes usadas identificadas. Relativamente ao teste de Moran  $I$  global, todos os testes levam à mesma conclusão, de que a hipótese nula de não existência de autocorrelação

espacial não deve ser aceite. No entanto, os valores de  $I$  levam a interpretações diferentes consoante a matriz usada, pelo que enquanto as que usam a distância em minutos evidenciam a presença de autocorrelação espacial positiva, a terceira aponta para a presença de autocorrelação espacial negativa. Quanto ao teste de Moran  $I$  aos resíduos, nenhum teste rejeita a hipótese nula de não existência de autocorrelação espacial, o que corrobora as expectativas *a priori* de que neste modelo há evidências para a estimativa de um modelo em *spatial lag*.

### 1.2. Teste do Multiplicador de Lagrange (LM)

Quando os testes de Moran  $I$  não aceitam a hipótese nula de não existência de efeitos espaciais, são geralmente efetuados os testes do Multiplicador de Lagrange (LM). A estatística LM-error testa a ausência de erros sistemáticos e correlacionados através do espaço (ausência de autocorrelação espacial através dos erros), ou seja, a hipótese nula é a de que não existe autocorrelação espacial ( $\lambda = 0$ ), dada a assunção de que não há autocorrelação espacial espacialmente desfasada ( $\rho = 0$ ). Se esta hipótese não é aceite, então há evidências de existência de autocorrelação espacial ( $\lambda \neq 0$ ), pelo que se deverá estimar um modelo de *spatial error* (4).

A estatística LM-lag testa a ausência da influência de variáveis vizinhas entre si (ausência de autocorrelação espacial desfasada), pelo que neste caso a hipótese nula é a de não existência de autocorrelação espacial ( $\rho = 0$ ), dada a assunção de que não há autocorrelação espacial dos erros ( $\lambda = 0$ ). Se esta hipótese não é aceite, temos evidências de existência de autocorrelação espacial ( $\rho \neq 0$ ), pelo que devemos estimar um modelo de *spatial lag* (2).

Em alternativa, ou como complemento a estes testes, procede-se ao cálculo dos testes robustos destas estatísticas, cujos procedimentos são idênticos aos testes descritos acima. A diferença é que neste caso o LM-error robusto (RLM-error) corrige a presença local de dependência espacial desfasada ( $\rho \neq 0$ ), assim como o LM-lag robusto (RLM-lag) corrige a presença local de dependência espacial dos erros ( $\lambda \neq 0$ ). Estes testes são assintoticamente distribuídos por  $\chi^2(1)$ .

Os resultados da tabela II evidenciam a presença de autocorrelação espacial desfasada e dos erros.

TABELA II  
TESTES DO MULTIPLICADOR DE LAGRANGE AO MODELO BASE

	<b>Matriz Distância (1/d<sup>2</sup>) em Minutos do Centro Comercial</b>	<b>Matriz Distância (1/d) em Quilómetros da Freguesia</b>	<b>Matriz Distância (1/d<sup>2</sup>) em Minutos da Freguesia</b>
LM-error	29,50 (0,00)	8,63 (0,00)	7,69 (0,01)
LM-lag	11,82 (0,00)	5,86 (0,02)	5,05 (0,03)
RLM-error	28,45 (0,00)	9,02 (0,00)	8,01 (0,01)
RLM-lag	10,77 (0,00)	6,25 (0,01)	5,36 (0,02)
$\rho$			0,058 (0,02)
$\lambda$			0,116 (0,04)

Nota: Os testes LM são assintoticamente distribuídos por  $\chi^2(1)$ , com um valor crítico de 3,84 a um nível de 5%. *P-values* em parêntesis.  $\rho$  e  $\lambda$  correspondem aos parâmetros estimados nos modelos de *spatial lag* e *spatial error* respetivamente. *Z-values* em parêntesis neste caso.

Desta forma, os testes apontam para a estimativa de modelos de *spatial lag*. A escolha entre estes, consoante as matrizes de distância  $W$  usadas, pode ser efetuada através da comparação do correspondente valor do Critério de Informação de Akaike (*AIC*):

$$(9) \quad AIC = 2k - 2 \ln(L)$$

onde:  $k$  é o número de parâmetros estimados; e  $L$  representa o valor maximizado da função de verosimilhança do modelo estimado. Escolhe-se o modelo que apresentar um menor *AIC*, o que neste caso corresponde ao modelo usando como matriz a distância inversa ao quadrado em minutos do centróide das freguesias dos centros comerciais, cujo *AIC* do modelo *spatial lag* é de 67,24.

## 2. EFEITOS DIRETOS E INDIRETOS

No modelo de *spatial lag*, conforme demonstrado na equação (3), a interpretação dos resultados obtidos difere do modelo base na medida em que neste caso cada  $\beta_k$  resulta da variação estimada da variável  $x_k$  e dos efeitos induzidos pela variação desta variável nas restantes localizações. Desta forma, no modelo base há um enviesamento dos coeficientes estimados, os quais são causados pela dependência espacial presente, neste caso, representada por  $\rho$ . Uma vez que a comparação direta entre os coeficientes estimados no modelo base e no modelo de *spatial lag* não pode ser efetuada, será aplicada a metodologia apresentada por LeSage e Pace (2009).

Segundo os autores, num modelo de *spatial lag* a variação de uma variável na região  $i$  irá afetar essa mesma região (efeitos diretos) e potencialmente afetar as regiões  $i \neq j$  indiretamente (efeitos indiretos ou *spillovers*). Considerando  $S(W) = (I - \rho W)^{-1} X\beta$ , os

elementos da diagonal principal de  $S(W)$  com as derivadas parciais representam os efeitos diretos e os elementos fora da diagonal principal representam os efeitos indiretos. O efeito total da variação de qualquer das variáveis explicativas resulta da soma dos dois efeitos.

Uma vez que o impacto da variação de uma variável explicativa difere entre todas as observações, LeSage e Pace (2009) sugerem que se resuma o impacto das variações, calculando-se a média dos elementos da diagonal principal de  $S(W)$  (média dos efeitos diretos) e média das somas dos elementos não diagonais em cada uma das linhas ou colunas de  $S(W)$  (média dos efeitos indiretos ou *spillovers*). A soma dos dois efeitos corresponde à média dos efeitos totais. No caso deste trabalho, podemos definir o efeito direto médio (*EDM*) e o efeito total médio (*ETM*) como:

$$(10) \quad EDM_j = n^{-1} \text{traço}(S_j(W))$$

$$(11) \quad ETM_j = n^{-1} \sum_{i=1}^n \sum_{j=1}^n S_j(W)_{ij}$$