



UNIVERSIDADE CATÓLICA PORTUGUESA

O Impacto da Taxa Turística no Norte de Portugal

Trabalho Finalna modalidade de Dissertação
apresentado à Universidade Católica Portuguesa
para obtenção do grau de mestre em Business Economics

por

Ana Cláudia Barbosa Leão da Costa

sob orientação de
Prof. Dr. Ricardo Ribeiro

Universidade Católica do Porto, Católica Porto Business School
Março 2016

Agradecimentos

Agradeço, com profundo reconhecimento, ao Senhor Professor Doutor Ricardo Ribeiro a generosidade com que assumiu a orientação desta dissertação e todo o incentivo que recebi da sua parte, ao longo da elaboração da mesma.

Resumo

Esta tese procura quantificar os *trade-offs* resultantes da aplicação de uma eventual taxa turística nos estabelecimentos hoteleiros da região Norte de Portugal, nomeadamente no que diz respeito ao impacto no número de dormidas e nas receitas das entidades hoteleiras, bem como ao imposto gerado pela taxa.

Esta quantificação assenta na estimação de um modelo de procura para o mercado hoteleiro e, nesse sentido, foram consideradas duas especificações de um modelo de escolha discreta logit: o multinomial logit e o nested multinomial logit, ambos desenvolvidos por McFadden (1978, 1981).

Estas duas especificações foram estimadas com dados mensais referentes ao mercado hoteleiro de Portugal Continental, no período de 2013 a 2015, que se segmenta em 5 regiões (Alentejo, Algarve, Centro, Lisboa e Norte), compostas por 5 tipologias de estabelecimentos hoteleiros (hotéis de 1* e 2*, 3*, 4*, 5* e hotéis-apartamentos) e uma opção externa (referente às restantes opções).

Os resultados sugerem que os consumidores são sensíveis ao preço e que a implementação da taxa turística implica, em média, uma diminuição de 12,1% no número de dormidas mensais de cada tipologia. No agregado estima-se uma diminuição de 48.207 dormidas e 1.660.549,75 € na receita total. Por outro lado, e contrabalançando, a implementação da taxa turística gera um imposto total de 350.298,96 € por mês para o total das autarquias pertencentes à região.

Palavras-chave: Taxa Turística; Modelos de Estimação da Procura; Multinomial Logit; Nested Multinomial Logit.

Abstract

This thesis aims to quantify the resulting trade-offs of the application of a tourist tax on hotel establishments in the northern region of Portugal, mainly the impact on the number of stays (per night) and on the revenues of the hotels, as well as the income generated from the tax.

This quantification is based on the estimation of a demand model, for the hotel market and, for this purpose, were considered two specifications of logit discrete choice model: Multinomial Logit (MNL) and Nested Multinomial Logit (NMNL), both developed by McFadden (1978, 1981).

These two specifications were estimated with monthly data for the hotel market of Continental Portugal, from 2013 to 2015, which is segmented into 5 groups per region (Alentejo, Algarve, Centro, Lisboa and Norte), which in turn are composed by 5 typologies of hotel establishments (1* and 2* hotels, 3*, 4*, 5* and hotel-apartments) and an outside option (referring to other options).

The results suggest that consumers are price sensitive and that the implementation of a tourist tax implies, on average, a decrease of 12,1% of the number of monthly stays (per night), in each typology. When aggregated, it is possible to estimate a decrease of 48.207 stays (per night) and 1.660.549,75€ in total revenue. On the other hand, and countering the previously stated negative effects, the implementation of a tourist tax generates a total income of 350.298,96€ per month for the total autarkies in the region.

Key-words: Tourist Tax; Demand Estimation Models; Multinomial Logit; Nested Multinomial Logit.

Índice

Agradecimentos	iii
Resumo	v
Abstract	vii
Índice	ix
Índice de Tabelas e Gráficos	xii
Introdução.....	14
Capítulo 1	18
1. REVISÃO DE LITERATURA	18
1.1 Modelos de Estimação da Procura	18
1.1.1 Modelos Contínuos	21
1.1.2 Modelos Discretos	22
1.1.3 Extensões	26
1.2 O Mercado Hoteleiro.....	28
1.2.1 Modelos, Variáveis e Características.....	28
1.1.2 Resultados	34
Capítulo 2	36
2. MODELO TEÓRICO	36
2.1 Variáveis e Pressupostos.....	36
2.2 Método de Estimação	40
Capítulo 3	45
3. DADOS.....	45
3.1 Estatísticas Descritivas	46
Capítulo 4.....	50
4. ANÁLISE PRELIMINAR.....	50

Capítulo 5.....	52
5. RESULTADOS DE ESTIMAÇÃO.....	52
Capítulo 6.....	59
6. O IMPACTO DA TAXA TURÍSTICA	59
6.1 Impacto na Procura.....	59
6.2 Impacto nos Estabelecimentos Hoteleiros.....	61
6.3 Imposto gerado pela Taxa Turística.....	63
CONCLUSÃO	64
Bibliografia.....	67

Índice de Tabelas e Gráficos

Tabela 1. Variáveis Conhecidas	47
Tabela 2. Variáveis Compostas	48
Tabela 3. Variáveis Instrumentais	49
Tabela 4. Resultados por Tipologia	51
Tabela 5. Resultados dos Modelos Logita	55
Tabela 6. Resultados MNL: VI	57
Tabela 7. Preço e Quota de Mercado por Tipologia.....	59
Tabela 8. Elasticidade-Preço por Tipologia (%).....	60
Tabela 9. Impacto da Taxa Turística por Tipologia.....	62
Tabela 10. Imposto obtido por Tipologia	63
Gráfico 1. Dormidas Vendidas.....	50

Introdução

Atualmente, Portugal é um país acentuadamente vocacionado para o turismo, sendo este um setor indubitavelmente relevante, tendo em conta o seu largo contributo para o desenvolvimento da economia nacional, através da geração de riqueza e criação de emprego. Assim, a atividade turística reflete-se com grande impacto na composição do PIB português.

Nos últimos anos, a relevância da região norte do país tem vindo a destacar-se no que concerne ao setor do turismo, tendo sido em Portugal Continental a zona que mais cresceu, em termos de estadias/dormidas, no ano de 2015. Dado o elevado potencial turístico desta região, está no centro das questões públicas o debate atinente à implementação de uma taxa turística nos estabelecimentos hoteleiros situados na zona descrita, factualidade esta que originou a presente dissertação.

Esta tese ambiciona avaliar de que forma a taxa turística proposta influencia o processo de procura dos visitantes, bem como quantificar os consequentes *trade-offs*, de curto prazo, nomeadamente no que diz respeito ao impacto no número de dormidas e nas receitas das entidades hoteleiras e ao imposto gerado, na eventual implementação da referida taxa. Através deste projeto, pretende-se uma aplicação prática concreta para as organizações do turismo, permitindo-lhes quantificar o imposto obtido com esta política, determinando, em última análise, a sustentabilidade da implementação da referida taxa.

Por outro lado, e não estando na livre disponibilidade das entidades hoteleiras controlar a eventual imposição futura de tal taxa (tal como sucedido no município de Lisboa), este estudo pretende figurar um benefício ao dispor daquelas, que poderão, através dele, prever os seus eventuais efeitos negativos,

bem como arquitetar antecipadamente medidas de reação com vista à atenuação desse possível impacto prejudicial.

Assim, de acordo com o estudo de caso exposto, é considerada a seguinte questão de investigação, à qual esta dissertação pretende responder: *Qual o impacto da nova taxa turística na região norte de Portugal?*

De acordo com a taxa turística em vigor no Município de Lisboa, assume-se que a taxa considerada para a região norte tomará o valor monetário de 1 € por dormida, e que esta irá ser refletida no preço praticado pelos estabelecimentos hoteleiros. Para estimar o impacto deste aumento de preço no mercado hoteleiro em estudo, é proposta a estimação de um modelo de procura.

Os capítulos seguintes desta tese encontram-se organizados da seguinte forma. No capítulo 1, é realizada uma revisão literária. O primeiro subcapítulo aborda dois grupos de modelos de procura dominantes na literatura, diferenciados pela natureza das escolhas, que assumem, que os consumidores fazem (Davis e Garcés, 2010, p. 436): *os modelos de escolha contínua e modelos de escolha discreta*. O segundo subcapítulo aborda a aplicação destes modelos de procura ao mercado hoteleiro, em que se evidencia a utilização de modelos de escolha discreta.

Neste sentido, é desenvolvida no capítulo 2, uma abordagem de resposta à questão de investigação com base em modelos de escolha discreta logit, designadamente o modelo Multinomial Logit (MNL) e Nested Multinomial Logit (NMNL), desenvolvidos por McFadden (1978, 1981).

Apesar de a literatura apresentar modelos com um maior grau de flexibilidade na estimação dos padrões de substituição dos consumidores entre alternativas, a escolha dos modelos *supra* mencionados deve-se, fundamentalmente, à sua facilidade de computação - que não requer o uso de técnicas de simulação e diminui substancialmente as exigências de cálculo.

Após a descrição do modelo teórico, procede-se, no capítulo 3, à descrição do mercado hoteleiro de Portugal Continental e dos dados utilizados. O mercado é

segmentado em 5 regiões (Alentejo, Algarve, Centro, Lisboa e Norte) e definido pela soma da procura servida por 5 tipologias de estabelecimentos hoteleiros (hotéis de 1* e 2*, 3*, 4*, 5* e hotéis-apartamentos) e a opções externas (respeitante às restantes alternativas, que ficarão fora do foco em estudo).

Os dados utilizados no presente estudo referem-se ao período decorrente entre os anos 2013 e 2015, e foram obtidos a partir dos quadros estatísticos publicados pela organização *Turismo de Portugal*. A amostra final é composta por 900 observações. Por sua vez, cada observação corresponde a uma tipologia específica, em determinada região e por determinado período de tempo (mês).

O capítulo 4 apresenta uma análise preliminar, independente da aplicação de modelos de estimação da procura, que analisa a relação entre o número de dormidas vendidas, na região Norte, e o preço médio praticado.

No capítulo 5 são apresentados os resultados de estimação, obtidos da aplicação dos modelos de estimação da procura, referidos anteriormente. Conclui-se que o modelo MNL, estimado através de variáveis instrumentais, seria o ideal para explicar o comportamento dos consumidores no mercado hoteleiro de Portugal. Analisados os resultados obtidos, verifica-se que o potencial consumidor, do mercado em estudo, é sensível ao preço, e estima-se que o aumento do preço em 1%, provoque, em média, uma diminuição da quantidade procurada (número de dormidas) em 4,52%, por tipologia, região e mês, decorrente da elasticidade-preço direta da procura média obtida: 4,52%, em valor absoluto. Adicionalmente, são descritos, neste capítulo, os rankings de preferências dos consumidores, relativos à tipologia, região e mês, que apresentam uma elevada percepção do comportamento dos consumidores no mercado hoteleiro de Portugal.

Por último, é apresentada, no capítulo 6, a análise de impacto da taxa turística na região Norte. Os resultados sugerem que o aumento do preço de 1€ em todas as tipologias da região norte origina, em média, um impacto negativo de 12,1% no número de dormidas por tipologia, por cada mês do ano.

Adicionalmente, verifica-se que a aplicação da taxa turística tem um impacto negativo mais elevado para a tipologia 4* e um menor impacto para a tipologia hotéis-apartamentos. Neste seguimento, no agregado estima-se uma diminuição de, aproximadamente, 48.207 dormidas, assim como uma diminuição das receitas em 1.660.549,75 €.

Por outro lado, a implementação da taxa turística em todas as tipologias gera um imposto total no valor de 350.298,96 €, por mês, para o total das autarquias pertencentes à região. Este imposto, poderá, eventualmente, ser vantajoso, na lógica de que o valor gerado tem aplicação na dinamização da região Norte, isto é, na promoção do turismo, e renovação de material urbano - o sucesso da atividade turística, da região Norte, instala um aumento de pressão sob o espaço urbano e derivados custos.

Capítulo 1

1. REVISÃO DE LITERATURA

No capítulo que ora se inicia, procuramos desenvolver uma base teórica relativamente aos modelos de estimação da procura, tendo como objetivo último a seleção e aplicação do modelo mais adequado ao caso em estudo, sendo realizada a conseqüente interpretação dos resultados obtidos.

Neste sentido, num primeiro subcapítulo será realizada uma revisão dos modelos dominantes na literatura (modelos contínuos e discretos), identificando-se os respetivos pressupostos, vantagens e desvantagens. No segundo subcapítulo, será concretizada uma revisão de literatura quanto aos modelos usualmente aplicados ao âmbito de estudo aqui em causa: o mercado hoteleiro.

1.1 Modelos de Estimação da Procura

Os modelos de estimação da procura são a maior ferramenta utilizada para a análise e previsão de resultados do mercado. Akerberg *et. al* (2007, p. 4174-4175) destacam três grandes vantagens que surgem da utilização destes modelos. Primeiramente, os mesmos revelam-se recursos facilitadores para análises de comparação estática de qualquer alteração no mercado, que não tenha um impacto imediato nos custos (como por exemplo os efeitos de fusões e alterações de preços), permitindo analisar o impacto da alteração da quantidade procurada e receitas das empresas. Em segundo lugar, os modelos de estimação da procura permitem analisar o efeito resultante da alteração no bem-estar dos consumidores. Por último, facilitam ainda a análise de problemas de *input/output*, possibilitando aproximar a rentabilidade de potenciais novos

produtos.

De uma forma geral, um modelo de estimação da procura permite caracterizar as receitas que as empresas irão obter a partir dos seus produtos que, por sua vez, desempenham um papel fundamental na determinação da sua rentabilidade e resultados do mercado.

Visto que as receitas dependem das preferências dos consumidores, a procura do consumidor é, claramente, considerada como um elemento crucial na configuração dos resultados do mercado.

Estimar a procura para um mercado de produtos homogéneos é, em princípio, relativamente simples, especialmente porque existe apenas uma equação de procura de mercado a ser estimada (Davis e Garcés, 2010, p. 437). No entanto, apesar de o pressuposto de bens homogéneos ser conveniente, a verdade é que na prática, este é frequentemente violado pelo facto de a maioria dos produtos serem diferenciados (Davis e Garcés, 2010, p. 446). Por esta razão, este capítulo foca a sua atenção em modelos de estimação da procura para produtos diferenciados.

Segundo Nevo (2000), a abordagem mais direta para estimar a procura para um determinado conjunto de produtos semelhantes mas não idênticos é especificar um sistema com várias equações de procura – uma para cada produto, em função do seu preço, do preço dos outros produtos, e outras variáveis. Um exemplo deste tipo de abordagem é o modelo de despesa linear (Stone, 1954), no qual as quantidades são funções lineares dos preços. Estudos subsequentes preocuparam-se em particularizar a relação entre quantidades e preços, de forma a que esta seja flexível e consistente com a teoria económica.¹ São exemplos os seguintes modelos: *Rotterdam* (Theil, 1965; Barten, 1966),

¹O sentido para a expressão flexível é o de que a forma funcional (relação entre as variáveis dependente e independentes) não impõe restrições prévias que os dados não conseguem ultrapassar. Já a denominação consistente é utilizada no sentido de que esta mantém uma relação com a teoria económica (Davis e Garcés, 2010, p. 119). As formas funcionais flexíveis são utilizadas em econometria pelo facto de permitirem ao analista modelar os efeitos de segunda ordem, como as elasticidades de substituição entre produtos (Greene, 2011, p. 12).

Translog (Christensen et al., 1975) e *Almost Ideal Demand System* (AIDS) (Deaton e Muellbauer, 1980).²

Esta abordagem, apesar de intuitiva, conduz a duas preocupações. A primeira atribui-se à questão da dimensionalidade, devido à elevada quantidade de produtos e conseqüente número de parâmetros a ser estimado (Nevo, 2000). Por exemplo, num mercado que existem 200 produtos diferenciados sem quaisquer restrições adicionais, a abordagem descrita acima implicaria estimar no mínimo 40,000 parâmetros - 200 equações de procura: uma para cada produto, com 200 preços em cada.

De acordo com Davis e Garcés (2010), a segunda preocupação refere-se aos vários problemas empíricos que surgem na estimação de modelos da procura: os preços dos produtos estritamente definidos são, na sua maioria, altamente colineares, o que torna difícil identificar separadamente os efeitos dos preços de cada produto. Ademais, este problema é potencializado pelo facto de os preços estarem geralmente correlacionados com determinantes não observáveis da procura, o que requer uma variável instrumental (VI) para cada preço.³ Sendo que encontrar uma única VI não se revela tarefa fácil, é quase impossível encontrar instrumentos suficientes que sejam exógenos e não gerem condições de momento quase colineares.

No que se refere ao problema da dimensionalidade, uma possível solução seria agregar os produtos individuais em agregados mais gerais. A título de exemplo, no caso do mercado de bebidas é possível agregar todas as águas, todos os refrigerantes, etc. Contudo, esta abordagem é apenas útil para casos em que o foco de análise seja um dos agregados gerais – por exemplo, uma análise de quanto espaço conceder em loja para as águas e refrigerantes. Já se nos reportarmos às situações em que o foco de análise são produtos específicos,

²Citado em Nevo (2000).

³Os requisitos básicos para uma VI são que esta seja correlacionada com o regressor potencialmente endógeno e não correlacionada com a componente não observável da procura (Davis e Garcés, 2010, p. 103).

a literatura apresenta dois tipos de modelos de procura, sendo que a sua distinção advém da natureza das escolhas, que se assume, que os consumidores fazem (Davis e Garcés, 2010, p. 436):

- 1) Modelos de Escolha Contínua: situações em que os consumidores têm que decidir a quantidade a comprar de um determinado produto (digase, a quantidade de refrigerantes com que vão abastecer a sua dispensa durante uma semana);
- 2) Modelos de Escolha Discreta: situações nas quais o consumidor deve optar por uma entre as várias alternativas (produtos) disponíveis (como a compra de um automóvel da marca A, B ou C).

1.1.1 Modelos Contínuos

Seguindo Davis e Garcés (2010), os modelos de escolha contínua, também designados por modelos de procura no espaço de produto, resolvem o problema da dimensionalidade partindo do pressuposto que é possível dividir os produtos em grupos (segmentos) mais pequenos e estimar um sistema de procura dentro de um grupo e entre grupos. Estes modelos são denominados multi-etapas, na medida em que o problema de maximização da utilidade dos consumidores é traduzido numa sequência de problemas de decisão separados mas relacionados.

Hausman *et al.* (1994), que estima a procura para a cerveja e Hausman (1996), que estima da procura para cereais (prontos-a-comer), constituem exemplos desta abordagem. Ambos apresentam uma estrutura similar, constituída por três níveis que admitem uma forma funcional flexível dentro de cada grupo.⁴O nível mais alto capta a procura por categoria, o nível médio

⁴ A utilização de um sistema de procura mais flexível impõe menos restrições nas preferências dos consumidores, permitindo assim obter padrões de elasticidade-preço intrínsecas e cruzadas menos restritos ao nível do segmento, que resultam em estimativas mais realistas (Hausman, 1996).

corresponde à procura por segmentos específicos (por exemplo: cerveja premium ou cereais para crianças) e o nível baixo representa a procura por determinadas marcas dentro de um segmento (como Budweiser ou Cheerios). Em cada uma das etapas identificadas, a decisão de alocação é uma função apenas da despesa total na respetiva etapa e dos índices de preços correspondentes à mesma.

Para solucionar o problema de correlação entre preços e a componente não observável da procura,⁵ são utilizadas técnicas de VI em ambos os artigos. Quanto ao nível mais alto, por exemplo, Hausman (1996) utiliza dados de custos com os diferentes ingredientes, embalagens e mão-de-obra. Para o nível médio e baixo, Hausman *et al.* (1994) e Hausman (1996) propõem a utilização de preços de outros mercados (neste caso cidades) como instrumentos para os preços num determinado mercado (cidade). Contudo, são estes instrumentos de risco.⁶ De acordo com a literatura, a realidade é que não existe uma solução inteiramente persuasiva para este problema.

Apesar de o modelo multi-etapas ser útil para ocasiões em que se pretende delinear a procura para a categoria ou segmentos, tal como outros modelos contínuos, este apresenta também uma grande limitação: o nível mais baixo não pode ser estimado com base em dados com significativa entrada e saída de produtos. Ora, esta desvantagem impossibilita a utilização do modelo em mercados que este facto se verifique (Davis e Sloan, 2001).

1.1.2 Modelos Discretos

⁵ Existem outros fatores (não relacionados com o preço) da componente não observável, que causam a variação da procura e podem levar ao aumento dos preços do mercado, resultando numa correlação entre “choques na procura” e preços (Davis e Garcés, 2010, p. 441). De acordo com Davis e Garcés (2010, p. 449) é necessário, pelo menos, um instrumento para cada produto para o qual o preço seja considerado potencialmente endógeno.

⁶ Considere-se, a título de exemplo, o caso das campanhas publicitárias e promoções - componentes não observáveis mais significativas - quando estas são aplicadas a nível nacional e os efeitos causados nos choques da procura, não sendo, por esse motivo, independentes entre mercados (Davis e Garcés, 2010, p. 450).

Uma metodologia alternativa para a resolução do problema da dimensionalidade é a utilização de modelos de escolha discreta, historicamente baseados nas características dos produtos – espaço de características (Davis e Garcés, 2010, p. 463).

Esta abordagem, proposta por Gorman (1980) e Lancaster (1966), projeta os produtos no conjunto das suas características, através das quais são derivadas as suas relações de substituição (Davis e Garcés, 2010, p. 465).

Segundo Davis e Garcês (2010, p. 463), o fundamento dos modelos de procura de escolha discreta assemelham-se à abordagem usual da maximização da utilidade dos consumidores. A grande diferença é que estes impõe restrições ao conjunto de escolhas dos consumidores. Neste sentido, os consumidores escolhem se compram 1 determinado produto dentro do grupo de produtos disponíveis (opções internas), ou se gastam todos os seus recursos numa outra alternativa externa.

O modelo de procura *multinomial logit* (*MNL*), desenvolvido por McFadden (1978, 1981) é, provavelmente, o modelo de escolha discreta mais popular entre os investigadores de econometria. Destarte, este modelo, para além de permitir um elevado número de dimensões de tipos de consumidor,⁷ apresenta ainda a vantagem de as funções de procura resultantes serem inteiramente analíticas, proporcionando uma análise e estimação relativamente simples (Davis e Garcés, 2010, p. 475).

O *MNL* é geralmente escrito, em termos da quota de mercado (proporção de indivíduos que escolhem um determinado produto) e tem como base dois pressupostos (Davis e Garcés, 2010, p. 475-476): 1) os consumidores apenas diferem nas suas preferências em termos aditivos, sendo estas específicas aos indivíduos e produtos; 2) os diferentes tipos de consumidores distribuem-se

⁷ O tipo de consumidor pode incluir uma ou várias dimensões de heterogeneidade do consumidor (Davis e Garcés, 2010, p. 475). Por exemplo: idade, género, rendimento, etc.

dentro da totalidade da população de forma idêntica e independente, e possuem uma função densidade de probabilidade de valores extremos tipo I. De acordo com McFadden (1981), a escolha cuidadosa da função densidade de probabilidade de valor extremo tipo I, produz uma expressão analítica para todas as funções da procura agregada, constituindo esta a vantagem chave do MNL. Estas funções estabelecem uma relação entre quantidade procurada e utilidade média de todos os produtos no mercado.

Berry (1994), especifica uma forma funcional linear para a utilidade média de cada produto, dividindo as suas características em dois conjuntos: observáveis e não observáveis. Ao contrário dos sistemas de estimação anteriores, esta nova abordagem fornece uma interpretação explícita ao termo erro e integra-o de forma completa no modelo de comportamento do consumidor (Davis e Garcés, p. 479).

Berry (1994) provou que, a partir do MNL, é possível recuperar as características não observáveis de cada produto e estimar-se um modelo de regressão linear simples nos seus parâmetros, o que permite a aplicação das técnicas usuais, incluindo VI (Davis e Garcés, 2010, p. 480).⁸ Berry *et al.* (1995) usam as características (incluindo a propriedade) de todos os outros produtos concorrentes como instrumentos para o preço de um produto em outra cidade.⁹

Apesar da introdução das características não observáveis dos produtos, o MNL mantém uma estrutura de padrões de substituição rígida e indesejada. Apresenta a propriedade de independência de alternativas irrelevantes (IAI), que implica, que uma alteração na composição do conjunto de escolhas não

⁸ O vetor da utilidade média do MNL é determinado apenas pelas quotas de mercado observáveis, que permitem estimar uma equação linear, em que o nível médio de utilidade observado se torna uma variável dependente.

⁹ Os autores defendem que esta abordagem é o resultado de uma particularidade natural dos preços de oligopólio: as características de todos os outros produtos estão correlacionadas com o preço porquanto produtos com características similares são bons substitutos, que tendem a ter *markups* baixos e, consequentemente baixos preços em termos de custos. Da mesma forma, a propriedade é correlacionada com o preço devido ao facto de os *markups* de Nash responderem de forma diferente ao próprio produto e a produtos rivais (Berry *et al.*, 1995).

afeta o rácio das probabilidades de escolha entre quaisquer duas alternativas (Davis e Garcés, 2010, p. 422).¹⁰ Desta forma, as elasticidades-preço diretas e cruzadas são determinadas apenas pela quota de mercado e preço do produto cujo valor se alterou, o que implica que os padrões de substituição entre produtos não dependam das suas características semelhantes (Davis e Garcés, 2010, p. 482). Por este motivo, não é recomendada a utilização do *MNL* em estudos nos quais o padrão de substituição seja essencial para a tomada de decisões quanto à apropriação de políticas.

Como alternativa ao modelo *MNL*, McFadden (1978, 1981) desenvolve o modelo *nested multinomial logit (NMNL)*, que apresenta uma estrutura mais flexível do que o modelo anterior (relaxa o pressuposto de IAI) e retém ainda a sua tratabilidade. Como exemplo da aplicação desta metodologia podemos referir Verboven (1996).¹¹ Assim, de uma forma semelhante aos modelos de Hausman *et al.* (1994) e Hausman (1996), o *NMNL* baseia-se na presunção de que o consumidor escolhe um produto por etapas. Contudo, esta nova abordagem permite estimar a procura para os vários produtos com uma única regressão (Davis e Garcés, 2010, p.484).

Segundo McFadden (1978), o *NMNL* assume que os consumidores realizam o processo de tomada de decisão em duas etapas: em primeiro lugar, escolhem a categoria (grupo) de produtos e, seguidamente, escolhem entre os vários produtos disponíveis dentro do grupo previamente selecionado. Os grupos são constituídos por compilações exaustivas de produtos, sendo também mutuamente exclusivos (Davis e Garcés, p. 484). Tendo por exemplo, Ivaldi e Verboven (2005), estes autores aplicam o modelo *NMNL* ao mercado de

¹⁰ A propriedade de IAI implica que todas as alternativas têm a mesma medida de semelhança, o que, na realidade, é frequentemente violado, visto que apenas algumas alternativas partilham atributos (são semelhantes).

¹¹ Verboven (1996), analisa a discriminação de preços praticada no mercado automóvel e conclui que, através do modelo *NMNL*, as preferências dos indivíduos por um carro específico num determinado mercado estão correlacionadas com outros carros do mesmo grupo e subgrupo. Desta forma, são geradas funções de procura agregada com padrões de substituição mais plausíveis e ainda economicamente tratáveis, permitindo a competição local entre carros do mesmo grupo e subgrupo.

camiónes pesados, sugerindo que este pode ser segmentado nos seguintes grupos: camiónes rígidos (dos quais não é possível separar o semirreboque), camiónes-trator (separáveis) e um terceiro grupo para o produto externo.

De acordo com Davis e Garcés (2010, p. 485), Cardell (1997), demonstra que a distribuição das escolhas preferenciais dos consumidores por um grupo específico depende de um parâmetro que controla a correlação das preferências, dentro dos grupos, sendo restringido a valores entre zero e um. À medida que este parâmetro se aproxima do valor um, aumenta o peso relativo das preferências dos consumidores por um grupo específico e, conseqüentemente a correlação entre preferências por produtos dentro do mesmo grupo. Inversamente, quando se aproxima de zero, haverá uma menor correlação entre as preferências dos consumidores por produtos do mesmo grupo. Desta forma, são gerados padrões de substituição mais realistas: quando o preço de um produto do grupo sobe, os consumidores com maior preferência pelo grupo, têm tendência a fazer substituições para produtos dentro desse mesmo grupo.

Verboven e Brenkers (2006) deixaram o seu contributo demonstrando que o NMNL pode ser estimado com parâmetros de preferência específicos a cada grupo - o parâmetro que controla a correlação de preferências entre grupos torna-se específico a cada grupo. A estimação através destes parâmetros torna os padrões de substituição (entre os produtos dentro de cada grupo) ainda mais flexíveis, visto que estes deixam de ser restringidos a serem os mesmos entre grupos (Davis e Garcés, 2010, p. 487).

1.1.3 Extensões

De acordo com Davis e Ribeiro (2010), os renomados economistas que estudam modelos de escolha discreta, têm vindo a desenvolver modelos com padrões de substituição mais flexíveis através da introdução da heterogeneidade (observável e não observável) dos consumidores. Exemplos

reconhecidos deste tipo de metodologia são o *multinomial probit* (Hausman e Wise, 1978) e o *mixed* ou *random-coefficients multinomial logit* (RC-MNL), que evoluiu a partir de McFadden (1981). O modelo BLP (Berry *et al.*, 1995), é um exemplo reconhecido pela literatura, resultante da aplicação da metodologia de RC-MNL, e que tendo vindo a ser utilizado em inúmeros estudos ao longo dos anos.

No entanto, Berry e Pakes (2007) realçam as limitações dos modelos *logit* (incluindo o modelo BLP) no que concerne ao “choque de preferência” ao nível dos produtos, propondo um novo modelo com base nas características puras dos produtos – *pure characteristic model* (PCM).¹² Song (2010) vai mais longe e propõe um modelo de escolha discreta híbrido – utilizando-se o modelo BLP para descrever a escolha de uma marca e modelo PCM para descrever a escolha de um produto dentro dessa marca.¹³ Outro caso de extensão dos modelos *logit*, também com um maior grau de flexibilidade nos padrões de substituição, é o modelo *generalized nested logit* (GNL), desenvolvido por Wen e Koppelman (2001).¹⁴

A este propósito, impõe-se, contudo, referir que a grande desvantagem dos modelos apresentados nesta secção relaciona-se com o típico *trade-off* entre os requisitos de flexibilidade e de computação. É certo que os modelos proporcionam uma maior flexibilidade, porém, requerem o uso de técnicas de simulação que, por sua vez, aumentam substancialmente as exigências de cálculo, tornando-se assim mais dispendiosos, em termos de tempo.

¹² O modelo descreve a heterogeneidade dos consumidores baseando-se apenas nas diferentes preferências quanto às características individuais dos produtos, sem considerações sobre as preferências por determinados produtos como um todo (por exemplo: preferência de marca).

¹³ Os grupos de produtos são diferenciados horizontalmente, enquanto os produtos dentro de um determinado grupo são diferenciados verticalmente, ou seja, condicionados às características dos produtos. Os dois tipos de diferenciação são distinguidos por um “choque de preferência”, que representa a preferência específica de cada consumidor por um grupo de produtos que não é capturado pelas características (quer observadas quer não observadas) dos produtos.

¹⁴ O GNL acomoda diferentes elasticidades cruzadas entre pares de alternativas através de um parâmetro (*logsum*) que mede quão similares são os produtos dentro de um grupo – sendo que cada grupo tem o seu próprio parâmetro. A combinação dos diferentes parâmetros *logsum* para cada grupo e alocação dos parâmetros das alternativas entre grupos fornece diferentes níveis de correlação entre alternativas dentro e entre grupos.

1.2 O Mercado Hoteleiro

Tal como em inúmeras outras indústrias de produtos ou serviços, a indústria hoteleira é composta por um conjunto de produtos, diga-se, estabelecimentos hoteleiros, diferenciados. Destarte, os estudos que procuram analisar este mercado em concreto recorrem a modelos de estimação da procura que têm por base a escolha do consumidor. Neste subcapítulo serão analisados exemplos de estudos aplicados ao mercado hoteleiro. A primeira secção (1.2.1) procura identificar os modelos de estimação da procura, variáveis e características (dos estabelecimentos), geralmente utilizados. Seguidamente, na secção 1.2.2, são apresentados os resultados obtidos em alguns destes estudos, especificamente as elasticidades-preço diretas e cruzadas.

1.2.1 Modelos, Variáveis e Características

Venkataraman e Kadiyali (2005) procuram demonstrar a capacidade do modelo GNL (Wen e Koppelman, 2000) de modo a capturar a escolha do consumidor sob condições em que não são previamente conhecidos, de forma óbvia, as heurísticas de escolha, e os dados sobre as características dos hotéis mais valorizadas pelos consumidores são incompletos. Nesta senda, o modelo utilizado é uma extensão do modelo GNL, denominado de *aggregate generalized nested logit* (AGNL).¹⁵

Neste estudo, é assumido que a indústria hoteleira caracteriza os hotéis com base no nível de preços, e que os consumidores têm dois pontos de ideias quanto à localização (centro da cidade e aeroporto). Os dados utilizados consistem em dados mensais sobre as vendas de quartos por noite de cada hotel, e a estrutura de agrupamento aplicada tem como base a combinação

¹⁵ O modelo AGNL, utiliza dados agregados mais facilmente disponíveis, considerando ao mesmo tempo a heterogeneidade do consumidor e endogeneidade das variáveis relativas à escolha de hotéis.

localização-preço, com sobreposição de uma estrutura de agrupamento que classifica os hotéis segundo determinados limites de distância dos pontos ideais de localização. As distâncias referidas denotam os atributos de localização, que capturaram as preferências heterogêneas dos consumidores relativas à mesma.

Quanto ao processo de escolha do consumidor, é considerado que este escolhe um hotel dentro do conjunto de hotéis disponíveis numa determinada cidade com base na marca, preço do quarto por noite e localização geográfica. O nível de procura de mercado é definido pela soma da procura gerada pelos hotéis pertencentes a três setores – sendo as marcas dos hotéis caracterizadas por serviço completo, serviço limitado ou extensão de estadia – e a opção externa. Os dados utilizados na descrição dos hotéis contêm informação quanto ao nome, capacidade, receita, taxas médias diárias (TMD),¹⁶ filiação da marca, número de dias que o hotel está aberto por ano e variáveis de custo das empresas locais – informações ao nível da cidade que incluem índices de utilização de eletricidade e serviços *in-house*. As duas últimas séries de dados são utilizadas, conjuntamente com outras variáveis exógenas, como instrumentos para corrigir a endogeneidade do lado da procura.

Ghose *et al.* (2010) desenvolvem um estudo, com o intuito de perceber o significado de excesso de utilidade ou excedente do consumidor na escolha de um hotel específico.¹⁷ Esta abordagem é conduzida com base num modelo de escolha discreta híbrido (Song, 2010), considerando dois níveis de heterogeneidade dos consumidores: diferentes categorias de viajantes (familiar, romance ou negócio) e diferentes características dos hotéis.

Para calcular a utilidade esperada associada a cada hotel é utilizado um modelo de coeficientes aleatórios, de modo a que seja possível identificar a

¹⁶ As TMD referem-se ao preço pré-venda imposto. Tal como em Berry *et al.* (1995), este indicador mede o preço médio/agregado oferecido aos consumidores num mercado específico.

¹⁷ O excedente do consumidor relativo a um hotel com determinado tipo de categoria é definido como a soma da sua utilidade média dividida pela elasticidade-preço média de todos os mercados.

distribuição oculta do peso que os consumidores atribuem a cada característica dos hotéis. Assim, estas características são divididas nas seguintes categorias: dados relativos às transações (preço do quarto, por noite e número de quartos vendidos por noite); localização (número de concorrentes locais, taxa de crime anual da cidade e comodidades externas como a proximidade de restaurantes, praia, transportes públicos, vias rápidas, rio e centro da cidade), serviços (número de estrelas e número de comodidades internas, apreciações extraídas através das classificações dos consumidores: qualidade da cama e comida, staff, comodidades da casa de banho e facilidade de estacionamento) e avaliação de *ratings* (atribuídos pelos consumidores no total e quanto às várias características isoladas referidas) e comentários.

Adicionalmente, os hotéis são distribuídos de acordo com oito tipos de categorias de viagem (familiar, negócios, romance, turística, viagem com crianças, viagem com adultos, aceitação de animais e adaptação para deficientes motores), e é assumido que o consumidor escolhe o hotel com a localização e características que melhor servem o propósito da sua viagem.

A heterogeneidade do propósito de viagem é capturada através da introdução de um “choque de preferência” idiossincrático ao nível da categoria, semelhante ao choque de preferência ao nível do produto, utilizado no modelo BLP.

Cada categoria contém um hotel que maximiza a utilidade do consumidor, denominado como o “melhor hotel”. Para eleger, assim, o melhor hotel, é utilizado o PCM, que apresenta a capacidade de capturar a diferenciação vertical entre produtos da mesma categoria. Consequentemente, é assumido que cada consumidor escolhe a categoria em que a utilidade obtida com o melhor hotel é superior a todas as outras alternativas. O choque ao nível da categoria é independente e distribuído de forma idêntica pelos vários consumidores e categorias – consistente com o modelo híbrido de Song (2010).

Por último, o mercado é definido como a combinação “cidade-noite”. Assim, a quota de mercado para cada hotel é calculada com base nas receitas obtidas pelo hotel na respectiva cidade e noite, sendo posteriormente divididas pelo total da receita gerada por todos os hotéis naquela cidade, naquela noite.¹⁸

Anderson e Xie (2011) aplicam um estudo de gestão de receitas ao mercado hoteleiro com dados obtidos através das transações diárias (venda de quartos de hotel) que ocorrem através de um mecanismo de venda opaca,¹⁹ com o intuito de obter conhecimento sobre as preferências dos consumidores na escolha dos hotéis.

Os hotéis são divididos em grupos por vizinhança (localização), e o conjunto de escolha de alternativas consiste em todos os hotéis identificados como um par dos atributos vizinhança e nível de estrelas. Segundo estes autores, “a comparação entre um hotel de 3 e 4 estrelas na mesma vizinhança é um *trade-off* diferente quando comparado com a escolha entre dois hotéis de 3 estrelas em vizinhanças diferentes”. Por conseguinte, é utilizado o modelo de escolha *NMNL*, que relaxa o pressuposto de independência de alternativas irrelevantes (IAI) agrupando alternativas semelhantes.

O processo de escolha do consumidor é constituído por dois níveis: escolha da vizinhança, seguida pelo nível de estrelas do hotel. A utilidade aleatória de um consumidor reservar um hotel específico, com determinado nível de estrelas em determinada vizinhança, é gerada pelo vetor de atributos dos hotéis e respectiva preferência dos consumidores pelos mesmos. Consequentemente, o vetor das características dos hotéis é composto pelo preço da alternativa

¹⁸ No contexto em estudo, a função da quota de mercado de um hotel dentro de uma categoria é semelhante à probabilidade de escolha no modelo BLP.

¹⁹ Fixação de preços em que determinadas características dos produtos não são observáveis pelo consumidor, permitindo às empresas vender os seus produtos diferenciados a preços mais altos para clientes leais à sua marca e, simultaneamente, praticar preços mais baixos para aqueles que não são (Anderson e Xie, 2011).

(vizinhança, estrela) e variáveis que indicam o tipo de qualidade dos hotéis (nº de estrelas).

Li *et al.* (2011) procuram desenvolver um sistema de ranking de hotéis que permita identificar aqueles que geram maior excedente para um consumidor, sendo os dados obtidos através de um sistema de reservas de hotéis que permite calcular a quota de mercado de cada hotel no mercado local (por exemplo, numa dada área metropolitana), para cada dia. Para a concretização desta abordagem, os autores recorrem ao modelo de estimação da procura de coeficiente aleatório, na linha de BLP (1995), tendo em consideração a maximização da utilidade por parte dos consumidores, aspetos demográficos e a heterogeneidade das preferências dos consumidores, e a variação de preços dos produtos.

Os tipos de consumidor são representados por um vetor que especifica a demografia de cada consumidor: propósito da viagem (familiar, romance, negócios, amigável) e idade. Neste sentido, o modelo assume que os coeficientes que medem o impacto do preço e preferências dos consumidores por determinadas características são específicos às demografias dos consumidores.

Para a caracterização dos hotéis são consideradas as seguintes variáveis: características externas da localização do hotel (taxa de crime da localização, concorrentes, e proximidade de lojas, bares, transportes, praias, lagos, centro da cidade e vias rápidas) e características do próprio hotel (nível de estrelas, *rating* total e das características internas,²⁰ e revisão de comentários).

Lederman *et al.* (2014) desenvolvem um estudo com o objetivo de permitir identificar os principais concorrentes (hotéis substitutos) num mercado com

²⁰ Designadamente, quanto ao quarto, localização, limpeza, serviço, check-in, centro de negócios, piscina, spa, etc.

oferta fixa de produtos. Esta metodologia é desenvolvida com base num modelo de escolha discreta RC-MNL,²¹ tendo em conta a identificação de segmentos de consumidores com preferências heterogéneas.

Os *inputs* considerados para o modelo estudado são as características e totais de ocupação diários dos hotéis,²² e uma lista de eventos com impacto na procura do mercado em estudo. Através destes *inputs*, são identificados cinco segmentos latentes de consumidores, correspondendo a uma localização atrativa específica e associado a um tipo de evento de negócios ou lazer. Os tipos de consumidores são divididos por propósito de viagem (negócios/lazer), e são também utilizadas como fonte de variação do tamanho dos segmentos as diferenças nos padrões de procura sazonal entre os tipos de viajantes, principalmente quanto aos fins-de-semana *vs.* dias de semana.

A variável localização é dividida em 3 regiões ideais (central, centro de convenções e *foggy bottom/Georgetown*) e são utilizadas noções de distância dos pontos de procura idealizados que se diferenciam entre segmentos e introduzem a heterogeneidade nas preferências relativas à localização.

O conjunto de produtos é constituído por 21 hotéis em Washington D.C. e a opção de escolher nenhum hotel desse mesmo conjunto. Adicionalmente, cada hotel é descrito por um conjunto de métricas de pontuação para o espaço disponível para reuniões, pela variedade de opções de quartos e comodidades nele oferecidas, e por uma localização única, relacionada com a distância dos 3 pontos potenciais de procura.

Posto isto, e tendo em conta o exposto, realizada uma análise dos artigos presentes nesta secção, verifica-se que são geralmente aplicados modelos de escolha discreta no estudo do mercado hoteleiro.

²¹ Ao contrário dos modelos de coeficiente aleatório geralmente utilizados (BLP, 1995), o modelo aplicado neste estudo utiliza segmentos discretos pré-estabelecidos em que é possível o seu tamanho variar ao longo do tempo.

²² A data de ocupação é recolhida ao nível diário e posteriormente agregada ao nível de meias-semanas (fim-de-semana *vs.* semana). São também incluídos os preços para os quais foram calculados a taxa média diária (receita média obtida por quarto ocupado durante metade da semana).

1.1.2 Resultados

Nesta secção, é feita uma análise baseada única e exclusivamente nos artigos da secção anterior, que se reportam os resultados de estimação referentes à elasticidade-preço.

De acordo com Anderson e Xie (2011) e Venkataraman e Kadiyali (2005), os estudos realizados para o mercado hoteleiro de Washington DC e Austin, respetivamente, permitem obter resultados consistentes quanto às elasticidades-preço, direta e cruzada,²³ a partir das quais é possível obter uma medida de sensibilidade dos consumidores face ao preço dos estabelecimentos hoteleiros.

Anderson e Xie (2011) obtêm uma elasticidade-preço direta média de -3,71% e uma elasticidade-preço cruzada média de 0,66% através da aplicação do modelo NMNL. Com base nestes autores, das diversas elasticidades-preço relativas a cada classe de hotel, tendo em conta o número de estrelas, retiram-se as seguintes conclusões:

- 1) Os consumidores são sensíveis ao preço, e quanto maior o nível de estrelas dos hotéis, maior é a elasticidade da procura (maior elasticidade-preço direta, em valor absoluto);
- 2) Os hotéis são considerados produtos substitutos tendo em conta o valor positivo de todas as elasticidades-preço cruzadas, e o hotel com o nível de estrelas mais elevado tem o maior impacto na probabilidade de escolha das restantes alternativas (maior efeito na elasticidade preço-

²³ A elasticidade-preço direta mede a variação percentual da quantidade procurada de um produto, dada uma variação de 1%, no seu preço. De acordo com a lei da procura, o aumento do preço de um produto j origina uma redução da quantidade procurada do produto em causa, e vice-versa, o que se traduz no valor negativo da elasticidade-preço direta (Davis e Garcés, 2010, p. 7). A elasticidade preço-cruzada expressa o efeito da alteração do preço de um outro produto k em 1%, na quantidade procurada do produto j . Esta elasticidade apresenta um valor positivo quando os produtos são considerados substitutos, isto é, o aumento do preço do produto k incentiva os consumidores a diminuírem a quantidade procurada do produto k , e a aumentarem a quantidade procurada do produto j . No caso em que os produtos são considerados complementares verifica-se uma diminuição da quantidade procurada de k e j (Davis e Garcés, 2010, p. 8).

cruzada).

Segundo Venkataraman e Kadiyali (2005), obtêm através da estimação do modelo de agregação AGNL valores mais altos: -4,22% (média da elasticidade-preço direta) e 0,993% (média da elasticidade-preço cruzada). Ora, mais uma vez, verifica-se que os consumidores são sensíveis ao preço e que os hotéis são considerados produtos substitutos. Assim, através das elasticidades-preço mencionadas, pretende-se obter um termo de comparação para uma posterior discussão dos resultados obtidos nesta tese.

Capítulo 2

2. MODELO TEÓRICO

Neste capítulo, é apresentado o modelo de procura utilizado para o estudo do mercado hoteleiro Português, que tem como base os modelos de escolha discreta *multinomial logit*(MNL) e *nested multinomial logit*(NMNL), desenvolvidos por McFadden (1978, 1981).

A escolha dos modelos é sustentada pelas conclusões retiradas no capítulo anterior. De acordo com a secção 1.2.1, é evidenciada a aplicação de modelos de escolha discreta em estudos sobre o mercado hoteleiro, que se justifica pela possibilidade de entrada e saída de estabelecimentos hoteleiros no mercado e conseqüente impossibilidade de utilização de modelos de escolha contínua. Apesar de terem sido analisados, no capítulo anterior (secção 1.1.3), modelos com um maior grau de flexibilidade no que concerne aos padrões de substituição entre alternativas, a escolha dos modelos logit referidos deve-se, essencialmente, à sua facilidade de computação – que facilita substancialmente as exigências de cálculo e torna o modelo menos dispendioso, em termos de tempo.

2.1 Variáveis e Pressupostos

De acordo com, Ivaldi e Verboven (2005), e tendo em linha de conta o estudo em questão, o consumidor i refere-se ao viajante em Portugal, e a unidade de procura corresponde à reserva de um quarto, de uma determinada entidade hoteleira (estabelecimento), por noite. Assume-se ainda, que existem N potenciais viajantes, que podem escolher entre reservar um dos produtos J (estabelecimentos hoteleiros de Portugal), $j=1, \dots, J$, ou escolher a opção externa 0

(por exemplo: não viajar para Portugal ou optar por outros meios de estadia como amigos ou família).²⁴ Os estabelecimentos são classificados em G grupos ($g = 1, \dots, G$) constituídos por conjuntos exaustivos e mutuamente exclusivos, e um grupo adicional para a opção externa (grupo 0). Neste caso, os estabelecimentos são divididos em grupos de acordo com a sua localização.

Segundo o modelo *nested multinomial logit* (NMNL), partimos do princípio de que o processo de escolha de um viajante se divide em duas fases: em primeiro lugar é escolhida a categoria (grupo), de acordo com a localização dos estabelecimentos, e posteriormente é escolhido um estabelecimento entre as várias alternativas disponíveis no grupo selecionado, com base num conjunto de características inerentes a esses estabelecimentos.

Através de Ivaldi e Verboven (2005), define-se que a utilidade do viajante i , no período de tempo t , resultante da reserva de um estabelecimento j ou escolha da opção externa, é dada por:

$$\begin{aligned} \mu_{ijt} &= \delta_{jt} + \zeta_{igt} + (1 - \sigma)\varepsilon_{ijt} \text{ para } j \in g, \text{ e} \\ \mu_{i0t} &= \zeta_{i0t} + (1 - \sigma)\varepsilon_{i0t} \text{ para } j = 0, \end{aligned} \tag{1}$$

em que t denota um determinado período de tempo que, *in casu*, é um mês específico.²⁵ O primeiro termo da equação, δ_{jt} , representa a utilidade média associada ao estabelecimento j no mês t , comum a todos os consumidores e irá assumir-se que depende do preço do estabelecimento j (p_{jt} - preço médio, por quarto ocupado, para cada estabelecimento j , de acordo com a tipologia do mesmo, no mês t), do vetor x_{jt} das características observáveis do estabelecimento j no mês t (tais como tipologia, nível de estrelas, marca, comodidades internas - piscina, ginásio, spa, etc. - e comodidades externas como, por exemplo, existência de praia, comércio nas redondezas, pontos

²⁴Escolha do turista fora do foco em estudo.

²⁵No caso em análise, são utilizados dados mensais para a estimação do modelo, de acordo com a tipologia dos estabelecimentos hoteleiros.

turísticos, transportes, etc.), e do termo erro ξ_{jt} que reflete as características dos estabelecimentos não observáveis pelo analista, no mês t (nomeadamente a qualidade do staff, limpeza do estabelecimento hoteleiro, qualidade das refeições, serviços de transporte, flexibilidade de check-in e check-out, publicidade, etc.):

$$\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}, \quad (2)$$

em que, α e β , são os parâmetros a serem estimados.

Os termos ζ_{igt} e ε_{ijt} , presentes na expressão (1), são variáveis aleatórias que refletem os desvios individuais dos viajantes da utilidade média. O termo ζ_{igt} denota a utilidade do viajante i no mês t , comum a todos os estabelecimento pertencentes ao grupo g , enquanto o termo ε_{ijt} representa a utilidade do viajante especificamente em relação ao estabelecimento j no mês t .

O parâmetro σ é restringido a valores entre 0 e 1 e mede a correlação da utilidade dos viajantes para os estabelecimentos dentro do mesmo grupo. Se $\sigma=1$, existe uma correlação perfeita de preferências por estabelecimentos, dentro do mesmo grupo (localização), sendo estes percebidos como substitutos perfeitos. Por conseguinte, à medida que σ diminui, diminui também a correlação de preferências para estabelecimentos do mesmo grupo. Já no caso em que, $\sigma=0$, não existe qualquer correlação de preferências: em resposta a um aumento de preços, os viajantes são igualmente propensos a mudar para estabelecimentos do mesmo grupo ou de um grupo diferente. Neste caso o modelo NMNL converte-se no modelo MNL.

$\theta_i = (\zeta_{i0t}, \zeta_{i1t}, \dots, \zeta_{iG}, \varepsilon_{i0t}, \varepsilon_{i1t}, \dots, \varepsilon_{iJt})$ denota o tipo de consumidor de acordo com G grupos pré-especificados, e é assumido que cada viajante i escolhe o estabelecimento j que maximiza a sua utilidade. Para calcular a probabilidade de um viajante i escolher o estabelecimento j , o NMNL assume que as variáveis aleatórias ζ_{igt} e ε_{ijt} , têm uma distribuição que de forma a que ζ_{igt} e $\zeta_{igt} + (1 -$

$\sigma)\varepsilon_{ijt}$ tenham uma distribuição de valor extremo tipo I.²⁶ De acordo com este pressuposto, e seguindo Berry (1994), a quota de mercado de um estabelecimento $j \in g$, no mês t , é expressa como uma fração do total da quota de mercado do grupo g , no mês t :

$$s_{j|gt}(\delta_t) = \frac{e^{\delta_{jt}/(1-\sigma)}}{D_{gt}} = \frac{e^{\delta_{jt}/(1-\sigma)}}{\sum_{j \in \varphi_g} e^{\delta_{jt}/(1-\sigma)}} \quad (3)$$

em que, φ_g refere-se ao conjunto de estabelecimentos pertencentes ao grupo g e, da mesma forma, a quota de mercado de um grupo g no mês t é expressa como uma fração do total da quota de mercado de todos os grupos, no mês t :

$$s_{gt}(\delta_t) = \frac{D_{gt}^{(1-\sigma)}}{\left[\sum_g D_{gt}^{(1-\sigma)} \right]}. \quad (4)$$

Através da combinação das equações (3) e (4) obtém-se a quota de mercado total para cada estabelecimento j no mês t :

$$s_{jt}(\delta_t) = s_{j|gt}(\delta_t) s_{gt}(\delta_t) = \frac{e^{\delta_{jt}/(1-\sigma)}}{D_{gt}^\sigma \left[\sum_g D_{gt}^{(1-\sigma)} \right]}. \quad (5)$$

bem como a quota de mercado total para a opção externa, no mês t :

$$s_{0t}(\delta_t) = \frac{1}{\left[\sum_g D_{gt}^{(1-\sigma)} \right]}. \quad (6)$$

na medida em que $\delta_{0t} \equiv 0$ (por normalização) e conseqüentemente $D_{0t} = 1$.

²⁶ Nos modelos *multinomial logit*, os erros das variáveis aleatórias (resultantes da comparação/diferença entre as utilidades obtidas por cada hotel j e opção externa) seguem uma distribuição Gumbel (valor extremo tipo I). Nesta esteira, a diferença entre erros (que seguem uma distribuição generalizada de valor extremo tipo I) tem uma distribuição logística, que se assemelha à distribuição normal. Conseqüentemente, é assumido que os erros seguem uma distribuição muito próxima da normal, que permite produzir uma expressão analítica para todas as funções de procura agregada.

2.2 Método de Estimação

Para estimar o modelo NMNL são necessários os seguintes dados: quotas de mercado (ou quantidades), preços, e as características atinentes a cada estabelecimento $j > 0$. Isto conduz ao seguinte problema debatido por Berry (1994): existem determinadas características que diferenciam os produtos e conduzem as vendas que são observadas pelos consumidores, mas não pelo analista. Logo, as características dos estabelecimentos, especificadas na função da utilidade média (δ_{jt}), dividem-se em dois grupos: observáveis (x_{jt}) e não observáveis (ξ_{jt}). Neste seguimento, é utilizado o procedimento de transformação proposto por Berry (1994) para estimar o NMNL, através do qual é possível recuperar as características não observáveis de todos os produtos e calcular uma equação linear em todos os seus parâmetros. Dada a normalização do nível de utilidade média da opção externa para 0 (δ_{0t}), o vetor das utilidades comuns entre indivíduos é representado por $\delta_t = (0, \delta_{1t}, \dots, \delta_{jt})$. Assumindo que s_{jt}^* representa a quota de mercado observada do estabelecimento j no mês t , e escolhendo-se δ_t de forma a que a quota de mercado prevista no mês t pelo NMNL seja exatamente igual à quota de mercado real, no mês t , obtém-se a seguinte fórmula:

$$s_{jt} = s_{jt}^* \text{ para } j = 1, \dots, J, \quad (7)$$

Por conseguinte, a quota de mercado do produto externo, no mês t , será:²⁷

$$s_{0t} = s_{0t}^*. \quad (8)$$

Ademais, dividindo as equações de cada estabelecimento j pela equação da opção externa e utilizando as equações, (5) e (6), derivadas acima, obtém-se:

²⁷A soma da quota de mercado da opção interna ($j > 0$) e opção externa ($j = 0$) tem de ser sempre igual a 1.

$$\frac{s_{jt}}{s_{0t}} = \frac{s_{jt}^*}{s_{0t}^*} = \frac{e^{\delta_{jt}/(1-\sigma)}}{D_{gt}^\sigma} \text{ para } j = 1, \dots, J. \quad (9)$$

A partir da equação (9), é derivada uma expressão analítica simples para os níveis de utilidade média, no mês t , aplicando logaritmos (logs) às quotas de mercado:

$$\begin{aligned} \ln s_{jt}(\delta_t) - \ln s_{0t}(\delta_t) &= \delta_{jt}/(1-\sigma) - \sigma \ln(D_{gt}) \\ &= \ln s_{jt}^* - \ln s_{0t}^* \text{ para } j = 1, \dots, J, \end{aligned} \quad (10)$$

a qual depende do valor desconhecido de D_{gt} .

De acordo com a equação (7), as quotas de mercado previstas no mês t para os diferentes grupos e para todas as quotas dentro de um grupo também igualam a quota de mercado real referente ao mês t :

$$s_{gt}(\delta_t) = \frac{D_{gt}^{(1-\sigma)}}{[\sum_g D_{gt}^{(1-\sigma)}]} = s_{gt}^* \text{ para } g = 0, \dots, G, \text{ e} \quad (11)$$

$$s_{j|gt}(\delta_t) = \frac{e^{\delta_{jt}/(1-\sigma)}}{D_{gt}} = s_{j|gt}^* \text{ para cada } j \in g. \quad (12)$$

Aplicando os logaritmos a ambos os lados da equação (12), obtém-se:

$$\ln s_{j|gt}(\delta_t) = \delta_{jt}/(1-\sigma) - \ln D_{gt} = \ln s_{j|gt}^* \text{ para cada } j \in g. \quad (13)$$

Sendo que, a partir desta última, é possível derivar-se o valor de D_{gt} :

$$\ln D_{gt} = \delta_{jt}/(1-\sigma) - \ln s_{j|gt}^*,$$

e substituindo-se este resultado na equação (10), obtém-se:

$$\begin{aligned} \ln s_{jt}(\delta_t) - \ln s_{0t}(\delta_t) &= \delta_{jt} + \sigma \ln s_{j|gt}^* \\ &= \ln s_{jt}^* - \ln s_{0t}^* \text{ para } j = 1, \dots, J. \end{aligned} \quad (14)$$

Como resultado final, o NMNL pode ser estimado através de uma equação de regressão linear nos seus parâmetros, que, por sua vez, são estimados através de variáveis instrumentais (VI):

$$\ln s_{jt}^* - \ln s_{ot}^* = x_{jt}\beta - \alpha p_{jt} + \sigma \ln s_{j|gt}^* + \xi_{jt}. \quad (15)$$

A expressão (15) fornece uma interpretação explícita ao termo erro (ξ_{jt}), integrando-o de modo completo no modelo de comportamento do consumidor. Devido ao facto de existirem características não observáveis dos estabelecimentos que os diferenciam e influenciam as vendas, o termo erro (ξ_{jt}) encontra-se positivamente correlacionado com as variáveis explicativas p_{jt} (pela lógica de que melhores características implicam preços mais caros) e $s_{j|gt}^*$ (estabelecimentos com melhores características atraem mais clientes que, conseqüentemente, aumentam a quota de mercado). Por esta razão, estas variáveis são consideradas endógenas e as estimativas dos parâmetros do modelo que se obtêm com o método dos mínimos quadrados (MQM) será enviesada.

O MQM exige que as variáveis explicativas do modelo não estejam correlacionadas com o termo erro (características não observáveis dos estabelecimentos hoteleiros). De acordo com a literatura analisada, o meio idóneo para ultrapassar este problema é, geralmente, a utilização de VI, que requerem o preenchimento de dois requisitos básicos: 1) o instrumento escolhido tem que estar correlacionado com a variável potencialmente endógena e 2) não pode estar correlacionado com a componente não observável da procura (Davis e Garcés, 2010, p. 102 e 103). Neste caso, será necessário construir-se, pelo menos, duas variáveis instrumentais, uma para cada variável endógena (Davis e Garcés, 2010, p. 102).

Com base no modelo BLP (1995), são utilizadas as características observáveis, de todos os outros estabelecimentos, no mês t , como instrumentos para o preço

de um estabelecimento (p_{jt}), e as características observáveis, de todos os outros estabelecimentos em determinada localização (grupo), no mês t , como instrumentos para o preço de um estabelecimento em outra localização ($s_{j|gt}^*$).

De acordo com Berry *et al.* (1995), as características dos restantes produtos estão correlacionadas com o preço, porquanto produtos com características similares são bons substitutos, que tendem a ter *markups* baixos e, conseqüentemente, baixos preços em termos de custos, estando, deste modo, correlacionadas com as variáveis endógenas (p_{jt} e $s_{j|g}^*$).

Adicionalmente, as características não observáveis de cada estabelecimento j não respondem, em princípio, às características observadas dos estabelecimentos concorrentes, na medida em que os atributos de um estabelecimento não são facilmente alteráveis. Parece plausível afirmar-se, então que as características observáveis dos outros produtos não estão correlacionadas com o termo erro (ξ_{jt}). Ora, são deste modo preenchidos os dois requisitos básicos, que permitem a utilização de uma VI.

Segundo Ivaldi e Verboven (2005), sabe-se que ao nível agregado, a probabilidade de escolha de um estabelecimento j , no mês t , coincide com a quota de mercado do estabelecimentos j , no mês t (s_{jt}). Por esta razão, a quantidade vendida pelo hotel j , no mês t , q_{jt} , é dada por:

$$q_{jrt} = s_{jt}N, \quad (16)$$

em que q_{jt} se refere ao número de dormidas mensal, por quarto, para cada estabelecimento j no mês t , de acordo com a tipologia do estabelecimento.

Refira-se ainda que, neste estudo, as escolhas dos consumidores são simuladas a nível individual, e posteriormente agregadas de forma a gerar o nível de quota de mercado dos estabelecimentos. Adicionalmente, para que o modelo seja consistente com a maximização da utilidade aleatória, o valor de α tem que ser positivo e σ tem que pertencer ao intervalo entre 0 e 1. No caso

em que o valor de σ se aproxima de zero, o modelo NMNL aproxima-se do modelo MNL, em que a correlação entre produtos de um grupo específico se torna zero.

Por último, as elasticidades-preço direta e cruzada, do NMNL, são representadas da seguinte forma:

$$\begin{aligned}\varepsilon_{jj} &= \alpha p_j \left[s_j^* - \frac{1}{1-\sigma} + \frac{\sigma}{1-\sigma} s_{j|g}^* \right], \text{ e} \\ \varepsilon_{jk} &= \alpha p_k \left[s_k^* + \frac{\sigma}{1-\sigma} s_{k|g}^* \right] \text{ se } j, k \in g, \text{ ou} \\ \varepsilon_{jk} &= \alpha p_k s_k^* \text{ se } j \in g, k \notin g.\end{aligned}$$

No caso do modelo MNL, em que $\sigma = 0$, as elasticidades são representadas pelas seguintes expressões:

$$\begin{aligned}\varepsilon_{jj} &= \alpha p_j s_j^*, \text{ e} \\ \varepsilon_{jk} &= \alpha p_k s_k^* \text{ se } j \in g, k \notin g \text{ e } j \in g, k \in g.\end{aligned}$$

Capítulo 3

3. DADOS

Pese embora o modelo teórico apresentado no capítulo anterior e os estudos analisados no subcapítulo 1.2 da revisão de literatura desenvolvam uma abordagem de estimação da procura para o mercado hoteleiro *ao nível de cada entidade hoteleira*, foi necessário, para o caso a que nos propusemos estudar, arquitetar uma abordagem mais agregada. Pelo facto de a generalidade das organizações que detêm dados estatísticos sobre as entidades hoteleiras estarem vinculadas por um termo de confidencialidade com as mesmas, não se afigurou possível a obtenção de dados específicos ao nível de cada entidade hoteleira pertencente à região norte de Portugal, o que, fatalmente, impediu a realização de uma abordagem mais detalhada.

Neste sentido, o modelo de estimação da procura descrito anteriormente é aplicado ao mercado hoteleiro de Portugal Continental, que é, como *supra* referido, fraccionado em 5 regiões (grupos): Alentejo, Algarve, Centro, Lisboa e Norte. Para cada um dos segmentos são consideradas 5 alternativas, que correspondem às 5 tipologias de estabelecimentos hoteleiros com maior peso na quantidade de dormidas vendidas, em Portugal: hotéis de 1* e 2*, 3*, 4*, 5* e hotéis-apartamentos (H-A).

O intervalo de tempo considerado para este efeito varia entre os anos de 2013 e 2015, para os quais existem dados completos e disponíveis, resultando num total de 36 meses. A amostra final é composta por 900 observações (5x5x36), sendo que a cada observação corresponde uma tipologia específica em determinada região e período de tempo (mês).

O consumidor *ié* definido como o indivíduo, de nacionalidade portuguesa ou estrangeira, que viaja em Portugal, e a procura ao nível do mercado é definida

como a soma da procura servida pelas 5 tipologias de estabelecimentos hoteleiros (opções internas: $j > 0$) e opção externa ($j = 0$).²⁸

Sendo assim, a unidade de procura no mercado corresponde à aquisição de uma dormida (quarto por noite), de uma determinada tipologia (j),²⁹ numa região (r) e mês (t) específicos, em que a tipologia e localização (região) são consideradas as características observáveis das opções internas. Para estimar, consistentemente, o modelo descrito no capítulo anterior, são necessários os seguintes dados: quantidades (ou quotas de mercado), preços e características das opções internas, de acordo com a respetiva tipologia, região e mês. Neste sentido, é apresentado na secção que se segue um sumário estatístico das principais variáveis utilizadas, bem como das suas fontes.

3.1 Estatísticas Descritivas

A Tabela 1 (cfr. *infra*) sumariza as estatísticas referentes às variáveis precedentes obtidas através dos quadros estatísticos publicados pela organização *Turismo de Portugal*. A variável “Proveitos” denota as receitas obtidas por tipologia (j), região (r) e mês (t),³⁰ e a variável “Dormidas”, q_{jrt} , alude às quantidades vendidas (número de quartos, por noite) por tipologia, região e mês.³¹ A “Taxa de Ocupação” indica a taxa de ocupação por quarto em Portugal por mês,³² sendo que são apenas englobadas, nesta variável, as seguintes tipologias: hotéis, hotéis-apartamentos e pousadas (de acordo com os dados publicados pela *Turismo de Portugal*). O número total de potenciais consumidores (N), para o mercado em estudo é definido como o número de

²⁸A opção externa refere-se, a situações em que o consumidor opta por não viajar em Portugal ou por outras opções de estadia tal como família ou amigos.

²⁹Ao contrário do modelo anteriormente descrito, a variável j não identifica um estabelecimento específico, mas sim um aglomerado (tipologia).

³⁰Fonte: Turismo de Portugal 2013, 2014 e 2015.

³¹Fonte: Turismo de Portugal 2013, 2014 e 2015.

³²Fonte: Turismo de Portugal 2013, 2014 e 2015.

quartos disponíveis de hotéis, hotéis-apartamentos e pousadas. A variável “Total Dormidas” indica a média total das vendas (dormidas) em Portugal (em todas as tipologias e todas as regiões) entre os anos de 2013 e 2015,³³ em cada período de tempo (mês: t). Reitera-se, ainda, que são englobadas na composição total desta variável somente as tipologias hotéis, hotéis-apartamentos e pousadas.

Tabela 1. Variáveis Conhecidas

Variável	Média	Mediana	Desvio Padrão	Mínimo	Máximo
Proveitos (€): R_{jrt}	3.716.601	1.786.773	5.177.532	50.438	3.1600.000
Dormidas (unidades): q_{jrt}	98.385	58.994	118.385	1.469	671.250
Taxa de ocupação (%): tx_t	56	59	16	29	83
Total Dormidas (unidades): Q_t	3.046.826	3.142.082	1.207.324	1.268.548	5.671.996

A Tabela 2 (cfr. *infra*) descreve as estatísticas concernentes às variáveis compostas a partir das variáveis presentes na tabela 1. A variável p_{jrt} respeita ao preço de cada dormida para cada tipologia, por região e mês, obtida através das variáveis “Proveitos” e “Dormidas”: $p_{jrt} = R_{jrt}/q_{jrt}$. O número total de quartos disponíveis em Portugal por mês, NQ_{jrt} , resulta das variáveis, “Taxa de ocupação” e “Total de Dormidas”: $NQ_t = Q_t/tx_t$. A variável s_{jrt} representa a quota de mercado de cada tipologia por região e mês, gerada a partir do número de quartos vendidos por noite (“Dormidas”) e número de quartos disponíveis: $s_{jrt} = q_{jrt}/NQ_t$. Através do somatório da quota de mercado de cada opção interna pertencente ao mesmo grupo (região) é obtida a quota de mercado de cada região por mês: $s_{rt} = \sum s_{jrt}$, para $j \in r$. A partir das últimas duas séries de variáveis referidas é calculada a quota de mercado de cada tipologia ($j > 0$) no respetivo grupo, por mês: $s_{j|rt} = s_{jrt}/s_{rt}$. A quota de mercado da opção externa ($j = 0$), por região e mês, representada por s_{0rt} , é obtida a partir do

³³Fonte: Turismo de Portugal 2013, 2014 e 2015.

somatório das quotas de mercado das opções internas (5 tipologias: $j > 0$), isto é: $s_{0t} = 1 - \sum s_{jrt}$. Por fim, a variável y_{jrt} , que representa a variável dependente, é alcançada através da diferença entre os logaritmos da quota de mercado de cada tipologia e a opção externa: $y_{rt} = \ln s_{jrt} - \ln s_{0t}$.

Tabela 2. Variáveis Compostas

Variável	Média	Mediana	Desvio Padrão	Mínimo	Máximo
p_{jrt} (€)	35,21	32,17	13,66	13,32	99,52
NQ_t (unidades)	5.240.666	52.49.397	717.659	3.988.564	6.870.727
s_{jrt} (%)	18,12	1,19	2,05	0,03	10,25
s_{rt} (%)	9,06	7,19	6,71	0,79	26,96
$s_{j rt}$ (%)	20,00	19,04	13,19	1,17	47,61
s_{0rt} (%)	54,69	52,16	13,81	30,86	77,76
y_{jrt} (unidades)	-4,02	-3,91	1,33	-7,68	-1,12

Para corrigir a endogeneidade das variáveis, preço e quota de mercado de cada tipologia na região, foram construídos os seguintes instrumentos apresentados na Tabela 3 (cfr. *infra*): z_1 , z_2 e z_3 . O instrumento z_1 ,³⁴ que revela o número de hotéis concorrentes por tipologia, região e mês, foi obtido através dos quadros estatísticos publicados pela organização *Turismo de Portugal*. O instrumento z_2 , produzido a partir da diferença entre total de hotéis disponíveis em Portugal e o instrumento z_1 , representa o número de hotéis concorrentes em Portugal, por mês. Ambas as variáveis instrumentais, z_1 e z_2 , são utilizados como instrumentos da variável preço: p_{jrt} . Finalmente, o instrumento z_3 , que resulta da diferença entre o número de hotéis disponíveis por região e o instrumento z_1 , denota o número de hotéis concorrentes por região. Este último é utilizado para a variável endógena: $s_{j|rt}$ (quota de mercado de cada tipologia na região).

A seleção das variáveis instrumentais expostas alicerça-se no pressuposto de que uma maior oferta (maior concorrência) de estabelecimentos hoteleiros

³⁴Fonte: Turismo de Portugal 2013, 2014 e 2015.

implica, em princípio, uma diminuição no preço praticado ao consumidor (e vice-versa).³⁵ Deste modo, espera-se que as características observáveis dos produtos (tipologias) não estejam correlacionadas com o termo erro (características não observáveis das tipologias), partindo do pressuposto de que estas não são, em princípio, facilmente alteráveis. Ora, parece poder afirmar-se que se encontram preenchidos os dois requisitos básicos para que seja permitida a utilização dos referidos instrumentos.

Tabela 3. Variáveis Instrumentais

Variável	Média	Mediana	Desvio Padrão	Mínimo	Máximo
z1 (unidades)	43	34	33	4	115
z2 (unidades)	1042	1039	52	920	1128
z3 (unidades)	174	177	64	54	298

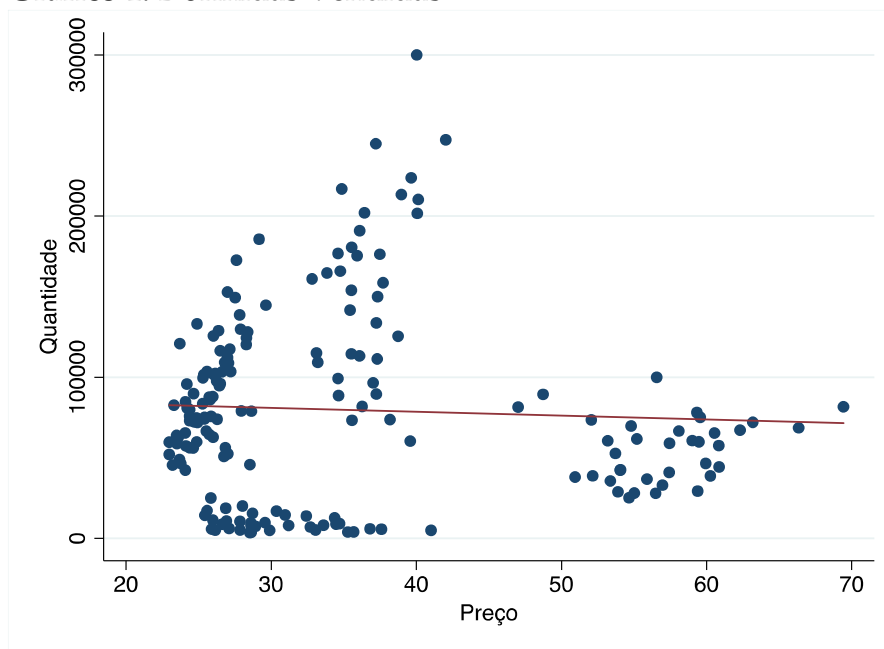
³⁵Correlação entre os instrumentos e variáveis endógenas (ρ_{jrt} e $s_{j|rt}$).

Capítulo 4

4. ANÁLISE PRELIMINAR

Como meio de resposta à questão de investigação proposta, foi realizada uma análise preliminar, que analisa a relação entre o número de dormidas vendidas e o preço médio praticado, na região Norte. Os resultados sugerem que um aumento do preço resulte numa diminuição do número de dormidas vendidas, na região Norte de Portugal (ver Gráfico 1). Em particular, espera-se que o aumento do preço médio em 1€ (valor da taxa turística, aplicada por dormida) conduza a uma diminuição de, aproximadamente 237 dormidas, por tipologia e mês.

Gráfico 1. Dormidas Vendidas



Através do preço médio, de cada tipologia e do declive da relação estabelecida (-237,45), obteve-se uma variação média das quantidades e receitas

para cada tipologia, por mês (ver Tabela 4, *infra*). Numa análise geral, conclui-se que o aumento do preço de 1€, decorrente da implementação da taxa turística, provoca uma diminuição total de, aproximadamente, 1.187 dormidas por mês, na quantidade média vendida, e uma diminuição da receita total no valor de 42.145€, por mês, na região Norte. A acrescer ao explanado, espera-se que a taxa tenha um maior impacto na tipologia 5*, e um menor impacto na tipologia 1* e 2*.

Contudo, na análise desenvolvida não são consideradas determinadas variáveis que poderão ter impacto nas quantidades vendidas, tais como: as preferências dos consumidores (por uma determinada tipologia e/ou mês) e os efeitos choque na procura, resultantes das características das tipologias que as diferenciam e conduzem as vendas. Assume-se, então, que os resultados obtidos, nesta leitura, poderão estar enviesados. É, portanto, elaborada, nos seguintes capítulos, uma análise mais discriminada, a partir de modelos de estimação da procura *logit*, com o intuito de obter resultados menos enviesados, para responder à questão de investigação proposta.

Tabela 4. Resultados por Tipologia

Tipologia	Preço médio (€)	Impacto no nº de Dormidas (unidades)	Impacto na Receita (€)
1*e 2*	24,68	-237,45	-5.860,27
3*	26,86	-237,45	-6.377,91
4*	36,69	-237,45	-8.712,04
5*	57,01	-237,45	-13.537,02
H-A	32,25	-237,45	-7.657,76
Total		-1.187,25	-42.145,00

Capítulo 5

5. RESULTADOS DE ESTIMAÇÃO

No presente capítulo, são apresentados e discutidos os resultados de estimação obtidos dos modelos de procura *logit*, *Multinomial Logit* (MNL) e *Nested Multinomial Logit* (NMNL), descritos na Tabela 5, *infra*. Em todas as especificações (colunas), presentes na tabela, é considerada a seguinte variável dependente: $y_{jrt} = \ln s_{jrt} - \ln s_{0t}$.

Inicialmente, foram estimadas as especificações (i), (ii), (iii) e (iv), que se reportam aos resultados obtidos para o modelo MNL através de regressões estimadas pelo método dos mínimos quadrados (MMQ). Na regressão da especificação (i) são incluídos, como variáveis independentes, somente os preços das tipologias por região e mês. De acordo com a lei da procura, o aumento do preço de um produto resulta na diminuição da quantidade procurada do mesmo, e vice-versa. Isto significa, que se espera, que o valor do coeficiente do preço seja negativo. No entanto, este facto não se verifica na especificação (i): o coeficiente do preço é positivo (0,0141), na medida em que todas as características dos produtos são não observáveis e com grande probabilidade estarão correlacionadas com o preço.

Assim sendo, nas seguintes especificações foram introduzidas variáveis indicador para controlar e medir os efeitos específicos de cada tipologia, região e mês.³⁶ A partir dos coeficientes obtidos para estas variáveis, é possível identificar-se a preferência dos consumidores por uma determinada tipologia,

³⁶ A variável indicador específica de um grupo (diferentes tipologias, regiões e meses) controla o efeito das variáveis omitidas, que são consideradas constantes entre membros pertencentes ao mesmo grupo, mas que variam entre grupos. Por outras palavras, o efeito fixo de um determinado grupo corresponde a uma variável indicador, que assume o valor 1 para todas as observações pertencentes a esse mesmo grupo, ou o valor 0 em caso contrário. Desta forma, são controladas as características não observáveis de um grupo específico, que podem afetar as variáveis dependente e explicatórias (Davis e Garcés, 2010, p. 98).

região e mês.

Assim, na especificação (ii) é introduzido um conjunto de variáveis indicador para as diferentes tipologias (1* e 2*, 3*, 4*, 5*, e H-A). Seguidamente, na especificação (iii), é ainda adicionado um conjunto de variáveis indicador para as regiões de Portugal (Alentejo, Algarve, Centro, Lisboa e Norte). Por último, dada a procura sazonal, característica inerente ao setor em estudo (mercado hoteleiro), acresce à especificação (iv), um conjunto de variáveis indicador que permite capturar o efeito de cada mês (Janeiro a Dezembro).

Através da Tabela 5(cfr. *infra*)é possível observar-se uma diminuição gradual do coeficiente do preço e um aumento do R^2 , resultante da introdução dos efeitos fixos, que manifestam um impacto significativo no preço e conseqüente diminuição do enviesamento. Não obstante, o coeficiente do preço mantém-se positivo.

Por este motivo, foram estimadas as especificações (v), (vi), (vii) e (viii) do modelo MNL através de variáveis instrumentais (VI). Para controlar a correlação entre o preço e termo erro (características não observáveis das diferentes tipologias), foram utilizadas, em cada uma destas especificações, os seguintes instrumentos: número de concorrentes por tipologia, região e mês (z1) e número de concorrentes em Portugal, por mês (z2).

Ora, frisando as conclusões acima descritas, depreende-se que a especificação (viii) apresenta melhores resultados do que as anteriores, dada a introdução dos efeitos fixos das variáveis tipologia, região e mês. Já o coeficiente do preço, nesta especificação, apresenta um valor negativo, corrigido pelas VI.

Todavia, devido à estrutura de padrões de substituição, rígida e indesejada, produzida pelo modelo MNL, foram estimadas especificações do modelo NMNL, que produz padrões de substituição mais flexíveis. De acordo com a equação (13), apresentada no subcapítulo 2.2, é introduzida uma variável independente: o logaritmo da quota de mercado de cada produto, em cada grupo, por mês ($\ln s_{j|rt}$). Tendo em conta que esta variável estará correlacionada

com as características não observáveis das tipologias (chamando-se aqui à colação a lógica de que melhores características atraem mais clientes, que conseqüentemente geram maior quota de mercado), é utilizado um terceiro instrumento: o número de hotéis concorrentes por região e mês (z3).

Aplicando-se, novamente, a lógica das conclusões retiradas nas análises anteriores, afirma-se que a especificação (xvi) produz resultados mais plausíveis quando comparada com as restantes especificações do modelo NMNL. Contudo, de acordo com os pressupostos do NMNL referidos no subcapítulo 2.2, para que o modelo seja consistente com a maximização da utilidade aleatória, o coeficiente da variável $\ln s_{j|rt}$, tem que pertencer ao intervalo entre 0 e 1. Para além de este coeficiente não apresentar um valor entre o intervalo referido, o coeficiente do preço volta a apresentar um valor positivo (ver tabela 5, *infra*), pelo que se conclui que a utilização do modelo NMNL não é adequada para estimar a procura no mercado hoteleiro de Portugal.

Deste modo, serão utilizados os resultados alcançados pela especificação (xii), do modelo MNL através de VI para retirar as principais conclusões quanto ao comportamento dos consumidores na procura de estabelecimentos hoteleiros, em Portugal.

Tabela 5. Resultados dos Modelos Logit^a

Variável	MNL: MMQ				MNL: VI				NMNL: MMQ				NMNL: VI			
	(i)	(ii)	(iii)	(iv)	(ix)	(x)	(xi)	(xii)	(v)	(vi)	(vii)	(viii)	(xiii)	(xiv)	(xv)	(xvi)
Preço	0,0141**	0,0528**	0,0397**	0,0184**	-0,1661**	0,3398**	-0,1300**	-0,1307**	0,0163**	0,0445**	0,0308**	0,0091**	0,3133	0,3989**	0,1516**	0,1516**
	(0,0026)	(0,0047)	(0,0040)	(0,0033)	(0,0150)	(0,1138)	(0,0604)	(0,0494)	(0,0017)	(0,0034)	(0,0027)	(0,0011)	(0,2192)	(0,0961)	(0,0435)	(0,0416)
$lns_{j rt}^b$									0,9574**	0,8530**	0,9798**	0,9856**	3,4052**	1,1765**	1,0837**	1,0837**
									(0,0345)	(0,0499)	(0,03)	(0,0111)	(1,7210)	(0,2019)	(0,0584)	(0,0552)
Efeitos Fixos:	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia	Tipologia
			Região	Região			Região	Região			Região	Região			Região	Região
				Mês				Mês								Mês
Instrumentos ^c						z1; z2								z1; z2; z3		
R ² ^d	0,0209	0,2894	0,5640	0,7178	.	.	.	0,1340	0,4375	0,4888	0,8206	0,9783	.	.	0,3825	0,4352

a: Em cada uma das especificações (colunas) apresentadas na tabela, é considerada a variável independente y_{rt} ($lns_{j|rt} - lns_{0t}$), 900 observações, uma constante e os erros-padrão robustos obtidos para cada uma das variáveis independentes na respetiva especificação, que se apresentam entre parênteses. Todos os valores apresentados na tabela estão arredondados a 4 casas decimais.

b: Logaritmo da quota de mercado de cada produto (tipologia: j), em cada grupo (região: r), por mês (t).

c: z1 denota o n.º de hotéis concorrentes por tipologia, região e mês; z2 representa o n.º de hotéis concorrentes em Portugal Continental, por mês; z3 indica o n.º de hotéis concorrentes por região e mês.

d: Indica a proporção da variância total, que é explicada pelo modelo numa escala de 0 a 1, significando uma melhoria proporcional na previsão de resultados a partir do modelo de regressão. Em algumas das especificações, este indicador é apresentado com um "." pelo facto de o software Stata suprimir a impressão deste indicador quando o seu valor é negativo – a soma dos quadrados do modelo é negativa - numa regressão, em que são introduzidas variáveis instrumentais.

**variáveis significativas na estimação da procura: p-Value < 0,05

Na Tabela 6 (cfr. *infra*) são descritos todos os resultados obtidos na especificação (xii). Tal como nas restantes especificações, foram omitidos os efeitos fixos da tipologia 1* e 2*, região Alentejo e mês de Janeiro. Assim, estas variáveis indicador tomam o valor 0, a partir do qual são comparados os efeitos fixos das restantes variáveis. As variáveis omitidas foram selecionadas de acordo com o critério de menor número de dormidas vendidas (em média) e consequentes (menores) proveitos gerados.

Ora, de acordo com os resultados apresentados na tabela 6 (*infra*), retiram-se as seguintes conclusões:

- 1) O consumidor é sensível ao preço, dado o coeficiente negativo desta variável (-0,1307), pelo que um aumento do preço provoca, inevitavelmente, uma diminuição na procura.
- 2) O ranking de preferências por tipologia é dado por: 5*, 4*, 3*, H-A e 1* e 2*. No entanto, a diferença entre hotéis-apartamentos e hotéis de 1* e 2* não é significativa.
- 3) O ranking de preferências por região é dado por: Lisboa, Norte, Algarve, Centro e Alentejo.
- 4) O ranking de preferências por mês é dado por: Agosto, Setembro, Julho, Junho, Maio, Outubro, Abril, Novembro, Março, Fevereiro, Dezembro e Janeiro. Refira-se, contudo, que a diferença entre os últimos três meses referidos não é significativa.
- 5) A elasticidade-preço direta média da procura,³⁷ por tipologia, região e mês é 4,52 em valor absoluto, o que significa que um aumento do preço em 1% provoca, em média, uma diminuição da quantidade procurada (número de dormidas) em 4,52%.

³⁷Obtida através do coeficiente do preço (-0,1306562), preço médio praticado (35,21355) e média da quota de mercado (0,0181253) em Portugal: $-0,1306562 \times 35,21355 \times (1 - 0,0181253)$.

Tabela 6. Resultados MNL: VI^a

Variável	Coefficiente	Erro-Padrão Robusto
Preço	-0,1307**	(0,0494)
Efeitos Fixos ^b :		
3*	0,9713**	(0,1097)
4*	2,3904**	(0,4384)
5*	3,9089**	(1,4906)
H-A	0,5917	(0,3220)
Algarve	1,4949**	(0,2119)
Centro	0,7641**	(0,1438)
Lisboa	3,5216**	(0,4938)
Norte	1,5408**	(0,1293)
Abril	0,9727**	(0,1799)
Agosto	2,6540**	(0,3849)
Dezembro	0,1618	(0,1925)
Fevereiro	0,1709	(0,1933)
Julho	1,9932**	(0,3124)
Junho	1,7127**	(0,2527)
Maiο	1,5620**	(0,2326)
Março	0,4527**	(0,1948)
Novembro	0,4850**	(0,1900)
Outubro	1,4468**	(0,2080)
Setembro	2,2555**	(0,3002)
Instrumentos		z1; z2
R ² ^d		0,1340

a: Na especificação apresentada na tabela é considerada a variável independente y_{rt} ($\ln s_{jrt} - \ln s_{0t}$), 900 observações, uma variável constante e os erros-padrão obtidos para cada uma das variáveis independentes, que se apresentam entre parênteses. Todos os valores apresentados na tabela estão arredondados a 4 casas decimais.

b: Variáveis omitidas: tipologia de 1* e 2*, região do Alentejo e mês de Janeiro.

c: z1 denota o n.º de hotéis concorrentes por tipologia, região e mês; z2 representa o n.º de hotéis concorrentes em Portugal Continental por mês; z3 indica o n.º de hotéis concorrentes por região e mês.

d: Indica a proporção da variância total que é explicada pelo modelo numa escala de 0 a 1, o que significa uma melhoria proporcional na previsão de resultados a partir do modelo de regressão.

** Variáveis significativas na estimação da procura: p-value < 0,05.

Analisando as conclusões retiradas acima, deduz-se que os resultados produzidos pelo modelo MNL através de VI são plausíveis e apresentam o sinal esperado, dada a percepção do comportamento dos consumidores no mercado hoteleiro de Portugal, mais concretamente aludindo à sensibilidade ao preço e preferências por tipologia, região e mês. Adicionalmente, comparando o valor obtido, para a elasticidade-preço direta média (4,52%), com os valores da elasticidade-preço direta média (3,71% e 4,22%), reportados nos estudos analisados, na secção 1.2.2, do subcapítulo “O Mercado Hoteleiro”, conclui-se que a estimação realizada produz uma estimativa plausível para esta componente da procura, dado que os valores apresentados são semelhantes.

Assim, conclui-se que os resultados de estimação, gerados pelo modelo MNL, com VI, são, na sua generalidade, confiáveis.

Capítulo 6

6. O IMPACTO DA TAXA TURÍSTICA

Tendo por base os resultados de estimação obtidos no capítulo anterior, prossegue a análise sobre o impacto da taxa turística, na região Norte, de Portugal, de forma a quantificar os *trade-offs* resultantes da implementação da referida taxa, no curto prazo.

6.1 Impacto na Procura

Para estimar o impacto do aumento de 1€ no preço praticado por todas as tipologias na região Norte, é necessário determinar, inicialmente, a sensibilidade dos consumidores a um aumento do preço neste mercado, isto é, a elasticidade-preço.

A Tabela 7, sumariza os valores médios dos preços e quotas de mercado, correspondentes a cada uma das tipologias na região norte, a partir dos quais foram calculadas as elasticidades-preço diretas e cruzadas representadas na Tabela 8(cfr. *infra*), assumindo um coeficiente do preço igual a -0,1307 (consistente com os resultados obtidos, no capítulo anterior).

Tabela 7. Preço e Quota de Mercado por Tipologia

Tipologia	Preço (€)	Quota de Mercado (%)
1* e 2*	24,38	1,49
3*	26,83	1,83
4*	36,35	2,92
5*	56,77	1,04
H-A	28,81	0,16

Tabela 8. Elasticidade-Preço por Tipologia (%)

Tipologia	1* e 2*	3*	4*	5*	H-A
1* e 2*	-3,176	0,064	0,137	0,076	0,007
3*	0,049	-3,445	0,137	0,076	0,007
4*	0,049	0,064	-4,657	0,076	0,007
5*	0,049	0,064	0,137	-7,373	0,007
H-A	0,049	0,064	0,137	0,076	-3,956

Os valores (negativos) na diagonal, apresentados na Tabela 8, correspondem às elasticidades-preço diretas, e os valores (positivos) fora da diagonal indicam as elasticidades-preço cruzadas. Exemplificando, na primeira linha, o número de dormidas da tipologia 1* e 2* diminui 3,176% com um aumento de 1% do respetivo preço, aumenta 0,064% com o aumento de 1% no preço da tipologia 3*, e assim sucessivamente.

Neste seguimento, e com fundamento nas elasticidades-preço apresentadas, é composta a média (total) das elasticidades-preço direta e cruzada, que tomam os seguintes valores (respetivamente): -4,521% e 0,067%.

Em média, espera-se que um aumento do preço de 1%, em todas as tipologias, da região norte, tenha um impacto negativo de 4,521% no número de dormidas por tipologia, por mês, repetindo-se a enunciação das conclusões referidas no capítulo anterior: a elasticidade-preço direta obtida apresenta um valor de confiança, um valor semelhante aos estudos analisados na secção 1.2.2, referente ao capítulo “O Mercado Hoteleiro”.

Acresce que, verifica-se que as tipologias com um nível de estrelas mais alto tendem a ser mais elásticas (elasticidade-preço direta com valores mais elevados), o que vai de encontro à esteira defendida no estudo de Anderson e Xie (2010), analisada na secção 1.2.2.

Para a elasticidade-preço cruzada, mantém-se a conclusão de que os

estabelecimentos hoteleiros são produtos substitutos, o que é condizente com os estudos aqui descritos – com a particularidade de que os padrões de substituição produzidos pelo modelo de estimação utilizado (MNL) nesta questão apresentarem uma estrutura mais rígida, refletida nas elasticidades-preço cruzadas apresentadas na Tabela 8 (cfr. *supra*). Assim motivado, o valor da média (total) da elasticidade-preço cruzada aponta para um menor grau de semelhança para com os valores obtidos nos estudos analisados.

Não obstante, as diferenças das elasticidades entre os modelos não resultam exclusivamente dos modelos de estimação da procura aplicados: existem características inerentes aos mercados hoteleiros, que diferem de acordo com a localização (cidade/país) a que os estudos são aplicados.

6.2 Impacto nos Estabelecimentos Hoteleiros

Na Tabela 9 (cfr. *infra*) são apresentados o impacto na quantidade e receita dos estabelecimentos hoteleiros do Norte de Portugal, por tipologia, decorrente da implementação da taxa turística.

Através do impacto percentual do valor da taxa (1€ por dormida) no preço médio de cada tipologia, calculado na primeira coluna, e das elasticidades exibidas na Tabela 8 (cfr. *supra*), é possível vislumbrar o impacto percentual no número de dormidas por tipologia. Estima-se uma diminuição da quantidade procurada em 12%, por mês, que corresponde à diminuição do número de dormidas vendidas, por tipologia, apresentadas na 3.^a coluna, e consequente diminuição das respetivas receitas, apresentadas na última coluna da Tabela 9. O impacto percentual na quantidade, por exemplo, para a tipologia 5*, é obtido através da expressão:

$$\Delta q_{5*} = e_{5*1*e2*}(w_{1*e2*}) + e_{5*3*}(w_1) + e_{5*4*}(w_{4*}) + e_{5*5*}(w_{5*}) + e_{5*H-A}(w_{H-A}),$$

em que a variável dependente Δq indica a variação percentual na quantidade, a variável e representa as elasticidades-preço direta e cruzadas da tipologia, e w concretiza o impacto percentual no preço médio da tipologia. Multiplicando Δq pelo preço médio de cada tipologia, obtém-se o resultado do impacto gerado nas receitas, por tipologia.

Tabela 9. Impacto da Taxa Turística por Tipologia

Tipologia	Impacto de 1€ no preço (%)	Impacto (%) no nº de Dormidas	Impacto no nº de Dormidas (unidades)	Impacto na Receita (€)
1* e 2*	4,05	-12,10	-9.879,89	-2.43.857,17
3*	3,72	-12,10	-11.907,76	-3.19.873,70
4*	2,73	-12,10	-18.610,36	-682.869,77
5*	1,75	-12,10	-6.638,22	-378.447,46
H-A	3,30	-12,10	-1.170,49	-35.501,65
Total			-48.206,72	-1.660.549,75

Em geral, verifica-se que a implementação da taxa turística gera, para a região Norte, uma diminuição total do número de dormidas vendidas, por mês, de aproximadamente 48.207 dormidas e consequente diminuição das receitas em 1.660.549,75€.

Comparando os valores apresentados na tabela acima, com os da análise preliminar (capítulo 4) constata-se que existe um grande diferencial nos resultados obtidos para o impacto nas quantidades e receitas, por tipologia.

Por outro lado, e ao contrário da conclusão retirada na análise preliminar, a implementação da taxa gera um impacto mais elevado para a tipologia 4*, e menor para a tipologia hotéis-apartamentos.

Destarte, em jeito de conclusão, afere-se que os resultados obtidos na análise preliminar estão enviesados, e que a estimação a partir do modelo *logit* gera resultados mais confiáveis, devido à introdução de variáveis que vão além do preço na função procura.

6.3 Imposto gerado pela Taxa Turística

Na Tabela 10, é quantificado o imposto gerado através da implementação da taxa turística por tipologia, determinando-se um imposto total, para a região Norte, no valor de 350.298,96 €, por mês.

Realça-se, que a implementação de tal taxa poderá, eventualmente, ser vantajosa, quando aplicada à promoção do turismo e renovação de material urbano, na lógica de que o sucesso da atividade turística, da região Norte, instala um aumento de pressão sob o espaço urbano e derivados custos. Em ultima análise, o imposto gerado poderá contribuir para a dinamização da região Norte, de Portugal.

Tabela 10. Imposto obtido por Tipologia

Tipologia	Imposto obtido (€)
1* e 2*	7.1793,22
3*	86.528,91
4*	135.234,04
5*	48.237,28
H-A	8.505,51
Total	350.298,96

CONCLUSÃO

O presente estudo foi realizado com o objetivo de quantificar o impacto da implementação de uma eventual taxa turística na região Norte, de Portugal, querendo-se este como um meio ao dispor das entidades hoteleiras, autarquias e organizações do turismo, para, através dele, obter estimativas dos efeitos decorrentes da implementação da referida taxa.

A análise desenvolvida para responder à questão de investigação proposta fundamentou-se na aplicação de modelos de estimação da procura de escolha discreta, nomeadamente, os modelos *logit*: *Multinomial Logit* (MNL) e *Nested Multinomial Logit* (NMNL), desenvolvidos por McFadden (1978, 1981).

A apropriação do modelo utilizado, bem como as respetivas variáveis incorporadas no mesmo, são sustentadas pelos resultados de estimação obtidos nas várias hipóteses testadas no capítulo 5, concluindo-se que o modelo MNL, estimado através de regressões com variáveis instrumentais, gera resultados mais plausíveis.

Analisando os resultados produzidos pelo modelo, verifica-se que as variáveis introduzidas são significativas, e que os diferentes coeficientes apresentam a sinalização esperada, sendo comprovadas as vantagens de utilização do mesmo, dada a percepção do comportamento dos consumidores no mercado hoteleiro de Portugal. Os consumidores são sensíveis ao preço, pelo que é expectável uma diminuição da quantidade procurada, dado um aumento do preço (resultado consistente com a Lei da Procura) e o ranking de preferências obtidas por tipologia, região e mês, representa, de um modo geral uma elevada aproximação à realidade.

Aplicando-se os resultados obtidos através da estimação já mencionada na quantificação do impacto da taxa turística na região Norte, de Portugal, podemos enumerar as seguintes conclusões finais:

- 1) É esperado, em média, que o aumento do preço em 1€ em todas as tipologias da região Norte, resulte num impacto negativo de 12,1% na quantidade procurada por tipologia, por mês;
- 2) As quantidades e receitas dos estabelecimentos hoteleiros, na região Norte, sofrerão, eventualmente, com a aplicação da taxa turística, um decréscimo significativo do número de dormidas vendidas e receitas obtidas por mês, com uma diminuição total destas variáveis, para o total da região Norte, em 48.207 dormidas e 1.660.549,75€, respetivamente;
- 3) A tipologia de 4* será a mais afetada, negativamente, pela taxa e a tipologia hotéis-apartamentos terá o menor impacto negativo;
- 4) O imposto gerado, através da implementação da taxa turística, por tipologia, para a região Norte, terá capacidade para gerar a quantia de 350.298,96€, por mês, para o total das autarquias, pertencentes à região norte de Portugal Continental.

Aquando da realização deste estudo, não nos foi possível o acesso a determinados dados específicos, ao nível de cada entidade hoteleira. No entanto, poderá ser do interesse de posteriores estudos a estimação da procura para o mercado analisado, através de dados ao nível de cada entidade hoteleira, com incorporação de variáveis específicas às mesmas, tais como a marca, comodidades internas (piscina, ginásio, spa, etc.) e comodidades externas (proximidade de praia, comércio, pontos turísticos, transportes, etc.), que permitem uma previsão de resultados mais precisa.

No âmbito de estudo desta tese, foi considerado apenas o impacto da taxa turística a curto prazo, pese embora este possa ser, num segundo momento, utilizado como base e/ou exame complementar para uma análise quanto aos seus efeitos a longo prazo.

Através do imposto gerado, é possível uma melhoria na dinamização da região norte de Portugal, podendo este resultar num aumento do nível de atração para os visitantes e, concludentemente, no incremento do bem-estar social a longo prazo. De resto, e ainda que não seja esse o resultado alcançado, este estudo constitui um benefício ao dispor das entidades hoteleiras, que poderão, através dele, prever eventuais efeitos negativos resultantes da implementação da taxa turística e, de acordo com a sua tipologia, traçar antecipadamente medidas de reação com vista à atenuação desse possível impacto prejudicial.

Bibliografia

- Akerberg, D., Benkard, L., Berry S. & Pakes, A. J. J. Heckman & E.E. Leamer (Eds.). 2007. Econometric Tools for Analyzing Market Outcomes. In *Handbook of Econometrics*: 4171–4276.
- Anderson, C. K., & Xie, X. 2012. A choice-based dynamic programming approach for setting opaque prices. *Production and Operations Management*: 21(3), 590–605.
- Berry, S. 1994. Estimating discrete-choice models of product differentiation. *The RAND Journal of Product Differentiation*, 25(2): 242–262.
- Berry, S., Levinsohn, J., & Pakes, A. 1995. Automobile Prices in Market Equilibrium. *Econometrica* 63(4): 841–890.
- Berry, S., & Pakes, A. 2007. The pure characteristics demand model. *International Economic Review*, 48(4): 1193–1225.
- Davis, P. & Garcés, E. 2009. *Quantitative Techniques for Competition and Antitrust Analysis*. Princeton University Press. ©
- Davis, P., & Ribeiro, R. 2010. *A Simple Globally Consistent Continuous Demand Model for Market Level Data*. SSRN Electronic Journal.
- Davis, P., & Sloan, M. 2001. *Demand models for market-level data working manuscript*, MIT Sloan School.
- Ghose, A., Ipeirotis, P. G., & Li, B. 2010. Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. *Marketing Science*, 31(3): 493–520.
- Greene, W. H. 2011. *Econometric Analysis*. Prentice Hall.

- Hausman, J. A., & Wise, D. A. 1978. A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica, Journal of the econometric society*: 403-426.
- Hausman, J. A. 1996. Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*: 207-248. University of Chicago Press.
- Hausman, J. A., Leonard, G., & Zona, J. D. 1994. Competitive analysis with differentiated products. *Annales d'Economie et de Statistique*: 159-180.
- Ivaldi, M., & Verboven, F. 2005. Quantifying the effects from horizontal mergers in European competition policy. *International Journal of Industrial Organization*, 23(9-10): 669–691.
- Lederman, R., Olivares, M., & Van Ryzin, G. 2014. **Identifying competitors in markets with fixed product offerings**. Working Paper No. 14-10. Columbia Business School.
- Li, B., Ghose, A., & Ipeirotis, P. G. 2011. Towards a theory model for product search. In *Proceedings of the 20th international conference on World wide web*: 327-336. ACM.
- Manski, C. F., & McFadden, D. (Eds.). 1981. Econometric models of probabilist choice. Em *Structural analysis of discrete data with econometric applications*: 202-204. Cambridge, MA: Mit Press.
- McFadden, D. L. Spatial A. Karlquist et al.(Eds).1978. Modeling The Choice of Residential Location. In *Interaction Theory and Planning Models*.
- Nevo, A. (2000). A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand. *Journal of Economics & Management Strategy*,9(4): 513–548.

- Song, M. 2010. A hybrid discrete choice model of differentiated product demand with an application to personal computers. *International Economic Review*, 56(1): 265-301.
- Turismo de Portugal (2013) Dormidas, Disponível em www.turismodeportugal.pt (14/08/2014)
- Turismo de Portugal (2013) N° Estabelecimentos, Disponível em www.turismodeportugal.pt (16/02/2016)
- Turismo de Portugal (2013) Proveitos de Aposento, Disponível em www.turismodeportugal.pt (14/08/2014)
- Turismo de Portugal (2013) Taxa de Ocupação, Disponível em www.turismodeportugal.pt (01/04/2015)
- Turismo de Portugal (2014 e 2015) Dormidas, Disponível em www.turismodeportugal.pt (16/02/2016)
- Turismo de Portugal (2014 e 2015) N° Estabelecimentos, Disponível em www.turismodeportugal.pt (16/02/2016)
- Turismo de Portugal (2014 e 2015) Proveitos de Aposento, Disponível em www.turismodeportugal.pt (16/02/2016)
- Turismo de Portugal (2014 e 2015) Taxa de Ocupação, Disponível em www.turismodeportugal.pt (16/02/2016)
- Venkataraman, S., & Kadiyali, V. 2005. An Aggregate Generalized Nested Logit Model of Consumer Choices: An Application to the Lodging Industry. *SSRN Electronic Journal*.
- Verboven, F. 1996. The nested logit model and representative consumer theory. *Economics Letters*, 50(1): 57-63.
- Wen, C. H., & Koppelman, F. S. 2001. The Generalized Nested Logit Model. *Transportation Research Part B*, 35(7): 627-641.