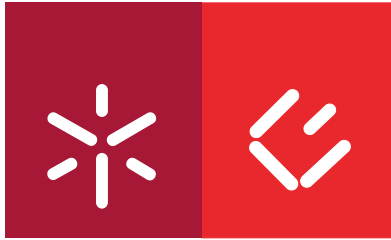


Universidade do Minho
Escola de Economia e Gestão

Maria José Andrade Pais Valente

**Tributação do património e determinantes
económicas no sucesso político dos
governantes**



Universidade do Minho

Escola de Economia e Gestão

Maria José Andrade Pais Valente

**Tributação do património e determinantes
económicas no sucesso político dos
governantes**

Tese de Doutoramento em Ciências Económicas
Área de Especialização em Teoria e Política Económica

Trabalho efectuado sob a orientação da
Professora Doutora Linda Rosa F. Gonçalves Veiga
e do
Professor Doutor Rui Nuno G. P. N. Baleiras

É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

Universidade do Minho, ___/___/_____

Assinatura: _____

Ao Fernando

E aos meus pais

Agradecimentos

A todas as pessoas a quem a Vida me concedeu o privilégio de conhecer e que por via da sua nobreza, da sua confiança, do seu estímulo ajudaram a que se tornasse realidade este projecto. A todas, um grande Bem-Haja.

Em pensamento, o meu reconhecimento é uma firme certeza. Gratidão entende-se com dever de reserva. Porém tal discrição sobreviria absurda por mais nobre que seja o propósito.

Antagonismo que solicita a manifestação do meu reconhecimento.

A ser assim, não posso esquecer o Professor Doutor José Da Silva Costa, pelas palavras de estímulo e de confiança, pelos horizontes que me apontou, à Professora Doutora Ana Paula Barreira pela sua amizade e enorme generosidade, ao Professor Doutor Rui Nuno Baleiras e à Professora Doutora Linda Rosa Veiga por confiarem no meu esforço e por me apoiarem nos momentos de maior perplexidade, ao Professor Doutor José Cadima Ribeiro por ter acolhido a minha intenção de prosseguir estudos de doutoramento na EEG e contribuído para que a mesma se viesse a oficializar institucionalmente.

À Fundação Portuguesa para a Ciência e a Tecnologia, por, no âmbito do projecto de investigação PTDC/ECO/65711/2006, co-financiado pelo Fundo Europeu de Desenvolvimento Regional, ter permitido à Professora Doutora Linda Rosa Veiga, enquanto co-orientadora da presente dissertação, beneficiar de apoio financeiro.

TRIBUTAÇÃO DO PATRIMÓNIO E DETERMINANTES ECONÓMICAS NO SUCESSO POLÍTICO DOS GOVERNANTES

RESUMO

A presente dissertação na área científica de Ciências Económicas, submetida à Escola de Economia e Gestão, da Universidade do Minho, para o cumprimento dos requisitos conducentes à obtenção do grau de doutoramento, é constituída por dois ensaios: um sobre *tributação do património* e outro sobre *determinantes económicas no sucesso político dos governantes*.

1º ensaio

A motivação para o tema do primeiro ensaio, *tributação do património*, decorre, simultaneamente, da importância do imposto patrimonial no sistema de financiamento público local e da miscelânea de valores matriciais dos imóveis, em Portugal, ao ano de 2001. Acresce, ainda, como desafio associado à elevada visibilidade do imposto patrimonial, *a necessidade de um equilíbrio, permanente, entre eficiência e equidade fiscal*.

Concretizar o objectivo significou eleger como questão nuclear do ensaio: *a avaliação da base do imposto sobre o património*. Relativamente aos métodos de avaliação, coincidimos com a diversa bibliografia analisada, quanto à necessidade do processo de estimação dos valores das propriedades reflectir o valor de mercado. Nesta perspectiva os modelos de preços hedónicos surgiram como uma boa alternativa para o aperfeiçoamento dos processos de avaliação patrimonial.

A contextualização do bem *habitação*, enquanto bem diferenciado, na teoria neoclássica permitiu a análise de mercado, de modo a que se consolidassem fundamentos para suportar um modelo teórico robusto ao propósito da *habitação* vir avaliada na vizinhança do valor de mercado. Dotados do modelo base da estimação do valor da propriedade imobiliária, em termos teóricos e em termos empíricos, procurou-se canalizá-lo para a avaliação de uma amostra do parque imobiliário da cidade da Guarda.

Salvaguardada a fiabilidade, a correcção e o rigor dos valores estimados a partir do modelo econométrico, entretanto construído, chegou-se a uma metodologia de avaliação imobiliária, para a cidade da Guarda.

Finalmente, porque o imposto sobre o património, como realidade que nos propúnhamos estudar, conheceu uma alteração em termos legislativos, foi estabelecido um

paralelismo entre o modelo de avaliação base deste ensaio e o modelo resultante do recente Imposto Municipal sobre Imóveis (IMI). Tendo-se verificado uma vantagem, relativamente, superior para a metodologia de avaliação imobiliária sustentada no modelo de preços hedónicos face ao regime transitório do IMI em vigor. Vantagem ligeiramente atenuada caso a avaliação imobiliária para efeitos de IMI correspondesse à aplicação do regime definitivo.

2º ensaio

A motivação para o tema do segundo ensaio, *determinantes económicas no sucesso político dos governantes*, decorre da avaliação probabilística dos resultados pós-eleição, enquanto consequência do bem-estar pós-eleitoral do político local e nesse contexto pretendeu-se abordar a probabilidade de reeleição do autarca incumbente numa perspectiva diversa da usualmente utilizada nos modelos de ciclos político-económicos. O objectivo a que nos propusemos foi o de equacionar esse parâmetro como uma variável contínua e não como uma variável *dummy*. Para o concretizar havia que desenhar um modelo teórico que suportasse o trabalho empírico subsequente. Essa modelização, fundamentada na relação de agência e no jogo de sinais, exigiu-nos uma adaptação que se ajustasse à escassez de informação disponível relativamente ao desempenho e popularidade do político incumbente.

O desenvolvimento do modelo empírico, a partir de informação político/económica, entre 1979 e 2001, para 275 municípios portugueses, permitiu, numa primeira fase, estimar a probabilidade de reeleição do autarca, de forma a sustentar, numa segunda fase, a observação da tendência (ou não) para o envolvimento dos governos locais em ciclos político-económicos. Neste capítulo, escolheu-se o serviço de dívida municipal para apuramento da relação (ou não) com as probabilidades estimadas de reeleição, de forma a perscrutar (ou não) eventual efeito da probabilidade de reeleição na política de endividamento autárquico.

As conclusões do ensaio permitem coincidir com as soluções de equilíbrio do modelo teórico, mostrando que as preocupações com a reeleição podem impelir um político competente a não implementar uma política socialmente desejável, mas tão só aquela política que lhe permite resolver a preocupação com o seu futuro pessoal; bem como podem evitar que um político incompetente concretize uma política socialmente indesejável em prol de uma política que lhe garanta solucionar o respectivo futuro político e sustentar o propósito de se alcançar valores estimados para a probabilidade de reeleição dos autarcas portugueses.

PROPERTY TAXATION AND THE POLITICIANS' PROBABILITY OF SUCESS FROM ECONOMICS DETERMINANTS

ABSTRACT

This dissertation in Political Economics, submitted to the Graduate School of Economics and Management of Minho University in partial fulfillment of the requirements for the degree of doctor of philosophy, brings together two essays: the first one about property taxes and the second about political business cycles and the probability of an incumbent's reelection.

First Essay:

Property taxation is the backbone of municipal finance in a great number of developed countries. Over the years, however, it has equally not been without controversy in a great number of implementation issues. For example, discussion has often emerged over the role that property taxes should play in financing municipal services, as well as in identifying taxable properties, choosing an appropriate tax base, setting up a proper assessment system and establishing property tax rates. The main purpose of the first essay is to bring efficiency, accountability and fairness into the implementation of a property tax system.

One of the main empirical tools used to assign an assessed value to the property are house-value regressions, also called hedonic regressions. To make this approach possible, the study analyzes and determines the patterns of the market values paid by consumers for housing features. Hedonic price models, based on Rosen's analytical framework, were developed to examine the information of housing market transactions in Guarda. For this propose we created a cross-sectional data collected from real estate offices. These models explained how housing features (i.e., size, localization, environment, comfort) affected housing prices. Results suggest different housing values for different attributes. This analysis provides fiscal authorities with relevant information to prepare assessment programs that underlie property taxes.

Finally, if application of property taxes is to be fair, it has to be based on assessments that are uniform within each taxing jurisdiction so we should evaluate this

proposed market value assessment base against the assessment base used nowadays by the Portuguese fiscal system. A set of housing assets within the data-base were registered in Guarda at the local services of the Directorate General for Taxation (DGCI) and provide an opportunity to look for this objective. To examine the impact of assessment on property tax equity, we used the empirical tax assessment data and the new property tax (IMI). Comparing the two bases we can then identify the strengths and weaknesses of each one. The findings revealed statistically significant evidence that both horizontal equity and vertical equity improved with the hedonic results if judged against the new property tax.

Second Essay:

This essay explores incumbency advantage as the result of a signalling game and supposes that incumbency advantage crucially depends on the level of information that voters hold about the incumbent candidate and the alignment of preferences between the electorate and the incumbent. From the principal-agent model we specify a two period electoral model with a continuum of voters and two parties, the incumbent and the challenger. An interim election period decides by majority vote whether the incumbent is re-elected or replaced in the government by the challenger. However, not all candidates are equal. They vary in their competence or ability. The one who holds office determines the implemented policy. Every equilibrium of the model shows that elections at least serve as a disciplining device.

Conventional wisdom is that good economic conditions and expansionary fiscal policy help incumbents get reelected and empirical research on pre-electoral manipulation has mainly focused on the existence of pre-electoral expansions themselves rather than on whether expansionary policies or good economic outcomes succeed in attracting votes. So we used the information about reputation (personality, competency, ability) of local incumbents, measured by number of years in power and the vote share at the last election, together with information about economic performance outcomes, considering as a proxy the investment expenses to estimate a value to the probability of reelection. This approach is applied to a very large data panel from Portuguese local governments, which offers a number of attractive features and the results show that the likelihood of reelection of local leaders increases with their reputation and economic performance.

Finally, with the estimated value of the likelihood of reelection we test its impact on the financial costs of municipal debt and we find that an increase in the incumbent party's reelection probability is associated with a reduction of financial costs.

Índice

LISTA DE QUADROS E TABELAS	XV
LISTA DE GRÁFICOS E FIGURAS	XVII
INTRODUÇÃO	1
 PARTE I	
PREFÁCIO	10
1. ENQUADRAMENTO DA TRIBUTAÇÃO DO PATRIMÓNIO	14
1.1. BREVE ENQUADRAMENTO METODOLÓGICO	14
1.2. BREVE ENQUADRAMENTO HISTÓRICO	19
1.3. ENQUADRAMENTO ECONÓMICO	24
1.3.1. Efeitos sobre o momento da promoção/urbanização	24
1.3.2. Incidência económica: imposto sobre a propriedade versus imposto sobre a terra.....	26
2. BENS DIFERENCIADOS	29
2.1. BENS DIFERENCIADOS E A TEORIA DO CONSUMIDOR	29
2.2. O MERCADO DOS BENS DIFERENCIADOS	34
2.2.1. Enquadramento do problema.....	34
2.2.2. O mercado propriamente dito	37
2.3. GENERALIZAÇÃO DA PERSPECTIVA HEDÓNICA AOS MERCADOS DE PRODUTOS DIFERENCIADOS	49

2.3.1. Enquadramento do problema.....	49
2.3.2. Função preço.....	53
2.3.3. Identificação da função preço e das características não observadas.....	57
2.3.3.1. Identificação utilizando a independência.....	57
2.3.3.2. Identificação baseada em pacotes opcionais.....	62
2.3.4. Identificação com um conjunto relevante de funções preço.....	65
2.3.5. Identificação das preferências.....	67
2.3.5.1. Conjunto escolha contínuo com um número finito de observações por indivíduo.....	67
2.3.5.2. Conjunto escolha discreto.....	70
2.3.5.3. Decisões de não-aquisição.....	74
APÊNDICE A. Prova do Equilíbrio de Pareto.....	75
3. O MODELO-BASE DE ESTIMAÇÃO.....	79
3.1. MODELO TEÓRICO.....	79
3.2. MODELO EMPÍRICO.....	86
3.3. CONCEPTUALIZAÇÃO DO MODELO DE PREÇOS HEDÓNICOS.....	93
APÊNDICE B. Função de Licitação do Preço do Consumidor.....	99
4. MODELO EMPÍRICO E AVALIAÇÃO ECONOMÉTRICA.....	105
4.1. METODOLOGIA E BASE DE DADOS.....	105
4.2. FORMAS FUNCIONAIS DOS MODELOS DE PREÇOS HEDÓNICOS.....	110
4.2.1. Forma funcional linear.....	110
4.2.2. Forma funcional logarítmica.....	111
4.2.3. Forma funcional quadrática.....	113
4.2.4. Forma transformada de Box-Cox.....	114
4.3. COMPARAÇÃO DAS MEDIDAS DE ERRO (EX POST) DE PREVISÃO.....	116
4.4. O MODELO DE PREÇOS HEDÓNICO SEMI-LOGARÍTMICO.....	117
4.5. ANÁLISE DA CONSISTÊNCIA ESTATÍSTICA DO MODELO ESCOLHIDO.....	119

4.5.1. A Especificação do Modelo	119
4.5.2. A Escolha das Variáveis Independentes e a Especificação do Modelo.....	123
4.5.3. A Heteroscedasticidade e a Consistência do Modelo Semi-logarítmico	127
4.5.4. Correção da heteroscedasticidade e a estimação generalizada dos mínimos quadrados	132
4.6. MULTICOLINEARIDADE ENTRE OS REGRESSORES.....	134
4.7. AUTOCORRELAÇÃO ESPACIAL	139
4.8. ESTIMAÇÃO SEMIPARAMÉTRICA	148
APÊNDICE C. Inquérito.	156
APÊNDICE D. Índice e Estatística Descritiva das Variáveis	157
APÊNDICE E. Programas Eviews para Estimação Paramétrica.....	159
APÊNDICE F. Matriz das Correlações das Variáveis Independentes	162
APÊNDICE G. Resultados de Testes de Multicolinearidade	163
APÊNDICE H. Programa Eviews para Estimação Semiparamétrica.....	165
5. DO MODELO DE AVALIAÇÃO BASE DESTE ENSAIO AO MODELO SAÍDO DO IMI	172
5.1. CONSIDERAÇÕES PRÉVIAS.....	172
5.2. DOS VALORES MATRICIAIS DO ANO DE INSCRIÇÃO AOS VALORES MATRICIAIS CORRIGIDOS	175
5.3. DOS VALORES DE AVALIAÇÃO HEDÓNICA AOS VALORES DE MERCADO	178
5.4. DO RÁCIO DE AVALIAÇÃO AO COEFICIENTE DE DISPERSÃO	180
5.5. DA RECEITA FISCAL À ELASTICIDADE DA RECEITA FISCAL	185
5.6. DAS TAXAS EFECTIVAS DO IMPOSTO PATRIMONIAL	188
5.7. DA APLICAÇÃO DA PORTARIA Nº 982/2004 À BASE DE DADOS.....	189
5.8. DA PESQUISA DO ENVIESAMENTO AO AJUSTAMENTO	195
5.9. CONSIDERAÇÕES FINAIS	202

APÊNDICE I. Valores das Avaliações Realizadas	206
BIBLIOGRAFIA	208
 PARTE II	
PREFÁCIO	226
1. COMPORTAMENTO DOS GOVERNANTES NO CONTEXTO DE UM JOGO DE SINAIS	229
1.1. INTRODUÇÃO.....	229
1.2. MODELO.....	230
1.2.1. Pressupostos Básicos.....	230
1.2.2. Probabilidades Condicionais	235
1.2.3. Estratégia do político.....	239
1.2.4. Equilíbrio.....	242
1.2.4.1. Definição	242
1.2.4.2. Equilíbrio Agregado “Pooling”	243
1.2.4.2. Equilíbrio Híbrido	246
APÊNDICE A. Dedução da Equação (1.2.23)	249
2. APLICAÇÃO AO MODELO DE CICLO POLÍTICO-ECONÓMICO	252
2.1. PRESSUPOSTOS.....	253
2.2. EQUILÍBRIO	258
APÊNDICE B. Preferências Monotónas dos Votantes	278
APÊNDICE C. Prova da Proposição 1	285
3. MODELO EMPÍRICO E FONTES DE DADOS.....	288
3.1. MODELO EMPÍRICO	288
3.2. FONTES DOS DADOS ESTATÍSTICOS.....	295

4. ANÁLISE ECONOMETRICA.....	297
4.1. AVALIAÇÃO DA ESTACIONARIEDADE.....	297
4.1.1. Introdução	297
4.1.2. Caracterização dos testes: identificação de semelhanças e diferenças	300
4.1.3. Testes de raiz unitária: aplicação e resultados.....	308
4.2. ESPECIFICAÇÃO DO MODELO.....	310
4.2.1. Questões Metodológicas	310
4.2.2. Validação Cruzada e a Heterogeneidade dos Dados.....	312
4.3. ESTIMADORES DE EFEITOS FIXOS.....	319
4.3.1. Abordagem Teórica da Estimação	319
4.3.2. Abordagem técnica da Estimação	326
4.3.2.1. Exploração da estrutura em painel do modelo.....	326
4.3.2.2. Teste à presença de dependência seccional no modelo agregado/pooled	332
4.3.2.3. Estimação do modelo e resultados.....	335
4.3.2.4. Apuramento da melhor especificação	343
4.4. ESTIMADORES DE EFEITOS ALEATÓRIOS	350
4.4.1. Abordagem teórica da estimação	350
4.4.2. Abordagem técnica da estimação	357
4.5. ESCOLHA ENTRE O PRESSUPOSTO DOS EFEITOS FIXOS E DOS EFEITOS ALEATÓRIOS	359
4.5.1. Testes de especificação de Hausman	362
4.5.2. Aplicação de testes de especificação de Hausman.....	365
4.6. ESTIMAÇÃO DA PROBABILIDADE DO INCUMBENTE VENCER AS ELEIÇÕES	368
4.7. IDENTIFICAÇÃO DE CICLOS POLÍTICO-ECONÓMICOS A NÍVEL LOCAL	371
4.8. CONCLUSÕES.....	377
APÊNDICE D. Descrição Sumária dos Testes de Raiz Unitária.....	380
APÊNDICE E. Total de Municípios Estacionários	388

APÊNDICE F. Transformação <i>dentro e entre</i> grupos.....	389
APÊNDICE G. Programa em EViews 5 para o Teste de Hausman	393
APÊNDICE H. Valores das Probabilidades de Reeleição por Município.....	394
5. O MODELO DE ESCOLHA BINÁRIA E A PROBABILIDADE ESTIMADA DE REELEIÇÃO	399
5.1. COLOCAÇÃO TEÓRICA DO PROBLEMA	399
5.2. MODELIZAÇÃO DO PROBLEMA	403
5.3. MÉTODOS DE ESTIMAÇÃO: LOGIT <i>VERSUS</i> PROBIT	406
5.4. ABORDAGEM TÉCNICA DA ESTIMAÇÃO	414
5.4.1. Modelo logit.....	415
5.5. MEDIDAS DE AVALIAÇÃO DO AJUSTAMENTO DO MODELO À BASE DE DADOS...	426
5.5.1. McFadden Pseudo- R^2	426
5.5.2. Testes de Hosmer/Lemeshow.....	427
5.5.3. Percentagem de Previsões Correctas	428
5.6. TESTES DE ESPECIFICAÇÃO DO MODELO	429
5.6.1. Teste do rácio verosimilhança	429
5.6.2. Heteroscedasticidade nos modelos logit	430
5.7. IDENTIFICAÇÃO DE CICLOS POLÍTICO-ECONÓMICOS A NÍVEL LOCAL	431
5.8. CONCLUSÃO	435
BIBLIOGRAFIA.....	437
COMENTÁRIOS FINAIS	448

Lista de Quadros e Tabelas

PARTE I

Quadro 1.1-1:	Teorias Normativas da Tributação	16
Quadro 3.2-1:	Tipologia Box-Cox de especificações	92
Tabela 4.3-1:	Medidas de erro médio por especificação	116
Tabela 4.3-2:	Medidas de erro por especificação	117
Tabela 4.5-1:	Medidas de erro médio (<i>MQG</i>)	133
Tabela 4.7-1:	Alinhamento dos Bairros	144
Tabela 4.7-2:	Resultados da Autocorrelação	147
Tabela 5.2-1:	Valores inscritos nas Matrizes e respectiva Correção	175
Tabela 5.3-1:	Valores relativos à Avaliação Hedónica	178
Tabela 5.3-2:	Valores de Mercado	179
Tabela 5.4-1:	Valores dos Rácios de Avaliação	181
Tabela 5.4-2:	Coefficiente de Dispersão dos Rácios de Avaliação	183
Tabela 5.5-1:	Receita Fiscal do Imposto Patrimonial	187
Tabela 5.6-1:	Taxa efectiva do Imposto Patrimonial	188
Tabela 5.7-1:	Bases de Incidência reflectindo as avaliações utilizadas	190
Tabela 5.7-2:	Rácios de Avaliação decorrentes da Portaria nº 982/2004	191
Tabela 5.7-3:	Receitas do Imposto Patrimonial em função das diversas bases de incidência	193
Tabela 5.7-4:	Taxa efectiva após avaliação conforme Portaria nº 982/2004	194

PARTE II

Tabela 3.2.1:	Estatísticas Descritivas	296
Tabela 4.1.3:	Testes Painel de Raiz Unitária	309
Tabela 4.3.2-1:	Resultados de regressões em painel, modelo efeitos fixos	341
Tabela 4.4.2-1:	Resultados de regressões em painel, modelo efeitos aleatórios	358
Tabela 4.7.1:	Resultados de regressões em painel	373
Tabela 5.4.1-1:	Coeficientes Estimados nas Regressões	420
Tabela 5.4.1-2:	Valores <i>AIC</i> e <i>BIC</i>	422
Tabela 5.5.3:	Valores das Previsões	428
Tabela 5.6.2-1:	Valores do erro padrão para os coeficientes estimados	431
Tabela 5.6.2-2:	Valores do erro padrão para os coeficientes estimados	431
Tabela 5.7-1:	Valores da estatística descritiva	432
Tabela 5.7-2:	Resultados das estimações relativas à equação 4.7.1	433

Lista de Gráficos e Figuras

PARTE I

Gráfico 1:	Famílias Homógeneas	101
Gráfico 2:	Preferências Quase-lineares	103
Gráfico 3:	Teste de predição N-etapas	123
Figura 1:	Distribuição espacial dos bairros	142

PARTE II

Figura 1:	Árvore de Decisão	234
Figura 2-1:	Sequência de Acontecimentos	271
Figura 2-2:	Soluções de Equilíbrio	277
Gráfico 1:	Curvas de Indiferença de três jogadores	280
Gráfico 2:	Conjunto de políticas exequíveis	282
Gráfico 3:	Preferências do votante	284

Introdução

It is the theory that decides what can be observed

Albert Einstein

O presente trabalho de dissertação é constituído por dois ensaios, versando dois temas distintos: tributação do património e determinantes económicas no sucesso político dos governantes, que se podem enquadrar nos domínios da economia pública e/ou da economia política. Metodologicamente, o trabalho de investigação, que o suporta, desenvolveu-se numa perspectiva normativa e numa perspectiva positiva, já que a economia pública é, talvez, um dos ramos da economia com maior apetência para a combinação dessas duas metodologias.

Assim, no primeiro ensaio perseguiu-se um enquadramento normativo aquando da apresentação do *desenho* institucional e comparativo da tributação do património; para, posteriormente, se assumir uma perspectiva positiva na análise teórica e empírica desenvolvida para esclarecimento das vantagens de um processo de avaliação, norteados pela gestão de um equilíbrio, permanente, entre eficiência e equidade fiscal.¹ Já, no segundo ensaio, a introdução da questão de endogeneização da probabilidade de um político ser reeleito, ao vir sustentada em teoremas relativos à expressão e realização de valores, fundamentados em postulados, geralmente aceites como verdadeiros, desenvolveu-se no universo da teoria normativa, e, posteriormente, é a vez da perspectiva positiva da teoria económica se revelar em contexto de métodos de estimação de uma variável directa de probabilidade de reeleição.

Os temas tratados em cada um dos dois ensaios, à partida, parecem distintos, no entanto uma incursão mais detalhada pelo domínio das finanças públicas revela a temática da tributação como um território de excelência para incursões da escolha pública. Uma compreensão perfeita do processo das finanças públicas requer uma análise dos

¹ Benton, A. e Vernor J., *in*, Seeking equity in the ad valorem taxation of anchor department stores, *Appraisal Journal*, 1997, reflectem essa perspectiva.

processos de tomada de decisão política, os quais determinam como os governos aumentam e gastam as receitas. Ou seja, uma completa compreensão das finanças públicas requer que as mesmas sejam analisadas numa perspectiva de escolha pública. A importância das considerações de escolha pública para o estudo das finanças públicas apresenta-se razoável a partir da redefinição da problemática das finanças públicas em termos de uma visão contratualista do Estado, isto é, desde que a abordagem das finanças públicas se perspectivou em termos de troca voluntária (*voluntary exchange theory of public finance*).² Há hoje uma grande convergência relativamente à escolha pública aplicada à finança pública, na medida em que a escolha pública se perspectiva como ramo de ciência política e como ramo de finanças públicas. O ramo da ciência política utiliza o paradigma da escolha racional para analisar os fenómenos políticos, o ramo das finanças públicas aplica o paradigma da escolha racional às decisões de tributação e de despesa tomadas pelos governos.

Antes da escolha pública raramente se considerava o ambiente político em que uma política fiscal se efectuava. Ora, acontece que uma legislatura, sob os auspícios de um órgão legislativo, ao desenhar a estrutura fiscal e ao conferir-lhe a forma de lei, corre sérios riscos de se aproximar de uma qualquer forma de mercado, em que as propostas fiscais competem entre si e as decisões políticas são tomadas sobre o equilíbrio das forças em contenda. Enquanto o sistema fiscal estiver sujeito a negociação permanente, os contribuintes terão sempre um incentivo a incorrer em custos políticos para acautelar alterações de impostos que os possam afectar perversamente.³ Decorrendo, daí, vantagens numa constituição fiscal que defina a estrutura fiscal base e apenas possa ser alterada se existir um consenso alargado para o efeito.

A aplicação dos princípios de escolha pública à política fiscal sugere a fixação de impostos de acordo com o princípio de benefício da tributação e a incorporação da estrutura fiscal numa constituição fiscal como forma de minimizar os custos políticos da

2 De Viti de Marco e Pantaleoni (escola italiana) e Wicksell e Lindhal (escola sueca) constituem um bom ponto de partida pois as suas ideias são muitas vezes citadas como inspiradoras do moderno paradigma da escolha pública. Acredita-se que a compreensão das actividades governamentais (despesas e receitas/tributação) exige a compreensão do processo de tomada da decisão política que determina essas actividades.

3 Desde que os contribuintes possam ganhar por conseguirem a redução de impostos, ponderarão entrar no mercado político do *lobbying* para a descida dos impostos. O contribuinte efectivo e potencial tem sempre um incentivo a incorrer em custos políticos para acautelar mudanças nos impostos que o possam afectar de forma adversa. Por isso, tornar mais difícil a alteração do sistema fiscal reduz o incentivo a suportar custos políticos para o conseguir.

tributação. É neste enquadramento teórico que se procurou desenvolver o tema do primeiro ensaio, já que ao nos propormos reflectir sobre um conjunto de características estruturais da propriedade imobiliária e, ainda, a característica espacial no preço da propriedade para efeitos fiscais, perseguimos a construção de uma base objectiva para distribuir a carga excedentária fiscal do imposto de forma equitativa. A prossecução deste objectivo tem subjacente que, em termos da economia política das finanças públicas, a elevada visibilidade do imposto sobre a propriedade proporciona uma elevada consciencialização dos cidadãos sobre o custo dos programas públicos locais.

Curiosamente, é esta “capitalização” das características da propriedade imobiliária, sejam estruturais, sejam espaciais, sejam ambientais, sejam sociais, sejam acessibilidades, no preço da propriedade, que convida a reflectir a escolha pública numa outra dimensão. Efectivamente, por essa via, as escolhas resultantes do processo de decisão política acabam reflectidas no preço da propriedade. Definitivamente, os valores da propriedade, no quadro deste procedimento, acabarão afectados por realizações de escolha pública que potenciem (ou não) acessibilidades, que descongestionem (ou não) serviços de saúde, de segurança e de educação, que (re)qualifiquem (ou não) os espaços urbanos, que reduzam (ou não) a tributação fiscal, entre outras.

Empreendendo, agora, incursão idêntica sobre o segundo ensaio, impulsionados pelo mesmo objectivo, estabelecimento de pontes entre a escolha pública e as finanças públicas, acabamos cativos da literatura sobre votos e despesa pública. Com o intuito de se observar a tendência (ou não) para o envolvimento dos governos em ciclos político-económicos, em regimes democráticos, sem restrições constitucionais, impunha-se o apuramento da origem de quaisquer défices e de quaisquer ciclos político-económicos, de forma a imputar a responsabilidade dos mesmos ao processo democrático em si e não a uma política económica de inspiração keynesiana, adoptada pelo governo. Ora, o desenvolvimento do trabalho nessa perspectiva acaba por colocar no domínio da análise, quer o processo da decisão política, quer a teoria das finanças públicas. As despesas de investimento, reais, *per capita*, municipais, constituem, uma das variáveis exógenas que suporta o cálculo da probabilidade de reeleição; variável esta que, posteriormente, suporta a determinação do ciclo político-eleitoral materializado sobre o rácio de endividamento do município.

O trabalho empírico desenvolvido é suportado por um modelo teórico delineado em linha com a teoria da agência política e o respectivo enquadramento, quer se faça sobre modelos de 1ª geração [Barro (1973) e Ferejohn (1986)],⁴ quer sobre modelos de 2ª geração [Banks e Sundaram (1998), Coate e Morris (1995), Fearon (1999) e Rogoff, (1990)],⁵ perspectiva as eleições como um instrumento de incentivo ao político, mas também como um instrumento de selecção dos políticos. Esta concepção das eleições é intrínseca à óptica da escolha pública,⁶ os políticos, quer persigam causas egoístas, quer persigam causas altruístas, oferecem pacotes de políticas contra recompensas, basicamente, privadas, incluindo poder, prestígio e dinheiro e constitui motivo para que tudo venha resumido a um voto económico. Esta responsabilização dos políticos suportada no voto económico, ainda que constituindo uma visão minimalista das eleições, mesmo assim legitima a democracia como forma de governo que contemporiza o exercício de controlo pelo eleitorado, conferindo-lhe o poder de decidir *o que fazer com o político*, por via do voto.

Neste capítulo importa referir que o voto económico pode comprometer a responsabilização (*accountability*) democrática, já que a recompensa/punição atribuída pelos votantes ao representante político pode estar errada, o que pode decorrer do desconhecimento por aqueles da estrutura institucional do sistema político e respectivas competências. Mesmo que a percepção e a avaliação económica dos votantes sejam rigorosas e mesmo que pretendam atribuir responsabilidade às instituições competentes, ainda assim podem não ser capazes de o fazer (a cadeia da responsabilização, *accountability*, do voto económico é quebrada).⁷ Em suma, a intensidade do impacto da economia no comportamento do votante depende da facilidade de observação das políticas e da possibilidade de imputação aos diversos actores políticos, por sua vez

4 As eleições são vistas como instrumento sancionatório, os políticos fazem o que os votantes pretendem.

5 As eleições, além de sancionatórias, são instrumento de selecção.

6 Dada a equivalência das motivações comportamentais do *homo politicus* e do *homo economicus*, proposta pelos pioneiros da escolha pública (Downs, A., 1957, *AN ECONOMIC THEORY OF DEMOCRACY*, New York: Harper and Row e Buchanan e Tullock, 1962, *THE CALCULUS OF CONSENT*, Ann Arbor: University of Michigan Press.).

7 Responsabilização (*accountability*) difere de responsabilidade (*responsibility*): mas atribuir responsabilidade aos actores políticos pelo estado da economia é uma condição necessária para que exista responsabilização, ainda que o contrário não seja sustentável. A diversidade de estruturas institucionais faculta maior ou menor “transparência de responsabilidade”, facilitando ou dificultando a tarefa dos votantes na culpabilização ou no reconhecimento (Anderson, 1995, *The Dynamics of Public Support for Coalition Governments*, *Comparative Political Studies*, Vol. 28(3), pp. 350-383; Powell, 2000, *ELECTIONS AS INSTRUMENTS OF DEMOCRACY: MAJORITARIAN AND PROPORTIONAL VISIONS*, New Haven, CT : Yale University Press e Powell e Whitten, 1993, *A cross-national analysis of economic voting: taking account of the political context*, *American Journal of Political Science*, Vol. 37, pp. 391-414).

explicadas pelo contexto institucional [Chappell & Veiga (2000),⁸ Dorussen & Taylor (2001)⁹].

No presente trabalho, o mecanismo de recompensa/punição subjacente ao processo eleitoral vem operacionalizado através da base de apoio do incumbente e das condições económicas, ainda que a ligação entre esses dois elementos acabe moderada pelo contexto político.¹⁰ É nessa óptica, com o objectivo de se conseguir sustentar a estimação da probabilidade da reeleição do político, que se explora a relação da percentagem de votos do político ganhador de eleições com os valores da permanência no poder (potenciais indicadores de aprovação/reprovação) e, simultaneamente, com resultados económicos.

Após breve sumarização da relação entre os dois ensaios, segue-se directamente para o primeiro deles, apresentando, desde logo, a motivação para o tema da tributação do património, a qual decorre do cenário (em 2001) apurado em Portugal. Cenário ilustrado, em breves linhas, por uma completa miscelânea dos valores constantes das matrizes (uns estavam desactualizados, outros decorriam de declarações falsas, outros correspondiam a valores reais e actualizados, outros correspondiam a actualizações administrativas); pela inexistência de um sistema de cadastro de todos os prédios com os respectivos valores patrimoniais actualizados e pela actualidade do valor inscrito na matriz ser tanto maior, quanto mais recente a última transacção do imóvel, resultando para imóveis idênticos grandes diferenças nos valores matriciais. Se a este cenário associarmos o facto do imposto sobre o património constituir um exemplo paradigmático de um imposto *real e objectivo*, em que o sujeito passivo é determinado pela qualidade de ser *titular de um direito real sobre a coisa*, então a busca de um modelo que objective o valor da propriedade (base tributável) constitui valor

⁸ *In*, Economics and elections in western Europe: 1960–1997, *Electoral Studies*, 19, pp. 183–197.

⁹ *In*, The political context of issue-priority voting: coalitions and economic voting in the Netherlands, 1970–1999, *Electoral Studies*, 20, pp. 399–426.

¹⁰ Mesmo no contexto de uma estrutura institucional estável, observa-se mutação na amplitude da capacidade dos votantes atribuírem responsabilidade aos actores políticos, de eleição para eleição, por via da dinâmica política gerada pelo sistema eleitoral, pelo sistema partidário, pelo processo de formação dos governos (Anderson, C., 1995-ver nota 7- e Anderson, C., 2000, Economic voting and political context: a comparative perspective, *Electoral Studies* 19, pp. 151–170; Bengtsson A., 2004, Economic voting: the effect of political context, volatility and turnout on voters' assignment of responsibility, *European Journal of Political Research*, Vol. 43, pp. 749–67 e Nadeau, *et al*, 1999, Elite economic forecasts, economic views, mass economic judgments, and presidential approval, *Journal of Politics*, Vol. 61, pp. 109-135).

fundamental para garantir os princípios da igualdade e da justiça, bem como garantir a salvaguarda e valorização do património paisagístico e arquitectónico, de modo a observar o interesse social. Nesta perspectiva, a Parte I apresenta o desenho de um modelo para avaliação da base de incidência da tributação do património imobiliário em Portugal, desenvolvido a partir de uma amostra significativa do equipamento imobiliário transaccionado na cidade da Guarda, no ano de 2002.

O desenvolvimento da *parte I*, após uma *introdução* sintética e um breve enquadramento histórico, equaciona questões como: *efeitos na promoção urbana; incidência económica; avaliação e visibilidade do imposto sobre o património*; para, finalmente, se centrar na questão nuclear: *análise e modelização do mercado da habitação, na cidade da Guarda*. Esse modelo de avaliação vem ensaiado sobre uma amostra relativa ao parque imobiliário da cidade da Guarda e para avaliar o rigor e a credibilidade dessa avaliação, *comparam-se os valores resultantes desse processo de avaliação com os valores do processo de avaliação decorrente da aplicação do Código do Imposto Municipal Imobiliário (CIMI)*.

Direccionando, agora, o percurso desta nota introdutória para o segundo ensaio, importa apresentar a motivação para o tema, *determinantes económicas no sucesso político dos governantes*, a qual decorre da avaliação probabilística dos resultados pós-eleição, enquanto consequência do bem-estar pós-eleitoral do político e nessa perspectiva avaliar o papel dos indicadores de popularidade na perturbação da probabilidade de sucesso eleitoral. Como, entretanto, esses níveis de popularidade aparecem explicados pelas variáveis económicas¹¹ e a probabilidade de sucesso eleitoral é susceptível de sofrer alterações, ao nível da sua componente exógena, via indicadores de popularidade, há espaço para desenvolver trabalho enquadrado na hipótese da avaliação do desempenho da actividade política ser função da influência e da responsabilidade das entidades políticas sobre a política económica¹² no âmbito do modelo de ciclos político-económicos, mas perspectivado para o cenário autárquico português.

11 Em Portugal, Veiga e Veiga (2004) encontra evidência empírica para a hipótese de responsabilização - com o desemprego e, em menos extensão, a inflação, a afectarem os níveis de popularidade.

12 Hipótese, sustentada por Veiga e Veiga, The impact of local and national economic conditions on legislative election results, aceite para publicação na *Applied Economics*.

A motivação referida convidava à construção de uma medida directa de probabilidade de reeleição, particularmente dirigida para os ciclos políticos nos governos das jurisdições locais, de forma a conciliar, a perspectiva subjacente aos modelos de ciclos político-económicos (as variáveis económicas explicadas por variáveis políticas) com a perspectiva intrínseca aos modelos de ciência política (a explicação para os resultados eleitorais advém de variáveis económicas exógenas). Efectivamente, o político, para além da dimensão do orçamento, também se preocupa com a possibilidade de perder as eleições, na verdade, enfrenta um custo de oportunidade ao decidir candidatar-se a um lugar de poder, elementos que suportam a ideia de que os respectivos níveis de popularidade vêm afectados pela hipótese de responsabilização.

Perseguindo esse propósito, a *parte II*, começa por situar o problema no quadro da literatura existente sobre economia política e ciência política; para desenvolver estimativas baseadas em análise *logit*, entre outras, de forma a se obter uma medida directa de probabilidade de reeleição. Finalmente, testa-se, empiricamente, o modelo de ciclos político-económicos para os municípios portugueses, no qual a probabilidade de reeleição, entretanto estimada, vem equacionada como variável exógena num contexto de endogeneização da variável económica.

PARTE I

TRIBUTAÇÃO DO PATRIMÓNIO

Palavras-chave: Bens Diferenciados; Preços implícitos; Avaliação Patrimonial; Imposto sobre a Propriedade Imobiliária.

JEL's Classification System for Journal Articles:

C14 - Semiparametric and Nonparametric Methods

C21 - Cross-Sectional Models; Spatial Models

C51 - Model Construction and Estimation

C53 - Forecasting and Other Model Applications

H71 - State and Local Taxation

H21 - Efficiency; Optimal Taxation

H22 - Incidence

R21 - Housing Demand

R31 - Housing Supply and Markets

Prefácio

O ensaio que nos propomos apresentar no âmbito da tributação da propriedade, após ponderar a problemática teórica envolvente e desenvolver as questões conceptuais suscitadas no seio da teoria das finanças públicas locais, centrar-se-á no elemento fiscal – base de incidência do imposto sobre a propriedade – e respectiva avaliação. O processo de avaliação constitui o objectivo a perseguir, norteado pela gestão de um equilíbrio, permanente, entre eficiência e equidade fiscal.

Em Portugal o imposto sobre a propriedade, denominado por Contribuição Autárquica (CA) até à entrada em vigor da reforma do imposto patrimonial, actualmente denominado por Imposto Municipal sobre Imóveis (IMI), é um imposto *real*, visa a tributação de bens económicos, de riqueza, em termos objectivos, sem que nessa tributação se reflectam as condições pessoais do contribuinte. Sendo o sujeito passivo determinado pela qualidade de ser *titular de um direito real sobre a coisa*, a CA, hoje IMI, é um exemplo paradigmático de um imposto *real e objectivo*.

Em sede de CA, hoje IMI (e porque se está perante um verdadeiro paradigma de imposto *real e objectivo*), a justiça fiscal impõe que a uniformidade se torne relevante e que se traduza em igualdade horizontal num modelo novo que *objective* o imposto, em função das características físicas e de utilização da coisa.

A constituição dos municípios como sujeitos activos de CA, actualmente de IMI, pode fundamentar-se no carácter marcadamente local das suas atribuições e competências, o que permite invocar argumentos próximos do princípio do benefício. Estar-se-ia a assumir o IMI como a contrapartida monetária dos serviços públicos locais recebidos, o que colocaria o imposto sobre o património muito próximo de um preço de mercado, e assim, implicitamente, promover-se-ia a eficiência dos serviços públicos fornecidos e nesta perspectiva o imposto/benefício não teria impacto sobre a distribuição do rendimento.

O cenário, em Portugal, há data de 2001, caracterizava-se por, (i) miscelânea dos valores constantes das matrizes, uns encontram-se desactualizados, outros decorrem de declarações falsas, outros correspondem a valores reais e actualizados, outros correspondem a actualizações administrativas; (ii) inexistência de um sistema de cadastro de todos os prédios com os respectivos valores patrimoniais actualizados e (iii) actualidade do valor inscrito na matriz vir função da última transacção realizada, resultando para prédios idênticos grandes diferenças nos valores matriciais.

No preâmbulo do Código do Imposto Municipal sobre Imóveis (IMI) equaciona-se o cenário acima referido e as distorções e iniquidades daí resultantes para demonstrar o imperativo do novo imposto, cujo objectivo fundamental é o de criar um novo sistema de determinação do valor patrimonial dos imóveis, o de actualizar os seus valores e o de repartir de forma mais justa a tributação da propriedade imobiliária. O cálculo do valor patrimonial dos edificios terá por base o custo médio da construção, as áreas de localização, o tipo de construção, as características intrínsecas aos edificios, as infra-estruturas marginantes e a envolvente urbanística.¹

Considerar como base tributável o valor patrimonial de cada prédio pode não permitir garantir os princípios da igualdade e da justiça, tão pouco atender ao interesse social na salvaguarda e valorização do património paisagístico e arquitectónico. É neste contexto que nos propomos introduzir um modelo, de base econométrica, que *objective* o valor da propriedade em função dos atributos físicos. Relativamente ao “Estudo sobre a Contribuição Autárquica”, coordenado pelo Prof. Sidónio Costa Pardal (1998) e ao novo IMI, coincidimos no objectivo de que o valor tributável dos prédios, para efeitos de imposto sobre a propriedade imobiliária, seja um valor patrimonial em função da localização, utilização e área, mas divergimos no processo de cálculo – abandonamos a presunção,² para ensaiarmos técnicas econométricas que nos permitam determinar

¹ Conforme arts.38º, 39º, 40º, 41º, 42º, 43º, 44º, Capítulo IV, Secção II, do Decreto Lei nº287/2003, de 12 de Novembro, que aprova o novo código do Imposto Municipal sobre o Património.

² Parece-nos existir algum paralelismo entre o método implícito na presunção seguida pelo Prof. S. C. Pardal e o método de avaliação da base de tributação patrimonial do “*advisor*”. No método do “*advisor*” os elementos constantes do cadastro da propriedade são parametrizados de forma a expressarem o valor da propriedade padrão. O valor de uma dada propriedade n , estima-se apurando os desvios dos respectivos elementos do cadastro em relação a esses parâmetros. A definição dos parâmetros,

coeficientes por atributos físicos considerados. Com esse modelo esperamos poder contribuir para a consolidação de uma metodologia para avaliação da base de incidência da tributação do património imobiliário em Portugal, convictos que sistemas de avaliação de imóveis orientados para o valor de mercado suportados em métodos estatísticos e numéricos geram valores para avaliação imobiliária de qualidade superior.³

O desenvolvimento da Parte I, imporá uma apresentação do imposto sobre o património, e para o cumprimento desse objectivo replica-se com o conteúdo do Capítulo 1, o qual após uma introdução sintética e um breve enquadramento histórico, centra o imposto sobre o património em termos económicos, equacionando questões como: efeitos na promoção urbana; incidência económica; avaliação e visibilidade desse imposto.

Seguidamente, e já no contexto do Capítulo 2, o tema do imposto patrimonial será canalizado para a respectiva base de incidência: *os bens imóveis*, em particular, a *habitação*. E, numa abordagem consentânea com a perspectiva da escola neoclássica, esse capítulo identificará a *habitação* como um bem diferenciado e nesse cenário analisará o mercado do bem diferenciado, de modo a consolidarem-se fundamentos que sustentem o desenho de um modelo teórico capaz de responder ao objectivo de se avaliar a *habitação* ao valor de mercado. Ultrapassada essa etapa e buscando uma coerência metodológica com o objectivo proposto, seguir-se-á o modelo base da estimação do valor da propriedade imobiliária, em termos teóricos e em termos empíricos, o qual será desenvolvido num espaço para o efeito, o do Capítulo 3.

Naturalmente, que a questão nuclear da presente proposta de investigação - *a avaliação da base do imposto sobre o património* - terá que ser analisada, quer em termos da forma funcional do modelo, quer em termos de consistência estatística do modelo, de

bem como do impacto dos desvios, em relação a esses parâmetros, no valor estimado da propriedade, compete a uma autoridade fiscal. Também o art. 62º, nº 1, alíneas a), b), c), d) e e), do Decreto Lei nº 287/2003, que aprova o novo código do imposto municipal sobre o património, parece revestir o espírito subjacente ao método do “advisor”.

³ Como presume o IAAO, International Association of Assessing Officers, 1999, em, Standard on ratio studies, *Assessment Journal*, Vol. 6(5), pp. 23-65, pesquisado em, <http://www.iaao.org/IAAO>, em Fevereiro, 2005. Esta crença foi confirmada pelo recente trabalho de Moore, J. W., 2006, Performance comparison of automated valuation models. *Journal of Property Tax Assessment & Administration*, Vol. 3(1), pp. 43-59, que utiliza diferentes modelos de avaliação para gerar valores para o mesmo conjunto de propriedades imobiliárias no mesmo momento do tempo.

modo a salvaguardar a fiabilidade, a correcção e o rigor dos valores estimados a partir do modelo econométrico, entretanto construído, e neste âmbito compete ao Capítulo 4 a descrição dos procedimentos desenvolvidos para garantir a robustez científica do modelo econométrico.

Finalmente, porque o imposto sobre o património, como realidade que nos propúnhamos estudar, conheceu uma alteração em termos legislativos e, conseqüentemente, uma grande viragem em termos do cálculo do valor da base de incidência, importava fazer um paralelismo entre o modelo de avaliação base deste ensaio e o modelo saído do IMI, e essa tarefa foi concretizada e descrita através do Capítulo 5.

1. Enquadramento da Tributação do Património

The power of taxing people and their property is essential to the very existence of government.

James Madison (1751-1836)

1.1. Breve Enquadramento Metodológico

O tema da tributação do património poder-se-á situar no domínio da economia política das finanças locais, um sub-conjunto da economia pública e o seu desenvolvimento metodológico enquadrar-se-á numa perspectiva normativa e positiva, já que a economia pública é, talvez, o ramo da economia mais sujeito à combinação das metodologias positiva e normativa.

Para a literatura das finanças públicas, o contributo da escola italiana, em particular De Viti de Marco (1936) e Pantaleoni (1883), bem como dos trabalhos de Lindhal (1919) e Wicksell (1896), traduzidos na relevância do custo de oportunidade das despesas públicas e na concepção de finanças públicas no quadro de um processo político de instituições específicas e de procedimentos de tomada de decisão, revela-se fundamental. Assim se estabelecendo os aspectos normativos e positivos da teoria.

A emergência da economia do bem-estar, primeiro na tradição *pigouviana*, depois na tradição de Samuelson (1954), reforça os aspectos normativos da teoria da economia pública. Mas o desenvolvimento da investigação no domínio da teoria da economia política das finanças públicas, embora enfatizando, inicialmente, *normas* orientadoras de políticas públicas, contribui, via Downs (1957) e Olson (1971), para uma análise positiva da economia pública.

O cenário metodológico do desenvolvimento da investigação, exigirá um enquadramento normativo aquando da apresentação do *desenho* institucional e comparativo da tributação do património; mas a análise teórica e empírica, que nos propomos realizar, para a clarificação das vantagens de eficiência e de equidade associadas à reforma fiscal, assumirá uma perspectiva positiva.

No domínio das finanças públicas a assumpção das funções do Estado [Musgrave (1973)], tem implícita o imperativo de mecanismos que assegurem recursos necessários à sua prossecução. Entre esses mecanismos, emissão de moeda, criação de dívida pública, é o sistema fiscal o mais complexo e de difícil compreensão. De facto o principal instrumento para transferir recursos do sector privado para o sector público é a fiscalidade. A tributação explícita, ainda que podendo contribuir para o abrandamento do crescimento económico, é um meio mais eficiente do que a inflação como técnica de transferência de recursos, já que os cidadãos terão melhor consciência de que estão a contribuir para a causa pública.

Ora, o imposto sobre a propriedade ao ser abordado numa perspectiva normativa, implica encarar a tributação como uma motivação económica dos governos para concretizarem essas funções Musgravianas do Estado. Com especial relevância para as funções de afectação e de estabilização.

Assim sendo perspectivar os impostos a partir da função de afectação implica encarar a tributação como meio de financiamento do aprovisionamento dos bens públicos, passando a carga fiscal a ser função dos saldos orçamentais. Já perspectivá-los partindo da função de estabilização conduz a equacioná-los como moderadores dos ciclos económicos, passando a carga fiscal a ser função do crescimento económico.

Estas duas hipóteses de análise normativa dos impostos podem ser encaradas, no entanto, como visões sub-óptimas dos impostos, já que os impostos surgem como mecanismo de reacção quer ao ciclo económico, quer ao ciclo orçamental, negligenciando-se as cargas fiscais excedentárias e os ajustamentos de custos decorrentes das constantes alterações fiscais que lhe estão subjacentes. Ora, valerá a pena equacionar a tributação, ainda numa perspectiva normativa, mas sem que seja

influenciada directamente pelo andamento da economia, elegendo para o efeito a função “Alisamento Fiscal”,⁴ aproximando-nos à conceptualização de Barro (1979, 1986).

A perspectiva normativa da tributação pode, em termos sintéticos, vir sumariada,

Função Afectação das Finanças Públicas: tributação como meio de financiar o aprovisionamento do bem público.	Função Estabilização das Finanças Públicas: tributação como meio de moderar o ciclo económico.	Função Alisamento Fiscal
Função da Tributação: “Financiamento” *Carga fiscal cresce com deficit orçamental; *Carga fiscal decresce com excedente orçamental.	Função da Tributação: “Estimulação” *Carga fiscal reduz em períodos de recessão económica; *Carga fiscal aumenta em períodos de expansão económica.	Função da Tributação: “Minimizar a carga fiscal excedentária” *Tributação globalmente independente da evolução cíclica da economia; *Alterações raras da base e das taxas fiscais.
Segundo Musgrave, 1973	Segundo Musgrave, 1973	Segundo Barro 1979, 86

Quadro 1.1-1: Teorias Normativas da Tributação

Assumindo o mecanismo fiscal como, economicamente, o mais eficiente, compreender-se-á o objectivo permanente de se potenciar o bem-estar nacional, por via de uma afectação mais eficiente dos recursos locais, o que se tem vindo a traduzir numa crescente delegação de competências do governo central para níveis de governo hierarquicamente inferiores, e consequentemente numa descentralização da fiscalidade.⁵ A assumpção dos cidadãos locais como contribuintes responsáveis e exigentes⁶ é um contributo importante para a eficiência na afectação de recursos, pois, tal como nas suas

⁴ O que conduziria a menos distorções económicas.

⁵ Para dar resposta à necessidade crescente de fontes de receitas locais gerando, também, alteração na relevância dos impostos sobre a propriedade.

⁶ Apesar da ênfase colocada na igualdade (coesão económica e social) a liberdade de acção e de opção é uma reivindicação constante do princípio de cidadania e aí o imposto sobre a propriedade pode surgir como uma fonte de força da independência local. O imposto sobre a propriedade concede ao cidadão/eleitor voz mais directa, quer na determinação do respectivo imposto, quer nas despesas dos governos locais. Há um elemento de justiça no sistema fiscal da propriedade, na perspectiva *quid pro quo*, pois, geralmente, as jurisdições com cargas fiscais mais elevadas fornecem melhores serviços.

escolhas individuais, serão capazes de conduzir a decisões de escolha pública que igualizam de forma rigorosa os benefícios e os custos marginais.

Esta descentralização fiscal tem trazido para a ordem do dia a discussão de modelos de federalismo fiscal *versus* finanças multi-jurisdicionais, e inevitavelmente preocupações com a arquitectura e implementação de sistemas fiscais locais⁷ e de transferências intergovernamentais. Perseguindo a perspectiva de eficiência, que está no cerne da descentralização, vasta literatura defende que o desenho da fiscalidade local deve estar o mais possível relacionado com os benefícios obtidos a partir dos serviços públicos locais, neste contexto os impostos sobre a propriedade apresentam-se atractivos,⁸ e fazem, ainda, sentido como parte de um sistema fiscal como um todo. E, embora relativamente pesados em termos administrativos respondem razoavelmente em termos de equidade e de eficiência.⁹

Julgar um sistema fiscal sobre a sua bondade exige avaliação de vários factores: i) relação entre o imposto/taxa e a receita fiscal gerada; ii) eficiência e equidade e iii) custos com a sua administração. As teorias subjacentes à avaliação da equidade do imposto baseiam-se no princípio da tributação do benefício e no princípio da tributação da capacidade de pagar, contribuindo para tornar a estrutura fiscal mais compreensiva. No que concerne ao imposto sobre a propriedade, seguindo a metodologia de Oates (1999) importará observar as respectivas particularidades de, eficiência, incidência e administração.

⁷ Um imposto seria verdadeiramente local caso, (1) fosse introduzido pelo governo local, (2) a base fiscal e as taxas fossem fixadas localmente, (3) fosse cobrado por esse governo e (4) a respectiva receita fiscal revertesse para esse governo local. A realidade revela que um imposto é local desde que satisfaça uma ou duas destas características.

⁸ A popularidade dos impostos sobre a propriedade como instrumento de financiamento dos orçamentos locais tem sido largamente evidenciada, p.e. Kitchen (1992) refere que 17 dos 22 países membros da OCDE os adoptaram e constituem a mais importante fonte de receitas em 9 deles e de acordo com o IAAO (1990), International Association of Assessing Officers, o imposto sobre a propriedade existe em mais de 130 países, ainda que com níveis de importância diversos. Esta popularidade não obsta as pesadas críticas decorrentes das desigualdades geradas pela sua aplicação.

⁹ Equidade traduzida por, (i) justiça baseada nos benefícios recebidos: relação impostos/benefícios recebidos face às despesas do governo; (ii) justiça baseada na capacidade de pagar: impostos similares para propriedades comparáveis e (iii) transparência: impostos desenhados de forma a que sejam claros para os contribuintes permitindo responsabilização dos políticos pelos custos governamentais. Eficiência traduzida por (i) neutralidade: imposto não distorcivo da actividade económica, incluindo decisões sobre onde viver e onde trabalhar e que tipo de equipamento construir; (ii) estabilidade: impostos não flutuam de ano para ano e (iii) administração fácil: simplicidade do sistema/facilidade de administração.

Relativamente à eficiência há que avaliar a carga excedentária da tributação sobre a propriedade, que variará em função da perspectiva que se tenha do imposto sobre a propriedade. Para a perspectiva que argumenta que a tributação da propriedade local conjugada com regulamentações urbanísticas (*zoning ordinances*), gera um sistema que é efectivamente um sistema de tributação/benefício, as decisões de aquisição, de investimento e de localização da propriedade das famílias são eficientes,¹⁰ para a corrente que identifica a tributação sobre a propriedade local com um imposto especial sobre o consumo – *excise tax* – o comportamento sobre decisões locais é distorcido e tende a desencorajar a utilização do capital.¹¹ Mesmo para a perspectiva tradicional,¹² a carga excedentária do imposto só não é suportada pelos proprietários no caso de propriedades para rendimento, e para tal o imposto sobre a propriedade tem que vir reflectido no preço do arrendamento.

A unanimidade, também, não é sobre a incidência do imposto sobre a propriedade que vai ser conseguida. Entre correntes que o impelem para um imposto regressivo, posteriormente contrariadas por estudos, relativos à elasticidade da procura/rendimento da habitação, que o aproximam de um imposto proporcional ao rendimento, chega-se a uma perspectiva que o vê essencialmente como um imposto sobre o capital, tornando-se pois, um imposto bastante progressivo na sua incidência.

Em termos da administração este imposto requer alguma sofisticação, já que não revestindo os actos de compra e/ou de venda de propriedades/habitações um carácter regular, as responsabilidades fiscais têm que, fundamentalmente, se basearem em avaliações ou em valores estimados. Ora, têm sido, precisamente, as práticas de avaliação da base de incidência do imposto a fonte de alguma injustiça na aplicação do imposto, os rácios Valor Avaliado/Valor de Mercado variam amplamente dentro da mesma jurisdição, constatando-se que reformas e melhorias nas práticas de avaliação estão longe de reduzir o problema. A importância da avaliação do património imobiliário reside no facto de constituir factor determinante para o apuramento da *carga fiscal* do imposto sobre a propriedade; expressa na *taxa efectiva* do imposto sobre cada

¹⁰ Esta a perspectiva do imposto sobre a propriedade como benefício, *benefit view*, Hamilton (1976) e Fischel (1992).

¹¹ A perspectiva *new view*, Mieszkowski (1972) e Mieszkowski e Zodrow (1989).

¹² A *tradicional view*, popularizada por Netzer (1966), que assenta numa análise de equilíbrio parcial.

contribuinte, que relaciona a factura fiscal do contribuinte decorrente do imposto sobre a propriedade com o valor de mercado dessa propriedade e na *carga fiscal*, em sentido estrito, que relaciona essa mesma factura fiscal com o rendimento da família. Assim, garantir uma base objectiva para que a *carga fiscal* do imposto seja distribuída de forma equitativa,¹³ constituirá o propósito da avaliação do património imobiliário.

Em termos da economia política das finanças públicas, a elevada visibilidade¹⁴ do imposto sobre a propriedade permite uma perfeita consciencialização dos cidadãos sobre o custo dos programas públicos locais, facto que explica alguma da elevada impopularidade deste imposto,¹⁵ mas que, também, pode servir aos decisores políticos para apelar ao julgamento dos seus programas públicos face à factura fiscal devida pelo imposto. Sendo assim esses decisores devem encontrar considerável mérito em políticas que reduzam as desvantagens aparentes do imposto sobre a propriedade, através da implementação de estruturas administrativas que visem a avaliação ao valor de mercado.¹⁶

1.2. Breve Enquadramento Histórico

No contexto dos sistemas fiscais, a história do imposto sobre a propriedade é veneranda e repleta de pergaminhos, quase tão antiga quanto a da civilização humana, pelo menos 5000 anos. Mesmo antes da existência da moeda já os governantes cobravam aos cidadãos prestações sobre a forma de colheitas ou cabeças de gado, fixadas na base das

¹³ Estudos desenvolvidos por Engle (1975), Black (1977), Ihlanfeldt (1982) e Netzer (1980), em universos seleccionados para o efeito, mostraram que, mesmo nos EUA, pioneiros nos sistemas de avaliação e com um percurso idóneo nesse domínio, as diferenças inter-jurisdicionais do rácio de avaliação, da taxa nominal de imposto e da taxa efectiva de imposto eram substanciais. Constataram que, em regra a taxa efectiva de imposto, definida como a razão entre a factura fiscal do imposto e o valor de mercado da base de incidência, era sempre inferior à taxa nominal do imposto, simplesmente porque o valor avaliado estava sempre abaixo do valor de mercado. Verificaram, ainda, diferenças inter-propriedades no rácio de avaliação dentro da mesma jurisdição; chegando à situação extrema do rácio de avaliação inversamente proporcional ao valor da propriedade, documentando a regressividade do imposto.

¹⁴ Mesmo equacionando a possibilidade do regime de propriedade poder ser fonte de alguma ilusão fiscal, já que a factura fiscal ao recair legalmente sobre os proprietários, não confere visibilidade do imposto ao arrendatário, ainda que alguma *carga fiscal* seja transferida.

¹⁵ Wassmer (1993) responsabiliza mesmo a visibilidade do imposto sobre a propriedade como factor de contenção para a sua aplicação generalizada.

¹⁶ Como referiu Clapp (1990) a bondade do imposto sobre a propriedade está dependente da avaliação do valor da propriedade.

propriedades exploradas. No trabalho de investigação de Adams C. (1993) sobre a história mundial na perspectiva da política de tributação e sua prática, são identificados impostos sobre a propriedade no Antigo Egipto, defendendo-se a tese de que o declínio do Antigo Egipto está associado a uma política de isenções nos impostos sobre a terra, desde que controlada e ocupada pelos templos e pelos sacerdotes.

No séc. VI AC, o imposto sobre a propriedade encontra paralelo no *tributum* romano. Até à Revolução Industrial, uma vez que grande parte da riqueza estava concentrada na terra para fins agrícolas, os impostos eram fixados directa ou indirectamente sobre essa riqueza - algumas ilustrações desse facto, são curiosas, conforme relata Adams C. (1993), p.e.: a punição, por evasão fiscal, da “roda”, para os proprietários da terra, perpetrada por Pedro, o Grande (1682-1725) e o entabuamento de janelas em edifícios existentes e a redução de janelas nas novas construções como meio de fugir ao imposto sobre as janelas, em Inglaterra nos anos 1700. Obviamente que a relevância desta base fiscal decorria do facto dos fluxos monetários envolvendo salários, dividendos e juros serem negligenciáveis à época. A origem do imposto sobre a propriedade, numa perspectiva histórica, ao nível da Europa, é imputada aos romanos, ainda que vestígios arqueológicos da civilização suméria a reportem a milhares de anos atrás. Já os americanos situam as origens históricas do seu imposto sobre a propriedade no período colonial, quando os senhores coloniais pagavam tributos ao Reino Unido.

A explicação para estes impostos sobre a propriedade, considerados, também numa perspectiva económica, como uma das mais antigas formas de tributação, encontra raízes nos fisiocratas franceses dos meados do séc. XVIII, incluindo o próprio Quesnay, onde se concluía que a terra era a única base lógica de tributação. Facto que decorria de os fisiocratas perspectivarem a terra agrícola como a base do único sector produtivo da economia, capaz de gerar um excedente sobre os custos reais de produção, enquanto todos os outros sectores da economia simplesmente forneciam serviços ou transformavam bens de uma forma para outra. O economista clássico D. Ricardo também sustentou que a terra constituía um factor de produção apropriado para ser tributado. Baseava a sua tese na necessidade de se irem afectando maiores parcelas de terra à produção de bens agrícolas para responder ao aumento da procura desses bens decorrente do crescimento da população. Ora, esses acréscimos de terra cultivada,

marginalmente menos produtivos, exigiriam maior quantidade de trabalho para as tornar mais produtivas, induzindo custos marginais superiores, que se reflectiriam em preços mais elevados para os produtos agrícolas. Esses preços mais altos, por sua vez fariam aumentar as receitas daqueles que possuem as terras mais produtivas e que não suportam custos marginais crescentes. Este aumento de receita, para Ricardo, representava “o poder original e indestrutível do solo”. Ricardo considerava que este excedente deveria ser tributado já que ele não representava o custo real da produção.

A perspectiva dos fisiocratas e de Ricardo, no séc. XIX, encontrou eco nos EUA, num movimento favorável a um imposto singular suportado no rendimento do factor terra, liderado por Henry George. George (1879) considerava que o retorno do factor terra representava uma renda económica que poderia ser tributada livremente sem distorcer a afectação de recursos. Na terminologia actual uma renda económica é definida como o diferencial de retorno relativo a uma qualquer utilização de um factor de produção *face* ao retorno que esse factor poderia receber no seu melhor emprego. Nesta óptica, para um qualquer factor de produção, tributar a renda económica não resulta em transferir a utilização desse factor para um outro qualquer uso. Os georgistas consideravam que qualquer aumento no valor da terra resultante do crescimento da economia seria visto como um excedente criado pela sociedade e podia ser correctamente tributado, defendiam que um imposto sobre o valor da terra era não distorcivo e esta perspectiva foi defendida modernamente por Wildasin (1982), Vickrey (1970) e Tideman (1982) argumentando que o imposto incide sobre a renda pura da terra, definida como o valor da terra independentemente do que se faça com essa terra ao contrário de Bentick (1982) e de Mills (1981) que defendiam que o valor de mercado da terra reflecte o uso presente e futuro da terra.

Infelizmente os fisiocratas, Ricardo e George falharam no reconhecimento de que as rendas se podiam atribuir a todos os factores de produção. Sobre a análise marginal todos os factores de produção – trabalho, capital, bem como a terra – são remunerados pelo valor do seu produto marginal numa economia concorrencial. A terra não pode, isoladamente, ser vista como recebendo o valor residual de toda a produção. Como o notou Fischel (1985, pp. 16-17) “A única diferença económica entre a terra e os outros factores de produção é que [terra] tem uma oferta rígida decorrente da sua imobilidade”

Esta característica oferta rígida em consequência do seu imobilismo é uma característica importante a reter para a compreensão dos efeitos económicos da tributação do património.

A evolução gradual do sistema de propriedade conduziu, contemporaneamente, a que a base de tributação se alargasse, para além da terra e das construções, às propriedades pessoais tangíveis – carros de luxo, jóias, barcos de recreio, entre outros – e às propriedades intangíveis – instrumentos financeiros como acções, obrigações, depósitos bancários – hoje fala-se cada vez mais em *real estate property* e *personal property*.

Em Portugal, até 1988, existia uma ambiguidade relativamente à incidência dos sucessivos impostos prediais. Desde a velha jugada, à "décima de prédios" e à Contribuição Predial, o que efectivamente se tributou foi o *rendimento* auferido pelo proprietário através da exploração de imóveis, ainda que de forma indirecta. Durante séculos tributou-se o proprietário, mas não se considerou o imóvel "*per se*", base directa de incidência do imposto. A Lei nº 106/88, de 17 de Setembro e o Decreto Lei nº 442-C/88, de 30 de Novembro ao autorizarem a revisão e ao instituírem um novo imposto, a Contribuição Autárquica (CA), revelaram uma inflexão nas orientações relativamente à tributação de bens imóveis, já que a base de incidência da CA passa a ser o valor patrimonial dos prédios. Tem-se, então, a partir dessa data, um novo imposto predial, cuja receita é afectada, por decisão legislativa, aos municípios, os quais são o sujeito activo deste imposto.

Um dos elementos chave da Contribuição Autárquica (CA), como de qualquer outro imposto, é a base de incidência que, no sistema actual¹⁷ é o "valor patrimonial do prédio" - cujo cálculo é remetido pela lei para o Código de Avaliações. Enquanto o referido Código não for estabelecido, o imposto incide sobre o "valor tributável do prédio", o qual resulta de uma avaliação feita de acordo com as regras do Código da Contribuição Predial e Imposto sobre a Indústria Agrícola. Passa-se a estar perante dois referenciais bem distintos: o valor patrimonial e o valor tributável, que só são coincidentes se, essa, for a vontade expressada pelo legislador. O valor patrimonial, base para o cálculo da colecta, é o valor inscrito na matriz fiscal. O valor médio dos

¹⁷ À data da redacção deste ensaio.

prédios reflecte o grau de desactualização das matrizes, agravada ainda pela distribuição irregular. Em 1995, 53,6% do total dos contribuintes registados estavam abrangidos pela isenção técnica, o mesmo é dizer o valor dos respectivos patrimónios estavam avaliados na matriz por menos de 100 mil escudos. Para a actualização dos valores tributáveis tem-se vindo a aplicar regras de avaliação que têm permitido a discricionariedade e arbitrariedade, que podem ser exercidas aos mais diversos níveis do aparelho administrativo.

Entretanto, o Código de Avaliações, ferramenta crucial para a consagração de critérios objectivos para o apuramento do valor patrimonial dos imóveis, ou seja para a determinação objectiva da base de incidência da CA, nunca foi publicado. Esse facto garantia actualidade às razões expostas, aquando da reforma de 1988. É pois num cenário de iniquidades e de distorções fiscais de toda a ordem, decorrentes da espera eterna pelo famigerado Código, que surge o novo Imposto Municipal sobre o Património (IMI), definitivamente decidido a criar um novo sistema de determinação do valor patrimonial dos imóveis¹⁸ e a actualizar os valores patrimoniais.¹⁹

Actualmente, poder-se-ão inventariar basicamente duas metodologias de avaliação. Uma que se fundamenta em informações directas do mercado, relativamente às propriedades, em processo de avaliação.²⁰ A outra baseia-se em métodos de extrapolação, a partir da observação de uma amostra de propriedades, na prática denominadas de *mass appraisal*.²¹ Em qualquer uma das perspectivas de aplicação do sistema de avaliação está subjacente a construção de um modelo, processo que inclui um modelo básico de especificação e estrutura e um modelo de calibração. A

¹⁸ Para a determinação do valor patrimonial, constante do Código do IMI, foram acolhidas as recomendações do relatório da Comissão de desenvolvimento da Reforma Fiscal, bem como os critérios do anteprojecto do Código de Avaliações, elaborado em 1991, actualizados mais tarde no âmbito da Comissão de Reforma da Tributação do Património.

¹⁹ Para a actualização dos valores patrimoniais, o Código do IMI introduz mecanismos para concretizá-la, a partir de um processo declarativo.

²⁰ Metodologia que se afigura, em termos potenciais, como sendo a mais rigorosa, mas que, em termos práticos, pode gerar situação inversa, já que o teste de (i) objectividade se revela falível, na ausência de uma fórmula de avaliação, deixando o avaliador vulnerável a subornos e (ii) exaustividade se revela débil, pelo facto da sua aplicabilidade se limitar a propriedades efectivamente transaccionadas, não se aplicando às demais.

²¹ Esta metodologia ao exigir a análise de uma amostra de propriedades, transaccionadas recentemente, para identificação e quantificação das características físicas que pareçam influenciar o valor da propriedade, permite, mediante informação das características físicas de um qualquer imóvel, em processo de avaliação, produzir o respectivo valor. Permitindo um processo de avaliação quase isento de discricionariedade e menos dependente da honestidade dos contribuintes.

arquitectura do modelo de especificação baseia-se na teoria económica e de avaliação e na análise de mercado (p.e., tem implícita, a selecção de variáveis da oferta e da procura no mercado da habitação e a definição das suas relações, quer com o valor de mercado, quer entre si).

1.3. Enquadramento Económico

1.3.1. Efeitos sobre o momento da promoção/urbanização

A propriedade imobiliária, *real property*, constitui uma base fiscal, onde o terreno e o imóvel/construção surgem como uma entidade una e indivisa. Embora a propriedade imobiliária, seja pensada para fins fiscais, de forma singular, revela-se teoricamente importante distinguir esses dois elementos da base. Harriss, L. (1983) relativamente ao imposto sobre a propriedade americano, afirma estar-se na presença de dois impostos num só. A preocupação com esta distinção encontra explicação na moldura histórica do imposto sobre a terra, desde sempre defendido por ser neutro, não distorcer o comportamento da economia, e por ser justo. De acordo com Henry George (séc. XIX) o imposto sobre a terra não distorce a utilização da terra e é a forma mais eficiente de financiar o serviço público. O binómio terra/política fiscal surgia assim naturalmente. Isto porque o factor de produção terra difere de outros, como o capital e o trabalho, por ter uma oferta perfeitamente rígida.

Os economistas, embora concordem acerca da neutralidade dos efeitos do imposto sobre o rendimento anual da terra, já não são consensuais sobre os efeitos do imposto sobre o valor da terra.²² Chegaram a defender que o imposto sobre o valor da terra distorce o cronograma das decisões de promoção imobiliária da terra, induzindo a uma aceleração do investimento sobre a terra. Bentick (1982), Mills (1981) e, mais recentemente, Lim, Duck-Ho (1992) mostraram que se o imposto sobre a terra dependesse do valor de mercado da terra e os promotores imobiliários tivessem que escolher entre projectos de urbanização, mutuamente exclusivos, com fluxos de retorno diferentes no tempo, os

²² Ainda que muito do valor da terra seja acrescentado pela comunidade e daí a legitimidade de constituir base fiscal, uma outra parte do valor da terra decorre das despesas dos governos locais em estradas, escolas e outras facilidades.

impostos premiariam o investimento imobiliário presente em detrimento do futuro. Para Arnott (2005) as decisões dos promotores imobiliários relativamente ao momento de desenvolvimento dos projectos de urbanização e à densidade de infra-estruturas/equipamentos nesses projectos são influenciadas pelo imposto sobre a propriedade

Na perspectiva de Henry George, os impostos sobre a terra constituiriam os impostos adequados para financiar os serviços públicos locais e poder-se-ia admitir um sistema fiscal singular, isto é um sistema fundamentado, apenas, na base fiscal terra, relegando para segundo plano as infra-estruturas e os equipamentos.²³ Em parte, pelo elemento fiscal, terra, base de incidência fiscal, constituir a única excepção à regra de que *uma elevada taxa de imposto reduz a quantidade daquilo que está sujeito a imposto*. Importa referir, no entanto, que uma pesada tributação da terra acompanhada por baixa tributação sobre as infra-estruturas e os equipamentos construídos sobre ela, aumentaria a atracção por estes últimos investimentos, acelerando a urbanização, o que poderia traduzir-se em desagregação/dispersão urbana e, inevitavelmente, em custos mais elevados com o fornecimento de serviços públicos locais.

O indicador de mercado, preço, tem uma função económica para além da de obter bens, pois constitui um guia de uso, previne prejuízos no consumo, afectando recursos de acordo com a respectiva produtividade e escassez. Neste contexto, um elevado preço da terra²⁴ seria uma forma de encorajar uma melhor utilização da mesma. Alguma terra, sobretudo aquela que está subutilizada, por vezes vaga ou semi-vaga, parecendo destinada a encorajar a especulação e a estimular a subida dos preços, passaria a ser gerida com maior eficiência.

²³ Além de que uma pesada tributação sobre as infra-estruturas e equipamentos localizados sobre a terra também pode constituir obstáculo ao progresso das cidades, desencorajando a manutenção, conservação e modernização.

²⁴ Um imposto com base no valor actual da terra conduzirá, em termos administrativos, a uma actualização nas avaliações da mesma, tendencialmente, próximo do valor de mercado.

1.3.2. Incidência económica: imposto sobre a propriedade *versus* imposto sobre a terra

A referência das controvérsias teóricas, relativas ao imposto sobre a propriedade, o revivalismo do imposto sobre a terra, neutro segundo Henry George, por oposição ao imposto sobre a propriedade, distorcivo, mesmo que transformado num imposto-benefício após equacionadas, as referências de Fischel (1992) relativamente à capacidade do loteamento e urbanização, ainda que sumária parece-nos importante para introduzir o tema do imposto sobre a propriedade.

A realidade é que o imposto sobre a propriedade não é, apenas um imposto sobre a terra, na medida em que se aplica quer às estruturas/construções, quer ao terreno. E é este imposto que, consensualmente, entre economistas, se sustenta constituir um elemento distorcivo das decisões de investimento e de consumo. Admitindo que os sujeitos passivos do imposto sobre a propriedade dispõem de investimentos alternativos, tributados a taxas diversas, poderão evitar o imposto sobre a propriedade alterando o seu comportamento. É esperado que o investimento em habitação e outras propriedades decresça à medida que se acentue a receita fiscal gerada pelo imposto sobre a propriedade, ao contrário do que aconteceria se essa mesma receita fiscal tivesse origem na base tributária – terra, Brueckner (1986).

Pelo facto do imposto sobre a propriedade conseguir uma maior distorção nas decisões relativas à utilização da terra, do que o imposto sobre a terra, têm vindo a surgir posições advogando a reestruturação do primeiro no segundo, sobretudo nas áreas urbanas, deprimidas, de forma a promover o desenvolvimento económico (nos EUA, Pittsburgh tem vindo a direccionar a sua política fiscal nesse sentido). Quase todos os analistas concordam que esta reestruturação da base de incidência do imposto reduzirá em primeira instância o efeito distorcivo, resultando um aumento no investimento em edifícios/construções, e conseqüentemente um acréscimo, no longo prazo, do rácio capital/terra (obviamente que este efeito terá que ser ponderado pelo acompanhamento ou não das outras jurisdições na reorientação da base de incidência do imposto). No entanto esta apetência para uma re-direccionação do imposto sobre a propriedade

produzirá consequências mais complexas do que possa parecer. Os efeitos positivos de distribuição e de utilização da terra sobre o desenvolvimento económico, daí resultantes, emergirão clara e acentuadamente, quando a reestruturação da base de incidência do imposto se aplicar a uma pequena parte da área metropolitana, ao invés do que acontecerá se se considerar toda a área e, mesmo assim, reflectem a redução no imposto sobre a propriedade e não o aumento do imposto sobre o valor da terra.

Ao longo do tempo, economistas têm vindo a disputar, academicamente, a tipologia do imposto sobre a propriedade e a procurar evidência empírica para as sucessivas caracterizações do imposto. Há, no entanto, uma perspectiva, com alguma relevância no mundo real, que se centra na forma como o cidadão/eleitor percebe o imposto e como a mesma condiciona decisões políticas e o contributo dos economistas para tais decisões. Nesta linha, Fishel e Campbell (1996), encontram evidência substancial de que os cidadãos sentem os governos locais de forma diferente do governo central, a mesma factura fiscal não é indiferente ao cofre-destino (administração local ou administração central).

Na essência os votantes depositam um interesse particular nos seus governos locais de forma inaplicável ao governo central; daí que se acredite, controvérsias à parte, que a persistência a nível local do imposto sobre a propriedade, financie um pacote de bens/serviços públicos susceptível de valorizar a propriedade nessa jurisdição. Esta presunção decorre do imposto sobre a propriedade ser assumido como imposto-benefício na perspectiva do modelo Tiebout - Hamilton, desenhado no pressuposto da existência de planos de urbanização (*zoning*) perfeitos. Este requisito, mesmo não significando fixar valores de propriedade mínimos, como exige o modelo original de Tiebout (1956), mas apenas a fixação de atributos físicos e de utilização do imobiliário, dificilmente se cumpre – as instituições económicas que o defendem não operam nos moldes exigidos pelo modelo.

Efectivamente, a interdependência entre os preços da habitação, os serviços locais e as taxas do imposto sobre a propriedade deveria ser equacionada, pois assim a avaliação da base de incidência poderia reflectir o verdadeiro valor económico dos activos a serem tributados, por via da transformação do fluxo de benefícios/serviços locais em aumento

dos valores activos e do fluxo de custos/taxa do imposto em redução desses valores activos.²⁵

Em suma o imposto sobre a propriedade é, institucionalmente, um imposto local, embora longe de ser perfeito pelas inter-acções entre a terra/propriedade e a política fiscal, não importando o nível jurisdicional (central e/ou local) de decisão relativamente à base fiscal, ao regime de isenções e às taxas de incidência desse imposto sobre a propriedade.

Entretanto se o objectivo do imposto sobre a propriedade é promover a distribuição justa e equitativa da carga fiscal gerada, quer pelas jurisdições locais, quer pelos agentes económicos tributados, então a avaliação da base de incidência do imposto sobre a propriedade deverá produzir resultados rigorosos e equitativos. Podendo e devendo afirmar-se que o propósito da avaliação da base de incidência do imposto sobre a propriedade é o fornecimento de uma base objectiva para distribuir a carga fiscal do imposto de forma equitativa.

A objectividade da avaliação será um propósito do presente ensaio e para desenhar uma metodologia que permita esse alcance importa alertar para a especificidade do *bem* que vai ser avaliado, *o imóvel*, e em particular a *habitação*. Trata-se pois de um *bem* diferenciado que pode ser decomposto em várias características/atributos e dado que o valor patrimonial deve perseguir a meta do valor de mercado reúnem-se condições para se estar no cenário dos preços hedónicos, permitindo uma estimação do valor patrimonial, endógeno e contingencial no sentido estatístico, baseado na informação cadastral disponível.

Ora, identificar e caracterizar as especificidades dos bens diferenciados, bem como o respectivo enquadramento no seio da teoria microeconómica, será o objectivo do capítulo 2.

²⁵ Induziria, ainda, este *trade-off*, relativo ao valor da propriedade, uma monitorização do poder político por parte dos cidadãos/eleitores.

2. Bens Diferenciados

The individual can disappear, provided he leaves us [a]

photograph of His tastes.

Manual of Political Economy,

Vilfredo Pareto (1971)

2.1. Bens Diferenciados e a Teoria do Consumidor

Este capítulo pretende situar, de forma breve, a problemática dos *bens* diferenciados, concretamente o *bem* habitação, na teoria do consumidor convencional e, simultaneamente, propõe um enquadramento numa nova perspectiva.

Efectivamente, os mercados da “habitação” constituem fenómenos, algo, complexos, e nem sempre encontram suporte na aplicação das ferramentas padrão da teoria do preço. Na verdade, as especificidades do *bem* habitação, relativamente à homogeneidade e à divisibilidade²⁶, exigem, da microeconomia, alternativas à análise convencional.²⁷ E, uma delas reside na definição das preferências de consumo em função das características do *bem* habitação – p.e. área do lote, número de quartos, qualidade do espaço de localização, etc. Assume-se que a escolha do consumidor depende, apenas, das características do bem,²⁸ e assim um problema de dimensão infinita²⁹ é reduzido a

²⁶ O pressuposto da indivisibilidade nas mercadorias diferenciadas pode parecer invulgar mas é decisivo em termos técnicos e conceptuais. Mas-Colell (1975) refere que a maior parte dos exemplos, onde a diferenciação do produto se torna interessante, envolve indivisibilidade e daí a probabilidade de aceitação desse pressuposto.

²⁷ A teoria convencional da escolha do consumidor não fornece qualquer mecanismo para identificar qualidades diferentes de um mesmo *bem* como sendo *bens* distintos.

²⁸ Os *bens* deixam de ser o objecto directo da utilidade, passando as características desses *bens* a argumentos da função de utilidade. Assume-se que o consumo é uma actividade em que os *bens*, singularmente ou em combinação, constituem *inputs* e a colecção de características/atributos os *outputs*.

²⁹ A teoria da utilidade, tradicional, revela-se particularmente insuficiente, no cenário de uma economia complexa, no que concerne ao comportamento dos consumidores face a um número tendencialmente infinito de mercadorias. Assim, sempre que um espaço de mercadorias de dimensão n se expande para uma dimensão $n + 1$, a função utilidade é substituída por uma nova, já que não consegue utilizar a informação adicional. Mesmo que se recorresse a uma técnica supletiva, em que a função utilidade tomasse

uma dimensão finita e passível de aplicação do cálculo subjacente à teoria neoclássica. Com as funções utilidade deduzidas a partir das características, passa-se a um conjunto finito e as mercadorias, ao identificarem-se como vectores de características, passam de um número potencialmente infinito e variável com o tempo a um número finito e invariável com o tempo – assim o aparecimento de novas mercadorias não exige nova reorganização de preferências, mas uma nova solução para o problema de minimização de custos envolvendo as características que lhe estão intrinsecamente subjacentes. Para os economistas urbanos a racionalidade da escolha do consumidor torna-se óbvia.

A teoria do comportamento do consumidor em situações determinísticas definida, quer por Debreu (1959³⁰, 1960), quer por Uzawa (1960), tal como Lancaster (1966) e Johnson (1958) sugeriam, tinha sido deduzida a partir da premissa que “*bens são bens*”, as especificidades intrínsecas a qualquer bem, que fazem com que *um diamante seja diferente de um pedaço de pão*, foram omissas. Na verdade os bens são, apenas, o que os consumidores mais preferem³¹ e verifica-se neutralidade face às diferenças de gostos, em última análise *os bens são o que se assume como bem*.

Apesar da teoria pura ignorar a relevância das especificidades intrínsecas aos bens, houve sempre uma convicção subversiva sugerindo que o comportamento dos economistas tinha em conta essas especificidades³², aliás é esse conflito, entre a doutrina e a crença, que está na génese da pesquisa para a definição da “complementaridade intrínseca”, sucesso conseguido apenas quando Morishima (1959) abandona a teoria tradicional em prol de uma nova perspectiva, similar à que Lancaster (1966) viria a desenvolver. Para se apreender um dos aspectos mais relevantes desta nova perspectiva da teoria do consumidor, que assenta na assumpção de que a utilidade do consumidor depende das particularidades ou características do *bem* em ruptura com a

como argumentos todos os bens potenciais, incluindo os ainda não existentes, e se encarasse os preços dos bens não existentes como infinito, suprema prova de elasticidade na imaginação dos consumidores, ainda assim não se conseguiria capacidade previsionial.

³⁰ Como é sabido, a teoria de equilíbrio geral, desenvolvida em *THE THEORY OF VALUE: AN AXIOMATIC ANALYSIS OF ECONOMIC EQUILIBRIUM*, Yale University Press, contempla um mercado com um número finito de mercadorias homogéneas e perfeitamente divisíveis, onde os agentes económicos interagem através de um sistema de preços, tomados por cada um deles como um dado.

³¹ “The individual can disappear, provided he leaves us [a] photograph of his tastes.” (Vilfredo Pareto, 1906, pp.120)

³² Os exemplos de substituição, presentes nos compêndios de economia, incidem sempre sobre bens como manteiga e margarina, em vez de bens como sapatos e barcos, como se, os autores, acreditassem haver algo intrínseco à manteiga e à margarina que os torna bons substitutos. Implicitamente, identifica-se uma crença relativamente à importância das especificidades dos bens na reacção dos consumidores a esses bens.

ideia clássica de que o objecto directo da utilidade do consumidor é o *bem*, há que sumariar três pressupostos, a essência desta nova perspectiva, (cada um deles representando uma ruptura com a tradição): i) o *bem*, *per se*, não confere utilidade ao consumidor – essa responsabilidade é imputável às características desse *bem*; ii) um *bem* possui, em geral, mais do que uma característica e muitas dessas características são partilhadas por mais do que um *bem* e iii) *bens* combinados podem possuir características diferentes das que estão presentes em *bens* separados. Neste contexto o problema de escolha enfrentado pelo consumidor, relativamente à maximização da respectiva função de utilidade, na presença de uma restrição orçamental linear, traduzir-se-ia por,

$$\text{Max}U = U(\mathbf{z})$$

$$\text{s. a} \quad \mathbf{z} = \mathbf{g}(\mathbf{x})$$

$$\mathbf{p}\mathbf{x} = M$$

$$\mathbf{z}, \mathbf{x} \geq 0$$

em que M é o rendimento do consumidor, \mathbf{p} é um vector preço, \mathbf{z} é o vector dos atributos/características, e \mathbf{x} são as mercadorias adquiridas no mercado. E a função $\mathbf{z} = \mathbf{g}(\mathbf{x})$, revela como as mercadorias, \mathbf{x} , se transformam nos atributos/características, \mathbf{z} , por via da função, \mathbf{g} .³³

³³ A função \mathbf{g} , que relaciona o vector das características \mathbf{z} com o vector dos bens \mathbf{x} resulta dos pressupostos “Lancasterianos”, (1) o consumo total de *bens*, identificado no vector \mathbf{x} , é a resposta às exigências de uma dada actividade de consumo, através da forma $\mathbf{x} = \mathbf{A}\mathbf{y}$, em que \mathbf{y} é o vector dos níveis de actividade/consumo e \mathbf{A} é uma matriz que traduz as relações lineares entre \mathbf{x} e \mathbf{y} , determinadas pelas propriedades intrínsecas aos bens e o nível de conhecimento tecnológico da sociedade (p.e. o consumo de imóveis/habitação depende dos níveis de actividades desenvolvidas com eles, habitação própria, habitação rendimento, escritórios, comércio, etc.); (2) o vector dos níveis de actividade/consumo, \mathbf{y} , produz um vector de características, \mathbf{z} , através da forma $\mathbf{z} = \mathbf{B}\mathbf{y}$, de tal forma que a quantidade de características responde às actividades de consumo através da matriz linear \mathbf{B} ; e (3) há uma correspondência entre *bens* e características de “um para um”.

A solução para o problema da escolha do consumidor, no universo “Lancasteriano”, dominado por actividades de consumo, características/atributos dos *bens* de consumo e *bens* exigiria uma transformação entre o espaço-*bens* e o espaço-*características*.³⁴ Optando por transformar a função utilidade do consumidor num espaço-*bens*, o modelo apresentado passaria a,³⁵

$$\text{Max}U = U(g(\mathbf{x})) = V(\mathbf{x}),$$

$$\text{s. a} \quad \mathbf{p}\mathbf{x} = M$$

$$\mathbf{x} \geq 0$$

Este modelo introduziu no seio da teoria do consumidor uma análise fundamentada em actividades, uma estrutura de tecnologia de consumo, pressupondo desde logo que os bens possuem ou dão origem a múltiplas características, em proporções fixas, e que as preferências dos consumidores são exercidas sobre essas características. As possibilidades de relações entre actividades e características que este modelo permite, faz com que seja, na maior parte das vezes, mais rico em explicações heurísticas e em capacidade previsional do que os modelos convencionais de comportamento do consumidor.

Os resultados de Lancaster (1966) são relevantes para a teoria de preços hedónicos, quer a decomposição do preço de um *bem* composto, quer a revelação dos preços implícitos dos atributos/características, por via de métodos regressivos, encontram um bom

³⁴ A função utilidade, objectivo do problema de escolha do consumidor, acima modelada, está definida no espaço-características, Z-espaço, enquanto a função restrição orçamental está definida no espaço-bens, X-espaço.

³⁵ As propriedades de transformação entre os Z-espaço e os X-espaço, de que (i) um conjunto convexo no X-espaço é transformado num conjunto convexo no Z-espaço; (ii) uma transformação inversa não tem que necessariamente existir entre espaços e (iii) a convexidade da função utilidade em \mathbf{z} é preservada quando essa função utilidade se expressa em \mathbf{x} , decorre dos elementos da matriz \mathbf{B} serem constantes e da transformação linear $\mathbf{z} = g(\mathbf{x})$ e são suficientes para garantir que o problema de maximização de utilidade restrito à questão orçamental terá solução.

suporte teórico na escolha do consumidor expressa por uma função utilidade dependente dos atributos/características que compõem um *bem*.

Para uma maior ênfase do comportamento do consumidor, via explicitação da utilidade em função das características dos bens, Rosen (1974) não negligenciando os contributos de, Becker (1965), Lancaster (1966) e Muth (1966), aponta para a necessidade de se analisarem as propriedades do equilíbrio de mercado. Pois existe um mercado de bens diferenciados mas não existe um mercado de características/atributos e os preços hedónicos são definidos como preços implícitos dessas características/atributos e são revelados aos agentes económicos a partir dos preços observados no mercado de bens diferenciados. Entretanto esses preços implícitos podem não reflectir a avaliação do consumidor a uma dada característica/atributo mas sim a dificuldade do produtor em oferecer essa característica/atributo.

A análise hedónica de preços encontrou na perspectiva Lancasteriana da teoria do consumidor importante substrato teórico, no entanto os preços de mercado observados quando dizem respeito a *bens* diferenciados (tratados como cabazes de características/atributos) devem também ser comparados nesses termos e o conteúdo económico da relação entre os preços observados do *bem* diferenciado e as características/atributos observadas torna evidente que as diferenças entre preços servem para igualizar as diferenças entre cabazes alternativos que incorporam esse *bem*. Haveria, pois, necessidade de resolver o problema de identificação³⁶ subjacente à utilização dos preços hedónicos. A secção 2.2. procurará ensaiar reflexões sobre essa problemática.

³⁶ A função de preços hedónica não identifica nem a procura, nem a oferta de um *bem* diferenciado; de facto as observações de preços do *bem* diferenciado e das características/atributos descrevem apenas uma função envelope conjunta, não podendo identificar nem a estrutura das preferências de consumo nem as tecnologias de produção que a induziram.

2.2. O Mercado dos Bens Diferenciados

2.2.1. Enquadramento do problema

O mercado do *bem* diferenciado, especificamente o da habitação, é invulgar, já que o preço de equilíbrio não identifica um qualquer preço singular, traduzindo o retorno de um qualquer atributo/ característica desse *bem*, em vez disso identifica um *continuum* de preços. O que, *a priori*, se traduz na impossibilidade física de as famílias poderem isolar um atributo/característica de uma qualquer habitação e de o poderem usufruir independentemente dos outros atributos/características. Esta indivisibilidade intrínseca ao *bem* habitação, no espaço e no tempo, permite a Day (2001) falar em impossibilidade de arbitragem no mercado do *bem* “habitação”, via inter-acção das forças de mercado e, conseqüentemente, em inexistência de um preço de equilíbrio constante.

Assim começar-se-á por tecer algumas considerações sobre o tipo de mercado em que ocorrem as transacções de *bens* diferenciados, particularmente as transacções de habitação. Dada a natureza única do *bem* diferenciado a ser transaccionado, quer os vendedores, quer os compradores dispõem de algum poder de mercado. Entretanto, no longo prazo, os *bens* diferenciados evoluem, tendencialmente, para bens substitutos próximos, ainda que não idênticos, conduzindo a um cenário de lucros económicos nulos. O cenário seria em tudo semelhante a um modelo de concorrência monopolística de Chamberlain. Os economistas, em geral, assumem mercados de habitação, tendencialmente, concorrenciais e assumem que os intervenientes nesse mercado (consumidores e produtores) enfrentam preços exógenos. Assim, os consumidores são tomadores de preços e tomam a função de preços hedónica como um dado, obtendo utilidade a partir das características do produto diferenciado.

Estar-se-ia no quadro de um equilíbrio concorrencial, ainda que subsistam reservas acerca do realismo do mesmo, na medida em que se assume a inclusão de um vector composto pelas características do *bem* e o mesmo é negligenciado na maior parte das análises de equilíbrio geral. A pertinência destas reticências reside no facto da extensão da teoria do consumidor aos bens diferenciados introduzir, simultaneamente, duas complexidades: a indivisibilidade do bem (o consumidor vive numa casa ou não vive) e

a continuidade evolutiva do bem (o conjunto de características do *bem*, assume-se em geral como não finito).

Resolver este problema passa por considerar as referidas complexidades, uma de cada vez, assim: H_1 - as características da habitação z assumem-se finitas e nesse caso se o número de consumidores é finito é possível provar a existência de um equilíbrio concorrencial “aproximado” [Broome (1972)], cuja aproximação melhora à medida que o número de consumidores aumenta – no limite com um número contínuo de consumidores, o equilíbrio concorrencial existe [Mas-Colell (1977)]. Neste cenário as características do *bem*, utilizadas pela análise, servem apenas para identificar o tipo de habitação – não são necessárias para estabelecer a existência de equilíbrio, funcionam como veículo de parametrização das funções de utilidade e, conseqüentemente, descrevem a procura por vários tipos de habitação de uma forma que sustenta melhor a estimação empírica; H_2 – cada tipo de *bem* indivisível assume-se identificado com um ponto z , no conjunto compacto de atributos/características K desse *bem*, anulando-se a hipótese de poderem falhar os teoremas da existência e da equivalência decorrente de um número infinito de *bem* indivisível – o conjunto de afectações com propriedade de “core” pode ser superior ao conjunto das afectações concorrenciais. Heuristicamente o que se exige para a equivalência do “core” é que os mercados das diversas mercadorias/*bens* sejam “fortes”, mas a presença de um número infinito de *bens* faz falhar esse requisito. A introdução do conceito *características*,³⁷ como o faz a extensão da teoria do consumidor, permite demonstrar a existência de equilíbrio concorrencial³⁸ e a equivalência com a eficiência.

Em síntese o alargamento da teoria da escolha do consumidor assume que a função hedónica é uma função evolutiva (*continuous*) das características da habitação; o que se traduz na imposição de uniformidade, relativamente à escala das preferências do

³⁷ As *características* são elementos de um conjunto, definido como um espaço métrico compacto, K , associado à classe de *bens* indivisíveis, de tal forma que cada ponto $z \in K$, representa a descrição de um determinado tipo de bem indivisível (x). Se K é finito, por exemplo $K = \{z_1, z_2, \dots, z_K\}$ e se um consumidor escolhe uma unidade do bem indivisível tipo 4, então esse bem indivisível pode ser expresso por $x = (0, 0, 0, 1, \dots, 0)$.

³⁸ Como é demonstrado por Ellikson B.(1980).

consumidor e, ainda, no pressuposto que casas, com características muito semelhantes, são substitutas, quase perfeitas, para todos os consumidores. Entretanto esse *continuum* de características pode ser aproximado a um conjunto finito de características,³⁹ por exemplo no caso da característica *acessibilidades*, essa aproximação pode ser encontrada através da definição de um número finito de “zonas”, e posterior tratamento de todos os imóveis dentro de uma mesma “zona” como equivalentes.

A teoria dos preços, na perspectiva hedónica, é formulada como um problema de equilíbrio da economia espacial, na medida em que as decisões localizacionais, quer dos consumidores, quer dos produtores, são tomadas com base num conjunto completo de preços implícitos no espaço das características, o que, em termos gerais, pode ser encarado como uma extensão dos modelos básicos de localização residencial.

O reforço do contexto espacial decorre da (i) *bid price function*,⁴⁰ função de preço/licitação, $\theta(z_1, \dots, z_K; U, M)$, em que z_k traduz as características da habitação, U o nível de satisfação que se pretende atingir e M o rendimento, que permite ilustrar uma superfície de curvas de indiferença, relacionando as características z_k com a “moeda/bem compósito” “ c ” renunciado,⁴¹ amplamente utilizadas em economia urbana, conforme Alonso (1964); (ii) função-objectivo utilidade indirecta do consumidor, $V = \theta(\mathbf{z}, M, U)$,⁴² que ao identificar-se com a função preço/licitação do consumidor (*bid price function*) estabelece um paralelismo com a denominada formulação de Muth e, ainda, com a formulação de Alonso e (iii) tendência à segmentação de mercado que, no mercado implícito das características incorporadas nos produtos diferenciados, decorre de distintos grupos de rendimentos e de gostos, já que

³⁹ Mas-Colell resolve este problema com a convergência em estrela fraca.

⁴⁰ Definida por Rosen (1974).

⁴¹ Já que a decisão do consumidor se pode definir implicitamente por, $U = U(M - \theta, z_1, \dots, z_K)$, partindo de $U = U(c, z_1, z_2, \dots, z_K)$.

⁴² Deduzida a partir da função objectivo do consumidor, $Max_z U = U(p_c, \mathbf{z}, M - \mathbf{p}(\mathbf{z}))$, em que $\mathbf{p}(\mathbf{z})$ é a função de preços hedónica/a função de preços, dado as características da habitação, \mathbf{z} ; c é o vector dos bem compósito e p_c é o preço padrão/unitário desse bem, depois de se assumir que, $V = \mathbf{p}(\mathbf{z})$.

consumidores com funções de preço/licitação similares compram produtos com especificações semelhantes - resultado constante dos modelos de equilíbrio espacial.

2.2.2. O mercado propriamente dito

Tendo por base o trabalho pioneiro de Rosen (1974), onde se fez a análise conjunta da procura e da oferta e se mostrou que a função de preço hedónica de mercado, de cada uma das características, é uma curva/envelope da família de curvas de preço oferta/licitação (*bid*) e da família de curvas de oferta de produção (*offer*), passível de traduzir um equilíbrio de mercado⁴³ e na senda dos trabalhos de Epple (1987) e Bartik (1988), e mais recentemente de Tauchen e Witte (2001) e de Ekeland, Heckman e Nesheim (2002, 2004) vai-se procurar desenhar a arquitectura de um modelo para o mercado da habitação.

Do lado da procura do *bem* “habitação” ir-se-ão buscar fundamentos à teoria neoclássica, baseada na utilidade do consumidor e para tal, equacionar-se-ão, essencialmente, os contributos de:

(i) Lancaster (1966), que produziu uma extensão à prova de existência da função de utilidade.⁴⁴ Provando que numa economia complexa (com mais actividades do que características)⁴⁵ a escolha de eficiência satisfaz sempre, (\mathcal{I}) o axioma fraco da preferência revelada,⁴⁶ (o qual reduzido a simples axioma indutivo, pode ser satisfeito

⁴³ Assume-se que o agente económico é, a este nível de desagregação, individualmente apenas um *tomador de preços* de equilíbrio, já que os preços de mercado, revelados pela função de preços hedónicos, estabilizam quando procura e oferta de habitação são coincidentes. Embora a compreensão do equilíbrio alcançado exija a observação da procura e da oferta, agregadas.

⁴⁴ Debreu e Arrow (1954) foram os primeiros a provar a existência da utilidade, baseando-se nos quatro axiomas da preferência: exaustividade, reflexividade, transitividade e continuidade.

⁴⁵ No modelo de comportamento do consumidor, de Lancaster, o bem ou conjunto de bens consumidos é uma actividade de consumo. E, através do vector dos níveis de actividade/consumo, \mathbf{y} , é possível estabelecer-se de forma inequívoca uma relação entre $U(\mathbf{z})$ e $U(x)$, em que \mathbf{z} características e x bens.

⁴⁶ O axioma fraco da preferência revelada (WARP) não vai muito para além do teorema de Weierstrass, já que para uma dada função de utilidade, U_0 , desse teorema se requer, apenas, a continuidade da função de utilidade e o conjunto compacto, não vazio, C^h , não se necessitando conhecer as propriedades de diferenciabilidade e de dupla diferenciabilidade da função de despesa

mesmo que os consumidores consumam características em proporções fixas e, possivelmente, mesmo que tenham preferências côncavas, daí que a adjectivação “revelada” possa, apenas, apontar escolha eficiente, em vez de convexidade); chegando mesmo a satisfazer (2) o axioma forte para um número suficientemente vasto de preços – daí, a satisfação deste axioma forte não revelar a convexidade da função de preferência. Registando, ainda, que, o perfil da fronteira de eficiência e a existência de efeitos de substituição eficientes podem gerar situações de procura, caracterizadas pelas particularidades, tradicionalmente, assumidas, mesmo que não se esteja na presença de funções de utilidade convexas, alisadas e tradicionais;⁴⁷

e

(ii) Mas-Collel (1974) que demonstrou que os axiomas da exaustividade e da transitividade são supérfluos no teorema da existência de equilíbrios de Debreu, sem necessitar de recorrer a um número infinito de consumidores e sem esforços indevidos com representações da função de utilidade.

Para se deduzir a procura do *bem* diferenciado (habitação), seguindo a metodologia desenvolvida em Rosen (1974) e assimilando os contributos acima referidos, procurou-se adaptar, à realidade do mercado da habitação, o modelo hedónico, desenvolvido por Ekeland, Heckman e Nesheim (2004) para o mercado de trabalho.

Para o efeito haveria que equacionar o comportamento dos consumidores face a um cabaz de *bens*: \mathbf{z} *bem* hedónico (vector das características \mathbf{z}_k) e c *bem* compósito (assumido como numerário) e parametrizar as preferências⁴⁸ desses consumidores, de forma a reflectir a dimensão espacial subjacente aos seus comportamentos, introduzindo os parâmetros, \mathcal{G} e A na função utilidade do consumidor, para captar essas diferenças

$E(p, U_0)$, desde que se pressuponha a existência dos pontos $b \in h(p_b, U_0) = \arg \min p_b$, s.a $b \in \{d \in C^h \mid U^h(d) \geq U_0^h\}$ e

$d \in h(p_d, U_0) = \arg \min p_d$, s.a $d \in \{b \in C^h \mid U^h(b) \geq U_0^h\}$, bem definidos.

⁴⁷ Por exemplo uma função de utilidade, simples, em que as características são consumidas em proporções constantes – possivelmente as proporções variando com o rendimento – pode vir substituída pela função utilidade convencional.

⁴⁸ Facto que decorre do pressuposto de uma função de preços hedónicos, $p(\mathbf{z})$, convexa e suficientemente regular no seu domínio, para a verificação das condições de 2ª ordem do problema de $Max U$, induzindo a uma tendência natural para a segmentação do mercado.

socioeconómicas e/ou de gostos que distinguem umas famílias das outras, os quais traduzem, respectivamente, a heterogeneidade e a homogeneidade das preferências, através de um problema de,

$$\text{maximização de utilidade, } \text{Max } U = U(\mathbf{z}; c, \mathcal{G}, A)$$

sujeito à restrição orçamental,

$$M = c + p(\mathbf{z}),^{49}$$

em que \mathcal{G} é o parâmetro das preferências, passíveis de variar inter-consumidores e A é o parâmetro das preferências comuns inter-consumidores. Assume-se que o preço do bem composto, c , é unitário e a quantidade do bem diferenciado, expresso pelo vector das características, \mathbf{z} , é, também, unitária.

Do lado da oferta os intervenientes, empresas, visam a maximização do lucro, ou seja procuram,

$$\text{Max } \Pi = \bar{F} p(\mathbf{z}) - C(\bar{F}, \mathbf{z}; \nu, B),^{50}$$

em que $\bar{F} p(\mathbf{z})$ é a receita total e $C(\dots)$ é a função de custos, deduzida a partir do problema de minimização de custos, sujeito à restrição da produção, $\bar{F} = F(\mathbf{z}; \nu, B)$. Os preços dos factores de produção e as combinações tecnológicas subjacentes ao problema de minimização de custos estão reflectidos no vector ν de parâmetros tecnológicos

⁴⁹ As condições de 1ª e 2ª ordem do problema de $\text{Max } U(\mathbf{z}; M - p(\mathbf{z}), \mathcal{G}, A)$ são, respectivamente, $U_{\mathbf{z}}(\mathbf{z}; M - p(\mathbf{z}), \mathcal{G}, A) - U_c(\mathbf{z}; M - p(\mathbf{z}), \mathcal{G}, A) p_{\mathbf{z}}(\mathbf{z}) = 0$ e $U_{zz'} - U_c p_{zz'} + p_{\mathbf{z}} U_{cc'} p_{\mathbf{z}}'$ definida negativa.

⁵⁰ As condições de 1ª e 2ª ordem para se verificar um máximo, são, respectivamente, $p_{\mathbf{z}}(\mathbf{z}) \bar{F} - C_{\mathbf{z}}(\bar{F}, \mathbf{z}, \nu, B) = 0$ e $p_{zz'} - C_{zz'}$ definida negativa.

susceptíveis de variarem entre empresas e no parâmetro tecnológico B partilhado por todas as empresas.

Quer os consumidores, quer as empresas diferem nas suas preferências, \mathcal{G} , e nas suas produtividades, ν , respectivamente. Assumindo que \mathcal{G} e ν , são variáveis perfeitamente aleatórias e contínuas, pode-se identificar \mathcal{G} com uma função de distribuição da densidade de probabilidade, $f_{\mathcal{G}} = f_{\mathcal{G}}(\mathcal{G}(\mathbf{z}, p_z, M - p(\mathbf{z}), A))$ e ν com uma função de distribuição da densidade de probabilidade, $f_{\nu} = f_{\nu}(\nu(\mathbf{z}, p_z, B))$. Desde que se admita a aplicação do teorema da função implícita, dado as condições de 1ª ordem poderem ser invertidas, essas funções podem vir expressas por, $\mathcal{G} = \mathcal{G}(\mathbf{z}, p_z, M - p(\mathbf{z}), A)$ e por, $\nu = \nu(\mathbf{z}, p_z, B)$.

Definidas estas funções de distribuição de densidade de probabilidade, utilizando a transformação de coordenadas,⁵¹ pode-se encontrar a densidade de \mathbf{z} ⁵² procurado, dado $p(\mathbf{z})$, o parâmetro dos gostos A e a heterogeneidade de gostos \mathcal{G} , $f_z^d |_{p(\mathbf{z}), \mathcal{G}, A}$, e a densidade de \mathbf{z} oferecido, dado $p(\mathbf{z})$, o parâmetro tecnológico B e a diversidade de tecnologia ν , $f_z^s |_{p(\mathbf{z}), \nu, B}$, assim,

$$f_z^d |_{p(\mathbf{z}), \mathcal{G}, A} = \int f_{\mathcal{G}}(\mathcal{G}(\mathbf{z}; p_z, M - p(\mathbf{z}), A)) \det \left[\frac{\partial \mathcal{G}(\mathbf{z}, p_z, M - \mathbf{p}(\mathbf{z}), A)}{\partial \mathbf{z}} \right] d\mathcal{G}$$

⁵¹ Se \mathcal{G} tem uma função de distribuição de densidade de probabilidade, $f_{\mathcal{G}}(\mathcal{G})$ e se $z = f(\mathcal{G}, \dots)$ é uma função monótona e contínua em \mathcal{G} , então a densidade de \mathbf{z} , após a transformação de coordenadas, vem, $f_{\mathcal{G}}(\mathcal{G}) |_{\mathcal{G}'}$.

⁵² Rosen (1974) estudou o caso em que, $\dim(\mathcal{G}) \geq \dim(z)$ e $\dim(\nu) \geq \dim(z)$, porém não há perda de generalidade pressupor igualdades estritas, em vez de desigualdades.

$$f_z^s |_{p(\mathbf{z}),v,B} = \int f_v(v(\mathbf{z}, p_z, B)) \det \left[\frac{\partial v(\mathbf{z}, p_z, B)}{\partial z} \right] dv. \quad ^{53}$$

O equilíbrio de mercado exigirá que $f_z^d |_{p(\mathbf{z}),g,A} = f_z^s |_{p(\mathbf{z}),v,B}$ e a perspectiva mais directa para a sua determinação passará pela resolução dessa equação para $p(\mathbf{z})$, em termos da função de lucro das empresas, da função de utilidade dos consumidores e das distribuições f_v e f_g . E, no contexto da teoria económica, há que ter subjacente que, (1) as produtividades e as utilidades marginais são positivas, isto é, $P_m \geq 0$ e $U_m \geq 0$ e que, (2) a utilidade e o lucro devem ser superiores a um valor/reserva, isto é $U \geq \bar{U}$ e $\Pi \geq \bar{\Pi}$.⁵⁴

Dado que $p(\mathbf{z})$ ⁵⁵ se obtém a partir do ajustamento entre procura e oferta é possível encontrar paralelismo com a forma como os preços relativos são ajustados nos modelos de Arrow/Debreu sem moeda (ou, mais especificamente sem o pressuposto da utilidade quase-linear). Não é, por isso, difícil mostrar que este tipo de equilíbrio “hedónico” é um óptimo de Pareto mesmo na ausência de moeda e de preferências quase-lineares. (Prova em Apêndice A)

Esta perspectiva, num cenário de equações diferenciais, torna-se de resolução complexa e não permite uma identificação transparente dos parâmetros estruturais. Mesmo que se admita que as distribuições de densidade do parâmetro das preferências, passíveis de

⁵³ Os termos do Jacobiano são ambos positivos, dadas as condições de 2ª ordem referidas em 49 e 50, daí a densidade de z ser diferente de zero, significando que as probabilidades associadas a intervalos na amplitude de z estão associadas a intervalos na amplitude de \mathcal{G} .

⁵⁴ Estes pressupostos, também, se revelam importantes na análise do problema de identificação da solução. Relativamente aos produtores, a fixação do valor $\bar{\Pi} = 0$ permitirá a endogeneização da distribuição de densidade dos produtores f_v , por via da entrada e saída de produtores até que se alcance as condições de equilíbrio de longo prazo.

⁵⁵ A forma funcional do preço hedónico de equilíbrio é ajustada, dentro de cada subconjunto do conjunto de características, até que se verifique a igualdade entre o conjunto de consumidores que procuram uma qualquer característica e o conjunto de produtores que oferecem essa característica.

variar inter-consumidores (\mathcal{G}) e dos parâmetros tecnológicos susceptíveis de variarem entre empresas (ν), são normais, isto é $\mathcal{G} \sim N(\mu_{\mathcal{G}}, \Sigma_{\mathcal{G}})$ e $\nu \sim N(\mu_{\nu}, \Sigma_{\nu})$, o que permitiria deduzir, para qualquer ponto z , uma densidade de procura e de oferta de distribuição normal, ainda assim o equilíbrio de mercado,⁵⁶ exigiria que se verificassem as igualdades entre: (i) a média da procura e a média da oferta de z e (ii) a variância da procura e a variância da oferta de z . E no final, a determinação dos parâmetros das preferências e das tecnologias só se torna possível em certas circunstâncias – apenas se conseguiram identificar, caso houvesse apenas um tipo de consumidores ou, então, apenas um tipo de empresas.

Para ilustrar a complexidade presente à resolução da equação de equilíbrio de mercado, $f_z^d |_{p(\mathbf{z}), \mathcal{G}, A} = f_z^s |_{p(\mathbf{z}), \nu, B}$, valerá a pena apresentar, numa adaptação ao mercado da habitação, um exemplo quase-linear desenvolvido a partir do ensaio sobre modelos hedónicos de Ekeland, Heckman e Nesheim (2004) - primeiramente, analisado por Tinbergen (1956) e, também, por Epple (1987) e Tauchen e Witte (2001).

Assim, seguindo a metodologia apresentada no início desta secção, para se deduzir a procura pelo *bem* diferenciado/habitação, começa-se por pressupor uma função utilidade, quadrática em \mathbf{z} e linear em c , verificando a condição, $\dim(\mathbf{z}) = \dim(\mathcal{G})$, expressa por,

$$U(c, \mathbf{z}, \mathcal{G}, \mathbf{A}) = M - p(\mathbf{z}) - \frac{(\mathbf{z} - \mathcal{G})' \mathbf{A} (\mathbf{z} - \mathcal{G})}{2} \quad (2.2.1)$$

⁵⁶ Dado estar-se na presença de variáveis aleatórias normais.

⁵⁷ As condições necessárias e suficientes, para se determinar o óptimo do consumidor, são, $1^a -(\mathbf{z} - \mathcal{G})\mathbf{A} - p_z = 0$ e $2^a -\mathbf{A} - p_{zz}$ definida negativa. As N funções, admitindo que $z \in R^N$, que expressam as condições de 1ª ordem, traduzem as funções de preço/licitação dos consumidores.

Assume-se, ainda, que a matriz \mathbf{A} , representativa dos gostos dos consumidores, é simétrica e definida positiva e que é a mesma para todos os consumidores, e que o vector \mathcal{G} , representativo das características das preferências, segue uma distribuição normal, $\mathcal{G} \sim N(\mu_{\mathcal{G}}, \Sigma_{\mathcal{G}})$.

Já, do lado da oferta, pressupõe-se uma função lucro,

$$\Pi(\mathbf{z}; \nu, \mathbf{B}, p(\mathbf{z})) = P(\mathbf{z}) - (\mathbf{v}'\mathbf{z} + \frac{1}{2}\mathbf{z}'\mathbf{B}\mathbf{z})^{58} \quad (2.2.2)$$

em que os custos de produzir uma unidade do bem diferenciado são expressos através de uma função quadrática em \mathbf{z} , verificando a condição $\dim(\mathbf{z}) = \dim(\mathbf{v})$. A matriz \mathbf{B} , representativa dos preços dos factores de produção (*inputs*), é a mesma para todas as empresas e o vector \mathbf{v} , representativo das características tecnológicas, segue uma distribuição normal, $\mathbf{v} \sim N(\mu_{\nu}, \Sigma_{\nu})$.

Uma vez que se admite que as diferenças relativamente às escolhas de atributos/características, entre consumidores e entre empresas, são atribuíveis, respectivamente, aos parâmetros das preferências, passíveis de variar inter-consumidores (\mathcal{G}) e aos parâmetros tecnológicos susceptíveis de variarem entre empresas (ν);⁵⁹ os quais, por sua vez, se assumem seguirem uma distribuição normal, virá então que uma qualquer função de preço, arbitrária, induz uma densidade da procura e da oferta, para qualquer z , também normal.

⁵⁸ As condições necessárias e suficientes, para se determinar o óptimo da empresa, são, 1ª $p_z - \nu - \mathbf{B}\mathbf{z} = 0$ e 2ª $p_{zz} - \mathbf{B}$ definida negativa. As N funções, admitindo que $z \in R^N$, que expressam as condições de 1ª ordem, traduzem a oferta de produção das empresas.

⁵⁹ Os pressupostos de que \mathbf{A} e \mathbf{B} não variam entre os agentes económicos tornam-se necessários para uma solução explícita para a função do preço hedónico.

Dado este exemplo estar suportado em funções do tipo, quadrático-linear-normal, poder-se-á admitir uma solução para o equilíbrio de mercado, quadrática em \mathbf{z} , expressa por,

$$p(\mathbf{z}) = \boldsymbol{\pi}_1' \mathbf{z} + \frac{1}{2} \mathbf{z}' \boldsymbol{\pi}_2 \mathbf{z} \quad (2.2.3)$$

em que os parâmetros constantes do vector $\boldsymbol{\pi}_1$ e da matriz $\boldsymbol{\pi}_2$ estão relacionados com os parâmetros das funções utilidade e custos.

Para se apurar o rigor da solução admitida, começar-se-á por substituí-la nas condições de 1ª ordem dos problemas de maximização da utilidade do consumidor e de maximização do lucro da empresa,⁶⁰ as quais, quando resolvidas em ordem a \mathbf{z} , após verificação das condições de 2ª ordem, permitem definir uma função de equilíbrio, por interacção entre a procura e a oferta dirigida a cada uma das características, expressa por,

$$(\boldsymbol{\pi}_2 + \mathbf{A})^{-1} (\boldsymbol{\vartheta} \mathbf{A} - \boldsymbol{\pi}_1) = (\boldsymbol{\pi}_2 - \mathbf{B})^{-1} (\boldsymbol{\nu} - \boldsymbol{\pi}_1) \quad (2.2.4)$$

O equilíbrio vem assim caracterizado pelo vector $\boldsymbol{\pi}_1$, e pela matriz $\boldsymbol{\pi}_2$, também, presentes na hipótese de solução proposta.

Quer as características constantes do vector $\boldsymbol{\vartheta}$, quer as características constantes do vector $\boldsymbol{\nu}$ se assumem como distribuições de densidade normais, e ao aparecerem como termos lineares da função de equilíbrio, fazem com que as distribuições dos \mathbf{z} vectores

⁶⁰ Ficando assim, do lado da procura, $\boldsymbol{\vartheta} \mathbf{A} - \mathbf{A} \mathbf{z} - \boldsymbol{\pi}_1 - \boldsymbol{\pi}_2 \mathbf{z} = 0 \Leftrightarrow -(\mathbf{A} + \boldsymbol{\pi}_2) \mathbf{z} = \boldsymbol{\pi}_1 - \boldsymbol{\vartheta} \mathbf{A}$ e do lado da oferta, $\boldsymbol{\pi}_1 + \boldsymbol{\pi}_2 \mathbf{z} - \boldsymbol{\nu} + \mathbf{B} \mathbf{z} = 0 \Leftrightarrow (\boldsymbol{\pi}_2 + \mathbf{B}) \mathbf{z} = \boldsymbol{\nu} - \boldsymbol{\pi}_1$.

procurado e oferecido venham, também, normais. Assim sendo, a função de equilíbrio está a igualizar variáveis aleatórias normais, o que requer, como condições necessária e suficiente, a igualização da média da procura com a média da oferta,

$$(\boldsymbol{\pi}_2 + \mathbf{A})^{-1}(\mathbf{A}\boldsymbol{\mu}_g - \boldsymbol{\pi}_1) = (\boldsymbol{\pi}_2 - \mathbf{B})^{-1}(\boldsymbol{\mu}_v - \boldsymbol{\pi}_1) \quad (2.2.5)$$

e a igualização das variâncias da procura e da oferta, traduzida por,

$$(\boldsymbol{\pi}_2 + \mathbf{A})^{-1}\boldsymbol{\Sigma}_g(\boldsymbol{\pi}_2 + \mathbf{A})^{-1} = (\boldsymbol{\pi}_2 - \mathbf{B})^{-1}\boldsymbol{\Sigma}_v(\boldsymbol{\pi}_2 - \mathbf{B})^{-1} \quad (2.2.6)$$

Resolvendo para os parâmetros, $\boldsymbol{\pi}_1$ e $\boldsymbol{\pi}_2$, presentes na função que iguala a média da procura com a média da oferta da variável aleatória normal, obter-se-á a partir de (2.2.5) uma expressão explícita para $\boldsymbol{\pi}_1$, do tipo,

$$\boldsymbol{\pi}_1 = [(\boldsymbol{\pi}_2 - \mathbf{B})^{-1} - (\boldsymbol{\pi}_2 + \mathbf{A})^{-1}]^{-1}[(\boldsymbol{\pi}_2 - \mathbf{B})^{-1}\boldsymbol{\mu}_v - (\boldsymbol{\pi}_2 + \mathbf{A})^{-1}\mathbf{A}\boldsymbol{\mu}_g] \quad (2.2.7)$$

E para se explicitar $\boldsymbol{\pi}_2$ recorrer-se-á à igualdade das variâncias,⁶¹ e a partir de (2.2.6) obtém-se,

$$\boldsymbol{\pi}_2 = \frac{\mathbf{B}\boldsymbol{\Sigma}_g^{1/2} + \mathbf{A}\boldsymbol{\Sigma}_v^{1/2}}{\boldsymbol{\Sigma}_g^{1/2} - \boldsymbol{\Sigma}_v^{1/2}} \quad (2.2.8)$$

⁶¹ As matrizes $(\boldsymbol{\pi}_2 + \mathbf{A})$ e $(\boldsymbol{\pi}_2 - \mathbf{B})$, sendo consistentes com as condições de 2ª ordem para a optimização dos consumidores e dos produtores, apresentam-se como negativas definidas.

onde $\Sigma_g^{1/2}$ e $\Sigma_v^{1/2}$ representam matrizes diagonais, constituídas pelas raízes quadradas positivas dos termos correspondentes das matrizes de variância, Σ_g e Σ_v , as quais, por sua vez, são matrizes diagonais, positivas definidas.

Uma solução indeterminada para π_1 e π_2 pode entretanto ocorrer, caso Σ_g e Σ_v venham iguais. A identificação directa de, π_1 e π_2 , só se consegue com os parâmetros, quer das preferências, quer das tecnologias, em contexto de um único tipo de consumidores, $\Sigma_g = 0$; ou então, de um único tipo de empresas $\Sigma_v = 0$. Neste cenário, para $\Sigma_g = 0$, viria $\pi_2 = -\mathbf{A}$ e $\pi_1 = \mathfrak{g}$; e para $\Sigma_v = 0$, viria $\pi_2 = \mathbf{B}$ e $\pi_1 = \mathbf{v}$. Em termos económicos, significaria ter os preços hedónicos a coincidir com, as avaliações marginais dos consumidores, sempre que na presença de, apenas, um tipo de consumidores, $\Sigma_g = 0$; e/ou as produtividades marginais das empresas, se ao invés houvesse apenas um tipo de empresas $\Sigma_v = 0$.

É neste contexto que Rosen (1974) reforça a importância da conceptualização teórica a partir dos dados observados, considerando-os necessários para recuperar os parâmetros da tecnologia e da preferência. Analisando um mercado em que se possa calcular $p(\mathbf{z})$ e em que não haja atributos negligenciados e utilizando as condições de 1ª ordem e de 2ª ordem dos problemas de maximização da utilidade do consumidor e de maximização do lucro da empresa, propôs um método de estimação dos parâmetros de preferência e de tecnologia a dois passos/estádios.

Num primeiro estágio, estimar-se-ia, a função de preços hedónica, $p(\mathbf{z})$, para as características/atributos da habitação, em função dos dados de mercado, já que ao identificar-se habitação com uma mercadoria diferenciada se assume como um cabaz de atributos/características, \mathbf{z} , para os quais não há um mercado explícito, passando o preço dessa mercadoria a depender das quantidades dos diversos atributos/características que a compõem, e por essa via conseguir-se-ia recuperar os valores para os parâmetros de preferência e tecnológicos – base para a dedução dos

preços marginais, p_z . Posteriormente, no estágio 2 utilizar-se-iam as condições de 1ª ordem, substituindo os preços marginais, p_z ,⁶² pelos deduzidos no passo anterior, para se conhecer a curvatura desses parâmetros. Especificamente, estimar-se-iam: (i) as preferências comuns \mathbf{A} e a média das preferências diversas \mathcal{G} , μ_g , pela regressão de mínimos quadrados, $(\hat{p}_z = \hat{\pi}_1 + \hat{\pi}_2 \mathbf{z}) = \mu_g + \mathbf{A}\mathbf{z} + \varepsilon_g$, com, $\mathcal{G} = \mu_g + \varepsilon_g$; e (ii) a tecnologia comum inter-empresas, \mathbf{B} , e a média do parâmetro ν , representativo de tecnologia diversa inter-empresas, μ_ν , através de uma regressão de mínimos quadrados, $(\hat{p}_z = \hat{\pi}_1 + \hat{\pi}_2 \mathbf{z}) = \mu_\nu + \mathbf{B}\mathbf{z} + \varepsilon_\nu$, em que, $\mathbf{v} = \mu_\nu + \varepsilon_\nu$. Entretanto, assumia-se que μ_g e μ_ν eram, por sua vez, funções dos regressores (x) e (y) ,⁶³ respectivamente

Os coeficientes estimados, no 1º estágio, sendo coeficientes da função envelope, que reflecte, quer as curvas de *bid*, quer as curvas de *offer*,⁶⁴ não são, nalgumas situações, directamente interpretados, por isso torna-se importante que as equações regressivas desse 1º estágio sejam capazes de fornecer preços implícitos marginais importantes para os atributos do cabaz habitação.

Ora, acontece que as estimações regressivas foram interpretadas como aproximações lineares das condições de 1ª ordem, tendo subjacente as distribuições de densidade das preferências, ϑ , e das tecnologias, ν , e assim sendo esses parâmetros de preferência e de tecnologia continuavam a vir não identificados, à excepção dos casos especiais, já referidos. Entretanto se o preço construído, $p(\mathbf{z})$, fosse uma função não linear de \mathbf{z} , o problema da identificação desses parâmetros cessava e os problemas de colinearidade

⁶² As derivadas parciais da função de preços hedónica, estimada no 1º estágio, são tratadas como observações nas funções de procura e oferta estimadas neste 2º estágio.

⁶³ Neste caso as distribuições dos parâmetros \mathcal{G} e ν vêm determinadas pelas distribuições de (x, ε_g) e (y, ε_ν) , respectivamente.

A aplicação das variáveis instrumentais, x e y , para a versão paramétrica subjacente à estimação é genericamente válida.

⁶⁴ A análise de Rosen indica rigorosamente a relação entre as funções de *offer* das empresas, as funções de *bid* dos consumidores e a estrutura hedónica dos preços.

entre \mathbf{z} e $\hat{p}_z(\mathbf{z})$ não se colocariam, porém esta não linearidade seria assumida como uma fonte artificial de identificação e portanto arbitrária.⁶⁵ Mesmo que o problema da arbitrariedade não fosse colocado e se utilizasse a não linearidade em $\hat{p}_z(\mathbf{z})$ sem restrições, ainda assim enfrentar-se-iam os problemas de endogeneidade padrão, para se identificar os parâmetros da preferência e da tecnologia.

As críticas padrão no que concerne à arbitrariedade da forma funcional e à endogeneidade das variáveis, sugeridas pelo trabalho pioneiro de Rosen (1974), relativamente ao procedimento de estimação a dois estádios/passos dos modelos hedónicos, têm-se vindo a revelar vazias de conteúdo. Por um lado, porque se apresenta a segmentação de mercados (por cidades diferentes e/ou por zonas diferentes de uma mesma cidade) como solução,⁶⁶ e por outro lado porque se apresenta a identificação integral dos modelos hedónicos com pressupostos moderados relativamente à forma funcional e com dois procedimentos de estimação: (a) métodos de transformação não paramétrica e (b) variáveis instrumentais num quadro não linear, mas paramétrico; para minorar esses problemas de identificação e de endogeneidade.⁶⁷

A estimação da função de preços hedónica, $p(\mathbf{z})$, coloca-nos novas questões relativamente à endogeneidade das variáveis, que residem no facto do preço de cada uma das características não ser directamente observável. Ao estimar-se a função de preços hedónica é comum a inclusão de um termo estocástico, pretendendo reflectir os erros de medida, ou os erros de especificação (decorrentes da forma funcional e da estimação) e, também, se consideram como argumentos, as características das empresas e as características dos consumidores (as observáveis; as não observáveis/mas correlacionadas inter-empresas e inter-consumidores e as não observáveis/não correlacionadas). Num cenário de preferências quase-lineares⁶⁸, é comum e razoável⁶⁹

⁶⁵ Embora, Ekeland, Heckman e Nesheim (2004) provassem que a não linearidade é genérica no modelo hedónico.

⁶⁶ Palmquist (1984).

⁶⁷ Ekeland, Heckman e Nesheim (2004).

⁶⁸ Preferências quase-lineares encontram correspondência geométrica em curvas de indiferença que são meras translações verticais, umas das outras.

⁶⁹ Como o que se pretende identificar é o preço de licitação para diferentes níveis de uma dada característica z_k , de forma a manter o nível de utilidade u_i , ou seja a inversa da procura compensada, que não é observável pelo comportamento de mercado, então só

assumir-se que os termos de erro da função de preços hedónica não estão correlacionados com os atributos escolhidos, caso contrário o termo erro poderia afectar as quantidades de características/atributos que maximizam a utilidade. Também é importante avaliar se o termo de erro está correlacionado, entre empresas e entre consumidores de uma comunidade, para avaliar a fiabilidade das variáveis instrumentais no processo de estimação.

Após a apresentação das dificuldades analíticas subjacentes ao cálculo do preço de equilíbrio de mercado e após se ter constatado, para o exemplo quadrático em apreciação, que só em circunstâncias específicas os parâmetros da função de preços identificam parâmetros da função utilidade do consumidor (parâmetros de preferências e de gostos) ou parâmetros da função custos (parâmetros produtivos e tecnológicos), o método regressivo, base da estimação da função de preços hedónicos, suporte do primeiro estágio da estimação hedónica apresentada por Rosen, ganha relevância e revela-se capaz de responder ao objectivo de inferir o valor da propriedade imobiliária para efeitos fiscais.

Neste enquadramento importa ainda, antes de expor o modelo teórico e econométrico, apresentar um conjunto de pressupostos suficientes à identificação de função de preços hedónica,⁷⁰ de forma a construir um suporte teórico que sustente a estimação hedónica deste ensaio, cujo desenvolvimento a secção 2.3. fará relato.

2.3. Generalização da Perspectiva Hedónica aos Mercados de Produtos Diferenciados

2.3.1. Enquadramento do problema

O modelo hedónico trata produtos diferenciados como cabazes dos atributos z que lhe estão subjacentes. As decisões das famílias sobre estes produtos baseiam-se na procura

resta a dedução a partir da inversa da procura ordinária, no pressuposto de preferências quase-lineares e preços hedónicos constantes.

⁷⁰ Benkard e Bajari (2005).

pelos atributos que lhes são oferecidos ordenando-os em função dos tipos de preferências. Esta ordenação resulta numa função de preço de equilíbrio, $p(\mathbf{z})$, onde os preços são função dos atributos e determinados pela procura e oferta desses atributos. O que implica que o declive da função de preço hedónico num dado ponto é igual à disposição marginal a pagar pelas famílias nesse ponto. É, ainda, exactamente igual à respectiva função preço/licitação, se as famílias são idênticas e é um envelope a essas funções preço/licitação, se as famílias são heterogéneas.

O modelo hedónico tem sido utilizado para estimar, quer os preços baseados nos atributos/características (utilizando a função preço), quer os preços implícitos desses atributos/características (utilizando as derivadas).

O modelo hedónico requer alguns pressupostos simplificadores fortes. A produção de produtos deve ser convexa nos atributos/características, por isso quantidades contínuas devem estar disponíveis mesmo que condicionadas a outras quantidades. Como a literatura da organização industrial frequentemente refere isso significa não haver espaço para a criação de novos produtos diferenciados. Ou, então o modelo implica que qualquer produto que preencha o espaço produto não tem valor, nalguns casos os economistas foram motivados a utilizar modelos de escolha discreta para estimar em vez dos hedónicos.

A essência do problema causado pela não convexidade é o agrupamento de consumidores heterogéneos em função de um único produto. Quando a escolha é convexa e cada produto está continuamente disponível, cada família tipo escolherá o seu produto favorito no conjunto, que em princípio poderá ser diferente do escolhido por outro tipo de famílias. O resultado é um agrupamento tipo Tiebout em cada dimensão, onde *ceteris paribus*, aqueles que preferem mais de um atributo comprarão mais dele.

Quando a escolha não é convexa tal ordenação não é possível. As características das casas, mesmo aparentando continuidade, impõem que se reflectam quatro aspectos: 1º a distribuição dos atributos/características, em si, tem que ser contínua, não basta que essas características venham distribuídas sobre uma base contínua (a habitação); 2º mesmo que os atributos/características fossem gerados a partir de uma distribuição

contínua, qualquer amostra finita de casas não o será; 3º mesmo que a densidade marginal de um atributo/característica fosse distribuída continuamente, os problemas à estimação hedónica podiam subsistir, desde que a distribuição conjunta das outras características não viesse contínua (isto é, se as famílias viessem ordenadas em função das suas preferências por uma característica contínua, condicionadas pelas características discretas, mesmo assim haveria dificuldade do efeito preço vir identificado via formas funcionais convencionais e até modelos semi-paramétricos) e 4º algumas características (p.e. poluição), ainda que geradas de forma contínua, acabam, muitas vezes, modeladas como integrando preferências numa forma discreta.

No modelo delineado por Rosen assume-se a concorrência perfeita e um *continuum* de produtos e estes pressupostos dificilmente se adequam à realidade dos mercados reais. Em geral, mercados de concorrência imperfeita, não garantem a existência da função de preços hedónica e, nos casos em que possa existir, é possível que não tenha uma forma simples, separável e aditiva. Também, o espaço de produtos contínuos parece pouco plausível, mesmo em mercados onde se verifique um número suficientemente elevado de produtos.

Há, ainda, o estrito pressuposto de que todos os atributos/características do produto são perfeitamente observáveis, mas efectivamente há características que não se conseguem observar. Como resultado é comum este pressuposto da observação perfeita conduzir para violações ao axioma da preferência revelada.

Relaxar os pressupostos da concorrência perfeita, da separabilidade da função preço, dos espaços contínuos de produtos e dos atributos/características observados tornaria mais fácil a aplicação do modelo hedónico ao mundo real.

Assim num primeiro estágio equacionam-se os pressupostos da separabilidade da função preço e da observação das características introduzindo alguns pressupostos adicionais, como,

(i) os atributos/características do produto passíveis de observação são independentes das características que não são passíveis de observação, apresentando-se como uma versão ligeiramente mais poderosa do que o pressuposto de i.i.d. comumente utilizado na

literatura empírica; (ii) a maximização de utilidade pelo consumidor processa-se sequencialmente por uma escolha inicial do “modelo de produto” e, só posteriormente, pela escolha do “pacote de opções” – é um comportamento não paramétrico análogo aos efeitos fixados num modelo linear; (iii) existe pelo menos um mercado onde os preços não são função das características observadas – p.e. casas de diferentes tipos recebem o mesmo preço, enquanto casas do mesmo tipo em diferentes urbanizações ou de diferentes construtores apresentam preços diferentes; e (iv) variáveis instrumentais não paramétricas são separáveis.

Se o conjunto de escolha é contínuo então os parâmetros do nível de preferências das famílias devem satisfazer o conjunto de condições de primeira ordem que exigem que a taxa marginal de substituição (*TMS*) entre o *continuum* de atributos/características e a mercadoria compósita seja igual ao preço implícito desses atributos/características.

Se a forma funcional da função utilidade é conhecida e o vector de parâmetros é de dimensão igual ou inferior ao vector dos atributos/características então as condições de primeira ordem podem ser utilizadas para obter os níveis dos coeficientes aleatórios das famílias. Agregando os coeficientes aleatórios, a distribuição desses coeficientes pode ser obtida não parametricamente.

Na perspectiva de Rosen para se obterem aproximações de ordem superior para a função de utilidade tornava-se necessário impor a homogeneidade entre indivíduos de forma a acautelar os potenciais problemas de identificação decorrentes dessas aproximações. No entanto, a heterogeneidade caracterizadora dos comportamentos individuais e existentes entre indivíduos (ilustrada através de parâmetros diferentes para diferentes funções de utilidade) pode ser equacionada, desde que se confie em restrições paramétricas, sem que se coloque em perigo a identificação fornecida pela função de utilidade.

Estas condições são suficientes para, seguidamente, identificar a função de preço, p , em função dos atributos/características, observáveis, z e não observáveis, ξ ; bem como o conjunto das características não observáveis $\{\xi_x\}$. Se a função de procura é conhecida

para todo $p \in I$, então os resultados de Mas-Colell (1977) fornecem condições suficientes para identificação da relação fraca das preferências do consumidor i , (\succeq_i) . Em Mas-Colell, as restrições orçamentais são lineares, por isso é suficiente conhecer a função de procura do consumidor, que é linear em (z, ξ) para aplicar esses resultados. Considera-se a identificação das preferências dos consumidores em casos onde dados sobre a função de procura não estão disponíveis.

2.3.2. Função preço

Começa-se por mostrar que, no caso de a procura ser obtida via modelo hedónico, então, mesmo num cenário de concorrência imperfeita e de um número finito de produtos, existe um domínio de preços hedónicos.

Para o efeito considera-se o cenário em que o conjunto escolha é contínuo mas existe um número de observações finitas por indivíduo.

Num modelo, em que,

- um produto $x \in X$ é um vector de atributos/características de dimensão finita, (\mathbf{z}_x, ξ_x) , onde, por sua vez, $\mathbf{z}_x = (z_{x1}, \dots, z_{xK})$ é um vector de dimensão K de atributos/características observados, quer pelos consumidores, quer pelos analistas e ξ_x é um escalar que traduz o atributo/característica do produto observado apenas pelo consumidor;

e

- o conjunto $Z = \cup_{x \in X} (\mathbf{z}_x, \xi_x) \subseteq R^{K+1}$ representa o universo de produtos disponíveis no mercado para os consumidores.

Seja p_{xt} o preço de um produto $x \in X$ num dado mercado $t \in T$. Os elementos de T podem ser pensados como mercados separados espacial ou temporalmente. Os consumidores são maximizadores da utilidade, seleccionando entre um produto

$x \in X$ e uma mercadoria compósita $c \in R^+$. Cada consumidor i , tem uma função de utilidade, dada por, $u_i(z_x, \xi_x, c): Z \times R^+ \rightarrow R$. O preço da mercadoria compósita é normalizado à unidade. Os consumidores dispõem de um rendimento m_i e o conjunto restrição do consumidor i no mercado t , $B(m_i, t)$ deve satisfazer,

$$B(m_i, t) = \{(x, c) \in X \times R^+ : p_{xt} + c \leq m_i\}$$

O consumidor i no mercado t resolve o problema de maximização,

$$\underset{(x, c) \in B(m_i, t)}{\text{Max}} u_i(z_x, \xi_x, c) \quad (2.3.1)$$

Em equilíbrio, sobre condições fracas, este modelo implica que os preços em cada mercado tenham as seguintes propriedades: (i) há um preço para cada cabaz de atributos/características; (ii) a superfície do preço é crescente com os atributos/características não observados e (iii) a superfície do preço satisfaz uma condição de Lipschitz.⁷¹ Ora, o teorema que se desenvolve depende apenas do problema de maximização do consumidor, do facto dos preços serem dados e de alguns simples pressupostos sobre as funções de utilidade dos consumidores. E o mais importante é que é independente dos pressupostos do lado da oferta.

Tomemos os seguintes pressupostos,

⁷¹ Dada a região em R^2 definida por $D = \{(z_x, \xi_x), d \leq z_x \leq e, f \leq \xi_x \leq g\}$ admita-se que $u_i(z_x, \xi_x)$ é contínua em D . Diz-se que $u_i(\cdot)$ satisfaz a condição de Lipschitz na variável z_x em D desde que haja uma constante $L > 0$ que verifique a propriedade seguinte,

$$|u_i(z_x, \xi_x) - u_i(z_{x'}, \xi_x)| \leq L |z_x - z_{x'}|, \text{ sempre que } (z_x, \xi_x), (z_{x'}, \xi_x) \in D.$$

P1 $u_i(\mathbf{z}_x, \xi_x, c)$ é continuamente diferenciável em c e estritamente crescente em c ,

com $\frac{\partial u_i(\mathbf{z}_x, \xi_x, c)}{\partial c} > \varepsilon$ para valores de $\varepsilon > 0$ e para qualquer valor de $c \in [0, m_i]$.

P2 u_i é contínua de Lipschitz em, (\mathbf{z}_x, ξ_x) .

P3 u_i é estritamente crescente em, ξ_x .

O pressuposto **P3** é o mais restritivo dos três. Implica que não há saturação no atributo/característica não observado, no entanto sem este pressuposto não se garantiria que a função preço viesse crescente com, ξ_x .

Teorema 1. Supondo que os pressupostos **P1-P3** se verificam para qualquer individuo, em qualquer mercado. Então, para dois quaisquer produtos x e x' , com procura positiva, em qualquer mercado t , vem,

(i) Se $\mathbf{z}_x = \mathbf{z}_{x'}$ e $\xi_x = \xi_{x'}$, então $p_{xt} = p_{x't}$.

(ii) Se $\mathbf{z}_x = \mathbf{z}_{x'}$ e $\xi_x > \xi_{x'}$, então $p_{xt} > p_{x't}$.

(iii) $|p_{xt} - p_{x't}| \leq J(|\mathbf{z}_x - \mathbf{z}_{x'}| + |\xi_x - \xi_{x'}|)$, para um qualquer $J < \infty$.

Prova:

A intuição para o teorema é que se as propriedades (i) – (iii) não fossem satisfeitas pelo equilíbrio de preços, então alguns dos bens podiam não ter procura positiva. \square

A função de preço de equilíbrio para o mercado t denomina-se $p_t(\mathbf{z}_x, \xi_x)$. É uma transformação do conjunto de características do produto em preços, que satisfazem

$p_{xt} = p_t(\mathbf{z}_x, \xi_x)$ para todo $x \in X$, e assume-se que (i) – (iii) se mantêm. Uma vez que (iii) se mantêm para todos os pares de produtos, no limite a função preço deve ser contínua de Lipschitz.

O Teorema 1 baseia-se, apenas, em argumentos do lado da procura, daí ser um teorema geral para vários tipos de equilíbrio, quer estáticos, quer dinâmicos. O teorema aborda unicamente os preços dos produtos actualmente observados com procura positiva. Uma consequência é que, para algumas funções de custo e algumas procuras padrão, certos cabazes poderão não ser observados. Por exemplo, esse pode ser o caso, de funções de custo descontínuas. Nesses casos é provável que haja um problema de selecção na estimação da função preço.

A função de preço em cada mercado é uma função de equilíbrio que está dependente dos mercados primários. Se um bem já não está disponível no mercado, essa função não informa sobre o preço desse bem. Se um novo bem é adicionado ao mercado, em geral, regista-se um ajustamento a um novo equilíbrio de todos os outros bens e a função de preço altera-se. A função preço muda se qualquer mercado primário muda, tal como se mudassem as preferências dos consumidores e os custos de produção marginal, ou, ainda, se um produto já no mercado passasse a ser produzido noutra empresa multi-produto. Em suma, o que uma função preço, num mercado específico, traduz é a relação, entre atributos/características e preços, percebida pelos consumidores nesse mercado.

Mesmo modelos simples de concorrência sugerem que a função preço de equilíbrio não é linear e separável em todos as características. Por exemplo a elasticidade inversa para produtos simples padrão implica funções de preço não separáveis mesmo para funções de custo marginal lineares. Não seria sensato assumir uma função preço aditivamente separável em características do produto não observáveis, optando-se em vez disso por manter a forma geral acima referida.

Provada que fica a existência de uma função preço há que mostrar a identificação da mesma e a identificação de atributos/características, não observados. Numa primeira

etapa procurar-se-á a identificação utilizando independência, conforme descrito na secção que se segue.

2.3.3. Identificação da função preço e das características não observadas

2.3.3.1. Identificação utilizando a independência

Vamos demonstrar que a função preço e os atributos/características não observados do produto $\{\xi_x\}$ são identificados, caso esses atributos/características não observados do produto, ξ , sejam independentes dos atributos/características observados, z . Identificação que se mantém mesmo que os preços sejam estimados.

Começemos por identificar o domínio dos preços no caso em que estes são observados sem erro. Assim assume-se que,

P4 ξ é independente de z .

P5 Para todos os mercados t e para todo z , $p_t(z, \cdot)$ é estritamente crescente, com

$$\frac{\partial p_t(z, \xi)}{\partial \xi} > \delta \text{ para todo o } (z, \xi), \text{ para todo o } t \text{ e para alguns } \delta > 0.^{72}$$

O pressuposto **P4** ao requerer independência plena reforça o significado do pressuposto de independência utilizado no contexto da Economia Industrial. Entretanto, ao equacionar-se a não separabilidade, o pressuposto da independência vem enfraquecido. Nos modelos onde os erros são considerados linearmente a independência exclui problemas de heteroscedasticidade não observada, apurados, normalmente, em aplicações empíricas. A não separabilidade permite que a independência subjacente aos

⁷² O limite inferior para a derivada é necessário para assegurar que à medida que o número de mercados se torna maior a função de preço não se torna arbitrariamente próxima de uma função crescente fraca. O teorema principal requer apenas $\delta \geq 0$. Quando se introduz a medida de erro requer $\delta > 0$.

erros interaja com os dados observados de forma a replicar modelos de heteroscedasticidade.

Se a independência se mantém então a base dos atributos/características não observados não depende dos atributos/características observados, daí que $p_t : A \times E \rightarrow R$, onde $A \subseteq R^K$ é o suporte de z , e $E \subseteq R$ é o suporte de ξ . O pressuposto **P5** decorre dos pressupostos iniciais sobre preferências dos consumidores (veja o Teorema 1).

No caso de um mercado singular, $T = \{1\}$ e na ausência de medida de erro os resultados de Matzkin (2003) podem ser utilizados para mostrar sob condições fracas que, quer a forma funcional $p_1(\cdot)$, quer a distribuição dos atributos/características não observados do produto, $\{\xi_x\}$ são identificadas segundo uma normalização sobre ξ . A primeira parte da nossa prova de identificação segue Matzkin, as únicas diferenças residem nos resultados que são alargados para cobrir o caso de muitos mercados e utiliza-se uma normalização alternativa para facilitar a estimação.

Seja I o conjunto de funções preço satisfazendo **P5**,

$$I = \left\{ p' : A \times E \rightarrow R \mid \text{para todo } z \in Z, p'(z, \cdot) \text{ é estritamente crescente} \right\} \quad (2.3.2)$$

Uma vez que os atributos não observados, não têm unidades inerentes é apenas possível identificá-los através de uma transformação monotónica. Sem perda de generalidade assume-se que a normalização feita para ξ é tal que a distribuição marginal de ξ é $U[0,1]$. Tecnicamente a normalização de ξ utiliza a sua função de distribuição.

Define-se identificação como sendo a identificação dentro de um conjunto satisfazendo a normalização feita acima,

Definição 1. A função \mathbf{p} é identificada em I se,

i. $\mathbf{p} \in I$, e

ii. Para todo $\mathbf{p}' \in I$,

$$\left[F_{pz}(\cdot; \mathbf{p}) = F_{pz}(\cdot; \mathbf{p}') \right] \Rightarrow \left[\mathbf{p} = \mathbf{p}' \right]$$

Mostra-se que a identificação se mantém caso os preços sejam medidos sem erro.

Teorema 2. Se os preços são observados sem erro e os pressupostos **P4** e **P5** se mantêm, então p_t é identificado em I , para todo o t . Além disso, $\{\xi_x\}$ é também identificado.

Prova: Primeiro há que mostrar como se obtêm os atributos/características não observados a partir da distribuição condicional de preços,

$$\begin{aligned} F_{p_t|z=z_x}(p_{xt}) &= \Pr(\mathbf{p}_t(\mathbf{z}, \xi) \leq p_{xt} | \mathbf{z} = \mathbf{z}_x) \\ &= \Pr(\xi \leq \mathbf{p}_t^{-1}(\mathbf{z}, p_{xt}) | \mathbf{z} = \mathbf{z}_x) \\ &= \Pr(\xi \leq \mathbf{p}_t^{-1}(\mathbf{z}_x, p_{xt})) \\ &= \mathbf{p}_t^{-1}(\mathbf{z}_x, p_{xt}) \\ &= \xi_x \end{aligned}$$

Para construir a função preço para cada mercado necessita-se, apenas, de inverter a relação acima,

$$\mathbf{p}_t(\mathbf{z}_0, \xi_0) = F_{p_t|z=z_0}^{-1}(\xi_0) \quad (2.3.3)$$

□

A partir da demonstração do teorema pode-se ver que, na ausência do termo erro, a identificação de atributos/características não observáveis pode ser obtida através de uma amostra simples. Ainda que a identificação se obtenha, mesmo que os produtos sejam observados em apenas um mercado.

Quando os preços são medidos com erro, a variação entre mercados pode ser utilizada para obter a identificação.

Consideremos pois o caso em que os preços são observados com erro. Especificamente assume-se que p_{xt} não é observado. Em vez disso observa-se y_{xt} , em que,

$$y_{xt} = p_{xt} + \varepsilon_{xt} \equiv \mathbf{p}_t(\mathbf{z}_x, \xi_x) + \varepsilon_{xt}$$

Assume-se a clássica medida do erro:

B1 ε_{xt} é i.i.d. e $E[\varepsilon | z, \xi] = 0$.

Para o propósito da identificação não é necessário que ε_{xt} seja i.i.d. Tudo o que importa é que para qualquer z e ξ , a lei dos grandes números se verifique para ε_{xt} através de cada x e de cada t .

A identificação com medida de erro é necessária para a observação de produtos em diversos mercados como pode ser visto na prova do seguinte teorema.

Teorema 3. Se os preços são observados com erro e se verificam **P4**, **P5** e **B1**, então p_t é identificado em I para todo o t . Para além de que $\{\xi_x\}$ é identificado.

Prova: Seja

$$\bar{\mathbf{p}}^T(z, \xi) = \frac{1}{T} \sum_{t=1}^T \mathbf{p}_t(z, \xi)$$

e seja,

$\bar{\mathbf{p}}_j^T \equiv \bar{\mathbf{p}}^T(\mathbf{z}, \xi)$. Para cada produto, pode-se observar $\bar{\mathbf{p}}_x^T$ a partir da média dos preços observados, y_{xt} , em cada mercado. Uma vez que o erro tem uma média condicional zero, $E[\varepsilon | z, \xi] = 0$, para qualquer (z, ξ) , vem que a média para valores grandes de T é, também, zero.

Para cada produto, x , define-se o conjunto

$$X_x = \left\{ w \in X \mid \mathbf{z}_w = \mathbf{z}_x \text{ e } \lim_{T \rightarrow \infty} (\bar{\mathbf{p}}_x^T - \bar{\mathbf{p}}_w^T) = 0 \right\}$$

o conjunto X_x é o conjunto de todos os produtos com as mesmas características das do produto x , quer sejam observáveis, quer não. O valor da função preço para cada produto x , p_{xt} é identificado pelos preços médios dentro de cada mercado t no universo dos produtos X_x :

$$p_{xt} = E[y_{wt} \mid w \in X_x]$$

O erro médio é de novo zero.

Uma vez identificado o valor da função preço para cada produto em cada mercado, o resto da prova de identificação segue o Teorema 2.

Finalmente, $\varepsilon_{xt} = y_{xt} - \mathbf{p}_t(\mathbf{z}_x, \xi_x)$, por isso ε_{xt} e a distribuição conjunta de ε , de z e de ξ são, também, identificadas. \square

Acontece que no mercado da habitação se pode escolher um modelo de habitação, correspondente a uma dada tipologia (p.e. T1, T2, ...) e, ainda, um pacote de opções associadas a esse modelo (p.e. acabamentos, orientação solar, ...). Assim sendo torna-se necessário introduzir um conjunto de pressupostos que permitam identificar a função preço e identificar as características do produto não observadas, de forma a equacionar-se esta particularidade dos mercados.

2.3.3.2. Identificação baseada em pacotes opcionais

Ponderando esta realidade implícita ao processo de escolha da habitação, procura-se demonstrar que, se a característica não observável ξ_j corresponde a um modelo e se z_j corresponde a um pacote opcional, então é possível identificar uma função de preço e as características do produto não observáveis.

Seja μ a notação de um modelo e \mathcal{M} a denominação do conjunto de todos os modelos. O conjunto de modelos induz uma partição no conjunto X . A transformação $\pi: X \rightarrow \mathcal{M}$ associa produtos (x) com modelos (μ). A imagem inversa de μ sob π é o conjunto de produtos identificados com o modelo μ , em que cada produto tem um potencial pacote diversificado de opções z . O modelo μ é observável e z e μ têm uma distribuição conjunta, $F_{z,\mu}: A \times \mathcal{M} \rightarrow R$.

O primeiro pressuposto introduzido, diz que ξ é partilhado por produtos identificados com o mesmo modelo:

P6. Para todo $x_1, x_2 \in \mathcal{M}$, se $\pi(x_1) = \pi(x_2)$, então $\xi_{x_1} = \xi_{x_2}$.

De forma a identificar o produto não observável, também se necessita de um pacote padrão de opções que está disponível para todos os modelos μ . Este requisito vem formalizado no seguinte pressuposto,

P7. Existe um $\bar{z} \in A$, tal que para todo $\mu \in \mathcal{M}$, $f(\bar{z}|\mu) > 0$

Na ausência de unidades implícitas de ξ , a identificação de ξ e da função preços faz-se a partir de um procedimento de normalização. Neste caso, normaliza-se ξ , tal que $F_{\xi|z=\bar{z}}$ é $U[0,1]$. A identificação para o caso em que os preços são observados sem erro segue o teorema.

Teorema 4. Se os preços são observados sem erro e **P5-P7** se verificam, então p_t é identificado em I para todo o t . Para além disso $\{\xi_x\}$ é identificado.

Prova: Para cada produto x , seja x^* um produto tal que $\pi(x) = \pi(x^*)$ e $z_{x^*} = \bar{z}$. Caso exista um produto para qualquer modelo $\pi(x)$ por via de **P7**. Então, paralelamente à secção anterior,

$$\xi_x = F_{p_t|z=\bar{z}}(p_{x^*t})$$

Esta equação identifica $\{\xi_x\}$.

A função de preço em cada mercado é dada pelos preços dos pacotes não-padrão, Para qualquer ponto $(z_0, \xi_0) \in A \times E$,

$$p_t(z_0, \xi_0) = p_{wt}, \text{ para } w \in X \text{ tal que } \xi_w = \xi_0 \text{ e } z_w = z_0 \quad (2.3.4)$$

□

De novo, neste caso, a identificação da característica não observada é obtida a partir de uma amostra singular. Entretanto, apesar da independência, a identificação pode ser obtida a partir de uma amostra singular ainda que os preços venham medidos com erro (conforme se verá a seguir). A razão para esta diferença é que se observam, agora, diversos produtos em cada mercado que são conhecidos por ter o mesmo valor de, ξ .

Outra consequência decorrente de se observarem, em cada mercado, diversos produtos, que se assumem com o mesmo valor para a característica não observável está na sobre identificação do modelo. Se não há erro nos preços, então o modelo é rejeitado no sentido em que o **P6** pode ser violado.

Resta-nos pois provar a identificação caso os preços sejam estimados. Tarefa trivial, desde que os modelos sejam observados.

Teorema 5 Se os preços são observados com erro, **P5-P8** e **B1** se verificam, então p_t é identificado em I para todo o t . Para além disso $\{\xi_x\}$ é identificado.

Prova: Seja $X_x = \{w \in \pi^{-1}(\pi(x)) | \mathbf{z}_w = \mathbf{z}_x\}$. Tal como no caso anterior X_x é o conjunto de todos os produtos com as mesmas características do produto x . Então,

$$\mathbf{p}_t(\mathbf{z}_x, \xi_x) = E[y_{wt} | w \in X_x],$$

Onde a média do erro é de novo zero. O resto da prova segue a do Teorema 4.

□

2.3.4. Identificação com um conjunto relevante de funções preço

Está-se perante uma outra perspectiva para identificação, única no sentido em que não requer pressupostos adicionais sobre a distribuição conjunta de z e de ξ . Em vez disso confia-se em dois pressupostos sobre o conjunto de funções preço que são observadas.

Primeiro supõe-se que a informação é, razoavelmente, poderosa, que há um mercado em que os preços não dependem muito das características observadas. Não se assume que os investigadores conheçam o mercado em questão.

P8. Existe um mercado, t , tal que $\mathbf{p}_t(\mathbf{z}, \xi) = f(\xi)$, com $f_\xi > 0$.

O pressuposto **P8** dificilmente se manterá na maioria das aplicações, mas, em circunstâncias especiais, pode manter-se. **P8** tem maior probabilidade em manter-se em mercados onde “qualidade” é a característica determinante na diferenciação do produto em termos de preço. Por exemplo em muitos mercados de bens deste tipo, mesmo que os consumidores possam ter preferências fortes sobre “atributos”, que seriam tipicamente observados como dimensões de z , todos os atributos de uma dada linha de produto têm muitas vezes o mesmo preço, enquanto linhas de produtos diferentes têm preços diferentes.⁷³

Segundo também se necessita da monotonicidade fraca dos preços em todas as características,

P9 Para todos os mercados t , $\mathbf{p}_t(\mathbf{z}, \xi)$ é fracamente crescente em todas as características observadas, \mathbf{z} e estritamente crescente nas características não observadas, ξ .

⁷³ Especificamente, suponha que num mercado a função preço é $\mathbf{p}_t(\mathbf{z}, \xi) = w\mathbf{z} + \xi$, onde \mathbf{z} é um vector de variáveis *dummy flavor* e o parâmetro w é constante.

O pressuposto **P9** mantém-se em muitas aplicações. Se todos os indivíduos têm preferências monótonas sobre todas as características, então **P9** sustenta-se com um argumento similar ao do Teorema 1, podendo mesmo manter-se se não for o caso. Por exemplo se os custos marginais fossem suficientemente crescentes em todas as características então também **P9** seria verificado.

Teorema 6. Se os preços são observados sem erro, se **P8** e **P9** se verificam, e (z, ξ) tem completo suporte em $A \times E$, então p_t é identificado em I para todo o t . Além disso $\{\xi_x\}$ é identificado.

Prova: Sejam $x \equiv (z_{x1}, \dots, z_{xk}, \xi_x)$ e $w \equiv (z_{w1}, \dots, z_{wk}, \xi_w)$ dois pontos no espaço de mercadorias. Para provar que $\{\xi_x\}$ está identificado, mostra-se que, unicamente, a ordenação de ξ_x e de ξ_w é determinada. Seja $x^* \equiv (\min(z_{x1}, z_{w1}), \dots, \min(z_{xk}, z_{wk}))$ a componente dos mínimos componentes das características observadas dos dois produtos. Defina-se $X' \subseteq X$ como se segue:

$$X' = \left\{ x' \in X : (z_{x'1}, \dots, z_{x'k}) = x^*, \text{ e } p_{x't} \leq p_t(x) \text{ para todo o } t \right\} \quad (2.3.5)$$

Decorre de **P8** e **P9** que existe um elemento $x' \in X'$ e um mercado t tal que $p_{x't} > p_t(w)$ sse $\xi_x > \xi_w$.

Isto identifica a ordem de $\{\xi_x\}$. Uma normalização identifica $\{\xi_x\}$ e $F_{z, \xi}$. A identificação de $p(z, \xi)$ decorre directamente.⁷⁴ □

Como se verificou existem condições suficientes para identificar as funções de preço e as características não observadas. Entretanto o potencial da informação disponível pode

⁷⁴ Esta prova requer que todos os produtos sejam observados em muitos mercados.

determinar outra realidade e num cenário em que essa informação, sobre a função de procura, não seja suficiente pode-se ter que considerar a identificação das preferências dos consumidores.

2.3.5. Identificação das preferências

2.3.5.1. Conjunto escolha contínuo com um número finito de observações por indivíduo

O primeiro caso a considerar é o de um conjunto escolha contínuo mas com um número finito de observações por indivíduo – caso similar ao de Rosen.

Normalmente, só, um pequeno número de escolhas é observado pelo consumidor, muitas vezes apenas uma. Pelo argumento padrão, em cada cabaz de escolha, a taxa marginal de substituição entre características é igual ao declive do conjunto restrição orçamental do consumidor nesse ponto. Esta informação fornece apenas uma informação local sobre preferências em cada um dos cabazes escolhidos.

Uma forma simples de reduzir o conjunto de possibilidades, contudo, é colocar restrições paramétricas sobre as curvas de indiferenças dos consumidores. Estas restrições podem ser encaradas, quer como um pressuposto de identificação, quer como uma aproximação local para a função de utilidade. Muitos modelos de escolha discreta na literatura assumem que a utilidade é linear ou log-linear em (\mathbf{z}, ξ, c) , como por exemplo,

$$u_{ix} = \beta_{i1} \log(z_{x1}) + \dots + \beta_{iK} \log(z_{xK}) + \beta_{i\xi} \log(\xi_x) + c \quad (2.3.6)$$

Na equação acima, a utilidade da família i pelo produto x depende dos parâmetros de preferências, $\beta_i = (\beta_{i1}, \dots, \beta_{iK}, \beta_{i\xi})$. Se há um máximo interior, então as condições de 1ª ordem para a maximização da utilidade são,

$$\frac{\beta_{ik}}{z_{xk}} = \frac{\partial \mathbf{p}_t}{\partial z_{xk}} \quad \text{para } k = 1, \dots, K \quad (2.3.7)$$

$$\frac{\beta_{i\xi}}{\xi_x} = \frac{\partial \mathbf{p}_t}{\partial \xi_x}$$

Estas condições de primeira ordem podem ser resolvidas simplesmente para os parâmetros desconhecidos das preferências do indivíduo,

$$\beta_{ik} = z_{xk} \frac{\partial \mathbf{p}_t}{\partial z_{xk}} \quad \text{para } k = 1, \dots, K \quad (2.3.8)$$

$$\beta_{i\xi} = \xi_x \frac{\partial \mathbf{p}_t}{\partial \xi_x}$$

Se a função de preço, p_t e as características não observadas $\{\xi_x\}$ são conhecidas, então neste exemplo os parâmetros de preferências do consumidor i , $\beta_i = (\beta_{i1}, \dots, \beta_{iK}, \beta_{i\xi})$, podem ser obtidos mesmo no caso em que, apenas, é observada uma única escolha (z_x, ξ_x) . Pela agregação das decisões de todos os consumidores no mercado t , a distribuição acumulada dos coeficientes das preferências, $F_t(\beta)$, pode ser deduzida.

Em geral, caracteriza-se um agente por um vector de parâmetros $\beta_i \in R^B$ de dimensão B . Uma vez que, na secção anterior, se mostrou que as características não observáveis, $\{\xi_x\}$, são identificadas pela função preço, procede-se como se ξ fosse conhecido e escreve-se a função de utilidade como,

$$u_i(\mathbf{z}, c) = u(\mathbf{z}, m_i - p(\mathbf{z}); \beta_i) \quad (2.3.9)$$

onde a dependência da utilidade relativamente a ξ foi negligenciada por simplificação notacional.

Os consumidores escolhem os elementos $z \in Z$ que maximizam a utilidade. Se u e $p(\mathbf{z})$ são diferenciáveis, então as condições necessárias de primeira ordem são,

$$\frac{\partial}{\partial z_k} \{u(\mathbf{z}, m_i - p(\mathbf{z}); \beta_i)\} = 0 \quad \text{para } k = 1, \dots, K \quad (2.3.10)$$

Denomine-se a escolha óptima de z condicionada por β , por $z(\beta)$. As condições de primeira ordem podem ser implicitamente diferenciadas para produzir,

$$z'(\beta) = -[D_{z,z}\bar{u}]^{-1} D_{z,\beta_i}\bar{u} \quad (2.3.11)$$

onde
$$\bar{u}(\mathbf{z}; \beta) = u(\mathbf{z}, m_i - p(\mathbf{z}); \beta_i) \quad (2.3.12)$$

Teorema 7. Suponha $\beta_i \in B \subseteq R^B$, onde B é um subconjunto aberto convexo e $\mathbf{z} \in R^K$. Então se $z'(\beta)$ é localmente definida positiva ou definida negativa, vem que β_i é localmente definida. Se $K = B$, e $z'(\beta)$ é globalmente definida positiva ou definida negativa, então $z(\beta)$ é bijectiva.

Prova: A primeira parte do teorema decorre da versão local do teorema da função inversa. A segunda parte decorre do teorema global da função inversa, pois se **(2.3.11)** é definida positiva ou definida negativa, então $z(\beta)$ é bijectiva, por isso as preferências são globalmente identificadas.⁷⁵ □

O Teorema 7 coloca restrições apertadas sobre os tipos de funções de utilidade que podem ser identificadas utilizando dados de escolha. Condicionados ao conhecimento do domínio dos p , podemos identificar K coeficientes aleatórios por observação de escolha.

Importa referir que um modelo de escolha contínua, em termos reais, não fornece uma boa aproximação para o comportamento de escolha, podendo-se elencar, pelo menos, três razões para o facto. **(R₁)** O número de produtos que constitui o conjunto de escolha pode não ser suficientemente grande para permitir uma aproximação a um conjunto contínuo. **(R₂)** Muitas características dos produtos são fundamentalmente discretas (p.e. o número de quartos, o número de casas de banho). **(R₃)** Alguns consumidores podem escolher produtos por limites (p.e. a habitação mais próxima do local de trabalho).

2.3.5.2. Conjunto escolha discreto

Quando o espaço do produto é discreto tem-se, em vez das condições marginais constantes de **(2.3.10)**, um conjunto de desigualdades restritivas como condições necessárias à maximização da satisfação do consumidor. Se o consumidor i escolhe um produto $x \in \{1, \dots, X\}$ então,

⁷⁵ De acordo com Gale e Nikaido (1965) referido por Bajari (2005).

$$u(\mathbf{z}_x, \xi_x, m_i - \mathbf{p}(\mathbf{z}_x, \xi_x); \beta_i) \geq u(\mathbf{z}_w, \xi_w, m_i - \mathbf{p}(\mathbf{z}_w, \xi_w); \beta_i), \text{ para todo } w \neq x \quad (2.3.13)$$

Para além disso $\beta_i \in A_{ix}$, onde,

$$A_{ix} = \{ \beta_i : \beta_i \text{ satisfaz (2.3.13)} \} \quad (2.3.14)$$

Se o conjunto de escolha é finito, os conjuntos A_{ix} , normalmente, não virão singulares, daí resultando a não identificação dos parâmetros β_i . O que não significa, contudo, estar-se na presença de dados não informativos. Se o conjunto de escolha é rico, os conjuntos A_{ix} podem ser pequenos.

Sobre esta problemática passamos a mostrar que, se todas as características são contínuas e se o conjunto de escolha é compacto, então à medida que o número de produtos aumenta, o conjunto A_{ix} converge para os coeficientes de preferências individuais β_i . Nas aplicações onde a dimensão suficientemente grande do conjunto A_{ix} torna relevante a falta de identificação, mostra-se que é possível proceder de duas formas. Primeiro, o conjunto A_{ix} pode ser utilizado para construir limites sobre a distribuição agregada dos coeficientes das preferências. Segundo é possível utilizar técnicas Bayesianas para identificar uma distribuição de interesse agregada.

Mostrou-se que, quando o conjunto de produtos é discreto, os parâmetros de preferências β_i não são identificados. Entretanto, à medida que o número de escolhas se torna suficientemente grande, β_i pode deduzir-se no limite. Para além disso o conjunto A_{ix} vai reduzindo a dimensão à taxa $\frac{1}{X}$.

Para simplificar a notação, a atenção é restringida ao caso em que as características dos produtos são observáveis, quer para o consumidor, quer para o observador. A utilidade do consumidor i traduz-se por, $u_{ix} = u(\mathbf{z}_x, p_x, \beta_i)$. Introduce-se, também, $p(\mathbf{z})$ que representa preço para qualquer produto x em função das características/atributos z e assume-se essa função independente do número de produtos X presentes no mercado, embora os resultados possam ser modificados para cobrir este caso. Estabelecem-se três pressupostos sobre o X -espaço e utilidade:

Pressuposto 1. Todas as características do produto \mathbf{z}_x são elementos do conjunto aberto convexo e limitado Z . Também, todos os β_i são elementos do conjunto B , igualmente, convexo limitado e aberto.

Pressuposto 2. Para qualquer β_i , quando o conjunto escolha é todo o Z , o jacobiano $z'(\beta)$, como definido em (2.3.11) é sempre definido positivo ou definido negativo.

Suponha que se retira uma sequência de características/produtos $z^{(1)}, z^{(2)}, \dots, z^{(n)}, \dots$ a partir de \mathbf{z} . Seja $S^{(n)} = \{z^{(1)}, z^{(2)}, \dots, z^{(n)}\}$ o conjunto de escolhas disponíveis para o consumidor i . Seja $U(n)$ a utilidade óptima do consumidor i quando pode escolher a partir de $S^{(n)}$. Seja o conjunto $B^{(n)} \subseteq B$ o conjunto de coeficientes de preferências que tornam $U(n)$ uma escolha óptima retirada de $S^{(n)}$. Note-se que, como implicação dos pressupostos 1 e 2, o teorema da função inversa global pode ser aplicado e $z(\beta)$ é bijectiva.

Teorema 8. Suponha que os pressupostos 1-2 se verificam. Então com a probabilidade

$$\text{um, } \lim_{n \rightarrow \infty} B^{(n)} = \beta_i .$$

Prova: Seja z^* a característica/produto que maximiza a utilidade de uma família com coeficientes aleatórios β_i^* quando o conjunto de produtos Z está disponível. Como $n \rightarrow \infty$, $\lim_{n \rightarrow \infty} U(n) = z^*$. Seja $B^* = \bigcap B^{(n)}$. Seja $\{\beta^{(n)}\}$ uma qualquer sequência em que $\beta^{(n)} \in B^{(n)}$. Suponha que $\beta' \neq \beta^*$ está dentro de B^* . Então para todo n e todo $\tilde{z}^{(n)} \in S^{(n)}$, $u(\tilde{z}^{(n)}, \mathbf{p}(\tilde{z}^{(n)}), \beta') \leq u(U(n), \mathbf{p}(U(n)), \beta')$. Dado $n \rightarrow \infty$, decorre para todo $\mathbf{z} \in Z$, $u(\mathbf{z}, \mathbf{p}(\mathbf{z}), \beta') \leq u(\mathbf{z}^*, \mathbf{p}(\mathbf{z}^*), \beta')$. Mas isto restringe o facto de $z(\beta)$ ser bijectiva.

□

Para estabelecer que no limite os parâmetros de preferências podem ser unicamente repostos, pode-se também estabelecer uma taxa de convergência. Seja A_{ix} definida como em (2.3.14). Obviamente que $\{A_{ix}\}_{x=1}^X$ forma uma partição de B . Se l representar uma medida de Lebesgue, decorre imediatamente que:

$$\sum_{x=1}^X l(A_{ix}) = l(B)$$

$$\frac{\sum_{x=1}^X l(A_{ix})}{X} = \frac{l(B)}{X} .$$

Uma vez que o conjunto B é limitado, pode acontecer que $\frac{l(B)}{X} \rightarrow 0$, o que implica, por sua vez que a média da medida de Lebesgue de A_{ix} convirja para zero à taxa proporcional de, $\frac{1}{X}$.

2.3.5.3. Decisões de não-aquisição

Finalmente importa referir que o modelo assume que os indivíduos compram os produtos em análise, já que aqueles que decidem não comprar qualquer produto são tratados similarmente aos que compram. A decisão de não comprar é equivalente à decisão de afectar todo o rendimento à mercadoria compósita c . Na realidade, é como se estivesse a comprar um cabaz que proporciona zero unidades de qualquer característica a um preço zero. Quer se esteja num espaço discreto, quer num espaço contínuo, isto implicaria um conjunto de desigualdades para o não comprador da forma,

$$u(0, 0, m_i; \beta_i) \geq u(\mathbf{z}_w, \xi_w, m_i - \mathbf{p}(\mathbf{z}_w, \xi_w); \beta_i), \text{ para todo } w. \quad (2.3.15)$$

Estas desigualdades podiam ser utilizadas para localizar as preferências dos não compradores destes bens, similarmente ao ocorrido com as constantes de **(2.3.13)**. Note-se que, mesmo que o espaço de produtos fosse contínuo, existiria um problema de identificação para os não compradores, já que **(2.3.15)** fornece, apenas, desigualdades.

Depois de se apresentarem os pressupostos suficientes à identificação das funções preço e à identificação das características não observadas, assumindo que os preços e as características se observam para $x = 1, \dots, X$ produtos através de $t = 1, \dots, T$ mercados e que há independência entre as características observadas, z , e as características não observadas, ξ , estão reunidas condições que permitem o esboço dos modelos teórico e econométrico que fundamentam o percurso do ensaio proposto. Para o efeito dedicar-se-á o capítulo **3** a esse objectivo.

Apêndice A

Prova do Equilíbrio de Pareto (secção 2.2.2.):

Como já se referiu no corpo do texto, secção 2.2.2., as preferências, \mathcal{G} , dos consumidores e as produtividades, ν , das empresas diferem. Por isso, assume-se que \mathcal{G} e ν , são variáveis perfeitamente aleatórias e contínuas, vindo \mathcal{G} identificado com uma função de distribuição da densidade de probabilidade, $f_{\mathcal{G}} = f_{\mathcal{G}}(\mathcal{G}(\mathbf{z}, p_z, M - p(\mathbf{z}), A))$ e ν com uma função de distribuição da densidade de probabilidade, $f_{\nu} = f_{\nu}(\nu(\mathbf{z}, p_z, B))$.

Entretanto, também se assume que, quer os consumidores, quer as empresas, definem as respectivas funções preço/licitação e preço/aceitação. Em que a função preço/licitação representa o montante que o consumidor está disposto a pagar por um conjunto de atributos/características e vem expressa por $\theta(\mathbf{z}, \bar{U}, M)$ e em que a função preço/aceitação do bem, traduz o preço para esse conjunto de atributos/características que a empresa está disposta a aceitar de forma a assegurar um determinado lucro, traduzida por, $\phi(\mathbf{z}, \bar{\Pi}, \beta)$.

As empresas pretendem obter lucro, para o efeito definem a função preço/aceitação que é arbitrada com a função preço/licitação definida pelos consumidores, daí que o resultado da empresa seja dado por, $\Pi : \phi \times \theta \times \nu$. O correspondente resultado para os consumidores equacionará também as funções preço licitação e aceitação, mas ponderando a diversidade de preferências, vindo $U : \phi \times \theta \times \mathcal{G}$.

Assume-se que para cada (ν, θ) a função $\Pi(\cdot, \nu, \theta)$ é contínua e crescente em, θ . Seja,

$$\phi^*(\theta, \nu) = \arg \underset{\phi}{\text{Max}} \Pi(\theta, \phi, \nu)$$

um nível de preço/aceitação de equilíbrio bilateral de *Nash*, para uma empresa do tipo ν interagindo com um consumidor tipo \mathcal{G} . Assume-se que há pelo menos tantos consumidores, quantas empresas. Um consumidor que não interage recolhe um resultado $u^0(\theta, \nu) = u(\theta, 0, \nu)$. Assume-se que $\phi > 0 \Rightarrow u(\theta, \phi, \nu) > u(\theta, 0, \nu)$, daí que os consumidores preferam sempre interagirem estritamente com uma qualquer empresa, desde que elas apresentem um preço/aceitação positivo.

Uma vez que se assume que as funções preço/licitação e preço/aceitação são definidas, respectivamente por consumidores (i) e por produtores (e), pode-se então assumir que os pontos de tangência entre essas funções definem a função de preços hedónica.

Defina-se $Z: \phi \times \theta$, como o conjunto de características hedónicas resultantes da interacção entre as funções preço/aceitação e preço/licitação. Uma afectação é um par de transformações mensuráveis, $(\theta^e(\cdot), \phi^e(\cdot)): \nu \rightarrow Z$ e $(\theta^i(\cdot), \phi^i(\cdot)): \mathcal{G} \rightarrow Z$ e este par de transformações é identificado com α . Uma afectação α é uma afectação possível se os valores que induz sobre $Z^+ = \{(\theta, \phi) \in Z : \phi > 0\}$ são os mesmos. Uma afectação α é eficiente de Pareto se não existe uma possível afectação alternativa para a qual os consumidores e empresas não ficam pelo menos melhor.

Uma afectação possível $\alpha = \{(\theta^e(\cdot), \phi^e(\cdot)); (\theta^i(\cdot), \phi^i(\cdot))\}$ é um equilíbrio concorrencial se existir uma função hedónica $p(\mathbf{z}) : \phi \rightarrow R$, tal que para qualquer ν ,

$$\Pi(\theta^e(\nu), \phi^e(\nu), \nu) \geq \Pi(\theta^i, \phi^i, \nu)$$

para todo (θ', ϕ') tal que,

$$p(\mathbf{z}) - \theta' \geq 0$$

e

$$u(\theta^i, \phi^i, \mathcal{G}) \geq u(\theta', \phi', \mathcal{G})$$

para todo (θ', ϕ') tal que,

$$\theta' - p(\mathbf{z}) \geq 0$$

Se as empresas pretendem obter uma função preço/licitação $\theta'(\mathbf{z}, \bar{U}, M)$ devem produzir uma unidade de produto final cujo valor hedónico, $p(\mathbf{z})$, seja pelo menos igual a essa função. Por sua vez os consumidores se querem comprar uma dada quantidade de produto final devem apresentar uma função preço/licitação pelo menos de valor igual ao preço hedónico, $p(\mathbf{z})$. A forma funcional de $p(\mathbf{z})$ pode-se ajustar a partir de preços relativos não precisando de traduzir um preço monetário.

Proposição 1. Supondo as preferências monotónicas na função preço/licitação dos consumidores. Então qualquer equilíbrio concorrencial é um óptimo de Pareto.

Prova: Suponha-se que para uma afectação concorrencial α , existe uma afectação alternativa possível α' , onde quase todos os consumidores (i) e quase todas as empresas (e) estão, pelo menos, tão bem quanto estivessem em α , enquanto um conjunto estrito está melhor em α' . Admita-se uma empresa qualquer, do tipo ν , pertencente a este

conjunto estrito, que está estritamente melhor em $(\theta_{\alpha'}^e(\cdot), \phi_{\alpha'}^e(\cdot))$ do que se estivesse em $(\theta_{\alpha}^e(\cdot), \phi_{\alpha}^e(\cdot))$, pela definição de equilíbrio concorrencial vem,

$$p(z(\phi_{\alpha'}^e)) - \theta_{\alpha'}^e < 0, \text{ em que } p(\mathbf{z}) \text{ é o preço funcional que}$$

sustenta o equilíbrio inicial.

De outra forma a empresa estaria estritamente melhor a produzir segundo a função preço/aceitação $\phi(\mathbf{z}_{\alpha'}, \bar{\Pi}, \beta)$ para um preço hedónico inicial $p(\mathbf{z})$. Então, ter-se-ia,

$$\theta_{\alpha'}^e - p(z(\phi_{\alpha'}^e)) > 0$$

O que significa que o cabaz $(\theta_{\alpha'}^e(\cdot), \phi_{\alpha'}^e(\cdot))$ permanece estritamente abaixo da restrição orçamental para qualquer consumidor. Sendo assim há uma função preço/licitação inferior, θ' , tal que,

$$\theta' - p(z(\phi_{\alpha'}^e)) \geq 0$$

Vindo para cada consumidor (i) com preferências \mathcal{G} ,

$$u(\theta^i(\mathcal{G}), \phi^i(\mathcal{G}), \mathcal{G}) \geq u(\theta', \phi_{\alpha'}^e(v), \mathcal{G}) > u(\theta_{\alpha'}^e(v), \phi_{\alpha'}^e(v), \mathcal{G})$$

Por isso nenhum consumidor pode estar melhor fora da afectação $(\theta_{\alpha'}^e(v), \phi_{\alpha'}^e(v))$.

□

3. O Modelo-Base de Estimação

Observed product prices and the specific amounts of characteristics associated with each good define a set of implicit or "hedonic" prices.

Sherwin Rosen (1974, p.34)

3.1. Modelo Teórico

No intuito de desenhar um suporte para a definição de um modelo teórico que sustente a avaliação hedónica, base deste ensaio, apresentar-se-ão um conjunto de premissas, no domínio da teoria económica, de forma a enquadrá-lo no contexto da realidade económica.

Começar-se-á por supor que a satisfação das famílias, no que concerne à habitação ocupada, não se encontra otimizada, estando pois reunidas condições para que as famílias procurem maximizar a escolha residencial face à respectiva restrição orçamental. Para tanto, considera-se um mercado de habitação, em que um número razoável de famílias com características variadas (p. e., quanto ao número de membros, composição e rendimento do agregado familiar) ocupa um vasto leque de habitações caracterizadas por uma ampla variedade de diferentes características/atributos; e, considera-se, ainda, que a distribuição das famílias pelas habitações é aleatória.

As famílias para licitarem um preço que lhes permita maximizarem a sua utilidade na presença da restrição orçamental,⁷⁶ concorrem, umas com as outras, desde que estejam em condições de avaliar o "valor" que podem retirar das diferentes propriedades imobiliárias disponíveis no mercado. Como é definido pelo próprio mercado, os preços dos diferentes imóveis, decorrem da confluência de intenções das famílias que

⁷⁶ Assumem-se preferências, separadamente fracas, entre o bem habitação e as respectivas características, o que permite que a procura por características seja independente do preço dos outros bens.

concorrem na licitação de um preço e dos agentes económicos (sejam eles, imobiliárias, famílias, e/ou construtoras) que aceitam as propostas de compra. Eventualmente, pode-se imaginar que esse mercado se estabelece sobre um conjunto de preços que definem um equilíbrio – assim sendo para esses preços, o “valor” que uma qualquer família atribui à sua habitação só pode aumentar com a deslocação para outro local. Os preços alcançados quando este equilíbrio é conseguido serão os preços justos de mercado para o parque habitacional existente e respectivas características – e estes preços definem a escala de preços hedónicos.

Alterações nas condições de procura (p. e. migrações de famílias inter-locações) ou no lado da oferta (p. e. construção de novos edifícios) induzirão ajustamentos no mercado, com as famílias a reavaliarem a sua localização residencial óptima, daí resultando nova escala de preços.

Obviamente que se está perante um modelo simplificado das inter-acções que constituem o mercado da habitação e que permitem estabelecer um conjunto de preços referência para as propriedades imobiliárias e um conjunto de preços implícitos para as suas características.

A realidade do mercado de habitação é algo mais complexa. As características do parque de habitação, bem como das famílias que o virão a habitar variam, por vezes bastante, de mercado para mercado – na verdade um vasto número de analistas económicos colocaram o mercado de habitação num contexto lato de “desejabilidade” por uma área particular. Se uma área particular representa um “local desejável” para viver e trabalhar, haverá uma pressão sobre a procura por habitação e o preço da propriedade imobiliária subirá – por fim após um processo de alteração de preços e de expansão do parque imobiliário o equilíbrio de mercado ajustar-se-á. Neste cenário, será de esperar encontrar duas escalas de preços hedónicos diferentes correspondentes a dois mercados distintos, indicando preços absolutos distintos para imóveis quase idênticos – porém este facto não significa que os preços implícitos subjacentes à função hedónica de preços não sejam relativamente semelhantes.

O modelo de preços hedónicos constitui uma simplificação dos processos que acontecem no mundo real,⁷⁷ por isso é importante ter presente os pressupostos em que assenta.

Um, fundamental, é que o mercado da habitação está em equilíbrio, e para tal deve satisfazer três condições: *i*) as famílias dispõem de informação perfeita, ou seja assume-se que as famílias conhecem as características e os preços de todas as propriedades existentes no mercado – se as famílias não estiverem conscientes destes elementos é provável que o preço implícito pago pelas diferentes características varie de venda para venda, traduzindo-se em funções preço hedónicos mal definidas; *ii*) os custos de transacção são negligenciados, supõem-se nulos, e são entendidos como os custos em que uma família incorre quando muda de casa; e *iii*) as escalas de preço hedónicos ajustam-se, instantaneamente, às variações nas condições de procura ou oferta do mercado de habitação.⁷⁸

Outro, que se reporta à recolha de informação estatística e construção de base de dados relativos ao parque habitacional, e que confere a possibilidade de agregação de dados ao longo do tempo, na ausência de choques significativos sobre o parque habitacional em análise e com a salvaguarda de técnicas estatísticas para testar essa hipótese.

E, finalmente, a assumpção de mercados segmentados por cidades, ainda que este pressuposto surja associado a significativas barreiras geográficas, geradoras de custos para obtenção de informação relativamente ao mercado imobiliário de outra cidade e, também, pelos custos de movimentação entre cidades.

Equacionando as considerações acima referidas e procurando ser consistente com a especificidade do *bem* habitação e a forma como a teoria económica o enquadrou numa nova abordagem da teoria neoclássica do consumidor e numa perspectiva de equilíbrio

⁷⁷ No mundo real as funções de preço hedónica não são lineares - assim, na presença de preços implícitos não constantes e sem preferências quase-lineares, a inversa da curva da procura compensada não existiria. Na verdade, com preços marginais não-constantos, não se pode esperar qualquer relação entre a disposição marginal a pagar por um qualquer atributo e a quantidade escolhida desse atributo.

⁷⁸ Freeman (1993), refere a satisfação destas três condições para que o mercado esteja em equilíbrio *constante*.

de mercado, conforme apresentado no capítulo 2, esboçar-se-á uma arquitectura para o modelo teórico, centrada no problema de maximização da utilidade do consumidor.

Assim, um consumidor padrão defronta-se com a resolução do problema,

$$\underset{(c, z_1, \dots, z_K)}{\text{Max}} \quad U(c, \mathbf{z})$$

$$\text{sujeito a,} \quad c + p(\mathbf{z}) = M$$

com c a corresponder a um cabaz composto de mercadorias, que se assumem com um preço unitário, com M a identificar o rendimento, com $p(\mathbf{z})$ a reflectir o preço da habitação, sendo z_1, \dots, z_K os atributos/características dessa habitação, elementos do vector \mathbf{z} .

E assume-se que, (1) a função utilidade do consumidor, $U(\dots)$, é transitiva, completa e contínua e é estritamente côncava nos seus argumentos; (2) as utilidades marginais são positivas e (3) a taxa de crescimento das utilidades marginais é decrescente e (4) U é semi-definida negativa, o que significa que todas as curvas de indiferença são convexas e não tocam nos eixos.

O lagrangeano do problema de maximização da utilidade é,

$$L(c, z_1, \dots, z_K) = U(c, \mathbf{z}) + \lambda(M - c - p(\mathbf{z}))$$

Admitindo que se verificam as condições suficientes de 2ª ordem e que há uma solução interior, obtém-se, a partir das condições necessárias de 1ª ordem, a relação,

$$p_{z_k} = U_{z_k} / U_c \quad (3.1.1)$$

Entretanto, introduz-se no modelo o conceito de função “valor” ou função “preço/licitação”, de Rosen, $\theta(\mathbf{z}, \bar{U}, M)$, que representa o montante que o consumidor está disposto a pagar por um conjunto de atributos/características, mantendo um dado nível de satisfação/utilidade, implicitamente definida por,

$$U(M - \theta(\dots), \mathbf{z}) = \bar{U} \quad (3.1.2)$$

A diferenciação total desta expressão,⁷⁹ assumindo que o nível de satisfação/utilidade e de rendimento do consumidor é constante, resulta em,

$$\theta_{z_k} = U_{z_k} / U_c \quad (3.1.3)$$

Constata-se que numa situação de óptimo do consumidor o declive da tangente à função de preços hedónicos, nesse ponto, ou seja ao preço implícito de uma característica/atributo z_k , coincide com o declive da tangente à função de licitação do preço do consumidor, já que as expressões (3.1.1) e (3.1.3) são iguais. Para salvaguarda da estabilidade desta identidade assumem-se preferências quase-lineares, uma vez que, em alguns ensaios sobre teoria do consumidor aplicada a bens diferenciados, subsiste a ideia de que o declive da função de licitação do preço do consumidor será multi-dimensional, tendo tantas dimensões quantos os níveis de utilidade considerados, conforme se ilustrará em apêndice (Apêndice B).

⁷⁹ Expressa por, $\frac{\partial U}{\partial c} \frac{\partial c}{\partial M} dM - \frac{\partial U}{\partial c} \frac{\partial c}{\partial \theta} \frac{\partial \theta}{\partial M} dM - \sum_{k=1}^K \frac{\partial U}{\partial c} \frac{\partial c}{\partial \theta} \frac{\partial \theta}{\partial z_k} dz_k - \frac{\partial U}{\partial c} \frac{\partial c}{\partial \theta} \frac{\partial \theta}{\partial U} dU + \frac{\partial U}{\partial z_k} dz_k = dU$.

Acontece que, em equilíbrio, a coincidência dos declives das tangentes às funções de preços hedónicos e de licitação do consumidor, constantes das expressões (3.1.1) e (3.1.3), também é extensível ao declive da função de custos, quando o lado da oferta do *bem* é equacionado.⁸⁰

Assim sendo, o esboço do modelo teórico só estará completo se for equacionado o problema do produtor/empresa do *bem* habitação, ou seja o objectivo de maximização do lucro, passando, também, a ponderar-se o lado da oferta. Assumindo uma função de custos convexa e definindo implicitamente uma função preço/aceitação do *bem* no mercado,⁸¹ face a um dado lucro, a coincidência referida é uma realidade. Em termos analíticos, o problema que, uma qualquer, empresa defronta é,

$$\text{Max } \Pi = \bar{F} p(\mathbf{z}) - C(F, \mathbf{z}, \beta)$$

com $p(\mathbf{z})$ a reflectir o preço da habitação e $C(\dots)$ a identificar a função de custos que é função da quantidade, F , dos atributos/características da habitação, z_1, \dots, z_K , elementos do vector \mathbf{z} , e dos preços dos factores de produção e tecnologias, β .

As condições de 1ª ordem permitem, entre outras, a interpretação de que o lucro máximo exige que o preço marginal de cada característica seja igual ao custo marginal unitário de a produzir, expresso por,

$$p_{z_k} = \frac{C_{z_k}}{\bar{F}} \quad (3.1.4)$$

⁸⁰ Ponderação que reflecte o importante contributo de Rosen.

⁸¹ *Offer function*, na terminologia de Rosen.

Entretanto, tal como ocorreu no lado da procura, introduz-se uma função, $\phi(\mathbf{z}, \bar{\Pi}, \beta)$, que pode ser denominada de função preço/aceitação do *bem*, por traduzir o preço que a empresa está disposta a aceitar de forma a assegurar um determinado lucro, passível de ser definida implicitamente por,

$$\Pi = \bar{F}\phi(\mathbf{z}, \bar{\Pi}, \beta) - C(F, \mathbf{z}, \beta) \quad (3.1.5)$$

A diferenciação total da equação (3.1.5), no cenário de valores de produção e de lucro constantes, isto é $d\bar{F} = d\bar{\Pi} = 0$, gera a seguinte relação,

$$\phi_{z_k} = C_{z_k} / \bar{F} \quad (3.1.6)$$

A observação das equações (3.1.4) e (3.1.6) revela que, em equilíbrio, os declives das tangentes, nesse ponto, à função de preços hedónicos e à função preço/aceitação coincidem.

Da conjugação das condições de equilíbrio constantes, quer do lado da procura, quer do lado da oferta, resulta,

$$p_{z_k} = \theta_{z_k} = \phi_{z_k} = C_{z_k} / \bar{F} = U_{z_k} / U_c \quad (3.1.7)$$

Esta relação entre os declives das funções de preços hedónicos, os declives das funções preço/licitação dos consumidores e os declives das funções preço/aceitação dos produtores, fundamenta a investigação no domínio dos mercados hedónicos.

Uma vez que se assume que as funções preço/licitação e preço/aceitação, são definidas, respectivamente por consumidores e por produtores, pode-se, então, assumir que os pontos de tangência entre essas funções definem a função de preços hedónica. Os resultados constantes da proposição (3.1.7), justificam a perspectiva hedónica da análise de mercados, pela qual se pode estimar o preço hedónico para um dado conjunto de características/atributos. O que permite a este problema da escolha do consumidor, no pressuposto da optimização, adquirir contornos similares aos da análise padrão do comportamento do consumidor, em que escolhas observadas e preços finais informam sobre as preferências desses consumidores, capazes de inferências rigorosas na presença de informação suficiente.

A estimação exigida pela observação da relação entre o preço da habitação e as respectivas características conduzir-nos-á ao desenvolvimento de um modelo econométrico na secção seguinte.

3.2. Modelo Empírico

Para se concretizar com sucesso a estimação de uma função de preços hedónica para o mercado da habitação, há que inventariar um conjunto assinalável de atributos/características presentes nesse *bem*, de forma a garantir uma explicação fiável dos preços de venda da habitação. Nos estudos de preços hedónicos o recurso ao maior número de variáveis explicativas possível deve ser uma realidade, ainda que essa extensão de variáveis explicativas na estimação da função de preços hedónica possa introduzir problemas de multicolinearidade, pois aumenta a possibilidade de duas ou mais variáveis explicativas terem uma relação muito semelhante com a variável

dependente.⁸² No entanto, negligenciar algumas dessas variáveis pode conduzir a enviesamentos sérios na estimação dos parâmetros das variáveis consideradas.⁸³

Para melhor concretizar este objectivo de exaustividade dos atributos/características da habitação poder-se-á recorrer à taxionomia de Tinch (1995) na identificação das variáveis explicativas do preço da habitação. Assim, variáveis como, número de quartos, número de casas de banho, existência ou não de garagem, dimensão do jardim, presença ou não de aquecimento central, entre outras, passam a ser identificadas com características classificadas como *variáveis estruturais*; outras variáveis como, distância às paragens de transporte público, ao centro da cidade, às escolas, etc, passam a integrar-se no conjunto de características classificadas como *variáveis acessibilidade*. Já, na categoria *variáveis vizinhança/envolvente socioeconómica (neighbourhood)* passam a enquadrar-se características como a idade média dos edifícios, a taxa de crime, a composição racial e a qualidade das escolas nesse espaço de localização da habitação. Finalmente, características como níveis de ruído, poluição do ar, qualidade das vistas oferecidas pelo imóvel, agrupam-se na categoria, *variáveis ambientais*.

Caso o modelo seja estimado com o único propósito de prever os preços totais das habitações, isto é com o objectivo da simples inferência do valor das propriedades individuais, então a perspectiva paramétrica, onde se escolhe primeiro a forma funcional e, posteriormente, se estimam os parâmetros que definem essa forma funcional, com o melhor nível de significância, revela-se bastante apropriada

Relativamente à forma funcional, que estabelece a natureza da relação entre a variável dependente e as variáveis explicativas, há que ajustar a forma correcta de modo a impedir o enviesamento decorrente de uma má especificação. A teoria económica

⁸²Na presença deste problema os parâmetros estimados podem apresentar valores exagerados e na pior das hipóteses sinais errados. É possível diluir o problema construindo índices, combinando as variáveis independentes altamente correlacionadas, com base na *análise dos componentes principais*; mesmo assim, continua a subsistir outra dificuldade – a estimação do efeito isolado sobre o preço da propriedade imobiliária de variáveis individuais, de interesse. Embora Atkinson e Crocker (1987, p. 29), empiricamente, identifiquem, “(...) *that the specification uncertainty caused by collinearity is small for structural attributes (e. g. floor space, age, and lot size) but substantial for neighbourhood attributes (e.g. air pollution, school quality, and crime)*”.

⁸³ Harrison e Rubinfeld (1978) mostraram que o parâmetro relativo à poluição do ar se alterava significativamente quando as variáveis de acessibilidade eram retiradas do estudo de preços hedónicos; enviesamento tanto mais relevante, quanto maior a correlação entre as características ambientais e de acessibilidade.

difícilmente fornece meios capazes para seleccionar uma forma funcional adequada. Muitas vezes, a evolução da especificação inicial de uma forma funcional faz-se pelo processo da tentativa e do erro. As formas funcionais que melhoram o ajustamento do modelo aos dados são, supostamente, aquelas que melhor representam a verdadeira função de preços hedónica.

A escolha da forma funcional é importante, especialmente quando se suspeita de não linearidade. O método Box-Cox tem sido escolhido por ser flexível, já que a adopção de uma forma flexível para a função de preço hedónica é, consensualmente, justificada pelas poucas restrições, colocadas pela teoria económica relativamente a essa questão. É uma realidade que algumas dessas restrições, como sejam, o valor da terra em função da localização, no contexto da teoria económica urbana; a procura contínua por características/atributos, em equilíbrio e os bens públicos e amenidades ambientais passíveis de caracterizar a habitação, exigem ponderação na forma funcional, tradução em funções de preço hedónicas convexas⁸⁴ e revelam limitações da especificação paramétrica, respectivamente.

Por esta via os preços transformados passam a expressar-se como função linear ou quadrática das características, não dicotómicas, transformadas - com parâmetros de transformação, possivelmente, diferentes entre as variáveis transformadas, os quais seriam estimados por via de outros parâmetros. A estimação pela máxima verosimilhança tem-se revelado poderosa, mesmo na presença de um número elevado de variáveis exógenas, atributos/características.

Tentativas para contemporizar esses modelos vão, desde a estimação do modelo Box-Cox por métodos da máxima verosimilhança ignorando a correlação espacial, até à linearização da transformação Box-Cox. Entretanto a linearização é uma aproximação que é válida, apenas, para valores de parâmetros específicos. A má especificação da forma funcional e/ou a negligência da dependência espacial⁸⁵ podem resultar em

⁸⁴ Jones (1988).

⁸⁵ Para testar, conjuntamente, a forma funcional e a dependência espacial, autores, como Baltagi e Li (2001), introduziram o teste do Multiplicador Lagrangeano (*LM*). Anselin e Bera (1998), já, tinham testado a dependência espacial, assumindo previamente uma forma funcional específica, definida pelo modelo Box-Cox, e a partir dessa forma funcional, escolhem o termo erro para prever a

inferências erradas e aumentar as dúvidas sobre a verdade e a precisão dos dados estimados.

Optando pela especificação hedónica tradicional, onde os efeitos espaciais não são incorporados, testes espaciais, baseados no multiplicador lagrangeano (*LM*) têm que ser utilizados para detectar efeitos espaciais⁸⁶ nos resíduos da regressão estimada. O teste *LM* é construído sobre as propriedades assintóticas do estimador da máxima verosimilhança (*MLE*) e o seu desempenho em amostras finitas é, ainda, uma questão sobre escrutínio. Se as propriedades assintóticas da distribuição dos termos de erro falharem então o teste *LM*⁸⁷ pode não ser suficientemente válido.

Relativamente à especificação da forma funcional da função de preços hedónicos é importante reter a estrutura espacial do mercado de habitação, a qual implica variabilidade espacial, quer via efeitos locacionais, quer via externalidades, negativas ou positivas, sobre os preços de habitação, traduzível em heterogeneidade nos parâmetros da função de preços hedónica. A entrada directa dos efeitos *vizinhança/proximidade*⁸⁸ como determinantes dos preços da habitação não deve implicar a existência de um mercado competitivo no longo prazo, já que se considera que, a longo prazo, haverá apenas uma escala de preços. Porém, as diferenças de efeitos *vizinhança/envolvente socioeconómica*, ao conduzirem a preços implícitos (preços marginais de atributos) diferentes, indiciarão a presença de escalas de preços diferentes, em suma a presença de um mercado segmentado. Essa existência de sub-mercados geográficos introduz o risco de se violar o pressuposto de equilíbrio de longo prazo nos mercados de habitação; já que, dentro da mesma área metropolitana, haverá diferentes escalas de preços hedónicos, reflectindo as respectivas estruturas de procura e oferta nesses sub-mercados.

correlação espacial, definindo $u = \psi \mathbf{W}u + v$, em que ψ é o coeficiente auto-regressivo espacial, \mathbf{W} é a matriz de ponderações espaciais e $v \sim N(0, \sigma_v^2 I)$, independente de, u .

⁸⁶ A presença de efeitos espaciais violará o pressuposto do erro padrão relativamente à normalidade, ou seja da distribuição dos termos de erro independente e identicamente (i.i.d.).

⁸⁷ Anselin e Rey (1991), num estudo baseado em experiências de simulação de Monte Carlo, relativamente ao comportamento do teste *LM*, fizeram as seguintes observações: i) o teste, para detecção da presença de dependência espacial nos termos de erro, é altamente sensível ao pressuposto da normalidade; ii) os testes para identificar a presença do desfasamento espacial relativamente à variável dependente revelam-se mais robustos para revelar a presença da heteroscedasticidade – daí sugerirem uma interpretação crítica do teste *LM*.

⁸⁸ Denominados por efeitos *neighbourhood*.

A perspectiva econométrica para incorporar a heterogeneidade dos coeficientes na especificação da função de preço hedónica e, simultaneamente, enfrentar a violação do pressuposto do equilíbrio de mercado de longo prazo, passa por assumir, *a priori*, que a relação entre as despesas de habitação e as características estruturais da habitação, é diferente entre espaços geográficos sujeitos aos mesmos efeitos *vizinhança/envolvente socioeconómica*. Nesta óptica, o mercado vem estratificado em segmentos geográficos discretos e uniformes, passando as funções de preço hedónicas a estimarem-se, simultaneamente, para cada um desses segmentos e para o mercado global.⁸⁹

A concretização desta estimação simultânea para o mercado global da habitação e para cada um dos seus segmentos, conduz à estimação da tradicional função de preços hedónica,

$$\mathbf{p} = f(\mathbf{s}, \mathbf{n}, \mathcal{G}, \mathbf{A}) + \boldsymbol{\varepsilon} ,$$

onde \mathbf{p} é o vector das despesas com habitação observadas no mercado (preço da habitação); \mathbf{s} e \mathbf{n} são, respectivamente, vectores de características estruturais e de características *vizinhança/envolvente socioeconómica*, pertencentes ao vector das características/atributos, \mathbf{z} ; \mathcal{G} e \mathbf{A} são os vectores de parâmetros correspondentes a gostos e preferências dos consumidores e $\boldsymbol{\varepsilon}$ o vector dos termos aleatórios de erro, e, também, à estimação, para cada segmento de mercado, $t = 1, \dots, T$, de,

$$\mathbf{p}_t = f_t(\mathbf{s}, \mathcal{G}) + \boldsymbol{\varepsilon}_t .$$

Entretanto, a utilização do teste de Chow, fornecerá, ou não, evidência estatística relativamente à ocorrência de instabilidade estrutural.

⁸⁹ Concretizando a técnica subjacente às *switching regressões*.

Tendo em conta as questões introdutórias, a este ponto, não nos parece razoável impor, *a priori*, uma estrutura definida para os dados recolhidos, e na senda de uma linha comum a ensaios desenvolvidos sobre o tema, vamos avançar com um modelo, suficientemente lato, capaz de incluir muitas das mais comuns especificações, sustentado na transformação Box-Cox,⁹⁰ assim,

$$P(\Theta) = \alpha + \sum_{k=1}^K \beta_k \mathbf{z}_k(\lambda_k) + \sum_{k'=1}^{K'} \beta_{k'} \mathbf{z}_{k'}(\lambda_{k'}) + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^{K'} \beta_{kk'} \mathbf{z}_k(\lambda_k) \mathbf{z}_{k'}(\lambda_{k'}) + \sum_{l=1}^L \gamma_l \mathbf{D}_l + \varepsilon$$

$$P(\Theta) = \begin{cases} \frac{P^\Theta - 1}{\Theta}, & \Theta > 0 \\ \ln(P), & \Theta = 0 \end{cases} \quad \text{e} \quad \mathbf{z}_u(\lambda_u) = \begin{cases} \frac{\mathbf{z}_u^{\lambda_u} - 1}{\lambda_u}, & \lambda_u > 0 \\ \ln(\mathbf{z}_u), & \lambda_u = 0 \end{cases}, \quad \text{para } u = k \vee k'$$

os β_k 's, $\beta_{k'}$'s, $\beta_{kk'}$'s, γ_l 's são parâmetros a estimar; os Θ , λ_k e $\lambda_{k'}$ são parâmetros de transformação Box-Cox⁹¹; \mathbf{z}_k , $\mathbf{z}_{k'}$ são atributos/características medidas de forma contínua, neste caso para além da área e do ano de construção, considerou-se a tipologia, o nº de casas de banho e o número de garagens; \mathbf{D}_l os atributos/características medidas de forma dicotómica, P o preço da habitação e ε a distribuição normal do termo de erro estocástico.

⁹⁰ Tratando-se da forma quadrática Box-Cox, a forma funcional mais lata, incluirá todas as formas funcionais e estarão disponíveis testes de hipóteses (p.e. o teste do rácio de verosimilhança).

⁹¹ Na forma linear Box-Cox, apresentada em Cropper *et al.* (1988, 1993), todas as variáveis independentes têm parâmetros de transformação λ_k e $\lambda_{k'}$ iguais; as variáveis dicotómicas não são transformáveis e o parâmetro de transformação da variável dependente, Θ , é diferente do das variáveis independentes.

Os parâmetros Θ , λ_k e $\lambda_{k'}$, podem ser determinados por via da máxima verossimilhança, maximizando a função,

$$L_{\max} = (\Theta - 1) \sum_{n=1}^N \ln P_n - \frac{N}{2} \ln \left(\frac{SSR \lambda_k \lambda_{k'}}{N} \right)$$

Para maximizar a função de verossimilhança pode-se implementar uma técnica a dois tempos. Primeiro escolhem-se valores para Θ , λ_k e $\lambda_{k'}$ e transformam-se os dados observados. Em função das restrições assumidas pelos parâmetros podem-se considerar três pares de especificações, assim:

<div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;"> Quadrática Box-Cox Θ e λ assumem \forall valor </div>		
Translog $\Theta = \lambda = 0$	Quadrática $\Theta = \lambda = 1$	Semilog Quadrática $\Theta = 0, \lambda = 1$
loglinear $\beta_{kk'} = 0, \forall k, k'$	Linear $\beta_{kk'} = 0, \forall k, k'$	Semilog $\beta_{kk'} = 0, \forall k, k'$

Quadro 3.2-1: Tipologia Box-Cox de especificações

Posteriormente, passa-se a avaliar a função logarítmica de verossimilhança,

$$L(\Theta, \lambda) = (\Theta - 1) \sum_{n=1}^{400} \ln P_n - \frac{400}{2} \ln \left(\frac{SSR \lambda_k \lambda_{k'}}{400} \right),$$

em que $N = 400$ é a dimensão da amostra, SSR é a variância dos valores transformados e P_n é o preço observado na observação n . Depois de obtido o valor máximo dessa função, aplicam-se testes de hipótese relativamente às formas funcionais mais restritivas.

Em suma, a estimação hedónica dos preços confronta os investigadores com um exemplar, bastante razoável, de dificuldades padrão, característico de trabalhos de estimação sustentados em bases de dados seccionais. Dificuldades que passam pela escolha de especificação paramétrica apropriada - quer da escolha funcional, quer da selecção de variáveis a incluir - pela convivência com a colinearidade e com dados mal condicionados; pela detecção de potencial heteroscedasticidade e erros não normais; por regressores sujeitos a erros de medida e por estimadores de máxima verosimilhança para relações não lineares.

Para além dos problemas que a estimação, a partir da observação da relação entre o preço do bem diferenciado/habitação e respectivas características, coloca, persistem, ainda, campos de pesquisa abertos a investigação futura, decorrentes das especificidades caracterizadoras dos mercados da habitação, que se prendem com: (i) a particularidade do acto de compra de habitação identificável com um acto envolvendo uma procura continuada, que cessa, apenas, quando o acréscimo de utilidade esperada por essa procura é inferior ao custo da mesma; (ii) a relação intrínseca da característica espacial com a habitação, podendo tomar vantagem sobre todas as outras características equacionadas e (iii) a convivência de habitações “novas” e “velhas” no mercado, colocando dúvidas quanto à incorporação deste elemento no lado da oferta.

3.3. Conceptualização do Modelo de Preços Hedónicos

A fonte para obtenção de informação necessária à estimação do modelo econométrico, na ausência de informação estatística desagregada (ou mesmo agregada), quer do INE,⁹²

⁹² Instituto Nacional de Estatística.

quer da APEMI,⁹³ teve que ser direccionada para a observação directa do mercado, neste caso o da cidade da Guarda, via processo de recolha de informação sustentado em inquérito dirigido às imobiliárias. A elaboração desse inquérito teve, subjacente, o objectivo de recolher dados relativos a transacções no mercado imobiliário da Guarda, actuais e individuais. No presente ensaio, a inexistência de bases de dados institucionais foi determinante pela opção por esta metodologia, no entanto Freeman (1993) utiliza-a por convicção, em detrimento dos dados agregados do *census* norte-americano (sejam dados fornecidos por avaliadores, sejam valores conhecidos para fins fiscais), justificada no pressuposto de que os dados, relativos às transacções, reflectem melhor o “verdadeiro” preço da transacção. É evidente que o “verdadeiro” preço de mercado pressupõe que todos os agentes económicos dispõem de informação perfeita e simétrica e esgotam todas as oportunidades de ganhos adicionais de troca.

A satisfação do objectivo de recolha de dados relativos a transacções no mercado imobiliário da Guarda impõe obter informação sobre o preço de mercado da habitação, isto é sobre o preço da transacção, bem como informação sobre um conjunto de características/atributos que caracterizam essa habitação, de forma a construir-se uma base de dados, que permita identificar as variáveis observadas, quer endógenas, quer exógenas, necessárias à estimação do modelo econométrico.

Identifica-se, de seguida, o conjunto de informação auscultada e recolhida, já com a denominação variável *endógena* e/ou *exógena*, necessária à modelização,

(i) Variável *endógena* (explicada/dependente)

O Valor de mercado da propriedade imobiliária/habitação, entendido, na perspectiva utilitarista, como a utilidade ou a desutilidade conferida pelos atributos/características da habitação para além da detenção integral do *bem*/habitação, há muito, que a moderna teoria económica, reconheceu a impossibilidade de o identificar, restando, apenas, conhecer o preço de mercado.

⁹³ Associação Portuguesa de Empresas de Mediação Imobiliária.

Assim, os preços actuais das transacções constituem os preços de mercado, mais prováveis, a procurar pela maior parte dos investigadores, Miller (1982) e Cheschire, P. e Sheppard, S., (1995), com a vantagem de, geralmente, coincidirem com os preços observados.⁹⁴

Para se conseguir observar o preço de transacção da habitação, solicitou-se informação sobre o Valor de Mercado e o Valor de Venda. O Valor de Mercado é subentendido como o valor do imóvel no mercado (o preço pedido) e, sem que as imobiliárias o assumam, acaba por ser o preço a que a transacção se realiza e o Valor de Venda o preço por que se regista contabilisticamente essa transacção (o preço a que supostamente se “fecha” a transacção).⁹⁵ A opção por esta terminologia, talvez pouco ortodoxa no contexto da teoria económica, decorre do facto de existir, há data, no mercado imobiliário em geral, e no da Guarda, em particular, por norma, um desvio entre o preço efectivo de transacção e o preço registado (o preço registado pode variar entre 70% a 90% do preço de transacção). Essa terminologia ao permitir ganhar a confiança das imobiliárias garantia maior fiabilidade à informação recolhida já que consente maior sinceridade de resposta.⁹⁶

(ii) Variáveis *exógenas* (explicativas/independentes)

I - Variáveis estruturais:

Contínuas em termos métricos:

Área da habitação: informação medida em m².

⁹⁴ Cheschire e Sheppard (1995), pp. 247-267, num estudo relativo ao valor da terra na cidade de Reading, equacionando a dificuldade de obter informação sobre preços de transacção, consideram os *asking prices* (preços pedidos) como uma boa aproximação aos preços de transacção, em períodos de estabilidade dos mercados. Constataram que, para as casas transaccionadas, no período da observação, o preço de transacção médio estava na vizinhança de 1% do *asking price*.

⁹⁵ Este valor, em regra, coincide com o valor constante do registo notarial, conforme se pôde comprovar ao perseguir se o rasto de quase 90% das transacções informadas pelas imobiliárias em sede notarial, após autorização do Ministério da Justiça.

⁹⁶ Esta sensibilidade suporta-se na sugestão de um mediador imobiliário, corroborada por um engenheiro avaliador junto de uma instituição bancária, ambos coincidiram neste léxico para um maior sucesso na recolha de elementos conducentes ao efectivo preço da transacção.

Idade: informação apurada a partir do confronto entre o Ano de Construção e o Ano da Venda.

Casas de banho: informação quantificada em termos unitários, 1, 2, 3.

Quartos: informação quantificada, em termos unitários, 1, 2, 3, 4, 5

Dicotómicas/Dummy:

Garagem: a informação recolhida pretendia avaliar essa existência. E em caso afirmativo, o número de lugares, 1 ou 2 carros. A modelação, para tornar problemas de matriz singular decorrentes da contemplação exaustiva do universo dicotómico, incidiu apenas sobre o nível afirmativo, vindo, assim *GSUM*, para representar, Sim=1, Carros=1 e *GSDOS*, para representar, Sim=1, Carros=2. Por esta via a não existência de garagem também é equacionada, quando *GSUM* =0 e *GSDOS* =0.

Gás Canalizado: a informação pretendia apenas identificar essa característica, traduzindo-se o Sim por 1; e o Não por 0.

Aquecimento Central: o mesmo objectivo e a mesma metodologia de modelização, traduzida por, Sim = 1; Não = 0.

Lareira: o mesmo objectivo e a mesma metodologia de modelização, traduzida por, Sim = 1; Não = 0.

Arrecadação: o mesmo objectivo e a mesma metodologia de modelização, traduzida por, Sim = 1; Não = 0.

Elevador: o mesmo objectivo e a mesma metodologia de modelização, traduzida por, Sim = 1; Não = 0.

Janelas: a informação recolhida visava identificar o tipo de janelas, categorizando-as em Duplas e Simples, modeladas assim, Duplas = 1; Simples = 0.

Piso: para a recolha de informação consideraram-se as seguintes possibilidades, cave; r/chão; 1º; 2º; 3º; 4º; superior ao 4º; a respectiva modelação realiza-se por via de 7 variáveis *dummy*. Considera-se que uma propriedade só ocupa um piso, assim, conforme a correspondência na posição da estrutura, a variável “piso” (cave; r/chão; 1º; 2º; 3º; 4º; superior ao 4º) assume o valor de 1; todas as outras alternativas de posição assumem o valor 0.

Exposição ao sol: para se obter esta informação, teve-se que impor parcimónia ao leque de hipóteses a considerar,⁹⁷ sob pena de não se conseguir concretizar o objectivo de recolha de dados e ficou-se pela simples caracterização de Boa e Má, modelizada assim Boa = 1; Má = 0.

II – Variáveis geográficas/espaciais:

Localização: A recolha de informação pretendeu conhecer os locais onde a habitação se situa, elencando para o efeito bairros da cidade, a saber, Centro; Sr^a Remédios; Luz; Castelos Velhos; Bonfim; Torrão; Lameirinhas; Ferrinho; Rio Diz; Póvoa de Mileu; S. Domingos; Estação; Sequeira. A modelização destas variáveis consiste na atribuição do valor 1 ao bairro onde se localiza a habitação e, simultaneamente, valor 0 a todos os outros bairros.

Uma vez recolhida a informação, recorrendo a cinco agências imobiliárias, conseguiu-se construir uma base de dados com 400 observações de transacções de andares/habitação, realizadas na cidade da Guarda, para o período de 2002, estava-se em condições de estimar a função de preços hedónicos,

⁹⁷ Escrutinar as possibilidades de janelas e ou frentes viradas a nascente e/ou a poente, bem como orientações a Norte e/ou a Sul avaliariam melhor a qualidade de exposição solar. Porém perseguir este objectivo significaria aumentar a complexidade de resposta em detrimento da disponibilidade de colaboração das imobiliárias.

$$VMERC = f(AREA, QUARTOS, CBANHO, ANOS, GARAG, LAR, JANEL, EXPSOL, \\ ELEV, AQCENTRAL, ARREC, GCANAL, PISO, LOCAL)$$

Sobre a estimação e todas as questões relativas à inferência estatística se dará desenvolvimento no capítulo 4.

Apêndice B

Função de licitação do preço do consumidor e multidimensionalidade do seu declive

A função de preços hedônicos estima-se, aplicando técnicas de regressão, a partir de informação recolhida sobre os preços de venda da propriedade imobiliária, sobre os detalhes dos atributos/características dessas propriedades e, ainda se possível, sobre as características das preferências das famílias. Este processo é muitas vezes identificado como análise hedónica do 1º estágio, de acordo com a nomenclatura de Rosen.

O investigador pretende observar o comportamento de uma família, num dado mercado, num processo de escolha da habitação que lhe maximize a sua utilidade. Assim, admitindo que essa família enfrenta uma função de preços implícita para um

determinado atributo z_k , $p_{z_k}(z_k) = \frac{\partial p(\mathbf{z})}{\partial z_k}$, ela escolherá a habitação que lhe confere

a máxima utilidade, quando o pagamento adicional por uma unidade adicional de z_k , tudo o resto constante, coincide com o preço que o consumidor está disposto a pagar por uma unidade adicional de z_k , de forma a manter um dado nível de utilidade, u_i . Em

termos geométricos tratar-se-á de um ponto de tangência entre $p_{z_k}(z_k) = \frac{\partial p(\mathbf{z})}{\partial z_k}$

(função de preços implícita) e $b_{z_k}(u_i) = \frac{\partial \theta(\mathbf{z})}{\partial z_k}$ (função preço/licitação marginal). Os

problemas de identificação, decorrentes da ampliação da teoria da utilidade aos bens diferenciados, residem neste ponto de tangência, já que conhecê-lo não é suficiente para identificar a função preço/licitação marginal total.

A identificação da função preço/licitação marginal exigiria mais informação. Precisar-se-ia, especificamente, de conhecer mais sobre as funções preço/licitação marginal para outros níveis de z_k , mas conservando o nível de utilidade u_i . Uma possibilidade

passaria por analisar o comportamento de uma outra família, num outro mercado, admitindo que os rendimentos e as características socioeconómicas de ambas as famílias são idênticos, o que implicaria as mesmas preferências. Significando pois, que ambas as famílias escolheriam o mesmo cabaz de atributos/características a estar presente na escolha de habitação óptima, desde que enfrentassem a mesma escala de preços hedónicos.

Acontecem, no entanto, eventuais diferenças nas condições de oferta e/ou procura entre ambos os mercados, que conduzem a que a função de preços hedónicos de equilíbrio seja diferente entre os dois mercados. Desta forma, em vez de um ponto óptimo,

$$p_{z_k}(z_k) = \frac{\partial p(\mathbf{z})}{\partial z_k} = b_{z_k}(u_i) = \frac{\partial \theta(\mathbf{z})}{\partial z_k},$$

passar-se-ia a ter um outro ponto óptimo e nesse cenário a forma da função de preço/licitação marginal poder-se-ia estabelecer.

Porém, há que ter em conta que, sendo a função de preços hedónicos diferente no segundo mercado, a escolha óptima da segunda família pode não atingir o mesmo nível de utilidade, podendo ser maior ou menor em função dos preços. No segundo mercado o que se observaria seria,

$$p_{z_k}(z_k) = \frac{\partial p(\mathbf{z})}{\partial z_k} = b_{z_k}(u_j) = \frac{\partial \theta(\mathbf{z})}{\partial z_k} \neq b_{z_k}(u_i) = \frac{\partial \theta(\mathbf{z})}{\partial z_k}$$

e não se ganharia informação sobre a forma da função preço/licitação marginal, a não ser que se estivesse na presença de preferências quase-lineares.

Além desta restrição, há ainda uma outra realidade, que tem a ver com o facto de as famílias serem incapazes se desagregar uma habitação nas suas partes constituintes e

desfrutar do benefício de cada uma dessas partes separadas do todo, daí resultando que a função de preços hedônicos não é linear. A ser assim, o preço pago por uma unidade extra de um atributo/característica é função do nível desse atributo/característica, e o preço implícito de uma característica/atributo declina quando o nível dessa característica aumenta.

Estas limitações são passíveis de se transporem admitindo preços implícitos constantes e considerando três mercados e três famílias, com o mesmo rendimento e com as mesmas características socioeconómicas, para que essas três famílias sejam observadas como se fossem apenas uma família no processo de escolha de uma habitação em cada um dos três mercados. A análise será direccionada para a escolha de um dado atributo z_1 , conforme se ilustra no gráfico 1.

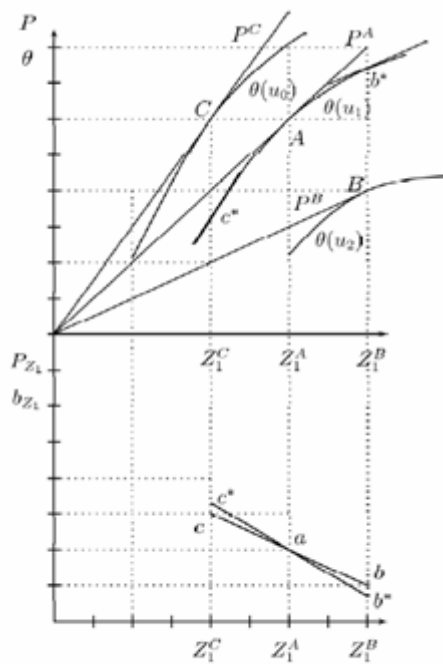


Gráfico 1: Famílias Homogêneas

Num dos mercados considerados, denominado por mercado A, a família enfrenta uma função de preços hedônicos, P_A , linear, o que significa um declive constante, quando definida para um dado parâmetro, z_1 , traduzido por,

$$p_{z_1}^A = \frac{\partial p(\mathbf{z})}{\partial z_1}$$

ponto A onde a função de preço/licitação é tangente à função de preços hedônicos, com conseqüente correspondência num dado nível óptimo de atributo/característica, z_1^A . Significa,

$$\text{pois que } p_{z_1}^A = b_{z_1}^A(z_1, u_1), \text{ ficando-se a}$$

conhecer, também, o nível de utilidade u_1 conferido pela habitação.

Alargando este raciocínio ao mercado B, com uma função de preços hedônicos P_B , com declive diferente de P_A , ilustrando as diferenças nas condições de procura e/ou

oferta dos mercados, vem um preço implícito, $p_{z_1}^B$, menor do que o verificado no mercado A. Como o preço de cada unidade de z_1 é agora mais baixo, a família é capaz de alcançar um nível de utilidade mais elevado. Como se pode constatar graficamente, a família maximiza a sua utilidade escolhendo um nível de atributo, z_1^B , e nesse ponto B a família atinge uma curva preço/licitação mais elevada, $\theta(z_1, u_2)$, onde concretiza um nível de utilidade u_2 . De novo se obtém um ponto passível de representação no quadro da análise marginal, b , mas trata-se de um ponto pertencente à curva preço/licitação marginal, $b_{z_1}(z_1, u_2)$ e não à curva preço/licitação $b_{z_1}(z_1, u_1)$.

Fica-se pois sem informação acerca da forma da curva preço/licitação marginal, a não ser que se tente localizar sobre a curva preço/licitação inicial, $\theta(z_1, u_1)$, a disposição a pagar para se ter a quantidade z_1^B de atributo, sendo este um comportamento não observado do mercado. Verifica-se que o declive neste ponto de tangência, b^* , não observado, é ligeiramente inferior ao do observado.

Este será o caso para qualquer atributo/característica desde que se comporte como um bem normal, apenas se as famílias tiverem preferências quase-lineares esses dois declives coincidirão e nesse caso valores observados e não observados serão idênticos.

Finalmente, o mercado C, e novamente preço implícito constante, $p_{z_1}^C$, agora um pouco mais elevado do que o registado nos outros dois mercados, novamente por razões que se prendem com as características da oferta e/ou procura nesse mercado, o que se traduzirá em menos utilidade para a família a actuar neste mercado. Repetindo o mesmo raciocínio, subjacente à optimização da função utilidade da família, já descrito aquando da análise do comportamento da família nos outros mercados, observar-se-á o ponto, C, de coordenadas, $(z_1^C, p_{z_1}^C)$, a transportar, também, para o quadro da análise marginal, ilustrado por, c .

Mais uma vez se regista ausência de informação relativamente à forma da curva preço/licitação marginal, a não ser que se tente localizar a disposição a pagar para se ter a quantidade z_1^C de atributo, com o nível de satisfação inicial u_1 , verificando, para tanto, o efeito sobre a curva preço/licitação inicial, $\theta(z_1, u_1)$; esta situação, no entanto, não é observável no mercado. Este cálculo auxiliar permite constatar que o declive deste ponto de tangência, c^* , não observado, é ligeiramente superior, ao do observado.

Esta análise traz à discussão pontos observados e pontos não observados conforme se ilustra graficamente. Verifica-se que, para definir a função/curva de preço/licitação marginal, apenas se dispõe de um ponto a , o que geometricamente se torna impossível, e é o recurso aos pontos não observados, b^* e c^* , que permite essa definição.

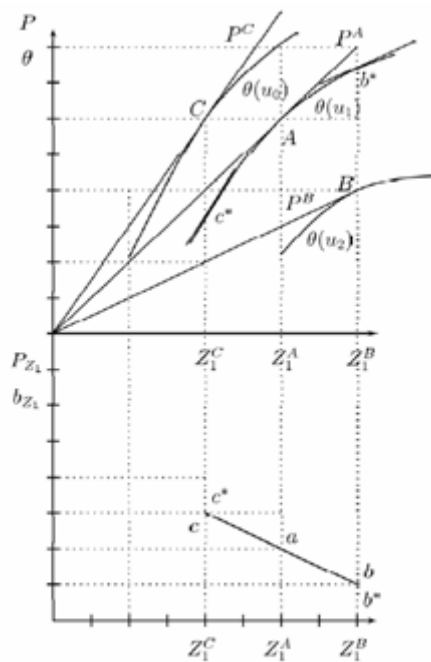


Gráfico 2: Preferências quase-lineares

Como já foi referido, caso se assumisse preferências quase-lineares, nesse caso havia coincidência entre os pontos observados e os não observados, conforme se pode visualizar no gráfico 2, podendo identificar-se a inversa da procura marshalliana da habitação com a curva preço/licitação marginal.

As preferências quase-lineares são ilustradas por curvas preço/licitação que são translações verticais umas das outras, fazendo com que o efeito líquido de um aumento de rendimento, traduzido na deslocação para outra curva, reflectindo um nível de utilidade mais elevado,

seja nulo em termos da quantidade procurada. Este tipo de preferências representam um caso especial em que a elasticidade procura/rendimento por atributo/característica é nula, embora a variação no rendimento influencie directamente a variação na função preço/licitação, já que aumentos no rendimento fazem com que as curvas preço/licitação se desloquem verticalmente e para níveis de utilidade superiores.

No gráfico 2 esta situação traduz-se por declives à curva $\theta(z_1, u_1)$, respectivamente nos pontos b^* e c^* , coincidentes com $p_{z_1}^B$ e $p_{z_1}^C$ e conseqüentemente com $b_{z_1}(z_1, u_2)$ e com, $b_{z_1}(z_1, u_0)$.

Acontece que no mundo real as preferências quase-lineares são a exceção e não a regra,⁹⁸ no entanto este pressuposto, ainda que teórico, permite estabelecer uma relação, entre a elasticidade rendimento e a forma dessas curvas – assim, quanto maior a elasticidade rendimento da procura por um atributo/característica maior a diferença entre a procura ordinária e a curva preço/licitação marginal.

⁹⁸ Dado que só têm acesso à aquisição de habitação famílias com um certo nível de rendimento, poder-se-á assumir que, relativamente a determinadas características/atributos, como seja o número de quartos, o nº de casas de banho, as elasticidades procura/rendimento por esses atributos tendam a ser rígidas, permitindo uma aproximação real ao pressuposto das preferências quase-lineares.

4. Modelo Empírico e Avaliação Econométrica

*The Mathematician is perfect only in so far as he is a perfect being,
In so far as he perceives the beauty of truth, only then his work be thorough,
transparent, comprehensive, pure, clear, attractive and even elegant.*

Johann Wolfgang von Goethe (1810)

4.1. Metodologia e Base de Dados

A estimação econométrica do modelo empírico especificado no capítulo 3 constitui o núcleo central deste capítulo. No entanto esse modelo suporte da estimação que o capítulo se propõe apresentar, como já foi descrito, fundamenta-se na teoria de preços hedónica, o que impõe, a jeito de ressalva, uma referência histórica à sua conceptualização. Assim, num primeiro momento, a teoria de preços hedónica surge como uma simples técnica hedónica,⁹⁹ omissa em fundamentações epistemológicas da teoria económica. Limitações científicas superadas, apenas, num segundo momento, por Lancaster (1966), o qual reorienta a teoria tradicional do consumidor com a introdução de três importantes pressupostos.¹⁰⁰ As técnicas hedónicas encontram aí suporte axiomático para fundamentar o edifício da teoria de preços hedónica.

⁹⁹ Reportada a Andrew Court (1939) que, ao construir um índice de preços para os automóveis, constatou que o automóvel servia vários propósitos e que o seu preço não devia ser apurado da mesma forma que um qualquer bem servindo apenas um propósito, mas sim pela combinação das diversas especificações desse automóvel. E, posteriormente, a Griliches (1971) que, ao aplicar técnicas hedónicas para derivar índices de preços para fertilizantes e automóveis, despertou a curiosidade da literatura sobre “diferenças qualitativas” na mensuração de salários, de amenidades ambientais, de avaliação do património, entre outras.

¹⁰⁰ A saber, “o bem, *per se*, não confere utilidade ao consumidor, mas sim as características que lhe estão intrínsecas”; “um bem possui mais do que uma característica, e muitas características são partilhadas por mais do que um bem” e “uma combinação de bens pode possuir características diferentes daquelas que caracterizam o bem isolado”

Já no domínio de uma teoria de preços hedónicos, o facto de esses preços virem definidos como preços implícitos dos atributos/características e de se obterem a partir dos preços observados dos bens diferenciados e da quantidade das características que lhe estão intrínsecas,¹⁰¹ fazia persistir um problema de identificação. Problema, a que Rosen (1974) tentou dar resposta, descortinando para o efeito mecanismos de comportamento, quer dos consumidores, quer dos produtores, relativamente às características/atributos de um bem composto, deduzindo, daí, funções de licitação (*bid*) para os consumidores e funções de oferta (*offer*) para os produtores, que, em equilíbrio, geravam a função de preços hedónica. É, no entanto, o próprio Rosen, quem admite que esses preços revelam pouco sobre as funções de procura e de oferta que lhe estão subjacentes, o que não permite estabelecer a forma funcional do modelo de preços hedónico, que traduz a dependência do preço de um bem composto em relação ao conjunto das suas características.

Como, em última instância, os preços hedónicos e os preços marginais das características/atributos vão depender da forma funcional dos preços hedónicos, a literatura sobre preços hedónicos¹⁰² tem, de forma recorrente, abordado as questões relativas à forma funcional dos modelos hedónicos e discutido a forma funcional que melhor estima os preços hedónicos. Contemporizando esta restrição, optou-se, para a concretização da estimação econométrica, por equacionar um leque de possibilidades para a forma funcional e, posteriormente, escolher a forma apropriada da função de preços através de um critério de qualidade do ajustamento (*goodness of fit*),¹⁰³ fundamentando a escolha em resultados econométricos.

A comparação dos modelos alternativos fez-se, não só, sustentada nos dados da amostra, com recurso às medidas estatísticas de avaliação do ajustamento do modelo à

¹⁰¹ O que se pode observar é o preço da mercadoria composta e não o preço dos atributos, não há um mercado dos atributos, embora os coeficientes da função de preços hedónica estimada reflectam, quer os custos dos produtores, quer as preferências dos consumidores.

¹⁰² Witte, A *et al.* (1979); Halvorsen R. e H. Pollakowski (1981); Cassel E. e Mendelsohn (1985); Cropper *et al.* (1988), por exemplo.

¹⁰³ Concretamente, Cropper, *et al.* (1988), para seleccionarem a forma da função de preços hedónica que, com maior rigor, permitisse estimar os preços marginais dos atributos/características, apuraram a variação dos erros de mensuração dos preços marginais em função da forma funcional.

amostra,¹⁰⁴ mas, também, sustentada nos valores da previsão, com recurso a medidas de rigor previsionais, como sejam por exemplo, o erro absoluto médio (*MAE*)¹⁰⁵, a raiz quadrado do erro quadrado médio (*RMSE*)¹⁰⁶ e o coeficiente de desigualdade *U* de *Theil*,¹⁰⁷ que procuram equacionar todos os desvios entre os valores previstos e os valores observados, de forma a reflectirem a qualidade da previsão. Se estes valores diferirem significativamente entre si, então as previsões diferem radicalmente da sua realização.¹⁰⁸

Na presença das medidas de rigor referidas, haveria que assegurar a eficiência relativa dos modelos de previsão. Para tanto recorreu-se a uma ferramenta padrão em estatística, conhecida como validação cruzada (*cross validation*),¹⁰⁹ a qual forneceria procedimentos equacionáveis para a análise, permitindo validar os modelos para um conjunto de dados, diferentes dos utilizados para estimar os parâmetros. A aplicação desta ferramenta implicaria dividir a amostra em duas sub-amostras:¹¹⁰ uma que serviria para determinar os coeficientes da regressão, denominada de amostra de estimação ou

¹⁰⁴ Como é o caso das estatísticas, R^2 , R^2 Ajustado e critério Akaike, que poderiam revelar grande poder explicativo dos regressores escolhidos, mas que poderiam não ser robustos a pequenas variações nos dados. Além de que, a utilização simultânea da amostra, para estimar e medir a qualidade de ajustamento (*goodness of fit*), conduziria a avaliações optimistas (naturalmente, que a regressão fornece melhores resultados para os dados da amostra sobre os quais se baseou, do que para quaisquer outros).

¹⁰⁵ $MAE = \frac{1}{N} \sum_{n=1}^N |\hat{P}_n - P_n|$, em que \hat{P}_n é o preço previsto, P_n é o preço observado e N é o número de dados previstos.

¹⁰⁶ $RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{P}_n - P_n)^2}$, em que \hat{P}_n é o preço previsto, P_n é o preço observado e N é o número de dados previstos.

¹⁰⁷ $U_{Theil} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{P}_n - P_n)^2} / \sqrt{\frac{1}{N} \sum_{n=1}^N P_n^2}$, em que \hat{P}_n é o preço previsto, P_n é o preço observado e N é o número de dados previstos.

¹⁰⁸ O IMAD, *Institut for Macroeconomic Analysis and Development*, utiliza este critério para avaliar o rigor das previsões macroeconómicas, conforme avaliação apresentada em, *Autumn Report, 2002 – “A comparative Analysis of the Quality of Forecasts”*.

¹⁰⁹ O conceito tem sido muitas vezes referido como o da amostra predictiva reutilizável, *predictive sample reuse*, Stone (1974, 1977). Importa, ainda, referir que o *output* da *cross validation* se situa no domínio do *erro de previsão*, quer como erro absoluto médio (*MAE*), quer como raiz do erro quadrado médio (*RMSE*).

¹¹⁰ Resulta do facto da validação cruzada exigir duas amostras e da dificuldade em ter, prontamente, disponíveis duas amostras. É evidente a perda de informação decorrente da divisão da amostra, sobretudo se atentarmos na dificuldade de se conseguir amostras grandes. Obviamente que temos que assumir que, dado a dimensão das amostras de calibração, os coeficientes de estimação terão que ser menos precisos do que aconteceria caso se trabalhasse com a amostra inteira. Note-se, no entanto, que a validação cruzada não tem que estar limitada a duas amostras, já que os dados se podem dividir em várias sub-amostras e, ainda, se podem recombinar sucessivamente.

de calibração do modelo; e outra, a amostra excluída (*out-of-sample*), que permitiria testar o modelo, isto é, aplicar os coeficientes obtidos a partir da amostra de estimação para avaliar os valores previstos (*predictor scores*). Pelas suas características estaríamos perante a amostra de validação ou teste.

A validação é um aspecto crítico de qualquer modelo econométrico e, embora não haja uma metodologia teórica ou, suficientemente bem formulada para a validação de modelos, a prática usual é basear a validação dos modelos em medidas específicas de desempenho expressas em função de dados não utilizados na construção do modelo.

Neste contexto, optou-se por comparar as previsões (*ex-post*) dos preços de venda da habitação utilizando três medidas de avaliação estatística, *MAE*, *RMSE* e *U* de *Theil*, de forma a assegurar que a inferência em relação à eficiência relativa dos modelos de previsão não pudesse ser afectada por, um qualquer, critério particular presente nessas comparações.¹¹¹ Para além destas três medidas individuais de previsão estatística *ex-post*, estabelecendo paralelismo com trabalho de Kick (2002), também se calculou a média dessas medidas, denominadas $MAE_{Média}$, $RMSE_{Média}$ e $U_{Média}$, tendo para o efeito pressuposto uma ordenação aleatória das variáveis independentes e dependente para um universo de 1000 regressões.¹¹²

Dirimidas as limitações decorrentes da teoria dos preços hedónicos, importa um reporte à base de dados necessária ao processo de estimação. Neste particular, a escolha da cidade da Guarda para a recolha de informação, que permitisse identificar uma amostra das vendas de habitação nessa cidade, apresentou-se óbvia, já o tipo de informação pretendida (valor de venda) deixava adivinhar alguma resistência à sua obtenção, a crer no desvio entre os valores da transacção e os valores constantes das escrituras de registo de propriedade, aferidos à data da recolha de dados (2003), o que se veio a confirmar.

¹¹¹ Mesmo correndo o risco de contrariar o princípio da parcimónia de Occam's Razor, que nos conduziria para a escolha de, apenas, uma medida para avaliar a qualidade (*goodness*) do desempenho de um qualquer modelo. Não havendo, no entanto, consenso na bibliografia consultada, relativamente à medida a utilizar, arrisca-se a complitude de medidas.

¹¹² A escolha de 1000 regressões tem como objectivo a optimização do binómio correcção da discricionariedade na entrada de dados/capacidade computacional. Este critério de avaliação das previsões, no entanto não pode determinar se um dado modelo de previsão é significativamente melhor que outros. Para avaliar a significância estatística dos modelos de previsão rivais ter-se-ia de utilizar testes de abrangência [Donaldson e Kamstra (1996, 1997)].

Ultrapassados os entraves, com potencial de maior bloqueio à recolha de informação, conseguiu-se assegurar uma amostra das vendas de habitação realizadas no ano 2002, apoiada na colaboração de empresas imobiliárias a operar na cidade da Guarda. A informação foi obtida a partir da ficha de identificação dos imóveis transaccionados, constante do apêndice C, e cruzada com os dados das escrituras realizadas em sede do cartório notarial da Guarda, o que permitiu construir uma base de dados (400 observações) com todas as informações relativas, quer ao preço de venda, quer aos atributos/características da habitação.

A ficha de identificação dos imóveis transaccionados foi gizada tendo subjacentes dois objectivos, um de atenuar a desconfiança suscitada pelo mesmo ao nível das empresas imobiliárias, o outro de conseguir recolher informação passível de quantificar as características dos imóveis que se pretendiam observadas. A concretização desse propósito garantia simultaneamente informação e fidedignidade, permitindo desenvolver as variáveis, independentes e dependente, requeridas pela especificação paramétrica convencional. Variáveis essas que vêm explicadas no apêndice D, dele constando 28 variáveis independentes: 24 variáveis dicotómicas (*dummy*) e 4 variáveis discretas/contínuas. Tal como se refere este requisito para a especificação paramétrica convencional materializa-se com a informação da base de dados, porém essa especificação só fica concretizada com a apresentação da forma funcional que define a relação determinística entre a variável dependente e as diversas variáveis independentes.

Relativamente à forma funcional, de acordo com a metodologia perseguida, houve que equacionar um leque de possibilidades para a função de preços hedónica, conduzindo a que se escolhesse as formas: linear, semi-logarítmica (log-linear), duplo-logarítmica (log-log), quadrática e as formas transformadas Box-Cox, linear e quadrática. E é sobre essa família de modelos lineares e linearizáveis que se vão aplicar técnicas de regressão¹¹³ e se vai procurar o erro de previsão mais baixo, aplicando a ferramenta

¹¹³ É crucial para qualquer regressão utilizando o método dos mínimos quadrados a não violação dos pressupostos subjacentes ao modelo clássico de regressão linear, a saber, (1) linearidade das variáveis; (2) ausência de correlação linear perfeita entre variáveis; (3) Valor esperado da perturbação aleatória zero, $E(\varepsilon_i) = 0 \forall i$; (4) Homoscedasticidade, variância das perturbações aleatórias constante, $\text{Var}(\varepsilon_i) = \sigma^2 \forall i$ e não autocorrelação, $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \forall i, j, i \neq j$; (5) regressores não estocásticos e (6) normalidade.

validação cruzada e medindo os valores médios do $MAE_{Média}$, do $RMSE_{Média}$ e do U de $Theil_{Média}$, em 1000 regressões.

A técnica de validação cruzada permite a utilização de todos os dados da amostra, mas para se poder aplicar essa técnica havia que dividir a amostra dos dados observados, 400, em duas sub-amostras.¹¹⁴ Uma, de 350 observações, sobre a qual se estimam os parâmetros do modelo e a outra, de 50 observações, sobre a qual se calculam as estatísticas de validação, ou seja as medidas de erro ou rigor da previsão. Optou-se por gerar séries aleatórias de forma a se poderem apurar valores médios das medidas de erro, $MAE_{Média}$, $RMSE_{Média}$ e U de $Theil_{Média}$, em 1000 regressões, utilizando o software EVIEWS 4, conforme programas constantes do apêndice E.

4.2. Formas Funcionais dos Modelos de Preços Hedónicos

4.2.1. Forma funcional linear

O modelo linear de preços hedónicos pode ser expresso na forma,

$$\mathbf{P} = \alpha + \sum_{k=1}^4 \beta_k \mathbf{z}_k + \sum_{l=1}^{24} \gamma_l \mathbf{D}_l + \boldsymbol{\varepsilon}$$

Onde, \mathbf{P} é um vector preços de habitação $[400 \times 1]$;

\mathbf{z}_k é um vector por cada k características/habitação, variável contínua, de dimensão $[400 \times 1]$, existem 4 vectores;

¹¹⁴ De acordo com Santos e Azevedo (2005) na validação cruzada K -desdobrável, as amostras são divididas em K subconjuntos mutuamente exclusivo de dimensões aproximadamente iguais (K -folds). Os modelos são ajustados e validados K -vezes e o erro final da generalização obtém-se pela média dos erros de validação obtidos durante K -vezes. Os valores de K tendem para o número mágico de 10 (*Ten-Fold Cross Validation*), no caso em apreço, 10 subconjuntos de 40, o que teve de ser ajustado para 8 de 50, por questões técnicas da regressão.

\mathbf{D}_l é um vector por cada l características/habitação, variável *dummy*, de dimensão $[400 \times 1]$, existem 24 vectores;

$\boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$ é um vector de resíduos, de dimensão $[400 \times 1]$

α e β_k e γ_l , são o termo constante, os k coeficientes e os l coeficientes, respectivamente.

O preço da habitação é estimado em euros e por essa via não decorre inferência de qualquer enviesamento. A desvantagem deste modelo reside na impossibilidade de interacção entre variáveis independentes, fazendo com que o valor adicional de um qualquer atributo/característica seja o mesmo, independentemente do valor de um outro atributo/característica. Os coeficientes β_k , correspondem a $\frac{\partial \mathbf{P}}{\partial z_k}$, podendo ser interpretados como os preços marginais dos k atributos/características de uma habitação.

4.2.2. Forma funcional logarítmica

A forma logarítmica já permite a interacção entre as variáveis independentes e, adicionalmente, reduz (ou elimina) eventuais problemas de heterocedasticidade. A estimação do logaritmo de preços é não enviesada, mas a transformação do logaritmo dos preços em preços, como já foi referido, induzirá enviesamento. No entanto, o elevado poder preditivo associado a estas formas funcionais compensa esse enviesamento induzido [Goldberger (1968)]. Consideraram-se duas formas logarítmicas, as mais convencionais e mais convenientes, assim:

- **Modelo de Preços Hedónicos semi-logarítmico**, que pode ser escrito como,

$$\mathbf{P} = \exp\left(\alpha + \sum_{k=1}^4 \beta_k \mathbf{z}_k + \sum_{l=1}^{24} \gamma_l \mathbf{D}_l + \boldsymbol{\varepsilon}\right)$$

ou

$$\ln \mathbf{P} = \alpha + \sum_{k=1}^4 \beta_k \mathbf{z}_k + \sum_{l=1}^{24} \gamma_l \mathbf{D}_l + \boldsymbol{\varepsilon}$$

Na especificação semi-logarítmica apenas a variável dependente é transformada, mas a regressão continua linear nos parâmetros, mantendo-se para $\ln \mathbf{P}$ o melhor estimador linear não enviesado, *Best Linear Unbiased Estimator*, BLUE.

Os coeficientes β_k são agora expressos por $\frac{\partial \ln \mathbf{P}}{\partial z_k}$, passando a medir a variação percentual nos preços de habitação relativamente a pequenas variações numa qualquer característica da habitação, passíveis de se assumirem a uma semi-elasticidade do preço relativamente à característica k . E neste modelo a variável idade, denominada “ANOS”, pode ser interpretada como uma taxa de depreciação da habitação.

- **Modelo de Preços Hedónicos duplo logarítmico**, que se pode expressar por,

$$\mathbf{P} = \exp(\alpha) + \prod_{k=1}^4 \mathbf{z}_k^{\beta_k} + \prod_{l=1}^{24} \mathbf{D}_l^{\gamma_l} + \exp(\boldsymbol{\varepsilon})$$

ou

$$\ln \mathbf{P} = \alpha + \sum_{k=1}^4 \beta_k \ln(\mathbf{z}_k) + \sum_{l=1}^{24} \gamma_l \ln(\mathbf{D}_l) + \boldsymbol{\varepsilon}$$

Esta especificação exige algumas modificações ao nível das variáveis independentes dicotómicas (*dummy*), que se traduziria em considerar o dígito “dois” na presença de um atributo/característica e o dígito “um” na ausência dessa característica.¹¹⁵

Passando os coeficientes β_k , a medir elasticidades, já que $\beta_k = \frac{\partial \ln \mathbf{P}}{\partial \ln z_k}$.

4.2.3. Forma funcional quadrática

O modelo quadrático dos preços hedónicos pode expressar-se por,

$$\mathbf{P} = \alpha + \sum_{k=1}^4 \beta_k \mathbf{z}_k + \frac{1}{2} \sum_{k=1}^4 \sum_{k'=1}^4 \beta_{kk'} \mathbf{z}_k \mathbf{z}_{k'} + \sum_{l=1}^{24} \gamma_l \mathbf{D}_l + \varepsilon$$

A grande vantagem é permitir a interacção das variáveis independentes, ilustrada no 3º

termo do 2º membro da forma quadrática, $\frac{1}{2} \sum_{k=1}^4 \sum_{k'=1}^4 \beta_{kk'} \mathbf{z}_k \mathbf{z}_{k'}$, tornando possível a

mensuração da variação dos preços em relação a um dado atributo/característica, em função de outras características/atributos.

Esta forma funcional incorpora a não linearidade, teoricamente, esperada sob os pressupostos de, (1) uma utilidade marginal decrescente associada ao consumo de um bem heterógeno por via da variação de uma qualquer característica/atributo desse bem (determinada, por sua vez, pela produtividade marginal decrescente associada à tecnologia disponível) e de (2) um preço do bem heterógeno determinado, simultaneamente, a partir de todas as característica/atributos (o que permite assumir que

¹¹⁵ O EVIEWS 4 permite esse ajustamento, simplesmente, acrescentando “um” a todas as variáveis *dummy* presentes na regressão.

a relação entre o preço implícito e a quantidade de uma qualquer característica é positiva). Em contexto de modelos hedónicos de elevada dimensão, com atributos/características múltiplos, estes pressupostos da teoria económica exigem a resolução de equações diferenciais complexas para a determinação do equilíbrio de mercado, o que pode ser superada por via desta especificação.

4.2.4. Forma transformada de Box-Cox

O modelo geral transformado de preços hedónicos, Box-Cox, é passível da seguinte expressão,

$$\mathbf{P}(\Theta) = \alpha + \sum_{k=1}^3 \beta_k \mathbf{z}_k(\lambda_1) + \beta_4 \mathbf{z}_4(\lambda_2) + \frac{1}{2} \sum_{k=1}^3 \sum_{k'=1}^3 \beta_{kk'} \mathbf{z}_k(\lambda_1) \mathbf{z}_{k'}(\lambda_1) + \frac{1}{2} \sum_{k=1}^3 \delta_{k4} \mathbf{z}_k(\lambda_1) \mathbf{z}_4(\lambda_2) + \sum_{l=1}^{24} \gamma_l \mathbf{D}_l + \varepsilon$$

$$\Psi(\wp) = \begin{cases} \frac{\Psi^\wp - 1}{\wp}, & \wp(\Theta \vee \lambda_1 \vee \lambda_2) > 0 \\ \ln(\Psi), & \wp(\Theta \vee \lambda_1 \vee \lambda_2) = 0 \end{cases} \quad \mathbf{z}_{k,4}(\lambda_1, \lambda_2) = \begin{cases} \frac{\mathbf{z}_{k,4}^{\lambda_1, \lambda_2} - 1}{\lambda_1, \lambda_2}, & \lambda_1, \lambda_2 > 0 \\ \ln(\mathbf{z}_{k,4}), & \lambda_1, \lambda_2 = 0 \end{cases}$$

$$\Psi = \mathbf{P} \vee \mathbf{z}_k \vee \mathbf{z}_4$$

Em que Θ , λ_1 e λ_2 , são os parâmetros de transformação Box-Cox.

A transformação Box-Cox é uma das formas funcionais mais flexíveis, o que permite ultrapassar todas as restrições teóricas que se possam colocar. A amplitude da sua flexibilidade é tal que, mesmo utilizando, apenas, um parâmetro de transformação para a variável dependente e para todas as variáveis independentes (à excepção das variáveis *dummy*, não transformadas), ainda assim, obter-se-iam 40 formas funcionais, incluindo a linear e a semi-logarítmica. Neste caso concreto optou-se por utilizar 3 parâmetros de transformação diferentes,¹¹⁶ um para a variável dependente (Θ), outro para todas as variáveis independentes (λ_1), não dicotómicas e, ainda, um outro para a variável independente “ANOS” (λ_2).

Consideraram-se duas situações. Na primeira, $\beta_k \neq 0$ e $\beta_{kk'} \neq 0$ enfrenta-se a especificação de preços hedónicos Box-Cox transformada quadrática, com interacção das variáveis independentes. Na segunda, $\beta_k \neq 0$ mas $\beta_{kk'} = 0$, a transformação Box-Cox reveste a forma linear.

Para cada uma destas formas funcionais, Box-Cox, linear e quadrática, maximiza-se uma função logarítmica de verosimilhança, que tem subjacente a transformação das variáveis após escolha prévia de valores para os parâmetros de transformação, Θ , λ_1 e λ_2 .¹¹⁷ Assim para o modelo Box-Cox linear os parâmetros de

¹¹⁶ Como as transformações são potencialmente não lineares, estes parâmetros de transformação não podem ser estimados directamente. Para serem obtidos, ter-se-á que maximizar a função concentrada de verosimilhança,

$$L_{\max} = (\Theta - 1) \sum_{n=1}^{400} \ln \mathbf{P}_n - \frac{N}{2} \ln(SSR_{\Theta, \lambda_1, \lambda_2} / N). \text{ Para a maximização dessa função pode-se recorrer a uma técnica a dois momentos.}$$

Primeiro, no momento 1, escolhem-se valores para, Θ , λ_1 , λ_2 e transformam-se os dados observados. Posteriormente, no

momento 2, avalia-se a função logarítmica de verosimilhança, $L(\Theta, \lambda_1, \lambda_2) = (\Theta - 1) \sum_{n=1}^{400} \ln \mathbf{P}_n - \frac{N}{2} \ln(SSR_{\beta, \Theta, \lambda_1, \lambda_2} / N)$, em que N é

a dimensão da amostra, $SSR/N = \hat{\sigma}_{(\Theta, \lambda_1, \lambda_2)}^2$, a soma do quadrado dos resíduos é a variância estimada dos valores transformados e

\mathbf{P}_n é o preço observado na observação n . Depois de se obter o valor máximo dessa função, aplicam-se, então, testes de hipóteses relativamente às formas funcionais mais restritivas.

¹¹⁷ O que em termos do programa EVIEWS 4, adaptado para o efeito, significou escolher os valores desses parâmetros a partir de um intervalo de números, entre $[-1,95; 1,95]$, em incrementos de 0,1, e para cada terno de valores assumidos por esses parâmetros estima-se, pelo ajustamento dos mínimos quadrados à função hedónica de preços Box-Cox, simultaneamente, os coeficientes dos regressores e a variância, $\sigma^2(\beta, \Theta, \lambda_1, \lambda_2)$. Os valores estimados da variância serão inseridos na função logarítmica de

transformação obtidos foram, $\Theta^{\max} = 0,05$, $\lambda_1^{\max} = 0,35$, $\lambda_2^{\max} = 1,05$, e para o modelo Box-Cox quadrático foram, $\Theta^{\max} = 1,85$, $\lambda_1^{\max} = 0,15$, $\lambda_2^{\max} = 0,75$, e foi com eles que se obtiveram as regressões, pelo método dos mínimos quadrados, para os dois modelos.¹¹⁸

4.3. Comparação das Medidas de Erro (ex post) de Previsão

Servindo-nos da ferramenta estatística validação cruzada, *cross-validation*, como já se referiu, e utilizando três medidas estatísticas de fiabilidade vai-se procurar avaliar o desempenho de cada modelo na previsão dos preços de habitação e escolher o que se revelar mais poderoso.

As medidas utilizadas, os valores médios das medidas de erro, $MAE_{Média}$, $RMSE_{Média}$ e $U_{Média}$ ¹¹⁹ podem resumir-se no quadro que se segue,

	Modelo Linear	Modelo Log-linear	Modelo Log-log	Modelo Quadrático	Modelo BoxCox linear	Modelo BoxCox Quadrático
$MAE_{Média}$	12552,2434	10793,2321	12323,0394	11294,0272	10983,1897	11110,8668
$RMSE_{Média}$	23805,7722	21011,7570	22618,6098	21218,1440	21632,9389	21532,3870
$U_{Média}$	0,03356785	0,03175237	0,03294268	0,03154068	0,03153867	0,03168446

Tabela 4.3-1: Medidas de erro médio por especificação

verosimilhança até se determinarem os valores que conduzem à maximização desta função. Finalmente, com esses valores estima-se o modelo hedónico Box-Cox.

¹¹⁸ A transformação subjacente aos modelos Box-Cox induz enviesamento e para algumas transformações Box-Cox a variável dependente pode mesmo ter distribuição não normal.

¹¹⁹ Estas medidas são as médias, do erro absoluto médio, da raiz quadrada do erro médio ao quadrado e do coeficiente de desigualdade de Theil, que foram obtidas a partir de 1000 regressões, resultantes da geração aleatória de n sequências de dados. Esta salvaguarda pretende corrigir eventuais enviesamentos na estimação dos parâmetros da regressão, decorrentes da ordem de introdução de dados.

De acordo com a metodologia seguida para a escolha da forma funcional de preços hedónica, a especificação de preços hedónica semi-logaritmica é a que apresenta melhores resultados, ainda que dentro da mesma escala de grandeza.

Pode-se ainda complementar esta metodologia, fundamentada na validação cruzada e sustentada em medidas de erro *ex post*, com o critério de avaliação do desempenho (*performance*) das previsões macroeconómicas seguido pelo IMAD,¹²⁰ que consiste em avaliar a proficiência das previsões, em função da proximidade dos valores das medidas de erro *MAE* e *RMSE*. Assim, observemo-las para os diversos modelos,

	Modelo Linear	Modelo Log-linear	Modelo Log-log	Modelo Quadrático	Modelo BoxCox Linear	Modelo BoxCox Quadrático
<i>MAE</i>	13464,7448	8452,5122	11167,0478	11152,3614	8736,8407	11525,5492
<i>RMSE</i>	16401,1911	10325,5086	15150,7163	14247,4822	12168,684	18455,6497

Tabela 4.3-2: Medidas de erro por especificação

Os melhores registos para essa proximidade observam-se no modelo log-linear, o qual, segundo o critério do IMAD, consistentemente, gera as melhores previsões fora da amostra, *out of sample*, de acordo com a evidência resultante da proximidade de valores do *MAE* e do *RMSE*.

4.4. O Modelo de Preços Hedónico Semi-logaritmico

Uma vez cumprido o objectivo de escolha de uma forma funcional para o modelo de preços hedónico, através da comparação entre os *benchmark* dos erros de previsão obtidos a partir de diferentes especificações do modelo, passar-se-á a apresentar o modelo de preços semi-logaritmico estimado,

¹²⁰ IMAD, ver nota 108.

$$\begin{aligned}
\ln VMERCADO = & 10,733 - 0,014 * ANOS + 0,0029 * AQCENTRAL + \\
& (86,743) \quad (-5,654) \quad (0,123) \\
& + 0,0035 * AREA + 0,068 * QUARTOS + 0,0361 * CBANHO + \\
& (3,859) \quad (2,467) \quad (0,923) \\
& + 0,0200 * ARRECA - 0,0167 * ELEV + 0,004 * EXPSOL + \\
& (0,608) \quad (-0,916) \quad (0,058) \\
& + 0,0618 * GCANAL + 0,1246 * GSDOS + 0,0623 * GSUM - \\
& (2,398) \quad (2,515) \quad (2,531) \\
& - 0,0093 * JANDUPLAS + 0,0646 * LAR - 0,0578 * LCNTRRO + \\
& (-0,436) \quad (1,841) \quad (-1,048) \\
& + 0,0317 * LCVLHOS - 0,1091 * LGGARE - 0,0138 * LLAMNH + \\
& (0,468) \quad (-2,087) \quad (-0,220) \\
& + 0,0020 * LLUZ - 0,0827 * LPMILEU - 0,0777 * LPNHR + \\
& (0,038) \quad (-1,474) \quad (-1,334) \\
& + 0,0113 * LRDIZ - 0,0233 * LSREMD - 0,0407 * CV - \\
& (0,197) \quad (-0,426) \quad (-0,611) \\
& - 0,0281 * RCH + 0,0329 * PPRIM - 0,0399 * PQRT - \\
& (-0,753) \quad (1,206) \quad (-0,734) \\
& - 0,0455 * PSEG + 0,0233 * PTERC \\
& (-1,136) \quad (0,726)
\end{aligned}$$

$$R^2 = 0,62582$$

$$R^2_{\text{Ajustado}} = 0,59758$$

Os números entre parênteses reportam-se aos valores do *t*-estatístico. Os quais permitem inferir uma elevada probabilidade de os coeficientes estimados para as variáveis, *AQCENTRAL*, *CBANHO*, *ARRECA*, *ELEV*, *EXPSOL*, *JANDUPLAS*, *LCNTRRO*, *LCVLHOS*, *LLAMNH*, *LLUZ*, *LPMILEU*, *LPNHR*, *LRDIZ*, *LSREMD*,

CV, *RCH*, *PPRIM*, *PQRT*, *PSEG*, *PTERC*, virem nulos, dado o baixo nível de significância estatística dos mesmos. Explicação para esta ocorrência pode ser ensaiada na frequência de observações com a característica que essas variáveis identificam no total das 400 observações. A título de exemplo, o caso das variáveis que identificam a localização espacial, *LCNTRO*, *LCVLHOS*, *LLAMNH*, *LLUZ*, *LPMILEU*, *LPNHR*, *LRDIZ*, *LSREMD*, onde se verifica uma média muito próxima de zero (o que em contexto de variáveis *dummy* significa baixa frequência dessa característica na amostra observada).¹²¹ O mesmo paralelismo se pode estabelecer para as variáveis que identificam o posicionamento na estrutura do imóvel, *CV*, *RCH*, *PPRIM*, *PQRT*, *PSEG*, *PTERC*. Já para a variável *CBANHO* a potencial correlação com a variável *QUARTOS* pode consentir uma explicação.

A leitura dos *t-estatísticos* permite elencar, ainda, como estatisticamente significantes os parâmetros estimados para as variáveis, *ANOS(-)*, *AQCENTRAL(+)*, *AREA(+)*, *QUARTOS(+)*, *GCANAL(+)*, *GSDOS(+)* e *GSUM(+)*. Os sinais entre parênteses colocados à frente da notação das variáveis reflectem os sinais esperados para os respectivos coeficientes estimados. Ora na presente estimação, pelo método dos *MQO*, aplicada ao modelo semi-logarítmico, não só se aceita como diferentes de zero, com probabilidade de quase 100%, os coeficientes estimados dessas variáveis, como se infere, ainda, que apresentam sinal de acordo com o esperado.

4.5. Análise da Consistência Estatística do Modelo Escolhido

4.5.1. A especificação do modelo

A especificação do modelo linear centra-se no vector das perturbações, ε , e na matriz de variáveis independentes, \mathbf{x} . Neste contexto podem surgir problemas com o vector

¹²¹ Por exemplo, *LCVLHOS* e *LRDIZ*, apresentam uma média de 0,03 e de 0,01, respectivamente e por via do desvio padrão um intervalo de dispersão onde apenas caem as observações neutras para o cálculo da média, 388 e 396, respectivamente.

das perturbações e/ou com as variáveis independentes e/ou com o vector dos coeficientes, β . Assim, relativamente,

- ao vector das perturbações, pode-se verificar que, não seguem uma distribuição i.d.d. $(0, \sigma^2)$, $i = 1, \dots, n$; ou que violam o pressuposto da homoscedasticidade, $E(\epsilon' \epsilon) = \text{diag}[\sigma_1^2 \dots \sigma_n^2]$; ou, ainda, que estão correlacionados dois a dois, $E(\epsilon_t \epsilon_{t-s}) \neq 0$, $s \neq 0$;
- à matriz das variáveis independentes pode-se assistir à omissão de variáveis relevantes; ou à inclusão de variáveis irrelevantes; ou a forma funcional incorrecta; ou a uma característica da matriz \mathbf{x} inferior ao número de colunas (colinearidade); ou a correlação não nula entre regressores e termo de perturbação (violação de $E(\mathbf{x}_{it} \epsilon_s) \neq 0$, para todo o $i = 1, \dots, k$ e $t, s = 1, \dots, n$), ou, ainda, a variáveis não estacionárias.
- ao vector dos coeficientes β pode-se não verificar a invariabilidade dos parâmetros, contrariando a estabilidade do vector dos coeficientes dentro e fora da amostra.

Começar-se-á por testar a validade dos pressupostos fundamentais do método dos mínimos quadrados para a amostra considerada, isto é testar-se-á a especificação do modelo. Mais concretamente, ir-se-á testar a especificação linear do modelo, hipótese nula, contra a especificação não linear, por via do teste RESET de Ramsey.¹²² Este teste produz como resultados, uma regressão alargada às potências da variável dependente estimada, um F estatístico e um rácio de verosimilhança LR .¹²³ A observação desses

¹²² Se o modelo estimado é $\hat{P}_i = \hat{\alpha} + \sum_{i=2}^{400} \hat{\beta}_i z_{ki}$ então o teste de RESET, testa a hipótese, $\gamma = 0$, sobre a regressão aumentada

$\hat{P}_i = \hat{\alpha} + \sum_{i=2}^{400} \hat{\beta}_i z_{ki} + \gamma \mathbf{P}$, onde $\mathbf{P} = [\hat{P}^2, \hat{P}^3, \hat{P}^4, \dots]$ é o vector das potências dos valores preditos da variável dependente (podendo compreender potências de valores preditos e produtos cruzados de variáveis independentes). Assim sendo a hipótese testada é a da linearidade.

¹²³ O teste F , obtém-se a partir de $F_{[M, N-K-1]} = \frac{(SSR_R - SSR_U) / M}{(SSR_U) / (N - K)}$, em que SSR é a soma dos quadrados dos resíduos, e os índices R e U , identificam, respectivamente a regressão com restrições e a regressão sem restrições. M indica o número de restrições – número de termos ajustados a incluir no teste de regressão, N , o número de observações e K , o número de parâmetros estimados na regressão.

resultados permite, caso o F estatístico seja superior ao F crítico, constante das tabelas da distribuição F , para níveis de significância de 95% e 99%, rejeitar-se a hipótese nula. Se esse não for o caso, então a verdadeira especificação do modelo é linear e conclui-se que passa o teste de RESET Ramsey.

No caso concreto do modelo semi-logarítmico escolhido para a função de preços hedónicos do mercado da Guarda, tem-se um $F_{[2,369]} = 2,7187$, pela leitura das tabelas da distribuição F , constata-se que $F_{[2,369]} = 2,7187 < F_{[2,369]}$ crítico, aceitando-se a hipótese de especificação linear para ambos os níveis de significância estatística, de 95% e de 99%.

Para testar a estabilidade dos parâmetros do modelo de preços hedónicos, pode-se utilizar testes à capacidade preditiva¹²⁴ desse modelo, sustentando-se, alguns deles, numa técnica empírica de partição do conjunto de observações em dois ou mais sub-grupos, um para estimar o modelo, o outro para testar o modelo. Esta técnica tem como pressuposto o conceito de invariabilidade dos parâmetros, que se concretiza no facto do vector β ter que ser o mesmo, quer dentro da amostra, quer fora da amostra.

Elegeram-se os testes de Chow para testar a estabilidade dos parâmetros do modelo de preços hedónicos, para o efeito procedeu-se à partição da amostra em duas,¹²⁵ a partir da variável quantitativa “preço da habitação”, por se estar na presença de dados seccionais. Os testes de Chow, de acordo com o EVIEWS 4, são descritos por dois testes estatísticos: o F estatístico,¹²⁶ que segue uma distribuição normal, caso os erros sejam

¹²⁴ Grandes erros de predição põem em causa a hipótese de invariabilidade dos parâmetros, acontecendo o inverso para pequenos erros de predição.

¹²⁵ Na ausência de uma razão apriorística de descontinuidade nos dados observados, pode-se aplicar uma regra de partição, comumente utilizada, em que 85 a 90% das observações servem à estimação do modelo e as restantes ao teste do modelo.

¹²⁶ O ajustamento do modelo de preços hedónico à totalidade da amostra (N) e a cada uma das duas sub-amostras (N_i) produz as Somas dos Quadrados dos Resíduos, SSR_R restrita e a SSR_{UR} não restrita, respectivamente. Da comparação destas Somas dos Quadrados dos Resíduos, SSR_R e SSR_{UR} , restrita e não restrita, resulta o F estatístico. Caso se equacione a Soma dos Quadrados

variáveis aleatórias normais idêntica e independentemente distribuídas; e o LR , estatístico que segue uma distribuição assintótica do χ^2 .

Aplicando os testes de Chow, às observações relativas ao mercado de habitação na cidade da Guarda, e procedendo à partição, na observação 320, da amostra global, obtêm-se os F estatísticos, $F_{[29,342]} = 0,3860$ e $F_{[80,291]} = 0,2651$, respectivamente para os testes de descontinuidade e de previsão. A consulta das tabelas da distribuição F mostra a evidência de qualquer um destes valores obtidos ser inferior ao F crítico, para ambos os níveis de significância, 95% e 99%, não se rejeitando a hipótese nula de que o vector β é o mesmo dentro e fora da amostra.

O teste de previsão de Chow, para detecção de qualquer ruído na estabilidade estrutural do modelo, poderá vir complementado com testes baseados em resíduos recursivos resultantes de estimação recursiva.¹²⁷ Optou-se pelo teste de predição N -etapas à frente, por constituir uma sequência de testes de previsão de Chow e por, também, permitir ultrapassar a necessidade de especificação da dimensão da amostra de previsão, já que neste teste a amostra é, progressivamente, ajustada pela inclusão, sucessiva de mais uma observação.¹²⁸ O resultado gráfico gerado pelo teste de predição N -etapas,

dos Resíduos, na amostra total (SSR_R) e em cada uma das sub-amostras (SSR_{UR_i}), está-se a aplicar o teste de descontinuidade de

Chow, expresso por, $F_{[K,N-2K]} = \left(SSR_R - \left(\sum_{i=1}^2 SSR_{UR_i} \right) \right) / K \left/ \left(\sum_{i=1}^2 SSR_{UR_i} \right) / (N - 2K) \right.$, em que K representa o número de

coeficientes estimados e N o total de observações. Caso se equacione a Soma dos Quadrados dos Resíduos, na amostra total (SSR_R) e apenas em uma amostra (SSR_{UR1}), está-se a aplicar o teste de previsão de Chow, expresso por,

$$F_{[N_2, N_1 - K]} = ((SSR_R - SSR_{UR1}) / N_2) / (SSR_{UR1} / (N_1 - K)).$$

¹²⁷ Na estimação recursiva o ajustamento regressivo inicia-se com a sub-amostra mais pequena possível, ou seja a partir da amostra com um número de observações igual ao número de regressores e, seguidamente, repete-se a estimação sobre nova sub-amostra que acrescenta uma observação à sub-amostra anterior e, assim, sucessivamente, até cobrir o total da amostra. Os erros resultantes dessa previsão denominam-se resíduos recursivos.

¹²⁸ Esta metodologia pode fragilizar as conclusões do teste, já que o impacto de uma nova observação nos coeficientes estimados é cada vez menor à medida que a amostra cresce. Para atenuar esta limitação, recorre-se ao teste um passo à frente, e a semelhança dos resultados obtidos permite conferir consistência aos resultados do primeiro.

mostra, com base no F estatístico, os erros de previsão recursiva e as probabilidades significantes, conforme se pode visualizar,

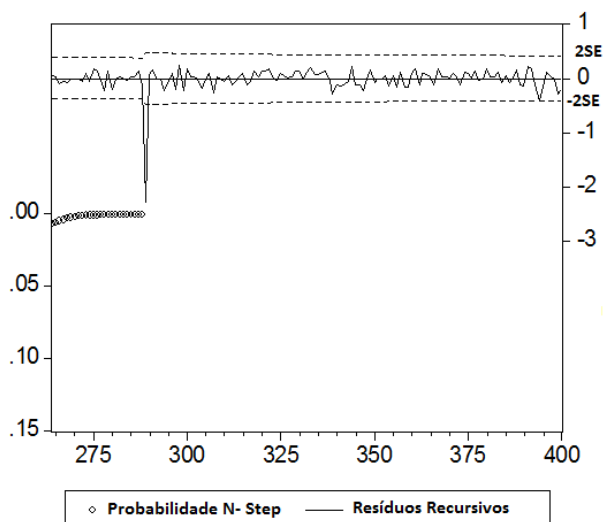


Gráfico 3: Teste de predição N -etapas

Uma vez que, o eixo da direita identifica a escala dos resíduos recursivos e dos respectivos erros padrão ($\pm 2SE$) e o eixo da esquerda a escala dos p -valores relativos às observações da amostra, em que a hipótese de estabilidade do modelo é rejeitada, para níveis de significância estatística de, 5%, 10% e 15%; podia-se suspeitar da instabilidade dos parâmetros devido ao movimento vertical para um nível fora dos intervalos de confiança previamente definidos, entre as observações 280 e 290 da amostra representativa da cidade da Guarda, porém os níveis de significância dessa ocorrência permitem afastar a suspeição de instabilidade.

4.5.2. A escolha das variáveis independentes e a especificação do modelo

Analisando os resultados relativos à regressão, pelo método dos mínimos quadrados, do modelo semi-logarítmico do mercado da habitação, apresentado na secção 4.4., constata-se que um elevado número de variáveis/características da habitação apresenta

valores da estatística t relativamente baixos. Esta ocorrência pode evidenciar que os respectivos coeficientes não são significativamente diferentes de zero, e, ainda, reduz os valores da estatística t associados às outras variáveis, aumentando por isso a probabilidade de *erros do tipo II*.¹²⁹ A inclusão de potenciais variáveis irrelevantes na especificação do modelo, ainda que permita estimar coeficientes não enviesados, gerará uma redução no R^2 -ajustado da regressão e uma sobrestimação na variância dos coeficientes estimados, significativamente diferentes de zero.

Tendo em conta que, (i) um aumento do número de variáveis explicativas num modelo de regressão linear produz uma transferência entre ganhos de R^2 e perda de graus de liberdade; (ii) a parcimónia é uma virtude; (iii) grandes modelos são mais vulneráveis à multicolinearidade; (iv) *erros do tipo II* estão associados à redução de graus de liberdade, poder-se-á ensaiar uma especificação alternativa (15 variáveis independentes) ao modelo semi-logarítmico com 28 variáveis independentes, assim,¹³⁰

$$\begin{aligned} \ln VMERCADO = & 10,726 - 0,0140* ANOS + 0,0036* AREA + 0,0694* QUARTOS \\ & (179,656) \quad (-5,938) \quad (4,049) \quad (1,966) \\ & + 0,0348 * CBANHO - 0,0196* ELEV + 0,0539* GCANAL \\ & (0,867) \quad (-0,764) \quad (1,591) \end{aligned}$$

¹²⁹ São erros resultantes da aceitação de uma hipótese falsa.

¹³⁰ Knight *et al.* (1993) introduziram a regra de Stein na estimação dos coeficientes das características da habitação, recorrendo ao método *MQO*, com o objectivo de melhorar o rigor da estimação. O procedimento subjacente a essa regra consiste em, primeiro estimar o modelo, via *MQO*, considerando todas as variáveis e respectivas correlações parciais (i.e. os beta coeficientes, factores de ajustamento marginal), a seguir proceder à remoção das variáveis, para as quais o nível de significância estatística permite aceitar a hipótese de valores beta igual a zero, ou seja remover aquelas com efeito insignificante sobre o preço e estimar o modelo restrito daí decorrente. Após o que se comparam os coeficientes estimados. De acordo com a regra de Stein se os coeficientes estimados com o modelo restrito vêm próximos dos coeficientes estimados com o modelo não restrito então os coeficientes estimados no contexto do modelo restrito ganham maior peso. No caso em estudo o valor dos coeficientes, estimados em cada uma das circunstâncias, denota uma tendência mais favorável aos resultados facultados pelo modelo restrito.

$$+0,1207^* \text{GSDOS} + 0,0589^* \text{GSUM} + 0,07152^* \text{LAR}$$

(2,002) (1,961) (2,479)

$$-0,0553^* \text{LCNTRO} - 0,0993^* \text{LGGARE} - 0,0771^* \text{LPMILEU}$$

(-1,980) (-3,475) (-1,619)

$$-0,0658^* \text{LPNHR} + 0,0436^* \text{PPRIM} - 0,0379^* \text{PSEG}$$

(-1,119) (1,680) (-1,457)

$$R^2 = 0,62134$$

$$R^2_{\text{Ajustado}} = 0,60655$$

Introduziram-se dois métodos, o Critério de Informação de Akaike (*AIC*) e o Critério de Schwarz (*SC*), para comparar a especificação do modelo, acima representada, com a especificação do modelo constante da secção 4.4.. Esses critérios permitem comparar as especificações alternativas, com base na soma do quadrado dos resíduos (*SSR*) para uma dada dimensão da amostra (*N*) e para um determinado número de variáveis independentes (*K*).¹³¹ Judge *et al.* (1988), mostraram que uma especificação, escolhida em função do melhor (maior) valor do coeficiente R^2 de uma regressão, tem uma probabilidade superior a 50% de incluir variáveis irrelevantes relativamente a uma escolha em função dos valores melhores (menores) do *AIC* ou do *SC*.¹³²

Comparando os valores de *AIC* e de *SC* para as duas especificações, tem-se $AIC_R = -0,29407$ e $SC_R = -0,13441$, para a especificação restrita apenas a 15 variáveis/características da habitação e $AIC = -0,24097$ e $SC = 0,04841$ para a especificação apresentada na secção 4.4.. Constata-se que na especificação restrita a 15

¹³¹ $AIC = \log(SSR/N) + 2(K + 1)/N$ e $SC = \log(SSR/N)N + \log(N)(K + 1)/N$

¹³² Da mesma forma, uma escolha da especificação do modelo que persiga o objectivo do *AIC* ou do *SC* mínimo tenderá a omitir, com mais frequência, variáveis relevantes do que aconteceria caso a escolha perseguisse o objectivo da maximização de R^2 .

variáveis explicativas (características da habitação) os valores de AIC e de SC são inferiores aos obtidos para a especificação alargada às 28 variáveis e, simultaneamente, o valor do coeficiente R^2 - Ajustado aumenta, já que os resultados se traduzem num R^2 Ajustado_R(= 0,60655) > R^2 Ajustado(= 0,59758) Estes resultados podem evidenciar erro de especificação no modelo de preços hedónico com 28 variáveis.

Para clarificar a evidência referida, introduz-se o teste das variáveis redundantes, para testar a significância estatística de um conjunto de variáveis incluídas num modelo. Formalmente, o teste visa avaliar se um determinado subconjunto de variáveis, presentes num determinado modelo, têm coeficientes nulos e, conseqüentemente, se pode ser retirado do modelo. Este teste gera um F -estatístico e uma estatística LR (razão de verosimilhança). O F -estatístico segue uma distribuição F , desde que os erros sejam variáveis aleatórias distribuídas idêntica e independentemente (i.i.d.), e nesse caso tem-se $F[13,371] = 0,34175$, que é inferior ao $F[13,371]$ crítico, facto que permite não rejeitar a hipótese nula. A estatística LR é um teste assintótico, com uma distribuição do χ^2 , assumindo o valor de $\chi^2_{[13]} = 4,76155$, que é inferior ao valor crítico para qualquer nível de significância, não se rejeitando, pois, a hipótese nula.

A não rejeição da hipótese nula significa aceitar que as 13 variáveis retiradas são, efectivamente, redundantes.

No que diz respeito aos coeficientes estimados para as 15 variáveis que permanecem nesta especificação alternativa, as conclusões relativamente aos níveis de significância e respectivos sinais coincidem com as alcançadas no contexto da estimação concretizada em 4.4..

4.5.3. A heteroscedasticidade e a consistência do modelo semi-logarítmico

A heteroscedasticidade coloca sérios problemas à inferência estatística de estimações fundamentadas no método de mínimos quadrados, dificultando a obtenção de previsões rigorosas, o que significaria no modelo em estudo, previsão de preços da habitação pouco fiáveis. Há pois que despistá-la e corrigi-la, caso exista.

Vários tipos de testes se podem utilizar para o efeito. A maior parte dos testes baseia-se no pressuposto que os coeficientes estimados, pela regressão dos mínimos quadrados, para o conjunto de regressores são consistentes, mesmo na presença de heteroscedasticidade. Ora, a presença de variabilidade na amostra, ou seja de heteroscedasticidade nas variâncias das perturbações reais da amostra, mascara a consistência dos resíduos, ainda que, de forma imperfeita. Os testes podem distinguir-se pelo grau de generalidade, o qual está associado ao poder do teste; assim quanto mais genérico, menos poderoso e vice-versa.

Principiando pelo teste assintótico de White, o seu carácter genérico, não exige a especificação das variáveis que se presumem causarem heteroscedasticidade, dispensando a assumpção de pressupostos específicos sobre a natureza e estrutura da heteroscedasticidade, o que constitui uma virtude, mas ao mesmo tempo pode constituir um potencial problema, pois se a hipótese de homoscedasticidade tiver que ser rejeitada, o teste não fornece indicação dos procedimentos seguintes. Muitas vezes, o teste revela problemas de heteroscedasticidade que podem, simplesmente, reflectir erros de especificação prevalentes no modelo (p.e. omissão de uma variável relevante ou má especificação da forma), daí a fragilidade do teste caso se rejeite a hipótese nula.¹³³

¹³³ O teste de White é obtido a partir do R^2 da regressão do quadrado dos resíduos, ε_i^2 , da função de preços hedónica, inicialmente estimada, sobre a constante e todas as variáveis independentes, características/atributos da habitação, multiplicado pelo nº de observações da amostra, n . Esta medida estatística segue uma distribuição assintótica do χ^2 , com $K-1$ graus de liberdade, em que K traduz o número de regressores, incluindo a constante. A hipótese a testar é da forma $H_0 : \sigma_i^2 = \sigma^2$, para qualquer i , e aceitá-la significa verificar que o estimador convencional, $V = s^2 [\mathbf{x}'\mathbf{x}]^{-1}$, é um estimador consistente da $\text{Var}[b]$

($\text{Est. Var}[b]$), isto é se, $V \equiv \text{Est. Var}[b] = [\mathbf{x}'\mathbf{x}]^{-1} \left[\sum_{i=1}^n e_i^2 \mathbf{x}_i \mathbf{x}_i' \right] [\mathbf{x}'\mathbf{x}]^{-1}$.

O teste de White, neste caso concreto, produz um $n * R^2$ de 26,6789. Observando as tabelas da distribuição do χ^2 , para 32 graus de liberdade, qualquer que seja o nível de significância estatística, não se rejeita a hipótese nula, ou seja a hipótese de existência de homoscedasticidade não se rejeita.¹³⁴ É evidente que os resultados obtidos, ao não excluírem a hipótese de homoscedasticidade, validam o poder do teste de White e diluem as fragilidades atrás referidas.

Mesmo assim, optou-se por aplicar ao modelo semi-logarítmico de preços hedônico, suporte deste trabalho, os testes Goldfeld-Quandt e Breusch-Pagan de forma a aumentar o poder das conclusões obtidas com o teste geral de White.¹³⁵

O teste de Goldfeld-Quandt pressupõe uma partição das observações em dois subgrupos, sustentada numa qualquer variável independente $x_i \equiv z_k$, suspeita de provocar distúrbio sistemático nas variâncias das perturbações, de tal forma que, sobre a hipótese nula de homoscedasticidade, as variâncias das perturbações serão as mesmas nesses dois sub-grupos. O teste estatístico, sobre a hipótese nula de homoscedasticidade, segue uma distribuição de $F[N_1 - K, N_2 - K]$, com $N_1 - K$ e $N_2 - K$ graus de liberdade, traduzindo N_1 e N_2 a dimensão de cada uma das sub-amostras e K o número de regressores.¹³⁶ A aplicação do teste ter-se-á que fazer variável a variável, caso se suspeite de mais do que uma variável perturbadora da constância das variâncias, já que as observações em cada uma das sub-amostras/suporte do teste se irão ordenar em função da ordem crescente da variável suspeita de causar heteroscedasticidade.

¹³⁴ Note-se que este teste é fornecido pelo menu EVIEWS 4, o qual por insuficiência de observações recusou a opção teste de White, com termos cruzados.

¹³⁵ Vários estudos sobre modelos hedônicos de preços da habitação identificaram heteroscedasticidade relacionada com as variáveis independentes, daí a necessidade de se implementarem testes mais poderosos de homoscedasticidade.

¹³⁶ $F_{[N_1-K, N_2-K]} = \frac{SSR_1 / (N_1 - K)}{SSR_2 / (N_2 - K)}$, o que significa estimar separadamente a variância da perturbação em cada amostra e assumir-se que a variância da perturbação é maior na primeira amostra, caso assim não aconteça dever-se-á inverter os índices.

Para o modelo, em estudo, aplicando o teste à variável que traduz a idade dos imóveis, “ANOS”, obteve-se um $F_{[203,71]} = 1,2507$,¹³⁷ o que permite não rejeitar a hipótese de homoscedasticidade, para qualquer nível de significância estatística.

Como este teste incide, apenas, sobre uma variável, de cada vez, para se ser exaustivo, ter-se-ia que apresentar tantos testes de Goldfeld-Quandt quantas as variáveis suspeitas de produzir heteroscedasticidade. Aplicando sobre a variável “AREA”, a amostra em estudo apresentar-se-á ordenada, segundo a ordenação crescente dessa variável e gerará o resultado $F_{[218,55]} = 2,53869$,¹³⁸ a rejeição da hipótese nula de homoscedasticidade será uma realidade.

Alargando o universo de aplicação do teste a nova variável, a amostra em estudo virá ordenada segundo a forma ascendente da variável “QUARTOS”, produzindo um resultado de $F_{[197,77]} = 2,47512$.¹³⁹ Também aqui se verificaria um F significativamente maior que o F crítico, rejeitando-se, pois, a hipótese de homoscedasticidade relativamente a “QUARTOS”.

Finalmente, aplicando à variável “CBANHO”, e passando a amostra a ordenar-se de acordo com a ordenação crescente dessa variável, suscitada pelo teste Goldfeld-Quandt,

¹³⁷ De acordo com Goldfeld e Quandt, omitiu-se um conjunto de observações centrais da amostra, para aumentar o poder do teste, equacionando-se, também, o efeito reverso que a diminuição dos graus de liberdade da estimação induz na redução do poder do teste. Dada a subjectividade do número de observações centrais a retirar, poder-se-ia optar pela evidência de Harvey e Phillips (1974) que sugere um número não superior a $\frac{1}{3}$ para as observações centrais retiradas, o que significaria, neste caso, deixar de fora não mais de 133 observações. Acontece que, ao estimar-se a regressão sobre a amostra, depois de ordenada, em função da ordem ascendente da variável “ANOS” (suspeita de heteroscedasticidade), dado as primeiras 231 linhas, do vector coluna “ANOS”, serem sempre iguais a 1 (pelo peso de imóveis novos na transacção imobiliária da Guarda), a partição de Harvey e Phillips conduziria a uma matriz singular para a sub-amostra relativa às primeiras 134 observações. Como a subjectividade da partição é uma característica do teste, havia liberdade para conjugar os valores da partição com os valores que eliminam a dependência linear das colunas da matriz x das variáveis independentes, avançando-se com 232 observações e 100 observações para cada uma das sub-amostras.

¹³⁸ A partição neste caso teve, também, que se fazer conciliando dimensão da amostra central, dependência linear das colunas da matriz x , e graus de liberdade utilizados no teste Goldfeld-Quandt sobre a variável “ANOS”.

¹³⁹ Neste caso a variável “QUARTOS” admitia uma partição mais generosa com o número de observações centrais passíveis de ficar de fora do teste, porém a comparabilidade do estudo, recomenda a manutenção do número de observações centrais retiradas (68).

ter-se-ia $F_{[197,77]} = 2,4018$. Novamente um F superior ao F crítico, rejeição da hipótese nula de homoscedasticidade.

Como se verificou, o teste de Goldfeld-Quandt apresenta limitações, na medida em que só contempla, de cada vez, uma variável, ainda que se suspeite de várias produzirem heteroscedasticidade. Entretanto, ao permanecer a suspeita de que pode haver variação na variância da perturbação, devido a um conjunto de variáveis independentes, como é o caso (“*QUARTOS*”, “*CBANHO*” e “*AREA*”), a aplicação de um outro teste, o de Breusch-Pagan, ultrapassará a restrição subjacente ao de Goldfeld-Quandt.

O teste de Breusch-Pagan é um teste de Multiplicador de Lagrange, LM teste, sobre a hipótese da variância da perturbação σ_i^2 , ser função de um conjunto de variáveis independentes, suspeitas de serem responsáveis dessa heteroscedasticidade.¹⁴⁰ O teste estatístico LM é ilustrado pela metade da soma dos quadrados explicada da regressão e segue uma distribuição do $\chi_{[g]}^2$, em que g , graus de liberdade, traduz o número de variáveis suspeitas de causar instabilidade na variância da perturbação. Com a aplicação do teste às variáveis, *a priori*, suspeitas de produzir heteroscedasticidade, “*ANOS*”, “*QUARTOS*”, “*CBANHO*” e “*AREA*”, obteve-se $LM = 7,22185$, ou seja uma distribuição do $\chi_{[4]}^2 = 7,22185$, que, apenas ao nível de significância de 1%, consegue ser inferior ao valor crítico de $\chi_{[4]}^2 = 7,779$. O teste é pouco robusto, pois a hipótese nula de não heteroscedasticidade, só pode ser aceite para um intervalo de confiança de 90%, daí que possam subsistir dúvidas quanto ao facto dessas variáveis conduzirem os resíduos a variarem sistematicamente.

¹⁴⁰ Aplica-se uma regressão de mínimos quadrados ao modelo hedónico completo. Com os resíduos estimados, $\hat{\varepsilon}_i$, faz-se correr uma regressão com $\frac{\hat{\varepsilon}_i^2}{SSR/N}$ sobre o vector das variáveis independentes suspeitas de causar heteroscedasticidade, com o objectivo de obter a soma dos quadrados explicada, suporte ao *output* do teste, $LM = \frac{1}{2}ESS \sim \text{assimpt.} \chi_g^2$.

A prevalência de alguma heterogeneidade nos resultados dos testes utilizados para identificar a presença de heteroscedasticidade, poderá conduzir a uma resposta inconclusiva no que concerne ao problema de heteroscedasticidade. Valerá pois, a pena avançar para o teste de Glesjer, o qual exige formulação específica para a variância dos resíduos, já que se pretende identificar a presença de heteroscedasticidade com a forma funcional da variância dos resíduos. Só conhecendo a forma estrutural da heteroscedasticidade se pode tentar corrigir a estimação dos mínimos quadrados, por via da estimação generalizada dos mínimos quadrados, de forma a conseguir-se melhorar a eficiência dos estimadores. O teste de Glesjer segue uma distribuição $F[G, N - G - 1]$,¹⁴¹ em que G é o número de variáveis independentes suspeitas de causar heteroscedasticidade e N é o número de observações da amostra. No caso em estudo e aplicando o teste às 3 variáveis suspeitas de causar heteroscedasticidade,¹⁴² “*QUARTOS*”, “*CBANHO*” e “*AREA*”, ter-se-ia um $F_{[3,396]} = 1,1631$, quando a regressão se faz sobre a forma positiva linear dos resíduos; um $F_{[3,396]} = 1,1148$, para uma regressão sobre a forma quadrática dos resíduos e finalmente um $F_{[3,396]} = 0,7103$, para uma regressão sobre a forma logarítmica dos resíduos. O valor crítico do $F_{[3,396]}$, para os diferentes níveis de significância, é sempre superior aos valores do $F_{[3,396]}$ obtidos via teste de Glesjer, o que permite rejeitar a hipótese de heteroscedasticidade relacionada com qualquer uma das formas funcionais consideradas.

No decurso desta secção a não uniformidade na rejeição da homoscedasticidade, admitida pelo teste de Goldfeld-Quandt, ponderada pelas limitações do mesmo,

¹⁴¹ $F[G, N - G - 1] = \frac{R^2/G}{(1 - R^2)/(N - G - 1)}$, em que o R^2 é obtido a partir da regressão dos resíduos estimados $\hat{\epsilon}_i$ (resultantes da estimação pelos mínimos quadrados do modelo de preços hedónicos geral) sobre o vector das variáveis independentes suspeitas de causar heteroscedasticidade. Ter-se-á tantos F quanto as formas funcionais consideradas, conforme a regressão de $\hat{\epsilon}_i$ sobre as variáveis suspeitas de causar heteroscedasticidade, se faz considerando $\hat{\epsilon}_i^2$, ou $|\hat{\epsilon}_i|$, ou $\ln|\hat{\epsilon}_i|$.

¹⁴² Escolha feita em função do resultado dos testes Goldfeld-Quandt, mas se avançássemos com as 4 variáveis, *a priori*, suspeitas de causar heteroscedasticidade, o teste de Glesjer conduziria às mesmas conclusões.

conduziu a que se procurasse obter mais informação sobre a estrutura do termo erro. Nessa perspectiva aplicou-se os testes de Glesjer, às 3 variáveis suspeitas de causar heteroscedasticidade, “*QUARTOS*”, “*CBANHO*” e “*AREA*”, e os resultados ao contemplarem a aceitação da homoscedasticidade permitiriam encerrar a problemática da heteroscedasticidade, ainda assim optou-se por apresentar na secção seguinte um exercício pedagógico de correcção de uma *pseudo*-heteroscedasticidade.

4.5.4. Correcção da heteroscedasticidade e a estimação generalizada dos mínimos quadrados

Assumindo-se a suspeita de heteroscedasticidade, seríamos confrontados com a presença de elementos de perturbação que violam os pressupostos clássicos do modelo de regressão múltiplo e nessa circunstância a estimação eficiente dos coeficientes, β_i , devia ser garantida através de um modelo de mínimos quadrados generalizado.¹⁴³

A estimação desse modelo permite a manutenção da especificação linear, mas exige o conhecimento de uma matriz simétrica definida positiva Ω , que garanta a eficiência do modelo clássico, dos mínimos quadrados.¹⁴⁴ O modelo inicial dos mínimos quadrados ordinários¹⁴⁵ é re-estimado com mínimos quadrados ponderados utilizando os recíprocos dos valores previstos normalizados a partir do modelo de resíduos

¹⁴³ Seguindo a metodologia de Goodman e Thibodeau (1995, 1997) baseada numa perspectiva do modelo iterativo proposto por Davidian and Carroll (1987).

¹⁴⁴ De modo a que, $\hat{\beta} = (\mathbf{x}'\Omega^{-1}\mathbf{x})^{-1}\mathbf{x}'\Omega^{-1}\mathbf{y}$. Maior detalhe para a dedução deste estimador pode ser encontrado em Greene (2003).

¹⁴⁵ O modelo inicial é o que decorre do ajustamento do logaritmo do preço da habitação a todas as variáveis independentes/características da habitação, como acontecia no modelo semi-logarítmico apresentado na secção 4.4. e, ainda, aos polinómios de 2ª ordem das variáveis suspeitas de causar heteroscedasticidade, “*QUARTOS*”, “*CBANHO*” e “*AREA*”, expresso por, $\ln p = \mathbf{x}\beta + \phi(CBANHO) + \varpi(AREA) + \gamma(QUARTOS) + \varepsilon$, em que, $\ln p$ é um vector 400*1 dos logaritmos dos preços de habitação, \mathbf{x} é a matriz 400*25 das características da habitação, à excepção das consideradas nos polinómios de 2ª ordem $\phi(\cdot)$, $\varpi(\cdot)$ e $\gamma(\cdot)$. A escolha da ordem dos polinómios procura reflectir a gravidade da heteroscedasticidade a corrigir. No caso em estudo, em que esta correcção toma essencialmente o papel de um exercício pedagógico, escolheu-se uma forma polinomial de 2ª ordem, vindo p.e., $\phi(CBANHO) = \phi_1 + \phi_2 CBANHO + \phi_3 CBANHO^2$, daí decorrendo $\partial \ln p / \partial CBANHO_i = \phi_2 + 2\phi_3 CBANHO_i$, forçando uma taxa de variação linear.

absolutos.¹⁴⁶ Este procedimento é repetido até que a variação (em valor absoluto) em qualquer parâmetro estimado venha inferior a **0,0001**.

Aplicando a metodologia, sustentada na ferramenta *validação cruzada*, ao modelo generalizado dos mínimos quadrados, a amostra dos dados observados, 400, continua a dividir-se em duas sub-amostras: uma de 350 observações para a estimação dos parâmetros do modelo e outra de 50 observações sobre a qual se calculam as estatísticas de validação, ou seja as medidas de erro ou rigor da previsão, *MAE* (Erro Absoluto Médio), *RMSE* (Raiz Quadrada do Erro Quadrado) e *U de Theil*,¹⁴⁷ utilizando para o efeito o software *EViews 4*.

Comparando os valores médios, do erro absoluto médio ($MAE_{Média}$), da raiz quadrada do erro quadrado ($RMSE_{Média}$) e do *U – Theil* ($U_{Média}$) do modelo log-linear (semi-logaritmo) e do modelo iterativo generalizado dos mínimos quadrados (*MQG*) para avaliar a superioridade residual existente, tem-se,

	Modelo Log-linear	Modelo <i>MQG</i>
$MAE_{Média}$	10793,2321	10989,5995
$RMSE_{Média}$	21011,7570	21534,6288
$U_{Média}$	0,03175237	0,0310197

Tabela 4.5-1: Medidas de erro médio (*MQG*)

¹⁴⁶ O modelo dos resíduos absolutos pressupõe uma regressão do valor absoluto dos resíduos estimados, sobre um polinómio de 2ª ordem, por cada uma das variáveis suspeitas de causar heteroscedasticidade, expressa por, $|\varepsilon_i| = \phi(CBANHO) + \varpi(AREA) + \gamma(QUARTOS) + \nu$, em que ν é a parte do termo erro não relacionado com as variáveis *CBANHO*, *AREA*, *QUARTOS*. O quadrado dos valores estimados dos resíduos, $|\hat{\varepsilon}_i|^2$, depois de normalizados constituem os elementos da diagonal $\sigma^2\Omega$.

¹⁴⁷ Continuou a optar-se por gerar séries aleatórias de forma a poderem-se apurar valores médios das medidas de erro *MAE*, *RMSE* e *U de Theil*, em 1000 regressões.

As diferenças entre as medidas de fiabilidade da previsão são relativamente insignificantes não aportando a estimação *MQG* qualquer melhoria, quando aplicada ao modelo do desvio padrão para um polinómio do 2º grau, para correcção de eventual heteroscedasticidade. A ser assim, mesmo na ausência de uniformidade nos resultados dos testes utilizados para detectar a presença de heteroscedasticidade, deve-se reter a conclusão de não heteroscedasticidade obtida a partir dos testes de Glesjer.

4.6. Multicolinearidade entre os Regressores

A multicolinearidade é uma propriedade dos dados da amostra, tem a ver com as características específicas da matriz de dados, \mathbf{x} , e não com aspectos estatísticos do modelo de regressão linear. Não conduz à violação dos pressupostos do modelo clássico de regressão linear, excepto para o caso extremo em que um dos regressores, x_j , é uma função linear dos outros regressores, x_i ; com $i \neq j$ (a acontecer, a matriz $\mathbf{x}'\mathbf{x}$ seria singular e o cálculo dos coeficientes de estimação falharia). Algum grau de multicolinearidade subsistirá, no entanto, entre a situação referida e uma outra, também extrema, em que os regressores x_j são ortogonais¹⁴⁸ e a matriz $\mathbf{x}'\mathbf{x}$ é uma matriz diagonal.

Como a multicolinearidade é um problema da amostra (e não da população) não existem testes formais para detectar a sua presença, ou melhor para detectar o seu grau. Porém alguns sintomas de multicolinearidade, relacionados com os coeficientes estimados da regressão, podem ser identificados, p.e., elevados erros padrão das estimações; t rácios bastante pequenos; acréscimo da probabilidade *de erros do tipo II*.¹⁴⁹ Num cenário de suspeita de alto grau de multicolinearidade poder-se-á, numa regressão com K variáveis independentes, proceder à regressão de uma qualquer k -ésima variável independente, x_k , sobre as restantes variáveis independentes, de forma exaustiva e

¹⁴⁸ Traduzido por variáveis exógenas ou explicativas perfeitamente independentes.

¹⁴⁹ Ainda que não sejam condição suficiente nem condição necessária para a existência de multicolinearidade.

comparar os sucessivos R_k^2 com o R^2 da regressão sobre o modelo geral de preços hedónico.¹⁵⁰

Uma vez que o mercado da habitação em Portugal segue uma tipologia mais ou menos padronizada, em que o número de quartos aparece relacionado com o número de casas de banho e em que essa tipologia tem subjacente uma determinada área, este facto poder-se-á traduzir em variáveis altamente correlacionadas, umas com as outras, o que matematicamente poderá induzir a uma característica da matriz \mathbf{x} inferior ao número de colunas (colinearidade).

Para determinar o grau de multicolinearidade, ainda antes de se recorrer à metodologia de regressão auxiliar, onde, sucessivamente, cada uma das variáveis independentes é estimada sobre as outras variáveis independentes, pode começar-se pela análise da matriz de correlação, de ordem zero, entre as variáveis independentes, para avaliar a dimensão das correlações entre as variáveis explicativas, estar-se-ia a aplicar o teste *ZOCMAOTIV* (*Zero-Order Correlation Matrix Among Only The Independent Variables*). Conforme matriz (em Apêndice **F**) e seguindo uma regra aproximada e prática de mensuração,¹⁵¹ apenas se observam três coeficientes de correlação, superiores a 0,7, o que denotaria, nesses casos, que mais de 49% da variância gerada pelas duas variáveis independentes em análise é variância partilhada. Esta ocorrência não tem que ser, necessariamente, um problema, mas deverá constituir uma preocupação.

Já no cenário das regressões auxiliares entre as variáveis independentes, sabendo que quanto mais elevado o R^2 nessas regressões, maior a “dose” de multicolinearidade presente nos dados e sabendo, ainda, que essa é uma questão de grau, passar-se-á a

¹⁵⁰ Subsistem dois problemas, no entanto, é apenas uma condição suficiente para existência de elevada multicolinearidade e não há um valor definido para R_k^2 relativamente a R^2 para identificar a potencial multicolinearidade. Klein (1962) refere que a seriedade do problema de multicolinearidade depende da relação entre os coeficientes de correlação simples entre variáveis independentes e o coeficiente de determinação da regressão contendo todas essas variáveis. Sobre esta questão, sugere que, caso $R^2 < R_k^2$, a multicolinearidade baixaria significativamente os níveis do t-rácio dos coeficientes estimados e colocaria sérios problemas aos resultados da regressão.

¹⁵¹ Para maior detalhe ver Belsley, D.A. (1991).

comparar os R^2 dessas regressões auxiliares com o R^2 do modelo de regressão geral dos preços hedónicos. Se este último for, pelo menos, duas vezes superior à média dos primeiros, então dever-se-á rejeitar a hipótese da dependência linear das variáveis em observação.¹⁵²

No caso concreto em estudo, as regressões auxiliares produziram os R^2 e R^2_{Ajustado} , constantes da tabela 1, apêndice G. A média dos, R^2 e R^2_{Ajustado} , para as 28 regressões auxiliares é $R^2 = 0,45522$ e $R^2_{\text{Ajustado}} = 0,41568$, respectivamente. Entretanto o R^2 e o R^2_{Ajustado} , para o modelo semi-logarítmico de preços hedónicos, é 0,62329 e 0,59485, respectivamente. Ora, não se verificando uma superioridade do R^2 da regressão, relativamente ao dobro da média dos R^2 das regressões auxiliares, estar-se-á, perante um problema sério de multicolinearidade. Com as regressões auxiliares, envolvendo as 28 variáveis independentes, pode-se, ainda, testar a hipótese nula de $R_k^2 = 0$, aplicando um F teste.¹⁵³ Os $F_{[28,371]}$ obtidos e, também, presentes na tabela 1, apêndice G, conduzem à rejeição da hipótese nula para todas as regressões auxiliares, à excepção das que dizem respeito à regressão de “*EXPSOL*”, “*GCANAL*”, “*LLUZ*”, “*LPNHR*”, “*LRDIZ*”, “*LSREMED*”, respectivamente, sobre todas as outras variáveis independentes.

Aproveitando os R^2 das regressões auxiliares, ainda, se pode recorrer a um outro indicador, para detectar a presença de multicolinearidade, o VIF,¹⁵⁴ Factor de Inflação da Variância, um rácio entre a variância da variável estimada, na situação real e a variância estimada, na situação ideal (colinearidade zero). Valores elevados de VIF

¹⁵² Para maior detalhe ver Belsley, D.A. (1991).

¹⁵³ Traduzido por $F[K-1; N-K] = (R_k^2 / (K-1)) / ((1-R_k^2) / (N-K))$, em que R_k^2 é o R-quadrado da regressão auxiliar de x_k sobre todas as outras variáveis independentes, há K variáveis independentes mais a constante e N observações da amostra.

¹⁵⁴ $VIF_k = \frac{1}{1-R_k^2}$, em que R_k^2 é o R-quadrado obtido a partir da regressão da variável independente k -ésima sobre todas as outras variáveis independentes.

revelam problemas severos de colinearidade, embora não haja um valor absoluto que o identifique, há alguma unanimidade para colocar a fasquia no limiar do valor 10. No caso em estudo e de acordo com os VIF constantes da tabela 1, referida, o valor mais elevado VIF (*AREA*) é 8,3827,¹⁵⁵ que, ainda assim, é indubitavelmente inferior a 10.

Poder-se-ão introduzir outros meios para determinar a presença de multicolinearidade na regressão, uma possibilidade é a medida de multicolinearidade de Theil, que utiliza os R^2 das j regressões múltiplas parciais.¹⁵⁶ Aplicando ao modelo de preços hedónicos do mercado da Guarda, obteve-se um $M_{theil} = 0,5606$. O intervalo de variação do M_{theil} está compreendido entre zero e um, e o caso ideal de não multicolinearidade induziria para um valor de zero, ou muito próximo de zero. O presente resultado evidencia a presença de multicolinearidade, no entanto, na ausência de uma regra prática para avaliar da gravidade do grau de multicolinearidade, não se revela conclusivo

Belsley *et al.* (1980) sugeriram o Número de Condição (φ) da matriz¹⁵⁷ como um potencial indicador de multicolinearidade, que no caso em estudo, toma os valores de $\varphi = 13592215,08$ ou $\sqrt{\varphi} = 3686,76$ e com base neles confrontando-os com regras de mensuração,¹⁵⁸ comumente, utilizadas estar-se-ia perante um grau de

¹⁵⁵ Significando que a variância da regressão da variável “*AREA*” sobre todas as outras variáveis independentes é cerca de 8 vezes superior à que decorreria caso os regressores fossem ortogonais.

¹⁵⁶ As regressões múltiplas parciais consistem na estimação, pelo método dos mínimos quadrados, do modelo de preços hedónicos omitindo, sucessivamente uma variável independente, de forma exaustiva. O $M_{theil} = R^2 - \sum_{j=1}^K (R^2 - R_j^2)$, em que R^2 é o coeficiente de determinação da regressão completa e R_j^2 é o R-quadrado de cada uma das j -ésima regressões parciais do logaritmo do preço da habitação sobre todas as variáveis independentes, à exceção de uma, x_j .

¹⁵⁷ O Número de Condição da matriz é o rácio entre o maior e o menor valor próprio da matriz, $CN = \varphi = \frac{\lambda_{\max}}{\lambda_{\min}}$ ou $\sqrt{CN} = \sqrt{\varphi} = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$, λ é o símbolo convencional para o valor próprio. O valor próprio é obtido a partir da matriz simétrica $\mathbf{x}'\mathbf{x}$, invertida, dado que é sobre $[\mathbf{x}'\mathbf{x}]^{-1}$ que se estimam os coeficientes de regressão.

¹⁵⁸ Myers (1990) sugere que $\varphi > 1000$ é um grau de multicolinearidade mau e outros autores, entre eles, Olson (1971) referem $30 \leq \sqrt{\varphi} < 100$ como um grau moderado de multicolinearidade e $\sqrt{\varphi} \geq 100$ como um grau severo de multicolinearidade.

multicolinearidade severa. De acordo com Gujarati (1992) o grau de multicolinearidade associado a cada valor próprio da matriz pode ser diagnosticado utilizando o Índice de Condição ou Rácio do Valor Próprio, os quais pressupõe a determinação dos valores próprios da matriz inversa de $\mathbf{x}'\mathbf{x}$, simétrica, matriz das variáveis explicativas (independentes).¹⁵⁹ No caso em estudo o cálculo do Índice de Condição conduziu aos resultados, constantes da Tabela 2, apêndice G. Constata-se que, para cerca de 90% dos valores próprios, o índice de condição é igual ou inferior a 20, denotando fraca ou moderada colinearidade, existindo, porém, 3 valores próprios para os quais o Índice de Condição é superior a 30 que alertam para um grau de multicolinearidade sério, isto é para uma elevada percentagem de variância partilhada entre as variáveis independentes.

Como a multicolinearidade parece envolver vários regressores o teste de Haitovsky,¹⁶⁰ baseado num teste do qui-quadrado, para testar se a quantidade de variância partilhada entre as variáveis independentes é significativamente maior que zero, poderá dar alguma resposta. Da aplicação do teste à amostra de dados, base do estudo do mercado da habitação da Guarda, obteve-se um $\chi^2_{H[378]} = 0,000162$, inferior ao valor crítico de $\chi^2_{H[378]}$ para qualquer nível de significância e a hipótese nula de que a quantidade de variância partilhada entre as variáveis independentes é zero deve ser aceite.

Relativamente à multicolinearidade os diferentes testes, ainda que inconclusivos, acusam a presença de algum grau de colinearidade, a oscilar entre fraca a severa. Esta constatação associada ao facto das soluções¹⁶¹ para remediar o problema da multicolinearidade não conseguirem ultrapassar a origem desse problema, que reside na

¹⁵⁹ O índice de condição ou rácio do j -ésimo valor próprio é expresso por $\varphi_j = \sqrt{\frac{\lambda_{\max}}{\lambda_j}}$.

¹⁶⁰ O teste de Haitovsky é um teste baseado no qui-quadrado, $\chi^2_{H(g)} = h \ln(1 - |R|)$, em que, $h = 1 + [(2K + 5)/6] - N$ é a constante de Haitovsky, com K representando o número de variáveis independentes e N o número de observações da amostra; g representa os graus de liberdade, traduzidos por $K(K - 1)/2$ e $|R|$ é o determinante da matriz de correlação das variáveis independentes.

¹⁶¹ Regressão de Ridge, regressão restrita dos mínimos quadrados ordinários, com a particularidade da diagonal principal da matriz $\mathbf{x}'\mathbf{x}$ vir aumentada por uma pequena quantidade positiva ζ , conhecida como índice de enviesamento. O efeito líquido é gerar reduzidos erros quadrados médios de previsão.

informação obtida a partir dos dados observados e, ainda, induzirem enviesamento nas estimações e/ou má-especificação do modelo; associada, ainda, à circunstância da multicolinearidade não violar nenhum dos pressupostos subjacentes à inferência estatística nos modelos de regressão linear e, finalmente, associada aos baixos erros padrão dos coeficientes estimados, levar-nos-á a optar por conviver com essa multicolinearidade.

Importa alertar para a circunstância dos modelos de preços hedónicos se caracterizarem pela inclusão de variáveis dependentes desfasadas espacialmente, podendo pois gerar problemas de autocorrelação,¹⁶² o que passará a ser contemplado na secção que se segue.

4.7. Autocorrelação Espacial

Os dados da amostra em estudo, são *seccionais* e, em princípio, por esse facto não se espera autocorrelação, mas há circunstâncias em que os erros estocásticos podem aparecer autocorrelacionados. Esta ocorrência pode surgir quando observações *seccionais* adjacentes entre si partilham efeitos aleatórios similares, por exemplo, a possibilidade dos preços da habitação no bairro Lameirinhas e em alguns locais da zona denominada Centro poderem ser afectados pela exposição solar e orientação a Norte e, também, pela idêntica proximidade ao parque escolar. Ou, ainda, quando se omite na forma funcional um importante regressor (p.e. coordenadas geográficas, distância às escolas)¹⁶³ e, nesse caso, estar-se-ia no cenário de autocorrelação “aparente”.

¹⁶² A autocorrelação que resulta da correlação entre as observações ser função da *distância* no espaço e não da distância no tempo.

¹⁶³ Em geografia é fundamental estar consciente do espaço e ponderá-lo. Facto que levou, em 1979, Tobler a formular a 1ª lei da geografia, estabelecendo “todas as *coisas* estão relacionadas com todas as *coisas*, mas as *coisas* próximas estão mais relacionadas do que as *coisas* distantes.” A relevância dos efeitos espaciais expande-se para além da geografia e está presente em muito do domínio das Ciências Sociais. Os economistas, no entanto, têm sido tradicionalmente mais relutantes a equacionar o espaço como um factor relevante. Em 1890, Marshall reconhece o papel do espaço, sustentando que o funcionamento dos mercados depende, “principalmente das variações espaciais e das variações temporais relativamente ao mercado em questão; revelando-se, no entanto, a influência do tempo mais fundamental do que a do espaço” e assim se manteve a ideia até 1950. Não sem que, antes, Walter Isard se opusesse ao que apelidou de “enviesamento anglo-saxónico”, que repudiava o factor espaço, reduzindo tudo a um ponto na economia, confinando-a a um mundo sem dimensão espacial. Persiste, desde essa altura, um debate recorrente sobre se, o espaço é apenas um facilitador geográfico ou meio de movimento ou se tem uma função explicativa intrínseca. A revolução, na teoria

Neste contexto, o modelo de regressão clássico, corre sérios riscos de violar o pressuposto de não correlação entre os termos de erro (perturbação), ou seja não consegue garantir a satisfação da condição, $cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$. A presença de autocorrelação, traduzida em valores subestimados de erros padrão para os coeficientes estimados ($SE(\hat{\beta})$), em coeficientes estimados ($\hat{\beta}$) não enviesados, em t -rácios sobre-estimados e em valores de F , R^2 e R^2 -Ajustado sobre-estimados, aumenta a probabilidade de *erros do tipo II*.

Equacionar os efeitos espaciais nos modelos de regressão pode ser induzido por motivações diferentes. Há métodos que sustentam a dependência espacial como um ruído (perturbação) dos dados da amostra contra outros que suportam uma dependência espacial substantiva (conduzida pela teoria).¹⁶⁴ Formalmente estas perspectivas resultam em técnicas que modelam a dependência espacial no termo erro da regressão (ou, então, que transformam as variáveis no modelo, de forma a eliminar a correlação espacial) *versus* métodos que explicitamente acrescentam uma variável de inter-acção espacial como regressor, passando a considerar-se variáveis desfasadas espacialmente.

Seguindo o padrão da correlação espacial induzida pelos valores dos termos de perturbação da regressão expressa por, $\varepsilon = \rho \mathbf{W}\varepsilon + \mathbf{v}$, em que ε é o vector $N * 1$ dos termos aleatórios de perturbação, de distribuição normal $N(0, \sigma_\varepsilon^2 \mathbf{I})$, \mathbf{W} é a matriz $N * N$ das ponderações espaciais não-negativas, com $w_{ii} = 0, (i = 1, \dots, N)$, o escalar ρ , desconhecido, apurará o grau de correlação espacial entre os termos ε , o que se vai testar será a hipótese nula, $H_0 : \rho = 0$, para se aceitar ou rejeitar a hipótese de não correlação espacial.

económica, Dixit-Stiglitz, aumentou a consciência para a concorrência imperfeita e para os rendimentos crescentes das economias de escala, que é subsequentemente aparente na Nova Economia de Fujita e Krugman. Hoje o espaço é ponderado nos estudos de crescimento económico, das inovações tecnológicas, da economia urbana, da produtividade do sector público da interdependência da política fiscal e das trocas internacionais.

¹⁶⁴ Para maior detalhe ver Anselin (1988).

Os testes para apurar a presença de potenciais efeitos espaciais são complexos pelo número de factores que envolvem. Por exemplo os efeitos espaciais induzem problemas de heteroscedasticidade, tornando-se, pois, difícil distingui-la da verdadeira heteroscedasticidade, verificando-se, também, o reverso, daí que os testes de dependência espacial sejam sensíveis à heteroscedasticidade podendo conduzir a resultados errados. Constata-se, ainda, que os desfasamentos espaciais e as especificações dos erros espaciais estão altamente relacionados, fazendo com que testes poderosos a uma forma de dependência, também, o sejam à outra, complicando-se a pesquisa da especificação. Depois, todos os testes de efeitos espaciais, pressupõem as propriedades das grandes amostra (assimptóticas) e o seu desempenho em pequenas amostras pode ser suspeito.

Apesar das limitações referidas, o recurso a testes do Multiplicador de Lagrange (*LM* testes) sobre os resíduos de regressões dos mínimos quadrados ordinários, é frequente, para, alternativamente, testar a presença de desfasamentos espaciais ou de erros espaciais. Testes mais poderosos baseiam-se na aplicação do Índice de Moran¹⁶⁵ aos resíduos da regressão e, ainda, no teste *KR* (Kelejian e Robinson).¹⁶⁶ Estes testes são sensíveis à forma como se definem a estrutura linear ou a matriz de ponderação espacial, o que torna os modelos espaciais mais complexos que os modelos econométricos padrão.

¹⁶⁵ A primeira medida de autocorrelação espacial foi introduzida por Moran (1950), de modo a estudar-se o fenómeno estocástico decorrente de espaços a duas ou mais dimensões. Como coeficiente de correlação o índice de Moran varia entre $[-1,1]$ e expressa-se por,
$$I = \left(\frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} \right) \left(\frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \right)$$
, em que w_{ij} são elementos da matriz das ponderações espaciais, $(x_i - \bar{x})(x_j - \bar{x})$ é o produto das distâncias na localização i e na localização j em relação à localização média.

¹⁶⁶ O teste desenvolvido por Kelejian e Robinson não depende do pressuposto de normalidade da distribuição do termo erro, ou da linearidade. O teste baseia-se numa regressão auxiliar *MQO* dos produtos cruzados dos resíduos i e j , espacialmente correlacionados, sobre o produto cruzado das variáveis exógenas x_i e x_j , traduz-se por, $KR = \hat{\gamma}' \mathbf{Z}' \mathbf{Z} \hat{\gamma} / \hat{\sigma}^4$, em que $\hat{\gamma}$ é o vector dos parâmetros estimados a partir da regressão auxiliar e \mathbf{Z} é a matriz contendo os produtos cruzados das variáveis exógenas. O teste *KR* é distribuído assimptoticamente como $\chi_{(z)}^2$, em que z representa o número de variáveis em \mathbf{Z} .

Neste caso particular, a arquitectura da matriz **W** pode desenhar-se a partir da contiguidade localizacional, reflectindo a posição relativa de uma dada observação face às outras no espaço. As medidas de contiguidade pressupõem o conhecimento da dimensão e da forma das observações esboçadas sobre um mapa. A partir daí, podem-se localizar as observações vizinhas (com fronteiras comuns), com graus de dependência espacial superiores aos apresentados por observações localizadas longe umas das outras. A definição de contiguidade espacial para os bairros considerados, na cidade Guarda, teve por base a partilha de um lado comum com o bairro de referência,¹⁶⁷ isto é uma contiguidade tipo “*torre*”, e é passível da ilustração que se segue,

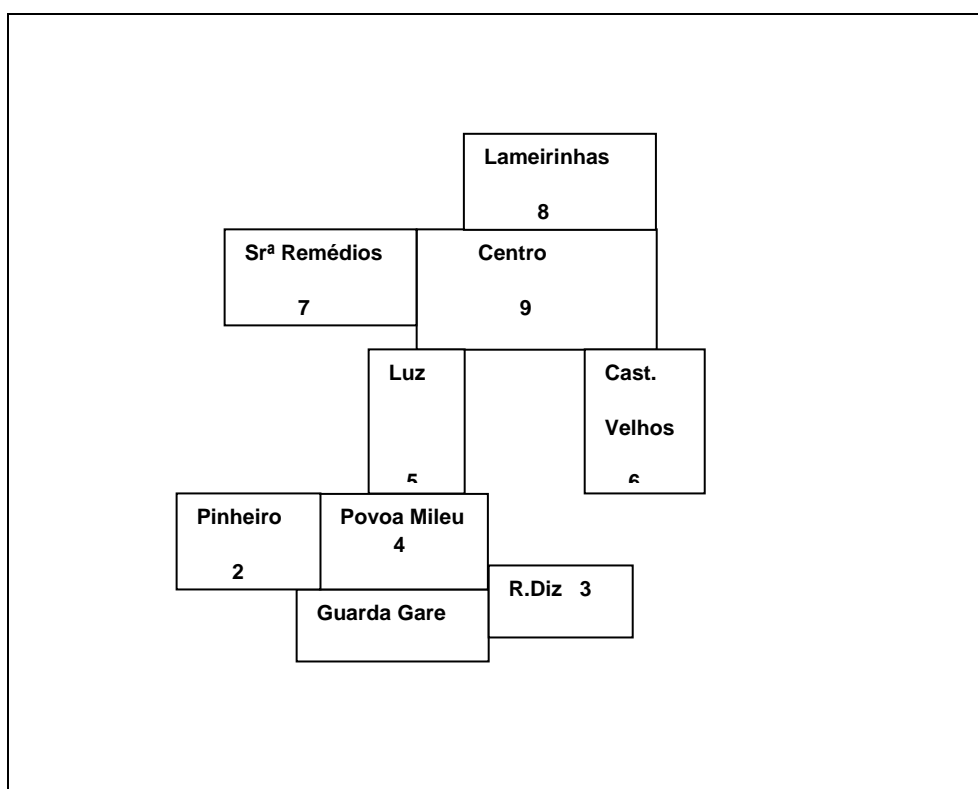


Figura 1: Distribuição espacial dos bairros

¹⁶⁷ Existem inúmeras formas de definir a contiguidade espacial, p.e., se, i) há a partilha de um limite comum, apenas, à direita ou esquerda da região de referência, há contiguidade linear; ii) há partilha de um lado comum com a região de referência, há contiguidade de “*torre*”; iii) há a partilha de um vértice comum com a região de referência, há contiguidade de “*bispo*”; iv) há a partilha de um vértice ou de um lado com a região de referência, há contiguidade de “*rainha*”, por paralelismo com as regras do xadrez. O princípio subjacente à qualificação da contiguidade também deverá ter presente a natureza do problema a ser modelado e, talvez, informações extra-amostra disponíveis (p.e. saber que duas regiões partilhando, apenas, um vértice estão ligadas por uma boa estrada, ou por um eficiente transporte colectivo).

A matriz \mathbf{W} que daí resulta, é conforme,

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

A matriz é simétrica e, por convenção, a respectiva diagonal principal é constituída por zeros. Geralmente, esta matriz é transformada, obrigando a que a soma das suas linhas seja sempre igual a um, passa-se a estar no cenário de uma matriz normalizada.¹⁶⁹ A motivação para esta transformação, uma vez que se assume dependência espacial entre as observações individuais da amostra, reside no facto de se conseguir explicar alguma da variação pela dependência das observações vizinhas e/ou contíguas.

No entanto, num modelo sustentado em **400** observações, distribuídas de forma aleatória pelos diferentes bairros, construir a matriz de ponderação \mathbf{W} não é tarefa fácil, obrigaria à construção de um bloco de matrizes até esgotar todas as observações pelo espaço. A acrescer a esta dificuldade tem-se, ainda, a não incorporação de métodos econométricos espaciais nos pacotes de *software* econométrico comercializados. Estimação de erros espaciais, aparecem no S+SpatialStats acrescentado ao S-Plus, e em

¹⁶⁹ Poder-se-ia apresentar assim, $\mathbf{W} = \begin{bmatrix} 0 & 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 0 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 \end{bmatrix}$

rotinas do Limdep, do Gauss e do Shazam, desenvolvidas por Anselin *et al.* (2000) para o efeito.

Na tentativa de ultrapassar as limitações referidas, a partir de uma ideia desenvolvida por Lykke e Reis (1997), procurou-se uma adaptação para as variáveis observadas de forma a permitir a aplicação dos testes de autocorrelação padrão disponibilizados pelo EVIEWS 4.

A adaptação consiste em reduzir o espaço bidimensional, correspondente à localização dos bairros considerados na cidade da Guarda, num espaço unidimensional, tornando-o semelhante ao espaço temporal. A concretização dessa transformação exige o alinhamento de todos os bairros de acordo com a sua localização e, posteriormente, a respectiva codificação, de um extremo ao outro, seguindo sequencialmente sempre para o bairro mais próximo,¹⁷⁰ assim,

Centro	Lameiri_nhas	Sr ^a Remédios	Castelos Velhos	Luz	Povoa Mileu	Rio Diz	Pinheiro	Guarda Gare
9	8	7	6	5	4	3	2	1

Tabela 4.7-1: Alinhamento dos Bairros

Acautelando esta transformação poder-se-á utilizar o teste estatístico *LM* de Correlação em Série de Breusch-Godfrey como um *proxy* dos testes de correlação espacial.

¹⁷⁰ Este método poderia abrir espaço à ambiguidade, porém a actual codificação postal, já tem implícito esse princípio de Vizinhança numa dimensão.

Entretanto, como já foi referido, a matriz \mathbf{W} deve ser normalizada, garantindo que o somatório dos valores constantes em cada uma das linhas da matriz venha sempre igual a um, o que permitirá reescrever o Índice de Moran,¹⁷¹ expresso por, $I = (N/S_0)(e' \mathbf{W} e / e' e)$, como $I = e' \mathbf{W} e / e' e$, uma vez que $S_0 = \sum_i \sum_j w_{ij} = 1$. Por esta via surge uma similaridade formal com a estatística de Durbin-Watson,¹⁷² passando, também, a apresentar propriedades similares, assintoticamente equivalentes aos LR e LM estatísticos.

Após se terem apresentado argumentos justificativos da adaptação dos testes de autocorrelação em amostras temporais a amostras seccionais, ir-se-ão observar os resultados obtidos para o mercado em análise. A partir do modelo de regressão pelo método dos mínimos quadrados ordinários (MQO), após o alinhamento dos dados da amostra em função da codificação espacial atribuída aos bairros, pode-se ler para o Durbin Watson (DW) estatístico o valor de 1,9981; sobre a hipótese nula de autocorrelação o DW estatístico deveria aproximar-se de 2, o que permitiria concluir pela não existência de associação linear entre os resíduos vizinhos de um modelo de regressão MQO . Assim sendo, estar-se-ia em condições de não rejeitar a hipótese nula de *não correlação em série*.

Existem, no entanto, limitações ao teste DW como teste de correlação em série, que se prendem com, *i*) a dependência da distribuição do DW estatístico dos valores da matriz \mathbf{x} , das variáveis independentes (as características da habitação) e conseqüente limitação das regiões críticas, gerando regiões onde os testes são inconclusivos; *ii*) a existência de variáveis dependentes desfasadas no lado direito do modelo de regressão; *iii*) a hipótese

¹⁷¹ Ver nota 165.

¹⁷² $DW = e' \mathbf{A} e / e' e$, sendo que \mathbf{A} matriz de banda, $-1, 2, -1$.

nula de não correlação em série ser testada contra a hipótese alternativa de correlação em série de 1ª ordem .

Para ultrapassar as limitações reveladas pelo *DW* estatístico, apresenta-se o teste do multiplicador de Lagrange (*LM*) para correlação em série, de Breusch-Godfrey. A hipótese nula deste teste é que não há correlação em série nos resíduos acima da ordem especificada. No presente caso em estudo, uma vez que as relações de contiguidade espacial, poderiam associar, no máximo 4 bairros, optou-se por introduzir um desfasamento (*lag*) no valor de 4, e os resultados obtidos sobre a forma de um “*F*-estatístico” e de um “Nº Observações**R*²” (*NR*²),¹⁷³ foram, respectivamente, 0,1565 e 0,6811. Da leitura das tabelas da distribuição do χ^2 , retira-se o respectivo valor crítico, em função do nível de significância, para 4 graus de liberdade, ora com um $\chi^2_{[4]} = 0,6311$ não se pode rejeitar a hipótese nula deste teste a um nível de significância de 5%, ou seja aceita-se a inexistência de correlação em série acima da ordem 4.¹⁷⁴

Poder-se-ia, ainda, complementar o domínio de testes de correlação em série, com o teste do *Q*-estatístico. O *EVIIEWS* 4, relativamente a este teste, disponibiliza informação sobre as funções de autocorrelação e autocorrelação parcial dos resíduos,

¹⁷³ O *NR*² estatístico tem uma distribuição, assintótica, do χ^2 . A distribuição do *F*-estatístico não é conhecida, mas é muitas vezes utilizada para um teste informal da hipótese nula.

¹⁷⁴ Se em vez de um desfasamento de 4, definido em função do valor máximo de contiguidades espaciais se tivesse introduzido os valores de 1, ou de 2, ou de 3, no teste *LM* de Breusch-Godfrey continuar-se-ia a não rejeitar a hipótese de inexistência de correlação em série acima da ordem 1, ou da ordem 2, ou da ordem 3, respectivamente.

conjuntamente, com Ljung-Box Q -estatísticos¹⁷⁵ e os seus valores de probabilidade. Os resultados obtidos são,

Autocorrelação	ParcialCorrelação		AC	PAC	Q -Est.	Probab.
. .	. .	1	0.037	0.037	0.5414	0.462
. .	. .	2	0.007	0.006	0.5631	0.755
. .	. .	3	0.007	0.007	0.5856	0.900
. .	. .	4	-0.010	-0.010	0.6232	0.960

Tabela 4.7-2: Resultados da Autocorrelação

Da leitura da tabela, constata-se que os valores da autocorrelação e da autocorrelação parcial, são muito próximos do zero para todos os resíduos desfasados considerados. Relativamente aos valores de Q -estatístico, apresentam-se insignificantes, para todos os resíduos considerados, com grandes valores de probabilidade associados. Em suma poder-se-á afirmar que estes resultados são de molde a não rejeitar a hipótese nula, indicando de forma significativa, a ausência de correlação em série nos resíduos.

A conclusão apresentada fundamenta-se no facto do Q -estatístico seguir uma distribuição assintótica do χ^2 , daí a leitura da tabela de resultados, em termos do

¹⁷⁵ O Q -estatístico para k resíduos desfasados testa a hipótese nula de que *não há autocorrelação acima dessa ordem k* e é calculado

da seguinte forma, $Q_{LB} = T(T+2) \sum_{j=1}^k \tau_j^2 / (T-J)$, baseia-se nos, τ_j^2 , quadrados dos primeiros k coeficientes de autocorrelação

dos resíduos MQO , com $\tau_j^2 = \frac{\sum_{t=j+1}^T e_t e_{t-j}}{\sum_{t=1}^T e_t^2}$, e T a representar o número de observações. Sob a hipótese de autocorrelação

nula para os resíduos, Q terá uma distribuição do χ^2 , assintótica, com um número de graus de liberdade igual ao número de autocorrelações.

χ^2 , se traduzir por um $\chi^2_{[1]} = 0,5414$, um $\chi^2_{[2]} = 0,5631$, um $\chi^2_{[3]} = 0,5856$ e um $\chi^2_{[4]} = 0,6232$, que, quando comparados com os respectivos valores críticos, se revelam significativamente menores que esses valores críticos, para elevados níveis convencionais de significância, permitindo, pois, não rejeitar a hipótese nula.

4.8. Estimação Semiparamétrica

O modelo de preços hedónicos, representativo do mercado imobiliário da cidade da Guarda, especificado sob a forma semi-logarítmica, apresentado na secção 4.4., foi escolhido, a partir de um conjunto de formas funcionais alternativas, para a estimação de preços hedónica, através da técnica de validação cruzada, desenvolvida na secção 4.2.. O processo atrás referido teve subjacente técnicas paramétricas, já que os parâmetros da(s) equação(ões) de preços hedónica foram estimados utilizando o método dos mínimos quadrados ordinários.

Ora a estimação paramétrica exige atenção, quer às variáveis independentes e dependente, quer à forma funcional e, ainda, implica assumir que os resíduos seguem uma distribuição independente e idêntica (i.i.d.), com média de zero, variância constante e covariância nula.

Entretanto a teoria económica raramente especifica a forma funcional dos modelos, em vez disso especifica, tipicamente, um conjunto de variáveis potencialmente relacionadas e umas propriedades gerais funcionais da relação; por exemplo a teoria económica permite aduzir sobre **(i)** o impacto, positivo ou negativo, de uma qualquer variável (monotonicidade); **(ii)** a homogeneidade de grau zero e/ou homogeneidade de grau um; **(iii)** o efeito de uma variável explicativa e a não dependência do nível de outra (separabilidade aditiva); **(iv)** certas propriedades de curvatura, tais como concavidade ou convexidade, presentes na relação das variáveis; **(v)** os padrões de comportamento

observados resultantes de processos de optimização sujeitos a restrições.¹⁷⁶ Ora neste contexto, a investigação empírica é imprescindível para avaliar se uma variável ou qualquer outra variável é significativa ou se um conjunto de propriedades particulares são sustentadas pelo modelo - a possível rejeição de hipóteses, daí resultante, pode advir da especificação seleccionada para a forma funcional, não tendo que ser implicada pela teoria económica.¹⁷⁷

Para tornar a especificação da função de regressão robusta pode-se recorrer a modelos não paramétricos que, *a priori*, não implicam restrições relativamente à forma funcional.¹⁷⁸ Os estimadores não paramétricos reduzem o erro de especificação, mas à custa da eficiência: a sua maior fragilidade reside na sua baixa velocidade de convergência. O desempenho destes estimadores decresce à medida que a dimensão do problema aumenta (número de variáveis independentes não-constantes), propriedade identificada, na literatura econométrica, como “*curse of dimensionality*”.¹⁷⁹

Uma vez que a estimação de preços hedónica é suportada em amostras relativamente pequenas e dada a circunstância de as regressões puramente não paramétricas não produzirem resultados superiores, cria-se espaço para incorporar uma parte não paramétrica na convencional regressão hedónica paramétrica. Esta forma semi-paramétrica resultante da inclusão de uma parte não paramétrica numa regressão paramétrica, ao permitir reduzir a dimensão das variáveis independentes não-constantes,

¹⁷⁶ No caso da teoria de preços hedónica algumas destas induções só foram possíveis depois dos artigos de referência de Lancaster (1966) e Rosen (1974), embora os preços hedónicos se reportassem a 1935.

¹⁷⁷ Simplesmente pode acontecer que a função escolhida não traduz a realidade da teoria económica, conduzindo à rejeição incorrecta de restrições.

¹⁷⁸ Anglin e Gençay (1996) referem várias técnicas, “*cubic splines*”, “*nearest neighbours*”, “*series approximators*” e “*Kernel estimates*”, para estimar os modelos não paramétricos.

¹⁷⁹ Yatchew (1998) identifica a distância entre N observações distribuídas uniformemente num espaço unidimensional, como $1/N$; como $1/N^{1/2}$ se essas N -observações estiverem distribuídas uniformemente num espaço bi-dimensional e como $1/N^{1/K}$ caso se alargue a um espaço K -dimensional. O erro de aproximação é função da distância entre observações, à medida que a densidade de pontos aumenta (i.e. à medida que a distância entre pontos se reduz) o erro de aproximação reduz-se à taxa $O(N^{1/K})$. Por exemplo, para $N = 400$, esse erro de aproximação vem 20 e cerca de 220 vezes superior, ao erro de aproximação num espaço unidimensional, respectivamente para os espaços bi-dimensional e décimo-dimensional. O erro de aproximação pode vir, substancialmente, reduzido, aplicando algumas restrições ao modelo não paramétrico, tais como pressupostos de alisamento baseados em parâmetros de alisamento, denominados por “*bandwith*” ou “*window*”, e, ainda, pressupostos de estrutura linear parcial e de separabilidade aditiva.

aumenta a eficiência da estimação. Os estimadores semi-paramétricos combinam a flexibilidade da regressão não paramétrica com a eficiência dos modelos paramétricos padrão. Anglin e Gençay (1996), Pace (1995), Thorsnes e McMillen (1998), Diáz, Gómez e Alvarez (2003), Park e Jyoung (2002) apresentam vários exemplos de modelos semi-paramétricos para os preços das casas. O modelo semi-paramétrico pode-se expressar por,

$$\ln \mathbf{p} = \mathbf{x}\boldsymbol{\beta} + f(\mathbf{z}) + \boldsymbol{\varepsilon}$$

em que $\ln \mathbf{p}$ é vector, $400 * 1$, do logaritmo do preço da habitação, $\mathbf{x}\boldsymbol{\beta}$ é a parte paramétrica da estimação, em que \mathbf{x} é a matriz $400 * 24$ das características paramétricas da habitação, e $f(\mathbf{z})$ é a parte não paramétrica, é função de \mathbf{z}_k , uma matriz de $400 * 4$ das características não paramétricas das habitações, e assume-se como uma função contínua e alisada. Os termos de erro $\boldsymbol{\varepsilon}$, constituem um vector $400 * 1$ de elementos de perturbação, assumem-se como independentes, mas podem ser heteroscedásticos e terem média diferente de zero. Nesta especificação optou-se por considerar as variáveis *dummy*, dicotómicas, como integrantes da parte paramétrica, uma vez que se fossem incluídas na função, $f(\mathbf{z})$, nunca afectariam a curvatura dessa função, induzindo, apenas, efeitos de escala. Além disso as variáveis dicotómicas são modeladas de forma aditiva comparativamente a todas as outras variáveis discretas que devem ser modeladas na função desconhecida, $f(\mathbf{z})$.¹⁸⁰ Esta classificação de variáveis favorece o modelo semi-paramétrico, permitindo, simultaneamente, explorar a generalidade da especificação não paramétrica e a eficiência das restrições paramétricas.

O procedimento para a estimação semi-paramétrica, seguindo Robinson (1988), concretiza-se por, (i) estimar, num primeiro estágio, por via não paramétrica, os valores esperados condicionados por \mathbf{z} do logaritmo do preço de habitação $E(\ln \mathbf{p} | \mathbf{z})$, e das

¹⁸⁰ Conforme é referido por Anglin e Gençay (1996), pp.640.

características dicotómicas da habitação, $E(\mathbf{x} | \mathbf{z})$, retendo-se os valores estimados $\widehat{\ln \mathbf{p}}$ e $\widehat{\mathbf{x}}$; **(ii)** aplicar, num segundo estágio, o método dos mínimos quadrados ordinários à regressão estabelecida entre os resíduos, $\boldsymbol{\varepsilon}_p (\equiv (\ln \mathbf{p} - \widehat{\ln \mathbf{p}}))$ e $\boldsymbol{\varepsilon}_x (\equiv (\mathbf{x} - \widehat{\mathbf{x}}))$, estimando-se os coeficientes β da regressão.

Para a estimação da densidade não paramétrica¹⁸¹ utiliza-se um procedimento de kernel,¹⁸² assim para se estimar $E(\ln \mathbf{p} | \mathbf{z})$ e $E(\mathbf{x} | \mathbf{z})$, vem,

$$E(\ln \mathbf{p}_n | \mathbf{z}_n) \equiv \frac{\sum_{n' \neq n}^N k_h(\mathbf{z}_{kn} - \mathbf{z}_{kn'}) * \ln \mathbf{p}_{n'}}{\sum_{n' \neq n}^N k_h(\mathbf{z}_{kn} - \mathbf{z}_{kn'})}$$

e

$$E(\mathbf{x}_n | \mathbf{z}_n) \equiv \frac{\sum_{n' \neq n}^N k_h(\mathbf{z}_{kn} - \mathbf{z}_{kn'}) * x_{n'}}{\sum_{n' \neq n}^N k_h(\mathbf{z}_{kn} - \mathbf{z}_{kn'})}$$

Em que \mathbf{z}_n é uma determinada observação, um valor particular de uma observação, por exemplo $CBANHO / 2$, o denominado ponto de interesse, *point of interest*, e $\mathbf{z}_{n'}$ as restantes outras observações. Sendo que,

¹⁸¹ A técnica de estimação não paramétrica da densidade está, proximamente, relacionada com o histograma de uma amostra, atribuindo um peso de 1 ou de 0 às observações, conforme caem ou não num dado intervalo. Para uma única variável x , a densidade estimada num ponto x_0 , relativamente ao histograma, $\widehat{pdf}_0(x)_h$, simplesmente, conta o número de observações individuais x_i que caem dentro de cada intervalo.

¹⁸² As funções de kernel positivas devem preencher três requisitos, (1) valor esperado finito; (2) valores não negativos para a relação entre um qualquer ponto do domínio e um valor particular e (3) a função de núcleo ser unitária, isto é $\int_{-\infty}^{+\infty} K(x)dx = 1$, $K(x)$ é a função de kernel para um valor particular x .

$$k_h(\mathbf{z}_n - \mathbf{z}_{n'}) = K\left(\frac{\mathbf{z}_n - \mathbf{z}_{n'}}{h}\right), \text{ onde } K(\cdot) \text{ é a função de kernel e } h$$

é a amplitude do intervalo de frequências.

Para a estimação utiliza-se uma função padrão normal multivariada de kernel, que tem implícito o cálculo da decomposição de Choleski da inversa da matriz de covariância da amostra para as variáveis não paramétricas z , denominada \mathbf{S} , de forma a construir-se um vector $\mathbf{v} \equiv \mathbf{S}\mathbf{z}$, tal que,

$$K\left(\frac{\mathbf{z}_{kn} - \mathbf{z}_{kn'}}{h}\right) = \prod_{z=1}^m \phi\left(\frac{\mathbf{v}_{zkn} - \mathbf{v}_{zkn'}}{h}\right),$$

em que ϕ é a função de densidade normal padrão e m é o número de elementos z (que, no caso em análise, é de 4).¹⁸³

Para escolher o valor de h , a amplitude do intervalo de frequências,¹⁸⁴ utilizou-se a proposta de Scott (1992), que tem subjacente a relação inversa entre variância e enviesamento, à medida que um maior número de observações acelera a convergência, daí que mais observações exijam maior amplitude do intervalo e vice-versa.¹⁸⁵

¹⁸³ O estimador de kernel permite a interacção entre as m variáveis que constituem a matriz z das variáveis não paramétricas, já que a densidade normal multivariada de kernel tem em conta as correlações entre as variáveis. O parâmetro h , controla o peso relativo dado às observações vizinhas – se h é pequeno o peso das observações vizinhas é maior relativamente à situação em que h é maior.

¹⁸⁴ No contexto da avaliação imobiliária a estimação de h , controla em que extensão as observações da amostra afectam o valor previsto. No caso de uma estimação de kernel, baseada numa *pdf* normal, as observações extremamente distantes do ponto de interesse, pouco contribuirão para o valor previsto.

¹⁸⁵ Um desvio padrão maior e um maior número de variáveis não constantes (não paramétricas) também aumentam a dimensão do intervalo de frequências.

$$h_{n'} = \omega \sigma_{n'} N^{\frac{-1}{(4+\rho)}},$$

onde, $\omega = 1$, para uma estimação de kernel baseada numa função de densidade da probabilidade, *pdf*, normal;

$\omega = 0,381$, para uma estimação baseada numa função de densidade da probabilidade, *pdf*, bponderada de Tukey;

$\sigma_{n'} = 1$, para variáveis dependentes estandardizadas,¹⁸⁶

N é o número de observações

e

ρ é o número de variáveis não constantes.

Para uma estimação de kernel baseada numa função de densidade da probabilidade, *pdf*, normal, com variáveis não estandardizadas, a formulação de Scott para a dimensão

óptima dos intervalos de frequência, virá, $h_{n'} = \sigma_{n'} N^{\frac{-1}{(4+\rho)}}$.

Com base nas premissas, anteriormente referidas, modelou-se o mercado da habitação da Guarda em termos semi-paramétricos, elegendo-se para variáveis não constantes, não paramétricas, as variáveis “ANOS”, “AREA”, “CBANHO” e “QUARTOS”, ou seja todas as variáveis da amostra que não são dicotómicas, e utilizando o EVIEWS 4 adaptou-se um programa para o efeito, constante do apêndice **H**.

O coeficiente de correlação múltipla R^2 , que mede a *bondade do ajustamento*, é obtido, no modelo semiparamétrico, ajustando uma regressão entre o valor esperado e o valor

¹⁸⁶ Se as variáveis não estiverem estandardizadas, a amplitude do intervalo de frequências para cada regressor deverá ser ajustado pelo respectivo desvio padrão, o que significará para os preços da habitação previstos, e para os preços implícitos uma não afectação pelas unidades de medida.

observado, entre $\widehat{\ln \mathbf{p}}$ e $\ln \mathbf{p}$, ou seja $R^2 = \widehat{\ln \mathbf{p}}' \widehat{\ln \mathbf{p}} / \ln \mathbf{p}' \ln \mathbf{p}$. Este pseudo- R^2 é idêntico ao R^2 convencional num modelo paramétrico linear.

Neste caso o R^2 da estimação semi-paramétrica é 0,6252, quase idêntico ao R^2 da estimação paramétrica, 0,62582. No entanto se compararmos os erros MAE (erro absoluto médio) e $RMSE$ (raiz quadrada do erro quadrado médio), suportados em previsões fora da amostra (*out of sample*), a vantagem tenderá para o modelo semi-paramétrico, $MAE_{SP} = 17,3767$ e $RMSE_{SP} = 187,0813$, versus $MAE_{MQO} = 8452,5122$ e $RMSE_{MQO} = 10325,5086$.

Para se compararem os modelos paramétricos e semi-paramétricos recorreu-se ao teste de Hausman (1978), para verificar se a função das variáveis não constantes, $f(\mathbf{z})$, é estatisticamente diferente da especificação paramétrica nula, ou seja para testar a linearidade de $f(\mathbf{z})$, implicando a rejeição da hipótese nula a inconsistência da estimação pelos mínimos quadrados ordinários. Assim, caso $\hat{\boldsymbol{\beta}}_{SP}$ e $\hat{\boldsymbol{\beta}}_{MQO}$ sejam os $(K-1) * 1$ vectores dos coeficientes estimados,¹⁸⁷ semi-paramétricos e paramétricos, respectivamente; e $\widehat{\mathbf{V}}_{SP}$ e $\widehat{\mathbf{V}}_{MQO}$ sejam as matrizes de covariância dos coeficientes estimados, o teste estatístico expressa-se por,

$$H = (\hat{\boldsymbol{\beta}}_{SP} - \hat{\boldsymbol{\beta}}_{MQO})' (\widehat{\mathbf{V}}_{SP} - \widehat{\mathbf{V}}_{MQO})^{-1} (\hat{\boldsymbol{\beta}}_{SP} - \hat{\boldsymbol{\beta}}_{MQO}),$$

onde H segue uma distribuição do χ^2 , com $K-1$ graus de liberdade.

¹⁸⁷ O termo independente, de intersecção, não é incluído.

O valor de H é de, 187,0813. Sendo superior ao valor crítico do $\chi^2_{[4]}$, rejeita-se pois a hipótese nula da linearidade. Em conformidade com o teste de Hausman, espera-se a não linearidade da função de preços hedónica do mercado imobiliário da Guarda, ainda que a teoria económica ofereça poucas orientações relativamente à não linearidade. A perspectiva semi-paramétrica permite preservar os graus de liberdade da estimação, mas aceitando um não conhecimento, com segurança, da forma funcional.

Apêndice C

Valor de Mercado	Ano de Venda	Ano Construção	Valor de Venda	Localização						Área	Tipologia					Nº Casas de Banho	Garagem	Gás Canalizado	Aquecimento Central	Janelas	Lareira	Piso				Elevador	Arrecadação	Exposição Solar																	
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má
				Centro	Luz	Sra Remed	Povoa Mileu	Guarda Gare	Pinheiro	Lameirinhas	Cast. Velhos	Rio Diz		T ₀	T ₁	T ₂	T ₃	T ₄	T ₅	1	2	3	S ₁	S ₂	N	S	N	S	N	Simplex	Duplas	S	N	cave	R/chão	1	2	3	4 ou +	S	N	S	N	Boa	Má

No sentido de se alargar a amostra de dados e consequente agregação, bem como garantir o propósito da investigação, o período temporal da informação incidirá sobre os anos 2000, 2001, 2002.

Apêndice D

1. Índice das variáveis utilizadas

<i>ANOS</i>	Variável quantitativa: Idade medida em anos
<i>AQCENTRAL</i>	Variável <i>dummy</i> : 1 se há aquecimento central, 0 se não
<i>AREA</i>	Variável quantitativa: área medida em m ²
<i>ARRECA</i>	Variável <i>dummy</i> : 1 se dispõe de arrecadação, 0 se não
<i>CBANHO</i>	Variável quantitativa: número de casas de banho
<i>CV</i>	Variável <i>dummy</i> : 1 se o piso é cave, 0 se não
<i>ELEV</i>	Variável <i>dummy</i> : 1 se dispõe de elevador, 0 se não
<i>EXPSOL</i>	Variável <i>dummy</i> : 1 se exposição solar boa, 0 se não
<i>GCANAL</i>	Variável <i>dummy</i> : 1 se dispõe de gás canalizado, 0 se não
<i>GSDOS</i>	Variável <i>dummy</i> : 1 se dispõe de duas garagens, 0 se não
<i>GSUM</i>	Variável <i>dummy</i> : 1 se dispõe de uma garagem, 0 se não
<i>JANDUPLAS</i>	Variável <i>dummy</i> : 1 se dispõe de janelas duplas, 0 se não
<i>LAR</i>	Variável <i>dummy</i> : 1 se dispõe de lareira, 0 se não
<i>LCNTRO</i>	Variável <i>dummy</i> : 1 se está localizada no Centro, 0 se não
<i>LCVLHOS</i>	Variável <i>dummy</i> : 1 se está localizada nos CastelosVelhos, 0 se não
<i>LGGARE</i>	Variável <i>dummy</i> : 1 se está localizada na Guarda Gare, 0 se não
<i>LLAMNH</i>	Variável <i>dummy</i> : 1 se está localizada nas Lameirinhas, 0 se não
<i>LLUZ</i>	Variável <i>dummy</i> : 1 se está localizada no Bairro Luz, 0 se não
<i>LNVMercado</i>	Variável quantitativa: Logaritmo do Valor de Mercado em euros
<i>LPMILEU</i>	Variável <i>dummy</i> : 1 se está localizada na Póvoa do Mileu, 0 se não
<i>LPNHR</i>	Variável <i>dummy</i> : 1 se está localizada no Bairro do Pinheiro, 0 se não
<i>LRDIZ</i>	Variável <i>dummy</i> : 1 se está localizada no Rio Diz, 0 se não
<i>LSREMD</i>	Variável <i>dummy</i> : 1 se está localizada na Sra Remédios, 0 se não
<i>PPRIM</i>	Variável <i>dummy</i> : 1 se o piso é 1º andar, 0 se não
<i>PQRT</i>	Variável <i>dummy</i> : 1 se o piso é 4º andar, 0 se não
<i>PSEG</i>	Variável <i>dummy</i> : 1 se o piso é 2º andar, 0 se não
<i>PTERC</i>	Variável <i>dummy</i> : 1 se o piso é 3º andar, 0 se não
<i>QUARTOS</i>	Variável quantitativa: número de quartos
<i>RCH</i>	Variável <i>dummy</i> : 1 se o piso é rés-chão, 0 se não

2. Estatística descritiva das variáveis utilizadas

	Média	Máximo	Mínimo	Desvio Padrão	Nº de Observações
<i>ANOS</i>	4,94	30	1	6,05	400
<i>AQCENTRAL</i>	0,51	1	0	0,5	400
<i>AREA</i>	114,34	300	42	30,97	400
<i>QUARTO</i>	2,67	5	1	0,77	400
<i>CBANHO</i>	1,72	3	1	0,5	400
<i>ARRECA</i>	0,89	1	0	0,31	400
<i>ELEV</i>	0,43	1	0	0,5	400
<i>EXPSOL</i>	0,84	1	0	0,37	400
<i>GCANAL</i>	0,75	1	0	0,43	400
<i>GSDOS</i>	0,05	1	0	0,22	400
<i>GSUM</i>	0,74	1	0	0,44	400
<i>JANDUPLAS</i>	0,58	1	0	0,49	400
<i>LAR</i>	0,72	1	0	0,45	400
<i>LCNTRO</i>	0,27	1	0	0,45	400
<i>LCVLHOS</i>	0,03	1	0	0,18	400
<i>LGGARE</i>	0,23	1	0	0,42	400
<i>LLAMNH</i>	0,1	1	0	0,3	400
<i>LLUZ</i>	0,13	1	0	0,34	400
<i>LPMILEU</i>	0,05	1	0	0,23	400
<i>LPNHR</i>	0,04	1	0	0,18	400
<i>LRDIZ</i>	0,01	1	0	0,11	400
<i>LSREMD</i>	0,12	1	0	0,32	400
<i>CV</i>	0,09	1	0	0,29	400
<i>RCH</i>	0,15	1	0	0,35	400
<i>PPRIM</i>	0,25	1	0	0,43	400
<i>PQRT</i>	0,06	1	0	0,23	400
<i>PSEG</i>	0,25	1	0	0,43	400
<i>PTERC</i>	0,2	1	0	0,4	400

Apêndice E

1. Regressão linear, método dos mínimos quadrados, com aplicação do EVIEWS 4

```
scalar i
scalar j
scalar u
scalar u2=0
scalar mae=0
scalar average_mae=0
scalar mse=0
scalar rmse=0
scalar average_rmse=0
scalar average_u=0
for i=1 to 1000
  smpl 1 400
  series order=rnd
  sort order
  smpl 1 350
  equation model.ls VMERCADO C ANOS AQCENTRAL AREA QUARTOS CBANHO ARRECA
  ELEV EXP SOL GCANAL GSDOS GSUM JANDUPLAS LAR LCNTRO LCVLHOS LGGARE
  LLAMNH LLUZ LPMILEU LPNHR LRDIZ LSREMD CV RCH PPRIM PQRT PSEG PTERC
  smpl 351 400
  model.fit vmercadof
  mae=0
  mse=0
  u2=0
  for j=351 to 400
    mae=mae+abs(vmercadof(j)-vmercado(j))/50
    mse=mse+((vmercadof(j)-vmercado(j)))^2/50
    u2=u2+((vmercado(j))^2/50)^0.5
  next j
  average_mae=average_mae+mae/1000
  rmse=mse^0.5
  u=rmse/u2
  average_rmse=average_rmse+rmse/1000
  average_u=average_u+u/1000
next i
show average_mae
show average_rmse
show average_u
```

2. Modelo Box-Cox Quadrático, utilizando EVIEWS 4.

```
scalar i
scalar j
scalar k
scalar imax
scalar jmax
```

continua

```

scalar kmax
scalar sumlogvmercado
scalar loglik
scalar loglikmax=-1e16
matrix(1,1) msumlogvmercado
rowvector one=@filledrowvector(400,1)
msumlogvmercado=one*@convert(log(vmercado))
sumlogvmercado=msumlogvmercado(1,1)
for i=-1.95 to 1.95 step .1
for j=-1.95 to 1.95 step .1
for k=-1.95 to 1.95 step .1
equation BoxCox.ls (vmercado^i-1)/i (quartos^j-1)/j (cbanho^j-1)/j (anos^k-1)/k (area^j-1)/j
0.5*((quartos^j-1)/j)^2 0.5*((cbanho^j-1)/j)^2 0.5*((anos^k-1)/k)^2 0.5*((area^j-1)/j)^2 ((quartos^j-
1)/j)*((cbanho^j-1)/j) ((anos^k-1)/k)*((area^j-1)/j) ((quartos^j-1)/j)*((anos^k-1)/k) ((area^j-
1)/j)*((cbanho^j-1)/j) ((quartos^j-1)/j)*((area^j-1)/j) ((cbanho^j-1)/j)*((anos^k-1)/k) gsum gsdos lar cv rch
PPRIM PQRT PSEG PTERC LGGARE LCNTRO LCVLHOS LLAMNH LLUZ ELEV EXPSOL
loglik=-0.5*BoxCox.@regobs*log(BoxCox.@se^2)+(i-1)*sumlogvmercado
if loglik>loglikmax then
imax=i
jmax=j
kmax=k
loglikmax=loglik
endif
next k
next j
next i
show imax jmax kmax loglikmax
scalar mae
scalar average_mae=0
scalar mse
scalar rmse
scalar average_rmse=0
scalar u
scalar u2
scalar average_u=0
for i=1 to 1000
smp1 1 400
series order=rnd
sort order
smp1 1 350
equation BoxCox.ls (vmercado^imax-1)/imax (quartos^jmax-1)/jmax (cbanho^jmax-1)/jmax (anos^kmax-
1)/kmax (area^jmax-1)/jmax 0.5*((quartos^jmax-1)/jmax)^2 0.5*((cbanho^jmax-1)/jmax)^2
0.5*((anos^kmax-1)/kmax)^2 0.5*((area^jmax-1)/jmax)^2 ((quartos^jmax-1)/jmax)*((cbanho^jmax-
1)/jmax) ((quartos^jmax-1)/jmax)*((area^jmax-1)/jmax) ((anos^kmax-1)/kmax)*((area^jmax-1)/jmax)
((quartos^jmax-1)/jmax)*((anos^kmax-1)/kmax) ((area^jmax-1)/jmax)*((cbanho^jmax-1)/jmax) gsum gsdos lar
cv rch PPRIM PQRT PSEG PTERC LGGARE LCNTRO LCVLHOS LLAMNH LLUZ ELEV EXPSOL
smp1 351 400
BoxCox.fit vmercadof
mae=0
mse=0
u2=0
for j=351 to 400
mae=mae+abs(vmercadof(j)-vmercado(j))/50
mse=mse+((vmercadof(j)-vmercado(j)))^2/50
u2=u2+((e^lnvmercado(j))^2/50)^0.5
next j

```

continua

```
average_mae=average_mae+mae/1000
rmse=mse^0.5
u=rmse/u2
average_rmse=average_rmse+rmse/1000
average_u=average_u+u/1000
next i
show average_mae
show average_rmse
show average_u
```


Apêndice G

TABELA 1

Regressões Auxiliares e Indicadores de Multicolinearidade

Variável dependente	R-quadrado	R-quadrado Ajustado	VIF	F - estatístico
<i>ANOS</i>	0,43817	0,39739	1,7799	10,334
<i>AQCENRAL</i>	0,46397	0,42506	1,8656	11,468
<i>AREA</i>	0,88071	0,87205	8,3827	89,613
<i>QUARTOS</i>	0,86978	0,86033	7,6793	88,501
<i>CBANHO</i>	0,74671	0,72833	3,9481	39,062
<i>ARRECA</i>	0,19138	0,13269	1,2367	3,145
<i>ELEV</i>	0,38976	0,34546	1,6387	8,4627
<i>EXPSOL</i>	0,07736	0,01039	1,0838	1,111
<i>GCANAL</i>	0,05291	-0,01583	1,0559	0,740
<i>GSDOS</i>	0,46436	0,42548	1,8669	11,487
<i>GSUM</i>	0,46115	0,42204	1,8558	11,339
<i>JANDUPLAS</i>	0,39697	0,35321	1,6583	8,722
<i>LAR</i>	0,41369	0,37114	1,7056	9,349
<i>LCNTRO</i>	0,50572	0,46984	2,0231	13,557
<i>LCVLHOS</i>	0,27590	0,22335	1,3810	5,049
<i>LGGARE</i>	0,39396	0,34997	1,6501	8,613
<i>LLAMNH</i>	0,26832	0,21522	1,3667	4,859
<i>LLUZ</i>	0,31432	0,26455	1,4584	6,074
<i>LPMILEU</i>	0,07645	0,00941	1,0828	1,097
<i>LPNHR</i>	0,06041	-0,0078	1,0643	0,852
<i>LRDIZ</i>	0,07130	0,00389	1,0767	1,017
<i>LSREMD</i>	0,07339	0,00614	1,0792	1,053
<i>CV</i>	0,76812	0,75129	4,3125	43,892
<i>RCH</i>	0,82659	0,8140	5,7667	63,159
<i>PPRIM</i>	0,86768	0,8581	7,5572	80,126
<i>PSEG</i>	0,87687	0,8679	8,1217	94,359
<i>PTERC</i>	0,84732	0,8362	6,5494	73,533
<i>PQRT</i>	0,67293	0,6492	3,0574	27,261

TABELA 2**Valores próprios, Números e Índices de Condição associados**

Nº	Valor Próprio λ	Número de Condição	Índice de Condição
1	0.00000	13564256.88682	3682.96849
2	0.00006	38750.14889	196.85058
3	0.00034	6949.81330	83.36554
4	0.00593	403.27104	20.08161
5	0.00855	279.49873	16.71822
6	0.00978	244.37184	15.63240
7	0.01077	221.89440	14.89612
8	0.01199	199.30133	14.11741
9	0.01394	171.39947	13.09196
10	0.01517	157.57887	12.55304
11	0.01642	145.52921	12.06355
12	0.01868	127.94504	11.31128
13	0.01986	120.34886	10.97036
14	0.02097	113.99822	10.67699
15	0.02304	103.71256	10.18394
16	0.02536	94.22764	9.70709
17	0.02564	93.21023	9.65454
18	0.03172	75.35741	8.68086
19	0.03251	73.51374	8.57402
20	0.03984	59.99056	7.74536
21	0.04523	52.83909	7.26905
22	0.04790	49.89741	7.06381
23	0.05230	45.69645	6.75992
24	0.06588	36.28054	6.02333
25	0.09102	26.25902	5.12436
26	0.10883	21.96043	4.68620
27	0.19236	12.42462	3.52486
28	0.58149	4.11010	2.02734
29	2.39493	0.99794	0.99897

Apêndice H

ESTIMAÇÃO SEMIPARAMÉTRICA UTILIZANDO O EIEWS 4.

```
rowvector one=@filledrowvector(350,1)
scalar h1=6.0478*400^(-1/8)
scalar h2=30.972*400^(-1/8)
scalar h3= 0.498*400^(-1/8)
scalar h4= 0.773*400^(-1/8)
scalar anos0
scalar area0
scalar cbanho0
scalar quartos0
scalar t
scalar i
scalar j
scalar z
scalar w
scalar ksum
scalar predict
scalar uanos
scalar uarea
scalar ucbanho
scalar uquartos
series k
series variable
matrix(1,1) mksum
matrix(1,1) mpredict
for z=1 to 26
if z=1 then
variable=lnvmercado
series lnvmercadof
endif
if z=2 then
variable=aqcentral
series aqcentralf
```



```
endif
if z=3 then
variable=arreca
series arrecaf
endif
if z=4 then
variable=elev
series elevf
endif
if z=5 then
variable=expsol
series expsol
endif
if z=6 then
variable=gcanal
series gcanalf
endif
if z=7 then
variable=gsdos
series gsdosf
endif
if z=8 then
variable=gsum
series gsumf
endif
if z=9 then
variable=janduplas
series janduplasf
endif
if z=10 then
variable=lar
series larf
endif
if z=11 then
variable=lcntro
series lcntrof
endif
```

```
if z=12 then
variable=lcvlhos
series lcvlhosf
endif
if z=13 then
variable=lggare
series lggaref
endif
if z=14 then
variable=llamnh
series llamnhf
endif
if z=15 then
variable=lluz
series lluzf
endif
if z=16 then
variable=lpmileu
series lpmileuf
endif
if z=17 then
variable=lpnhr
series lpnhrf
endif
if z=18 then
variable=lrdiz
series lrdizf
endif
if z=19 then
variable=lsremd
series lsremdf
endif
if z=20 then
variable=pprim
series pprimf
endif
if z=21 then
```

```

variable=pseg
series psegf
endif
if z=22 then
variable=pterc
series ptercf
endif
if z=23 then
variable=pqrt
series pqrtf
endif
if z=24 then
variable=rch
series rchf
endif
if z=25 then
variable=cv
series cvf
endif
if z=26 then
for t=1 to 500
smp1 1 400
series order=rnd
sort order
smp1 1 350
equation linear.ls (Invmercado- Invmercado) (aqcentral-aqcentral) (arrecas-arrecas) (elev -elev) (expsol-
expsol) (gcanal- gcanal) (gsdos-gsdos) (gsum-gsum) (janduplas- janduplas) (lar- lar) (lcntro-lcntro)
(lcvlhos- lcvlhos) (lggare-lggare) (llamnh- llamnh) (lluz-lluz) (lpmileu-lpmileu) (lpmhr- lpmhr) (lrdiz-
lrdiz) (lsremd-lsremd) (pprim-pprim) (pseg- pseg) (pterc-pterc) (pqrt- pqrt) (rch- rch) (cv-cv)
for j=1 to 350
variable (j) = Invmercado(j) - c(1)*aqcentral(j) - c(2)*arrecas(j)-c(3)*elev(j) -c(4)*expsol(j)-c(5)*gcanal(j)-
c(6)*gsdos(j)-c(7)*gsum(j)-c(8)*janduplas(j)- c(9)*lar(j)-c(10)*lcntro(j)-c(11)*lcvlhos(j)- c(12)*lggare(j)-
c(13)*llamnh(j)- c(14)*lluz(j)- c(15)*lpmileu(j)-c(16)*lpmhr(j)-c(17)*lrdiz(j)- c(18)*lsremd(j)-
c(19)*pprim(j)-c(20)*pseg(j)- c(21)*pterc(j)-c(22)*pqrt(j)- c(23)*rch(j)-c(24)*cv(j)
next t
next j
series an_ar_cb_qtf

```

```

endif
for j=351 to 400
scalar u
scalar u2=0
scalar mae=0
scalar average_mae=0
scalar mse=0
scalar rmse=0
scalar average_rmse=0
scalar average_u=0
mae=mae+abs(e^lnvmercadof(j)-e^lnvmercado(j))/50
mse=mse+((e^lnvmercadof(j)-e^lnvmercado(j)))^2/50
u2=u2+((e^lnvmercado(j))^2/50)^0.5
rmse=mse^0.5
u=rmse/u2
next j
for i=1 to 350
anos0=anos(i)
area0=area(i)
cbanho0=cbanho(i)
quartos0=quartos(i)
k=@dnorm((anos0-anos)/h1)*@dnorm((area0-area)/h2)*@dnorm((cbanho0-
cbanho)/h3)*@dnorm((quartos0-quartos)/h4)
mksum=one*@convert(k)
ksum=mksum(1,1)
mpredict=@transpose(@convert(variable))*@convert(k)/ksum
predict=mpredict(1,1)
if z=1 then
lnvmercadof(i)=predict
endif
if z=2 then
aqcentralf(i)=predict
endif
if z=3 then
arrecaf(i)=predict
endif
if z=4 then

```

```
elevf(i)=predict
endif
if z=5 then
expsof(i)=predict
endif
if z=6 then
gcanalf(i)=predict
endif
if z=7 then
gsdosf(i)=predict
endif
if z=8 then
gsumf(i)=predict
endif
if z=9 then
janduplasf(i)=predict
endif
if z=10 then
larf(i)=predict
endif
if z=11 then
lcntrof(i)=predict
endif
if z=12 then
lcvlhosf(i)=predict
endif
if z=13 then
lggaref(i)=predict
endif
if z=14 then
llamnhf(i)=predict
endif
if z=15 then
lluzf(i)=predict
endif
if z=16 then
lpmileuf(i)=predict
```

```
endif
if z=17 then
lpnhf(i)=predict
endif
if z=18 then
lrdizf(i)=predict
endif
if z=19 then
lsremdf(i)=predict
endif
if z=20 then
pprimf(i)=predict
endif
if z=21 then
psegf(i)=predict
endif
if z=22 then
ptercf(i)=predict
endif
if z=23 then
pqrtf(i)=predict
endif
if z=24 then
rchf(i)=predict
endif
if z=25 then
cvf(i)=predict
endif
if z=26 then
an_ar_cb_qtf(i)=predict
endif
next i
next z
```

5. Do Modelo de Avaliação Base deste Ensaio ao Modelo Saído do IMI

Science is intimately bound up with the story of measurement; each drives the other forward. But measurement, either conscious or instinctive, is a part of every human activity—choosing the right color, drawing in perspective, correctly valuing a house for sale, or fitting the correct number of syllables in a line of poetry.

In a Measure of Everything (2005)

5.1. Considerações Prévias

O objectivo deste ensaio relativo à tributação do património tem por alvo a base de incidência do imposto sobre o património, atendendo, a preocupações de justiça, à percepção de justiça e de equidade, a questões de eficiência fiscal, num imposto que, tal como outros, está longe de ser perfeito. Neste contexto procurou-se uma metodologia de avaliação da base fiscal, que tivesse em conta as características específicas da propriedade, incluindo a localização, de forma a conseguir-se um sistema intermédio de avaliação entre os que se sustentam nos valores de mercado e os que o ignoram.

A análise e modelização do mercado da habitação, na cidade da Guarda, para o ano 2002, desenvolvida nos capítulos anteriores, permitem uma avaliação das propriedades imobiliárias desta cidade, fazendo pois crer na concretização do objectivo proposto. Importa, no entanto, avaliar o rigor e a credibilidade de tal avaliação, para tanto há que

comparar os valores resultantes desse processo de avaliação com os valores do processo de avaliação decorrente da aplicação do Código do Imposto Municipal Imobiliário (CIMI).

Acontece que a aplicação das novas regras de avaliação¹⁸⁸ previstas no quadro do IMI, mesmo depois de publicada a Portaria 982/2004, de 4 de Agosto, que aprova e publicita os coeficientes dentro dos limites estabelecidos no CIMI, o custo médio de construção e os coeficientes de capitalização da renda anual, não abrangerá automaticamente todo o universo de prédios urbanos existentes no país; ficando restrita, inicialmente, às propriedades imobiliárias transaccionadas ou àquelas cujos proprietários requeiram a respectiva avaliação. Mais, o CIMI avança com a data de 2013 como prazo optimista para a conclusão do processo de avaliação alargado a todo o universo da propriedade imobiliária urbana.

Neste cenário e procurando contemporizar a dinâmica de avaliação contemplada na legislação relativa à reforma do património, procurou-se, *numa primeira fase*, comparar os valores da avaliação, decorrentes do modelo preconizado no presente ensaio, com os valores da avaliação resultantes da aplicação do regime transitório de avaliação, constante da Portaria nº 1337/2003, de 5 de Dezembro¹⁸⁹ Posteriormente, *numa segunda fase*, e na sequência da publicação da Portaria nº 982/2004,¹⁹⁰ procurar-se-á simular um valor de avaliação patrimonial para as propriedades analisadas, conforme preconizado

¹⁸⁸ O articulado constante dos artigos 38º a 46º, secção II (Das operações de avaliação), do decreto-lei que institui o CIMI indicia estar-se na presença de um método de avaliação subjectivo dos valores do mercado. A definição anual do valor base dos prédios edificados é obtida via auscultação de especialistas do sector imobiliário e os factores responsáveis pela diferenciação, como os coeficientes de afectação e de localização, parecem ter subjacente um modelo gravitacional (utilizando sistemas de informação geográfica (SIG)) para gerar índices territoriais. Já a pré-definição de valores referência para os coeficientes de qualidade e conforto e de vetustez reflectem combinações de critérios objectivos e subjectivos que se presumem afectar a desejabilidade de todas as comunidades. Estes métodos de avaliação apresentam vantagens, na medida em que permitem ultrapassar falhas relativas à obtenção de dados do mercado, mas uma grande dependência deles pode conduzir à descreditação do mercado, já que as ordenações e classificações dos coeficientes podem não reflectir as evidências das preferências de mercado.

¹⁸⁹ O que decorre da base de dados, suporte do ensaio, contemplar algumas propriedades imobiliárias que não constituem prioridade para efeitos de avaliação, pelos serviços da administração fiscal, no quadro das novas regras de avaliação, por se encontrarem, há data, avaliadas pela repartição de Finanças. Em sede da Portaria em apreço perspectiva-se uma avaliação gradual de todos os prédios urbanos, que se iniciará por todos os imóveis declarados à Direcção Geral de Impostos depois de 12 de Novembro de 2003 e dos transmitidos a partir de 1 de Dezembro desse ano, logo que a Portaria que fixa os coeficientes constantes do IMI seja publicada.

¹⁹⁰ Portaria nº 982/2004, de 4 de Agosto, aprova e publicita os coeficientes a fixar para a avaliação a decorrer no âmbito do CIMI.

pelo CIMI, abandonando-se a avaliação que decorre do regime transitório, para de novo comparar com a avaliação decorrente do modelo proposto.

Antes de se proceder à avaliação patrimonial, propriamente dita, e à comparação de processos de avaliação, há que referir que, na fase de recolha de dados, paralelamente ao processo de auscultação às imobiliárias a operar na cidade da Guarda, se procedeu à observação dos registos notariais, para o ano 2002, relativos a actos de compra e venda de habitação. Estes últimos dados revelaram-se interessantes pelo facto de fornecerem o valor matricial da propriedade a ser transaccionada, informando sobre o número do artigo de registo na Direcção Geral de Finanças e o montante (valor patrimonial) inscrito e isso verificava-se para todos os imóveis (92) com idade superior a, pelo menos, dois a três anos; os outros surgiam como omissos nos serviços da Direcção Geral de Finanças. Ora, esses dados relativos à venda de imóveis, desde que não omissos, permitiram estruturar uma base de dados de forma a responder ao objectivo de avaliar o rigor do modelo de avaliação hedónico face ao regime transitório de avaliação do IMI, e face à implementação das novas regras de avaliação constantes do CIMI.

A partir dos valores matriciais das propriedades imobiliárias,¹⁹¹ constantes da base de dados construída, a correcção matricial aconteceria pela aplicação de um coeficiente de desvalorização da moeda, em função do ano de inscrição do imóvel na matriz predial urbana, constante da Portaria nº 1337/2003, de 5 de Dezembro, com o intuito de se actualizar os valores patrimoniais tributários dos prédios urbanos não arrendados e dos prédios rústicos. Acontece, porém, que a informação disponível nada nos podia acrescentar sobre a coincidência do valor matricial conhecido com o valor matricial do ano de inscrição do imóvel na matriz predial, este último a base para a actualização do valor patrimonial dos prédios não arrendados; no entanto a informação relativa ao

¹⁹¹ Assume-se, por falta de informação que permita ir noutro sentido, que os imóveis não estão arrendados, o que nem é muito irreal se atentarmos nas especificidades e particularidades do actual mercado de arrendamento.

número do artigo de registo matricial veio a revelar-se fulcral para uma pesquisa directa na base de dados da repartição de finanças da Guarda.¹⁹²

5.2. Dos Valores Matriciais do Ano de Inscrição aos Valores Matriciais Corrigidos

Para se ficar com uma ideia dos valores constantes dessas matrizes, apresentar-se-á uma tabela com algumas medidas de estatística descritiva relativamente aos valores iniciais inscritos na matriz e aos valores corrigidos, por aplicação do coeficiente de desvalorização ao valor matricial,

	Valor Matricial	Correcção Matricial
	(em €)	(em €)
Média	16 788,84	24 740,39
Mediana	15 158,47	24 355,31
Máximo	48 819,84	79 537,95
Mínimo	358,78	523,81
$ Max - Min /Max$	0,993	0,993
Max/Min	136,072	151,845
Desvio Padrão	12 618,98	14 724,92
Assimetria	0,390594	0,647522
Curtose	1,932902	3,777167

Tabela 5.2-1: Valores inscritos nas Matrizes e respectiva Correção

A observação dos valores matriciais é reveladora de uma substancial disparidade nesses valores, sobretudo se atentarmos nas características físicas, de localização e do ano de registo dos imóveis em análise, facto possivelmente imputável ao valor de referência para a avaliação imobiliária até agora utilizado, a presunção do “rendimento do imóvel”.¹⁹³ Para se ilustrar os anacronismos decorrentes da avaliação patrimonial

¹⁹² O conhecimento do número do artigo da inscrição matricial permitiu aceder, via meios informáticos da repartição de Finanças da Guarda, ao valor matricial do ano de inscrição e ao valor, entretanto, actualizado por aplicação da Portaria n° 1337/2003.

¹⁹³ Presunção que se consolidou pelo facto do Código da Avaliação, referenciado no Código da Contribuição Autárquica, nunca ter conhecido a luz do dia. Na verdade a base de incidência da Contribuição Autárquica era o “valor patrimonial do prédio” e para o

utilizada, poder-se-á recorrer a alguns elementos da amostra, componentes da base de dados utilizada para este objectivo de apuramento do rigor do modelo de avaliação hedónico, constantes do apêndice I.

Sem se pretender a exaustividade, destaca-se a situação de,

- propriedades imobiliárias com a idade comum de dezoito anos, onde se identificam seis apartamentos, caracterizados pela mesma tipologia, T2 e com áreas médias semelhantes, aproximadamente, de 90 m², com valores matriciais a variarem entre 632,38€ a 22 468,5€. Imputar responsabilidade por esta discrepância à localização destes apartamentos não parece razoável pois que, neste episódio, todos estão localizados no Centro e, também, não parece suficiente atribuí-la à presença (ou não) de garagem. Ainda, com a mesma idade, mas, agora, com uma tipologia T3, têm-se três apartamentos, com áreas aproximadas de 130 m², com valores matriciais a variarem entre 1 612,51€ e 5 858,38€.
- propriedades imobiliárias de idade igual a quinze anos, onde se identificam cinco apartamentos, caracterizados por tipologia T2 e com áreas aproximadas de 90 m², com valores matriciais a oscilarem entre os 4 374,32€ e os 19 527,94€ e seis apartamentos caracterizados por tipologia T3 e com áreas aproximadas de 120 m², com valores matriciais a oscilarem entre os 5 207,45€ e os 16 969,1€.
- propriedades imobiliárias com a mesma idade de cinco anos, onde se identificam três apartamentos, caracterizados por tipologia T2 e com áreas médias de 90 m², com valores matriciais a variarem entre 15 338,04€ e 31 903,11€ e, ainda, dois,

cálculo desse valor, a lei remetia para um Código de Avaliação, equacionava-se, porém, a eventualidade do imposto incidir sobre “o valor tributável do prédio”, enquanto esse código não fosse estabelecido e entrasse em vigor, o que, aliás, nunca viria a acontecer. Ora, esse “valor tributável do prédio” resultaria da avaliação feita de acordo com as regras do Código de Contribuição Predial e do Imposto sobre Indústria Agrícola, sendo apurado por aplicação do disposto nos artigos 6º e 7º do Decreto Lei nº 442-C/88, de 30 de Novembro.

caracterizados por uma tipologia T3 e com áreas médias de 130 m², os valores matriciais são de 25 386,32€ e 28 236,61€.

Esta ilustração da disparidade, patente nos valores matriciais da propriedade imobiliária, tem por objectivo reforçar a realidade de uma evidência, desde logo, expressa nos números que traduzem o peso percentual da amplitude de variação dos valores matriciais da amostra sobre o valor matricial máximo, cerca de 99,3% , revelando uma amplitude de variação quase coincidente com o valor matricial máximo, o qual por sua vez é cerca de 136,1 vezes superior ao valor matricial mínimo da amostra em estudo.

A divergência desses valores/base matricial vem ampliada pela introdução dos coeficientes de actualização, decorrente da aplicação do regime transitório de avaliação do valor patrimonial tributável, até que todos os imóveis possam ser avaliados de acordo com as novas regras do IMI. Realidade, facilmente, observável no apêndice I . Para ilustrar este impacto da correcção matricial, pode-se recorrer a uma das situações, referidas acima, p.e. a relativa a seis apartamentos de tipologia T2, com áreas médias semelhantes (aproximadamente, de 90 m²) e com valores matriciais a variarem entre 632,38€ a 22 468,5€, que passam, após aplicação do mesmo coeficiente de desvalorização da moeda,¹⁹⁴ a variar entre 1 694,77€ e 79 537,95€. Também, o facto do peso percentual da amplitude de variação dos valores matriciais em relação ao valor matricial máximo continuar nos mesmos 99,3% e do valor matricial máximo passar a ser cerca de 151,9 vezes superior ao valor matricial mínimo confirmam o empolamento da disparidade constante da situação de partida.

O recurso às medidas de simetria (ou de falta de simetria) e de curvatura (curtose) permite observar que a distribuição dos valores matriciais actualizados, nesta fase transitória de aplicação do IMI, se acentua à direita do valor médio, já que a medida de

¹⁹⁴ O coeficiente é o mesmo para os 6 apartamentos, já que têm a mesma idade, 18 anos, e toma o valor de 3,54.

assimetria passa de aproximadamente 0,391 para 0,648 e que se apresenta tendencialmente mais inclinada, já que o valor da curtose passa de um valor negativo, -1,067, (distribuição “plana”) para um valor positivo, 0,777 (distribuição “unimodal”).

5.3. Dos Valores de Avaliação Hedónica aos Valores de Mercado

Entretanto, por aplicação do modelo de preços hedónicos às propriedades imobiliárias, constantes desta base, obtém-se um conjunto de valores patrimoniais, cujas medidas de estatística descritiva se apresentam,

	Avaliação Hedónica (€)
Média	82 267,30
Mediana	83 137,81
Máximo	143 446,9
Mínimo	46 113,59
$ Max - Min /Max$	0,6785
Max/Min	3,111
Desvio Padrão	18 426,55
Assimetria	0,309860
Curtose	3,084467

Tabela 5.3-1: Valores relativos à Avaliação Hedónica

Constata-se que nos valores patrimoniais decorrentes da avaliação, base hedónica, a amplitude dos valores hedónicos em relação ao valor máximo de avaliação hedónica é de cerca de 67,9% e que o valor avaliado máximo é cerca de três vezes superior ao valor avaliado mínimo. Os valores da avaliação hedónica tendem a seguir uma

distribuição normal, já que os valores da assimetria e da curtose se aproximam dos valores zero e três, respectivamente.

Já no que concerne aos valores de venda das propriedades imobiliárias, constantes da amostra/base do ensaio, os valores relativos às respectivas medidas de estatística descritiva são,

	Valor Mercado (€)
Média	83 873,49
Mediana	86 800,00
Máximo	126 647,0
Mínimo	32 334,00
$ Max - Min /Max$	0,7447
Max/Min	3,917
Desvio Padrão	21 019,83
Assimetria	-0,455271
Curtose	2,599936

Tabela 5.3-2: Valores de Mercado

É interessante verificar que o valor máximo de venda de um imóvel é cerca de 3,9 vezes o valor mínimo de venda de um imóvel e que essa relação é, aproximadamente, de três vezes nos valores avaliados – base hedónica, contra as relações de 136,1 e de 151,9 vezes na avaliação matricial original e na avaliação matricial corrigida, respectivamente. Dilui-se um pouco a tendência para uma distribuição normal dos valores de venda dos imóveis; note-se, no entanto, a prevalência para esta tendência, quer nos valores hedónicos, quer nos valores de venda.

A análise relativa às medidas estatísticas dos dados relativos, aos valores matriciais originais, aos valores matriciais corrigidos, aos valores avaliados - base hedónica e aos valores de venda dos imóveis, constantes da base de dados em estudo, tem um suporte

relativamente robusto, com melhor desempenho para os dois últimos, pois que, em qualquer uma das situações, esses dados estão distribuídos mais simetricamente sobre o valor médio e não concentrados na proximidade dos valores extremos e isto porque cerca de 65% a 66% das observações têm valores dentro de um desvio padrão do valor médio (isto é, o intervalo $[\bar{X} - S_X; \bar{X} + S_X]$) e, similarmente cerca de 98% a 99% das observações têm valores dentro de dois desvios padrão da média (isto é, o intervalo $[\bar{X} - 2S_X; \bar{X} + 2S_X]$), o que confere um considerável poder interpretativo ao desvio padrão destes dados como medida de dispersão.

5.4. Do Rácio de Avaliação ao Coeficiente de Dispersão

Para se medir a capacidade de um sistema fiscal avaliar consistentemente a base de incidência do imposto há que introduzir um indicador que relacione o valor avaliado com o preço de venda do imóvel, que se pode denominar de rácio de avaliação e expressar-se por, $RA_n = \frac{ValorAvaliado_n}{ValorVenda_n}$, em que n identifica um qualquer imóvel.

A construção deste rácio, ao exigir o valor de venda dos imóveis, aliás a variável dependente do modelo empírico proposto no capítulo 3, sub-secção 3.2., coloca de novo questões relativamente à sua fiabilidade. No entanto, o recurso às informações auscultadas junto das imobiliárias da cidade da Guarda relativas ao Valor de Mercado, a garantia obtida de traduzirem o valor da transacção imobiliária e a percepção pessoal do mercado imobiliário, permitem garantir um rigor, próximo dos cem por cento, para os valores de venda constantes da base de dados construída.

O rácio de avaliação permite acomodar a taxa nominal do imposto, de tal forma que quanto mais próximo vier da unidade, maior a vizinhança dos valores das taxas nominal e efectiva do imposto, sinalizando, nesse cenário, a observância da equidade exigida a um qualquer imposto. Assim a escolha política, pelos governos, de um determinado

valor para a taxa nominal do imposto sobre o património devia ser uma decisão conjunta com a escolha dos valores do rácio de avaliação.¹⁹⁵

Naturalmente que um rácio de avaliação tendencialmente igual a um é indicador de uma avaliação da propriedade imobiliária próxima do valor de mercado.¹⁹⁶ Mais, numa situação ideal o rácio de avaliação deveria ser o mesmo para todas as propriedades imobiliárias pertencentes à mesma jurisdição. A uniformidade deste rácio é um garante da consistência do processo de avaliação. Nesta perspectiva, procurou-se encontrar o rácio de avaliação, quer para a base corrigida do IMI, RA_{IMI} , quer para a base hedónica resultante da equação de preços hedónicos estimados, $RA_{Hedónica}$, tendo por base o valor de venda declarado pelas imobiliárias, e das respectivas medidas de concentração e dispersão estatísticas se dá conta,

	RA_{IMI}	$RA_{Hedónica}$
Média	0,289850	0,995887
Mediana	0,273117	0,989071
Máximo	1,104694	1,83398
Mínimo	0,0162	0,595382
$ Max - Min /Max$	0,985	0,675
Max/Min	68,2	3,1
Desvio-Padrão	0,175324	0,164263
Assimetria	1,414669	1,797348
Curtose	7,045049	9,90665

Tabela 5.4-1: Valores dos Rácios de Avaliação

Constata-se que a média e mediana do rácio de avaliação, base hedónica, se fixa nos valores 0,995887 e 0,989071, respectivamente, contra os 0,289850 e 0,273117 da média e da mediana do rácio de avaliação, base IMI. É notória a aproximação tendencial à unidade para o rácio de avaliação, base hedónica. Também, o valor mínimo atingido

¹⁹⁵ Como a receita fiscal (T) é resultado da taxa nominal do imposto (t) sobre a base fiscal, valor avaliado, (Va) tem-se $t = T/Va$ e como a taxa efectiva do imposto (τ) decorre da relação de T com o valor de venda (Vv), $\tau = T/Vv$, chega-se à relação $\tau = RA.t$.

¹⁹⁶ Em Washington é exigido, por lei, aos avaliadores imobiliários uma avaliação patrimonial igual à do valor de mercado, refere Goolsby, W. (1997).

por este rácio se cifra em 0,595382, contra o valor mínimo de 0,01620 alcançado pelo rácio de avaliação, base IMI, indicando a presença de subavaliação, a qual atinge contornos mais problemáticos na avaliação, base IMI. Na base IMI, o valor máximo do rácio de avaliação é 68,2 vezes, superior ao valor mínimo desse rácio de avaliação, contra um valor de 3,1 vezes, na base hedónica, revelando uma maior uniformidade para os valores do rácio de avaliação desta última base. Tese corroborada pelo facto da amplitude do intervalo entre os valores máximo e mínimo do rácio de avaliação, na base IMI, ter quase a dimensão do valor máximo desse intervalo, enquanto, na base hedónica, se fica pelos 67,5% do valor máximo do rácio de avaliação.

Os valores do rácio de avaliação permitem identificar enviesamentos, mais ou menos, consistentes na avaliação, assim:

- para valores do rácio de avaliação iguais ou inferiores a 0,5 assume-se subavaliação grave, quer para o Tesouro (execução de receita fiscal) quer para o contribuinte (violação do princípio de equidade fiscal). Este episódio tem uma ocorrência de 0% na avaliação hedónica contra 91,3% na avaliação IMI. Esta situação exige particular atenção das autoridades fiscais pois que, ao não sofrer qualquer contestação da parte do sujeito passivo do imposto não permitirá recolher informação sobre a presença de eventuais erros sistemáticos presentes no processo de avaliação;
- para valores de rácios de avaliação, menores ou iguais a 1 e maiores ou iguais a 0,8, pode-se assumir avaliação mais ou menos justa. Nesta situação ficam 37% dos rácios na avaliação hedónica contra 0% na base IMI. Caso o limite inferior do intervalo do rácio de avaliação, considerado mais ou menos justo, descesse para 0,7, então as posições \pm justas passavam a 56,5% na avaliação hedónica e a 2,2% na avaliação IMI;

- para valores do rácio de avaliação superiores a um, assume-se sobreavaliação. Inventariando as situações de sobreavaliação decorrentes da aplicação das bases de avaliação IMI e hedónica, tem-se, respectivamente 1% e 42,4% dessas situações. Embora seja expectável que as situações de sobreavaliação venham a ser alvo da contestação dos contribuintes, impedindo que a autoridade fiscal as negligencie, o que induzirá um mecanismo de auto-correcção no processo de avaliação, mesmo assim esta ocorrência deve ser ponderada na presente análise do modelo de avaliação hedónica e deve conduzir ao despiste de eventuais enviesamentos constantes deste processo de avaliação.

Para cada uma das bases de incidência fiscal, os valores dos rácios de avaliação apurados não permitem inferir da sua uniformidade, daí que uma análise mais precisa da consistência do processo de avaliação imobiliária imponha observar o grau de variância do rácio de avaliação, através de um coeficiente de dispersão,¹⁹⁷ passível de se expressar

por, $CD = \frac{\sigma_{RA}}{\mu_{RA}}$.¹⁹⁸ Os valores a reter são assim,

	RA_{IMI}	RA_{Hedónica}
CD	0,60488	0,16494

Tabela 5.4-2: Coeficiente de Dispersão dos Rácios de Avaliação

O coeficiente de dispersão (**CD**) vai medir a percentagem do desvio do rácio de avaliação em relação ao respectivo valor médio, fornecendo informação sobre a confiança associada a cada um dos procedimentos de avaliação. Assim, quanto mais baixo o valor do coeficiente de dispersão melhor indicação de uma avaliação imobiliária

¹⁹⁷ Na linha do indicador sugerido por Bowman e Mikesell (1978).

¹⁹⁸ Em que σ_{RA} é o desvio padrão do rácio de avaliação, definido por $\sigma = \sqrt{\sum_{n=1}^{92} (RA_n - \mu)^2 / 92}$, com $\mu = \mu_{RA}$ a representar a

média do rácio de avaliação, traduzida por, $\mu = \sum_{n=1}^{92} RA_n / 92$.

mais consistente e de um tratamento mais igualitário dos sujeitos passivos do imposto. Neste enquadramento os valores obtidos vão no sentido de conferir uma maior confiança aos procedimentos de avaliação constantes do modelo de preços hedónicos, face aos procedimentos seguidos no quadro transitório de aplicação do IMI.

Constata-se, também, que, em termos percentuais, a média dos desvios absolutos dos rácios de avaliação individual relativamente ao valor mediano desses rácios de avaliação, se situa nos 47,2% e nos 11,5%, do rácio de avaliação mediano, significando que, em média, os imóveis estão avaliados na vizinhança de 47,2% e 11,5% do nível de avaliação mediana, respectivamente para a base IMI e para a base hedónica. Segundo valores indicativos, seguidos por, Gloudemans (2001), apenas valores superiores a 15% ou 20% constituem motivo de preocupação, pois significam que um imóvel subavaliado 15% a 20% em relação ao valor mediano paga 35% a 50% menos de imposto patrimonial do que um imóvel sobreavaliado 15% a 20% em relação ao mesmo valor mediano. Neste caso, ter-se-ia, para a base IMI, imóveis subavaliados 47,2% em relação ao valor avaliado mediano a pagarem 179% menos de imposto patrimonial do que um imóvel sobreavaliado em 47,2% e para a base hedónica, imóveis subavaliados 11,5% em relação ao valor avaliado mediano a pagarem 26% menos de imposto patrimonial do que um imóvel sobreavaliado em 11,5% – a equidade fiscal é preocupante na base IMI enquanto na base hedónica, não sendo perfeita é aceitável para os padrões de ajustamento e de fiabilidade característicos dos processos de avaliação.

A variabilidade dos valores de venda dos imóveis exige a mensuração dos enviesamentos que esse diferencial de preços pode provocar na avaliação desses imóveis, tornando-se importante verificar se os imóveis são avaliados ao mesmo nível, quer se identifiquem com valores de venda elevados, quer se identifiquem com valores

de venda inferiores. Para tanto introduz-se um índice diferencial de preços, que se obtém dividindo a média simples dos rácios de avaliação pela média ponderada dos mesmos rácios,¹⁹⁹ vindo para a base IMI o valor de 93 e para a base hedónica o valor de 101.

Os parâmetros orientadores para este índice apontam dois limites, um para valores do índice diferencial de preços superiores a 103, sugerindo uma subavaliação dos imóveis de valor elevado e outro para valores do índice diferencial de preços inferiores a 98 sugerindo uma subavaliação dos imóveis de valor baixo. Neste contexto, a avaliação resultante da correcção do IMI ao conduzir a um valor de 93 para o índice diferencial de preços predispõe para uma situação de subavaliação nos imóveis com valor de mercado mais baixo.

5.5. Da Receita Fiscal à Elasticidade da Receita Fiscal

Apurar o rigor e o desempenho do modelo de avaliação patrimonial hedónica proposto face ao que decorre dos valores matriciais corrigidos pela aplicação do coeficiente de desvalorização da moeda, constantes da Portaria nº 1337/2003, passa, também, por apurar a receita fiscal do imposto patrimonial. Para o efeito, a taxa de 0,8% de IMI, fixada pela autarquia da Guarda, para 2004, passa a ser aplicada à base de dados, suporte da análise ao rigor e precisão do modelo proposto, quer aos valores patrimoniais decorrentes da avaliação hedónica quer aos valores matriciais corrigidos pela aplicação do coeficiente de desvalorização da moeda.²⁰⁰ A receita fiscal, decorrente da aplicação

¹⁹⁹ Média ponderada dos rácios de avaliação, $\sum_{n=1}^N \frac{A_n}{V_n} * V_n / \sum_{n=1}^N V_n$, em que A_n é o valor avaliado de cada imóvel n , V_n é o valor

de venda de cada imóvel n .

²⁰⁰ Estes últimos, reflectem a base de incidência para efeitos do IMI 2004 na autarquia da Guarda, relativamente às observações constantes da base de dados em estudo, caso não se tivesse em conta a cláusula de salvaguarda, constante da Portaria já referida. Por

da taxa de 0,8% à base IMI, das observações em estudo, cifra-se em 18 208,93€. Caso essa taxa se aplicasse à base de avaliação hedónica, das mesmas observações, a receita fiscal passaria a ser de 60 548,73€.

Caso se aplicasse uma taxa de IMI de apenas 0,2% ²⁰¹ à base de valores patrimoniais resultante, quer da correcção matricial do regime transitório do IMI, quer da aplicação da equação de preços hedónicos, viria uma receita fiscal no valor de 4 552,231€ e no valor de 15 137,18€, respectivamente.

Admitindo a hipótese constante para a base dos valores patrimoniais a que se chegou, via correcção matricial e via aplicação hedónica, então o efeito perfeitamente elástico da variação da receita fiscal a uma qualquer variação da taxa nominal de IMI, permite concluir que, para se conseguir a mesma receita fiscal de 18 208,93€, obtida pela aplicação da taxa de 0,8% à base IMI corrigida, bastaria aplicar uma taxa nominal de 0,241% à base de avaliação hedónica. Captar a mesma receita fiscal com uma taxa efectiva de imposto inferior sinaliza potencial para maior eficiência fiscal, característica já perceptível através do valor médio do rácio de avaliação/base hedónica (0,995887).

A receita fiscal da autarquia, decorrente da aplicação da taxa do imposto municipal à base IMI, constitui para os contribuintes (os munícipes) a carga fiscal do imposto patrimonial. Ora, é esta perspectiva dualista do imposto, benefício/custo, que torna mais simpática a base hedónica, pois permite alcançar a mesma receita fiscal, 18 208,93€, com uma taxa nominal de apenas 0,241% e com uma distribuição mais uniforme de carga fiscal entre os contribuintes.

uma questão de coerência metodológica, optou-se por não considerar a cláusula de salvaguarda e comparar as receitas fiscais efectivas decorrentes da aplicação da taxa nominal a qualquer uma das bases patrimoniais.

²⁰¹ Taxa nominal mínima do IMI, prevista no CIMI, a aplicar após avaliação dos imóveis segundo as novas regras de avaliação, a concluir num prazo previsto de 10 anos.

Como se constatou a factura fiscal (custo) do contribuinte pode ser diferente e, ainda assim, gerar a mesma receita fiscal (benefício) para a autarquia, é o caso da factura fiscal decorrente da base IMI após aplicação da taxa nominal de 0,8% , e da factura fiscal decorrente da base hedónica após aplicação da taxa nominal 0,241% , cujas medidas de tendência central e de dispersão se apresentam,

	IMI	IMI
	Corrigido (€)	Hedónico(€)
Média	197,9231	197,9269
Mediana	194,8425	200,0213
Máximo	636,3036	345,1188
Mínimo	4,190480	110,9447
Max/Min	151,845	3,11
DesvioPadrão	117,7993	44,33243
Assimetria	0,647522	0,309860
Curtose	3,777167	3,084467

Tabela 5.5-1: Receita Fiscal do Imposto Patrimonial

A disparidade da factura fiscal, base IMI corrigida, é uma realidade expressa pelo valor máximo da factura fiscal, IMI, que supera, cerca de 152 vezes, o valor mínimo da factura fiscal, IMI. Já a factura fiscal, base hedónica, é caracterizada pela uniformidade, ilustrada pela relação entre os valores máximo e mínimo da factura fiscal, base hedónica, o primeiro é 3,1 vezes o segundo.

A introdução de outros indicadores, como o peso percentual do desvio absoluto médio da factura fiscal de cada contribuinte face à factura fiscal mediana no valor da factura fiscal mediana, com os valores de 50,2% e de 18% , respectivamente para a base IMI e para a base hedónica; e a relação percentual entre o desvio padrão da factura fiscal e a

factura fiscal média, com os valores de 59,5% e de 22,4% , respectivamente para a base IMI e para a base hedónica, corroboram a presença de uma maior uniformidade para a factura fiscal decorrente da base hedónica, já que valores mais baixos traduzem um tratamento mais igualitário dos sujeitos passivos do imposto patrimonial.

5.6. Das Taxas Efectivas do Imposto Patrimonial

Uma avaliação uniforme, equitativa e não discriminatória do valor patrimonial das propriedades imobiliárias constitui um dos requisitos para que se verifique uniformidade nas taxas efectivas do imposto sobre o património. Ora, observando as taxas efectivas do imposto sobre o património,²⁰² no quadro do IMI corrigido e no quadro da base hedónica, pode obter-se os seguintes indicadores estatísticos,

	Taxa Efectiva IMI_{Corrigido} 0,8%	Taxa Efectiva IMI_{Hedónica} 0,8%
Média	0,002319	0,007967
Mediana	0,002185	0,007913
Máximo	0,008838	0,014672
Mínimo	0,000130	0,004763
Max/Min	67,98	3,08
Desvio Padrão	0,001403	0,001314
Assimetria	1,414669	1,797348
Curtose	7,045049	9,906650

Tabela 5.6-1: Taxa efectiva do Imposto Patrimonial²⁰³

²⁰² Num cenário de uma taxa nominal de 0,8%, taxa aplicada à base patrimonial da cidade da Guarda, para apuramento do IMI de 2004.

²⁰³ Os valores máximos da taxa efectiva do imposto patrimonial ao serem superiores a 0,8% indiciam que a base de incidência do imposto (valor avaliado) está sobrevalorizada, o que não surpreende, pois está em linha com os valores máximos das tabelas 5.4-1., os quais ao serem superiores a um traduziam sobreavaliação patrimonial face ao valor de mercado.

Efectivamente as taxas, efectiva e nominal, do imposto patrimonial, para uma base fiscal avaliada a preços hedónicos são, em termos médios e medianos, quase iguais, pois 0,007967 e 0,007913, são uma boa aproximação de 0,008. Já o mesmo não acontece com as taxas, efectiva e nominal, na base IMI corrigida, onde a média e a mediana rondam os 0,2% contra os 0,8% nominais. Além desta disparidade entre taxas, nominal e efectiva, na base IMI corrigida, também a amplitude entre os valores mínimo e máximo da taxa efectiva, é bem acentuada, com o valor máximo a ultrapassar cerca de 68 vezes o mínimo. Ao invés, na base de avaliação hedónica, o valor máximo de taxa efectiva do imposto sobre o património é, apenas, 3,1 vezes, superior ao valor mínimo da taxa efectiva do imposto.

Em suma, na base de avaliação hedónica, verifica-se uma quase coincidência entre as taxas, nominal e efectiva, do imposto patrimonial, bem como uma tendência para uma maior uniformidade relativa na distribuição da taxa efectiva do imposto, facto conducente a uma maior equidade já que em termos nominais a taxa é absolutamente uniforme.

5.7. Da Aplicação da Portaria nº 982/2004 à Base de Dados

No âmbito da secção 5.1. decidiu-se escolher os imóveis constantes da base de dados em estudo que eram portadores de inscrição matricial para suportar a análise comparativa objecto deste capítulo e para o efeito procedeu-se à correcção dos valores matriciais no quadro do regime transitório de avaliação preconizado em sede da Portaria regulamentadora nº 1337/2003. Ainda assim, decidiu-se avaliar esses imóveis segundo as regras do IMI e de acordo com os valores fixados pela Portaria nº 982/2004,²⁰⁴ de

²⁰⁴ A qual permite informação sobre: o zonamento e os limites, mínimo e máximo, para os coeficientes de localização aprovados (na Guarda [0,35;1,3], para a habitação); as percentagens correspondentes à área de implantação, para apuramento do valor

forma a simular o efeito do IMI sobre uma base de incidência corrigida e reflectindo o impacto da fórmula de avaliação estabelecida em sede do decreto-lei n.º 287/2003 e dos parâmetros entretanto estabelecidos.

Consumando o propósito, da aplicação da fórmula de avaliação e da utilização dos parâmetros propostos, vieram novos valores para os imóveis em estudo, que podem ser resumidos segundo as medidas de concentração e dispersão relativas à distribuição desses valores, assim,

	Avaliação Nova_{IMI} (€)	Avaliação Hedónica (€)	Valor Mercado (€)
Média	74 520,28	82 267,3	83 873,49
Mediana	73 070,4	83 137,81	86 800,00
Máximo	128 622,0	143 446,9	126 647,0
Mínimo	39 780,00	46 113,59	32 334,00
 Max – Min /Max	0,69	0,6785	0,7447
Max/Min	3,2(3)	3,111	3,917
Desvio Padrão	15 601,64	18 426,55	21 019,83
Assimetria	0,458375	0,30986	-0,455271
Curtose	3,573857	3,084467	2,599936

Tabela 5.7-1: Bases de Incidência reflectindo as avaliações utilizadas

Desta avaliação decorre uma maior aproximação às medidas de estatística descritiva da **Avaliação Hedónica** e dos **Valores de Mercado** praticados. Por exemplo, relativamente à amplitude dos valores da Avaliação Nova como percentagem do valor máximo e à relação entre os limites inferior e máximo dessa distribuição, obtêm-se

patrimonial tributário dos terrenos para construção, bem como as respectivas áreas de aplicação; os coeficientes majorativos aplicáveis às moradias unifamiliares; as directrizes relativas à apreciação da qualidade construtiva, da localização excepcional e do estado deficiente de conservação; o custo médio de construção por metro quadrado em (euros) 480, a vigorar em 2003 e 2004.

valores de 69% e 3,2(3), respectivamente, face aos valores de 67,85% e 3,111 da distribuição da **Avaliação Hedónica** e aos valores de 74,47% e de 3,917 da distribuição dos **Valores de Mercado**. Também a ordem de grandeza dos valores relativos à média, mediana e desvio padrão da distribuição dos valores dos imóveis decorrentes da **Avaliação Nova** se aproxima mais das correspondentes medidas da distribuição dos valores dos imóveis resultante, quer da **Avaliação Hedónica**, quer dos **Valores de Mercado**.

Se compararmos os valores de avaliação das propriedades imobiliárias, resultantes da aplicação das novas regras do CIMI e dos valores e limites fixados na Portaria nº 982/2004 com os valores de mercado, apurar-se-á o rácio de avaliação de cada uma das propriedades e, efectivamente, o desempenho deste rácio melhora notoriamente em relação ao rácio de avaliação que decorre da avaliação transitória aplicada nesta fase, assim,

	RA Nova_{IMI}	RA_{IMI}
Média	0,914950	0,289850
Mediana	0,864865	0,273117
Máximo	2,062535	1,104694
Mínimo	0,482664	0,016200
$ Max - Min /Max$	0,77	0,985
Max/Min	4,3	68,2
Desvio Padrão	0,214426	0,175324
Assimetria	2,284492	1,414669
Curtose	12,08745	7,045049

Tabela 5.7-2: Rácios de Avaliação decorrentes da Portaria nº 982/2004

Como se visualiza os valores médios e medianos passam a aproximar-se da unidade contra uma prestação de cerca de 0,28 na avaliação transitória, denotando uma maior proximidade entre valores de avaliação e valores de mercado. O valor máximo torna-se

significativamente mais próximo do valor mínimo, no entanto o desvio padrão da distribuição dos valores dos imóveis avaliados de acordo com as regras do CIMI aumenta face ao regime transitório.

Observando o desvio absoluto médio do rácio de avaliação de cada imóvel face ao rácio de avaliação mediano como percentagem do rácio de avaliação mediano, tem-se um valor de 15,2%, que permite afirmar que os bens imobiliários, com este rácio de avaliação após aplicação das novas regras de CIMI, passam a estar avaliados, em média, na vizinhança de 15,2%, do nível mediano de avaliação, ficando-se no limiar dos valores que constituem motivo de preocupação, como já foi referido.

Apurando o coeficiente de dispersão (**CD**) do rácio de avaliação, agora em estudo, verifica-se uma relação percentual de 23,4% entre o desvio padrão e a média desse rácio de avaliação, o que confere uma melhoria significativa à percentagem do desvio em relação ao rácio de avaliação médio face ao que decorre da avaliação transitória, o que vai no sentido da presença de uma maior consistência nos procedimentos de avaliação, já que valores mais baixos traduzem um tratamento mais igualitário dos sujeitos passivos do imposto patrimonial e por isso uma maior confiança a este procedimento de avaliação. Este coeficiente de dispersão (**CD**) é, no entanto, mais elevado do que o que resulta da **Avaliação Hedónica**, o que significa que existe maior variância entre valor avaliado e valor de venda no quadro do CIMI do que no âmbito da **Avaliação Hedónica**.

Ainda se pode calcular o índice diferencial de preços, que se obtém dividindo a média simples dos rácios de avaliação pela média ponderada dos mesmos rácios, vindo para esta base o valor de **88**, indiciando uma situação tendencial de subavaliação nos imóveis com valor de mercado mais baixo face aos parâmetros orientadores deste índice, conforme referido na secção **5.4.**

Suportados nesta nova avaliação dos bens imóveis e tendo por base, (i) a taxa nominal de 0,8% , fixada pelo município da Guarda, para o ano de 2004; (ii) o valor da receita fiscal resultante da aplicação dessa taxa à base de avaliação transitória; (iii) a variação perfeitamente elástica da receita fiscal relativamente à variação da taxa nominal do imposto; conseguir-se-á essa receita fiscal, no montante 18 208,93€, aplicando, às observações em estudo, uma taxa nominal de 0,26% .

Agora, no cenário desta nova taxa nominal, os valores da factura fiscal devidos pelos contribuintes no que concerne ao imposto patrimonial sobre imóveis, podem ser sintetizados em,

	IMI Nova_{IMI} (€)	IMI Corrigido (€)	IMI Hedónico(€)
Média	197,9231	197,9231	197,9269
Mediana	194,0723	194,8425	200,0213
Máximo	341,6153	636,3036	345,1188
Mínimo	105,6542	4,190480	110,9447
Max/Min	3,2(3)	151,845	3,11
Desvio Padrão	41,43740	117,7993	44,33243
Assimetria	0,458375	0,647522	0,309860
Curtose	3,573857	3,777167	3,084467

Tabela 5.7-3: Receitas do Imposto Patrimonial em função das diversas bases de incidência

De novo se verifica uma aproximação notória aos valores da factura fiscal que decorrem da avaliação hedónica. Desde a aplicação do regime transitório até à aplicação das novas regras do CIMI passou-se de uma factura fiscal máxima a superar a factura fiscal mínima de 152 vezes para 3 vezes, em harmonia com o que sucede em termos da relação entre os valores máximo e mínimo da base de tributação, claramente um andamento no sentido de uma maior uniformidade no que concerne à distribuição da carga fiscal entre os contribuintes.

Observando o desvio absoluto médio da factura fiscal de cada contribuinte face à factura fiscal mediana como percentagem do valor da factura fiscal mediana, com um valor de 16,7% , e a relação percentual entre o desvio padrão da factura fiscal e a factura fiscal média, com um valor de 20,9% , não há dúvida quanto à presença de uma maior uniformidade na factura fiscal, já que valores mais baixos traduzem um tratamento mais igualitário dos sujeitos passivos do imposto patrimonial.

Para um melhor e mais rigoroso apuramento da uniformidade do imposto patrimonial há que relacionar a factura fiscal paga por cada contribuinte com o valor de mercado da propriedade imobiliária e apurar a taxa efectiva do referido imposto, quando se está a aplicar a taxa nominal de 0,8% à base de incidência fiscal. Ora, para a base de incidência coincidente com a avaliação decorrente da aplicação generalizada das novas regras constantes do CIMI, tem-se

	Taxa Efectiva IMI_{AvaliaçãoNova} 0,8%
Média	0,007320
Mediana	0,006919
Máximo	0,016500
Mínimo	0,003861
Max/Min	4,27
Desvio Padrão	0,001715
Assimetria	2,284492
Curtose	12,08745

Tabela 5.7-4: Taxa efectiva após avaliação conforme Portaria nº 982/2004

Com a generalização das regras de avaliação a todas as propriedades imobiliárias a taxa efectiva do imposto patrimonial aproxima-se significativamente da taxa nominal 0,8% , o que não acontece com a avaliação transitória, e também o valor máximo da taxa efectiva passa a ser cerca de 4 vezes o valor mínimo dessa taxa. A coincidência entre as

taxas nominal e efectiva do imposto patrimonial, com a base de avaliação decorrente das novas regras do CIMI, torna-se quase uma realidade e desenha-se uma tendência concreta para uma uniformidade relativa da distribuição da taxa efectiva na vizinhança reduzida da taxa nominal.

5.8. Da Pesquisa do Enviesamento ao Ajustamento

Da exposição anterior decorre que há uma vantagem, relativamente, superior para a metodologia de avaliação imobiliária sustentada no modelo de preços hedónicos face ao regime transitório do IMI em vigor e uma evidente aproximação à avaliação imobiliária decorrente da aplicação generalizada das novas regras do IMI; subsistem, no entanto, questões de uniformidade na sua aplicação. Factos a ter em conta, já que a observância da uniformidade no processo de avaliação é elemento fundamental para garantir o máximo de equidade fiscal, sobretudo se se atentar na elevada visibilidade do imposto patrimonial (talvez uma virtude perversa) e na percepção pública de que esse imposto é largamente regressivo.

Embora haja autores, tais como Fisher (1996), que afirmam que a sobrevivência do imposto patrimonial advém da necessidade dos governos locais em gerar receitas e pouco tem a ver com a prossecução de uma política de equidade na tributação da propriedade, reforçando assim a tese de que as decisões fiscais são decisões políticas e de que os sistemas de receitas fiscais e as estruturas governamentais estão interdependentes, mesmo assim vale a pena perseguir o objectivo da equidade fiscal.

Ora a imposição legal de uniformidade relativamente à taxa do imposto e à base de incidência fiscal, ao limitar, não só, a forma como as cargas fiscais do imposto são distribuídas, mas, também, ao desencorajar, potenciais manipulações fiscais orientadas para determinados grupos de cidadãos (os “votantes não decisivos” na perspectiva do

decisor político), constitui um meio para atingir esse objectivo. Neste âmbito valerá a pena introduzir um estudo conducente a avaliar a existência (ou não) de um enviesamento consistente no processo de avaliação proposto.²⁰⁵

Para tal há que recorrer a uma regressão entre o rácio de avaliação e as características/atributos dos imóveis habitacionais, aplicando o método dos mínimos quadrados, quer à base patrimonial avaliada por aplicação do modelo hedónico proposto, quer à base patrimonial decorrente da aplicação das novas regras do IMI.²⁰⁶

Negligencia-se a base IMI corrigida, pois ao ser gerada a partir dos valores matriciais originais dos imóveis, coincidentes com “o valor tributável do prédio”, pelo facto de nunca se ter estabelecido o Código de Avaliação, resultam de uma avaliação feita de acordo com as regras do Código de Contribuição Predial e do Imposto sobre Indústria Agrícola, o que torna a análise relativa à existência de enviesamentos despicienda - nas avaliações feitas nesse período dificilmente foram ponderados o Valor de Venda e as características dos imóveis.

Seja qual for a base patrimonial, a regressão vem expressa por,

$$\ln RA = f(\ln VMERCADO, \ln ANOS, \ln AREA, QUARTOS, CBANHO, GCANAL, \\ GSDOS, GSUM, LCNTR, LGGARE, LPMILEU, LPNH, PPRIM, PSEG)$$

²⁰⁵ Vários ensaios têm mostrado enviesamento consistente a favor dos imóveis com preço de venda mais elevado [Netzer (1966); Oldman e Aaron (1965); Paglin e Fogarty, (1972); Kochin e Parks (1984)], podendo mesmo os enviesamentos resultarem de outras características dos imóveis, como mostraram Berry e Bednarz (1975).

²⁰⁶ Parece mais plausível introduzir essa base não só porque é o objectivo último da plena execução do CIMI, mas também porque é aquela cujos rácios de avaliação são reveladores de uma maior proximidade aos valores de venda dos imóveis.

²⁰⁷ Optou-se por trabalhar com os logaritmos naturais das variáveis, *RA*, rácio de avaliação; *VMERCADO*, representativa do Valor de Venda do Imóvel; *ANOS*, traduzindo a idade do imóvel; e *AREA*, representativa da dimensão do imóvel, para acautelar possíveis enviesamentos de escala.

Da estimação dos coeficientes dos regressores, pelo método dos mínimos quadrados, resulta para a base de avaliação hedónica, a equação,

$$\begin{aligned} \ln RA_{Hed} = & 8,197 - 0,8489 * \ln VMERCADO - 0,0825 * \ln ANOS + 0,2624 * \ln AREA + \\ & (11,822) \quad (-18,967) \qquad \qquad \qquad (-6,493) \qquad \qquad \qquad (2,05) \\ & + 0,0825 * QUARTOS + 0,0586 * CBANHO + 0,0457 * GCAN \\ & (1,706) \qquad \qquad \qquad (1,562) \qquad \qquad \qquad (2,622) \\ & + 0,1158 * GSDOS + 0,0632 * GSUM - 0,0735 * LCNTRO - \\ & (1,553) \qquad \qquad \qquad (3,331) \qquad \qquad \qquad (-4,123) \\ & - 0,063 * LGGARE - 0,0668 * LPMILEU - 0,050 * LPNHR + \\ & (-2,418) \qquad \qquad \qquad (-2,630) \qquad \qquad \qquad (-1,370) \\ & + 0,0271 * PPRIM - 0,0111 * PSEG \\ & (1,418) \qquad \qquad \qquad (-0,600) \end{aligned}$$

$$R^2 = 0,852847$$

$$R^2 - Ajustado = 0,826092$$

$$Erro Padrão Estimado = 0,0637262$$

Verifica-se que, em termos estatísticos, todos os coeficientes, à excepção do que se refere à variável *PSEG*, identificada com o piso “segundo” onde se localiza o andar, são significativamente diferentes de zero, facto consistente com a metodologia utilizada para a avaliação hedónica. Naturalmente, também, explica os valores de R^2 e de R^2 _Ajustado.

No que concerne aos sinais dos coeficientes de cada uma das variáveis, caso sejam positivos indicam efeitos de sobreavaliação e caso sejam negativos denotam efeitos de subavaliação. Assim sendo verifica-se que, em relação ao rácio de avaliação hedónica, a tendência é para que habitações com maior valor de venda e com mais idade venham subavaliadas, enquanto as habitações com maior dimensão tendem a ser sistematicamente sobreavaliadas, o que está em linha com as estimativas dos estudos já referidos.

Já para a base imobiliária avaliada no quadro das novas regras do CIMI e dos valores e limites dos coeficientes de localização fixados na Portaria nº 982/2004, a equação relativa aos valores estimados vem,

$$\begin{aligned}
 \ln RA_{Nova} = & 7,664 + 0,936 * \ln VMERCADO - 0,156 * \ln ANOS - 4,851 * \ln AREA + \\
 & (1,055) \quad (1,996) \qquad \qquad \qquad (-1,173) \qquad \qquad \qquad (-3,617) \\
 & + 1,240 * QUARTOS + 0,023 * CBANHO - 0,179 * GCANAL + \\
 & (2,448) \qquad \qquad \qquad (0,057) \qquad \qquad \qquad (-0,978) \\
 & + 0,003 * GSDOS + 0,355 * GSUM + 0,076 * LCNTRO - \\
 & (0,004) \qquad \qquad \qquad (1,787) \qquad \qquad \qquad (0,406) \\
 & - 0,155 * LGGARE + 0,105 * LPMILEU + 0,410 * LPNHR - \\
 & (-0,570) \qquad \qquad \qquad (0,393) \qquad \qquad \qquad (1,073) \\
 & - 0,033 * PPRIM + 0,009 * PSEG \\
 & (-0,166) \qquad \qquad \qquad (0,047)
 \end{aligned}$$

$$R^2 = 0,314609$$

$$R^2 - \text{Ajustado} = 0,189993$$

$$\text{Erro Padrão Estimado} = 0,667881$$

Agora as variáveis independentes com coeficientes significativamente diferentes de zero conheceram assinalável redução. No conjunto das variáveis com coeficientes estimados significativamente diferentes de zero, destacam-se, $\ln VMERCADO$ e $\ln AREA$ pelo facto dos respectivos coeficientes estimados terem sofrido alterações de sinal de uma para a outra regressão. Para $\ln VMERCADO$, a passagem de um sinal (–) a um sinal (+) revela que, no cenário da avaliação hedónica, habitações com maior valor de venda tendem a ser subavaliadas, passando depois, no quadro das novas regras do IMI, a ser sobreavaliadas; ao invés para $\ln AREA$, no quadro da avaliação hedónica, dir-se-ia que habitações com maior dimensão tendem a estar sobreavaliadas, o que já não se verifica no quadro do novo IMI. Já as variáveis, $QUARTOS$ e $GSUM$, mantêm o sinal (+), permitindo prosseguir com a interpretação de que o número de quartos e a presença de uma garagem contribui para sobreavaliar o valor da avaliação.

Relativamente à variável $\ln ANOS$, ainda que o nível de significância estatística para o coeficiente estimado aponte para a aceitação da hipótese de um valor nulo em contexto das novas regras do IMI, importa registar a manutenção do sinal (–) nesse coeficiente estimado, nas duas regressões, o que permite continuar a interpretar a característica *mais idade* como potencial factor de subavaliação da habitação.

Ao nível das variáveis de localização, $LCNTRO$, $LGGARE$ e $LPMILEU$, que ilustram a localização no Centro, na Guarda-Gare e na Póvoa do Mileu, respectivamente, verifica-se a perda de significância estatística para os correspondentes coeficientes estimados em cenário de avaliação baseada nas novas regras do IMI. Esta ocorrência

torna despicendo a observação dos sinais desses coeficientes, no entanto arrisca-se a apontar a mudança do sinal (-) para o sinal (+) das localizações, *LCNTRO* e *LPMILEU*, desenhando a hipótese de as localizações no Centro e na Póvoa do Mileu deixarem de serem elementos contribuidores para uma subavaliação e passarem a dispor de potencial para contribuir para a sobreavaliação patrimonial.

Finalmente para melhor comparar o desempenho dos dois métodos de avaliação e determinar se as diferenças entre os rácios de avaliação são estatisticamente significativos, vai-se testar a hipótese de que a variância dos dois rácios de avaliação é a mesma. O *F* - estatístico para o teste virá,

$$F(92;92) = \frac{(0,214426)^2}{(0,164263)^2} = 1,704$$

A hipótese é rejeitada para um intervalo de confiança de 99%, o que significa que a variabilidade do rácio de avaliação, para a base de avaliação IMI é significativamente diferente da variabilidade do rácio de avaliação, para a base de avaliação hedónica.

Finalmente se assumirmos que os valores de mercado das propriedades imobiliárias se podem inferir a partir dos valores de avaliação dessas propriedades utilizando um método traduzível pela equação,

$$VMP_n = \frac{VA_n}{E(RA)}, \text{ em que } VMP_n \text{ é o valor de mercado}$$

previsto para uma dada propriedade imobiliária n , VA_n é o valor de avaliação de uma qualquer propriedade imobiliária n e $E(RA)$ é o valor médio do rácio de avaliação, também se pode deduzir um rácio de avaliação a partir de VMP_n , assim,

$$RA_{VMP_n} = \frac{VMP_n}{VV_n}.$$

Comparando este rácio de avaliação, para os dois métodos de avaliação em confronto, pode-se determinar qual dos dois fornece melhor estimativa do valor de mercado. Assim, uma vez que o rácio de avaliação estimado está centrado na unidade,²⁰⁸ verifica-se que cerca de 97% dos valores de venda previstos a partir da avaliação hedónica se encontram no intervalo $[0,6;1,4]$, com um valor de 2,9 para o somatório dos quadrados do desvio padrão, contra 93,5% dos valores de venda previstos a partir da avaliação IMI para o mesmo intervalo e um valor de 5,6 para o somatório dos quadrados do desvio padrão. Caso o intervalo de variação do rácio de avaliação estimado diminua para $[0,8;1,2]$ os valores de venda previstos a partir da avaliação hedónica, aí situados, passam a 88% com um valor de 9,7 para o somatório dos quadrados do desvio padrão; enquanto os que decorrem da avaliação IMI passam a 81,5% com um valor de 13,9 para o somatório dos quadrados do desvio padrão.

Seja qual for o método de avaliação utilizado para inferir os valores de mercado dos imóveis uma realidade é permanente e prende-se com a pequena percentagem de

²⁰⁸ Devido à divisão pelo valor médio do rácio de avaliação.

observações fora do intervalo, quer dos 40% , quer dos 20% , da distribuição dos rácios de avaliação previstos; apenas os valores relativos ao somatório dos quadrados do desvio padrão denotam alguma variação relativa. Estes resultados apontam para a possibilidade dos valores relativos à avaliação das propriedades imobiliárias fornecerem informação importante sobre os valores de mercado, com alguma vantagem para os que decorrem da avaliação hedónica.

A presença, nas duas metodologias de avaliação, de enviesamentos consistentes, aponta para a necessidade de introduzir ajustamentos a esses enviesamentos de forma a melhorar o rigor das previsões dos valores de mercado. Trabalhar no domínio desses ajustamentos constituirá um esforço positivo na busca de um instrumento eficiente e económico para assegurar a regularidade dos processos de avaliação.

5.9. Considerações Finais

As novas regras de avaliação introduzidas pelo CIMI melhoram significativamente os valores avaliados do parque habitacional como foi revelado na secção anterior. O manuseamento da fórmula para apuramento do valor patrimonial tributário,²⁰⁹ dada a relativa uniformidade das características do parque habitacional, decorrente das regras de concorrência do mercado da habitação, revela o coeficiente de localização como o factor com maior capacidade de diferenciação dos valores patrimoniais.

Neste domínio acontece que, apesar da excepcional interacção com o Sistema de Informação Geográfica (*SIG*), no mesmo perímetro habitacional, a variabilidade do

²⁰⁹ A expressão para avaliação do valor patrimonial constante do artigo 38º traduz-se por, $V_t = V_c * A * C_a * C_l * C_q * C_v$, em que V_t é o valor patrimonial tributário, V_c o valor base dos prédios edificados, C_a o coeficiente de afectação, C_l o coeficiente de localização, C_q o coeficiente de qualidade e conforto e C_v o coeficiente de vetustez.

coeficiente de localização é uma realidade em ruas paralelas e perpendiculares, ou seja em ruas que oferecem as mesmas acessibilidades ao centro da cidade e às escolas e que apresentam uma envolvimento socioeconómica semelhante.

Algumas situações se podem identificar na cidade da Guarda - a título de exemplo, a Rua 31 de Janeiro com um coeficiente de localização de 1,3 e duas ruas secundárias adjacentes (Rua Paiva Couceiro e Rua do Sol) com o coeficiente de localização 1,1, enquanto uma terceira rua secundária, também adjacente, mantém o coeficiente 1,3; ainda, a Rua António Sérgio com um coeficiente de 1,3, entroncada pela Rua 5 de Outubro, uma paralela, com um coeficiente de localização de 1,1. As diferenças induzidas no valor patrimonial tributário por este coeficiente de localização, dificilmente, reflectem os valores decorrentes do mercado, podendo mesmo conduzir a enviesamentos perversos na avaliação patrimonial.

Entretanto o papel principal nesta avaliação pertencerá ao valor base dos prédios edificados (a propor anualmente) e aos coeficientes de afectação, localização, qualidade e conforto e vetustez caberá o papel de diferenciação dessa avaliação. Ocorre que, enquanto a informação relativa ao valor base dos prédios edificados é obtida anualmente por via da auscultação das entidades oficiais e das associações privadas do sector imobiliário urbano, a informação relativa ao zonamento e respectivos coeficientes de localização é trienal e baseia-se em propostas de peritos locais e regionais e das entidades representadas no CNAPU.²¹⁰ Ora, este procedimento é passível de introduzir duas velocidades²¹¹ na actualização dos parâmetros constitutivos da fórmula de

²¹⁰ A CNAPU, Comissão Nacional de Avaliação de Prédios Urbanos, ao ser constituída por um presidente e 10 vogais, pode colocar questões de operacionalidade, de funcionalidade e de celeridade no que respeita ao cumprimento dos prazos de apresentação das propostas relativas aos zonamentos e respectivos coeficientes e localização, e, ainda às directrizes relativas à apreciação da qualidade construtiva, da localização excepcional e do estado deficiente de conservação. Conforme artigos 60º, 61º e 62º do CIMI.

²¹¹ Não se pode negligenciar que, desde 1988, com a aprovação do Código da Contribuição Autárquica, a apresentação do Código de Avaliação ficou em aberto e várias Comissões foram trabalhando para a concretização desse objectivo, daí que a actual proposta relativa aos diversos coeficientes de avaliação equacione, naturalmente, conclusões dessas comissões – e propostas trienais não se contemporizam com tais espaços temporais.

apuramento do valor patrimonial tributário, podendo constituir-se num potencial embrião de inequidade fiscal.

As virtudes democráticas intrínsecas ao modelo de constituição da CNAPU e ao seu funcionamento não são discutíveis, porém os factores de inércia que daí podem advir não devem ser negligenciados.

A concretização da equidade fiscal exige uma (re)avaliação da base de incidência fiscal actualizada e, quer se siga uma perspectiva de vendas comparadas,²¹² e/ou uma perspectiva de custos,²¹³ e/ou uma perspectiva de rendimento,²¹⁴ o que importa é a celeridade e regularidade da sua execução. Neste âmbito os modelos de preços hedónicos, pese embora os ruídos econométricos que lhe estão adstritos, podem constituir-se como uma boa alternativa ao desenvolvimento de um programa de avaliação patrimonial,²¹⁵ capaz de utilizar os poderosos meios computacionais disponíveis. E, ainda, a seu favor o reconhecimento de potencial para captar apreciação/depreciação de valor decorrente de alterações ambientais e de políticas ambientais.²¹⁶

²¹² A perspectiva de vendas comparadas (*market valuation design*) compara a propriedade em análise (*subject property*) com outras (*comparable properties*) vendidas recentemente. Esta perspectiva baseia-se na ideia de que propriedades semelhantes se reflectem os mesmos serviços/benefícios para o comprador então devem ser avaliadas similarmente no mercado. O valor de mercado da propriedade em estudo determina-se pela comparação com propriedades semelhantes transaccionadas na mesma zona. Considerações a ter em conta para avaliação são: p.e., localização, dimensão/área, condições de habitabilidade, qualidade e tempo de venda.

²¹³ A perspectiva de custo (*cost approach design*) tem por base o custo corrente da terra, os custos de substituição dos edifícios e de outros equipamentos. Os custos de substituição são ajustados de forma a reflectirem a idade, as condições e a utilidade dos imóveis em apreciação, o valor da terra é determinado separadamente.

²¹⁴ A perspectiva do rendimento (*income capitalization*) baseia-se na ideia de que o valor da propriedade é determinado pela quantidade de rendimento que a propriedade é capaz de gerar num tempo futuro. Há que converter o valor futuro esperado em rendimento actual. Esta perspectiva é de importância crucial para avaliar propriedades industriais, comerciais e de investimento.

²¹⁵ É o caso dos sistemas CAMA, *Computer Assisted Mass Appraisal*, que suportam as actividades de (re)avaliação do património na existência de uma poderosa base de dados e respectiva gestão, bem como na análise do desempenho e na administração dos procedimentos de avaliação, generalizados a vários estados americanos, como New Hampshire, Kansas, Minnesota, Montana, entre outros.

²¹⁶ The Hypoport Group converteu-se ao cálculo hedónico do índice de preços da habitação, atribuindo-lhe maior flexibilidade para lidar com a volatilidade de preços decorrente das alterações ambientais e da legislação para defesa do ambiente. Por exemplo,

O centro nevrálgico de um sistema informático de avaliação patrimonial reside na base de dados, indispensável ao seu bom desempenho. É, aliás, a construção dessa base de dados que acaba por introduzir uma monitorização das actividades nos mercados imobiliários, permitindo a detecção de tendências de preços nesses mercados, que, por sua vez facilitam o ajustamento de transacções antigas aos dados da avaliação actual. A inconsistência e/ou omissão de dados relativos à venda de imóveis, bem como a extensão do seu domínio, pode fazer perigar o sucesso da concretização dessa base de dados, no entanto a ficha técnica de habitação das propriedades imobiliárias, respeitando o articulado do Decreto Lei nº 68 / 2004 , já implementada, poderia servir de um bom ponto de partida, após aperfeiçoamentos no sentido da sua simplicidade e funcionalidade.

A flexibilidade e os procedimentos estatísticos que decorrem de um sistema informático de avaliação patrimonial deslocam o centro do problema fiscal, passando a questão da equidade fiscal a ser substituída pela questão do custo/eficiência fiscal. Em termos conceptuais uma análise custo/eficiência do imposto patrimonial exige que todas as propriedades imobiliárias estejam virtualmente identificadas e que os procedimentos de avaliação e os erros daí decorrentes estejam minimizados, de tal forma que a coincidência entre previsão e execução de receita fiscal seja uma realidade e que os custos com a administração fiscal sejam mínimos.

reportam a pressão negativa sobre o preço das habitações nas zonas verdes circundantes às grandes cidades induzida pelas introdução de taxas verdes de acesso ao centro das cidades e, ainda, pela subida dos preços dos combustíveis.

Apêndice I

Valores relativos às avaliações realizadas, ordenados por idade, de forma decrescente

MATRIZ	VALORHEDONICO	ANOS	CORRMATRIZ	AREA
2573,80	72256,88	24	30370,84	110
799,07	57924,74	24	7423,36	118
1331,79	74317,36	23	12372,32	125
2513,94	75078,00	21	17220,48	124
2528,31	54386,05	19	11478,52	92
14149,25	55170,46	19	51044,14	80
22468,35	60443,41	18	79537,95	90
5858,38	79787,32	18	20738,66	130
2801,25	62174,48	18	9916,42	92
11043,39	53544,21	18	32467,56	90
849,35	53995,87	18	2497,08	90
632,38	53245,20	18	1694,77	90
5202,44	46113,59	18	18416,63	60
4808,74	75265,93	18	17022,93	128
1612,51	82398,28	18	13959,73	135
9800,78	56823,20	17	26266,09	90
976,40	74160,20	17	2616,75	120
3368,50	99337,94	16	9027,58	194
4994,21	77568,75	16	12185,87	124
6509,31	54982,10	16	17444,95	90
6256,42	86320,28	16	16767,20	130
7324,85	86230,82	15	8350,32	120
6334,23	86182,28	15	14061,94	130
17421,61	62735,51	15	38675,97	90
4393,79	50772,99	15	9754,21	82
9122,76	77661,89	15	20252,52	120
19527,94	56262,46	15	43352,02	80
16969,10	73129,19	15	37671,40	120
5532,92	72679,88	15	12283,08	120
4374,32	57472,56	15	9651,05	90
5207,45	110031,27	15	11560,53	130
16476,79	67709,58	15	40203,37	90
2304,45	57789,53	14	5115,85	90
3600,32	93077,05	14	12745,13	128
10935,65	118233,87	14	24277,14	120
5532,92	70807,48	13	10899,85	138
2880,56	80104,15	13	5098,59	128
5473,06	63036,10	13	9687,31	90
6625,28	60087,85	13	11726,74	95
8080,53	92237,30	12	14302,53	130
5858,38	54543,99	12	10369,33	94
8641,67	55838,56	12	13481,00	78
12420,07	104864,03	12	19375,30	153
15622,35	78520,88	11	22808,63	110
839,78	56565,97	11	1226,07	80
34970,72	88558,61	10	47210,47	122
2528,31	70032,86	10	3413,21	90

Valores relativos às avaliações realizadas, ordenados por idade, de forma decrescente

MATRIZ	VALORHEDONICO	ANOS	CORRMATRIZ	AREA
358,78	59299,90	10	523,81	90
7200,20	78379,67	10	12120,19	120
26501,13	90906,15	10	35776,52	118
18098,88	91292,13	10	24433,48	126
26501,13	94326,73	10	35776,52	130
32435,08	92891,08	10	43787,35	124
14978,90	51863,10	10	20221,51	80
15622,35	83518,20	9	19996,60	126
15951,56	75178,45	9	20417,99	120
11522,23	65131,45	9	14748,45	96
24735,39	79468,81	9	31661,29	118
11716,76	84518,01	9	12888,44	126
39265,37	88709,29	9	50259,67	118
26501,13	78730,26	9	35776,52	95
26501,13	85655,00	9	33921,44	118
26037,25	84531,78	8	33327,68	128
48819,84	88886,88	8	62489,40	125
25194,38	97209,13	8	30939,88	124
12674,45	82670,65	8	16223,29	90
21480,73	98133,38	8	27495,33	128
20163,90	95208,06	8	24801,59	130
29291,91	104007,67	8	36029,04	126
25060,85	97034,31	8	30824,84	125
31424,27	98398,70	7	38651,85	126
22445,91	95882,89	7	27608,46	120
34566,69	75734,03	7	41134,36	90
34566,69	70506,69	7	41134,36	90
25588,33	97306,38	7	31473,64	118
29291,91	102602,74	7	36029,04	130
25386,32	99520,59	6	29701,99	120
33743,68	97069,54	6	39480,10	130
25386,32	100651,55	5	29701,99	126
21473,25	78387,51	5	25123,70	92
31903,11	83718,89	5	37326,63	92
15338,04	76634,29	5	17485,36	90
28236,95	80831,33	5	31343,01	132
32546,56	99749,76	4	37103,07	130
34357,20	109504,39	4	39167,20	115
31005,28	106508,80	4	34415,86	132
37035,75	137546,85	3	39998,59	156
33773,61	99785,97	2	35124,55	126
40357,74	99520,59	2	43586,35	120
32052,75	79492,65	2	34616,97	92
41480,03	79739,46	1	41480,03	120
32546,66	109154,53	1	32546,66	134
29621,11	110031,27	1	29621,11	130

Bibliografia

Adams, Charles, 1993, FOR GOOD AND EVIL: THE IMPACT OF TAXES ON THE COURSE OF CIVILIZATION, May, Madison Books.

Alonso, William, 1964, LOCATION AND LAND USE, Cambridge: Harvard University Press.

Anglin, P. M. e R. Gençay, 1996, Semiparametric Estimation of a Hedonic Price Function, *Journal of Applied Econometrics*, Vol. 11, pp. 633–648.

Anselin, L. e A. K. Bera, 1998, Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics, in A. Ullah e D. E. A. Giles, eds, HANDBOOK OF APPLIED ECONOMIC STATISTICS, Marcel Dekker: New York.

Anselin, L. e S. Rey, 1991, Properties of Tests for Spatial Dependence in Linear Regression Models, *Geographical Analysis*, Vol. 23, pp. 112–131.

Anselin, L., 1988, SPATIAL ECONOMETRICS: METHODS AND MODELS, Kluwer: Dordrecht.

Anselin, L., J. Cohen *et al.*, 2000, Spatial Analysis of Crime, Measurement and Analysis of Crime and Justice, CRIMINAL JUSTICE, Vol. 4, pp. 213-262.

Arnott, R., 2005, Neutral Property Taxation, *Journal of Public Economic Theory*, Vol. 7 (1), pp. 27-50.

Atkinson, S. E. e T. D. Crocker, 1987, A Bayesian Approach to Assessing the Robustness of Hedonic Property Value Studies, *Journal of Applied Econometrics*, Vol. 2 (1), pp. 27-45.

Bajari, P. e M. Khan, 2005, Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities, *Journal of Business and Economics Statistics*, Vol. 23 (1), pp. 20-33.

Baltagi, B. W. e Dong Li, 2001, LM Tests for Functional Form and Spatial Error Correlation, *International Regional Science Review*, Vol. 24 (2), pp. 194-225.

Barro, Robert J, 1979, On the Determination of the Public Debt, *Journal of Political Economy*, Vol. 87 (5), pp. 940–971.

Barro, R. J, 1986, U.S. Deficits since World War I, *Scandinavian Journal of Economics*, Vol. 88 (1), pp. 195–222.

Bartik, T. J., 1988, Measuring the Benefits of Amenity Improvements in Hedonic Price Models, *Land Economics*, Vol. 64 (2), pp. 172–183.

Becker, G., 1965, A theory of the allocation of time, *Economic Journal*, Vol. 75 (299), pp. 493-519.

Belsley, D. A., 1991, CONDITIONING DIAGNOSTICS – COLLINEARITY AND WEAK DATA IN REGRESSION, John Wiley, New York.

Belsley, D. A., E. Kuh e R. Welsch, 1980, REGRESSION DIAGNOSTICS: IDENTIFYING DATA AND SOURCES OF COLLINEARITY, 1st edition, Wiley Series in Probability and Mathematical Series, John Wiley.

Benkard, C. L. e P. Bajari, 2005, Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach, *Journal of Political Economy*, Vol. 113 (6), 1239-1276.

Bentick, B. L., 1982, A Tax on Land Value May Not Be Neutral, *National Tax Journal*, Vol. 35 (1), pp. 113.

Berry, B. J. L. e R. S. Bednarz, 1975, A Hedonic Model of Prices and Assessments for Single Family Homes: Does the Assessor Follow the Market or the Market Follow the Assessor? *Land Economics*, Vol. 51 (1), pp. 21-50.

Black, D., 1977, Property Tax Incidence: The Excise-Tax Effect and Assessment Practices, *National Tax Journal*, Vol. 30 (4), pp. 429-434.

Bowman, J. H. e J. L. Mikesell, 1978, Uniform Assessment of Property: Returns from Institutional Remedies, *National Tax Journal*, Vol. 31 (2), pp. 137–152.

Broome, J., 1972, Approximate equilibrium in economies with indivisible commodities, *Journal of Economic Theory*, Vol. 5 (2), pp. 224-249.

Brueckner, J. K., 1986, A Modern Analysis of the Effects of Site Value Taxation, *National Tax Journal*, Vol. 39 (1), pp. 49–58.

Cassel, E. e R. Mendelsohn, 1985, The Choice of Funcional Forms for Hedonic Price Equations, *Journal of Urban Economics*, Vol. 18, pp. 135-142.

Cheshire, P. e S. Sheppard, 1995, On the Price of Land and the Value of Amenities, *Economica*, Vol. 62, pp. 247–267.

Cheshire, P. e S. Sheppard, 1997, Welfare Economics of Land Use Regulation, *Research Papers in Environmental and Spatial Analysis*, Vol. 42, London School of Economics.

Clapp, J. M., 1990, A New Test for Equitable Real Estate Tax Assessment, *Journal of Real Estate Finance and Economics*, Vol. 3 (9), pp. 233-249.

Court A. T, 1939, Hedonic price indexes with automotive examples, in THE DYNAMICS OF AUTOMOBILE DEMAND, papers presented at a joint meeting of the American Statistical Association and the Econometric Society, in Detroit, Michigan, 27 December 1938, General Motors Corporation, New York.

Cropper, M. L. *et al.*, 1988, On the Choice of Functional Form for Hedonic Price Functions, *The Review of Economics and Statistics*, Vol. 70 (4), pp. 668 –675.

Cropper, M. L. *et al.*, 1993, Valuing Product Attributes Using Single Market Data: A Comparison of Hedonic and Discrete Choice Approaches, *The Review of Economics and Statistics*, Vol. 75 (2), pp. 225–232.

Davidian, M. and , R. J. Carroll, 1987, Variance function estimation, *Journal of the American Statistical Association*, Vol. 82, pp. 1079–1091.

Day, B. H., 2001, The Theory of Hedonic Markets: Obtaining Welfare Measures for Changes in Environmental Quality Using Hedonic Market Data, *Report for the EU Working Group on Noise*.

De Viti de Marco, A., 1936, FIRST PRINCIPLES OF PUBLIC FINANCE, London, Jonathan Cape.

Debreu, G. e K. J. Arrow, 1954, Existence of an Equilibrium for a Competitive Economy, *Econometrica*, Vol. 22 (3), pp. 265-290.

Debreu, G., 1960, Topological Methods in Cardinal Utility Theory, in Arrow, Karlin and Suppes, eds, *Mathematical Methods in the Social Sciences*.

Díaz M. Á., M. G. Gómez e A. Álvarez, 2003, Using datadriven prediction methods in a hedonic regression problem, *Documento de Trabalho*, 0302 Universidade de Vigo, Departamento de Economía Aplicada.

Donaldson, R G. e M. Kamstra, 1996, A New Dividend Forecasting Procedure That Rejects Bubbles in Asset Prices: The Case of 1929's Stock Crash, *Review of Financial Studies*, Vol. 9 (2), pp. 333-383.

Donaldson, R. G. e M. Kamstra, 1997, An artificial neural network-GARCH model for international stock return volatility, *Journal of Empirical Finance*, Vol. 4 (1), pp. 17-46.

Downs, A., 1957, AN ECONOMIC THEORY OF DEMOCRACY, New York, Harper and Row.

Ekeland, I., J. Heckman e L. Nesheim, 2002, Identifying Hedonic Models, *AEA Papers and Proceedings: Market Implications of Peer and Neighborhood Effects*, Vol. 92 (2), pp. 304-309.

Ekeland, I., J. Heckman e L. Nesheim, 2004, Identification and Estimation of Hedonic Models, *Journal of Political Economy*, Vol. 112 (S1), S60-S109.

Ellickson, B., 1980, An alternative test of the Hedonic Theory of Housing Markets, *Journal of Urban Economics*, Vol. 9, pp. 56-79.

Engle, Robert, 1975, De facto Discrimination in Residential Assessment Boston, *National Tax Journal*, Vol. 28, pp. 445-451.

Epple, D., 1987, Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differential Products, *Journal of Political Economy*, Vol. 95 (1), pp. 59-80.

Fischel, W. A. e C. Campbell, 1996, Preferences for School Finance Systems: Voters versus Judges, *National Tax Journal*, Vol. 49, pp. 1-15.

Fischel, W. A., 1985, THE ECONOMICS OF ZONING LAWS: A PROPERTY RIGHTS APPROACH TO AMERICAN LAND USE CONTROLS, Baltimore: Johns Hopkins University Press.

Fischel, W. A., 1992, Communication: Property Taxation and the Tiebout Model: Evidence for the Benefit View from Zoning and Voting, *Journal of Economic Literature*, Vol. 30, pp. 171–177.

Fisher, G. W., 1996, THE WORST TAX? A HISTORY OF THE PROPERTY TAX IN AMERICA, Lawrence: University Press of Kansas.

Freeman, A. Myrick, 1993, The Measurement of Environmental and Resource Values: Theory and Methods, in RESOURCES FOR THE FUTURE, Washington D.C., pp. 104-105.

Gale, D. e H. Nikaido, 1965, The Jacobian matrix and global univalence of mappings. *Mathematische annalen*, Vol. 2, pp. 81-93.

George, Henry, 1879, PROGRESS AND POVERTY, Robert Schalkenbach Foundation, <http://www.schalkenbach.org/library/george.henry/ppcont.html>.

Goldberger, A. S., 1968, The Interpretation and Estimation of Cobb-Douglas Functions, *Econometrica* Vol. 36, pp. 464-472.

Goodman, A. e T. Thibodeau, 1995, Dwelling age heteroscedasticity in hedonic house price equations, *Journal of Housing Resources*, Vol. 6, pp. 25–42.

Goodman, A. e T. Thibodeau, 1997, Dwelling age heteroscedasticity in hedonic house price equations: an extension, *Journal of Housing Resources*, Vol. 8, pp. 299–317.

Goodman, A. e T. Thibodeau, 1998, Dwelling age heteroscedasticity in repeat sales house price equations, *Real Estate Economics*, Vol. 26, pp. 151–171.

Goolsby, W. C., 1997, Assessment Error in the Valuation of Owner-Occupied Housing, *Journal of Real Estate Research*, Vol. 13 (1), pp. 33–45.

Gouldemans, R. J., 2001, Confidence intervals for the coefficient of dispersion. *Assessment Journal*, Vol. 8 (6), pp. 23-27.

Greene, W., 2003, *ECONOMETRIC ANALYSIS*, 5th Edition, New York Macmillan.

Griliches, Z. (ed.), 1971, *PRICE INDEXES AND QUALITY CHANGE*, Cambridge, Massachusetts: Harvard University Press.

Gujarati, D., 1992, *ESSENTIALS OF ECONOMETRICS*, 1st Edition, McGraw-Hill International Editions.

Halvorsen, R. e H. Pollakowski, 1981, Choice of Functional Form for Hedonic Price Equations, *Journal of Urban Economics*, Vol. 10, pp. 37-49.

Hamilton, B. W., 1976, The Effects of Property Taxes and Local Public Spending on Property Values, *Journal of Political Economy*, Vol. 84 (3), pp. 647–650.

Harrison D. J. e D. L. Rubinfeld, 1978, Hedonic Housing Prices and the Demand for Clean Air, *Journal of Environmental Economics and Management*, Vol. 5, pp. 81–102.

Harriss, C. L., 1983, THE PROPERTY TAX AND LOCAL FINANCE, Academy of Political Sciences, New York.

Harvey, A.C. e G. D. A. Phillips, 1974, A comparison of the power of some tests for heteroscedasticity in the general linear model, *Journal of Econometrics* Vol. 2, pp. 307–316.

Hausman, J. A., 1978, Specification Tests in Econometrics, *Econometrica*, Vol. 46 (6), pp. 1251-1271.

Ihlanfeldt, K. R., 1982, Property Tax Incidence on Owner Occupied Housing: Evidence from the Annual Housing Survey, *National Tax Journal*, Vol. 35, pp. 89-97.

International Association of Assessing Officers (IAAO), 2003, STANDARD ON AUTOMATED VALUATION MODELS, Chicago.

International Association of Assessing Officers (IAAO), 2003, STANDARD ON ADMINISTRATION OF MONITORING AND COMPLIANCE RESPONSIBILITIES, Chicago.

International Association of Assessing Officers (IAAO), 1990, PROPERTY APPRAISAL AND ASSESSMENT ADMINISTRATION, Chicago.

International Association of Assessing Officers (IAAO), 1978, IMPROVING REAL PROPERTY ASSESSMENT: A REFERENCE MANUAL, Chicago, Illinois.

Johnson, H. G., 1958, Demand Theory Further Revised or Goods Are Goods, *Economica*, N. S. 25.

Jones, L. E., 1988, The Characteristics Model, Hedonic Prices, and the Clientele Effect, *Journal of Political Economy*, Vol. 96, pp. 551–567.

Judge, G., W. Griffiths *et al.*, 1988, INTRODUCTION TO THE THEORY AND PRACTICE OF ECONOMETRICS, John Wiley and Sons, New York.

Kick, Thomas, 2002, Determinants of One-Family House Prices in the Detroit Area: An Econometric Analysis Based on the Hedonic Price Approach, *Freie Wissenschaftliche Arbeitb zur Erlangung des Akademischen Grades "Diplom-Handelslehrer"*.

Kitchen, H., 1992, Property Taxation in Canada, *Canadian Tax Paper*, Vol. 92, Canadian Tax Foundation, Toronto.

Klein, L. R. e M. Nakamura, 1962, Singularity in the Equation Systems of Econometric: Some Aspects of the Problem of Multicollinearity, *International Economic Review*, Vol. 3, pp. 274-99.

Knight, R., R. C. Hill e C. F. Sirmans, 1993, Stein-rule estimation in real estate appraisal, *The Appraisal Journal*, pp. 539-544.

Kochin, L. A., e R. W. Parks, 1984, Testing for assessment uniformity: A reappraisal. *Property Tax Journal*, Vol. 3 (1), pp. 27–54.

Lancaster K. J., 1966, A New Approach to Consumer Theory, *Journal of Political Economy*, Vol. 74, pp. 132-157.

Lim, Duck-Ho, 1992, The Nonneutrality of the Land Value Tax: Impacts on Urban Structure, *Journal of Urban Economics*, Vol. 32 (2), pp. 186–194.

Lindahl, E., 1919, Just Taxation - A Positive Solution, English translation of portion of *Die Gerechtigkeit der Besteuerung*, in CLASSICS IN THE THEORY OF PUBLIC FINANCE, ed. by R. Musgrave and A. Peacock, New York: MacMillan, 1958, pp. 168-176.

Lykke, A. e E. Reis, 1997, Deforestation, Development, and Government Policy in the Brazilian Amazon: An Econometric Analysis, Trabalho desenvolvido no âmbito do NEMESIS, www.nemesis.org.br.

Mas-Colell, A., 1975, A Model of Equilibrium With Differentiated Commodities, *Journal of Mathematical Economics* Vol. 2, pp.263–295.

Mas-Colell, A., 1977, Indivisible Commodities and General Equilibrium Theory, *Journal of Economic Theory*, Vol. 16 (2), pp. 443-456.

Mieszkowski, P. e G. R. Zodrow, 1989, Taxation and the Tiebout Model: The Differential Effects of Head Taxes, Taxes on Land Rents, and Property Taxes, *Journal of Economic Literature*, Vol. 27 (3), pp. 1098-1146.

Mieszkowski, P., 1972, The Property Tax: An Excise Tax or a Profit Tax?, *Journal of Public Economic*, Vol. I (1), pp. 73–96.

Miller, N. G., 1982, Residential Property Hedonic Pricing Models: A Review, *Research in Real Estate*, Vol. 2, pp. 31–56.

Mills, D. E., 1981, The Non-Neutrality of Land Value Taxation, *National Tax Journal*, Vol. 34 (1), pp. 125–129.

Moran, P. A. P., 1950, Notes on continuous stochastic phenomena, *Biometrika*, Vol. 37, pp. 17-23.

Morishima, M., 1959, The Problem of Intrinsic Complementarity and Separability of Goods, *Metroeconomica*, Vol. XI.

Musgrave, R. e P. Musgrave, 1989 (1^{ed.} 1973), PUBLIC FINANCE IN THEORY AND PRACTICE, 5th ed., New York: McGrawHill.

Muth, R. F., 1966, Household Production and Consumer Demand Functions, *Econometrica*, Vol. 34, pp. 699–708.

Muth, R. F., 1969, CITIES AND HOUSING, Chicago: University of Chicago Press.

Myers, R.H., 1990, CLASSICAL AND MODERN REGRESSION WITH APPLICATIONS, 2nd ed., PWS-KENT, Boston.

Netzer, D., 1966, ECONOMICS OF THE PROPERTY TAX, Brookings Institute.

Netzer, D., 1980, REAL PROPERTY TAX POLICY FOR NEW YORK CITY, New York University.

Oates, W., 1999, An Essay on Fiscal Federalism, *Journal of Economic Literature*, Vol. 37, pp. 1120-1149.

Oldman, O., e H. Aaron, 1965, Assessment-Sales Ratios under the Boston Property Tax. *National Tax Journal*, Vol. 18 (1), pp. 37-45.

Olson, M., 1971, THE LOGIC OF COLLECTIVE ACTION, Cambridge, Harvard University.

Pace, R.K., 1995, Parametric, semiparametric, and nonparametric estimation of characteristic values within mass assessment and hedonic pricing models, *Journal of Real Estate Finance and Economics*, Vol. 11, pp.195-217.

Paglin, M. e M. Fogarty, 1972, Equity and the Property Tax: A New Conceptual Focus, *National Tax Journal*, Vol. 25, pp. 557-565.

Palmquist, R. B., 1984, Estimating The Demand For The Characteristics Of Housing, *The Review of Economics and Statistics*, Vol. 66 (3), pp. 394-404.

Pantaleoni, M., 1883, Contributo alla teoria del riparto delle spese pubbliche, Partial English translation quoted in CLASSICS IN THE THEORY OF PUBLIC FINANCE Musgrave, ed. by R. A. Musgrave e A. T. Peacock, A.T., 1967, London: Macmillan.

Pardal, Sidónio C., 1998, ESTUDO SOBRE A CONTRIBUIÇÃO AUTÁRQUICA, Ministério das Finanças.

Pareto, V., 1906, English transl., 1971, MANUAL OF POLITICAL ECONOMY, Edited by Ann S. Schwier and Alfred N. Page, New York: A.M. Kelley.

Park, H. e S. Y. Jyoung, 2002, Parametric and Semiparametric Estimation of Land Value, presented at AREUEA/ASRES, International Meeting, Seoul, 3-6 Julho.

Robinson, P., 1988, Root N Consistent Semiparametric Regression, *Econometrica*, Vol. 56, pp. 931-954.

Rosen, S., 1974, Hedonic prices and implicit markets: Product differentiation in pure competition, *Journal of Political Economy*, Vol. 82, pp. 34–55.

Samuelson, P. A., 1954, The Pure Theory of Public Expenditures. *Review of Economic and Statistics*, Vol. 36, pp. 387–389.

Santos, M. F. e C. Azevedo, 2005, DATA MINING – DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS, FCA: Lisboa.

Scott, D. W., 1992, MULTIVARIATE DENSITY ESTIMATION, John Wiley & Sons, New York.

Stone, M., 1974, Cross-validators choice and assessment of statistical predictions (with discussion), *Journal of Royal Statistic Society, B*, Vol. 36, pp. 111-146.

Stone, M., 1977, An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of Royal Statistic Society, B*, 39, pp. 44-47.

Tauchen, H. e A. D. Witte, 2001, Estimating Hedonic Models: Implications of The Theory, *Technical Working Paper 271*, NBER.

Thorsnes, P. e D. P. McMillen, 1998, Land Value and Parcel Size: A Semiparametric Analysis, *The Journal of Real Estate Finance and Economics*, Vol. 17(3), pp. 233-244.

Tideman, N., 1982, A Tax on Land is Neutral, *National Tax Journal*, Vol. 35, pp. 109-111.

Tiebout, C., 1956, A Pure Theory of Local Expenditures, *The Journal of Political Economy*, Vol. 64 (5), pp. 416-424.

Tinbergen, Jan, 1956, On The Theory of Income Distribution, *Weltwirtschaftliches Archiv*, Vol. 77, pp. 155-173.

Tinch, R., 1995, Valuation of Environmental Externalities, *Report of Department of Transport*, London.

Uzawa, H., 1960, Preference and Rational Choice in the Theory of Consumption, in K. J. Arrow, S. Karlin and P. Suppes (eds.), MATHEMATICAL METHODS IN THE SOCIAL SCIENCES, Stanford.

Vickrey, W., 1970, Defining Land Value for Tax Purposes, in D. M. Holland, ed., THE ASSESSMENT OF LAND VALUE, Madison: University of Wisconsin Press, pp. 25-36.

Wassmer, R. W., 1993, Property Taxation, Property Base, and Value: An Empirical Test of “The New View”, *National Tax Journal*, Vol. 46 (2), June, pp. 135–159.

Wicksell, K., 1896, A New Principle of Just Taxation, in CLASSICS IN THE THEORY OF PUBLIC FINANCE, ed. by R. A. Musgrave and Alan T. Peacock, 1958, London: Macmillan, pp. 72-118.

Wildasin D. E., 1982, More on the Neutrality of Land Taxation, *National Tax Journal*, Vol. 35 (1), pp. 105-108.

Witte, A., H. Sumka e H. Ereckson, 1979, An Estimate of a Structural Hedonic Price Model of the Housung Market: An Application of Rosen’s Theory of Implicit Markets, *Econometrica*, Vol. 47 (5), pp.1151-1173.

Yatchew, A, 1998, Nonparametric Regression Techniques in Economics, *Journal of Economic Literature*, Vol. XXXVI, June, pp. 669 - 721.

Zodrow, G. R., 2001, The Property Tax as a Capital Tax: A Room with Three Views, *National Tax Journal*, Vol. 54 (1), pp. 139–156.

PARTE II

DETERMINANTES ECONÓMICAS NO SUCESSO POLÍTICO DOS GOVERNANTES

Palavras Chave: Ciclos Políticos Orçamentais; Ciclos Político-Económicos; Reeleição; Manipulação Fiscal; Modelo Agente Principal; Reputação; Sinalização; Tomada de Decisão Política.

JEL's Classification System for Journal Articles:

C51 - Model Construction and Estimation

C52 - Model Evaluation and Selection

C53 - Forecasting and Other Model Applications

C7 - Game Theory and Bargaining Theory

D7 - Analysis of Collective Decision-Making

D78 - Positive Analysis of Policy-Making and Implementation

D82 - Asymmetric and Private Information

D83 - Search; Learning; Information and Knowledge; Communication; Belief

Prefácio

A probabilidade de um político ser reeleito, dirigida, particularmente, para os ciclos políticos nos governos das jurisdições locais, constitui o objectivo do trabalho desenvolvido neste ensaio. Na linha perspectivada por Baleiras (1997) a probabilidade de sucesso eleitoral é endógena, dado vir função da avaliação, ex-ante, feita pelo eleitorado ao desempenho futuro do político no poder, e, ainda que contenha uma componente exógena de probabilidade, mesmo essa acaba induzida pela componente endógena, na medida em que toma valores pertencentes a um domínio desenhado precisamente pela componente endógena. Esta modelização, implicitamente, sugere a utilização de indicadores observáveis para avaliar a competência do político no poder.

A probabilidade de reeleição, nos modelos empíricos de ciclo político-económico, tem sido equacionada como uma variável *proxy*, uma variável de recandidatura dicotómica, que pelas suas características poderá gerar sintomas de multicolinearidade nesses modelos. A interpretação dos resultados, daí decorrentes, deve ser cautelosa, particularmente, no que concerne à proposição relativa à probabilidade de reeleição e o ciclo de despesa política. Uma forma de contornar este problema é construir uma medida directa de probabilidade de reeleição que substitua a variável *proxy*, já referida. A sua concretização permitiria conciliar a perspectiva subjacente aos modelos de ciclos político - económicos (as variáveis económicas explicadas por variáveis políticas) com a perspectiva intrínseca aos modelos de ciência política (a explicação para os resultados eleitorais advém de variáveis económicas exógenas).

Baleiras (1997, 2004) aponta para a possibilidade de se estimar directamente a probabilidade de reeleição nos modelos de ciclos político-económicos, aplicando técnicas de estimação *probit*. O cruzamento desta linha de investigação com os resultados encontrados por Veiga e Veiga (2004a, 2004b e 2004c),¹ configura a

¹ Sustentam a hipótese de Anderson (2000) e Powell e Witten (1993) de que o desempenho da actividade política é avaliado tendo em conta o poder e responsabilidade das entidades políticas sobre a política económica.

moldura que se pretende desenvolver, no âmbito do modelo de ciclos político-económicos, em termos do poder local português.

O enquadramento teórico para o desenvolvimento de um modelo que suporte a análise empírica que nos permita inferir resultados que concretizem o objectivo proposto para este ensaio, conduz-nos para a prossecução de uma linha metodológica que interpreta os ciclos políticos económicos como parte de um jogo de sinais;² onde a prioridade fundamental vai para a assumpção de informação assimétrica nesse jogo de sinais.

Importa equacionar a informação assimétrica nesse jogo de sinais, já que os votantes ao não observarem a competência do político incumbente, nem o carácter desses políticos, enfrentam uma dupla incerteza. Ora é esta incerteza que constitui o objecto a esculpir pelo incumbente, o político no “poder”, de forma a melhorar as suas possibilidades de reeleição.

Estas preocupações com a reeleição podem impelir a que um político competente não implemente uma política socialmente desejável, mas tão só aquela política que lhe permite resolver a sua preocupação com o futuro; bem como podem evitar que um político incompetente concretize uma política socialmente indesejável em prol de uma política que lhe garanta solucionar o respectivo futuro político. De certa forma, a necessidade de modelar essa incerteza faz com que as eleições funcionem como disciplinadoras do comportamento dos incumbentes, empurrando-os para actuações tendencialmente “niveladas” para comportamentos no quadro do institucional e do tradicionalmente concretizado e, sobretudo visualizadas pelos eleitores como a “obra” do político.

Este é o quadro de um equilíbrio “politicamente correcto”, que se escolheu modelizar com o objectivo de se esboçar um modelo teórico que sustentasse o propósito de se alcançar valores estimados para a probabilidade de reeleição dos autarcas portugueses. Ora, a arquitectura desse modelo teórico que aparece desenvolvida no capítulo 2, aquando da aplicação ao ciclo político-económico, vai buscar fundamentação teórica ao comportamento dos políticos e votantes no quadro de um jogo de sinais apresentado,

² Para Rogoff (1990) e Lohmann (1998) os ciclos sinalizam a competência do incumbente, percebida pelos votantes racionais.

previamente, no capítulo 1. Posteriormente, procura-se, com base num *trade-off* entre a informação obtida a partir da base de dados e os preceitos das fundamentações teóricas do modelo, apresentar, no capítulo 3, um modelo denominado de econométrico que suporte uma análise empírica que permita uma razoável inferência estatística que valide reflexão político-económica que se possa desenvolver sobre o processo eleitoral nos municípios portugueses. A avaliação empírica e os requisitos técnicos para a opção e estimação do modelo econométrico que sustenta esse ensaio serão objecto do capítulo 4. Nesse capítulo far-se-á, também, a apresentação dos testes de inferência estatística utilizados para avaliação do modelo econométrico. Finalmente, o capítulo 5, procura perspectivar a reflexão político-económica desenvolvida no capítulo 4, na óptica de sistemas alternativos de regras pelos quais os cidadãos, por via da votação, podem directamente agregar as suas preferências individuais em escolhas sociais e nesse enquadramento a escolha como um comportamento ou decisão de comportamento que tem subjacente a selecção de uma das alternativas disponíveis alicia equacionar o objectivo deste ensaio através de um modelo de escolha binária, recorrendo-se à estimação dos valores para a probabilidade de reeleição dos autarcas portugueses pelo método *logit*.

1. Comportamento dos Governantes no Contexto de um Jogo de Sinais

*Voici mon secret. Il est très simple: on ne voit bien qu'avec le coeur.
L'essentiel est invisible pour les yeux.
Le Petit Prince
Antoine de Saint Exupéry (1943)*

1.1. Introdução

Uma importante tradição na teoria da democracia está na concepção do acto eleitoral como um mecanismo de responsabilização política. Nesta perspectiva, as eleições constituiriam meios de sancionamento que induziriam os políticos a prosseguir políticas a contento dos eleitores. Porém, os votantes não necessitam do acto eleitoral para exercerem o direito à responsabilização dos governantes, em vez disso podem ver nesse acto a possibilidade de escolherem um político competente, um que se comportasse independentemente dos incentivos à reeleição. Obviamente que este acto de selecção tem implícita a sanção. Seleccionar e sancionar não são actos incompatíveis ou mutuamente exclusivos, a selecção de um político competente é, implicitamente, a sanção do político incompetente, facto que incentivará os incompetentes a pretender aparecer como competentes – para o efeito orientam as suas escolhas políticas na direcção pretendida pelo eleitorado, tal como na teoria clássica de responsabilização.

O mecanismo de selecção, na medida em que os políticos competentes e incompetentes se confundem, torna-se mais problemático. A interacção deste mecanismo com o mecanismo de sanção acaba reflectida na componente estratégica, pelo menos no que concerne ao comportamento dos políticos.

O efeito da reeleição no comportamento estratégico do governante traz para a actualidade questões relativas a preocupações com a carreira. Quer Barro (1973), quer Ferejohn (1986) desenvolvem modelos em que os eleitores, no final de cada período

decidem se reelegem ou não o incumbente, com base no desempenho geral percebido. A reeleição induz o governante a esforçar-se mais para obter um melhor desempenho, daí a sua preocupação em emitir sinais que possam ser percebidos correctamente pelos eleitores. Estes sinais resultam de decisões não triviais, daí serem emitidos entre períodos pré e pós eleitorais, para responderem à ambição de reeleição. A sequência temporal implícita à emissão de sinais perspectiva-se numa variante dos modelos de ciclos político-económicos.

Ora, assim sendo, talvez mereça a pena começar por situar o problema da reeleição do governante no âmbito de um jogo de sinais, em que se tem de um lado os votantes que dispõem de informação imperfeita sobre o tipo de político que os governa e do outro lado políticos que escolhem as respectivas estratégias de acção com o objectivo de conseguirem ser reeleitos e, a partir daí, encontrar o equilíbrio do jogo – situação que ocorrerá quando as estratégias óptimas das decisões políticas são, mutuamente, consistentes com as crenças dos votantes.

Para o efeito procedeu-se a uma adaptação dos trabalhos, de Besancenot *et al.* (2009), de Krishna (2007) e de Armstrong-Taylor (2005), desenvolvidos no cenário da ciência empresarial, à ciência política no intuito de modelar os comportamentos de políticos e de eleitores num jogo de “sinais”. Adaptação, calibrada posteriormente, por confrontação com o modelo de ciclo político-económico desenvolvido por Wrasai, P. (2005). Para que essa adaptação se concretize passa-se, desde já, a apresentar os pressupostos básicos e as definições fundamentais para analisar o equilíbrio desse jogo.

1.2. Modelo

1.2.1. Pressupostos básicos

Há dois tipos de políticos, os competentes tipo H , com um valor fundamental (o qual se pode identificar com o rendimento explícito do incumbente mais as compensações

por estar e por se manter no “poder”) R ,³ em que $R > 0$ e os incompetentes tipo L , com valor fundamental igual a zero, $R = 0$.

O número total de políticos é normalizado à unidade e a frequência de políticos competentes na população total de políticos recebe a notação q .

Os votantes admitem-se homogéneos nas preferências por políticos com elevada capacidade governativa, daí que tenham sempre a opção de substituir o incumbente pelo opositor.⁴

Os votantes são, apenas, capazes de avaliar o desempenho dos governantes a partir do seu próprio bem-estar, já que só conseguem observar se as políticas propostas são (ou não) implementadas. A incapacidade do votante em verificar e/ou prever as consequências dessas políticas impede uma avaliação a partir do bem-estar social. A estratégia dos votantes é fornecida a partir de um limite ideal de bem-estar fixado, *cutoff level*, k , e apenas para níveis de satisfação, $U_V(T, e) > k$, há reeleição do incumbente.⁵

Em que o argumento T traduz o tipo de político e por isso definir-se como $T \in \{L, H\}$ e o argumento e identifica o esforço desenvolvido pelo político.

Os votantes, que se assumem neutros ao risco, conhecem a distribuição relativa ao tipo de políticos,⁶ mas não conhecem o tipo de cada político e o político conhece o valor do seu próprio tipo.

³ Este valor fundamental, R , que acontece por se estar no poder, fundamenta potenciais comportamentos oportunistas. Quanto maior for R , maior a motivação para se estar no poder, o que decorre directamente da função de utilidade do político ser monótona crescente em R ; e, por isso, também, uma variação do esforço na mesma proporção que o valor R .

⁴ O opositor não desempenha qualquer papel activo no jogo. A importância do opositor reside na possibilidade de permitir identificar uma força alternativa. Qualquer votante, seja qual for o envolvimento com o incumbente, sabe sempre que há uma alternativa a esse poder.

⁵ O limiar k gera mais esforço e maior desempenho por parte do político, funcionando por isso, como um incentivo ao político. Assim, quanto maior k , melhor, mas dentro de limites pois se esse ideal assumir valores irrealistas será impossível satisfazê-lo

⁶ Acreditam que segue uma distribuição normal, $T \sim N(\mu, \sigma_T^2)$.

Os políticos, também, são neutros ao risco. A generalização deste pressuposto a todos os intervenientes no jogo permite que questões de partilha de risco entre votantes e políticos não se coloquem. E, ainda que, um cenário divergente pudesse ser mais real, mesmo assim torna-se desnecessário complicar a análise presente.

O espaço temporal que suporta a interacção dos participantes no jogo e a racionalidade subjacente às respectivas decisões de optimização exige o equacionamento da taxa de desconto, ζ . Ora, sem perda de generalidade assume-se, para todos os intervenientes no jogo, uma taxa de desconto unitária, reveladora de que os jogadores não são impacientes.

Os políticos durante um período de tempo relevante devem fornecer aos votantes informação⁷ sobre o seu desempenho e os sinais emitidos podem decorrer de, apenas, um esforço natural (e_n) do político ou de um esforço oportunista (e_o). Um esforço natural tem um custo $c(e_n)$ e sinalizará com uma precisão $p \in]\frac{1}{2}, 1[$ um sinal $\tau \in \{l, h\}$ relativamente ao tipo de políticos, formalmente ter-se-á,

$$\Pr[l|L, e_n] = \Pr[h|H, e_n] = p$$

e

$$\Pr[l|H, e_n] = \Pr[h|L, e_n] = 1 - p$$

Um esforço oportunista custa $c(e_o)$, em que $c(e_o) > c(e_n)$, qualquer que seja o tipo de político. Neste caso o sinal fornecido será sempre h , de competente, implicando uma

⁷ O político no poder pode utilizar a vantagem de informação que detém sobre os eleitores, o que lhe permite reduzir o esforço político sempre que a conjuntura económica e social é favorável.

probabilidade unitária. Seja qual for o tipo de político que opte por este esforço ter-se-á, formalmente,

$$\Pr[h|H, e_o] = \Pr[h|L, e_o] = 1$$

e

$$\Pr[l|L, e_o] = \Pr[l|H, e_o] = 0$$

A sequência das decisões acontecerá assim,

- No momento zero, $t = 0$, a natureza determina um dos dois tipos de políticos, ou o político competente (motivado pelo bem estar social) com uma frequência q , ou o político incompetente (motivado pela posição) com uma frequência $1 - q$.
- No momento um, $t = 1$, o político decide se vai desenvolver um esforço natural (e_n) ou um esforço oportunista (e_o).
- No momento dois, $t = 2$, a natureza decide se o esforço natural é percebido como um sinal verdadeiro ou como um sinal falso. Já o esforço oportunista emitirá sempre um sinal de competência, h . Os votantes observam os sinais e formam as suas expectativas relativamente à competência dos políticos.

O que pode ser ilustrado através da árvore de decisão que se segue,

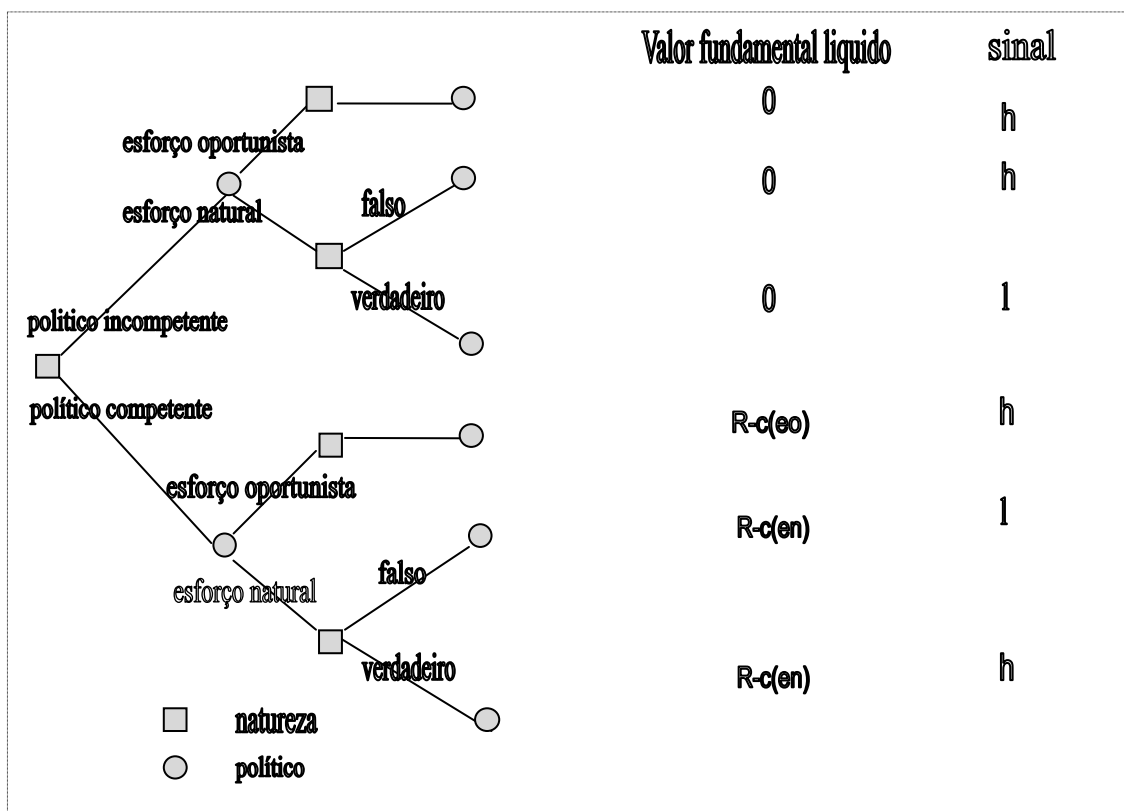


Figura 1: Árvore de Decisão

O valor fundamental de um político incompetente é zero, qualquer que seja o custo do esforço dispendido, o valor líquido do político só não se torna negativo pelo facto dos políticos pertencerem a partidos que detêm responsabilidade no quadro institucional e estarem preocupados com a respectiva reputação.

Como $c(e_o) > c(e_n)$, decorre que o valor líquido fundamental de um político competente é superior quando recorre a um esforço natural.

No fim do jogo os votantes observam um sinal que tanto pode significar político competente, como político incompetente. Como mostra a figura o sinal não permite inferir acerca do ponto terminal do jogo. Veja-se, por exemplo, um sinal de competência pode resultar de um esforço natural de um político competente, mas também pode resultar de um esforço oportunista de um político incompetente.

1.2.2. Probabilidades condicionais

Começa-se por assumir que os votantes, no que concerne ao esforço desenvolvido pelo político, têm, *a priori*, a crença de que é um esforço oportunista, independentemente de serem ou não competentes. Neste cenário ter-se-á, sempre, a crença da competência, h , de um político ou da incompetência, l , dado o esforço real desenvolvido, expressa por,

$$\pi_H^{e_o} = \Pr[e_o | H] \quad (1.2.1)$$

$$\pi_L^{e_o} = \Pr[e_o | L] \quad (1.2.2)$$

Uma vez assumidas as crenças relativas à competência dos políticos, pode-se, determinar $\Pr[\tau | T]$, isto é a probabilidade condicional de que o sinal é $\tau \in \{l, h\}$, quando os políticos são do tipo $T \in \{L, H\}$:⁸

$$\begin{aligned} \Pr[h | H] &= \Pr[h, e_o | H] + \Pr[h, e_n | H] \\ &= \Pr[h | e_o, H] \Pr[e_o | H] + \Pr[h | e_n, H] \Pr[e_n | H] \\ &= \pi_H^{e_o} + p(1 - \pi_H^{e_o}) \\ &= p + (1 - p)\pi_H^{e_o} \end{aligned} \quad (1.2.3)$$

⁸ Partindo da regra de Bayes tem-se: $\Pr[A, B, C] = \Pr[A | B, C] \Pr[B, C] = \Pr[A | B, C] \Pr[B, C] \Pr[C] =$
 $= \Pr[A, B | C] \Pr[C] \Leftrightarrow \Pr[A | B, C] \Pr[B, C] = \Pr[A, B | C].$

$$\begin{aligned}
\Pr[l|H] &= \Pr[l, e_o|H] + \Pr[l, e_n|H] \\
&= \Pr[l|e_o, H] \Pr[e_o|H] + \Pr[l|e_n, H] \Pr[e_n|H] \\
&= (1-p)(1-\pi_H^{e_o})
\end{aligned} \tag{1.2.4}$$

$$\begin{aligned}
\Pr[h|L] &= \Pr[h, e_o|L] + \Pr[h, e_n|L] \\
&= \Pr[h|e_o, L] \Pr[e_o|L] + \Pr[h|e_n, L] \Pr[e_n|L] \\
&= \pi_L^{e_o} + (1-p)(1-\pi_L^{e_o}) \\
&= (1-p) + p\pi_L^{e_o}
\end{aligned} \tag{1.2.5}$$

$$\begin{aligned}
\Pr[l|L] &= \Pr[l, e_o|L] + \Pr[l, e_n|L] \\
&= \Pr[l|e_o, L] \Pr[e_o|L] + \Pr[l|e_n, L] \Pr[e_n|L] \\
&= p(1-\pi_L^{e_o})
\end{aligned} \tag{1.2.6}$$

Finalmente, está-se em condições de determinar a probabilidade dos votantes avaliarem um político como sendo do tipo $T \in \{L, H\}$ quando observam o sinal $\tau \in \{l, h\}$. De novo, pela aplicação da regra de Bayes, obtêm-se as expressões de $\Pr[T|\tau]$ que se seguem,

$$\begin{aligned}
\Pr[H|h] &= \frac{\Pr[h|H]\Pr[H]}{\Pr[h|H]\Pr[H] + \Pr[h|L]\Pr[L]} = \\
&= \frac{q[p + (1-p)\pi_H^{e_o}]}{q[p + (1-p)\pi_H^{e_o}] + (1-q)[(1-p) + p\pi_L^{e_o}]}
\end{aligned}
\tag{1.2.7}$$

$$\begin{aligned}
\Pr[L|h] &= \frac{\Pr[h|L]\Pr[L]}{\Pr[h|H]\Pr[H] + \Pr[h|L]\Pr[L]} = \\
&= \frac{(1-q)[(1-p) + p\pi_L^{e_o}]}{q[p + (1-p)\pi_H^{e_o}] + (1-q)[(1-p) + p\pi_L^{e_o}]}
\end{aligned}
\tag{1.2.8}$$

$$\begin{aligned}
\Pr[H|l] &= \frac{\Pr[l|H]\Pr[H]}{\Pr[l|H]\Pr[H] + \Pr[l|L]\Pr[L]} = \\
&= \frac{q(1-p)(1-\pi_H^{e_o})}{q(1-p)(1-\pi_H^{e_o}) + (1-q)p(1-\pi_L^{e_o})}
\end{aligned}
\tag{1.2.9}$$

$$\begin{aligned}
\Pr[L|l] &= \frac{\Pr[l|L]\Pr[L]}{\Pr[l|H]\Pr[H] + \Pr[l|L]\Pr[L]} = \\
&= \frac{(1-q)p(1-\pi_L^{e_o})}{q(1-p)(1-\pi_H^{e_o}) + (1-q)p(1-\pi_L^{e_o})}
\end{aligned}
\tag{1.2.10}$$

O valor do desempenho do político, numa dada jurisdição λ , num dado período de tempo t , define-se por $x_{\lambda t}$ e relaciona-se com o sinal retirado do esforço desenvolvido pelo político. Assim quando os votantes observam um dado sinal, o valor do desempenho político esperado vem,

(i) para um sinal de incompetência do político, o que só acontece a um político que desenvolve um esforço político natural,

$$\begin{aligned}
 E[x_{\lambda t} | l] &= (R - c(e_n)) \Pr[H | l] = \\
 &= (R - c(e_n)) \frac{q(1-p)(1-\pi_H^{e_o})}{q(1-p)(1-\pi_H^{e_o}) + (1-q)p(1-\pi_L^{e_o})}
 \end{aligned}
 \tag{1.2.11}$$

(ii) para um sinal de competência do político que pode resultar, quer de um esforço natural, quer de um esforço oportunista,

$$\begin{aligned}
 E[x_{\lambda t} | h] &= (R - c(e_n)) \Pr[H, e_n | h] + (R - c(e_o)) \Pr[H, e_o | h] \\
 &= \frac{q \left\{ \left[p + (1-p)\pi_H^{e_o} \right] (R - c(e_n)) p(1-\pi_H^{e_o}) - c(e_o) \pi_H^{e_o} \right\}}{q \left[p + (1-p)\pi_H^{e_o} \right] + (1-q)[(1-p) + p\pi_L^{e_o}]}
 \end{aligned}
 \tag{1.2.12}$$

1.2.3. Estratégia do político

O político procura que o seu desempenho, $x_{\lambda t}$, seja o mais elevado possível, ainda que condicionado pelo tipo de político que é, ou seja por $T \in \{L, H\}$, e pelo esforço desenvolvido, ou seja $e_{\lambda} \in \{e_n, e_o\}$, o que pode ser traduzido por, $E_P[x_{\lambda t} | T, e_{\lambda}]$.

Conhecedores deste legítimo objectivo do político vamos procurar compreender as diversas estratégias, passíveis de definir, em função do tipo do político. Importa, no entanto, inventariar os valores esperados do desempenho, para os diversos tipos e para os diversos esforços desenvolvidos, assim,

(1) o político, que decide desenvolver um esforço oportunista, qualquer que seja o seu tipo de competência, sinalizará sempre competência elevada, h , portanto com uma probabilidade de ocorrência unitária. Vindo,

$$E_P[x_{\lambda t} | T, e_o] = E[x_{\lambda t} | h] \quad (1.2.13)$$

(2) o político incompetente, que decide desenvolver um esforço natural, defrontar-se-á com uma probabilidade p de assinalar a sua incompetência, mas também poderá produzir um sinal de competência em $1 - p$ das ocorrências. Vindo,

$$\begin{aligned} E_P[x_{\lambda t} | L, e_n] &= E[x_{\lambda t} | l] \Pr[l | L, e_n] + E[x_{\lambda t} | h] \Pr[h | L, e_n] \\ &= pE[x_{\lambda t} | l] + (1 - p)E[x_{\lambda t} | h] \end{aligned} \quad (1.2.14)$$

(3) o político competente, que decide desenvolver um esforço natural, submete-se a uma probabilidade p de assinalar competência para os eleitores, mas também a uma possibilidade $1 - p$ do sinal ser percebido como incompetência. Vindo,

$$\begin{aligned}
 E_P[x_{\lambda t} | H, e_n] &= E[x_{\lambda t} | h] \Pr[h | H, e_n] + E[x_{\lambda t} | l] \Pr[l | H, e_n] \\
 &= pE[x_{\lambda t} | h] + (1 - p)E[x_{\lambda t} | l]
 \end{aligned}
 \tag{1.2.15}$$

Estando, agora, reunidos os requisitos necessários para se prever as regras de decisão dos políticos, vamos observar a condição necessária para que o político incompetente opte por um esforço oportunista e que é passível de se traduzir por,

$$\begin{aligned}
 E_P[x_{\lambda t} | L, e_o] &> E_P[x_{\lambda t} | L, e_n] \\
 \Leftrightarrow E[x_{\lambda t} | h] &> (1 - p)E[x_{\lambda t} | h] + pE[x_{\lambda t} | l] \\
 \Leftrightarrow E[x_{\lambda t} | h] &> E[x_{\lambda t} | l]
 \end{aligned}
 \tag{1.2.16}$$

Induzir-se-ia, para o político incompetente, a seguinte estratégia ótima, desenhada sobre um esforço,

$$\left\{ \begin{array}{l} e_o, \text{ se } E[x_{\lambda t} | h] > E[x_{\lambda t} | l] \\ e_o \text{ ou } e_n, \text{ se } E[x_{\lambda t} | h] = E[x_{\lambda t} | l] \\ e_n, \text{ se } E[x_{\lambda t} | h] < E[x_{\lambda t} | l] \end{array} \right. \quad (1.2.17)$$

Similarmente, a condição necessária para que o político competente opte por desenvolver um esforço oportunista, expressa-se por,

$$\begin{aligned} E_P[x_{\lambda t} | H, e_o] &> E_M[x_{\lambda t} | H, e_n] \\ \Leftrightarrow E[x_{\lambda t} | h] &> pE[x_{\lambda t} | h] + (1-p)E[x_{\lambda t} | l] \\ \Leftrightarrow E[x_{\lambda t} | h] &> E[x_{\lambda t} | l] \end{aligned} \quad (1.2.18)$$

A estratégia óptima do político competente implicaria

$$\text{um esforço,} \quad \left\{ \begin{array}{l} e_o, \text{ se } E[x_{\lambda t} | h] > E[x_{\lambda t} | l] \\ e_o \text{ ou } e_n, \text{ se } E[x_{\lambda t} | h] = E[x_{\lambda t} | l] \\ e_n, \text{ se } E[x_{\lambda t} | h] < E[x_{\lambda t} | l] \end{array} \right. \quad (1.2.19)$$

Constata-se que a estratégia óptima do político é igual, quer se trate de um político incompetente, quer se trate de um político competente.

1.2.4. Equilíbrio

1.2.4.1. Definição

Neste momento estão equacionados os elementos necessários ao apuramento do equilíbrio Bayesiano perfeito deste jogo. Em equilíbrio, os políticos (jogadores) definem (jogam) a sua estratégia óptima, tendo em conta as crenças dos eleitores;⁹ por sua vez, em equilíbrio essas crenças admitem-se correctas face à melhor estratégia do político.

A política implementada pelo político combinada com o tipo desse político pode perspectivar três tipos de equilíbrio. Um equilíbrio separado, caso as políticas implementadas e o tipo de políticos coincidam ou, então, discordem, isto é para combinações, “esforço natural/político competente”, “esforço oportunista/político incompetente” ou, então, para combinações “esforço oportunista/político competente” e “esforço natural/político incompetente”. Um equilíbrio agregado “*pooling*”, caso as políticas seguidas coincidam independentemente do tipo de político, sobrevém para as combinações, “esforço natural/político competente/político incompetente” e “esforço oportunista/político competente/político incompetente”. Um equilíbrio híbrido, caso o político optimize, escolhendo aleatoriamente entre as duas estratégias puras.

Acontece que, para este caso em estudo, apenas podem ocorrer equilíbrios agregados e híbridos.

Proposição 1: O jogo não induz equilíbrios separados.

Prova: Do grupo de estratégias dos políticos, competente e/ou incompetente, constantes das equações (1.2.17) e (1.2.19), decorre que,

⁹ Estas crenças apuram-se utilizando a regra de Bayes, sempre que possível.

* todos os políticos (competentes ou incompetentes) desenvolvem esforço oportunista

e_o , apenas se $E[x_{\lambda t} | h] > E[x_{\lambda t} | l]$;

e

* todos os políticos (competentes ou incompetentes) desenvolvem esforço natural, e_n ,

apenas se $E[x_{\lambda t} | h] < E[x_{\lambda t} | l]$. □

1.2.4.2. Equilíbrio agregado “Pooling”

Corresponde à situação em que todos os políticos seguem a mesma política, o que se traduz em dois equilíbrios, em função do tipo de político.

Proposição 2: Um equilíbrio *pooling* em que todos os políticos se comportam honestamente desenvolvendo um esforço natural, ($\pi_H^{e_o} = \pi_L^{e_o} = 0$) não pode ser solução do jogo.

Prova: Para ($\pi_H^{e_o} = \pi_L^{e_o} = 0$), os valores esperados para o desempenho do político, seriam,

$$\left\{ \begin{array}{l} E[x_{\lambda t} | l] = \frac{q(1-p)(R-c(e_n))}{(1-p)q + p(1-q)} \\ E[x_{\lambda t} | h] = \frac{qp(R-c(e_n))}{qp + (1-q)(1-p)} \end{array} \right. \quad (1.2.20)$$

Entretanto, a partir das estratégias óptimas, constantes das equações (1.2.17) e (1.2.19), para o político incompetente e para o político competente, respectivamente, sabe-se que

o desenvolvimento de um esforço natural pelo político só é a melhor estratégia se $E[x_{\lambda t} | l] > E[x_{\lambda t} | h]$, ou seja se,

$$\Leftrightarrow \frac{q(1-p)(R-c(e_n))}{(1-p)q+p(1-q)} > \frac{qp(R-c(e_n))}{qp+(1-q)(1-p)}$$

$$\Leftrightarrow (1-p)[qp+(1-q)(1-p)] > p[(1-p)q+p(1-q)]$$

$$\Leftrightarrow (1-p)^2 > p^2$$

Ora, como $p \in]0,5;1]$ vem que, esta desigualdade é impossível.

Por outras palavras, mesmo que os eleitores acreditem que todo o esforço desenvolvido pelos políticos é natural, o que acontece é que esses políticos, sejam competentes, sejam incompetentes, são fortemente incentivados a desenvolver um esforço oportunista. O que significa que essas crenças “bondosas” dos votantes não são sustentáveis.

Proposição 3: Um equilíbrio *pooling* em que todos os políticos se comportam demagogicamente desenvolvendo um esforço oportunista, ($\pi_H^e = \pi_L^e = 1$) pode ser uma solução do jogo.

Prova: Para $\pi_H^e = 1$, a probabilidade $\Pr[H | l]$ não está definida. Pode-se, no entanto, ensaiar uma solução, recorrendo à convenção assumida relativamente às crenças dos votantes. Considere-se, pois, que sendo os políticos todos desonestos eles procurarão desenvolver esforços passíveis de emitirem sempre sinais de competência para os votantes, já que um sinal em sentido contrário levaria os votantes a perceber-lo como político incompetente. Formalmente, esta crença, poder-se-ia escrever,

$$\begin{cases} \Pr[H|h] = q \\ \Pr[H|l] = 0 \end{cases} \quad (1.2.21)$$

O desempenho esperado do político condicionado pelo sinal emitido torna-se,

$$E[x_{\lambda t}|l] = (R - c(e_n)) \Pr[H|l] = 0 \quad (1.2.22)$$

e

$$\begin{aligned} E[x_{\lambda t}|h] &= (R - c(e_n)) \Pr[H, e_n|h] + (R - c(e_o)) \Pr[H, e_o|h] \\ &= (R - c(e_n)) \Pr[e_n|H, h] \Pr[H|h] + \\ &\quad + (R - c(e_o)) \Pr[e_o|H, h] \Pr[H|h] \\ &= q \left\{ R - \frac{e_o \Pr[h|e_o, H] \Pr[e_o|H] + e_n \Pr[h|e_n, H] \Pr[e_n|H]}{\Pr[h|e_o, H] \Pr[e_o|H] + \Pr[h|e_n, H] \Pr[e_n|H]} \right\} > 0 \end{aligned} \quad (1.2.23)^{10}$$

Dado que $E[x_{\lambda t}|h] > E[x_{\lambda t}|l]$, qualquer que seja o tipo de político, a estratégia óptima consiste em “jogar” desenvolvendo um esforço oportunista. Resulta que, para estas crenças, o equilíbrio $\pi_H^e = \pi_L^e = 1$ é viável. \square

¹⁰ Dedução constante do Apêndice A.

1.2.4.2. Equilíbrio híbrido

Neste tipo de equilíbrio, alguns políticos, do tipo competente e do tipo incompetente, desenvolvem um esforço oportunista, isto é $\pi_H^{e_o} \in]0,1[$ e $\pi_L^{e_o} \in]0,1[$.

Se ocorresse um equilíbrio desta natureza, os políticos seriam indiferentes à estratégia a seguir. A partir das estratégias óptimas, constantes das equações (1.2.17) e (1.2.19), viria que, $E_P[x_{\lambda t} | L, e_o] = E_P[x_{\lambda t} | L, e_n]$ e $E_P[x_{\lambda t} | H, e_o] = E_P[x_{\lambda t} | H, e_n]$ caso,

$$E[x_{\lambda t} | h] = E[x_{\lambda t} | l] \quad (1.2.24)$$

A equação (1.2.24) é uma condição geral para se obter um equilíbrio estratégico misto, substituindo os respectivos membros pelas equações, (1.2.11) e (1.2.12), pode reescrever-se assim

$$\begin{aligned} & \frac{q \left\{ \left[p + (1-p)\pi_H^{e_o} \right] (R - c(e_n)p(1 - \pi_H^{e_o}) - c(e_o)\pi_H^{e_o}) \right\}}{q \left[p + (1-p)\pi_H^{e_o} \right] + (1-q)[(1-p) + p\pi_L^{e_o}]} = \\ & = \frac{q(1-p)(1 - \pi_H^{e_o})(R - c(e_n))}{q(1-p)(1 - \pi_H^{e_o}) + (1-q)p(1 - \pi_L^{e_o})} \end{aligned} \quad (1.2.25)$$

Vem satisfeita por um qualquer par $(\pi_H^{e_o}, \pi_L^{e_o})$, com $\pi_H^{e_o} \in [0,1]$ e $\pi_L^{e_o} \in [0,1]$. Os valores do equilíbrio $\pi_H^{e_o}$ e $\pi_L^{e_o}$ podem ser interpretados como probabilidades objectivas ou subjectivas, já que em equilíbrio são idênticas.

Proposição 4: Há um *continuum* de pares $(\pi_H^{e_o}, \pi_L^{e_o})$ consistentes com o equilíbrio híbrido.

Prova: Formalmente, a relação entre os valores do equilíbrio pode expressar-se como uma função implícita, $\Phi(\pi_H^{e_o}, \pi_L^{e_o}) = const$, o que permite, após alguns cálculos, escrever $\pi_L^{e_o}$ como função convexa de, $\pi_H^{e_o}$.¹¹ Dado que, essa função é contínua em $[0,1]$ e que, $\pi_H^{e_o} = 1 \Rightarrow \pi_L^{e_o} = 1$ e $\pi_H^{e_o} = 0 \Rightarrow \pi_L^{e_o} \in]0,1[$, há uma infinidade de pares $(\pi_H^{e_o}, \pi_L^{e_o})$, com $\pi_H^{e_o} \in [0,1]$ e $\pi_L^{e_o} \in [0,1]$, que satisfazem a condição de equilíbrio.

Resulta que, o *segundo* equilíbrio *pooling* surge como um limite do equilíbrio híbrido e mesmo que todos os políticos competentes decidissem ser honestos, ainda assim haveria alguns políticos incompetentes a comportarem-se demagogicamente.¹² \square

A condição geral de equilíbrio híbrido, **(1.2.25)**, ainda que venha satisfeita por um número de pares, $(\pi_H^{e_o}, \pi_L^{e_o})$ infinito, não permite determinar o impacto de uma mudança simultânea nos valores do equilíbrio $\pi_H^{e_o}$ e $\pi_L^{e_o}$. Apenas se consegue saber como se altera um dos valores de equilíbrio, quando o outro valor de equilíbrio permanece constante.

Proposição 5: Num equilíbrio híbrido, para um valor constante de $\pi_H^{e_o}$, a probabilidade de um político incompetente agir oportunistamente ($\pi_L^{e_o}$) é crescente com a precisão do esforço e decrescente com a frequência de políticos competentes no universo de políticos.

¹¹ Vindo expressa, após alguns cálculos, por,

$$\pi_L^{e_o} = \frac{[q(1-p)(c(e_o) - c(e_n))](\pi_H^{e_o})^2 - [(c(e_o) - c(e_n))(p + q - 3pq) + (R - c(e_n))(p + q - 1) - (R - c(e_o))qp]\pi_H^{e_o} + (R - c(e_n))(2p - 1)(1 - q)}{p(1 - q)[R - \pi_H^{e_o}c(e_o) - (1 - \pi_H^{e_o})c(e_n)]}$$

¹² O que não surpreende, já que de acordo com os pressupostos, o político incompetente não suporta custos adicionais por comportamentos desonestos.

Prova: A partir da função convexa, compacta, $\pi_L^{e_o} = \pi_L^{e_o}(\pi_H^{e_o}, q, p)$, deduzida a partir da função implícita da relação entre os valores do equilíbrio, vem,

Sinal positivo para a derivada parcial de $\pi_L^{e_o}$ relativamente a p ,

$$\frac{\partial \pi_L^{e_o}}{\partial p} = \frac{(1 - \pi_H^{e_o})[q\pi_H^{e_o}(c(e_o) - c(e_n)) + (1 - q)(R - c(e_n))]}{p^2(1 - q)[R - \pi_H^{e_o}c(e_o) - (1 - \pi_H^{e_o})c(e_n)]} > 0$$

Sinal negativo para a derivada parcial de $\pi_L^{e_o}$ relativamente a q ,

$$\frac{\partial \pi_L^{e_o}}{\partial q} = \frac{\pi_H^{e_o}(1 - \pi_H^{e_o})(1 - p)(c(e_o) - c(e_n))}{p(1 - q)^2[R - \pi_H^{e_o}c(e_o) - (1 - \pi_H^{e_o})c(e_n)]} < 0 \quad \square$$

Proposição 6: Num equilíbrio híbrido, para um valor constante de $\pi_L^{e_o}$, a probabilidade de um político competente agir oportunamente ($\pi_H^{e_o}$) é decrescente com a precisão do esforço e crescente com a frequência de políticos competentes no universo de políticos.

Prova: A diferenciação total da função convexa, compacta, $\pi_L^{e_o} = \pi_L^{e_o}(\pi_H^{e_o}, q, p)$, para, $d\pi_L^{e_o} = 0$, permite ter-se,

$$\frac{\partial \pi_H^{e_o}}{\partial p} = - \frac{\frac{\partial \pi_L^{e_o}}{\partial p}}{\frac{\partial \pi_L^{e_o}}{\partial \pi_H^{e_o}}} \quad \text{e} \quad \frac{\partial \pi_H^{e_o}}{\partial q} = - \frac{\frac{\partial \pi_L^{e_o}}{\partial q}}{\frac{\partial \pi_L^{e_o}}{\partial \pi_H^{e_o}}}$$

No caso em escrutínio, $\frac{\partial \pi_L^{e_o}}{\partial \pi_H^{e_o}} > 0$, então $\frac{\partial \pi_H^{e_o}}{\partial p} < 0$ e $\frac{\partial \pi_H^{e_o}}{\partial q} > 0$. □

Em síntese, este jogo, apresenta equilíbrios múltiplos e em alguns deles os políticos competentes desenvolvem também esforços oportunistas, já que assim evitam o risco de avaliações erradas por parte dos votantes. É evidente que estas conclusões não são independentes dos pressupostos assumidos no jogo, e nesse particular há que recordar que o custo do esforço oportunista, embora superior ao custo do esforço natural, não está limitado superiormente. Apesar das limitações do jogo, mesmo assim é interessante analisá-lo, já que enfatiza, dentro de um trabalho analítico, como a informação imperfeita pode constituir um limite natural ao comportamento honesto do político, servindo de fundamentação teórica à variante de modelo de ciclos político-económico que se vai desenvolver no capítulo 2.

Apêndice A

Dedução da equação (1.2.23)

Relativamente ao sinal de competência que, tanto pode resultar de um esforço natural ou de um esforço oportunista, as deduções conducentes ao valor esperado do desempenho do político, $E[x_{\lambda t} | h]$ são expressas por,

$$E[x_{\lambda t} | h] = (R - c(e_n)) \Pr[H, e_n | h] + (R - c(e_o)) \Pr[H, e_o | h]$$

$$\begin{aligned}
&= (R - c(e_n)) \Pr[e_n | H, h] + (R - c(e_o)) \Pr[e_o | H, h] \\
&= \frac{q \left\{ \left[p + (1-p)\pi_H^{e_o} \right] (R - c(e_n))(1 - \pi_H^{e_o}) - e_d \pi_H^{e_o} \right\}}{q \left[p + (1-p)\pi_H^{e_o} \right] + (1-q)[(1-p) + p\pi_L^{e_o}]}
\end{aligned}$$

dado que,

$$\Pr[e_n | H, h] = 1 - \Pr[e_o | H, h]$$

pode reescrever-se o valor esperado do desempenho do político como,

$$\begin{aligned}
E[x_{\lambda t} | h] &= R \Pr[H | h] - c(e_n) \Pr[e_n | H, h] \Pr[H | h] - c(e_o) \Pr[e_o | H, h] \Pr[H | h] \\
&= \Pr[H | h] \{ R - c(e_n) \Pr[e_n | H, h] - c(e_o) \Pr[e_o | H, h] \}
\end{aligned}$$

Simultaneamente, sabe-se que,

$$\begin{aligned}\Pr[e_o|H, h] &= \frac{\Pr[h|e_o, H]\Pr[e_o, H]}{\Pr[h, H]} = \frac{\Pr[h|e_o, H]\Pr[e_o|H]\Pr[H]}{\Pr[h|H]\Pr[H]} \\ &= \frac{\Pr[h|e_o, H]\Pr[e_o|H]}{\Pr[h|H]}\end{aligned}$$

e que,

$$\Pr[h|H] = \Pr[h|e_o, H]\Pr[e_o, H] + \Pr[h|e_n, H]\Pr[e_n, H]$$

Neste enquadramento, o valor esperado do desempenho, quando se enfrenta um sinal de competência é,

$$\begin{aligned}E[x_{\lambda t}|h] &= \Pr[H|h] \left\{ R - \frac{c(e_n)\Pr[h|e_n, H]\Pr[e_n|H] + c(e_o)\Pr[h|e_o, H]\Pr[e_o|H]}{\Pr[h|e_o, H]\Pr[e_o|H] + \Pr[h|e_n, H]\Pr[e_n|H]} \right\} \\ &= \frac{q \left[p + (1-p)\pi_H^{e_o} \right]}{q \left[p + (1-p)\pi_H^{e_o} \right] + (1-q) \left[(1-p) + p\pi_L^{e_o} \right]} \left\{ R - \frac{c(e_o)\pi_H^{e_o} + c(e_n)(1-\pi_H^{e_o})}{\pi_H^{e_o} + p(1-\pi_H^{e_o})} \right\}\end{aligned}$$

2. Aplicação ao Modelo de Ciclo Político-Económico

Game theory is for providing theorems, not for playing games.

Reinhard Selten, quoted in Goerre e Holt (2001)

Como foi referido o ciclo político-económico pode ser interpretado como parte de um jogo de sinais, num cenário de informação assimétrica relativamente à competência e ao carácter dos políticos. Constitui objectivo desta secção a explicação do ciclo político-económico com base num modelo estrutural, em que o ciclo vem identificado através de instrumentos políticos, mais concretamente orçamentais.

Vai modelar-se a probabilidade de reeleição com base no pressuposto de que os políticos que procuram a reeleição têm um incentivo para exercer esforço de forma a parecerem mais competentes aos olhos dos eleitores e daí afectarem esforço ao desempenho das actividades desenvolvidas.

Esta modelação da probabilidade vai ocorrer num modelo de afectação óptima do esforço político, considerando um ambiente em que os políticos são incentivados a revelarem-se competentes,¹³ respondendo por isso a todos os pressupostos dos modelos de ciclo político-económico padrão – isto é vão escolher a despesa pública para explicar os ciclos, vão definir objectivos em função do interesse próprio,¹⁴ e vão eleger para esse objectivo as variáveis económicas que participam no indicador de avaliação do seu desempenho.

Neste capítulo existem diferenças motivacionais entre políticos que devem ser equacionadas e compete aos votantes o controlo dos comportamentos dos políticos e a selecção dos bons políticos. Assim as preocupações de reeleição podem induzir à não

¹³ Este modelo suporta-se na literatura da teoria de jogos relativa à construção e à preservação da reputação, socorrendo-se do contributo de Kreps e Wilson (1982) que revelaram que a empresa incumbente, no longo prazo, pode construir a sua reputação protegendo-se contra entradas de potenciais concorrentes, o que é consentâneo com posterior descrição de Persson e Tabellini (1990) relativamente à reputação do banqueiro central fundamentada no permanente combate à inflação.

¹⁴ O interesse dos eleitores só é equacionado na função objectivo do político na medida em que lhe permita manter-se no poder, daí que as variáveis económicas presentes na avaliação feita pelos eleitores sejam equacionadas na função objectivo do político.

implementação de políticas socialmente desejáveis, o receio do político em ser substituído pode conduzir a comportamentos diversos dos expectáveis em função da motivação original do político. As conclusões a que chegou Wrasai (2005) estimularam a adaptação do seu modelo ao desenvolvimento deste ensaio.

No sentido de se desenhar a arquitectura de um modelo estrutural para explicar os ciclos político-económicos sobre a despesa pública local, numa variante tipo oportunista, começar-se-á por enunciar os pressupostos que devem estar presentes nesse esboço do modelo.

2.1. Pressupostos

No sentido de se criar um ambiente propício ao desenvolvimento de um modelo teórico, como variante de um jogo de “sinais”, referido a dois momentos do tempo, consideram-se um conjunto de pressupostos relativamente aos intervenientes no jogo, os políticos (*agent*) e os eleitores (*principal*), às interacções desses intervenientes e aos períodos de tempo em que decorrem essas interacções.

Assim,

- Há um *continuum* de votantes, normalizados à quantidade unitária. Encontrando-se os votantes distribuídos por λ jurisdições, com $\lambda \in \{1, \dots, \Lambda\}$.
- A quantidade de votantes em cada jurisdição λ é definida por N_λ , com $\sum N_\lambda = 1$.
- Os votantes na jurisdição λ , num dado momento do tempo, t , preocupam-se com o fornecimento de bens públicos, $g_{\lambda t}$.¹⁵
- As preferências políticas dos votantes em cada jurisdição são semelhantes, mas heterogéneas numa dimensão ideológica.
- A utilidade do votante depende do desempenho do político, $x_{\lambda t}$, com $\lambda \in \{1, \dots, \Lambda\}$ e t o período temporal e caracteriza-se por preferências monótonas (ver apêndice **B**).

¹⁵ As decisões relativas ao fornecimento local de serviços públicos podem ocorrer a níveis mais ou menos centralizados, em função do sistema político vigente.

- A função de utilidade para os votantes na jurisdição λ é expressa por $U_\lambda = \alpha_\lambda x_{\lambda t}$, em que α_λ é uma variável exógena, que por sua vez é função de um vector de outras variáveis exógenas \mathbf{v} , podendo incluir variáveis como N_λ e $N_{-\lambda}$, que cumprem outro papel neste modelo.

- O desempenho do político $x_{\lambda t}$ é observável na quantidade e qualidade do fornecimento de bens/serviços públicos, $g_{\lambda t}$, que por sua vez é função do esforço, $e_{\lambda t}$, exercido pelos políticos na tentativa de aumentar a visibilidade da sua acção e traduzido em medidas para reduzir elementos perturbadores da percepção desses bens/serviços públicos¹⁶ e da competência em os produzir de forma eficiente, θ_λ . Vindo, $g_{\lambda t} = g(e_{\lambda t}, \theta_\lambda)$.

- Dois períodos de tempo, t_1 e t_2 , em que findo o período 1, t_1 , ocorre o acto eleitoral; pelo que t_1 é o período pré-eleitoral e t_2 é o período pós-eleitoral.

- O político no poder, o incumbente, não está consciente da leitura que é feita da sua competência θ_λ , no momento em que exerce esforço, $e_{\lambda t}$, por isso a calendarização dos seus actos é,

t_1 - o incumbente escolhe o nível de esforço, $e_{\lambda 1}$, em cada uma das jurisdições λ , não observável;¹⁷

t_1 - a competência do incumbente é realizada, θ_λ , também não observável, a qual conjuntamente com a decisão de esforço, determina o nível de fornecimento de bem público, $g_{\lambda 1}$, este já observável;

t_1 - os votantes decidem se reelegem ou não o incumbente;

t_2 - o político no poder, no período 2, quer seja o incumbente, quer seja o opositor, escolhe um nível de esforço, $e_{\lambda 2}$, que conjuntamente com a competência, θ_λ ,¹⁸ determina o nível de bem público, $g_{\lambda 2}$;

¹⁶ Evitar que, o ruído na comunicação e a fraude na execução, obscureçam a crença no trabalho realizado, este esforço ocorre sempre, pois negligenciá-lo sugeriria incompetência: “[She] is trapped in supplying the equilibrium level that is expected of him, because, as in a rat race, a lower [level of support] will bias the evaluation process against [her]” [Holmstrom (1982), pp.172].

¹⁷ O esforço aumenta a visibilidade do fornecimento de bens públicos.

¹⁸ A competência é constante nos dois períodos de tempo.

Este pressuposto de que o incumbente conhece tanto a sua competência quanto os votantes no momento em que escolhe o nível de esforço a desenvolver simplifica a análise, já que elimina a possibilidade de equilíbrios múltiplos, permitindo a focalização no nível de esforço a desenvolver pelo incumbente. Efectivamente a questão não está em saber se ele conhece ou não que é competente, mas sim no facto de desconhecer se o ambiente político-económico, no momento em que desenvolve esforço, evoluirá de molde a evidenciar a respectiva competência.

Não há por isso informação assimétrica entre o político no poder, o incumbente, e os votantes no que concerne à competência do político, daí que se possa generalizar a incerteza acerca da competência, *a priori*, do político a todos os envolvidos.

Assume-se que a competência, θ_λ , é distribuída de forma idêntica e independente, i.i.d; com um valor esperado de zero, $E(\theta_\lambda) = 0$ e com as funções de densidade $f(\cdot)$ e de distribuição $F(\cdot)$ definidas num espaço limitado, por $[l, h]$, onde $l < 0 < h$, em que l reflecte uma má competência à esquerda de zero e h traduz uma boa competência à direita de zero.

Caso o político incumbente seja vencido no términos do período 1, o político opositor toma o poder no período 2, com uma competência desconhecida e $E(\theta_\lambda) = 0$.

- O esforço desenvolvido pelo político, numa dada jurisdição, num dado período de tempo, $e_{\lambda t}$, como forma de melhorar o desempenho do político. Esforço que apresenta custos para o político, sendo que o custo do esforço $c(e_{\lambda t})$ é suficientemente convexo para sustentar uma solução interior, com $c(0) = 0$ ¹⁹ e $\lim_{x \rightarrow \infty} c(e_{\lambda t}) = \infty$.

¹⁹ Isto pode ser racionalizado de duas formas. Uma, pressupondo o político adversário do risco, e nesse caso a sua função de utilidade viria côncava, $v(R_{\lambda t}) = v(\bar{R} - e_{\lambda t})$, o que definiria, implicitamente, uma função de custo convexa $c(e_{\lambda t})$. Por exemplo se $v(R_{\lambda t}) = \sqrt{\bar{R} - e_{\lambda t}}$, então pode-se definir $c(e_{\lambda t})$ tal que $v(R_{\lambda t}) = \sqrt{\bar{R}} - c(e_{\lambda t})$, ou seja $c(e_{\lambda t}) = \bar{R} - \sqrt{\bar{R} - e_{\lambda t}}$, que é claramente convexa em $e_{\lambda t}$. Outra, admitindo um custo directo do rendimento retirado, $d(R_{\lambda t})$, (relativo, talvez, ao custo de esconder ganhos

Para se assumir um custo de esforço convexo, o esforço tem que vir definido como um rendimento antecipado $e_{\lambda t} = (\bar{R} - R_{\lambda t}) \in [0, \bar{R}]$. Em que $R_{\lambda t} \in [0, \bar{R}]$ é o nível de rendimento que o político retira a partir do orçamento para bens públicos, na jurisdição λ , no período t ,²⁰ e pelo qual retira uma utilidade $v(R_{\lambda t})$, estritamente côncava. E \bar{R} representa um rendimento exógeno²¹ que, em cada período, o político, no poder, recebe e que pode representar, tanto a prossecução de políticas importantes para si mas negligenciáveis para os votantes, como vantagens decorrentes do exercício do poder.

- O incumbente assume-se como o político que se reporta ao período 1 e por simplicidade notacional não se identificam formalmente os políticos, daí que se deixe cair os subscritos I (incumbente) e O (opositor) passíveis de identificar os políticos.
- Os meios financeiros disponíveis, $K_{\lambda t}$, para afectar ao fornecimento de bens públicos são financiados através de uma taxa de imposto fixa, $\bar{\tau}$, incidindo sobre todos os cidadãos, que por simplicidade se assumem como detentores de um rendimento *per capita*, Y , comum, depois de deduzido o rendimento do político, \bar{R} .

Num sistema descentralizado sem transferências inter-jurisdicionais, viria, $K_{\lambda t} = (N_{\lambda} \bar{\tau} Y - \bar{R})$.

Já, num sistema centralizado, em que a distribuição de receitas entre jurisdições pode depender, quer da afectação ao poder partilhado inter-jurisdicionalmente, quer da estrutura jurisdicional constitucionalmente definida, tudo o que se tem de exigir é que a distribuição de receitas obedeça a uma fórmula estável e publicamente conhecida, para

desonestos), com custos marginais crescentes, por isso $d''(\cdot) > 0$. Neste caso, se $d(R_{\lambda t}) = R_{\lambda t}^2$ e $v(R_{\lambda t}) = R_{\lambda t}$, viria um benefício líquido do rendimento obtido no período t ,

$$R_{\lambda t} - d(R_{\lambda t}) = (\bar{R} - e_{\lambda t}) - (\bar{R} - e_{\lambda t})^2 = \bar{R} - [e_{\lambda t} + (\bar{R} - e_{\lambda t})^2], \text{ que permitiria definir}$$

$$c(e_{\lambda t}) = [e_{\lambda t} + (\bar{R} - e_{\lambda t})^2], \text{ função convexa em } e_{\lambda t}.$$

²⁰ Este rendimento não é observável pelos votantes.

²¹ Limite superior relativamente ao nível de rendimento que pode ser retirado. Sendo que $\bar{R} < NY\bar{\tau}$, ou seja tem que ser inferior ao orçamento das receitas.

que os votantes, em cada jurisdição, possam imputar a competência, θ_λ , do incumbente ao fornecimento de bens públicos para essa jurisdição.

- O fornecimento de bens públicos numa localidade λ , pode definir-se através da expressão, $g_{\lambda t} = K_{\lambda t} + e_{\lambda t} + \theta_\lambda$.

Ao equacionar-se a competência, θ_λ , de forma aditiva, na função de produção do bem público está-se a assumir um modelo em que os políticos sabem que vão ser avaliados (recompensados ou penalizados) por essa competência. Entretanto, como se sabe que os votantes não observam a competência, θ_λ , mas sim o fornecimento de bens/serviços públicos, $g_{\lambda t}$, serão incentivados a exercerem esforço, $e_{\lambda t}$, para parecerem mais competentes. Naturalmente que um dos requisitos para o equilíbrio é que todos os envolvidos tenham crença correctas sobre as acções dos outros, implicando pois, que os votantes não se deixam iludir: conhecendo o esforço, $e_{\lambda t}$, e observando o serviço/bem público fornecido, $g_{\lambda t}$, serão capazes de inferir correctamente a competência, θ_λ . Mesmo assim o incumbente deve exercer o nível de esforço esperado, já que se não o fizer será julgado menos competente do que efectivamente é.

Ao considerar-se a competência como uma parcela da função de produção de bens/serviços públicos, facilita-se a dedução de soluções sempre que se trabalhe em cenários de eleições maioritárias.

- A matriz idiossincrática dos votantes denota, ao invés da uniformidade das preocupações políticas dos votantes, diferenças ideológicas, $\delta_{i\lambda}$, entre os votantes i , na jurisdição λ . Estas diferenças ideológicas respondem, em maior ou menor grau, pela sustentação no poder dos políticos, ou seja pela tolerância. A inclusão das preferências ideológicas permite considerar escolhas de esforço multi-dimensional pelo político incumbente interessado em conseguir a reeleição e responde, também, a uma necessidade técnica de que a probabilidade de reeleição de um político seja contínua no seu esforço.

Uma interpretação comum para as preferências ideológicas, no contexto dos modelos de concorrência eleitoral, consiste em identificá-las com preferências sobre políticas, em que a margem de liberdade de acção dos políticos é ínfima. A ser assim $\delta_{i\lambda}$ capta o grau de tolerância de um votante face a mudanças de orientação política dos políticos, mede o limite aceitável sem redundância em perda de credibilidade desse político.

Assume-se que a distribuição das preferências ideológicas em cada jurisdição é conhecida *ex ante* pelo político no poder, o incumbente, ou seja a incerteza do incumbente não deriva da incerteza na distribuição das preferências ideológicas, mas sim do facto do incumbente desconhecer a sua competência, θ_λ , quando escolhe o nível de esforço no momento 1, $e_{\lambda 1}$.

A distribuição das preferências ideológicas é independente entre jurisdições e é uniforme dentro da mesma jurisdição no intervalo $\left[-\frac{1}{2\gamma_\lambda}, \frac{1}{2\gamma_\lambda}\right]$, com um parâmetro

γ_λ suficientemente grande para assegurar que a percentagem de votos do incumbente, venha, estritamente, entre 0 e 1. Este parâmetro é um indicador da relevância do fornecimento de bens/serviços públicos face a preocupações de natureza ideológica. Assim, um elevado valor de γ_λ revela diferenças ideológicas diminutas entre os votantes, por isso as decisões de votar são tomadas em função da competência percebida do político pela melhoria na eficiência do fornecimento de bens públicos, pouco pesando as considerações ideológicas.

2.2. Equilíbrio

Para se apurar a responsabilidade política num sistema descentralizado, em que o fornecimento de bens/serviços públicos é controlado pelas autoridades locais vai-se encontrar uma solução, tipo “equilíbrio sequencial fraco”, que combina a noção de racionalidade sequencial com o requisito de que as crenças dos jogadores (votantes) no caminho de equilíbrio seguem a regra de Bayes. No entanto, os equilíbrios perfeitos

bayesianos são muitas vezes empregues em jogos de sinalização e não definidos para jogos com acções não observáveis. Daí que se restrinja o estudo a equilíbrios de estratégias puras.

Procura-se pois um equilíbrio perfeito de puras estratégias bayesianas, onde o principal objectivo é o equilíbrio no momento 1. Começa-se por considerar o que ocorre no momento 2; nesse período, o político no poder, quer seja o incumbente, quer seja o opositor do período 1, resolve o seguinte problema,

$$\max_{e_{\lambda 2}} R - c(e_{\lambda 2})$$

Em que $R \equiv v(\bar{R})$, ou seja representa o nível de utilidade que um político local num sistema descentralizado de poder local retira a partir de um valor máximo de rendimento.

A solução deste problema conduz-nos para um $\frac{\partial c}{\partial e_{\lambda 2}} = 0$, o que tendo em conta as características da função de custo de esforço, permite identificar o óptimo do problema com uma ausência de esforço do político no momento 2, $e_{\lambda 2}^* = 0$, para conseguir o máximo de rendimento possível a partir do orçamento para bens/serviços públicos. Daqui decorre que o fornecimento de bens/serviços públicos é inteiramente determinado pela competência do político no poder depois das eleições e pela variável exógena $K_{\lambda 2}$, isto é, $g_{\lambda 2} = K_{\lambda 2} + \theta_{\lambda}$. Consequentemente poder-se-á inferir que os votantes têm todo o interesse em repor no poder os políticos competentes.

Neste cenário, admitindo que a condição, para que um qualquer votante i , numa qualquer jurisdição, λ , reeleja o incumbente é dada por,

$$E(\alpha_\lambda x_{\lambda t} | \text{vence o incumbente}) > \delta_{i\lambda} + E(\alpha_\lambda x_{\lambda t} | \text{vence o opositor})^{22}$$

vem reduzida a,

$$E(\alpha_\lambda (k_{\lambda 2} + \theta_\lambda) | \text{vence o incumbente}) > \delta_{i\lambda} + E(\alpha_\lambda (k_{\lambda 2} + \theta_\lambda) | \text{vence o opositor}) \quad (2.2.1)$$

o que permite,

$$\alpha_\lambda E(\theta_\lambda | \text{vence o incumbente}) > \delta_{i\lambda}, \quad \text{uma vez que}$$

$E(\theta_\lambda) = 0$ para o opositor.

Entretanto, dado que os votantes não observam a competência do incumbente mas sim o desempenho, $x_{\lambda 1} \equiv (\theta_\lambda + e_{\lambda 1} + K_{\lambda 1})$, há um incentivo a que os incumbentes exerçam esforço numa tentativa de parecerem mais competentes e aumentarem a probabilidade de reeleição. Formalmente, poder-se-á dizer que o incumbente, no período 1, enfrenta o seguinte problema,

$$\max_{e_{\lambda 1} e_{\lambda 2}} R - c(e_{\lambda 1}) - c(e_{\lambda 2}) + \Pr(\text{incumbente vencer} | e_{\lambda 1}, e_{\lambda 2}) \cdot R \quad (2.2.2)$$

A chave para este problema está em encontrar uma expressão para $\Pr(\text{incumbente vencer} | e_{\lambda 1}, e_{\lambda 2})$. Efectivamente passado é passado, daí que os votantes

²² A preferência ideológica, $\delta_{i\lambda}$, de um votante i numa jurisdição λ , assume-se como referida ao político opositor.

estejam disponíveis para reeleger o incumbente apenas se esperarem, dele, acções futuras positivas, ainda que o desempenho passado constitua um bom indicador do desempenho futuro.

A ser assim torna-se necessário introduzir variáveis que traduzam a convicção dos votantes no que concerne à competência e ao esforço dos incumbentes. Para o efeito passar-se-á a identificar a variável $\tilde{\theta}_\lambda$, como a competência, θ_λ , atribuída ao incumbente pelos votantes, fundamentada na observação dos bens/serviços públicos fornecidos, $g_{\lambda 1}$, e na convicção das acções levadas a cabo para o efeito, $\tilde{e}_{\lambda 1}$, o que se traduz em, $\tilde{\theta}_\lambda = \tilde{\theta}_\lambda(g_{\lambda 1}, \tilde{e}_{\lambda 1})$. A competência do incumbente é, pois, uma variável aleatória, dependente da realização de uma outra variável aleatória. Em equilíbrio, qualquer desempenho $x_{\lambda 1}$, observável por via dos bens/serviços públicos fornecidos, $g_{\lambda 1}$, que esteja dentro do intervalo $[\tilde{e}_{\lambda 1} + K_{\lambda 1} + l, \tilde{e}_{\lambda 1} + K_{\lambda 1} + h]$ pode ser observado com probabilidade positiva, desde que a competência, θ_λ , venha dentro do intervalo $[l, h]$.

Para se determinar a $\Pr(\text{incumbente vencer} | e_{\lambda 1}, e_{\lambda 2})$ há que,

(1) deduzir a regra de votação dos votantes em cada jurisdição.

(2) Observar, que, no período 2, o desempenho esperado pelos votantes, é traduzível por,

(i) $(\tilde{\theta}_\lambda + e_{\lambda 2}^* + K_{\lambda 2}) = (\tilde{\theta}_\lambda + K_{\lambda 2})^{23}$, caso se reeleja o incumbente, e

(ii) $K_{\lambda 2}$, caso seja o opositor o eleito, já que não tendo estado a ser avaliado

apresenta um $E(\tilde{\theta}) = 0$.

²³ Recorde-se que nenhum esforço será despendido depois da eleição.

(3) Observar a condição para que um votante i , numa determinada jurisdição λ , reeleja o incumbente, considerando a preferência ideológica, $\delta_{i\lambda}$,

$$\alpha_\lambda(\tilde{\theta}_\lambda + K_{\lambda 2}) > \delta_{i\lambda}\alpha_\lambda K_{\lambda 2}, \quad (2.2.3)$$

significando que o desempenho esperado do incumbente terá que ser suficientemente grande para compensar qualquer enviesamento ideológico a favor do opositor.

Esta condição pode, ainda, apresentar-se como,

$$\delta_{i\lambda} < \alpha_\lambda \tilde{\theta}_\lambda^{24} \quad (2.2.4)$$

Dado que $\delta_{i\lambda}$ se distribui uniformemente em $\left[-\frac{1}{2\gamma_\lambda}, \frac{1}{2\gamma_\lambda}\right]$, vem que a percentagem de

votantes que sustentam o incumbente na jurisdição λ é,

$$\frac{1}{2} + \gamma_\lambda \alpha_\lambda \tilde{\theta}_\lambda^{25} \quad (2.2.5)$$

Para que o incumbente consiga a reeleição, tem que assegurar pelo menos a maioria dos votos, o que só acontece quando a probabilidade da percentagem de votos recebida é superior a $\frac{1}{2}$, ou seja,

²⁴ Permite inferir que, quanto maior a competência esperada do incumbente, maior o enviesamento ideológico necessário para que o votante vote contra o incumbente. Mas a uma menor utilidade dos votantes, numa jurisdição λ , a partir de um dado desempenho de governo, menor importância da competência esperada relativamente à consideração ideológica.

²⁵ γ_λ é a densidade de $\delta_{i\lambda}$ como já foi visto.

$$\Pr\left(\frac{1}{2} + \gamma_\lambda \alpha_\lambda \tilde{\theta}_\lambda > \frac{1}{2}\right) = \Pr(\tilde{\theta}_\lambda > 0) \quad (2.2.6)$$

Em suma, a resolução do problema de escolha do incumbente reside na dedução de uma expressão para, $\Pr(\tilde{\theta}_\lambda > 0)$. Na essência, o objectivo do incumbente passa por aumentar a probabilidade de vencer as eleições exercendo esforço no sentido de que as estimativas dos votantes sobre a sua competência aumentem – ele vai pretender convencer os votantes que é mais competente do que, efectivamente, é, desenvolvendo mais esforço do que aquele que os votantes acreditam ter sido desenvolvido. E, ainda que, em equilíbrio as crenças dos votantes serão correctas, significando que o incumbente nada pode fazer para induzir os votantes a apreciarem-no como mais competente. Mesmo assim ele pretende exercer esforço já que deixar de o fazer poderia sugerir incompetência.

Primeiro que tudo há que observar que a crença dos votantes relativamente à competência se constrói a partir dos serviços públicos fornecidos, $g_{\lambda 1}$, na convicção relativa ao esforço gasto pelo incumbente, $\tilde{e}_{\lambda 1}$, e no conhecimento do parâmetro exógeno $K_{\lambda 1}$,

$$\begin{aligned} g_{\lambda 1} &= K_{\lambda 1} + \tilde{e}_{\lambda 1} + \tilde{\theta}_\lambda \\ \Rightarrow \tilde{\theta}_\lambda &= g_{\lambda 1} - K_{\lambda 1} - \tilde{e}_{\lambda 1} \end{aligned} \quad (2.2.7)$$

Quer isto dizer que, para todo $g_{\lambda 1}$ que, com probabilidade positiva pode ser observado, dadas as convicções dos votantes sobre a acção do incumbente, $\tilde{e}_{\lambda 1}$, virá uma competência estimada para o incumbente crescente com o seu esforço:

$$\begin{aligned}\tilde{\theta}_\lambda &= (K_{\lambda 1} + e_{\lambda 1} + \theta_\lambda) - K_{\lambda 1} - \tilde{e}_{\lambda 1} \\ \Rightarrow \tilde{\theta}_\lambda &= \theta_\lambda + e_{\lambda 1} - \tilde{e}_{\lambda 1}\end{aligned}\tag{2.2.8}$$

Qualquer realização de $g_{\lambda 1}$ dentro do intervalo $[K_{\lambda 1} + e_{\lambda 1} + l, K_{\lambda 1} + e_{\lambda 1} + h]$ pode ser observado com probabilidade positiva, na medida em que a competência, θ_λ , tem um suporte em, $[l, h]$. Desde que as convicções dos votantes sobre o nível de esforço dos incumbentes esteja correcta (como qualquer equilíbrio de Nash impõe), isso implica que para $\forall g_{\lambda 1} \in [K_{\lambda 1} + e_{\lambda 1} + l, K_{\lambda 1} + e_{\lambda 1} + h]$, a convicção dos votantes sobre a competência do incumbente é expressa por, $\tilde{\theta}_\lambda = \theta_\lambda + e_{\lambda 1} - \tilde{e}_{\lambda 1}$.

Para obter o equilíbrio, entretanto, há que introduzir alguns pressupostos relativos às convicções dos votantes fora do caminho do equilíbrio. Uma conjectura realística tem a ver com a situação em que os votantes observam desempenhos *suficientemente elevados* para serem credíveis no quadro da estratégia de equilíbrio do incumbente e por isso assumem que a competência atingiu o valor mais elevado; similarmente se o desempenho é “*impossivelmente baixo*” assumem que o incumbente é do tipo “o menos competente”. Pode-se formalizar esta realidade no seguinte pressuposto:

Pressuposto 1. Para observações fora caminho de equilíbrio, isto é para $g_{\lambda 1} \notin [K_{\lambda 1} + \tilde{e}_{\lambda 1} + l, K_{\lambda 1} + \tilde{e}_{\lambda 1} + h]$, os votantes têm as seguintes convicções acerca do tipo de competência que enfrentam,

$$\begin{aligned}\tilde{\theta}_\lambda &= h \text{ se } g_{\lambda 1} > K_{\lambda 1} + \tilde{e}_{\lambda 1} + h \\ \tilde{\theta}_\lambda &= l \text{ se } g_{\lambda 1} < K_{\lambda 1} + \tilde{e}_{\lambda 1} + l\end{aligned}\tag{2.2.9}$$

O pressuposto 1 é, então, uma condição de monotonicidade, na medida em que implica que os votantes nunca assumiriam com competência h um desempenho de $g_{\lambda 1} = K_{\lambda 1} + \tilde{e}_{\lambda 1} + h$, enquanto um desempenho de $g_{\lambda 1} > K_{\lambda 1} + \tilde{e}_{\lambda 1} + h$ implicar competência inferior a h . Sem este pressuposto, equilíbrios implausíveis poderiam ser sustentados, tais como os votantes atribuírem baixas competências a elevados desempenhos, desincentivando os incumbentes a esforçarem-se mais do que o esperado.

Conjugando o pressuposto 1 com a equação (2.2.8) poder-se-á afirmar que os votantes têm, em relação ao tipo de competência que enfrentam, as seguintes crenças,

$$\begin{aligned}\tilde{\theta}_{\lambda} &= l \text{ se } g_{\lambda 1} < K_{\lambda 1} + e_{\lambda 1} + l \\ &= \theta_{\lambda} + e_{\lambda 1} - \tilde{e}_{\lambda 1} \text{ se } g_{\lambda 1} \in [K_{\lambda 1} + e_{\lambda 1} + l, K_{\lambda 1} + e_{\lambda 1} + h]\end{aligned}\quad (2.2.10)$$

$$\tilde{\theta}_{\lambda} = h \text{ se } g_{\lambda 1} > K_{\lambda 1} + e_{\lambda 1} + h$$

Uma vez que $l < 0 < h$, pode-se escrever $\Pr(\tilde{\theta}_{\lambda} > 0)$ como,

$$\Pr(\tilde{\theta}_{\lambda} > 0) = \Pr(\theta_{\lambda} + e_{\lambda 1} - \tilde{e}_{\lambda 1} > 0) = \Pr(\theta_{\lambda} > \tilde{e}_{\lambda 1} - e_{\lambda 1}) \quad (2.2.11)$$

Antes de resolver o nível de equilíbrio do esforço, numa jurisdição λ , no período 1, há que ter presente que sendo $\tilde{\theta}_{\lambda} = \theta_{\lambda} + e_{\lambda 1} - \tilde{e}_{\lambda 1}$, e dado que competência, θ_{λ} , é distribuída de forma idêntica e independente, ou seja uma distribuição i.i.d; com um valor esperado de zero, $E(\theta_{\lambda}) = 0$ e com as funções de densidade $f(\cdot)$ e de distribuição

$F(\cdot)$ definidas num espaço limitado, por $[l, h]$, tem-se para o valor esperado da competência percebida, $\tilde{\theta}$, a seguinte notação,

$$E(\tilde{\theta}_\lambda) = \bar{f}(\tilde{\theta}_\lambda) \equiv \int_{-\infty}^{+\infty} f(\tilde{\theta}_\lambda) f(\theta_\lambda) d\theta_\lambda$$

em que $\bar{f}(\tilde{\theta})$ representa apenas um tipo de “densidade média” da variável aleatória e por isso é sempre maior que zero.

O que por sua vez permite reescrever, $\Pr(\tilde{\theta}_\lambda > 0) = \Pr(\theta_\lambda + e_{\lambda 1} - \tilde{e}_{\lambda 1} > 0)$, em termos da função de distribuição cumulativa da variável aleatória θ_λ , $F(\tilde{e}_{\lambda 1} - e_{\lambda 1})$, vindo,

$$\Pr(\tilde{\theta}_\lambda > 0) = \Pr(\theta_\lambda + e_{\lambda 1} - \tilde{e}_{\lambda 1} > 0) = 1 - F(\tilde{e}_{\lambda 1} - e_{\lambda 1}) \quad (2.2.12)$$

Ora, introduzindo a equação (2.2.11) no problema de maximização, constante da equação (2.2.2), no período 1; diferenciando em relação a $e_{\lambda 1}$ e impondo a condição de equilíbrio de que as convicções dos votantes sobre o esforço dos incumbentes estão correctas, i.e. que $\tilde{e}_{\lambda 1} = e_{\lambda 1}$, obtêm-se um equilíbrio para o incumbente no período pré-eleitoral.

Proposição 1. Num sistema descentralizado, a única estratégia de puro equilíbrio Bayesiano, consistente com o pressuposto 1, é a que tem implícito um esforço no período 1 definido por,

$$c'(e_{\lambda 1}^*) = f(0).R, \quad (2.2.13)$$

onde $R = v(\bar{R})$, a utilidade do incumbente decorrente do máximo de rendimento que consegue obter num sistema descentralizado.

Prova: (em Apêndice C)

Esta solução permite os seguintes comentários, *primeiro*, o nível de responsabilidade no período 1 é crescente com a capacidade do incumbente para obter rendimento no período 2, $R = v(\bar{R})$, e com o seu grau de avaliação desse rendimento; *segundo*, o nível de responsabilidade é independente do valor, α_λ , que os votantes acrescentam à utilidade que tiram do fornecimento do bem público e da relevância das suas preferências políticas, γ . O mesmo é dizer que os políticos desenvolvem maior esforço nos períodos pré-eleitorais do que fora deles. Decorre pois, que a solução para o sistema descentralizado é idêntica à de Persson e Tabellini (2000, cap 9), ainda que se tenha equacionado a ideologia.

Relativamente à solução encontrada que aponta para uma responsabilização crescente do político, no poder, no período pré-eleitoral, em função do rendimento a obter no período 2, valerá a pena enquadrá-la numa perspectiva de eleições repetidas, em que o incumbente, quer seja competente, quer seja incompetente, tem consciência de que o seu desempenho está a ser avaliado pelos votantes, embora os votantes,²⁶ no curto prazo, observem as decisões políticas mas não os resultados políticos. Há informação assimétrica acerca da competência percebida do político e acerca do efeito das políticas executadas, está-se num ambiente perfeito para soluções arbitrárias. Conhece-se a política (p.e. identifica-se a construção de estradas e infra-estruturas e observa-se a recuperação do parque escolar) mas não se consegue avaliar o impacto sobre a carga fiscal do cidadão nem a eficiência dessas despesas e a votação ocorre no desconhecimento destes parâmetros.

Um político competente vai procurar que a sua função objectivo,

²⁶ Os votantes não estão informados acerca do tipo de políticos que ocupam o poder, quer no que concerne às motivações que o impeliram ao lugar, quer no efeito das políticas implementadas por eles.

$$\max_{e_{\lambda 2}} R - c(e_{\lambda 2})$$

seja alcançada ao nível mais elevado de esforço, para agir de acordo com o interesse público, enquanto o político incompetente se vai comportar de forma diametralmente oposta. Eleições sequenciais irão mostrar que estes comportamentos podem ser disciplinados, aumentando por isso a probabilidade de se conseguir eleger um político competente no período 2.

Recorrendo a um jogo simples a dois períodos, em que os intervenientes activos são o político no poder e o votante representativo, tem-se, para o período 1 e para o período 2,

(i) A natureza, a decidir se,

* o político no poder é competente ou incompetente, ou seja a determinar se $\theta_{\lambda t} \in \{l, h\}$

e

* a conjuntura política económica vai ser favorável a otimizar os resultados das políticas desenvolvidas ou não, isto é a determinar se o desempenho, $x_{\lambda t}$, vai ser percebido como bom, b , ou mau, $-b$, isto é, $x_{\lambda t} \in \{-b, b\}$.

(ii) O incumbente a observar a conjuntura política onde se desenvolverá o seu esforço, e a decidir a sua política,

$$g_{\lambda 1} \in \{g_{\lambda 1} = 0, g_{\lambda 1} = 1\},$$

em que $g_{\lambda 1} = 1$ significa que vai implementar a sua política, enquanto $g_{\lambda 1} = 0$ significa manter o *status quo*.²⁷ O político tem incertezas acerca dos resultados da sua política,

²⁷ Manter o *status quo* significa seguir uma linha “politicamente correcta”, tomando decisões na linha das tradicionalmente seguidas pelo poder local, ainda que possam ser social e economicamente ineficientes.

tanto pode produzir bons resultados, b , como maus resultados, $-b$, com igual probabilidade.

(iii) O votante a observar a decisão relativamente a $g_{\lambda t}$, e a escolher se votam ou não no incumbente, ou seja $v \in \{v = 0, v = 1\}$, ainda que não observe os resultados ($x_{\lambda t}$ e R).

Assume-se que,

(A1) uma política, se implementada por um político competente induz um resultado (*payoff*) para o votante representativo coincidente com o desempenho, $x_{\lambda t}$, desse político;²⁸ no entanto se a mesma política for implementada por um mau político, o resultado (*payoff*) que daí advém para o votante é do tipo $x_{\lambda t} - \alpha R$,²⁹ – agora a satisfação do votante não se fica só pelo nível de desempenho $x_{\lambda t}$, mas vem reduzida na medida da parcela de rendimento apropriada pelo mau político, R , ainda que afectada pelo parâmetro α , que reflecte a intensidade com que essa apropriação é percebida por ele. Caso os políticos decidam manter o *status quo*, $g_{\lambda t} = 0$, então há uniformidade nos resultados (*payoffs*), para o votante representativo e para o político incumbente, que se assumem normalizados a zero, independentemente de o político ser competente e/ou incompetente;

e, ainda, que,

(A2) um resultado (*payoff*), para a implementação de uma política, por um mau político, em conjuntura desfavorável à respectiva percepção pelo votante, sempre positivo, ou seja, $-b + R > 0$, o que implica um estímulo a que o político incompetente implemente, no período 2, sempre uma política,³⁰ ainda que socialmente indesejável.

²⁸ Há coincidência de resultados (*payoffs*), quer o político competente, quer o votante representativo, obtêm a mesma consequência, ou seja o resultado (*payoff*) de uma política que se preocupa com o interesse público, que será percebida como boa, b , se a conjuntura for favorável à percepção do resultado dessa política.

²⁹ Neste cenário, não há coincidência de resultados (*payoffs*) já que o mau político ao apropriar-se do rendimento obtém um resultado, $x_{\lambda t} + R$.

³⁰ Só assim consegue extrair um rendimento.

Mais, neste cenário, o político incompetente decidirá sempre agir, $g_{\lambda t} = 1$, no período 2, independentemente do estado da arte, contrariamente ao procedimento do político competente que só agirá nesse período se e só se a conjuntura determinar uma observância de, $x_{\lambda 2} = b$.

Estes pressupostos enviesam o comportamento dos políticos em função da respectiva competência, já que o mau político relativamente ao bom político é levado a agir, daí haver necessidade de introduzir, ainda, o pressuposto de que os votantes substituirão o político no poder, $v = 1$, se esse político no poder cumprir o seu programa, independentemente do ciclo eleitoral, ou seja se, $g_{\lambda t} = 1$. Decorre, deste último pressuposto, uma estratégia de acção para os votantes suportada na observação de $g_{\lambda t}$, a qual lhes permitirá decidir se votam ou não no incumbente, ou seja, $v \in \{v = 0, v = 1\}$, depois de observarem $g_{\lambda t}$.

No fim deste primeiro período, os resultados, para estes intervenientes, podem-se sintetizar, em termos de utilidade, assim,

$$U_h(g_{\lambda t} = 1) = b$$

$$U_l(g_{\lambda t} = 1) = b(-b) + R$$

$$U_h(g_{\lambda t} = 0) = U_l(g_{\lambda t} = 0) = 0$$

$$U_v(g_{\lambda t} = 1) = b(-b) - \alpha R$$

$$U_v(g_{\lambda t} = 0) = 0$$

representando, U_h , U_l e U_v , as utilidades respectivamente do político competente, do político incompetente e do votante.

No período 2, acontece a repetição das ocorrências do período 1.

A sequência destes acontecimentos é passível de ser ilustrada por,

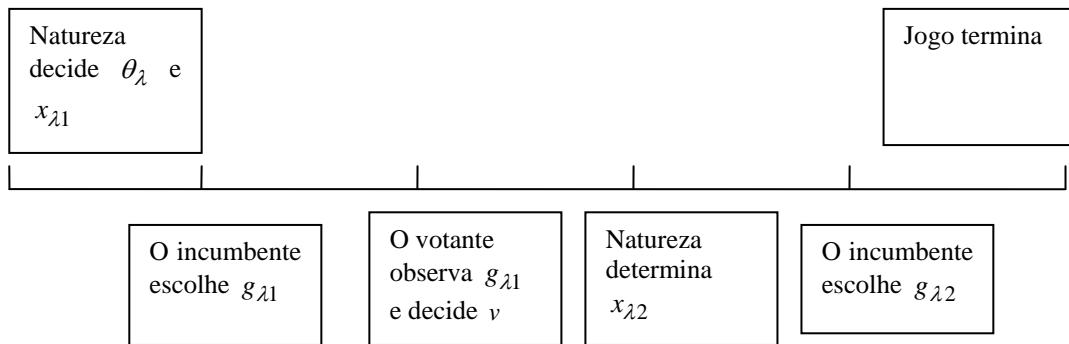


Figura 2-1: Sequência de Acontecimentos

As estratégias de acção para qualquer um dos tipos de político, no segundo período, decorrem directamente dos pressupostos acima apresentados.

Admitindo um político competente no período 1 e aceitando, ainda, que a conjuntura conduzirá a uma má percepção do seu desempenho, ilustrada por, $x_{\lambda 1} = -b$, então, seguramente, que o político não será incentivado a prosseguir com as suas linhas programáticas, já que serão percebidas como uma má política, e, certamente que, no período 2, virá substituído, com uma probabilidade, $1 - \omega$, por um político incompetente. Assim a resposta óptima deste político competente é decidir $g_{\lambda 1} = 0$. Ao invés, se admitirmos que a conjuntura é favorável à percepção do desempenho do político competente, vindo uma apreciação de um bom desempenho, expressa por, $x_{\lambda 1} = b$, então ele enfrentará um dilema: (1) desenvolve o seu programa, isto é, decide $g_{\lambda 1} = 1$, e arrisca que um político incompetente o substitua no período 2; ou (2) mantém o *status quo*, isto é, decide $g_{\lambda 1} = 0$, e evita que o político incompetente o venha a substituir no período 2.

Antecipando o comportamento no período 2, poder-se-á encontrar o resultado esperado do político competente nos dois cenários, assim,

$$U_h(g_{\lambda 1} = 0 \mid x_{\lambda 1} = b) = \frac{1}{2}b, \quad \text{caso decida}$$

manter o *status quo*;

$$U_h(g_{\lambda 1} = 1 \mid x_{\lambda 1} = b) = \frac{1}{2}b + \omega \frac{1}{2}b - (1 - \omega)\alpha R, \quad \text{caso decida}$$

avançar com a sua política.

Naturalmente se observa que uma atitude idealista do político, expressa por,

$$U_h(g_{\lambda 1} = 1 \mid x_{\lambda 1} = b) > U_h(g_{\lambda 1} = 0 \mid x_{\lambda 1} = b)$$

só conduzirá a um melhor resultado se,

$$\alpha R < \frac{1 + \omega}{2(1 - \omega)}b.$$

O mesmo é dizer que, uma elevada percepção, por parte do votante, do rendimento do político, uma baixa probabilidade de reeleição, ω , e um baixa perspectiva de conjuntura

favorável levam a que o político competente decida manter o *status quo*,³¹ de forma a prevenir que um político incompetente venha a ser eleito.

Considerando, agora, que o político no poder é incompetente, e admitindo que a conjuntura é desfavorável, tem-se um desempenho, $x_{\lambda 1} = -b$. Neste caso, antecipando o comportamento no período 2, os resultados vêm,

$$U_I(g_{\lambda 1} = 1 | x_{\lambda 1} = -b) = -b + R + \omega \frac{1}{2} b - (1 - \omega) \alpha R$$

e

$$U_I(g_{\lambda 1} = 0 | x_{\lambda 1} = -b) = R.$$

Consequentemente, agir em conformidade com o seu programa político, só produz melhores resultados, se,

$$\alpha R < \frac{\frac{1}{2} \omega - 1}{1 - \omega} b.$$

Facilmente se verifica que esta condição é permanentemente violada. O que quer dizer que, quando a conjuntura é desfavorável à produção de bons resultados políticos (isto é

³¹ Para que um político competente implementasse o seu programa, isto é $g_{\lambda 1} = 1$, ter-se-ia que verificar uma baixa visibilidade da apropriação do rendimento pelo político por parte dos votantes, uma elevada probabilidade de reeleição do bom político e uma elevada perspectiva de conjuntura favorável à percepção da política como *boa* política.

susceptível de gerar a percepção de um mau desempenho, $x_{\lambda 1} = -b$), mesmo para o político incompetente a resposta óptima é escolher $g_{\lambda 1} = 0$. Este resultado ilustra o potencial papel disciplinador das eleições. Neste caso, o votante sai também ganhador, não só não vem implementada a má política como também não há apropriação de rendimento, no período 1,³² por parte do político. Há um equilíbrio socialmente eficiente.

Ainda no cenário do parágrafo anterior, em que o político incumbente é incompetente, há que considerar a situação em que a conjuntura é favorável à percepção da política a implementar, ou seja vai determinar consequências socialmente desejáveis, o mesmo que pressupor, $x_{\lambda 1} = b$. Aplicando a mesma metodologia, implícita à produção dos resultados anteriores, para o caso da decisão “agir”, isto é, para $g_{\lambda 1} = 1$, tem-se,

$$U_I(g_{\lambda 1} = 1 | x_{\lambda 1} = b) = b + R + \omega \frac{1}{2}b - (1 - \omega)\alpha R$$

e, para a decisão “não agir”, $g_{\lambda 1} = 0$, tem-se,

$$U_I(g_{\lambda 1} = 0 | x_{\lambda 1} = b) = R$$

Consequentemente a escolha, $g_{\lambda 1} = 1$, por um político incompetente acontece se e só se,

³² A visibilidade do rendimento R transforma este parâmetro num factor inibidor da promoção de uma política socialmente indesejável pelo político incompetente, na medida em que o político não pretende ficar prisioneiro da fama de mau político. As eleições acabam por incentivar ao diferimento da apropriação do rendimento por parte do político incompetente.

$$\alpha R < \frac{2 + \omega}{2(1 - \omega)} b.$$

Esta condição revela, mais uma vez, que o político no poder, neste caso o político incompetente, mesmo com uma conjuntura favorável à produção de consequências benéficas, pode abster-se de agir no período 1 e de se apropriar da renda nesse período, para o efeito basta que estejam reunidas as circunstâncias de: (i) uma elevada percepção do rendimento do político; (ii) uma baixa probabilidade de eleição de um político competente e (iii) um reduzido valor para o desempenho percebido.

Ao esgotarem-se os cenários determinados pela natureza, (1) tipo de político no poder, competente ou incompetente e (2) conjuntura política económica, favorável ou desfavorável, à percepção das consequências das políticas implementadas, constata-se circunstâncias favoráveis a uma estratégia de *não acção* por parte do incumbente: o político competente resolve não correr riscos de ser substituído e por isso decide não implementar uma política socialmente desejável; e o político incompetente escolhe agir apenas se a conjuntura é favorável a uma leitura positiva das suas decisões políticas. Esta solução ocorre no quadro da assumpção que a manutenção do político acontece se e só se o político não agir³³ e no pressuposto que a condição $\alpha R < \frac{1 + \omega}{2(1 - \omega)} b$ é violada

e a condição $\alpha R < \frac{2 + \omega}{2(1 - \omega)} b$ não.

Este é o quadro de um equilíbrio “politicamente correcto”, que nos permite concluir que, por via das eleições, as actuações dos políticos são “alisadas” tendencialmente no sentido de comportamentos no quadro do institucional e do tradicionalmente concretizado e, sobretudo visualizadas pelos eleitores como a “obra” do político. Nesta perspectiva as eleições acabam por desempenhar um papel disciplinador para os maus políticos, na medida em que as políticas socialmente indesejáveis não são

³³ Registe-se que esta regra de votação assumida, face às estratégias escolhidas pelos políticos, revela-se uma resposta óptima.

implementadas no período 1; e um papel estabilizador, ao melhorarem a probabilidade de se ter um político competente no período 2 (a probabilidade de um bom político no poder, ω , com eleições viria $\omega + \frac{1}{2}(1 - \omega)$).

Havendo espaço para explorar um cenário em que as condições, $\alpha R < \frac{1 + \omega}{2(1 - \omega)}b$ e

$\alpha R < \frac{2 + \omega}{2(1 - \omega)}b$, são ambas violadas, importa conhecer as estratégias dos políticos

nesse contexto, conjugadas com a regra de votação “o incumbente é reeleito se e só se não agir”. De novo se cai no quadro de um equilíbrio em que os políticos permanecem no domínio do “*status quo*”, no período 1, mas agora esta posição defensiva persiste sempre, independentemente do tipo de político e do tipo de conjuntura ditado pela natureza. Políticos competentes e incompetentes agem da mesma forma no período 1, isto é, “*não agem*”, o que faz com que as eleições potenciem a probabilidade de maus políticos se manterem no poder (a probabilidade de um mau político se manter no poder, no período 2, com eleições, é $1 - \omega + \frac{1}{2}\omega_{\text{superior}} > \omega + \frac{1}{2}(1 - \omega)$, a probabilidade de um bom político se manter, no período 2, com eleições, desde que, $\omega < 0,5$).

Caso se pretenda analisar se a estratégia de sustentação do *status quo* continua a prevalecer como a melhor resposta para um político se manter no poder, independentemente do tipo e da conjuntura, em contexto de satisfação de ambas as

condições, $\alpha R < \frac{1 + \omega}{2(1 - \omega)}b$ e $\alpha R < \frac{2 + \omega}{2(1 - \omega)}b$, o resultado alcançado é afirmativo. Este

equilíbrio caracteriza-se, novamente, por disciplinar os maus políticos.

A ser assim pode-se sumariar que a probabilidade do político se manter, por via das eleições, é determinante nas decisões políticas a tomar no curto prazo, seja ele competente ou incompetente. A escolha óptima do esforço, pode conduzir o político a

não agir, uma vez que a reeleição pode estar melhor garantida por esta via. De certa forma este resultado vai de encontro ao comportamento oportunista do incumbente, caracterizado por o esforço ser desenvolvido para influenciar a *reputação futura*, e é convergente com o apurado na sub-secção 1.2.4, onde se concluiu que a informação imperfeita pode instituir limites a um comportamento sincero dos políticos, impelindo acções manipuladoras para que os votantes infiram competência desses actos.

As soluções reportadas anteriormente são passíveis de virem esquematizadas assim,

$$\begin{array}{c}
 \frac{(1+\omega)b}{2(1-\omega)} \qquad \qquad \qquad \frac{(2+\omega)b}{2(1-\omega)} \qquad \qquad \qquad \gamma R \\
 \hline
 H : (g_{\lambda 1}^* | b, g_{\lambda 1}^* | -b) = (1,0) \quad H : (g_{\lambda 1}^* | b, g_{\lambda 1}^* | -b) = (0,0) \quad H : (g_{\lambda 1}^* | b, g_{\lambda 1}^* | -b) = (0,0) \\
 L : (g_{\lambda 1}^* | b, g_{\lambda 1}^* | -b) = (1,0) \quad L : (g_{\lambda 1}^* | b, g_{\lambda 1}^* | -b) = (1,0) \quad L : (g_{\lambda 1}^* | b, g_{\lambda 1}^* | -b) = (0,0)
 \end{array}$$

Figura 2-2: Soluções de Equilíbrio

Em síntese, o pressuposto de substituição, por via das eleições, do político no poder caso decida agir, revela o poder disciplinador das eleições, já que em qualquer das soluções as políticas susceptíveis de serem percebidas como políticas socialmente indesejáveis não são implementadas, quer pelo político competente, quer pelo político incompetente, de forma a melhorar a probabilidade de se manterem no poder no período temporal subsequente.

Para se perceber esta atitude do político é necessário perceber que é o votante que prefere uma política menos interventora, para o efeito há que enquadrar o espaço de desempenho dos políticos no domínio das preferências dos votantes num cenário de teoria de jogos. Propósito a remeter para o conteúdo do Apêndice **B**.

Apêndice B

Preferências Monótonas dos Votantes

Os votantes,³⁴ caracterizam-se por preferências monótonas relativamente a um espaço de desempenho do político, $X = \mathbb{R}_+^2$, coincidente com o quadrante positivo do espaço bidimensional euclídeo. Esse desempenho é apenas observável, por via dos serviços públicos³⁵ fornecidos pelos políticos no poder, no período t , podendo definir-se através de coordenadas, vindo, $\mathbf{x} = (x_{g_t}, x_{-g_t})$ ³⁶, em que x_{g_t} identifica uma política pública mais interventora ($g_{\lambda t} = 1$) e x_{-g_t} identifica uma política pública menos interventora ($g_{\lambda t} = 0$).

Num modelo com dois políticos e um votante representativo, tem-se então, dois políticos $P = \{O, I\}$, em que I é o político no poder, o incumbente, e em que O é o político opositor. Ambos se apresentam a escrutínio a um lugar de poder, em cada período eleitoral, e há um votante v ³⁷ a seleccionar um, dos dois políticos. Assume-se que estes três jogadores se caracterizam, globalmente, por terem preferências não saciadas.

Assume-se que,

(i) cada político, no poder, preocupa-se com o desempenho percebido e com os valores defendidos. Se o desempenho, \mathbf{x}_t , é escolhido no período t , pelo tipo político pertencente ao universo P , no poder nesse período, então os resultados do político são,

³⁴ Os principais no modelo de eleições agente/principal.

³⁵ Em quantidade e qualidade.

³⁶ Optou-se por negligenciar a notação λ , que reflecte a diversidade jurisdicional, uma vez que o aporte de simplicidade à apresentação supera a quebra de preciosismo da apresentação.

³⁷ Considera-se somente um votante singular para enfatizar as dificuldades de controlar os políticos e, também, pelo facto de um acréscimo de votantes/principais introduzir apenas complexidade.

$$u_I(\mathbf{x}_t) + R$$

O termo específico $R \geq 0$ mede as rendas não políticas associadas à permanência no poder. Assume-se que a função de utilidade específica $u_I(\mathbf{x}_t) : X \rightarrow \mathbb{R}^1$ é diferenciável duas vezes;

(ii) os votantes preocupam-se apenas com o desempenho do político e têm uma função de utilidade duplamente diferenciável, $u_v(\mathbf{x})$.

Estamos interessados no caso em que o político no poder, denominado de incumbente, com a notação I , é um grande consumidor de políticas de dimensão g , políticas públicas interventoras, enquanto o político na oposição, O , é um grande consumidor de dimensão $-g$, e o votante, v , pretende o máximo de cada dimensão. Formalmente, pode-se definir a taxa marginal de substituição entre uma política pública mais interventora e uma política pública menos interventora, no espaço \mathbf{x} , para um qualquer jogador i , como,

$$TMS_i(\mathbf{x}) = \frac{\frac{\partial u_i(x)}{\partial x_g}}{\frac{\partial u_i(x)}{\partial x_{-g}}}$$

E assume-se que, para qualquer desempenho, $\mathbf{x} \in \mathbf{X}$, $TMS_O(\mathbf{x}) < TMS_v(\mathbf{x}) < TMS_I(\mathbf{x})$.

Para simplificar, assumem-se os casos extremos, os quais vêm ilustrados pelos valores adoptados pelas taxas marginais de substituição, $TMS_O(\mathbf{x}) = 0$ e $TMS_I(\mathbf{x}) = \infty$, para

$\forall \mathbf{x} \in \mathbf{X}$. Isto verifica-se quando $u_O(\mathbf{x}) = u_O(x_{-g})$ e $u_I(\mathbf{x}) = u_I(x_g)$, sendo as funções

$u_O(\cdot)$ e $u_I(\cdot)$ estritamente crescentes. Assume-se que o votante, v , tem preferências estritamente convexas.

Neste enquadramento teórico, as curvas de indiferença para os três jogadores considerados, que se assume serem o político incumbente, o político opositor e o votante, relativamente a um desempenho específico, podem ser ilustradas por,

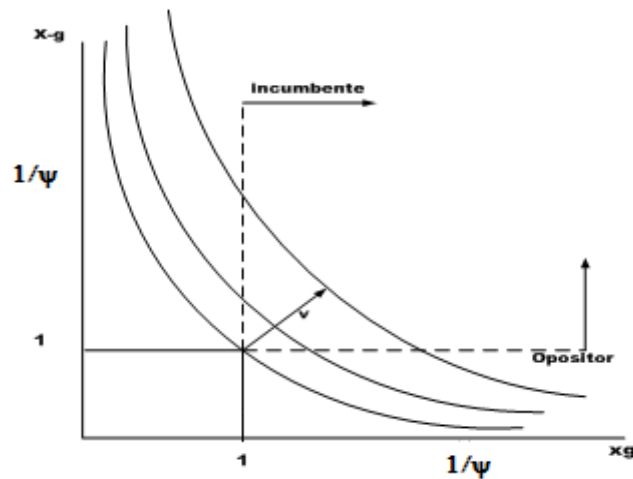


Gráfico 1: Curvas de Indiferença de três jogadores

Se a estas considerações teóricas acrescermos a assumpção de uma sequência infinita de períodos eleitorais em que se verifica ausência de limitação de mandatos, então para um dado período, t , tem-se,

* a natureza, um jogador não estratégico, a seleccionar aleatoriamente a restrição orçamental, $B_t \subset X$, com a qual, o político no poder se defrontará;

e

* o governo, $p_t \in P$, depois de observar a restrição orçamental B_t , a seleccionar um ponto de desempenho político $x_t \in B_t$.

O votante conhece x_t , mas não B_t , neste contexto apresenta uma função votação $v_t \in \{0,1\}$, em que um voto notado com o valor de 0 é um voto para manter o incumbente e um voto notado com o valor de 1 é um voto para substituir o incumbente, isto é para fazer variar o político no poder, $P \setminus p_t$.

Sequencialmente, no período, $t + 1$, enfrenta-se,

* uma nova restrição orçamental B_{t+1} , e

* um novo desempenho político, $\mathbf{x}_{t+1} \in B_{t+1}$, a ser seleccionado pelo governo, p_{t+1} e ocorrem novas eleições.

Sem perda de generalidade assume-se que, no período $t = 1$, decorre a selecção de um desempenho político $\mathbf{x}_1 \in B_1$, pelo partido no governo, $p_1 = I$.

Este tipo de jogo necessita de ser expandido às funções de utilidade $u_i(\mathbf{x})$ para um horizonte infinito. De acordo com uma sequência $\{\mathbf{x}_t, p_t\} = \{(\mathbf{x}_1, p_1), \dots, (\mathbf{x}_t, p_t)\}$ a utilidade de um qualquer político p é,

$$U_p(\{\mathbf{x}_t, p_t\}) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} [u_p(\mathbf{x}_t) + \eta_p 1_p(p_t)]$$

Onde $\delta \in (0,1)$ é uma taxa de desconto e $1_p(p^t)$ é um indicador que toma o valor de 1 se $p^t = p$ e o valor de 0 nas outras situações. Similarmente a utilidade do votante sobre esta sequência é,

$$U_v(\{\mathbf{x}_t, p_t\}) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_v(\mathbf{x}_t)$$

A competência do incumbente, observada na quantidade e na qualidade do serviço público fornecido, sustenta-se em políticas públicas mais interventoras, e implica enfrentar uma restrição orçamental exequível. O político no poder conhece essa restrição, mas os eleitores e o opositor não a conhecem. Assume-se que, qualquer restrição orçamental fiável, é do tipo,

$$B^t = \{x \in X : \beta_t x_g + (1 - \beta_t) x_{-g} \leq 1\}$$

em que β_t é um preço relativo, onde $\beta_t \in [\psi, 1 - \psi]$, sendo ψ uma medida no intervalo, $0 < \psi < \frac{1}{2}$, para sustentar a aleatoriedade do preço relativo. O conjunto possível de restrições é isomórfico para $B = [\psi, 1 - \psi]$. O conjunto de políticas que são exequíveis no quadro dessas restrições traduz-se por, $\mathbf{B} = \cup_{\beta \in [\psi, 1 - \psi]} B(\beta)$, e é passível de se ilustrar graficamente por,

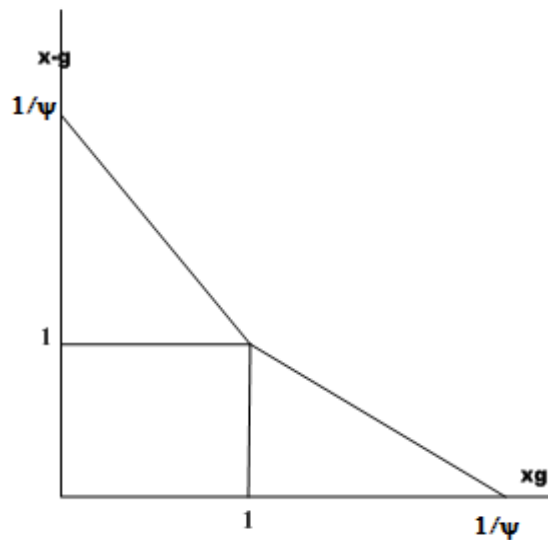


Gráfico 2: Conjunto de políticas exequíveis

O limite eficiente desse conjunto de políticas, \mathbf{B} , denomina-se, $\hat{\mathbf{B}}$, e consiste no conjunto de pontos que satisfazem $B(\beta)$, com igualdade para alguns $\beta \in [\psi, 1 - \psi]$.

Assume-se como crença comum que, para qualquer período de tempo t , o parâmetro, β_t , é obtido através de uma arbitrio aleatório a partir de uma função de distribuição contínua e estritamente crescente, $F_\beta(\cdot)$, no domínio, $[\psi, 1 - \psi]$. O valor do parâmetro,

β_t , é apenas observado pelo político no poder no momento t , quer se trate do político incumbente, I , quer se trate do político na oposição, O , ou seja que tem que estar no governo, isto é em p_t .

Uma ordenação parcial sobre X , permite interrelacionar espacialmente políticas diferentes, por exemplo, para uma política \mathbf{x} e para uma política \mathbf{y} , pode-se assumir que se tem, $\mathbf{x} \succ \mathbf{y}$, caso $x_{-g} > y_{-g}$ e $x_g < y_g$, intuitivamente $\mathbf{x} \succ \mathbf{y}$ significa que a política \mathbf{x} está localizada à esquerda e acima da política \mathbf{y} . Para uma qualquer restrição, $B(\beta)$, do conjunto fechado e convexo de restrições, a continuidade e convexidade estrita de preferências assegura que o conjunto de políticas induzidas para o agente i ,

$$\mathbf{x}_i^*(\beta) = \arg \max_{\mathbf{x} \in B(\beta)} u_i(\mathbf{x})$$

existe e é singular. Para além disso pelo teorema do máximo [Berge (1963)] esta função é contínua.

Impõem-se dois pressupostos adicionais relativamente às preferências do votante, v .

Pressuposto 1: Para um qualquer $\beta^* \in [\psi, 1 - \psi]$,

$$\mathbf{x}_v^*(\beta^*) = (1, 1).$$

Esta condição estabelece que, para alguns valores de β^* plausíveis, a política ótima para o votante v , sujeita ao conjunto restrição, $B(\beta^*)$, corresponde ao ponto $(1, 1)$, o

único ponto que se situa sobre o limite de qualquer restrição aceitável. Este pressuposto é satisfeito se,

$$\frac{\psi}{1-\psi} < TMS_v(1,1) < \frac{1-\psi}{\psi}.$$

Pressuposto 2: Se $\beta < \beta'$, então $x_v^*(\beta') \searrow x_v^*(\beta)$.

Este pressuposto requer que v responde a aumentos no preço relativo de uma política mais interventora, procurando menos dessa política e mais da política não interventora.

Estes dois pressupostos, asseguram, conjuntamente, que as preferências do votante v sejam bem comportadas. Graficamente este comportamento é passível de vir ilustrado assim,

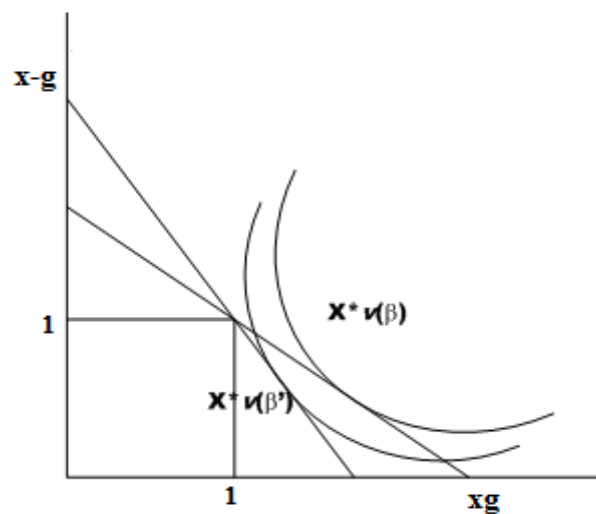


Gráfico 3: Preferências do votante

Apêndice C

Prova da Proposição 1

O pressuposto 1 formaliza as convicções que os votantes terão sobre o tipo de competência do político que os governa, dado o desempenho observado $x_{\lambda 1}$, e as convicções formadas relativamente às acções do político $\tilde{e}_{\lambda 1}$,

$$\begin{aligned}\tilde{\theta}_{\lambda} &= l, \text{ se } g_{\lambda 1} < k_{\lambda 1} + \tilde{e}_{\lambda 1} + l \\ &= g_{\lambda 1} - k_{\lambda 1} - \tilde{e}_{\lambda 1} = \theta_{\lambda} + e_{\lambda 1} - \tilde{e}_{\lambda 1}, \text{ se } g_{\lambda 1} \in [k_{\lambda 1} + \tilde{e}_{\lambda 1} + l, k_{\lambda 1} + \tilde{e}_{\lambda 1} + h] \\ &= h, \text{ se } g_{\lambda 1} > k_{\lambda 1} + \tilde{e}_{\lambda 1} + h\end{aligned}$$

Antecipando isto, o político pode formalizar a sua probabilidade de vencer, $\Pr(\tilde{\theta}_{\lambda} > 0)$.

Ora, tendo em conta as convicções dos votantes sobre o esforço do político, $\tilde{e}_{\lambda 1}$, e respectiva competência $\tilde{\theta}_{\lambda}$, pode expressá-la, para $e_{\lambda 1} \leq \tilde{e}_{\lambda 1}$, como,

$$\int_l^{l+\tilde{e}_{\lambda 1}-e_{\lambda 1}} \mathbf{I}(l \geq 0) f(\theta_{\lambda}) d\theta_{\lambda} + \int_{l+\tilde{e}_{\lambda 1}-e_{\lambda 1}}^h \mathbf{I}(\theta_{\lambda} + \tilde{e}_{\lambda 1} - e_{\lambda 1} \geq 0) f(\theta_{\lambda}) d\theta_{\lambda}$$

em que $\mathbf{I}(\cdot)$ indica uma função característica, que toma o valor de 1 se o pressuposto é verdadeiro e o valor de zero caso contrário. O primeiro termo desta expressão representa observações do desempenho do político, $x_{\lambda 1} (\equiv g_{\lambda 1})$ fora do caminho de equilíbrio, i.e. para $g_{\lambda 1} < \tilde{e}_{\lambda 1} + k_{\lambda 1} + l$, já o segundo termo representa as observações sobre o caminho de equilíbrio.

Uma vez que $\tilde{\theta}_\lambda = l$, para $g_{\lambda 1} < \tilde{e}_{\lambda 1} + k_{\lambda 1} + l$, virá que $\tilde{\theta}_\lambda > 0$, para todas as observações fora do caminho de equilíbrio, sse $l \geq 0$. Já para as observações sobre o caminho de equilíbrio, $\tilde{\theta}_\lambda = \theta_\lambda + e_{\lambda 1} - \tilde{e}_{\lambda 1}$, a vitória do incumbente verificar-se-á sse $\theta_\lambda + e_{\lambda 1} - \tilde{e}_{\lambda 1} \geq 0$.

Se $\tilde{\theta}_\lambda \geq l$ ³⁸ ter-se-ia a função característica a tomar o valor de 1 para todas as realizações de θ_λ , quer no 1º termo, quer no 2º. Implicando que, neste cenário, os políticos vencem com probabilidade igual a um.

Por contraste se $\tilde{\theta}_\lambda < l$, viria um valor de zero para a função indicador constante do 1º termo e, para o 2º termo só se conseguiria o valor de um sse $\theta_\lambda \geq \tilde{e}_{\lambda 1} - e_{\lambda 1} - l$.

Reescrevendo a expressão relativa à probabilidade de vencer, em termos da função de distribuição de θ_λ , vem,

$$\int_{\theta_\lambda < l} (1 - F_{\theta_\lambda}(\tilde{e}_{\lambda 1} - e_{\lambda 1} - \theta_\lambda)) f(\theta_\lambda) d\theta_\lambda + \int_{\theta_\lambda \geq l} 1 \cdot f(\theta_\lambda) d\theta_\lambda$$

em que F_{θ_λ} é a função de distribuição conjunta de θ_λ .

Para realizações de θ_λ suficientemente baixas, a probabilidade de vencer é estritamente inferior a um. Enquanto, realizações superiores de θ_λ levarão à vitória do político.

Similarmente pode-se derivar a $\Pr[\tilde{\theta}_\lambda > 0]$, para $e_{\lambda 1} \geq \tilde{e}_{\lambda 1}$, como,

³⁸ Há que reter que a competência esperada, $\tilde{\theta}_\lambda$, é uma função da competência real θ_λ , que é uma variável aleatória, distribuída de forma independente.

$$\int_{\theta_\lambda < h} 0 \cdot f(\theta_\lambda) d\theta_\lambda + \int_{\theta_\lambda \geq h} (1 - F_{\theta_\lambda}(\tilde{e}_{\lambda 1} - e_{\lambda 1} - \theta_\lambda)) f(\theta_\lambda) d\theta_\lambda$$

Conjuntamente estas duas equações definem uma função contínua e diferenciável em $e_{\lambda 1}$. Derivando esta função e aplicando a condição de equilíbrio $e_{\lambda 1} = \tilde{e}_{\lambda 1}$, bem como a de $\tilde{\theta}_\lambda = \theta_\lambda$, tem-se,

$$\begin{aligned} \frac{\partial \Pr(\theta_\lambda)}{\partial e_{\lambda 1}} &= \int_{\theta_\lambda \in [-h, l]} f(0, \theta_\lambda) f(\theta_\lambda) d\theta_\lambda \\ &= \int f(0) f(\theta_\lambda) d\theta_\lambda \\ &= \bar{f}(0) \end{aligned}$$

A segunda igualdade resulta do facto de θ_λ estar definido sobre o intervalo $[l, h]$, daí que a integração, sobre $[-h, l]$, não limite qualquer realização de θ_λ , para a qual $f(-\theta_\lambda) > 0$ e a terceira igualdade aplica a definição de \bar{f} . Multiplicando por R_d e fixando a igualdade com a derivada da função custo obtém-se o nível óptimo de esforço.

3. Modelo Empírico e Fontes de Dados

An economic theorist is one who says he has proved the existence of a black cat in a dark room (in which there is no cat). An econometrician is one who looks for a black cat in a dark room (in which there is no cat) and declares that he has found one. A Bayesian econometrician is one who looks for a black cat in a dark room (in which there is no cat) and declares he has found a cat and that the cat is white.

Adrian Pagan, cited in Maddala (1998)

3.1. Modelo Empírico

É nosso objectivo procurar evidência empírica para a eventual capacidade do político incumbente gerar ciclos político-económicos em função da respectiva probabilidade de reeleição. Em princípio os incentivos para gerar ciclos político-económicos serão sempre decrescentes com a probabilidade de reeleição, já que sinalizar competência num cenário de elevada probabilidade de ser reeleito induz mais perda de reputação esperada do que benefícios.

A probabilidade do político, no poder, vencer as próximas eleições não é observável e por isso torna-se necessário estimá-la. Existem estudos realizados, em que a probabilidade aparece estimada em função da popularidade e de estudos de opinião,³⁹ passíveis de serem enquadrados em duas perspectivas. Uma, seguida por Chappell e Keech (1988), que constrói a probabilidade de reeleição a partir dos valores estimados para a percentagem de votos obtida pelo incumbente e dos erros da previsão estimada. Outra, seguida por Cohen (1993), em que a probabilidade de reeleição é construída a partir da média e da variância estimada do caminho aleatório, que admite, seguido por uma série de sondagens.

³⁹ Para uma revisão sobre funções voto e popularidade consultar Paldam (2004) e Nannestad e Paldam (1994).

Neste caso específico, não se possui informação, em termos locais, relativa a estudos de opinião e/ou de popularidade dos políticos no poder, mas sabendo, a partir do capítulo 2, secção 2.2., que,

“o político pode formalizar a sua probabilidade de vencer, $\Pr(\tilde{\theta}_\lambda > 0)$, em termos de esforço desenvolvido para conseguir a reeleição, tendo em conta as convicções dos votantes sobre o esforço do político, $\tilde{e}_{\lambda 1}$, e respectiva competência $\tilde{\theta}_\lambda$ ”,

pode-se ensaiar uma adaptação à realidade concreta a partir das perspectivas referidas no parágrafo anterior. Para o efeito, fundamentamo-nos na **proposição 1 (2.2.13)**, apresentada nesse capítulo e que, sumariamente, refere que os políticos desenvolvem maior esforço nos períodos pré-eleitorais do que fora deles e, também, no facto das sondagens estarem relacionadas com o desempenho dos políticos no poder.⁴⁰

A ser assim vamos procurar estimar uma relação entre os resultados das eleições autárquicas,⁴¹ (potencial) indicador da aprovação eleitoral e diversas variáveis que procuram captar o desempenho do político. Para essa concretização, há que reter as variáveis capazes de conferir significado à relação que se pretende ver estimada, as quais podem vir identificadas como:

1. *variável dependente*: os resultados eleitorais alcançados pela força política em exercício de funções na data do acto eleitoral, nos sucessivos momentos eleitorais autárquicos, conforme dados oficiais do Secretariado Técnico dos Assuntos para o Processo Eleitoral (STAPE). Na especificação do modelo essa variável dependente aparecerá identificada com a designação $VV1_{it}$ e reportará a percentagem de votos obtida pelo partido que está no poder no ano t da eleição autárquica, no município i .

⁴⁰ Esta perspectiva permite ensaiar a possibilidade de estabelecer uma relação entre o desempenho do político, observado no respectivo desempenho orçamental, e a aprovação dos eleitores.

⁴¹ Já que o presidente do executivo autárquico coincide com o 1º membro da lista proposta pelo partido mais votado.

2. *variáveis independentes, por sua vez sistematizadas em três grupos:*

i) as variáveis *proxy* da convicção dos votantes acerca da competência do político e do esforço do político,

(1) O número de anos sucessivos no exercício do cargo de presidente da câmara no município i à data de eleição t , pode revestir as características de uma *proxy*, que, eventualmente, ilustra a confiança (ou não) que os votantes depositam no político, a qual de certa forma traduz uma crença positiva (negativa) na competência e no esforço desse político, nesta perspectiva, os dados relativos a esta observação, serão equacionados e identificados pela variável, $APRES_{it}$. Em suma, a escolha desta variável pretende captar a convicção que os votantes possam ter no desempenho do incumbente (pois que tornar-se incumbente já significa vencer os candidatos opositores, ser *seleccionado*). Mais, com esta variável, ao assumir-se que ela tem implícita as características individuais do incumbente e o fenómeno do enviesamento da incumbência, pode-se aspirar a observar uma *proxy* do comportamento dos políticos, que, em equilíbrio “politicamente correcto”, tenderão a prosseguir comportamentos que sejam observados pelos votantes como a “obra” do político.

(2) O número de anos sucessivos que o partido vencedor esteve (ou não) à frente dos destinos do município i à data de eleição t , variável designada por $APART_{it}$, na especificação do modelo introduz a interacção (ou não) com o partido a que pertence o presidente do executivo camarário e por essa via permite corrigir a avaliação do desempenho que é atribuída à variável $APRES_{it}$. Mais esta variável, ainda que correndo o risco de se apresentar excessivamente correlacionada com $APRES_{it}$, pode corrigir situações em que o candidato é reeleito com o apoio de outro partido, ocorrências há em que o autarca passa a ser apoiado por outro partido e consegue ser reeleito;

ii) as variáveis identificáveis com indicadores de gestão,

(1) As despesas de investimento, em valor real, feitas no município i no ano de eleição t , por habitante, identificadas no modelo por $DIREAL_{it}$,⁴² servem para explorar a ideia de uma eventual relação de causalidade entre investimentos e votos. Numa primeira fase, avançar-se-á com o valor agregado da despesa de investimento. A significância estatística do respectivo coeficiente estimado, pela aplicação do modelo de estimação, determinará se, numa segunda fase, a especificação do modelo deverá (ou não) evoluir para a desagregação das despesas de investimento;

(2) A afectação das despesas correntes à rubrica de despesas com pessoal, $DESPESSOAL_{it} / DESPCORRENTE_{it}$, podia, também, assumir-se como uma variável capaz de ilustrar actos de gestão a avaliar pelos votantes, no entanto a prevalência de uma pesada afectação das despesas correntes a despesas com pessoal, extensível de forma quase uniforme a todos os municípios portugueses, decorrente da observação das respectivas contas de gerência, sugere a opção de não inclusão dessa variável, satisfazendo o requisito de parcimónia que deve estar subjacente à especificação dos modelos;

iii) as variáveis macroeconómicas,

(1) Optou-se por considerar a taxa de desemprego nacional ($PDESEMP1_{it}$) e a taxa de inflação nacional ($PINFLAC1_{it}$), quando o partido incumbente coincide, no município e no governo central; caso contrário, assumem o valor zero. Estas variáveis, também podem captar alguma dose de ideologia, na medida em que permitem expurgar a avaliação realizada pelos votantes ao desempenho do partido que elegeu o incumbente local.

iv) finalmente, numa análise dinâmica, importaria fazer da variável dependente $VV1_{it}$, uma variável exógena, introduzindo-a no modelo como o valor relativo ao momento

⁴² A despesa de investimento considerada reporta-se ao ano em que ocorre o acto eleitoral; a uniformidade do momento dessa ocorrência (final do ano civil) permite aceitar o valor orçamentado para o efeito como valor executado.

eleitoral, imediatamente anterior ao observado, isto é ter-se-ia $VV1_{it-1}$, ou seja a percentagem de votos recolhida pelo partido em funções executivas no acto eleitoral imediatamente anterior, $t - 1$, no município i . A característica preponderante nos dados da amostra aponta para uma coincidência relevante entre essa informação e a percentagem de votos do partido ganhador da eleição anterior (afinal o resultado que faz esse partido estar em exercício de funções, na data do acto eleitoral t , no município i). Em conformidade, com a evidência dos dados observados a variável escolhida para a especificação do modelo foi esta última, vindo identificada pela notação, $VOTOSANT_{it}$.⁴³

A regressão ajustada a estas variáveis poderia vir especificada assim,

$$VV1_{it} = \alpha_{it} + \beta_1 * VOTOSANT_{it} + \beta_2 * APART_{it} + \beta_3 * APRES_{it} + \\ + \beta_4 * DIREAL_{it} + \beta_5 * PDESEMP1_{it} + \beta_6 * PINFLAC1_{it} + \varepsilon_{it}$$

em que i representa o município (1 a 275), t o ano eleitoral e ε_{it} o termo de erro.

Para estimar a probabilidade de reeleição, de acordo com a metodologia seguida em Chappell e Keech (1988), começa-se por,

- (1) estimar os valores esperados para a variável dependente, resultados eleitorais, denominada, $EVV1_{it}$, ajustados em função do valor estimado para a regressão, ou seja estimam-se os resultados eleitorais com base nos coeficientes da regressão estimada;
- (2) estimar os desvios padrão dos erros de previsão, $\hat{\sigma}_f$, a partir dos erros padrão da regressão e da matriz de covariância dos parâmetros estimados;

⁴³ Veiga. e Veiga (2007 a e 2007 b), encontraram uma correlação aproximadamente de 75% entre os votos obtidos pelo incumbente na eleição imediatamente anterior e os votos desfasados de um período.

para se obter,
$$\text{Pr} = F\left[(EVV1_{it} - 0,43822) / \hat{\sigma}_f\right],^{44}$$

em que F é uma função de distribuição cumulativa *Normal*.

Concretizada a etapa da estimação descrita no parágrafo anterior, uma vez no domínio dos valores para a probabilidade estimada, para o município i , no momento eleitoral t , Pr_{it} , é oportuno assumi-la com potencial para ser equacionada como variável independente, num modelo tipo auto-regressivo. Modelo esse, que virá gizado sobre uma variável dependente, pensada de forma a poder (ou não) revelar ciclos político-económicos, e nesse âmbito a escolha recai sobre uma variável susceptível de expressar o nível de endividamento do político no poder. Para identificar essa variável opta-se pela leitura do valor da rubrica *encargos financeiros*, constante das contas de gerência das autarquias. A esta rubrica obedecem todos os juros devidos por endividamento a terceiros. De forma a equacionarem-se as diferenças inter-municipais a variável é considerada em termos *per capita* e vem traduzida pela notação $CFcapita$.

Eleita que está a variável, passível de poder traduzir potenciais ciclos político-económicos, há que ajustá-la sobre os valores prévios dela própria, de forma a concretizar-se um modelo auto-regressivo tal como desenvolvido nalguns trabalhos empíricos.⁴⁵ Optou-se por avançar para a seguinte forma do modelo auto-regressivo:

⁴⁴ Escolheu-se o valor limite 0,43822 em vez do valor limite 0,5, por se considerar que o primeiro retrata melhor o panorama político-partidário em Portugal. Observa-se que, (i) as forças partidárias não se restringem a dois partidos políticos e (ii) as vitórias eleitorais acontecem logo que uma das forças partidárias obtém mais um voto do que a segunda força mais votada. Realidade que, ponderada por, outras duas realidades, *uma* que se prende com o facto de as duas forças políticas mais votadas recolherem, por norma, cerca de 65% a 75% dos votos e *outra* que tem a ver com uma vitória eleitoral autárquica, com a maioria ou não de mandatos, sugeriu a construção de um índice, $I = 51\% / \%Perc_mand$. O objectivo deste índice é conseguir que os resultados eleitorais efectivamente obtidos, ilustrem melhor as circunstâncias relativas ao sistema eleitoral português, conforme já referido. Concretização alcançada através da ponderação desses resultados eleitorais pelo índice em apreço. Decorrendo daí, que, o valor 0,43822, não é mais do que um valor médio desses resultados eleitorais ponderados pelo índice, que se presumiu susceptível de ajustar melhor a(s) realidade(s) reportada(s).

⁴⁵ Carlsen, F. (1997); Alesina, Cohen e Roubini (1992, 1993); Grier (1987, 1989) e Grier e Neiman (1987).

$$CFcapita_{it} = \gamma_0 + \gamma_1 Pr_{it} + \gamma_2 CFcapita_{it-\ell} + \varepsilon_{it}$$

Em que o termo $CFcapita_{it-\ell}$ traduz o nível de *encargos financeiros* no ℓ – ano anterior à data de eleição.

Optou-se por transformar moderadamente esta modelação,⁴⁶ de forma a contemporizar a dimensão temporal da base de dados consubstanciada na juventude da democracia portuguesa. Historicamente referenciada no marco do 25 de Abril de 1974, só em 1976 proporcionou as primeiras eleições municipais livres, e numa primeira fase, até 1985, com ciclos eleitorais de três anos, factos que restringiam a dimensão do ℓ – ano anterior à data de eleição para o termo $CFcapita_{it-\ell}$.

Em teoria, se esta especificação do modelo empírico reproduzisse a solução de equilíbrio denominada de “politicamente correcta”, deduzida no contexto dos equilíbrios desenvolvidos na secção 2.2., seria expectável um sinal negativo (–) para o coeficiente estimado associado à variável Pr_{it} , uma vez que essas soluções de equilíbrio permitem inferir que a probabilidade do político se manter no poder, por via das eleições, é determinante das decisões políticas tomadas no curto prazo. Em suma, o valor da probabilidade de reeleição do político no poder, independentemente da sua competência, constitui (ou não) estímulo para gerar ciclos político-económicos no intuito de garantir a respectiva reeleição

Além disso, uma observação unilateral dos pressupostos subjacentes ao modelo teórico, *a competência do político no poder é assumida como parcela da função de produção de bens/serviços públicos e a respectiva decisão de escolha do nível de esforço ocorre num momento em que o político desconhece a aceção dos eleitores relativamente a essa*

⁴⁶ Assim sobre o modelo auto-regressivo proposto, avançou-se com a possibilidade das regressões virem ajustadas sobre variáveis transformadas (a partir de $CFcapita_{it-\ell}$), tais como, taxas de variação do $CFcapita$ entre os sucessivos dois períodos consecutivos dentro de cada ciclo eleitoral, $\Delta Peso_{it-\ell}$ e taxas de variação do $CFcapita$ relativo aos diversos períodos antecedentes ao acto eleitoral face à média de cada um dos ciclos eleitorais, $\Delta Média_{it-\ell}$.

competência, apresentados na secção 2.1., conferem plausibilidade ao incentivo para o incumbente exercer mais esforço para sinalizar competência em cenário de baixos valores de probabilidade de reeleição, o que acaba por se traduzir na implementação de políticas tendencialmente manipuladoras da inferência de competência e, eventualmente, na contenção de políticas socialmente desejáveis.

3.2. Fontes dos Dados Estatísticos

A base de dados relativa aos resultados das eleições autárquicas, aos indicadores de Finanças Locais e, ainda, alguns indicadores socio-económicos e demográficos,⁴⁷ foi tratada de molde a poder testar-se, quer a possibilidade de se trabalharem os dados observados de forma agregada, como uma *pool data*, quer a possibilidade de se trabalharem com uma estrutura painel de dados.⁴⁸ As variáveis observadas referem-se aos 275 municípios de Portugal continental,⁴⁹ durante um período temporal iniciado em 1979 e finalizado em 2001, configurando por isso características espaciais e temporais, que oferecem um potencial de informação que importa não se negligenciar. Efectivamente, trabalhar com dados que equacionam as dimensões espaço/tempo permite mais informação do que a que decorreria se, apenas, se recorresse a dados seccionais, os quais não permitiriam avaliar as variações nas observações que acontecem pelo efeito tempo.

⁴⁷ A Base de Dados, suporte do trabalho empírico deste ensaio, foi, gentilmente, disponibilizada pela Professora Doutora Linda Veiga, e decorre dos trabalhos de investigação no âmbito dos projectos *Interacções entre a Economia e a Política em Portugal e Eleitores, governantes e a economia*, ambos financiados pela Fundação para a Ciência e a Tecnologia (POCTI/ECO/2001/37457 e POCI/EGE/58641/2004, respectivamente) conduzidos pelo Núcleo de Investigação de Política Económica, NIPE da EEG.

⁴⁸ Relativamente à estrutura painel dos dados importa introduzir a questão relativa à distinção entre dados em painel e dados em séries cronológicas de unidades seccionais “time series and cross section”, *TSCS*. Freedman e Peters (1984) registaram que os dados observados ao longo do tempo em diferentes unidades seccionais divergem relativamente à dimensão espacial, têm-se unidades seccionais fixas (não dependentes da amostra recolhida) e têm-se unidades seccionais variáveis (em função da recolha de dados obtida), sendo que as últimas, tradicionalmente, se observam para períodos mais curtos e assim os distinguiram como dados *TCSC* e dados em painel, respectivamente. Teoricamente, as diferenças entre esses tipos de dados, ilustram-se por um apelo assintótico em T para os dados *TCSC* e por assintóticas em N para os dados em painel. Finalmente, a identificação com um tipo de dados, revela-se crítica, na medida em que a literatura econométrica tende a confundir a análise de dados *TCSC* com dados em painel.

⁴⁹ Embora a base de dados se reporte a 278 municípios, optou-se por excluir os municípios de Odivelas, Vizela e Trofa por não disporem de observações relativas aos períodos eleitorais iniciais, dado terem sido criados posteriormente.

Finalmente para as variáveis utilizadas no trabalho empírico, construídas a partir da base de dados, apresentam-se na tabela que se segue as respectivas estatísticas descritivas.

Variáveis	Observações	Média	Desvio Padrão	Mínimo	Máximo
<i>VVI</i>	1925	48,905	12,522	29,2	92,18
<i>VOTOSANT</i>	1925	48,944	9,7783	26,98	91,74
<i>APART</i>	1925	9,215	6,173	1	25
<i>APRES</i>	1925	6,974	4,589	1	25
<i>DIREAL_t</i>	1925	180,955	136,886	5,037	1439,097
<i>Dummy</i>	1925	0,448	0,498	0	1
<i>DESEMP</i>	1925	7,2	1,832	4	10,4
<i>PDESEMPI</i>	1925	3,121	3,664	0	10,4
<i>ΔDESEMP</i>	1650	-0,617	2,876	-4,4	3,3
<i>INFLAC</i>	1925	13,086	8,338	2,2	24,2
<i>PINFLACI</i>	1925	5,811	8,589	0	24,2
<i>CFcapita_t</i>	1923	5,525	8,100	0	93,235
<i>CFcapita_{t-1}</i>	1652	6,051	9,557	0	171,817
<i>CFcapita_{t-2}</i>	1652	5,385	7,639	0	108,127
<i>ΔPeso_t</i>	1474	9,143	140,22	-128,07	4996
<i>ΔPeso_{t-1}</i>	1529	4,548	50,09	-1	1496
<i>ΔMédia_{t-1}</i>	1600	-0,106	0,476	-1	2,106
<i>ΔMédia_{t-2}</i>	1600	-0,071	0,66	-1	2,977

Tabela 3.2-1: Estatísticas Descritivas

4. Análise Econométrica

Econometrics differs from statistics in defining the identification problem (in terms of structural versus reduced form equations). "Cross-sectional" econometrics (as opposed to time-series) operationalizes Marshall's comparative statics idea (*ceteris paribus*) into its main notion of causality (compare to time-series analysis and its statistical Granger causality definition). The ultimate goal of econometrics is to provide policy evaluation.

In the classical paradigm of econometrics, economic models based on clearly stated axioms allow for a definition of well-defined structural "policy invariant" parameters. Recovery of the structural models allows for induction of causal parameters.

This paradigm was built within the work of the Cowles Commission starting in the 1930s.

J.J. Heckman (2000)

4.1. Avaliação da Estacionariedade

4.1.1. Introdução

A utilização de dados em painel, além de controlar efeitos de dimensão dos municípios e efeitos de ciclos económicos locais, equaciona a heterogeneidade em cada um dos municípios, confere maior variabilidade e, conseqüentemente, reduz o potencial de colinearidade entre variáveis; finalmente, sendo mais informativo gera um maior número de graus de liberdade. A informação amplificada, quer pelo elemento espacial, quer pelo elemento temporal dos dados da amostra, garante uma maior eficiência nos coeficientes estimados permitindo maior precisão nas inferências estatísticas.

A componente temporal presente nos dados em painel (*panel data*), introduz um problema recorrente em séries cronológicas (*time series*), que se prende com o facto do valor de uma variável observada, num certo momento, poder estar afectado pelos valores observados em momentos anteriores, contrariamente ao que acontece com os dados seccionais (*cross-section*) e, ainda, que a intensidade desse efeito é função do valor dos parâmetros relativos a termos desfasados. Ora, um modelo econométrico, ao

pretender testar ou quantificar relações causais entre variáveis, deve evitar regressões espúrias,⁵⁰ que simulam bons ajustamentos e que prevêm relações (não existentes) entre variáveis, estatisticamente significantes; assim num cenário de dados em painel, minimizar esse perigo implica assegurar que todas as variáveis utilizadas são estacionárias, depurando-as, quer da diferença estacionária, quer da tendência estacionária.

Exemplificando, com a base de dados utilizada, se, por hipótese, a percentagem de votos obtida pelo incumbente no acto eleitoral crescesse a uma taxa constante sustentada em factores idiossincráticos da população e no carisma do incumbente e se, simultaneamente, por via das alterações à Lei das Finanças Locais e do crescimento dos preços e de uma maior aprendizagem na utilização de fundos comunitários, as despesas de investimento crescessem, também, a uma taxa constante – ainda que essas duas taxas de crescimento fossem independentes entre si, já que as forças indutoras destes processos são completamente distintas – no final a regressão de uma variável sobre a outra colocar-nos-ia perante uma correlação perfeita, sem que isso signifique uma relação causal entre as duas variáveis, mas tão só uma relação obtida a partir da tendência determinística de cada uma das variáveis e que se imputa à variável independente. No caso mais simples, o comportamento das sucessões reduz-se a um passeio aleatório simples, identificado por $I(1)$.⁵¹

Dado que, um modelo baseado em dados estacionários, pode medir relações de longo prazo entre variáveis, importa avaliar o comportamento das variáveis observadas e caso se esteja perante sucessões não estacionárias, há que absorver o comportamento explosivo das mesmas, transformando-as em sucessões de dados estacionárias.⁵² É neste contexto que importa testar a presença de não estacionariedade nas variáveis utilizadas

⁵⁰ Muita da teoria assintótica convencional relativa à estimação dos mínimos quadrados (como sejam as demonstrações de consistência padrão e de normalidade dos estimadores dos mínimos quadrados) pressupõe estacionariedade das variáveis independentes, possivelmente à volta da tendência determinística. Testar essa estacionariedade torna-se importante para evitar regressões espúrias, as quais são ilustradas por elevados R^2 e baixas estatísticas DW .

⁵¹ Nelson e Plosser (1982) constataram que um grande número de séries temporais de agregados económicos apresenta características tipo caminho aleatório.

⁵² A estacionariedade normalmente é conseguida por diferenças de ordem 1, mas há situações onde se necessita de diferenças de ordem superior. As variáveis que necessitam de ser diferenciadas n vezes para se tornarem estacionárias, denominam-se variáveis $I(n)$ e dizem-se integradas de ordem n .

no modelo e a dimensão temporal, caracterizadora dos dados em painel, impõe a necessidade de aplicação de testes de raiz unitária⁵³ sobre esse painel de dados.

Os testes de raiz unitária, enquanto técnicas de estimação, características de séries temporais, apresentam baixo poder e dimensão distorcida⁵⁴ e seguem complicados processos de Wiener,⁵⁵ porém as conjugações recentes com métodos de estimação de dados em painel, com efeitos específicos fixos ou aleatórios, e, em particular, a introdução de técnicas de estimação para séries temporais modificadas, de forma a combinar-se a endogeneidade dos regressores e heteroscedasticidade e a correlação dos resíduos com métodos de estimação painel com efeitos fixos ou aleatórios,⁵⁶ aumentaram o poder e reduziram a dimensão das distorções dos testes estatísticos de raiz unitária, que passaram, ainda, a comportar-se como distribuições normais.

A utilização de testes de raiz unitária e de cointegração potencia o rigor na determinação do impacto, de longo prazo, do tempo de permanência do partido e do incumbente no poder e das despesas de investimento sobre a percentagem de votos obtidos pelo incumbente. Antes da estimação do modelo econométrico é importante determinar a ordem de integração das variáveis, utilizando testes de raiz unitária para dados em painel. A modelização das variáveis em níveis justifica-se na presença de variáveis com raízes unitárias, por exemplo num universo de variáveis $I(1)$, caso o nível das referidas variáveis seja capaz de formar um vector de cointegração.⁵⁷

⁵³ Já que, actualmente, se tornou impossível analisar série temporais sem que se abordem as raízes unitárias e os testes de raiz unitária e a não estacionariedade tende a ser uma fonte de preocupação à medida que aumenta a componente temporal (T) da base de dados.

⁵⁴ Os testes de hipótese clássicos especificam a hipótese nula e/ou a hipótese alternativa, para que o exame de dados possa rejeitar a hipótese nula (H_0) e aceitar a hipótese alternativa (H_A). Nesse momento, pode-se incorrer em dois tipos de erros: rejeitar incorrectamente a H_0 (erro tipo I) ou aceitar incorrectamente a H_0 (erro tipo II). A probabilidade de se aceitar erradamente a H_A , i. e. a probabilidade de cometer erros tipo I reflecte a dimensão do teste e a probabilidade de se aceitar correctamente a H_A , i.e. a probabilidade de se cometer um erro tipo II, depois de deduzida à unidade, traduz-se no poder do teste. Na generalidade dos testes de raiz unitária a hipótese nula é que as variáveis têm uma raiz unitária, a ser assim o baixo poder dos testes significa que são, muitas vezes, incapazes de rejeitar a hipótese nula, concluindo pela existência de raiz unitária.

⁵⁵ O processo Wiener, W_t , é um processo estocástico contínuo no tempo, que se caracteriza por, (i) $W_0 = 0$; (ii) W_t quase contínuo e (iii) W_t sofrer incrementos independentes com distribuição, $W_t - W_s \sim N(0, t - s)$, para $0 \leq s < t$.

⁵⁶ Como é o caso das técnicas de cointegração, método mínimos quadrados ordinários painel totalmente modificados (*panel, FM-MQO*) propostas por Pedroni (2000, 2001),

⁵⁷ Se a combinação linear de duas ou mais variáveis não-estacionárias é estacionária, as variáveis dizem-se cointegradas e pode-se estimar uma relação de longo prazo introduzindo variáveis por níveis.

4.1.2. Caracterização dos testes: identificação de semelhanças e diferenças

Num passado, relativamente, recente têm sido desenvolvidos, para os dados em painel, múltiplos métodos de testes de raiz unitária, bem como análises de cointegração. A maioria desses testes de raiz unitária baseia-se na estimação, de uma versão, de um modelo dinâmico padrão para dados em painel, do tipo:

$$y_{it} = \rho_i y_{it-1} + \delta_0 + \delta_1 t + \eta_i + \nu_t + \varepsilon_{it}, \quad i = 1, \dots, N; t = 1, \dots, T,$$

para, posteriormente, passarem a testar o valor unitário do coeficiente ρ .⁵⁸ Exemplos desses testes são t^* [Levin, Lin e Chu (2002)], B [Breitung (2000)], MW [Maddala e Wu (1999)] e \bar{t} [Im, Pesaran e Shin (2003)], que ao aplicarem e generalizarem o teste Dickey-Fuller (DF)⁵⁹ e o teste Dickey-Fuller Aumentado (ADF)⁶⁰ ilustram bem essa característica dinâmica do modelo.

⁵⁸ Os índices i identificam os 275 municípios incluídos no painel e t identificam os períodos eleitorais, δ_1 o coeficiente tendência e η_i, ν_t os efeitos específicos às unidades seccionais e ao tempo, respectivamente e $\varepsilon_{it} \sim \text{i.i.d.}(0, \sigma^2)$.

⁵⁹ Os testes DF são obtidos, por estimação, através de uma regressão, de um processo simples auto-regressivo, $y_{it} = \rho_i y_{it-1} + d'_{mt} \delta_{mi} + \varepsilon_{it}$, o qual depois de subtraído por y_{it-1} a ambos os membros, vem ilustrado por, $\Delta y_{it} = \alpha_i y_{it-1} + d'_{mt} \delta_{mi} + \varepsilon_{it}$, em que $\alpha_i = \rho_i - 1$, d'_{mt} vector das variáveis determinísticas (efeitos fixos, tendência determinística e *dummies* temporais) e δ_{mi} o correspondente vector dos coeficientes para o modelo específico m ($m \in \{1, 2, 3\}$, em que $m = 1$ identifica um DF sem constante, nem tendência; $m = 2$ identifica um modelo com termo constante e finalmente $m = 3$ identifica um DF com constante e tendência). As hipóteses, a nula e a alternativa, são: $H_0 : \alpha_i = 0$ e $H_A : \alpha_i < 0$. Um *t-rácio* pode ser utilizado para fazer a avaliação. Dickey e Fuller mostraram que, sob a hipótese nula, esta estatística não segue a distribuição *t-student* convencional. O teste DF é apenas válido se as variáveis satisfazem um processo simples $AR(1)$.

⁶⁰ Na presença de erros correlacionados em série, devem-se considerar os testes ADF ,

$$\Delta y_{it} = \alpha_i y_{it-1} + \sum_{j=1}^{pi} \theta_{ij} \Delta y_{it-j} + d'_{mt} \delta_{mi} + \varepsilon_{it},$$

pois que ao permitirem adicionar desfasamentos de Δy_{it} à regressão DF , sem que isso

afecte a já limitada distribuição do teste, eliminam a potencial correlação em série dos erros (o número de desfasamentos escolhido deve ser suficiente para garantir a ausência de autocorrelação nos resíduos da regressão suporte do referido teste).

Como se acabou de referir os testes, Levin, Lin e Chu (2002), doravante teste *LLC*, e Im, Pesaran e Shin (1977, 2003), teste *IPS* daqui em diante, encontram suporte na estimação de regressões *DF* e *ADF*,⁶¹ com a diferença de que o primeiro utiliza os dados agregados, enquanto o segundo aplica testes de raiz unitária, separadamente, a cada uma das 275 séries cronológicas com a mesma dimensão temporal (7 períodos eleitorais). Entretanto, o teste *LLC* até contempla a variação, entre municípios, dos coeficientes-intercepção, das tendências cronológicas, das variâncias residuais e de autocorrelações padrão de ordem superior, mas em contrapartida requer dados gerados independentemente entre indivíduos e com uma dimensão comum à amostra, deveras limitativo, na medida em que, implicitamente, exige igualdade entre todos os coeficientes individuais de autocorrelação de 1ª ordem, quer sob a hipótese nula, quer sob a hipótese alternativa.⁶² A relaxação deste pressuposto consegue-se através do teste *IPS*, ao assumir-se que sob a hipótese alternativa, pelo menos uma variável,⁶³ não necessariamente todas as variáveis, é estacionária.

O teste *LLC* baseia-se nos *t*-estatísticos dos coeficientes estimados a partir de um processo regressivo entre resíduos ortogonalizados e ajustados, passíveis de serem ilustrados por, $\tilde{\epsilon}_{it} = \beta \tilde{\nu}_{it-1} + \tilde{\epsilon}_{it}$. Quer o resíduo, $\tilde{\epsilon}_{it}$, quer o resíduo, $\tilde{\nu}_{it-1}$, foram $(t, \beta=0)$ primeiro gerados separadamente a partir de duas regressões auxiliares,⁶⁴ tipo *ADF* sem parâmetro auto-regressivo, tendo-se obtido $\hat{\epsilon}_{it}$ e $\hat{\nu}_{it}$, posteriormente, submeteram-se a um processo de normalização, consubstanciado na divisão de $\hat{\epsilon}_{it}$ e de $\hat{\nu}_{it-1}$ pelo erro

⁶¹ Este suporte confere aos testes uma correcção paramétrica dos erros correlacionados em série.

⁶² Este restritivo pressuposto implica que $\alpha_i = \alpha$, já que $\rho_i = \rho$, $\forall i$, na regressão *ADF*.

⁶³ Já que equaciona heterogeneidade entre os coeficientes α_i , na regressão *ADF*, o que faz com que a estatística \bar{t} venha definida como a média das *n* individuais *ADF* estatísticas. Agora, a hipótese nula é $\alpha_i = 0$, para todo o *i*, e é colocada contra a hipótese alternativa, $\alpha_i < 0$, para alguns *i* (ao invés de todos os *i*).

⁶⁴ Uma das regressões seria do tipo, $\Delta y_{it} = \delta_1 d_{1t} + \sum_{j=1}^{p_i} \theta_{1ij} \Delta y_{i,t-j} + e_{it}$, e a outra do tipo, $y_{it} = \delta_2 d_{2t} + \sum_{j=1}^{p_i} \theta_{2ij} \Delta y_{i,t-j} + v_{it}$, em que

d_{mt} engloba o conjunto das variáveis determinísticas.

padrão decorrente de uma regressão tipo *ADF*.⁶⁵ $\tilde{e}_{it} = \hat{e}_{it}/\hat{\sigma}_{\varepsilon i}$ e $\tilde{v}_{it-1} = \hat{v}_{it-1}/\hat{\sigma}_{\varepsilon i}$, de modo a controlar-se a heterogeneidade entre unidades seccionais do painel.⁶⁶ Os *t*-estatísticos convencionais para o coeficiente β têm uma distribuição normal limitada, caso o modelo subjacente não inclua efeitos fixos nem tendências individuais, daí o teste *LLC* corrigir essa estatística utilizando termos de ajustamento, relativos ao primeiro momento, média, $\mu_{\tilde{T}}$ e ao segundo momento, desvio padrão, $\sigma_{\tilde{T}}$,⁶⁷ e introduzindo um rácio, $s_i = \frac{\hat{\sigma}_{y_i}}{\hat{\sigma}_{\varepsilon i}}$, entre os desvios padrão de longo prazo e de curto

prazo,⁶⁸ é passível de vir traduzido por, $t_{\beta}^* = \frac{t_{\beta=0} - NTS\hat{\sigma}_{\beta}^2\mu_{\tilde{T}}}{\hat{\sigma}_{\varepsilon}^2\hat{\sigma}_{\tilde{T}}}$, em que \hat{S} é a média

dos s_i para todo o painel. O teste *LLC* pode equacionar dependência seccional homogénea desde que todas as variáveis sejam expressas como desvios em relação ao valor médio observado em todos os municípios num dado momento eleitoral.

Mais flexível, do que os testes *LLC* e *IPS*, é o teste proposto por Maddala e Wu (1999), *MW* daqui em diante; pois que, além de equacionar dados em painel não equilibrados, consegue utilizar procedimentos não paramétricos e adapta-se a qualquer tipo de teste, isto é a uma hipótese nula que tanto pode testar a integridade unitária, como pode testar a estacionariedade. Caso a hipótese nula formulada seja comum à formulada no teste *IPS*,⁶⁹ então o teste *MW* fundamenta-se nos testes *DF* e *ADF*. Quer o teste *MW*, quer o teste *IPS* se baseiam, para cada uma das variáveis observadas, em 275 testes estatísticos de raiz unitária correspondentes a cada um dos 275 municípios; assumindo, ambos, independência seccional. Apenas, o teste

⁶⁵ $\Delta y_{it} = \delta_{mi}d_{mt} + \alpha_i y_{i,t-1} + \sum_{j=1}^{p_i} \theta_{ij} \Delta y_{i,t-j} + \varepsilon_{it}$.

⁶⁶ O mesmo é dizer que a estandardização dos resíduos visa o controlo de diferentes variâncias entre os diferentes municípios.

⁶⁷ Tabelados por Levin *et al.* (2002), a partir de uma simulação de Monte Carlo.

⁶⁸ Decorrente das críticas ao teste inicialmente desenvolvido, uma versão painel do teste *DF* sem efeitos específicos, que se prendiam com a oscilação das médias e variâncias assintóticas dos testes estatísticos da raiz unitária em função das especificações da regressão, ainda que convergentes para a distribuição normal à medida que *N* e *T* tendem para infinito.

⁶⁹ A hipótese nula testa a presença de raiz unitária, isto é $H_0 : \alpha = 0$, nas variáveis observadas em cada um dos 275 municípios.

MW combina os níveis de significância observados,⁷⁰ isto é os p_i -valores da estimação individual das regressões *DF* e *ADF*,⁷¹ directamente disponíveis e de simples

utilização, de forma que se pode traduzir por, $\lambda = -2 \sum_{i=1}^N \ln P_i$,

o qual segue uma distribuição χ^2 com $2N$ graus de liberdade, sob a hipótese nula; enquanto o teste *IPS* combina os t -testes estatísticos, que decorrem das 275 estimações individuais (municipais) das regressões *DF*, permitindo a formulação,

$$t_{IPS} = \frac{\sqrt{N} \{ \bar{t} - E[t_i | \rho_i = 0] \}}{\sqrt{\text{Var}[t_i | \rho_i = 0]}} \rightarrow N(0,1),^{72}$$

que, em probabilidade, se aproxima do padrão de uma distribuição de probabilidade normal, à medida que N e T tendem para infinito.

Já o teste *Breitung*, assemelha-se ao teste *LLC* no que concerne à hipótese nula, de que cada série cronológica individual, no painel, é integrada *versus* a hipótese alternativa de que todas as séries cronológicas individuais são estacionárias; bem como à regressão *ADF* agregada (*pooled*), em que se baseia,

$$\Delta y_{it} = \delta_{0i} + \delta_{1i}t + \alpha y_{i,t-1} + \sum_{j=1}^{p_i} \theta_{ij} \Delta y_{i,t-j} + \varepsilon_{it},$$

onde se assume um $\alpha = \rho - 1$ comum, e se equacionam efeitos específicos municipais δ_{0i} e coeficientes tendência δ_{1i} .

Entretanto, enquanto o teste *LLC* se fundamenta numa técnica que remove a autocorrelação e a componente determinística, ou seja a intercepção individual e a tendência individual, quando se efectuam as padronizações relevantes, o teste estatístico proposto por *Breitung* utiliza uma técnica que remove a componente autoregressiva da

⁷⁰ Seguindo Fisher (1932).

⁷¹ Designados com a notação, P_i .

⁷² Em que \bar{t} é valor médio das t -estatísticas *ADF* para cada um dos municípios, e a $E[t_i | \rho_i = 0]$ o termo ajustado do momento média das estatísticas individuais *ADF*, e $\text{Var}[t_i | \rho_i = 0]$ o termo ajustado do momento de variância, dada a inexistência de raiz auto-regressiva ρ_i , tabelados por *IPS* (1997, 2003) a partir de uma simulação Monte Carlo.

equação *ADF*, mas não a componente determinística dessa equação.⁷³ O teste de Breitung num primeiro passo é o mesmo que o teste *LLC* com a diferença de que, para se obterem os resíduos \hat{e}_{it} e \hat{v}_{it-1} , a regressão se faz sobre uma equação *ADF*, apenas, especificada sobre o parâmetro $\Delta y_{i,t-j}$, posteriormente o ajustamento para correcção das variâncias específicas individuais segue o mesmo procedimento (como em *LLC*). Já com os resíduos estandardizados, isto é com, $\tilde{e}_{it} = \hat{e}_{it}/\hat{\sigma}_{\varepsilon i}$ e $\tilde{v}_{it-1} = \hat{v}_{it-1}/\hat{\sigma}_{\varepsilon i}$, Breitung dispensa os factores de correcção do enviesamento,⁷⁴ recorrendo a uma transformação baseada na ortogonalização dos resíduos,⁷⁵ passível de ser ilustrada por:

$$e_{it}^* = \sqrt{\frac{T-t}{T-t+1}} \left(\tilde{e}_{it} - \frac{\tilde{e}_{it+1} + \dots + \tilde{e}_{iT}}{T-t} \right) \text{ e } v_{it-1}^* = \tilde{v}_{it-1} - \tilde{v}_{i1} - \frac{t-1}{T} \tilde{v}_{iT}.$$
 Com os resíduos transformados, Breitung corre a regressão
$$e_{it}^* = \underset{(t, \beta=0)}{\beta} v_{it-1}^* + \varepsilon_{it}^*,$$
 tal como acontece no

teste *LLC*, de forma a obter os *t*-estatísticos, os quais segundo Breitung, no limite seguem uma distribuição normal padrão.

Na literatura empírica relativa a dados em painel é frequente a utilização, quer do teste *IPS*, quer do teste *LLC*, embora não isenta de críticas relativas à potência (poder) de cada um desses testes.⁷⁶ A(o) fraca(o) potência (poder) desses testes, para Hadri (2000) decorre da tendência para se aceitar (ou não rejeitar) a hipótese nula de raiz unitária, apesar da forte evidência da hipótese alternativa, ou seja com o potencial de se cometer uma forma de erros tipo *II*; passando a propor um teste em que a hipótese nula é de estacionariedade em qualquer uma das *N* séries cronológicas que constituem o painel. Teste este que, em vez de se basear na generalização dos testes cronológicos *ADF* aos dados em painel (como acontece com *LLC* e *IPS*), se vai suportar na generalização do

⁷³ O poder dos testes *IPS* e *LLC* é muito sensível à especificação dos termos determinísticos.

⁷⁴ Rácio entre os desvios padrão de curto prazo e de longo prazo e momentos (média e desvio padrão).

⁷⁵ Transformação utilizada por Arellano e Bover (1995).

⁷⁶ Bowman, (1999) mostra que há perda de poder do teste *IPS* relativamente ao teste *LLC*, quando há igualdade entre todos os coeficientes individuais de autocorrelação de 1ª ordem do painel. Mostra, ainda, que a velocidade de ajustamento da probabilidade de escolha errada da H_A (dimensão) à probabilidade de escolha correcta da H_A (poder) é superior para o teste *LLC* relativamente ao teste *IPS*, quando apenas um subconjunto das unidades seccionais é estacionária.

teste Kwiatkowski *et al.* (1992), daqui em diante teste *KPPS*,⁷⁷ aplicado às séries temporais do painel. Isto é vai utilizar-se um teste construído a partir dos resíduos de um processo regressivo entre as variáveis suspeitas de não estacionariedade, como é o caso da percentagem de votos do incumbente e/ou das despesas de investimento, e uma qualquer constante ou uma qualquer constante e uma tendência, para cada um dos 275 municípios.

Como já foi, acima, referido, o teste *IPS* permitiu relaxar o pressuposto do teste *LLC*, de que todos os parâmetros autoregressivos, α_i , são iguais, quer sob a hipótese nula, quer sob a hipótese alternativa⁷⁸ e o mesmo acontece com o teste de Hadri, com a diferença que este equaciona a hipótese nula da não existência de raiz unitária em qualquer um dos municípios, constantes do painel, nos sete momentos eleitorais analisados *versus* a hipótese alternativa da existência de raiz unitária no painel, enquanto o anterior testa a hipótese nula de existência de raiz unitária no painel *versus* a hipótese alternativa de não existência de raiz unitária em qualquer um dos municípios, enquanto séries cronológicas de um painel.

Baseados no pressuposto, comum a todos os testes referidos, de que há independência, ou pelo menos correlação zero entre as diferentes variáveis observadas em cada uma das unidades seccionais (municípios), ou seja que a matriz de covariância dos resíduos é uma matriz diagonal, pode-se, ainda, referir o procedimento *LM*⁷⁹ proposto por Hadri, para testar a hipótese nula de que as 275 séries cronológicas no painel, observadas individualmente, são estacionárias relativamente ao nível determinístico ou à tendência determinística, contra a hipótese alternativa de uma raiz unitária nos dados em painel,

⁷⁷ Este teste fundamenta-se nos resíduos de um processo regressivo de uma variável, y_{it} , sobre uma constante, ou uma constante e uma tendência, pelo método dos mínimos quadrados, *MQO*, traduzido por, $y_{it} = r_{it} + \beta_{it} + \varepsilon_{it}$, em que $r_{it} = r_{it-1} + \mu_{it}$, é um caminho aleatório.

⁷⁸ Foi sob este restritivo pressuposto que Levin, Lin e Chu (2002) numa simulação de Monte Carlo, deduziram uma distribuição assintótica, tipo Gaussiana, para o estimador α e o *t-estatístico* associado, depois de equacionados os ajustamentos da média e da variância.

⁷⁹ A estatística *LM* é, $LM = \frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T S_{it}^2 / \hat{\sigma}_\varepsilon^2$, em que S_{it} é a soma parcial dos resíduos, $S_{it} = \sum_{j=1}^t \hat{\varepsilon}_{ij}$, e $\hat{\sigma}_\varepsilon^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{\varepsilon}_{it}^2$ é um estimador consistente da variância do erro (corrigido pelos graus de liberdade).

que revela seguir uma distribuição normal assintótica, do tipo,

$$Z = \frac{\sqrt{N}(LM\hat{M} - \xi)}{\zeta} \rightarrow N(0,1).^{80}$$

Acontece que a violação deste pressuposto da independência entre as observações dos 275 municípios constantes do painel,⁸¹ transversal a todos os testes apresentados, induzirá enviesamentos profundos na dimensão dos resultados dos testes, ou seja, a relaxação do pressuposto de que todos elementos fora da diagonal da matriz de covariância dos resíduos são nulos, fragiliza estes testes, na medida em que deixam de ser válidas as distribuições assintóticas derivadas para os mesmos. A potencial correlação entre municípios, denominada correlação contemporânea pode ser corrigida utilizando métodos fiáveis de mínimos quadrados generalizados, $MQGF$, de forma a repor a ortogonalidade entre as unidades municipais, O'Connell (1998).

A variedade de factores que podem originar dependência entre as variáveis observadas em cada um dos municípios, tais como omissão de factores observados comuns ao painel, efeitos espaciais/externalidades,⁸² factores não observados comuns ao painel ou interdependência geral entre resíduos (a qual pode subsistir, mesmo que todos os efeitos comuns, observados e não observados, sejam equacionados) parecem negligenciáveis no presente estudo, o que nos pode levar a confiar um pouco mais nos testes até agora descritos, os denominados testes de 1ª geração. Todavia, esta assumption, potencialmente falsa, de independência entre municípios pode induzir à rejeição (ou não rejeição) da hipótese nula de raiz unitária, o que torna interessante o desenvolvimento de outro tipo de testes que equacionem a dependência dos municípios.

⁸⁰ Foram necessárias estandardizações baseadas em vectores de funções de movimento Browniano para se obter estatísticas aceitáveis ao procedimento de inferências. Hadri introduziu uma ponte de um movimento Browniano padrão,

$V(r) = W(r) - rW(1)$, em que $W(r)$ é um processo Wiener, para definir $LM\hat{M} \rightarrow E \left[\int_0^1 V(r)^2 dr \right] = \xi$ e $\zeta^2 = \text{var}(\int_0^1 V(r)^2 dr)$. A

média e a variância da variável Z , ξ e ζ , respectivamente, são calculadas usando a técnica da função característica, vindo

$\xi = 1/15$ e $\zeta^2 = 11/6300$.

⁸¹ Ou, pelo menos, de correlação zero entre os 275 municípios.

⁸² *Spacial spill over effects*.

Esta perspectiva parece ganhar relevância, quando se sujeitam as variáveis, suspeitas de não estacionariedade, dos dados painel em estudo, ao teste Hadri. Efectivamente o teste de estacionariedade de Hadri pode introduzir alguma inconsistência nos resultados. Um desencontro entre os resultados dos testes de raiz unitária (*LLC*, *IPS*, *MW*, *Breitung*) e os resultados dos testes de estacionariedade (Hadri), todos testes de primeira geração, é uma realidade comum na presença de interdependência entre os dados observados inter-municípios. Em suma, ser induzido a rejeitar a hipótese nula de raiz unitária com a primeira bateria de testes, que resultam da generalização do teste *ADF* e têm implícita a correcção paramétrica das correlações contemporâneas, e, simultaneamente, ser induzido a rejeitar a hipótese nula de estacionariedade com o teste de Hadri, uma generalização do teste *KPPS*, que não equaciona essa característica dos dados, é uma contradição que nos deve alertar para a presença de correlação contemporânea entre os dados dos diferentes municípios.

O alerta que possa resultar da aplicação simultânea destes diversos testes de 1ª geração deverá ser equacionado posteriormente aquando da estimação do modelo em termos da modelização de uma estrutura simplificada de covariância,⁸³ já que a reduzida dimensão da componente temporal nos dados em painel não é incentivadora de se avançar para testes de 2ª geração, robustos à dependência inter-município.⁸⁴

O presente ensaio apoia-se no pacote estatístico *EViews 5*, que disponibiliza os testes de raiz unitária, Levin, Lin e Chu (*LLC*); Im, Pesaran e Shim (*IPS*); Dickey-Fuller Aumentado (*ADF*) e Hadri, entre outros, para aplicação. Para um enquadramento

⁸³ Uma potencial correlação entre municípios pode ser equacionada se a perturbação ε_{it} passar a incorporar duas componentes: ξ_i , erro específico dos municípios, traduz o choque idiossincrático ($e_{it} \sim N(0, \sigma_i^2)$, *i.i.d.*) e f_t , comum a todos os municípios, equaciona o grau de dependência entre municípios devido a efeitos colectivos significantes (por exemplo, optimismo/pessimismo geral, condições macroeconómicas e condições institucionais). Composição ilustrada por, $\varepsilon_{it} = \xi_i f_t + e_{it}$. Usualmente, assume-se que os parâmetros ξ_i são invariantes entre municípios em cada momento eleitoral e o impacto do termo comum $\xi_i f_t$ pode ser removido simplesmente, subtraindo as médias inter-municipais dos dados, permitindo obter testes *ADF* por município utilizando dados transformados.

⁸⁴ Os testes de 2ª geração exploram erros de co-movimentos, de forma a definir novos testes estatísticos, onde a correlação entre unidades seccionais passa a estar modelada e deixa de ser considerada como parâmetro de perturbação.

teórico desses testes, propõe-se uma breve referência aos mesmos, conforme apresentado no Apêndice **D**.

4.1.3. Testes de raiz unitária: aplicação e resultados

O modelo proposto neste ensaio sob a forma de uma relação linear entre percentagem de votos obtida pelo incumbente na eleição autárquica, *VV1*, e o número de anos consecutivos do mesmo autarca no município, *APRES*; o número de anos consecutivos do mesmo partido no município, *APART*, e as despesas de investimento real, *per capita*, *DIREAL* a partir de uma base de dados constituída por 275 municípios (dimensão seccional) em sete actos eleitorais (dimensão temporal) coloca-nos perante a necessidade de equacionar a estrutura temporal das variáveis, quer exógenas, quer endógena que o compõem. Para se concretizar esse objectivo há que detectar a presença de uma raiz unitária e, caso todas as variáveis sejam não estacionárias, $I(1)$, encontrar a cointegração dos dados em painel.

Os resultados obtidos, a partir dos testes (*IPS*, *MW*, *Breitung* e *LLC*) aplicados⁸⁵ sobre os dados em painel, para as variáveis *APRES*, *APART*, *VV1* e *DIREAL* conforme tabela abaixo, são praticamente unânimes relativamente à rejeição da hipótese nula da existência de uma raiz unitária.⁸⁶ Esta conclusão é reforçada pelos resultados dos testes *ADF* aplicados município a município (ver apêndice **E**) reveladores de que numa fracção razoável de municípios, a hipótese nula de existência de raiz unitária pode ser rejeitada.

⁸⁵ As regressões que servem para testar a estacionariedade das variáveis em nível podem incluir uma constante e uma tendência linear. A inclusão ou a não inclusão da tendência linear têm sido objecto de um debate amplo: McCoskey e Selden (1998) mostraram que as regressões *ADF* não deveriam incluir a tendência linear, uma vez que o termo intercepção comporta, em si mesmo, a tendência, e daí poder-se negligenciá-la para dispor de testes mais poderosos em cenário de amostras restritas; já Hansen e King (1998) defendem a evidência da tendência temporal das variáveis para a inclusão da tendência linear. No caso em estudo os testes de raiz unitária foram efectuados utilizando uma equação que incorpora, quer a formulação com uma constante, quer a formulação com uma tendência temporal.

⁸⁶ As únicas excepções dizem respeito ao teste *LLC* para a variável *APART* (apenas quando não se considera a tendência, sendo que esta é estatisticamente significativa) e ao teste *Breitung* (com tendência) para a variável *DIREAL*.

	desf. ^(a)	t _{IPS} ^(b)	Maddala-Wu ^(b)	Breitung ^(b)	LLC ^(b)
<i>APART</i>		-9,611	1975,09	-2,527	32,153
intercepção sem tendência	[2]	(0,0000)	(0,0000)	(0,0057)	(1,0000)
<i>APART</i>		-432,465	3231,93	-18,232	-829,177
intercepção com tendência	[1]	(0,0000)	(0,0000)	(0,0000)	(0,0000)
<i>APRES</i>		-37,645	2763,1	-5,621	-6,764
intercepção sem tendência	[2]	(0,0000)	(0,0000)	(0,0000)	(0,0000)
<i>APRES</i>		-99,573	3098,23	-9,40	-264,205
intercepção com tendência	[1]	(0,0000)	(0,0000)	(0,0000)	(0,0000)
<i>VVI</i>		-72,109	1998,79	-8,635	-153,622
intercepção sem tendência	[2]	(0,0000)	(0,0000)	(0,0000)	(0,0000)
<i>VVI</i>		-271,326	4159,33	-10,826	-437,348
intercepção com tendência	[1]	(0,0000)	(0,0000)	(0,0000)	(0,0000)
<i>DIREAL</i>		-18,252	1557,28	-12,207	-21,357
intercepção sem tendência	[2]	(0,0000)	(0,0000)	(0,0000)	(0,0000)
<i>DIREAL</i>		-71,727	2827,22	0,446	-127,273
intercepção com tendência	[1]	(0,0000)	(0,0000)	(0,6723)	(0,0000)

Tabela 4.1.3: Testes Painel de Raiz Unitária

Notas: (a) O número entre parênteses corresponde à amplitude óptima de desfasamentos fixada pelo critério Akaike, *AIC*, corroborando Ng e Perron, 2005, na identificação desta selecção como a que melhor garante o desempenho dos testes em termos de poder e dimensão.

(b) O número entre parênteses corresponde à probabilidade do teste ser inferior ao valor crítico.

Em suma, os resultados dos testes em painel de raiz unitária permitem com alguma fiabilidade concluir que as variáveis observadas têm um comportamento estável de longo prazo.

4.2. Especificação do Modelo

4.2.1. Questões metodológicas

Ignorando completamente a estrutura dados em painel, têm-se os dados ordenados uns na sequência de outros, por exemplo têm-se as observações relativas ao município “Aveiro”, de 1979 até 2001, depois as observações relativas ao município “Castelo de Paiva”, de 1979 até 2001 e, assim, sucessivamente, até esgotar os 278 municípios.⁸⁷ Nesse caso, o modelo econométrico estimar-se-ia aplicando simplesmente o método dos mínimos quadrados, *MQO*, a um modelo do tipo,

$$\begin{matrix} \mathbf{y} & = & \mathbf{x} & \boldsymbol{\beta} & + & \boldsymbol{\varepsilon} & & (4.2.1) \\ [NT*1] & & [NT*K] & [K*1] & & [NT*1] & & \end{matrix}$$

$$\varepsilon_{it} \approx \text{i.i.d. } (0, \sigma_{\varepsilon}^2)$$

em que $\mathbf{x} = x_{kit}$, matriz $NT * K$ das k variáveis exógenas, $\mathbf{y} = y_{it}$, vector $NT * 1$ da variável dependente, $\boldsymbol{\beta}$ é o vector $K * 1$ dos coeficientes das k variáveis exógenas e $\boldsymbol{\varepsilon}$ o vector dos erros aleatórios. Pode-se, ainda, identificar a matriz variância-covariância dos erros, $NT * NT$, por $E(\varepsilon_{it}, \varepsilon_{js}) = \boldsymbol{\Omega}$.

⁸⁷ O intervalo temporal entre a observação inicial e a observação final apresenta a mesma amplitude em 275 municípios, para que os dados em painel viessem equilibrados optou-se por considerar apenas esses municípios. (Excluem-se os 3 municípios que não dispõem de observações iniciais, já que são aqueles que, entretanto, foram criados, é o caso de Odivelas, Vizela e Trofa).

Trata-se de um método relativamente simples, identificado com um estimador *pooled* (agregado), ainda que pouco útil e raramente apropriado,⁸⁸ dado o pressuposto da homogeneidade dos elementos de intercepção e dos coeficientes do declive que lhe está subjacente [Pesaran e Smith (1995)]. No cenário das observações municipais, no período de 1979 a 2001, significaria assumir que todas as unidades seccionais – municípios – são idênticas e que a única vantagem de se dispor de dados em painel reside num maior número de observações e conseqüentemente num maior número de graus de liberdade. De facto é pouco ambicioso, dado o manancial de informação disponível.

Ora, uma questão muito importante a considerar nos dados em painel é precisamente a heterogeneidade entre as unidades seccionais, a qual é negligenciada nas *pooled* regressões. No mundo real sabe-se que o pressuposto de uma unidade representativa não se mantém e é precisamente essa diferença de comportamentos inter-unidades que os dados em painel permitem captar.

A heterogeneidade pode verificar-se, quer no parâmetro constante (intercepções), quer no declive (coeficiente das variáveis independentes), quer em ambos, daí ser importante estabelecer se há heterogeneidade (nos casos raros onde não se verifique, a *pooled* regressão é, então, apropriada) e qual a fonte de heterogeneidade. Basta que, apenas, as intercepções sejam heterogêneas entre unidades seccionais, (ou entre períodos temporais) para que a *pooled* regressão venha enviesada.⁸⁹ Caso sejam heterogêneas as intercepções e os declives, neste caso o *pooling* (agregação) com $N * T$ observações só pode produzir resultados sem sentido [Hsiao (1986)].

Caso essa heterogeneidade seja modelada, espera-se que as variáveis exógenas expliquem muito do que é diferente sobre uma observação, quer por unidade seccional, quer por período de tempo, mas ainda assim alguma heterogeneidade ficará por modelar.

⁸⁸ De facto a estimação da regressão pelo método *MQO* será, muito provavelmente, enviesada, ineficiente e/ou inconsistente, quando aplicada sobre dados *pooled* (agregados) e isto porque os erros, decorrentes da estimação, tendem a, (i) não ser independentes em termos temporais (correlação em série), em termos seccionais (correlação contemporânea); (ii) conter componentes, quer cronológicas, quer seccionais, reflectindo efeitos temporais e efeitos espaciais; (iii) ser heteroscedásticos, diferentes variâncias entre diferentes unidades seccionais ou grupos de unidades seccionais e (iv) não ser aleatórios através do espaço e/ou do tempo.

⁸⁹ A direcção deste enviesamento é, *a priori*, impossível de determinar [Hsiao (1986)].

Não conseguindo modelar essa heterogeneidade, o termo do erro ε_{it} virá afectado por essa circunstância. O verdadeiro problema advém quando algumas unidades seccionais (ou, menos comumente, períodos de tempo) partilham alguma heterogeneidade não modelada. Torna-se necessário fazer algo para remover esta heterogeneidade sistemática e partilhada do termo erro. Ora, o enviesamento das variáveis omitidas, gerado pelos estimadores *pooling* na presença de heterogeneidade, pode ser reduzido pela introdução dos efeitos específicos, unidade seccional e/ou tempo, para os dados em painel.

4.2.2. Validação cruzada e a heterogeneidade dos dados

Um universo completamente homogéneo não é muito compatível com uma base de dados que se caracteriza por combinar um espectro espacial, i – unidades seccionais (que, no caso em estudo, serão os municípios, com $i = 1, \dots, 275$), e, simultaneamente, um espectro temporal, $t = 1, \dots, 7$;⁹⁰ em vez disso presumir uma maior ou menor perspectiva gradativa de heterogeneidade subjacente a esses dados e tentar modelá-la revelar-se-á mais enriquecedor do rigor e fiabilidade das inferências realizadas a partir dos modelos estudados. Assim, importa, antes de se avaliar a modelação da heterogeneidade, discutirem-se os testes de homogeneidade.⁹¹

Para o efeito há que introduzir especificações diferentes para o modelo original e continuando a identificar as k variáveis exógenas (explicativas) como x_{kit} , e a variável dependente como y_{it} , há que apurar se os valores de intercepção e/ou de declive são heterogéneos, para isso utilizam-se testes que permitem confrontar as diferentes especificações do modelo correspondentes a diferentes níveis de heterogeneidade com aquela(s) que equaciona(m) a heterogeneidade de ambos os parâmetros.

⁹⁰ No período de 1979 a 2001 considerou-se apenas o ano da realização do acto eleitoral, justificado pelo facto de, no ano das eleições, o orçamento autárquico ser completamente executado, já que o acto eleitoral tem decorrido no mês de Dezembro do ano das eleições.

⁹¹ Os testes de homogeneidade padrão são os F -testes.

Um modelo é considerado superior a outro se conduz para a redução do quadrado da soma dos resíduos (*SSR*), assim para avaliarmos e seleccionarmos a especificação mais apropriada há que estabelecer comparação entre o quadrado da soma de resíduos de diversos modelos que ponderam graus diferentes de heterogeneidade.

O modelo mais genérico é o que equaciona completa heterogeneidade, por isso, definido como modelo não restrito e ilustrado como,

$$M_1 : y_{it} = \alpha_i + \beta'_{ki} x_{kit} + \eta_{it} \quad (4.2.2)$$

em que para todo o $i \neq j$, são diferentes, quer os termos constantes, $\alpha_i \neq \alpha_j$, quer os declives, $\beta_i \neq \beta_j$.

A partir deste modelo, denominado modelo não restrito e anotado por M_1 , equacionam-se várias restrições, as quais reflectem um maior ou menor grau de homogeneidade a ser introduzido nesse modelo, que permitem o desenho de novas modelizações. Assim, caso se,

(i) equacione a restrição em que todos os declives são idênticos, ou seja $\beta_i = \beta_j$, para todo $i \neq j$, a especificação do modelo, vem,

$$M_2 : y_{it} = \alpha_i + \beta'_k x_{kit} + \eta_{it} \quad (4.2.3)$$

(ii) equacione a restrição em que todas as intercepções são idênticas, ou seja que $\alpha_i = \alpha_j$, para todo o $i \neq j$, o modelo original vem reescrito,

$$M_3 : y_{it} = \alpha + \beta'_{kit} x_{kit} + \eta_{it} \quad (4.2.4)$$

(iii) equacione a restrição em todos os declives e de todas as intercepções permanecem iguais, estar-se-ia no domínio do modelo *pooled*.

$$M_4 : y_{it} = \alpha + \beta'_k x_{kit} + \eta_{it} \quad (4.2.5)$$

Em todos os modelos, de M_1 , equação (4.2.2), a M_3 , equação (4.2.4), os coeficientes α e β são assumidos como integrantes de modelos de efeitos fixos. Se a homogeneidade for rejeitada, a heterogeneidade entre unidades/seccionais pode ser captada por quaisquer um desses modelos, representados pelas equações, (4.2.2) a (4.2.4). O propósito destes modelos é reduzir o enviesamento decorrente das variáveis omitidas.

O modelo mais utilizado com dados em painel assume que os declives são idênticos, mas as intercepções variam entre unidades/seccionais ou através do tempo. O efeito específico individual (ou específico temporal) pode ainda ser considerado num modelo de efeito fixo ou num modelo de efeito aleatório, cada um dos quais oferecendo diversos estimadores.

Como já foi referido, aplica-se um *F – teste* padrão para testar a presença de heterogeneidade, utilizando para o efeito o quadrado da soma dos resíduos (*SSR*) dos modelos que se pretendem testar. Assim,

(A) Numa primeira etapa, começa-se por contrapor o modelo 4, M_4 , e o modelo 1, M_1 , testa-se a hipótese de uma estrutura perfeitamente homogénea (intercepções e declives idênticos), isto é testa-se, $H_0^1 : \alpha_i = \alpha, \beta_i = \beta \quad \forall i \in \{1, \dots, N\}$, utilizando-

se uma estatística de Fisher para testar $[(N-1)(K+1)]$ restrições lineares.⁹² Esta estatística segue uma distribuição $F_{[(N-1)(K+1)][N(T-K-1)]}$, com $[(N-1)(K+1)]$ e $[N(T-K-1)]$ graus de liberdade. Se o valor do F – estatístico observado, sob a hipótese nula, é significativo, rejeita-se a hipótese nula, os dados não podem ser *pooled* para $N * T$ observações, a fonte de heterogeneidade deve ser mais investigada;

(B) Passa-se a uma segunda etapa, que consiste em determinar se a heterogeneidade é atribuível aos coeficientes β_i ; ou seja há que testar a igualdade para todas as unidades seccionais dos K vectores β_i , isto é a hipótese, $H_0^2 : \beta_i = \beta \quad \forall i \in \{1, \dots, N\}$. Em termos práticos, vai contrapor-se os *SSR* do modelo 2, M_2 , com os do modelo 1, M_1 . Sob a hipótese nula não há qualquer imposição sobre as constantes α_i , neste cenário e do mesmo modo se constrói a estatística de Fisher para testar estas $(N-1)K$ restrições lineares. Esta estatística, sempre sob a hipótese de independência e normalidade dos resíduos, segue uma lei de Fisher, com $[(N-1)K]$ e $[NT - N(K+1)]$ graus de liberdade. Se o F – estatístico é significativo, rejeita-se a hipótese nula de homogeneidade dos coeficientes β_i , rejeitando-se consequentemente a estrutura painel. Os parâmetros vectoriais β_i viriam estimados utilizando modelos individuais, município a município, conforme ilustrado em M_1 . Caso se aceitasse a hipótese nula de homogeneidade dos coeficientes β_i , a estrutura painel manter-se-ia e, num momento posterior, ir-se-ia apurar da dimensão individual das constantes α_i ;

(C) O procedimento, subjacente a esta terceira etapa, centra-se num teste à igualdade das N constantes individuais α_i , com imposição de coeficientes, β_i , comuns a todos os indivíduos, isto é, $H_0^3 : \alpha_i = \alpha \quad \forall i \in \{1, \dots, N\}$. De novo se está na presença de uma

⁹² No modelo em estudo cada vector β_i compreende K parâmetros. Para as N unidades seccionais do painel, obtêm-se NK parâmetros. A igualdade dos N vectores β_i impõe $NK - K$ restrições. Da mesma forma a igualdade das N constantes introduz $N - 1$ restrições. Em termos globais a hipótese nula acaba a impor $(K+1)(N-1)$ restrições lineares.

estatística seguindo a lei de Fisher com $[(N - 1)K]$ e $[N(T - 1) - K]$ graus de liberdade. Se a hipótese nula de homogeneidade das constantes α_i , vem rejeitada então o modelo que prevalece é o modelo painel com efeitos seccionais, M_2 . Caso a hipótese nula fosse aceite estar-se-ia numa estrutura painel totalmente homogénea (modelo *pooled*), M_4 . Este teste da hipótese, $H_0^3 : \alpha_i = \alpha \quad \forall i \in \{1, \dots, N\}$, serve para confirmar ou rejeitar as conclusões obtidas com o teste da hipótese, H_0^1 , já que o facto de apresentar redução nas restrições lineares aumenta-lhe poder.

Testes similares podem ser utilizados para investigar efeitos específicos temporais que são constantes entre unidades/seccionais, os quais se comportam de forma similar aos anteriores.

O procedimento geral para os testes de homogeneidade referidos (Hsiao, 1986), suportou-se, sempre, na estatística de Fisher; por sua vez essas diferentes estatísticas de Fisher, associadas aos diversos testes de homogeneidade, calculam-se com base na soma dos quadrados dos resíduos, SSR , dos modelos em confronto.

Para o efeito, importa identificar,

(1) a soma dos quadrados dos resíduos do modelo não restrito, M_1 , SSR_U , definido como somatório das 275 somas dos quadrados dos resíduos obtidos em cada uma das 275 regressões individuais, expresso por,

$$SSR_U = \sum_{i=1}^N \frac{\sum_{t=1}^T (y_{it} - \bar{y}_i)^2 - \sum_{t=1}^T (x_{kit} - \bar{x}_{ki})(y_{it} - \bar{y}_i)'}{\sum_{t=1}^T (x_{kit} - \bar{x}_{ki})(x_{kit} - \bar{x}_{ki})'}, \quad \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{x}_{ki} = \frac{1}{T} \sum_{t=1}^T x_{kit}$$

(2) a soma dos quadrados dos resíduos do modelo agregado (*pooled*), M_4 , SSR_4 , correspondente a uma transformação total da variância dos resíduos (*variância total*), obtidos a partir da estimação de um modelo único sobre os $275 * 7$ dados agregados, expressa por,

$$SSR_4 = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y})^2 - \left[\sum_{i=1}^N \sum_{t=1}^T (x_{kit} - \bar{x})(y_{it} - \bar{y})' \right] \left[\sum_{i=1}^N \sum_{t=1}^T (x_{kit} - \bar{x})(y_{it} - \bar{y})' \right]}{\sum_{i=1}^N \sum_{t=1}^T (x_{kit} - \bar{x})(x_{kit} - \bar{x})'}$$

$$\bar{y} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it}, \quad \bar{x} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{kit}$$

(3) a soma dos quadrados dos resíduos do modelo de efeitos individuais, M_2 , em que não se permite declives heterogéneos, apenas se admite intercepções diferentes, corresponde a um modelo *Within* em que os parâmetros estimados são obtidos através de um processo de centralização das variáveis sobre as respectivas médias individuais, traduzido em,

$$SSR_2 = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)^2 - \left[\sum_{i=1}^N \sum_{t=1}^T (x_{kit} - \bar{x}_{ki})(y_{it} - \bar{y}_i)' \right] \left[\sum_{i=1}^N \sum_{t=1}^T (x_{kit} - \bar{x}_{ki})(y_{it} - \bar{y}_i)' \right]}{\sum_{i=1}^N \sum_{t=1}^T (x_{kit} - \bar{x}_{ki})(x_{kit} - \bar{x}_{ki})'}$$

O pressuposto básico subjacente a qualquer um dos modelos, ilustrados pelas equações de, (4.2.2) a (4.2.4), é que, condicionado ao conjunto de k variáveis independentes x_{kit} ,

os efeitos de todas as variáveis omitidas ou excluídas podem resultar de três tipos de orientações:

- (i) Variações seccionais mas invariantes no tempo [é o caso da competência do autarca num município, desde que se verifique a respectiva recandidatura ao longo do período estudado].
- (ii) Variações temporais mas invariantes nas unidades/seccionais [é o caso da taxa de inflação ou da taxa de juro].
- (iii) Variações seccionais e variações temporais [é o caso do esforço desenvolvido pelo autarca]

Uma interpretação para os efeitos das variáveis omitidas, acima referidos, de (i) a (iii), é a de, ou permanecerem constantes através do tempo para uma dada unidade seccional, ou serem os mesmos para todas as unidades seccionais num dado momento do tempo, ou, então, uma combinação dos dois. Nestas circunstâncias, estes efeitos podem ser absorvidos pelo termo constante do modelo da regressão. Esse será o caso do modelo de efeitos fixos (*FEM*).

Outra interpretação para os efeitos das numerosas variáveis seccionais omitidas variantes ao longo do tempo descritas de (i) a (iii), é a de serem individualmente pouco relevantes, mas em conjunto incluírem uma variável aleatória significativa. Assumem-se não estarem correlacionados com o conjunto de variáveis independentes contidas no modelo, x_{kit} , nem com as variáveis excluídas. Esta interpretação é consistente com a especificação conhecida como modelo de efeitos aleatórios (*REM*).

Em suma, a admissão de efeitos específicos, seccionais e/ou temporais, nos dados em painel, permite reduzir o enviesamento e a inconsistência dos estimadores *pooling* na presença de heterogeneidade decorrente das variáveis omitidas. Estes efeitos específicos podem ser equacionados em modelos de efeitos fixos (*FEM*) ou em modelos de efeitos aleatórios (*REM*), cada um deles disponibilizando vários estimadores, os quais passarão a ser analisados nas secções seguintes. Finalmente, uma vez determinada a presença ou a não presença de heterogeneidade nos dados em painel e identificada a fonte da heterogeneidade observada, o modelo pode ser especificado.

Perseguindo o objectivo de proceder à especificação do modelo, ir-se-á procurar captar os efeitos não observados, sejam efeitos de unidades seccionais, sejam efeitos de períodos temporais e/ou de ambos. Para o propósito assumir-se-á, primeiro, diferenças nas intercepções entre unidades seccionais e/ou entre períodos temporais subjacente ao modelo de efeitos fixos, que será abordado na secção seguinte, onde, após uma abordagem teórica, se apresentará o resultado da aplicação às variáveis observadas a partir da base de dados. Evidentemente que, sob pena do nosso propósito ficar incompleto, impõe-se a abordagem, numa secção posterior, do modelo de efeitos aleatórios, de forma a explorar-se a presença de diferenças nas variâncias do erro.

4.3. Estimadores de Efeitos Fixos

4.3.1. Abordagem teórica da estimação

Para introduzir o modelo dos efeitos específicos importa ter presente que se pudéssemos assumir, relativamente aos estimadores *pooling*, os pressupostos de Gauss-Markov, subjacentes ao óptimo de um modelo de regressão linear pelo método dos mínimos quadrados, *MQO*, de que os erros se apresentam distribuídos idêntica e independentemente, isto é, $\varepsilon_{it} \approx \text{i.i.d. } (0, \sigma_\varepsilon^2)$, ou seja se pudéssemos assumir que a perturbação estocástica é homoscedástica e não está correlacionada em série, quer através das unidades seccionais (municípios), quer ao longo do tempo (períodos eleitorais), obteríamos os melhores coeficientes estimados lineares e não enviesados,⁹³ expressos pelo vector $\hat{\beta}$, de dimensão $K * 1$, $\hat{\beta} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}$,⁹⁴ pela simples aplicação do método dos mínimos quadrados, *MQO*, a um modelo equilibrado de dados painel.

Entretanto, acredita-se que estes pressupostos não são sustentáveis na presença de heterogeneidade não modelada resultante da natureza da base de dados, mais especificamente aquela que decorre da dimensão espacial (diferentes municípios) e da

⁹³ BLUE.

⁹⁴ Com uma covariância $\text{Cov}(\hat{\beta}) = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{\Omega}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}$, que no caso especial de $\mathbf{\Omega} = \sigma^2\mathbf{I}$ vem $\text{Cov}(\hat{\beta}) = \sigma^2(\mathbf{x}'\mathbf{x})^{-1}$.

dimensão temporal (diversos momentos eleitorais). Esta convicção vem ilustrada pela assumption de que o termo erro se decompõe da seguinte forma, $\varepsilon_{it} = \mu_i + \gamma_t + \eta_{it}$, onde μ_i é o efeito específico das unidades seccionais (municípios), γ_t é o efeito específico dos períodos temporais (actos eleitorais) e η_{it} é o efeito comum municípios/actos eleitorais, perturbação estocástica remanescente.⁹⁵

O modelo de efeitos fixos (*FEM*) assume que os efeitos específicos das unidades seccionais invariantes no tempo, μ_i , e/ou os efeitos específicos aos períodos temporais comuns às diferentes unidades seccionais, γ_t , devem ser captados como diferenças entre os termos da intercepção da regressão. Resulta pois que μ_i e/ou γ_t passam a ser parâmetros a equacionar aquando da especificação do modelo e para tal aparecem identificados como termo(s) intercepção desse modelo. Neste cenário a remoção dos efeitos específicos, seccionais e/ou temporais que o modelo de efeitos fixos nos propõe faz-se via transformação dos dados.⁹⁶

Ilustrar o processo de remoção dos efeitos específicos via transformação do modelo original/*pooled*, depois de considerados os efeitos específicos seccionais, μ_i ,⁹⁷ exige *em primeiro lugar* especificar esse modelo. A expressão desses efeitos específicos seccionais reflecte-se no termo α_i , sugerindo variabilidade nos termos da intercepção do modelo, vindo uma especificação do tipo da apresentada para o M_2 , assim,

$$y_{it} = \alpha_i + \beta_k' x_{kit} + \eta_{it},$$

⁹⁵ Assume-se $\eta_{it} \approx$ i.i.d. $(0, \sigma_\eta^2)$ e, ainda que as k variáveis exógenas, x_{kit} , são independentes de η_{it} , para todo o i e para todo o t .

⁹⁶ Para Johnson e DiNardo (1997) um estimador de efeitos fixos é aquele que se obtém por qualquer transformação que liberte um modelo de dados painel dos efeitos fixos.

⁹⁷ Dado que o efeito específico temporal obtém tratamento paralelo, em nome da simplicidade considera-se apenas os efeitos específicos seccionais.

tendo subjacente os pressupostos de que $E(\eta_{it}) = 0$; de que $E(\eta_{it}^2) = \sigma^2$, para todo o i ; e de que $E(\eta_{it}, \eta_{jt-s}) = 0$, para $s \neq 0$ e $i \neq j$. O mesmo que, pressuporem-se coeficientes do declive e variâncias iguais através das unidades seccionais, ainda que admitindo discordância nos coeficientes das intercepções e assumindo todas as covariâncias entre unidades seccionais nulas. Para, *posteriormente*, se introduzir a via da transformação, a qual pode ser desenvolvida de diversas formas, como,

(1) desvios relativamente às médias dos valores observados nas unidades seccionais, e/ou às médias dos valores observados num período de tempo, (transformação *dentro*, “*Within*”)

$$(y_{it} - \bar{y}_{i\cdot}) = \beta'_k(x_{kit} - \bar{x}_{i\cdot}) + (\eta_{it} - \eta_{i\cdot}),^{98}$$

e por essa via remove-se a intercepção específica;⁹⁹

(2) médias de unidades seccionais, (transformação *entre*, “*Between*”)

$$\bar{y}_{i\cdot} = \beta'_k \bar{x}_{i\cdot} + \eta_{i\cdot}$$

A perspectiva dos estimadores fixos suportada na transformação *dentro*, “*Within*”, W , é relevante quando é expectável que as médias das variáveis dependentes, para cada município e/ou para cada período eleitoral, venham diferentes, embora a variância dos erros o não venha. Estes estimadores, também, são conhecidos como estimadores dos mínimos quadrados de variáveis *dummy*, *MQVD*,¹⁰⁰ e, ainda, passíveis de se

⁹⁸ Ou $(y_{it} - \bar{y}_{i\cdot}) = \beta'_k(x_{kit} - \bar{x}_{i\cdot}) + (\eta_{it} - \eta_{i\cdot})$.

⁹⁹ Tal como Ahn e Schmidt (1995) referiram, o modelo de efeitos fixos trata o efeito individual como um ruído que é removido via transformação *dentro* (*Within transformation*).

¹⁰⁰ De acordo com Johnston e DiNardo (1997), Greene (2003), Hsiao (1986) e Pesaran e Smith, (1995).

identificarem como estimadores da covariância, CV , uma vez que se admite que os efeitos específicos, seccionais e/ou temporais, estão correlacionados com as variáveis explicativas, ou seja $cov(x_{kit}, \alpha_i) \neq 0$, obrigando a que o efeito específico, α_i e/ou γ_t , venha(m) tratado(s) como um parâmetro desconhecido [Hsiao (1986)].¹⁰¹ Já a baseada na transformação *entre*, “*Between*”, B , grupos ignora a variação *dentro* dos grupos, transformando os dados em médias de grupo. O estimador *entre* grupos produz um estimador consistente e não enviesado, quando os regressores são estritamente exógenos e os coeficientes diferem aleatória e independentemente dos regressores.

A metodologia para a remoção dos efeitos específicos subjacente, ao método dos estimadores fixos *dentro*, $MQVD$ e ao método dos estimadores fixos *entre*, B , consiste na pré multiplicação do modelo original/*pooled* por uma matriz de transformação simétrica e idempotente, \mathbf{M}_d ,¹⁰² e \mathbf{P} ,¹⁰³ respectivamente (para mais detalhe ver apêndice F).

Evoluir nesta questão da remoção dos efeitos específicos, seccionais e/ou temporais, impõe readmitir a suspeita (introduzida no início desta secção) de presença de heteroscedasticidade e de correlação contemporânea nos erros das estimações decorrentes da aplicação do método regressivo dos mínimos quadrados, MQO , a modelos com estrutura de dados painel e, suportados nessa tese, há que ponderar a perspectiva econométrica que assume as violações dos pressupostos Gauss-Markov como um ruído na estimação a ser corrigido pela utilização do método regressivo fiável dos mínimos quadrados generalizado, doravante MQG_F .¹⁰⁴

¹⁰¹ Este estimador é o melhor estimador linear não enviesado, $BLUE$, quando o efeito específico seccional é invariante no tempo [Hsiao (1986)].

¹⁰² É uma matriz que transforma valores observados em desvios relativamente às médias individuais (médias das observações em todas as unidades seccionais num dado período eleitoral). E que ao ser, (i) simétrica faz com que, $\mathbf{M}_d = \mathbf{M}_d'$ e (ii) idempotente faz com que, $\mathbf{M}_d = \mathbf{M}_d^2$.

¹⁰³ É uma matriz que transforma em médias os valores observados ao longo do tempo em cada unidade seccional. E que ao ser, (i) simétrica faz com que, $\mathbf{P} = \mathbf{P}'$ e (ii) idempotente faz com que, $\mathbf{P} = \mathbf{P}^2$.

¹⁰⁴ Uma estimação, denominada método generalizado dos mínimos quadrados, MQG , revelar-se-ia mais robusta às características, espacial e temporal, dos dados painel, porém a sua concretização exigiria conhecer os elementos da matriz de variância-covariância

Para dar sequência à perspectiva que apresenta o método MQG_F como meio para a correção do ruído que decorre de erros heteroscedásticos e correlacionados contemporaneamente vai assumir-se que as observações são inter-temporalmente independentes. A melhoria da eficiência é conseguida através de um sistema de ponderação que pondera as observações de acordo com a inversa da matriz variância-covariância dos erros, Ω , que se apresenta como uma matriz diagonal por blocos, de dimensão, $NT * NT$, em que cada matriz bloco é uma matriz de covariâncias contemporâneas, Σ ,¹⁰⁵ $N * N$, o que decorre do pressuposto do processo de erro ser caracterizado por,

$$E(\varepsilon_{i,t}, \varepsilon_{j,s}) = \begin{cases} \sigma_i^2, & \text{se } i = j \text{ e } s = t \\ \sigma_{i,j}, & \text{se } i \neq j \text{ e } s = t \\ 0, & \text{outros casos} \end{cases} .$$

O método MQG_F ultrapassa o problema do não conhecimento dos valores da matriz variância-covariância, Ω , através de um procedimento¹⁰⁶ que, primeiro, aplica o método regressivo dos mínimos quadrados, MQO , sobre o modelo original especificado na equação (4.2.1), para se obterem os resíduos necessários à estimação da

dos erros. Na prática essa matriz Ω nunca é conhecida, na maior parte das situações a matriz variância-covariância tem que ser estimada, com MQO . O pressuposto MQG de que a matriz é conhecida, vem contrariado, estar-se-ia no domínio de um método generalizado dos mínimos quadrados fiável, daí a denominação MQG_F .

$$^{105} \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_2^2 & \dots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \dots & \sigma_n^2 \end{bmatrix} .$$

¹⁰⁶ Este procedimento foi primeiro descrito por Parks (1967) e depois popularizado por Kmenta (1986), daí que, usualmente, seja conhecido como método Parks-Kmenta.

matriz $\hat{\Sigma}$;¹⁰⁷ segundo, transforma esse modelo por via da introdução dessa matriz $\hat{\Sigma}$; finalmente estima os parâmetros do modelo¹⁰⁸ e os erros padrão, fazendo correr sobre o modelo transformado, novamente, uma regressão *MQO*.

O método *MQGF* não é uma panaceia para corrigir a heteroscedasticidade painel e a correlação contemporânea dos erros nos dados painel já que a sua utilização requereria estimar um número elevado de parâmetros para as covariâncias do erro que se poderia tornar incomportável numa estrutura de dados painel dominada pela dimensão temporal ou, ainda pior, poderia conduzir a que se deixasse de estimar a matriz de variância-covariância dos erros e se passasse a assumi-la conhecida. Estar-se-ia perante uma subestimação dos erros padrão *MQGF*.

Persistindo na utilização do método *MQGF* para corrigir a heteroscedasticidade painel importa, contudo, ter presente as limitações referidas acima, para o efeito há que assumir que todas as covariâncias do erro entre as diferentes unidades seccionais são nulas, ainda que cada unidade tenha a sua própria variância do erro, passível da seguinte ilustração,

¹⁰⁷ Cada elemento desta matriz resulta, se na diagonal, da média do erro quadrado em cada unidade seccional i , isto é, $\sigma_i^2 = \frac{1}{T}(\hat{e}_{i1}^2 + \hat{e}_{i2}^2 + \dots + \hat{e}_{iT}^2)$ e, se fora da diagonal, da média do produto cruzado dos erros de duas unidades seccionais, ou seja, $\sigma_{ij} = \frac{1}{T}(\hat{e}_{i1}\hat{e}_{j1} + \hat{e}_{i2}\hat{e}_{j2} + \dots + \hat{e}_{iT}\hat{e}_{jT})$. Esta perspectiva fundamenta-se no pressuposto de que a correlação entre unidades seccionais, num dado período temporal, é a mesma.

¹⁰⁸ O valor estimado de β_{MQGF} é $(\mathbf{x}'\hat{\Omega}^{-1}\mathbf{x})^{-1}(\mathbf{x}'\hat{\Omega}^{-1}\mathbf{y})$, em que $\hat{\Omega}$ é o valor estimado da matriz de variância-covariância Ω , cujos elementos são s_{ij} em vez das (co)variâncias σ_{ij} , sendo desconhecidas vêm estimadas através de $s_{ij} = \frac{1}{T}\sum_{t=1}^T e_{it}e_{jt}$. Esta substituição tem consequências na construção dos intervalos de confiança. Em sentido estrito, esta substituição é, apenas, válida assintoticamente (isto é, à medida que $N \rightarrow \infty$ e, ainda, se os elementos de Ω são estimados com estimadores consistentes). Esta ambivalência reflecte-se na utilização, umas vezes, dos valores críticos z -estatíst. reflectindo a natureza assintótica dos estimadores e, outras vezes dos valores críticos t -estatíst., nos resultados *MQGF*, apresentados pelos pacotes estatísticos.

$$E(\varepsilon_{i,t}\varepsilon_{j,s}) = \begin{cases} \sigma_i^2, & \text{se } i = j \text{ e } s = t \\ 0, & \text{outros casos} \end{cases} .$$

Com esta assumção reduz-se a limitação que decorria da potencial impossibilidade de se estimar um número insustentável de parâmetros, uma vez que, agora, apenas se torna necessário estimar N parâmetros de variância de erros. O actual procedimento $MQGF$ segue o precedente, denominado por método Parks-Kmenta, já que, numa primeira ronda estima uma regressão MQO e numa segunda ronda uma regressão MQO ponderada, em que as ponderações vêm inversamente proporcionais ao valor de σ_i estimado para cada unidade seccional.

Este espaço, de breves referências metodológicas aos modelos de efeitos fixos de estimação dos dados painel, não pode terminar sem se reflectir nas restrições dos estimadores produzidos pelas duas variantes do método $MQGF$,¹⁰⁹ as quais, apesar de tudo, não são suficientes para justificar uma opção pelo estimador MQO e por isso talvez valha a pena considerar uma correcção dos erros padrão induzida pela estrutura painel, isto é assumir-se que o valor do coeficiente estimado, via MQO , que tem subjacente uma matriz de variância-covariância do erro que é, apenas, o produto de uma constante por uma matriz identidade, deve ser corrigido estimando uma matriz de variância-covariância, $\hat{\Omega}$,¹¹⁰ que tenha em conta a estrutura repetida no tempo dos dados painel. Esta técnica denominada de *PCSE* (*Panel Correct Standard Errors*) estima os coeficientes $\hat{\beta}$, do modelo ajustado, via MQO ; posteriormente, através dessa

¹⁰⁹ Beck e Katz (1995) mostraram que os ganhos de eficiência resultantes de se corrigir a correlação contemporânea são relativamente pequenos enquanto o enviesamento do erro padrão é relativamente elevado, por isso sugeriram estimar os erros padrão de acordo com um procedimento robusto que incorpore informação de potencial correlação contemporânea na matriz variância-covariância sem ajustamento do coeficiente estimado.

¹¹⁰A matriz estimada de variância-covariância, expressa por $\hat{\Omega} = \hat{\Sigma} \otimes \mathbf{I}_T$ implica o conhecimento dos resíduos decorrentes de se ajustar, ao modelo original definido pela equação (4.2.1), uma regressão MQO , de forma a estimar-se $\hat{\Sigma}$, a matriz bloco das covariâncias contemporâneas, traduzida por, $\hat{\Sigma} = (\mathbf{E}'\mathbf{E})/T$, em que \mathbf{E} reflecte a matriz, $T * N$, dos resíduos MQO .

matriz variância-covariância, corrige o erro padrão dos coeficientes estimados, tomando a raiz quadrada dos elementos da diagonal de,

$$Cov(\hat{\beta}) = (\mathbf{x}'\mathbf{x})^{-1} \left\{ \mathbf{x}'\hat{\Omega}\mathbf{x} \right\} (\mathbf{x}'\mathbf{x})^{-1} = (\mathbf{x}'\mathbf{x})^{-1} \left\{ \mathbf{x}' \frac{\mathbf{E}'\mathbf{E}}{T} \otimes \mathbf{I}_{T\mathbf{x}} \right\} (\mathbf{x}'\mathbf{x})^{-1}.$$

Em certa medida, consegue, previamente à estimação final, prevenir a correlação em série suportada em erros padrão robustos. O *PCSE* é uma técnica de estimação robusta em termos de White-Huber.

Após se terem analisado algumas questões relativas a procedimentos subjacentes à estimação dos efeitos, seccionais e/ou temporais, passíveis de serem captados por dados em painel, pelo modelo de efeitos fixos (*FEM*), valerá a pena transpô-las para o modelo empírico desenhado na secção 3.1.. Para a respectiva concretização proceder-se-á a uma aplicação que se caracteriza por ajustar regressões sobre esse modelo de forma a estimarem-se os coeficientes das variáveis observadas e os valores dos quadrados das somas dos resíduos (base dos testes *F* – padrão) para se seleccionar a regressão que melhor adapta o modelo em estudo, o que acontecerá na secção seguinte.

4.3.2. Abordagem técnica da estimação

4.3.2.1. Exploração da estrutura em painel do modelo

Concretizar o objectivo sugerido pelo título desta subsecção exige revisitar o modelo empírico que suporta o modelo teórico desenhado na secção 3.1., vindo:

$$VV1_{it} = \alpha_{it} + \beta_1 * VOTOSANT_{it} + \beta_2 * APART_{it} + \beta_3 * APRES_{it} + \\ + \beta_4 * DIREAL_{it} + \beta_5 * PDESEMP1_{it} + \beta_6 * PINFLAC1_{it} + \varepsilon_{it}$$

Em que,

$VV1_{it}$, é a percentagem de votos obtido pelo partido em exercício de funções no executivo autárquico no município i no ano de eleição t ;

$VOTOSANT_{it}$, é a percentagem de votos que o partido em exercício de funções no executivo autárquico no município i no ano de eleição t tinha obtido na eleição anterior;

$APRES_{it}$, traduz o número de anos sucessivos no exercício do cargo de presidente da câmara no município i à data de eleição t ;

$APART_{it}$, ilustra o número de anos sucessivos que o partido vencedor esteve ou não à frente dos destinos do município i à data de eleição t ;¹¹¹

$DIREAL_{it}$, identifica as despesas de investimento, em valor real, em termos *per capita*, feitas no município i no ano de eleição t e

¹¹¹ Esta variável, $APART_{it}$, surge razoavelmente correlacionada com a variável $APRES_{it}$, aproximadamente 0,6743. Valor que não significa, necessariamente, que uma qualquer variação numa dessas variáveis explique a variação na outra variável e vice-versa. É possível, pela natureza das observações em análise, que a correlação entre estas duas variáveis venha explicada por uma terceira variável: por exemplo a população (dimensão e composição), ou, ainda, a prevalência de um certo conservadorismo na população (traduzível num espírito avesso à mudança). Aliás, se avançarmos para o teste de hipótese *de não causalidade à Granger*, equacionando até 4 períodos de desfasamento, constata-se que esta hipótese é aceite com significância estatística até 3 períodos, em ambiente dados agregados (*pool*). Esta evidência comporta-se, ao contrário, em universo dados painel, onde a hipótese nula *de não causalidade à Granger* começa a aceitar-se a partir do 3º período de desfasamento (esta precedência estatística confirma o apelo intuitivo de que quaisquer uma dessas variáveis encontra explicação em si própria). A não convergência de resultados do teste de Granger, suscita reflexão acerca dos efeitos decorrentes de uma estrutura painel de dados, a qual também, servirá para fundamentar o equacionar das duas variáveis no modelo econométrico - uma vez que os, potenciais, efeitos de multicolinearidade, se introduzidos, acabam atenuados pela estrutura painel dos dados.

$PDESEMP1_{it}$ e $PINFLAC1_{it}$, são variáveis macroeconómicas, respectivamente, a taxa de desemprego nacional e a taxa de inflação nacional, quando o partido incumbente coincide no município i e no governo central, no momento t .

As vantagens que decorrem das características subjacentes à base de dados painel impõem que, em termos de metodologia de trabalho, se comece por explorar a estrutura painel do modelo, em observação. A ser assim, importa testar a *poolability* dos dados, no caso de se agregarem todas as unidades seccionais, concretamente, os municípios, tendo presente que a agregação temporal, também, pode ser realizada. Efectivamente, num cenário em que a dimensão seccional, N , tende a dominar face à componente temporal, T , ou seja para um painel caracterizado por um número relativamente grande de indivíduos observados, mas durante um período relativamente curto, há a tentação para agregar as séries temporais e os dados seccionais. E, aí, a questão relativa à possibilidade de se trabalharem ou não os dados como uma grande *pool* (dados agregados) poderia colocar-se, o que, naturalmente, decorreria dos dados em painel.

A estrutura painel dos dados caracterizada por, blocos ou dados seccionais sequenciais, com uma série cronológica no seio de cada um deles,¹¹² admite um tipo de modelo analítico painel em que os coeficientes são constantes entre os municípios e ao longo do tempo.¹¹³ Estar-se-ia a agregar todos os dados e a ajustar um modelo regressivo ordinário dos mínimos quadrados, *MQO*.¹¹⁴ Em suma, estar-se-ia perante um modelo restrito, o modelo *pooled* (agregado).

¹¹² No caso em análise, obter-se-ia para cada município, nos diversos momentos eleitorais, um agrupamento do tipo, $VV1_{Aveiro,1979}, \dots, VV1_{Aveiro,2001}$; $VOTOSANT_{Aveiro,1979}, \dots, VOTOSANT_{Aveiro,2001}$; $APRES_{Aveiro,1979}, \dots, APRES_{Aveiro,2001}$, $APART_{Aveiro,1979}, APART_{Aveiro,1979}, \dots, APART_{Aveiro,2001}$, $DIREAL_{Aveiro,1979}, \dots, DIREAL_{Aveiro,2001}$, para o município Aveiro, e assim sucessivamente para todos os outros municípios.

¹¹³ Assume-se que, para cada município, as observações não estão correlacionadas em série e para os municípios e o tempo, os erros são homoscedásticos.

¹¹⁴ Importa referir que o estimador dos mínimos quadrados em domínio *pooled* é, tão só, uma média ponderada dos estimadores *dentro* e *entre* unidades seccionais do modelo de efeitos fixos; ou seja de estimadores a partir de um modelo em que o valor observado em cada município é centralizado em relação ao valor médio de todas as observações e de estimadores a partir de um modelo em que o valor observado em cada município é o valor médio para o período de tempo observado, respectivamente. (Conforme virá desenvolvido no Apêndice F). Por essa via os coeficientes estimados por uma regressão *pooled* serão enviesados e inconsistentes, uma vez que combinam estimadores enviesados e estimadores não enviesados – realidade sustentada no modelo de

Entretanto, o modelo não restrito é aquele que traduz a mesma relação, mas com parâmetros diferentes entre municípios e ao longo do período analisado, ou seja um vector de parâmetros diferentes, δ_i , significando coeficientes, α_i , β_{1i} , β_{2i} , β_{3i} , β_{4i} , β_5 , β_6 , diferentes. O mesmo é dizer que, para a estimação do vector de parâmetros, δ_i , diferente para cada um dos 275 municípios, se assumem 275 regressões,¹¹⁵ cada uma delas passível da ilustração que se segue,

$$\underset{[T*1]}{\mathbf{y}_i'} = \underset{[T*K]}{\mathbf{Z}_i} \underset{[K*1]}{\boldsymbol{\delta}_i} + \underset{[T*1]}{\boldsymbol{\varepsilon}_i}, \quad i = 1, 2, \dots, N \quad 4.3.1$$

em que,

$\mathbf{y}_i' = (VV1_{i1}, \dots, VV1_{iT})$, é o vector que ilustra os valores observados da variável dependente, isto é o vector da percentagem de votos obtida pelo incumbente no município i , nos diversos períodos eleitorais t ;

$\mathbf{Z}_i = [t_T, \mathbf{x}_{ki}]$, é uma matriz simbólica constituída pelo vector de valores unitários t_T de dimensão T e pela matriz das k variáveis explicativas \mathbf{x}_{ki} de dimensão $T * K$.

Os modelos analíticos, restrito e não restrito, de dados painel, entretanto estimados, permitem conhecer o valor da soma dos quadrados dos resíduos em cada um deles, $e'e$ e

$\sum_{i=1}^{275} e_i'e_i$, respectivamente; e partir para a utilização de um teste de Chow, para que,

subjacente ao pressuposto que $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 I_{NT})$, se dê corpo ao teste da *poolability*

efeitos fixos, onde se assume que os efeitos específicos das unidades seccionais estão correlacionados com os regressores [Greene (2003) e Hsiao (1986)].

¹¹⁵ No caso em estudo, dado que, as séries cronológicas são curtas, apenas 7 períodos eleitorais; e os valores observados das variáveis exógenas, *APART*, *APRES*, configuram vectores linearmente dependentes, torna-se difícil o ajustamento de regressões ao modelo desenhado para cada município, dada a possibilidade de se estar na presença de matrizes singulares. Razão, mais do que pertinente, para se avaliar o interesse da *poolability* dos dados.

(teste de homogeneidade global) dos dados observados, para se testar a hipótese, $H_0 : \delta_i = \delta$, para todo o i ,^{116,117} ilustrado por,

$$F_{obs} = \frac{(e'e - e_1'e_1 - e_2'e_2 - \dots - e_N'e_N) / (N-1)K'}{(e_1'e_1 + e_2'e_2 + \dots + e_N'e_N) / N(T-K')} \quad 118$$

Neste cenário, ultrapassado o problema de uma eventual matriz singular, poder-se-ia equacionar aplicar o teste de Chow. A ser assim, a soma dos quadrados dos resíduos da forma restrita, é a que resulta da estimação pelo método dos mínimos quadrados ordinários (*MQO*) no modelo *pooled*, vindo $SSR_R = 196792,8$; e a soma dos quadrados dos resíduos da forma não restrita, é a que se obtém adicionando a soma dos quadrados dos resíduos das 275 regressões municipais, pelo método dos mínimos quadrados, cada um com 2 graus de liberdade, vindo $SSR_U = 71595,9$.

Para o caso em estudo tem-se um *F* – estatístico, $F_{(1370,550)} = 0,702$. Há 1370 restrições e o teste aceita a *poolability* entre os municípios para todos os coeficientes estimados, para qualquer nível de significância estatística. Este resultado revela-se interessante na medida em que evita que a agregação dos dados seja legitimada, apenas, pelo facto de haver uma linha de investigação em dados de painel que perspectiva a estimação, para cada variável, de um único coeficiente sobre o “espaço” e sobre o “tempo”, como uma razoável aproximação à realidade.

¹¹⁶ Sob essa hipótese o modelo anterior pode ser escrito na forma restrita, $y = Z\delta + \varepsilon$, em que $Z' = (Z_1', Z_2', \dots, Z_N')$ e $\varepsilon' = (\varepsilon_1', \varepsilon_2', \dots, \varepsilon_N')$.

¹¹⁷ Sob a $H_0 : \delta_i = \delta$, para todo o i , o F_{obs} apresenta-se como uma distribuição $F((N-1)K', N(T-K'))$, em que K' é $K+1$.

¹¹⁸ Importa referir que as variáveis macroeconómicas introduzidas no modelo, *PDESEMP1* e *PINFLAC1*, não foram consideradas, dado o facto de, tendencialmente, observarem valores nulos, o que faria perigar o ajustamento ao modelo não restrito, pelo método dos mínimos quadrados ordinários, *MQO*, para estimar os coeficientes δ_i , ao impossibilitar a inversão do produto matricial $(x'x)$. Além disso colocar-nos-ia perante um número k de variáveis exógenas quase igual ao número de períodos observados, $t=6$ e $t=7$, respectivamente.

As conclusões deste teste poderiam ser reforçadas ou não com um teste de igualdade das 275 constantes individuais α_i , isto é com a identificação de uma dimensão individual para todas as constantes seccionais do modelo. Para tanto havia que testar a $H_0 : \alpha_i = \alpha$, para todo o i , com a imposição de que os coeficientes $\beta_i = \beta$. Aplicando o teste ao modelo em estudo, no quadro da base dados utilizada, a imposição da homogeneidade dos declives do modelo, levar-nos-ia, previamente, a testar a $H_0^2 : \beta_i = \beta$, $\forall i \in \{1, \dots, N\}$, vindo uma estatística de Fisher, $F_{[1096,550]} = 0,9740$, a qual promoveria a aceitação da hipótese nula de que os coeficientes dos declives são homogêneos, apontando para a confirmação da estrutura de painel. Finalmente, o teste relativo às constantes individuais α_i , a realização da estatística de Fisher, $F_{[274,1646]} = 0,8456$, associada a este teste, permite aceitar a hipótese da igualdade dos coeficientes das intercepções do modelo.

Os resultados apurados são no sentido de uma estrutura painel totalmente homogênea, modelo *pooled*. O teste, último, de homogeneidade das constantes α_i , produziu um *F – estatístico* que reforça as conclusões do teste da homogeneidade global, com a vantagem de que testa 274 restrições, contra as 1370 restrições, então, testadas – essa redução do número de restrições lineares permite reforçar o poder do teste de Fisher.

Se decidirmos agregar a população dos 275 municípios, assumindo homogeneidade nos parâmetros da intercepção e do declive, aquando do ajustamento regressivo ao modelo, mas ignorando a dependência seccional (*cross-sectional*) pode acontecer que os ganhos de eficiência, que se espera alcançar por esta via, comparativamente aqueles que se obteriam caso se corresse regressões individuais, *MQO*, diminuam dramaticamente. A dependência seccional pode ter, também, um importante impacto sobre os procedimentos de correcção do enviesamento, subjacentes ao modelo de efeitos fixos, passando essas correcções do enviesamento a ter que, eventualmente, equacionar, ou a média, ou a mediana, da distribuição do estimador dos efeitos fixos.

4.3.2.2. Teste à presença de dependência seccional no modelo agregado

Evidentemente que o impacto da dependência inter-seccional, na estimação dos coeficientes do modelo agregado, vai depender de uma variedade de factores, tais como, a dimensão da correlação entre unidades seccionais e a natureza dessa própria dependência seccional. Se assumirmos que a dependência seccional é gerada pela presença de factores comuns que, não são observáveis (e por consequência, o efeito desses factores vem reflectido no termo do erro); estando esses factores, no entanto, não correlacionados com os regressores contidos no modelo, ocorre que os estimadores gerados pelo modelo de efeitos fixos, *FEM*, e/ou pelo modelo dos efeitos aleatórios, *REM*, são consistentes, embora não eficientes e os erros padrão estimados são enviesados.¹¹⁹

Importa pois testar a dependência seccional dos dados painel em estudo, tendo em conta que a dimensão temporal, *T* (7 períodos eleitorais), é francamente menor que a dimensão seccional, *N* (275 municípios). O teste *LM*, desenvolvido por Breusch e Pagan (1980), não goza, neste cenário, de qualquer das propriedades estatísticas desejáveis, pelo que a sua aplicação produz distorções substanciais de dimensão.¹²⁰ Resta pois recorrer a três testes de dependência seccional, válidos quando $T < N$, concretamente o teste de Pesaran (2004), a estatística de Friedman (1937) e o teste estatístico proposto por Frees (1995).

O pacote estatístico STATA disponibiliza suporte para que se implementem os três testes. A hipótese a ser testada é a hipótese nula de independência seccional, a qual vem consubstanciada via testes de Pesaran, Friedman e Frees, depois de estimado o modelo em estudo, através do método de efeitos fixos com duas componentes de erro. Começa-

¹¹⁹ Os coeficientes estimados β e respectivos erros padrão são consistentes se os erros seguirem uma distribuição idêntica e independente, i.i.d., ou seja se a matriz variância-covariância Ω , for $\Omega = \sigma^2 \mathbf{I}_{NT}$. Entretanto, na presença de correlação inter-seccional, a matriz diagonal Ω perde essa característica.

¹²⁰ Para o efeito observe-se Pesaran (2004) ou Sarafidis, Yamagata e Robertson (2006).

se pelo teste de Pesaran, que se suporta na correlação de pares de resíduos estimados,¹²¹ o qual, sob a hipótese nula de não dependência seccional, se comporta da seguinte forma, $CD \xrightarrow{d} N(0,1)$, isto é segue uma distribuição normal, para $N \rightarrow \infty$ e T suficientemente grande, vindo,

$$\text{Teste de Pesaran de independência seccional} = 5,991, \quad \text{Pr} = 0,000$$

$$\text{Valor médio absoluto de valores fora da diagonal} = 0,392$$

Como se pode ver a hipótese nula de não dependência seccional vem rejeitada, para níveis relevantes de significância estatística, pelo coeficiente de dependência de Pesaran, CD , perspectivando-se uma clara evidência de dependência seccional.

Estes resultados vêm corroborados pelo teste de Frees, que se baseia na soma dos quadrados dos coeficientes de correlação das característica dos resíduos estimados,¹²² os quais desde que retirados de forma independente de uma distribuição do χ^2 , permitem analisar o resultado do teste à luz do comportamento de uma Q distribuição.¹²³ Importa pois, analisar os valores produzidos pelo teste de Frees, assim,

$$\text{Teste de Frees de independência seccional} = 3,051$$

$$\text{Valores críticos a partir da distribuição } Q \text{ de Frees}$$

¹²¹ $CD = \sqrt{2T/N(N-1)} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right)$, em que $\hat{\rho}_{ij} = \hat{\rho}_{ji} = \sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt} / \left(\sum_{t=1}^T \hat{u}_{it}^2 \right)^{1/2} \left(\sum_{t=1}^T \hat{u}_{jt}^2 \right)^{1/2}$ é a correlação entre o valor estimado dos resíduos, \hat{u}_{it} .

¹²² Ou seja em \hat{r}_{ij}^2 , definindo-se $\{r_1, \dots, r_T\}$, $\{r_{j1}, \dots, r_{jT}\}$ como as características dos resíduos, $\{u_{11}, \dots, u_{1T}\}$ e $\{u_{j1}, \dots, u_{jT}\}$, respectivamente. A estatística de Frees vem ilustrada pela expressão, $FRE = N \left(\frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{r}_{ij} - (T-1)^{-1} \right)$.

¹²³ A Q distribuição é uma soma (ponderada) de duas variáveis aleatórias com uma distribuição χ^2 , que depende da dimensão temporal, T .

$alfa = 0,10 : 0,3583$

$alfa = 0,05 : 0,4923$

$alfa = 0,01 : 0,7678$

Efectivamente o teste de Frees permite rejeitar de novo a hipótese nula de independência seccional.

De forma a esgotar a bateria de testes disponíveis para testar a hipótese nula em análise, parte-se para a aplicação do teste de Friedman. O teste de Friedman é um teste não paramétrico baseado na média dos coeficientes de correlação das características dos resíduos de Spearman,¹²⁴ que segue uma distribuição assintótica do χ^2 , com $T - 1$ graus de liberdade, para valores fixos de T à medida que N cresce. Os valores obtidos,

$$\text{Teste de Friedman de independência seccional} = 22,489, \quad \text{Pr} = 1,000,$$

permitem aceitar a hipótese de independência seccional.

O resultado do teste de Friedman introduz um factor de instabilidade relativamente à conclusão de rejeição da hipótese nula de independência seccional decorrente dos testes de Pesaran e Frees. A propósito, a literatura econométrica refere que estes dois últimos testes partilham uma fragilidade comum, na medida em que se suportam nos coeficientes de pares de correlação da matriz dos resíduos, em vez de se fundamentarem nas somas dos quadrados dessas correlações (como acontece nos testes *LM*).¹²⁵

Independentemente da divergência de resultados e de uma identificação concludente de dependência seccional há que ter presente que o impacto da presença de dependência

¹²⁴ A relação entre a estatística de Friedman, FR , e a média da correlação de Spearman, R_{AVE} , vem ilustrada na expressão,

$$FR = [(T - 1)(N - 1)R_{AVE} + 1], \text{ sendo } R_{AVE} = 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{r}_{ij} / N(N - 1).$$

¹²⁵ O facto dos testes de Pesaran e Frees utilizarem, apenas, os coeficientes de pares de correlações implica que esses testes excluam da dependência seccional as correlações de resíduos onde o sinal é simétrico, tornando-os pouco realistas.

seccional na inferência estatística impõe a introdução de procedimentos de correcção de eventuais enviesamentos nos estimadores e a ser assim na secção que se segue apresentar-se-á o modelo suportado no método de efeitos fixos estandardizado, mas com erros padrão corrigidos.

4.3.2.3. Estimação do modelo e resultados

Em síntese, após se ter inventariado o potencial dos dados em painel e se ter concluído do interesse em trabalhá-los como um grande agregado, apurou-se a presença de dependência seccional; o suficiente para ensombrar a inferência estatística do modelo, e exigir técnicas de covariâncias robustas de White; daí, a estimação do modelo proposto adoptar, quer a metodologia implícita a um modelo *pooled*, quer a metodologia patente no modelo de efeitos fixos.^{126,127} A inclusão de variáveis *dummy* temporais no modelo de efeitos fixos, para estimação pelos mínimos quadrados ordinários, *MQO*, só por si, seria suficiente para captar os efeitos de perturbações comuns, porém esta técnica só seria apropriada num cenário de pressupostos muito fortes relativamente à forma das correlações espaciais; isto é, apenas para correlações contemporâneas iguais entre qualquer par de municípios e para valores nulos de correlações inter-seccionais desfasadas. Acontece que estas restrições fortes sobre a forma das correlações espaciais dificilmente se revelam correctas na maior parte das aplicações, já que o mais natural é uma reacção diferente dos diversos municípios a perturbações comuns e, também que essa reacção venha desfasada entre municípios. Nesse caso, a estrutura da correlação

¹²⁶ Para se gerarem os estimadores fixos, a técnica adoptada suporta-se na transformação dentro, *Within*, e, dado os parâmetros estimados dos regressores virem idênticos àqueles que se obteriam se fosse implementado o estimador mínimos quadrados ordinários com variáveis *dummies* (*MQVD*), materializa-se introduzindo variáveis *dummy*: seccionais (\mathbf{D}_μ) para traduzir os efeitos específicos das unidades seccionais e temporais (\mathbf{D}_γ) para equacionar os efeitos específicos temporais. Nesse caso, estar-se-ia face a um modelo de efeitos fixos com componente de erro a duas dimensões (cuja expressão, na forma vectorial, é passível de se traduzir por, $\varepsilon = \mathbf{D}_\mu\boldsymbol{\mu} + \mathbf{D}_\gamma\boldsymbol{\gamma} + \boldsymbol{\eta}$)

¹²⁷ A componente erro, decorrente dos efeitos fixos, para cada uma das estimações pode ser equacionada a uma dimensão, μ_i ou γ_t , e a duas dimensões μ_i e γ_t . Neste caso, dado que se verifica um $N > T$, começa-se por equacionar os efeitos fixos periódicos e observa-se o valor do critério Schwarz. Como esse valor é inferior ao da estimação *pooled* ($7,314461 < 7,394256$), mantém-se a especificação do modelo de efeitos fixos temporais. Posteriormente, introduzem-se os efeitos fixos seccionais.

espacial, implicitamente modelada pelas variáveis *dummy*, temporais e seccionais, poderia estar mal especificada¹²⁸ e, aí, o estimador resultante, apresentaria propriedades desconhecidas.

Uma vez que, na presença de correlação espacial, as restrições fortes subjacentes à especificação do modelo de efeitos específicos, temporais e seccionais, se podem revelar indesejáveis, a alternativa para que o problema da heteroscedasticidade painel e da correlação contemporânea venha equacionado pode estar na estimação pelos mínimos quadrados generalizados fiáveis, MQG_F .

Uma técnica padrão, já referida em 4.3.1., passaria por estimar uma especificação do tipo mínimos quadrados generalizados (MQG), capaz de equacionar uma correlação seccional arbitrária (contemporânea) e heteroscedasticidade periódica entre resíduos para uma dada unidade seccional. A estrutura desse tipo de covariância implicaria assumir que,

$$E(\varepsilon_{it}\varepsilon_{jt} \mid \mathbf{x}_t^*) = \sigma_{ij},$$

$$E(\varepsilon_{is}\varepsilon_{jt} \mid \mathbf{x}_t^*) = 0$$

para todo o i, j, s e t , com $s \neq t$, o mesmo que admitir correlação contemporânea entre resíduos, para i e para j , mas restringindo essa variância contemporânea a não variar ao longo de t . Efectivamente, este pressuposto, relativamente às covariâncias através das unidades seccionais, num dado momento do tempo, envolveria uma abordagem semelhante à que decorre da utilização de uma técnica de regressões aparentemente não relacionadas (SUR). Acontece que, sendo a dimensão temporal, T

¹²⁸ Problemas potenciais, resultantes donexo modelo-teoria, podem ocorrer e prendem-se com uma injustificada absorção da variância seccional e da variância temporal e uma injustificada assumpção de invariância temporal para as relações causais.

¹²⁹ Utilizando o vector dos resíduos específico do período temporal, pode-se reescrever este pressuposto como, $E(\varepsilon_t \varepsilon_t' \mid \mathbf{x}_t^*) = \Omega_N$, para todo o t .

(7), inferior à dimensão espacial, N (275), esta perspectiva revela-se inviável, não permitindo a estimação dos coeficientes do modelo,¹³⁰ já que a matriz bloco das covariâncias estimadas dos erros não é passível de ser invertida, dado a característica da matriz bloco das covariâncias contemporâneas vir potencialmente superior à ordem da matriz diagonal por blocos dos resíduos.¹³¹

Um processo para ultrapassar esta restrição, que decorre da dimensão dos dados, em estudo, é expandir o procedimento utilizado em modelos estimados pelo método dos mínimos quadrados ordinário, *MQO*, de introdução do cálculo das covariâncias robustas de White, aos modelos estimados pelo método dos mínimos quadrados generalizado possível, *MQGF*. Nesta perspectiva, um método de aplicação possível, é o White periódico, em que a regressão agregada é tratada como uma regressão multivariada (como se tivéssemos uma equação para cada período temporal) e os erros padrão robustos, tipo White, vêm calculados para esse sistema de equações. Os coeficientes estimados da matriz de covariância são robustos à correlação entre períodos temporais (em série), bem como às diferentes variâncias do erro em cada período temporal. Podem-se, no entanto, obter coeficientes estimados de covariância, apenas, robustos à heteroscedasticidade, utilizando um método mais geral e mais restritivo do que o anterior. Mais geral porque permite diferentes variâncias para observações da mesma unidade seccional e mais restritivo por serem nulas as variâncias fora da diagonal da matriz de covariâncias-variâncias.

Uma outra técnica, é o *PCSE* (*Panel Correct Standard Error*), uma variante da estatística White, os resíduos são substituídos por estimadores momentos para variâncias não condicionais – a estimação da variância do erro tem subjacente a média

¹³⁰ $\hat{\beta}_{MQGF} = (\mathbf{x}'\hat{\Omega}^{-1}\mathbf{x})^{-1}\mathbf{x}'\hat{\Omega}^{-1}\mathbf{y}$.

¹³¹ A matriz estimada de variância-covariância, $\hat{\Omega}$ é uma matriz diagonal por blocos, de dimensão $NT * NT$, com uma matriz $\hat{\Sigma}$, de dimensão $N * N$, nos blocos da diagonal, constituída pelos elementos, $\hat{\sigma} = \mathbf{e}_i'\mathbf{e}_j/T$, em que \mathbf{e}_i e \mathbf{e}_j representam os vectores, $N * 1$, dos resíduos *MQO*, da unidade i e j , respectivamente. Significa pois, que inverter a matriz $\hat{\Omega}$ depende de $\hat{\Sigma}$ ser invertível, exigindo que a característica de $\hat{\Sigma} = N \leq T$, o que não se verifica.

dos termos de erro padrão de cada observação.¹³² É robusto às variâncias não condicionais não restritas, mas coloca restrições adicionais sobre as matrizes de variância condicional.

As estimativas geradas pelos modelos acima referidos vão ser obtidas utilizando o pacote estatístico EVIEWS 5, o qual permite uma especificação $MQGF$, capaz de equacionar quatro padrões de estrutura de correlação entre resíduos, a saber, heteroscedasticidade específica seccional;¹³³ heteroscedasticidade específica temporal;¹³⁴ covariância contemporânea (*Cross Section SUR*)¹³⁵ e covariância entre períodos temporais (*Period SUR*).¹³⁶

No caso em estudo, a estrutura de covariância contemporânea não pode ser equacionada, já que a técnica $MQGF$, em ambiente *Cross Section SUR*, não pode ser utilizada, pelo facto da condição necessária, *de que a característica da matriz das covariâncias contemporâneas estimadas, $\hat{\Sigma}$,¹³⁷ de dimensão, $N * N$, é igual a N , e inferior a T* , não estar satisfeita; o que decorre da dimensão espacial, $N = 275$, e da

¹³² Para se ter uma ideia do que isto significa, para cada unidade, a variância do termo erro viria estimada como uma média dos respectivos resíduos, ter-se-ia, por exemplo, $\hat{\sigma}_1^2 = \{ \hat{e}_{11}^2 + \hat{e}_{12}^2 + \dots + \hat{e}_{1T}^2 \} / T$.

¹³³ As variâncias são estimadas a partir dos resíduos específicos das unidades seccionais, após estimação preliminar do modelo; resíduos esses, que permitem escrever o pressuposto chave da heteroscedasticidade seccional, como, $E(\varepsilon_i \varepsilon_i' | \mathbf{x}_i^*) = \sigma_i^2 \mathbf{I}_T$, onde \mathbf{x}_i^* contém todas as k variáveis exógenas $x_{ki,t}$ e os efeitos específicos, seccionais e temporais (α_i, λ_i) .

¹³⁴ Em termos análogos, ao que acontece na nota anterior, estas variâncias são estimadas a partir de resíduos específicos do período temporal, após estimação preliminar do modelo, e esses resíduos permitem que o pressuposto base, de que há uma variância de resíduos diferente, para cada período, venha expresso por, $E(\varepsilon_t \varepsilon_t' | \mathbf{x}_t^*) = \sigma_t^2 \mathbf{I}_N$.

¹³⁵ Esta denominação, *Seemingly Unrelated Regressions* seccional, explica-se por esta especificação envolver covariâncias entre unidades seccionais como se estivesse a trabalhar com regressões aparentemente não relacionadas, em que cada uma corresponderia a uma unidade seccional.

¹³⁶ O termo *Seemingly Unrelated Regressions* temporal aplica-se pelo facto da especificação envolver covariâncias entre períodos dentro de uma unidade seccional, como se estivesse a trabalhar com regressões específicas de cada período.

¹³⁷ Cada um dos elementos estimados, da matriz de covariâncias contemporâneas, $\hat{\Sigma}$, é calculado a partir dos vectores \mathbf{e}_i e \mathbf{e}_j , dos resíduos MQO das unidades seccionais i e j , respectivamente, de dimensão $T * 1$, analiticamente, traduzido por, $\sigma_{ij} = \mathbf{e}_i' \mathbf{e}_j / T$. Em termos matriciais, $\hat{\Sigma} = \mathbf{E}' \mathbf{E} / T$, desde que \mathbf{E} seja uma matriz $\mathbf{E} = (e_1, \dots, e_n)$, de dimensão $T * N$ e característica $(\hat{\Sigma}) = \text{característica}(\mathbf{E}) = N \leq T$.

dimensão temporal, $T = 7$ dos dados painel. Confrontar-nos-íamos com uma matriz de correlação estimada de resíduos não invertível. Esta impossibilidade torna evidente que a especificação $MQGF$ não vai equacionar a correcção por heteroscedasticidade seccional e por correlação contemporânea, resta pois equacionar a correcção por heteroscedasticidade periódica e por correlação geral dos resíduos dentro de uma unidade seccional, através da técnica $MQGF$, em ambiente *Period SUR*. Estão reunidas circunstâncias para que a condição necessária venha satisfeita, já que agora a matriz de covariâncias entre períodos temporais, $\hat{\Sigma}$, teria uma dimensão $T * T$, o que consentiria uma característica T inferior a N . Mesmo assim, somos confrontados com uma matriz de covariâncias dos erros estimados não invertível, pelo facto de estarmos perante uma matriz singular, o que pode decorrer de se ter uma matriz de 275 blocos de matrizes de $7 * 7$ e, também, do comportamento das variáveis, observadas em cada um das unidades seccionais, ser, tendencialmente, um potencial para vectores linearmente dependentes.

Assim sendo, através da especificação $MQGF$, conseguir-se-ia equacionar a heteroscedasticidade seccional e/ou a heteroscedasticidade entre períodos; mas não a correcção por heteroscedasticidade seccional e por correlação contemporânea nem a correcção por heteroscedasticidade periódica e por correlação geral dos resíduos dentro de uma unidade seccional. Naturalmente, que neste cenário, houve a preocupação de que as covariâncias estimadas¹³⁸ fossem robustas à correlação entre resíduos para as diferentes observações, especificando-se, para o efeito, uma matriz covariância consistente e heteroscedástica de White (*HCCM*). As matrizes de covariâncias robustas

¹³⁸ Para se conseguirem inferências válidas é necessário implementar, $\text{var } \beta = \sigma^2 (\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\mathbf{\Omega}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}$, com $\sigma^2 \mathbf{\Omega} = \text{diag} \{ \sigma_1^2, \sigma_2^2, \dots, \sigma_n^2 \}$. White consegue uma estimativa satisfatória ilustrada pela relação $\mathbf{x}'\sigma^2 \mathbf{\Omega}\mathbf{x} = \sum_{i=1}^N \sigma_i^2 \mathbf{x}_i \mathbf{x}_i'$. O estimador White substitui o parâmetro desconhecido, σ^2 ($t = 1, 2, \dots, n$), por \hat{e}_t^2 , onde \hat{e}_t representa os resíduos *MQO*, dados por, $\mathbf{y}_t - \mathbf{x}_t' \boldsymbol{\beta}$. Tem-se assim um estimador consistente da matriz de variância para o vector coeficiente e é particularmente útil porque não requer nenhum pressuposto específico sobre a forma de heteroscedasticidade, vindo $\text{est. var } \beta = \sigma^2 (\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\hat{\mathbf{\Omega}}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}$ e $\sigma^2 \hat{\mathbf{\Omega}} = \text{diag} \{ \hat{e}_1^2, \hat{e}_2^2, \dots, \hat{e}_n^2 \}$.

foram calculadas por diversos métodos de forma a atender, quer a correlação em série e as variâncias do erro em cada período de tempo (*White period method* e *PCSE*),¹³⁹ quer, apenas, a heteroscedasticidade (*White diagonal method*).¹⁴⁰

No que concerne aos resultados relativos à estimação do modelo em estudo, apresentam-se os que decorrem dos métodos:

(1) *MQO*, os potenciais comportamentos não esféricos dos termos erros não são equacionados mas os efeitos específicos a duas dimensões são-no, a eventualidade de uma distribuição de erros não-esférica vem corrigida via métodos de coeficiente covariância, *PCSE* /regressão (2), *White period method* / regressão (3) e *White diagonal method* / regressão (4);

(2) *MQGF*, a heteroscedasticidade temporal é equacionada, tendo subjacente as variâncias estimadas específicas do período temporal, os efeitos específicos são considerados a duas dimensões e continua-se a introduzir métodos de coeficiente covariância, *PCSE* /regressão (5), *White period method* / regressão (6) e *White diagonal method* / regressão (7), para corrigir a eventualidade de uma distribuição de erros não-esférica.

Conforme ilustrado na tabela que se segue,

¹³⁹ A diferença entre o método *White period* e o *PCSE* reside no cálculo da variância do erro padrão robusto. Para o *White period* tem-se $\sigma^2 \hat{\Omega} = \text{diag} \{ \hat{e}_1^2, \hat{e}_2^2, \dots, \hat{e}_n^2 \}$, enquanto para o *PCSE* vem, $\sigma^2 \Omega = \text{diag} \{ \hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_n^2 \}$. As estimativas na diagonal de *PCSE* são mais precisas, na medida em que resultam das médias dos termos de erro, por exemplo para $\hat{\sigma}_1^2$ tem-se $\hat{\sigma}_1^2 = \{ \hat{e}_{11}^2 + \hat{e}_{12}^2 + \dots + \hat{e}_{1T}^2 \} / T$.

¹⁴⁰ Computacionalmente qualquer estimador painel que pretenda equacionar comportamentos não esféricos no termo erro realiza cedências, *tradeoffs*, entre os vários tipos de comportamentos, por exemplo se assume as formas mais gerais de heteroscedasticidade e de correlação em série, “compensa” assumindo ausência de correlação seccional. Ou seja o *EViews 5* não comporta estimadores *White* (robustos) para a matriz *var-cov* do erro, Ω , que incorporem quer a correlação em série, quer a correlação seccional, presentes na matriz de ponderação W , mas permite que esses estimadores sejam utilizados para ajustar os coeficientes dos erros padrão (i.e. não podem ser utilizados para calcular $\hat{\beta}$, mas são-no para estimar $\text{Cov}(\hat{\beta})$). É neste contexto que a estimação painel, em ambiente *EViews 5*, introduz uma caixa de diálogo que permite essas compensações inter-supostos não esféricos, ilustrada por, uma janela para a opção “*MQG Weights*” e outra janela para a opção “*Coef covariance method*”.

	Pooled (1)	$MQO^{(a)(c)}$ PCSE (2)	$MQO^{(a)(c)(e)}$ White(3)	$MQO^{(a)(d)}$ White (4)	$MQG_F^{(a)(b)(c)}$ PCSE (5)	$MQG_F^{(a)(b)(e)}$ White (6)	$MQG_F^{(a)(b)(d)}$ White(7)
VOTOSANT	0,79134*** (33,6440)	0,7275*** (23,830)	0,7275*** (17,902)	0,7275*** (19,1525)	0,7750*** (26,8933)	0,7750*** (20,0489)	0,7750*** (21,4972)
APART	-0,17685*** (-2,68256)	-0,06076 (-0,8576)	-0,06076 (-0,89561)	-0,06076 (-0,91476)	-0,05605 (-0,82226)	-0,05605 (-0,861073)	-0,05605 (-0,89762)
APRES	0,26558*** (5,36658)	0,37612*** (6,08535)	0,37612*** (5,21427)	0,37612*** (5,67599)	0,33547*** (5,70445)	0,33547*** (4,76811)	0,33547*** (5,27762)
DIREAL	-0,00110 (-0,66119)	0,00613*** (2,50007)	0,00613*** (2,62542)	0,00613*** (2,77203)	0,00461** (2,00949)	0,00461*** (2,27556)	0,00461*** (3,58336)
PDESEMP1	-0,43335*** (-4,30041)	0,10069 (0,80738)	0,10069 (0,75134)	0,10069 (0,80663)	0,13164 (1,09642)	0,13164 (1,03058)	0,13164 (1,09375)
PINFLAC1	0,12492*** (2,77660)	-0,13104** (-2,04591)	-0,13104** (-1,94387)	-0,13104** (-2,04591)	-0,17355*** (-2,87585)	-0,17355*** (-2,80016)	-0,17355*** (-2,95771)
R^2 Ajust.	0,40664	0,44435	0,44435	0,44435	0,60619	0,60619	0,60619
N°Obs.	1925	1925	1925	1925	1925	1925	1925
N*T	275*7	275*7	275*7	275*7	275*7	275*7	275*7

Tabela 4.3.2-1: Resultados de regressões em painel, modelo efeitos fixos

Notas: Os valores dentro de parênteses dizem respeito aos valores do *t* - estatístico e, sempre que a notação *** está presente, revelam-se parâmetros estimados significantes a um nível de significância de 1%; já a presença do símbolo ** nos valores dos coeficientes estimados remete-nos para valores estimados com um nível de significância de 5%; finalmente a não presença de qualquer asterisco informa que os valores estimados não são significantes, isto é a probabilidade dos valores dos coeficientes estimados virem iguais a zero é elevada.

^(a) significa que a componente do erro foi considerada a duas dimensões, a componente específica seccional e a componente específica temporal; ^(b) estimação pelo método fiável dos mínimos quadrados generalizado, com ponderação temporal; ^(c) estimação com correcção painel dos erros padrão PCSE; ^(d) estimação com correcção *White diagonal*; ^(e) estimação com correcção *White period*.

Relativamente aos coeficientes estimados, convém referir que,

(i) o coeficiente estimado, β_1 , para a variável *VOTOSANT*,¹⁴¹ apresenta-se com o sinal esperado (+) e é estatisticamente significativo;

(ii) o coeficiente estimado, β_2 , para a variável *APART*,¹⁴² em domínio painel, apresenta-se com sinal (-), em linha com o esperado desgaste da governação, no entanto o facto de não ser estatisticamente significativo, admite que não lhe seja imputada valorização determinante na formulação do modelo;

(iii) a persistência da significância estatística e do sinal esperado para o coeficiente estimado, β_3 , da variável *APRES*, é uma boa predição, já que ao atribuímos-lhe capacidade de interagir com a variável *APART*, permitindo a correcção da avaliação do desempenho, podemos inferir que a convicção dos votantes acerca do desempenho do incumbente influi no resultado eleitoral;

(iv) o coeficiente estimado, β_4 , para a variável *DIREAL*, em domínio *pool*, apresenta sinal (-), contrário ao esperado, mas sem significância estatística. Este desempenho é invertido em domínio painel, o coeficiente passa a apresentar o sinal esperado (+) e é estatisticamente significativo;

(v) o sinal do coeficiente previsto, β_5 , para a variável macroeconómica desemprego em caso de coincidência partidária inter-poderes (local e central), *PDESEMP1*, revela-se contrário ao esperado,¹⁴³ mas a ausência de significância estatística, aconselha a uma potencial omissão da mesma no modelo,

e

¹⁴¹ No que concerne a *VOTOSANT*, importa referir que, caso a regressão viesse ajustada a um modelo restrito apenas a esta variável, de entre um conjunto possível de variáveis exógenas, os resultados daí decorrentes poderiam suscitar dúvidas relativamente à especificação do modelo, mas um teste de Wald aos coeficientes do modelo não restrito, ao permitir rejeitar a hipótese nula de que *os coeficientes das restantes variáveis são nulos*, introduz confiança à especificação do modelo proposto.

¹⁴² Este resultado quando cruzado com os potenciais problemas de multicolinearidade referidos na nota ¹¹⁰, sugerem ensaiar a estimação do modelo, apenas com uma das duas variáveis, ou *APRES*, ou *APART*. Caso a regressão seja ajustada sobre o modelo em estudo, sem a variável *APRES*, verifica-se que os sinais e os valores dos coeficientes estimados, bem como as respectivas significâncias estatísticas não sofrem alterações assinaláveis. O mesmo exercício, mas agora, negligenciando a variável, *APART*, regista comportamentos mais ou menos semelhantes, à excepção do sinal e do coeficiente esperado para a variável *APRES*, que passa a ter o sinal esperado e a ser estatisticamente significativa ao nível de 95%.

¹⁴³ Sobre este resultado importa referir que Kirchgassner (1985), Kramer (1971), Chappel (1990) e Suzuki (1992) encontraram evidência empírica de que a inflação é avaliada mais intensamente do que o desemprego pelos votantes.

(vi) o coeficiente estimado, β_6 , para a variável macroeconómica inflação em caso de coincidência partidária inter-poderes (local e central), *PINFLAC1*, comporta-se como o esperado, com sinal (–) e com significância estatística.

4.3.2.4. Apuramento da melhor especificação

Como foi referido nas secções anteriores, os parâmetros estimados, segundo o método *MQO*, sobre o modelo *pooled*, na presença de efeitos fixos, revelam-se enviesados e inconsistentes, uma vez que são omissas variáveis *dummy*, que são relevantes para explicar o modelo. Daí que se tenham equacionado os efeitos fixos, passando o modelo empírico a configurar essa especificação, conforme resultados expressos nas colunas (2), (3), (4), (5), (6) e (7). A estimação dos parâmetros constantes das colunas (2), (3) e (4) fez-se recorrendo ao método dos mínimos quadrados ordinários, *MQO*; enquanto na estimação dos parâmetros constantes das colunas (5), (6) e (7) se recorre ao método dos mínimos quadrados generalizado fiável, *MQGF*.

Seja qual for o método aplicado, a estrutura da nossa base de dados sugere que se inicie o processo de estimação por uma especificação de estimadores fixos temporais, $FE(T)$; a qual, pelo facto de se centrar sobre a dimensão seccional, faz com que as diferenças na percentagem de votos obtidos no acto eleitoral, em cada município, sejam captadas sem interferência da tendência geral,¹⁴⁴ apontando para efeitos mais pronunciados. Os valores do critério Schwarz permitem validar a opção pela especificação dos efeitos fixos. Entretanto, para que se possam remover todas as variações inter-seccionais, introduzem-se os efeitos específicos seccionais, $FE(N)$. Finalmente, o modelo empírico é estimado, assumindo a especificação que combina efeitos específicos, seccionais e temporais, $FE(NT)$, isto é os coeficientes estimados têm ambos os efeitos subjacentes, em termos de uma média ponderada das

¹⁴⁴ Tendência que, na especificação *pooled*, majora o valor das variâncias dos coeficientes estimados.

especificações $FE(T)$ e $FE(N)$, em que os pesos dependem da variação relativa em cada uma das dimensões [Baltagi (2005)].

Tendo em conta a metodologia perseguida importa testar a significância das variáveis *dummy* introduzidas para captar, quer os efeitos específicos dos períodos temporais, quer os efeitos específicos conjuntos, os temporais e os municipais.

Através do comportamento de um teste de Chow, pode-se começar por testar a hipótese nula de homogeneidade entre períodos eleitorais, $H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_{N-1} = 0$. Para o efeito recorre-se às somas dos quadrados dos resíduos da forma restrita (SSR_R) e da forma não restrita (SSR_U), as quais decorrem das estimações, pelo método dos mínimos quadrados (*MQO*) no modelo *pooled* e pelo método dos mínimos quadrados das variáveis *dummy*, *MQVD*, para efeitos específicos temporais, respectivamente, vindo,

$$F = \frac{(SSR_R - SSR_U) / (T - 1)}{SSR_U / (NT - T - K)} \sim F_{[T-1, T(N-1)-K]}$$

O valor produzido pelas estimações sobre o modelo em estudo, $F[6, 1912] = 34,7023$, permite rejeitar a hipótese nula para qualquer nível de significância, o mesmo que rejeitar homogeneidade inter-temporal.

Como já foi referido, no caso em estudo, tendo em conta as características da base de dados, optou-se por estimar os modelos de regressão de dados em painel, como modelos de efeitos fixos com componentes de erro a duas dimensões,¹⁴⁵ conforme nos dá conta a notação ^(a) da tabela **4.3.2-1** dos resultados.

¹⁴⁵ Wallace e Hussain (1969), Nerlove (1971) e Amemya (1971), entre outros, consideraram os modelos de regressão de dados em painel, com duas componentes perturbadoras, uma que decorre dos efeitos seccionais, μ_i , e outra que decorre dos efeitos temporais, γ_t , traduzindo-se essa componente do erro na seguinte expressão: $\varepsilon_{it} = \mu_i + \gamma_t + \eta_{it}$.

Para essas estimações, onde as componentes do erro μ_i e γ_t foram assumidas como parâmetros fixos a ser estimados e as remanescentes perturbações estocásticas foram assumidas como distribuídas idêntica e independentemente, isto é $\eta_{it} \approx \text{i.i.d. } (0, \sigma_\eta^2)$, vem uma matriz das variáveis *dummy*, \mathbf{D} , de dimensão $NT * T$, composta por *dummies* seccionais e por *dummies* temporais. Assim a hipótese nula, H_0 , a testar, assume, agora, quer a homogeneidade de efeitos entre unidades/seccionais, quer a homogeneidade de efeitos temporais, isto é, $H_0: \mu_1 = \mu_2 = \dots = \mu_{N-1} = 0$ e $\gamma_1 = \gamma_2 = \dots = \gamma_{T-1} = 0$

Novamente, recorrendo a um F -teste, procurar-se-á conhecer a significância da aceitação ou da não-aceitação da hipótese nula. Neste caso, com componente de erro a duas dimensões, a soma dos quadrados dos resíduos da forma restrita (SSR_R) continua a ser a que resulta da estimação pelo método dos mínimos quadrados (MQO) no modelo *pooled* e a soma dos quadrados dos resíduos da forma não restrita é, agora, a que resulta da estimação do modelo de efeitos fixos, para as unidades seccionais e para as temporais.

Acontece que a especificação do modelo de efeitos fixos com componente de erro a duas dimensões se suporta em estimações, onde se recorreu, quer ao método dos mínimos quadrados (MQO), produzindo os valores dos coeficientes estimados da coluna (2) (3) e (4) da tabela de resultados, quer ao método de mínimos quadrados generalizado fiável (MQG_F), com ponderação temporal,¹⁴⁶ ilustrado nas colunas (5), (6) e (7). Havia pois que expandir o cálculo do F -teste, para que viesse equacionado o valor da soma dos quadrados dos resíduos da forma não restrita (SSR_U) que decorre, quer da aplicação do método dos mínimos quadrados (MQO), quer do método dos mínimos quadrados generalizado possível (MQG_F).

¹⁴⁶ Como já foi referido a especificação MQG_F é mais robusta aos problemas de heteroscedasticidade e de correlação em série, e introduzi-la nas estimações realizadas, com a preocupação de rigor subjacente, implicou equacionar a dimensão da base de dados, decorrendo daí que estimar os parâmetros temporais seria mais preciso do que estimar parâmetros seccionais, tudo o resto constante.

Perseguindo esse objectivo, ter-se-ia, para a especificação do modelo de efeitos fixos com componente de erro a duas dimensões, que se suporta no método dos mínimos quadrados (*MQO*) sem qualquer tipo de ponderação, ilustrada nas regressões (2) (3) e (4) da tabela de resultados apresentada, um *F – teste*, identificado como F_1 e assim ilustrado,

$$F_1 = \frac{(SSR_R - SSR_U) / (N + T - 2)}{SSR_U / ((N - 1)(T - 1) - K)} = \frac{(178360,1 - 142809,1) / 280}{142809,1 / (274 * 6 - 6)} =$$

$$= 1,45630$$

$$F_1[N + T - 2; (N - 1)(T - 1) - K] \equiv F[280; 1638] = 1,45630 .$$

Acontece, face ao valor obtido, estar-se perante a rejeição da hipótese nula para todos os níveis de significância estatística, indicando a presença de heterogeneidade entre unidades seccionais e ao longo do período temporal.

No mesmo sentido, haveria que equacionar o valor da soma dos quadrados dos resíduos da forma não restrita (SSR_U), decorrente do método de mínimos quadrados generalizado fiável (MQG_F), com ponderação temporal, aplicado à especificação do modelo de efeitos fixos com componente de erro a duas dimensões, no cálculo do *F – teste*. Viria, para as regressões (5), (6) e (7), um mesmo F_2 ,

$$F_2 = \frac{(SSR_R - SSR_U) / (N + T - 2)}{SSR_U / ((N - 1)(T - 1) - K)} = \frac{(178360,1 - 143151,2) / 280}{143151,2 / (274 * 6 - 6)} =$$

$$= 1,4388$$

$$F_2[N + T - 2; (N - 1)(T - 1) - K] \equiv F[280; 1638] = 1,4388.$$

A leitura dos valores obtidos para F_2 permitem coincidir com as conclusões consentidas pela análise do valor de F_1 relativamente à rejeição da hipótese nula de homogeneidade espacial e temporal dos dados painel, para qualquer nível de significância.

Em síntese, os testes de significância conjunta das variáveis *dummy* para os efeitos seccionais e para os efeitos temporais fixos, sustentados nos resultados F_1 e F_2 , decorrentes de um *F-teste*, utilizado para testar a hipótese $H_0 : \mu_i = 0$ assumindo $\gamma_t = 0$, permitem rejeitar a hipótese nula para quaisquer níveis de significância estatística.

Tal como foi referido na secção anterior, o vector β , com heteroscedasticidade pode ser estimado por *MQO*, desde que se calcule uma matriz de covariâncias de White, daí que se tenha implementado as correcções dos erros padrão por *period White*, *diagonal White* e por estimação painel *PCSE*. Importa pois, testar a especificação da matriz de covariância-variância dos erros.¹⁴⁷ O que se concretiza através de um teste do rácio de verosimilhança, *likelihood ratio test*, o qual pretende medir em que extensão as

¹⁴⁷ A matriz de covariância dos erros, em contexto de estimação *PCSE* e *period White*, vem identificada como uma matriz não restrita, já que todos os parâmetros da diagonal e de fora da diagonal são valores a estimar; já, em contexto *diagonal White*, vem identificada como uma matriz restrita (21 restrições) apenas se estimam os valores da diagonal.

restrições impostas à matriz de variância-covariância dos erros reduz a estatística *log-likelihood*.

Da aplicação do teste aos resultados apurados neste estudo, obtém-se,

$$l = \frac{\hat{L}_R}{\hat{L}_{UR}} = \frac{-6876,543}{-6876,543} = 1^{148}$$

O resultado mostra que as duas formas, a restrita e a não restrita, não são diferentes. A evidência do valor unitário afasta o ruído relacionado com a distribuição do rácio de verosimilhança, l . Mesmo assim, sabendo que a distribuição de $-2l$ é uma distribuição assintótica do *Qui-quadrado*, $-2l \sim \chi_R^2$, e dada a dimensão relativamente grande da base de dados, vale a pena verificar se o valor realizado de $-2l$ é inferior/superior ao valor crítico do χ_{21}^2 . Vindo um valor inferior ao valor crítico do χ_{21}^2 , para qualquer nível de significância, aceita-se a hipótese nula de que os modelos não são diferentes.

Expandindo estas conclusões aos resultados obtidos a partir da estimação MQG_F , poder-se-á inferir que a imposição ou não imposição de restrições à matriz de covariâncias não é relevante para uma tomada de decisão a favor de uma qualquer técnica de erros padrão utilizada.

Aqui chegados reunir-se-iam condições para que a especificação do modelo a estimar fosse a dos efeitos fixos. Importa, no entanto, ponderar algumas das lacunas apresentadas pelo modelo de efeitos fixos, FEM , assim,

- (1) Os graus de liberdade podem vir reduzidos de forma drástica com a utilização dos efeitos fixos num cenário em que a dimensão temporal das observações é pequena

¹⁴⁸ Em que \hat{L}_R é a estatística *log-likelihood* da forma restrita e \hat{L}_{UR} é a estatística *log-likelihood* da forma não restrita. O teste *ratio likelihood*, l , virá, $0 \leq l \leq 1$.

(característica presente no painel de dados em estudo, 7 períodos, incluindo ano de eleição e ano pré-eleição);

(2) Os problemas de multicolinearidade podem ser potenciados, já que a inclusão de demasiadas variáveis neste modelo favorece o problema e isso pode conduzir para uma estimação imprecisa dos coeficientes.

(3) A influência dos indicadores invariantes ao longo do tempo, como por exemplo a área dos municípios, pode ser “depurada” pelo equacionamento dos efeitos específicos individuais fixos. Pode-se não ser capaz de identificar o impacto destes factores invariantes no tempo e de um ponto de vista político pode representar grandes restrições.

(4) O enviesamento para baixo vem exacerbado se a variável explicativa está altamente correlacionada ao longo do período temporal, no contexto dos estimadores fixos, pois que ao tentar resolver o problema das variáveis omitidas, removendo alguma da “variância” que “contamina” o estimador *MQO pooled*, a estimativa *MQO* seccional torna-se assintoticamente enviesada em direcção a zero, caso uma variável explicativa venha medida com erro (e isto é uma realidade na maior parte das aplicações de dados em painel).¹⁴⁹

(5) A estrutura do erro assumida em $M_2 : y_{it} = \alpha_i + \beta_k' x_{kit} + \eta_{it}$ pode ser inapropriada. Pode haver erros heteroscedásticos entre unidades seccionais, pode haver erros auto-correlacionados dentro de uma unidade seccional ou grupos através do tempo, pode haver correlações espaciais entre unidades seccionais, quer dentro de todo o período temporal, quer num período específico. Todos estes resultados são possíveis e a perspectiva adoptada no modelo de efeitos fixos ignora-os.¹⁵⁰

(6) As variáveis dependentes desfasadas, utilizadas nos modelos dinâmicos, colocam problemas de correlação entre elas e o termo erro, nas estimações pelo método dos efeitos fixos.¹⁵¹

¹⁴⁹ Entretanto as correlações entre x_t e x_{t-k} tendem a enfraquecer à medida que k aumenta (ou seja à medida que a ordem do desfasamento aumenta), vindo que, quanto maior o desfasamento utilizado mais atenuado o enviesamento devido ao erro de medida. O que poderia sugerir a utilidade de se introduzir modelos de efeitos fixos com diferenciação, ou com variáveis instrumentais.

¹⁵⁰ É possível corrigir a matriz de variância-covariância para a heteroscedasticidade utilizando os estimadores sandwich de White (1980) ou de Huber (1967). Também é possível equacionar um processo autoregressivo de diferença unitária *AR(1)* para estimação do erro.

¹⁵¹ O modelo dinâmico, expresso por, $y_{it} = \alpha_i + \gamma y_{i,t-1} + \beta x_{it} + u_{it}$, em que os efeitos fixos estão ilustrados em, α_i . Ao ser sujeito a uma transformação “dentro” virá sobre a forma, $y_{it} - \bar{y}_i = \gamma (y_{i,t-1} - \bar{y}_{i-1}) + \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i)$. Neste caso \bar{y}_{i-1} está

A metodologia perseguida para equacionar as lacunas, acima referidas, na decisão de escolher a especificação do modelo de efeitos fixos, para o desenvolvimento do nosso modelo empírico, exigiu a aplicação da especificação do modelo de efeitos aleatórios. Para o efeito, apresentar-se-á na secção 4.4. o estimador de efeitos aleatórios, com um breve enquadramento teórico seguido de aplicação concreta ao caso em estudo.

4.4. Estimadores de Efeitos Aleatórios

4.4.1. Abordagem teórica da estimação

A diferença chave entre modelos de efeitos fixos e o modelo de efeitos aleatórios reside nos pressupostos feitos acerca dos efeitos específicos, o das unidades seccionais invariante no tempo e o dos períodos cronológicos dentro de uma dada unidade seccional. Enquanto a perspectiva dos efeitos fixos os trata como constantes, a perspectiva dos efeitos aleatórios assume que se comportam de forma aleatória, de acordo com um qualquer processo especificado.

O modelo de efeitos fixos assume que os efeitos específicos individuais e temporais estão correlacionados com os regressores, $E[\eta_i, x_{it}] \neq 0$ e $E[\lambda_t, x_{it}] \neq 0$, respectivamente; enquanto o modelo de efeitos aleatórios os assume como independentes dos regressores, $E[\eta_i, x_{it}] = E[\lambda_t, x_{it}] = 0$.¹⁵² Essencialmente, transforma os efeitos específicos individuais em termos erro, a serem acrescentados aos termos de erro convencionais da regressão. Por esta razão o modelo de efeitos aleatórios é muitas vezes referido como modelo das componentes do erro.

correlacionado com $u_{it} - \bar{u}_i$ mesmo que u_{it} não apresente correlação em série. Isto porque \bar{y}_{i-1} está correlacionado com \bar{u}_i pela sua construção – a média \bar{u}_i contém u_{it-1} correlacionado com \bar{y}_{it-1} . Para pequenos T o enviesamento de γ é sempre negativo se $\gamma > 0$. A solução para o problema requer a utilização de variáveis instrumentais.

¹⁵² Este pressuposto permite assumir a exogeneidade restrita, isto é o pressuposto que $Cov(x_{it}, \varepsilon_{is}) = 0$, para todo o $t \neq s$. Se é violado há um problema de endogeneidade: a variável independente e o termo de erro idiosincrático estão correlacionados. Sob esta realidade o estimador dos efeitos fixos virá enviesado: endogeneidade, nesta perspectiva, constitui um problema mesmo com dados em painel.

O modelo pode ser especificado como se segue, com particular relevância para os pressupostos feitos relativamente ao termo erro:

$$\mathbf{y}_{it} = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it}, \quad \text{onde } \varepsilon_{it} = \mu_i + \lambda_t + v_{it}$$

$$\mu_i \approx \text{i.i.d. } (0, \sigma_\mu^2)$$

$$\lambda_t \approx \text{i.i.d. } (0, \sigma_\lambda^2)$$

$$v_{it} \approx \text{i.i.d. } (0, \sigma_v^2)$$

Assumindo-se que,

$$E[v_{it}, v_{js}] = 0 \quad (t \neq s, i \neq j); \quad E[\mu_i, \mu_j] = 0 \quad (i \neq j) \quad \text{e} \quad E[\lambda_t, \lambda_s] = 0 \quad (t \neq s)$$

$$E[v_{it}, \mu_j] = 0, \quad \text{para todo o } i, t \text{ e } j$$

$$E[v_{it}, \lambda_s] = 0, \quad \text{para todo o } i, t \text{ e } s$$

E relativamente ao termo composto da perturbação, decorre, obviamente que,

$$E(\varepsilon_{it}) = 0,$$

enquanto,

$$E(\varepsilon_{it}^2) = \text{var}(\varepsilon_{it}) = \sigma_{\mu}^2 + \sigma_{\lambda}^2 + \sigma_{\nu}^2 \quad \text{e} \quad E[\varepsilon_{it}, \varepsilon_{is}] = \sigma_{\mu}^2, \quad (t \neq s)$$

$$E[\varepsilon_{it}, \varepsilon_{jt}] = \sigma_{\lambda}^2, \quad (i \neq j)$$

e

$$E[\varepsilon_{it}, \varepsilon_{js}] = 0, \quad (i \neq j; t \neq s)$$

A assumção de aleatoriedade nos efeitos seccionais, μ_i , e nos efeitos temporais, λ_t , com a consequente assumção de que seguem distribuições idênticas e independentes, isto é $\mu_i \approx \text{i.i.d.}(0, \sigma_{\mu}^2)$, $\lambda_t \approx \text{i.i.d.}(0, \sigma_{\lambda}^2)$ e, ainda que esses efeitos, μ_i e λ_t , são independentes da perturbação estocástica ν_{it} ; adicionada, à independência das variáveis exógenas x_{it} de μ_i , de λ_t e de ν_{it} , para todo o i e para todo o t , evita a perda de graus de liberdade e a estimação de demasiados parâmetros que decorrem do recurso ao modelo de efeitos fixos.

Tratando-se do modelo de efeitos aleatórios há que considerar que: (i) o método dos mínimos quadrados ordinário (MQO) produz estimativas consistentes de β , mas com erros padrão pouco interessantes e (ii) o método ordinário dos mínimos quadrados (MQO) não é eficiente. Ora, sob o modelo de efeitos aleatórios, o estimador dos mínimos quadrados generalizados (MQG), baseado em componentes de variância do erro verdadeiras, é BLUE e todos os estimadores admissíveis dos mínimos quadrados generalizados (MQG_F) são assintoticamente eficientes, quer seja N , ou T a tender para infinito.

De forma a obterem-se coeficientes estimados pelo método generalizado dos mínimos quadrados, *MQG*, precisa-se da matriz de variância-covariância, baseada na estrutura do termo perturbação, $\varepsilon_{it} = \mu_i + \lambda_t + v_{it}$,¹⁵³ vindo,

$$\begin{aligned}\mathbf{\Omega} &= E(\varepsilon\varepsilon') = \mathbf{Z}_\mu E(\mu\mu')\mathbf{Z}_\mu' + \mathbf{Z}_\lambda E(\lambda\lambda')\mathbf{Z}_\lambda' + E(vv') \\ &= \sigma_\mu^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_\lambda^2(\mathbf{J}_N \otimes \mathbf{I}_T) + \sigma_v^2(\mathbf{I}_N \otimes \mathbf{I}_T)\end{aligned}$$

154

Isso implica variância homoscedástica, $\text{var}(\varepsilon_{it}) = \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2$, para todo o i e para todo o t , e uma matriz por blocos equicorrelacionada, que apresenta correlação em série apenas para perturbações da mesma unidade seccional.¹⁵⁵ Na realidade:

$$\begin{aligned}\text{cov}(\varepsilon_{it}, \varepsilon_{js}) &= \sigma_\mu^2, & \text{para } i = j, t \neq s \\ &= \sigma_\lambda^2, & \text{para } i \neq j, t = s\end{aligned}$$

e zero para os outros casos.

¹⁵³ Modelo de componente de erro a duas dimensões, que ainda pode ser reescrito como $\varepsilon = \mathbf{Z}_\mu\mu + \mathbf{Z}_\lambda\lambda + v$, em que, \mathbf{Z}_μ é a matriz das variáveis *dummy* seccionais que se incluem na regressão para estimar μ_i , também, expressa como um produto Kronecker de uma matriz identidade, de dimensão N , \mathbf{I}_N , por um vector de uns de dimensão T , $\mathbf{1}_T$, isto é $\mathbf{Z}_\mu = \mathbf{I}_N \otimes \mathbf{1}_T$, e \mathbf{Z}_λ é a matriz das variáveis *dummy* temporais que se incluem na regressão para estimar λ_t , também, expressa como um produto Kronecker de uma matriz identidade, de dimensão T , \mathbf{I}_T , por um vector de uns de dimensão N , $\mathbf{1}_N$, isto é $\mathbf{Z}_\lambda = \mathbf{I}_T \otimes \mathbf{1}_N$, caso se assumam parâmetros fixos.

¹⁵⁴ $\mathbf{Z}_\mu\mathbf{Z}_\mu' = \mathbf{I}_N \otimes \mathbf{J}_T$ e $\mathbf{Z}_\lambda\mathbf{Z}_\lambda' = \mathbf{I}_T \otimes \mathbf{J}_N$, em que \mathbf{J}_T e \mathbf{J}_N são matrizes, cujos elementos são a unidade, de dimensão T e N , respectivamente.

¹⁵⁵ O que vem ilustrado por, $\mathbf{\Omega}_{(NT \times NT)} = \begin{bmatrix} \Sigma & \ddots & \Sigma_\lambda \\ \ddots & \ddots & \ddots \\ \Sigma_\lambda & \dots & \Sigma \end{bmatrix}$, em que $\Sigma_\lambda = \begin{bmatrix} \sigma_\lambda^2 & \dots & 0 \\ \dots & \sigma_\lambda^2 & \dots \\ 0 & \dots & \sigma_\lambda^2 \end{bmatrix}$ e

$$\Sigma_{T \times T} = \begin{bmatrix} \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2 & \dots & \sigma_\mu^2 \\ \dots & \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2 & \dots \\ \sigma_\mu^2 & \dots & \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2 \end{bmatrix}$$

O procedimento da estimação *MQG*, segundo Fuller e Battese (1973, 1974), sugere a transformação da equação da regressão, por via da pré-multiplicação por $\sigma_v \mathbf{\Omega}^{-1/2}$, para posterior estimação pelo método dos mínimos quadrados ordinários, traduzido como se segue:

$$\sigma_v \mathbf{\Omega}^{-1/2} y_{it} = \sigma_v \mathbf{\Omega}^{-1/2} x_{it} \beta + \sigma_v \mathbf{\Omega}^{-1/2} \varepsilon_{it} \quad (4.4.1)$$

Entretanto a matriz de transformação $\mathbf{\Omega}^{-1/2}$ pode obter-se recorrendo à decomposição espectral da matriz de variância $\mathbf{\Omega}$,

$$\begin{aligned} \mathbf{\Omega} = \sum_{i=1}^4 \zeta_i \mathbf{Q}_i = & \sigma_v^2 (\mathbf{E}_N \otimes \mathbf{E}_T) + (T\sigma_\mu^2 + \sigma_v^2) (\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + (N\sigma_\lambda^2 + \sigma_v^2) (\bar{\mathbf{J}}_N \otimes \mathbf{E}_T) + \\ & + (T\sigma_\mu^2 + N\sigma_\lambda^2 + \sigma_v^2) (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T) \end{aligned} \quad 156$$

Onde os ζ_i correspondem aos diferentes valores próprios da matriz de covariância $\mathbf{\Omega}$ e os diversos \mathbf{Q}_i , $(\mathbf{E}_N \otimes \mathbf{E}_T)$, $(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T)$, $(\bar{\mathbf{J}}_N \otimes \mathbf{E}_T)$, $(\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)$,¹⁵⁷ são as correspondentes matrizes dos projectores dos valores próprios: simétricas e idempotentes com as respectivas características coincidentes com os respectivos traços.¹⁵⁸

¹⁵⁶ Esta decomposição espectral permite a extensão $\mathbf{\Omega}^r = \sum_{i=1}^4 \zeta_i^r \mathbf{Q}_i$, onde r é um escalar arbitrário e por isso o parâmetro de transformação $\sigma_v \mathbf{\Omega}^{-1/2}$ poder assumir a expressão, $\sum_{i=1}^4 \sigma_v \zeta_i^{-1/2} \mathbf{Q}_i$.

¹⁵⁷ As matrizes \mathbf{J}_T e \mathbf{J}_N , de dimensão T e N , respectivamente, cujos elementos são a unidade, vêm substituídas pelos produtos dos escalares T e N pelas correspondentes matrizes $\bar{\mathbf{J}}_T$ e $\bar{\mathbf{J}}_N$, decorrendo os elementos $1/T$ e $1/N$, em cada uma das matrizes $\bar{\mathbf{J}}_T$ e $\bar{\mathbf{J}}_N$. Já as matrizes, \mathbf{I}_T e \mathbf{I}_N vêm substituídas por $\mathbf{E}_T + \bar{\mathbf{J}}_T$ e $\mathbf{E}_N + \bar{\mathbf{J}}_N$, respectivamente.

¹⁵⁸ Para explicação mas detalhada ver Baltagi (2005).

A necessidade de se estimarem as componentes de variância é evidente, a resposta pode, naturalmente, residir numa estimação pelo método dos mínimos quadrados da regressão (4.4.1), a qual virá, $y^* = z^* + \varepsilon^*$, em que,

$$y^* = \sigma_v \mathbf{\Omega}^{-1/2} y_{it} = y_{it} - \theta_1 \bar{y}_i. - \theta_2 \bar{y}_t - \theta_3 \bar{y}..$$

$$z^* = \sigma_v \mathbf{\Omega}^{-1/2} x_{it} \beta = (x_{it} - \theta_1 \bar{x}_i. - \theta_2 \bar{x}_t - \theta_3 \bar{x}..) \beta$$

e

$$\varepsilon^* = \sigma_v \mathbf{\Omega}^{-1/2} \varepsilon_{it} = \varepsilon_{it} - \theta_1 \bar{\varepsilon}_i. - \theta_2 \bar{\varepsilon}_t - \theta_3 \bar{\varepsilon}..$$

Sendo que os parâmetros,

$$\theta_1 = 1 - \left(\frac{\sigma_v}{\sqrt{T\sigma_\mu^2 + \sigma_v^2}} \right); \quad \theta_2 = 1 - \left(\frac{\sigma_v}{\sqrt{N\sigma_\lambda^2 + \sigma_v^2}} \right) \quad e$$

$$\theta_3 = 1 - \left(\frac{\sigma_v}{\sqrt{T\sigma_\mu^2 + \sigma_v^2}} \right) + 1 - \left(\frac{\sigma_v}{\sqrt{N\sigma_\lambda^2 + \sigma_v^2}} \right) + \left(\frac{\sigma_v}{\sqrt{T\sigma_\mu^2 + N\sigma_\lambda^2 + \sigma_v^2}} \right) - 1,$$

reflectem factores de ponderação a reter para os estimadores *dentro* e *entre*.

Efectivamente a matriz de transformação $\mathbf{\Omega}^{-1/2}$ transforma dados observados em desvios de médias específicas, onde, as médias específicas das unidades seccionais são ponderadas utilizando a componente de variância θ_1 ; as médias específicas dos períodos temporais são ponderadas utilizando a componente de variância θ_2 e as médias específicas das unidades seccionais e dos períodos temporais são ponderadas utilizando a componente de variância θ_3 .

Como resultado obtêm-se as estimações das componentes de variância admissíveis por via da substituição das verdadeiras perturbações por resíduos estimados pelo método dos mínimos quadrados ordinários (*MQO*). Sob o modelo de efeitos aleatórios o estimador *MQO* é, ainda, consistente e não enviesado, mas é ineficiente, o que resulta em erros padrão e *t* – estatísticos enviesados. Swamy e Arora (1972) sugeriram que as componentes de variância resultassem da média dos erros quadrados decorrentes de três regressões pelo método dos mínimos quadrados; em que a primeira regressão correspondia à regressão *dentro*, a que se baseia no modelo original transformado por $\mathbf{Q}_1 = \mathbf{E}_N \otimes \mathbf{E}_T$; ¹⁵⁹ a segunda regressão equivalia a uma regressão *entre* unidades seccionais, suportada na transformação do modelo original por $\mathbf{Q}_2 = \mathbf{E}_N \otimes \bar{\mathbf{J}}_T$ ¹⁶⁰ e, por fim, a terceira regressão condizia com uma regressão entre períodos temporais, decorrendo da transformação do modelo original por $\mathbf{Q}_3 = \bar{\mathbf{J}}_N \otimes \mathbf{E}_T$. ¹⁶¹

Utilizando a decomposição de $\mathbf{\Omega}^{-1/2}$ pode-se mostrar que o coeficiente estimado β pelo método *MQG* é uma soma ponderada de estimadores *dentro*(*Within*) e *entre*(*Between*), ¹⁶² vindo,

$$\hat{\beta}_{MQG} = \left[\mathbf{x}'\mathbf{Q}_1\mathbf{x}/\sigma_v^2 + \mathbf{x}'\mathbf{Q}_2\mathbf{x}/\xi_2 + \mathbf{x}'\mathbf{Q}_3\mathbf{x}/\xi_3 \right]^{-1} \left[\mathbf{x}'\mathbf{Q}_1\mathbf{y}/\sigma_v^2 + \mathbf{x}'\mathbf{Q}_2\mathbf{y}/\xi_2 + \mathbf{x}'\mathbf{Q}_3\mathbf{y}/\xi_3 \right]$$

o qual pode ser reescrito como,

¹⁵⁹ Equivalente à regressão $(y_{it} - \bar{y}_i - \bar{y}_t - \bar{y}_{..}) = (x_{it} - \bar{x}_i - \bar{x}_t - \bar{x}_{..})\beta + (v_{it} - \bar{v}_i - \bar{v}_t - \bar{v}_{..})$ e permite estimar σ_v^2 .

¹⁶⁰ Equivalente à regressão $(\bar{y}_i - \bar{y}_{..}) = (\bar{x}_i - \bar{x}_{..})\beta + (\bar{v}_i - \bar{v}_{..})$ e permite estimar $T\sigma_\mu^2 + \sigma_v^2$.

¹⁶¹ Equivalente à regressão $(\bar{y}_t - \bar{y}_{..}) = (\bar{x}_t - \bar{x}_{..})\beta + (\bar{v}_t - \bar{v}_{..})$ e permite estimar $N\sigma_\lambda^2 + \sigma_v^2$.

¹⁶² Conforme apresentado no Apêndice F.

$$\hat{\beta}_{MQG} = \left[\mathbf{x}'\mathbf{Q}_1\mathbf{x} + \mathbf{x}'\mathbf{Q}_2\mathbf{x}(\sigma_v^2/(T\sigma_\mu^2 + \sigma_v^2)) + \mathbf{x}'\mathbf{Q}_3\mathbf{x}(\sigma_v^2/(N\sigma_\lambda^2 + \sigma_v^2)) \right]^{-1} * \\ * \left[\mathbf{x}'\mathbf{Q}_1\mathbf{y} + \mathbf{x}'\mathbf{Q}_2\mathbf{y}(\sigma_v^2/(T\sigma_\mu^2 + \sigma_v^2)) + \mathbf{x}'\mathbf{Q}_3\mathbf{y}(\sigma_v^2/(N\sigma_\lambda^2 + \sigma_v^2)) \right]$$

Se $\sigma_\mu^2 = \sigma_\lambda^2 = 0$, viria apenas, $\hat{\beta}_{MQG} = [\mathbf{x}'\mathbf{x}]^{-1} * [\mathbf{x}'\mathbf{y}]$, estar-se-ia face a um coeficiente estimado em tudo semelhante ao que decorre da estimação pelo método dos mínimos quadrados. Caso $N \rightarrow \infty$ e $T \rightarrow \infty$, ter-se-ia $(\sigma_v^2/(N\sigma_\lambda^2 + \sigma_v^2)) \rightarrow 0$ e $(\sigma_v^2/(T\sigma_\mu^2 + \sigma_v^2)) \rightarrow 0$ e o coeficiente estimado viria idêntico ao coeficiente estimado pelo método dos efeitos fixos, por via da técnica *dentro*. Resta explorar a situação em que $(\sigma_v^2/(N\sigma_\lambda^2 + \sigma_v^2)) \rightarrow \infty$, com $(\sigma_v^2/(T\sigma_\mu^2 + \sigma_v^2))$ finito, nesse caso o coeficiente estimado viria semelhante ao estimado pelo método *entre*, centrado nas unidades seccionais ou, então a situação em que $(\sigma_v^2/(T\sigma_\mu^2 + \sigma_v^2)) \rightarrow \infty$, para um valor finito de $(\sigma_v^2/(N\sigma_\lambda^2 + \sigma_v^2))$, o coeficiente estimado seria do mesmo tipo, mas agora centrado nos períodos temporais.

Uma vez que um número infinito de períodos e grupos homogêneos dificilmente convivem com a realidade prática, então, sob o pressuposto de efeitos aleatórios, o estimador *MQG* revela-se mais apropriado, constitui um feliz compromisso entre esses dois pressupostos. [Greene (2003), Hsiao (1986) e Hausman e Taylor (1981)].

4.4.2. Abordagem técnica da estimação

A especificação do modelo de efeitos aleatórios vai ser estimada recorrendo ao EVIEWS 5. Este pacote estatístico reclama a utilização de técnicas do método dos mínimos quadrados generalizado admissível. O primeiro passo para estimar a matriz covariância para o erro $\varepsilon_{it} = \mu_i + \lambda_t + v_{it}$ suporta-se num estimador de componentes de

variância que utiliza o método do estimador quadrático não enviesado (*QUE*), presente nos estimadores Swamy-Arora (SWAR),¹⁶³ Wallace-Hussain (WALHUS)¹⁶⁴ e Amemya (AMEMYA)¹⁶⁵, disponibilizados pelo pacote estatístico empregado. Logo que as componentes de variância sejam estimadas obtém-se um estimador composto de covariância para que o modelo original seja transformado, via *MQG*.

A síntese de resultados obtidos é a constante do quadro que se segue,

	$MQG_F^{(a)(b)(c)}$ SWAR	$MQG_F^{(a)(b)(c)}$ WALHUS	$MQG_F^{(a)(b)(c)}$ AMEMYA
<i>VOTOSANT</i>	0,80711*** (34,6308)	0,80733*** (34,6120)	0,80717*** (34,6261)
<i>APART</i>	0,47874*** (11,0561)	0,46977*** (10,85101)	0,47638*** (11,0021)
<i>APRES</i>	-0,067176** (-1,225029)	-0,071704 (-1,306745)	-0,068368** (-1,246556)
<i>DIREAL</i>	0,00060 (0,02878)	-0,0003 (-0,01409)	0,00037 (0,01752)
<i>PDESEMP1</i>	-0,12233*** (-2,08294)	-0,10897** (-1,86944)	-0,11880*** (-2,02675)
<i>PINFLAC1</i>	0,06481 (0,62258)	0,03967 (0,38291)	0,05817 (0,55947)
σ_μ	0,000000	0,000000	0,000000
σ_λ	4,39196	2,97989	3,84204
σ_ν	9,33729	9,40727	9,33729
$R^2_{ajust.}$	0,45227	0,45064	0,45184

Tabela 4.4.2-1: Resultados de regressões em painel, modelo efeitos aleatórios

¹⁶³ Neste procedimento o estimador das componentes de variância, resulta dos resíduos obtidos a partir de regressores *dentro e entre*.

¹⁶⁴ Neste procedimento utiliza-se, apenas, os resíduos *MQO*.

¹⁶⁵ Neste procedimento os resíduos são, apenas, os que decorrem da estimação *dentro*.

^(a) significa que a componente do erro foi considerada a duas dimensões, a componente específica seccional e a componente específica temporal; ^(b) estimação pelo método fiável dos mínimos quadrados generalizado, com ponderação temporal; ^(c) estimação com correcção painel dos erros padrão *PCSE*.

Para as três estimações estudadas, da componente de variância, para o modelo de efeitos aleatórios, obtêm-se os mesmos resultados no que se refere à variância dos efeitos específicos das unidades seccionais, os quais representam 0% da variância total. Por este facto, os valores de θ_1 e de θ_3 vêm nulos e a regressão *MQO* faz-se sobre o modelo, $y_{it} - \theta_2 \bar{y}_t = (x_{it} - \theta_2 \bar{x}_t) \beta + (\varepsilon_{it} - \theta_2 \bar{\varepsilon}_t)$, onde a transformação reveste apenas a forma de desvios relativamente às médias dos valores observados em cada um dos períodos de tempo, transformação *dentro* (*W*).

No que diz respeito à variância dos efeitos temporais na variância total o valor oscila entre os 9,1% para o estimador de variância WALHUS e os 18,1% para o estimador SWAR e pelo meio atinge os 14,5% para o estimador AMEMYA. Resta identificar os valores da variância da perturbação remanescente, estocástica, na variância total, vindo o valor de 91,9%, de 81,9% e de 85,5%, para os estimadores de variância de WALHUS, de SWAR e de AMEMYA, respectivamente.

O intervalo, em que oscila, a variância da perturbação estocástica, de 82% a 92%, sugere uma excepcional relevância desta componente na variância total da estimação, o que pode ir ao encontro do demonstrado por Maddala (1971) relativamente à igualdade entre os coeficientes estimados, quer pelo método ordinário dos mínimos quadrados, *MQO*, quer pelo método generalizado dos mínimos quadrados, *MQG*.

4.5. Escolha entre o Pressuposto dos Efeitos Fixos e dos Efeitos Aleatórios

Como já foi referido na secção anterior a diferença crucial, entre modelos de efeitos fixos e de efeitos aleatórios, está nos pressupostos feitos acerca da natureza dos efeitos específicos seccionais invariantes no tempo (ou dos efeitos específicos temporais invariantes nas unidades seccionais). Os termos específicos fixos das

unidades/seccionais assumem-se como fornecendo resultados apenas aplicáveis à amostra em análise, daí que as inferências possam, apenas, ser feitas condicionalmente sobre essa amostra. Enquanto isso, se o termo específico for tratado como aleatório, possibilitará, então, uma inferência incondicional relativamente a todo o tipo de amostras, para qualquer período de tempo, que possam ser obtidas a partir de uma dada população [Greene (2003) e Hsiao (1986)]. Neste sentido o modelo de efeitos fixos é um caso especial do modelo de efeitos aleatórios, onde a amostra em análise existe como sendo a população inteira. Esta ideia aparece nas secções anteriores onde os resultados dos efeitos fixos se apresentam como casos especiais de aplicações de efeitos aleatórios.

A escolha entre efeitos fixos e efeitos aleatórios tem que ter em conta o problema empírico em análise, bem como com a natureza dos dados; e, ainda, o método de recolha dos dados, bem como com a questão a ser respondida. Para o efeito, existem testes formais de especificação.

A óptica dos efeitos fixos liberta a regressão de todos os efeitos fixos, independentemente de estarem contidos como inobserváveis nos termos das intercepções, ou como observáveis nas matrizes dos regressores.

Os efeitos dos regressores específicos seccionais, não observáveis, como aqueles que possam decorrer do sexo ou do local de nascimento dos cidadãos eleitores em cada município, ou de qualquer talento do candidato a autarca percebido pelos cidadãos/eleitores em cada município, sobre a variável dependente, são captados por via das diferenças no termo constante da regressão,¹⁶⁶ já que não podem ser estimados sob a perspectiva dos efeitos fixos [Johnston e DiNardo (1997), Hausman e Taylor (1981) e Maddala (1971)].

Efectivamente, não há razão lógica para a especificação dos efeitos fixos não observáveis, μ_i , já que a perturbação induzida por esses efeitos, similarmente ao que acontece com o termo erro estocástico, decorre de alguma dose de ignorância presente à especificação do modelo e imputável aos pesquisadores; não havendo, por

¹⁶⁶ Denominando-se, por isso, “heterogeneidade não mensurada”, uma forma de enviesamento omitida.

isso, fundamento para que se distinga entre a "ignorância específica" decorrente da heterogeneidade não mensurada e a "ignorância geral" implícita nos termos de resíduos estimados [Maddala (1971) e Hsiao (1986)].

Os efeitos dos regressores específicos seccionais, observáveis, sob a perspectiva dos mínimos quadrados com variáveis *dummy*, *MQVD*, são integralmente captados pela estimação das $N - 1$ variáveis *dummy* introduzidas na especificação do modelo.

O método de estimação dos efeitos fixos é encarado como convivendo com problemas teóricos, já que a óptica dos mínimos quadrados com variáveis *dummy*, *MQDV*, tem sido amplamente criticada, sendo a perda de graus de liberdade, associada ao elevado número de parâmetros estimado, um dos pontos censuráveis; ainda que, em termos práticos, o problema, só, em função do número disponível de observações, poderá vir a colocar-se [Maddala (1971)].

Escolher uma especificação incorrecta pode resultar, quer em enviesamento, quer em perda de eficiência. O estimador dos efeitos *MQVD* é *BLUE*, quando o pressuposto $E[\varepsilon_{it} X_{it}] \neq 0$ se verifica, mas pode resultar num estimador assintoticamente ineficiente, consistente e não enviesado se o efeito individual é assumido contemporaneamente aleatório – isto é sob o pressuposto de que os μ_i seguem uma distribuição aleatória - ainda que se comporte como um modelo de efeitos fixos. Neste cenário utilizar *MQG* gera estimativas enviesadas e inconsistentes [Hsiao (1986)]. Quando a regressão é corrigida, *a priori*, utilizando o procedimento *MQG*, os resultados estimados tornam-se idênticos aos estimados por *MQVD* [Mundlak (1978)]. A diferença chave entre modelos de efeito fixo e de efeito aleatório não é, de acordo com Mundlak, se os efeitos individuais são fixos ou aleatórios, mas se estão correlacionados com os regressores [Mundlak (1978)]. É, então, útil, apenas, propor testes para discutir efeitos fixos *versus* efeitos aleatórios [Hausman e Taylor (1981)].

4.5.1. Testes de especificação de Hausman

A base para qualquer teste, descrito abaixo, tem a ver com a ideia de saber se, caso falhe o pressuposto $E[\varepsilon_{it}, x_{it}] = 0$, os estimadores, *MQG*, *dentro* e *entre*, vêm afectados diferentemente. Testar esse pressuposto fundamenta-se na aplicação desses três estimadores.

Nesse sentido, Hausman desenvolveu três testes de especificação, de forma a testar,

a hipótese nula, $H_0 = E[\varepsilon_{it}, x_{it}] = 0$,

contra,

a hipótese alternativa, $H_1 : E[\varepsilon_{it}, x_{it}] \neq 0$.

Seguindo-se, por isso,

Teste 1: Procura encontrar diferença entre os coeficientes estimados, $\hat{\beta}$, comparando as estimativas obtidas a partir, do método dos mínimos quadrados com variáveis *dummy*, *MQVD*, e do método dos mínimos quadrados generalizados, *MQG*. Ambos são consistentes sob a hipótese nula, mas cada um deles terá limites probabilísticos diferentes, caso a hipótese nula não seja verdadeira. O teste estatístico de Hausman suporta-se na diferença entre os coeficientes estimados, $\hat{\beta}_{MQVD}$ e $\hat{\beta}_{MQG}$, a qual se passa a identificar como um parâmetro \hat{q}_1 , vindo, $\hat{\beta}_{MQVD} - \hat{\beta}_{MQG} = \hat{q}_1$.¹⁶⁷ É esta diferença entre os coeficientes estimados, conjuntamente, com a diferença entre as matrizes de covariância dos coeficientes estimados, via *MQVD* e *MQG*,

¹⁶⁷ Sob a hipótese nula, $\text{plim}\hat{q}_1 = 0$ e $\text{cov}(\hat{q}_1, \hat{\beta}_{MQG}) = 0$.

$\text{var}(\hat{q}_1) = \text{var}(\hat{\beta}_{MQVD}) - \text{var}(\hat{\beta}_{MQG})$, que permite obter a estatística,

$$m_1 = \hat{q}'_1 [\text{var}(\hat{q}_1)]^{-1} \hat{q}_1 \approx \chi^2_k. \quad 168$$

Teste 2: Procura encontrar diferença entre estimativas *entre (B)* e *MQG*; agora o parâmetro \hat{q}_2 , paralelamente, ao que decorreu para o teste 1, identifica a diferença entre

os coeficientes estimados por via dos modelos *entre* e *MQG*, vindo, $\hat{q}_2 = \hat{\beta}_{MQG} - \hat{\beta}_B$.

A estatística de Hausman vem, $m_2 = \hat{q}'_2 [\text{var}(\hat{q}_2)]^{-1} \hat{q}_2 \approx \chi^2_k$, em que,

$$\text{var}(\hat{q}_2) = \text{var}(\hat{\beta}_{MQG}) - \text{var}(\hat{\beta}_B).$$

Teste 3: A hipótese nula a ser testada utiliza agora a diferença entre os valores estimados dos coeficientes, decorrente das estimativas do modelo de efeitos fixos *dentro (W)* e *entre (B)*, bem assim como a diferença entre as matrizes de covariância desses coeficientes, vindo,

$m_3 = \hat{q}'_3 [\text{var}(\hat{q}_3)]^{-1} \hat{q}_3 \approx \chi^2_k$, com $\hat{q}_3 = \hat{\beta}_W - \hat{\beta}_B$ e

$$\text{var}(\hat{q}_3) = \text{var}(\hat{\beta}_W) - \text{var}(\hat{\beta}_B).$$

Se m_1 , m_2 , ou m_3 ,¹⁶⁹ se revelam não significantes a H_0 pode ser rejeitada, por outras palavras tem-se $E[\varepsilon_{it}, x_{it}] \neq 0$ e as especificações de efeitos aleatórios não são

¹⁶⁸ A estatística de Hausman, sob a hipótese nula segue, assintoticamente, uma distribuição do *Qui-Quadrado*, onde k identifica o número de coeficientes, a estimar, elementos do vector β .

¹⁶⁹ Estas estatísticas fundamentam-se no lema de Hausman (1978), o qual sob a hipótese nula (*REM* – modelo de efeitos aleatórios), permite escrever,

$$\sqrt{NT}[\hat{\beta}_{MQG} - \beta] \xrightarrow{d} N[0, \text{var}_{MQG}] \quad (\text{eficiente})$$

$$\sqrt{NT}[\hat{\beta}_{MQVD} - \beta] \xrightarrow{d} N[0, \text{var}_{MQVD}] \quad (\text{ineficiente}).$$

$$(\hat{q} = (\hat{\beta}_{MQVD} - \beta) - (\hat{\beta}_{MQG} - \beta))$$

Este lema estabelece que na distribuição conjunta limite de $\sqrt{NT}[\hat{\beta}_{MQG} - \beta]$ e de $\sqrt{NT}\hat{q}$, a covariância limite, $\text{cov}_{q, MQG}$, é zero.

Mas, $\text{cov}_{q, MQG} = \text{cov}_{MQVD, MQG} - \text{var}_{MQG}$. Então, $\text{var}[q] = \text{var}_{MQVD} + \text{var}_{MQG} - \text{cov}_{MQVD, MQG} - \text{cov}'_{MQVD, MQG}$.

A partir do lema, $\text{cov}_{MQVD, MQG} = \text{var}_{MQG}$. Seguindo-se $\text{var}[q] = \text{var}_{MQVD} - \text{var}_{MQG}$. Com base no referido, obtém-se,

$$H = (\hat{\beta}_{MQVD} - \hat{\beta}_{MQG})' [\text{Var Est.}(\hat{\beta}_{MQVD}) - \text{Var Est.}(\hat{\beta}_{MQG})]^{-1} (\hat{\beta}_{MQVD} - \hat{\beta}_{MQG})$$

Com,

apropriadas. Os três testes são numericamente idênticos em valor. Não importa qual das duas estimações está disponível, mas sim que possam ser utilizadas para testar a especificação de efeitos aleatórios [Hausman e Taylor (1981)].

A extensão do teste de Hausman, ao modelo de componentes de erro a duas dimensões, pode, ainda, basear-se no procedimento padrão desse teste para discriminar entre estimadores de efeitos fixos (com *dummies*, quer seccionais, quer periódicas) e estimadores de efeitos aleatórios *MQG* a duas dimensões. O problema que se coloca é ao nível da equivalência entre os testes para a estimação do erro a uma dimensão e para a estimação do erro a duas dimensões, já que na estimação do erro a duas dimensões ter-se-ia dois estimadores *entre* (B), um *entre* períodos temporais ($\hat{\beta}_T$) e outro *entre* unidades seccionais ($\hat{\beta}_N$). Consequentemente o estimador dos efeitos aleatórios ($\hat{\beta}_{MQG}$) passa a ser uma combinação ponderada dos estimadores, *entre*, $\hat{\beta}_T$ e $\hat{\beta}_N$, e *dentro* $\hat{\beta}_W$. De acordo com Kang (1985) o teste de Hausman baseado em $\hat{\beta}_{MQVD}$ e $\hat{\beta}_{MQG}$ (*teste 1*) não é equivalente ao *teste 2*, agora suportado em $\hat{\beta}_{MQG} - \hat{\beta}_T$ e $\hat{\beta}_{MQG} - \hat{\beta}_N$, por via dos dois estimadores *entre*, avançando, no entanto, para outro tipo de equivalências através de cinco hipóteses testáveis:

(H1) Testa $E(\lambda_t | x_{it}) = 0$, suportado na diferença $\hat{\beta}_W - \hat{\beta}_T$, uma vez assumidos fixos os efeitos seccionais, μ_i ;

(H2) Testa $E(\lambda_t | x_{it}) = 0$, suportado na diferença $\hat{\beta}_T - \hat{\beta}_{MQG}$, já que se assumem aleatórios os efeitos seccionais, μ_i ;

(H3) Testa $E(\mu_i | x_{it}) = 0$, suportado na diferença $\hat{\beta}_W - \hat{\beta}_N$, pela assumpção de efeitos temporais, λ_t , fixos;

$$\text{Var Est.}(\hat{\beta}_{MQVD}) = \sigma_\varepsilon^2 \left[\sum_{i=1}^N \mathbf{x}_i' \left(\mathbf{I} - \frac{1}{T_i} \mathbf{ii}' \right) \mathbf{x}_i \right]^{-1}$$

$$\text{Var Est.}(\hat{\beta}_{MQG}) = \sigma_\varepsilon^2 \left[\sum_{i=1}^N \mathbf{x}_i' \left(\mathbf{I} - \frac{\tilde{\gamma}_i}{T_i} \mathbf{ii}' \right) \mathbf{x}_i \right]^{-1},$$

$$0 \leq \tilde{\gamma}_i = \frac{T_i \sigma_\mu^2}{\sigma_\varepsilon^2 + T_i \sigma_\mu^2} \leq 1.$$

À medida que $N \rightarrow \infty$, e σ_ε^2 e σ_μ^2 se tornam consistentes, $\text{Var Est.}(\hat{\beta}_{MQVD}) - \text{Var Est.}(\hat{\beta}_{MQG})$ torna-se definida não negativa.

(H4) Testa $E(\mu_i | x_{it}) = 0$, suportado na diferença $\hat{\beta}_N - \hat{\beta}_{MQG}$, já que os efeitos temporais, λ_t , se assumem aleatórios;

(H5) Compara dois estimadores, um que assume que os efeitos seccionais μ_i e os efeitos temporais λ_t são fixos e outro que assume que ambos são aleatórios, tal que, $E(\lambda_t | x_{it}) = E(\mu_i | x_{it}) = 0$. Este teste suporta-se em, $\hat{\beta}_{MQG} - \hat{\beta}_W$.

4.5.2. Aplicação de testes de especificação de Hausman

No âmbito do estudo em curso, após se ter estimado o modelo de acordo com os métodos de efeitos fixos e de efeitos aleatórios, importava avaliar qual dos métodos se revela mais poderoso, isto é qual o método que gera coeficientes estimados robustos e eficientes. Ora, como foi referido na secção anterior, o teste de Hausman apresenta-se como o procedimento standardizado, em dados painel, para distinguir entre parâmetros estimados via modelo de efeitos fixos e modelo de efeitos aleatórios.

Efectivamente, introduziu-se um teste de Hausman, para contrapor modelos de efeitos fixos a modelos de efeitos aleatórios, num ambiente EVIEWS 5. As limitações decorrentes do pacote estatístico impuseram o desenvolvimento de um programa que permitisse aplicar o teste de Hausman aos dados em estudo (conforme apêndice G), o qual permite aferir um resultado de $\chi^2_{[6]} = 82,4081$, com um p -valor = $1,11E - 15$. O recurso ao pacote estatístico STATA 9 para aplicação do teste de Hausman produz o resultado $\chi^2_{[6]} = 79,04$, com $Prob > \chi^2 = 0,0000$.

A coerência dos resultados obtidos, que se caracteriza por valores acima do limite superior da distribuição do *Qui-Quadrado* permite a rejeição da hipótese nula, $H_0 = E[\varepsilon_{it}, x_{it}] = 0$. O estimador dos efeitos aleatórios não é consistente. A hipótese de que os efeitos “municípios” não estão correlacionados com os outros regressores deveria ser rejeitada. Levando a concluir que o modelo de efeitos fixos se comporta bem, o que permite equacionar diferenças intrínsecas entre os municípios.

Ainda, recorrendo ao tipo de equivalência, proposto por Kang, no contexto da extensão do teste de Hausman ao modelo de componentes de erro a duas dimensões (seccionais e temporais), constante da hipótese que compara dois estimadores, um que assume que os efeitos seccionais μ_i e os efeitos temporais λ_t são fixos e outro que assume que ambos são aleatórios, tal que, $E(\lambda_t | x_{it}) = E(\mu_i | x_{it}) = 0$, podemos deter-nos na estimação do modelo de componentes de erro, pelo método dos mínimos quadrados ordinário, *MQO*, depois de uma transformação do tipo,

$$\sigma_v \Omega^{-1/2} y_{it} = \sigma_v \Omega^{-1/2} x_{it} \beta + \sigma_v \Omega^{-1/2} \varepsilon_{it} \quad ,$$

transformação que, de acordo com os pressupostos, pode ser reescrita,

$$y_{it} - \theta_1 \bar{y}_i - \theta_2 \bar{y}_t - \theta_3 \bar{y}_{..} = (x_{it} - \theta_1 \bar{x}_i - \theta_2 \bar{x}_t - \theta_3 \bar{x}_{..}) + \beta \varepsilon_{it} - \theta_1 \bar{\varepsilon}_i - \theta_2 \bar{\varepsilon}_t - \theta_3 \bar{\varepsilon}_{..} \quad .$$

Ora, acontece que a estimação *MQO* acaba a fazer-se sobre o modelo, $y_{it} - \theta_2 \bar{y}_t = (x_{it} - \theta_2 \bar{x}_t) \beta + (\varepsilon_{it} - \theta_2 \bar{\varepsilon}_t)$, uma vez que a componente da variância dos efeitos específicos das unidades seccionais representa 0% da variância total,¹⁷⁰ facto gerador de valores nulos para θ_1 e para θ_3 .

Neste cenário a transformação reveste apenas a forma de desvios relativamente às médias dos valores observados em cada um dos períodos de tempo, transformação *dentro* (*W*) e dado que, de acordo com a proposta de equivalência de Kang, testar a hipótese nula, $H_0 = E(\lambda_t | x_{it}) = E(\mu_i | x_{it}) = 0$, se suporta na diferença entre os estimadores de coeficientes pelo método generalizado dos mínimos quadrados, *MQG*,

¹⁷⁰ Importa referir que o facto de $\sigma_\mu = 0$, conduz a que a que o pacote estatístico STATA 9, aquando da solicitação directa do teste de Hausman, informe que o teste Wald a partir do teste Hausman não é válido, por o estimador dos efeitos aleatórios (*REM*) ter degenerado para um estimador *pooled-MQO*. Facto que induziu outro tipo de teste para se obter o valor referido no corpo do texto desta secção.

Importa referir que, pese embora o facto do coeficiente estimado para a variável *APRES* não ser estatisticamente significativa, decidiu-se incluir esta variável no modelo. O comportamento desta variável, *APRES*, face à variável *APART*, aquando dos testes de causalidade à Granger e das especificações restritas ensaiadas para ajustar a regressão, deixam em aberto essa possibilidade, e, ainda, a interacção assumida entre estas variáveis permite um potencial de correcção ao efeito o incumbente é reeleito suportado num apoio partidário diferente.

É oportuno, ensaiar-se a estimação da probabilidade do partido incumbente vencer as eleições no acto eleitoral em que se submete a escrutínio enquanto partido que está a ocupar a “cadeira” do poder autárquico, para tanto procurar-se-á adaptar a perspectiva de Chappell-Keech¹⁷¹ a este caso concreto na secção que se segue.

4.6. Estimação da Probabilidade do Incumbente Vencer as Eleições

O primeiro passo a concretizar para responder ao objectivo de estimação da probabilidade do incumbente vencer as eleições traduz-se na obtenção dos valores da matriz variância-covariância dos coeficientes estimados, a partir dos resultados da regressão que melhor se ajusta ao estudo desenvolvido, expressa por,

	β_0	β_1	β_2	β_3	β_4	β_5	β_6
β_0	3,2558						
β_1	-0,0591	0,0013					
β_2	-0,0239	-0,0002					
β_3	0,0030	-0,0003	-0,0014	0,0039			
β_4	-0,0004	-4,00E-06	-6,39E-06	-7,94E-06	3,64E-06		
β_5	0,0245	-0,0009	0,002	-0,0011	-6,14E-06	0,0145	
β_6	-0,0134	0,0004	-0,0009	0,0004	-5,22E-06	-0,0062	0,0034

¹⁷¹ Chappell e Keech (1988) estimaram a probabilidade do partido Democrático ganhar as eleições num dado ano, conhecendo a popularidade do partido.

Os valores constantes da matriz variância-covariância dos coeficientes estimados e o erro padrão da regressão estimada, permitem calcular a variância do erro de previsão, analiticamente, traduzida por,

$$\begin{aligned}
\hat{\sigma}_f^2 = & \text{var}(\hat{\beta}_0) + VOTOSANT^2 * \text{var}(\hat{\beta}_1) + APART^2 * \text{var}(\hat{\beta}_2) + APRES^2 * \text{var}(\hat{\beta}_3) + \\
& + DIREAL^2 * \text{var}(\hat{\beta}_4) + PDESEMP1^2 * \text{var}(\hat{\beta}_5) + PINFLAC1^2 * \text{var}(\hat{\beta}_6) + \\
& + 2 * VOTOSANT * \text{cov}(\hat{\beta}_0, \hat{\beta}_1) + 2 * APART * \text{cov}(\hat{\beta}_0, \hat{\beta}_2) + 2 * APRES * \text{cov}(\hat{\beta}_0, \hat{\beta}_3) + \\
& + 2 * DIREAL * \text{cov}(\hat{\beta}_0, \hat{\beta}_4) + 2 * PDESEMP1 * \text{cov}(\hat{\beta}_0, \hat{\beta}_5) + 2 * PINFLAC1 * \text{cov}(\hat{\beta}_0, \hat{\beta}_6) + \\
& + 2 * VOTOSANT * APART * \text{cov}(\hat{\beta}_1, \hat{\beta}_2) + 2 * VOTOSANT * APRES * \text{cov}(\hat{\beta}_1, \hat{\beta}_3) + \\
& + 2 * VOTOSANT * DIREAL * \text{cov}(\hat{\beta}_1, \hat{\beta}_4) + 2 * VOTOSANT * PDESEMP1 * \text{cov}(\hat{\beta}_1, \hat{\beta}_5) + \\
& + 2 * VOTOSANT * PINFLAC1 * \text{cov}(\hat{\beta}_1, \hat{\beta}_6) + 2 * APART * APRES * \text{cov}(\hat{\beta}_2, \hat{\beta}_3) + \\
& + 2 * APART * DIREAL * \text{cov}(\hat{\beta}_2, \hat{\beta}_4) + 2 * APART * PDESEMP1 * \text{cov}(\hat{\beta}_2, \hat{\beta}_5) + \\
& + 2 * APART * PINFLAC1 * \text{cov}(\hat{\beta}_2, \hat{\beta}_6) + 2 * APRES * DIREAL * \text{cov}(\hat{\beta}_3, \hat{\beta}_4) + \\
& + 2 * APRES * PDESEMP1 * \text{cov}(\hat{\beta}_3, \hat{\beta}_5) + 2 * APRES * PINFLAC1 * \text{cov}(\hat{\beta}_3, \hat{\beta}_6) + \\
& + 2 * DIREAL * PDESEMP1 * \text{cov}(\hat{\beta}_4, \hat{\beta}_5) + 2 * DIREAL * PINFLAC1 * \text{cov}(\hat{\beta}_4, \hat{\beta}_6) + \\
& + 2 * PDESEMP1 * PINFLAC1 * \text{cov}(\hat{\beta}_5, \hat{\beta}_6) + \hat{\sigma}^2
\end{aligned}$$

Onde $\hat{\sigma}_f^2$ é a variância estimada do erro de previsão e $\hat{\sigma}^2$ é uma estimativa da variância do termo erro da regressão (como seja, por exemplo, o quadrado do erro padrão).

A probabilidade de vitória do partido que, à data de eleições, está no poder, assumindo que são necessárias 43,8 % das intenções de voto para se vencer as eleições¹⁷² e que a percentagem de votos do incumbente ajustada a partir da regressão segue uma distribuição normal, vem expressa por,

$$F \left[\frac{(VV1f - 0,43822)}{\hat{\sigma}_f^2} \right],$$

em que $VV1f$ é o valor da percentagem de votos ajustada do incumbente, obtida a partir da regressão.

Os valores obtidos, município a município, são os constantes do apêndice **H**. Em termos globais, esses valores relativos à probabilidade do incumbente vencer as eleições parecem sustentáveis com o que a realidade permite intuir; um valor médio da probabilidade de reeleição na casa dos 52,2% , num intervalo de variação compreendido entre 68,5% e 39,9% , com um desvio padrão de 4,16, vai de encontro à ideia generalizada de que “estar no poder é quase garante de reeleição”, o chamado enviesamento do incumbente.

Estimada que foi a probabilidade de um autarca vir a ser (re)eleito, **Pr**, importa verificar se esta variável, enquanto variável exógena, apresenta potencial explicativo para a existência (ou não) de ciclos político-económicos, o que será objecto da secção que se segue.

¹⁷² A realidade eleitoral, em Portugal, permite concluir que o partido vencedor nas eleições é o partido mais votado, não importando o valor percentual alcançado. Mais, tendencialmente, as duas forças políticas, PS e PSD, têm recolhido conjuntamente cerca de 65% a 75% dos votos em qualquer acto eleitoral, daí que em termos de análise empírica se possa ensaiar um estudo restringido a dois partidos. Ora, ao pretender-se fixar um valor necessário para que se vença as eleições, estas particularidades deveriam vir equacionadas. Ainda que se pudesse apostar no conceito de voto *normal*, já que a distribuição normal ajusta bem a distribuição de votos, optou-se por construir um índice, a partir da informação da base de dados que informa da percentagem de vereadores do mesmo partido que o presidente da câmara. Por via deste índice, $I = \frac{51\%}{\%Perc_mand}$, ponderam-se os valores obtidos nas eleições pelos partidos vencedores. Os valores, então obtidos, podem-se considerar os resultados eleitorais filtrados e a partir da respectiva média chega-se ao valor 43.8 %, que se apresenta mais razoável, já que equaciona o resultado eleitoral mas também o potencial de governabilidade do executivo camarário.

4.7. Identificação de Ciclos Político-Económicos a Nível Local

Responder ao objectivo a que se propõe esta secção impõe que a probabilidade estimada em 4.6. passe a ser equacionada num modelo em que a variável dependente se prende com o nível de endividamento do político no poder e cujo desenvolvimento se fundamenta num modelo auto-regressivo. Relativamente à variável dependente, o nível de endividamento vem expresso a partir do conteúdo da rubrica de encargos financeiros das autarquias em termos *per capita*,¹⁷³ traduzido pela notação, $CFcapita$.

A aplicação do modelo auto-regressivo, base de inspiração da modelização proposta, ao ajustar-se à base de dados em análise, encontra restrições de informação decorrentes da série temporal curta e não uniforme que caracteriza os momentos eleitorais autárquicos,¹⁷⁴ o que impele a que se arrisquem adaptações para as séries de desfasamentos iguais ou superiores a dois períodos da variável nível de endividamento, de forma a vir minimizado o potencial efeito “*missing values*” gerado pelos desfasamentos numa série temporal limitada apenas a 7 períodos eleitorais. Nesta perspectiva, elege-se a taxa de variação de encargos financeiros entre 2 períodos sucessivos de um ciclo eleitoral e/ou a taxa de variação de encargos financeiros entre um dado período e o valor médio de encargos financeiros de cada um dos ciclos eleitorais, para suprir eventuais problemas decorrentes dos níveis de encargos financeiros desfasados a dois ou mais períodos.

Modelando a probabilidade de reeleição, Pr_{it} , na perspectiva de um modelo auto-regressivo adaptado, conforme sugerido, no parágrafo anterior, para a variável $CFcapita$, ter-se-á,

¹⁷³ Uma vez que o mercado de capitais a que podem recorrer os diversos municípios é o mesmo, parece aceitável considerar que não há variação espacial de preços, o que faz crer que esta variável pode constituir uma razoável *proxy* de desempenho técnico do incumbente.

¹⁷⁴ Importa reportar que a base se inicia com o momento (1979) do segundo acto eleitoral autárquico, o que impossibilita informação relativamente a desfasamentos de $CFcapita_{79}$ e, ainda, que o intervalo entre os dois primeiros actos eleitorais da base de dados é de 3 anos, o que faz com que $CFcapita_{82,3}$ coincida com $CFcapita_{79}$ (ano de eleições) e $CFcapita_{85,3}$ coincida com $CFcapita_{82}$ (ano de eleições).

$$CFcapita_{it} = \gamma_0 + \gamma_1 CFcapita_{it-\ell} + \gamma_2 Pr_{it} + \varepsilon_{it}$$

$$CFcapita_{it} = \gamma_0 + \gamma_1 \Delta Peso_{it-\ell} + \gamma_2 Pr_{it} + \varepsilon_{it} \quad (4.7.1)$$

$$CFcapita_{it} = \gamma_0 + \gamma_1 \Delta Média_{it-\ell} + \gamma_2 Pr_{it} + \varepsilon_{it}$$

A notação $CFcapita_{it-\ell}$ identifica em cada município i o nível de endividamento no momento, imediatamente, anterior ao ano de eleição e assim sucessivamente; a notação $\Delta Peso_{it-\ell}$ ¹⁷⁵ identifica em cada município i a taxa de variação do nível de encargos financeiros *per capita* verificada entre dois períodos consecutivos dentro de um mesmo ciclo eleitoral e a notação $\Delta Média_{it-\ell}$ ¹⁷⁶ identifica em cada município i a taxa de variação do nível de encargos financeiros *per capita* relativamente ao valor médio de encargos financeiros *per capita* do ciclo eleitoral nos sucessivos momentos ℓ que antecedem, o momento t do acto eleitoral, em cada ciclo eleitoral.

Uma vez que o modelo que suporta o cálculo da probabilidade de reeleição estimada consegue estabilizar a partir da regressão estimada pelo método dos estimadores dos mínimos quadrados generalizados, MQG_F , com assumpção de efeitos seccionais e de efeitos periódicos fixos, optou-se por expandir o método, à ocasião, adoptado, ao modelo auto-regressivo proposto. A prevalência, no entanto, de *missing values*, nomeadamente no que concerne aos valores desfasados a um período e aos valores médios da variável $CFcapita$, faz com que esta estimação só consiga vir compatível com um único pressuposto, o de efeitos fixos seccionais.

Os resultados decorrentes da estimação constam da tabela que se segue,

¹⁷⁵ $\Delta Peso_{it-1} = (CFcapita_{it} - CFcapita_{it-1}) / CFcapita_{it-1}$.

¹⁷⁶ $\Delta Média_{it-\ell} = (Percapita_{it-\ell} - Percapita_{Médiait}) / Percapita_{Médiait}$

	$MQGF^{(a)(b)(c)}$ White (1)	$MQGF^{(a)(b)(c)}$ White (2)	$MQGF^{(a)(b)(c)}$ White (3)	$MQGF^{(a)(b)(c)}$ White (1')	$MQGF^{(a)(b)(c)}$ White (2')	$MQGF^{(a)(b)(c)}$ White (3')
γ_1	0,6195*** (38,42)	0,6241*** (38,91)	0,4718*** (16,82)	-0,0014 (-0,87)	-0,0007 (-1,44)	-0,0196*** (-6,32)
γ_2				-0,0135*** (-3,13)	-0,0125*** (-2,67)	-0,06*** (-3,38)
γ_3	-0,1207*** (2,41)			-0,2432*** (-3,026)		
γ_4		0,6112* (1,68)	-0,6292 (-1,46)		0,4570** (2,31)	0,3373* (1,74)
γ_5			0,1965*** (5,21)			0,0194*** (6,22)
γ_6						0,048*** (2,64)
$R^2_{Ajust.}$	0,6681	0,5812	0,5890	0,2054	0,1098	0,1347
Teste Wooldridge	$F_{[1,274]} = 2,01$ ***	$F_{[1,274]} = 1,84$ ***	$F_{[1,274]} = 2,34$ ***	$F_{[1,263]} = 1,77$ ***	$F_{[1,263]} = 2,04$ ***	$F_{[1,263]} = 2,23$ ***
N°Obs.	1652	1652	1652	1425	1425	1425

	$MQGF^{(a)(b)(c)}$ White (1'')	$MQGF^{(a)(b)(c)}$ White (2'')	$MQGF^{(a)(b)(c)}$ White (3'')
γ_1	3,2415*** (18,77)	3,5246*** (22,69)	3,2453*** (9050)
γ_2	-0,2164*** (-2,17)	-0,3516*** (-4,24)	-0,5511*** (-2,69)
γ_3	-0,1810*** (-7,21)		
γ_4		0,37181*** (2,54)	0,4410** (2,39)
γ_5			0,3493 (0,89)
γ_6			0,2495 (1,07)
$R^2_{Ajust.}$	0,392778	0,311084	0,311537
Teste Wooldridge	$F_{[1,273]} = 5,26$ **	$F_{[1,273]} = 5,93$ **	$F_{[1,273]} = 5,19$ **
N°Obs.	1600	1600	1600

Tabela 4.7.1: Resultados das Regressões

Notas: ^(a) significa que a componente do erro foi considerada apenas a uma dimensão, a componente específica seccional; ^(b) estimação pelo método fiável dos mínimos quadrados generalizado, com ponderação seccional; ^(c) estimação com correcção *White diagonal* e entre parênteses os valores do *t*-estatístico. As notações, ***, **, reportam níveis de significância de 1% e de 5%, respectivamente, para a rejeição da hipótese nula ($\beta = 0$), significando, ainda, a omissão dessa notação a ausência de significância estatística para a rejeição da hipótese nula.

No teste de Wooldrige, *** traduz rejeição da hipótese de não auto-correlação de 1ª ordem ao nível de significância de 5% para todas as regressões à excepção das regressões (1'') (2'') e (3'') onde ** traduz nível de significância de 1%.

A sequência das regressões apresentadas segue a estrutura das equações que se seguem, assim,

$$E(1) \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \text{PERCapita}_{it-1} + \gamma_3 * \text{Pr}_{it} + \varepsilon_{it}$$

$$E(2) \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \text{PERCapita}_{it-1} + \gamma_4 * D_{it} + \varepsilon_{it}$$

$$E(3) \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \text{PERCapita}_{it-1} + \gamma_4 * D_{it} + \gamma_5 * \text{PERCapita}_{it-1} * D_{it} + \varepsilon_{it}$$

$$E(1') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Peso}_{it-1} + \gamma_2 * \Delta\text{Peso}_{it-2} + \gamma_3 * \text{Pr}_{it} + \varepsilon_{it}$$

$$E(2') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Peso}_{it-1} + \gamma_2 * \Delta\text{Peso}_{it-2} + \gamma_4 * D_{it} + \varepsilon_{it}$$

$$E(3') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Peso}_{it-1} + \gamma_2 * \Delta\text{Peso}_{it-2} + \gamma_4 * D_{it} + \gamma_5 * \Delta\text{Peso}_{it-1} * D_{it} + \gamma_6 * \Delta\text{Peso}_{it-2} * D_{it} + \varepsilon_{it}$$

$$E(1'') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Media}_{it-1} + \gamma_2 * \Delta\text{Media}_{it-2} + \gamma_3 * \text{Pr}_{it} + \varepsilon_{it}$$

$$E(2'') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Media}_{it-1} + \gamma_2 * \Delta\text{Media}_{it-2} + \gamma_4 * D_{it} + \varepsilon_{it}$$

$$E(3'') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Media}_{it-1} + \gamma_2 * \Delta\text{Media}_{it-2} + \gamma_4 * D_{it} + \gamma_5 * \Delta\text{Media}_{it-1} * D_{it} + \gamma_6 * \Delta\text{Media}_{it-2} * D_{it} + \varepsilon_{it}$$

$$\text{com, } \Delta\text{Peso}_{it-\ell} = (\text{Percapita}_{it-\ell} - \text{Percapita}_{it-\ell-1}) / \text{Percapita}_{it-\ell-1} \text{ e}$$

$$\Delta\text{Media}_{it-\ell} = (\text{Percapita}_{it-\ell} - \text{Percapita}_{\text{Media}_{it}}) / \text{Percapita}_{\text{Media}_{it}}$$

O objectivo deste ensaio era encontrar evidência empírica para as soluções de equilíbrio desenvolvidas no âmbito do capítulo 2, e para a respectiva concretização desenhou-se um modelo econométrico de forma a estimar-se o valor das probabilidades de reeleição. Ora é, precisamente essa variável, probabilidade estimada de reeleição, **Pr**, que está equacionada como variável explicativa nas equações (1) (1') e (1''). Decorrem das regressões (1) (1') e (1''), ajustadas a essas equações, uma coincidência entre o sinal (-), estatisticamente robusto, do coeficiente estimado para **Pr** e as inferências do modelo teórico. Resultados que motivaram a que se procurasse despistar o risco de coincidência virtuosa, para o efeito consideraram-se regressões sobre equações do tipo das supra-referidas, mas onde a variável **Pr** vem substituída por uma variável *dummy* ($D_{it} = 1$, caso se verifique reeleição e $D_{it} = 0$, caso não se verifique a reeleição). Esta expansão do modelo inicial coloca-nos perante um coeficiente estimado para a variável *dummy*

positivo, menos robusto estatisticamente, mas, ainda assim, recuperando significância estatística à medida que os valores dos encargos financeiros se equacionam em termos de taxa de variação (primeiro em relação ao valor do encargo financeiro do momento imediatamente anterior ao considerado e depois em relação ao encargo financeiro médio do ciclo entre períodos eleitorais). Ainda que não se procure a exaustividade, mesmo assim opta-se por considerar como variável explicativa um termo de interacção entre a variável encargos financeiros (seja desfasada, seja em termos de taxa de variação) e a variável reeleição D_{it} , de forma a captar-se eventual efeito *agência* do político no poder.

Dos resultados apurados o sinal (-) obtido para o coeficiente estimado γ_3 confirma as expectativas, na medida em que até sugere um perfil de político que é convergente com a perspectiva subjacente às soluções apuradas aquando do desenvolvimento do modelo teórico na subsecção 2.2., isto é o resultado empírico sustenta a solução de equilíbrio do modelo teórico. Decorre do modelo teórico, um equilíbrio, tendencialmente, favorável à definição de uma estratégia conducente à escolha de uma acção do tipo “politicamente correcto” – opção que suporta a medida do desempenho do político no poder, e nesse sentido o sinal (-) do coeficiente estimado γ_3 revela-se potencialmente interessante para reflexões neste domínio. O facto do coeficiente estimado γ_3 se revelar estatisticamente significativo permite intuir que, o político no poder, seja competente, seja incompetente, vai decidir a sua estratégia eleitoral em função da respectiva probabilidade de reeleição – quanto maior a certeza de se manter no poder nos próximos anos, menor o incentivo para comportamentos oportunistas. Os comportamentos oportunistas podem também ter a leitura de comportamentos arriscados, na perspectiva de que o político está disposto a correr mais riscos se não tiver muita certeza de que vai conseguir a reeleição, caso contrário esse comportamento de risco é racionalizado na medida em que o político sabe que no período pós-eleitoral as consequências desses comportamentos virão equacionadas nos respectivos orçamentos de gestão autárquica que eles têm de elaborar, fazer aprovar e executar.

Já no que diz respeito ao coeficiente γ_1 é interessante verificar o seu valor (+) e a sua relevância estatística nas regressões (1), (2), (3), (1’), (2’) e (3’). Nas primeiras três

regressões o coeficiente γ_1 reporta-se a encargos financeiros do ano imediatamente anterior ao ano de eleições e nas três últimas regressões o coeficiente γ_1 alude à taxa de variação desses encargos financeiros face à média de encargos financeiros do ciclo eleitoral, o que sugere um efeito positivo dos encargos financeiros do ano anterior ao das eleições no valor dos encargos financeiros *per capita* do ano de eleições. *A priori* este resultado faz sentido face ao conteúdo da rubrica em análise, uma vez que os encargos financeiros englobam juros devidos por endividamento face a terceiros é natural que, de um ano para o outro, ocorra um impacto positivo pela diversidade temporal das obrigações tomadas pelos municípios (curto/médio/longo prazo).

Nas regressões (1'), (2') e (3') o coeficiente γ_1 está associado à variável taxa de variação de encargos financeiros desfasados entre si de um período (o do ano eleitoral e o do ano imediatamente precedente) e o coeficiente γ_2 surge relacionado com a variável taxa de variação de encargos financeiros desfasados entre si de um período (mas agora entre os dois anos consecutivos antecedentes ao ano de eleições). Nesse contexto, o sinal (-) que apresentam pode ser susceptível de vir interpretado como traduzindo a gestão de uma agenda eleitoralista. Aliás o carácter regular, nas três regressões, da significância estatística do coeficiente estimado γ_2 pode corroborar esta conjectura, já que uma baixa taxa de variação dos encargos financeiros entre dois períodos sucessivos, desfasados de 3 e 2 anos relativamente ao ano de eleições, tem um impacto positivo sobre o encargo financeiro *per capita* do ano de eleições e vice-versa.

Uma observação dos coeficientes estimados para os termos de interacção entre as diversas variantes assumidas da variável encargos financeiros e a *dummy* reeleição (ou não), constantes das regressões (3), (3') e (3''), faculta informação para aceitarmos que, (i) o valor do encargo financeiro *per capita* do ano anterior à eleição e a taxa de variação desse encargo financeiro face ao encargo financeiro do ano transacto, para os autarcas reeleitos, explica positivamente o valor do encargo financeiro do ano das eleições e com significado estatístico (o que decorre do sinal e da significância estatística de γ_5 nas regressões (3) e (3')) e (ii) a taxa de variação desse encargo financeiro face ao encargo financeiro desfasado a dois períodos, para os autarcas reeleitos, continua a explicar positivamente o valor do encargo financeiro do ano das

eleições e com significado estatístico (conforme sinal e robustez estatística de γ_6 na regressão (3')). Resultado que pronuncia potencial capacidade de esses termos captarem o efeito de *agência*.

Finalmente o coeficiente estimado γ_4 da variável *dummy* de reeleição (ou não), apresenta-se, na generalidade das regressões, com sinal (+) e com significado estatístico dentro dos níveis convencionais de confiança, na maior parte das regressões. Facto que nos permite a leitura de que nos municípios onde o autarca foi reeleito o encargo *per capita* no ano eleitoral é superior ao dos restantes municípios.

Os resultados obtidos, para γ_3 e para γ_4 , permitem reflectir das vantagens de se governar com o conhecimento de funções de popularidade, uma *proxy* para avaliar as probabilidades de reeleição. Evidentemente que um cenário deste tipo apresenta *prós* na medida em que políticos, quer competentes, quer incompetentes, podem disciplinar a respectiva governação, por saberem que a probabilidade de se fazerem reeleger é elevada (eleitos e eleitores poderiam obter vantagens de uma governação mais responsável), mas também é fonte de *contras* já que pode alimentar decisões políticas irresponsáveis, face ao conhecimento de possibilidades escassas de reeleição.

4.8. Conclusões

Os resultados empíricos do modelo desenhado para estimar a percentagem de votos do autarca, base do ensaio para o cálculo da probabilidade de reeleição do autarca, apontam para, (i) uma relativa influência do coeficiente estimado para a variável *APART*, quando o modelo interage com a variável *APRES*; ou, então, (ii) a primazia do coeficiente estimado da variável *APRES*, caso se negligencie a variável *APART*, aquando da estimação. Motivo que, associado ao sinal dos coeficientes estimados e à significância estatística dos mesmos, sugere uma potencial correspondência com a premissa de que o votante é racionalmente ignorante. Premissa que, se alavancada nos valores estimados para o coeficiente estimado da variável *VOTOSANT*, permite assumir

os anos de permanência no poder, quer do partido, quer do político, como um activo determinante para a selecção do candidato a eleger.

A aparente vantagem decorrente de se estar no poder, do reconhecimento público da figura do autarca, muitas vezes indissociável do partido que o apoia, parece mais evidente nos resultados estimados a partir do modelo em que os valores das variáveis permanecem em estrutura agregada (*pool*), já que nesse cenário o coeficiente estimado da variável *DIREAL*, nem significância estatística apresenta. Mais, o aparente apelo para se negligenciar essa variável no modelo em estudo, produz resultados que pouco alterariam os resultados finais estimados.¹⁷⁷ Será suficiente esta constatação para se questionar a capacidade do votante eficiente? Para isso ter-se-ia que garantir que todas as variáveis relevantes tinham sido convocadas ao desenho do modelo e, embora o cuidado dispensado à especificação do modelo, estamos certos que existem áreas não cobertas por não se conseguir informação continuada e fiável, potenciada, ainda, pela extensão limitada da série temporal disponível.

Os valores obtidos para a probabilidade de reeleição do autarca, calculados, a partir da percentagem de votos, entretanto, estimada pelo tratamento empírico do modelo em estudo, apresentam um valor médio de **52,2%**, oscilando num intervalo entre **39,9%** e **68,5%**,¹⁷⁸ parecem não colidir muito com a sensibilidade existente sobre o que tem acontecido no país, em termos da escolha local dos governos autárquicos.

Finalmente, é curiosa a relação encontrada entre a variável probabilidade de reeleição e o nível de encargos de endividamento da autarquia, o sinal (-) do coeficiente estimado apresenta um potencial de reflexões a explorar, tanto mais pertinente quanto o sinal (+) estimado para o coeficiente da variável *dummy* de reeleição (= 1)/não reeleição (= 0) na relação com o nível de encargos de endividamento da autarquia. Esta discordância de sinais pode captar o efeito de responsabilidade/irresponsabilidade do autarca no poder caso conheça o valor da respectiva probabilidade de reeleição.

¹⁷⁷ Estimando o modelo em estudo, retirando a variável *DIREAL*, os resultados empíricos viriam em tudo semelhantes aos obtidos com o modelo original, um arredondamento desses coeficientes às décimas, conduziriam aos mesmos coeficientes.

¹⁷⁸ Com um desvio padrão de 0,04.

Assim, pode-se questionar, Afinal, o autarca acaba mesmo a escolher o comportamento politicamente correcto? Ou seja decide agir de forma oportunista! Será que não arrisca comportamentos mais activos (isto é estratégias de acção com consequências na futura gestão orçamental) quando tem mais convicções de ganhar? Ou ainda, será que o autarca teme a avaliação dos seus eleitores, e não quer ser penalizado? Isto é, o autarca receia que os seus eleitores possam perceber como irresponsáveis as acções decorrentes de uma opção estratégica mais vigorosa!

Este resultado empírico deve ser evidenciado, ainda, por ir ao encontro, (i) da percepção de que o político, não importa o tipo (competente ou incompetente), é aliciado, em cenário de informação assimétrica, a comportamentos oportunistas, para evitar leituras erradas por parte dos votantes (conforme exposto em **1.2.4**); (ii) da decisão do político, no curto prazo, de *não acção* como garantia da reeleição (conforme deduzido em **2.2.**); (iii) da proposição de Baleiras e Costa (2004) que um aumento da probabilidade (exógena) da reeleição não expande o ciclo político-económico e (iv) dos resultados recentes de Aidt, Veiga e Veiga (2007), que revelam existir uma relação inversa entre os comportamentos oportunistas do incumbente e as respectivas margens de vitória (num estudo que pretende fazer a ponte entre os ciclos político – económicos e as funções de popularidade, para a mesma base de dados).

Apêndice D

Descrição sumária dos testes de raiz unitária utilizados

No presente ensaio, utiliza-se o pacote estatístico EVIEWS 5, que disponibiliza os testes de raiz unitária Levin, Lin e Chu (*LLC*); Im, Pesaran e Shim (*IPS*); Dickey-Fuller Aumentado (*ADF*) e Hadri, entre outros; por isso importa fazer uma breve introdução a cada um desses testes, assim:

D.1. Testes Levin Lin e Levin, Lin e Chu (*LLC*)

As propriedades assintóticas da análise regressiva dos dados em painel têm sido deduzidas sob o pressuposto de que as séries temporais, de cada unidade seccional, no painel, são fracamente estacionárias, o que colide com o facto de grande parte das variáveis observadas nesses painéis apresentarem raízes unitárias. Num dos primeiros contributos neste domínio, Levin e Lin (1993) desenvolveram teoria assintótica para análise de regressões de dados em painel, sempre que a estacionariedade fraca¹⁷⁹ é violada pela presença de uma raiz unitária dentro de cada série temporal seccional. A estrutura da análise de Levin e Lin pode ser sumarizada na seguinte equação:

$$\Delta y_{i,t} = \alpha_i + \delta_i t + \theta_t + \rho_i y_{i,t-1} + \xi_{i,t}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

A qual permite equacionar os efeitos específicos das unidades/seccionais (α_i), de forma a controlar a heterogeneidade das unidades seccionais e, também, os efeitos específicos

¹⁷⁹ Um processo estocástico, x_t ; $t = 0, \pm 1, \dots$ é estacionário se a distribuição conjunta de x_t, \dots, x_{t+k} é a mesma da distribuição conjunta de x_0, \dots, x_k para todo o t e k (muitas vezes identificado como estacionariedade estrita). Um processo estocástico x_t ; $t = 0, \pm 1, \dots$ é estacionário fraco (ou de 2ª ordem) se, as respectivas, média e variância são constantes ao longo do tempo, isto é, $E(x_t) \equiv \mu$, para todo o t (isto é a média não depende do tempo) e $Cov(x_t, x_{t+k}) = Cov(x_0, x_k) \equiv C_x(k)$, é função apenas de k (e não depende de t).

temporais $(\theta_t)^{180}$ para evitar o problema da correlação em série e, ainda, os coeficientes de tendência (δ_i) . Entretanto o teste Levin, Lin e Chu, *LLC*, (2002) baseia-se em técnicas que removem as componentes de autocorrelação, bem como as componentes determinísticas, isto é a intercepção individual e a tendência individual, quando executa as padronizações relevantes.

A hipótese a testar é $H_0 : \rho_i = 0$, denominada hipótese nula, contra a hipótese alternativa, $H_A : \rho_i = \rho < 0$, para todo o i .

Levin e Lin retiraram duas conclusões a partir da sua análise: em primeiro lugar, este procedimento é mais poderoso do que os testes padrão das raízes unitárias baseados nas séries temporais seccionais; depois, e em segundo lugar, quando as dimensões, seccional e temporal, do painel crescem muito e de forma arbitrária ($T \rightarrow \infty, N \rightarrow \infty$) acontece que os estimadores da regressão de dados em painel e os *t-estatísticos* ficam restringidos a distribuições normais; passando a convergir com o número de períodos temporais a uma taxa mais rápida da que se verificaria se essa convergência se fizesse com o número de unidades seccionais.

D.2. Testes Im, Pesaram e Shin (*IPS*)

Posterior pesquisa, de Im, Pesaran e Shin, *IPS* daqui em diante, relaxa o pressuposto restritivo feito por Levin e Lin (1993) de que os valores de ρ_i são homogêneos. Facto que faz com que a hipótese nula e a hipótese alternativa sejam modificadas para:

$$H_0 : \rho_i = 0, \text{ para todo o } i, \quad \text{contra } H_A: \begin{cases} \rho_i < 0, & i = 1, 2, \dots, N_l, \\ \rho_i = 0, & i = N_{l+1}, N_{l+2}, \dots, N \end{cases}$$

¹⁸⁰ Os efeitos específicos temporais são tidos em conta nos testes de raiz unitária em painel, desprezando-se, para isso, dados como $\tilde{y}_{it} = y_{it} - \bar{y}_t$, em que \bar{y}_t é a média dos valores observados em cada um dos municípios, i , num dado período do tempo.

equaciona-se a possibilidade de algumas séries individuais (mas não todas) possuírem raiz unitária.

Im *et al.* (2003) propuseram um teste de verosimilhança baseado na média dos DF ¹⁸¹ estatísticos obtidos para cada unidade seccional no painel, denominando-o \bar{t} - teste. A estatística, é expressa por:

$$\bar{t}_{N,T} = \left(\frac{1}{N} \right) \sum_{i=1}^N t_{iT}$$

Onde t_{iT} é o DF estatístico da unidade seccional i no painel.

Primeiro, os autores assumem que os erros das regressões DF não estão correlacionados em série. Fixando que, para $T \rightarrow \infty$, seguido por $N \rightarrow \infty$, uma versão padrão de $\bar{t}_{N,T}$ estatístico converge em probabilidade para $N(0,1)$ traduzida por,

$$Z_{\bar{t}} = \frac{\sqrt{N} \{ \bar{t}_{NT} - E(t_{iT}) \}}{\sqrt{\text{var}(t_{iT})}} \rightarrow N(0,1),$$

onde $E(t_{iT})$ e $\text{var}(t_{iT})$ são a média e a variância de t_{iT} , respectivamente.

Ou de uma forma mais complexa, quando t_i difere entre unidades seccionais, tem-se,

¹⁸¹ DF , Dickey - Fuller. Os testes DF são obtidos, pela estimação, através de uma regressão, de um processo simples auto-regressivo, $y_t = \rho y_{t-1} + x_t' \delta + \varepsilon_t$, depois de se subtrair y_{t-1} a ambos os membros, traduzido por, $\Delta y_t = \alpha y_{t-1} + x_t' \delta + \varepsilon_t$, em que, $\alpha = \rho - 1$. As hipóteses, a nula e a alternativa, são: $H_0 : \alpha = 0$ e $H_A : \alpha < 0$. Um t -rácio pode ser utilizado para fazer a avaliação. Dickey e Fuller mostraram que sob a hipótese nula, esta estatística não segue a distribuição t -student convencional. O teste DF é apenas válido se as variáveis satisfazem um processo simples $AR(1)$.

$$Z_{\bar{t}} = \frac{\sqrt{N} \left\{ \bar{t}_{NT} - \frac{1}{N} \sum_{i=1}^N E(t_{iT}) \right\}}{\sqrt{\frac{1}{N} \sum_{i=1}^N \text{var}(t_{iT})}} \rightarrow N(0,1)$$

Os valores $E(t_{iT})$ e de $\text{var}(t_{iT})$ para os diferentes valores de t (s) estão presentes em Im *et al.* (2003, p.60).

As estatísticas do teste *IPS* revelam-se um suporte para a hipótese nula de não-estacionaridade. O teste *IPS* baseia-se no teste *ADF*, que se introduzirá a seguir, tomando a raiz unitária como hipótese nula.

D.3. Teste Dickey-Fuller Aumentado (*ADF*)

Um caso mais geral, em que os erros, da especificação padrão, de um processo simples auto-regressivo de grau 1 (*AR*(1)), definido por, $y_t = \rho y_{t-1} + x_t' \delta + \varepsilon_t$ ¹⁸², estão correlacionados em série com diferentes padrões de correlação em série entre unidades/seccionais é considerado na segunda parte de Im *et al.* (2003). Para corrigir os enviesamentos daí decorrentes o teste Dickey-Fuller aumentado (*ADF*) introduz, $p - 1$ termos desfasados,¹⁸³ assim ilustrados,

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \sum_{j=1}^{p_i} \rho_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

¹⁸² Em que y_t é a variável que se pretende explicar, no caso em estudo, ter-se-á *VV1*, percentagem de votos obtidos; *APRES*, anos como presidente; *APART* anos do partido no poder; *DIREAL*, despesas de investimento, *per capita*, em termos reais; x_t vector das variáveis explicativas que pode incluir a constante, ou a constante e tendência, ε_t , o elemento perturbação. Se $|\rho| \geq 1$, y_t é uma série não estacionária e a sua variância aumenta com o tempo e no limite tenderá para ∞ . Se $|\rho| < 1$, y_t é uma série estacionária.

¹⁸³ O número de desfasamentos é escolhido de forma a eliminar a autocorrelação dos resíduos e a minimizar o critério de informação de Akaike (*AIC*).

os t -estatísticos para testar $\beta_i = 0$ são agora função dos parâmetros de perturbação, $\rho_i = (\rho_{i1}, \rho_{i2}, \dots, \rho_{ip_i})'$ e p_i , isto é, $t_{iT} = t_{iT}(p_i, \rho_i)$ e $\bar{t}_{N,T} = \left(\frac{1}{N}\right) \sum_{i=1}^N t_{iT}(p_i, \rho_i)$. Quando T e N são suficientemente grandes é possível desenvolver testes assintóticos válidos.

Uma alternativa prática é completar os \bar{t} -estatísticos utilizando médias e variâncias de $t_{iT} = t_{iT}(p_i, 0)$, avaliados sob a hipótese de $\beta_i = 0$. O \bar{t} -estatístico padrão sob este pressuposto é,

$$W_{\bar{t}}(p, \rho) = \frac{\sqrt{N} \left\{ \bar{t}_{NT} - \frac{1}{N} \sum_{i=1}^N E[t_{iT}(p_i, 0) | \beta_i = 0] \right\}}{\sqrt{\frac{1}{N} \sum_{i=1}^N \text{var}[t_{iT}(p_i, 0) | \beta_i = 0]}} \rightarrow N(0,1)$$

Os valores da média e da variância para os diferentes valores de T e de p , obtidos por simulação estocástica, com 50 000 replicações, são disponibilizados por Im *et al.* (2003, Quadro 3).

Dado que nos testes de hipótese clássicos, a hipótese nula é sustentada, mesmo que haja forte evidência contra ela, é bastante comum nos testes de raiz unitária, utilizarem-se dois tipos de testes com duas diferentes hipóteses nulas para verificar se os resultados são robustos. Um desses testes é o teste *ADF*, que tem como hipótese nula a raiz unitária; e outro é o teste Kwiatkowski *et al.*, denominado *KPSS* doravante,¹⁸⁴ que tem a estacionaridade como hipótese nula, ou seja a hipótese deste teste é a tendência estacionária correspondente à hipótese de uma variância nula do caminho aleatório.

¹⁸⁴ A estatística *KPSS* tem por base os resíduos decorrentes de uma regressão pelo método linear dos mínimos quadrados (*MQO*) de y_t sobre as variáveis exógenas x_t , traduzida na expressão: $y_t = x_t' \delta + u_t$.

D.4. Teste de Hadri

A versão do teste *KPSS* para dados em painel foi proposta por Hadri (2000), com um teste de raiz unitária, cuja hipótese nula é estacionária. Dado que este teste se suporta nos resíduos de uma regressão pelo método linear dos mínimos quadrados ordinários (*MQO*) de y_{it} sobre uma constante, ou de y_{it} sobre uma constante e uma tendência, a sua modelização pode-se traduzir, analiticamente, em,

$$y_{it} = r_{it} + \varepsilon_{it}$$

ou

$$y_{it} = r_{it} + \beta_i t + \varepsilon_{it},$$

$$i = 1, \dots, N; \quad t = 1, \dots, T$$

com r_{it} a identificar um caminho aleatório, expresso por:

$$r_{it} = r_{i,t-1} + u_{it}, \quad t = 1, 2, \dots, T; \quad i = 1, 2, \dots, N$$

quer ε_{it} , quer u_{it} têm distribuições i.i.d., seja entre as unidades seccionais i , seja ao longo do tempo t , com $E[\varepsilon_{it}] = 0$ e $E[\varepsilon_{it}^2] = \sigma_\varepsilon^2 \geq 0$, $E[u_{it}] = 0$ e $E[u_{it}^2] = \sigma_u^2 \geq 0$.

Introduzindo o caminho aleatório no modelo, há que reescrevê-lo, vindo pois:

$$y_{it} = r_{i0} + e_{it}$$

ou

$$y_{it} = r_{i0} + \beta_i t + e_{it}$$

Para $i = 1, \dots, N$; $t = 1, \dots, T$.

Em que $e_{it} = \sum_{s=1}^t u_{is} + \varepsilon_{it}$, vindo, também,

$$E(e_{it}) = 0,$$

e

$$\begin{aligned} E(e_{it}, e_{js}) &= \min(t, s)\sigma_u^2 + \sigma_\varepsilon^2, & i = j, t = s \\ &= \min(t, s)\sigma_u^2, & i = j, t \neq s \\ &= 0, & i \neq j \end{aligned}$$

Se a série é estacionária, então $\sigma_u^2 = 0$ e nesse caso $e_{it} = \varepsilon_{it}$, nas outras situações $\sigma_u^2 \neq 0$.

Portanto, as hipóteses, a nula e a alternativa, do teste de Hadri, são: $H_0: \lambda = 0$, contra

$H_A: \lambda > 0$, onde $\lambda = \frac{\sigma_u^2}{\sigma_\varepsilon^2}$. Hadri (2000) demonstra que, para a hipótese nula da estacionaridade, o teste estatístico de um painel apresenta a seguinte distribuição limite se,

(1) equaciona apenas a regressão de y_{it} sobre uma constante,

$$Z_\tau = \frac{\sqrt{N}(LM_\tau - \frac{1}{6})}{\sqrt{\frac{1}{45}}} \rightarrow N(0,1);$$

ou se,

(2) equaciona a regressão de y_{it} sobre uma constante e a tendência

$$Z_{\tau} = \frac{\sqrt{N}(LM_{\tau} - \frac{1}{15})}{\sqrt{\frac{11}{6300}}} \rightarrow N(0,1)$$

Em que LM_{τ} é o multiplicador lagrangeano estatístico aplicado aos erros heterogêneos e correlacionados em série, já que é mais poderoso, para testar a hipótese nula da estacionariedade, expressa-se por,

$$LM_{\tau} = \frac{1}{N} \left(\sum_{i=1}^N \left(\frac{1}{T^2} \sum_{t=1}^T S_{it}^2 / \hat{\sigma}_{\varepsilon i}^2 \right) \right)$$

Em que $S_{it} = \sum_{s=1}^t \hat{\varepsilon}_{is}$ é a soma parcial dos resíduos da regressão linear simples e $\hat{\sigma}_{\varepsilon i}^2$ ¹⁸⁵ é

uma estimativa consistente de $\sigma_{\varepsilon i}^2$ sob a hipótese nula.

Quer o teste de Hadri, quer o teste *IPS* utilizam limites que têm subjacente que, $T \rightarrow \infty$, seguido por $N \rightarrow \infty$; isto é, pressupõem uma sequência assintótica. Se um base de dados em painel se caracteriza por uma dimensão temporal muito maior que a dimensão seccional, estes testes estão justificados, segundo Hadri (2000). Infelizmente os dados utilizados neste ensaio, têm um pequeno T relativamente ao N .

¹⁸⁵ Hadri equacionou $\hat{\sigma}_{\varepsilon i}^2 = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_{it}^2$, em vez de $\hat{\sigma}_{\varepsilon}^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{\varepsilon}_{it}^2$, para que a heteroscedasticidade entre unidades seccionais i viesse equacionada, tornando o teste mais poderoso.

Maddala and Wu (1999) propuseram a utilização de um teste de Fisher para testar a raiz unitária nos dados em painel. Este teste baseia-se nos p – valores de N – testes independentes e, conseqüentemente, podem ser utilizados como qualquer tipo de testes de raiz unitária. Esta é a sua maior vantagem quando comparado com os testes de raiz unitária de dados em painel, como os propostos por Im *et al.* (1997) e Pedroni (1997). O teste de Fisher segue uma distribuição do χ^2 com $2 * N$ graus de liberdade, em que N é o número de testes individuais. Entretanto, esta distribuição tem subjacente o pressuposto da independência. Quando há correlação entre as variáveis a ser estimadas a distribuição do teste de Fisher é desconhecida.¹⁸⁶

Apêndice E

Total dos municípios estacionários e não estacionários

	APART		APRES		DIREAL		VV1	
	const.	const.tend.	const.	const.tend.	const.	const.tend.	const.	const.tend.
I(1)	168	84	123	88	174	113	104	42
I(0)	107	191	152	187	101	162	171	233
	275	275	275	275	275	275	275	275

¹⁸⁶ Para equacionar este efeito Maddala and Wu (1999) sugerem o emprego de um método regressivo *bootstrap*.

Apêndice F

Transformação *dentro(Within)* e *entre(Between)* grupos

Como se referiu no corpo do texto, subsecção 4.3.1., a transformação *dentro(Within)* e *entre(Between)* grupos implica a pré-multiplicação do modelo original por uma matriz de transformação, neste cenário, observe-se,

Estimador, *dentro*, grupos

A transformação gerada pela matriz \mathbf{M}_d resulta no modelo,

$$\mathbf{M}_d \mathbf{y} = \boldsymbol{\beta}' \mathbf{M}_d \mathbf{x} + \mathbf{M}_d \boldsymbol{\varepsilon},$$

Ou em alternativa:

$$\mathbf{M}^0 \mathbf{y}_i = \mathbf{M}^0 \boldsymbol{\alpha}_i + \mathbf{M}^0 \mathbf{x}_i \boldsymbol{\beta} + \mathbf{M}^0 \boldsymbol{\eta}_i = \mathbf{M}^0 \mathbf{x}_i \boldsymbol{\beta} + \mathbf{M}^0 \boldsymbol{\eta}_i .$$

Em que $\mathbf{M}_d = \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$,¹⁸⁷ com $\mathbf{M}_d \mathbf{D} = \mathbf{0}$; em que \mathbf{I} é uma matriz identidade, $T * T$, e \mathbf{D} é uma matriz de variáveis *dummy* que toma, em cada uma das i -ésima observações seccionais, o valor de ι , expressa por, $\mathbf{D} = \mathbf{I}_n \otimes \iota_T = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n]$.¹⁸⁸

¹⁸⁷ \mathbf{M}_d é a matriz que produz os resíduos Frish-Waugh-Lovell. $\mathbf{M}_d \mathbf{x}$ é a matriz em que cada uma das colunas é o vector de resíduos que decorre de se ajustar uma regressão de \mathbf{x} sobre \mathbf{D} e $\mathbf{M}_d \mathbf{y}$ é a matriz em que cada coluna é um vector de resíduos resultante de se regredir \mathbf{y} sobre \mathbf{D} .

¹⁸⁸ \otimes , identifica um produto Kronecker, que se define como um produto de duas matrizes $\mathbf{A}_{p*q} = \{a_{ij}\}$ e $\mathbf{B}_{m*n} = \{b_{kj}\}$, em que,

$$\mathbf{A}_{p*q} \otimes \mathbf{B}_{m*n} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1q}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2q}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}\mathbf{B} & a_{p2}\mathbf{B} & \dots & a_{pq}\mathbf{B} \end{bmatrix}. \text{ Verificam-se, ainda, as seguintes regras, } (\mathbf{A}_{p*q} \otimes \mathbf{B}_{m*n})' = \mathbf{A}'_{p*q} \otimes \mathbf{B}'_{m*n};$$

$$(\mathbf{A}_{p*q} \otimes \mathbf{B}_{m*n})(\mathbf{C}_{q*r} \otimes \mathbf{B}_{n*s}) = \mathbf{AC} \otimes \mathbf{BD} \text{ e } (\mathbf{A}_{p*p} \otimes \mathbf{B}_{m*m})^{-1} = \mathbf{A}^{-1}_{p*p} \otimes \mathbf{B}^{-1}_{m*m} .$$

Observe-se o facto das colunas da matriz \mathbf{D} serem ortogonais, do que resulta, $\mathbf{d}_i' \mathbf{d}_j = 0$, para $i \neq j$, implicando que a matriz de transformação \mathbf{M}_d possa vir traduzida por,

$$\mathbf{M}_d = \begin{bmatrix} \mathbf{M}^0 & 0 & \dots & 0 \\ 0 & \mathbf{M}^0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{M}^0 \end{bmatrix} \quad \mathbf{M}^0 = \mathbf{I}_T - \frac{1}{T} \mathbf{u}'$$

Aplicando sobre esse modelo transformado o método dos mínimos quadrados, *MQO*, resulta um estimador,

$$\hat{\beta}_{MQVD} = \hat{\beta}_{CV} = \hat{\beta}_W = \left[\sum_{i=1}^N \mathbf{x}_i' \mathbf{M}^0 \mathbf{x}_i \right]^{-1} \left[\sum_{i=1}^N \mathbf{x}_i' \mathbf{M}^0 \mathbf{y}_i \right] = \left[\mathbf{x}' \mathbf{M}_d \mathbf{x} \right]^{-1} \left[\mathbf{x}' \mathbf{M}_d \mathbf{y} \right]^{189}$$

Ou, simplesmente,

$$\hat{\beta}_W = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)' (x_{it} - \bar{x}_i) \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)' (y_{it} - \bar{y}_i) \right] \quad (\text{F.1})$$

Estimador, *entre*, grupos

A transformação gerada pela matriz \mathbf{P} resulta no modelo,

¹⁸⁹ A matriz de variância-covariância para o estimador $\hat{\beta}_{MQVD}$ é $\text{Var}(\hat{\beta}) = \sigma_\eta^2 \left[\mathbf{x}' \mathbf{M}_d \mathbf{x} \right]^{-1}$, onde σ_η^2 é estimado por,

$$s^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{\alpha} - x_{kit} \hat{\beta})^2}{NT - N - K} = (\mathbf{y} - \mathbf{x}\hat{\beta})' \mathbf{M}_d (\mathbf{y} - \mathbf{x}\hat{\beta}) / (NT - N - K).$$

$$\mathbf{Py} = \alpha + \boldsymbol{\beta}'\mathbf{Px} + \mathbf{P}\varepsilon \quad ^{190}$$

Em que, $\mathbf{P} = \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$, sendo \mathbf{D} uma matriz de variáveis *dummy* que toma, em cada uma das i -ésima observações seccionais, o valor de t , para captar os efeitos específicos individuais, caso se assumam como parâmetros fixos. Tem-se por isso $\mathbf{P} = \left[\mathbf{I}_n \otimes \frac{1}{T} \mathbf{ii}' \right]$.

Aplicando sobre esse modelo transformado o método dos mínimos quadrados, *MQO*, resulta um estimador,

$$\hat{\beta}_B = \left[\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{x}_i - \bar{\bar{x}})' \right]^{-1} \left[\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{y}_i - \bar{\bar{y}})' \right] \quad ^{191} \quad (\text{F.2})$$

Entretanto, cada um dos termos deste estimador *entre(Between)* grupos, **(F.2)**, pode vir expresso como,

$$\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{x}_i - \bar{\bar{x}})' = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{\bar{x}})'(x_{it} - \bar{\bar{x}}) \right] - \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)'(x_{it} - \bar{x}_i) \right]$$

e

$$\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{y}_i - \bar{\bar{y}})' = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{\bar{x}})'(y_{it} - \bar{\bar{y}}) \right] - \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)'(y_{it} - \bar{y}_i) \right]$$

¹⁹⁰ A regressão não se liberta das intercepções específicas individuais, assim a possibilidade dessa intercepções estarem correlacionadas com alguns dos regressores na equação permite um estimador *entre* grupos enviesado e inconsistente.

¹⁹¹ A multiplicação pela matriz transformação, conduz a $\mathbf{Py} = \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$, $\mathbf{Px} = \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}$ e $\bar{\alpha}_i = \alpha_i$.

Caso se equacione o estimador *dentro(Within)* grupos, **(F.1)**, e se observe o coeficiente β , estimado pelo método dos mínimos quadrados ordinário, *MQO*, em domínio *pooled*, verificar-se-á que tem implícito uma soma ponderada dos estimadores *dentro(Within)* e *entre(Between)* grupos, assim,

$$\hat{\beta}_{MQO} = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})'(x_{it} - \bar{x}) \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})'(y_{it} - \bar{y}) \right]$$

Que se pode decompor em,

$$\begin{aligned} \hat{\beta}_{MQO} = & \left[\left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)'(x_{it} - \bar{x}_i) \right] + \left[\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{x}_i - \bar{\bar{x}})' \right] \right]^{-1} \\ & * \left[\left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)'(y_{it} - \bar{y}_i) \right] + \left[\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{y}_i - \bar{\bar{y}})' \right] \right] \end{aligned}$$

O que é passível da seguinte tradução,

$$\hat{\beta}_{MQO} = \Delta_1 \hat{\beta}_W + (1 - \Delta_1) \hat{\beta}_B$$

Com,

$$\Delta_1 = \frac{\left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)'(x_{it} - \bar{x}_i) \right]}{\left[\left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)'(x_{it} - \bar{x}_i) \right] + \left[\sum_{i=1}^N T(\bar{x}_i - \bar{\bar{x}})(\bar{x}_i - \bar{\bar{x}})' \right] \right]}$$

Apêndice G

Programa em EVIEWS 5 para o teste HAUSMAN

```
workfile exag 1
' set sample
smpl @first 13

' estimate fixed effects and store results
pool01.ls(cx=f) vv1? apres? apart? direal?
vector beta = pool01.@coefs
matrix covar = pool01.@cov

' keep only slope coefficients
vector b_fixed = @subextract(beta,2,1,4,1)
matrix cov_fixed = @subextract(covar,2,2,4,4)

' estimate random effects and store results
pool01.ls(cx=r) vv1? apres? apart? direal?
beta = pool01.@coefs
covar = pool01.@cov

' keep only slope coefficients
vector b_gls = @subextract(beta,2,1,4,1)
matrix cov_gls = @subextract(covar,2,2,4,4)

' compute Hausman test stat
matrix b_diff = b_fixed - b_gls
matrix var_diff = cov_fixed - cov_gls
matrix qform = @transpose(b_diff)*@inverse(var_diff)*b_diff

if qform(1,1)>=0 then
    ' set table to store results
    table(4,2) result
    setcolwidth(result,1,20)
    setcell(result,1,1,"Hausman test for fixed versus random effects")
    setline(result,2)

    !df = @rows(b_diff)
    setcell(result,3,1,"chi-sqr(" + @str(!df) + ") = ")
    setcell(result,3,2,qform(1,1))
    setcell(result,4,1,"p-value = ")
    setcell(result,4,2,1-@cchisq(qform(1,1),!df))
    setline(result,5)

    show result
else
    statusline "Quadratic form is negative"
endif
```

Apêndice H

Valores das Probabilidades de Reeleição por Município

	munic	probab.⁽¹⁾	probab.⁽²⁾		munic	probab.⁽¹⁾	probab.⁽²⁾		munic	probab.⁽¹⁾	probab.⁽²⁾		
	Alcoutim	0,53666			Gavião	0,56642			Manteigas	0,50994			
	Alcoutim	0,48983	0,93182		Gavião	0,51149	0,81259		Manteigas	0,49037	0,64922		
	Alcoutim	0,52031	0,78711		Gavião	0,55175	0,74249		Manteigas	0,51631	0,89723		
	Alcoutim	0,52028	0,69769		Gavião	0,56554	0,72862		Manteigas	0,60073	0,96810		
	Alcoutim	0,51587	0,60961		Gavião	0,55156	0,42741		Manteigas	0,52707	0,75364		
	Alcoutim	0,48051	0,95972		Gavião	0,60541	0,68627		Manteigas	0,48131	0,90647		
	Alcoutim	0,48777	0,85288		Gavião	0,59276	0,29863		Manteigas	0,47632	0,70315		
	Aljezur	0,54832			Góis	0,52012			Fronteira	0,53392			
	Aljezur	0,52348	0,90956		Góis	0,51680	0,87819		Fronteira	0,56870	0,95000		
	Aljezur	0,53298	0,73010		Góis	0,49570	0,86978		Fronteira	0,56072	0,81739		
	Aljezur	0,53080	0,69930		Góis	0,54493	0,90939		Fronteira	0,49135	0,83210		
	Aljezur	0,48305	0,93259		Góis	0,50594	0,66281		Fronteira	0,45234	0,81058		
	Aljezur	0,49651	0,86119		Góis	0,53207	0,75360		Fronteira	0,46963	0,90510		
	Aljezur	0,49659	0,65514		Góis	0,58918	0,65049		Fronteira	0,53751	0,88347		
	Crato	0,52953			Sátão	0,54256			Monforte	0,49128			
	Crato	0,46255	0,78076		Sátão	0,58647	0,94315		Monforte	0,54194	0,93476		
	Crato	0,50603	0,69700		Sátão	0,50059	0,71341		Monforte	0,51676	0,69127		
	Crato	0,51519	0,73430		Sátão	0,52930	0,68498		Monforte	0,50983	0,51747		
	Crato	0,45716	0,20827		Sátão	0,50324	0,33357		Monforte	0,49395	0,28012		
	Crato	0,47104	0,93689		Sátão	0,49021	0,91086		Monforte	0,49679	0,19064		
	Crato	0,47058	0,92899		Sátão	0,50274	0,75189		Monforte	0,49807	0,94843		
	Cuba	0,58166			Vila Viçosa	0,52521			Alvito	0,49207			
	Cuba	0,58103	0,94805		Vila Viçosa	0,51304	0,92078		Alvito	0,47659	0,94235		
	Cuba	0,59268	0,89865		Vila Viçosa	0,54266	0,87061		Alvito	0,48362	0,86097		
	Cuba	0,59793	0,82792		Vila Viçosa	0,50722	0,57402		Alvito	0,49221	0,93822		(¹) de acordo c/Cap4
	Cuba	0,55290	0,45269		Vila Viçosa	0,46815	0,90196		Alvito	0,42701	0,71246		
	Cuba	0,54757	0,23549		Vila Viçosa	0,42280	0,76192		Alvito	0,39896	0,78957		(²) de acordo c/Cap5
	Cuba	0,50997	0,92400		Vila Viçosa	0,42972	0,82338		Alvito	0,45466	0,81472		
	FreixoEspadaCinta	0,54117			Mesão Frio	0,56358			Barrancos	0,55187			
	FreixoEspadaCinta	0,48608	0,63372		Mesão Frio	0,53503	0,79365		Barrancos	0,58352	0,99034		Municípios com
	FreixoEspadaCinta	0,52649	0,76347		Mesão Frio	0,57290	0,85353		Barrancos	0,60150	0,98305		População<=5000
	FreixoEspadaCinta	0,50058	0,37383		Mesão Frio	0,55221	0,55629		Barrancos	0,60575	0,90955		
	FreixoEspadaCinta	0,50171	0,93520		Mesão Frio	0,48307	0,18919		Barrancos	0,53251	0,45846		
	FreixoEspadaCinta	0,53363	0,95153		Mesão Frio	0,56542	0,31816		Barrancos	0,50577	0,93944		
	FreixoEspadaCinta	0,50663	0,92686		Mesão Frio	0,57747	0,14205		Barrancos	0,50450	0,86208		continua

	munic	probab.⁽¹⁾	probab.⁽²⁾		munic	probab.⁽¹⁾	probab.⁽²⁾		munic	probab.⁽¹⁾	probab.⁽²⁾		
	Alcochete	0,51171			Almeida	0,58714			Castro Marim	0,49038			
	Alcochete	0,52650	0,95499		Almeida	0,63159	0,95805		Castro Marim	0,47011	0,82279		
	Alcochete	0,52244	0,86263		Almeida	0,63383	0,93840		Castro Marim	0,53354	0,78619		
	Alcochete	0,54203	0,83502		Almeida	0,52396	0,42435		Castro Marim	0,48426	0,55032		
	Alcochete	0,50205	0,53039		Almeida	0,47046	0,83087		Castro Marim	0,51803	0,51438		
	Alcochete	0,50863	0,27874		Almeida	0,49800	0,90555		Castro Marim	0,52337	0,36823		
	Alcochete	0,52108	0,13139		Almeida	0,57320	0,91491		Castro Marim	0,49465	0,96101		
	Ferreira do Alentejo	0,53722			Alvaiázere	0,50248			Castro Verde	0,58519			
	Ferreira do Alentejo	0,55761	0,94093		Alvaiázere	0,61642	0,98020		Castro Verde	0,61164	0,97645		
	Ferreira do Alentejo	0,54401	0,82524		Alvaiázere	0,56139	0,91760		Castro Verde	0,61184	0,93045		
	Ferreira do Alentejo	0,53420	0,63778		Alvaiázere	0,52095	0,90490		Castro Verde	0,60528	0,86281		
	Ferreira do Alentejo	0,50271	0,28906		Alvaiázere	0,59170	0,96314		Castro Verde	0,57822	0,54424		
	Ferreira do Alentejo	0,51347	0,95876		Alvaiázere	0,60657	0,92655		Castro Verde	0,54508	0,21215		
	Ferreira do Alentejo	0,51455	0,84438		Alvaiázere	0,61253	0,77216		Castro Verde	0,56517	0,09618		
	Ferreira do Zêzere	0,50562			Armamar	0,61365			Castanheira de Pêra	0,56758			
	Ferreira do Zêzere	0,60889	0,94683		Armamar	0,61263	0,92226		Castanheira de Pêra	0,54495	0,91764		
	Ferreira do Zêzere	0,59030	0,90124		Armamar	0,53958	0,70009		Castanheira de Pêra	0,58283	0,85835		
	Ferreira do Zêzere	0,59776	0,74172		Armamar	0,54601	0,89504		Castanheira de Pêra	0,59121	0,82240		
	Ferreira do Zêzere	0,53586	0,41630		Armamar	0,52136	0,80087		Castanheira de Pêra	0,49990	0,92430		
	Ferreira do Zêzere	0,50477	0,16686		Armamar	0,57207	0,82887		Castanheira de Pêra	0,53682	0,97476		
	Ferreira do Zêzere	0,55130	0,09653		Armamar	0,62060	0,77875		Castanheira de Pêra	0,60275	0,97448		
	Melgaço	0,49097			Belmonte	0,49496			Avis	0,56375			
	Melgaço	0,50051	0,80806		Belmonte	0,42791	0,68849		Avis	0,55484	0,93540		
	Melgaço	0,54199	0,92418		Belmonte	0,46165	0,78068		Avis	0,58907	0,90325		
	Melgaço	0,58926	0,95593		Belmonte	0,45603	0,88595		Avis	0,59474	0,82767		(1) de acordo c/Cap4
	Melgaço	0,57140	0,86472		Belmonte	0,42714	0,83774		Avis	0,56760	0,53995		
	Melgaço	0,57724	0,77687		Belmonte	0,47144	0,93416		Avis	0,56328	0,27237		(2) de acordo c/Cap5
	Melgaço	0,61659	0,77550		Belmonte	0,48089	0,84835		Avis	0,57836	0,17289		
	Mondim de Basto	0,53357			Celorico da Beira	0,58470			Vidigueira	0,51087			
	Mondim de Basto	0,46907	0,55461		Celorico da Beira	0,61859	0,93489		Vidigueira	0,52528	0,95973		Municípios com
	Mondim de Basto	0,49550	0,63413		Celorico da Beira	0,52230	0,58829		Vidigueira	0,58073	0,95902		População<=10000
	Mondim de Basto	0,57572	0,81375		Celorico da Beira	0,57626	0,57986		Vidigueira	0,56793	0,90261		e População>5000
	Mondim de Basto	0,50994	0,31858		Celorico da Beira	0,58036	0,50852		Vidigueira	0,53518	0,66245		
	Mondim de Basto	0,50729	0,13658		Celorico da Beira	0,52818	0,95436		Vidigueira	0,49842	0,23689		
	Mondim de Basto	0,50092	0,89125		Celorico da Beira	0,54250	0,90468		Vidigueira	0,49655	0,10611		continua

	munic	probab.⁽¹⁾	probab.⁽²⁾		munic	probab.⁽¹⁾	probab.⁽²⁾		munic	probab.⁽¹⁾	probab.⁽²⁾		
	Lamego	0,48565			Lagos	0,51384			Idanha-a-Nova	0,50643			
	Lamego	0,49772	0,66422		Lagos	0,48930	0,82038		Idanha-a-Nova	0,49848	0,83880		
	Lamego	0,48097	0,50395		Lagos	0,49586	0,54801		Idanha-a-Nova	0,56126	0,84826		
	Lamego	0,44121	0,69541		Lagos	0,46050	0,34404		Idanha-a-Nova	0,59357	0,88768		
	Lamego	0,42217	0,71876		Lagos	0,49415	0,93182		Idanha-a-Nova	0,58126	0,76248		
	Lamego	0,46118	0,72234		Lagos	0,46951	0,71868		Idanha-a-Nova	0,58429	0,60520		
	Lamego	0,48850	0,59249		Lagos	0,48610	0,62209		Idanha-a-Nova	0,49079	0,94511		
	Anadia	0,54589			Pinhel	0,51234			Alcanena	0,47772			
	Anadia	0,54991	0,80310		Pinhel	0,60466	0,91610		Alcanena	0,47813	0,76350		
	Anadia	0,55696	0,76173		Pinhel	0,52336	0,60869		Alcanena	0,50059	0,87258		
	Anadia	0,57104	0,57107		Pinhel	0,48285	0,71834		Alcanena	0,46138	0,72349		
	Anadia	0,52871	0,29232		Pinhel	0,48818	0,64707		Alcanena	0,53312	0,82583		
	Anadia	0,53726	0,13879		Pinhel	0,48769	0,89958		Alcanena	0,50822	0,54790		
	Anadia	0,55762	0,09938		Pinhel	0,54625	0,83606		Alcanena	0,53617	0,42923		
	Arcos de Valdevez	0,55085			Portalegre	0,51642			Chamusca	0,53240			
	Arcos de Valdevez	0,58399	0,87757		Portalegre	0,47249	0,70956		Chamusca	0,48760	0,87643		
	Arcos de Valdevez	0,52947	0,66966		Portalegre	0,50161	0,85555		Chamusca	0,57421	0,92621		
	Arcos de Valdevez	0,54945	0,46156		Portalegre	0,49114	0,78878		Chamusca	0,58642	0,88726		
	Arcos de Valdevez	0,55086	0,40167		Portalegre	0,48602	0,89836		Chamusca	0,52693	0,49108		
	Arcos de Valdevez	0,59868	0,43667		Portalegre	0,50517	0,80696		Chamusca	0,53908	0,32639		
	Arcos de Valdevez	0,59207	0,15611		Portalegre	0,47818	0,84788		Chamusca	0,53622	0,17922		
	Montemor-o-Velho	0,51636			Arganil	0,51040			Carregal do Sal	0,59954			
	Montemor-o-Velho	0,47983	0,77920		Arganil	0,53844	0,86018		Carregal do Sal	0,59887	0,92308		
	Montemor-o-Velho	0,55719	0,80170		Arganil	0,56688	0,89157		Carregal do Sal	0,56233	0,79748		
	Montemor-o-Velho	0,48030	0,35487		Arganil	0,51958	0,47858		Carregal do Sal	0,54096	0,45783		
	Montemor-o-Velho	0,48060	0,89258		Arganil	0,50842	0,94158		Carregal do Sal	0,45777	0,81561		(1) de acordo c/Cap4
	Montemor-o-Velho	0,48249	0,92547		Arganil	0,48375	0,92775		Carregal do Sal	0,49834	0,90274		
	Montemor-o-Velho	0,49884	0,75954		Arganil	0,48158	0,92268		Carregal do Sal	0,54825	0,84961		(2) de acordo c/Cap5
	Estarreja	0,50008			Alcácer do Sal	0,56723			Aljustrel	0,58683			
	Estarreja	0,51413	0,71651		Alcácer do Sal	0,56514	0,95417		Aljustrel	0,58370	0,95307		
	Estarreja	0,50683	0,60341		Alcácer do Sal	0,56102	0,84839		Aljustrel	0,58261	0,86616		Municípios com
	Estarreja	0,48019	0,23319		Alcácer do Sal	0,55614	0,75868		Aljustrel	0,57087	0,71413		População<=30000
	Estarreja	0,45620	0,11096		Alcácer do Sal	0,51616	0,36294		Aljustrel	0,54665	0,49441		e População>10000
	Estarreja	0,46961	0,89915		Alcácer do Sal	0,51037	0,19769		Aljustrel	0,56023	0,30402		
	Estarreja	0,48845	0,69478		Alcácer do Sal	0,56330	0,19114		Aljustrel	0,57987	0,20206		continua

munic	probab. (1)	probab. (2)	munic	probab. (1)	probab. (2)	munic	probab. (1)	probab. (2)
Abrantes	0,49326		Paços de Ferreira	0,51019		Bragança	0,49853	
Abrantes	0,46882	0,74475	Paços de Ferreira	0,50645	0,61823	Bragança	0,51361	0,81697
Abrantes	0,48857	0,51373	Paços de Ferreira	0,54482	0,72189	Bragança	0,47846	0,88129
Abrantes	0,47242	0,34084	Paços de Ferreira	0,57144	0,56164	Bragança	0,47649	0,79616
Abrantes	0,43078	0,78297	Paços de Ferreira	0,55685	0,47991	Bragança	0,45486	0,89218
Abrantes	0,48579	0,94409	Paços de Ferreira	0,56606	0,28875	Bragança	0,49561	0,89546
Abrantes	0,52324	0,90394	Paços de Ferreira	0,58505	0,12390	Bragança	0,47474	0,91308
Águeda	0,49376		Palmela	0,52934		Beja	0,53819	
Águeda	0,45607	0,45531	Palmela	0,55910	0,92977	Beja	0,53540	0,90481
Águeda	0,49106	0,57705	Palmela	0,56566	0,85215	Beja	0,53412	0,77530
Águeda	0,47053	0,22008	Palmela	0,54122	0,64279	Beja	0,52198	0,59070
Águeda	0,48897	0,21967	Palmela	0,51421	0,40142	Beja	0,48658	0,23189
Águeda	0,48941	0,10843	Palmela	0,50049	0,13407	Beja	0,50143	0,11934
Águeda	0,51746	0,05625	Palmela	0,53999	0,10164	Beja	0,50499	0,05496
Loulé	0,50108		Vila Real	0,49754		Seia	0,47766	
Loulé	0,47737	0,76479	Vila Real	0,55572	0,82433	Seia	0,56139	0,91785
Loulé	0,51028	0,83592	Vila Real	0,55201	0,76772	Seia	0,56478	0,73521
Loulé	0,52737	0,73884	Vila Real	0,54247	0,45193	Seia	0,54551	0,55465
Loulé	0,48441	0,90894	Vila Real	0,50119	0,19913	Seia	0,46210	0,72408
Loulé	0,50383	0,86496	Vila Real	0,50514	0,13168	Seia	0,51606	0,94045
Loulé	0,52404	0,62826	Vila Real	0,56238	0,11532	Seia	0,58024	0,93634
Lousada	0,54967		Montijo	0,47048		Fundão	0,50259	
Lousada	0,51956	0,69808	Montijo	0,47162	0,89223	Fundão	0,49937	0,78506
Lousada	0,53375	0,69823	Montijo	0,49082	0,80502	Fundão	0,51165	0,77932
Lousada	0,54003	0,44368	Montijo	0,48662	0,90747	Fundão	0,48972	0,46742
Lousada	0,48557	0,86615	Montijo	0,43484	0,59790	Fundão	0,51054	0,91423
Lousada	0,54870	0,92280	Montijo	0,44690	0,43935	Fundão	0,52378	0,87432
Lousada	0,57570	0,82792	Montijo	0,46850	0,91145	Fundão	0,54185	0,69666
Mafra	0,50960		Sesimbra	0,51744	0,94068	Santiago do Cacém	0,53079	
Mafra	0,50017	0,77750	Sesimbra	0,55682	0,97538	Santiago do Cacém	0,53637	0,91283
Mafra	0,51568	0,79182	Sesimbra	0,55212	0,93643	Santiago do Cacém	0,55005	0,81568
Mafra	0,48511	0,43023	Sesimbra	0,53683	0,88020	Santiago do Cacém	0,52319	0,58525
Mafra	0,51863	0,49642	Sesimbra	0,49583	0,54449	Santiago do Cacém	0,51906	0,39992
Mafra	0,50762	0,23997	Sesimbra	0,48434	0,27677	Santiago do Cacém	0,52481	0,20074
Mafra	0,54965	0,14449	Sesimbra	0,48307	0,96890	Santiago do Cacém	0,51438	0,06328

(1) de acordo c/Cap4

(2) de acordo c/Cap5

Municípios com População <= 50000 e População >30000

continua

5. O Modelo de Escolha Binária e a Probabilidade Estimada de Reeleição

It seems easy to accept that rationality involves many features that cannot be summarized in terms of some straightforward formula, such as binary consistency. But this recognition does not immediately lead to alternative characterizations that might be regarded as satisfactory, even though the inadequacies of the traditional assumptions of rational behaviour standardly used in economic theory have become hard to deny. It will not be an easy task to find replacements for the standard assumptions of rational behaviour that can be found in the traditional economic literature, both because the identified deficiencies have been seen as calling for rather divergent remedies, and also because there is little hope of finding an alternative assumption structure that will be as simple and usable as the traditional assumptions of self-interest maximization, or of consistency of choice.

Amartya Sen (1990, p. 206)

Na secção anterior seguiu-se uma abordagem convencional aos modelos de percentagem de votos, no entanto pode-se conceber a escolha do partido/candidato como um resultado implícito de preferências eleitorais ou utilidades¹ e não uma mera escolha partidária. As escolhas dos votantes podem ser conceptualizadas como tomadas de decisão entre alternativas não ordenadas (partidos, candidatos) em função das características, (i) do votante; (ii) das alternativas disponíveis (partido/candidato) e respectivos atributos e (iii) da regra de decisão.

5.1. Colocação Teórica do Problema

A introdução de uma variável dicotómica no modelo em estudo é uma tradução implícita dos modelos de escolha discreta, os quais se suportam nos pressupostos

¹ Esta alternativa não é trivial, na medida em que a escolha individual pode ser deduzida a partir das preferências individuais ou utilidades, enquanto o reverso não tem que ser necessariamente verdadeiro.

subjacentes à natureza da tomada de decisão individual.² Como consequência metodológica os modelos analíticos de escolha (binário, multinomial, *probit/logit*, condicional *logit*), ao incorporarem essa variável binária, contêm estimativas de preferências eleitorais e utilidades baseadas nas observações das escolhas partidárias e nos atributos dos partidos e dos votantes. De facto, quando se estimam os modelos de escolha discreta de votação e os resultados são interpretados, está-se a estimar mais do que a escolha realizada, efectivamente os parâmetros estimados identificam a utilidade percebida (atribuída) pelo eleitor pela escolha de uma determinada alternativa (partido/candidato) sob o pressuposto de que essa escolha maximiza essa utilidade.

Prever uma variável reeleição dicotómica, *dummy*; 1, se houver reeleição, 0, nos outros casos, em vez de se prever a percentagem de votos de cada um dos partidos concorrentes ao acto eleitoral, poderia constituir um propósito deste ensaio. Já que, relativamente aos modelos de percentagem de votos convencionais, uma medida de ajustamento exigiria a combinação de informação relativa à previsibilidade das forças de expansão ou contracção das intenções de votos com a informação sobre o “*timing*” das reeleições, *em si mesma*, enquanto, o modelo com a variável *dummy*, reeleição, dependente, isolaria o rigor com que se pode anunciar a reeleição do partido incumbente.

Métodos estatísticos comumente utilizados (binomial ou multinomial *logit/probit*) para estimar modelos de escolha multipartidária impõem muitas vezes pressupostos restritivos acerca dessa escolha, já que estimam relações entre variáveis independentes e utilidades.³ Uma forma de contornar o problema é gerar informação acerca da utilidade das alternativas a partir de comportamentos observados, neste caso a reeleição da alternativa (partido/candidato) significaria uma utilidade positiva e a não reeleição da alternativa (partido/candidato) significaria uma utilidade não positiva.⁴ Esta solução encontra suporte teórico em Lancaster (1966), ao sugerir que as escolhas sobre alternativas não são, *em si mesmas*, o objecto da utilidade, em vez disso a utilidade decorre dos atributos/características dessas opções de escolha relativamente às

² O processo de escolha avalia a utilidade de cada alternativa, para um conjunto fixo de alternativas e respectivos atributos, para uma posterior decisão, baseada no objectivo de maximização de utilidade.

³ A informação acerca da utilidade de escolhas alternativas passa a constituir um problema que se centra em tentar observá-las ou tentar deduzi-las.

⁴ Estar-se-ia num cenário de utilidades/preferências reveladas.

características/atributos dos indivíduos que decidem a partir dessas opções de escolha. A ser assim se assumirmos que todas as características/atributos das opções de escolha (partidos) e dos indivíduos que escolhem (votantes) são observadas, pode-se estimar a utilidade indirecta a partir das escolhas observadas (percentagem de votos do partido incumbente no acto eleitoral). Este procedimento de gerar utilidades estimadas é o que está subjacente à utilização de modelos analíticos de escolha discreta.

A escolha é um comportamento ou decisão de comportamento que tem subjacente a selecção de uma das alternativas disponíveis. O pressuposto de que os fluxos de escolha se fazem a partir das utilidades implica que as preferências podem ser parcialmente deduzidas a partir das escolhas, com a alternativa escolhida a ser preferida sobre todas as outras alternativas. Ou seja a escolha é apenas uma observação parcial das preferências/utilidades por apenas revelar a ordenação da alternativa escolhida face a todas as outras, sem que se evidencie a ordenação entre as alternativas que não foram escolhidas. De certa forma pode-se dizer que se está confrontado com um certo tipo de problema reverso de “Arrow”.⁵

Os modelos de escolha discreta assumem que a componente estocástica das utilidades para cada alternativa são independentes, isto é a independência de alternativas irrelevantes (*IIA*) tem que estar subjacente. Esta assumption tem a virtude de nos orientar para uma simples observação de uma função logarítmica de verosimilhança e uma interpretação relativamente directa dos parâmetros estimados em termos probabilísticos. Contudo, estas características do modelo assentam num pressuposto dúbio, que consiste em assumir que a componente estocástica da utilidade associada a cada resultado não está correlacionada com as componentes estocásticas de outros resultados, na prática, muitas vezes, não é o caso.⁶

⁵ Arrow (1951) postulou que as preferências individuais não são, necessariamente, convertidas adequadamente em escolhas públicas.

⁶ Uma experiência interessante para ilustrar este caso é observar se as probabilidades relativas das alternativas, j e k , permanecem inalteráveis, caso se acrescente ou não, uma alternativa, l , ao pacote de escolhas alternativas, identificada, frequentemente, como o problema “autocarro vermelho, autocarro azul”, uma vez que se ilustra assim: “ As escolhas individuais de transporte relativamente às alternativas {“automóvel”, “autocarro vermelho”}, com as probabilidades respectivas $\{3/5, 2/5\}$, induzem uma probabilidade relativa de $3/2$. Com a introdução, ao conjunto dessas alternativas, de uma terceira alternativa, semelhante e inócua, como seja a do “autocarro azul”, as probabilidades passam a ser, $\{3/5, 1/5, 1/5\}$; os indivíduos são indiferentes à cor e as probabilidades relativas entre “carro” e “autocarro vermelho” mudam para 3. Neste caso a independência entre o erro estocástico associado ao resultado “escolha de carro”, ε_1 , e o erro estocástico associado ao resultado “autocarro vermelho”, ε_2 , é um pressuposto sem sentido, já que

Embora a imposição do pressuposto da independência de alternativas irrelevantes (*IIA*) nos modelos comportamentais de escolha se apresente restritivo, a sua importância não deve ser exacerbada nas aplicações estatísticas desses modelos.

Primeiro, o *IIA* é uma propriedade lógica da tomada de decisão e não uma propriedade estatística como é a consistência e o não enviesamento. Como propriedade lógica, apenas, é expectável alguma dose de *IIA* para sustentar, quer os modelos teóricos de escolha, quer os comportamentos de escolha observados. Segundo, qualquer modelo que especifique a probabilidade de escolha e viole a propriedade *IIA* pode ser expresso na forma logística com as escolhas apropriadas como variáveis dependentes. As probabilidades estimadas pelo modelo logístico serão consistentes com a violação de *IIA* [Train (1993), pp. 18-24].⁷ Normalmente, isto envolve um pouco mais do que incluir intercepções específicas de escolha. Mais exactamente, se estimarmos um modelo logístico especificado com o termo independente de escolha específica (intercepção), as probabilidades obtidas captam as frequências de escolha observadas independentemente do processo de escolha subjacente ser consistente com *IIA*.⁸

Muitas das decisões de escolha realizadas pelos indivíduos têm subjacentes interesses de natureza política e/ou económica. Entretanto, apenas, os resultados do processo de decisão são observáveis, até lá o processo implícito a esses resultados não é observável, diz-se latente. No caso mais simples as decisões dos indivíduos são reduzidas a decisões binárias (sim/não), as quais, sem perda de generalidade, podem ser identificadas com 1 para decisões sim e identificadas com 0 para decisões não.

uma maior (menor) utilidade pelo autocarro vermelho implicaria uma maior (menor) utilidade pelo autocarro azul. A probabilidade $P_0 = P(U_0 > U_1, U_0 > U_2)$, viria subestimada face à verdadeira probabilidade, sob o pressuposto da independência (uma vez que o pressuposto ignora o facto de que o acontecimento $U_0 > U_1$ torna mais provável o acontecimento $U_0 > U_2$ ” [Amemiya (1985), pp. 298].

⁷ Estas variáveis, tipicamente, devem medir as características das alternativas, à excepção da alternativa para a qual a utilidade em questão é definida. Naturalmente, se o processo de escolha admite violação de *IIA*, a estimação da constante num modelo logístico de escolha específica retomarás as correctas probabilidades agregadas (frequência relativa).

⁸ Alvarez e Nagler (1998) observam que “*IIA* não se agrega”. Esta dedução mostra que, mesmo que as escolhas individuais reflectam *IIA*, a introdução de uma nova alternativa produzirá probabilidades médias que reflectem violações de *IIA*. A questão empírica mais genérica é que, mesmo que a escolha individual reflecta ou não reflecta *IIA*, os parâmetros estimados pelo modelo, retomarão as correctas probabilidades agregadas.

5.2. Modelização do Problema

Caso se assuma que a variável binária observada y_i , é a realização discreta de uma variável aleatória contínua (latente) y_i^* , pode-se especificar uma forma funcional linear para a variável latente, do tipo,

$$y_i^* = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \dots, N$$

em que \mathbf{x}_i , é o $K * 1$ vector das variáveis explicativas

$\boldsymbol{\beta}$, é o $K * 1$ vector dos coeficientes desconhecidos

ε_i , é o termo erro desconhecido.

A variável latente y_i^* , é contínua e não é observada, pode ser considerada como representando utilidade esperada do indivíduo i , neste caso poderia ser o benefício líquido implícito na decisão de reeleger o partido incumbente. Pressupondo o modelo de utilidade aleatória, os indivíduos escolherão uma dada alternativa (reeleição ou não reeleição) se esperarem que ela venha a maximizar a sua utilidade, benefício líquido, analiticamente, expressa por,

$$y_i = \begin{cases} 1, & \text{se } y_i^* > \bar{y} \\ 0, & \text{se } y_i^* < \bar{y} \end{cases}$$

em que \bar{y} representa o valor crítico que transforma o índice subjacente y_i^* numa decisão de *sim* ou de *não*. Vindo que,

$$\Pr(y_i = 1) = \Pr(y_i^* > \bar{y}) = \Pr(\varepsilon_i > -\mathbf{x}_i'\boldsymbol{\beta}) = F(\mathbf{x}_i'\boldsymbol{\beta}).^9$$

A forma funcional do modelo de escolha binária, isto é a forma funcional de $F(\mathbf{x}_i'\boldsymbol{\beta})$, pode diferir de acordo com os pressupostos assumidos para a estrutura do termo erro, ε_i . Assim se for assumida uma distribuição independente e idêntica (i.i.d.) de valor extremo do tipo I (log *Weibull*) para os erros, tem-se o modelo *logit*,¹⁰ caso se assuma que os erros seguem uma distribuição normal padrão, com média zero e variância unitária, independente e identicamente distribuída, (i.i.d.), tem-se o modelo *probit*.¹¹ As distribuições, seja a logística, seja a normal cumulativa, são muito próximas, dentro de determinados limites da dimensão da amostra.¹² Quer num caso, quer no outro, ignora-se qualquer heterogeneidade entre os indivíduos.

Acontece que, trabalhando com dados em painel, como é o caso em estudo, há que não neutralizar o manancial de informação disponível, em termos seccionais (município a município) e em termos temporais (sete momentos de escolha eleitoral), para o efeito ter-se-á que equacionar o modelo de escolha binária em cenário “dados em painel”. Ora, o modelo de escolha binária, com dados em painel, permite-nos equacionar essa heterogeneidade na forma de efeitos específicos individuais (e/ou efeitos específicos temporais),¹³ os quais surgem representados por μ_i (e/ou λ_t), conforme Hsiao (1986, 1996). Uma maior realização de μ_i (e/ou λ_t) aumenta a probabilidade de $y_{it} = 1$, para

⁹ A última igualdade é sustentada na medida em que a função de densidade cumulativa descrita por F é simétrica à volta de zero, o mesmo que $1 - F(-\mathbf{x}_i'\boldsymbol{\beta})$.

¹⁰ Tem-se, $F(\mathbf{x}_i'\boldsymbol{\beta}) = \exp(\mathbf{x}_i'\boldsymbol{\beta}) / (1 + \exp(\mathbf{x}_i'\boldsymbol{\beta}))$.

¹¹ Tem-se, $F(\mathbf{x}_i'\boldsymbol{\beta}) = \int_{-\infty}^{\mathbf{x}_i'\boldsymbol{\beta}} \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon$.

¹² Enquanto a probabilidade de resposta binária estiver entre (0,1;0,9) é difícil discriminar entre as funções logística e normal, apenas ao nível da inferência estatística se pode encontrar essa diferença.

¹³ A assumpção dos factores omitidos como aleatórios permite captar a heterogeneidade negligenciada. Isto mais na tradição da especificação do modelo do que na perspectiva dos efeitos fixos. A técnica de estimação dos efeitos fixos aplicada quando os efeitos específicos individuais são aleatórios resultará, na melhor das hipóteses, na perda de eficiência de estimação de β e, na pior, em estimativas inconsistentes atribuídas a pequenos T , na maior parte dos dados em painel, Hsiao (1996).

todo o $i = 1, \dots, N$ e $t = 1, \dots, T$. Com dados em painel, o modelo latente pode ser re-parametrizado de forma a incluir esses efeitos específicos individuais, vindo,

$$\mathbf{y}_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + \mu_i + \lambda_t + v_{it}, \quad (5.2.1)$$

Em que,

\mathbf{y}_{it}^* , representa o vector da variável dependente não observada numa i -unidade seccional (município), num dado período de tempo, t ;

\mathbf{x}_{it}' , é a matriz das variáveis exógenas numa i -unidade seccional (município), num dado período de tempo, t ;

$\boldsymbol{\beta}$, é o vector dos coeficientes associados às características exógenas identificadas em \mathbf{x} ;

μ_i , é o efeito específico individual não observável;

λ_t , é o efeito específico temporal não observável

e

v_{it} é o termo de erro aleatório do modelo latente, em dados painel

Em ambiente “dados em painel”, tal como sucede com a estimação de modelos de escolha binária em domínio de “dados seccionais”, utilizam-se os métodos de estimação *logit* e *probit*. A selecção do método de estimação continua a vir determinada pelos pressupostos assumidos para a distribuição do termo de erro aleatório v_{it} . Tal como nos modelos painel de variáveis dependentes contínuas e no que concerne aos efeitos específicos individuais, μ_i , (e/ou aos efeitos específicos temporais, λ_t), continua-se a assumir que, ou são fixos (parâmetros constantes), ou são aleatórios (seguem uma distribuição com uma média e variância associada). A diferença chave entre o modelo contínuo e o modelo dicotómico reside no facto de que a escolha, entre efeitos específicos fixos e efeitos específicos aleatórios, determinará os pressupostos relativos à distribuição do termo de erro aleatório, v_{it} , e consequentemente decidirá o método de estimação, ou *logit* ou *probit*.

5.3. Métodos de Estimação: Logit versus Probit

Qualquer que seja o método de estimação escolhido, estar-se-á sempre na presença de parâmetros desconhecidos: aos constantes do vector β , dos coeficientes associados às características exógenas identificadas em \mathbf{x} , acrescentar-se-á os efeitos específicos individuais, μ_i , (e/ou aos efeitos específicos temporais, λ_t) e, assim, desde que $N \rightarrow \infty$ para um T fixo, ter-se-á um número de parâmetros μ_i a crescer com o valor de N (e/ou um número de parâmetros λ_t a crescer com o valor de T). Isto significa que μ_i pode não vir estimado de forma consistente para um dado T (e/ou λ_t pode não vir estimado de forma consistente para um dado N).¹⁴

No ensaio em curso, suportado numa base de dados em que a dimensão T é curta e a dimensão N é, razoavelmente, grande, a perspectiva de estimação pelo método da máxima verosimilhança, onde os efeitos específicos μ_i (e/ou λ_t), são estimados, adicionalmente, à estimação dos coeficientes β relativos às variáveis exógenas constantes do vector \mathbf{x} , revelar-se-á inconsistente.¹⁵ Torna-se necessário um método

¹⁴ Situação identificada, em estatística, como problema de parâmetros acidentais, discutida por Neyman e Scott (1948) e mais recentemente revista por Lancaster (2000). Este problema estatístico decorre da não linearidade dos modelos *logit/probit*, não havendo forma de remover os efeitos específicos μ_i e λ_t , sem inconsistências, que se virão a reflectir na estimação dos coeficientes β . Sabe-se que, nos modelos lineares, de dados painel, com T fixo, os valores de β vêem estimados consistentemente, apenas, depois de removidos os efeitos específicos μ_i e λ_t , por exemplo, por via da transformação dentro, *within*, o que não acontece com os modelos não lineares, onde $f(y_{it}|\mathbf{x}_i)$ é função de μ_i (para simplificar a exposição assume-se $\lambda_t = 0$, para todo o t). Isto é, a independência assintótica, dos estimadores de máxima verosimilhança (*MLE*) de β e de μ_i , que suporta a consistência da estimação nos modelos lineares não pode ser assegurada nos modelos não lineares.

¹⁵ Quando μ_i (e/ou λ_t) são tratados como parâmetros, se T (e/ou N) é finito, as condições de regularidade, para obter estimadores consistentes dos parâmetros estruturais β , são violadas pela presença de parâmetros acidentais, ainda que β e μ_i sejam informação ortogonal, já que vem nula a média das derivadas cruzadas da função logarítmica de verosimilhança, condicionada a \mathbf{x}_i e a μ_i , isto é $E\left(\frac{\partial^2 \ell_i(\beta, \mu_i)}{\partial \beta \partial \mu_i} \middle| \mathbf{x}_i, \mu_i\right) = 0$. Por exemplo no modelo de escolha binária, onde a variável observada, y_{it} , toma o valor 0, ou 1, dependente da função de resposta latente e o termo de erro v_{it} segue uma distribuição idêntica e independente, tem-se $E(y_{it}|\mathbf{x}_{it}, \mu_i) = [1 - F(-\mathbf{x}_{it}'\beta - \mu_i)]$. Uma vez que μ_i afecta não linearmente $E(y_{it}|\mathbf{x}_{it}, \mu_i)$, se procurarmos removê-lo por transformação de “diferenças” não se consegue. Viria, $y_{it} - y_{i,t-1} = [1 - F(-\mathbf{x}_{it}'\beta - \mu_i)] - [1 - F(-\mathbf{x}_{i,t-1}'\beta - \mu_i)] + (v_{it} - v_{i,t-1})$ e a função de verosimilhança condicionada por \mathbf{x}_{it} e μ_i , base para a estimação dos parâmetros β e μ_i , por via da maximização,

para libertar o modelo dos efeitos específicos, similar ao utilizado para os modelos de variáveis dependentes contínuas, para que se produzam estimações consistentes, encontrando para o efeito transformações que remodelem o modelo original num modelo que não envolva parâmetros acidentais.¹⁶ O estimador de máxima verosimilhança, desenvolvido por Chamberlain (1980), evita o problema dos parâmetros acidentais, condicionando a função de verosimilhança a uma estatística, suficientemente mínima, para os efeitos específicos μ_i : propondo para o efeito que $\sum_{t=1}^T y_{it} = y_{i+}$ seja essa estatística suficientemente mínima.¹⁷ Está-se na presença de um modelo *logit* condicional de efeitos fixos,¹⁸ em que se comparam todas as observações dentro de um município (unidade seccional) quando há reeleição do incumbente, com todas as observações dentro de um município (unidade seccional) quando não há reeleição, condicionada à mesma soma dentro do conjunto.

Acresce, ainda, que a obtenção de resultados similares aos alcançados via modelo linear, quando a forma do modelo é conforme (5.2.1), constitui motivação e objectivo subjacente à escolha de um modelo de escolha binária, com efeitos fixos. Importa precisar que o termo “efeitos fixos” nada tem a ver com a natureza da amostra. Reporta,

tomaria a forma de $\prod_{i=1}^N \prod_{t=1}^T [F(-\mathbf{x}'_{it}\boldsymbol{\beta} - \mu_i)]^{1-y_{it}} [1 - F(-\mathbf{x}'_{it}\boldsymbol{\beta} - \mu_i)]^{y_{it}}$, produzindo parâmetros $\boldsymbol{\beta}$ inconsistentes, já que a característica finita da base de dados, limita informação relativamente a μ_i e por consequência desconhece-se $f(\mu_i | \mathbf{x}_{it})$.

¹⁶ Não existe uma regra geral aplicável a um qualquer modelo não linear. Acontece que se tem de explorar a estrutura específica de um modelo não linear para encontrar uma transformação. Por exemplo, no modelo ilustrado pela equação (5.2.1), se assumirmos que o erro aleatório é independente ao longo do tempo e segue uma distribuição logística, ter-se-á,

$\Pr(y_{it} = 1 | \mathbf{x}_{it}, \mu_i) = e^{\mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i} / (1 + e^{\mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i})$ (para simplificar assume-se $\lambda_t = 0$). Uma vez que no modelo *logit*, os denominadores da

$\Pr(y_{it} = 1 | \mathbf{x}_{it}, \mu_i)$ e da $\Pr(y_{it} = 0 | \mathbf{x}_{it}, \mu_i)$ são idênticos, e o numerador, de uma qualquer sequência $\{y_{i1}, \dots, y_{iT}\}$ com $\sum_{t=1}^T y_{it} = s$,

é sempre igual a $e^{\sum_{t=1}^T (\mathbf{x}'_{it}\boldsymbol{\beta}) y_{it}}$, a estimação do modelo pela maximização da função de verosimilhança, condicionada sobre

$\sum_{t=1}^T y_{it} = s$, não envolverá os parâmetros acidentais, μ_i .

¹⁷ Passar-se-á a utilizar $f(y_{it} | \mathbf{x}_{it}, y_{i+})$ não dependente de μ_i na estimação dos coeficientes $\boldsymbol{\beta}$.

¹⁸ Por isso, para se obterem os valores estimados de $\boldsymbol{\beta}$, Chamberlain sugere a maximização de uma função de verosimilhança

condicional, $L_C = \prod_{i=1}^N \Pr(Y_{i1} = y_{i1}, \dots, Y_{iT_i} = y_{iT_i} / \sum_{t=1}^{T_i} y_{it})$, a qual requer exogeneidade estrita, parcialmente relaxada, pelo facto

da distribuição logística ser a única função de distribuição conjunta em que (i) os efeitos específicos temporais e seccionais são independentes e (ii) a soma das variáveis binárias é a estatística suficiente.

apenas, um modelo que traduz o efeito de um conjunto de variáveis exógenas, ilustradas matricialmente em \mathbf{x} , sobre uma variável binária dependente, identificada vectorialmente em \mathbf{y} , dado as variáveis exógenas constantes de \mathbf{x} e os efeitos específicos de unidade assumidos em μ , em que os valores de \mathbf{x} e \mathbf{y} são observados, mas não os de μ , e onde não se impõe restrições sobre a distribuição de $\mu|x$. De acordo com a literatura sobre modelos painel de escolha binária, apenas se obtêm estimadores consistentes, sob a propriedade da suficiência, daí decorrendo uma inferência condicional, que impõe uma extensão do estimador de máxima verosimilhança de Chamberlain.¹⁹

Intuitivamente pode-se pensar que, se conhecermos o número total de “1s” a função de verosimilhança pode ser condicionada, para se estimarem coeficientes β , consistentes, relativos às variáveis exógenas constantes do vector \mathbf{x} ; neste caso os efeitos específicos, μ_i , determinariam a proporção total de “1s” nos dados, enquanto os valores de β determinariam quando é que aqueles “1s” ocorriam; porém se os valores da variável binária, y_{it} , viessem sempre nulos ou sempre unitários, numa dada unidade seccional i , isto é não variassem ao longo do tempo, não se podia condicionar a função de verosimilhança, a estimação dos efeitos específicos, μ_i , tornar-se-ia impossível, uma vez que tenderiam para infinito.²⁰

¹⁹ Chamberlain (1980) sugere condicionar cada conjunto de observações sobre o número de “1s” para uma dada unidade seccional

em particular. Isto é, para $T = 2$, apenas a estatística $\sum_{t=1}^T y_{it} = 1$ se revelaria interessante para a concretização da função condicional

salvaguardada de heterogeneidade, materializada na remoção dos efeitos específicos μ_i , ainda que a estatística $\sum_{t=1}^T y_{it}$ pudesse

tomar os valores 0, 1, 2. Acontece que, caso se condicionasse sobre a estatística $\sum_{t=1}^T y_{it} = 0$, significaria, $y_{i1} = y_{i2} = 0$ e caso se

condicionasse sobre a estatística $\sum_{t=1}^T y_{it} = 2$, significaria $y_{i1} = y_{i2} = 1$, vindo para ambos os casos

$\Pr(Y_{i1} = y_{i1}, \dots, Y_{iT_i} = y_{iT_i} / \sum_{t=1}^T y_{it}) = 1$, o que nada contribuiria para a função de verosimilhança condicionada.

²⁰ Não há estimador de máxima verosimilhança para μ_i , caso a variável dependente seja sempre 1 ou sempre 0 dentro de uma dada unidade seccional – a equação de verosimilhança $\log \ell_i$ não tem solução, caso y_{it} não sofra variação ao longo do tempo dentro de uma dada unidade seccional i , de acordo com Greene (2004).

Num modelo tipo Chamberlain, a dois períodos e a duas dimensões, as condições, necessária e suficiente, sobre a função de distribuição conjunta dos efeitos específicos, seccionais e temporais, permitem que a soma das variáveis binárias individuais através do tempo constitua estatística suficiente para os efeitos específicos seccionais;²¹ e que existam \sqrt{n} -estimadores consistentes de verosimilhança condicional. Críticas a esta perspectiva consubstanciam-se no facto dos pressupostos da independência e da distribuição logística das perturbações do erro serem demasiado fortes.²² A propriedade da suficiência pode ser interpretada como redutora da dimensão do modelo, transformando uma função de densidade conjunta dos efeitos específicos, seccionais e temporais, desconhecidos, em duas funções de densidade marginal univariada, tendo que verificar condições de extremidade, adicionais.²³ Para além disso a maximização da função de verosimilhança condicionada à estatística suficiente tem sido apresentada como equivalente à utilização de uma estimação que relaciona a expectativa da diferença entre o par das variáveis binárias e a diferença no índice linear das variáveis, numa amostra onde as variáveis binárias diferem através do tempo.²⁴ Ora, acontece que a base de dados que suporta este ensaio, apresenta um $T = 7$, e neste cenário a função de verosimilhança condicionada à estatística suficiente, $\sum_{t=1}^7 y_{it} = 1$, torna-se computacionalmente muito “pesada”. Esta particularidade, actualmente, não constitui

²¹ Para que a soma das variáveis binárias venha uma estatística S -suficiente para o parâmetro acidental μ , assumiram-se alguns pressupostos de regularidade que pressupõem completa variação de todas as probabilidades na dimensão “entre” unidades seccionais e completa variação de todas as probabilidades na dimensão “dentro” unidades seccionais; e, ainda, que as perturbações do erro v_1 e v_2 , seguem distribuições independentes e logísticas [Magnac (2005)].

²² Chamberlain (1980) prova que, se as perturbações do erro, v_1 e v_2 , são independentes, então o limite eficiente e semi-paramétrico do parâmetro β é igual a zero, a não ser que a função de distribuição da perturbação do erro seja logística.

²³ A partir do modelo de escolha binária, $y_{it} = 1\{\mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i + v_{it} \geq \bar{y}_i\}$, $t = 1, \dots, T; i = 1, \dots, N$, assumindo que y_{it} é independente de

x_{it} e de μ_i , tem-se a função de logarítmica de verosimilhança $\sum_{i=1}^N \ell_i(\beta, \mu_i)$. Apenas, os valores observados para (y_{i1}, y_{i2}) , de

(0,1) ou de (1,0), contribuem com o termo $2 \log F(\beta/2)$ e com o termo $2 \log [1 - F(\beta/2)]$, respectivamente, para a função

logarítmica de verosimilhança concentrada, vindo $2 \sum_{i=1}^N \{d_{10i} \log [1 - F(\beta/2)] + d_{01i} \log [F(\beta/2)]\}$, em que

$d_{10i} = 1(y_{i1} = 1, y_{i2} = 0)$ e $d_{01i} = 1(y_{i1} = 0, y_{i2} = 1)$. Neste cenário o estimador $\hat{\beta}$ de máxima verosimilhança (MLE) satisfaz,

$$p \lim_{N \rightarrow \infty} \hat{\beta} = 2F^{-1} \{ E_{\mu} [F(-\mu)F(\beta + \mu)] / E_{\mu} [F(-\mu)F(\beta + \mu)] + E_{\mu} [F(\mu)F(-\beta - \mu)] \}.$$

²⁴ O que decorre de $\Pr(y_{i1} = 1, y_{i2} = 0 | y_{i1} + y_{i2} = 1) = e^{\mathbf{x}'_{i1}\boldsymbol{\beta} + \mu_i} / e^{\mathbf{x}'_{i1}\boldsymbol{\beta} + \mu_i} + e^{\mathbf{x}'_{i2}\boldsymbol{\beta} + \mu_i} = e^{(\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\boldsymbol{\beta}} / 1 + e^{(\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\boldsymbol{\beta}}$.

óbice aos cálculos a efectuar, o mesmo já não acontece com o sacrifício das unidades seccionais (municípios), onde a variável binária não varia ao longo do tempo, podendo assistir-se a uma redução substancial de observações disponíveis, por não contribuírem para a função de máxima verosimilhança do estimador *logit* condicional.²⁵

Adoptar, à medida que o valor de T cresce, o modelo *probit*, pode revelar-se mais sustentável.²⁶ Entretanto, torna-se necessário assumir efeitos específicos aleatórios em vez de efeitos específicos fixos, na medida em que o modelo *probit* de efeitos fixos não garante um bom desempenho – não há forma fiável de remover a heterogeneidade e com grandes números de unidades seccionais a estimação dos valores μ_i é irresolúvel. Algum melhoramento é conseguido por via da especificação de efeitos aleatórios: maior parcimónia no número de parâmetros a estimar.²⁷ O *trade-off* é que esta especificação, exige pressupostos mais restritivos: (1) os efeitos específicos seccionais, μ_i , e os termos de erro aleatório, ν_{it} , seguem distribuições normais, idênticas e independentes, i.i.d., com média zero e variâncias σ_μ^2 e σ_ν^2 ;²⁸ (2) os efeitos específicos, μ_i , são independentes de x_{it} ;²⁹ e (3) uma distribuição normal para os valores de μ_i , condicionada aos valores de x_{it} , os quais são, ainda, independentes dos valores de x_{it} e de ν_{it} .³⁰

O método da máxima verosimilhança, normalmente, suporta a estimação dos modelos *probit*, dados em painel, (*PP*), com efeitos aleatórios, sob o pressuposto da normalidade de μ_i e de ν_{it} . Daí decorre, a função logarítmica de verosimilhança para o modelo, do tipo,

²⁵ Já que $\ln \left[\Pr \left(Y_{i1} = y_{i1}, \dots, Y_{iT_i} = y_{iT_i} / \sum_{t=1}^T y_{it} \right) \right] = 0$, para $y_{i1} = \dots = y_{i7} = 0$ e para $y_{i1} = \dots = y_{i7} = 1$.

²⁶ O estimador de máxima verosimilhança em modelos não lineares, dados em painel, com efeitos fixos, é amplamente percebido (com raras excepções) como um estimador enviesado e inconsistente quando T , o período temporal do painel, é curto e fixo.

²⁷ Em trabalhos desenvolvidos no âmbito do comportamento dos eleitores, tem sido atribuída importância ao estimador dos efeitos aleatórios, na medida em que permite estimar os efeitos das variáveis invariantes no tempo, tal como, por exemplo, filiação partidária, ideologia, “clientelismo”, que podem explicar alguma estabilidade de preferências ao longo do processo de decisão eleitoral.

²⁸ $\mu_i \sim N(0, \sigma_\mu^2)$ e $\nu_i \sim N(0, \sigma_\nu^2)$.

²⁹ O que pode acusar alguma falha de razoabilidade.

³⁰ Esta assumpção tem por objectivo derivar a função de verosimilhança.

$$l_{PP} = \sum_{i=1}^N \ln(P(y_{it}))$$

Em que $P(y_i)$, o contributo de cada um dos municípios i para a função de verosimilhança, seria a probabilidade conjunta para todas as 7 observações temporais em cada município i , ilustrada por,

$$P(y_i) = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{\mu_i^2}{2} \left\{ \prod_{t=1}^{T_i} \Phi\left(\frac{\mathbf{x}'_{it}\boldsymbol{\beta}}{\sigma_v} + \frac{\mu_i}{\sigma_v}(2y_{it} - 1)\right)\right\}\right\} d\mu_i \quad 31$$

A estimação da máxima verosimilhança (*MLE*) da função logarítmica de verosimilhança fornece estimativas consistentes e eficientes para β (sob as condições de regularidade fraca)³² no entanto é computacionalmente intensiva, chegando a ser impraticável em algumas situações. O problema reside no facto da função de verosimilhança para o indivíduo i , ou seja para o município i , ser um integral de um produto, em vez de ser apenas um produto e do operador logarítmico não poder ser aplicado sob o sinal de integral. E a situação é tanto mais complexa quanto a presença latente de correlação em série nas perturbações de erro para cada um dos municípios i . Nesse cenário a falta de independência entre as observações impede a possibilidade de se factorizar a função de verosimilhança contributo, dos T_i períodos, de cada município i , impondo um cálculo T -dimensional do integral da função de distribuição.

³¹ Φ , identifica a função de distribuição cumulativa (*CDF*) da variável.

³² Para estabelecer a existência de um estimador de máxima verosimilhança (*MLE*), gozando de propriedades assintóticas, são, apenas, necessários os pressupostos matemáticos das condições de regularidade fraca. Estes pressupostos podem ser sumariados, conforme King (1989, pp. 75), **1**. Os valores de y , para os quais a função de densidade $f(y|\theta) > 0$, não dependem de, θ ; **2**. $f(y|\theta)$ é uma função com 2ª derivada contínua em ordem a θ , $\forall \theta \in \Theta \subseteq \mathbb{R}^K$; **3**. A matriz de informação, é definida positiva e limitada e **4**. As primeiras três derivadas da função de verosimilhança em ordem a θ são limitadas e os limites da 3ª derivada não dependem de, θ .

Este método de estimação clássica ter-se-ia de abandonar, logo que se estivesse na presença de uma dimensão temporal superior a três ou quatro períodos, porém a estrutura de equicorrelação através dos períodos,³³ a qual é consequência dos pressupostos sobre o termo de heterogeneidade, μ_i , invariante no tempo e fortemente exógeno, certifica a firmeza da estimação de máxima verosimilhança (*MLE*). Esta estrutura tem subjacente a decomposição do termo erro em duas componentes diferentes e independentes, $\varepsilon_{it} = \mu_i + v_{it}$; em que uma das componentes é μ_i , o denominado efeito específico, constante ao longo do tempo, e não correlacionado com as variáveis, x_{it} , em todos os períodos,³⁴ e a outra das componentes é, v_{it} , o termo de perturbação remanescente, variável com o tempo, mas seguindo uma distribuição independente e idêntica (i.i.d.) através do tempo e das unidades seccionais.

Decorre destes pressupostos, uma matriz de covariância modificada, composta pelos

elementos: (i) fora da diagonal, $\sigma_{ts} = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_v^2} = \rho$,³⁵ para $t \neq s$; e (ii) na diagonal

principal $\sigma_{tt} = \sigma_\mu^2 + \sigma_v^2 = 1$.³⁶

Desta manipulação decorre que a função de verosimilhança passa a envolver, apenas, integrais de dimensão unitária e, por via da quadratura Gauss-Hermite, torna-se atractiva a sua estimação.³⁷ A estimação dos parâmetros β e do parâmetro de correlação ρ deixam de ser problema.

Este procedimento pretende que o modelo se transforme num modelo de escolha binária padrão, à medida que os termos da perturbação remanescente, v_{it} , vão sendo

³³ Proposta por Butler e Moffitt (1982).

³⁴ Ou seja, é independente através das unidades e invariante no tempo.

³⁵ O elemento ρ , não é senão a assumção, implícita, de uma correlação constante entre dois termos de erro sucessivo, para a mesma unidade seccional i , ou seja $\rho = corr(\varepsilon_{it}, \varepsilon_{is})$. O que justifica o enquadramento do modelo na classificação modelos de “equicorrelação”, ou de correlação uniforme.

³⁶ Significando ausência de qualquer informação acerca da escala de y_{it} .

³⁷ A perspectiva para a estimação do modelo probit efeitos aleatórios constante do pacote estatístico STATA tira vantagem da quadratura Gauss-Hermite sugerida em Butler e Moffitt (1982).

independentes entre períodos e unidades seccionais, para um dado valor de, μ_i . Nesta perspectiva, no caso particular dos termos de erro específico, μ_i , seguirem uma T -distribuição normal,³⁸ a função de verosimilhança, apropriada ao modelo *probit*, em relação a μ_i , expressa-se por,

$$\prod_i \left\{ \int_{-\infty}^{\infty} \prod_{t=1}^T \left[1 - \Phi \left(\mathbf{x}_{it}' \boldsymbol{\beta}^* + \sqrt{\frac{\rho}{1-\rho}} \mu^* \right) \right]^{1-y_{it}} \left[\Phi \left(\mathbf{x}_{it}' \boldsymbol{\beta}^* + \sqrt{\frac{\rho}{1-\rho}} \mu^* \right) \right]^{y_{it}} \phi(\mu^*) d\mu^* \right\}$$

Com $\beta^* = \beta / \sigma_v$ e $\mu^* = \mu / \sigma_\mu$, obtêm-se os valores estimados de β^* e ρ , os quais são normalizados sobre σ_v .³⁹ O problema de estimação vem, agora, reduzido a um integral simples, com o seu integrando a ser o produto das funções simples de densidade normal e as T diferentes funções de distribuição cumulativa normal.

Mesmo que, nesta função de verosimilhança, a estrutura de correlação ρ fosse ignorada,⁴⁰ ainda assim, os parâmetros estimados dos coeficientes β vinham consistentes, como refere Maddala (1987), embora ineficientes na medida em que os respectivos erros padrão assintóticos, estimados convencionalmente, se apresentavam

³⁸ Registe-se que, tal como nos modelos de escolha binária seccionais, as probabilidades de escolha marginal podem vir expressas em termos dos parâmetros do modelo latente: $\text{Prob}(y_{it} = 1 / \mu_i, x_{it}) = \text{Prob}\left(\frac{v_{it}}{\sigma_v} > \frac{-x_{it}'\beta - \mu_i}{\sigma_v}\right) = \Phi(z_{it})$, em que

$z_{it} = -(x_{it}'\beta + \mu_i) / \sigma_v$ e Φ é uma função de distribuição padrão normal. Para o efeito, atente-se no comportamento normal e independente da distribuição de y_{it}^* , condicionada por μ_i , Heckman (1981).

³⁹ Na estimação em curso, o pressuposto padrão da igualdade dos períodos temporais nos modelos painel exigível, é satisfeito, na medida em que a dimensão T é sempre 7, para todos os municípios.

⁴⁰ No limite, quando os termos de erro são independentes ao longo do tempo, o modelo painel converge num modelo seccional, com NT observações independentes, e o estimador de máxima verosimilhança é o estimador de escolha padrão, suportado na função de

verosimilhança logarítmica, $\prod_{t=1}^T \left[1 - \Phi(x_{it}'\beta^*) \right]^{1-y_{it}} \left[\Phi(x_{it}'\beta^*) \right]^{y_{it}}$.

inapropriados por ser ignorada a correlação através dos períodos.⁴¹ Nesta perspectiva um modelo *probit* com dados agregados (*pooled*) fornecerá parâmetros estimados consistentes.⁴² Mas, dada a natureza binária dos dados, uma estimação simples do modelo *probit* não permite estimar o parâmetro escala, estimando-se somente parâmetros de, β/σ_ε . Apenas, quando $\sigma_\mu^2 = 0$, se verifica igualdade entre o parâmetro estimado pelo *probit pooled* e o parâmetro estimado pelo modelo *probit efeitos aleatórios*. Caso $\sigma_\mu^2 \neq 0$, então a estimação de valores consistentes com o modelo *probit agregado* não produzirá resultados similares aos estimados pelo modelo *probit aleatório* por causa da normalização. Um cálculo simples mostrará que, para converter os coeficientes *probit aleatórios* β/σ_ε em β/σ_ν , se necessita apenas de multiplicar pelo factor $\sqrt{1-\hat{\rho}}$ ⁴³ esses coeficientes estimados.

5.4. Abordagem Técnica da Estimação

Na linha das questões metodológicas suscitadas na secção introdutória aos modelos painel de escolha binária, passíveis de serem sintetizadas no “problema dos parâmetros acidentais” do estimador de máxima verosimilhança na presença de efeitos fixos (*MLE / FEM*)⁴⁴ versus os “pressupostos restritivos de heterogeneidade” impostos pela

⁴¹ O programa estatístico STATA, permite o cálculo dos erros padrão a partir de um estimador denominado “cluster” estimador e que se suporta na matriz hessiana, \mathbf{H} , da função logarítmica de verosimilhança maximizada, e na derivada do termo individual na

função logarítmica de verosimilhança agregada, g_{it} , expresso por, $\hat{V}(\hat{\beta}) = \left(\frac{N}{N-1}\right)(-\mathbf{H}^{-1}) \left(\sum_{i=1}^N g_i g_i'\right) (-\mathbf{H}^{-1})$, em que

$g_i = \sum_{t=1}^T g_{it}$, na linha do que Guilkey e Murphy (1993) discutiram como estimativa consistente da matriz de covariância.

⁴² Robinson (1982) admitia a estimação de parâmetros consistentes, a partir dos dados agregados (*pooling data*), num modelo estático, ignorando a estrutura particular das correlações.

⁴³ Os pacotes estatísticos se encarregarão de obter resultados para β^* e para ρ e respectiva matriz de covariância Ω .

Mas $\hat{\beta} = \sqrt{1-\hat{\rho}}\beta^*$ implicará que, $\text{Covar}(\hat{\beta}) = \mathbf{A}\Omega\mathbf{A}'$. Em que, $\mathbf{A} = \begin{bmatrix} \sqrt{1-\hat{\rho}} & 0 & \dots & \dots & 0 & -\hat{\beta}_1/(2\sqrt{1-\hat{\rho}}) \\ 0 & \sqrt{1-\hat{\rho}} & \dots & \dots & \vdots & -\hat{\beta}_2/(2\sqrt{1-\hat{\rho}}) \\ \vdots & 0 & \dots & \dots & 0 & \vdots \\ 0 & 0 & \dots & 0 & \sqrt{1-\hat{\rho}} & -\hat{\beta}_k/(2\sqrt{1-\hat{\rho}}) \end{bmatrix}$.

⁴⁴ Refira-se, por exemplo, Wooldridge (2002, pp. 10-11) que, ao referir-se aos modelos de dados painel, estabelece “(...) with fixed T , it suffers from an incidental parameters problem: except in very special cases, the estimator of θ_0 is inconsistent.”

estimação do modelo de efeitos aleatórios (*MLE / REM*); e equacionando as restrições da base de dados em estudo, onde T é pequeno e N é grande, optou-se por desenvolver a estimação do modelo aplicando a técnica de máxima verosimilhança do modelo *logit*, efeitos fixos, condicional, mesmo assumindo que este método impede a estimação do(os) efeito(s) de(as) variável(eis) que não variam ao longo do tempo, bem como a perda de graus de liberdade.

5.4.1. Modelo logit

Para concretizar este objectivo, começou-se por desenhar um modelo de escolha binária que, simultaneamente, equaciona a base de dados do ensaio e a especificidade desse modelo,⁴⁵ vindo:

$$\begin{aligned}
 p_{it} = & \alpha_{it} + \beta_1 * VOTOSANT_{it} + \beta_2 * APART_{it} + \beta_3 * APRES_{it} + \\
 & + \beta_4 * DIREAL_{it} + \beta_5 * D_{it} + \beta_6 * DESEMP_t + \beta_7 * PDESEMP1_{it} + \\
 & + \beta_8 * INFLAC_t + \beta_9 * PINFLAC1_{it} + \varepsilon_{it}
 \end{aligned}$$

$$\varepsilon_{it} = \mu_i + v_{it}$$

⁴⁵ Em relação à especificação do modelo desenvolvido na secção 4.3.2.1, com o objectivo de se refinar a estimação e acautelando potenciais problemas de significância estatística, optou-se por considerar para além da variável desemprego ponderada pela *dummy* de coincidência partidária entre poderes, local e central ($PDESEMP1_{it}$) e da variável inflação ponderada pela *dummy* de coincidência partidária entre poderes, local e central ($PINFLAC1_{it}$), também as variáveis relativas, a desemprego sem quaisquer restrições (em nível ($DESEMP_t$) e em variação ($\Delta DESEMP_t$)), a inflação sem quaisquer restrições ($INFLAC_t$) e a *dummy* coincidência de “cor” política entre poderes, local e central, D_{it} .

Em que,
$$p_{it} = \begin{cases} 1, & \text{se } BL > a_1 \\ 0, & \text{se } BL < a_1 \end{cases},$$
 traduzindo,

p_{it} a probabilidade de reeleição, 1 se houver lugar à reeleição, 0 se não acontecer a reeleição; já o termo, BL , identifica uma variável não observada, e pode-se assumir como atribuível ao Benefício Líquido da decisão de imutabilidade do incumbente no poder; a_1 é o valor que permite a reeleição do partido do incumbente, em termos de percentagem de votos obtidos;

e,

$VOTOSANT_{it}$, é a percentagem de votos que o partido em exercício de funções no executivo autárquico no município i no ano de eleição t tinha obtido na eleição anterior;

$APART_{it}$, ilustra o número de anos sucessivos que o partido vencedor esteve ou não à frente dos destinos do município i à data de eleição t ;

$APRES_{it}$, traduz o número de anos sucessivos no exercício do cargo de presidente da câmara no município i à data de eleição t ;

$DIREAL_{it}$, representa as despesas de investimento *per capita*, em valor real, feitas no município i no ano de eleição t ;

D_{it} , capta a coincidência partidária entre governos, central e local, no município i no ano de eleição t ;

$DESEMP_t$, identifica a taxa de desemprego nacional, variável macroeconómica, que será considerada em nível e em variação ($\Delta DESEMP_t$ variação de desemprego inter eleições autárquicas);

$PDESEMP1_{it}$, é a denominação para a taxa de desemprego nacional, mas agora com restrições, ou seja reflecte a taxa de desemprego, apenas em municípios em que o partido do incumbente coincide com o do governo central;

e, ainda,

$INFLAC_t$, designa a taxa de inflação nacional, variável macroeconómica, à data do momento eleitoral t e

$PINFLAC1_{it}$, revela a taxa de inflação nacional, mas apenas quando o partido incumbente coincide no município e no governo central

Utilizando os procedimentos, disponibilizados pelo pacote estatístico STATA, ajustou-se uma regressão *logit* condicional (efeitos fixos). Nessa aplicação, apenas foram utilizados os dados dos municípios que experimentaram mudanças na probabilidade de reeleição ao longo do tempo, isto é os municípios que mantiveram sempre o autarca, ou que sempre o substituíram não contribuíram para a função logarítmica de verosimilhança.⁴⁶

A esta estimação condicionada à estatística suficiente para os efeitos específicos, μ_i , contrapôs-se uma regressão *logit* efeitos aleatórios.⁴⁷ O cálculo da função logarítmica de verosimilhança é difícil quando os erros não são independentes, mas, entretanto,⁴⁸

⁴⁶ O número de municípios sacrificados foi 68 (75, quando vem equacionada a taxa de desemprego em variação), já que nesses a distribuição condicional de y_{i1}, \dots, y_{iT} , se $\sum_{t=1}^T y_{it} = 0$ ou se $\sum_{t=1}^T y_{it} = 1$, não apresenta um comportamento padrão.

⁴⁷ Esta especificação requer a satisfação do pressuposto da exogeneidade estrita; ou seja reclama não só exogeneidade contemporânea, mas também exogeneidade passada e futura, isto é restrições sobre a relação entre as variáveis exógenas e o termo de erro no mesmo período, mas também em todos os períodos passados e futuros, no sentido que $F_{\varepsilon_t}(\varepsilon_{it} | x_i) = F_{\varepsilon_t}(\varepsilon_{it})$, em que $F_{\varepsilon_t}(\varepsilon_{it})$ identifica a função de distribuição marginal do termo de erro no período t . Pressuposto que, na ausência de erros independentes ao longo do tempo, exigirá uma condição mais pesada sobre a distribuição dos T -termos de erro ao longo do tempo, ilustrada por, $F_{\varepsilon}^{(T)}(\varepsilon_i | x_i) = F_{\varepsilon}^{(T)}(\varepsilon_i)$.

⁴⁸ Baseados numa estrutura em que os termos de erro ε_{it} são decompostos em duas componentes independentes, uma constante ao longo do tempo, denominada de efeitos específicos, μ_i , e outra variável ao longo do tempo, mas idêntica e independentemente distribuída através do tempo e das unidades seccionais, v_{it} .

obtem-se a densidade conjunta dos termos de erro remanescente, v_{it} , integrando os efeitos específicos μ_i fora da densidade conjunta de $(\varepsilon_{i1}, \dots, \varepsilon_{iT_i}, \mu_i)$, e por essa via os termos de erro ε_{it} condicionados sobre μ_i vêm independentes ao longo do tempo.⁴⁹

Na estimação *logit* desenvolvida, após prévio confronto entre o método dos efeitos fixos (*FEM*) e o método dos efeitos aleatórios (*REM*), ensaiaram-se diversas regressões, suportadas no método *FEM*, de modo a contemporizar, para além das variáveis, taxa de desemprego e taxa de inflação (constantes do modelo original), também as diversas cambiantes a essas grandezas. Da conseqüente exploração dos melhores desempenhos, em termos de significância estatística, resultaram as seis regressões que sustentam os resultados apresentados na tabela que se segue. Assim,

numa primeira regressão equaciona-se o modelo base sem qualquer uma das variáveis macroeconómicas, relativas ao desemprego e inflação e sem a *dummy* coincidência partidária;

numa segunda regressão, acresce-se ao modelo base constante da regressão anterior a *dummy* coincidência partidária entre governo local e governo central;

Já,

numa terceira regressão, o modelo base, suporte da regressão dois, passa a equacionar, também, a taxa de desemprego nacional sem quaisquer restrições;

numa quarta regressão, amplia-se o leque das variáveis contantes do modelo que suporta a regressão três, introduzindo o termo de interacção da variável taxa de desemprego nacional sem quaisquer restrições pela *dummy* coincidência partidária entre governo local e governo central;

⁴⁹ A generalização mais imediata da independência dos termos de erro, ε_{it} , é a estrutura de equicorrelação, a qual permite uma aproximação numérica ao integral da função de verosimilhança por um algoritmo de quadratura. Em suma, a eficiência do cálculo da função de verosimilhança vem garantida pela independência dos termos de erro e pela estrutura de equicorrelação, o que permite equacionar a razoabilidade técnica do ajustamento de uma especificação *logit* efeitos aleatório, ao modelo apresentado.

Este processo de inclusão sucessiva das variáveis propostas no modelo base só vem completo com,

uma quinta regressão, onde se introduz a variação da taxa de desemprego (definida a partir da relação entre a taxa de desemprego à data do ano da eleição autárquica e a taxa de desemprego à data do acto eleitoral autárquico imediatamente anterior)

e finalmente, com,

uma sexta regressão, de forma a enquadrar, ainda, as variáveis relativas à taxa de inflação nacional sem restrições e à taxa de inflação ponderada pela dummy de coincidência partidária nos governos, central e local.

Em todas as regressões, a estimação *logit* fez-se com recurso ao método dos efeitos fixos (*FEM*) e ao método dos efeitos aleatórios (*REM*), porém um teste de Hausman, que admite testar a hipótese nula de homogeneidade dos coeficientes estimados pelo eficiente estimador *REM* e pelo consistente estimador *FEM*,⁵⁰ permite escolher o estimador *FEM* em vez do estimador *REM*, por via da rejeição da hipótese nula.⁵¹ A estabilidade de resultados a favor do método dos efeitos fixos (*FEM*) para as seis regressões convidava à parcimónia na apresentação de resultados. Propósito concretizado na tabela de resultados, que se segue, onde a apresentação das duas estimações, *logit* efeitos fixos, condicional, (*FEM*) e *logit* efeitos aleatórios (*REM*), apenas, se verifica para a *primeira* regressão.

⁵⁰ Os parâmetros estimados, sob a hipótese nula, são consistentes, em qualquer uma das estimações (*logit* condicional efeitos fixos e *logit* efeitos aleatórios), no entanto o estimador condicional, sob esta hipótese nula, vêm ineficiente (uma vez que apenas utiliza informação dos municípios que sofreram alterações em termos da reeleição ou não reeleição). Já, sob a hipótese alternativa, o estimador condicional é consistente e eficiente, enquanto o estimador não condicionado é inconsistente.

⁵¹ Efectivamente, o valor obtido para a estatística de Hausman, em qualquer uma das situações, vem, respectivamente, 199,85; 210,37; 676,31; 2312,85; 166,19 e 182,35. Valores que, quando comparados com o valor crítico constante da distribuição dos χ^2 , permitem a rejeição da hipótese nula de homogeneidade, para o mais elevado nível de significância.

	Logit FE (1)	Logit RE (1)	Logit FE (2)	Logit FE (3)	Logit FE (4)	Logit FE (5)	Logit FE (6)
<i>VOTOSANT</i>	0,068*** (0,009) [0,005]	0,074*** (0,008)	0,074*** (0,009) [0,006]	0,074*** (0,009) [0,005]	0,074*** (0,009) [0,006]	0,064*** (0,009) [0,013]	0,069*** (0,011) [0,013]
<i>APART</i>	-0,149*** (0,021) [-0,012]	0,004 (-0,015)	-0,152*** (0,022) [-0,012]	-0,151*** (0,022) [-0,011]	-0,150*** (0,022) [-0,011]	-0,206*** (0,025) [-0,043]	-0,221*** (0,026) [-0,040]
<i>APRES</i>	-0,006 (0,024) [-0,0004]	-0,064*** (0,018)	-0,010 (0,024) [-0,0008]	-0,010 (0,024) [-0,0007]	-0,010 (0,025) [-0,0008]	-0,018 (0,028) [-0,004]	-0,034 (0,028) [-0,006]
<i>DIREAL</i>	0,002*** (0,0007) [0,0001]	0,0001 (0,0005)	0,002*** (0,0006) [0,0002]	0,002*** (0,0007) [0,0001]	0,002*** (0,0007) [0,0001]	0,002*** (0,0008) [0,0004]	0,001* (0,0009) [0,0003]
<i>D</i>			-0,646*** (0,138) [-0,053]	-0,635*** (0,140) [-0,046]	-0,772 (0,497) [-0,059]	-0,640 (0,535) [-0,133]	-0,689 (0,548) [-0,126]
<i>DESEMP</i>				0,018 (0,036) [0,001]	0,009 (0,048) [0,0007]	-0,076 (0,091) [-0,157]	-0,015 (0,107) [-0,003]
<i>PDESEMP1</i>					0,019 (0,067) [0,001]	0,038 (0,072) [0,008]	0,147* (0,079) [0,027]
<i>ΔDESEMP</i>						0,058 (0,067) [0,012]	-0,008 (0,006) [-0,014]
<i>INFLAC</i>							0,003 (0,019) [-0,0006]
<i>PINFLAC1</i>							-0,066*** (0,023) [-0,012]
<i>ρ</i>		0,072 (0,032)					
<i>Ln L</i>	-509,44	-969,96	-498,21	-498,08	-498,04	-374,70	-365,33
<i>Haus.Teste</i>		199,85					
<i>N°Obs.</i>	1449	1925	1449	1449	1449	1200	1200

Tabela 5.4.1-1: Coeficientes Estimados nas Regressões

Notas: As notações, ***, ** e *, reportam níveis de significância de 1%, de 5% e de 10%, respectivamente, para a rejeição da hipótese nula ($\beta = 0$), significando, ainda, a omissão dessa notação a ausência de significância estatística para a rejeição da hipótese nula.⁵² Os valores entre parênteses curvos correspondem aos erros padrão para cada um dos coeficientes estimados e os valores entre parênteses rectos reportam os efeitos marginais estimados.

Observar a rejeição das sucessivas estimações pelo método painel efeitos aleatórios, recomendada pelos resultados do teste de Hausman, referida previamente à apresentação da tabela de resultados 5.4.1-1, induz a encerrar a temática relativa ao modelo painel de efeitos aleatórios. Há, no entanto, uma informação disponibilizada por esse modelo, referente ao teste de Rácio de Verosimilhança, ($H_0 : \rho = 0$) que incita a uma inflexão nessa decisão. Na verdade a rejeição da hipótese nula admite, implicitamente, concluir que $\sigma_{\mu}^2 \neq 0$.⁵³ A relevância desta conclusão reside na possibilidade de se inferir que se rejeita a hipótese nula de homogeneidade dos efeitos específicos, ($\mu_i = \mu$), o que permite superar a omissão de informação relativa ao LR teste para o modelo *logit*, condicional. Eventualidade, que decorre de não ser possível utilizar o teste do Rácio de Verosimilhança, no cenário de uma estimação *logit* condicional, consequência do problema dos parâmetros acidentais, já que as funções de verosimilhança do *pooled logit* e do condicional *logit* não são comparáveis e daí, o papel substituto do teste de Hausman [Greene (2003), pp. 699].

Efectivamente, os valores do teste do Rácio de Verosimilhança, para os efeitos específicos aleatórios vêm, 5,32*** ; 3,76** , 2,74** , 2,53* 0 e 0 , relativamente a cada uma das regressões. Os primeiros quatro resultados ao permitirem rejeitar a hipótese nula de não existência de efeitos aleatórios, para o nível de significância de 1% , ***, para o nível de significância de 5% , ** e para o nível de significância de 10% , * , respectivamente, acabam por despistar o problema que decorreria de se ter que cotejar

⁵² O nível de significância é fundamentado nos valores dos *z-estatísticos*, calculados utilizando os erros padrão Hiber/White e a covariância.

⁵³ Já que, $\rho = \sigma_{\mu}^2 / (\sigma_{\mu}^2 + \sigma_v^2)$.

a estimação *pooled logit* (agregada) com a estimação *logit* condicional.⁵⁴ O mesmo não acontece com as duas últimas regressões, que ao não validarem a rejeição da hipótese nula impõem que se confronte a estimação *pooled logit* (agregada) com a estimação *logit* condicional.⁵⁵

Após este parêntese suscitado pela estatística do teste de Rácio de Verosimilhança para os efeitos específicos aleatórios, havia que tornar aos resultados dos coeficientes estimados para as diversas variáveis do modelo (constantes da tabela 5.4.1), que decorrem das seis regressões ensaiadas sob a técnica do método *logit* efeitos fixos e decidir qual delas ajustaria melhor o modelo em estimação.

A preferência acabou por recair sobre a regressão ensaiada no sexto momento, decisão suportada nos valores do critério de informação Akaike (*AIC*),⁵⁶ e nos valores do critério de informação Bayesiano (*BIC*).⁵⁷ A indicação dessa estimação *logit* efeitos fixos decorre do facto desses índices apresentarem, aí, o valor mais baixo, conforme se ilustra,

	Logit FE(1)	Logit FE(2)	Logit FE(3)	Logit FE(4)	Logit FE(5)	Logit FE(6)
<i>AIC</i>	1026,87	1006,41	1008,17	1010,09	765,40	750,65
<i>BIC</i>	1047,99	1032,80	1039,84	1047,04	806,12	801,56
Nº Obser.	1449	1449	1449	1449	1200	1200

Tabela 5.4.1-2: Valores *AIC* e *BIC*

⁵⁴ O teste de Hausman ao permitir a rejeição da hipótese nula, aponta a preferência pelo método dos efeitos fixos, o que ao deixar subjacente uma correlação significativa entre os efeitos específicos, μ_i e as variáveis exógenas x_{it} , possibilita a extensão da preferência do método dos efeitos fixos sobre o método ordinário.

⁵⁵ Da aferição entre essas duas estimações, resultam, para a estimação agregada, valores *AIC* e *BIC* de, 1627,82 e 1687,31, respectivamente. Valores bastante superiores aos valores *AIC* e *BIC* apurados para as regressões (5) e (6), tabela 5.4.2., que favorecem a opção pela estimação *logit* condicional.

⁵⁶ Índice utilizado para escolher entre modelos concorrentes, definido como, $AIC = 2(-\ln L + k)$, em que $-\ln L$ é a função de log-verosimilhança maximizada, em valores negativos e k o número de parâmetros estimados.

⁵⁷ Este índice, também, se baseia na maximização da função logarítmica de verosimilhança, vem definido por, $BIC = 2(-\ln L + \frac{k}{2} \ln N)$, representando N , o número de observações.

Ainda, antes de se analisarem os coeficientes estimados, importa relevar a solidez dos coeficientes estimados para as variáveis que prevalecem no decurso de qualquer uma das seis regressões estimadas, quer ao nível dos sinais, quer ao nível da significância estatística, como os casos da,

(i) variável *VOTOSANT*, em que o coeficiente estimado é sempre positivo e estatisticamente significativo;

(ii) variável *APART*, com um coeficiente estimado sempre negativo e estatisticamente significativo;

(iii) variável *APRES*, em que o coeficiente estimado é sempre negativo e sem significância estatística;

(iv) variável $DIREAL_{it}$, em que o coeficiente estimado é sempre positivo e estatisticamente significativo;⁵⁸

(v) variável D_{it} ,⁵⁹ em que o coeficiente estimado se mantém sempre negativo, mas estatisticamente significativo, unicamente, nas regressões (2) e (3); na primeira as variáveis macroeconómicas (desemprego e inflação) não vêm modeladas e na segunda, apenas, a taxa de desemprego vem ponderada.

Após ressalva da característica referenciada no parágrafo anterior, há espaço para se orientar a observação para a regressão *seis*, podendo-se sumariar que,

(i) o sinal (+) e a significância estatística *** do coeficiente estimado associado a $VOTOSANT_{it}$, vêm conforme a intuição e coincidem nesses itens com o obtido aquando da modelização suportada na variável dependente contínua *VV1*, percentagem de votos obtida pelo partido em exercício de funções executivas, na data de eleição autárquica (conforme subsecções 4.3.2.3. e 4.5.2.). O valor do efeito marginal mostra que por cada ponto percentual adicional na percentagem de votos obtida no acto eleitoral anterior a probabilidade do incumbente ganhar a eleição aumenta em 1,3 pontos percentuais;

⁵⁸ Ainda que se passe de uma significância de 1%, regressão (1), para uma significância de 10%, regressão (6).

⁵⁹ Ainda que não seja considerada na regressão (1) pode considerar-se de elevado grau de predominância nas estimações apresentadas.

(ii) o sinal do coeficiente estimado associado a $APART_{it}$ ao apresentar-se negativo (-) aponta para um impacto negativo da variável na probabilidade de reeleição (o que poderá ser explicado pelo desgaste associado ao tempo de permanência do partido no poder, conforme conclui Veiga e Veiga [2007a e 2007b]); facto, também, ocorrido na modelização em domínio de variável contínua (subsecções 4.3.2.3. e 4.5.2.), só que aí a ausência de significância estatística admitia que não se lhe imputasse uma valorização determinante na formulação do modelo.⁶⁰ O valor do respectivo efeito marginal mostra que para cada ano a mais do partido no poder a probabilidade de ganhar a eleição diminui em 4 pontos percentuais;

(iii) o coeficiente estimado associado à variável $APRES_{it}$ apresenta sinal negativo mas não é estatisticamente significativo;⁶¹

(iv) o coeficiente estimado para a variável $DIREAL_{it}$, despesa de investimento real, *per capita*, vem com sinal negativo (+), ainda que perdendo significância estatística face ao ocorrido nas outras regressões, indiciando um potencial de oportunismo a partir das despesas de investimento no ano das eleições. Na verdade a despesa de investimento real, *per capita*, influencia positivamente a probabilidade de reeleição. Por via do valor do respectivo efeito marginal pode-se concluir que, por cada euro adicional em despesas de investimento reais, *per capita*, a probabilidade do incumbente ganhar a eleição cresce em 0,03 pontos percentuais;

(v) o sinal (-) do coeficiente estimado relativo à variável *dummy* de coincidência do partido político nos poderes, local e central, deixa espaço para inferir um potencial de penalização sobre o incumbente local, caso pertença ao mesmo partido do incumbente nacional. Esta conjectura é tanto mais relevante quanto a robustez estatística alcançada por esta variável, na regressão (2), onde não foram consideradas as variáveis macroeconómicas relativas ao desemprego e à inflação, o que pode indiciar capacidade do votante para imputar a responsabilidade dos resultados da política económica aos níveis correctos de poder. A focagem sobre os valores da significância estatística,

⁶⁰ A multicolinearidade potencial entre esta variável $APART$ e a variável $APRES$, já abordada na modelização da variável dependente contínua, é equacionada. Correndo a regressão, com apenas uma das variáveis alternadamente, os valores dos coeficientes estimados, os sinais e as significâncias estatísticas não conhecem oscilações passíveis de referência.

⁶¹ Quando se equaciona o modelo apenas com esta variável importa assinalar que o respectivo coeficiente estimado apresenta-se estatisticamente significante.

entretanto, perdida no seio da regressão escolhida (6), que equaciona as variáveis macroeconómicas, pode, também, sinalizar a mesma aptidão do votante para a responsabilização da entidade exacta. Esta tese parece vir corroborada, por via da robustez estatística alcançada pelos coeficientes estimados para os termos de interacção dessas variáveis macroeconómicas com a variável *dummy* de coincidência do partido nos poderes, local e central. Nas regressões onde a variável é estatisticamente significativa, a coincidência partidária inter-poderes tem um impacto negativo de cerca de 5 pontos percentuais na probabilidade de reeleição;

(vi) os coeficientes estimados associados à taxa de desemprego, em nível e em variação, $DESEMP_{it}$ e $\Delta DESEMP_{it}$, respectivamente, apresentam-se com sinal (-), ainda que sem significância estatística;

(vii) o coeficiente estimado relativo à variável desemprego em interacção com a *dummy* de coincidência partidária dos governos, local e central, apresenta-se com sinal (+), coerente com o sinal do mesmo coeficiente estimado em cenário de variável dependente contínua (subsecções 4.3.2.3. e 4.5.2.). O sinal positivo (+) do coeficiente indicia que o desemprego majora a probabilidade de reeleição, caso o partido do autarca coincida com o partido do governo, contrariando a intuição - importa, no entanto, um aparte sobre a taxa de desemprego, a taxa de desemprego nacional pode não reflectir a taxa de desemprego local.⁶² Relativamente aos efeitos marginais, interessa anotar que um aumento de um ponto percentual na taxa de desemprego nacional nos municípios em que o autarca pertence ao partido no governo aumenta 2,7 pontos percentuais a probabilidade de reeleição do incumbente⁶³

e, finalmente,

(viii) os coeficientes estimados para a taxa de inflação e para a variável interacção da taxa de inflação e *dummy* coincidência partidária dos governos, central e local, apresentam sinais (+) e (-), respectivamente. Apenas o coeficiente para o termo de

⁶² Apresenta diversidade espacial decorrente da própria diversidade espacial do tecido empresarial.

⁶³ De novo este valor causa estranheza, *a priori*, até se desejaria que fosse negligenciável, ou mesmo não discernível, parece no entanto compatível com a realidade económica-geográfica do país, senão atente-se no facto dos desvios da taxa de desemprego local relativamente à taxa de desemprego nacional se fazerem sentir tanto mais positivamente quanto maior a dependência local do emprego público e do valor atingido pelo desemprego nacional. Municípios com baixa empregabilidade nos sectores de actividade económica privada serão menos afectados pela subida da taxa de desemprego, nesse momento o *handicap* da baixa empregabilidade funciona como elemento de consolação, que, talvez por efeito de informação assimétrica, acaba a reverter a favor do autarca desses municípios.

interacção da taxa de inflação e da *dummy* coincidência de “*cor*” política revela robustez estatística. Este cenário de resultados, permite a leitura de que a taxa de inflação, caso coincidam os partidos do autarca e do primeiro-ministro, prejudica a reeleição do autarca (o elemento espacial não se coloca na taxa de inflação, a variabilidade que, também, lhe está subjacente decorre da variabilidade dos cabazes de compras em função dos diferentes indivíduos). No que aos efeitos marginais respeita, os resultados sugerem que um aumento de um ponto percentual na taxa de inflação nacional diminui a probabilidade de reeleição dos autarcas com a mesma “*cor*” partidária que o governo central em 1,2 pontos percentuais.

5.5. Medidas de Avaliação do Ajustamento do Modelo à Base de Dados

5.5.1. McFadden Pseudo- R^2

Este indicador fornece uma medida de avaliação do ajustamento do modelo aos dados.⁶⁴ Vindo definido, como,

$$Pseudo - R^2 = 1 - \frac{L(\max imo)}{L(\beta = 0)}$$

Em que $L(\max imo)$ é o valor maximizado do logaritmo da função de verosimilhança para o modelo estimado e $L(\beta = 0)$ é o valor do logaritmo da função de verosimilhança quando os coeficientes/declive se assumem iguais a zero. A medida do $Pseudo - R^2$ sugere, intuitivamente, que o ajustamento possa ser tanto melhor quanto mais elevado o seu valor.⁶⁵ É, também, possível definir um valor para o $Mc - Fadden R^2$ _Ajustado [McFadden (1974)] como se segue,

⁶⁴ Também denominado de índice rácio de verosimilhança.

⁶⁵ Uma vez que o intervalo de variação do $Pseudo_R^2$ se situa no intervalo $[0,1]$: se todos os coeficientes estimados, do modelo, fossem zero, ter-se-ia $L(\max imo) = L(\beta = 0)$, atingindo-se o limite inferior desse intervalo, mas o limite superior só se alcançaria à

$$Pseudo - R^2 = 1 - \frac{L(\text{maximo}) - k}{L(\beta = 0)} \quad ^{66}$$

No modelo em estudo o valor do *Mc - Fadden* R^2 é de **0,23**, valor razoável, que não permite, mas também não impede, afirmação peremptória de que se está face a um mau ajustamento do modelo. O elemento importante a reter é que os coeficientes do modelo estimado não foram estimados para maximizar esta medida de ajustamento, como acontece nos modelos de regressão linear, em que os coeficientes β maximizam o R^2 .⁶⁷

5.5.2. Testes de Hosmer/Lemeshow

O teste de Hosmer/Lemeshow, *H - L* estatístico segue uma distribuição assintótica desconhecida, mas existe evidência empírica que mostra seguir uma distribuição do qui-quadrado com $J - 2$ ⁶⁸ graus de liberdade. Está-se a testar a hipótese nula de que os desvios entre expectativas (valores esperados) e as observações actuais tomam o valor de zero, isto é testa-se a hipótese de uma previsão perfeita concretizada pelo modelo. A rejeição da hipótese sugere que o modelo tem um desempenho pobre.

O software STATA utilizado para a estimação do modelo em estudo não gera este resultado, apenas o retrata em cenário de estimação *logit* agregada. Embora já se tenha concluído na subsecção 5.4.1., por via do *LR* teste, estar-se na presença de efeitos

medida que $L(\text{maximo}) \rightarrow 0$. Medidas observadas num intervalo deste tipo, $[0,1]$, não têm, infelizmente, uma interpretação natural, o que decorre da própria estrutura da equação do modelo *logit*.

⁶⁶ Em que k reflecte o número de parâmetros estimados no modelo.

⁶⁷ É razoável ter em mente que medidas de ajustamento em modelos de escolha binária não são importantes para avaliar a significância das variáveis explicativas, nem do ponto de vista estatístico, nem do ponto de vista económico [Wooldridge (2002)]. O teste de razão de verosimilhança e os testes individuais de significância, são tudo o que se necessita para concretizar essa avaliação (*LR* estatístico_[10] é de 217,83).

⁶⁸ J são os grupos em que se dividem os dados, por exemplo de acordo com os quantis desta estatística, apura-se o número de observações que tomam os valores $p=0$ e/ou os valores $p=1$, dentro de cada quantil.

específicos, o que aconselha a não estimação do modelo via agregada, mesmo assim ensaiou-se esse percurso. E, desse exercício de simulação, resultou um $H - L$ estatístico, com um valor de **5,64** e com **75%** de probabilidade de vir inferior ao valor crítico do $\chi^2_{[8]}$, desfecho que permite remeter para a aceitação da hipótese nula. Em suma, deste exercício de simulação, caso a estimação viesse suportada no modelo *pooled logit*, decorreria a possibilidade de se inferir que esse modelo ajustaria razoavelmente, pela comparação formal entre os valores observados e os valores esperados.

5.5.3. Percentagem de previsões correctas

Uma possível estratégia para avaliar o modelo estimado é analisar se o modelo classifica correctamente os indivíduos, já que muitos estudos recorrem ao número ou à percentagem de previsões geradas “correctamente” através do modelo empírico.

Para se obterem estes indicadores, começa-se por assumir um valor de 1, caso $F(\alpha + \beta X_i) \geq 0,5$ ⁶⁹ e um valor de 0, caso $F(\alpha + \beta X_i) < 0,5$, o que permite construir a tabela de resultados que se segue,

	$F(\alpha + \beta X_i) \geq 0,5$	$F(\alpha + \beta X_i) < 0,5$	Total
$p_i = 1$	1255	28	1283
$p_i = 0$	346	21	367
Total	1601	49	1650

Tabela 5.5.3: Valores das Previsões

⁶⁹ A fixação deste valor, 0,5 prende-se com o facto de nada se conhecer acerca da aleatoriedade da escolha seleccionada e assim assume-se, por defeito, o valor de 50% a probabilidade de um partido se fazer reeleger. Obviamente que esta opção induzirá dois tipos de erros, ao classificar incorrectamente zeros (0) como uns (1) e uns (1) como zeros (0). Uma eventual alteração no limite 0,5 reduziria a probabilidade de um tipo de erros, enquanto aumentava a probabilidade do outro tipo de erro. Não há uma resposta correcta sobre o melhor valor a escolher, depende do critério e da definição subjacente à regra de previsão.

O número de previsões correctas é retirado da diagonal principal⁷⁰ do quadro apresentado, perfazendo o total de 1276, o que em relação ao número de observações produz uma taxa de previsão correcta de, 77,33% , com uma sensibilidade de 97,82% e uma especificidade 5,72% .⁷¹

5.6. Testes de Especificação do Modelo

5.6.1. Teste do rácio verosimilhança

Para testar a significância total da relação estimada, pode, também, ser utilizado o *LR* - teste, o teste do Rácio Verosimilhança,⁷² que testa a hipótese nula, de que todos os coeficientes do modelo são nulos, isto é,

$$H_0 : \beta_i = 0.$$

Esta estatística segue uma distribuição do χ^2 com k graus de liberdade. Neste caso $k = 10$. O resultado, **153,88** , permite a rejeição da hipótese de omissão dos regressores, para um nível de significância estatística de 100%.⁷³

Dadas as preocupações com o potencial problema de multicolinearidade entre as variáveis *APART* e *APRES* optou-se por testar a hipótese de os respectivos coeficientes virem iguais, os resultados para o teste dos parâmetros vêm

⁷⁰ Corresponde às posições em que, a probabilidade prevista é maior ou igual a 0,5 e o valor observado $p_i = 1$ e a probabilidade prevista é menor que 0,5 e os valores observados $p_i = 0$.

⁷¹ A medida de sensibilidade resulta da $\Pr(F(\alpha + \beta X_i) \geq 0,5 | p_i = 1)$ e a de especificidade resulta da $\Pr(F(\alpha + \beta X_i) < 0,5 | p_i = 0)$.

⁷² $LR_teste = -2(L^R - L^U)$, em que L^R é o valor maximizado da função log-verosimilhança restrita, quando todos os regressores são omitidos, e L^U é o valor maximizado da função log-verosimilhança não restrita. O teste de Wald podia também ser utilizado, mas o *LR* teste permanece um dos mais populares para o efeito.

⁷³ O que corrobora a informação obtida a partir dos z -estatísticos.

consubstanciados numa distribuição do χ^2 com 1 grau de liberdade. O valor obtido de **15,13** permite rejeitar a hipótese nula em toda a dimensão estatística. Este resultado potencia a determinação da inclusão simultânea das duas variáveis no modelo desenhado. A alternativa de incluir, apenas, uma delas, poderia produzir um coeficiente inconsistente relativamente à variável incluída, ainda que a variável sobre a qual viesse a incidir a decisão de omissão não estivesse correlacionada com a variável sobre a qual viesse a incidir a decisão de inclusão.

5.6.2. Heteroscedasticidade nos modelos logit

A heteroscedasticidade, em contexto de estimação *logit*, causa problemas, não só ao nível da variância condicional, mas também ao nível da média condicional. Concretamente, se a heteroscedasticidade estiver presente e não for equacionada, então os coeficientes estimados (e os correspondentes efeitos parciais) virão inconsistentes. Um formato para que o problema seja considerado é o recurso a erros padrão robustos, isto é a erros que permaneçam consistentes mesmo na presença de heteroscedasticidade. Ora, esta opção, aquando da estimação do modelo *logit*, em estudo,⁷⁴ é acautelada.

O problema com esta estratégia é o da consistência lógica, pois, como foi referenciado, caso se esteja na presença de heteroscedasticidade, ainda que se recorra aos erros padrões robustos, a função estimada da média condicional é inconsistente. Consciente destas restrições optou-se por estimar o modelo de resposta binária (*regressão 6*) a partir do método *bootstrap*, alternativa não paramétrica à estimativa assintótica (erros padrão robustos).⁷⁵ Deste exercício de despistagem de eventuais problemas de heteroscedasticidade, passíveis de iludirem as prestações dos valores estimados dos erros padrão, resultaram para os mesmos valores de coeficientes estimados, valores de erro padrão muito próximos ou mesmo coincidentes, conforme se pode visualizar na tabela abaixo,

⁷⁴ O programa STATA dispõe do comando com a opção *robust*, que permite contemplar os erros padrão robustos.

⁷⁵ O verdadeiro valor dos erros padrão vêm sobrestimados por este método não paramétrico, enquanto a fórmula assintótica os subestima. Entretanto, para pequenas amostras, as estimações logísticas *bootstrap* vêm mais próximas do valor verdadeiro que as estimações logísticas assintóticas.

β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
0,069***	-0,221***	-0,034	-0,001*	0,689	0,015	0,147*	-0,008	0,003	-0,066***
(0,011)	(0,026)	(0,028)	(0,0009)	(0,548)	(0,107)	(0,079)	(0,006)	(0,019)	(0,023)
(0,014)	(0,033)	(0,036)	(0,001)	(0,616)	(0,104)	(0,099)	(0,066)	(0,016)	(0,022)

Tabela 5.6.2-1: Valores do erro padrão para os coeficientes estimados

Nota: Entre parênteses os valores dos erros padrão, a negrito os erros padrão *bootstrap* e a simples os erros padrão assintóticos.

Tendo em conta que a presença de qualquer tipo de heteroscedasticidade e de correlação contemporânea consegue ser controlada através de uma regressão aparentemente não relacionada (*seemingly unrelated regression, SUR*) e dado que o *software* estatístico, STATA, disponibiliza essa opção, já ao nível da pós-estimação, aplicou-se ao modelo essa estimação. Os resultados obtidos apresentam-se de molde a permitir afastar a hipótese de uma má especificação do modelo por causa da presença de heteroscedasticidade, conforme revela a tabela abaixo,

β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
0,069***	-0,221***	-0,034	-0,001*	0,689	0,015	0,147*	-0,008	0,003	-0,066***
(0,011)	(0,026)	(0,028)	(0,0009)	(0,548)	(0,107)	(0,079)	(0,006)	(0,019)	(0,023)
(0,011)	(0,028)	(0,031)	(0,001)	(0,550)	(0,098)	(0,096)	(0,056)	(0,018)	(0,023)

Tabela 5.6.2-2: Valores do erro padrão para os coeficientes estimados

Nota: Entre parênteses os valores dos erros padrão, a negrito os erros padrão *SUR* e a simples os erros padrão assintóticos.

5.7. Identificação de Ciclos Político-Económicos a Nível Local

Uma vez assumida a opção *logit* efeitos fixos (*FEM*) e feita a avaliação do ajustamento do modelo à base de dados e a aplicação dos testes de especificação passou-se a inferir

os resultados para as probabilidades de reeleição. Relativamente aos valores estimados da probabilidade apresenta-se uma descrição estatística na tabela abaixo e os valores, município a município, surgem no apêndice **H**, ao lado das probabilidades calculadas no capítulo **4**, secção **4.6.**.

Média	Máximo	Mínimo	Desvio Padrão
66,58	99,03	2,49	25,76

Tabela 5.7-1: Valores da estatística descritiva

Seguindo a metodologia utilizada, na secção **4.7.**, aquando da exploração do potencial explicativo da variável probabilidade de reeleição de um autarca, como variável exógena, para apuramento da existência (ou não) de ciclos político-económicos, revisitou-se o modelo estimado nessa secção. O modelo esboçado para esse efeito, equação **(4.7.1)**, mantém-se, apenas, a variável probabilidade de reeleição, \mathbf{Pr}_{it} , vem substituída pela agora calculada em contexto de modelo binário. Na tabela seguinte são apresentados os resultados obtidos na estimação das seguintes equações:

$$E(1) \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \text{PERCapita}_{it-1} + \gamma_3 * \mathbf{Pr}_{it} + \varepsilon_{it}$$

$$E(1') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Peso}_{it-1} + \gamma_2 * \Delta\text{Peso}_{it-2} + \gamma_3 * \mathbf{Pr}_{it} + \varepsilon_{it}$$

$$E(1'') \text{ PERCapita}_{it} = \gamma_0 + \gamma_1 * \Delta\text{Media}_{it-1} + \gamma_2 * \Delta\text{Media}_{it-2} + \gamma_3 * \mathbf{Pr}_{it} + \varepsilon_{it}$$

Em que, $\Delta\text{Peso}_{it-\ell} = (\text{PERCapita}_{it-\ell} - \text{PERCapita}_{it-\ell-1})/\text{PERCapita}_{it-\ell-1}$ e

$$\Delta\text{Media}_{it-\ell} = (\text{PERCapita}_{it-\ell} - \text{PERCapita}_{\text{Media}_{it}})/\text{PERCapita}_{\text{Media}_{it}} .$$

	$MQG_F^{(a)(b)(c)}$ White (1)	$MQG_F^{(a)(b)(c)}$ White (1')	$MQG_F^{(a)(b)(c)}$ White (1'')
γ_1	0,694*** (42,33)	-0,0005 (-1,24)	3,446*** (26,63)
γ_2		-0,012*** (-2,53)	-0,574*** (-5,24)
γ_3	-0,014*** (-7,29)	-0,034*** (-8,17)	-0,046*** (-14,62)
$R^2_{Ajust.}$	0,832	0,265	0,642
Teste Wooldridge	$F_{[1,274]} = 1,95$ ***	$F_{[1,263]} = 2,08$ ***	$F_{[1,274]} = 15,44$ *
NºObs.	1652	1425	1600

Tabela 5.7-2: Resultados das estimações relativas à equação 4.7.1

Nota: A apresentação das regressões (2), (3), (2'), (3'), (2'') e (3'') é dispensada, uma vez que nessas regressões nada se altera face ao apresentado na tabela 4.7.1.; ^(a) significa que a componente do erro foi considerada apenas a uma dimensão, a componente específica seccional; ^(b) estimação pelo método fiável dos mínimos quadrados generalizado, com ponderação seccional; ^(c) estimação com correcção *White diagonal* e entre parênteses os valores do *t*-estatístico; as notações, ***, **, reportam níveis de significância de 1% e de 5%, respectivamente para a rejeição da hipótese nula, $\beta = 0$, e a omissão desse tipo de notação identifica ausência de significância estatística, para a rejeição da hipótese nula.

O $R^2_{Ajust.}$ das equações (1') e (1'') vem mais baixo como era expectável por se incluírem diferenças entre variáveis nessas equações, no entanto o valor do *F* – estatístico é suficientemente alto para inferir que as regressões são estatisticamente significantes.

No teste de Wooldrige, *** traduz rejeição da hipótese de não auto-correlação de 1ª ordem ao nível de significância de 5% para as regressões (1) e (1') e ao nível de significância de 1% para a regressão (1'').

Observando os coeficientes estimados para as variável relativa à probabilidade de reeleição, **Pr**, equacionada como variável explicativa nas equações (1) (1') e (1''), regista-se um comportamento, para esses coeficientes γ_3 , quer ao nível do sinal, quer

ao nível do significado estatístico,⁷⁶ similar ao apurado aquando das estimações realizadas no âmbito da secção 4.7., permitindo que se continue a manter uma coincidência entre o sinal (-), estatisticamente robusto, do coeficiente estimado para \Pr e as inferências do modelo teórico.

Ao nível dos outros coeficientes, o padrão de comportamento, ao nível do sinal e da robustez estatística, segue na perfeição o observado nas estimações similares realizadas sobre as mesmas equações, no âmbito da secção 4.7.; ou seja para γ_1 , relativo à variável de encargos de endividamento, quer seja a do ano imediatamente anterior ao do acto eleitoral, equação (1), quer seja a da taxa de variação dos encargos financeiros do ano eleitoral face aos do ano imediatamente anterior, equação (1'), quer seja a da taxa de variação dos encargos financeiros desfasado de um período face ao valor médio do ciclo eleitoral, equação (1''), e para γ_2 , relativo à variável de encargos de endividamento, quer seja a da taxa de variação dos encargos financeiros do ano anterior ao eleitoral face aos do ano imediatamente anterior, equação (1'), quer seja a da taxa de variação dos encargos financeiros desfasado de dois períodos face ao valor médio do ciclo eleitoral, equação (1'').

Em suma, reconhece-se, em qualquer uma destas regressões, uma perfeita similaridade entre as características dos seus coeficientes e as características dos estimados nas regressões homólogas, desenvolvidas na secção 4.7.,⁷⁷ seja em termos quantitativos, seja a nível de sinais, seja a nível de significância estatística. A ser assim, neste cenário, onde os valores relativos à probabilidade de reeleição passaram a ser calculados em contexto de estimação *logit* e no que a estes coeficientes diz respeito, as conclusões a retirar em nada divergem das deduzidas na secção 4.7..

⁷⁶ O valor absoluto do coeficiente γ_3 decresce em cenário de probabilidades estimadas via *logit*, o que pode advir da variável binária dependente se construir a partir da ocorrência reeleição (não reeleição).

⁷⁷ Onde, apenas, diverge o procedimento de apuramento dos valores da variável probabilidade de reeleição.

5.8. Conclusão

Face ao exposto, os resultados obtidos, independentemente do método subjacente à estimação da probabilidade, apontam para que os encargos financeiros variem em sentido inverso ao valor da probabilidade. Perante uma elevada probabilidade de vir a ser reeleito o político no poder prefere contenção no valor dos encargos financeiros, denotando preocupação em não sobrecarregar financeiramente o seu próximo mandato. Com baixa expectativa de vitória, essa atitude preventiva dissipa-se, o político arrisca.

O raciocínio subjacente à metodologia que sustenta estes resultados perspectiva a probabilidade como instrumento de informação no processo de tomada de decisão do político relativamente à política económica. Essa probabilidade é assumida como variável exógena, ao invés de vir determinada, endogenamente, pela avaliação ex-ante do eleitorado acerca do desempenho futuro do incumbente.

É consensual aceitar que o(s) incumbente(s) utiliza(m) a política económica para potenciar as perspectivas de reeleição, embora a eficácia da mesma não esteja *a priori* garantida.⁷⁸ Esta é uma perspectiva de probabilidade que a toma como uma grandeza a ser sistematicamente manipulada pelos políticos, este ensaio, porém, procurou observar a probabilidade numa óptica diversa.

Explorar essa vertente do efeito da probabilidade de reeleição, exigia a respectiva estimação. Tarefa concretizada a partir dos resultados das eleições, inspirada no trabalho de Chappell e Keech (1988) e, ainda, através de um modelo de escolha binária, para, posterior, inserção como variável exógena num modelo em que os *encargos financeiros* da autarquia, em termos *per capita*, aparecem como variável dependente.

A coerência dos resultados empíricos torna pertinente que se continue a reflectir sobre as questões suscitadas pelo modelo teórico e, também, colocadas no âmbito da secção 4.7., por isso finalizarmos, questionando,

⁷⁸ Os resultados económicos decorrentes das políticas accionadas pelo(s) incumbente(s) podem ser observados como sinais de competência mas também como sinais de suspeição, passando a afectar negativamente a votação.

** Afinal, o autarca acaba mesmo a escolher o comportamento politicamente correcto?
Ou seja decide não agir de forma oportunista!*

** Será que não arrisca comportamentos mais activos (isto é estratégias de acção com
consequências na futura gestão orçamental) quando tem mais convicções de ganhar?*

** Ou ainda, será que o autarca teme a avaliação dos seus eleitores, e não quer ser
penalizado?*

** Isto é, o autarca receia que os seus eleitores possam perceber como irresponsáveis as
acções decorrentes de uma opção estratégica mais vigorosa!*

Bibliografia

Ahn, S. C. e P. Schmidt, 1995, Efficient estimation of models for dynamic panel data, *Journal of Econometrics*, Vol. 68, pp. 5–27.

Aidt, T. S., F. Veiga e L.Veiga, 2007, Election Results and Opportunistic Policies: An integrated Approach. NIPE WP 24/2007, EEG, Universidade do Minho.

Alesina, A., G. Cohen, e N. Roubini, 1992, Macroeconomic policy and elections in OECD democracies, *Economics and Politics*, Vol.4, pp. 1–30.

Alesina, A., G. Cohen e N. Roubini, 1993, Electoral business cycle in industrial democracies, *European Journal of Political Economy*, Vol. 9, pp. 1–23.

Alvarez, R.M. e J. Nagler, 1998, Economics, entitlements and social issues: Voter choice in the 1996 presidential election, *American Journal of Political Science*, Vol. 42, pp. 1349-1363.

Amemiya, T., 1971, The Estimation of the Variances in a Variance-Components Model, *International Economic Review*, Vol. 12, pp. 1-13.

Amemiya, T., 1985, *ADVANCED ECONOMETRICS*, Harvard University Press, Cambridge, MA.

Anderson C. J., 2000, Economic voting and political context: a comparative perspective. *Electoral Studies*, Vol. 19, pp. 151–170.

Arellano, M. e O. Bover, 1995, Another Look at the Instrumental Variable Estimation of Error Component Models, *Journal of Econometrics*, Vol. 68, pp. 29-51.

Armstrong-Taylor, P., 2005, Strategic Credibility, *Phd Economics thesis*, Harvard University, Cambridge, Massachusetts.

Arrow, K. J., 1963, (1951, 1st ed.) SOCIAL CHOICE AND INDIVIDUAL VALUES, 2nd ed., John Wiley & Sons, Inc., New York.

Baleiras, R. N. e J. S. Costa, 2004, To Be or Not To Be in Office Again: an Empirical Test of a Local Political Business Cycle Rationale, *European Journal of Political Economy*, Vol. 20 (3), pp. 655–671.

Baleiras, R. N., 1997, Political Economy in Local Governments, *Tese de Doutoramento em Economia*, Universidade Nova de Lisboa.

Baltagi, B.H., 2005, ECONOMETRIC ANALYSIS OF PANEL DATA, 3rd ed., John Wiley and Sons.

Banerjee A., M. Marcellino e C. Osbat, 2004, Some cautions on the use of panel methods for integrated series of macro-economic data, *Econometrics Journal*, Vol. 7, pp. 322–340.

Barro, Robert, 1973, The Control of Politicians: An Economic Model, *Public Choice*, Vol. 14, pp. 19-42.

Beck, N. e J. N. Katz, 1995, What to do (and not to do) with time series cross-section data, *American Political Science Review*, Vol. 89 (3), pp. 634-647.

Berge, C., 1963, TOPOLOGICAL SPACES, Oliver and Boyd, Edimburgh.

Besancenot, D. e R. Vranceanu, 2009, Strategic managerial dishonesty and financial distress, *Research in Economics*, Vol. 63 (1) pp. 11-21.

Bowman, D., 1999, Efficient Tests for Autoregressive Unit Roots in Panel Data, *manuscript*, Federal Reserve Board.

Breitung, J., 2000, The Local Power of Some Unit Root Tests for Panel Data, in B. Baltagi (ed.), NON STATIONARITY PANELS, PANEL COINTEGRATION AND DYNAMIC PANELS, ADVANCES IN ECONOMETRICS, Jai Press, Amsterdam, Vol. 15, pp. 161-178.

Breusch, T. e A. Pagan, 1980, The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies*, Vol. 47, pp. 239–253.

Butler, J.S e R. Moffitt, 1982, A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model, *Econometrica*, Vol. 50 (3), pp. 761-764.

Carlsen, F., 1997, Opinion Polls and political business cycles: Theory and evidence for the United States. *Public Choice*, Vol. 92, pp. 387-406.

Chamberlain, G., 1980, Analysis of Covariance with Qualitative Data, *Review of Economic Studies*, Vol. 47, pp. 225–238.

Chappel, H., 1990, Economic performance, voting, and political support: A unified approach. *Review of Economics and Statistics*, Vol. 72: 313–320.

Chappell, H. e Keech W, 1988, The unemployment rate consequences of partisan monetary policies, *Southern Economic Journal*, Vol. 55, pp. 107-122.

Cohen, G., 1993, Pre- and Post- Electoral Macroeconomic Fluctuations, Ph.D. *Dissertation*, Department of Economics, Harvard University.

Ferejohn, J., 1986, Incumbent performance and electoral control, *Public Choice*, Vol. 50, pp. 5-26.

Fisher R, 1932, STATISTICAL METHODS FOR RESEARCH WORKERS, 4th ed. Edinburgh.

Freedman, D. A. e S. C. Peters, 1984, Bootstrapping an economic model: some empirical results, *Journal of Business and Economic Statistics* Vol. 22, pp. 150–158.

Freedman, D.A. e S.C. Peters, 1984, Some notes on the bootstrap in regression problems, *Journal of Business and Economic Statistics*, Vol. 24, pp. 406–409.

Frees, E.W., 1995, Assessing cross-sectional correlation in panel data", *Journal of Econometrics*, Vol. 69, pp. 393-414.

Friedman, M., 1937, The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance, *Journal of the American Statistical Association*, Vol. 32, pp. 675-701.

Fuller W. A. e G. E. Battese, 1973, Transformations for estimation of linear models with nested error structure, *Journal of the American Statistical Association*, Vol. 68, pp. 626–632.

Fuller W. A. e G. E. Battese, 1974, Estimation of linear models with cross-error structure. *Journal of Econometrics*, Vol. 2, pp. 67–78.

Goeree, J. e C. Holt, 2001, Ten little treasures of game theory and ten intuitive contradictions, *American Economic Review*, Vol. 91, pp. 1402-1422.

Green, W., 2003, *ECONOMETRIC ANALYSIS*, 5th ed., Prentice Hall.

Green, W., 2004, Convenient Estimators for the Panel Probit Model: Further Results, *Empirical Economics*, Vol. 29, pp. 21-47.

Grier, K., 1987, Presidential elections and federal reserve policy: An empirical test, *Southern Economic Journal*, Vol. 54, pp. 475–486.

Grier, K., 1989, On the existence of a political monetary cycle. *American Journal of Political Science*, Vol. 33, pp. 376–389.

Grier, K. e H. Neiman, 1987, Deficits, politics and money growth, *Economic Inquiry*, Vol. 24, pp. 201–214.

Guilkey, D. K. e J. L. Murphy, 1993, Estimation and testing in the random effects probit model, *Journal of Econometrics*, Vol.59, pp. 301-317.

Hadri, K., 2000, Testing for stationarity in heterogeneous panel data, *Econometrics Journal, Royal Economic Society*, Vol. 3(2), pp. 148-161.

Hansen P. e A. King, 1998, Health care expenditure and GDP: panel data unit root test results – comment, *Journal of Health Economics*, Vol. 17, pp. 377–381.

Hausman, J. A., 1978, Specification Tests in Econometrics, *Econometrica*, Vol. 46(6), pp. 1251-1271.

Hausman, J. A. e W. E. Taylor, 1981, Panel Data and Unobservable Individual Effects, *Econometrica*, Vol. 49(6), pp. 1377-1398.

Heckman, J., 1981, STATISTICAL MODELS FOR DISCRETE PANEL DATA. IN STRUCTURAL ANALYSIS FOR DISCRETE DATA: WITH ECONOMETRIC APPLICATIONS, C. F. Manski e D. McFadden (eds.), The MIT Press, Cambridge, Mass.

Heckman, J., 2000, Causal parameters and policy analysis in economics: A twentieth century retrospective, *The Quarterly Journal of Economics*, Vol. 115, pp. 45-97.

Holmström, B., 1982, Managerial Incentive Problems – A Dynamic Perspective, In ESSAYS IN ECONOMICS AND MANAGEMENT IN HONOR OF LARS WAHLBECK, Helsinki: Swedish School of Economics.

Hsiao, C., 1996, Logit and Probit Models, Chapter 16 in: THE ECONOMETRICS OF PANEL DATA, L. Matyas e P. Sevestre eds., Kluwer Academic Publishers.

Hsiao, C., 1986, ANALYSIS OF PANEL DATA, 1st ed., Cambridge University Press, Cambridge.

Huber, P. J., 1967, The Behavior of Maximum Likelihood Estimation Under Nonstandard Conditions, in L. M. LeCam e J Neyman (eds.), PROCEEDINGS OF THE

FIFTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY,
University of California Press, Berkeley, CA.

Im, K, H. Pesaran e Y. Shin, 1997, Testing for unit roots in heterogeneous panels.
University of Cambridge, DAE working paper, n° 9526.

Im, K, H. Pesaran e Y. Shin, 2003, Testing for unit roots in heterogeneous panels.
Journal of Econometrics, Vol. 115, pp. 53–74.

Johnston, J., J. DiNardo, 1997, ECONOMETRIC METHODS, chapter 12, "*Panel Data*",
McGrawHill, Companies Ltd., Toronto.

Kang S., 1985, A note on the equivalence of specification tests in the two-factor
multivariate variance components model, *Journal of Econometrics*, Vol. 28, pp. 193–
203.

King, G., 1989, UNIFYING POLITICAL METHODOLOGY: THE LIKELIHOOD THEORY OF
STATISTICAL INFERENCE, Cambridge eds.

Kirchgassner, G., 1985, Rationality, causality, and the relation between economic
conditions and the popularity of parties: An empirical investigation for the Federal
Republic of Germany, 1971-1982, *European Economic Review*, Vol. 28(1-2), pp. 243-
268.

Kmenta, J., 1986, ELEMENTS OF ECONOMETRICS, 2nd ed., Macmillan, New York, NY.

Kramer, G. H., 1971, Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964,
American Political Science Review, Vol. 65, pp. 131–143.

Kreps, D. M. e R. B. Wilson, 1982, Reputation and imperfect information, *Journal of
Economic Theory*, Vol. 27, pp. 253–279.

Krishna Vijay, R., 2007, Communication in Games of Incomplete Information: the two-
player case, *Journal of Economic Theory*, Vol. 132, pp.584 – 592.

Kwiatkowski, D., P. Phillips, P. Schmidt e Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of unit root, *Journal of Econometrics*, Vol. 54, pp. 159-178.

Lancaster K. J., 1966, A New Approach to Consumer Theory, *Journal of Political Economy*, Vol. 74, pp. 132-157.

Lancaster, T., 2000, The incidental parameter problem since 1948, *Journal of Econometrics*, Vol. 95, pp. 391- 413.

Levin, A. e C.-F. Lin, 1993, Unit Root Tests in Panel Data: New Results, *Discussion Paper*, N° 93-56, Department of Economics, University of California.

Levin A., C.-F. Lin e C. Chu, 2002, Unit roots tests in panel data: Asymptotic and finite-sample properties, *Journal of Econometrics*, Vol. 108, pp. 1-24.

Lohmann S., 1998, Federalism and central bank independence: The politics of German monetary policy, 1957-1992, *World Politics*, Vol. 50, pp. 401-446.

Maddala G. S. e S. Wu, 1999, A Comparative Study of Panel Data Unit Root Tests and a Simple Alternative, *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 631-652.

Maddala, G. S., 1999, On the Use of Panel Data Methods with Cross-Country Data, *Annales d'Économie et de Statistique*, Vol. 55-56, pp. 429-449.

Maddala G. S., 1998, Recent Developments in Dynamic Econometric Modelling: A Personal Viewpoint, *Political Analysis*, Vol. 7 (1), pp. 59-87.

Maddala, G. S., 1987, Limited Dependent Variable Models Using Panel Data, *Journal of Human Resources*, Vol. 22 (3), pp. 305-338.

Maddala, G.S., 1971, The Use of Variance Components Models in Pooling Cross Section and Time Series Data, *Econometrica*, Vol. 39(2), pp. 341-358.

Magnac, T., 2005, Logit models of individual choices, prepared for the NEW PALGRAVE.

McCoskey, S. K. e M. Selden, 1998, Health care expenditures and GDP: panel data unit root test results, *Journal of Health Economics*, Vol. 17(3), pp. 369-376.

McFadden, D. F., 1974, Conditional Logit Analysis of Quantitative Choice Behavior, in: P. Zarembka, ed., *FRONTIERS IN ECONOMETRICS*, Academic Press, New York, pp. 105-142.

Mundlak, Y., 1978, On the Pooling of Time Series and Cross Section Data, *Econometrica*, Vol. 46, pp. 69-85.

Nannestad, P. e M. Paldam, 1994, The VP-function: A Survey of the Literature on Vote and Popularity Functions after 25 Years, *Public Choice*, Vol. 79(3-4), pp. 213-245.

Nelson C. R. e C. I. Plosser, 1982, Trends and random walk in macroeconomic time series, *Journal of Monetary Economics*, Vol. 10, pp. 139-162.

Nerlove, Marc, 1971, Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections, *Econometrica*, Vol. 39(2), pp. 359-382.

Neyman J. e E. L. Scott, 1948, Consistent estimates based on partially consistent observations, *Econometrika*, Vol. 16, pp. 1-31.

Ng S. e P. Perron, 2005, A Note on the Selection of Time Series Models, *Oxford Bulletin of Economics and Statistics*, Vol. 67(1), pp. 115-134.

O'Connell, P., 1998, The overvaluation of purchasing power parity, *Journal of International Economics*, Vol. 44, pp. 1-19.

Paldam, M., 2004, Are Vote and Popularity Functions Economically Correct? In Rowley, C. K. and Friedrich Schneider, Eds., *THE ENCYCLOPEDIA OF PUBLIC CHOICE*, Vol. I, pp. 49-59, Kluwer Academic Publishers.

Parks, R.W., 1967, Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated, *Journal of the American Statistical Association*, Vol. 62 (2), pp. 500-509.

Pedroni, P., 2000, Fully Modified OLS for Heterogeneous Cointegrated Panels, in Baltagi, B. e C.D. Kao, Eds., *ADVANCES IN ECONOMETRICS, NONSTATIONARY PANELS, PANEL COINTEGRATION AND DYNAMIC PANELS*, Elsevier Science, New York, pp. 93-130.

Pedroni, P., 2001, Purchasing Power Parity Tests in Cointegrated Panels, *The Review of Economics and Statistics*, Vol. 83, pp. 727-731.

Pedroni, P., 1997, On the Role of Cross Sectional Dependency in Panel Unit Root and Panel Cointegration Exchange Rate Studies, *Working paper*, Indiana University.

Persson, T. e G. Tabellini, 1990, *MACROECONOMIC POLICY, CREDIBILITY, AND POLITICS*, Harwood Academic Publishers.

Persson, T. e G. Tabellini, 2000, *POLITICAL ECONOMICS: EXPLAINING ECONOMIC POLICY*, Cambridge, MA: MIT Press.

Pesaran, M. H. e R. Smith, 1995, Estimating Long-Run Relationships form Dynamic Heterogeneous Panels, *Journal of Econometrics*, Vol. 68, pp. 79-113.

Pesaran, M. H., 2004, General Diagnostic Tests for Cross Section Dependence in Panels. *IZA Discussion Papers* 1240, Institute for the Study of Labor, IZA.

Powell G. B. e G. D. Whitten, 1993, A cross-national analysis of economic voting: taking account of the political context, *American Journal of Political Science*, Vol. 37, pp. 391-414.

Robinson, P. M., 1982, On the asymptotic properties of estimators of models containing limited dependent variables, *Econometrica*, Vol. 50 (1), pp. 27-42.

Rogoff, K., 1990, Equilibrium Political Budget Cycles, *American Economic Review*, Vol. 80, pp. 21-36.

Sarafidis, V., T. Yamagata e D. Robertson, 2006, A Test of Cross Section Dependence for a Linear Dynamic Panel Model with Regressors, *mimeo*, University of Cambridge.

Sen, A., 1990, Rational Behavior, In Eatwell, J., M. Milgate e P. Newman, UTILITY AND PROBABILITY, New York: W. W. Norton & Company, pp. 1998-2216.

Suzuki, M., 1992, Political business cycles in the public mind, *American Political Science Review*, Vol. 86, pp. 989-996.

Swamy, P. e S. Arora, 1972, The Exact Finite Sample Properties of the Estimates Components Regression Models, *Econometrica*, Vol. 40. pp. 261–275.

Train, K., 1993, QUALITATIVE CHOICE ANALYSIS, MIT Press, Cambridge, MA.

Veiga, F. e L. Veiga, 2004a, Economia, Popularidade e Intenções de Voto em Portugal: uma Análise Longitudinal com Dados Agregados, in Freire, A., M. Costa Lobo e P. Magalhães, PORTUGAL A VOTOS: ELEIÇÕES LEGISLATIVAS DE 2002, Imprensa de Ciências Sociais, Lisboa, pp. 125-157.

Veiga, F. e L. Veiga, 2004b, Popularity functions, partisan effects and support in Parliament. *Economics & Politics*, Vol. 16(1), pp. 101-115.

Veiga, F. e L. Veiga, 2004c, The Determinants of Vote Intentions in Portugal. *Public Choice*, Vol. 118(3-4), pp. 341-364.

Veiga, L. e F. Veiga, 2007a, Political Business Cycles at the Municipal Level. *Public Choice*, Vol. 131 (1/2), pp. 45-64.

Veiga, L. e F. Veiga, 2007b, Does Opportunism pay off?, *Economics Letters*, Vol. 96, pp. 177-182.

Wallace, D. e A. Hussain, 1969, The use of error components model in combining cross-section and time-series data, *Econometrica*, Vol. 37, pp. 55-72.

White, H., 1980, A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity, *Econometrica*, Vol. 48, pp. 817–838.

Wooldridge, J.M., 2002, *ECONOMETRIC ANALYSIS OF CROSS SECTION AND PANEL DATA*, MIT Press: Cambridge, MA.

Wrasai, P.T., 2005, Agency Problems in Political Decision Making, *Doctor Thesis* from the Erasmus University Rotterdam, ISBN 90 5170 614 6, Book n° 368.

Wrasai, P., 2005, Politician's Motivation, Role of Elections, and Policy Choices, Tinbergen Institute, *Discussion Paper No. 2005-050/1*.

Comentários Finais

Este trabalho, que as vicissitudes da Vida ajudaram a dilatar no tempo, parece chegado ao limite de uma etapa. De propósito a terminologia *fim* não foi utilizada, sustentada em duas razões, sintetizadas numa só: nada na Vida está terminado, enquanto a Vida não finda, os *inícios* progridem sobre os *fins*, com a cadência de um mecanismo perfeito e aperfeiçoado pelo conhecimento, pela experiência e pela sensatez, decorrentes da chama da inquietação, privilégios da nossa existência.

É nesta óptica que encaro este longo processo de crescimento científico, por isso ocuparei algumas linhas com alguns rumos para futuras incursões, com o intuito de perspectivar o amadurecimento de *hoje* em trabalho do *amanhã*.

O motor para o desenvolvimento da parte I capturou, a título de empréstimo, a ideia original de Ely (1888, 131)¹ *“The fundamental idea of our tax system is a democratic one. It is, that all should contribute to the support of government in proportion to their capacity or “respective abilities,” as Adam Smith expresses it. It is, however, assumed that one’s ability to contribute to the support of government is measured by the actual selling value of all one’s property, real and personal; then it is further assumed that it is possible in each case to discover the selling value of all the property of citizens”*

Ainda que, a determinação do valor e do preço das propriedades imobiliárias tenha, nas três últimas décadas, sido objecto de um vasto conjunto de estudos e nessa proliferação de estudos empíricos, mais de 10.000 artigos econométricos, se suportem na teoria hedónica dos preços implícitos, mesmo assim este sucessivo fluxo de pesquisa melhorou o poder preditivo dos modelos de avaliação ou, no mínimo, permitiu uma melhor compreensão dos factores específicos que influenciam o valor da propriedade. Neste enquadramento, um balanço sobre os preços hedónicos revela-o como um método analítico relativamente poderoso sugerindo mais vantagens do que desvantagens. Na verdade, o método desenvolvido por Rosen, há mais de 30 anos, tem sabido evoluir e

¹ Ely, R. T., in TAXATION IN AMERICAN STATES AND CITIES, New York, Crowell & Co.

tem sabido adaptar-se a novas mudanças metodológicas, muitas vezes emprestadas por outras disciplinas, ganhando pois eficiência e consistência. O sucesso dos preços hedónicos na economia urbana e na economia do imobiliário (*real estate*) decorre da repercussão do trabalho conceptual de Rosen, bem como da transparência desta perspectiva.

Ainda que, a força motivadora do trabalho desenvolvido na parte I possa ter amortecido pela entrada em vigor do novo imposto sobre o património (CIMI), mesmo assim subsiste o interesse em redireccioná-lo para um trabalho com uma amplitude maior, qual gérmen de um índice de preços para a habitação. A habitação apresenta-se, cada vez mais, como um bem de consumo, cuja oscilação de valor reage cada vez mais a efeitos qualidade, intrínsecos e extrínsecos, e a efeitos mercado, daí que um índice de preços suportado no modelo hedónico se apresente mais poderoso.

A análise temporal de um índice de preços do imobiliário apenas pode produzir observação comparativa consistente caso a qualidade desse imobiliário permaneça fixa, a não ser assim a evolução desse índice de preços decorre de alterações na composição dos imóveis e de alterações no preço dos imóveis.² Ora é na captação destas alterações que o modelo hedónico se pode qualificar como um bom suporte para a criação de um índice de preços capaz de ilustrar com rigor o desenvolvimento do mercado imobiliário em Portugal.³ Parece-me residir nesta valência o potencial futuro desta metodologia agora estudada.

Dado que os modelos hedónicos surgiram como necessidade de apontar um caminho para a avaliação da base fiscal de um imposto, o imposto sobre o património, pode-se por via da tributação fazer a ponte para a parte II deste trabalho. A tributação identificada como um tópico crucial das finanças públicas é um lugar-comum, mas a

² O Relatório Boskin (1996) sugere que medições incorrectas dos efeitos de qualidade num período de rápido progresso tecnológico são indutoras de sobrestimação nos valores da inflação.

³ Na Alemanha, sob os auspícios do Federal Office for Building and Regional Planning (BBR), a pedido do Bundesbank, está a desenvolver-se um moderno método hedónico de cálculo de preços do imobiliário, que começará a ser publicado a partir de Agosto, conjuntamente com o tradicional HPX índice, Hypoport House Price Index. Em Portugal, o Índice Confidencial Imobiliário, a mais longa série sobre imobiliário, dado remontar a Janeiro de 1988, apresenta-se como indicador de inflação do valor da habitação em oferta no Continente, e refere a utilização de procedimentos de ajustamento de qualidade (via estratificação e estimação de preços hedónicos) para minimizar o impacto da alteração temporal do mix de oferta.

tributação é resultado de princípios políticos (Holcombe, 1998) e ignorar este facto conduz a uma incompleta compreensão da tributação. De acordo com Wicksell (1896), a literatura da escolha pública encara a tributação como a contrapartida do fornecimento de serviços e bens públicos, em que um equilíbrio de Nash é atingido sob o mecanismo de decisão de escolha colectiva. Como resultado, diferentes factores institucionais (políticos), nomeadamente, a democracia, representativa ou directa, e as correspondentes regras de votação geram diferentes sistemas fiscais. Assim os modelos de votação concedem suporte para a compreensão teórica das decisões fiscais e das estruturas fiscais.

É com recurso a North, que se norteiam trilhos futuros, no âmbito do trabalho desenvolvido nesta parte II. Assim conquistados pelo conceito instituições, “*Institutions are the humanly devised constraints that structure political, economic and social interaction, which consist of formal rules (constitutions, laws, and property rights), informal constraints (sanctions, taboos, customs, traditions and codes of conduct) and their enforcement characteristics.*”,⁴ e após reflexão sobre os resultados obtidos ao nível dos valores da probabilidade de reeleição, consolidamos a ideia de que estudos desta natureza podem fortalecer uma dinâmica de reformulação das instituições do país (um novo desenho do poder local, redimensionado de forma a reflectir a (re)centralização facultada pelas novas tecnologias de informação, a (re)concentração do tecido demográfico, a eficiência do serviço público e a eficácia da estrutura fiscal.

Embora, o termo “instituição” tenha uma longa história de pesquisa, para sobre essa noção alguma nebulosidade. Numa analogia do processo económico a um jogo, as instituições têm sido encaradas como, regras do jogo, participantes no jogo ou estratégias de equilíbrio para os participantes no jogo. Neste sentido a estrutura fiscal é indubitavelmente o centro da instituição formal do Estado, desde que os impostos não ajudem apenas a criar Estado mas o ajudem a aperfeiçoar, em suma o sistema fiscal é o órgão que acomoda todos os outros órgãos (Schumpeter, 1918:108).

⁴ North, D. C., 1991 e 1994, Institutions, Institutional Change and Economic Performance, *Journal of Economic Perspectives* Vol. 5, pp. 97-112 e Economic Performance through Time, *American Economic Review*, Vol. 84, pp. 359-368.