ISSN 0873-4275

INSTITUTO NACIONAL DE ESTATÍSTICA

PORTUGAL

REVISTA DE ESTATÍSTICA

STATISTICAL REVIEW

VOLUME III

3º QUADRIMESTRE 2002

REVISTA DE ESTATÍSTICA

STATISTICAL REVIEW

ÍNDICE

INDEX

MENSAGEM DO DIRECTOR

MESSAGE FROM THE DIRECTOR

MENSAGEM DO DIRECTOR

O lançamento em 1996, por iniciativa do Instituto Nacional de Estatística (INE), da REVISTA DE ESTATÍSTICA, veio preencher uma evidente lacuna.

Na verdade, sendo o INE o ponto focal da produção das estatísticas oficiais portuguesas, compreendia-se mal que não dispusesse de uma publicação periódica adequada à divulgação, entre profissionais do mesmo ofício, de trabalhos de investigação e análise no domínio da Estatística, quer enquanto ramo das ciências matemáticas de larga aplicação na gestão e na tomada de decisões pelos agentes económicos, quer enquanto instrumento ao serviço do conhecimento da realidade social e económica.

Desejou assim o INE que a REVISTA DE ESTATÍSTICA viesse a constituir um veículo prestigiado e, portanto, procurado pelos autores de trabalhos daquele tipo, para os tornar amplamente conhecidos pela comunidade dos estaticistas.

Daí que o INE tenha rodeado o lançamento da REVISTA DE ESTATÍSTICA dos cuidados considerados necessários para lhe assegurar uma Direcção profissional e um Conselho Editorial constituído por especialistas de reconhecido mérito capaz de fazer prevalecer critérios de exigência científica que se pretenderam, a um tempo, rigorosos e sensatos.

De salientar que a REVISTA DE ESTATÍSTICA ao longo dos seus 7 anos de existência editou 21 números, tendo publicado 115 artigos.

Como tive oportunidade de vir anunciando nos dois últimos números, com o presente número termina o INE a edição da REVISTA DE ESTATÍSTICA, mas por uma boa causa, para iniciar a edição de uma nova revista, verdadeiramente internacional, a **REVSTAT – STATISTICAL JOURNAL**.

É, pois, meu indeclinável dever deixar aqui, o que faço com honra extrema, uma lápide singela de reconhecido agradecimento à colaboração dos ilustres membros do Conselho Editorial da Revista de Estatística que tiveram a generosa benevolência de se dignarem aceitar o convite que em nome do INE lhes formulei, e que são:

- Prof. Dinis Duarte Ferreira Pestana
- Prof. João António Branco
- Prof. João Ferreira do Amaral
- Prof. Óscar Soares Barata
- Prof. Pedro Miguel Girão Nogueira Ramos
- Dr. António Daniel Correia dos Santos
- Dr. Francisco José Neto Melro
- Dr. Pedro Jorge Nunes da Silva Dias
- Dr. Sérgio Manuel Bacelar e Silva

Finalmente, aproveito para relembrar os principais elementos caracterizadores do formato da nova revista **REVSTAT - STATISTICAL JOURNAL**.

- A política editorial centra-se na publicação de artigos de investigação nos domínios das Probabilidades e Estatística, com ênfase na originalidade e importância da investigação, sendo os artigos objecto de avaliação científica por pelo menos duas pessoas, uma do Corpo de Editores e outra externa.
- A língua de trabalho é o Inglês.
- A periodicidade da edição é quadrimestral, Março, Julho e Novembro, sendo o primeiro número publicado no final de Março de 2003, sendo publicados em média 6 artigos por cada número.
- O **Corpo de Editores**, de índole internacional, é constituído por:
  - **Editor-Chefe**, Prof.ª *M. Ivette Gomes*, Faculdade de Ciências da Universidade de Lisboa, Portugal
  - **Co-Editor**, Prof.ª *M. Antónia Amaral Turkman*, Faculdade de Ciências da Universidade de Lisboa, Portugal
  - **Editor Executivo**, Dr. *Ferreira da Cunha*, Consultor da Direcção do Instituto Nacional de Estatística, Portugal
  - **Editores Associados**:
    - Prof. *António Pacheco*, Instituto Superior Técnico de Lisboa, Portugal
    - Prof. *Barry Arnold*, Universidade da Califórnia, Riverside, EUA.
    - Prof. *Dani Gamerman*, Universidade Federal do Rio de Janeiro, Brasil
    - Sir *David Cox*, Universidade de Oxford, Reino Unido
    - Prof. *Dinis Pestana*, Faculdade de Ciências da Universidade de Lisboa
    - Prof. *Edwin Diday*, Universidade de Paris IX Dauphine, França
    - Prof. *Gilbert Saporta*, Conservatório Nacional das Artes e Ofícios, Paris, França
    - Prof.ª *H. Bacelar Nicolau*, Faculdade de Psicologia da Universidade de Lisboa, Portugal
    - Prof. *Isaac Meilijson*, Universidade de Tel-Aviv, Israel
    - Prof. *Jef Teugels*, Universidade Católica de Lovaina, Bélgica
    - Prof. *João Branco*, Instituto Superior Técnico de Lisboa, Portugal
    - Prof. *Ludger Rüschendorf*, Instituto de Matemática Estocástica de Freiburg, Alemanha
    - Prof.ª *M. Lucília Carvalho*, Faculdade de Ciências da Universidade de Lisboa, Portugal
    - Prof.ª *Marie Husková*, Universidade Carlos de Praga, República Checa
    - Prof.ª *Nazaré Mendes-Lopes*, Faculdade de Ciências e Tecnologia da Universidade de Coimbra, Portugal
    - Prof. *Radu Theodorescu*, Universidade de Laval, Sainte-Foy, QC, Canadá
    - Prof.ª *Susie Bayarri*, Universidade de Valência, Espanha
  - **Secretária**, *Liliana Martins*, Instituto Nacional de Estatística, Lisboa, Portugal

- A submissão de artigos para publicação na REVSTAT significa que contêm um trabalho original ainda não publicado nem em vias de ser publicado sob qualquer forma algures.
- Os artigos deverão ser escritos em Inglês e poderão ser submetidos de três maneiras:
  - Enviando 3 cópias em papel ao Editor Executivo (ferreira.cunha@ine.pt), que estará em contacto permanente com o Editor-Chefe e o Co-Editor,
  - Enviando 2 cópias em papel ao Editor Executivo e uma cópia ao Editor-Chefe ou ao Co-Editor ou a um dos Editores Associados, cuja opinião o autor queira que seja tomada em linha de conta,
  - Enviando 1 cópia em papel ao Editor Executivo, conjuntamente com o envio do correspondente ficheiro PDF ou Postscript para o E-mail do Editor-Chefe: revstat@fc.ul.pt .
- Os manuscritos devem ser escritos numa só página em duplo espaço, com uma margem de pelo menos 3cm do lado esquerdo, e não podem ter mais de 30 páginas. A primeira página deverá incluir o nome, instituição e endereço do(s) autor(es), e um resumo com o máximo de 100 palavras, seguido das palavras-chave num máximo de 6, e a classificação do assunto AMS 2000.
- Os autores deverão escrever a versão final dos artigos aceites para publicação utilizando o LaTeX, no estilo da REVSTAT. O estilo (REVSTAT.sty), exemplos de ficheiros (REVSTAT.tex) e instruções adicionais aos autores poderão ser obtidas no endereço web da REVSTAT: http://www.ine.pt./revstat/.
- Os autores dos artigos aceites para publicação deverão enviá-los em ficheiro LaTeX, utilizando o REVSTAT.sty e o REVSTAT.tex, para a Secretária da REVSTAT (liliana.martins@ine.pt).
- Aceites os artigos para publicação, os respectivos autores serão solicitados a transferir o *copyright* dos mesmos para o Instituto Nacional de Estatística, a fim de permitir a sua maior divulgação possível, nomeadamente através do Website do Instituto (http://www.ine.pt).

MESSAGE FROM THE DIRECTOR

The launch, in 1996, at the initiative of the National Statistics Institute (INE), of the REVISTA DE ESTATÍSTICA (Journal of Statistics), filled an obvious gap.

In truth, since the INE is the focal point for the output of official Portuguese statistics, it was difficult to understand why it did not have a periodical to disseminate papers on statistical research and analysis to professionals of the same trade, both as a branch of the mathematical sciences for extensive use in the management and decision-taking process of economic agents, and as an instrument enabling a better understanding of social and economic reality.

The INE thus intended Revista de Estatística to be a prominent vehicle, sought by authors of papers on statistical research and analysis for widespread dissemination to the community of statisticians.

It follows that the INE put considerable time and effort into the launch of REVISTA DE ESTATÍSTICA to ensure that it had a professional Management and an Editorial Board made up of specialists of recognised merit capable of guaranteeing standards of scientific rigour, which are both accurate and sensible.

During its 7 years of existence, REVISTA DE ESTATÍSTICA published 21 issues and 115 articles.

As I have already had the opportunity to announce in the last two issues, this is the last issue of Revista de Estatística that the INE will publish, but there is a good reason for this, which is to commence publication of a new and truly international journal, **REVSTAT – Statistical Journal**.

It is therefore not only my indeclinable duty but also an extreme honour for me to say a simple word of appreciation in recognition of the contribution made by the illustrious members of the Editorial Board of Revista de Estatística who were generous enough to accept the invitation which I, on behalf of the INE, extended to them and who are:

- Prof. Dinis Duarte Ferreira Pestana
- Prof. João António Branco
- Prof. João Ferreira do Amaral
- Prof. Óscar Soares Barata
- Prof. Pedro Miguel Girão Nogueira Ramos
- Dr. António Daniel Correia dos Santos
- Dr. Francisco José Neto Melro
- Dr. Pedro Jorge Nunes da Silva Dias
- Dr. Sérgio Manuel Bacelar e Silva

Finally, I would like to take this opportunity to recall the main features characterising the format of the new publication, **REVSTAT - Statistical Journal**.

- The editorial policy of REVSTAT - Statistical Journal will now be centred on publishing research articles at the highest level in the domains of Probability and Statistics, stressing the originality and importance of the research, with all articles being subject to scientific evaluation by at least two persons, one from the Editorship and another, external.

- The working language is English.

- The journal will be published every four months, in March, July and November, with the first issue coming out at the end of March 2003 and an average of 6 articles being published in each issue.

- The Editorial Board, international in nature, will comprise:
  - **Editor-in-Chief**, Prof. *M. Ivette Gomes*, Faculty of Science at the University of Lisbon, Portugal
  - **Co-Editor**, Prof. *M. Antónia Amaral Turkman*, Faculty of Science, University of Lisbon, Portugal
  - **Executive Editor**, Dr. *Ferreira da Cunha*, Consultant to the Board of the National Statistics Institute, Portugal
  - **Associate Editors**:
    - Prof. *António Pacheco*, Higher Technical Institute of Lisbon, Portugal
    - Prof. *Barry Arnold*, University of California, Riverside, USA.
    - Prof. *Dani Gamerman*, Federal University of Rio de Janeiro, Brazil
    - Sir *David Cox*, Oxford University, United Kingdom
    - Prof. *Dinis Pestana*, Faculty of Science, University of Lisbon, Portugal
    - Prof. *Edwin Diday*, University of Paris IX Dauphine, France
    - Prof. *Gilbert Saporta*, National Conservatory of Arts and Crafts, Paris, France
    - Prof. *H. Bacelar Nicolau*, Faculty of Psychology, University of Lisbon, Portugal
    - Prof. *Isaac Meilijson*, University of Tel Aviv, Israel
    - Prof. *Jef Teugels*, Catholic University of Louvain (Leuven), Belgium
    - Prof. *João Branco*, Higher Technical Institute of Lisbon, Portugal
    - Prof. *Ludger Rüschendorf*, Institute of Stochastic Mathematics, Freiburg, Germany
    - Prof. *M. Lucília Carvalho*, Faculty of Science, University of Lisbon, Portugal
    - Prof. *Marie Husková*, Charles University of Prague, Czech Republic
    - Prof. *Nazaré Mendes-Lopes*, Faculty of Science and Technology, University of Coimbra, Portugal
    - Prof. *Radu Theodorescu*, University of Laval, Sainte-Foy, QC, Canada
    - Prof. *Susie Bayarri*, University of Valencia, Spain
  - **Secretary**, *Liliana Martins*, National Statistics Institute, Lisbon, Portugal

- In order to submit an article for publication in REVSTAT, it must contain original work that has not been published nor is it about to be published elsewhere in any form whatsoever.

- **The articles must be written in English and may be submitted in the following three ways:**

    - By sending 3 copies in paper form to the Executive Editor (ferreira.cunha@ine.pt), who will be in permanent contact with the Editor-in-Chief and the Co-Editor,

    - By sending two copies in paper form to the Executive Editor and one copy to the Editor-in-Chief or to the Co-Editor or to one of the Associate Editors, whose opinion the author would like to be taken into account,

    - By sending a paper copy to the Executive Editor, together with the corresponding PDF or postscript file to the following e-mail address: revstat@fc.ul.pt.

- Manuscripts must be written on one side only with double spacing, a left margin of at least 3 cm and must not contain more than 30 pages. The first page must include the name, institution and address of the author(s) and a summary of not more than 100 words, followed by a maximum of 6 key words and the AMS 2000 subject classification.

- The authors are required to write the final version of the articles accepted for publication using LaTeX, in the REVSTAT style. The style (REVSTAT.sty), examples of files (REVSTAT.tex) and additional instructions for the authors may be obtained at the REVSTAT web address: http://www.ine.pt/revstat.

- The authors of the articles accepted for publication must send them in a LaTeX file using REVSTAT.sty and REVSTAT.tex, to the REVSTAT Secretary (liliana.martins@ine.pt).

- Once the articles are accepted for publication, the respective authors will be asked to transfer copyright thereof to the National Statistics Institute, in order to enable the widest possible dissemination, namely through the Institute's website (http://www.ine.pt).

ARTIGOS

Revista de
Estatística
17ª
Página
3º Quadrimestre de 2002

O Uso de Regressores Dummy na Especificação de Modelos com Parâmetros Variáveis

Autores:
Patrícia Oom do Valle
Efigénio Rebelo

O USO DE REGRESSORES DUMMY NA ESPECIFICAÇÃO DE MODELOS COM PARÂMETROS VARIÁVEIS

THE USE OF DUMMY VARIABLES IN SPECIFYING MODELS WITH CHANGING COEFFICIENTS

Autores: Patrícia Oom do Valle
- Assistente da Faculdade de Economia da Universidade do Algarve; Área Científica de Métodos Quantitativos.

Efigénio Rebelo
- Professor Associado da Faculdade de Economia da Universidade do Algarve; Área Científica de Métodos Quantitativos.

Resumo:

- A formulação de um modelo com parâmetros variáveis constitui uma aplicação importante mas pouco divulgada dos regressores dummy na análise econométrica. O presente estudo sistematiza e analisa em detalhe três situações típicas em que o uso desta técnica se revela pertinente (o ajustamento de modelos com variações descontínuas nos parâmetros, a análise de sazonalidade e o ajustamento de modelos com variações contínuas nos parâmetros) e generaliza a sua utilização a modelos mais complexos que incluam um número qualquer de regressores quantitativos.

Palavras-chave:

- *Regressão linear, variáveis dummy, modelos com parâmetros variáveis.*

Abstract:

- The specification of a regression model with changing coefficients represents an important but less recognized application of dummy variables in econometrics. The current study summarizes and assesses the use of this technique into three standard contexts: the adjustment of models with discontinuous variations in parameters; the seasonality analysis; and the specification of models with continuous variations in parameters. The generalization of this technique to more complex models is also provided.

Key-Words:

- *Linear regression, dummy variables, models with Changing Coefficients.*

1. INTRODUÇÃO

A problemática da estimação de um modelo com parâmetros variáveis levanta-se sobretudo em estudos cronológicos que abranjam períodos de tempo de dimensão considerável em que, consequentemente, é pouco sustentável a hipótese de estabilidade dos parâmetros para todas as observações de uma determinada amostra. Aliás, o exemplo tradicionalmente referido para ilustrar a formulação de um modelo deste tipo consiste na análise da relação entre consumo e rendimento disponível durante um dado período de tempo que abranja uma guerra ou uma depressão económica profunda. Neste caso, é de esperar que o consumo autónomo e/ou a propensão marginal ao consumo no período que antecede estes acontecimentos sofra alguma alteração após o seu término. Outro exemplo clássico, traduz-se na modelação de um fenómeno em que exista sazonalidade, situação em que faz sentido admitir que o efeito das variáveis quantitativas na variável dependente seja variável com a estação do ano. Contudo, e como facilmente se compreende, também num estudo seccional pode ser adequado permitir a possibilidade de variabilidade nos parâmetros entre grupos de observações. Em estudos desta natureza, as variações nos valores dos coeficientes de regressão são atribuíveis às características dos indivíduos que constituem a amostra, como sejam a sua área de residência, profissão, sexo, grupo etário ou classe social a que pertençam.

Este trabalho apresenta e desenvolve diversas situações em que a definição de regressores dummy possibilita a estimação de modelos com parâmetros variáveis, evidenciando, deste modo, um contexto importante mas pouco divulgado destas variáveis na análise econométrica. Especificamente, três situações serão objecto de investigação: o ajustamento de modelos com variações descontínuas nos parâmetros, a análise de sazonalidade e o ajustamento de modelos com variações contínuas nos parâmetros.

2. VARIAÇÕES DESCONTÍNUAS NOS PARÂMETROS

Considere-se o estudo de um fenómeno económico (seccional ou cronológico) em que, numa primeira fase, se assume que a relação a entre variável dependente e a variável quantitativa é estável para todas as observações de uma dada amostra. Seja (1) a equação de regressão correspondente

$$y_i = \alpha + \beta\ X_i + u_i ; i = 1,2,\dots,n ; u_i \sim \mathrm{IIN}(0,\sigma^2) \qquad (1)$$

em que X_i é uma variável quantitativa.

Suponha-se, contudo, que uma reflexão mais cuidada acerca do fenómeno em causa sugere a existência de variabilidade nos parâmetros entre alguns grupos de observações. A primeira etapa na formulação de uma equação de regressão com parâmetros variáveis consubstancia-se na tomada de posição acerca da forma como se processam essas variações. Neste contexto, admita-se que da estimação do modelo de regressão (1), com base numa amostra de n observações, resultava um valor significativo para a estatística t associada ao coeficiente da variável X e, simultaneamente, um valor para o coeficiente de determinação (R^2) relativamente baixo ou um valor da D.W. longe de 2. Estes resultados fazem pensar que embora X constitua de facto um factor determinante no comportamento de y, existe ainda uma forte componente não explicada na variabilidade de y, ou dito de outra forma, que o modelo (1) pode encontrar-se mal especificado por incorrecta omissão de variáveis explicativas. Suponha-se que a figura 1 retracta a relação entre a variável dependente e o regressor X.

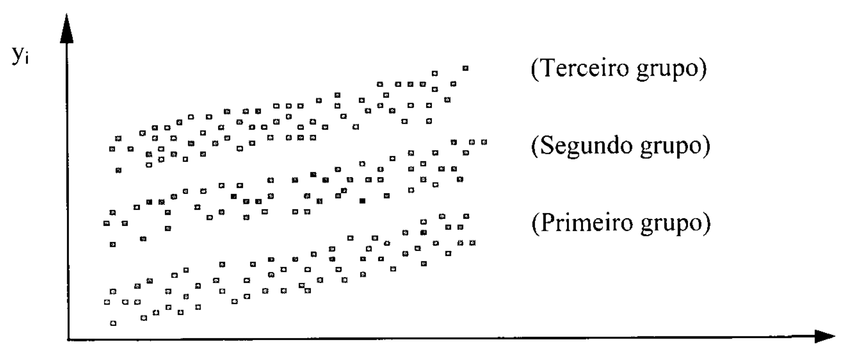

Figura 1. Observações de um hipotético estudo econométrico

A observação da figura não oferece dúvidas acerca do relacionamento (neste caso positivo) existente entre as variáveis em análise. Todavia, parece existir uma relação distinta entre as duas variáveis para as observações que pertençam a cada um dos grupos[1]. Esta circunstancia explica porque é que ao se ajustar uma única equação de regressão com todas as observações ocorre um valor relativamente baixo para R^2. É que, neste caso, a recta de regressão ajustada vai posicionar-se algures entre as três nuvens de pontos ocasionando resíduos com uma dimensão relativamente alta. Mais, uma situação de grande dispersão das observações em torno da recta, traduz-se obviamente numa estimativa de valor elevado para a variância da variável residual o que pode originar uma estatística t associada ao coeficiente da variável X não significativa. Nesta situação, concluía-se erradamente que a variável X deve ser desprezada como factor relevante do comportamento de y.

Uma vez justificada a insuficiente adequação do modelo (1) ao problema em análise, importa perspectivar procedimentos que permitam ter em atenção o gap entre os três grupos na relação entre a variável explicada e o regressor quantitativo.

Uma forma de solucionar o problema seria considerar em separado cada um dos grupos de observações e utilizá-los para ajustar três modelos de regressão distintos.

Como mostra a figura 1, as rectas de regressão que melhor se ajustam às nuvens de pontos parecem diferir apenas no termo independente pelo que, em termos formais, a sua estrutura deverá ser:

$$y_i = \alpha_1 + \beta X_i + u_i \quad \text{para o primeiro grupo} \qquad (2)$$

$$y_i = \alpha_2 + \beta X_i + u_i \quad \text{para o segundo grupo} \qquad (3)$$

$$y_i = \alpha_3 + \beta X_i + u_i \quad \text{para o terceiro grupo.} \qquad (4)$$

Contudo, da estimação dos modelos (2), (3) e (4) não resultará certamente o mesmo valor para β, o parâmetro que se assume ser comum às três especificações. Se de facto for razoável assumir que os três grupos reagem de forma similar a uma variação unitária na variável X, a atitude mais sensata do investigador será reunir todas as observações para ajustar um modelo de regressão que produza três termos independentes diferentes mas uma estimativa única para o coeficiente inclinação.

A definição de regressores dummy apresenta-se como o procedimento adequado à prossecução deste objectivo. Com efeito, a definição das variáveis

$$D_{2i} = \begin{cases} 1, & \text{se a observação verifica a característica que define o segundo grupo} \\ 0, & \text{caso contrário} \end{cases}$$

$$D_{3i} = \begin{cases} 1, & \text{se a observação verifica a característica que define o terceiro grupo} \\ 0, & \text{caso contrário} \end{cases}$$

e o posterior ajustamento da equação de regressão,

$$y_i = \alpha_1 + (\alpha_2 - \alpha_1)D_{2i} + (\alpha_3 - \alpha_1)D_{3i} + \beta X_i + u_i \,;\, i = 1,2,\ldots,n \,;\, u_i \sim IIN(0,\sigma^2) \qquad (5)$$

permitirá uma única estimativa para β e, simultaneamente, três ordenadas na origem distintas. Note que, quando $D_{2i} = D_{3i} = 0$, o modelo (5) reduz-se a:

$$y_i = \alpha_1 + \beta X_i + u_i \qquad \text{(primeiro grupo);}$$

quando $D_{2i} = 1$ e $D_{3i} = 0$,

$$y_i = \alpha_2 + \beta X_i + u_i \qquad \text{(segundo grupo);}$$

e, finalmente, quando $D_{2i} = 0$ e $D_{3i} = 1$

$$y_i = \alpha_3 + \beta X_i + u_i \qquad \text{(terceiro grupo).}$$

Quer dizer, a equação (5) é uma forma alternativa de representar conjuntamente as equações (2), (3) e (4) e torna evidente de que forma a aplicação de regressores dummy possibilita a estimação de um modelo com parâmetros variáveis, neste caso particular os termos independentes. É importante não esquecer que ao agrupar-se as três equações numa só está-se não só a assumir que o parâmetro β tem o mesmo valor para os três grupos mas também que o vector dos desvios, u_i, tem a mesma distribuição nas três situações. É necessário assumir também que existe o número de observações suficiente para estimar cada modelo per si.

O coeficiente da variável dummy D_2 representa a diferença entre os termos independentes das equações de regressão relativas aos dois primeiros grupos. Similarmente, o coeficiente da variável D_3 representa a diferença entre os termos independente dos modelos que dizem respeito ao primeiro e terceiro grupos. Esta situação seria ainda mais clara se o modelo (5) estiver representado na forma equivalente

$$y_i = \alpha_1 + \delta_2 D_{2i} + \delta_3 D_{3i} + \beta X_i + u_i \,;\, i = 1,2,\dots,n \,;\, u_i \sim \text{IIN}(0,\sigma^2) \qquad (6)$$

em que $\delta_2 = \alpha_2 - \alpha_1$ e $\delta_3 = \alpha_3 - \alpha_1$. Neste caso, ter-se-ia para cada grupo de observações o seguinte modelo de regressão linear

$$y_i = \alpha_1 + \beta X_i + u_i \quad \text{se} \quad D_{2i} = D_{3i} = 0 \qquad (7)$$

$$y_i = (\alpha_1 + \delta_2) + \beta X_i + u_i \quad \text{se} \quad D_{2i} = 1 \text{ e } D_{3i} = 0 \qquad (8)$$

$$y_i = (\alpha_1 + \delta_3) + \beta X_i + u_i \quad \text{se} \quad D_{2i} = 0 \text{ e } D_{3i} = 1 \qquad (9)$$

o que não deixaria dúvidas de que o termo independente assume um valor distinto consoante o grupo em análise (presumivelmente inferior para o primeiro grupo)[2]. Um teste de significância conjunta aos parâmetros δ_2 e δ_3 equivale a testar que não existe uma diferença significativa entre os termos independentes dos três grupos. A tradução geométrica da estrutura estimada do modelo (6) pode agora ser apresentada:

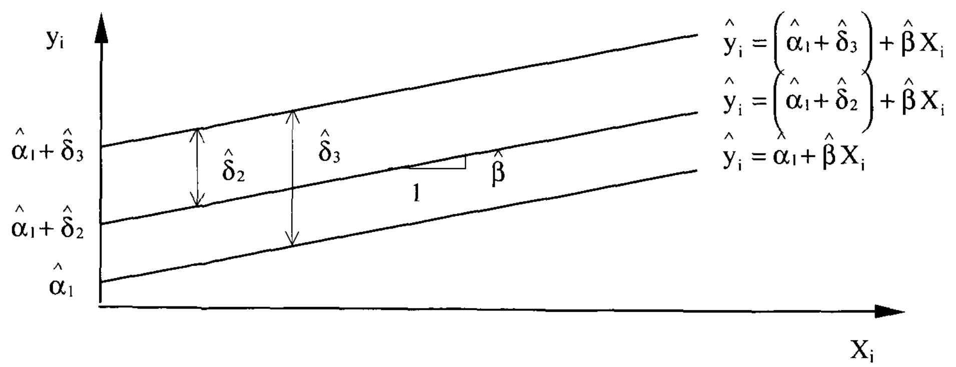

Figura 2. Tradução geométrica da estrutura estimado do modelo 6

O efeito da pertença a cada um dos grupos deverá ser incorporado de forma diferente na relação entre y e X se se assumir, por exemplo, que a ordenada na origem é fixa nas três situações e que apenas o declive das rectas de regressão apresenta variabilidade. Nesta conformidade, as rectas de regressão que melhor se adaptam às hipotéticas nuvens de pontos correspondentes definem-se como:

$$y_i = \alpha + \beta_1 X_i + u_i \qquad \text{para o primeiro grupo} \qquad (10)$$

$$y_i = \alpha + \beta_2 X_i + u_i \qquad \text{para o segundo grupo} \qquad (11)$$

$$y_i = \alpha + \beta_3 X_i + u_i \qquad \text{para o terceiro grupo} \qquad (12)$$

Os problemas enunciados acima relativamente ao ajustamento das três equações de regressão separadas são também aplicáveis neste caso particular pelo que a solução mais indicada é, mais uma vez, recorrer à definição de regressores dummy e ajustar um único modelo de regressão a partir da globalidade das observações que produza, por um lado, uma estimativa para o termo independente e, por outro, três coeficientes inclinação distintos. Este modelo é o seguinte

$$\begin{aligned} y_i &= \alpha + \beta_1 X_i + (\beta_2 - \beta_1)(D_{2i} X_i) + (\beta_3 - \beta_1)(D_{3i} X_i) + u_i \; ; \\ &; \; i = 1,2,\ldots n \; ; \; u_i \sim \mathrm{IIN}(0,\sigma^2) \end{aligned} \qquad (13)$$

em que D_2 e D_3 são as variáveis dummy anteriormente definidas, ou de forma equivalente

$$y_i = \alpha + \beta_1 X_i + \gamma_2 (D_{2i} X_i) + \gamma_3 (D_{3i} X_i) + u_i \; ; \; i = 1,2,\ldots n \; ; \; u_i \sim \mathrm{IIN}(0,\sigma^2) \qquad (14)$$

onde $\gamma_2 = \beta_2 - \beta_1$ e $\gamma_3 = \beta_3 - \beta_1$, tornando claro que o coeficiente de cada regressor dummy mede a diferença entre os declives de dois modelos de regressão.

A atribuição dos valores adequados às variáveis dummy permite determinar a equação de regressão relativa a cada grupo de observações. Na verdade, para o primeiro grupo,

$$y_i = \alpha + \beta_1 X_i + u_i$$

uma vez que neste caso $D_{2i} = D_{3i} = 0$, enquanto que para o segundo grupo

$$y_i = \alpha + \beta_2 X_i + u_i$$

dado que $D_{2i} = 1$ e $D_{3i} = 0$ e, por último, para o terceiro grupo,

$$y_i = \alpha + \beta_3 X_i + u_i$$

uma vez que $D_{2i} = 0$ e $D_{3i} = 1$. Estas duas últimas equações podem ainda ser apresentadas como:

$$y_i = \alpha + (\beta_1 + \gamma_2) X_i + u_i$$

e

$$y_i = \alpha + (\beta_1 + \gamma_3) X_i + u_i$$

se atender à formulação equivalente (14). Portanto, um teste à hipótese nula $H_0 : \gamma_2 = \gamma_3 = 0$ equivale a testar que não existem diferenças significativas entre os coeficientes inclinação das equações de regressão relativas aos três grupos de observações. A representação gráfica da estrutura estimada, que se adapta à situação em análise, é a que consta na figura 3.

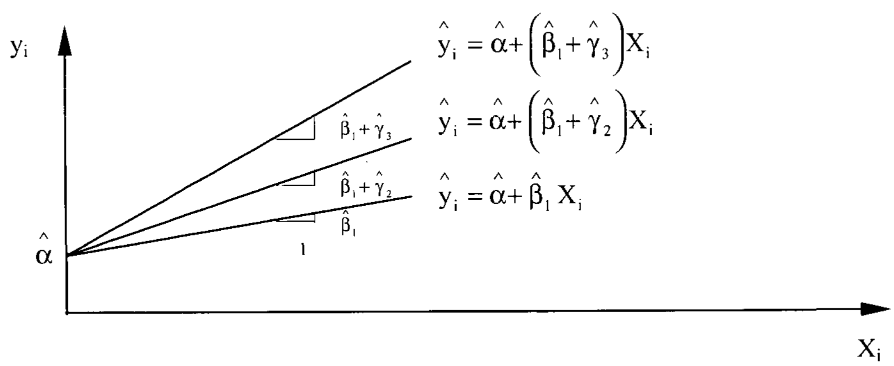

Figura 3. Tradução geométrica da estrutura estimada do modelo (14)

A figura 3 deixa claro que o coeficiente de cada regressor dummy mede a diferença entre os declives de duas equações de regressão separadas.

A análise precedente pode ser combinada para permitir que três equações de regressão com termos independentes e declives diferentes sejam ajustáveis através da especificação de um só modelo. Como facilmente se imagina, o modelo com regressores dummy que oferece esta possibilidade é

$$\begin{aligned} y_i &= \alpha_1 + (\alpha_2 - \alpha_1) D_{2i} + (\alpha_3 - \alpha_1) D_{3i} + \\ &+ \beta_1 X_i + (\beta_2 - \beta_1)(D_{2i} X_i) + (\beta_3 - \beta_1)(D_{3i} X_i) + u_i \; ; \\ &; i = 1,2,\ldots n \; ; \; u_i \sim IIN(0, \sigma^2) \end{aligned} \qquad (15)$$

ou, atendendo às definições acima apresentadas para $\delta_2, \delta_3, \gamma_2$ e γ_3,

$$y_i = \alpha_1 + \delta_2 D_{2i} + \delta_2 D_{3i} + \beta_1 X_i + \gamma_2 (D_{2i} X_i) + \gamma_3 (D_{3i} X_i) + u_i \; ; \quad ; \; i = 1,2,...n \; ; \; u_i \sim IIN(0, \sigma^2). \tag{16}$$

Mediante a afectação de valores aos regressores dummy, tem-se para cada grupo de observações:

$$y_i = \alpha_1 + \beta_1 X_i + u_i \qquad \text{quando } D_{2i} = D_{3i} = 0 \tag{17}$$

$$y_i = (\alpha_1 + \delta_2) + (\beta_1 + \gamma_2) X_i + u_i \quad \text{quando } D_{2i} = 1 \text{ e } D_{3i} = 0 \tag{18}$$

$$y_i = (\alpha_1 + \delta_3) + (\beta_1 + \gamma_3) X_i + u_i \quad \text{quando } D_{2i} = 0 \text{ e } D_{3i} = 1 \tag{19}$$

As estruturas estimadas correspondentes a esta situação encontram-se representadas na figura seguinte:

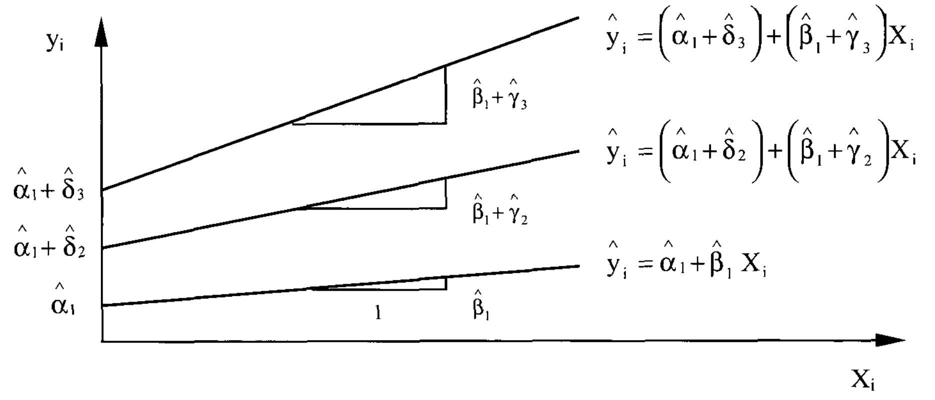

Figura 4. Tradução geométrica da estrutura estimada do modelo (16)

A partir dos E.M.Q.O produzidos pelo modelo (16) é possível obter com exactidão os E.M.Q.O. resultantes do processo de ajustamento das três equações de regressão separadas (17), (18) e (19), isto é:

$$\hat{\alpha}_1 + \hat{\delta}_2 = \hat{\alpha}_2, \quad \hat{\alpha}_1 + \hat{\delta}_3 = \hat{\alpha}_3, \quad \hat{\beta}_1 + \hat{\gamma}_2 = \hat{\beta}_2 \text{ e } \hat{\beta}_1 + \hat{\gamma}_3 = \hat{\beta}_3$$

As três situações específicas analisadas não esgotam o domínio de aplicação dos regressores dummy na especificação de modelos com parâmetros variáveis. Assim, admita-se, por hipótese, que existem fortes razões para pensar que o termo independente apenas varia do primeiro para o segundo grupo e que o declive só apresenta variabilidade entre os dois últimos grupos. Formalmente:

$$y_i = \alpha_1 + \beta_1 X_i + u_i \qquad \text{para o primeiro grupo} \qquad (20)$$

$$y_i = \alpha_2 + \beta_1 X_i + u_i \qquad \text{para o segundo grupo} \qquad (21)$$

$$y_i = \alpha_2 + \beta_2 X_i + u_i \qquad \text{para o terceiro grupo} \qquad (22)$$

A criação de regressores dummy pode ser utilizada para permitir que os coeficientes das equações de regressão (20), (21) e (22) sejam determináveis através da estimação de um só modelo de regressão. Esse modelo é o seguinte

$$y_i = \alpha_1 + (\alpha_2 - \alpha_1)D_{2i} + \beta_1 X_i + (\beta_2 - \beta_1)(D_{3i}X_i) + u_i \qquad (23)$$
$$i = 1,2,...n \; ; \; u_i \sim IIN(0,\sigma^2)$$

onde

$$D_{2i} = \begin{cases} 1, & \text{se a i-ésima observação verifica a característica que define o segundo ou o terceiro grupos} \\ 0, & \text{caso contrário} \end{cases}$$

$$D_{3i} = \begin{cases} 1, & \text{se a i-ésima observação verifica a característica que define o terceiro grupo} \\ 0, & \text{caso contrário} \end{cases}$$

ou, de forma equivalente:

$$y_i = \alpha_1 + \delta_2 D_{2i} + \beta_1 X_i + \gamma_2 (D_{2i}X_i) + u_i \; ; \; i = 1,2,...n \; ; \; u_i \sim IIN(0,\sigma^2) \qquad (24)$$

cuja estrutura estimada aparece representada na figura 5.

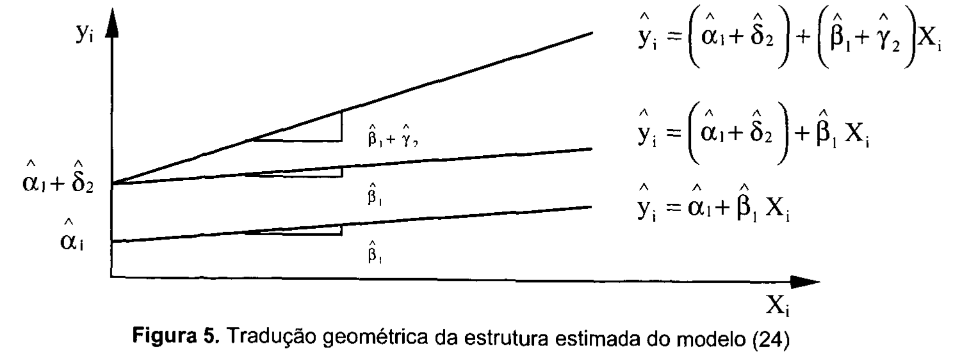

Figura 5. Tradução geométrica da estrutura estimada do modelo (24)

Neste caso, testar simultaneamente $H_0: \delta_2 = 0 \wedge \gamma_2 = 0$ equivale a testar a estabilidade dos parâmetros entre os três grupos de observações.

Outro tipo de modelo com parâmetros variáveis, aparece definido na literatura econométrica como "modelo de resposta assimétrica". Como refere Kmenta (1986, p.421), o objectivo de especificação de um modelo de resposta assimétrica é testar se "*...a variação na variável dependente que resulta de um aumento unitário de uma dada variável explicativa quantitativa é diferente em magnitude da variação que decorre de uma diminuição unitária dessa variável*". Como exemplo de uma situação que justifica a formulação de um modelo com esta particularidade, considere-se que pretendemos avaliar o poder explicativo do rendimento disponível (y_d) no consumo (C) de um conjunto de famílias num determinado ano. Ora, é plausível admitir que no ano em análise existam famílias que assistiram a um incremento do seu rendimento disponível relativamente ao ano anterior e outras que, pelo contrário, viram os seus rendimentos decrescerem. Existe uma teoria que defende que os consumidores não reagem de uma forma uniforme a uma variação unitária do rendimento, isto é, a dimensão do acréscimo de C que resulta de um aumento unitário de y_d é superior à magnitude da diminuição de C que decorre na sequência de uma redução unitária de y_d. Perante este cenário, a equação usualmente seleccionada para representar a relação entre as duas variáveis em causa, $C = \alpha + \beta y_d + u$, revela-se inadequada na medida em que, independentemente da natureza da variação unitária ocorrida em y_d (subida ou descida), o efeito em C tem dimensão constante (β). Para incorporar um comportamento de resposta assimétrica na equação de regressão basta definir o regressor dummy

$$D_i = \begin{cases} 1, & \text{se o } y_d \text{ da i-ésima família aumentou em relação ao ano anterior} \\ 0, & \text{caso contrário} \end{cases}$$

e ajustar o modelo:

$$C_i = \alpha + \beta y_{di} + \gamma(D_i y_{di}) + u_i \, ; i = 1,2,\ldots,n \, ; u_i \sim \text{IIN}(0,\sigma^2) . \tag{25}$$

Deste modo, a propensão marginal ao consumo das famílias que beneficiaram de um aumento do seu rendimento disponível é dada por $\beta + \gamma$ e a propensão marginal ao consumo das restantes famílias restringe-se a β. Para testar se de facto existe uma situação de resposta assimétrica é suficiente realizar um teste de significância individual ao parâmetro γ.

Como se constata, o modelo de resposta assimétrica não é mais do que um caso particular de um modelo que apenas admite variabilidade no coeficiente inclinação. Contudo, o contexto interessante em que se insere, em que a característica que acciona a variabilidade é a própria natureza da variação ocorrida no regressor quantitativo, justifica o destaque que lhe foi proporcionado neste trabalho.

3. ANÁLISE DE SAZONALIDADE

Outra aplicação muito comum dos regressores dummy na análise econométrica e que conduz, também, à especificação de um modelo com parâmetros variáveis, consubstancia-se na remoção de flutuações sazonais em séries cronológicas[3]. Com efeito, é frequente encontrar em estudos económicos temporais, sobretudo mensais ou trimestrais, movimentos sazonais significativos. Por exemplo, as vendas de uma grande diversidade de lojas tendem a disparar na época de Natal, a procura de dinheiro e de viagens tende a ser mais elevada nos períodos de férias e o consumo de refrescos é, em geral, superior no Verão.

Como salienta Gujarati (1992, p.273), "*o processo de remover a componente sazonal de uma série cronológica é conhecido como dessazonalização ou ajustamento sazonal e a série cronológica resultante diz-se dessazonalizada ou ajustada sazonalmente*". Como será evidenciado, a especificação de um modelo de regressão linear com regressores dummy constitui um meio de ajustamento sazonal.

Várias hipóteses podem ser assumidas relativamente a um conjunto de dados. A mais simples consiste em admitir que nem a variável dependente nem a variável explicativa possuem comportamento sazonal. Neste caso, um modelo de regressão do tipo (1) representa de forma apropriada a relação entre as duas variáveis. A situação em que o comportamento sazonal de y é totalmente explicado pelo comportamento sazonal de X constitui outro enquadramento em que a equação (1) pode ser aplicada (Stewart,1991).

A criação de um modelo de regressão com variáveis dummy justifica-se, designadamente, quando a variável dependente exibe um comportamento sazonal não totalmente captado pelo comportamento de X, ainda que este também possa ser sazonal. O mesmo é dizer que y possui uma componente sazonal aditiva pelo que a especificação de uma equação de regressão que permita a variabilidade do termo independente consoante a periodicidade das observações (mensais, trimestrais, etc.) apresenta-se como o procedimento adequado para fazer face a uma situação desta natureza. Se se assumir o caso frequente em que os dados de uma amostra se reportam a períodos trimestrais, é possível isolar a componente sazonal de y especificando o seguinte modelo de regressão linear

$$y_i = \alpha + \delta_2 D_{2i} + \delta_3 D_{3i} + \delta_4 D_{4i} + \beta X_i + u_i \; ; i = 1,2,...,n \; ; u_i \sim \mathrm{IIN}(0, \sigma^2 I_n) \qquad (26)$$

a qual, possibilita que o termo independente varie em função do trimestre do ano. Em (26), D_{ti} $(t = 2,3,4$ e $i = 1,2,...,n)$ assume o valor 1 para as observações que pertençam ao t-ésimo trimestre. Está-se portanto a assumir que o primeiro trimestre funciona como categoria base. Estes regressores são conhecidos na literatura econométrica como "dummies sazonais", precisamente por permitirem captar o efeito do factor explicativo "sazonalidade" na variabilidade da variável explicada. Da afectação de valores aos regressores dummy vem:

$$y_i = \alpha + \beta X_i + u_i \qquad \text{(primeiro trimestre)} \qquad (27)$$

$$y_i = (\alpha + \delta_2) + \beta X_i + u_i \qquad \text{(segundo trimestre)} \qquad (28)$$

$$y_i = (\alpha + \delta_3) + \beta X_i + u_i \qquad \text{(terceiro trimestre)} \qquad (29)$$

$$y_i = (\alpha + \delta_4) + \beta X_i + u_i \qquad \text{(quarto trimestre)} \qquad (30)$$

Assim, se admitir, por exemplo, que $y =$ vendas de um produto (em unidades monetárias) e $X =$ gastos em publicidade, os estimadores dos coeficientes do modelo (26) são interpretáveis como se segue:

i) $\hat{\alpha}$ estima as vendas esperadas no primeiro trimestre quando as despesas com publicidade são nulas;

ii) $\hat{\delta}_t$ $(t = 2,3,4)$ estima a diferença esperada entre o valor das vendas do t-ésimo trimestre e as vendas do primeiro trimestre, para um mesmo montante gasto em publicidade (qualquer que ele seja);

iii) $\hat{\beta}$ estima a variação esperada nas vendas que decorre na sequência de uma variação unitária no investimento em publicidade, qualquer que seja o trimestre do ano em causa.

Esta ilustração deixa claro em que medida a definição de "dummies sazonais" permite expurgar a sazonalidade da relação entre y e X. De facto, a estimativa do coeficiente de regressão de X na equação (26) reflecte a genuína relação entre essa variável independente e a variável explicada enquanto que o efeito da sazonalidade nesta última variável é captado pelo conjunto de regressores dummy. É neste sentido que Miller (1981, p.105) sublinha que "*...quando a equação estimada contem dummies sazonais como variáveis explícitas, (...) os coeficientes de regressão são libertados das influências sazonais*". Observe que, para testar a relevância da influência da sazonalidade no modelo (26), basta testar a hipótese de δ_2, δ_3 e δ_4 serem conjuntamente nulos.

Alternativamente, pode introduzir-se o factor sazonalidade no modelo de regressão (1) da seguinte forma

$$\begin{aligned} y_i &= \alpha + \beta X_i + \gamma_2 (D_{2i} X_i) + \gamma_3 (D_{3i} X_i) + \gamma_4 (D_{4i} X_i) + u_i \,; \\ &; i = 1,2,\ldots,n \,; u_i \sim IIN(0, \sigma^2 I_n) \end{aligned} \qquad (31)$$

o que sugere que a forma como y responde a variações de X varia de trimestre para trimestre. Efectivamente, a atribuição de valores aos regressores dummy resulta em:

$$y_i = \alpha + \beta X_i + u_i \qquad \text{(primeiro trimestre)} \qquad (32)$$

$$y_i = \alpha + (\beta + \gamma_2)X_i + u_i \quad \text{(segundo trimestre)} \tag{33}$$

$$y_i = \alpha + (\beta + \gamma_3)X_i + u_i \quad \text{(terceiro trimestre)} \tag{34}$$

$$y_i = \alpha + (\beta + \gamma_4)X_i + u_i \quad \text{(quarto trimestre)} \tag{35}$$

o que significa que a importância do efeito sazonal pode ser avaliada através de um teste à hipótese de que γ_2, γ_3 e γ_4 são conjuntamente nulos.

A possibilidade final consiste em admitir que quer o termo independente quer o coeficiente inclinação variam entre trimestres. O modelo de regressão com variáveis dummy correspondente é:

$$\begin{aligned} y_i &= \alpha + \delta_2 D_{2i} + \delta_3 D_{3i} + \delta_4 D_{4i} + \beta X_i + \gamma_2(D_{2i}X_i) + \gamma_3(D_{3i}X_i) + \gamma_4(D_{4i}X_i) + u_i\,; \\ &; i = 1,2,\ldots,n\,;\, u_i \sim \text{IIN}(0, \sigma^2 I_n) \end{aligned} \tag{36}$$

Neste enquadramento, para cada trimestre de cada ano, tem-se:

$$y_i = \alpha + \beta X_i + u_i \quad \text{(primeiro trimestre)} \tag{37}$$

$$y_i = (\alpha + \delta_2) + (\beta + \gamma_2)X_i + u_i \quad \text{(segundo trimestre)} \tag{38}$$

$$y_i = (\alpha + \delta_3) + (\beta + \gamma_3)X_i + u_i \quad \text{(terceiro trimestre)} \tag{39}$$

$$y_i = (\alpha + \delta_4) + (\beta + \gamma_4)X_i + u_i \quad \text{(quarto trimestre)} \tag{40}$$

Portanto, $\hat{\delta}_t$ $(t = 2,3,4)$ mede a diferença entre o termo independente da recta de regressão relativa ao t-ésimo trimestre e o intercepto do primeiro trimestre e $\hat{\gamma}_t$ $(t = 2,3,4)$ a diferença entre a inclinação no trimestre t e a inclinação no primeiro trimestre. Considerando novamente o exemplo da relação entre vendas, publicidade e trimestre do anos, verifica-se que:

i) $\hat{\alpha}$ estima as vendas esperadas no primeiro trimestre quando as despesas com publicidade são nulas;

ii) $\hat{\delta}_t$ $(t = 2,3,4)$ estima a diferença entre o valor esperado das vendas do t-ésimo trimestre e as vendas do primeiro trimestre, quando o montante gasto em publicidade é nulo ;

iii) $\hat{\beta}$ estima a variação esperada nas vendas do primeiro trimestre que resulta de uma variação unitária no investimento em publicidade;

iv) $\hat{\gamma}_t$ $(t = 2,3,4)$ estima a diferença no efeito da publicidade nas vendas entre o t-ésimo trimestre e o primeiro.

Neste caso, a influência de X em y é variável com o trimestre em causa e de magnitude β, $\beta+\gamma_2$, $\beta+\gamma_3$ e $\beta+\gamma_4$, respectivamente, em cada um dos trimestres. Para testar a hipótese de que a relação entre a variável dependente e a variável explicativa X é a mesma nos quatro trimestres, é necessário realizar um teste de significância conjunta aos coeficientes $\hat{\delta}_t$ e $\hat{\gamma}_t$ $(t = 2,3,4)$.

Para finalizar, refira-se que a definição de regressores dummy no modelação de fenómenos com sazonalidade só faz sentido sob a hipótese de que a influência do factor sazonal na variável explicada é constante ao longo de todo o período de tempo em que incide a análise. Esta hipótese esteve naturalmente presente em todas as formulações analisadas. Por exemplo, na especificação (26) admitiu-se não só a possibilidade do termo independente variar com o trimestre do ano mas também, embora implicitamente, que qualquer que fosse o ano considerado, o efeito de cada trimestre particular na variável dependente era sempre dado pelas magnitudes δ_2, δ_3 e δ_4. Como última nota, Abeysinghe (1992) demonstra que apenas neste caso específico é seguro utilizar variáveis dummy na modelação de fenómenos com sazonalidade. Este autor mostra que o uso deste tipo de regressores na modelação de séries cronológicas caracterizadas por raízes unitárias sazonais produz resultados sem qualquer validade.

4. VARIAÇÕES CONTÍNUAS NOS PARÂMETROS

A utilização de regressores dummy revela-se igualmente muito útil quando se pretende modelar fenómenos económicos em que a variação no declive da recta de regressão, embora abrupta, não é descontínua. Esta categoria de modelos que passamos a analisar, é geralmente conhecida na literatura econométrica como "modelos spline" "broken-line" ou "piecewise".

Um exemplo típico de um modelo que se insere nesta categoria baseia-se na relação existente entre os impostos pagos pelas famílias e o seu rendimento bruto. Como é do conhecimento geral, a taxa marginal de imposto assume um valor constante nas várias bandas de rendimento mas sofre uma variação de magnitude entre quaisquer dois escalões que consideremos. Neste enquadramento, a relação entre impostos pagos e rendimento, pode ter a seguinte configuração:

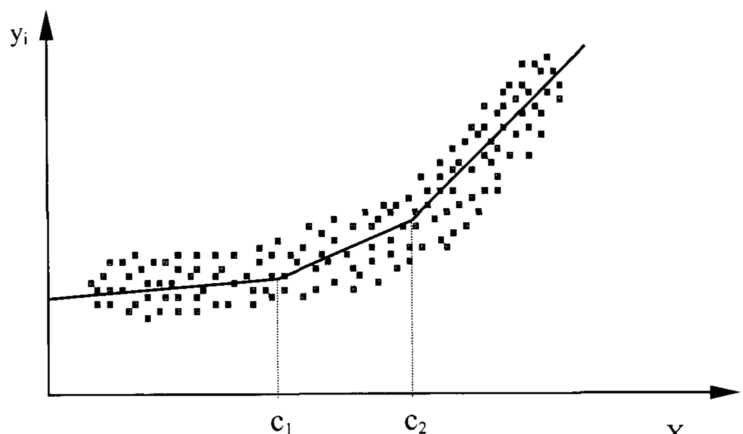

Figura 6. Observações de um hipotético estudo econométrico

A figura 6 ilustra que a nuvem de pontos concentra-se em torno da taxa marginal de imposto que vigora para cada nível de rendimento. Por hipótese, admita-se que os valores de c_1 e c_2 são conhecidos. Face à situação apresentada, ajustar um modelo de regressão linear standard do tipo $I = \alpha + \beta y + u$ parece de todo contraproducente. Efectivamente, tendo em atenção a figura 6, o que se deve estimar é:

$$I_i = \alpha_1 + \beta_1 y_i + u_i \quad \text{para} \quad y_i \leq c_1$$

$$I_i = \alpha_2 + \beta_2 y_i + u_i \quad \text{para} \quad c_1 \leq y_i \leq c_2$$

$$I_i = \alpha_3 + \beta_3 y_i + u_i \quad \text{para} \quad y_i \geq c_2$$

Assim, e numa primeira análise, pode pensar-se que a definição de duas variáveis dummy, por exemplo,

$$D_{2i} = \begin{cases} 1, & \text{se } c_2 \leq y_i \leq c_2 \\ 0, & \text{caso contrário} \end{cases} \qquad \text{e} \qquad D_{3i} = \begin{cases} 1, & \text{se } y_i > c_2 \\ 0, & \text{caso contrário} \end{cases}$$

e a especificação do modelo

$$\begin{aligned} I_i = {} & \alpha_1 + (\alpha_2 - \alpha_1)D_{2i} + (\alpha_3 - \alpha_1)D_{3i} + \\ & + \beta_1 y_i + (\beta_2 - \beta_1)(D_{2i} y_i) + (\beta_3 - \beta_1)(D_{3i} y_i) + u_i \; ; \; i = 1,2,\ldots n \; ; \; u_i \sim IIN(0, \sigma^2) \end{aligned} \qquad (41)$$

permite uma aproximação adequada ao problema em causa. Todavia, da aplicação deste procedimento resultarão não só três equações de regressão totalmente distintas mas também dois valores diferentes quer para $E(I_i \mid y_i = c_1)$ quer para $E(I_i \mid y_i = c_2)$. Quer dizer, a formulação do modelo (41) não garante a continuidade da linha de regressão nos pontos de mudança c_1 e c_2 (também designados por "nós"). Para forçar a linha de regressão a ser contínua, é necessário definir as seguintes restrições lineares

$$\alpha_1 + \beta_1 c_1 = \alpha_2 + \beta_2 c_1$$

$$\alpha_2 + \beta_2 c_2 = \alpha_3 + \beta_3 c_2$$

donde resulta

$$\alpha_2 - \alpha_1 = -c_1(\beta_2 - \beta_1) \qquad (42)$$

e

$$\alpha_3 - \alpha_2 = -c_2(\beta_3 - \beta_2)\,. \qquad (43)$$

Se impuser as restrições (42) e (43) no modelo (41), este transforma-se em:

$$\begin{aligned} I_i &= \alpha_1 + \beta_1 y_i + (\beta_2 - \beta_1) D_{2i}(y_i - c_1) + (\beta_3 - \beta_1) D_{3i}(y_i - c_2) + u_i\,; \\ &;i = 1,2,\ldots n\,;\ u_i \sim \mathrm{IIN}(0,\sigma^2) \end{aligned} \qquad (44)$$

Para obter a linha de regressão que melhor se adapta à nuvem de pontos representada na figura 6, basta regredir I na variável associada ao termo independente α_1, em y e nas variáveis produto $D_2(y - c_1)$ e $D_3(y - c_2)$. Observe-se também que o modelo (44) pode expressar-se como

$$\begin{aligned} I_i &= \alpha_1 + \beta_1 y_i + \gamma_2 D_{2i}(y_i - c_1) + \gamma_3 D_{3i}(y_i - c_2) + u_i\,; \\ &;i = 1,2,\ldots n\,;\ u_i \sim \mathrm{IIN}(0,\sigma^2) \end{aligned} \qquad (45)$$

ou, tendo em atenção as três combinações possíveis de valores das variáveis dummy,

$$I_i = \begin{cases} \alpha_1 + \beta_1 y_i + u_i & \text{se } y_i \le c_1 \\ \alpha_1 - c_1\gamma_2 + (\beta_1 + \gamma_2) y_i + u_i & \text{se } c_1 \le y_i \le c_2 \\ \alpha_1 - c_2\gamma_3 + (\beta_1 + \gamma_3) y_i + u_i & \text{se } y_i \ge c_2 \end{cases}\,. \qquad (46)$$

Portanto, o processo de estimação produzirá um coeficiente inclinação distinto para cada uma das categorias de rendimento consideradas e assegurará, simultaneamente, a continuidade pretendida da linha de regressão. Neste contexto,

$\hat{\beta}_1$ estima o declive da linha de regressão para níveis de rendimento inferiores a c_1, $\hat{\beta}_1 + \hat{\gamma}_2$ o declive para os valores de y que oscilem entre c_1 e c_2 e, finalmente, $\hat{\beta}_1 + \hat{\gamma}_3$ a inclinação da linha de regressão para os valores de y que excedam c_2. Assim, por exemplo, para níveis de rendimento superiores a c_2, estima-se que o montante de imposto pago aumente $\hat{\beta}_1 + \hat{\gamma}_3$ unidades monetárias por cada unidade monetária de aumento do rendimento. Um teste F à hipótese $\gamma_2 = \gamma_3 = 0$ permite testar a significância das mudanças estimadas nos declives.

5. GENERALIZAÇÕES E CONSIDERAÇÕES FINAIS

Neste estudo demonstra-se que dois tipos de modelos com parâmetros variáveis são normalmente ajustáveis mediante a definição de regressores dummy. A situação mais comum e que se analisou em primeiro lugar é aquela em que se assume uma variação descontínua dos parâmetros em causa, seja ela incidente apenas na ordenada na origem da recta de regressão, apenas no seu declive ou em ambos. Dos autores que referem este tipo de modelos, pode referir-se, por exemplo, Maddala (1992), Griffiths (1992), Ramanathan (1989), Greene (1993), Johnston (1984), Kmenta (1986) e Gujarati (1992). A segunda situação, abordada designadamente por Darnell (1994), Ramanathan (1989), Greene (1993), Fomby (1984), Sen e Srivastava (1990) e Gujarati (1992), ocorre quando a variação nos parâmetros se processa de uma forma contínua num ou mais pontos conhecidos à priori. Como se referiu, estes últimos modelos designam-se por "modelos spline", "broken line" ou "piecewise" e foram analisados na parte final deste estudo.

Os modelos especificados para ilustrar este trabalho assentaram na suposição de que apenas uma variável quantitativa era considerada relevante para explicar o comportamento da variável dependente. A visualização gráfica possibilitada por estes modelos justificou a sua selecção. Naturalmente, a utilização de regressores dummy na criação de modelos com parâmetros variáveis encontra também campo de aplicação em modelos mais complexos que incluam um número qualquer de regressores quantitativos. Por exemplo, para incorporar a possibilidade de variabilidade na totalidade dos coeficientes (termo independente e inclinações parciais) do seguinte modelo de regressão com k variáveis explicativas quantitativas (incluindo a variável associada ao termo independente)

$$y_{ij} = \alpha_i + \beta_{1i} X_{1j} + \beta_{2i} X_{2j} + \cdots + \beta_{(k-1)i} X_{(k-1)j} + u_{ij}$$
$$; i = 1,2,\ldots,p \,;\, j = 1,2,\ldots,n_i \,;\, u_i \sim \mathrm{IIN}(0,\sigma^2) \qquad (47)$$

de acordo com p categorias de uma determinada variável nominal, o seguinte modelo com regressores dummy deve ser formulado

$$\begin{aligned} y_i &= \alpha_1 + (\alpha_2 - \alpha_1)D_{2i} + (\alpha_3 - \alpha_1)D_{3i} + \cdots + (\alpha_p - \alpha_1)D_{pi} \\ &+ \beta_{11}X_{1i} + (\beta_{12} - \beta_{11})(D_{2i}X_{1i}) + (\beta_{13} - \beta_{11})(D_{3i}X_{1i}) + \cdots + (\beta_{1p} - \beta_{11})(D_{pi}X_{1i}) + \cdots + \\ &+ \beta_{(k-1)1}X_{(k-1)i} + (\beta_{(k-1)2} - \beta_{(k-1)1})(D_{2i}X_{(k-1)i}) + \cdots + (\beta_{(k-1)p} - \beta_{(k-1)1})(D_{pi}X_{(k-1)i}) + u_i \\ &= \alpha_1 + \sum_{t=2}^{p}(\alpha_t - \alpha_1)D_{ti} + \beta_{11}X_{1i} + \sum_{t=2}^{p}(\beta_{1t} - \beta_{11})(D_{ti}X_{1i}) + \cdots + \beta_{(k-1)1}X_{(k-1)i} + \\ &+ \sum_{t=2}^{p}(\beta_{(k-1)t} - \beta_{(k-1)1})(D_{ti}X_{(k-1)i}) + u_i; \\ &; i = 1,2,\ldots n;\ u_i \sim IIN(0,\sigma^2) \end{aligned} \tag{48}$$

ou, de forma equivalente,

$$\begin{aligned} y_i &= \alpha_1 + \delta_2 D_{2i} + \delta_3 D_{3i} + \cdots + \delta_p D_{pi} \\ &+ \beta_{11}X_{1i} + \gamma_{12}(D_{2i}X_{1i}) + \gamma_{13}(D_{3i}X_{1i}) + \cdots + \gamma_{1p}(D_{pi}X_{1i}) + \cdots + \\ &+ \beta_{(k-1)1}X_{(k-1)i} + \gamma_{(k-1)2}(D_{2i}X_{(k-1)i}) + \cdots + \gamma_{(k-1)p}(D_{pi}X_{(k-1)i}) + u_i \\ &= \alpha_1 + \sum_{t=2}^{p}\delta_t D_{ti} + \beta_{11}X_{1i} + \sum_{t=2}^{p}\gamma_{1t}(D_{ti}X_{1i}) + \cdots + \beta_{(k-1)1}X_{(k-1)i} + \sum_{t=2}^{p}\gamma_{(k-1)t}(D_{ti}X_{(k-1)i}) + u_i; \\ &;\ i = 1,2,\ldots n;\ u_i \sim IIN(0,\sigma^2) \end{aligned} \tag{49}$$

em que:

$$D_{ti} = \begin{cases} 1, & \text{se a i-ésima observação pertence ao t-ésimo grupo } (i = 1,2,\ldots,n;\ t = 2,\ldots,j) \\ 0, & \text{caso contrário} \end{cases}$$

Naturalmente, outros cenários são admissíveis. Nomeadamente, pode considerar-se a situação de apenas o termo independente e parte dos coeficientes inclinação apresentarem diferenças entre os p grupos. Assim, suponha-se que as k variáveis explicativas do modelo (47) são divididas em duas componentes: as q primeiras, cujos coeficientes se acredita serem distintos entre os p grupos e os restantes $r = (k-1) - q$, com parâmetros hipoteticamente iguais para esses três grupos. O modelo com regressores dummy apropriado a este contexto é o seguinte:

$$\begin{aligned} y_i &= \alpha_1 + (\alpha_2 - \alpha_1)D_{2i} + (\alpha_3 - \alpha_1)D_{3i} + \cdots + (\alpha_p - \alpha_1)D_{pi} \\ &+ \beta_{11}X_{1i} + (\beta_{12} - \beta_{11})(D_{2i}X_{1i}) + (\beta_{13} - \beta_{11})(D_{3i}X_{1i}) + \cdots + (\beta_{1p} - \beta_{11})(D_{pi}X_{1i}) + \cdots + \\ &+ \beta_{q1}X_{qi} + (\beta_{q2} - \beta_{q1})(D_{2i}X_{qi}) + \cdots + (\beta_{qp} - \beta_{q1})(D_{pi}X_{qi}) + \\ &+ \beta_{(q+1)}X_{(q+1)i} + \cdots + \beta_{(k-1)}X_{(k-1)i} + u_i \\ &= \alpha_1 + \sum_{t=2}^{p}(\alpha_t - \alpha_1)D_{ti} + \beta_{11}X_{1i} + \sum_{t=2}^{p}(\beta_{1t} - \beta_{11})(D_{ti}X_{1i}) + \cdots + \beta_{q1}X_{qi} + \\ &+ \sum_{t=2}^{p}(\beta_{qt} - \beta_{q1})(D_{ti}X_{qi}) + \beta_{(q+1)}X_{(q+1)i} + \cdots + \beta_{(k-1)}X_{(k-1)i} + u_i; \\ &;\ i = 1,2,\ldots n;\ u_i \sim IIN(0,\sigma^2) \end{aligned} \tag{50}$$

ou, mais simplesmente

$$\begin{aligned} y_i &= \alpha_1 + \delta_2 D_{2i} + \delta_3 D_{3i} + \cdots + \delta_p D_{pi} + \beta_{11} X_{1i} + \gamma_{12}(D_{2i}X_{1i}) + \gamma_{13}(D_{3i}X_{1i}) + \cdots + \gamma_{1p}(D_{pi}X_{1i}) + \cdots + \\ &+ \beta_{q1} X_{qi} + \gamma_{q2}(D_{2i}X_{qi}) + \cdots + \gamma_{qp}(D_{pi}X_{qi}) + \beta_{(q+1)} X_{(q+1)i} + \cdots + \beta_{(k-1)} X_{(k-1)} + u_i \\ &= \alpha_1 + \sum_{t=2}^{p} \delta_t D_{ti} + \beta_{11} X_{1i} + \sum_{t=2}^{p} \gamma_{1t}(D_{ti}X_{1i}) + \cdots + \beta_{q1} X_{qi} + \sum_{t=2}^{p} \gamma_{qt}(D_{ti}X_{qi}) + \\ &+ \beta_{(q+1)} X_{(q+1)i} + \cdots + \beta_{(k-1)} X_{(k-1)} + u_i; \\ &; \ i = 1,2,\ldots n \ ; \ u_i \sim IIN(0,\sigma^2) \end{aligned} \tag{51}$$

Observe-se que, para cada grupo de observações, o modelo (51) é equivalente a:

$$y_i = \begin{cases} \alpha_1 + \beta_{11} X_{1i} + \beta_{21} X_{2i} + \cdots + \beta_{(k-1)1} X_{(k-1)i} + u_i, & (i = 1,2,\ldots,n_1) \\ (\alpha_1 + \delta_2) + (\beta_{11} + \gamma_{12}) X_{1i} + (\beta_{21} + \gamma_{22}) X_{2i} + \cdots + (\beta_{q1} + \gamma_{q2}) X_{qi} + & \\ + \beta_{(q+1)} X_{(q+1)i} + \cdots + \beta_{(k-1)1} X_{(k-1)i} + u_i, & (i = n_1 + 1, n_1 + 2, \ldots, n_1 + n_2) \\ \vdots & \\ (\alpha_1 + \delta_p) + (\beta_{11} + \gamma_{1p}) X_{1i} + (\beta_{21} + \gamma_{2p}) X_{2i} + \cdots + (\beta_{q1} + \gamma_{qp}) X_{qi} + & \\ + \beta_{(q+1)} X_{(q+1)i} + \cdots + \beta_{(k-1)1} X_{(k-1)i} + u_i, & (i = n_1 + n_2 + \cdots + n_{(p-1)} + 1, \ldots, n) \end{cases} \tag{52}$$

O modelo "spline" acima descrito pode igualmente ser generalizado a um qualquer número de nós conhecido. Assim, se admitir, por hipótese, a presença de r nós relevantes, o modelo

$$y_i = \alpha_t + \beta_t X_i + u_i \qquad \text{para} \qquad c_{t-1} \le X_i \le c_t, \ t = 1,2,\ldots,r+1 \tag{53}$$

em que c_0 pode ser igual a $-\infty$ e c_{r+1} pode corresponder a $+\infty$, é ajustável mediante a definição de r regressores dummy

$$D_{ti} = \begin{cases} 1, \ \text{se}\ c_{t-1} \le X_i \le c_t & (i = 1,2,\ldots,n\,;\, t = 1,2,\ldots,r+1) \\ 0, \ \text{caso contrário} \end{cases}$$

Nesta circunstância, o modelo que permitirá obter uma linha de regressão contínua com r + 1 "segmentos" é

$$\begin{aligned} y_i &= \alpha_1 + \sum_{t=2}^{r} (\alpha_t - \alpha_1) D_{ti} + \beta_1 X_i + \sum_{t=2}^{j} (\beta_{ti} - \beta_1)(D_{ti} X_{1i}) + u_i; \\ &; i = 1,2,\ldots n; \ u_i \sim IIN(0,\sigma^2) \end{aligned} \tag{54}$$

após a sua reparametrização com as restrições lineares:

$$\alpha_t + \beta_t c_t = \alpha_{t+1} + \beta_{t+1} c_t, \quad t = 1,2,\ldots,r \tag{55}$$

É importante salientar que os modelos "spline", ajustáveis mediante a criação de regressores dummy, pressupõem que se conheça à priori o valor do nós. Se esta hipótese não se verificar, entra-se no domínio da especificação e estimação de modelos de regressão não lineares cuja análise não se enquadra nos objectivos deste estudo.

REFERÊNCIAS BIBLIOGRÁFICAS

ABEYSINGHE, T. (1992), "Deterministic Seasonal Models and Spurious Regressions", *Journal of Econometrics*, Vol. 61, pp. 259-272.

CHATTERJEE, S. (1991), Regression Analysis by Example, segunda edição. New York (John Wiley & Sons, Inc.).

DARLINGTON, R. B. (1990), Regression and Linear Models. New York (McGraw-Hill).

DARNELL, A. C. (1994), A Dictionary of Econometrics. Aldershot (Edward Elgar).

FOMBY, T., CARTER, R. e JOHNSON, S. (1984), Advanced Econometric Methods. New York (Springer-Verlag).

GREENE, W. H. (1993), Econometric Analysis, segunda edição. New York (Macmillan Publishing Company).

GRIFFITHS, W. E., CARTER H. e JUDGE, G. (1992), Learning and Practicing Econometrics. New York (John Wiley & Sons, Inc.).

GUJARATI, D. (1970), "Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Note", *The American Statistician*, Vol. 24, n. 1, pp 50-52.

NOTAS:

[1] Por grupo entenda-se um conjunto de indivíduos que têm um comportamento uniforme face a uma determinada variável nominal (por exemplo, que possuem o mesmo estado civil ou grau de formação escolar), no caso do estudo em análise ser do tipo seccional, ou um conjunto de observações que digam respeito a um determinado período de tempo, se o estudo em causa for cronológico.

[2] Observe também que omitindo o termo independente do modelo (5) ou da especificação equivalente (6) é possível criar um regressor dummy para cada grupo de observações e obter assim, de forma imediata, o valor de cada um dos termos independentes através dos E.M.Q.O dos parâmetros associados a cada um desses regressores. A formulação que conduz directamente a esse objectivo é

$$y_i = \alpha_1 D_{1i} + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \beta X_i + u_i$$

em que D_{1i} assume o valor 1 para o primeiro grupo de observações.

[3] Uma possível definição de sazonalidade é a proposta por Hylleberg (1992), segundo a qual "*sazonalidade é um movimento intra-anual sistemático, não necessariamente regular, provocado pela variação do clima, do calendário, da periodicidade das decisões, directa ou indirectamente através da produção e consumo de decisões tomadas pelos agentes económicos*".

GUJARATI, D. (1992), Essentials of Econometrics. New York (McGraw- Hill).

HARDY, M. A.(1993), Regression With Dummy Variables (Sage University Paper on Quantitative Applications in the Social Sciences, series nº 07-093). Newbury Park (Sage Publications, Inc.).

HAYS, W. L. (1994), Statistics, quinta edição. Philadelphia (Harcourt Brace College Publishers).

HYLLBERG, S. (1992), Modelling Seasonality. Oxford (Oxford University Press).

JOHNSTON, J. (1984), Econometric Methods, terceira edição Auckland (McGraw-Hill).

KMENTA (1986), Elements of Econometrics, segunda edição. New York (The Macmillan Publishing Company).

LONG, J, S. e MIETHE, T. D. (1988), The Statistic Comparation of Groups. Newbury Park (Sage Publications, Inc.).

MADDALA, G. S. (1992), Introduction to Econometrics, segunda edição. New York (Macmillan Publishing Company).

MILLER (1981), Applied Econometrics. New York (The Macmillan Publishing Company).

SEN, A. e SRIVASTAVA, M. (1990), Regression Analysis: Theory, Methods and Applications. New York (Springer-Verlag),.

STEWART, J. (1991), *Econometrics*. Cambridge (Philip Allan).

Revista de Estatística
41ª Página
3º Quadrimestre de 2002

Testes à Estabilidade dos Parâmetros de um Modelo de Regressão: *Uma Aplicação Especial dos Regressores Dummy*

Autores:
Patrícia Oom do Valle
Efigénio Rebelo

TESTES À ESTABILIDADE DOS PARÂMETROS DE UM MODELO DE REGRESSÃO: UMA APLICAÇÃO ESPECIAL DOS REGRESSORES DUMMY

TESTING STABILITY OF REGRESSION COEFFICIENTS: A SPECIAL APPLICATION OF DUMMY VARIABLES

Autores: Patrícia Oom do Valle
- Assistente da Faculdade de Economia da Universidade do Algarve; Área Científica de Métodos Quantitativos.

Efigénio Rebelo
- Professor Associado da Faculdade de Economia da Universidade do Algarve; Área Científica de Métodos Quantitativos.

***Resumo*:**

- O objectivo deste trabalho é discutir e comparar entre si duas metodologias que permitem testar se os parâmetros de um modelo de regressão são ou não diferentes entre duas ou mais sub-amostras: o teste tradicional de Chow e a técnica das variáveis dummy. Em particular, demonstra-se analiticamente que a técnica das variáveis dummy substitui de forma muito eficaz o teste de Chow, apresentando até algumas vantagens relativamente a este teste.

***Palavras-chave*:**

- *Regressão linear, variáveis dummy, teste de Chow.*

***Abstract*:**

- This study investigates and compares two methods that look to test if parameters of a linear regression model differ between two or more sub-samples: the classic Chow test and the dummy variables technique. More specifically, this paper analytically shows that the latter approach replaces very efficiently and with important advantages the Chow test.

***Key-Words*:**

- *Linear regression, dummy variables, Chow test.*

1. INTRODUÇÃO

A necessidade de testar a hipótese de que parte ou a totalidade dos coeficientes de uma equação de regressão permanece estável para subconjuntos de observações de uma amostra coloca-se, quer o estudo em causa seja do tipo cronológico ou do tipo seccional. Na primeira situação, os subconjuntos de observações dizem respeito a períodos de tempo diferentes e, no segundo caso, referem-se a grupos de indivíduos que diferem em alguma característica qualitativa (classe social, grupo etário, sexo, formação escolar, etc.). Um bom exemplo do que se referiu é a modelação da relação entre poupança e rendimento disponível. De facto, pode ter interesse verificar se a propensão marginal a poupar apresenta alguma mutação de valor quando se transita de um período de recessão económica para um período de prosperidade (estudo cronológico) ou se a referida propensão marginal varia consoante as famílias que constituem uma determinada amostra sejam consideradas de classe alta ou de classe baixa (estudo seccional).

Inicia-se esta análise com uma descrição sumária do teste de Chow (1960) à igualdade dos parâmetros de um modelo de regressão entre duas sub-amostras. As notações que serão utilizadas neste estudo para definir os vários vectores, matrizes, parâmetros, estimadores e estatísticas envolvidos na exposição diferem, em geral, das definidas originalmente pelos autores abordados. Isto aplica-se não só ao teste de Chow (1960) mas também a todos os restantes testes que forem apresentados. O objectivo deste estudo é sistematizar os contributos de autores como Gujarat (1970), Valentine e Erlat (1978; 1985) no uso de regressores dummy na realização de testes à estabilidade dos parâmetros de um modelo de regressão e apresentar as demonstrações matemáticas – em geral muito abreviadas nos papers originais – que sustentam a equivalência entre este método e os clássicos testes de Chow. Desta análise comparativa resultará uma síntese das principais vantagens da abordagem segundo as variáveis dummy em relação à metodologia tradicional de Chow para testar variações nos parâmetros de um modelo.

2. TESTE DE CHOW QUANDO $k < n_i$ $(i = 1,2)$. DUAS SUB-AMOSTRAS

Como enquadramento necessário à exposição que se pretende efectuar, considere-se a existência de dois conjuntos independentes de observações de dimensão n_1 e n_2, respectivamente. Sejam (1) e (2) as equações de regressão correspondentes

$$y_{1i} = \beta_{10} + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \cdots + \beta_{1(k-1)}X_{(k-1)i} + u_{1i} \qquad i = 1,2,\ldots,n_1 \tag{1}$$

$$y_{2i} = \beta_{20} + \beta_{21}X_{1i} + \beta_{22}X_{2i} + \cdots + \beta_{2(k-1)}X_{(k-1)i} + u_{2i} \qquad i = n_1+1, n_1+2,\ldots,n \tag{2}$$

em que o índice 1. identifica as observações e parâmetros relativos ao primeiro grupo e o índice 2. as observações e parâmetros associados ao segundo grupo. Observe que os modelos (1) e (2) podem expressar-se na forma matricial como

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0_1 \\ 0_2 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{3}$$

onde

y_i: vector de n_i observações da variável dependente (i = 1,2)

X_i: matriz de n_i observações dos k regressores (incluindo a variável associada ao termo independente (i = 1,2)

0_i: matriz nula to tipo $(n_i \times k)$ (i =1,2)

β_i: vector de k parâmetros relativos ao i-ésimo grupo de observações (i =1,2)

u_i: vector de n_i variáveis residuais (i =1,2)

A aplicação do teste de Chow requer a verificação simultânea de dois conjuntos de pressupostos:

1) $u_1 \sim N(0, \sigma^2 I_{n_1})$ e $u_2 \sim N(0, \sigma^2 I_{n_2})$;

2) $r(X_i) = k < n_i$ (i =1,2) em que o símbolo r representa a característica da matriz.

A hipótese de igualdade dos parâmetros das duas populações de onde foram extraídas as duas sub-amostras pode ser formalizada como se segue

$$H_0 : \beta_{10} = \beta_{20} \wedge \beta_{11} = \beta_{21} \wedge \ldots \wedge \beta_{1(k-1)} = \beta_{2(k-1)}$$

ou, mais simplesmente,

$$H_0 : \beta_1 = \beta_2 = \beta.$$

Segundo Chow, se esta hipótese não for rejeitada, o procedimento mais adequado será reunir os dois conjuntos de dados e utilizá-los para estimar uma única equação de regressão, a saber:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{4}$$

Caso contrário, as duas equações de regressão separadas (1) e (2) deverão ser ajustadas para cada uma das sub-amostras. Designe-se (3) como regressão não restrita e (4) como regressão restrita.

A realização do teste de Chow pode sintetizar-se nas quatro etapas que seguintes:

(1) Reunir todas as observações ($n_1 + n_2 = n$) e estimar o modelo restrito (4). Seja S_1 o S.Q.R. com $n_1 + n_2 - k$ graus de liberdade resultante desse processo de ajustamento;

(2) Estimar em separado cada um dos modelos de regressão (1) e (2) e identificar os respectivos somatórios dos quadrados dos resíduos (S.Q.R). Seja S_2 o S.Q.R. com $n_1 - k$ graus de liberdade do modelo (1) e S_3 a SQR com $n_2 - k$ graus de liberdade do modelo (2). De seguida, adicionar S_2 e S_3 para obter S_4 com $n_1 + n_2 - 2k$ graus de liberdade;

(3) Determinar $S_5 = S_1 - S_4$ com $(n_1 + n_2 - k) - (n_1 + n_2 - 2k) = k$ graus de liberdade;

(4) Aplicar a estatística

$$F = \frac{S_5 / k}{S_4 /(n_1 + n_2 - 2k)} \tag{5}$$

que sob a hipótese nula segue uma distribuição de F com k e $n_1 + n_2 - 2k$ graus de liberdade. Se F > F crítico, a hipótese nula de igualdade das duas equações de regressão é rejeitada.

Na sua análise, Chow reconhece também a importância de existir um teste estatístico que possibilite a comparação de apenas parte dos coeficientes dos dois modelos de regressão. Assim, suponha-se que possui informação que permite assegurar que algum ou alguns dos parâmetros das duas equações de regressão são efectivamente diferentes, de tal modo que apenas interessa testar se os restantes coeficientes dos dois modelos apresentam diferenças entre si. Os parâmetros cujo valor nos dois modelos se assume à priori ser diferente serão identificados através do índice *a* e os demais coeficientes cuja igualdade se pretende mediante a afectação do índice *b*. Sejam (6) e (7) as respectivas equações de regressão na forma matricial

$$y_1 = X_{1a}\beta_{1a} + X_{1b}\beta_{1b} + u_1 \tag{6}$$

$$y_2 = X_{2a}\beta_{2a} + X_{2b}\beta_{2b} + u_2 \tag{7}$$

ou, de forma equivalente,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_{1a} & 0_1 & X_{1b} & 0_{11} \\ 0_2 & X_{2a} & 0_{22} & X_{2b} \end{bmatrix} \begin{bmatrix} \beta_{1a} \\ \beta_{2a} \\ \beta_{1b} \\ \beta_{2b} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{8}$$

onde

y_i: vector de n_i observações da variável dependente (i = 1,2)

X_{ia}: matriz de n_i observações dos p regressores cujos parâmetros correspondentes são à priori não constantes nos dois modelos (i = 1,2)

X_{ib}: matriz de n_i observações dos q regressores ($q = k - p$) cuja estabilidade dos parâmetros correspondentes pretendemos testar (i = 1,2)

0_i: matriz nula $(n_i \times p)$ (i = 1,2)

0_{ii}: matriz nula $(n_i \times q)$ (i = 1,2)

β_{1a}: vector de p parâmetros relativos ao primeiro grupo de observações cujo conteúdo se sabe ser diferente do de β_{2a} (vector de parâmetros relativos ao segundo grupo)

β_{1b}: vector de q parâmetros relativos ao primeiro grupo de observações cujo conteúdo se pretende verificar se é ou não igual ao do correspondente vector associado ao segundo grupo, β_{2b}

u_i: vector de n_i variáveis residuais (i =1,2)

Os regressores contidos em X_{1a} podem ou não ser diferentes dos que constam em X_{2a}. Para uma maior facilidade na exposição, assuma-se que esses regressores são idênticos nos dois modelos e que o seu número é p. Por hipótese, assuma-se também que se conhece à partida que os coeficientes dessas variáveis diferem nos dois modelos, pelo que apenas importa saber se os parâmetros associados às variáveis contidas em X_{1b} e X_{2b} são ou não iguais. Nesta circunstância, a hipótese nula a ensaiar é, naturalmente:

$$H_0 : \beta_{1b} = \beta_{2b} = \beta_b \,.$$

Se H_0 não for rejeitada, o seguinte modelo com todas as observações deverá ser estimado:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_{1a} & 0_1 & X_{1b} \\ 0_2 & X_{2a} & X_{2b} \end{bmatrix} \begin{bmatrix} \beta_{1a} \\ \beta_{2a} \\ \beta_b \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{9}$$

Como demonstra Chow, a estatística com distribuição F, sob H_0, com base na qual será realizado o teste é:

$$F = \frac{S_5^{'} / q}{S_4^{'} /(n_1 + n_2 - 2k)}. \tag{10}$$

À semelhança do que se disse relativamente à estatística (5), $S_4^{'}$ representa a adição dos S.Q.R. resultantes do ajustamento em separado dos modelos (6) e (7) e $S_5^{'}$ a diferença entre o S.Q.R. da relação (9) e $S_4^{'}$. A estatística (10) tem q e $n_1 + n_2 - 2k$ graus de liberdade.

3. TESTE DE GUJARATI QUANDO $k < n_i$ $(i = 1,2)$. DUAS SUB-AMOSTRAS

Gujarati (1970) revela em que medida a especificação de um modelo de regressão com variáveis dummy constitui uma técnica alternativa ao teste de Chow como forma de verificar se os coeficientes de um modelo de regressão diferem significativamente entre duas sub-amostras. A importância da aplicabilidade dos regressores dummy neste contexto, é sublinhada posteriormente por Maddala (1992), Erlat (1985), Darnell (1994) e Dufour (1981), entre outros. Todavia, nenhum destes autores apresenta as demonstrações matemáticas completas que sustentam a equivalência das duas abordagens.

Para tornar clara a similitude entre as duas técnicas, considere-se novamente os modelos de regressão (1) e (2) cujos parâmetros pretendemos comparar. Os S.Q.R. necessários à construção do teste estatístico à igualdade dos parâmetros das duas equações de regressão podem ser obtidos através da estimação do seguinte modelo com regressores dummy a partir da totalidade das observações ($n_1 + n_2$)

$$
\begin{aligned}
y_i &= \beta_{10} + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \cdots + \beta_{1(k-1)}X_{(k-1)i} + \\
&+ (\beta_{20} - \beta_{10})D_i + (\beta_{21} - \beta_{11})(D_i X_{1i}) + (\beta_{22} - \beta_{12})(D_i X_{2i}) + \ldots + \\
&+ (\beta_{2(k-1)} - \beta_{1(k-1)})(D_i X_{(k-1)i}) + u_i \ ; \ i = 1,2,\ldots,n \ ; \ u_i \sim IIN(0,\sigma^2)
\end{aligned} \tag{11}
$$

em que

$$
D_i = \begin{cases} 0, & i = 1,2,\ldots,n_1 \\ 1, & i = n_1 + 1,\ldots,n \end{cases}
$$

ou, de forma equivalente,

$$
\begin{aligned}
y_i &= \beta_{10} + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \cdots + \beta_{1(k-1)}X_{(k-1)i} + \\
&+ \gamma_0 D_i + \gamma_1(D_i X_{1i}) + \gamma_2(D_i X_{2i}) + \ldots + \gamma_{(k-1)}(D_i X_{(k-1)i}) + u_i ; \\
&; i = 1,2,\ldots,n \ ; \ u_i \sim IIN(0,\sigma^2)
\end{aligned} \tag{12}
$$

onde $\gamma_i = \beta_{2i} - \beta_{1i}$ $(i = 0,1,2,\ldots,k-1)$. Admitindo que as observações relativas ao primeiro grupo são inscritas em primeiro lugar, é fácil perceber que (12) pode expressar-se em notação matricial como se segue

$$
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0_1 \\ X_2 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \gamma \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{13}
$$

em que $\gamma = \beta_2 - \beta_1$. Portanto, a hipótese nula que formulada a propósito do teste de Chow, $H_0 : \beta_2 = \beta_1 = \beta$, pode ser representada em função de γ. Na verdade, se $\beta_2 = \beta_1$ então $\beta_2 - \beta_1 = 0$, o que corresponde a $H_0 : \gamma = 0$, em que 0 representa um vector nulo com k elementos. Quer dizer, o teste de Chow encontra correspondência no modelo de regressão com variáveis dummy (12) num teste de significância conjunta aos coeficientes γ_i $(i = 0,1,2,\ldots,k-1)$. Uma vez que o que se pretende é testar a significância conjunta de k parâmetros, a estatística que permite conduzir este teste é

$$
F = \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right)/g}{\hat{u}'\hat{u}/(n-T)} \sim F_{(g,n-T)} \quad \Bigg| \quad H_0: \gamma = 0 \tag{14}
$$

em que $\hat{u}'\hat{u}$ é o S.Q.R. do modelo (13) e $\hat{u}_r'\hat{u}_r$ o S.Q.R. deste modelo uma vez imposta a restrição $\gamma = 0$. g representa o número de parâmetros cuja significância pretendemos testar, isto é, g = k e T representa o número de regressores do modelo, ou seja, T = 2k. Portanto, a estatística (14) pode expressar-se da seguinte forma

$$F = \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right)/k}{\hat{u}'\hat{u}/(n_1 + n_2 - 2k)} \sim F_{(k, n_1+n_2-2k)} \quad \Bigg| \; H_0: \gamma = 0 \qquad (15)$$

Para demonstrar que a estatística (15) coincide com a apresentada em (5) para realizar o teste de Chow, importa, numa primeira fase, determinar o vector de E.M.Q.O. do modelo (3). Assim, aplicando a fórmula tradicional para estes estimadores , $\hat{\beta} = (X'X)^{-1}X'y$, vem, após a utilização de alguma álgebra matricial:

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1}X_1'y_1 \\ (X_2'X_2)^{-1}X_2'y_2 \end{bmatrix}. \qquad (16)$$

A observação da expressão (16) não deixa dúvidas de que a combinação das duas equações (1) e (2) em (3) não tem consequências ao nível dos parâmetros estimados, isto é, ao fazer correr a regressão (3) chega-se exactamente às mesmas estimativas para os vários coeficientes que se obteriam a partir do ajustamento de cada equação per si. Como se demonstrará, este procedimento de combinar os dois modelos revela-se útil na análise que se segue.

Para determinar a expressão que define o vector de E.M.Q.O. do modelo (13), pode aplicar-se também a fórmula $\hat{\beta} = (X'X)^{-1}X'y$. Neste caso, obter-se-á:

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1}X_1'y_1 \\ (X_2'X_2)^{-1}X_2'y_2 - (X_1'X_1)^{-1}X_1'y_1 \end{bmatrix} \qquad (17)$$

ou, atendendo a (16)

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 - \hat{\beta}_1 \end{bmatrix}. \qquad (18)$$

Consequentemente, o vector dos resíduos do modelo (13) define-se como

$$\begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 & 0_1 \\ X_2 & X_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 - \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} y_1 - X_1 \hat{\beta}_1 \\ y_2 - X_2 \hat{\beta}_2 \end{bmatrix} \tag{19}$$

e $\hat{u}'\hat{u}$ é dado pelo quadrado da norma de $\hat{u}$

$$\left\| \begin{matrix} \hat{u}_1 \\ \hat{u}_2 \end{matrix} \right\|^2 = \left\| \begin{matrix} y_1 - X_1 \hat{\beta}_1 \\ y_2 - X_2 \hat{\beta}_2 \end{matrix} \right\|^2 = \sum_{i=1}^{n_1} \left(y_i - \hat{\beta}_{10} - \hat{\beta}_{11} X_{1i} - \cdots - \hat{\beta}_{1(k-1)} X_{(k-1)i} \right)^2 +$$

$$+ \sum_{i=1}^{n_2} \left(y_i - \hat{\beta}_{20} - \hat{\beta}_{21} X_{1i} - \cdots - \hat{\beta}_{2(k-1)} X_{(k-1)i} \right)^2 =$$

$$= S_2 + S_3 = S_4 \tag{20}$$

tendo em atenção a segunda etapa do teste de Chow acima descrita.

Por seu turno, para se encontrar a fórmula que traduz $\hat{u}_r'\hat{u}_r$, é suficiente impor a restrição $\gamma = 0$ no modelo (13). Este transforma-se então no modelo (4) cujo vector de E.M.Q.O. é dado simplesmente por

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

e, relativamente ao qual, o vector de resíduos se define como

$$\begin{bmatrix} \hat{u}_{1r} \\ \hat{u}_{2r} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \hat{\beta} = \begin{bmatrix} y_1 - X_1 \hat{\beta} \\ y_2 - X_2 \hat{\beta} \end{bmatrix}. \tag{21}$$

Em resultado,

$$\hat{u}_r'\hat{u}_r = \left\| \begin{matrix} \hat{u}_{1r} \\ \hat{u}_{2r} \end{matrix} \right\|^2 = \left\| \begin{matrix} y_1 - X_1 \hat{\beta} \\ y_2 - X_2 \hat{\beta} \end{matrix} \right\|^2 = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \cdots - \hat{\beta}_{(k-1)} X_{(k-1)i} \right)^2 =$$

$$= S_1 \tag{22}$$

considerando agora à primeira etapa do teste de Chow.

As igualdades (20) e (22) possibilitam que se expresse a estatística (15) da seguinte forma

$$F=\frac{(S_1-S_4)/k}{S_4/(n_1+n_2-2k)}\sim F_{(k,n_1+n_2-2k)}\Bigg|_{H_0:\ \gamma=0}$$

ou, atendendo ao terceiro passo do teste de Chow,

$$F=\frac{S_5/k}{S_4/(n_1+n_2-2k)}\sim F_{(k,n_1+n_2-2k)}\Bigg|_{H_0:\ \gamma=0} \tag{23}$$

o que comprova a equivalência entre o teste de Chow e o teste das variáveis dummy à estabilidade dos coeficientes de dois modelos de regressão.

Similarmente, o teste de Chow à igualdade de apenas parte dos coeficientes de regressão de dois modelos apresentado em (10) pode também ser conduzido de uma forma muito simples através da especificação da equação de regressão adequada com variáveis dummy. Neste enquadramento, para ensaiar a hipótese nula $H_0:\beta_{1b}=\beta_{2b}=\beta_b$ na relação (9) através do método das variáveis dummy, a seguinte equação de regressão deverá ser estimada a partir da totalidade das observações

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}=\begin{bmatrix} X_{1a} & 0_1 & X_{1b} & 0_1 \\ 0_2 & X_{2a} & X_{2b} & X_{2b} \end{bmatrix}\begin{bmatrix} \beta_{1a} \\ \beta_{2a} \\ \beta_{1b} \\ \gamma_2 \end{bmatrix}+\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{24}$$

em que $\gamma_2=\beta_{2b}-\beta_{1b}$ é do tipo $(q\times 1)$.

A validade da hipótese nula em causa pode ser testada mediante a realização de um teste de significância aos coeficientes que constituem o vector γ_2. A demonstração de que a estatística necessária à realização deste teste coincide com a apresentada em (10) no contexto do teste de Chow é facilmente deduzível mediante o mesmo raciocínio que utilizado para provar a igualdade dos rácios (5) e (15). Esta demonstração não será, portanto, apresentada.

4. TESTE DE CHOW QUANDO $k \geq n_2$. DUAS SUB-AMOSTRAS

No início da exposição do teste de Chow, referiu-se que um dos pressupostos necessários à sua aplicação era que o número de observações em cada sub-amostra fosse de ser suficiente para permitir a estimação de cada uma das equações de regressão em separado, isto é, $r(X_1) = k < n_1$ e $r(X_2) = k < n_2$. Efectivamente, só sob a verificação desta condição é possível determinar os S.Q.R. de cada um dos modelos, elementos fundamentais para a construção do teste.

Chow (1960) sugere que outro teste seja utilizado quando uma das sub-amostras não possui o número de observações necessário ao ajustamento do modelo de regressão correspondente. Admita-se que a sub-amostra nestas condições é a segunda, isto é, $r(X_2) = n_2 \leq k$ e que se pretende testar a igualdade dos parâmetros dos modelos (1) e (2). Este teste designa-se por teste de previsão (ex-post). Esta designação resulta da forma como é conduzido o teste: estende-se o período de estimação inicial (com base apenas em n_1 observações) ao período de previsão (usando também as restantes n_2 observações) e testa-se se os erros de previsão são em média nulos; Isto é, testa-se a significância dos erros de previsão. A designação 'ex-post' resulta do facto de não se tratar de uma verdadeira previsão, ou previsão ex-ante, uma vez que os valores de y são conhecidos para o período de previsão. A sequência adequada de procedimentos, por forma a realizar o teste, é a seguinte:

(1) Determinar o S.Q.R. resultante da regressão apenas com as n_1 observações. Designe-se por S_1^* este S.Q.R., o qual terá, obviamente, $n_1 - k$ graus de liberdade;

(2) Reunir todas as observações $(n_1 + n_2)$ para voltar a estimar o mesmo modelo de regressão. Seja S^* o S.Q.R. com $n_1 + n_2 - k$ graus de liberdade que resulta desta regressão. Esta é considerada a versão restrita do modelo já que obriga os parâmetros a assumirem o mesmo valor no período correspondente a n_2;

(3) Determinar $S_d^* = S^* - S_1^*$ com $(n_1 + n_2 - k) - (n_1 - k) = n_2$ graus de liberdade;

(4) Aplicar a estatística:

$$F^* = \frac{S_d^*/n_2}{S_1^*/(n_1 - k)} \tag{25}$$

que sob a hipótese nula $H_0 : \beta_1 = \beta_2 = \beta$ segue uma distribuição F com n_2 e $n_1 - k$ graus de liberdade.

Chow demonstra que se a hipótese a testar for $H_0 : \beta_{1b} = \beta_{2b}$, no contexto das equações de regressão (6) e (7) em que parte dos parâmetros é variável à priori, a estatística com base na qual deverá ser conduzido o teste de previsão é

$$F^* = \frac{S_d^{*\prime}/(n_2 - k + q)}{S_1^{*\prime}/(n_1 - k)} \tag{26}$$

com distribuição F com $n_2 - k + q$ e $n_1 - k$ graus de liberdade sob H_0. $S_1^{*\prime}$ representa o S.Q.R. que resulta da estimação do primeiro modelo, (7), e $S_d^{*\prime}$ representa a diferença entre o S.Q.R. da equação de regressão (9) e $S_1^{*\prime}$.

5. TESTES DE VALENTINE E ERLAT (1978) QUANDO $k \geq n_2$. DUAS SUB-AMOSTRAS

Assim como o primeiro teste de Chow descrito pode ser realizado via estimação de um modelo de regressão linear com regressores dummy, o teste de previsão pode igualmente ser conduzido através da definição deste tipo de regressores. A formulação do teste de previsão de Chow através de variáveis dummy é da autoria de Valentine (1971) para a situação em que se pretende testar se todos os parâmetros do modelo relativos às duas sub-amostras apresentam diferenças entre si. Posteriormente, as suas conclusões foram generalizadas por Erlat (1978) ao caso em que se admite à partida que algum ou alguns dos coeficientes são efectivamente diferentes entre as equações de regressão relativas às duas populações. Nesta secção serão apresentadas as demonstrações matemáticas que evidenciam a correspondência entre o segundo teste de Chow e o teste à estabilidade dos parâmetros usando regressores dummy.

Assim, considere-se novamente os modelos (1) e (2), representados na forma matricial como:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0_1 \\ 0_2 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

em que $u_1 \sim N(0,\sigma^2 I_{n_1})$ e $u_2 \sim N(0,\sigma^2 I_{n_2})$. Admita-se que $r(X_1) = k < n_1$ e $r(X_2) = n_2 \leq k$.

Valentine demonstra que, nestas condições, o modelo com variáveis dummy definido em (13) não pode ser estimado. De facto, é possível expressar o modelo (13) na forma equivalente:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_1 + \begin{bmatrix} 0_1 \\ X_2 \end{bmatrix} \gamma + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{27}$$

ou ainda como:

$$y = X\beta_1 + D\gamma + u \ ; u \sim N(0, \sigma^2 I_n) \ . \tag{28}$$

Em termos gerais, os E.M.Q.O. para β_1 e γ de um modelo do tipo (28) são dados pelas expressões

$$\hat{\beta}_1 = (X'M_D X)^{-1} X'M_D y \tag{29}$$

$$\hat{\gamma} = (D'D)^{-1} D' \left(y - X\hat{\beta}_1 \right) \tag{30}$$

onde $M_D = I_n - D\left(D'D\right)^{-1} D'$. Contudo, a matriz X_2 é do tipo $(n_2 \times k)$, de tal forma que $r(X_2) = n_2 \leq k$. Nestas condições, $D'D = X_2' X_2$ não é invertível[1] se $n_2 < k$.

Por forma a solucionar este problema, Valentine sugere que se elimine o "excesso" de variáveis do modelo (27) e que se baseie o teste de permanência de estrutura no modelo resultante. Nesta conformidade, demonstra-se seguidamente que a estatística que permite realizar o teste identifica-se com o rácio F (25) do teste de previsão de Chow. De acordo com Valentine, $k - n_2$ colunas da matriz D devem ser desprezadas no processo de estimação. Uma vez efectuado este corte, a matriz X_2 fica com uma estrutura do tipo $(n_2 \times n_2)$ e, consequentemente, a matriz $D'D$ torna-se regular e $\hat{\beta}_1$ e $\hat{\gamma}$ são determináveis através das fórmulas (29) e (30). Designe-se por X_2^* $(n_2 \times n_2)$ e $D^*(n \times n_2)$, respectivamente, as matrizes X_2 e D após a realização do referido corte. Note que, ao se reduzir o número de variáveis do modelo diminui-se também o número de parâmetros contidos em γ. Este vector, que se passa a designar por γ^*, é agora do tipo $(n_2 \times 1)$. Demonstra-se em apêndice que os vectores $\hat{\beta}_1$ e $\hat{\gamma}^*$ são dados pelas seguintes expressões:

$$\hat{\beta}_1 = (X_1'X_1)^{-1} X_1 y_1 \tag{31}$$

$$\hat{\gamma}^* = \left(X_2^{*-1}\right)\left(y_2 - X_2\hat{\beta}_1\right). \tag{32}$$

Os resíduos desta regressão definem-se portanto como:

$$\begin{aligned}\begin{bmatrix}\hat{u}_1\\ \hat{u}_2\end{bmatrix} &= \begin{bmatrix}y_1\\ y_2\end{bmatrix} - \begin{bmatrix}X_1\\ X_2\end{bmatrix}\hat{\beta}_1 - \begin{bmatrix}0_1^*\\ X_2^*\end{bmatrix}\left\{\left(X_2^{*-1}\right)\left(y_2 - X_2\hat{\beta}_1\right)\right\} = \\ &= \begin{bmatrix}y_1 - X_1\hat{\beta}_1\\ 0_2^{**}\end{bmatrix}\end{aligned} \tag{33}$$

em que 0_1^* é uma matriz nula $(n_1 \times n_2)$ e 0_2^{**} é um vector nulo com n_2 elementos.

A hipótese nula de igualdade dos parâmetros dos dois modelos, formulada como $H_0 : \gamma^* = 0$, pode ser testada através da aplicação da estatística apresentada em (14). Neste caso, $\hat{u}'\hat{u}$ é o S.Q.R. do modelo (27), após o corte de variáveis, e $\hat{u}_r'\hat{u}_r$ o S.Q.R. deste modelo uma vez imposta a restrição $\gamma^* = 0$. Uma vez que γ^* é $(n_2 \times 1)$, o número de parâmetros cuja significância se pretende testar é n_2, isto é, $g = n_2$. Por sua vez, o valor de T no rácio (14) é agora dado por $k + n_2$, atendendo a que $\begin{bmatrix}X_1\\ X_2\end{bmatrix}$ é $(n \times k)$ e D^* é $(n \times n_2)$. Portanto, para o caso concreto em análise, a estatística (14) define-se como:

$$\begin{aligned} F &= \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right)/n_2}{\hat{u}'\hat{u}/(n_1 + n_2 - k - n_2)} \sim F_{(n_2, n_1 - k)} \Bigg|_{H_0:\gamma^* = 0} \\ &= \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right)/n_2}{\hat{u}'\hat{u}/(n_1 - k)} \sim F_{(n_2, n_1 - k)} \Bigg|_{H_0:\gamma^* = 0} \end{aligned} \tag{34}$$

A partir de (33), é fácil concluir que $\hat{u}'\hat{u}$ coincide com o S.Q.R. resultante da aplicação do método dos mínimos quadrados ordinários a (1), isto é, corresponde a S_1^* do teste de previsão de Chow (Vide a segunda etapa do teste de Chow). Por sua vez, sob H_0, o modelo (27), pós o corte de variáveis, reduz-se ao modelo (4), cuja estimação produz um S.Q.R. igual a S^* do teste de previsão de Chow (Vide a segunda etapa deste teste). Assim, a estatística (34) pode escrever-se como:

$$F = \frac{(S^* - S_1^*)/n_2}{S_1^*/(n_1 - k)} \sim F_{(n_2, n_1 - k)} \qquad \Bigg| \; H_0: \gamma^* = 0$$

Fica deste modo demonstrado em que medida a especificação de um modelo de regressão linear com variáveis dummy, e o seu ajustamento segundo a sugestão de Valentine, conduz à mesma estatística do teste de previsão de Chow.

Erlat (1978), parte da análise de Valentine e generaliza-a à situação em que se conhece à partida que uma parcela dos coeficientes dos dois modelos é efectivamente distinta. Nesta caso, as equações de regressão (6) e (7), em que $n_2 < k$, podem ser utilizadas para se expor o contributo de Erlat. Pelo mesmo processo descrito acima, mostra-se que a matriz de variáveis explicativas do modelo (24) não é regular e, por isso, a técnica de Gujarati (1970) não é aplicável. O primeiro procedimento a adoptar para que seja possível utilizar um modelo com variáveis dummy para obter os mesmos resultados que o teste de previsão de Chow e, em particular, a estatística (26), consiste em escrever o modelo (24) na forma equivalente:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_{1b} & X_{1a} \\ X_{2b} & 0_2 \end{bmatrix} \begin{bmatrix} \beta_{1b} \\ \beta_{1a} \end{bmatrix} + \begin{bmatrix} 0_1 & 0_{11} \\ X_{2a} & X_{2b} \end{bmatrix} \begin{bmatrix} \beta_{2a} \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{35}$$

Seguidamente, $p + q - n_2$ colunas de $\begin{bmatrix} 0_{11} \\ X_{2b} \end{bmatrix}$ deverão ser eliminadas de modo a que os estimadores para que os vários parâmetros do modelo resultante sejam determináveis através das fórmulas (29) e (30). Designe-se as matrizes X_{2b}, $\begin{bmatrix} 0_1 \\ X_{2b} \end{bmatrix}$ e $\begin{bmatrix} 0_1 & 0_{11} \\ X_{2a} & X_{2b} \end{bmatrix}$ resultantes deste corte, por X_{2b}^*, $\begin{bmatrix} 0_{11}^* \\ X_{2b}^* \end{bmatrix}$ e $\begin{bmatrix} 0 \\ V \end{bmatrix}$, respectivamente. Seja também γ^* o vector de coeficientes que resulta da supressão de $p + q - n_2$ parâmetros de γ_2. Note-se que γ^* é do tipo $(n_2 - p) \times 1$ e β_{2a} é $(p \times 1)$. Logo, $\begin{bmatrix} \beta_{2a} \\ \gamma^* \end{bmatrix}$ é $(n_2 \times 1)$.

É importante ter presente que X_{2b}^* é $(n_2 \times (n_2 - p))$, V é $(n_2 \times n_2)$ e $\begin{bmatrix} 0 \\ V \end{bmatrix}$ é $(n \times n_2)$ em que 0 é uma matriz nula do tipo $(n_1 \times n_2)$. Como observa Erlat (1978), testar uma mudança de estrutura na situação em análise, conduz exactamente aos mesmos procedimentos que Valentine propõe na sua abordagem. De facto, $p + q - n_2 = k - n_2$, o que significa que não existe qualquer diferença na aplicação do método de Valentine quando se passa de uma situação em que se pretende testar a possibilidade de todos os parâmetros variarem para o enquadramento mais geral, que caracteriza a análise de Erlat. Na presente situação, pode demonstrar-se que a fórmula (30) produz o seguinte resultado (a demonstração encontra-se em apêndice)

$$\begin{bmatrix} \hat{\beta}_{2a}^{*} \\ \hat{\gamma} \end{bmatrix} = V^{-1}\left(y_2 - X_{2b}\hat{\beta}_{1b}\right) \tag{36}$$

pelo que o vector de resíduos é dado por

$$\begin{aligned} \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_{1b} & X_{1a} \\ X_{2b} & 0_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_{1b} \\ \hat{\beta}_{1a} \end{bmatrix} - \begin{bmatrix} 0 \\ V \end{bmatrix} \left\{ V^{-1}\left(y_2 - X_{2b}\hat{\beta}_{1b}\right) \right\} = \\ &= \begin{bmatrix} y_1 - X_{1a}\hat{\beta}_{1a} - X_{1b}\hat{\beta}_{1b} \\ 0_2^{**} \end{bmatrix} \end{aligned} \tag{37}$$

onde, mais uma vez, 0_2^{**} representa um vector nulo com n_2 elementos.

Neste caso, a estatística apropriada para testar H_0 é:

$$F = \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right) / (n_2 - k + q)}{\hat{u}'\hat{u} / (n_1 - k)} \sim F_{(n_2-k+q,\, n_1-k)} \Bigg|_{H_0:\ \gamma^* = 0} \tag{38}$$

De facto, pretende-se testar a significância conjunta de $n_2 - p = n_2 - (k - q) = n_2 - k + q$ coeficientes e o número de graus de liberdade do modelo (35), uma vez efectuado o corte de variáveis, é $(n_1 + n_2) - (q + p + n_2) = n_1 - q - p = n_1 - k$. Em (38), $\hat{u}'\hat{u}$ representa o S.Q.R. da regressão (35), após a redução de variáveis, e $\hat{u}_r'\hat{u}_r$ o S.Q.R. dessa regressão uma vez imposta a restrição $\gamma^* = 0$. Como (37) deixa claro, $\hat{u}'\hat{u} = \hat{u}_1'\hat{u}_1$. Assim, a estatística (38) coincide com a apresentada em (26) no âmbito do teste de previsão de Chow.

6. GENERALIZAÇÃO A TRÊS SUB-AMOSTRAS

Já se demonstrou que a técnica das variáveis dummy pode ser utilizada em alternativa ao teste de Chow como forma de verificar se os parâmetros de uma equação de regressão apresentam ou não diferenças entre duas sub-amostras. Obviamente, qualquer uma destas técnicas pode ser generalizada ao caso em que existem mais do que dois conjuntos de observações cuja igualdade dos parâmetros no modelo de regressão se pretende verificar. Seguidamente, analisa-se de que forma a técnica das variáveis dummy é aplicável numa situação em que se pretende testar a igualdade dos parâmetros de um modelo de regressão entre três sub-amostras.

Assim, considere-se três modelos de regressão, um para cada grupo de observações, na forma matricial

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} X_{1a} & 0_1 & 0_1 & X_{1b} & 0_{11} & 0_{11} \\ 0_2 & X_{2a} & 0_2 & 0_{22} & X_{2b} & 0_{22} \\ 0_3 & 0_3 & X_{3a} & 0_{33} & 0_{33} & X_{3b} \end{bmatrix} \begin{bmatrix} \beta_{1a} \\ \beta_{2a} \\ \beta_{3a} \\ \beta_{1b} \\ \beta_{2b} \\ \beta_{3b} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{39}$$

em que y_i, X_{ia}, X_{ib}, β_{ia}, β_{ib}, 0_i, 0_{ii} e u_i são os vectores/matrizes já definidos considerando agora i=1,2,3. Comece por admitir que $n_i > k,\ \forall i$. Segundo a técnica de Gujarati (1970), para testar se as três equações de regressão são ou não iguais, deve especificar-se o seguinte modelo com regressores dummy e estimá-lo a partir de todas as observações

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} X_{1a} & 0_1 & 0_1 & X_{1b} & 0_1 & 0_1 \\ 0_2 & X_{2a} & 0_2 & X_{2b} & X_{2b} & 0_2 \\ 0_3 & 0_3 & X_{3a} & X_{3b} & 0_3 & X_{3b} \end{bmatrix} \begin{bmatrix} \beta_{1a} \\ \beta_{2a} \\ \beta_{3a} \\ \beta_{1b} \\ \gamma_2 \\ \gamma_3 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{40}$$

em que $\gamma_2 = \beta_{2b} - \beta_{1b}$ e $\gamma_3 = \beta_{3b} - \beta_{1b}$. Observe que as variáveis dummy

$$D_{2i} = \begin{cases} 1, & i = n_1 + 1, n_1 + 2, ..., n_1 + n_2 \\ 0, & \text{caso contrário} \end{cases}$$

$$D_{3i} = \begin{cases} 1, & i = n_1 + n_2 + 1, n_1 + n_2 + 2, ..., n \\ 0, & \text{caso contrário} \end{cases}$$

estão na base da formulação (40). De facto, os dois últimos conjuntos de colunas da matriz de regressores do modelo (40) resultam do produto das variáveis dummy D_2 e D_3 pelas variáveis quantitativas contidas em X_b. A hipótese nula de permanência de estrutura, pode ser formalizada como $H_0 : \beta_{1b} = \beta_{2b} = \beta_{3b} = \beta_b$ e a estatística que permite testar a sua validade é:

$$F = \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right) / 2q}{\hat{u}'\hat{u} \Big/ \sum_{i=1}^{3}(n_i - p - q)} \sim F_{\left(2q, \sum_{i=1}^{3}(n_i - q - p)\right)} \Bigg|_{H_0} . \quad (41)$$

No presente caso, o número de parâmetros cuja significância conjunta se pretende testar é 2q (uma vez que quer γ_2 quer γ_3 são do tipo $(q \times 1)$) e $n - T = n_1 + n_2 + n_3 - 3p - 3q = \sum_{i=1}^{3}(n_i - p - q)$. $\hat{u}'\hat{u}$ é o S.Q.R. da regressão (21) e $\hat{u}_r'\hat{u}_r$ o S.Q.R. da regressão (40) sob $H_0 : \gamma_2 = \gamma_3 = 0$.

Erlat (1985), generaliza o contributo de Valentine e a sua própria análise de 1978, ao caso em que uma das três sub-amostras não possui observações em número suficiente. Admitindo que a amostra nestas condições é a terceira, isto é, $r(X_3) = n_3 \le k$, Erlat sugere que se escreva a equação de regressão (40) na forma equivalente

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} X_{1b} & X_{1a} & 0_1 \\ X_{2b} & 0_2 & X_{2a} \\ X_{3b} & 0_3 & 0_3 \end{bmatrix} \begin{bmatrix} \beta_{1b} \\ \beta_{1a} \\ \beta_{2a} \end{bmatrix} + \begin{bmatrix} 0_{11} & 0_1 & 0_{11} \\ X_{2b} & 0_2 & 0_{21} \\ 0_{31} & X_{3a} & X_{3b} \end{bmatrix} \begin{bmatrix} \gamma_2 \\ \beta_{3a} \\ \gamma_3 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad (42)$$

em que γ_2 e γ_3 são vectores de parâmetros do tipo $(q \times 1)$.

Erlat propõe que se retirem $q+p-n_3$ colunas de $\begin{bmatrix} 0_{11} \\ 0_{22} \\ X_{3b} \end{bmatrix}$ por forma a que X_{3b} fique do tipo $(n_3 \times (n_3 - p))$. Designe-se a matriz resultante desta supressão de colunas como $\begin{bmatrix} 0^*_{11} \\ 0^*_{22} \\ X^*_{3b} \end{bmatrix}$ e também $V = [X_{3a} \quad X^*_{3b}]$ em que V é $(n_3 \times n_3)$. Seja γ^*_3 o vector $((n_3 - p)\times 1)$ que decorre do corte de $q+p-n_3$ parâmetros de γ_3. Após algumas simplificações matemáticas, demonstra-se que[2]

$$\hat{\gamma}_2 = \hat{\beta}_{2b} - \hat{\beta}_{1b} \tag{43}$$

$$\begin{bmatrix} \hat{\beta}_{3a} \\ \hat{\gamma}^*_3 \end{bmatrix} = V^{-1}\left(y_3 - X_{3b}\hat{\beta}_{1b} \right) \tag{44}$$

o que significa que os resíduos da regressão (42), após o corte de variáveis, são dados por

$$\begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} X_{1b} & X_{1a} & 0_1 \\ X_{2b} & 0_2 & X_{2a} \\ X_{3b} & 0_3 & 0_3 \end{bmatrix} \begin{bmatrix} \hat{\beta}_{1b} \\ \hat{\beta}_{1a} \\ \hat{\beta}_{2a} \end{bmatrix} + \begin{bmatrix} 0_{11} & 0_{111} \\ X_{2b} & 0_{222} \\ 0_{31} & V \end{bmatrix} \begin{bmatrix} \hat{\beta}_{2b} - \hat{\beta}_{1b} \\ V^{-1}(y_3 - X_{3b})\hat{\beta}_{1b} \end{bmatrix} =$$

$$= \begin{bmatrix} y_1 - X_{1a}\hat{\beta}_{1a} - X_{1b}\hat{\beta}_{1b} \\ y_1 - X_{2a}\hat{\beta}_{2a} - X_{2b}\hat{\beta}_{2b} \\ 0^{**}_3 \end{bmatrix} \tag{45}$$

em que 0_{111}, 0_{222} são matrizes nulas do tipo $((n_i \times n_3)$ com $i = 1,2$ e 0^{**}_3 é um vector nulo com n_3 elementos.

A estatística adequada para testar H_0, define-se agora como:

$$F = \left. \frac{\left(\hat{u}_r'\hat{u}_r - \hat{u}'\hat{u}\right) \Big/ (q+n_3-p)}{\hat{u}'\hat{u} \Big/ \sum_{i=1}^{2}(n_i - q - p)} \sim F_{(q+n_3-p,\, \sum_{i=1}^{2}(n_i-q-p))} \right|_{H_0} . \tag{46}$$

Nesta situação, $g = q + n_3 - p$ dado que γ_2 é do tipo $(q \times 1)$ e γ_3^* do tipo $((n_3 - p) \times 1)$. Por sua vez,

$$n - T = n_1 + n_2 + n_3 - (q + 2p + q + p + n_3 - p) = n_1 + n_2 - 2q - 2p = \sum_{i=1}^{2}(n_i - q - p).$$

$\hat{u}'\hat{u}$ é o S.Q.R. da regressão (42), após o corte de variáveis. Como (45) permite constatar, $\hat{u}'\hat{u} = \sum_{i=1}^{2} \hat{u}_i'\hat{u}_i$ e, por seu turno, $\hat{u}_r'\hat{u}_r$ é o S.Q.R. dessa regressão sob a restrição $H_0 : \gamma_2 = \gamma_3^* = 0$.

Qualquer uma das situações envolvendo três sub-amostras pode facilmente ser adaptada para ilustrar os procedimentos que devem ser adoptados quando se pretende testar a hipótese de que todos os parâmetros do modelo variam entre os três conjuntos de observações. Neste caso, deve considerar-se a relação de regressão evidenciada em (39) mas assumindo agora que os vectores de observações relativos aos regressores contidos em X_{1a}, X_{2a} e X_{3a} não existem. Tudo o que se disse relativamente à forma de testar a estabilidade dos parâmetros dos modelos em (39) seria agora válido considerando, obviamente p = 0, razão porque não se apresenta o desenvolvimento dos referidos procedimentos.

7. CONSIDERAÇÕES FINAIS

A análise precedente permite identificar algumas das vantagens que a técnica das variáveis dummy apresenta relativamente aos testes de Chow. Em primeiro lugar, a aplicação do primeiro teste de Chow apenas permite testar se parte ou a totalidade dos parâmetros do modelo de regressão diferem ou não entre dois conjuntos de observações, mediante a não rejeição ou a rejeição da hipótese nula. Quer dizer, a utilização deste teste não especifica em que parâmetros se consubstancia a variabilidade, no caso da hipótese nula ser rejeitada. Pelo contrário, mediante a especificação de uma equação de regressão com variáveis dummy é sempre possível saber, através da realização de testes t ou F, qual é a origem da variabilidade, isto é, em que parâmetros ela se faz sentir em concreto. Num estudo econométrico, a identificação dos parâmetros que sofrem uma diferença de valor entre duas equações de regressão revela-se, em geral, tão importante como o simples conhecimento de que de facto existe uma diferença entre a estrutura estimada de duas populações. Por exemplo, se na equação de regressão (12), as estatísticas t associadas aos E.M.Q.O. $\hat{\gamma}_0$ e $\hat{\gamma}_2$ forem significativa do ponto de vista estatístico, conclui-se, com um nível de significância de $100 \times \alpha\%$, que existe uma diferença entre os termos independentes dos dois planos de regressão e que, simultaneamente, o coeficiente inclinação parcial associado à variável X_2 também muda entre os dois grupos. Atendendo a que normalmente cada um destes parâmetros tem um significado económico, uma informação desta natureza pode ser extremamente útil.

Em segundo lugar, estimar uma única equação de regressão com todas as observações constitui uma forma mais eficiente de testar a estabilidade dos parâmetros de uma equação de regressão entre várias sub-amostra do que recorrer ao teste de Chow. Como se deixou claro nas quatro etapas que caracterizam a realização do primeiro teste de Chow, a sua aplicação no contexto de duas equações de regressão obriga ao ajustamento de três equações de regressão distintas: uma para cada grupo de observações e uma terceira reunindo todos os dados. Do mesmo modo, o teste de previsão requer, necessariamente, a estimação de pelo menos dois modelos de regressão (um para cada período de estimação). Em qualquer dos contextos e independentemente do número de modelos de regressão cuja igualdade dos parâmetros se pretenda testar, através da técnica das variáveis dummy apenas é necessário estimar um único modelo de regressão linear para obter toda a informação relevante:

i) Se existe ou não igualdade entre parâmetros dos modelos associados aos vários grupos;

ii) No caso de existir variabilidade entre os parâmetros dos vários modelos, qual é a fonte ou quais são as fontes dessa variabilidade;

iii) Qual é a estrutura estimada para cada um dos conjuntos de observações (determinável através da atribuição de valores aos regressores dummy).

Por último, reunir todas as observações para estimar um único modelo de regressão, aumenta o número de graus de liberdade e, consequentemente, melhora a precisão relativa da estimação dos vários parâmetros. Por exemplo, se o primeiro teste de Chow rejeitar a hipótese nula, então uma equação de regressão para cada grupo de observações deverá ser ajustada. Se se pretender prever o comportamento da variável dependente para uma observação que se insira num determinado grupo particular, deve utilizar-se a estrutura estimada para esse grupo de modo a encontrar a previsão pretendida, perdendo assim um número de graus de liberdade igual ao número de observações que existem para o outro grupo.

REFERÊNCIAS BIBLIOGRÁFICAS

CHOW, G. C. (1960), "Tests of Equality Between Sets of Coefficients in two Linear Regressions", *Econometrica*, Vol. 28, n. 3, pp. 591-605.

CHATTERJEE, S. (1991), Regression Analysis by Example, segunda edição. New York (John Wiley & Sons, Inc.).

DARNELL, A. C. (1994), A Dictionary of Econometrics. Aldershot (Edward Elgar).

DUFOUR, J. M. (1980), "Dummy Variables and Predictive Tests for Structural Changes: A coordinate Free Approach", *International Economic Review*, Vol. 23, pp. 565-575.

DUFOUR, J. M. (1981), "Dummy Variables and Predictive Tests for Structural Change", *Economic Letters*, Vol. 6, pp. 241-247.

DUFOUR, J. M. (1982), "Generalised Chow Tests for Structural Change: A Coordinate- Free Approach", *International Economic Review*, Vol. 23, n. 3, pp.565-575.

ERLAT, H. (1978), "On the Chow Test when the Degrees of Freedom are Inadequate", *METU Studies in Development*, n. 21, pp. 17-48.

NOTAS:

[1] Um resultado importante da álgebra matricial relativamente à característica de uma matriz, diz que se o produto de duas matrizes A e B se encontra definido, então r(AB) não é maior do que r(A) ou r(B) (ver Maddala (1992) pp. 49). Este resultado pode aplicar-se às matrizes X_2' e X_2 para provar que a matriz $X_2'\ X_2$ não é invertível. De facto, o produto $X_2'\ X_2$ existe porque X_2' é $(k \times n_2)$ e X_2 é $(n_2 \times k)$. Uma vez que $r(X_2) = n_2 \leq k$, $r(X_2'X_2) \leq k$ o que significa que $X_2'\ X_2$ não é regular. Neste caso, o determinante de $X_2'\ X_2$ é nulo e, consequentemente, a matriz não é invertível.

[2] A demonstração das igualdades (43) e (44) é semelhante à que apresentamos em apêndice para provar as igualdades (31), (32) e (36), pelo que não será apresentada.

ERLAT, H. (1985), "Testing for Structural Change at More than One Switch Point: Inadequate Degrees of Freedom and Dummy Variables", *Oxford Bulletin of Economics and Statistics*, Vol. 47, n. 3, pp 293-302.

FOMBY, T., CARTER, R. e JOHNSON, S. (1984), Advanced Econometric Methods. New York (Springer-Verlag).

GREENE, W. H. (1993), Econometric Analysis, segunda edição. New York (Macmillan Publishing Company).

GRIFFITHS, W. E., CARTER H. e JUDGE, G. (1992), Learning and Practicing Econometrics. New York (John Wiley & Sons, Inc.).

GUJARATI, D. (1970), "Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Note", *The American Statistician*, Vol. 24, n. 1, pp 50-52.

GUJARATI, D. (1970), "Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Note", *The American Statistician*, Vol. 24, n. 1 , pp 50-52.

GUJARATI, D. (1970), "Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Generalisation", *The American Statistician*, Vol. 24, n. 5, pp 18-21.

HARDY, M. A. (1993), Regression With Dummy Variables (Sage University Paper on Quantitative Applications in the Social Sciences, series nº 07-093). Newbury Park (Sage Publications, Inc.).

JOHNSTON, J. (1984), Econometric Methods, terceira edição. Auckland (McGraw-Hill).

KMENTA (1986), Elements of Econometrics, segunda edição. New York (The Macmillan Publishing Company).

MADDALA, G. S. (1992), Introduction to Econometrics, segunda edição. New York (The Macmillan Publishing Company).

STEWART, J. (1991), *Econometrics*. Cambridge (Philip Allan).

APÊNDICE

Demonstração das Igualdades (31) e (32)

Partindo do modelo (28) na forma equivalente:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_1 + \begin{bmatrix} 0_1^* \\ X_2^* \end{bmatrix} \gamma^* + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad (47)$$

vem

$$\hat{\beta}_1 = (X'M_{D^*}X_1)^{-1}X'M_{D^*}y \qquad (48)$$

$$\hat{\gamma}^* = (D^{*\prime}D^*)^{-1}D^{*\prime}(y - X\hat{\beta}_1) \qquad (49)$$

Assim, tem-se que:

$$D^{*\prime}D^* = \begin{bmatrix} 0_1^{*\prime} & X_2^{*\prime} \end{bmatrix} \begin{bmatrix} 0_1^* \\ X_2^* \end{bmatrix} = X_2^{*\prime}X_2^* \quad {}_{(n_2 \times n_2)} \qquad (50)$$

e, consequentemente,

$$(D^{*\prime}D^*)^{-1}D^{*\prime} = (X_2^{*\prime}X_2^*)^{-1}\begin{bmatrix} 0_1^{*\prime} & X_2^{*\prime} \end{bmatrix} = \begin{bmatrix} 0_1^{*\prime} & (X_2^{*\prime}X_2^*)^{-1}X_2^{*\prime} \end{bmatrix} {}_{(n_2 \times n)} \qquad (51)$$

Observe-se contudo que X_2^* e $X_2^{*\prime}$ são matrizes regulares pelo que

$$(X_2^{*\prime}X_2^*)^{-1}X_2^{*\prime} = (X_2^*)^{-1}(X_2^{*\prime})^{-1}X_2^{*\prime} = (X_2^*)^{-1} \qquad (52)$$

dado que $(X_2^{*\prime})^{-1}X_2^{*\prime} = I_{n_2}$. Pode então escrever-se (51), alternativamente, como:

$$(D^{*\prime}D^*)^{-1}D^{*\prime} = \begin{bmatrix} 0_1^{*\prime} & (X_2^*)^{-1} \end{bmatrix} \qquad (53)$$

Por sua vez,

$$y - X\hat{\beta}_1 = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\hat{\beta}_1 = \begin{bmatrix} y_1 - X_1\hat{\beta}_1 \\ y_2 - X_2\hat{\beta}_1 \end{bmatrix}_{(n \times 1)} \tag{54}$$

o que significa que

$$\begin{aligned} \hat{\gamma}^* &= (D^{*'}D^*)^{-1}D^{*'}(y - X\hat{\beta}_1) = \begin{bmatrix} 0_1^{*'} & (X_2^*)^{-1} \end{bmatrix} \begin{bmatrix} y_1 - X_1\hat{\beta}_1 \\ y_2 - X_2\hat{\beta}_1 \end{bmatrix} = \\ &= (X_2^*)^{-1}\begin{bmatrix} y_2 - X_2\hat{\beta}_1 \end{bmatrix} \end{aligned} \tag{55}$$

como se pretendia demonstrar.

Para encontrar a expressão que define $\hat{\beta}_1$ determine-se:

$$\begin{bmatrix} 0_1^* \\ X_2^* \end{bmatrix} \begin{bmatrix} 0_1^{*'} & (X_2^*)^{-1} \end{bmatrix} = \begin{bmatrix} 0_{11} & 0_1^* \\ 0_2^* & I_{n_2} \end{bmatrix}_{(n \times n)} \tag{56}$$

$$M_{D^*} = I_{n_2} - D^*(D^{*'}D^*)^{-1}D^{*'} = \begin{bmatrix} I_{n_1} & 0_1^* \\ 0_2^* & I_{n_2} \end{bmatrix} - \begin{bmatrix} 0_{11} & 0_1^* \\ 0_2^* & I_{n_2} \end{bmatrix} = \begin{bmatrix} I_{n_1} & 0_1^* \\ 0_2^* & 0_{22} \end{bmatrix}_{(n \times n)} \tag{57}$$

em que $0_1^*, 0_2^*, 0_{11}$ e 0_{22} são matrizes nulas do tipo $(n_1 \times n_2)$, $(n_2 \times n_2)$ e $(n_2 \times n_2)$, respectivamente. Por sua vez, e tendo em atenção (57),

$$X'M_{D^*}X = \begin{bmatrix} X_1' & X_2' \end{bmatrix} \begin{bmatrix} I_{n_1} & 0_1^* \\ 0_2^* & 0_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1' & 0_2' \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = X_1'X_1 \quad_{(k \times k)} \tag{58}$$

e, consequentemente,

$$(X'M_{D^*}X)^{-1} = (X_1'X_1)^{-1} \tag{59}$$

Do mesmo modo,

$$X'M_{D^*}y = \begin{bmatrix} X_1' & 0_2' \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \underset{(k \times 1)}{X_1' y_1} \tag{60}$$

Finalmente,

$$\begin{aligned} \hat{\beta}_1 &= (X'M_{D^*}X_1)^{-1} X'M_{D^*}y = (X_1'X_1)^{-1}\, X_1'y_1 \\ &= (X_1'X_1)^{-1}\, X_1'y_1 \end{aligned} \tag{61}$$

como se pretendia demonstrar.

Demonstração da Igualdade (36)

Adaptando a igualdade (49) às notações das matrizes/vectores do modelo (16) vem

$$\hat{\gamma}_2' = (D^{*'}D^{*})^{-1}D^{*'}(y - X\hat{\beta}_1) \tag{62}$$

onde, neste caso específico,

$$X = \begin{bmatrix} X_{1b} & X_{1a} \\ X_{2b} & 0_2 \end{bmatrix} \qquad \text{e} \qquad D^{*} = \begin{bmatrix} 0 \\ V \end{bmatrix}$$

Portanto,

$$D^{*'}D^{*} = \begin{bmatrix} 0' & V' \end{bmatrix}\begin{bmatrix} 0 \\ V \end{bmatrix} = \underset{(n_2 \times n_2)}{V'V} \tag{63}$$

e

$$(D^{*'}D^{*})^{-1}D' = (V'V)^{-1}\begin{bmatrix} 0' & V' \end{bmatrix} = \underset{(n_2 \times n)}{\begin{bmatrix} 0' & (V'V)^{-1}V' \end{bmatrix}} \tag{64}$$

Uma vez que V é regular, $(V'V)^{-1}V' = V^{-1}(V')^{-1}V' = V^{-1}$, o que significa que

$$(D^{*'}D^{*})^{-1}D' = \begin{bmatrix} 0' & V' \end{bmatrix} = \begin{bmatrix} 0' & V^{-1} \end{bmatrix}. \tag{65}$$

Logo

$$\begin{bmatrix} \hat{\beta}_{2a} \\ \hat{\gamma}^{*} \end{bmatrix} = \begin{bmatrix} 0' & V^{-1} \end{bmatrix}\left\{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_{1b} & X_{1a} \\ X_{2b} & 0_2 \end{bmatrix}\begin{bmatrix} \hat{\beta}_{1b} \\ \hat{\beta}_{1a} \end{bmatrix}\right\} = V^{-1}\left(y_2 - X_{2b}\hat{\beta}_{1b}\right) \tag{66}$$

como se pretendia demonstrar.

REVISTA DE ESTATÍSTICA
71ª PÁGINA
3º QUADRIMESTRE DE 2002

On Portuguese Investment Fund Performance: I. An Application of a *Generalized Varying Parameter model*

Autor:
Peter de Sousa

ON PORTUGUESE INVESTMENT FUND PERFORMANCE: I. AN APPLICATION OF A *GENERALIZED VARYING PARAMETER MODEL*

AVALIAÇÃO DA PERFORMANCE DE FUNDOS DE INVESTIMENTO MOBILIÁRIO EM PORTUGAL: I. A APLICAÇÃO DE UM MODELO GENERALIZADO DE PARÂMETROS VARIÁVEIS

Autor: Peter de Sousa
- Técnico Superior de Estatística da Direcção Regional do Algarve do Instituto Nacional de Estatística.

ABSTRACT:

- Previous research on portfolio behavior fall mainly into two categories: fixed systematic risk level and varying systematic risk level studies. In recent develpments, scholars have favored the second approach since systematic risk level of a portfolio could be considered as a decision variable and/or a random coefficient. This study, following the second approach, adopts a generalized varying parameter model in investigating portuguese security investment fund performance by taking into account both timing decisions of funds and random behavior of fund's systematic risk level.

  The empirical results suggest that portuguese security investment fund managers in general terms do not show ability in selectivity, although some tend to oppose this trend in one moment or another, and do not show ability in adjusting their fund's systematic risk level to market movements.

KEY-WORDS:

- *Portfolios, Security Investment Funds, Systematic Risk, Selectivity, Market Timing.*

RESUMO:

- Os estudos sobre o comportamento dos portafólios podem ser categorizadas basicamente em dois tipos de abordagem: uns assumindo níveis de risco sistemático fixos e outros assumindo níveis de risco sistemático variáveis. Nos desenvolvimentos recentes tem sido favorecido a segunda abordagem, uma vez que o nível de risco sistemático dos portafólios pode ser visto como uma variável de decisão e/ou como um coeficiente aleatório. No presente estudo, seguindo a segunda linha de orientação, foi adoptado um modelo generalizado de parâmetros variáveis na investigação empírica da performance dos fundos de investimento mobiliários em Portugal, com o intuito de captar não só as decisões de *market timing* dos gestores, mas também o comportamento aleatório dos níveis de risco sistemático.

Os resultados empíricos sugerem que os gestores dos fundos de investimento em termos globais não evidenciam capacidades de selectividade, embora alguns num momento ou outro contrariem esta tendência, assim como não evidenciam capacidades no ajustamento do nível do risco sistemático aos movimentos do mercado.

***PALAVRAS-CHAVE*:**

- *Portafólios, Fundos de Investimento Mobiliário, Risco Sistemático, Selectividade, Market Timing..*

1. INTRODUCTION

Stability of the systematic risk level, or beta coefficient,[2] for portfolios over market conditions has been considerably debated in the literature.[3] The issue is particularly relevant when evaluating the selectivity and the market timing abilities of a portfolio manager. While selectivity (or microforecasting) involves the ability to identify and select individual securities that are over or undervalued, market timing (or macroforecasting) involves the ability to predict and anticipate price movements for the market as a whole. If, for example, a manager can make better than average forecasts of future realizations on market factors, he is expected to adjust portfolio risk in anticipation of market changes. The level of systematic risk can be changed substantially in either direction by altering the proportions of high and low risk securities in the portfolio. Thus, the nonstationarity of systematic risk from portfolio manipulation is a violation of Ordinary Least Squares (OLS) model specification. Consequently, the estimated parameters and statistical inference obtained do not have the appropriate sampling properties that are generally expected in a traditional fixed coefficient model.

We can ascertain that previous research on portfolio behavior fall mainly into two categories: analysis with fixed systematic risk level and analysis with varying systematic risk level. This study, following the second approach, investigates portuguese security investment funds'[4] selectivity, beta stationarity and timing decision simultaneously by using a generalized varying parameter procedure.

The model specification is explained in section two of this paper. In section three, the estimation procedure is presented. In section four the data base is described. The empirical results are presented in section five. Finally, section six contains concluding remarks.

---

[2] Risk associated with investment in marketable securities consists of two types, specific risk and systematic risk. Most of the specific risk can be diversified away by holding diversified portfolios. Relatively little random diversification will eliminate nearly all specific risk, leaving a rump of systematic risk which is impervious to diversification. Systematic risk affects securities differentially and the extent of particular securities' risk depends on the level of covariance with risky investments generally.

[3] Several studies have examined the intertemporal behaviour or stability of a portfolio's systematic risk level. Among others, see Jensen (1969), Levy (1971), Blume (1971, 1975), Klemkosky and Martin (1975), Klemkosky and Maness (1978), Fabozzi and Francis (1978, 1979), Kon and Jen (1978, 1979), Alexander and Stover (1980), Miller and Gressis (1980), Lee and Chen (1982), Kon (1983), Chen and Stockum (1986), Ferson and Schadt (1996) and Kryzanowski, Lalancette and To (1997).

[4] Portuguese Security Investment Funds, as other types of Collective Investment Schemes (e.g., US Mutual funds, UK Unit trusts, french Sicav's, or spanish Fondos de Inversíon), are to be considered as portfolios. See Sousa (1999) for a detailed examination on Portuguese Collective Investment Schemes.

2. THE MODEL SPECIFICATION

Many studies on performance evaluation of securities and/or portfolios consider the systematic risk level (beta coefficient) to be stationary, thus the estimation for the *ex post* Capital Asset Pricing Model (CAPM), developed by Sharpe (1964), Lintner (1965) and Mossin (1966), can be expressed as

$$\left(R_{p,t} - R_{f,t}\right) = \alpha_p + \beta_{p,t}\left(R_{m,t} - R_{f,t}\right) + \mu_{p,t} \qquad [1]$$

where

$R_{p,t}$ = rate of return for portfolio **p** at time t,

$R_{m,t}$ = rate of return for the market portfolio **m** at time t,

$R_{f,t}$ = rate of return for the riskless asset at time t,

α_p = measure of selectivity performance for portfolio **p**,

$\beta_{p,t}$ = measure of systematic risk level for portfolio **p** at time t,

$\mu_{p,t}$ = residual error.

However, abandoning the research orientation introduced by Treynor (1965) and Jensen (1968, 1969), some studies have appointed to the possibility of systematic risk level (beta coefficient) being nonstationary mainly when attempting to capture market timing efforts of portfolios. Therefore, beta coefficient should be regarded as a decision variable instead of a fixed coefficient. In this perspective, estimates of the parameters generated from the CAPM and, consequently, the traditional performance evaluation measures to evaluate mutual funds are biased and inefficient.

Chen and Stockum (1986), when reexamining issues on the performance of mutual funds, developed an alternative econometric procedure, known as *Generalized Varying Parameter* (GVP) *model.* The specification of this procedure results from an extension of Hidreth and Houck's (1968) *Pure Random Coefficient model* and Singh, Nagar, Choudbry and Raj's (1976) *Variable Mean Response Regression model.*

Following Chen and Stockum, the beta coefficient for portfolios, in general, and mutual funds, in particular, can be expressed by

$$\beta_{p,t} = \bar{\beta}_p + \lambda_p\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{p,t} \qquad [2]$$

The equation indicates that the systematic risk level for a portfolio **p** at time t can be decomposed into:

a) mean beta or *target beta*, i.e., the systematic risk level by which the fund manager will tolerate in average thru time and corresponds to the risk level in the absence of market timing, $\bar{\beta}_p$;

b) changes due to market timing, $\lambda_p(R_{p,t} - R_{f,t})$; and

c) random error, $\varepsilon_{p,t}$.[5]

According to Chen and Stockum, the market timing activity can be either active or passive. While active market timing requires the fund manager to shift the portfolio risk level attending to their expectations, passive market timing may occur, even if the fund managers are not actively engaged in market timing, just by the low beta and high beta securities different responses to market changes (e.g., stocks and bonds).[6] The random error component, $\varepsilon_{p,t}$, allows the risk level (beta coefficient) of a mutual fund to vary due to nonsystematic factors since the beta of a portfolio may change over time, in the absence of market timing, if the fund manager does not rebalance the fund's portfolio.[7]

The specification of equation [2] can be considered as a case of previous research. There are four reasons why this is so.

First, if the estimated values for λ_p and var($\varepsilon_{p,t}$) are not significantly different from zero, then, in the absence of market timing activities, the fixed-coefficient ordinary least squares (OLS) estimation will be an acceptable method for estimating the betas. Thus, $\beta_{p,t} = \bar{\beta}_p$, and the traditional performance evaluation measures, (e.g., Treynor (1965) and Jensen (1968)), are adequate.

If, in the other hand, the λ_p coefficient is not statistically significant and var($\varepsilon_{p,t}$) $\neq 0$, then the model becomes a *pure random coefficient model*. In this case, changes in the portfolios' betas are caused by non-market related factors.[8]

Third, if the portfolio is engaged in market timing, actively or passively, while the beta does not follow any random behavior, then the variability of $\varepsilon_{p,t}$ should not be different from zero and the λ_p coefficient should be statistically significant. A positive (negative) sign of the λ_p coefficient will suggest that the portfolio has a successful (unsuccessful) active market timing activity.

---

[5] The estimated variance of $\varepsilon_{p,t}$ can be used to measure the degree of uncertainty for $\beta_{p,t}$.

[6] Veit and Cheney (1982), among other scholars, also argued the idea that there can be active and passive market timing activities. According to these authors "*timing activities result from changing the proportion of funds devoted to risk classes as well as from the proportion of funds devoted to individual securities within a single risk class. This can be done directly by actively managed funds or indirectly by passively managed funds*". *Cit.* Veit and Cheney (1982), p. 42.

[7] Studies on random betas, such as Sunder (1980), support this argument.

[8] Fabozzi and Francis (1980) in their study conclude that only a minority of funds exhibit random betas.

But if λ_p and var($\varepsilon_{p,t}$) are statistically significant, then a *varying parameter model* should be used since the nonstationarity beta results not only from market timing activities, but also from a random behavior.

Thus, the model presented by Chen and Stockum should be considered as a generalized portfolio performance model, which could be used to estimate explicitly the constant component, the timing component and the random component of beta coefficients.[9]

Substituting equation [2] into equation [1] provides

$$\left(R_{p,t} - R_{f,t}\right) = \alpha_p + \bar{\beta}_p\left(R_{m,t} - R_{f,t}\right) + \lambda_p\left(R_{m,t} - R_{f,t}\right)^2 + \omega_{p,t} \qquad [3]$$

where

$$\omega_{p,t} = \mu_{p,t} + \varepsilon_{p,t}\left(R_{m,t} - R_{f,t}\right) \qquad [4]$$

As can be seen, this specification is very similar to the one used in Treynor and Mazuy's (1966) study to test the market timing ability of mutual funds. The major difference between the Chen and Stockum model, expressed by equation [3], and the Treynor and Mazuy model resides on the residual variable. While Treynor and Mazuy use a traditional regression analysis and assume normally distributed residuals with zero mean and constant variance (homoscedasticity assumption), Chen and Stockum admit to the possibility of heteroscedasticity. Therefore, this stochastic model is more consistent with the argument of Alexander, Benson and Eger (1982), according to which the systematic risk level of mutual funds may be nonstationary even if the funds do not engage in market timing.

3. THE ESTIMATION PROCEDURE

Chen e Stockum (1986), when estimating the selectivity and the market timing capabilities of mutual funds, make, in addition to equation [3], the following assumptions:

$$\mu_{p,t} \sim IID\left(0,\sigma_\mu^2\right);\ \varepsilon_{p,t} \sim IID\left(0,\sigma_\varepsilon^2\right) \text{ and } E\left(\mu_{p,t},\varepsilon_{p,t}\right) = 0 .$$

[9] Some studies, such as Kau, Lee and Chen (1983), Lee and Chen (1982), and Kau and Johnson (1980), use models with a similar specification, but are to be considered merely special cases of the Chen and Stockum model.

Consequently,

$$E(\omega_t, \omega_s) = 0, \quad \text{for } t \neq s \qquad [5]$$

$$E(\omega_t^2) = \sigma_\mu^2 + \sigma_\varepsilon^2 (R_{m,t} - R_{f,t})^2 \qquad [6]$$

For simplicity of analysis, let Ω be the variance-covariance matrix of ω where the diagonal elements of Ω are $E(\omega_t^2) = \Omega_t$ and the off-diagonal elements are zero. Furthermore, let $\dot{R}_m$ represent an (n×2) matrix of one's and $(R_{m,t}-R_{f,t})^2$, and Σ represent a (2×1) column vector with elements $(\sigma_\mu^2, \sigma_\varepsilon^2)$. Then

$$(\Omega_1, \Omega_2, \ldots \Omega_n)' = \dot{R}_m \Sigma \qquad [7]$$

In addition, let Y represent a column vector of elements $(R_{i,t}-R_{f,t})$, X represent an (n×3) matrix of regressors, $\{1, (R_{m,t}-R_{f,t}), (R_{m,t}-R_{f,t})^2\}$, and Θ denote a vector of parameters, $\{\alpha, \overline{\beta}, \lambda\}$. Thus, in matrix notation, equation [3] can be written as

$$Y = X\Theta + \omega \qquad [8]$$

The generalized varying parameter (GVP) model estimator Θ will be

$$\hat{\Theta} = (X'\hat{\Omega}^{-1}X)^{-1} X'\hat{\Omega}^{-1}Y, \qquad [9]$$

with the following variance-covariance matrix

$$Var(\hat{\Theta}) = (X'\hat{\Omega}^{-1}X)^{-1} \qquad [10]$$

Since

$$\begin{aligned} Cov((R_{m,t}-R_{f,t}),\omega_t) &= E[(R_{m,t}-R_{f,t}).\omega_t] - E[(R_{m,t}-R_{f,t})].E(\omega_t) \\ &= E\{(R_{m,t}-R_{f,t}).(\mu_t + (R_{m,t}-R_{f,t}).\varepsilon_t)\} - \{E(R_{m,t}-R_{f,t}).E[(\mu_t + (R_{m,t}-R_{f,t})\varepsilon_t)]\}, \end{aligned}$$

the covariance will be zero if $(R_{m,t}-R_{f,t})$, μ_t and ε_t are independent – even if the excess returns of the market portfolio, $(R_{m,t}-R_{f,t})$, is stochastic. Therefore, the least square estimator is unbiased. However, Chen and Stockum (1986) alert to the fact, as has been discussed by Lee and Jen (1978), that if $(R_{m,t}-R_{f,t})$ is subject to measurement error, then the estimated parameter vector may be biased. Also, according to Chen and Stockum (1986), the degree of bias depends on the magnitude of $Var(\omega_t)$ relative to

the magnitude of $Var[(R_{m,t}-R_{f,t})]$; hence, for a mutual fund (as a diversified portfolio) instead of an individual security, it is more likely that the variability of ω_t will be relatively small to the variability of $(R_{m,t}-R_{f,t})$, and therefore the bias could be negligible.
The estimation of Ω requires that σ_μ^2 e σ_ε^2 be estimated first. Therefore, applying the ordinary least squares (OLS) method to equation [8] yields the following

$$\hat{\omega} = Y - X\hat{\Theta} = M\omega \quad [11]$$

where M is a symmetric, idempotent matrix of order n such that

$$M = I - X(X'X)^{-1}X' . \quad [12]$$

Since the variance-covariance matrix of $\hat{\omega}$ can be expressed as

$$E(\hat{\omega}\hat{\omega}') = E(M\omega\omega' M) = M\Omega M , \quad [13]$$

the diagonal elements of $E(\hat{\omega}\hat{\omega}')$ can be expressed as

$$E(\dot{\hat{\omega}}) = \dot{M}\dot{R}_m\Sigma , \quad [14]$$

where $\dot{\hat{\omega}}$ and $\dot{M}$ are vectors and matrices of the squared elements of $\hat{\omega}$ and M, respectively. Hence, equation [14] can be written as

$$\dot{\hat{\omega}} = \dot{M}\dot{R}_m\Sigma + v = \Gamma\Sigma + v , \quad [15]$$

where v is a vector of random errors that is $v \sim IID\left(0, \sigma_v^2\right)$. Consequently, the column vector Σ can be estimated by

$$\hat{\Sigma} = (\Gamma'\Gamma)^{-1}\Gamma'\dot{\hat{\omega}} . \quad [16]$$

Finally, with the estimation of Ω, the generalized varying estimator Θ can be estimated.

4. THE DATA

As in previous years, Security Investment Funds (SIF) in Portugal were very active in 1997. There was a significant growth both in number and in value of the assets under management.[10] As shown by table 1, on December 31st 1997 there were 204 funds in activity.

TABLE 1 - NUMBER OF SECURITY INVESTMENT FUNDS

	December 31st 1997
Domestic Funds	**189**
Funds with Equities	39
Stock (or Equity) Funds	25
Balanced Funds	14
Bond Funds	51
Money Market Funds	34
Personal Retirement Plan Funds	8
Personal Equity Plan Funds	16
Funds of Funds	41
Umbrella Funds	1
International Funds	**15**
Funds with Equities	11
Stock Funds	11
Balanced Funds	0
Bond Funds	0
Money Market Funds	4
Total	**204**
N.º SIF Management Companies	**21**

Source: APFIN

Focusing our attention to domestic security investment funds, we ascertain that the funds with equities category can be partitioned into two subcategories: one incorporating 25 stock funds and another incorporating 14 balanced funds. Within the subcategory of stock funds there are 5 closed-end funds. Recall that, regarding to capital variability classification, funds can be open-end or closed-end.[11]

---

[10] For a detailed description of the Portuguese Investment Fund Industry see Sousa (1999).

[11] For a brief view on Portuguese Investment Fund classification criteria see Sousa (1999).

With the purpose of improving the usefulness of the portfolio management analysis, it is important to consider a few methodology issues.

First, our analysis is centered in two samples of security investment funds in activity on December 31st 1997. Each sample was determined by funds that belong to the same category and/or subcategory.[12] Therefore, on one hand, an analysis is done with a sample of stock funds and, on the other, an analysis is done with a sample of balanced funds. Within the sample of stock funds, five are closed-end and twenty are opened-end. In this study, we only considered the opened-end stock funds, due to the fact that a performance evaluation on closed-end funds would require a few addition assumptions. All of the selected funds are domestic. Thus, a list of security investment funds for each sample is presented:

TABLE 2 – STOCK FUNDS

Security Investment Fund Management Companies	Security Investment Funds	Code	Activity Date
BARCLAYS FUNDOS	Barclays Premier Acç. Portugal	BPAP	7-Jul.-94
BCI - SGFIM	BCI Acções Portugal	BCIAP	12-Jul.-93
BCP INVESTIMENTOS	BCP Acções	BCPA	20-Jun.-95
BCP INVESTIMENTOS	Fundo Índice BVL	IBVL	31-Mar-91
BCP INVESTIMENTOS	Novofundo Capital	NFC	31-Dez.-90
BPI FUNDOS	BPI Acções	BPIA	11-Jun.-91
BPI FUNDOS	Eurocapital	EC	3-Jan.-94
CAIXAGEST	Caixagest Acções Portugal	CXAP	20-Jun.-96
CAIXAGEST	Caixagest Valorização	CXV	5-Mar.-91
CENTRAL FUNDOS	Raiz Valorização	RV	28-Mai.-97
CPG	FIPOR Poupança Investimento	FIPOR	26-Fev.-87
FINIVALOR	Finicapital	FC	15-Abr.-97
ESAF - SGFIM	Espírito Santo Portugal Acções	ESPA	15-Set.-97
INVESTIL	BNU Acções	BNUA	24-Jun.-91
M FUNDOS	M Capital	MC	7-Jun.-95
MG FNDOS	MG Acções	MGA	1-Fev.-94
PRIMOGEST	Primus Capital	PC	20-Fev.-89
TOTTAFUNDOS	Totta Acções	TA	20-Jun.-87
TOTTAFUNDOS	Unicapital	UC	1-Out.-91
TOTTAFUNDOS	Capital Portugal	CP	30-Out.-89

[12] The classification of funds based on a specific criteria and/or its characteristics is *per si* an important issue to take into account. However, we note for the possibility that a fund may present eventually hybrid characteristics. In this case, depending on the requirements and thoroughness of the study in question, a fund classification should be accounted in determining a sample for analysis.

TABLE 3 – BALANCED FUNDS[13]

Security Investment Fund Management Companies	Security Investment Funds	Code	Activity Date
BBV GEST	BBV Global	BBVG	9-Sset.-97
BCP INVESTIMENTO	BCP Obrigações	BCPO	3-Jul.-89
BCP INVESTIMENTO	Novofundo Obrigações	NFO	31-Dez.-90
BPI FUNDOS	BPI Global	BPIG	31-Jan.-97
CAIXAGEST	Caixagest Multivalor	CXM	8-Set.-97
CPG	Mealheiro FIPOR	MFIP	19-Mar.-91
CPG	Poupança FIPOR	PFIP	11-Dez.-95
CPG	Rendimento FIPOR	RFIP	19-Jun.-89
DB FUNDOS	DB Investimento	DBI	13-Out.-95
INVESTIL	Invest	INVES	21-Mai.-86
M FUNDOS	Postal Acções	PA	23-Jun.-87
PLURIFUNDOS	Sotto Capital	SC	1-Set.-94
PRIMOGEST	Aforro FIPOR	AFIP	1-Ago.-90

Secondly, we assume that the observation time horizon coincides with the "true" investment time horizon for the investors.[14] We also assume that a week is the investors appropriate investment horizon. Thus, time series of weekly observations are used in this study.[15]

Thirdly, since the majority of the selected security investment funds begin their activity in different time dates, it was necessary to take into account this aspect in the performance analysis. For the entire sample period, going from January 1st 1993 to December 31st 1997, a "multi-case" study was carried out for each selected category of funds. The "multi-case" structure consisted in building sample time periods, within the entire period, as the funds started on their activity to December 31st 1997. Hence, we obtained for the stock fund category eleven sample time periods and for the balanced fund category six sample time periods. To each of these sample time periods (or panels) a label was assigned as the funds integrated them.[16]

---

13 Although the fund FINIGLOBAL, managed by the management company FINIVALOR, started its activity on November 11th of 1997 it was not included in the sample because it would present a very small number of observations.

14 The concept of the "true" investment horizon implies that investors will all share the same horizon. The justification and the implication of this assumption can be found in Levy (1972) and Lee (1976), as Chen and Lee (1986) recall.

15 According to Miller and Gressis (1980) the use of weekly observations seems more appropriate when attempting to detect shifts in risk-return relationships.

16 The identification of the funds in each panel are in the Appendix.

TABLE 4

Panels	Sample Time Period	N.º Obs.[17]
AEstim1	04-Jan.-1993 a 29-Dez.-1997	260
AEstim2	12-Jul.-1993 a 29-Dez.-1997	233
AEstim3	03-Jan.-1994 a 29-Dez.-1997	208
AEstim4	07-Fev.-1994 a 29-Dez.-1997	203
AEstim5	11-Jul.-1994 a 29-Dez.-1997	181
AEstim6	12-Jun.-1995 a 29-Dez.-1997	133
AEstim7	17-Jul.-1995 a 29-Dez.-1997	128
AEstim8	24-Jun.-1996 a 29-Dez.-1997	79
AEstim9	21-Abr.-1997 a 29-Dez.-1997	36
AEstim10	02-Jun.-1997 a 29-Dez.-1997	30
AEstim11	15-Set.-1997 a 29-Dez.-1997	15
MEstim1	04-Jan.-1993 a 29-Dez.-1997	260
MEstim2	15-Set.-1994 a 29-Dez.-1997	172
MEstim3	16-Out.-1995 a 29-Dez.-1997	146
MEstim4	11-Dez.-1995 a 29-Dez.-1997	107
MEstim5	03-Fev.-1997 a 29-Dez.-1997	47
MEstim6	15-Set.-1997 a 29-Dez.-1997	15

Thus, having determined our panel samples, we then collected data from January 1st 1993 to December 31st 1997. Data collection for the funds was done through the Stock Exchange Interactive Information System (SIIB)[18] and the Lisbon Stock Exchange Quotation Bulletin.

The rates of return for the selected funds were obtained from the following expression

$$R_{p,t}=\left(\frac{NAV_{p,t}-NAV_{p,t-1}+RD_{p,t}}{NAV_{p,t-1}}\right)\times 100 \qquad [17]$$

where

$R_{p,t}$ = rate of return for fund **p** at time **t**;

$NAV_{p,t}$ = net asset value for fund **p** at time **t**;

[17] In each panel the number of observations corresponds to the number of weekly returns for the funds that it groups.

[18] SIIB for *Sistema Interactivo de Informação de Bolsa.*

$NAV_{p,t-1}$ = net asset value for fund **p** at time **t-1**;

$RD_{p,t}$ = dividends and/or capital gains received and reinvested at time **t**.

Data for the General BVL Index was also obtained from the Stock Exchange Interactive Information System. The General BVL Index was used to calculate the one week rate of return as a proxy for the market portfolio:

$$R_{m,t} = \left(\frac{I_{m,t} - I_{m,t-1}}{I_{m,t-1}} \right) \times 100 \qquad [18]$$

where

$R_{m,t}$ = rate of return for the market portfolio at time **t**;

$I_{m,t}$ = value of the market index at time **t**;

$I_{m,t-1}$ = value of the market index at time **t-1**.

The rate of returns for the Interbank Money Market was used to calculate a proxy of the risk-free rate of interest for a one week horizon. Data for the Interbank Money Market was obtained from the Portuguese Central Bank.

5. THE EMPIRICAL EVIDENCE

Following the methodology approach previously mentioned, we first ascertained on the adequacy of the *Generalized Varying Parameter* (GVP) *model* suggested by Chen and Stockum (1986) for a number of portuguese security investment funds, in the attempt to investigate on the existence or not of selectivity and/or market timing.

Taking into account that, in regression analysis, it is essential to use stationary time series[19], we ran a stationary test on the variables that would be included in the our model, namely on the adjusted returns of the security investment funds, $(R_{i,t}-R_{f,t})$, and the General BVL Index, $(R_{m,t}-R_{f,t})$. To test the stationarity of these series, we used the *Argumented Dickey-Fuller* (ADF) *Test* .[20] According to the results of the ADF test, we

[19] A time series is considered to be stationary when it is generated by a stochastic process with the following properties:

a) $E(Y_t) = \mu$ constant for all t.
b) $Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$ constant for all t.
c) $Cov(Y_t,Y_{t+s}) = E[(Y_t - \mu)(Y_{t+s} - \mu)]$ constant for all $t \neq s$.

[20] In general terms, the test can be applied to a regression run in the following form

$$dY_t = \alpha + \delta^* Y_{t-1} + \delta^*_1 dY_{t-1} + \delta^*_2 dY_{t-2} + \ldots + \varepsilon_t,$$

testing the hypothesis

H_0: $\delta^* = 0$ (nonstationary hypothesis) H_1: $\delta^* < 0$ (stationary hypothesis).

According to the stationary criterion: $|\tau$ (tau) statistic $| > |\tau$ (tau) critical $|$.

concluded that, for all sample time periods in analysis, the returns of every fund in the study, as well as the returns of the General BVL Index are stationary.

Then the adequacy of the model was verified by the multiple coefficient of determination (R^2) and by the F test, which is a measure of the overall significance of the estimated regression.

Observing the values of R^2, summarized on tables 5 and 6, one can verify that the estimated model reveals its inadequacy for some funds, namely for balanced funds. Subsequently, it was confirmed by the F test the inadequacy of the estimated model for the same funds. The analysis of these results allows us to assert that, for various panels, at the chosen level of significance of 5%, as well as 10%, the estimated model for the balanced funds BCPO, NFO, MFIP, RFIP, AFIP, PFIP and CXM is not adequate. The inadequacy of the mentioned balanced funds mainly results from the composition of their portfolios. On one hand, the stock component is null or very small in weight and, on the other, the bond component is dominant in weight. It would then be necessary, when constructing the market portfolio, as a benchmark, to consider also a bond index and not only a stock index as was done.

TABLE 5 - (R^2) for the GVP model

Code	AEstim1	AEstim2	AEstim3	AEstim4	AEstim5	AEstim6	AEstim7	AEstim8	AEstim9	AEstim10	AEstim11
IBVL	0,4513	0,4508	0,4829	0,4997	0,6195	0,6255	0,6291	0,6586	0,6598	0,7325	0,6301
NFC	0,4695	0,4911	0,4826	0,4820	0,5673	0,5828	0,5873	0,6350	0,6972	0,7350	0,6420
BPIA	0,7095	0,7420	0,7971	0,7914	0,8600	0,9087	0,9083	0,9293	0,9729	0,9696	0,9521
CXV	0,5303	0,6114	0,6918	0,6823	0,6805	0,7167	0,7143	0,7287	0,6879	0,6591	0,5259
FIPOR	0,2902	0,3626	0,4191	0,4421	0,5281	0,7905	0,8293	0,9297	0,6823	0,7298	0,6539
BNUA	0,9972	0,8587	0,7437	0,7472	0,7375	0,7618	0,7560	0,7650	0,7558	0,7291	0,6647
PC	0,3729	0,3874	0,3758	0,3688	0,5948	0,5907	0,5911	0,6375	0,6810	0,7337	0,6461
TA	0,5508	0,5514	0,8308	0,5377	0,4809	0,5794	0,5844	0,6536	0,6772	0,7161	0,6456
UC	0,4524	0,4768	0,4401	0,4509	0,4373	0,5859	0,5900	0,6474	0,6863	0,7607	0,7843
CP	0,4757	0,4515	0,4025	0,3892	0,5340	0,5424	0,5429	0,5912	0,5615	0,5899	0,6069
BCIAP	*	0,4914	0,5148	0,5374	0,8779	0,7472	0,7490	0,7710	0,7681	0,7505	0,5693
EC	*	*	0,6671	0,7039	0,7301	0,7322	0,7336	0,7683	0,7827	0,7660	0,6063
MGA	*	*	*	0,5317	0,5904	0,6552	0,6613	0,7306	0,7243	0,6952	0,5619
BPAP	*	*	*	*	0,6533	0,7257	0,7278	0,7558	0,7828	0,7501	0,5902
MC	*	*	*	*	*	0,5846	0,5899	0,5653	0,5423	0,5636	0,4488
BCPA	*	*	*	*	*	*	0,5982	0,6391	0,6730	0,7121	0,6316
CXAP	*	*	*	*	*	*	*	0,7320	0,6857	0,6634	0,5828
FC	*	*	*	*	*	*	*	*	0,6382	0,6688	0,4970
RV	*	*	*	*	*	*	*	*	*	0,6490	0,0325
ESPA	*	*	*	*	*	*	*	*	*	*	0,5340
R^2											
]0,0;0,5[	**6**	**7**	**6**	**6**	**2**	**0**	**0**	**0**	**0**	**0**	**2**
]0,5;0,75[	**3**	**3**	**4**	**6**	**10**	**12**	**13**	**11**	**13**	**15**	**16**
]0,75;1,0[	**1**	**1**	**2**	**1**	**2**	**3**	**3**	**6**	**5**	**4**	**2**

If the null hypothesis is not rejected, then the time series is stationary.

TABLE 6 - (R^2) for the GVP model

Code	MEstim1	MEstim2	MEstim3	MEstim4	MEstim5	MEstim6
BCPO	0,1892	0,1418	0,3972	0,5958	0,7118	0,7094
NFO	0,4056	0,1142	0,6071	0,6210	0,5361	0,2811
MFIP	0,0177	0,0741	0,2722	0,3368	0,2918	0,2017
RFIP	0,0232	0,0268	0,3249	0,4072	0,4570	0,7659
INVES	0,2738	0,5332	0,7372	0,7501	0,7581	0,6381
PA	0,2777	0,5294	0,4140	0,4464	0,4258	0,5723
AFIP	0,0044	0,0469	0,0869	0,0937	0,1340	0,3624
SC	*	0,4183	0,7273	0,6609	0,4893	0,6715
DBI	*	*	0,7615	0,7544	0,7747	0,7413
PFIP	*	*	*	0,6264	0,3315	0,6394
BPIG	*	*	*	*	0,8077	0,9220
CXM	*	*	*	*	*	0,1895
BBVG	*	*	*	*	*	0,5093
R^2						
[0,0;0,5[	7	6	5	4	6	7
[0,5;0,75[	0	2	3	4	2	5
[0,75;1,0[	0	0	1	2	3	1

Thereafter, with the purpose of determining if the funds in the study experimented eventually a structural change in their systematic risk level, namely in the components *target beta,* $\overline{\beta}_p$, and *market timing* , λ_p , Chow's structural stability tests were performed. The Chow test, for all the funds in each of the panels, involved the following steps:

a) the global estimation period was divided into two subperiods with N_1, the number of observations for the first subperiod, and N_2, the number of observations for the second subperiod;

b) the model was estimated for the global period with N observations, and its residual sum of squares (SQR_N) was obtained with (N-k) degrees of freedom[21];

c) the model was estimated for each of the subperiods obtaining the residual sum of squares SQR_{N1} and SQR_{N2} , respectively, with (N_1-k) and (N_2-k) degrees of freedom;

d) the test following the F distribution was expressed as:

$$F_{est.} = \frac{\left[SQR_N - \left(SQR_{N_1} + SQR_{N_2}\right)\right] / k}{\left(SQR_{N_1} + SQR_{N_2}\right) / \left(N_1 + N_2 - 2k\right)} \quad \sim \quad F_{k,N1+N2-2k} \qquad [19]$$

e) the stability criterion was analyzed.[22]

---

[21] k corresponds to the total number of estimated coefficients.

[22] If the F computed from [19] exceeds the critical F at the chosen level of significance (i.e., Fest. > Fcrit.), then the null hypothesis would be rejected. In this case, there would be evidence that during the global period structural change had occured. If, in the other hand, the F obtained from [19] does not exceed the critical F (i.e., Fest. < Fcrit.), then the null hypothesis of structural stability is not rejected.

The results of the stability tests are summarized on tables 7 and 8 for stock funds and balanced funds, respectively. In the tables, it should be taken into account that, at the chosen level of significance, only the filled cells represent the situations were the null hypothesis of structural stability is not rejected for the model. According to the results, we can ascertain that whenever the sample time horizon in analysis corresponds to panels AEstim5 through AEstim11, for stock funds, and to panels MEstim2 through MEstim6, for balanced funds, the estimated model is relatively stable, therefore significant changes in the fund's systematic risk level components did not occur.

TABLE 7 - Stability Test Results for Stock funds

Code	AEstim1	AEstim2	AEstim3	AEstim4	AEstim5	AEstim6	AEstim7	AEstim8	AEstim9	AEstim10	AEstim11
IBVL					1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
NFC					5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
BPIA					1%	5%, 1%	5%, 1%		5%, 1%	5%, 1%	1%
CXV					1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	1%
FIPOR					1%	5%, 1%	5%, 1%	5%, 1%	1%	5%, 1%	5%, 1%
BNUA					5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
PC	5%, 1%	5%, 1%	1%	1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	1%	5%, 1%	5%, 1%
TA						5%, 1%	5%, 1%	5%, 1%	1%	5%, 1%	1%
UC						1%	1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
CP					1%	5%, 1%	5%, 1%	1%	1%	5%, 1%	5%, 1%
BCIAP	*					5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
EC	*	*				5%, 1%	5%, 1%	5%, 1%	1%	5%, 1%	5%, 1%
MGA	*	*	*			5%, 1%	5%, 1%	5%, 1%	1%	1%	1%
BPAP	*	*	*	*	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
MC	*	*	*	*	*	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
BCPA	*	*	*	*	*	*	5%, 1%	5%, 1%	1%	5%, 1%	5%, 1%
CXAP	*	*	*	*	*	*	*	5%, 1%	5%, 1%	5%, 1%	5%, 1%
FC	*	*	*	*	*	*	*	*	5%, 1%	5%, 1%	1%
RV	*	*	*	*	*	*	*	*	*	5%, 1%	5%, 1%
ESPA	*	*	*	*	*	*	*	*	*	*	1%

TABLE 8 - Stability Test Results for Balanced funds

Code	MEstim1	MEstim2	MEstim3	MEstim4	MEstim5	MEstim6
BCPO	1%	5%, 1%	5%, 1%		5%, 1%	5%, 1%
NFO		5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
MFIP		5%, 1%	5%, 1%	5%, 1%		5%, 1%
RFIP	5%, 1%		5%, 1%	5%, 1%	1%	5%, 1%
INVES		5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
PA		5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
AFIP	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%	5%, 1%
SC	*	5%, 1%	1%	5%, 1%	5%, 1%	5%, 1%
DBI	*	*	1%	5%, 1%	1%	1%
PFIP	*	*	*		5%, 1%	5%, 1%
BPIG	*	*	*	*	5%, 1%	5%, 1%
CXM	*	*	*	*	*	1%
BBVG	*	*	*	*	*	5%, 1%

Secondly, summary results of the *Generalized Varying Parameter* (GVP) estimates for the portuguese security investment funds being investigated, in each panel, are presented in tables 9 through 25.

Column 1 presents results for σ_ε^2, which measures the variability of ε_t. It measures a source of nonstationarity for the systematic risk level (beta coefficient) of a security investment fund. The inclusion of ε_t in equation [2] allows the beta nonstationarity of security investment funds to be stochastic. In fact, this flexibility is desirable since a security investment fund's systematic risk level (beta coefficient) may be nonstationarity, even if the fund manager is not engaged in market timing decisions, as argued by Alexander, Benson and Eger (1982). According to some scholars, the nonstationarity captured by the variability of ε_t result from non-market related factors.

On the other hand, columns 2, 3 and 4 report fund's selectivities (α_p), target systematic risk ($\bar{\beta}_p$) and market timing performance (λ_p), respectively.

SUMMARY RESULTS FOR THE GVP MODEL

TABLE 9

AEstim1	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	9	6	10	5
N.º of negative coefficients	1	4	0	5
N.º of significant positive coefficients	8	0	10	3
N.º of significant negative coefficients	0	0	0	0
Mean	*	0,0619	0,5837	-0,0005
Maximum	*	0,1938	0,8507	0,0191
Minimum	*	-0,1155	0,4731	-0,0114
Median	*	0,1152	0,5291	-0,0009
Standard deviation	*	0,1191	0,1210	0,0087

TABLE 10

AEstim2	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	10	7	11	6
N.º of negative coefficients	1	4	0	5
N.º of significant positive coefficients	9	0	11	3
N.º of significant negative coefficients	0	0	0	0
Mean	*	0,0669	0,5923	0,0007
Maximum	*	0,1952	0,7175	0,0188
Minimum	*	-0,1168	0,492	-0,0111
Median	*	0,1044	0,5662	0,0008
Standard deviation	*	0,104	0,0803	0,0076

TABLE 11

AEstim3	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	11	9	12	5
N.º of negative coefficients	1	3	0	7
N.º of significant positive coefficients	9	0	12	1
N.º of significant negative coefficients	0	0	0	0
Mean	*	0,0662	0,5980	0,0011
Maximum	*	0,1726	0,7256	0,0108
Minimum	*	-0,0755	0,4938	-0,0061
Median	*	0,1080	0,5690	-0,0007
Standard deviation	*	0,0925	0,0796	0,0054

TABLE 12

AEstim4	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	12	9	13	6
N.° of negative coefficients	1	4	0	7
N.° of significant positive coefficients	10	0	13	1
N.° of significant negative coefficients	0	0	0	0
Mean	*	0,0647	0,5995	0,0008
Maximum	*	0,1812	0,7376	0,0092
Minimum	*	-0,0877	0,4840	-0,0060
Median	*	0,1023	0,5718	-0,0001
Standard deviation	*	0,0901	0,0826	0,0046

TABLE 13

AEstim5	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	10	14	14	0
N.° of negative coefficients	4	0	0	14
N.° of significant positive coefficients	1	2	14	0
N.° of significant negative coefficients	0	0	0	10
Mean	*	0,1287	0,7024	-0,0118
Maximum	*	0,3778	0,8410	-0,0010
Minimum	*	0,0262	0,5881	-0,0204
Median	*	0,0999	0,6832	-0,0121
Standard deviation	*	0,0998	0,0849	0,0048

TABLE 14

AEstim6	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	8	14	15	0
N.° of negative coefficients	7	1	0	15
N.° of significant positive coefficients	0	1	15	0
N.° of significant negative coefficients	0	0	0	10
Mean	*	0,1221	0,7552	-0,0135
Maximum	*	0,3770	0,9041	-0,0020
Minimum	*	-0,0639	0,6142	-0,0229
Median	*	0,1013	0,7448	-0,0152
Standard deviation	*	0,1123	0,0812	0,0064

TABLE 15

AEstim7	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	6	15	16	0
N.° of negative coefficients	10	1	0	16
N.° of significant positive coefficients	0	1	16	0
N.° of significant negative coefficients	0	0	0	11
Mean	*	0,1407	0,7522	-0,0143
Maximum	*	0,3976	0,9053	-0,0021
Minimum	*	-0,0514	0,6115	-0,0233
Median	*	0,1282	0,7353	-0,0154
Standard deviation	*	0,1094	0,0778	0,0067

TABLE 16

AEstim8	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	4	16	17	1
N.° of negative coefficients	13	1	0	16
N.° of significant positive coefficients	0	2	17	1
N.° of significant negative coefficients	0	0	0	9
Mean	*	0,2379	0,7720	-0,0193
Maximum	*	0,4512	0,9114	0,0005
Minimum	*	-0,0737	0,5902	-0,0325
Median	*	0,2632	0,7791	-0,0197
Standard deviation	*	0,1238	0,0845	0,0104

TABLE 17

AEstim9	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	0	17	18	0
N.° of negative coefficients	18	1	0	18
N.° of significant positive coefficients	0	0	18	0
N.° of significant negative coefficients	0	0	0	8
Mean	*	0,2675	0,8036	-0,0351
Maximum	*	0,4836	0,9322	-0,0008
Minimum	*	-0,1649	0,6594	-0,0549
Median	*	0,2923	0,8154	-0,0355
Standard deviation	*	0,1668	0,0692	0,0159

TABLE 18

AEstim10	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	0	17	19	0
N.° of negative coefficients	19	2	0	19
N.° of significant positive coefficients	0	0	19	0
N.° of significant negative coefficients	0	0	0	3
Mean	*	0,2051	0,7927	-0,0307
Maximum	*	0,3837	0,9325	-0,0010
Minimum	*	-0,2027	0,5206	-0,0448
Median	*	0,251	0,8153	-0,0340
Standard deviation	*	0,1478	0,0984	0,011

TABLE 19

AEstim11	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.° of positive coefficientes	0	4	20	8
N.° of negative coefficients	20	16	0	12
N.° of significant positive coefficients	0	0	20	0
N.° of significant negative coefficients	0	0	0	0
Mean	*	-0,119	0,856	-0,0125
Maximum	*	0,2000	1,0104	0,0654
Minimum	*	-0,7229	0,6644	-0,0675
Median	*	-0,0701	0,8490	-0,0045
Standard deviation	*	0,2304	0,1024	0,0378

TABLE 20

MEstim1	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	4	0	4	2
N.º of negative coefficients	3	7	3	5
N.º of significant positive coefficients	2	0	3	0
N.º of significant negative coefficients	0	2	0	0
Mean	*	-0,3617	0,062	-0,0002
Maximum	*	-0,0609	0,2238	0,0008
Minimum	*	-1,0786	-0,0172	-0,0028
Median	*	-0,2137	0,0051	0,0000
Standard deviation	*	0,3859	0,1035	0,0012

TABLE 21

MEstim2	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	3	2	4	2
N.º of negative coefficients	5	6	4	6
N.º of significant positive coefficients	0	0	4	0
N.º of significant negative coefficients	0	4	2	2
Mean	*	-0,3157	0,1416	-0,0018
Maximum	*	0,0877	0,4503	0,0000
Minimum	*	-1,0576	-0,0073	-0,0082
Median	*	-0,3197	0,0207	-0,0002
Standard deviation	*	0,3738	0,1936	0,0029

TABLE 22

MEstim3	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	6	3	6	2
N.º of negative coefficients	3	6	3	7
N.º of significant positive coefficients	0	0	4	0
N.º of significant negative coefficients	0	5	1	3
Mean	*	-0,5285	0,2102	-0,0031
Maximum	*	0,1858	0,8171	0,0000
Minimum	*	-1,4218	-0,0050	-0,0123
Median	*	-0,5686	0,0023	0,0000
Standard deviation	*	0,6174	0,2885	0,0044

TABLE 23

MEstim4	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	7	3	6	1
N.º of negative coefficients	3	7	4	9
N.º of significant positive coefficients	0	0	5	0
N.º of significant negative coefficients	0	6	2	5
Mean	*	-0,5825	0,1917	-0,0032
Maximum	*	0,2083	0,8381	0,0000
Minimum	*	-1,4970	-0,0066	-0,0153
Median	*	-0,5793	0,0040	0,0000
Standard deviation	*	0,6651	0,2857	0,0052

TABLE 24

MEstim5	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	4	4	8	1
N.º of negative coefficients	7	7	3	10
N.º of significant positive coefficients	1	0	6	0
N.º of significant negative coefficients	0	6	0	3
Mean	*	-0,5346	0,2120	-0,0069
Maximum	*	0,3375	0,9010	0,0001
Minimum	*	-1,9928	-0,0052	-0,0297
Median	*	-0,6812	0,0042	0,0000
Standard deviation	*	0,7117	0,2927	0,0115

TABLE 25

MEstim6	epsilon variability	alfa (selectivity)	target beta	lambda (market timing)
N.º of positive coefficientes	2	1	11	6
N.º of negative coefficients	11	12	2	7
N.º of significant positive coefficients	0	0	6	0
N.º of significant negative coefficients	1	3	0	0
Mean	*	-0,6609	0,2684	0,0086
Maximum	*	0,0340	0,8573	0,1166
Minimum	*	-2,0534	-0,0355	-0,0538
Median	*	-0,5484	0,0163	0,0000
Standard deviation	*	0,6431	0,3422	0,0379

Hence, we present a summary analysis of the results for each panel[23]:

The empirical results for panel AEstim1 (with sample period from Jan. 4th 1993 to Dec. 29th 1997) suggest that 6 funds have a positive selectivity measure (alpha coefficient) and 5 have a positive timing performance (lambda coefficient). However, none of the alpha coefficients are significant, but there are 3 statistically significant lambda coefficients. Results also suggest that 8 out of 10 stock funds present a nonstationarity systematic risk level due to non-market related factors, which is captured by the variability of $\varepsilon_{p,t}$.

The results from panel AEstim2 (with sample period from Jul. 12th 1993 to Dec. 29th 1997) show that 7 out of 11 stock funds have a positive selectivity measure, although not statistically significant, and that 3 out of 6 have a significant positive timing performance. There are 9 stock funds that present a nonstationarity beta.

For panel AEstim3 (with sample period from Jan. 3rd 1994 to Dec. 29th 1997), the results show that 9 out of 12 stock funds have a positive selectivity measure, although not statistically significant, and that only 1 out of 5 has a significant positive timing performance. 9 funds present a nonstationarity systematic risk level related to non-market factors.

For panel AEstim4 (with sample period from Feb. 7th 1994 to Dec. 29th 1997), the empirical results reveal that 9 stock funds have a positive alpha coefficient, although none statistically significant, and that only 1 has a significant positive timing

[23] Recall that panels labeled with AEstim_ incorporate stock funds, while panels labeled with MEstim_ incorporate balanced funds.

performance. 10 funds present a nonstationarity systematic risk level, captured by the variability of $\varepsilon_{p,t}$.

The results for panel AEstim5 (with sample period from Jul. 11th 1994 to Dec. 29th 1997) show that 14 stock funds have a positive selectivity, although only 2 are statistically significant, and that 10 have a negative market timing performance.

The results for panel AEstim6 (with sample period from Jun. 12th 1995 to Dec. 29th 1997) suggest that only 1 fund has a statistically significant positive selectivity measure and that 10 have a negative timing performance which are statistically significant. None of the funds in this sample show a nonstationarity beta coefficient.

For panel AEstim7 (with sample period from Jul. 17th 1995 to Dec. 29th 1997), the empirical results reveal that only 1 out of 15 positive alpha coefficients has a significant value and that 11 out of 16 negative lambda coefficients are significant values. All the funds present a stationarity systematic risk level since $var(\varepsilon_{p,t}) \approx 0$.

For panel AEstim8 (with sample period from Jun. 24th 1996 to Dec. 29th 1997), the empirical results show that only 2 funds have positive selectivity which is statistically significant and that 9 funds present a negative timing performance, measured by the lambda coefficient. Again all the funds present a stationarity systematic risk level with $var(\varepsilon_{p,t}) \approx 0$.

The results for panel AEstim9 (with sample period from Apr. 21st 1997 to Dec. 29th 1997) reveal that none of the 17 positive alpha coefficients are significant and that 8 funds present a negative market timing measure which are statistically significant. None of the funds in this sample show a nonstationarity beta coefficient, since $var(\varepsilon_{p,t}) \approx 0$.

The empirical results for panel AEstim10 (with sample period from Jun. 2nd 1997 to Dec. 29th 1997) suggest that 17 stock funds have a positive selectivity measure, but none are statistically significant, and that 3 out of 19 negative timing performance are significant. The funds in this sample reveal a stationarity beta coefficient.

For panel AEstim11 (with sample period from Set. 15th 1997 to Dec. 29th 1997), the results indicate that 4 funds have positive values for alpha coefficients while 16 have negative values, although none are significant. 12 stock funds have negative lambda coefficients, all of which are not significant. None of the funds in this sample suggest a nonstationarity systematic risk level (beta coefficient).

According to the results for panel MEstim1 (with sample period from Jan. 4th 1993 to Dec. 29th 1997) there are 7 balanced funds in the sample, none have positive selectivity and only 2 present a positive timing measure. Results also reveal that 2 funds have nonstationarity systematic risk level due to non-market related factors.

For panel MEstim2 (with sample period from Set. 15th 1994 to Dec. 29th 1997), the results show 4 funds with negative and significant selectivity measures, and 2 with negative and significant timing. All funds present a null variability of $\varepsilon_{p,t}$.

The empirical results for panel MEstim3 (with sample period from Oct. 16th 1995 to Dec. 29th 1997) indicate that 5 out of 6 funds with negative alpha coefficient are statistically significant and that 3 funds out of 7 funds with negative lambda coefficient are significant. All funds in this sample present a null variability of $\varepsilon_{p,t}$.

For panel MEstim4 (with sample period from Dec. 11th 1995 to Dec. 29th 1997), the results suggest that 6 funds have a negative and significant selectivity measure, and that 5 funds have a negative and significant timing measure. Once again, all funds do not show evidence of nonstationarity systematic risk level.

For panel MEstim5 (with sample period from Feb. 3rd 1997 to Dec. 29th 1997), the results suggest that 6 funds have statistically significant negative alpha values and that 3 funds have significant negative market timing performance. There is a fund that presents a nonstationarity systematic risk level.

The results for panel MEstim6 (with sample period from Set. 15th 1997 to Dec. 29th 1997) show that 3 out of 12 negative alpha coefficients are statistically significant and that none of the funds show significant lambda coefficients.

During the analysis, we have considered a 5% level of significance in the significance tests.

It should also be noted that following one of Hidreth and Houck's (1968) suggestions, negative σ_ε^2 estimates are assigned a zero value. Therefore, the GVP estimates will be essentially the same as the OLS estimates for those funds with negative σ_ε^2.

From the results shown throughout the panel analysis, we can identify a few characteristics for the estimates on selectivity (α_p), target beta ($\overline{\beta}_p$) and market timing (λ_p). Hence, we present the following issues:

According to the mean values for the selectivity measure estimates, the stock funds in almost all panels obtained on average a superior performance compared with the market portfolio, namely the General BVL Index. The balanced funds, however, obtained on average an inferior performance than the market portfolio, showing negative mean values for the selectivity estimates.[24] A closer examination on the statistical significance of the selectivity (alpha coefficient) estimates was performed. Thus, statistical significance tests, also known as t-tests, were carried out. Evidence from the tests suggest that, globally for stock funds, the estimated values of alpha coefficient (α_p) are not statistically significant, i.e., $\alpha_p \approx 0$. We can, therefore, ascertain that, for the all funds, managers in general terms are not able to select over and undervalued securities, although a few in a particular moment tend to oppose this trend.

When analysing the target beta, $\overline{\beta}_p$, estimates, we encountered two interesting aspects. First, we verified that target beta estimates on average are higher for stock

[24] As we expected, the balanced funds DBI, SC and PA with a high proportion of stocks in their compositions, being more closely related to stock funds, present positive alpha coefficients.

funds than for balanced funds, on one hand, and that the variability of target beta estimates on average are also higher for stock funds than for balanced funds, on the other. Secondly, stock funds revealed a larger sensitivity to market movements when compared to balanced funds, as can be shown by the mean, maximum, minimum and median values of the target beta estimates in all sample time periods. These results suggest that the composition of the investment funds must be intimately related to the composition of the market portfolio, since for all the funds in analysis that did not reject the adequacy of the model the target beta estimates are statistically significant (at a 5% level of significance).

Examining the mean values of the market timing estimates, the stock funds, not to mention the balanced funds, in practically all panels obtained on average a weak performance for active timing activities. In testing the individual significance on timing performance, we verified that the majority of its statistically significant estimates presented negative values very close to zero.

Another important implication of this study deals with the question of whether portuguese security investment funds engage in market timing activities by shifting their (systematic) risk levels to take advantage of market movements. The empirical evidence suggests that exceptionally security investment funds are engaged in somewhat (un)successful timing activities. In general, the unexistence of successful market timing activities may result because investment managers simply do not attempt timing activity or because they are unsuccessful at timing. We believe that the decision of investment managers to make no attempt of timing results essentially from one or a combination of the following factors: (a) the inability to forecast market movements in general; (b) the unwillingless to directly change the fund's systematic risk level and/or (c) the unwillingless to bear transaction costs which would be superior to potential gains due from changing the composition of the funds.

However, even for security investment funds not practicing active timing strategies, the systematic risk level can change. In effect, if, in general, we conceive the (global) systematic risk level, β_p, of a portfolio as a weighted average of the systematic risk of individual securities comprising the portfolio, then it is easy to understand how market movements can influence the systematic risk level of portfolios, by changing the (systematic) risk levels of its components.[25] This situation will occur in any investment fund that is "passively" managed, since the systematic risk (beta) level will not depend exclusively on the manager revising the portfolio. For example, being a fund composed by high and low beta securities in its portfolio, the high beta securities will decline in price more than the low beta securities during a bear market; hence, the low beta securities will represent a larger proportion of the value of the investment fund than they did at the market peak. Therefore, investment funds that do not attempt timing may present a nonstable (global) beta because of their passive management, by indirectly forming a low beta portfolio prior to a bull market or a high beta portfolio prior to a bear market. In this study, the results also suggest that, for the sample time periods with a large number of weekly observations, there might be changes in the investment funds systematic risk level due to passive management, namely when in analysis we consider more then a 180 observations.

---

[25] Veit and Cheney (1982) present this idea on (global) systematic risk level of a portfolio.

6. CONCLUSION

The paper adopts a general model to investigate security investment funds' performances by taking into account both timing decisions of funds and random behavior of funds' systematic risk levels. The empirical results we get by using the *Generalized Varying Parameter* (GVP) *model* from Chen and Stockum (1986) indicate that portuguese security investment fund managers in general do not show ability in selectivity, although a few in a particular moment tend to oppose this trend, and do not show ability in successful (active) timing.

In spite of our particular attempt to evaluate portfolio managers, in this case security investment fund managers, we feel that the important question of fund performance evaluation cannot be completely answered by this study. Thus, we believe that further research in this area is needed.

REFERENCES

ALEXANDER, G. and STOVER, R. (1980), "Consistency of mutual fund performance during varying market conditions", *Journal of Economics and Business*, Spring, pp. 219-226.

ALEXANDER, G. J.; BENSON, P. G. and EGER, C. E. (1982), "Timing decisions and the behavior of mutual fund systematic risk", *Journal of Financial and Quantitative Analysis* , vol. 17, nº.4, November, pp. 579-602.

BLUME, M. (1971), "On the assessment of risk", *The Journal of Finance*, vol. 26, nº1, March, pp. 1-10.

BLUME, M. (1975), "Betas and their regression tendencies", *The Journal of Finance*, vol. 30, nº3, June, pp. 785-795.

CHEN, C. and STOCKUM, S. (1986), "Selectivity, market timing, and random beta behavior of mutual funds: a generalized model", *The Journal of Financial Research*, vol. 9, nº.1, Spring, pp. 87-96.

CHEN, S. and LEE, C. (1986), "The effects of the sample size, the investment horizon and market conditions on the validity of composite performance measures: a generalization", *Management Science*, vol. 32, n.º11, November, pp. 1410-1421.

CHOW, G. C. (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions", *Econometrica*, vol. 28, no. 3, pp. 591-605.

DOAN, T. A. (1992) *RATS user's manual - version 4*, Estima, Evanston, Illinois.

FABOZZI, F. and FRANCIS, J. C. (1978), "Beta as a random coefficient", *Journal of Financial Quantitative Analysis*, vol. 13, nº1, March, pp. 101-116.

FABOZZI, F. and FRANCIS, J. C. (1979), "Mutual fund systematic risk for bull and bear markets: an empirical examination", *The Journal of Finance*, vol. 34, nº5, December, pp. 1243-1250.

FABOZZI, F. and FRANCIS, J. C. (1980), "Stability of Mutual Fund Systematic Risk Statistics", *Journal of Business Research*, June, pp. 263-275.

FERSON, W. and SCHADT; R. (1996), "Measuring fund strategy and performance in changing economic conditions", *The Journal of Finance*, vol. 51, nº2, June, pp. 425-461.

HILDRETH, C. and HOUCK, J. (1968), "Some Estimators for a linear Model with Random Coefficient", *Journal of the American Statistical Association*, June, pp. 584-595;

JENSEN, M. C. (1968), "The performance of mutual funds in the period 1945-1964", *The Journal of Finance*, vol. 23, nº.2, May, pp. 389-416.

JENSEN, M. C. (1969), "Risk, the pricing of capital assets, and the evaluation of investment portfolios", *Journal of Business*, vol. 42, n.º2, April, pp. 167-247.

KAU, J. and JOHNSON, S. (1980), "Urban Spatial Structure: an Analysis with Varying Coefficient Model", *Journal of Urban Economics*, vol. 7, pp. 141-154.

KAU, J., LEE, C. and CHEN, C. (1983), " Structural Shifts in Urban Population Density Gradients: an Empirical Investigation", *Journal of Urban Economics*, vol. 12, pp. 364-377;

KLEMKOSKY, R. and MANESS, T. (1978), "The predictability of real portfolio risk levels", *The Journal of Finance*, vol. 33, May, pp. 631-639.

KLEMKOSKY, R. C. and MARTIN, J. D. (1975), "The adjustment of beta factors", Journal of Finance, vol. 30, pp. 1123-1128

KON, S. (1983), "The market-timing performance of mutual fund managers", *Journal of Business*, vol. 56, nº.3, July, pp. 323-347.

KON, S. J. and JEN, F. C. (1978), "Estimation of the time-varying systematic risk and performance for mutual fund portfolios: an application of switching regression", *The Journal of Finance*, vol. 33, nº2, pp. 457-475.

KON, S. J. and JEN, F. C. (1979), "The investment performance of mutual funds: an empirical investigation of timing, selectivity and market efficiency", *Journal of Business*, vol. 52, nº2, April, pp. 263-289.

KRYZANOWSKI, L.; LALANCETTE, S. and TO, MINH C. (1997), "Performance Attribution using a APT with Prespecicied Macrofactors and Time-Varying Risk Premia and Betas", Journal of Finance and Quantative Analysis, vol.32, n.º2, June, pp.205-224.

LEE, C. and CHEN, C. (1982), "Beta Stability and Tendency: an Application of a Variable Mean Response Regression Model", *Journal of Economics and Business*, vol. 34, July, pp. 201-206;

LEE, C. and JEN, F. (1978), "Effects of measurement errors on systematic risk and performance measure of a portfolio", *Journal of Financial and Quantitative Analysis*, vol. 13, nº2, June, pp. 299-312.

LEE, C. F. (1976) "Investment Horizon and the Functional Form of the Capital Asset Pricing Model", Review of Economics and Statstics, vol. 58, August, pp. 356-363.

LEVHARI, D. e LEVY, H. (1977), "The Capital Asset Pricing Model and the Investment Horizon", *The Review of Economics and Statistics*, February, pp. 92-104.

LEVY, H. (1972), "Portfolio Performance and Investment Horizon", *Management Science*, vol. 18, nº.12, August, pp. 645-653.

LEVY, R. (1971), "On short-term stationary of beta coefficients", *Financial Analysts Journal*, November-December, pp. 55-62.

LINTNER, J. (1965), "The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics*, vol. 47, nº1, February, pp. 13-37.

MILLER, T. e GRESSIS, N. (1980), "Nonstationarity and evaluation of mutual fund performance", *Journal of Financial and Quantitative Analysis*, vol. 15, nº.3, September, pp. 639-654.

MOSSIN, J. (1966), "Equilibrium in a capital asset market", *Econometrica*, vol. 34, nº4, October, pp. 768-783.

SHARPE, W. F. (1964), "Capital asset prices: a theory of market equilibrium under conditions of risk", *The Journal of Finance*, vol. 19, nº3, September, pp. 425-444.

SINGH, B., NAGAR, L., CHOUDBRY, N. K. and RAJ, B. (1976), "On the Estimation of Structral Change: A Generalization of the Random Coefficient Regression Model", *International Economic Review*, pp. 340-361.

SOUSA, P. (1999), *Uma Avaliação da Performance de Fundos de Investimento Mobiliário em Portugal*, Tese de Mestrado em Economia Financeira, Faculdade de Economia da Universidade de Coimbra, Março.

TREYNOR, J. L. (1965), "How to rate management of investment funds", *Harvard Business Review*, vol. 43, nº1, January-February, pp. 63-75.

TREYNOR, J. L. and MAZUY, K. K. (1966), "Can mutual funds outguess the market", *Harvard Business Review*, vol. 44, n.º4, July-August, pp. 131-136.

VEIT, E. T. and CHENEY, J. M. (1982), "Are mutual funds market timers?", *The Journal of Portfolio Management*, vol. 9, Winter, pp. 35-42.

WHITE, H. (1980), "A heteroscedasticity - consistent covariance matrix estimation and a direct test for heteroscedasticity", *Econometrica*, vol. 48, nº4, May, pp. 817-838.

APPENDIX - SECURITY INVESTMENT FUNDS PER PANEL

Table A.1 - Security Investment Funds in Panel AEstim1

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	04 - Jan. - 1993 a 29 - Dez. - 1997	260
BPIA	BPI Acções	11 - Jun. - 1991	04 - Jan. - 1993 a 29 - Dez. - 1997	260
CXV	Caixagest Valorização	05 - Mar. - 1991	04 - Jan. - 1993 a 29 - Dez. - 1997	260
CP	Capital Portugal	30 - Out. - 1989	04 - Jan. - 1993 a 29 - Dez. - 1997	260
IBVL	Fundo IBVL	31 - Mar. - 1991	04 - Jan. - 1993 a 29 - Dez. - 1997	260
NFC	Novofundo Capital	31 - Dez. - 1990	04 - Jan. - 1993 a 29 - Dez. - 1997	260
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	04 - Jan. - 1993 a 29 - Dez. - 1997	260
PC	Primus Capital	20 - Fev. - 1989	04 - Jan. - 1993 a 29 - Dez. - 1997	260
TA	Totta Acções	22 - Jun. - 1987	04 - Jan. - 1993 a 29 - Dez. - 1997	260
UC	Unicapital	01 - Out. - 1991	04 - Jan. - 1993 a 29 - Dez. - 1997	260

Table A.2 - Security Investment Funds in Panel AEstim2

Code	*Funds*	*Activity Date*	*Sample Time Period*	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	12 - Jul. - 1993 a 29 - Dez. - 1997	233
BPIA	BPI Acções	11 - Jun. - 1991	12 - Jul. - 1993 a 29 - Dez. - 1997	233
CXV	Caixagest Valorização	05 - Mar. - 1991	12 - Jul. - 1993 a 29 - Dez. - 1997	233
CP	Capital Portugal	30 - Out. - 1989	12 - Jul. - 1993 a 29 - Dez. - 1997	233
IBVL	Fundo IBVL	31 - Mar. - 1991	12 - Jul. - 1993 a 29 - Dez. - 1997	233
NFC	Novofundo Capital	31 - Dez. - 1990	12 - Jul. - 1993 a 29 - Dez. - 1997	233
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	12 - Jul. - 1993 a 29 - Dez. - 1997	233
PC	Primus Capital	20 - Fev. - 1989	12 - Jul. - 1993 a 29 - Dez. - 1997	233
TA	Totta Acções	22 - Jun. - 1987	12 - Jul. - 1993 a 29 - Dez. - 1997	233
UC	Unicapital	01 - Out. - 1991	12 - Jul. - 1993 a 29 - Dez. - 1997	233
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	12 - Jul. - 1993 a 29 - Dez. - 1997	233

Table A.3 - Security Investment Funds in Panel AEstim3

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	03 - Jan. - 1994 a 29 - Dez. - 1997	208
BPIA	BPI Acções	11 - Jun. - 1991	03 - Jan. - 1994 a 29 - Dez. - 1997	208
CXV	Caixagest Valorização	05 - Mar. - 1991	03 - Jan. - 1994 a 29 - Dez. - 1997	208
CP	Capital Portugal	30 - Out. - 1989	03 - Jan. - 1994 a 29 - Dez. - 1997	208
IBVL	Fundo IBVL	31 - Mar. - 1991	03 - Jan. - 1994 a 29 - Dez. - 1997	208
NFC	Novofundo Capital	31 - Dez. - 1990	03 - Jan. - 1994 a 29 - Dez. - 1997	208
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	03 - Jan. - 1994 a 29 - Dez. - 1997	208
PC	Primus Capital	20 - Fev. - 1989	03 - Jan. - 1994 a 29 - Dez. - 1997	208
TA	Totta Acções	22 - Jun. - 1987	03 - Jan. - 1994 a 29 - Dez. - 1997	208
UC	Unicapital	01 - Out. - 1991	03 - Jan. - 1994 a 29 - Dez. - 1997	208
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	03 - Jan. - 1994 a 29 - Dez. - 1997	208
EC	Eurocapital	03 - Jan. - 1994	03 - Jan. - 1994 a 29 - Dez. - 1997	208

Table A.4 - Security Investment Funds in Panel AEstim4

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	07 - Fev. - 1994 a 29 - Dez. - 1997	203
BPIA	BPI Acções	11 - Jun. - 1991	07 - Fev. - 1994 a 29 - Dez. - 1997	203
CXV	Caixagest Valorização	05 - Mar. - 1991	07 - Fev. - 1994 a 29 - Dez. - 1997	203
CP	Capital Portugal	30 - Out. - 1989	07 - Fev. - 1994 a 29 - Dez. - 1997	203
IBVL	Fundo IBVL	31 - Mar. - 1991	07 - Fev. - 1994 a 29 - Dez. - 1997	203
NFC	Novofundo Capital	31 - Dez. - 1990	07 - Fev. - 1994 a 29 - Dez. - 1997	203
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	07 - Fev. - 1994 a 29 - Dez. - 1997	203
PC	Primus Capital	20 - Fev. - 1989	07 - Fev. - 1994 a 29 - Dez. - 1997	203
TA	Totta Acções	22 - Jun. - 1987	07 - Fev. - 1994 a 29 - Dez. - 1997	203
UC	Unicapital	01 - Out. - 1991	07 - Fev. - 1994 a 29 - Dez. - 1997	203
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	07 - Fev. - 1994 a 29 - Dez. - 1997	203
EC	Eurocapital	03 - Jan. - 1994	07 - Fev. - 1994 a 29 - Dez. - 1997	203
MGA	MG Acções	01 - Fev. - 1994	07 - Fev. - 1994 a 29 - Dez. - 1997	203

Table A.5 - Security Investment Funds in Panel AEstim5

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	11 - Jul. - 1994 a 29 - Dez. - 1997	181
BPIA	BPI Acções	11 - Jun. - 1991	11 - Jul. - 1994 a 29 - Dez. - 1997	181
CXV	Caixagest Valorização	05 - Mar. - 1991	11 - Jul. - 1994 a 29 - Dez. - 1997	181
CP	Capital Portugal	30 - Out. - 1989	11 - Jul. - 1994 a 29 - Dez. - 1997	181
IBVL	Fundo IBVL	31 - Mar. - 1991	11 - Jul. - 1994 a 29 - Dez. - 1997	181
NFC	Novofundo Capital	31 - Dez. - 1990	11 - Jul. - 1994 a 29 - Dez. - 1997	181
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	11 - Jul. - 1994 a 29 - Dez. - 1997	181
PC	Primus Capital	20 - Fev. - 1989	11 - Jul. - 1994 a 29 - Dez. - 1997	181
TA	Totta Acções	22 - Jun. - 1987	11 - Jul. - 1994 a 29 - Dez. - 1997	181
UC	Unicapital	01 - Out. - 1991	11 - Jul. - 1994 a 29 - Dez. - 1997	181
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	11 - Jul. - 1994 a 29 - Dez. - 1997	181
EC	Eurocapital	03 - Jan. - 1994	11 - Jul. - 1994 a 29 - Dez. - 1997	181
MGA	MG Acções	01 - Fev. - 1994	11 - Jul. - 1994 a 29 - Dez. - 1997	181
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	11 - Jul. - 1994 a 29 - Dez. - 1997	181

Table A.6 - Security Investment Funds in Panel AEstim6

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	12 - Jun. - 1995 a 29 - Dez. - 1997	133
BPIA	BPI Acções	11 - Jun. - 1991	12 - Jun. - 1995 a 29 - Dez. - 1997	133
CXV	Caixagest Valorização	05 - Mar. - 1991	12 - Jun. - 1995 a 29 - Dez. - 1997	133
CP	Capital Portugal	30 - Out. - 1989	12 - Jun. - 1995 a 29 - Dez. - 1997	133
IBVL	Fundo IBVL	31 - Mar. - 1991	12 - Jun. - 1995 a 29 - Dez. - 1997	133
NFC	Novofundo Capital	31 - Dez. - 1990	12 - Jun. - 1995 a 29 - Dez. - 1997	133
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	12 - Jun. - 1995 a 29 - Dez. - 1997	133
PC	Primus Capital	20 - Fev. - 1989	12 - Jun. - 1995 a 29 - Dez. - 1997	133
TA	Totta Acções	22 - Jun. - 1987	12 - Jun. - 1995 a 29 - Dez. - 1997	133
UC	Unicapital	01 - Out. - 1991	12 - Jun. - 1995 a 29 - Dez. - 1997	133
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	12 - Jun. - 1995 a 29 - Dez. - 1997	133
EC	Eurocapital	03 - Jan. - 1994	12 - Jun. - 1995 a 29 - Dez. - 1997	133
MGA	MG Acções	01 - Fev. - 1994	12 - Jun. - 1995 a 29 - Dez. - 1997	133
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	12 - Jun. - 1995 a 29 - Dez. - 1997	133
MC	M Capital	07 - Jun. - 1995	12 - Jun. - 1995 a 29 - Dez. - 1997	133

Table A.7 - Security Investment Funds in Panel AEstim7

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	17 - Jul. - 1995 a 29 - Dez. - 1997	128
BPIA	BPI Acções	11 - Jun. - 1991	17 - Jul. - 1995 a 29 - Dez. - 1997	128
CXV	Caixagest Valorização	05 - Mar. - 1991	17 - Jul. - 1995 a 29 - Dez. - 1997	128
CP	Capital Portugal	30 - Out. - 1989	17 - Jul. - 1995 a 29 - Dez. - 1997	128
IBVL	Fundo IBVL	31 - Mar. - 1991	17 - Jul. - 1995 a 29 - Dez. - 1997	128
NFC	Novofundo Capital	31 - Dez. - 1990	17 - Jul. - 1995 a 29 - Dez. - 1997	128
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	17 - Jul. - 1995 a 29 - Dez. - 1997	128
PC	Primus Capital	20 - Fev. - 1989	17 - Jul. - 1995 a 29 - Dez. - 1997	128
TA	Totta Acções	22 - Jun. - 1987	17 - Jul. - 1995 a 29 - Dez. - 1997	128
UC	Unicapital	01 - Out. - 1991	17 - Jul. - 1995 a 29 - Dez. - 1997	128
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	17 - Jul. - 1995 a 29 - Dez. - 1997	128
EC	Eurocapital	03 - Jan. - 1994	17 - Jul. - 1995 a 29 - Dez. - 1997	128
MGA	MG Acções	01 - Fev. - 1994	17 - Jul. - 1995 a 29 - Dez. - 1997	128
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	17 - Jul. - 1995 a 29 - Dez. - 1997	128
MC	M Capital	07 - Jun. - 1995	17 - Jul. - 1995 a 29 - Dez. - 1997	128
BCPA	BCP Acções	20 - Jun. - 1995	17 - Jul. - 1995 a 29 - Dez. - 1997	128

Table A.8 - Security Investment Funds in Panel AEstim8

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	24 - Jun. - 1996 a 29 - Dez. - 1997	79
BPIA	BPI Acções	11 - Jun. - 1991	24 - Jun. - 1996 a 29 - Dez. - 1997	79
CXV	Caixagest Valorização	05 - Mar. - 1991	24 - Jun. - 1996 a 29 - Dez. - 1997	79
CP	Capital Portugal	30 - Out. - 1989	24 - Jun. - 1996 a 29 - Dez. - 1997	79
IBVL	Fundo IBVL	31 - Mar. - 1991	24 - Jun. - 1996 a 29 - Dez. - 1997	79
NFC	Novofundo Capital	31 - Dez. - 1990	24 - Jun. - 1996 a 29 - Dez. - 1997	79
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	24 - Jun. - 1996 a 29 - Dez. - 1997	79
PC	Primus Capital	20 - Fev. - 1989	24 - Jun. - 1996 a 29 - Dez. - 1997	79
TA	Totta Acções	22 - Jun. - 1987	24 - Jun. - 1996 a 29 - Dez. - 1997	79
UC	Unicapital	01 - Out. - 1991	24 - Jun. - 1996 a 29 - Dez. - 1997	79
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	24 - Jun. - 1996 a 29 - Dez. - 1997	79
EC	Eurocapital	03 - Jan. - 1994	24 - Jun. - 1996 a 29 - Dez. - 1997	79
MGA	MG Acções	01 - Fev. - 1994	24 - Jun. - 1996 a 29 - Dez. - 1997	79
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	24 - Jun. - 1996 a 29 - Dez. - 1997	79
MC	M Capital	07 - Jun. - 1995	24 - Jun. - 1996 a 29 - Dez. - 1997	79
BCPA	BCP Acções	20 - Jun. - 1995	24 - Jun. - 1996 a 29 - Dez. - 1997	79
CXAP	Caixagest Acç. Portugal	20 - Jun. - 1996	24 - Jun. - 1996 a 29 - Dez. - 1997	79

Table A.9 - Security Investment Funds in Panel AEstim9

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	21 - Abr. - 1997 a 29 - Dez. - 1997	36
BPIA	BPI Acções	11 - Jun. - 1991	21 - Abr. - 1997 a 29 - Dez. - 1997	36
CXV	Caixagest Valorização	05 - Mar. - 1991	21 - Abr. - 1997 a 29 - Dez. - 1997	36
CP	Capital Portugal	30 - Out. - 1989	21 - Abr. - 1997 a 29 - Dez. - 1997	36
IBVL	F. IBVL	31 - Mar. - 1991	21 - Abr. - 1997 a 29 - Dez. - 1997	36
NFC	Novofundo Capital	31 - Dez. - 1990	21 - Abr. - 1997 a 29 - Dez. - 1997	36
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	21 - Abr. - 1997 a 29 - Dez. - 1997	36
PC	Primus Capital	20 - Fev. - 1989	21 - Abr. - 1997 a 29 - Dez. - 1997	36
TA	Totta Acções	22 - Jun. - 1987	21 - Abr. - 1997 a 29 - Dez. - 1997	36
UC	Unicapital	01 - Out. - 1991	21 - Abr. - 1997 a 29 - Dez. - 1997	36
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	21 - Abr. - 1997 a 29 - Dez. - 1997	36
EC	Eurocapital	03 - Jan. - 1994	21 - Abr. - 1997 a 29 - Dez. - 1997	36
MGA	MG Acções	01 - Fev. - 1994	21 - Abr. - 1997 a 29 - Dez. - 1997	36
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	21 - Abr. - 1997 a 29 - Dez. - 1997	36
MC	M Capital	07 - Jun. - 1995	21 - Abr. - 1997 a 29 - Dez. - 1997	36
BCPA	BCP Acções	20 - Jun. - 1995	21 - Abr. - 1997 a 29 - Dez. - 1997	36
CXAP	Caixagest Acç. Portugal	20 - Jun. - 1996	21 - Abr. - 1997 a 29 - Dez. - 1997	36
FC	Finicapital	15 - Abr. - 1997	21 - Abr. - 1997 a 29 - Dez. - 1997	36

Table A.10 - Security Investment Funds in Panel AEstim10

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	02 - Jun. - 1997 a 29 - Dez. - 1997	30
BPIA	BPI Acções	11 - Jun. - 1991	02 - Jun. - 1997 a 29 - Dez. - 1997	30
CXV	Caixagest Valorização	05 - Mar. - 1991	02 - Jun. - 1997 a 29 - Dez. - 1997	30
CP	Capital Portugal	30 - Out. - 1989	02 - Jun. - 1997 a 29 - Dez. - 1997	30
IBVL	F. IBVL	31 - Mar. - 1991	02 - Jun. - 1997 a 29 - Dez. - 1997	30
NFC	Novofundo Capital	31 - Dez. - 1990	02 - Jun. - 1997 a 29 - Dez. - 1997	30
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	02 - Jun. - 1997 a 29 - Dez. - 1997	30
PC	Primus Capital	20 - Fev. - 1989	02 - Jun. - 1997 a 29 - Dez. - 1997	30
TA	Totta Acções	22 - Jun. - 1987	02 - Jun. - 1997 a 29 - Dez. - 1997	30
UC	Unicapital	01 - Out. - 1991	02 - Jun. - 1997 a 29 - Dez. - 1997	30
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	02 - Jun. - 1997 a 29 - Dez. - 1997	30
EC	Eurocapital	03 - Jan. - 1994	02 - Jun. - 1997 a 29 - Dez. - 1997	30
MGA	MG Acções	01 - Fev. - 1994	02 - Jun. - 1997 a 29 - Dez. - 1997	30
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	02 - Jun. - 1997 a 29 - Dez. - 1997	30
MC	M Capital	07 - Jun. - 1995	02 - Jun. - 1997 a 29 - Dez. - 1997	30
BCPA	BCP Acções	20 - Jun. - 1995	02 - Jun. - 1997 a 29 - Dez. - 1997	30
CXAP	Caixagest Acç. Portugal	20 - Jun. - 1996	02 - Jun. - 1997 a 29 - Dez. - 1997	30
FC	Finicapital	15 - Abr. - 1997	02 - Jun. - 1997 a 29 - Dez. - 1997	30
RV	Raíz Valorização	28 - Mai. - 1997	02 - Jun. - 1997 a 29 - Dez. - 1997	30

Table A.11 - Security Investment Funds in Panel AEstim11

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BNUA	BNU Acções	24 - Jun. - 1991	15 - Set. - 1997 a 29 - Dez. - 1997	15
BPIA	BPI Acções	11 - Jun. - 1991	15 - Set. - 1997 a 29 - Dez. - 1997	15
CXV	Caixagest Valorização	05 - Mar. - 1991	15 - Set. - 1997 a 29 - Dez. - 1997	15
CP	Capital Portugal	30 - Out. - 1989	15 - Set. - 1997 a 29 - Dez. - 1997	15
IBVL	F. IBVL	31 - Mar. - 1991	15 - Set. - 1997 a 29 - Dez. - 1997	15
NFC	Novofundo Capital	31 - Dez. - 1990	15 - Set. - 1997 a 29 - Dez. - 1997	15
FIPOR	Fipor Poupança Investimento	26 - Fev. - 1987	15 - Set. - 1997 a 29 - Dez. - 1997	15
PC	Primus Capital	20 - Fev. - 1989	15 - Set. - 1997 a 29 - Dez. - 1997	15
TA	Totta Acções	20 - Jun. - 1987	15 - Set. - 1997 a 29 - Dez. - 1997	15
UC	Unicapital	01 - Out. - 1991	15 - Set. - 1997 a 29 - Dez. - 1997	15
BCIAP	BCI Acções Portugal	12 - Jul. - 1993	15 - Set. - 1997 a 29 - Dez. - 1997	15
EC	Eurocapital	03 - Jan. - 1994	15 - Set. - 1997 a 29 - Dez. - 1997	15
MGA	MG Acções	01 - Fev. - 1994	15 - Set. - 1997 a 29 - Dez. - 1997	15
BPAP	Barclays Premier Acç. Port.	07 - Jul. - 1994	15 - Set. - 1997 a 29 - Dez. - 1997	15
MC	M Capital	07 - Jun. - 1995	15 - Set. - 1997 a 29 - Dez. - 1997	15
BCPA	BCP Acções	20 - Jun. - 1995	15 - Set. - 1997 a 29 - Dez. - 1997	15
CXAP	Caixagest Acç. Portugal	20 - Jun. - 1996	15 - Set. - 1997 a 29 - Dez. - 1997	15
FC	Finicapital	15 - Abr. - 1997	15 - Set. - 1997 a 29 - Dez. - 1997	15
RV	Raíz Valorização	28 - Mai. - 1997	15 - Set. - 1997 a 29 - Dez. - 1997	15
ESPA	Espírito Santo Portugal Acç.	15 - Set. - 1997	15 - Set. - 1997 a 29 - Dez. - 1997	15

Table A.12 - Security Investment Funds in Panel MEstim1

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BCPO	BCP Obrigações	03 - Jul. - 1989	04 - Jan. - 1993 a 29 - Dez. - 1997	260
NFO	Novofundo Obrigações	31 - Dez. - 1990	04 - Jan. - 1993 a 29 - Dez. - 1997	260
MFIP	Mealheiro FIPOR	19 - Mar. - 1991	04 - Jan. - 1993 a 29 - Dez. - 1997	260
RFIP	Rendimento FIPOR	19 - Jun. - 1989	04 - Jan. - 1993 a 29 - Dez. - 1997	260
INVES	Invest	21 - Mai. - 1986	04 - Jan. - 1993 a 29 - Dez. - 1997	260
PA	Postal Acções	23 - Jun. - 1987	04 - Jan. - 1993 a 29 - Dez. - 1997	260
AFIP	Aforro FIPOR	01 - Ago. - 1990	04 - Jan. - 1993 a 29 - Dez. - 1997	260

Table A.13 - Security Investment Funds in Panel MEstim2

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BCPO	BCP Obrigações	03 - Jul. - 1989	15 - Set. - 1994 a 29 - Dez. - 1997	172
NFO	Novofundo Obrigações	31 - Dez. - 1990	15 - Set. - 1994 a 29 - Dez. - 1997	172
MFIP	Mealheiro FIPOR	19 - Mar. - 1991	15 - Set. - 1994 a 29 - Dez. - 1997	172
RFIP	Rendimento FIPOR	19 - Jun. - 1989	15 - Set. - 1994 a 29 - Dez. - 1997	172
INVES	Invest	21 - Mai. - 1986	15 - Set. - 1994 a 29 - Dez. - 1997	172
PA	Postal Acções	23 - Jun. - 1987	15 - Set. - 1994 a 29 - Dez. - 1997	172
AFIP	Aforro FIPOR	01 - Ago. - 1990	15 - Set. - 1994 a 29 - Dez. - 1997	172
SC	Sotto Capital	01 - Set. - 1994	15 - Set. - 1994 a 29 - Dez. - 1997	172

Table A.14 - Security Investment Funds in Panel MEstim3

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BCPO	BCP Obrigações	03 - Jul. - 1989	16 - Out. - 1995 a 29 - Dez. - 1997	146
NFO	Novofundo Obrigações	31 - Dez. - 1990	16 - Out. - 1995 a 29 - Dez. - 1997	146
MFIP	Mealheiro FIPOR	19 - Mar. - 1991	16 - Out. - 1995 a 29 - Dez. - 1997	146
RFIP	Rendimento FIPOR	19 - Jun. - 1989	16 - Out. - 1995 a 29 - Dez. - 1997	146
INVES	Invest	21 - Mai. - 1986	16 - Out. - 1995 a 29 - Dez. - 1997	146
PA	Postal Acções	23 - Jun. - 1987	16 - Out. - 1995 a 29 - Dez. - 1997	146
AFIP	Aforro FIPOR	01 - Ago. - 1990	16 - Out. - 1995 a 29 - Dez. - 1997	146
SC	Sotto Capital	01 - Set. - 1994	16 - Out. - 1995 a 29 - Dez. - 1997	146
DBI	DB Investimento	13 - Out. - 1995	16 - Out. - 1995 a 29 - Dez. - 1997	146

Table A.15 - Security Investment Funds in Panel MEstim4

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BCPO	BCP Obrigações	03 - Jul. - 1989	11 - Dez. - 1995 a 29 - Dez. - 1997	107
NFO	Novofundo Obrigações	31 - Dez. - 1990	11 - Dez. - 1995 a 29 - Dez. - 1997	107
MFIP	Mealheiro FIPOR	19 - Mar. - 1991	11 - Dez. - 1995 a 29 - Dez. - 1997	107
RFIP	Rendimento FIPOR	19 - Jun. - 1989	11 - Dez. - 1995 a 29 - Dez. - 1997	107
INVES	Invest	21 - Mai. - 1986	11 - Dez. - 1995 a 29 - Dez. - 1997	107
PA	Postal Acções	23 - Jun. - 1987	11 - Dez. - 1995 a 29 - Dez. - 1997	107
AFIP	Aforro FIPOR	01 - Ago. - 1990	11 - Dez. - 1995 a 29 - Dez. - 1997	107
SC	Sotto Capital	01 - Set. - 1994	11 - Dez. - 1995 a 29 - Dez. - 1997	107
DBI	DB Investimento	13 - Out. - 1995	11 - Dez. - 1995 a 29 - Dez. - 1997	107
PFIP	Poupança FIPOR	11 - Dez. 1995	11 - Dez. - 1995 a 29 - Dez. - 1997	107

Table A.16 - Security Investment Funds in Panel MEstim5

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BCPO	BCP Obrigações	03 - Jul. - 1989	03 - Fev. - 1997 a 29 - Dez. - 1997	47
NFO	Novofundo Obrigações	31 - Dez. - 1990	03 - Fev. - 1997 a 29 - Dez. - 1997	47
MFIP	Mealheiro FIPOR	19 - Mar. - 1991	03 - Fev. - 1997 a 29 - Dez. - 1997	47
RFIP	Rendimento FIPOR	19 - Jun. - 1989	03 - Fev. - 1997 a 29 - Dez. - 1997	47
INVES	Invest	21 - Mai. - 1986	03 - Fev. - 1997 a 29 - Dez. - 1997	47
PA	Postal Acções	23 - Jun. - 1987	03 - Fev. - 1997 a 29 - Dez. - 1997	47
AFIP	Aforro FIPOR	01 - Ago. - 1990	03 - Fev. - 1997 a 29 - Dez. - 1997	47
SC	Sotto Capital	01 - Set. - 1994	03 - Fev. - 1997 a 29 - Dez. - 1997	47
DBI	DB Investimento	13 - Out. - 1995	03 - Fev. - 1997 a 29 - Dez. - 1997	47
PFIP	Poupança FIPOR	11 - Dez. 1995	03 - Fev. - 1997 a 29 - Dez. - 1997	47
BPIG	BPI Global	31 - Jan. - 1997	03 - Fev. - 1997 a 29 - Dez. - 1997	47

Table A.17 - Security Investment Funds in Panel MEstim6

Code	Funds	Activity Date	Sample Time Period	Nº. Obs.
BCPO	BCP Obrigações	03 - Jul. - 1989	15 - Set. - 1997 a 29 - Dez. - 1997	15
NFO	Novofundo Obrigações	31 - Dez. - 1990	15 - Set. - 1997 a 29 - Dez. - 1997	15
MFIP	Mealheiro FIPOR	19 - Mar. - 1991	15 - Set. - 1997 a 29 - Dez. - 1997	15
RFIP	Rendimento FIPOR	19 - Jun. - 1989	15 - Set. - 1997 a 29 - Dez. - 1997	15
INVES	Invest	21 - Mai. - 1986	15 - Set. - 1997 a 29 - Dez. - 1997	15
PA	Postal Acções	23 - Jun. - 1987	15 - Set. - 1997 a 29 - Dez. - 1997	15
AFIP	Aforro FIPOR	01 - Ago. - 1990	15 - Set. - 1997 a 29 - Dez. - 1997	15
SC	Sotto Capital	01 - Set. - 1994	15 - Set. - 1997 a 29 - Dez. - 1997	15
DBI	DB Investimento	13 - Out. - 1995	15 - Set. - 1997 a 29 - Dez. - 1997	15
PFIP	Poupança FIPOR	11 - Dez. 1995	15 - Set. - 1997 a 29 - Dez. - 1997	15

VOLUME III

3° QUADRIMESTRE DE 2002