

## DEFINIÇÃO, IDENTIFICAÇÃO E ESTIMAÇÃO DO MODELO ARCH E COMPARAÇÃO COM OUTROS MODELOS DE VOLATILIDADE (1)

João C. H. C. Nicolau (2)

### 1 — Introdução

Até ao início da década de 80 a maior parte da econometria financeira das sucessões cronológicas centrava a sua análise na média condicionada (3) do processo em estudo, propondo, normalmente, especificações lineares (ARMA ou ARMAX). Estes modelos assumem que a volatilidade é constante ao longo do tempo. Esta hipótese, contudo, raramente se verifica na prática. Para estimar a volatilidade é necessário considerar a variância condicionada. Assim, presentemente, considera-se que a via de análise mais correcta para analisar o comportamento estatístico de uma sucessão cronológica financeira consiste em modelizar não só a média (4) como também a variância condicionada (5).

Neste contexto, foi decisivo o trabalho de Engle (1982). Este autor mostrou que era possível considerar um modelo paramétrico no qual a variância condicionada pela álgebra pode ser explicitamente escrita e modelizada. Engle designou o modelo por *autoregressive conditional heteroscedasticity* (ARCH). Assume-se que os períodos de estabilidade (e de instabilidade) se sucedem entre si. Isto é, um período de baixa (alta) volatilidade é precedido por períodos de baixa (alta) volatilidade.

Procura-se neste trabalho caracterizar os modelos mais importantes na modelização da volatilidade com especial ênfase no modelo ARCH. Concretamente, os objectivos deste trabalho são: i) apresentar os principais resultados já estabelecidos pela teoria dos modelos ARCH; ii) apresentar desenvolvimentos recentes na área dos modelos ARCH (modelo GARCH com alterações estocásticas de regime [Gray(1996)]; iii) apresentar outros modelos concorrentes do modelo ARCH que recentemente têm merecido muita atenção na literatura (modelo de volatilidade estocástica); iv) mostrar afinidades originais entre certos modelos a tempo contínuo e os modelos ARCH.

O trabalho apresenta-se da seguinte forma. No ponto 2 discutem-se as principais regularidades empíricas das sucessões cronológicas financeiras; no

(1) Os primeiros cinco pontos baseiam-se na tese de mestrado — Nicolau (1994).

(2) ISEG-UTL.

(3) No presente texto a média condicionada de  $y_t$  reporta-se ao valor esperado  $E[y_t | F_{t-1}]$ , onde  $F_{t-1}$  é o conjunto de informação disponível até ao momento  $t-1$ , isto é,  $F_{t-1} = \{y_{t-1}, y_{t-2}, \dots; x_{t-1}, x_{t-2}, \dots\}$  sendo  $x_{t-i}$  o conjunto das variáveis explicativas do modelo  $Y$  no momento  $t-i$ . Da mesma forma, a variância condicionada de  $y_t$  corresponde a  $\text{Var}[y_t | F_{t-1}]$ .

(4) Certos modelos não lineares, baseados apenas na modelização da média, como, por exemplo, o modelo bilinear ou o limiar autoregressivo podem, eventualmente, modelizar com êxito sucessões financeiras.

(5) E por que não outros momentos superiores?

ponto 3 define-se o modelo ARCH e as suas principais derivações e mostra-se em que medida cada um dos modelos pode estimar as principais regularidades empíricas discutidas no ponto 2; nos pontos 4 e 5 discute-se o problema da identificação e da estimação do modelo ARCH; no ponto 6 introduz-se o modelo de volatilidade estocástica e mostra-se que o principal problema deste modelo é o da sua estimação; no ponto 7 estabelecem-se algumas comparações originais entre certas equações diferenciais estocásticas e os modelos ARCH. O trabalho termina com uma conclusão.

## 2 — Regularidades empíricas das sucessões cronológicas financeiras

São geralmente aceites as seguintes regularidades empíricas das sucessões cronológicas financeiras [para maior desenvolvimento v., por exemplo, Taylor (1986), Engle, Bollerslev e Nelson (1993) e Bollerslev, Chou e Kroner (1992)].

### Distribuições leptocúrticas

As distribuições são em geral leptocúrticas e, nalguns casos, as variâncias são infinitas. Nas distribuições leptocúrticas as caudas das distribuições são relativamente altas, evidenciando que existe uma massa de probabilidade alta para as observações que se encontram relativamente afastadas da média. Variâncias infinitas podem ocorrer em distribuições onde as caudas das distribuições não converjam rapidamente para o eixo das abcissas.

### Volatilidade

«[...] large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes [...]» [Mandelbrot, citado em Engle, Bollerslev e Nelson (1993).] Esta característica está presente na maioria das sucessões cronológicas financeiras, como podemos observar, por exemplo, para as variações diárias do marco alemão (fig. 1). O padrão de volatilidade (fig. 2) resulta claro a partir da fig. 1.

FIGURA 1

Variações diárias do marco alemão em escudos no período de Janeiro de 1987 a Dezembro de 1996

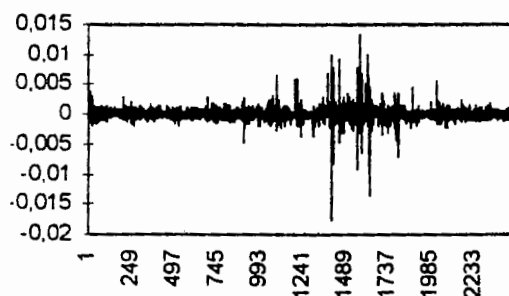
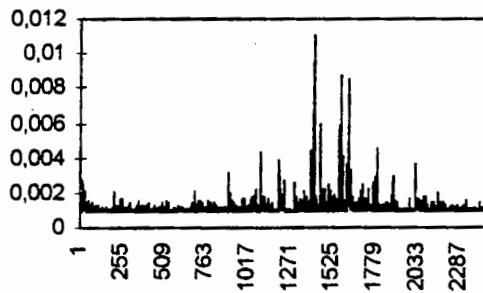


FIGURA 2

Volatilidade estimada <sup>(6)</sup> das variações diárias**Efeito assimétrico**

Em certas aplicações observa-se existir uma correlação negativa entre as variações do preço das acções e a sua volatilidade [Black (1976)]. Mais concretamente, a volatilidade tende a aumentar mais em resposta a «más notícias» (rendimentos abaixo do esperado) do que a «boas notícias» (rendimento acima do esperado).

**Períodos de não transacção**

Regista-se frequentemente nos mercados financeiros os preços reflectirem a informação acumulada durante os períodos de não transacção. Por exemplo, o preço das acções na segunda-feira deve apresentar um comportamento diferente dos outros dias da semana, pois o mercado na segunda-feira reflecte a informação acumulada durante o fim-de-semana.

**Frequência das observações**

Constata-se frequentemente que o aumento da frequência das observações aumenta significativamente a volatilidade [v., por exemplo, Nijman e Drost (1993)].

**Acontecimentos previsíveis**

Cornell (1978) mostra que quando são anunciados os rendimentos das empresas a volatilidade associada à cotação dessas empresas aumenta. Harvey e Huang (1991) mostram também que a volatilidade das taxas de câmbio aumenta quando os bancos centrais incrementam as suas transacções ou quando existe alguma notícia macroeconómica relevante.

<sup>(6)</sup> A volatilidade foi estimada a partir do modelo ARMA(1,1)-GARCH(1,1) com distribuição condicionada normal.

**Co-movimentos na volatilidade**

Vários autores [Engle, Victor e Rothschild (1990), Schwert (1989)] mostram que há efeitos comuns de volatilidade dentro de um mercado, assim como de um mercado para outro mercado.

O sucesso dos modelos ARCH resulta, em certa medida, da sua capacidade para modelizar as regularidades empíricas discutidas neste ponto.

**3 — Definição do modelo**

Seja  $\{u_t, t=1,2,\dots\}$  um processo estocástico a tempo discreto, definido da seguinte forma:  $u_t = \varepsilon_t \sigma_t$ , onde  $\varepsilon_t$  é i. i. d.,  $E[\varepsilon_t] = 0$ ,  $\text{Var}[\varepsilon_t] = 1$ , e  $\sigma_t$  é uma função mensurável com respeito a  $\{u_{t-1}, u_{t-2}, \dots\}$  estritamente positiva (isto é,  $\sigma_t$  depende dos acontecimentos gerados por  $u_\tau$  com  $\tau \leq t-1$ ). Assim:

$$E[u_t | F_{t-1}] = 0 \text{ e } \text{Var}[u_t | F_{t-1}] = \sigma_t^2(u_{t-1}, u_{t-2}, \dots) > 0$$

A variável  $u_t$  (que é uma diferença de martingala com respeito a  $F_{t-1}$ ) pode ser interpretada como a variável residual de um modelo ARMAX (7).

A função  $\sigma_t^2$ , quando convenientemente parametrizada, pode contemplar os principais efeitos descritos no ponto anterior. Vejamos alguns modelos que podem ser estabelecidos para  $\sigma_t^2$ .

**Modelo ARCH**

No modelo ARCH(q) [Engle (1982)] a variância condicionada corresponde  $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 > 0$ . Por construção, facilmente se verifica que «[...] large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes [...]» Além disso, se  $\varepsilon_t$  tem distribuição  $N(0,1)$  então  $u_t$  tem distribuição marginal leptocúrtica. Com efeito:

$$\frac{E[u_t^4]}{E[u_t^2]^2} = \frac{E[\varepsilon_t^4] E[\sigma_t^4]}{E[\varepsilon_t^2]^2 E[\sigma_t^2]^2} \geq \frac{E[\varepsilon_t^4] E[\sigma_t^2]^2}{E[\varepsilon_t^2]^2 E[\sigma_t^2]^2} = 3$$

Podemos ainda constatar que, definindo  $v_t = u_t^2 - \sigma_t^2$ , o modelo pode ser rescrito na forma:

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 + v_t$$

e, como  $v_t$  é uma diferença de martingala, o processo  $u_t^2$  admite uma representação AR(q). Esta observação permite, em aplicações empíricas, detectar e identificar a ordem  $q$  através da FAC e FACP do processo  $u_t^2$  desde que exista  $E[u_t^4]$  (esta condição é restritiva; por exemplo, para um ARCH(1) pode-se pro-

(7) Por exemplo, com apenas um regressor  $A(L)y_t = B(L)x_t + u_t$  onde  $\Phi(L)u_t = \Theta(L)\varepsilon_t$ , com as notações habituais.

var que  $E[u_t^4]$  existe se  $3\alpha^2 < 1$ , isto é se  $\alpha_1 < 0.58$  — assim, é necessário proceder com prudência à análise das funções de autocorrelação).

É fácil contemplar na especificação da variância condicionada efeitos decorrentes de acontecimentos totalmente previsíveis. Por exemplo, num modelo de dados diários, para atender ao efeito dos dias da semana sobre a volatilidade, podem-se acrescentar quatro variáveis artificiais [ARCH(1)]:

$$\sigma_t^2 = \alpha_0 + \delta_1 \text{Ter}_t + \delta_2 \text{Qua}_t + \delta_3 \text{Qui}_t + \delta_4 \text{Sex}_t + \alpha_1 u_{t-1}^2$$

onde  $\alpha_0$  representa o efeito da segunda-feira,  $\alpha_0 + \delta_1$  o efeito de terça-feira, etc.

É curioso verificar que o modelo auto-regressivo de coeficientes aleatórios  $u_t = \sum_{j=1}^q \phi_j u_{t-j} + \varepsilon_t$ , sendo  $\phi_j$  uma sucessão de variáveis aleatórias de média e variâncias finitas é, (sob certas condições) observacionalmente equivalente ao modelo ARCH (i. e., conduz à mesma função de verosimilhança).

#### Modelo GARCH

No modelo GARCH( $p, q$ ) [Bollerslev (1986)] a variância condicionada corresponde a  $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 + \delta_1 \sigma_{t-1}^2 + \dots + \delta_p \sigma_{t-p}^2 > 0$ . O GARCH é um ARCH de ordem infinita com restrições sobre os parâmetros. Por exemplo no GARCH(1,1), a variância condicionada é:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2$$

e, é fácil verificar, por substituição sucessiva, que esta expressão é equivalente a:

$$\sigma_t^2 = \frac{\alpha_0}{1 - \delta_1} + \alpha_1 u_{t-1}^2 + \alpha_1 \delta_1 u_{t-2}^2 + \alpha_1 \delta_1^2 u_{t-3}^2 + \dots$$

Quando o processo ARCH tende a apresentar desfasamentos longos o modelo GARCH é preferível, pois é parcimonioso e não perde a informação mais atrasada.

Pode-se provar que, neste caso, o processo  $u_t^2$  admite uma representação ARMA(max{ $p, q$ },  $p$ ).

#### Modelo IGARCH, persistência e co-persistência

Em aplicações empíricas com o modelo GARCH(1,1) constata-se, por vezes, que  $\hat{\alpha}_1 + \hat{\delta}_1 \approx 1$ . No caso  $\alpha_1 + \delta_1 = 1$ , o modelo é designado por GARCH integrado ou IGARCH. A análise da estacionariedade do modelo GARCH e IGARCH revelou resultados curiosos. Por exemplo, para o modelo GARCH(1,1)  $E[\sigma_t^2]$  é finito e  $\{u_t\}$  é estacionário de segunda ordem se  $\alpha_1 + \delta_1 < 1$ . Esta é uma condição suficiente mas não necessária para a estacionariedade estrita. Nelson (1990) provou para o caso GARCH(1,1) que  $\{u_t\}$  e  $\{\sigma_t^2\}$  são estritamente estacionários se e só se  $E[\log(\alpha_1 \varepsilon_t^2 + \delta_1)] < 0$ . Pode-se provar que a condição  $\alpha_1 + \delta_1 < 1$  é mais exigente do que  $E[\log(\alpha_1 \varepsilon_t^2 + \delta_1)] < 0$ , isto é

$\alpha_1 + \delta_1 < 1 \Rightarrow E[\log(\alpha_1 \varepsilon_t^2 + \delta_1)] < 0$  (não se verificando a implicação contrária). Assim, o modelo IGARCH(1,1) viola a condição  $\alpha_1 + \delta_1 < 1$  e  $E[\sigma_t^2]$  não é finito, no entanto, é estritamente estacionário e  $\sigma_t^2 < \infty$  quase certamente. Apesar de o IGARCH não ser estacionário de segunda ordem Lumsdaine (1991) mostra que os testes assintóticos usuais são válidos no contexto do modelo IGARCH, embora Hong (1988) sugira que a amostra deva ser razoavelmente grande.

Ligado a esta questão encontra-se o problema da persistência dos choques na volatilidade. A definição de persistência depende do conceito de convergência que se considerar. Se considerarmos a convergência em  $L^1$ , um processo é persistente na variância se  $E[\sigma_{t+h}^2 | \sigma_t^2] \rightarrow \infty$  ( $h \rightarrow \infty$ ); se considerarmos a convergência quase certa (ou com probabilidade 1) um processo é persistente na variância se  $\sigma_t^2 \rightarrow \infty$  quase certamente. Estas duas definições nem sempre se verificam simultaneamente e, claramente, não se verificam simultaneamente para o modelo IGARCH(1,1), pois  $E[\sigma_{t+h}^2 | \sigma_t^2] = \alpha_0 h + \sigma_t^2 \rightarrow \infty$  quando  $h \rightarrow \infty$  mas  $\sigma_t^2 < \infty$  quase certamente. Considerando a definição de persistência  $E[\sigma_{t+h}^2 | \sigma_t^2] \rightarrow \infty$  ( $h \rightarrow \infty$ ), Engle e Bollerslev (1993) definem co-persistência na variância de uma forma similar à da co-integração na média: pode suceder que várias sucessões sejam persistentes na variância, mas uma certa combinação linear dessas sucessões pode exibir não persistência.

#### Modelo GARCH com alterações de regime

Em aplicações empíricas, por vezes, a situação  $\hat{\alpha}_1 + \hat{\delta}_1 \approx 1$  num modelo GARCH(1,1) deve-se a alterações exógenas na variância não condicionada que não foram contempladas na especificação do modelo. Por exemplo, Nicolau (1994) analisa a taxa de câmbio do marco (em escudos) no período Março de 1989 a Maio de 1993. Durante este produziram-se modificações substanciais no regime monetário (*crawling-peg*, «pré-SME» e adesão ao SME). Quando estas alterações são explicitamente consideradas na especificação da variância (através de variáveis artificiais) a soma das estimativas dos coeficientes  $\alpha_1$  e  $\delta_1$  diminui substancialmente. A introdução *ad hoc* de variáveis artificiais deve ser, no entanto, evitada, sobretudo quando é especificada para acontecimentos incertos que se podem repetir (pelo contrário, a especificação de variáveis artificiais para acontecimentos que nunca mais se repetem ou que se repetem e são totalmente certos, como por exemplo, os dias da semana, é justificável).

A abordagem dos modelos com alterações estocásticas de regime [de acordo com a exposição de Hamilton (1994)] é uma boa solução para processos com regimes diferenciados (sobretudo quando existem acontecimentos extraordinários e incertos que geram volatilidades extremas, como crises bolsistas, ataques especulativos, etc.). Os valores extremos da amostra não serão considerados como *outliers* mas como fazendo parte do processo gerador dos dados (estes acontecimentos têm probabilidade não nula de ocorrerem e devem ser contemplados na função de previsão). Por exemplo, admita-se que um dado processo é governado por dois regimes (regime 1:  $S_t = 1$ ; regime 2:  $S_t = 2$ ) com médias e variâncias condicionadas diferentes (poderíamos ter funções de densidade de probabilidade distintas). No período  $t$  o processo, dado a  $\sigma$ -álge-

bra  $F_{t-1}$ , pode encontrar-se em cada um dos regimes com uma dada probabilidade:

$$\begin{cases} y_t | F_{t-1}, S_t = 1 \cap N(\mu_{t,1}, \sigma_{t,1}^2), & \text{com probabilidade } P[S_t = 1 | F_{t-1}] \\ y_t | F_{t-1}, S_t = 2 \cap N(\mu_{t,2}, \sigma_{t,2}^2), & \text{com probabilidade } P[S_t = 2 | F_{t-1}] \end{cases}$$

Por definição, tem-se:

$$P[S_t = i | F_{t-1}] = \sum_{j=1}^2 P[S_t = i | S_{t-1} = j, F_{t-1}] P[S_{t-1} = j | F_{t-1}]$$

e:

$$P[S_{t-1} = j | F_{t-1}] = P[S_{t-1} = j | y_{t-1}, F_{t-2}] = \frac{f(y_{t-1} | S_{t-1} = j, F_{t-2}) P[S_{t-1} = j | F_{t-2}]}{\sum_{k=1}^2 f(y_{t-1} | S_{t-1} = k, F_{t-2}) P[S_{t-1} = k | F_{t-2}]}$$

onde  $f(y_{t-1} | S_{t-1} = j, F_{t-2})$  é a função de densidade de probabilidade condicionada de  $y_{t-1}$ . Dada uma estimativa para  $f(y_{t-1} | S_{t-1} = j, F_{t-2})$  (basta dispor de uma estimativa para os parâmetros desconhecidos) e para  $P[S_t = i | S_{t-1} = j, F_{t-1}]$ , a expressão  $P[S_t = i | F_{t-1}]$  calcula-se facilmente, dada a sua estrutura recursiva.

Diferentes hipóteses têm sido estabelecidas para a probabilidade  $P[S_t = i | S_{t-1} = j, F_{t-1}]$ . A mais simples consiste em admitir que  $P[S_t = i | S_{t-1} = j, F_{t-1}] = P[S_t = i | S_{t-1} = j] = P[S_t = i]$ , i. e., o regime em  $t$  não depende nem do estado no momento anterior nem de  $F_{t-1}$  ( $P[S_t = i]$  podem ser interpretadas como as probabilidades estacionárias de uma cadeia de Markov). Esta hipótese (no contexto do modelo ARCH) encontra-se em Palm e Vlaar (1993) sob a designação de distribuição Normal-Bernoulli. Hamilton e Susmel (1994) consideram o modelo ARCH e admitem  $P[S_t = i | S_{t-1} = j, F_{t-1}]$  que pode ser caracterizado por uma cadeia de Markov (a tempo discreto e com espaço de estados discreto) constante ao longo do tempo, i. e.,  $P[S_t = i | S_{t-1} = j, F_{t-1}] = P[S_t = i | S_{t-1} = j]$ . Gray (1996) resolve o problema da estimação do modelo GARCH no contexto em que  $P[S_t = i | S_{t-1} = j, F_{t-1}]$  é caracterizada por uma cadeia de Markov variável ao longo do tempo. Concretamente, Gray admite  $P[S_t = i | S_{t-1} = j, F_{t-1}] = \Phi(c_{ij} + d_{ij} y_{t-1})$ , onde  $\Phi(\cdot)$  é a função de distribuição normal e os parâmetros  $c_{ij}$  e  $d_{ij}$  e são estimados pelo método da máxima verosimilhança. Note-se que, em qualquer caso, a função de verosimilhança  $L_n(\theta)$  deve ser definida a partir das densidades  $f(y_t | F_{t-1})$  e não a partir das densidades  $f(y_t, S_t | F_{t-1})$  pois os regimes não são observados. Assim:

$$\log L_n(\theta) = \sum_{t=1}^n \log f(y_t | F_{t-1})$$

com:

$$f(y_t | F_{t-1}) = \sum_{i=1}^2 f(y_t | S_t = i, F_{t-1}) P[S_t = i | F_{t-1}]$$

#### Outros modelos

Outros modelos importantes são o modelo EARCH [Nelson (1991)] e o ARCH-M de Engle, Lilien e Robinson (1987). O primeiro permite estimar o efeito assimétrico discutido no ponto 2; o segundo permite que a variância condi-

cionada influencia explicitamente a média condicionada (por exemplo, a volatilidade influencia o preço do activo financeiro).

Outros modelos têm sido apresentados: NARCH, AARCH, TARCH<sup>(8)</sup>, QARCH, etc. Estes modelos propõem outras especificações para a variância condicionada, mas o princípio é o mesmo do modelo ARCH original: a variância condicionada é escrita como uma função estritamente positiva das variáveis  $u_{t-1}, u_{t-2}, \dots$ , isto é,  $\sigma_t^2 = g(u_{t-1}, u_{t-2}, \dots)$ , onde  $u_t \equiv y_t - E[y_t | F_{t-1}]$ .

Os modelos ARCH admitem uma representação multivariada e permitem estimar os co-movimentos na volatilidade [v. Bollerslev, Chou e Kroner (1992) e Engle, Bollerslev e Nelson (1993)]. Nicolau (1994) mostra uma aplicação de um sistema de equações com efeito ARCH para as taxas de câmbio, com as seguintes equações: marco em escudos, peseta em escudos e dólar americano em marcos (as variáveis estão definidas nas primeiras diferenças dos logaritmos e os dados são diários de Março de 1989 a Maio de 1993). São ainda estimados os coeficientes de correlação condicionados entre as diferentes moedas e discutido, com base nesses coeficientes, a constituição de carteiras de investimento monetário com níveis diferenciados de risco.

#### 4 — Identificação do modelo ARCH

São necessárias as seguintes etapas para identificar o modelo ARCH: primeiro, identifica-se o efeito ARCH, isto é, verifica-se se a variância condicionada depende ou não de  $F_{t-1}$ ; segundo, se o efeito ARCH está presente, identifica-se, na classe dos modelos ARCH (GARCH, EARCH, etc.) o modelo mais apropriado e a ordem dos parâmetros associada ao modelo seleccionado.

Consideremos a primeira questão<sup>(9)</sup>. Seja uma v. a. estacionária na média e expurgada de efeitos lineares. Quando o efeito ARCH está presente em  $u_t$  tem-se:

- 1) A distribuição de  $u_t$  é leptocúrtica;
- 2)  $u_t^2$  está correlacionado com  $u_{t-1}^2, \dots, u_{t-q}^2$ ;
- 3)  $u_t$  não é i. i. d;
- 4)  $E[u_t | F_{t-1}] = 0$ .

Como podemos verificar 1) a 4)? Para verificar 1) calcula-se o coeficiente de curtose e testa-se se é significativamente superior ao valor 3; 2) verifica-se através da FAC da variável  $u_t^2$ ; 3) pode verificar-se, por exemplo, pelo teste BDS [v., por exemplo, Nicolau (1994)]; e 4) analisa-se, por exemplo, através do teste de Hsieh (1989). Como salienta Hsieh (1989) é importante analisar 4)

<sup>(8)</sup> Para um desenvolvimento sobre as propriedades e comportamento das previsões do processo GTARCH v. Gonçalves e Lopes (1995).

<sup>(9)</sup> Em certas circunstâncias o modelo ARCH é aparente e resulta de uma deficiente especificação na média condicionada. Por exemplo, um modelo AR(1) exhibe efeitos similares a um ARCH(1) se a média condicionada não for modelizada, i. e., se assumirmos que o AR(1) é um ruído branco.

porque certos modelos não lineares do tipo MA não linear, limiar AR e bilinear apresentam algumas características similares às dos modelos ARCH e satisfazem 1) 2) e 3). No entanto, estes modelos não lineares não satisfazem 4), pois são do tipo  $u_t = \varepsilon_t + f_t$  onde  $\varepsilon_t$  é i. i. d. e tem média nula,  $f_t$  é não linear em  $(u_{t-1}, u_{t-2}, \dots)$  e  $E[u_t | F_{t-1}] \neq 0$ .

Consideremos a segunda questão. Uma vez confirmada a presença do efeito ARCH, é necessário seleccionar o modelo (ARCH, GARCH, etc.). Não aprofundaremos esta questão. Note-se, no entanto, que *a priori* o investigador deverá poder caracterizar genericamente a forma funcional da variância condicionada, nos seguintes aspectos: relação funcional entre os valores passados de  $u_t$  e  $\sigma_t^2$  e memória temporal da variância condicionada. Na relação funcional, é necessário distinguir os principais efeitos discutidos na literatura: efeito magnitude, efeito assimétrico e efeito de interacção. O efeito magnitude (valores altos de  $u_{t-1}^2, \dots, u_{t-q}^2$  implicam um valor alto de  $\sigma_t^2$ ) está presente na generalidade das sucessões financeiras e é estimado por todos os modelos ARCH; o efeito assimétrico (os valores e sinais de  $u_{t-1}, \dots, u_{t-q}$  influenciam  $\sigma_t^2$ ) é estimado pelos modelos EARARCH, QARCH e TARARCH; o efeito de interacção (os termos cruzados  $u_{t-i}u_{t-j}$  para  $i, j = 1, 2, \dots, q$  e  $i \neq j$ , influenciam  $\sigma_t^2$ ) é estimado pelos modelos AARCH, QARCH, etc.

Outra questão é a extensão da memória temporal da variância condicionada, que está relacionada com o período máximo de desfasamento entre a sucessão  $u_t$  e  $\sigma_t^2$ . Sempre que este período for elevado é preferível considerarmos o modelo generalizado (GARCH, EGARCH, TGARCH, etc.). Em geral, sempre que se aumenta a frequência das observações, o período máximo tende a aumentar [v. Nijman e Drost (1993) sobre o problema da agregação temporal].

A par da interpretação do investigador sobre o modelo mais apropriado, é conveniente realizar ensaios estatísticos assintóticos apropriados (o mais utilizado é o teste multiplicador de Lagrange). Estão bem estabelecidos os principais testes para os diferentes modelos como, por exemplo, ARCH, GARCH, NARCH, AARCH, etc. [v., por exemplo, Nicolau (1994) <sup>(10)</sup>]. Se vários modelos são simultaneamente não rejeitados pode-se proceder, como em Hsieh (1989), utilizando os critérios AIC e BIC.

## 5 — Estimação do modelo ARCH

Considere-se o seguinte modelo de regressão ou ARMA:

$$y_t = \mu_t(\beta_0) + \sigma_t(\alpha_0)\varepsilon_t$$

onde  $\beta_0$  e  $\alpha_0$  representam os vectores dos verdadeiros parâmetros desconhecidos, especificados, respectivamente, na média condicionada,  $\mu_t(\cdot)$ , e no desvio padrão condicionado,  $\sigma_t(\cdot)$ . O objecto é estimar consistentemente  $\beta_0$  e  $\alpha_0$ .

<sup>(10)</sup> O teste GARCH encontra-se em errata de Nicolau (1994).

Vamos admitir as seguintes hipóteses:

$\varepsilon_t$  é i. i. d. e tem distribuição  $N(0,1)$ ;  
 $\{y_t\}$  é um processo estacionário.

Dado que  $p(y_n, y_{n-1}, \dots, y_1) = p(y_n | F_{n-1}) p(y_{n-1} | F_{n-2}) \dots p(y_1)$  e, como  $y_t | F_{t-1}$  tem distribuição  $N(\mu_t(\beta_0), \sigma_t^2(\alpha_0))$ , o logaritmo da função de verosimilhança corresponde à expressão:

$$\log L_n(\alpha, \beta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^n \log \sigma_t^2(\alpha) - \frac{1}{2} \sum_{t=1}^n \frac{(y_t - \mu_t(\beta))^2}{\sigma_t^2(\alpha)}$$

Pode-se provar, sob as condições usuais de regularidade [v. por exemplo Gourieroux e Monfort (1995), cap. 7], que os estimadores de máxima verosimilhança são fortemente consistentes e:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, J_{\theta_0}^{-1})$$

onde:

$$J_0 = -E \left[ \frac{\partial^2 \log l_t}{\partial \theta \partial \theta'} \right] = E \left[ \frac{\partial \log l_t}{\partial \theta} \frac{\partial \log l_t}{\partial \theta'} \right] \text{ e } \log l_t = -\frac{1}{2} \log \sigma_t^2(\alpha) - \frac{1}{2} \frac{(y_t - \mu_t(\beta))^2}{\sigma_t^2(\alpha)}$$

É importante registar que os estimadores de máxima verosimilhança de  $\alpha_0$  e  $\beta_0$  são assintoticamente não correlacionados (o modelo ARCH-M é a exceção), i. e., a matriz  $J_0$  é diagonal por blocos relativamente aos vectores  $\alpha_0$  e  $\beta_0$  [Engle (1982)]. Desta forma é possível estimar os vectores  $\alpha_0$  e  $\beta_0$  separadamente sem perda de eficiência. A estimação de cada um deles pode basear-se numa estimativa consistente do outro. Além disso, como a matriz de informação é diagonal por blocos a sua inversão é fácil de obter. Assim, é possível aplicar, por exemplo, o algoritmo scoring, separadamente, para  $\alpha_0$  e  $\beta_0$ .

Em várias aplicações tem-se verificado que a hipótese de  $\varepsilon_t$  ter distribuição  $N(0,1)$  não se ajusta convenientemente aos dados, pois verifica-se, por vezes, que a distribuição de  $\varepsilon_t$  é leptocúrtica e (menos frequentemente) assimétrica. Nestes casos, a utilização (forçada) da hipótese de  $\varepsilon_t$  ter distribuição  $N(0,1)$  conduz seguramente a uma perda da eficiência dos estimadores. Quando  $\varepsilon_t$  é leptocúrtica a solução é procurar distribuições com essa característica. A distribuição GED( $\nu$ ) [Generalized error distribution, Nelson (1991) e Bollerslev e Baillie (1989)]:

$$f(\varepsilon_t) = \nu \exp \left\{ -\frac{1}{2} \left| \frac{\varepsilon_t}{\lambda} \right|^\nu \right\} / \lambda 2^{(1-1/\nu)} \Gamma \left( \frac{1}{\nu} \right), \quad 0 < \nu \leq +\infty$$

onde  $\Gamma(\cdot)$  é a função gama e  $\lambda = [2^{-2/\nu} \Gamma(1/\nu) / \Gamma(3/\nu)]^{1/2}$ , é interessante porque permite incorporar, no vector dos parâmetros a estimar, um parâmetro  $\nu$  que fornece uma indicação do achatamento da distribuição. Algumas distribuições importantes são casos particulares da distribuição GED( $\nu$ ). Assim, se  $\nu = 2$  a distribuição GED corresponde à distribuição normal; se  $\nu > 2$  a distribuição GED é platicúrtica e, em particular, tende para a distribuição uniforme no intervalo  $[-3^{1/2}, 3^{1/2}]$  quando  $\nu \rightarrow +\infty$ ; finalmente, se  $\nu < 2$  a distribuição GED é leptocúrtica e, em especial, com  $\nu = 1$ , a distribuição GED corresponde à distribuição de

Laplace ou exponencial dupla. A estimativa (de máxima verosimilhança) do parâmetro  $v$  fornece uma indicação do grau de achatamento de  $f(\varepsilon_t)$ .

O caso em que  $\varepsilon_t$  tem distribuição assimétrica é analisado, por exemplo, em, Palm e Vlaar (1993). É importante notar [Taylor (1986)] que um ou dois *outliers* podem ser responsáveis por coeficientes de *skewness* estimados aparentemente diferentes de zero. Nestes casos, a assimetria é apenas aparente, e resulta, em última análise, da natureza leptocúrtica da distribuição.

Todavia, em certos modelos, a assimetria está presente. Observa-se, por exemplo, para a distribuição das variações das taxas de câmbio de algumas moedas do SME, um comportamento assimétrico, particularmente forte. E a razão é esta: certas moedas do SME estão frequentemente sujeitas a realinhamentos e ou a ataques especulativos que implicam, normalmente, fortes depreciações (isto é especialmente válido para as moedas cotadas face ao marco). Nestes casos, uma das abas da distribuição tende a «alongar-se» mais e a distribuição torna-se assimétrica.

Nestas circunstâncias, Palm e Vlaar (1993) sugerem a utilização das distribuições Normal-Bernoulli ou Normal-Poisson. Estas distribuições admitem a possibilidade de a variável sofrer «saltos» aleatórios (no contexto das taxas de câmbio estes saltos podem corresponder a realinhamentos ou ataques especulativos).

A f. d. p. Normal-Bernoulli <sup>(11)</sup> tem por expressão:

$$f(u_t | F_{t-1}) = (1 - \lambda) \frac{1}{\sigma_t \sqrt{2\pi}} \exp \left\{ -\frac{(u_t + \lambda v)^2}{2\sigma_t^2} \right\} + \lambda \frac{1}{\sqrt{\sigma_t^2 + d^2} \sqrt{2\pi}} \exp \left\{ -\frac{(u_t - (1 - \lambda)v)^2}{2(\sigma_t^2 + d^2)} \right\}$$

onde  $\lambda$  é a probabilidade de ocorrer um salto e  $d^2$  é o efeito sobre a variância decorrente do salto. Quando  $v \neq 0$  e  $\lambda \neq 0$  a f. d. p. é assimétrica, como podemos constatar através do terceiro momento não condicionado:  $E[u_t^3] = (\lambda - \lambda^2) v[(1 - 2\lambda)v^2 + 3d^2]$ . O parâmetro  $v$  esclarece ainda o sinal da assimetria (tendo-se em geral  $\lambda < 0.5$ ).

Uma forma mais expedita de resolver o problema da distribuição de  $\varepsilon_t$  é estimar os parâmetros através do método da pseudo máxima verosimilhança, considerando, por exemplo, como pseudo verdadeira distribuição, a distribuição normal [v. Gouriéroux e Monfort (1995)] para um exposição sobre o método). Um método de estimação não paramétrico é analisado em Engle e Riviera (1991).

## 6 — Modelos de volatilidade estocástica e os modelos ARCH

Recentemente tem havido algum interesse nos modelos designados por volatilidade estocástica (VE), introduzidos por Clark (1973) que, na sua forma mais simples, são descritos por:

$$\begin{cases} y_t = \sigma_t \varepsilon_t \\ \log \sigma_t^2 = \alpha + \delta \log \sigma_{t-1}^2 + \sigma_v v_t \end{cases}$$

onde  $\varepsilon_t$  são  $v_t$  v. a. i. i. d. com distribuição  $N(0,1)$ .

<sup>(11)</sup> Deve-se notar que o modelo subjacente a esta distribuição corresponde a um caso particular do modelo com alterações de regime já discutido no ponto 3.

No modelo ARCH a variância define-se como uma especificação determinística de variáveis aleatórias; no modelo de VE a variância incorpora uma v. a. residual.

De acordo com Clark, o mercado de títulos encontra-se num de dois estados: equilíbrio Walrasiano ou transição para o equilíbrio. O que faz o mercado deslocar-se de um equilíbrio para o outro é a chegada de informação. À medida que a informação chega aos mercados os agentes heterogéneos reavaliam os portfólios desejados e transaccionam até que um novo equilíbrio seja alcançado. No modelo de VE a fonte de volatilidade do preço dos títulos deve-se à chegada de informação. Embora a teoria não sugira nenhuma forma funcional para a chegada de informação, Clark modeliza  $\sigma_t$  como um processo log-normal com inovações i. i. d e conclui que essa especificação é preferível à especificação de Poisson. Outros autores chegam à mesma conclusão [por exemplo, Gallant, Hsieh e Tauchen (1991)].

O modelo de VE pode ser visto ainda como uma discretização do processo a tempo contínuo (discretização de Euler):

$$\begin{cases} dy_t = -y_t dt + \sigma_t dw_{t,1} \\ d\log\sigma_t^2 = (\alpha + \delta - 1)\log\sigma_t^2 dt + \sigma_v dw_{t,2} \end{cases}$$

onde  $w_{t,1}$  e  $w_{t,2}$  são processos brownianos independentes (em rigor, a estimação dos parâmetros deve realizar-se no contexto do modelo a tempo contínuo).

Ao contrário do modelo ARCH, o modelo de VE levanta sérios problemas de estimação. Considere-se, por exemplo, o modelo discutido em Danielsson (1994):

$$\begin{cases} y_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \cap \text{i. i. d. } N(0,1) \\ \log\sigma_t^2 = \omega + \gamma_1 y_{t-1} + \gamma_2 |y_{t-1}| + \delta \log\sigma_{t-1}^2 + \sigma_v v_t, \quad v_t \cap \text{i. i. d. } N(0,1) \end{cases}$$

seja  $\theta = (\omega, \delta, \sigma_v, \gamma_1, \gamma_2)^T$  o vector a estimar,  $Y_n = \{y_t\}_{t=1, \dots, n}$  e  $\Sigma_n = \{\sigma_t^2\}_{t=1, \dots, n}$ . A densidade conjunta  $f(Y_n, \Sigma_n | \theta)$  obtém-se facilmente a partir de:

$$\begin{aligned} f(Y_n | \Sigma_n, \theta) &= \prod_{t=1}^T \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left\{-\frac{y_t^2}{2\sigma_t^2}\right\} \\ f(\Sigma_n | \theta) &= \prod_{t=1}^T \frac{1}{v_t \sqrt{2\pi}} \exp\left\{-\frac{1}{2v^2} \frac{(\log\sigma_t^2 - \omega - \gamma_1 y_{t-1} - \gamma_2 |y_{t-1}| - \delta \log\sigma_{t-1}^2)}{2\sigma_t^2}\right\} \\ f(Y_n, \Sigma_n | \theta) &= f(Y_n | \Sigma_n, \theta) f(\Sigma_n | \theta) \end{aligned}$$

Se  $\Sigma_n$  é observável, o método da máxima verosimilhança baseia-se na maximização da função objectivo  $f(Y_n, \Sigma_n | \theta) = f(Y_n | \Sigma_n, \theta) f(\Sigma_n | \theta)$  em ordem a  $\theta$ , i. e.:

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} f(y_n, \Sigma_n | \theta)$$

No entanto, como a variável  $\sigma_t$  é latente (não observável),  $\sigma_t$  não pode figurar na função objectivo e, portanto, o método da máxima verosimilhança requer que a densidade conjunta  $f(Y_n, \Sigma_n | \theta)$  seja integrada em ordem a  $\Sigma_n$ , isto é:

$$f(Y_n | \theta) = \int_{R^n} f(Y_n | \Sigma_n, \theta) f(\Sigma_n | \theta) d\Sigma_n$$

e o estimador de máxima verosimilhança obtém-se a partir de:

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} f(Y_n | \theta)$$

O problema está em que não é fácil deduzir  $f(Y_n | \theta)$ , sobretudo quando a especificação de  $\sigma_t$  é dinâmica. Existem várias estratégias para lidar com este problema. Por exemplo, Danielsson (1994), considera a simulação da função de verosimilhança, Harvey, Ruiz e Shephard (1994) e Ruiz (1994) empregam métodos baseados em filtros lineares com vista a produzirem o estimador de pseudo máxima verosimilhança, Taylor (1986), Melino e Turnbull (1990) consideram o GMM, Jacquier, Polson e Rossi (1994) consideram métodos bayesianos.

Será interessante no futuro comparar detalhadamente as vantagens e desvantagens dos modelos ARCH com os modelos de volatilidade estocástica. Engle (1994) é bastante crítico com o modelo de VE.

## 7 — Equações diferenciais estocásticas e os modelos ARCH

A utilização das equações diferenciais estocásticas (EDE) na área dos modelos financeiros tem aumentado significativamente [Melino (1991) estima que a proporção de artigos nas quatro mais importantes revistas na área financeira passou de 0,8 % em 1970 para 17,1 % em 1989]. Apesar do interesse crescente nas EDE subsiste ainda a ideia da clivagem entre os métodos discretos e contínuos [v. Nelson (1990a)]. No entanto, algum trabalho tem sido realizado com vista a conciliar esses métodos, nomeadamente nos seguintes aspectos: estudo das condições de convergência de sistemas discretos para sistemas contínuos (assim como o estudo do problema da discretização de EDE) e estimação de EDE.

Nelson (1990a), a partir da especificação discreta GARCH(1,1)-M, deduz, sob certas condições de regularidade, um sistema de EDE compatível com a especificação discreta e estabelece, para o caso contínuo, a distribuição estacionária de  $\sigma_t^2$ -gama invertida (de notar que não existe para a especificação discreta uma expressão definida para a distribuição estacionária).

O efeito ARCH pode ser modelizado por equações diferenciais estocásticas univariadas. Por exemplo, o efeito assimétrico descrito no ponto 2 pode ser estimado, pela EDE [proposta por Duffie-Kan, v. Aït-Sahalia (1996)]:

$$dy_t = \beta(\theta - y_t)dt + \sqrt{\alpha + \gamma y_t} dw_t$$

onde  $\{w_t; t \geq 0\}$  é um processo browniano *standard*. Pode-se provar que a expressão exacta da variância condicionada em tempo discreto corresponde a:

$$\text{Var}[y_{t_i} | F_{t_{i-1}}] = \frac{1}{2\beta} (1 - e^{-\beta\Delta})(\theta\gamma - e^{-\beta\Delta}\gamma\theta + \alpha + e^{-\beta\Delta}\alpha + 2e^{-\beta\Delta}\gamma y_{t_{i-1}})$$

onde  $t_i (i = 1, \dots, n)$  são os momentos em que  $Y$  é observado e  $\Delta = t_i - t_{i-1}$  (constante, por hipótese). Note-se que a variância condicionada é não constante e depende do valor e do sinal de  $y_{t_{i-1}}$ .

O efeito de magnitude e assimétrico pode ser estimado, por exemplo, através da EDE:

$$dy_t = \beta(\theta - y_t) dt + \alpha y_t dw_t$$

e neste caso pode-se provar <sup>(12)</sup> que:

$$\text{Var}[y_{t_i} | F_{t_{i-1}}] = \varphi_1 + \varphi_2 y_{t_{i-1}} + \varphi_3 y_{t_{i-1}}^2$$

onde  $\varphi_i, i = 1, 2, 3$ , são expressões não lineares dos parâmetros  $\alpha$ ,  $\beta$  e  $\theta$  do tempo de discretização  $\Delta$ .

Relativamente a:

$$dy_t = \theta y_t dt + \alpha y_t dw_t$$

tem-se:

$$\text{Var}[y_{t_i} | F_{t_{i-1}}] = \varphi y_{t_{i-1}}^2$$

com  $\varphi = e^{(2\theta + \alpha^2)\Delta} - e^{2\theta\Delta}$  e, recorrendo a Johnson e Kotz (1970), p. 115, pode-se provar que o coeficiente de curtose da distribuição de  $y_t$  é estritamente maior do que 3 [note-se que  $y_t$  tem distribuição log normal, dado a  $\sigma$ -álgebra  $F_s (s < t)$ ].

Deve-se sublinhar que a variância condicionada dos modelos ARCH não depende explicitamente do nível da sucessão mas sim da v. a. residual (desta forma podem ocorrer divergências significativas entre as estimativas de volatilidade dos dois modelos — ARCH e EDE). No entanto, o princípio é o mesmo: a variância condicionada é, em ambos os casos, uma função de  $F_{t_{i-1}}$ , i. e.,  $\text{Var}[y_{t_i} | F_{t_{i-1}}] = \sigma^2(F_{t_{i-1}})$ .

## 8 — Conclusões

O problema da estimação da volatilidade não é recente, mas só depois do trabalho de Engle (1982) foi possível definir rigorosamente um modelo paramétrico para a sua estimação. Muitos trabalhos foram publicados no contexto

<sup>(12)</sup> É prática corrente considerar a variância condicionada do processo de difusão como o coeficiente de difusão discretizado de acordo com o esquema de Euler. Em certos casos, esta prática conduz a divergências importantes entre a expressão exacta da variância condicionada em tempo discreto e a variância condicionada aproximada de acordo com o esquema de Euler, como se vê no presente exemplo.

dos modelos ARCH, mas actualmente os investigadores procuram novas formas de interpretar e estimar a volatilidade. É necessário, na minha opinião, discutir dois aspectos essenciais: a volatilidade no momento  $t$  é uma função das variáveis  $\{y_{t-1}, y_{t-2}, \dots, x_{t-1}, x_{t-2}, \dots\}$ , tal como está especificado no modelo ARCH, ou eventualmente, deve ser caracterizada também através de uma variável aleatória residual, tal como está definido no modelo de volatilidade estocástica? O segundo aspecto tem a ver com a forma como interpretamos a evolução das variáveis no tempo. O paradigma dominante em econometria é o de que as variáveis económicas evoluem através de saltos discretos ao longo do tempo. Este ponto de vista tem sido criticado pelo menos desde A. R. Bergstrom [v. a sua entrevista em Bergstrom (1993)] e actualmente, no domínio dos modelos financeiros, o interesse pelos modelos a tempo contínuo tem vindo a crescer.

Seria interessante no futuro discutir, com aplicações concretas, as vantagens e desvantagens dos modelos do tipo ARCH *versus* modelos de volatilidade estocástica quer em tempo discreto quer em tempo contínuo.

Um passo importante na abordagem dos modelos de volatilidade seria estimar consistentemente os parâmetros do modelo a tempo contínuo com volatilidade estocástica (variável latente) apenas com a hipótese de que o número das observações tenda para infinito, independentemente de o intervalo entre as observações tender ou não para zero. Procuraremos abordar este tema num próximo trabalho.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AÏT-SAHALIA, Y. (1996), «Nonparametric Pricing of Interest Rate Derivative Securities», *Econometrica*, vol. 64, n.º 3.
- BERGSTROM, A. R. (1993), «The ET Interview: A. R. Bergstrom», in Phillips P. B. (ed.), *Models, Methods and Applications of Econometrics*, B. Blackwell, Cambridge, Mass.
- BLACK, F. (1976), «Studies of Stock Markets Volatility Changes», *1976 Proceedings of the American Statistical Association, Business And Economic Statistics Section*, 177-181.
- BOLLERSLEV, T. (1986), «Generalized Autoregressive Conditional Heteroskedasticity», *Journal of Econometrics*, vol. 31, 307-327.
- BOLLERSLEV, T., e BAILLIE, R. T. (1989), «The Message in Daily Exchange Rates: A Conditional-Variance Tale», *Journal of Business & Economic Statistics*, vol. 7, n.º 3, 297-305.
- BOLLERSLEV, T., CHOU, R. Y., e KRONER, K. F. (1992), «Arch Modeling in Finance: A Review of the Theory And Empirical Evidence», *Journal of Econometrics*, vol. 52, 5-59.
- CLARK, P. (1973), «A subordinated Stochastic Process Model with finite Variance for Speculative Process», *Econometrica*, 41, 135-155.
- CORNELL, B. (1978), «Using the Options Pricing Model to Measure the Uncertainty Producing Effect of Major Announcements», *Financial Management*, 7, 54-59.
- DANIELSSON, J. (1994), «Stochastic Volatility in Asset Prices — Estimation with Simulated Maximum Likelihood», *Journal of Econometrics*, 64, 375-400.
- ENGLE, R. F. (1982), «Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation», *Econometrica*, vol. 50, n.º 4, 987-1008.
- ENGLE, R. F., BOLLERSLEV, T., e NELSON, D. B. (1993a), «Arch Models», prepared for the *Handbook of Econometrics*, vol. 4.
- ENGLE, R. F. (1994), «Comment», *Journal of Business & Economic Statistics*, vol. 12, n.º 4, 395-396.
- ENGLE, R. F., LILIEN, D. M., e ROBINS, R. P. (1987), «Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model», *Econometrica*, vol. 55, n.º 2, 391-407.
- ENGLE, R. F., e BOLLERSLEV, T. (1993), «Common Persistence in Conditional Variances», *Econometrica*, vol. 61, n.º 1, 167-186.
- ENGLE, R. F., VICTOR, K. N. G., e ROTHSCHILD, M. (1990), «Asset Pricing with a Factor-ARCH Covariance — Empirical Estimates for Treasury Bills», *Journal of Econometrics*, vol. 45, 213-237.
- GALLANT, A. R., HSIEH, D. A., e TAUCHEN, G. E., «On fitting a Recalcitrant Series: The Pound/dollar Exchange Rate, 1973-83», in Barnett W., Powell J., e Tauchen G. (ed.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press, Cambridge.
- GONÇALVES, E., e LOPES, N. M. (1995), «Séries Temporais com Erros Condicionamente Heteroscedásticos», in Branco, J., Gomes, P., e Prata, J., (eds.), *Bom Senso e Sensibilidade*, Novas Tecnologias, Edições Salamadra.
- GOURIEROUX, C., e MONFORT, A. (1995), *Statistics and Econometric Models*, vol. 1, Cambridge University Press.
- GRAY, S. F. (1996), «Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process», *Journal of Financial Economics*, 42, 25-62.
- HAMILTON, J. D. (1994), *Time Series Analysis*, Princeton University Press.
- HAMILTON, J. D., e SUSMEL, R. (1994), «Autoregressive Conditional Heteroscedasticity and Change in Regime», *Journal of Econometrics*, 64.
- HARVEY, A., RUIZ, E., e SHEPHARD, N. (1994), «Multivariate Stochastic Variance Models», in Engle, R. (ed.), *ARCH — Selected Readings*, Advanced Texts in Econometrics, Oxford University Press.
- HARVEY, C. R., e HUANG, R. D. (1991), «Volatility in the Foreign Currency Futures Markets», *Review of Financial Studies*, 4, 543-569.
- HONG, C. H. (1988), «The Integrated Generalized Autoregressive Conditional Heteroskedastic Model: The Process, Estimation, and Monte Carlo Experiments», trabalho não publicado.
- HSIEH, A. D. (1989a), «Testing for Nonlinear Dependence in Daily Foreign Exchange Rates», *Journal of Business*, vol. 62, n.º 3, 339-368.
- JOHNSON, N. L., e KOTZ, S., (1970a), *Continuous Univariate Distributions*, vol. 1, John Wiley & Sons.

- LUMSDAINE, R. L. (1991), «Asymptotic Properties of the Maximum Likelihood Estimator in Garch(1,1) and IGARCH(1,1) Models», trabalho não publicado.
- MELINO, A. (1991), «Estimation of Continuous time Models in Finance», University of Toronto, Department of Economics and Institute for Policy Analysis, working paper 9115.
- MELINO, A., e TURNBULL, S. M. (1990), «Pricing Foreign Currency Options with Stochastic Volatility», *Journal of Econometrics*, vol. 45, 239-265.
- NELSON, D. B. (1990a), «Arch Models as Diffusion Approximations», *Journal of Econometrics*, vol. 45, 7-38.
- (1990b), «Stationarity and Persistence in the GARCH(1,1) Model», *Econometric Theory*, n.º 6, 318-334.
- (1991), «Conditional Heteroskedasticity in Asset Returns: A New Approach», *Econometrica*, vol. 59, n.º 2, 347-370.
- NICOLAU, J. C. (1994), «Modelos ARCH», dissertação apresentada como requisito parcial para a obtenção do grau de mestre em Matemática Aplicada à Economia e Gestão (sob orientação do Prof. Doutor Carlos Silva Ribeiro), ISEG-UTL, Lisboa.
- NIJMAN, T. E., e DROST, F.C. (1993), «Temporal Aggregation of GARCH Processes», *Econometrica*, vol. 61, n.º 4, 909-927.
- PALM, F. C., e VLAAR, P. J. G. (1993), «The Message in Weekly Exchange Rates in The European Monetary System: Mean Reversion, Conditional Heteroscedasticity and Jumps», *Journal of Business & Economic Statistics*, vol. 11, n.º 3.
- RUIZ, E. (1994), «Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models», *Journal of Econometrics*, vol. 63, 287-306.
- SCHWERT, G. W. (1989), «Why Does Stock Market Volatility Change Over Time», *Journal of Finance*, 44, 1115-1153.
- TAYLOR, S. J. (1986), *Modelling Financial Time Series*, John Wiley & Sons.

