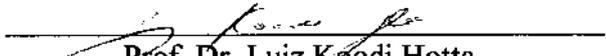


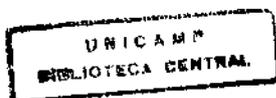
# **Volatilidade nos Modelos ARCH e Variância Estocástica: um estudo comparativo**

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. Mauricio Enrique Zevallos Herencia e aprovada pela Comissão Julgadora

Campinas, 25 de Março de 1997

  
Prof. Dr. Luiz Koodi Hotta  
(Orientador)

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica - (IMECC), UNICAMP, como requisito parcial para a obtenção do título de MESTRE em ESTATÍSTICA.



UNIDADE BC  
N.º CHAMADA: 77 UNICAMP  
Z61v  
V. Ex  
TOMBO BU/ 30453  
PROC. 28 A 197  
C  D   
PREÇO R\$ 11,00  
DATA 23/02/197  
N.º CPD

CM-00098840-3

**FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DO IMECC DA UNICAMP**

Zevallos Herencia, Mauricio Enrique

Z61v      Volatilidade nos modelos ARCH e variância estocástica: um estudo comparativo / Mauricio Enrique Zevallos Herencia -- Campinas, [S.P. :s.n.], 1997.

Orientador : Luiz Koodi Hotta

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Análise de series temporais. 2. Teoria da previsão. I. Hotta, Luiz Koodi. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Dissertação de Mestrado defendida e aprovada em 20 de fevereiro de 1997

pela Banca Examinadora composta pelos Profs. Drs.



Prof (a). Dr (a). CRISTIANO AUGUSTO COELHO FERNANDES



Prof (a). Dr (a). DIÓGENES MANOEL LEIVA MARTINS



Prof (a). Dr (a). LUIZ KOODI HOTTA

A mis padres y hermanos  
A Nora

## Agradecimentos

- Ao professor Luiz Koodi Hotta, pela orientação, paciência, dedicação e amizade.
- Ao professor Pedro L. Valls Pereira, por participar como co-orientador, respondendo às dúvidas, criticar o trabalho e estar disposto a me atender.
- A CNPq, FAEP e BDLA pelo apoio financeiro.
- Aos membros da banca, Prof. Dr. Cristiano Coelho Fernandes e Dr. Diógenes Leiva Martins pelas sugestões e correções.
- Aos professores e colegas do mestrado, especialmente a Rissa, André e Rui .
- Aos meus professores e amigos do Perú.
- À minha família, sempre presente.
- À Nora, por todos os momentos...

# SUMÁRIO

Lista de Tabelas.....	xiv
Lista de Figuras.....	xviii
Resumo.....	xxii

## 1. CAPÍTULO 1: Introdução

1.1 Séries Financeiras.....	1
1.2 Volatilidade.....	2
1.3 Retornos e <i>Fatos Estilizados</i> .....	4
1.4 Modelos de Volatilidade abordados nesta dissertação.....	6
1.5 Descrição das Séries Financeiras a serem estudadas: <i>Séries Exemplo</i> .....	8
1.6 Nomenclatura e Definições Básicas.....	9
1.7 Breve Esboço da Dissertação.....	11

## 2. CAPÍTULO 2: Características Empíricas das Séries Financeiras: os *Fatos Estilizados*

2.1 Introdução.....	16
2.2 Conglomerados de Volatilidade.....	17
2.3 Simetria.....	17
2.4 Cauda Pesada.....	20
2.5 Medidas de Locação e Variabilidade.....	22
2.6 Estacionariedade.....	25
2.7 Autocorrelação, Independência e Não Linearidade.....	27
2.7.1 Testes de Autocorrelação.....	28
2.7.2 Independência e Não Linearidade.....	30
2.7.3 Outros Testes de Não linearidade.....	32
2.8 Normalidade.....	33
2.9 Outras Características Empíricas dos Retornos.....	35
2.10 Conclusões.....	36

### 3. CAPÍTULO 3: Modelos ARCH

3.1 Introdução.....	46
3.2 Definições e Propriedades Básicas: <i>Fatos Estilizados</i> .....	47
3.3 Quatro Versões dos Modelos ARCH.....	50
3.3.1 ARCH(q) .....	51
3.3.2 GARCH(p,q).....	53
3.3.2.1 GARCH(1,1).....	56
3.3.3 IGARCH(p,q).....	58
3.3.4 EGARCH.....	60
3.4 Estimação.....	64
3.4.1 Método de Máxima Verossimilhança.....	64
3.4.2 Métodos de Quase-Máxima Verossimilhança.....	72
3.5 Testes de Hipóteses e Diagnóstico.....	74
3.6 Previsão .....	75
3.6.1 Previsão nos Modelos GARCH e IGARCH.....	76
3.6.2 Previsão nos Modelos EGARCH.....	81
3.7 Aplicações nas <i>Séries Exemplo</i> .....	82
3.7.1 <i>Ações</i> .....	84
3.7.2 <i>Taxa de Câmbio</i> .....	100

### 4. CAPÍTULO 4: Modelos de Variância Estocástica

4.1 Introdução.....	132
4.2 Abordagem em Espaço de Estados dos Modelos de Variância Estocástica.....	133
4.3 Propriedades Básicas: Reprodução dos <i>Fatos Estilizados</i> .....	136
4.4 Estimação.....	143
4.4.1 Estimação Pontual.....	143
4.4.2 Incerteza na Estimação dos Parâmetros e na Estimação da Volatilidade.....	147
4.4.3 Eficiência comparativa dos Estimadores QML.....	148
4.5 Diagnóstico e Bondade de Ajuste.....	150

4.5.1 Diagnóstico.....	150
4.5.2 Bondade de Ajuste.....	152
4.6 Previsão.....	152
4.7 Extensões.....	155
4.7.1 Cauda Pesada.....	156
4.7.2 Correlação Serial e <i>Efeitos Calendário</i> .....	157
4.7.3 Intervenções.....	157
4.8 Aplicações nas <i>Séries Exemplo</i> .....	158
4.8.1 <i>Ações</i> .....	161
4.8.2 <i>Taxa de Câmbio</i> .....	170

## **5. CAPÍTULO 5: Comparação dos Modelos ARCH e Variância Estocástica.**

5.1 Introdução.....	192
5.2 Comparação empírica dos Modelos .....	193
5.3 Comparação empírica das Previsões um-passo à frente <i>ex-ante</i> .....	199
5.4 Comparação dos Modelos GARCH(1,1) e AR(1)-SV estacionário através de simulações.....	201
5.4.1 Modelos Equivalentes GARCH(1,1) e AR(1)-SV estacionário.....	202
5.4.2 Comparação das Estimativas de Volatilidade e Previsão um-Passo à frente.....	205
5.4.3 Comparação dos Modelos GARCH(1,1) e AR(1)-SV estacionário quando não existem Modelos Equivalentes.....	212
5.5 Conclusões.....	216

## **6. REFERÊNCIAS BIBLIOGRÁFICAS.....235**

# LISTA DE TABELAS

## Capítulo 1

Tabela 1.1: Descrição das <i>Séries Exemplo</i> .....	9
---	---

## Capítulo 2

Tabela 2.1: Estatísticas dos retornos nas <i>Séries Exemplo</i> .....	18
Tabela 2.2: Simetria: estimativas da regressão $v=\alpha+\beta u+\varepsilon$ .....	19
Tabela 2.3: Estatísticas da série <i>Ações</i> .....	23
Tabela 2.4: Estatísticas da série <i>Taxa de Câmbio</i> .....	23
Tabela 2.5: <i>Efeito Calendário</i> nas <i>Séries Exemplo</i> .....	26
Tabela 2.6: Valores- P do Teste Box-Ljung nos retornos das <i>Séries Exemplo</i> .....	28
Tabela 2.7: Percentagem das autocorrelações no Modelo AR(1)-SV estacionário que se encontram fora dos limites usuais de Bartlett.....	30

## Capítulo 3

Tabela 3.1: Valores de $\frac{f'(\varepsilon)}{f(\varepsilon)}$ para as densidades normal, <i>t-Student</i> e GED.....	68
Tabela 3.2: Identificação na série <i>Ações</i> .....	85
Tabela 3.3: Estimativas dos modelos GARCH, IGARCH e variância condicional constante na série <i>Ações</i> .....	86
Tabela 3.4: Análise dos resíduos e dos quadrados dos resíduos ARCH na série <i>Ações</i> .....	88
Tabela 3.5: Resíduos extremos GARCH(1,2) na série <i>Ações</i> .....	90
Tabela 3.6: Previsão um-passo à frente <i>ex-ante</i> considerando níveis na volatilidade estimada GARCH(1,2) na série <i>Ações</i> .....	91
Tabela 3.7: Ajustes EGARCH na série <i>Ações</i> .....	92

Tabela 3.8: Resíduos extremos EGARCH(2,2) na série <i>Ações</i> .....	94
Tabela 3.9: Previsão um-passo à frente <i>ex-ante</i> considerando níveis na volatilidade estimada EGARCH(2,2) na série <i>Ações</i> .....	94
Tabela 3.10: Resíduos extremos GARCH(1,1) na série <i>Ações</i> .....	96
Tabela 3.11: Previsão um-passo à frente <i>ex-ante</i> considerando níveis na volatilidade estimada GARCH(1,1) na série <i>Ações</i> .....	96
Tabela 3.12: Identificação na série <i>Taxa de Câmbio</i> .....	101
Tabela 3.13: Estimativas dos modelos GARCH, IGARCH e variância condicional constante na série <i>Taxa de Câmbio</i> .....	102
Tabela 3.14: Análise dos resíduos e dos quadrados dos resíduos ARCH na série <i>Taxa de Câmbio</i> .....	104
Tabela 3.15: Resíduos extremos GARCH(2,2) na <i>Taxa de Câmbio</i> .....	105
Tabela 3.16: Estimativas dos modelos EGARCH na série <i>Taxa de Câmbio</i> .....	106
Tabela 3.17: Resíduos extremos EGARCH(2,0) na série <i>Taxa de Câmbio</i> .....	108
Tabela 3.18: Resíduos extremos GARCH(1,1) na série <i>Taxa de Câmbio</i> .....	109
Tabela 3.19: Previsão um-passo à frente <i>ex-ante</i> considerando níveis na volatilidade estimada GARCH(2,2) na <i>Taxa de Câmbio</i> .....	111
Tabela 3.20: Previsão um-passo à frente <i>ex-ante</i> considerando níveis na volatilidade estimada EGARCH(2,0) na <i>Taxa de Câmbio</i> .....	111

#### Capítulo 4

Tabela 4.1: <i>Ações</i> : estimativas dos modelos com <i>Efeitos dos dias da semana</i> na volatilidade.....	163
Tabela 4.2: <i>Ações</i> : valores-P do Teste Box-Ljung nos QRP do AR(1)-SV estacionário...	165
Tabela 4.3: <i>Ações</i> : estabilidade das estimativas e previsão um-passo à frente <i>ex-ante</i> .....	166
Tabela 4.4: <i>Ações</i> : estimativas dos modelos AR(1)-SV não estacionários.....	168
Tabela 4.5: <i>Taxa de Câmbio</i> : valores-P do Teste Box-Ljung nos QRP do AR(1)-SV estacionário.....	171
Tabela 4.6: <i>Taxa de Câmbio</i> : estabilidade das estimativas e previsão um-passo à frente <i>ex-ante</i> .....	173

## Capítulo 5

Tabela 5.1: Diferenças entre as estimativas de volatilidade ARCH e AR(1)-SV estacionário nas <i>Ações</i> .....	195
Tabela 5.2: Diferenças entre as estimativas de volatilidade ARCH e AR(1)-SV estacionário na série <i>Taxa de Câmbio (%)</i> .....	196
Tabela 5.3: Persistências estimadas nas <i>Séries Exemplo</i> .....	197
Tabela 5.4: Previsão um-passo à frente na série <i>Ações</i> . Número de retornos fora dos intervalos de confiança 95% (# esperado=4,75).....	200
Tabela 5.5: Previsão um-passo à frente na série <i>Taxa de Câmbio</i> . Número de retornos fora dos intervalos de confiança 95% (# esperado=4,75).....	200
Tabela 5.6: Equivalência de alguns modelos GARCH(1,1) e AR(1)-SV estacionário.....	205
Tabela 5.7: Modelos utilizados nas simulações.....	206
Tabela 5.8: Simulações realizadas.....	206
Tabela 5.9: Modelo simulado GARCH(1,1): percentagem de superestimação da volatilidade.....	207
Tabela 5.10: Magnitude da subestimação e superestimação da volatilidade estimada (vs) AR(1)-SV com respeito à volatilidade verdadeira (vg) GARCH(1,1).....	208
Tabela 5.11: Modelo simulado GARCH(1,1): percentagem de retornos fora dos intervalos de confiança 95%.....	209
Tabela 5.12: Modelo simulado GARCH(1,1): efeito das observações muito pequenas nos quadrados dos retornos.....	209
Tabela 5.13: Magnitude da subestimação e superestimação da volatilidade estimada (vg) GARCH(1,1) com respeito à volatilidade verdadeira (vs) AR(1)-SV.....	211
Tabela 5.14: Modelo simulado AR(1)-SV: percentagem de retornos fora dos intervalos de confiança 95%.....	211
Tabela 5.15: Parâmetros do modelo simulado GARCH(1,1) e estimativas do modelo AR(1)-SV no caso quando não existe equivalência.....	213
Tabela 5.16: Magnitude da subestimação e superestimação da volatilidade estimada (vs) AR(1)-SV com respeito à volatilidade verdadeira (vg) GARCH(1,1) no caso de modelos não equivalentes.....	213
Tabela 5.17: Percentagem de retornos fora dos intervalos de confiança 95% se o modelo simulado é GARCH(1,1) e não existe equivalência.....	214
Tabela 5.18: Modelos AR(1)-SV simulados e estimativas GARCH(1,1) quando não existe equivalência.....	214

Tabela 5.19: Magnitude da subestimação e superestimação da volatilidade estimada (vg) GARCH(1,1) com respeito à volatilidade verdadeira (vs) AR(1)-SV quando não existe equivalência.....	215
Tabela 5.20: Percentagem de retornos fora dos intervalos de confiança 95% quando o modelo simulado é AR(1)-SV e não existe equivalência.....	215

# LISTA DE FIGURAS

## Capítulo 1

Figura 1.1: Gráfico dos preços de fechamento das <i>Séries Exemplo</i> .....	13
Figura 1.2: Box-plot dos preços de fechamento das <i>Séries Exemplo</i> .....	14
Figura 1.3: Gráfico dos retornos das <i>Séries Exemplo</i> .....	15

## Capítulo 2

Figura 2.1: Gráfico dos quadrados dos retornos das <i>Séries Exemplo</i> .....	38
Figura 2.2: Histograma, densidade ajustada e Gráfico de Simetria nas <i>Ações</i> .....	39
Figura 2.3: Histograma, densidade ajustada e Gráfico de Simetria na <i>Taxa de Câmbio</i> .....	40
Figura 2.4: Gráfico de Probabilidade Normal dos retornos das <i>Séries Exemplo</i> .....	41
Figura 2.5: Gráficos de Desvio Padrão Móvel e CUSUM dos quadrados dos retornos nas <i>Séries Exemplo</i> .....	42
Figura 2.6: Variâncias Recursivas das <i>Séries Exemplo</i> .....	43
Figura 2.7: Dinâmica dos retornos e dos quadrados dos retornos nas <i>Séries Exemplo</i> .....	44
Figura 2.8: Função de autocorrelação amostral do modelo AR(1)-SV com parâmetros $\gamma = 0,95$ , $q=0,05$ , $k=5,5 \times 10^{-7}$ .....	45

## Capítulo 3

Figura 3.1: Funções de autocorrelação dos resíduos ARCH e dos quadrados dos resíduos nas <i>Ações</i> .....	114
Figura 3.2: Gráfico de Probabilidade Normal dos resíduos padronizados GARCH(1,2) nas <i>Ações</i> .....	115
Figura 3.3: Volatilidades estimadas e i.c 95% para os retornos no período da <i>Mudança da Banda Cambial</i> .....	116

Figura 3.4: Volatilidades estimadas e i.c 95% para os retornos no período da <i>Queda da Ministra Zélia Cardoso de Mello</i> .....	117
Figura 3.5: Volatilidades estimadas e i.c 95% para os retornos no período da implementação do <i>Plano Collor I</i> .....	118
Figura 3.6: Gráfico de Probabilidade Normal dos resíduos padronizado EGARCH(2,2) nas <i>Ações</i> .....	119
Figura 3.7: Comparação das volatilidades estimadas GARCH(1,1) e GARCH(1,2) nas <i>Ações</i> .....	120
Figura 3.8: Volatilidade estimada GARCH(1,2) nas <i>Ações</i> .....	121
Figura 3.9: Volatilidade estimada EGARCH(2,2) nas <i>Ações</i> .....	122
Figura 3.10: Comparação das volatilidades estimadas GARCH(1,2) e EGARCH(2,2) nas <i>Ações</i> .....	123
Figura 3.11: Gráfico das diferenças entre as volatilidades estimadas GARCH(1,2) e EGARCH(2,2) nas <i>Ações</i> .....	124
Figura 3.12: Volatilidades estimadas GARCH(2,2) e EGARCH(2,0) na <i>Taxa de Câmbio</i> ...	125
Figura 3.13: Funções de autocorrelação dos resíduos ARCH e dos quadrados dos resíduos na <i>Taxa de Câmbio</i> .....	126
Figura 3.14: Gráfico de Probabilidade Normal dos resíduos GARCH(2,2) na <i>Taxa de Câmbio</i> .....	127
Figura 3.15: Volatilidades estimadas ARCH e i.c 95% para os retornos no período da ocorrência do maior valor do retorno na <i>Taxa de Câmbio</i> .....	128
Figura 3.16: Volatilidades estimadas ARCH e i.c 95% para os retornos da <i>Taxa de Câmbio</i> no período de Fevereiro a Junho de 1986.....	129
Figura 3.17: Comparação das volatilidades estimadas GARCH(1,1), GARCH(2,2) e EGARCH(2,0) na <i>Taxa de Câmbio</i> .....	130
Figura 3.18: Volatilidades estimadas ARCH e i.c 95% para os retornos da <i>Taxa de Câmbio</i> no período de Fevereiro a Maio de 1985.....	131

#### Capítulo 4

Figura 4.1: Componentes estimados $T_t$ e $K_t$ nas <i>Ações</i> .....	176
Figura 4.2: Componentes estimados $h_{f_t}$ e $h_{s_t}$ nas <i>Ações</i> .....	177
Figura 4.3: Componente estimado irregular filtrado nas <i>Ações</i> .....	178

Figura 4.4: Volatilidades estimadas filtrada e suavizada nas <i>Ações</i> .....	179
Figura 4.5: Análise dos Quase-resíduos padronizados nas <i>Ações</i> .....	180
Figura 4.6: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período da <i>Queda da Ministra Zélia Cardoso de Mello</i> .....	181
Figura 4.7: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período da <i>Mudança da Banda Cambial</i> .....	182
Figura 4.8: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período do <i>Impeachment Collor</i> .....	183
Figura 4.9: Comparação das volatilidades estimadas filtrada e suavizada nas <i>Ações</i> .....	184
Figura 4.10: Gráfico do componente $K_t$ na <i>Taxa de Câmbio</i> .....	185
Figura 4.11: Volatilidades estimadas suavizada e filtrada na <i>Taxa de Câmbio</i> .....	186
Figura 4.12: Análise dos quase-resíduos padronizados na <i>Taxa de Câmbio</i> .....	187
Figura 4.13: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período da ocorrência do maior retorno na <i>Taxa de Câmbio</i> .....	188
Figura 4.14: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período de Março a Julho de 1986 na <i>Taxa de Câmbio</i> .....	189
Figura 4.15 Comparação das volatilidades estimadas suavizada e filtrada na <i>Taxa de Câmbio</i> .....	190
Figura 4.16: Volatilidades estimadas AR(1)-SV no período do maior conglomerado de retornos grandes na <i>Taxa de Câmbio</i> .....	191

## Capítulo 5

Figura 5.1: Volatilidades estimadas GARCH(1,2) e AR(1)-SV nas <i>Ações</i> .....	219
Figura 5.2: Volatilidades estimadas EGARCH(2,2) e AR(1)-SV nas <i>Ações</i> .....	220
Figura 5.3: Volatilidades estimadas GARCH(2,2) e AR(1)-SV na <i>Taxa de Câmbio</i> .....	221
Figura 5.4: Volatilidades estimadas EGARCH(2,0) e AR(1)-SV na <i>Taxa de Câmbio</i> .....	222
Figura 5.5: Comparação das volatilidades estimadas ARCH e Variância Estocástica nas <i>Ações</i> .....	223
Figura 5.6: Comparação das volatilidades estimadas ARCH e Variância Estocástica na <i>Taxa de Câmbio</i> .....	224
Figura 5.7: Influência das observações extremas dos retornos nas volatilidades estimadas ARCH e Variância Estocástica.....	225
Figura 5.8: Volatilidades estimadas ARCH e Variância Estocástica em dois períodos da <i>Taxa de Câmbio</i> .....	226

Figura 5.9: Comparação dos intervalos de confiança 95% para os retornos obtidos pelos modelos ARCH e Variância Estocástica nas <i>Ações</i> .....	227
Figura 5.10: Comparação dos intervalos de confiança 95% para os retornos obtidos obtidos pelos modelos ARCH e Variância Estocástica na <i>Taxa de Câmbio</i> .....	228
Figura 5.11: Regiões de equivalência dos modelos GARCH(1,1) e AR(1)-SV estacionário sob alta persistência.....	229
Figura 5.12: Comparação da volatilidade GARCH(1,1) (verdadeira) e a volatilidade estimada AR(1)-SV.....	230
Figura 5.13: Comparação das volatilidades geradas e estimadas quando existem modelos equivalentes.....	231
Figura 5.14: i.c 95% para os retornos em um período de Sim 1.....	232
Figura 5.15: volatilidades GARCH(1,1) e AR(1)-SV em um período de Sim 1.....	232
Figura 5.16: Volatilidades GARCH(1,1) e AR(1)-SV em um período de Sim2a.....	233
Figura 5.17 : i.c 95% para os retornos em um período de Sim2a.....	233
Figura 5.18: Comparação das volatilidades geradas e estimadas quando não existem modelos equivalente .....	234

## RESUMO

Entre os modelos utilizados para modelar a volatilidade apresentada pelos retornos das séries financeiras, dois dos mais estudados na literatura são os modelos *ARCH* e de *Variância Estocástica*. Ambos dos modelos são capazes de reproduzir teoricamente as características empíricas observadas nos retornos das séries financeiras, mas fazem distinção quanto ao tratamento da volatilidade. Nos modelos *ARCH* considerada como observável e nos modelos de *Variância Estocástica* como não-observável. Por sua vez, nos modelos *ARCH*, um dos modelos mais utilizados é o *GARCH(1,1)*, o qual apresenta em alguns casos, características semelhantes ao modelo de *Variância Estocástica AR(1)-SV* estacionário. O objetivo básico desta dissertação é fazer uma comparação dos modelos *ARCH* e de *Variância Estocástica* em termos das estimativas de volatilidade e da previsão da volatilidade um-passo à frente. No Capítulo 1 são apresentadas as idéias básicas e a nomenclatura. O Capítulo 2 está dedicado ao estudo das características empíricas das séries de retornos dos ativos da TELEBRÁS e da taxa de câmbio marco alemão com relação ao dólar norte-americano, que serão analisadas nesta dissertação, dando ênfase à não-linearidade apresentada pelos retornos. Nos capítulos 3 e 4 são apresentados os modelos *ARCH* e *Variância Estocástica*, suas principais propriedades e métodos para estimação, teste de hipóteses, diagnóstico e previsão. Finalmente, no Capítulo 5 realizamos a comparação dos modelos *ARCH* e *Variância Estocástica* utilizando as séries financeiras mencionadas e, adicionalmente, para os modelos *GARCH(1,1)* e *AR(1)-SV* através de simulações.

# CAPÍTULO 1

## INTRODUÇÃO

### 1. 1 Séries Financeiras

Ações, taxas de câmbio, taxas de juros, etc, são negociadas nos mercados financeiros, como por exemplo as Bolsas de Valores e a sequência dos valores negociados de cada um desses “produtos” forma uma *Série Financeira*. Na tentativa de modelar essas séries com as metodologias tradicionais de séries temporais lineares, inúmeros estudos têm mostrado que esses modelos não são capazes de reproduzir as características peculiares que estas séries apresentam. Essas características são conhecidas como *Fatos Estilizados (Stylized Facts)*. Nas últimas duas décadas houve uma mudança significativa nessa situação, com o desenvolvimento de *modelos não-lineares não-Gaussianos*, viabilizados através de um aumento considerável da capacidade computacional disponível com o avanço da informática.

Os dois objetivos principais no estudo das séries financeiras são *interpretação e previsão*. Há uma grande preocupação em entender a evolução das séries, isto é, entender o comportamento dos preços. A tradução estatística desse problema é a *identificação, estimação e teste de hipóteses*. Por outro lado, e talvez seja o aspecto de maior interesse para os investidores, queremos prever o valor da série ou conhecer as características da distribuição do estimador no futuro do processo. Esses aspectos estão relacionados com a *previsão*. Os modelos propostos terão que satisfazer dois requisitos, reproduzir os *Fatos Estilizados* e fornecer boas previsões.

Os preços dos produtos negociados nos mercados financeiros são determinados pelas condições de mercado (oferta e demanda). Os dias em que o mercado permanece aberto são chamados de *trading days* e os restantes de *non trading days*. No mês temos usualmente 22 dias de funcionamento do mercado e no ano 252 dias. Geralmente, os modelos teóricos (de séries temporais ou processos de difusão) ao analisar a evolução diária das séries financeiras, consideram como indexador do tempo os dias em que existe movimento no mercado, não considerando como observações faltantes (*missing*) os dias sem movimento.

Existem várias teorias que propõem modelos que têm como objetivo determinar o valor do *preço justo* ou *valor intrínseco* dos produtos e cada uma delas é desenvolvida precisando considerar o tipo de produto, sendo conhecidas, para as opções, como *Teorias de Precificação de Opções*. Na implementação desses modelos usualmente é fundamental conhecer o valor de uma variável que é *não observável*, a **volatilidade**. Informalmente podemos definir a volatilidade como o grau de variabilidade que apresentam as variações dos preços (retornos), o que é interpretada como uma *medida do risco* associado a cada produto. Por exemplo, na Figura 1.1a é apresentado a evolução dos preços de fechamento<sup>1</sup> das ações da TELEBRÁS PN em Dólar norte-americano, no qual é observado que a variabilidade dos preços evolui no tempo.

Os modelos que se ocupam da medição da volatilidade são conhecidos como *Modelos de Volatilidade*, sendo os dois modelos estatísticos mais utilizados para estimá-los: os modelos *Autoregressivos de Variância Condicional Heteroscedástica (ARCH)* e suas extensões e os **Modelos de Variância Estocástica**. Dentro dos modelos do tipo ARCH, o mais utilizado nas aplicações empíricas é o GARCH (1,1) e, dentro dos Modelos de Variância Estocástica o AR(1)-SV.

Este trabalho tem como primeiro objetivo discutir as propriedades teóricas mais importantes dos modelos ARCH e de Variância Estocástica mostrando que são capazes de reproduzir o comportamento empírico típico apresentado por certas séries financeiras. O segundo objetivo é fazer uma comparação dos modelos em termos das estimativas da volatilidade e da previsão da volatilidade um-passo à frente. Estas comparações serão realizadas utilizando duas séries financeiras, as quais são nomeadas como *Séries Exemplo* e, adicionalmente, para os modelos GARCH(1,1) e AR(1)-SV através de simulações.

## 1.2 Volatilidade

A volatilidade é uma variável fundamental para várias fórmulas de Precificação de Opções. A grande maioria desses modelos é formulado em tempo contínuo e frequentemente os modelos propostos assumem que os preços das opções podem ser descritos por *processos de difusão*<sup>2</sup>. A fórmula mais conhecida e estudada é a proposta por Black e Scholes (1973), a

<sup>1</sup>Preço de fechamento (abertura) é o preço que tem um produto quando o mercado fecha (abre).

<sup>2</sup>O primeiro antecedente conhecido de utilização de processos de difusão é o trabalho de Bachelier (1900).

qual permite calcular o *preço justo* do prêmio para a compra de uma opção de compra do tipo europeia (C) sem dividendos, com taxa de juros constante diária (J). Uma opção de compra europeia é um contrato que dá à pessoa que o possui, o direito de comprar uma quantidade específica de um título-objeto (ações, ouro, dólar, etc) a um preço específico, chamado valor de exercício (VE), numa data determinada, chamada data de exercício. Nessa fórmula assume-se que a dinâmica dos preços P(t) pode ser modelada através da equação diferencial estocástica:

$$dP = \alpha P dt + \sigma P d\omega ,$$

onde  $\alpha$  é o retorno esperado instantâneo do ativo e  $\sigma$  é o *desvio padrão instantâneo por unidade de tempo ou volatilidade*, a qual é especificada como constante. A solução da equação sob determinadas condições, que não serão dadas aqui, é:

$$C = P\Phi(\delta) - \frac{VE}{(1+J)^t} \Phi(\delta - \sigma\sqrt{t}) , \quad \delta = \frac{\ln\left(\frac{P(1+J)^c}{VE}\right)}{\sigma\sqrt{t}} + \frac{\sigma\sqrt{t}}{2} , \quad (1.2.1)$$

onde  $t$  é o tempo a decorrer até a data do exercício (em dias de funcionamento do mercado) e  $\Phi(.)$  é a função acumulada normal. Na expressão (1.2.1) todas as variáveis são conhecidas com exceção da volatilidade, daí o interesse em estimar essa variável. Na prática, é bastante comum os investidores utilizarem um procedimento empírico simples para prever a volatilidade como o desvio padrão das últimas 22 observações (aproximadamente um mês de funcionamento do mercado).

Por outro lado, invertendo a fórmula de Black e Scholes, isto é, expressando a volatilidade como função das outras variáveis<sup>3</sup>, é possível determinar o valor da volatilidade passada uma vez conhecidos os valores dos preços. Essa volatilidade é conhecida como *volatilidade implícita* e é considerada pelos investidores como uma estimativa adequada da volatilidade.

---

Esse trabalho dá um tratamento matemático aos processos de difusão e introduziu o conceito de movimento browniano, antes dos trabalhos de Einstein e Wiener.

<sup>3</sup>Mediante aproximações.

A fórmula de Black e Scholes (1.2.1) serve para ilustrar dois aspectos na modelagem dos preços; o primeiro, é que ela, assim como a grande maioria de modelos teóricos, é formulada a tempo contínuo. Em segundo lugar, a volatilidade é suposta constante no tempo em que é negociada a opção. Entretanto, diversos estudos tem mostrado que essa hipótese não é adequada sendo conveniente a especificação de modelos em que a volatilidade evolui ao longo do tempo. Apesar disto, segundo Ghysels *et alii* (1995), a fórmula (1.2.1) ainda é a mais utilizada pelos investidores<sup>4</sup>. Os modelos que consideram a evolução da volatilidade são conhecidos como modelos de *Volatilidade Estocástica* e na Seção 1.4 serão apresentadas duas classes desses modelos.

### 1.3 Retornos e *Fatos Estilizados*

As séries financeiras possuem a característica que as cotações ou preços apresentam médias e variabilidade que mudam consideravelmente ao longo do tempo. Isso pode ser constatado na Figura 1.1 e, com maior detalhe, na Figura 1.2, para as séries de ações da TELEBRÁS PN em Dólar norte-americano e taxa de câmbio Marco Alemão/Dólar norte-americano. Adicionalmente, preços consecutivos são altamente correlacionados. O tratamento estatístico das séries é facilitado quando analisamos uma função da variação do preço de dois dias consecutivos. Essas quantidades são conhecidas como *retornos* e podem ser definidos de várias formas, sendo a mais utilizada os *retornos compostos*. Sejam  $\{P_t\}_{t=1, \dots, T}$  preços de fechamento diários, os retornos compostos *diários*<sup>5</sup> são definidos como:

$$Y_t = \Delta \log(P_t) = \log(P_t) - \log(P_{t-1}), \quad (1.3.1)$$

sendo  $\log(\cdot)$  o logaritmo neperiano. Os objetivos de considerar essa função são tornar a série estacionária, mas pode ser utilizada qualquer outra da família Box-Cox<sup>6</sup> e, por outro lado, porque representa aproximadamente a variação percentual do preço, isto é,  $\Delta \log(P_t) \cong \frac{P_t - P_{t-1}}{P_{t-1}}$ . Para definir o retorno de um dia após um período de dias nos quais o

mercado não trabalha, utilizamos o último preço disponível. Por exemplo, o retorno da terça feira 2 de Janeiro é calculado utilizando o preço da sexta feira 29 de Dezembro.

<sup>4</sup>Recentemente tem sido desenvolvidas versões da fórmula BS incorporando a volatilidade como evoluindo no tempo, por exemplo em Hull e White (1987).

<sup>5</sup> Podem ser utilizados preços de abertura e/ ou preços mensais, em cujo caso teríamos retornos mensais.

<sup>6</sup>Ver por exemplo Cao e Tsay (1993).

Na Figura 1.3 apresentamos o gráfico dos retornos da série das ações da TELEBRÁS PN em Dólar norte-americano e da taxa de câmbio Marco Alemão com relação ao Dólar norte-americano, sendo observado que apresentam um nível médio constante, próximo a zero e que existem várias situações com valores muito altos ou muito baixos.

Outra característica importante observada e que define em geral as séries financeiras é que normalmente os retornos são não correlacionados. Supondo que o valor esperado dos retornos é constante, essas duas características compõem a denominada *Hipótese de Passeio ao Acaso* (Random Walk Hypothesis). Nesse sentido, um modelo para os preços é o modelo *martingale*<sup>7</sup>, no qual  $E[P_{t+1} / I_t] = P_t$ , onde  $I_t$  representa a informação<sup>8</sup> disponível até o instante  $t$  que inclui o conhecimento dos preços passados  $P_{t-j}$ ,  $j > 0$ . Assim, neste modelo as previsões dos preços não podem ser melhoradas utilizando a informação dos preços anteriores.

A observação empírica da hipótese do Passeio ao Acaso serve de base, dentro da Teoria Financeira Econômica, para justificar a *Hipótese de Mercado Eficiente* (Efficient Market Hypothesis) segundo a qual seja qual for a forma de utilizar a informação passada dos preços é *impossível construir estratégias de investimento com esperança de obter lucro maior que o custo de oportunidade*.

Admitindo-se que os retornos são não correlacionados e se a hipótese de normalidade dos retornos fosse correta então os retornos seriam *independentes*. Nesse sentido a Hipótese de Passeio ao Acaso muitas vezes é entendida como que os retornos são independentes identicamente distribuídos. No entanto, como mostram diversos estudos, a cauda da distribuição dos retornos é mais pesada que a normal. Essa característica foi formulada por Mandelbrot (1963) como sendo a *Hipótese de Pareto* (Paretian Hypothesis). Ele foi também o primeiro em observar que nos retornos (1963, p.418):

*“...large changes tend to be followed by large changes-of either sign- and small changes tend to be followed by small changes...”*

o que é conhecido como *Conglomerados de Volatilidade* (Volatility Clustering). Ambas as características, cauda pesada e conglomerados de volatilidade fazem parte do que é conhecido

<sup>7</sup> Outro modelo é o modelo *submartingale*.

<sup>8</sup> Em Fama (1970) são discutidos os tipos de informação.

como os *Fatos Estilizados*. Modelos de volatilidade satisfatórios têm que ser capazes de reproduzir essas características. A seguir apresentamos duas classes de modelos.

#### 1.4 Modelos de Volatilidade abordados nesta dissertação

Existem duas formas de abordar a modelagem da volatilidade. Na primeira é assumido um modelo contínuo e realiza-se a estimação via aproximação discreta. Nessa abordagem existem muitas dificuldades e é pouco explorada literatura<sup>9</sup>. Na segunda, o modelo é especificado em tempo discreto, usualmente como a versão discretizada de modelos a tempo contínuo. É nesta abordagem que encontramos os modelos estudados nesta dissertação.

Os Modelos de Volatilidade Estocástica a Tempo Discreto podem ser classificados dentro de duas abordagens<sup>10</sup>, dependendo da natureza da volatilidade, seja como variável observável ou como variável não observável. Uma variável é (não é) observável se os valores dela são (não são) conhecidos quando os parâmetros do modelo são conhecidos. Os modelos estatísticos que consideram a volatilidade como observável, recorrem ao procedimento *ad-hoc* de expressá-la como função das observações (retornos). Por exemplo, nos modelos *Autoregressivos de Variância Condicional Heteroscedástica*; **ARCH**, propostos por Engle (1982) temos que:

$$Y_t / \varepsilon_t \approx N(0, \sigma_t^2)$$

$$\sigma_t^2 = f(Y_{t-1}, Y_{t-2}, \dots),$$

onde  $\{\varepsilon_t\}$  é um processo de inovações, “ / ” significa distribuição condicional e  $f$  significa função. Definindo  $I_t$  como a informação disponível até o instante  $t$ , isto é,  $I_t = \{Y_t, Y_{t-1}, \dots\}$ , temos que:

$$Y_t / I_{t-1} \approx N(0, \sigma_t^2).$$

Assim, neste caso a volatilidade é construída como a previsão um-passo à frente da variância.

---

<sup>9</sup>Ver por exemplo Shephard (1994b) e Ghysels *et alii* (1995) para referências.

<sup>10</sup>Recentemente Andersen (1994) propôs uma abordagem mais geral que inclui os dois tipos de modelos tratados nesta dissertação. Para uma breve ilustração das propriedades ver Ghysels *et alii* (1995).

As vantagens da utilização dos modelos construídos como previsão um-passo à frente têm sido assinaladas por Shephard (1994b):

- i) Do ponto de vista estatístico, é simples obter a verossimilhança via a decomposição do erro de previsão (*prediction error decomposition*) o qual facilita o processo de inferência.
- ii) Do ponto de vista econômico, modelos baseados em momentos um-passo à frente com respeito à informação passada (pelos agentes econômicos) possuem um apelo muito grande.
- iii) Especificando a volatilidade como modelos ARIMA dispomos de toda a teoria a esse respeito, a qual está bastante desenvolvida.

A segunda abordagem é considerar a volatilidade como variável latente. Por exemplo, no modelo de **Variância Estocástica Log-normal** introduzido por Taylor (1980) e denominado AR(1)-SV:

$$Y_t / h_t \approx N(0, \sigma_t^2) \quad , \quad h_t = \ln(\sigma_t^2) \\ h_t = \delta + \gamma h_{t-1} + \eta_t \quad , \quad \eta_t \approx N(0, \sigma_\eta^2) \text{ i.i.d.}$$

Esse modelo pode ser colocado facilmente na forma de espaço de estados, em cujo caso dispomos de bastantes resultados (ver por exemplo West e Harrison, 1989 e Harvey, 1989). No entanto, essa representação não é Gaussiana e, nesse caso, a inferência é difícil devido ao fato das distribuições da previsão um-passo à frente das observações não serem normais e de mais de um passo serem calculadas apenas numericamente. Consequentemente, temos que recorrer a procedimentos que fornecem soluções aproximadas ou implementar métodos computacionais intensivos<sup>11</sup>. Não obstante as dificuldades, os modelos de Variância Estocástica apresentam a vantagem que suas propriedades são derivadas facilmente e a extensão para a análise multivariada é simples.

Como referido anteriormente, neste trabalho serão estudados os modelos ARCH, com ênfase no modelo GARCH(1,1) e o modelo de Variância Estocástica AR(1)-SV, o qual será estimado mediante o procedimento aproximado proposto por Harvey *et alii* (1994).

Existem vários Modelos de Precificação de Opções com volatilidade estocástica. Entre os trabalhos recentes temos os de Taylor (1986, Cap. 9), Hull e White (1987), Scott (1987),

---

<sup>11</sup> Entre as exceções temos Shephard (1994a) que propõe um modelo com soluções analíticas.

Stein e Stein (1991), Wiggins (1987), Melino e Turnbull (1990) e Heston (1993) utilizando Modelos de Variância Estocástica. Por sua vez, com modelos ARCH temos: Engle e Mustafa (1992), Engle *et alii* (1993a), Engle *et alii* (1993b), Noh *et alii* (1994) e Duan (1995).

### **1.5 Descrição das Séries Financeiras a serem estudadas: *Séries Exemplo*.**

As metodologias descritas neste trabalho serão ilustradas através da sua aplicação em duas séries financeiras, as quais nomeamos como *Séries Exemplo*. Com o objetivo de enriquecer os resultados, são escolhidos séries de tipos diferentes; uma de ações e outra de taxa de câmbio.

A primeira série é a correspondente aos preços de fechamento das ações da empresa TELEBRÁS PN em Dólar norte-americano no período de 3 de Janeiro de 1989 até 17 de Julho de 1995 (1599 retornos). Nesse período foram implementados vários Planos de Estabilização na economia e aconteceram vários episódios políticos que repercutiram no comportamento da série. Temos, por exemplo, em Março de 1990 o *Plano Collor I*, em Fevereiro de 1991 o *Plano Collor II*, a *Queda da Ministra Zélia Cardoso de Mello* em Janeiro de 1992, o início do processo do *Impeachment Collor* em Junho de 1992 e, em Março de 1995, a *Mudança da Banda Cambial* devida à crise do México. Tais efeitos são claramente observados na Figura 1.3 e correspondem aos valores grandes (em valor absoluto) dos retornos. Nesta série, durante o período estudado não observamos presença de tendência dominante crescente porém subidas e descidas (ver Figura 1.1). A série foi estudada em Herencia *et alii* (1995).

A série de taxa de câmbio utilizada é a cotação do Marco Alemão em relação ao Dólar norte-americano no mercado futuro<sup>12</sup>, no período de 3 de Janeiro de 1984 a 31 de Dezembro de 1990 (1767 retornos). Essa série é talvez a mais estudada do seu tipo na literatura (considerando períodos e negociações diferentes). Observando a Figura 1.1 percebemos que existe uma tendência geral de aumento das cotações sendo identificadas duas etapas de crescimento. A primeira delas começa ao redor de maio de 1985 (época na qual o Bundesbank implementou medidas para desregular o mercado financeiro) e termina no dia 19 de Outubro de 1987 (*Quebra* da Bolsa de New York). A segunda etapa de crescimento começa no final de

<sup>12</sup>Apesar de que no mercado futuro os retornos são calculados seguindo um procedimento diferente ao apresentado na Seção 1.3, nesta dissertação e com propósitos ilustrativos, utilizamos a fórmula (1.3.1).

1989. Esta série de taxa de câmbio tem sido analisada em vários estudos, entre eles Harvey e Shephard (1993b), Harvey *et alii* (1994), Ruiz (1994), Taylor (1994) com modelos de Variância Estocástica e, utilizando modelos ARCH temos: Hsieh (1991a), Bollerslev e Engle (1993), Taylor (1994), Bollerslev *et alii* (1994), Shephard (1994b), Liu e Brorsen (1995) e Bollerslev e Ghysels (1996).

Na tabela a seguir é apresentada uma breve descrição das características da cada uma das *Séries Exemplo*:

**Tabela 1.1**  
**Descrição das *Séries Exemplo***

Série	Tipo de Série	Mercado	Início	Final	Retornos Diários
TELEBRÁS	Ações	BOVESPA	03/01/89	17/07/95	1599
DM/US\$	Taxa de Câmbio	Internacional	03/01/84	31/12/90	1767

### 1.6 Nomenclatura e Definições Básicas.

Nesta dissertação o processo de retornos é denotado por  $\{Y_t\}$  e as amostras (de tamanho T) por:  $\{y_1, y_2, \dots, y_T\}$ . Por outro lado, nas aplicações empíricas com as *Séries Exemplo*, trabalhamos com os retornos compostos diários, calculados através da fórmula (1.3.1) e chamados simplesmente de retornos. Da mesma forma, a série de retornos das ações da TELEBRÁS será referenciada simplesmente como a série das *Ações* enquanto que a série de retornos correspondente a taxa de câmbio DM/US \$ é nomeada como de *Taxa de Câmbio*.

Nos modelos de volatilidade, normalmente a volatilidade é o desvio padrão condicional do processo dos retornos. Neste trabalho consideramos a volatilidade como a **variância condicional do processo** e ambos os termos; volatilidade e variância condicional, serão usados indistintamente.

A seguir apresentamos alguns conceitos que serão de utilidade na compreensão do trabalho.

### *Processo Covariante Estacionário*

O processo  $\{Y_t\}$  é covariante estacionário (de segunda ordem) ou simplesmente estacionário, se possui média constante, variância finita e as autocovariâncias independem do tempo, mas apenas da defasagem. Matematicamente:

$$i) E[|Y_t|^2] < \infty \quad \forall t \in Z$$

$$ii) E[Y_t] = \mu \quad \forall t \in Z$$

$$iii) \phi_{Y(r,s)} = \text{Cov}[Y_r, Y_s] = \text{Cov}[Y_{t+r}, Y_{t+s}] = \phi_{Y(r+t, s+t)} \quad \forall r, s, t \in Z$$

ou equivalentemente,  $\phi_Y(\tau) = \phi_Y(\tau, 0) \quad \tau = r - s$ .

### *Processo Estritamente Estacionário*

Se as distribuições conjuntas de  $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k})$  e  $(Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_k+h})$  são idênticas  $\forall t_1, t_2, \dots, t_k$ , e  $h \in Z$ , então o processo  $\{Y_t\}$  é estritamente estacionário.

Todo processo estritamente estacionário que possui segundo momento finito é covariante estacionário, sendo o contrário falso, a menos que a distribuição seja normal.

### *Ruído Branco (WN)*

Dizemos que o processo  $\{Y_t\}$  é ruído branco se possui média constante;  $E[Y_t] = \mu$ , e as autocovariâncias são nulas;  $\phi_{Y(r,s)} = \text{Cov}[Y_r, Y_s] = 0 \quad \forall r, s \in Z$ .

Para que o processo seja covariante estacionário a variância tem que ser finita e constante,  $E[|Y_t|^2] = \sigma^2 < \infty \quad \forall t \in Z$ .

### *Ruído Branco Estrito (SWN)*

Se em um ruído branco  $\{Y_t\}$  as variáveis são independentes então ele é dito ser ruído branco estrito. Por exemplo, as variáveis independentes e identicamente distribuídas (*i.i.d*).

### *Processos Martingale Diferença (MD)*

O processo  $\{Y_t\}$  é martingale diferença com respeito a uma seqüência de  $\sigma$ -álgebras  $D_1 \subset D_2 \subset \dots \subset D_t \subset \dots$  se satisfaz:

- i)  $Y_t$  é uma variável aleatória com respeito à  $D_t$ .
- ii)  $E[|Y_t|] < \infty \quad \forall t \in Z$
- iii)  $E[Y_t / D_{t-1}] = 0 \quad \forall t \in Z$ .

Como consequência desta definição:  $E[Y_t] = E[E[Y_t / D_{t-1}]] = 0$  e, além disso,  $E[Y_t Y_{t-\tau}] = E[E[Y_t Y_{t-\tau} / D_{t-1}]] = E[Y_{t-\tau} E[Y_t / D_{t-1}]] = 0, \tau > 0$ . Portanto, os processos MD são ruído branco.

### *Processo de Inovações*

Se o processo  $\{Y_t\}$  é martingale diferença com respeito à sigma álgebra  $D_t$  gerada por  $X_t, X_{t-1}, \dots, X_0$  onde  $X_t = \sum_{i=0}^t Y_i$  e satisfaz:

- i)  $E[|Y_t|^2] < \infty \quad \forall t \in Z$ ,
- ii)  $E[Y_t Y_{t-\tau}] = 0 \quad \tau > 0$ ,

então é um processo de inovações. Por exemplo, o processo definido como  $Y_t = X_t - E[X_t / D_{t-1}]$  é um processo de inovações e, se  $E[|Y_t|^2] < \infty$  então ele é ruído branco covariante estacionário.

### *Processo Linear*

O processo  $\{Y_t\}$  é dito linear, se pode ser expressado na seguinte forma:  $Y_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$ , onde o processo  $\{\varepsilon_t\}$  é ruído branco estrito e  $\psi_i, i=0,1,2,\dots$  uma seqüência de constantes tal que  $\sum_{i=0}^{\infty} \psi_i^2 < \infty$ .

## **1.7 Breve Esboço da Dissertação.**

Este trabalho está estruturado em 5 Capítulos, sendo o primeiro dedicado a apresentar o contexto no qual são estudadas os modelos de séries financeiras, assim como as idéias básicas, objetivos e nomenclatura básica.

O Capítulo 2 está dedicado ao estudo das características empíricas das séries financeiras (denominados *Fatos Estilizados*), através das *Séries Exemplo*. O objetivo é ressaltar as principais características a serem incorporadas nos modelos teóricos apresentados posteriormente, como a *não-linearidade*. Por sua vez enfatizamos porque não é correto utilizar os modelos tradicionais de séries temporais.

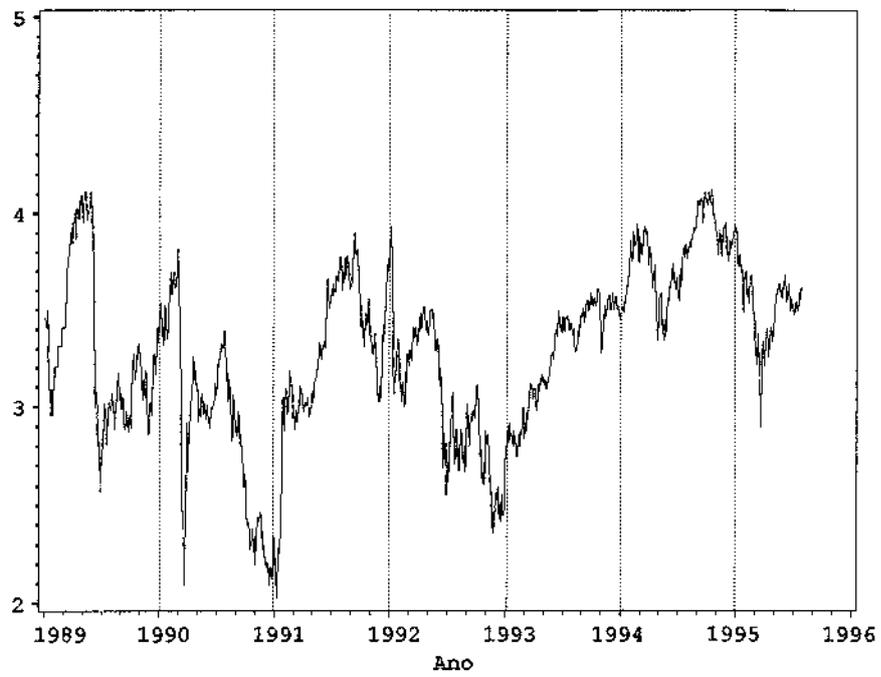
No Capítulo 3 são estudados com algum detalhe os *Modelos ARCH* e suas extensões. Começamos mostrando como esses modelos são capazes de reproduzir teoricamente alguns *Fatos Estilizados* apresentados pelas séries financeiras. A seguir discutimos o processo de estimação, testes de hipóteses, diagnóstico e previsão do modelo. Posteriormente, ajustamos esses modelos às *Séries Exemplo*.

Uma outra família de modelos é apresentada no Capítulo 4; os *Modelos de Variância Estocástica*. Mostramos que são capazes de reproduzir teoricamente os *Fatos Estilizados* e ajustamos esses modelos às *Séries Exemplo*. Para a estimação da volatilidade é utilizada a *Quase-Verossimilhança* como proposto por Harvey *et alii* (1994). Dentro dessa abordagem é discutido o diagnóstico e a análise do modelo, assim como a previsão.

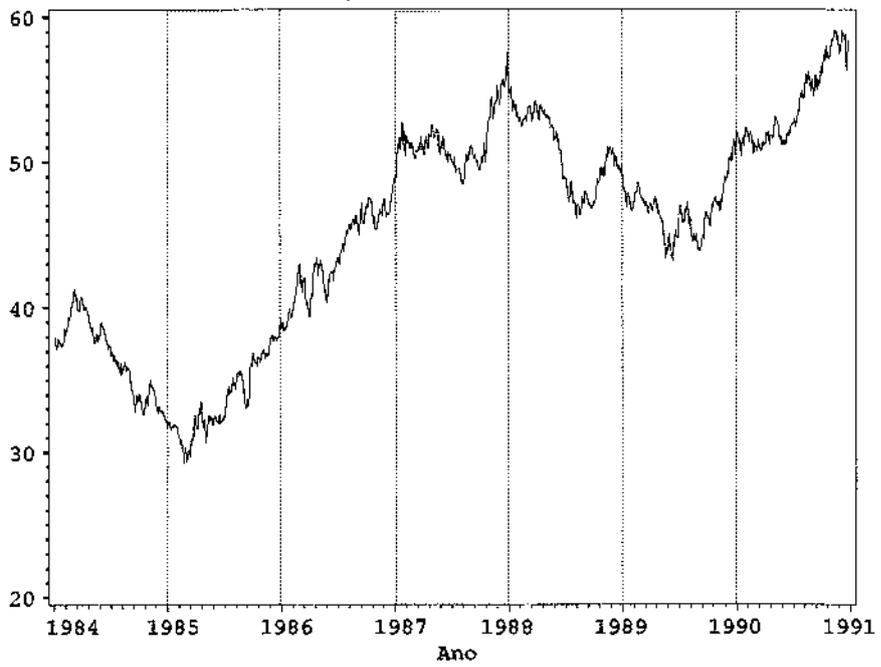
No Capítulo 5 é realizada a comparação dos modelos modelos ARCH e de Variância Estocástica, com ênfase nos modelos GARCH(1,1) e AR(1)-SV estacionário, através do ajuste das *Séries Exemplo*. Os aspectos considerados na comparação são estimação da volatilidade e a previsão um-passo à frente. A seguir, apresentamos a comparação dos modelos GARCH(1,1) e AR(1)-SV através de simulações. Considerando o número reduzido de simulações esses resultados devem ser tomados como ilustrativos. Finalmente, na Seção 5.5 são apresentados alguns comentários e as conclusões finais da comparação dos modelos de volatilidade estudados nesta dissertação.

As referências sobre modelos de volatilidade em séries financeiras são numerosas. Resumos atualizados dos modelos trabalhados nesta dissertação e sobre outros são: Bollerslev *et alii* (1992, 1994) , Shephard (1994b) e Ghysels *et alii* (1995). Nesses estudos podem ser encontradas também referências sobre aspectos que, por delimitação do trabalho, não são tratados nesta dissertação; como a sazonalidade, agregação (dentro da abordagem paramétrica), a modelagem não paramétrica e a abordagem multivariada.

a) *Ações*

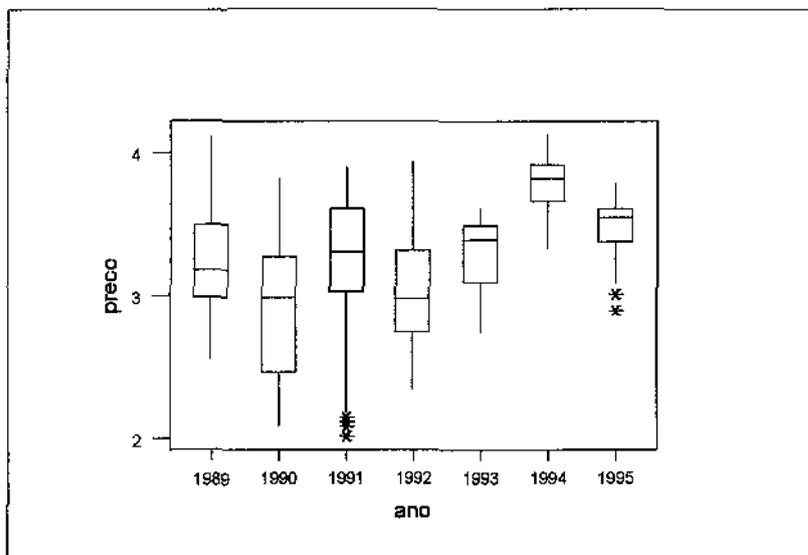


b) *Taxa de Câmbio*



**Figura 1.1: Gráfico dos preços de fechamento das *Séries Exemplo***

a) *Ações*



b) *Taxa de Câmbio*

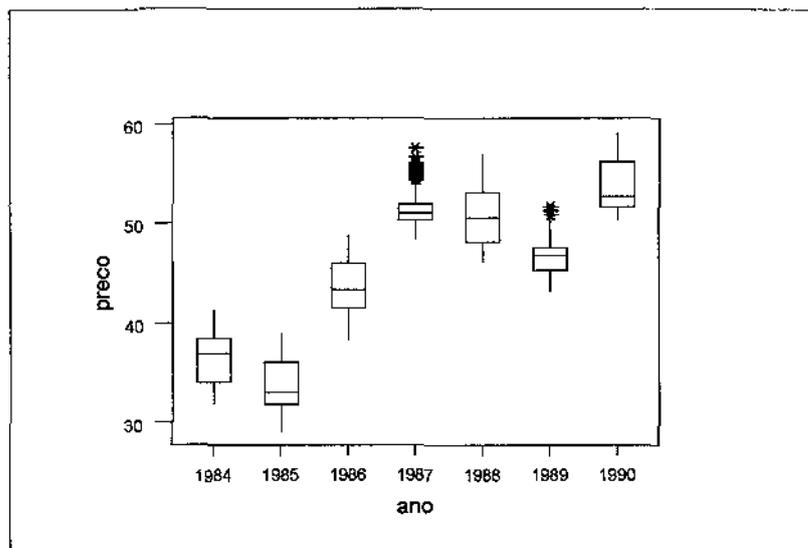
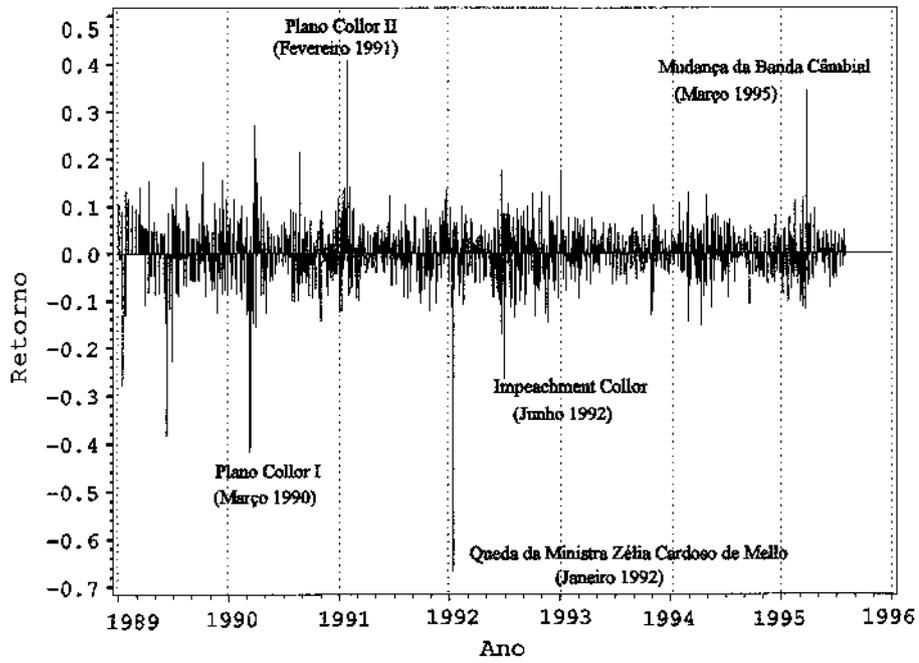


Figura 1.2: Box-plot dos preços de fechamento das *Séries Exemplo*

a) *Ações*



b) *Taxa de Câmbio*

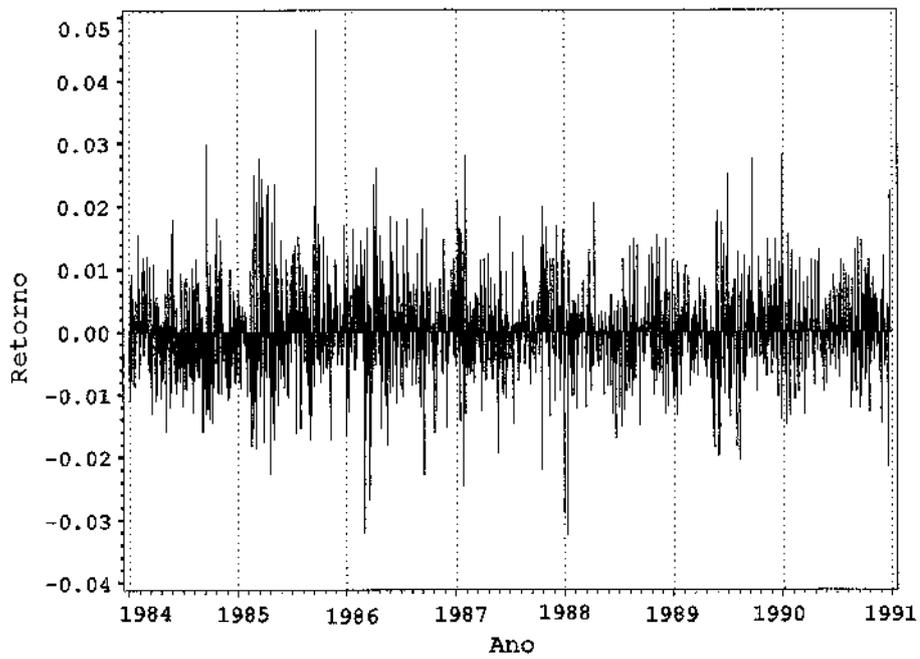


Figura 1.3: Gráfico dos retornos das *Séries Exemplo*

## CAPÍTULO 2

# CARACTERÍSTICAS EMPÍRICAS DAS SÉRIES FINANCEIRAS: OS *FATOS ESTILIZADOS*

### 2.1 Introdução

Durante muito tempo os pesquisadores de finanças trabalharam com a hipótese básica que *os retornos são Gaussianos independentes e identicamente distribuídos*. Tal característica era incorporada em teorias tal como a do Mercado Eficiente e um modelo estatístico para os preços era simplesmente o passeio ao acaso (Random Walk). A consequência básica dessa abordagem, do ponto de vista da estatística, era a seguinte: a última observação constituía a melhor previsão para o preço futuro. Uma revisão detalhada das implicações teóricas da hipótese de mercado eficiente e das evidências empíricas até o ano 1970 pode ser encontrada em Fama (1970).

O panorama mudou na década de 1960 através de trabalhos isolados como os de Mandelbrot (1963) e Fama (1965). O estudo pioneiro de Benoit Mandelbrot teve uma importância crucial para a modelagem das séries financeiras introduzindo uma abordagem original e rejeitando a hipótese de normalidade por dois motivos. Primeiro, porque a distribuição dos retornos apresentava cauda pesada e em segundo lugar pela existência de um comportamento peculiar nas observações: os *Conglomerados de Volatilidade*. Ambas características empíricas e muitas outras próprias das séries financeiras compõem os *Fatos Estilizados*. Posteriormente, na década de setenta o volume de trabalhos incrementou consideravelmente, mas é a partir dos anos oitenta que houve uma explosão de estudos na literatura sobre as características empíricas das séries financeiras. Uma revisão seletiva desses estudos é encontrada, entre outros, em Fama (1970), Taylor (1986), Bollerslev *et alii* (1992), Pagan (1994), Shephard (1994b), Ghysels *et alii* (1995).

Explorar as características empíricas dos retornos é fundamental para a implementação dos modelos de volatilidade. Visando a isso, o objetivo do presente capítulo é discutir, através da análise das *Séries Exemplo*, algumas características empíricas que apresentam os retornos

das séries financeiras. A Seção 2.2 trata dos Conglomerados de Volatilidade. Nas seções 2.3 e 2.4 discutimos a simetria e a cauda pesada mostrando a necessidade de utilizar métodos robustos à influência de observações extremas. Essa discussão será útil para se ter uma idéia correta da forma funcional das distribuições dos retornos, que permite, por sua vez, na Seção 2.5, interpretar corretamente as medidas de locação e dispersão. A seguir, na Seção 2.6, discutimos brevemente a estacionariedade dos retornos e na Seção 2.7, apresentamos metodologias para desvendar a estrutura de autocorrelação dos retornos e a existência de dependência, concluindo pela não linearidade dos retornos. Na seção seguinte discutimos a normalidade, e na Seção 2.9 é apresentado um resumo de outras características dos retornos e sua volatilidade, vinculadas a sua natureza econômica. Finalmente, na Seção 2.10, apresentamos as conclusões. O desenvolvimento deste Capítulo está baseado fundamentalmente nos trabalhos de Pagan (1994), Taylor (1986) e Bollerslev *et alii* (1994).

## 2.2 Conglomerados de Volatilidade

O conceito de Conglomerados de Volatilidade foi introduzido na seção 1.3. A presença de períodos com variações pequenas e períodos com variações grandes (em valor absoluto) é um fenômeno sempre presente nos retornos, segundo consta na literatura revista. Em particular, nas *Séries Exemplo* (Figura 2.1) observamos a presença desses conglomerados, sendo esses mais evidentes para a série das *Ações*. Modelos tradicionais de séries temporais (lineares) são incapazes de gerar Conglomerados de Volatilidade. Nos Capítulos 3 e 4 mostramos como modelos não lineares com variância condicional heteroscedástica são capazes de reproduzir essa característica.

## 2.3 Simetria

Um recurso gráfico para avaliar a simetria é a utilização dos histogramas. Como observado nas Figuras 2.2a e 2.3a, existe uma forte evidência de que a distribuição dos retornos da *Taxa de Câmbio* é aproximadamente simétrica enquanto que a série das *Ações* apresenta certa assimetria negativa. No entanto, nos histogramas não podemos observar a grande quantidade de observações extremas, característica que salta à vista como mostrado posteriormente. Devido à presença dessas, a utilização do coeficiente de simetria ( $s$ ) como medida de simetria requer cuidado. Na Tabela 2.1 apresentamos as estimativas de  $s$  para as

*Séries Exemplo.* Os valores observados: -1,12, para as *Ações* e 0,29 para a *Taxa de Câmbio* estão próximos de zero, mas não temos forma de avaliá-los adequadamente através do desvio padrão da estimativa, por causa das observações extremas. Sob a hipótese que os retornos são independentes e identicamente distribuídos normais, intervalos de confiança de 95% (iguais a  $s + \sqrt{6/T}$ ) não incluem o valor de zero em nenhum dos dois casos, rejeitando a hipótese de simetria. Mas, ao retirar o 0,5% das observações extremas (em valor absoluto) obtemos resultados contrários nas *Ações* que evidenciam distribuições aproximadamente simétricas.

**Tabela 2.1**  
**Estatísticas dos retornos nas *Séries Exemplo***

Série	T	Média <sup>1</sup>	Mediana <sup>1</sup>	Trimédia <sup>1</sup>	Desvio Padrão <sup>2</sup>	Q <sub>3</sub> -Q <sub>1</sub> <sup>2</sup>	Coefficiente de simetria	Excesso de curtose
<i>Ações</i>	1599	1,56	0	5,57	60,79	62,66	-1,12	14,8
<i>Taxa de Câmbio</i>	1767	2,39	0	2,23	7,55	8,52	0,29	2,12

<sup>1</sup> (x10<sup>4</sup>)  
<sup>2</sup> (x10<sup>3</sup>)

Considerando o exposto, é aconselhável utilizar um procedimento gráfico robusto para testar a simetria da distribuição. Podemos utilizar, por exemplo, um *Gráfico de Simetria* que se baseia na comparação das distâncias dos quantis  $Q(\alpha)$  e  $Q(1-\alpha)$  com respeito à mediana. Sob a hipótese de simetria, as distâncias mencionadas possuem o mesmo valor. Dada uma amostra  $y_1, y_2, \dots, y_n$  podemos estimar essas distâncias da seguinte forma: sejam as estatísticas de ordem  $y_{(1)}, y_{(2)}, \dots, y_{(n)}$  e a mediana  $m$ , então  $u_i = m - y_{(i)}$ , é estimador de  $Q_{(i/n)}$  e  $v_i = y_{(n-i+1)} - m$  estimador de  $Q_{(1-i/n)}$ , onde  $i=1, 2, \dots, [T/2]$ <sup>1</sup>. Finalmente construi-se o gráfico dos  $n$  pares  $u_i$  e  $v_i$ . Dessa forma, se a distribuição dos valores amostrais for simétrica então os pontos ( $u_i$  e  $v_i$ ) estariam numa reta que passa pela origem com coeficiente angular igual a um.

Os Gráficos de Simetria para as *Séries Exemplo* são apresentados nas Figuras 2.2b e 2.3b, sendo que a linha continua corresponde à reta da situação ideal de simetria. Podemos dizer que, em geral, nas duas séries, a distribuição dos retornos é aproximadamente simétrica com exceção dos valores extremos (correspondentes as caudas das distribuições) que apresentam grandes desvios em relação à reta ideal de simetria. A proporção dessas

<sup>1</sup>  $[x]=n$  se e somente se  $n \leq x < n+1$

observações é pequena. Por exemplo, existem somente 1.1% dos pontos na Figura 2.2b que satisfazem  $u > 17$  e, na *Taxa de Câmbio* 1,9% dos pontos satisfazem  $u > 1,8$  (Figura 2.3b).

Na Tabela 2.2 apresentamos os resultados da regressão  $v = \alpha + \beta u + \varepsilon$ . Embora incorretos, calculamos os erros padrão das estimativas para ter uma idéia aproximada da precisão. Nas duas séries, a estatística F, assim como as estatísticas *t-Student* são significativas com valores-P inferiores à 0,0001.

**Tabela 2.2**  
**Simetria: estimativas da regressão  $v = \alpha + \beta u + \varepsilon$**

	<i>Ações</i>			<i>Taxa de Câmbio</i>		
	I	II	III	I	II	III
$\hat{\alpha}$	9,3x10 <sup>-3</sup> (3,9x10 <sup>-4</sup> ) <sup>a</sup>	7,2x10 <sup>-3</sup> (3,3x10 <sup>-4</sup> )	3,7x10 <sup>-3</sup> (1,4x10 <sup>-4</sup> )	-3,75x10 <sup>-4</sup> (2,7x10 <sup>-5</sup> )	-3,02x10 <sup>-4</sup> (1,96x10 <sup>-5</sup> )	-3,52x10 <sup>-4</sup> (9,4x10 <sup>-6</sup> )
$\hat{\beta}$	0,786 (0,006)	0,840 (0,003)	0,940 (0,003)	1,159 (0,0038)	1,143 (0,0028)	1,153 (0,0015)
R (adj)	0,9531	0,9659	0,9931	0,9904	0,9948	0,9986

I: considerando todas as observações.

II: retirando os dois maiores retornos (em valor absoluto) nas *Ações* e, na *Taxa de Câmbio*, o maior retorno.

III: retirando 1,1% dos pontos nas *Ações* e, na *Taxa de Câmbio*, 1,9% dos pontos

<sup>a</sup> Desvio padrão das estimativas

Ao utilizar a série completa (situação I), os resultados apontam que existe simetria ou uma pequena assimetria a direita, na *Taxa de Câmbio* sendo a estimativa de  $\beta$  maior do que um. Por sua vez, nas *Ações* a estimativa do  $\beta$  evidencia ausência de simetria. Mas, considerando que a regressão é sensível aos pontos de *alavanca*, recalculamos as retas de regressão deixando fora as observações extremas. Os resultados podem ser observados na tabela anterior. Assim, nas *Ações*, retirando os dois maiores retornos (II) percebemos que o ajuste da reta melhora sensivelmente e, retirando o 1,1% das observações obtemos estimativas que evidenciam uma pequena assimetria à esquerda. Por sua vez, na *Taxa de Câmbio*, ao considerar partes da amostra nas situações II e III, encontramos que as estimativas de regressão são muito similares.

Concluindo, podemos dizer que nas *Séries Exemplo* existe uma evidência muito forte de distribuições aproximadamente simétricas. Esses resultados coincidem com os encontrados na literatura, na qual são utilizados modelos com distribuições simétricas, que, por sua vez, possuem um apelo muito grande na Teoria de Finanças.

## 2.4 Cauda Pesada

Na subseção anterior já chamamos a atenção ao fato das *Séries Exemplo* apresentarem caudas mais pesadas que as correspondentes a uma distribuição normal. Esse comportamento é confirmado ao analisar os histogramas (Figuras 2.2a e 2.3a) e Gráficos de Probabilidade Normal na Figura 2.4.

Por outro lado, a presença de observações extremas (como possíveis *outliers* e/ou quebras) nas séries dos retornos, exige a utilização de procedimentos robustos para medir o peso das caudas. Um índice não paramétrico que pode ser utilizado é o *Índice de Cauda* (Hoaglin *et alii*, 1983 p.320) o qual é definido da seguinte forma:

$$\tau_D(F, \alpha) = \left( \frac{1}{C} \right) \frac{F^{-1}(1 - \alpha) - F^{-1}(0,5)}{F^{-1}(0,75) - F^{-1}(0,5)},$$

$$\tau_E(F, \alpha) = \left( \frac{1}{C} \right) \frac{F^{-1}(0,5) - F^{-1}(\alpha)}{F^{-1}(0,5) - F^{-1}(0,25)} \quad 0 < \alpha < 0,5.$$

A letras D (E) corresponde à cauda direita (esquerda) da distribuição acumulada dos retornos (F) e a constante C é exatamente igual à expressão fora do parêntesis quando a distribuição é normal padrão. Assim definido, esse índice é igual a um quando a distribuição é normal e exibe valores cada vez maiores na medida que a cauda da distribuição se torna mais pesada. Uma propriedade interessante desse índice é a invariância por locação e escala. A magnitude do efeito das observações extremas é controlada através de  $\alpha$ . Desta forma, o valor assumido pelo  $100\alpha\%$  percentil não influi o índice por maior que seja. Assim, nas *Séries Exemplo*, considerando  $\alpha = 0,0025$ , obtemos os seguintes resultados:

<i>Ações:</i>	$\tau_E = 2,34$	$\tau_D = 1,45,$
<i>Taxa de Câmbio:</i>	$\tau_E = 1,39$	$\tau_D = 1,56.$

Como podemos observar, na série *Taxa de Câmbio* o tamanho das caudas são próximos (simétricos). Por sua vez, na série das *Ações* a cauda esquerda é maior que a cauda direita, mas, ao aumentar o valor de  $\alpha$  para 0,005, obtemos que  $\tau_E = 1,43$  e  $\tau_D = 1,36$ . Isso significa que 99,5% das observações fazem com que o tamanho das caudas seja similar e poucos valores fazem com que a cauda esquerda seja grande. Com o intuito de obter uma idéia da correspondência desses números com os valores teóricos de algumas distribuições conhecidas temos que, para  $\alpha=0,0025$ , os valores do Índice de Cauda são:

*t-Student*, 6 graus de liberdade: 1,45,

*t-Student*, 10 graus de liberdade: 1,23,

Cauchy : maior que 8,4.

Em consequência, as caudas das *Séries Exemplo* são mais pesadas do que a da normal mas não tanto quanto as da distribuição de Cauchy.

Finalmente, observemos os valores do excesso de curtose na Tabela 2.1. A estimativa do excesso de curtose nas *Ações* é muito grande: 14,89. Utilizando esse valor como um indicador obteríamos uma noção bastante deturpada do tamanho da cauda, já que o excesso de curtose é extremamente sensível as observações extremas. Por sua vez, no caso da *Taxa de Câmbio* a estimativa é igual a 2,12, em consequência, para as duas séries concluímos que as distribuições são leptocúrticas.

Cauda pesada é uma característica presente em quase toda série financeira. Na literatura, alguns pesquisadores como Mandelbrot (1963) e Fama (1963,1965) encontraram que a cauda dos retornos da série do algodão e de ações de empresas norte-americanas, respectivamente, era extremamente pesada, parecida à da Cauchy. Nas *Séries Exemplo*, através da metodologia proposta, os resultados sugerem que mesmo ao retirar os valores extremos, os retornos apresentam cauda mais pesadas do que a normal, similares as da distribuição *t* com graus de liberdade entre 6 e 10.

## 2.5 Medidas de Locação e Variabilidade

Uma hipótese razoável que pode ser assumida ao observar os gráficos dos retornos na Figura 1.3, é que o nível médio da série é igual a zero. Na Tabela 2.1 apresentamos as estimativas da média, mediana e a Trimédia (calculada retirando 0,5% das observações extremas). Para as *Ações*, essas três estimativas são diferentes enquanto que para a *Taxa de Câmbio* a média e a Trimédia são similares. Todos esses valores são positivos e estão em torno de zero sugerindo que o parâmetro de locação é aproximadamente zero. Sob a hipótese incorreta que os retornos são independentes identicamente distribuídos normais<sup>2</sup>, intervalos de confiança a 95% incluem o valor zero nas duas séries. Não obstante, na prática é conveniente modelar os retornos como processos com média diferente de zero. Assim, tanto na série das *Ações* quanto na *Taxa de Câmbio*, do ponto de vista financeiro as médias 0,000156 e 0,000239 são *significativamente diferentes de zero*. Para mostrar isso basta observar que os valores da taxa histórica anual  $G$  (historic anual compound rate) apresentadas pelas *Ações* e a *Taxa de Câmbio* são iguais, respectivamente, a 4,01% e 6,2%. Essa taxa é calculada da seguinte forma: seja  $P_i$  o preço no instante  $i$  e, se no tempo  $T=0$  investimos 1 real, então após  $n$  dias de mercado o ganho é igual a:  $\frac{P_n}{P_0} = \exp(y_1 + y_2 + \dots + y_n) = \exp(n\bar{y})$ , logo a taxa que dá o mesmo retorno após  $n/N$  dias ( $N=252$  dias de funcionamento do mercado num ano) é definido como:  $(1 + G)^{n/N} = \exp(n\bar{y})$ , dessa forma  $G = \exp(N\bar{y}) - 1$ , que normalmente é dada em porcentagem.

Ao fazer uma análise separada ano a ano, como apresentada nas Tabelas 2.3 e 2.4 observa-se que, nas duas séries, existem diferenças no nível em cada ano.

Por outro lado, com relação à variabilidade dos retornos, na Tabela 2.1 podemos observar que o desvio padrão amostral da *Taxa de Câmbio* é muito menor do que o desvio padrão das *Ações*. Sendo que o desvio padrão é grandemente influenciado pelas observações extremas, retiramos o efeito do 0,5% dessas. Nesse caso, para as *Ações* o desvio padrão diminui bastante, em 11%, sendo igual a  $53,54 \times 10^{-3}$ . Por sua vez, para a *Taxa de Câmbio*, o desvio padrão é igual a  $7,56 \times 10^{-3}$  sendo similar ao encontrado utilizando toda a amostra.

<sup>2</sup> Como mostrado na Seção anterior as distribuições são leptocúrticas.

**Tabela 2.3**  
**Estatísticas da série *Ações***

Ano	N	Média <sup>1</sup>	Mediana <sup>1</sup>	Desvio Padrão <sup>2</sup>	Q <sub>3</sub> -Q <sub>1</sub> <sup>2</sup>	Coefici. Simetria	Excesso Curtose
1989	241	-0,13	0	6,88	5,88	-0,97	4,84
1990	241	-4,95	-3,83	7,25	7,89	-1,06	8,47
1991	247	6,56	3,85	5,82	6,86	1,33	8,60
1992	247	-3,82	1,12	7,64	8,48	-2,79	22,78
1993	246	2,86	0,11	3,58	4,52	-0,29	1,06
1994	245	1,14	2,42	4,45	5,91	-0,24	0,69
1995	132	-1,38	-1,15	5,63	6,57	1,68	9,98

<sup>1</sup> (x10<sup>3</sup>)<sup>2</sup> (x10<sup>2</sup>)

**Tabela 2.4**  
**Estatísticas da série *Taxa de Câmbio***

Ano	N	Média <sup>1</sup>	Mediana <sup>1</sup>	Desvio Padrão <sup>2</sup>	Q <sub>3</sub> -Q <sub>1</sub> <sup>2</sup>	Coefici. Simetria	Excesso Curtose
1984	251	-7,1	-0,1	7,24	10,11	0,48	0,60
1985	252	7,7	-3,1	9,40	9,40	0,81	2,46
1986	253	9,0	11,83	8,14	8,75	-0,32	1,60
1987	253	6,53	0	7,17	7,92	0,29	1,58
1988	253	-5,7	-3,6	6,64	6,56	-0,56	3,25
1989	252	1,2	-4,1	7,13	8,52	0,25	1,38
1990	253	5,1	1,7	6,68	8,63	0,23	1,26

<sup>1</sup> (x10<sup>3</sup>)<sup>2</sup> (x10<sup>2</sup>)

Entretanto, o comportamento mais interessante ocorre ao fazer a análise a cada ano. Do exame informal das Tabelas 2.3, 2.4 podemos concluir que o desvio padrão não é constante, implicando que a variabilidade depende do período de tempo considerado. Podemos obter a mesma conclusão se usamos uma medida robusta como a distância entre quartis (Q<sub>3</sub>-Q<sub>1</sub>). Assim, para as *Ações*, ao observar os valores do desvio padrão e Q<sub>3</sub>-Q<sub>1</sub> encontramos que os anos de menor variabilidade foram 1993 e 1994 e o ano de maior variabilidade foi 1992. Por

sua vez, para a *Taxa de Câmbio*, o ano que apresentou maior variabilidade foi 1985 (1984) considerando o desvio padrão ( $Q_3-Q_1$ ) como medida de variabilidade e os anos de menor variabilidade foram 1988, 1990 considerando o desvio padrão como medida de variabilidade e 1988, 1983 considerando  $Q_3-Q_1$ .

No lugar de analisar a variabilidade ano a ano, podemos refinar a análise calculando o *Desvio Padrão Móvel* (DPM) utilizando uma *janela* de 22 dias (aproximadamente um mês de mercado). Os gráficos do DPM nas *Séries Exemplo* são mostrados nas Figuras 2.5a e 2.5b, sendo observado que em ambas séries existe uma evidência muito forte de heteroscedasticidade na variância, isto é, a variância varia com o tempo. Em particular, na série das *Ações*, vemos que existem várias subidas abruptas. Confrontando esse gráfico com o correspondente da Figura 2.1 observamos que essas subidas são consequência das observações extremas.

É ilustrativo também examinar o gráfico do CUSUM dos quadrados dos retornos (Brown *et alii*, 1975), para detectar heteroscedasticidade e/ou presença de mudanças estruturais na variância. Observando as Figuras 2.5c e 2.5d percebemos que tanto as *Ações* quanto a *Taxa de Câmbio*, apresentam um afastamento da hipótese de homoscedasticidade.

O comportamento da variabilidade exibido pelos retornos, pode ser explicado através de dois modelos. No primeiro assumimos que a variância (não condicional) não é constante. No segundo é suposto que a variância (não condicional) é constante e que a *variância condicional à informação passada* é heteroscedástica. Os modelos estudados nos Capítulos 3 e 4 são deste tipo.

Baseados na estimativa da variância, os investidores acostumam calcular o *Retorno Futuro Esperado* (RFE). Esse índice é calculado como  $RFE = \exp(N(\bar{x} + 0.5s^2)) - 1$  ( $N=252$ ) e usualmente é dado em porcentagem (ver Taylor, 1986 p.35). Para as *Séries Exemplo* os resultados são: 65,69% para as *Ações* e 6,97 % para a *Taxa de Câmbio*. Esse indicador apresenta dois problemas. Ele é calculado sob a suposição errada de que a distribuição dos retornos é normal e, em segundo lugar, a variância amostral superestima a variância na presença de observações extremas. Conseqüentemente, a utilização desse índice deve ser considerado com cautela.

Por outro lado, na análise empírica das séries financeiras, alguns autores como Taylor (1986), Nelson (1991), entre outros, mostraram que os dias da semana influenciam o valor do

retorno. Isso é conhecido na literatura como *Efeito Calendário*. Uma forma simples de avaliar esse efeito é através da comparação das medidas de locação e variabilidade em cada dia da semana. Na Tabela 2.5 são apresentados esses resultados para as *Séries Exemplo*. Ao calcular intervalos de confiança 95% (i.c.) para a média<sup>3</sup> de cada dia da semana encontra-se que, nas duas séries, os i.c. incluem as médias em todos os casos. Entretanto, pelos valores G mostrados, é conveniente modelar os retornos como tendo média diferente de zero. Por outro lado, nas *Ações* os níveis na Segunda e Quarta Feira são negativos. A característica observada nas *Ações* da Segunda Feira apresentar valor negativo é concordante com resultados encontrados na literatura para séries de ações (ver Taylor, 1986). Na série *Taxa de Câmbio*, a Quinta Feira apresenta a maior média e a Sexta Feira a menor média.

Com relação as variabilidades, na série das *Ações* a maior variabilidade se dá na Segunda Feira (utilizando tanto o desvio padrão quanto  $Q_3-Q_1$  como medidas de variabilidade) e, na *Taxa de Câmbio*, os desvios padrão correspondentes à Segunda e Terça Feira são os maiores, enquanto que a Quinta Feira apresenta o menor valor  $Q_3-Q_1$ .

Em conclusão: parece razoável supor que o *nível médio dos retornos é melhor modelado como sendo diferente de zero, que a variabilidade dos retornos depende do período de tempo considerado e, finalmente, que ambas, o nível e variabilidade são influenciados pelos dias da semana.*

## 2.6 Estacionariedade

O problema da estacionariedade nas séries financeiras está centrado no fato de assumir *variância finita* (não condicional). Na subseção anterior foi discutida a possibilidade de que a variância não condicional dos retornos varie temporalmente, sendo detectada evidências nesse sentido para as *Séries Exemplo*. Resta saber se a variância não condicional é finita. Nesse sentido, observando a Figura 2.5 do desvio padrão móvel temos a impressão que sem. Além disso, na subseção 2.4 concluímos que as *Séries Exemplo* exibiam cauda pesada, mas não tão grandes quanto à Cauchy. Pelos tamanhos de cauda obtidos é razoável assumir que exista quarto momento, inclusive. Em consequência as séries são covariante estacionárias.

---

<sup>3</sup>Intervalos de confiança calculados sob a suposição de que os retornos são *i.i.d* Gaussianos.

**Tabela 2.5**  
*Efeito Calendário nas Séries Exemplo*

*Ações*

	Média <sup>1</sup> (G%) <sup>a</sup>	Mediana <sup>1</sup>	Desvio Padrão <sup>2</sup>	Q <sub>3</sub> -Q <sub>1</sub> <sup>2</sup>
Segunda Feira	-74.30 (-31,2)	-48.50	73.37	71.57
Terça Feira	45.97 (26,1)	5.68	57.33	60.36
Quarta Feira	-29.50 (-13,8)	0	56.21	59.80
Quinta Feira	49.16 (28,1)	6.84	54.81	61.35
Sexta Feira	16.25 (8,5)	0	60.03	58.01

*Taxa de Câmbio*

	Média <sup>1</sup> (G%) <sup>a</sup>	Mediana <sup>1</sup>	Desvio Padrão <sup>2</sup>	Q <sub>3</sub> -Q <sub>1</sub> <sup>2</sup>
Segunda Feira	2.49 (13,4)	1.97	8.02	8.80
Terça Feira	3.88 (21,6)	0	7.88	8.56
Quarta Feira	-0.06 (-0,3)	0	7.42	8.93
Quinta Feira	5.91 (34,7)	5.70	7.37	8.30
Sexta Feira	-1.70 (-8,2)	6.40	7.38	8.70

<sup>1</sup> (x10<sup>4</sup>)

<sup>2</sup> (x10<sup>3</sup>)

<sup>a</sup> calculada considerando N=252/5.

Uma ferramenta adicional para avaliar a hipótese de variância finita, originalmente proposta por Mandelbrot (1963) e enfatizada por Pagan e Schwert (1990)<sup>4</sup> é o *Gráfico das Variâncias Recursivas*. A variância recursiva pode ser definida como<sup>5</sup>:

$$vr_t = \frac{1}{t} \sum_{i=1}^t Y_i^2, \quad i=1,2,3...T \quad (2.6.1)$$

Na presença de heteroscedasticidade, observaríamos saltos na seqüência dos {vr<sub>t</sub>} e, se os valores dessa não convergem, obtemos evidência de que a variância pode não existir. Na Figura

<sup>4</sup> Nesse trabalho é realizado um apelo a favor da utilização de modelos que respeitem a natureza da variância das séries como sendo finita ou infinita. Assim, modelos covariante estacionários que apresentam bom ajuste dos dados mas que apresentam variâncias recursivas que explodem com o tempo não são adequados.

<sup>5</sup> Alternativamente podem ser utilizados os retornos corrigidos pela média, mas isso não faz muita diferença.

2.6 mostramos os gráficos da Variância Recursiva para as *Séries Exemplo*. A análise é realizada considerando a última parte (digamos 3/4) de cada série. Observando a parte média em cada uma, vemos que existem várias subidas e descidas, evidenciando algum tipo de heteroscedasticidade. Em particular, na série das *Ações* existe um grande salto no começo de 1992, o qual é produzido pela observação extrema, que corresponde à *Queda da Ministra Zélia Cardoso de Mello*. No entanto, nas duas séries, as seqüências convergem e, portanto, existe evidência que a variância é finita.

Como sugerem Phillips e Loretan (1990), o comportamento de saltos na seqüência  $\{v_t\}$  pode ser consequência da inexistência de quarto momento. Esse comportamento é confirmado por Pagan (1994) através de simulações com modelos de variância condicional heteroscedástica que não apresentam quarto momento finito.

Outras formas de avaliar a hipótese de variância finita são através da modelagem e dos Testes de Hipóteses. Na literatura, encontramos evidência de variância infinita através de ajuste, em Mandelbrot (1963) utilizando *Famílias Estáveis*<sup>6</sup>, Bollerslev e Engle (1986) com modelos IGARCH (desenvolvidos no Capítulo 3), entre outros. Por sua vez, alguns pesquisadores como Pagan (1994) e Mills (1993) realizam testes de Raiz Unitária sob a hipótese nula que o modelo possui raiz unitária.

Finalmente, do ponto de vista intuitivo e da Teoria de Finanças é difícil acreditar na hipótese que os retornos possuam variância (não condicional) infinita. Em particular, para as *Séries Exemplo*, a hipótese de estacionariedade é certamente adequada.

## 2.7 Autocorrelação, Independência e Não Linearidade

Na Seção 2.1 foi comentado que um dos modelos mais utilizados no passado para modelar os retornos foi o Passeio ao Acaso, no qual era assumido que os retornos eram independentes ou não correlacionados. Nesse sentido, nesta seção serão discutidos os problemas da autocorrelação e independência dos retornos. A seção está dividida em três partes: na primeira, apresentamos algumas metodologias para tratar o problema do teste da autocorrelação; a seguir tratamos o assunto da independência das observações vinculada à não

---

<sup>6</sup> Embora não apresentados, os ajustes de distribuições estáveis para as *Séries Exemplo* não foram satisfatórios.

linearidade e, finalmente, apresentamos uma breve discussão de outros testes de não linearidade.

### 2.7.1 Testes de Autocorrelação

Ao observar os resultados do Teste Box-Ljung para os retornos na Tabela 2.6, encontramos que na *Taxa de Câmbio* há evidência de ruído branco (ou aproximadamente ruído branco), o qual é verificado na Figura 2.7 através do gráfico da função de autocorrelação (f.a.c). Nas *Ações*, pelo contrário, os resultados da Tabela 2.6 apontam que a hipótese de ruído branco nos retornos não é plausível. Observando a Figura 2.7 notamos que a primeira autocorrelação está muito distante dos limites usuais, sugerindo que a estrutura da série pode ser AR(1) ou MA(1). O valor apresentado pela primeira autocorrelação nas *Ações* é positivo e relativamente pequeno; 0,083. Essas características são freqüentemente observadas em séries de ativos (ver Taylor, 1986). Por sua vez, na série *Taxa de Câmbio* a estimativa da primeira autocorrelação não é significativa.

**Tabela 2.6**  
Valores P do Teste Box-Ljung nos retornos das *Séries Exemplo*

Série	Lag			
	6	12	18	24
<i>Ações</i>	0,012	0,017	0,014	0,015
<i>Taxa de Câmbio</i>	0,412	0,323	0,060	0,085

O argumento usado na discussão anterior para obter evidência de ausência de autocorrelação através da f.a.c tem alicerces na distribuição assintótica dos estimadores das autocorrelações (Brockwell e Davis p. 221). A primeira suposição para que a distribuição assintótica seja válida é que a série seja *estacionária*. Como foi discutido na seção 2.6, essa hipótese é razoável para muitas séries financeiras incluindo as *Séries Exemplo*. Outra suposição considera os retornos como sendo *identicamente distribuídos* mas existem resultados assintóticos para a distribuição desses estimadores supondo momentos finitos (Brockwell e Davis p. 222). Entretanto, a hipótese mais importante é a existência de um *processo gerador de dados linear*, mas, como mostraremos nas subseções seguintes, existe uma evidência muito forte contra essa suposição.

Na modelagem de séries financeiras, o uso da função de autocorrelação (f.a.c) e função de autocorrelação parcial (f.a.c.p) na especificação do modelo, supondo que o processo gerador de dados é *linear*, pode conduzir a uma super-parametrização desse. Como é documentado em Taylor (1986), Weiss (1984), Mills (1993) e Bollerslev *et alii* (1994), entre outros, modelos não-lineares com variância condicional heterocedástica apresentam a característica que as variâncias dos estimadores das autocorrelações serem maiores que  $T^{-1}$ . Em consequência, os intervalos de confiança tradicionais, usando modelos lineares, subestimam a largura verdadeira dos intervalos. Para ilustrar isso, apresentamos na Tabela 2.7 os resultados de 100 simulações para um modelo de variância condicional heteroscedástica; o modelo de Variância Estocástica AR(1)-SV estacionário (ver Capítulo 4) com séries de tamanho 1000. Como será demonstrado no Capítulo 4, os retornos que seguem esse modelo são não-correlacionados. Os resultados da Tabela 2.7, ilustrados na Figura 2.8, mostram que as primeiras autocorrelações são maiores que os limites usuais com uma frequência maior à esperada se o modelo fosse linear. Adicionalmente é observado que a distribuição das autocorrelações é assimétrica à esquerda. Os resultados mostrados valem em geral para várias especificações do modelo AR(1)-SV, de forma que, sob a suposição de que o processo gerador das *Ações* é dado pelo modelo AR(1)-SV, a grande primeira autocorrelação observada poderia ser não significativa e, conseqüentemente, essa série seria não correlacionada.

Não existem resultados gerais para calcular a distribuição assintótica dos estimadores das autocorrelações quando o processo é não-linear. Nesse sentido, uma exceção é o trabalho de Taylor (1986, Cap 5) no qual são derivados resultados assintóticos baseados no método dos momentos, sob a (única) hipótese de simetria na distribuição dos retornos.

A presença de autocorrelação nas séries financeiras (*Mean-Reversing*) tem sido encontrada em índices de ativos, por exemplo em Scholes e Williams (1977), que ajustam modelos MA(1) e Lo e McKinlay (1988) e Poterba e Summers (1988) que ajustam AR(1).

Finalmente, supondo normalidade, não correlação implica independência. A partir do discutido na Seção 2.4, a cauda das distribuições é mais pesada que a normal, portanto normalidade não é uma suposição válida. Outras formas alternativas de testar a independência dos retornos, são apresentadas nas duas subseções seguintes.

**Tabela 2.7**  
**Percentagem das Autocorrelações no Modelo AR(1)-SV estacionário<sup>a</sup> que se encontram fora dos limites usuais de Bartlett**

Autocorrelação	% maior que 1 desvio <sup>1</sup>			% maior que 2 desvios		
	< 1 desvio	> 1 desvio	total	< 1,96desvios	>1,96desvios	total
Primeira	39	30	69	24	15	39
Segunda	28	35	63	16	19	35
Terceira	40	26	66	26	10	36
Quarta	38	23	61	22	13	35
Quinta	33	33	66	14	18	32
Sexta	30	34	64	18	19	37
Sétima	23	28	51	12	12	24
Oitava	27	18	45	6	10	16
Nona	21	28	49	11	11	22
Decima	27	19	46	8	9	17
Esperado (%)	15,87	15,87	32	2,5	2,5	5

<sup>a</sup> Parâmetros do modelo:  $\gamma=0,95$ ,  $q = 0,05$ ,  $\kappa=5,5 \times 10^{-7}$ .

<sup>1</sup> desvio= $1/1000^{0,5}$

### 2.7.2 Independência e Não Linearidade.

Se o processo gerador de dados é um ruído branco estrito, então as observações e quaisquer função delas, incluindo *os quadrados das observações são não correlacionadas*. Essa idéia foi implementada pela primeira vez por Granger e Andersen (1978), para mostrar que a hipótese de independência utilizada em várias séries estudadas na literatura era incorreta. Da mesma forma, sob a suposição de independência, os valores absolutos das observações também são não correlacionados. Assim, podemos obter evidência da independência dos retornos através das funções de autocorrelação das observações, os quadrados das observações e os valores absolutos das observações. Como mostraram McLeod e Li (1983), supondo um processo gerador de dados *linear*, o valor da variância assintótica das autocorrelações dos quadrados das observações continua sendo  $T^{-1}$ . Considerando isso e o fato que as autocorrelações são sempre positivas, obtemos os limites para os correlogramas. Adicionalmente, nesse trabalho é proposta uma estatística portmanteau similar à estatística proposta por Box-Ljung utilizada nos modelos ARIMA.

Na Figura 2.7 são apresentadas as funções de autocorrelação dos quadrados dos retornos das *Séries Exemplo*. Para ambas séries, considerando que existem mais retornos dos esperados fora dos limites, concluímos que a hipótese de independência nos retornos é implausível. Por sua vez, utilizando o teste portmanteau de McLeod e Li (1983) obtemos valores-P inferiores a 0.0001 para os primeiros 10 lags, confirmando as nossas conclusões.

Autocorrelação nos quadrados das observações constitui uma evidência do caráter não linear do processo gerador de observações. Essa afirmação como observa Pagan (1994, p. 14) “*is a somewhat imprecise description and one that arise more by default than anything else*”. O sentido da afirmação é o seguinte: se os retornos  $\{Y_t\}$  são gerados por um processo linear, então quaisquer dependência nos momentos de  $Y_t$  *pode ser produto da estrutura não-linear*. Dessa forma, na literatura frequentemente, *não-linearidade é entendido como dependência no segundo momento*. Portanto, o interesse dos pesquisadores para tratar séries financeiras é modelar não somente a média, mas também a variância, considerando a natureza *heterocedástica*. Nesse sentido, várias propostas fazem distinção entre variância não condicional e variância condicional. A literatura sobre modelos com variância heteroscedástica é numerosa. Dentro dessa abordagem, nos Capítulos 3 e 4 discutimos em detalhe duas famílias de modelos que incorporam a hipótese de heterocedasticidade na variância.

É interessante discutir o comportamento das autocorrelações dos quadrados dos retornos para poder construir modelos capazes de reproduzi-los. Em primeiro lugar as séries financeiras apresentam frequentemente a *primeira autocorrelação pequena e estatisticamente diferente de zero*. Isso acontece nas *Séries Exemplo*, onde a primeira autocorrelação de  $\{Y_t^2\}$  é igual a 0,11 para as *Ações* e 0,055 para a *Taxa de Câmbio*. Esses valores são similares aos encontrados na literatura para outras séries de ativos e taxas de câmbio, ver Taylor (1986 p.55). A primeira autocorrelação significativa e positiva, pode ser explicada pelo fato de que retornos grandes (em valor absoluto) tendem a suceder retornos grandes, enquanto que retornos pequenos tendem a suceder a retornos pequenos. Em consequência, *os conglomerados de volatilidade e autocorrelação nos quadrados dos retornos estão intimamente relacionados*.

Outra característica importante, frequentemente observada nas séries dos quadrados dos retornos, é que as autocorrelações apresentam decaimento lento. Esse comportamento pode ser observado para a série das *Ações* na Figura 2.7. Da mesma forma, para a série *Taxa de Câmbio*

(Figura 2.7), as autocorrelações também decaem, mas só a partir do lag 6, sendo as 5 anteriores menores.

Sintetizando, encontramos que os quadrados dos retornos das séries financeiras possuem a característica de apresentar a primeira autocorrelação pequena e as outras decaindo lentamente. Ambos fatos podem ser reproduzidos pelos modelos ARMA(1,1) ou por outros processos integrados. No caso das séries *Taxa de Câmbio* e *Ações* é razoável supor que o f.a.c dos quadrados dos retornos apresenta um comportamento similar a ARMA(1,1). Nos Capítulos 3 e 4 apresentamos duas famílias de modelos que exibem f.a.c similares à ARMA(1,1).

Em conclusão, *as Séries Exemplo são não lineares e os conglomerados de volatilidade são evidência da não-linearidade nos retornos.*

A partir do que foi abordado nesta subseção e na seção 2.4, torna-se difícil manter a hipótese de que os retornos sejam independentes. Na literatura de séries financeiras, a grande maioria de modelos de volatilidade considera os retornos como não-lineares e não correlacionados ou no máximo com estruturas AR de ordem pequeno.

### 2.7.3 Outros Testes de Não Linearidade.

A hipótese de independência foi rejeitada com base nas autocorrelações das funções das observações. Entretanto, é possível testar tal hipótese diretamente através da metodologia proposta por Brock, Dechert e Scheinkman (1987) conhecida como Teste BDS, ver Brock *et alii* (1987) e Brock *et alii* (1996). Entre as aplicações dessa metodologia em séries financeiras temos Hsieh (1989b), em taxas de câmbio, Scheinkman e LeBaron (1989) em ações (de retornos diários e mensais), Kearns (1990) em taxa de juros. Nos três estudos a hipótese nula de independência é rejeitada.

Por outro lado, Taylor (1986, p.57) desenvolve uma metodologia para testar o caráter não linear das observações e aplica essa metodologia para 40 séries, entre ações, taxas de câmbio, mercadorias, etc., concluindo que as séries são não lineares. Tsay (1986) propõe um teste que é poderoso contra alternativas de séries MA ou bilineares, mas basicamente detecta linearidades somente na média condicional, e por esse fato não é de muita utilidade para séries financeiras (Pagan, 1994). Cao e Tsay (1993) apresentam um resumo de outros testes não

lineares e os aplicam em retornos obtendo presença de não-linearidade. Finalmente, uma fonte para consultar sobre testes aplicados não somente a finanças é Tong (1990).

## 2.8 Normalidade

A presença de cauda pesada nos retornos invalida a hipótese de normalidade. Por sua vez, se o processo  $\{X_t\}$  é normal multivariado martingale diferença, então o processo  $\{X_t^2\}$  é não correlacionado. Os resultados empíricos encontrados para as *Séries Exemplo* mostram que as séries dos quadrados dos retornos apresentam correlação. Portanto já que os retornos da série *Taxa de Câmbio* e (muito provavelmente) das *Ações* são não correlacionados, podemos concluir que elas não podem ser normais

As evidências encontradas são suficiente para rejeitar a hipótese de normalidade na distribuição dos retornos das *Séries Exemplo*. Mas, como complemento, nesta seção é apresentada uma metodologia adicional para confirmar a ausência de distribuição normal das observações, o Teste  $W$  de Shapiro e Wilk (1965). Testes de momentos, como o proposto por Bowman e Shenton (1975), baseados na simetria e curtose, devem ser utilizados com cautela. Assim por exemplo, altíssimas estimativas de curtose como as encontradas para a série das *Ações* (devida a uns poucos valores extremos) conduzem a rejeitar a hipótese de normalidade. Por outro lado Pagan (1994), propõe um teste de frequências esperadas que incorpora a hipótese de heteroscedasticidade na variância condicional, a qual é a característica principal dos modelos desenvolvidos nos Capítulos 3 e 4. Adicionalmente, com o objetivo de visualizar a forma da distribuição dos retornos, será estimada a densidade utilizando métodos não-paramétricos.

### *Teste W de Shapiro e Wilk*

O Teste  $W$  é definido como o quociente do melhor estimador da variância baseado numa combinação linear de estatísticas de ordem. Portanto a influência das observações extremas é menor. Assim, sejam  $y_{(1)}, y_{(2)}, \dots, y_{(T)}$  as estatísticas de ordem, com  $m_i = E[y_{(i)}]$ ,  $m' = (m_1, m_2, \dots, m_T)$  e  $v_{i,j} = \text{Cov}[y_{(i)}, y_{(j)}]$ ,  $V = [v_{i,j}]$  matriz  $T \times T$ . Logo, a estatística  $W$  é definida como:

$$W = \frac{\left( \sum_{i=1}^T a_i y(i) \right)^2}{\sum_{i=1}^T (y(i) - \bar{y})^2} \quad (2.8.1)$$

onde  $a' = (a_1, a_2, \dots, a_T) = \frac{m' V^{-1}}{(m' V^{-1} V^{-1} m)^{1/2}}$  e  $\bar{y} = \frac{1}{T} \sum_{i=1}^T y(i)$ .

Os valores de  $W$  estão entre 0 e 1 sendo que valores pequenos desse conduzem à rejeição da hipótese de normalidade. A distribuição do teste  $W$  pode ser encontrada na referência. Experimentos Monte Carlo com amostras pequenas (20) sugerem que o teste tem poder nos casos de assimetria e cauda pesada. Aplicando o Teste  $W$  nas *Séries Exemplo* obtemos que: nas *Ações*,  $W=0,933$  (valor  $P=0,0001$ ) e na *Taxa de Câmbio*  $W=0,982$  (valor  $P=0,0001$ ), portanto rejeitamos a hipótese de normalidade.

#### *Estimação Não-paramétrica da Densidade*

Na teoria de estimação de densidades, a maior parte dos resultados referem-se à observações independentes e identicamente distribuídas (ver Silverman, 1986). Para esse tipo de dados, um dos estimadores mais utilizados é o estimador de Núcleo (*Kernel*):

$$\hat{f}(y) = \frac{1}{Th} \sum_{i=1}^T K\left(\frac{y_i - y}{h}\right) = \frac{1}{Th} \sum_{i=1}^T K(\varphi_i). \quad (2.8.2)$$

Para calcular efetivamente (2.8.2) é necessário escolher o valor de  $h$  e definir a função  $K(\varphi)$ . Escolha ótimas em determinadas condições (ver Silverman, 1986) são:  $K(\varphi)$  igual a  $\frac{1}{2\pi} \exp\left(-\frac{1}{2} \varphi^2\right)$ , e  $h = 1,096 T^{-1/5}$  vezes o desvio padrão amostral das observações. Adicionalmente, as estimativas da densidade apresentam robustez à escolha do  $K(\varphi)$ .

Utilizando o pacote *STAMP 5.0* (Koopman e Harvey, 1995) encontramos as estimativas da densidade para as *Séries Exemplo*. Na Figura 2.2 observamos que as densidades estimadas são aproximadamente simétricas para as duas séries.

Por outro lado, na estimação de densidades para séries financeiras (dados dependentes) Ait-Sahalia (1992,1996) demonstrou que sob determinadas condições de regularidade é possível obter uma seqüência de estimadores de núcleo que convergem em distribuição a uma distribuição normal (dessa forma podemos obter limites de confiança para os estimadores) e que o estimador da densidade é robusto à escolha de  $K(\varphi)$ . Delgado (1995) utiliza os resultados de Ait-Sahalia (1992) para estimar a densidade dos retornos de ações da TELEBRÁS<sup>7</sup> (num período diferente ao considerado neste trabalho e com retornos deflacionados) chegando à conclusão de que a densidade estimada é simétrica e apresenta caudas *menos* pesadas do que a distribuição normal. Esse resultado é uma das exceções na literatura sob séries financeiras. As séries financeiras têm densidades com caudas mais pesadas que a normal devido à presença de *picos*, isto é, valores mais altos ou baixos dos valores esperados em uma distribuição normal. Hoje existe na literatura um amplo consenso com respeito à natureza não normal da distribuição dos retornos.

## 2.9 Outras Características Empíricas dos Retornos

Vários outros *Fatos Estilizados* que apresentam os retornos das séries financeiras podem ser encontrados nos estudos de Bollerslev *et alii* (1994) e Ghysels *et alii* (1995). Fazemos a seguir, uma breve apresentação desses.

Estudos como os de Black (1976), Christie (1982), Schwert (1989a), ilustram a presença de autocorrelação negativa entre os retornos dos ativos e a volatilidade, isto é, os choques positivos produzem uma diminuição da volatilidade enquanto que os choques negativos produzem um aumento da volatilidade. Esse fenômeno é conhecido como *Leverage Effect*. O mesmo acontece para as taxas de câmbio (Kim, 1989), e para a série do índice CRSP (Nelson, 1991; Harvey e Shephard, 1993c). Le Baron (1992) mostrou que a relação inversa entre a volatilidade e a correlação serial nas ações é inusitadamente robusta à eleição de variáveis como período amostral, índice de mercado, medida de volatilidade, entre outras.

Por outro lado, diversos estudos ilustram que os períodos de fechamento do mercado, ou *Non Trading Periods*, produzem modificações no comportamento dos retornos. Por

---

<sup>7</sup>Neste trabalho também são estimadas as densidades de outras 3 séries brasileiras.

exemplo, Fama (1965) e French e Roll (1986), encontraram que a informação (*news*) acumula-se mais lentamente no período de fechamento comparado ao período de atividade.

Como todo componente econômico, a volatilidade está sujeita à influência de determinados acontecimentos. Por exemplo, acontecimentos como o anúncio de pagamento de dividendos ocasionam o aumento da volatilidade nas ações (Pattel e Wolfson, 1979,1981). Por sua vez, Baillie e Bollerslev (1991) encontraram que a volatilidade é maior na abertura e fechamento, tanto para as ações quanto para as taxas de câmbio. Por outro lado, *variáveis macroeconômicas* também produzem mudanças na volatilidade. Assim, Schwert (1989ab) encontrou que a volatilidade nas ações aumenta bruscamente nos períodos de recessão ou crises, porém diminui nos períodos de expansão. Em outro estudo, Ziegelmann (1996) mostrou que quanto maior é o volume negociado, maior é a volatilidade do índice IBOVESPA no dia seguinte. Finalmente, Glosten *et alii* (1993) mostraram que existe uma relação positiva entre a volatilidade das ações e a taxa de juros.

Finalmente, alguns autores mostram que podemos explicar grande parte das volatilidades de um bem negociado em vários mercados através de uns poucos *fatores comuns*. O mesmo acontece ao considerar vários bens dentro de um mercado. Aplicações para a taxa de câmbio podem ser encontrados em King *et alii* (1994) e Harvey *et alii* (1992).

## 2.10 Conclusões

As regularidades apresentadas pelos retornos das séries financeiras podem ser classificadas de dois tipos. O primeiro corresponde às características próprias das distribuições, como forma e momentos, enquanto que a segunda corresponde à evolução temporal retratada pela estrutura de autocorrelação e da natureza não-linear do processo gerador de dados. O estudo dessas características foi o objetivo básico deste Capítulo.

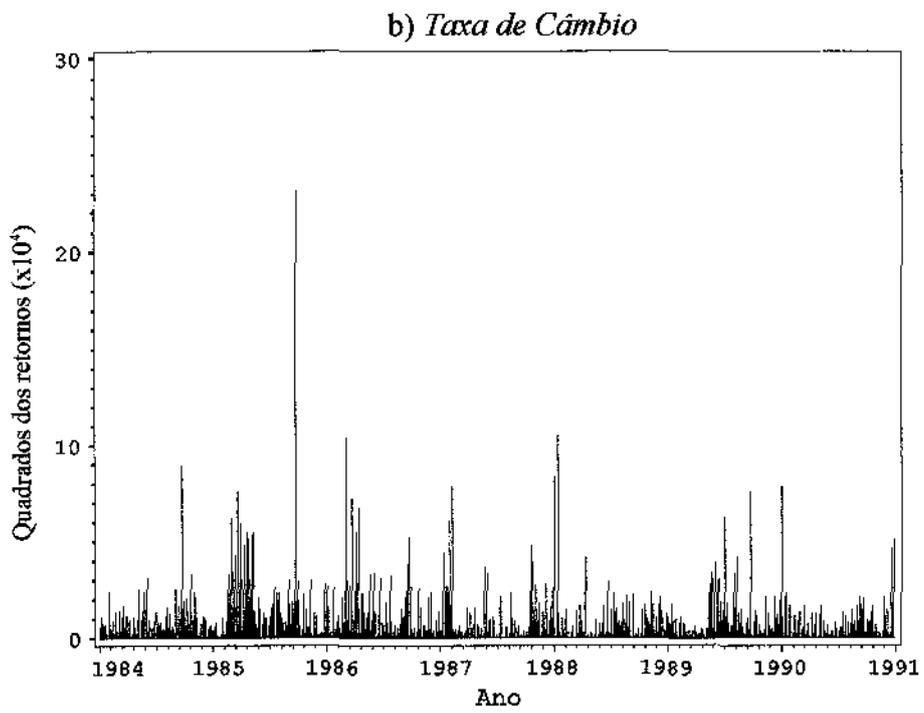
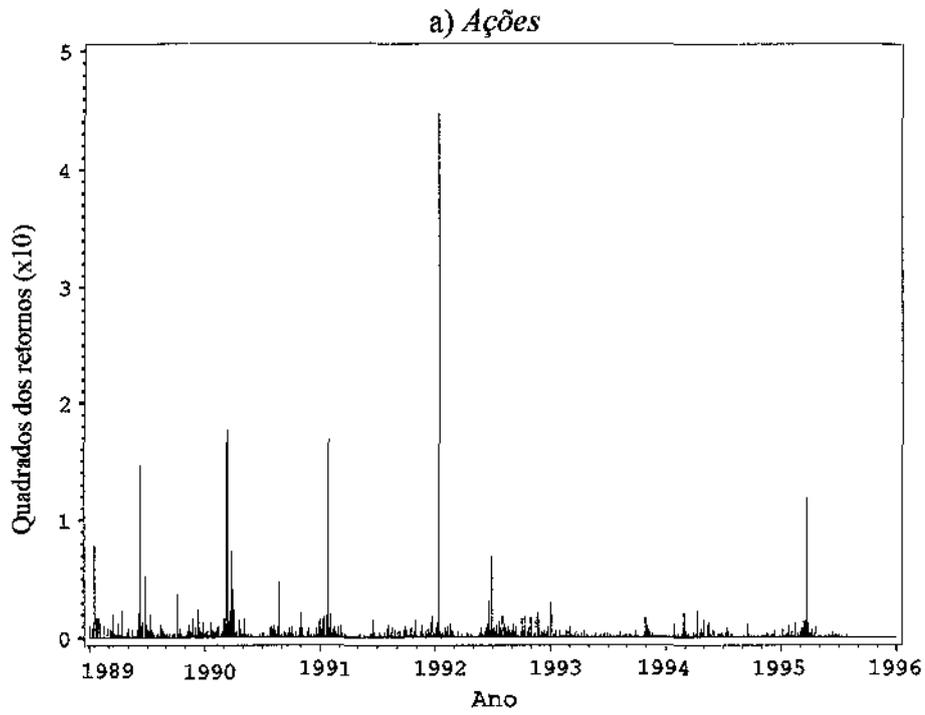
Os *Fatos Estilizados* que caracterizam os retornos de séries financeiras não podem ser explicados pelos modelos tradicionais de séries temporais (lineares). O tratamento dessas séries requer a implementação de outro tipo de modelos que sejam capazes de reproduzir essas características empíricas. A seguir apresentamos uma discussão dos *Fatos Estilizados* analisados neste capítulo e suas implicações na modelagem.

- i) A distribuição dos retornos geralmente é simétrica, o que implica que os momentos de ordem ímpar são iguais a zero.
- ii) Embora as médias amostrais sejam próximas de zero, é conveniente modelar os retornos com nível diferente de zero já que mesmo valores muito pequenos nos retornos podem ser importantes do ponto de vista financeiro.
- iii) A distribuição dos retornos não é normal mas leptocurtica. A cauda não é muito pesada, o que garante a existência de pelo menos quarto momento.
- iv) Os retornos apresentam estrutura de autocorrelação de ruído branco (ou no máximo autocorrelação de ordem um usando os testes das metodologias lineares), mas não são independentes, evidenciando o caráter não-linear dos retornos.
- v) As f.a.c da série dos quadrados dos retornos geralmente apresentam um comportamento similar a ARMA(1,1), sendo a primeira autocorrelação pequena, mas estatisticamente significativa e as outras decaindo lentamente.
- vi) A variância apresentadas pelas *Séries Exemplo* é finita.

Dada a existência de valores extremos nas séries financeiras, uma característica adicional desta dissertação é testar a utilização de procedimentos robustos na análise e advertir para o perigo de se obter conclusões deturpadas ao utilizar procedimentos convencionais.

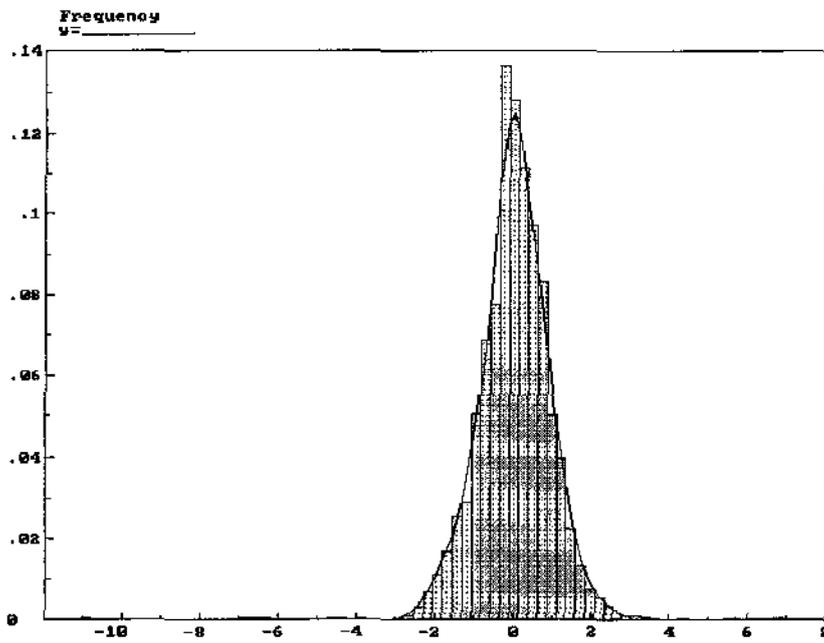
Adicionalmente, e dentro dos modelos não-lineares, é sugerido o uso de modelos de variância condicional heteroscedástica como alternativa para modelar os retornos. O tratamento detalhado da metodologia é assunto dos Capítulos posteriores, nos quais mostramos como esses modelos ajustam satisfatoriamente as *Séries Exemplo*.

Finalmente, os resultados obtidos neste capítulo para a *Taxa de Câmbio* são similares aos encontrados na literatura (ver referências no Capítulo 1).



**Figura 2.1:** Gráficos dos quadrados dos retornos das *Séries Exemplo*

a) Histograma e densidade ajustada



b) Gráfico de Simetria

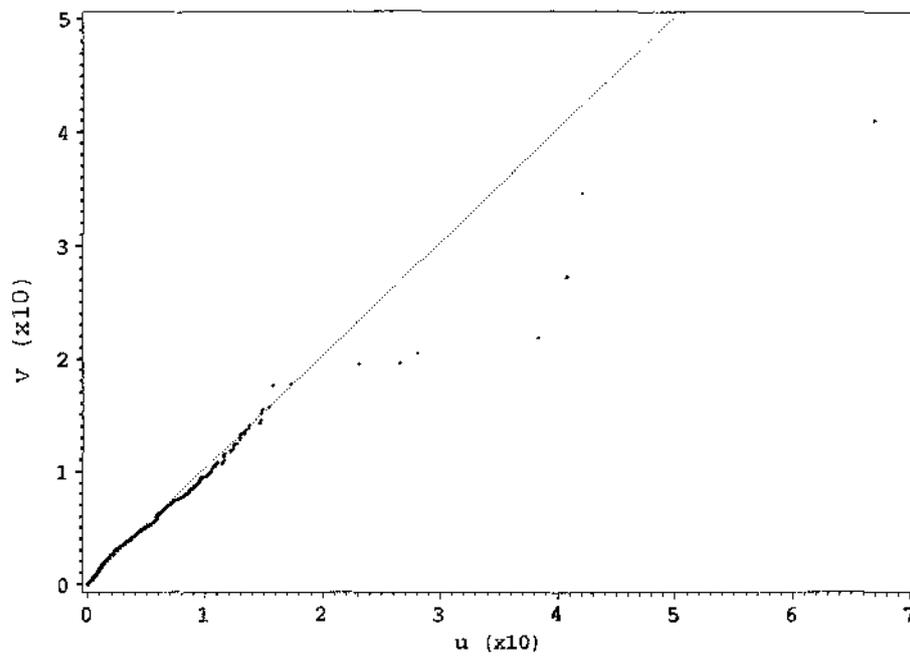
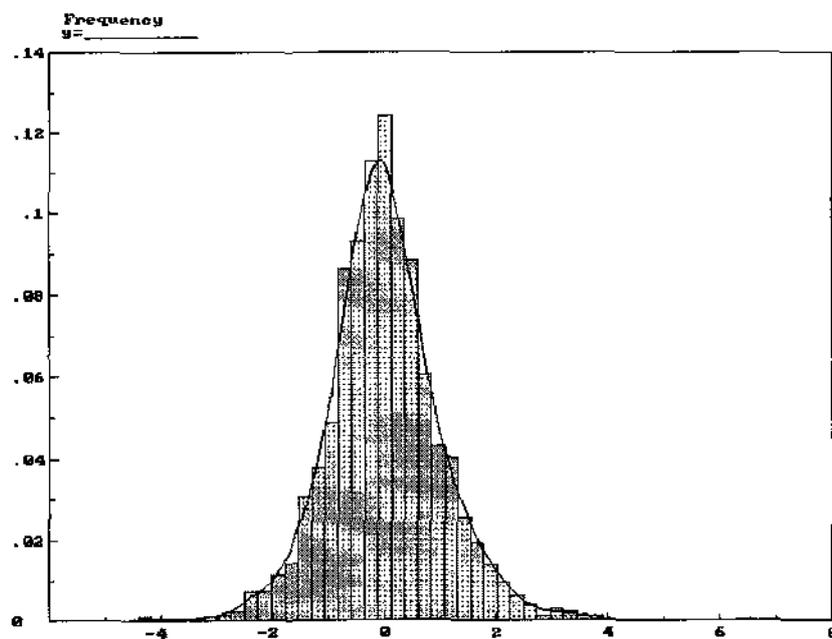


Figura 2.2: Histograma, densidade ajustada e Gráfico de Simetria nas *Ações*

a) Histograma e densidade ajustada



b) Gráfico de Simetria

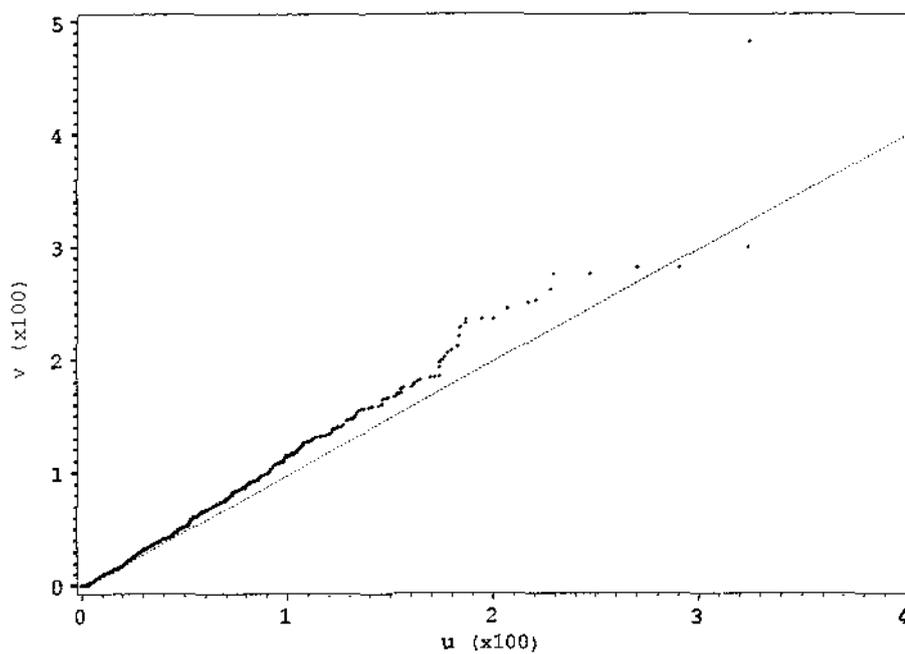
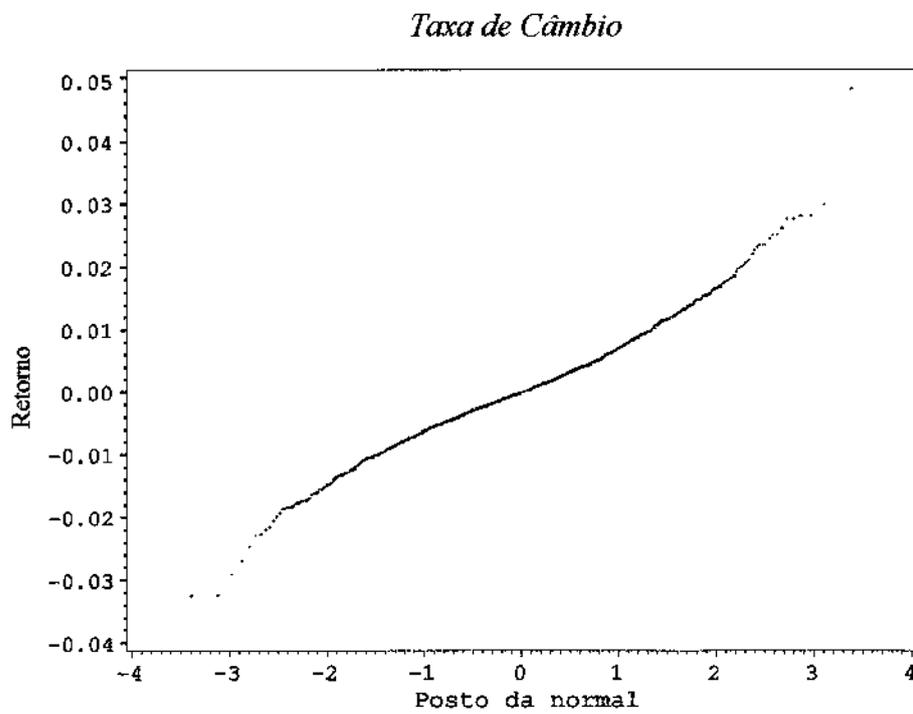
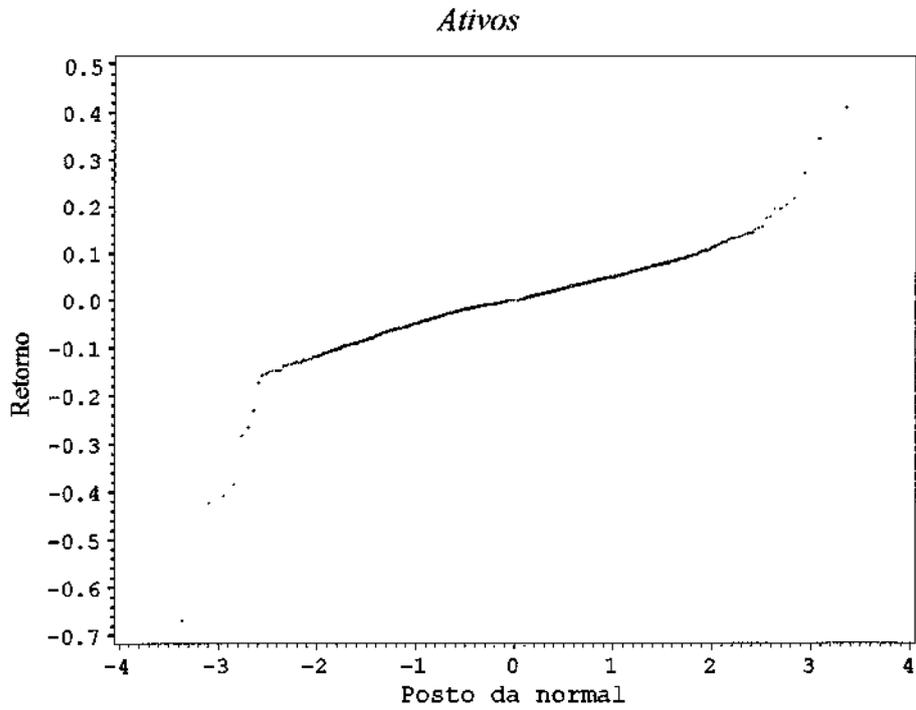
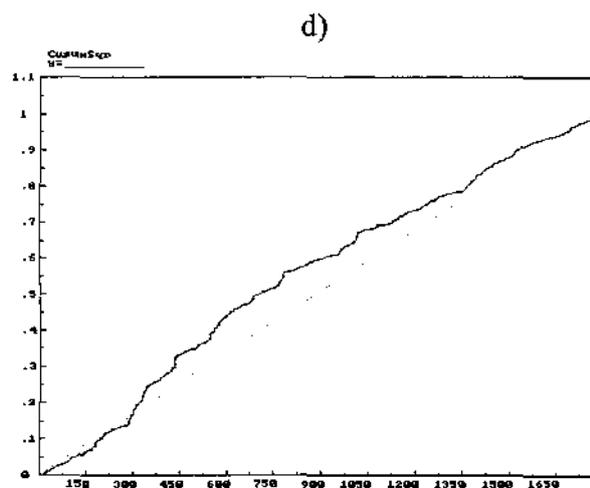
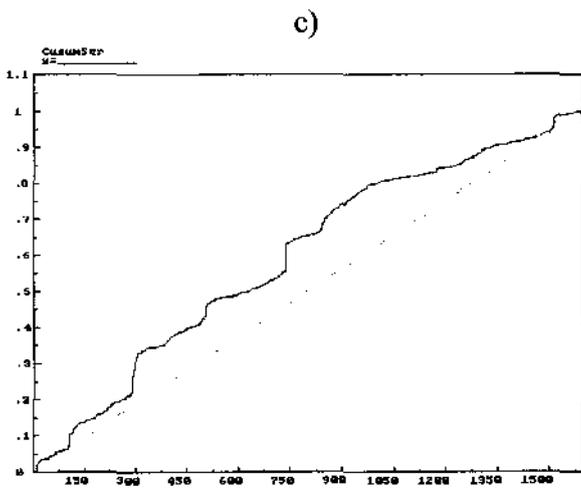
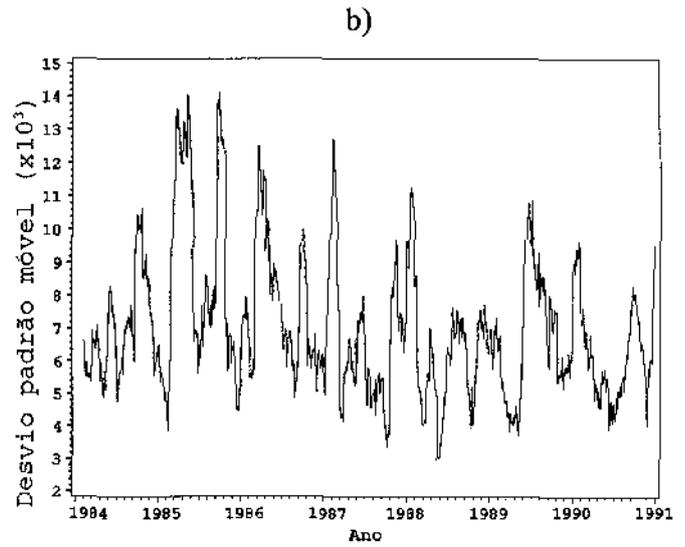
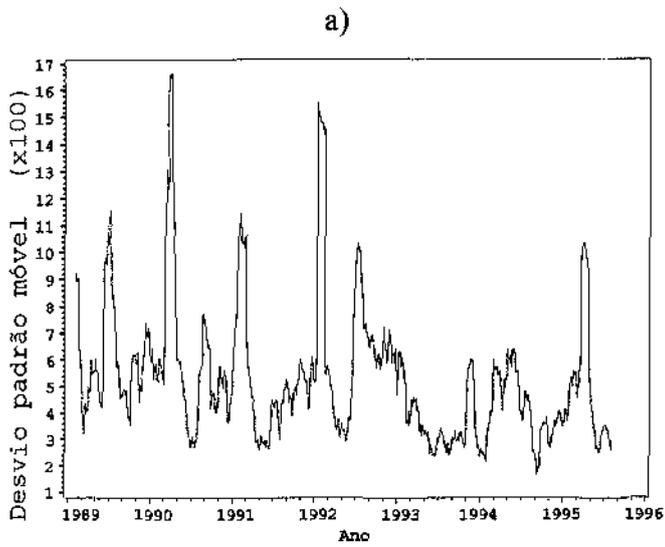


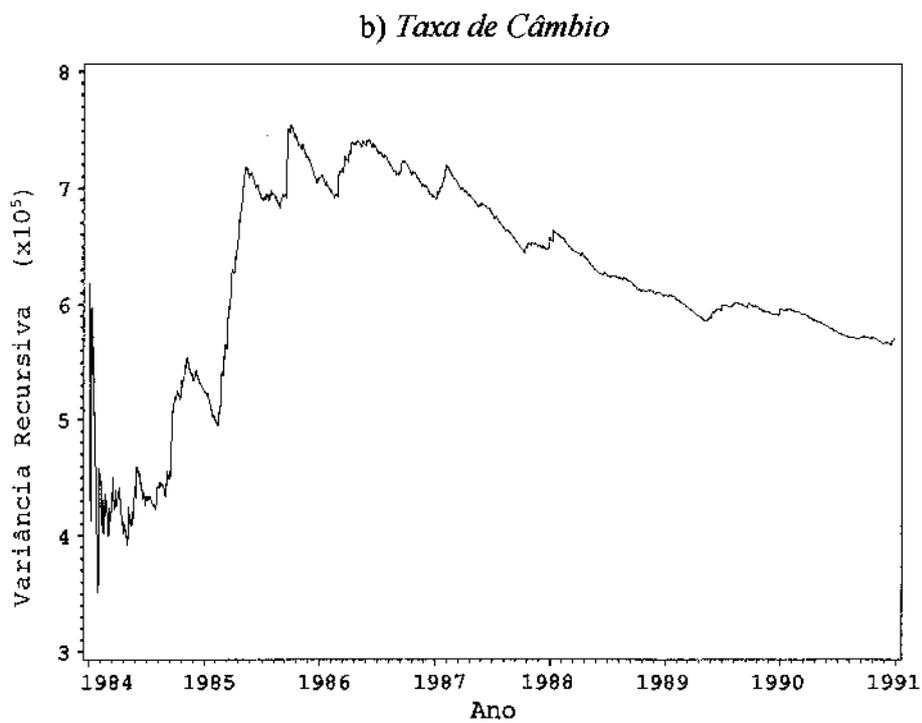
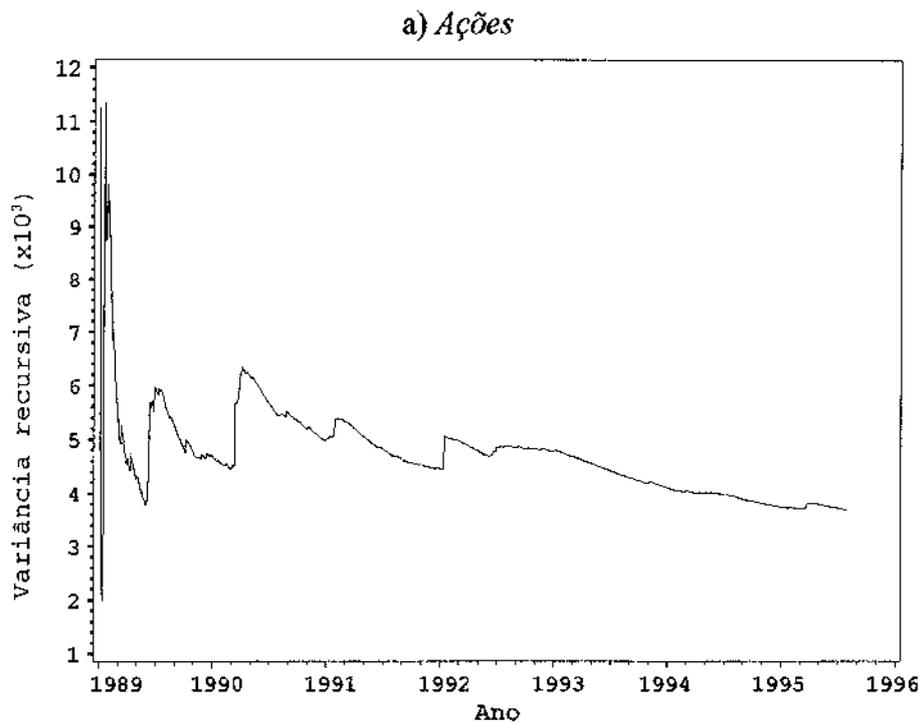
Figura 2.3: Histograma, densidade ajustada e Gráfico de Simetria na *Taxa de Câmbio*



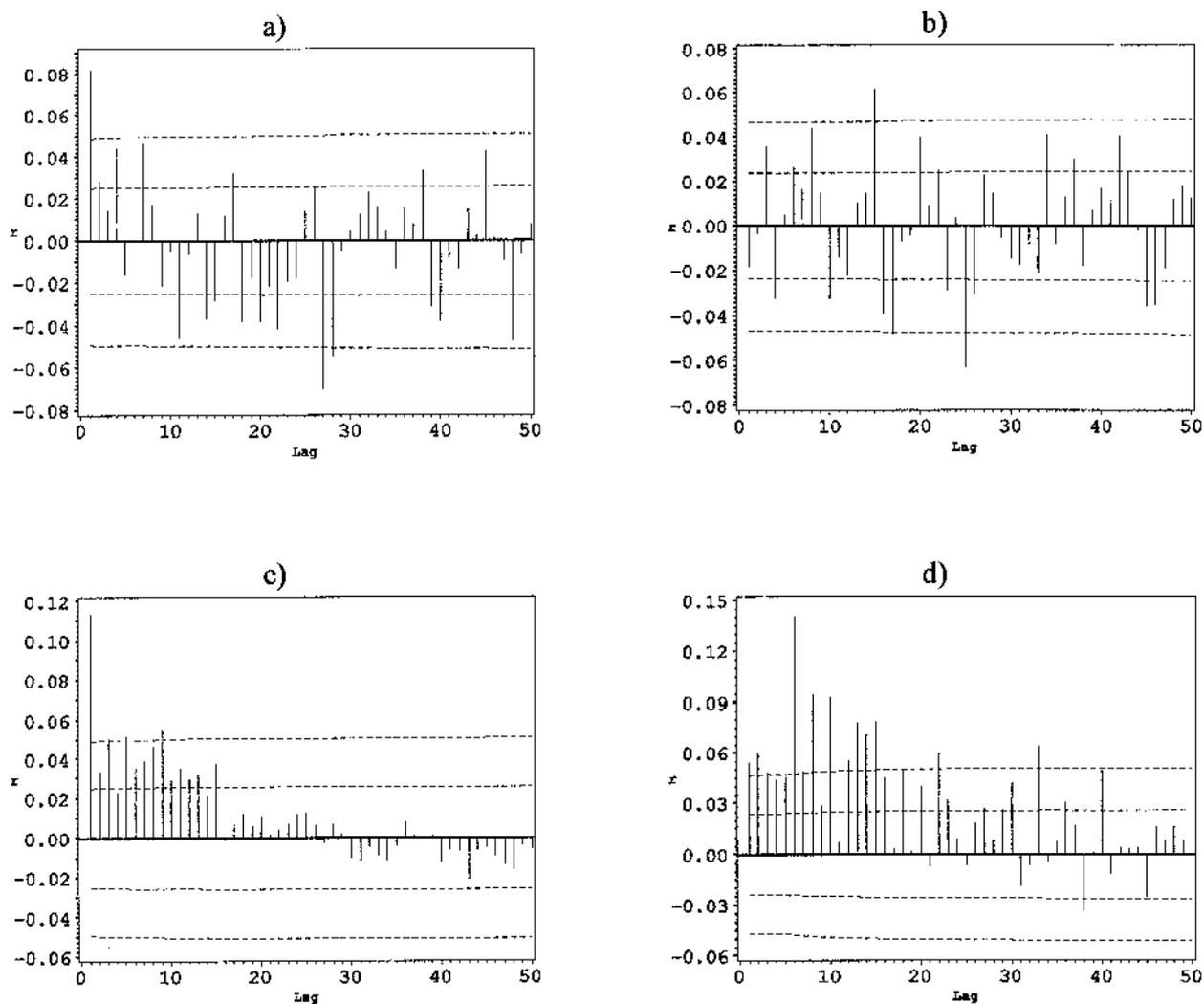
**Figura 2.4:** Gráfico de Probabilidade Normal dos retornos das *Séries Exemplo*



**Figura 2.5: Gráficos de Desvio Padrão Móvel e CUSUM dos quadrados dos retornos nas *Séries Exemplo***  
 a) Desvio Padrão Móvel nas *Ações*, b) Desvio Padrão Móvel na *Taxa de Câmbio*  
 c) CUSUM nas *Ações*, d) CUSUM na *Taxa de Câmbio*.



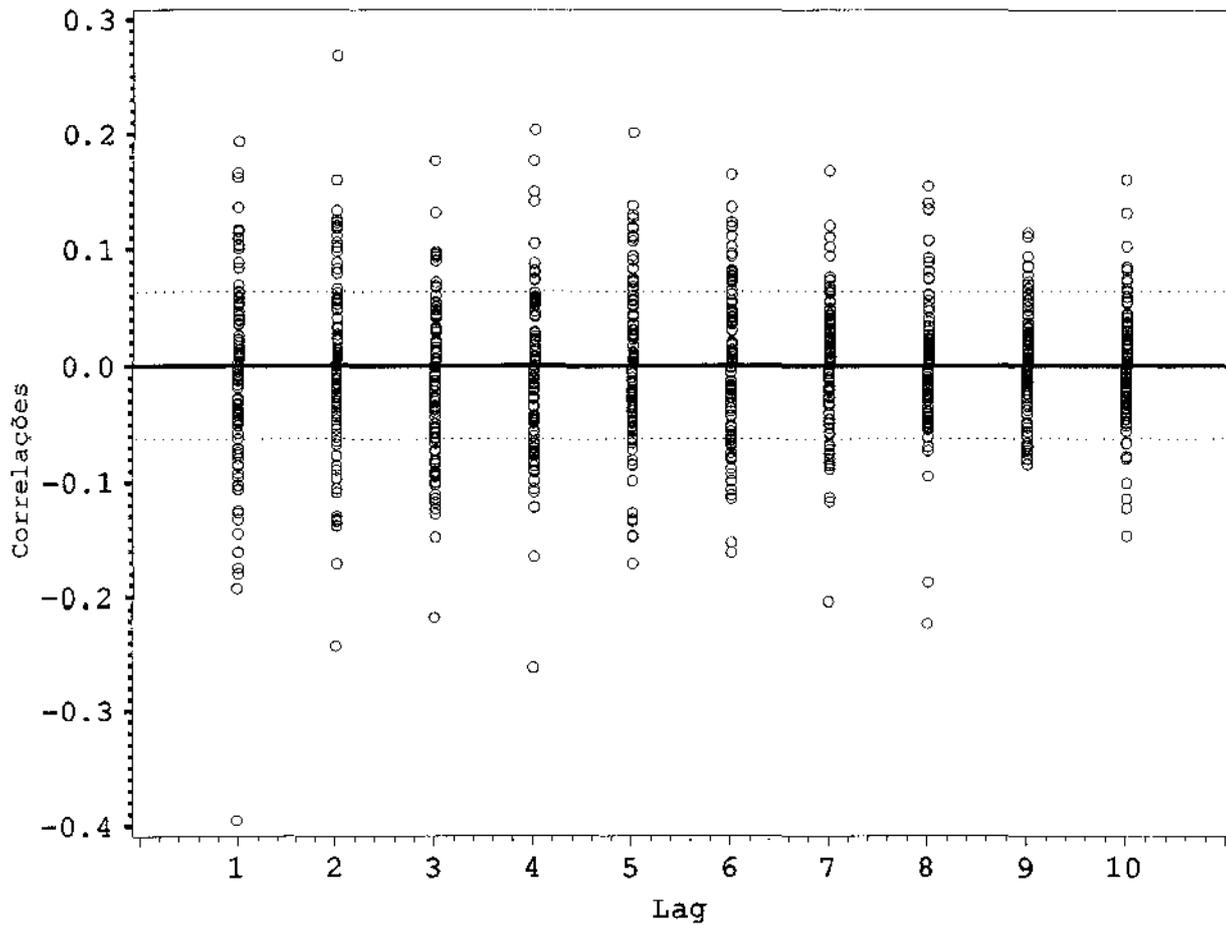
**Figura 2.6: Variâncias Recursivas das *Séries Exemplo***



**Figura 2.7: Dinâmica dos retornos e dos quadrados dos retornos nas *Séries Exemplo***

- a) Função de autocorrelação dos retornos nas *Ações*
- b) Função de autocorrelação dos retornos na *Taxa de Câmbio*
- c) Função de autocorrelação dos quadrados dos retornos nas *Ações*
- d) Função de autocorrelação dos quadrados dos retornos na *Taxa de Câmbio*

**Legenda:** (---) limites usuais de Bartlett de 1 e 2 desvios padrão



**Figura 2.8:** Função de autocorrelação amostral do modelo AR(1)-SV com parâmetros  $\gamma=0,95$ ,  $q=0,05$ ,  $k=5,5 \times 10^{-7}$

**Legenda:** (---) limites usuais de Bartlett de 2 desvios padrão.

# CAPÍTULO 3

## MODELOS ARCH

### 3.1 Introdução

Os modelos Autoregressivos de Variância Condicional Heteroscedástica (*ARCH*), foram originalmente introduzidos por Engle (1982) para modelar a inflação do Reino Unido. Em princípio, o interesse foi tentar modelar os *ruidos* ou perturbações de modelos mais complexos como os ARMA ou Regressão, com o objetivo de melhorar as estimativas e obter intervalos de confiança mais precisos. Posteriormente, essa abordagem, com suas várias extensões, passaram a ser fortemente associadas à modelagem de séries financeiras e, em 1992 Bollerslev *et alii* (1992), constatam mais de 300 aplicações bem sucedidas nesta área. Entre os modelos mais importantes temos: GARCH (Bollerslev,1986)<sup>1</sup>, IGARCH (Engle e Bollerslev,1986), EGARCH (Nelson,1991). Em particular, o modelo GARCH(1,1) passou a ser considerado como série *benchmark* para a análise da volatilidade nas séries financeiras (Bollerslev *et alii*, 1992). Referências atualizadas desses modelos são encontradas em Bollerslev *et alii* (1992) e Bollerslev *et alii* (1994).

Nos modelos ARCH a variância condicional e variância não condicional são distintas, sendo que a variância condicional depende apenas dos parâmetros e dos valores passados. Devido a esse fato, a verossimilhança é simples de calcular e a inferência estatística é facilitada. Mesmos assim, nas aplicações tem surgido vários problemas na estimação dos parâmetros em modelos de ordens maiores, pela presença de um grande número de restrições nos parâmetros.

Dentro da abordagem ARCH, o modelo mais simples é o ARCH(1), o qual é definido como:

$$Y_t / I_t \approx N(0, \sigma_t^2)$$
$$\sigma_t^2 = \omega + \alpha Y_{t-1}^2 ,$$

---

<sup>1</sup>Taylor (1986) propôs independentemente o modelo GARCH(1,1) chamando-o ARMACH

onde, como foi estabelecido,  $Y_t$  são os retornos e  $I_t$  a informação disponível até o instante  $t$ ,  $I_t = \{ Y_t, Y_{t-1}, \dots \}$ . Neste modelo *percebemos que grandes desvios (com respeito à média igual a zero) ou um grande valor do quadrado do retorno produz um aumento da variância condicional no instante posterior*. Modelar a variância condicional como dependente dos quadrados das observações tem um apelo natural e constitui uma forma muito elegante e simples de gerar conglomerados de volatilidade.

Como comentado anteriormente, nos modelos ARCH, a *volatilidade é dita observável* pois ela é conhecida quando os parâmetros são conhecidos. Essa é uma das características principais destes modelos. Uma vez estimados os parâmetros, por exemplo, pelo método de máxima verossimilhança, obtemos as estimativas da volatilidade.

A estrutura deste Capítulo é a seguinte: na Seção 3.2 são definidos os modelos ARCH de forma geral e apresentadas suas propriedades básicas, mostrando que são capazes de reproduzir os *Fatos Estilizados*. A seguir, na Seção 3.3, apresentamos 4 versões (extensões) dos modelos ARCH que são produto de especificações diferentes da variância condicional. Na seção seguinte, discutimos como é realizada a Inferência utilizando os métodos de Máxima Verossimilhança e Quase-Máxima Verossimilhança. Os mecanismos para verificar a qualidade dos modelos ajustados são assuntos da Seção 3.5 com a discussão dos Testes de Hipóteses e Diagnóstico. Na seção 3.6 apresentamos o processo de previsão, com ênfase na previsão um-passo à frente. Finalmente, na Seção 3.7 é realizada a aplicação da metodologia descrita no ajuste das *Séries Exemplo*.

### 3.2 Definições e Propriedades Básicas: *Fatos Estilizados*

Seja  $\{Y_t(\theta_0)\}$  um processo discreto (de retornos) com média e variância condicionais parametrizadas pelo vetor  $\theta \in \Theta \subset \mathbb{R}^m$ , com  $\theta_0$  sendo o valor verdadeiro valor<sup>2</sup>. Bollerslev *et alii* (1994) apresentam a seguinte definição para os modelos ARCH:

<sup>2</sup> Em todo o Capítulo consideramos  $\{Y_t\}$  no lugar de  $\{Y_t(\theta_0)\}$ .

*Definição 3.2.1*

Seja  $I_t = \{Y_t, Y_{t-1}, \dots, Y_1\}$  a informação disponível até o instante  $t$ . O processo  $\{Y_t\}$  segue um modelo ARCH se satisfaz:

- i)  $E[Y_t / I_{t-1}] = 0$ ,
- ii)  $\sigma_t^2 = \text{Var}[Y_t / I_{t-1}] = E[Y_t^2 / I_{t-1}]$  depende não trivialmente da  $\sigma$ -álgebra gerada pelas observações passadas ou informação anterior, isto é de  $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \dots, Y_1\}$ .

Se o modelo é ARCH, é conveniente expressá-lo como modelo produto:

$$Y_t = \sigma_t \varepsilon_t \tag{3.2.1}$$

no qual as variáveis  $\varepsilon_t = \frac{Y_t}{\sigma_t}$  constituem um *processo de inovações padronizado*. Com efeito, fazendo uso extensivo da propriedade:  $E[\cdot] = E[E[\cdot / I_{t-1}]]$ , é fácil mostrar isso:

- 1) Pela definição de  $\sigma_t^2$  como função de  $I_{t-1}$ , então:

$$E[\varepsilon_t / I_{t-1}] = E\left[\frac{Y_t}{\sqrt{\sigma_t^2}} / I_{t-1}\right] = \frac{1}{\sqrt{\sigma_t^2}} E[Y_t / I_{t-1}] = 0, \text{ portanto } E[Y_t] = 0.$$

$$2) E[\varepsilon_t \varepsilon_{t-\tau}] = E\left[E\left[\frac{Y_t}{\sigma_t} \frac{Y_{t-\tau}}{\sigma_{t-\tau}} / I_{t-1}\right]\right] = E\left[\frac{Y_{t-\tau}}{\sigma_t \sigma_{t-\tau}} E[Y_t / I_{t-1}]\right] = 0, \text{ pela definição 3.2.1.}$$

$$3) \text{Var}[\varepsilon_t / I_{t-1}] = E[\varepsilon_t^2 / I_{t-1}] = E\left[\frac{Y_t^2}{\sigma_t^2} / I_{t-1}\right] = \frac{1}{\sigma_t^2} E[Y_t^2 / I_{t-1}] = 1.$$

Dentro da análise dos modelos ARCH, *todo o processo de inferência está baseado no processo das inovações*  $\{\varepsilon_t\}$ . Usualmente é assumido que essas seguem uma distribuição invariante no tempo com quarto momento finito, como por exemplo, a normal. A suposição de invariância na distribuição é assumida em todo o Capítulo.

A seguir, apresentamos algumas propriedades que mostram como os modelos ARCH são capazes de reproduzir teoricamente alguns dos *Fatos Estilizados* apresentados pelas séries financeiras.

*Teorema 3.2.1*

Se o processo  $\{Y_t\}$  segue um modelo ARCH e as inovações  $\{\varepsilon_t\}$  possuem distribuição invariante no tempo com quarto momento finito então:

- i)*  $\{Y_t\}$  é martingale diferença.
- ii)* A autocovariância do processo  $\{Y_t\}$  é igual a zero.
- iii)* Se a variância não condicional do processo  $\{Y_t\}$  é finita então o processo  $\{Y_t\}$  é ruído branco.
- iv)* Se  $E[\varepsilon_t^4] > 3$ , a distribuição de  $\{Y_t\}$  tem cauda mais pesada do que a distribuição normal.
- v)* Quando os momentos de ordem par de  $\{Y_t\}$  existem, então a distribuição de  $\{Y_t\}$  é simétrica se a distribuição de  $\varepsilon_t$  também é.

*Prova:*

*i)* Do item *ii)* da Definição 3.2.1 temos que  $E|Y_t|$  é finito, logo o item *i)* da Definição 3.2.1 completa a prova.

*ii)* Como  $E[Y_t Y_{t-s}] = E[E[Y_t Y_{t-\tau} / I_{t-1}]] = E[Y_{t-\tau} E[Y_t / I_{t-1}]] = 0$ , e pelo item *i)* da definição 3.2.1 então  $Cov[Y_t, Y_{t-\tau}] = 0$ .

*iii)* Segue diretamente de *ii)* e da definição de ruído branco.

*iv)*  $E[Y_t^4] = E[E[\sigma_t^4 \varepsilon_t^4 / I_{t-1}]] = E[\sigma_t^4 E[\varepsilon_t^4 / I_{t-1}]] = E[\sigma_t^4] E[\varepsilon_t^4]$ . Por sua vez, pela desigualdade de Jensen temos que:  $E[\sigma_t^4] E[\varepsilon_t^4] \geq \{E[\sigma_t^2]\}^2 E[\varepsilon_t^4]$ , portanto  $E[Y_t^4] \geq \{E[\sigma_t^2]\}^2 E[\varepsilon_t^4]$ . Por outro lado:

$Var[Y_t] = E[Y_t^2] = E[E[\sigma_t^2 \varepsilon_t^2 / I_{t-1}]] = E[\sigma_t^2 E[\varepsilon_t^2 / I_{t-1}]] = E[\sigma_t^2]$ , em consequência a curtose de  $Y_t$  satisfaz:  $\frac{E[Y_t^4]}{\{E[Y_t^2]\}^2} \geq E[\varepsilon_t^4]$ . Finalmente, já que  $E[\varepsilon_t^4] > 3$  obtemos o resultado

desejado.

V) Por hipótese  $E[Y_t^{2m-1}] < E[Y_t^{2m}] < \infty$  para  $m=1,2,3,\dots$ . Por sua vez, como a distribuição de  $\varepsilon_t$  é simétrica temos que  $E[\varepsilon_t^{2m-1}] = 0$  para  $m=1,2,3,\dots$ . Em conseqüência:

$$E[Y_t^{2m-1}] = E[E[\sigma_t^{2m-1} \varepsilon_t^{2m-1} / I_{t-1}]] = E[\sigma_t^{2m-1} E[\varepsilon_t^{2m-1} / I_{t-1}]] = 0.$$

□

Do teorema anterior percebemos que, especificando adequadamente as inovações  $\varepsilon_t$ , os modelos ARCH podem gerar processos não correlacionados com média zero e com distribuição simétrica. Além disso, quando a variância é finita (o que depende da especificação da variância condicional) obtemos processos de ruído branco. Por sua vez, a propriedade de cauda pesada vale para quaisquer distribuição de  $\varepsilon_t$  sempre que possua quarto momento finito, a menos que a variância não condicional seja constante. Com efeito, na prova de iv) a igualdade na fórmula de Jensen só é satisfeita se  $E[\sigma_t^2]$  é constante.

A não-linearidade é mostrada examinando a f.a.c dos quadrados dos retornos, para o qual precisamos especificar o processo da variância condicional ou volatilidade. Na seção seguinte são apresentados os resultados para alguns modelos propostos.

Por outro lado, mesmo no caso mais simples do modelo ARCH(1), é  *muito difícil encontrar expressões para a distribuição conjunta dos retornos*. Essa é uma desvantagem desses modelos e constitui uma das características mais importantes com graves conseqüências na predição e inferência.

### 3.3 Quatro Versões dos Modelos ARCH

Nesta seção são apresentados 4 versões dos modelos ARCH junto com a discussão das propriedades mais importantes. Cada modelo é construído utilizando a expressão (3.2.1) onde o processo  $\{\varepsilon_t\}$  é assumido *i.i.d* com média zero e variância unitária e especificando a equação da volatilidade. Na literatura encontramos inúmeras propostas para a variância condicional, por exemplo em Glosten *et alii* (1993), Geweke (1986), Pantula (1986), Milloj (1987), Higgins e Bera (1992), Engle *et alii* (1987) e outras em Bollerslev *et alii* (1994).

### 3.3.1 ARCH (q)

Nos modelos ARCH (q) a variância condicional é definida como:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i Y_{t-i}^2, \quad \omega > 0 \quad \alpha_i > 0 \quad i=1,2,3,\dots, q. \quad (3.3.1.1)$$

onde as restrições sobre os parâmetros são necessárias para garantir que a variância condicional seja positiva. Por outro lado, definindo as inovações<sup>3</sup>:

$$v_t = Y_t^2 - E[Y_t^2 / I_{t-1}] = Y_t^2 - \sigma_t^2, \quad (3.3.1.2)$$

e substituindo-as em (3.3.1.1) obtemos:

$$Y_t^2 = \omega + v_t + \sum_{i=1}^q \alpha_i Y_{t-i}^2,$$

$$(1 - \sum_{i=1}^q \alpha_i L^i) Y_t^2 = \omega + v_t,$$

$$A(L)Y_t^2 = \omega + v_t, \text{ onde } A(L) = 1 - \sum_{i=1}^q \alpha_i L^i \text{ e } L \text{ é o operador de defasagem usual.}$$

Dessa forma, o processo dos quadrados dos retornos apresenta uma estrutura AR(q) com *drift*, o que constitui uma evidência da *não-linearidade* nos retornos. Podemos observar que se  $\alpha_i = 0$  para  $i=1,2,3,\dots,q$ , então a variância condicional é constante e, nesse caso o processo  $\{Y_t^2\}$  é ruído branco.

Sob a suposição que as inovações  $\{\varepsilon_t\}$  são normais, Engle (1982) demonstra que o processo dos retornos é estacionário, se e somente se, a equação característica associada  $A(L)=0$  possui raízes fora do círculo unitário. Nesse caso:

$$\text{Var}[Y_t] = E[Y_t^2] = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i}.$$

<sup>3</sup>Parametrização proposta por Pantula (1986)

O cálculo dos momentos de ordens maiores envolve expressões recursivas que podem ser encontradas em Engle (1982). A modo de ilustração apresentamos o resultados para o modelo ARCH(1):

$$\text{Var}[Y_t] = E[Y_t^2] = \frac{\omega}{1-\alpha_1}, \quad E[Y_t^4] = \frac{3\omega^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)},$$

em conseqüência,  $\text{Curtose}[Y_t] = \frac{3(1-\alpha_1^2)}{1-3\alpha_1^2} > 3$  para  $\alpha_1 > 0$ .

Observe que essa expressão não depende de  $\omega$  e por outro lado só é válida quando  $3\alpha_1^2 < 1$ , caso contrário, a variância é infinita.

Por outro lado, considerando que no ARCH(1) a série  $\{Y_t^2\}$  segue um processo AR(1), então as autocorrelações satisfazem:  $\rho_j = \alpha_1^j \leq (1/3)^{j/2} \leq 0,58$  (pois  $3\alpha_1^2 < 1$ ). Em conseqüência, mesmo que a primeira autocorrelação seja grande, os valores das outras autocorrelações tendem a zero rapidamente. Portanto, o modelo ARCH(1) não pode imitar o comportamento empírico dos quadrados dos retornos (ver Capítulo 2), os quais apresentam uma primeira autocorrelação pequena e as outras decaindo lentamente.

Já que o processo  $\{Y_t^2\}$  segue um AR(q), podemos utilizar a metodologia de Box-Jenkins para a identificação do modelo através da f.a.c e p.a.c.f. Nesse sentido, experimentos de simulação efetuados por Bollerslev (1988) baseados em 1000 replicações mostram que, mesmo com amostras pequenas de tamanho 100 e 200 a análise do p.a.c.f conduz à eleição correta da ordem do modelo ARCH. Entretanto, utilizando essa metodologia corre-se o risco de freqüentemente super-parametrizar o modelo. Similarmente ao que ocorre nos modelos AR(1)-SV, como discutido na Seção 2.7, os intervalos de confiança (i.c.) tradicionais subestimam a verdadeira largura dos i.c. das autocorrelações.

Como é referido em Bollerslev *et alii* (1992), nas aplicações tem-se encontrado duas dificuldades. A primeira está relacionada com a dificuldade de estimar os parâmetros com as restrições de forma a garantir que a variância condicional seja positiva. Nesse sentido, uma solução pouco satisfatória é a proposta por Engle (1982), na qual é imposta a seguinte estrutura para os parâmetros:  $\alpha_i = a_i \alpha$ , sendo que os valores de  $a_i$  são especificados de forma que a importância das observações passadas  $Y_t^2$  diminui na medida que o tempo aumenta,

$a_1 > a_2 > \dots > a_q$  com  $\sum_{i=1}^q a_i = 1$ . Mediante esse processo só temos que estimar um parâmetro;  $\alpha$ .

Estimar com restrições é uma das principais desvantagens dos modelos ARCH.

O segundo problema é que, na prática, tem-se encontrado a necessidade de especificar valores grandes de  $q$ , de forma a capturar o comportamento das f.a.c dos quadrados dos retornos que apresentam a primeira autocorrelação pequena e as outras diminuindo lentamente. Portanto, ao ajustar modelos ARCH( $q$ ), eles serão pouco parcimoniosos. Isso acarreta outra dificuldade que é a de estimar os parâmetros sob um grande número de restrições. Para contornar esse problema, precisamos encontrar modelos que possuam memória mais longa. Uma especificação natural constitui o modelo GARCH proposto por Bollerslev (1986), que é assunto da subseção seguinte.

### 3.3.2 GARCH (p,q)

Com o objetivo de obter processos com memória mais longa, modificamos a equação (3.3.1.1) da seguinte forma:

$$\sigma_t^2 = \kappa + \sum_{i=1}^{\infty} \pi_i Y_{t-i}^2 = \kappa + D(L)Y_t^2, \quad (3.3.2.1)$$

onde  $D(L) = \sum_{i=1}^{\infty} \pi_i L^i$ . Ao mesmo tempo, podemos parametrizar  $D(L)$  de forma a obter uma representação finita e parcimoniosa na forma:

$$D(L) = \frac{A(L)}{1 - B(L)}, \quad (3.3.2.2)$$

com  $A(L) = \sum_{i=1}^q \alpha_i L^i$ ,  $B(L) = \sum_{j=1}^p \beta_j L^j$  e as raízes de  $1 - B(L)$  fora do círculo unitário. Logo,

substituindo (3.3.2.2) em (3.3.2.1) obtemos:  $\sigma_t^2 = \kappa + \frac{A(L)}{1 - B(L)} Y_t^2$ , ou equivalentemente:

$[1 - B(L)]\sigma_t^2 = \kappa[1 - B(L)] + A(L)Y_t^2$ , isto é:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (3.3.2.3a)$$

$$\text{onde } p > 0 \quad q \geq 0 \quad \omega > 0 \quad \alpha_i \geq 0 \quad \beta_j \geq 0 \quad i=1,2,\dots,p \quad j=1,2,\dots,q. \quad (3.3.2.3b)$$

Esse modelo denominado *Modelo Generalizado Autoregressivo de Variância Condicional Heteroscedástica*; GARCH (p,q) foi proposto por Bollerslev (1986). Por sua vez, o GARCH(1,1) foi proposto independentemente por Taylor (1986) com o nome de ARMACH.

Neste modelo, a variância condicional é função tanto dos quadrados das observações quanto das variâncias condicionais passadas. Nesse sentido, é um modelo adaptativo mais sofisticado que o ARCH(q). Por sua vez, na expressão (3.3.2.3a) é fácil ver que se  $p=0$  então obtemos o modelo ARCH (q) definido em (3.3.1.1), e se  $p=q=0$  então o processo  $\{Y_t\}$  é ruído branco com variância condicional (e não condicional) constante e igual a  $\omega$ .

As condições (3.3.2.3b) foram impostas por Bollerslev (1986) para garantir que a variância condicional seja positiva. Entretanto, baseados na representação infinita do modelo GARCH (3.3.2.1), Nelson e Cao (1992) demonstraram que existem condições menos restritivas para algumas especificações. Em consequência, a presença de estimativas negativas dos parâmetros não significa que o modelo esteja mal especificado ou que as estimativas sejam ruins. Por exemplo, Nelson e Cao (1992) ajustaram o modelo GARCH (1,2) à série de taxa de câmbio Franco Francês/US\$, e encontraram uma estimativa negativa para o parâmetro  $\alpha_1$ , que está dentro da região admissível:  $\omega \geq 0$ ,  $0 \leq \beta < 1$ ,  $\beta\alpha_1 + \alpha_2 \geq 0$ ,  $\alpha_1 \geq 0$ . O modelo ajustado apresenta AIC e BIC maiores que o obtido quando utilizadas as condições de Bollerslev. Por sua vez, é importante ressaltar que para o caso dos modelos ARCH (q) e GARCH (1,1) não é possível relaxar as condições de Bollerslev.

Por outro lado, as condições para a estacionariedade do modelo GARCH(p,q) foram encontradas por Bollerslev (1986). Temos assim que o modelo GARCH(p,q) definido por (3.2.1) e (3.3.2.3) é covariante estacionário com variância igual a:

$$\text{Var}[Y_t] = \frac{\omega}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j}, \quad \text{se e somente se,} \quad \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1.$$

Por sua vez, Bougeroll e Picard (1992) apresentam condições para a estacionariedade e estacionariedade estrita dos modelos GARCH (p,q).

A evidência de não linearidade é encontrada através do estudo da série dos quadrados das observações. Assim, substituindo (3.3.2.3a) em (3.3.1.2) obtemos:

$$Y_t^2 = \omega + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + v_t$$

$$Y_t^2 = \omega + \sum_{i=1}^m (\alpha_i + \beta_i) Y_{t-i}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t, \quad (3.3.2.4)$$

$$Y_t^2 = \omega + [A(L) + B(L)]Y_t^2 + [1 - B(L)]v_t, \quad (3.3.2.4a)$$

onde  $m = \max\{p, q\}$ . Dessa forma,  $\{Y_t^2\}$  segue um processo ARMA(m,p) com *drift*. Portanto, a série dos retornos é não correlacionada mas não independente.

Por outro lado, se fazemos  $q=0$  em (3.3.2.4) obtemos:

$$Y_t^2 = \omega + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + v_t,$$

$$[1 - \sum_{j=1}^p \beta_j L^j] Y_t^2 = \omega + [1 - \sum_{j=1}^p \beta_j L^j] v_t,$$

isto é,  $Y_t^2 = \omega [1 - \sum_{j=1}^p \beta_j L^j]^{-1} + v_t$ , em consequência, o modelo é não identificável. Portanto, a menos que  $p=0$ ,  $q$  tem que ser maior que zero.

Tal como nos modelos ARCH(q), o processo de identificação pode ser realizado aproveitando a estrutura de autocorrelações e autocorrelações parciais dos quadrados das observações que exibem um comportamento similar aos modelos ARMA(p,q). No caso do GARCH (1,1) experimentos de simulação efetuados por Bollerslev (1988) mostram que os resultados são satisfatórios. No entanto, na literatura é freqüente a escolha do modelo através de critérios como o AIC ou BIC, ver Seção 3.5.

Como será discutido na subseção 3.5, o diagnóstico do modelo é realizado através da estimativa dos resíduos, definidos como  $\hat{\varepsilon}_t = \frac{\hat{Y}_t}{\hat{\sigma}_t}$ . Nas aplicações empíricas com modelos GARCH, supondo inovações normais, tem-se encontrado que a distribuição dos resíduos

apresenta cauda pesada. Para contornar esse problema Bollerslev (1987) introduz o uso de inovações com distribuição *t-Student Padronizada*<sup>4</sup>, cuja função de densidade é:

$$f(\varepsilon_i) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu-2}} \frac{1}{\left[1 + \frac{\varepsilon_i^2}{\nu-2}\right]^{\frac{\nu+1}{2}}} \quad (3.3.2.5)$$

onde  $\nu > 4$  para garantir cauda pesada.

Outro aspecto observado empiricamente nos ajustes de modelos GARCH em séries financeiras é que os valores de  $p$  e  $q$  raramente são maiores que dois e que a contribuição dos termos de ordem superior a um são insignificantes (Pagan, 1994). Nesse sentido, de longe o modelo mais utilizado e com bastante sucesso é o modelo GARCH(1,1). A seguir apresentamos uma breve discussão deste modelo.

### 3.3.2.1 GARCH (1,1)

No modelo GARCH(1,1), a variância condicional segue a fórmula:

$$\sigma_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (3.3.2.1.1)$$

O cálculo dos momentos envolve operações recursivas. Bollerslev (1986) mostrou que uma condição necessária e suficiente para a existência de momentos de ordem  $2m$  é <sup>5</sup>:

$$\eta = \sum_{j=0}^m \binom{m}{j} a_j \alpha^j \beta^{m-j} < 1, \quad a_0 = 1, \quad a_j = \prod_{i=1}^j (2i-1) \quad j=1,2,\dots$$

nesse caso, temos que:

$$E[Y_t^{2m}] = a_m \left[ \sum_{n=0}^{m-1} a_n^{-1} E(Y_t^{2n}) \alpha_o^{m-n} \binom{m}{m-n} \eta \right] [1-\eta]^{-1}.$$

<sup>4</sup>Possue média zero e variância unitária.

<sup>5</sup>Assumindo que o processo começa no passado com os primeiros  $2m$  momentos finitos.

Em particular, se  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$  então o quarto momento existe e obtemos:

$$E[Y_t^2] = \frac{\omega}{1 - \alpha - \beta}, \quad E[Y_t^4] = \frac{3\omega^2(1 + \alpha + \beta)}{(1 - \alpha - \beta)(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)}. \text{ Portanto, a curtose é igual a:}$$

$$\text{Curt}[Y_t] = \frac{3(1 - (\alpha + \beta)^2)}{1 - \beta^2 - 2\alpha\beta - 3\alpha^2} \text{ que é maior que 3 para todo } \alpha \geq 0.$$

Nelson (1990)<sup>6</sup> e Bougeroll e Picard (1992) derivam uma condição necessária e suficiente para provar a ergodicidade e estacionaridade estrita do modelo GARCH(1,1). Sendo  $\omega > 0$ , o modelo GARCH(1,1) definido em (3.3.2.1.1) é ergódico estritamente estacionário se e somente se  $E[\log(\beta + \alpha\varepsilon_t^2)] < 0$  onde  $\{\varepsilon_t\}$  é o processo de inovações padronizado definido em (3.2.1). Adicionalmente, já que a função  $\log(\cdot)$  é côncava, então pela desigualdade de Jensen temos que:  $E[\log(\beta + \alpha\varepsilon_t^2)] < \log(E[\beta + \alpha\varepsilon_t^2]) = \log(\beta + \alpha)$ . Logo, se  $\alpha + \beta < 1$ , então  $E[\log(\beta + \alpha\varepsilon_t^2)] < 0$  e, conseqüentemente, o modelo GARCH(1,1) é estritamente ergódico e covariante estacionário.

Pela expressão (3.3.2.4) o processo  $\{Y_t^2\}$  no GARCH(1,1) é da forma ARMA(1,1) com parâmetro autoregressivo  $(\alpha + \beta)$  e parâmetro de média móvel  $(-\beta)$ . Portanto, supondo que o quarto momento de  $\{Y_t\}$  é finito, a p.a.c.f é em geral finita e dominada por uma exponencial e, a f.a.c é:

$$\rho_{Y_t^2}(\tau) = \begin{cases} \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} & \tau=1 \\ (\alpha + \beta)^{\tau-1} \rho_1 & \end{cases} \quad (3.3.2.1.2)$$

Supondo que o quarto momento existe é possível mostrar (Bollerslev, 1988) que as duas primeiras autocorrelações satisfazem:

$$\begin{aligned} \rho_2 &\geq \rho_1^2 \\ \rho_2 &< \rho_1, \quad 0 \leq \rho_1 \leq 1/3 \\ \rho_2 &< (2/3 + 2^{3/2} \rho_1 (1/3 - \rho_1^2)^{1/2} - \rho_1^2)^{1/2} \rho_1 \quad 1/3 \leq \rho_1 \leq \sqrt{1/3} = 0,58. \end{aligned}$$

<sup>6</sup>fornece fórmulas para avaliar as condições de estacionaridade no caso que a distribuição condicional dos retornos seja normal e Cauchy.

Analogamente ao ARCH(1), a primeira autocorrelação é menor que 0,58.

Quando  $\alpha$  é pequeno e  $\alpha + \beta$  grande, a f.a.c dos quadrados das observações num modelo GARCH(1,1) apresenta um comportamento que aparece freqüentemente nas séries financeiras (assim como também as *Séries Exemplo*), isto é, a presença de uma primeira autocorrelação pequena e as outras diminuindo lentamente. Esse é um dos motivos do sucesso do modelo GARCH(1,1) na modelagem das séries financeiras. Nas aplicações, freqüentemente o ajuste dos modelos apresenta valores para  $\alpha + \beta$  muito próximos de 1<sup>7</sup>. As conseqüências desse fato são exploradas na subseção seguinte.

### 3.3.3 IGARCH (p,q)

Levando em consideração o fato de que nas aplicações empíricas com modelos GARCH(p,q) as estimativas de  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j$  são muito próximas de um, Engle e Bollerslev (1986) propõem o modelo *GARCH Integrado*: IGARCH, o qual é definido como (3.3.2.3a) sujeito à restrição  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$ . O modelo mais utilizado nas aplicações empíricas é o IGARCH(1,1).

Existem duas conseqüências importantes da definição desse modelo. A primeira é que pelo Teorema de Bollerslev, o modelo IGARCH não é covariante estacionário e, a segunda é que o efeito dos *choques* na variância condicional permanece indefinidamente. Este fato está relacionado ao conceito de *Persistência*, que será discutido logo a seguir. Com efeito, pela expressão (3.3.2.3a) com a condição  $A(L)+B(L)=1$ , temos que:

$$Y_t^2 = \omega + A(L)Y_t^2 + B(L)\sigma_t^2,$$

$$\sigma_t^2 = \frac{\omega}{1 - B(L)} + Y_t^2,$$

<sup>7</sup> Ver Bollerslev *et alii* (1992) e Taylor (1986). Nesse último estudo são ajustados modelos GARCH(1,1) ou ARMACH para 40 séries financeiras. Com exceção de 6 séries, os valores de  $\alpha + \beta$  são maiores ou iguais a 0,97. Entretanto, como mostram Lamoreaux e Lastrapes (1990) no estudo de 30 séries de ações, ao incorporar na variância condicional do GARCH(1,1) mudanças estruturais, as estimativas de  $\alpha + \beta$  são bem mais baixas quando comparadas com as estimativas sem considerar as mudanças estruturais.

e, considerando que  $Y_t^2 = \varepsilon_t^2 \sigma_t^2$  obtemos finalmente,

$$\sigma_t^2 = \frac{\omega}{(1 - B(L))(1 - \varepsilon_t^2)},$$

portanto o efeito dos choques na variância permanece indefinidamente.

Apesar dos modelos IGARCH não serem covariante estacionários, Geweke (1986) e Nelson (1990) mostraram o importante resultado que esses modelos são estritamente estacionários e ergódicos. A existência de modelos com essas características dificulta uma definição única de persistência, sendo necessário, portanto, distinguir dois tipos de persistência<sup>8</sup>, os quais são definidos considerando a natureza da convergência dos processos  $\{\sigma_t^2\}$ .

*i)* Convergência em distribuição: considere uma sequência  $\{\sigma_t^2\}$  e denote por  $F(\sigma_t^2)$  a função de distribuição acumulada (f.d.a.) de  $\{\sigma_t^2\}$  e por  $F_s(\sigma_t^2)$  a f.d.a. de  $\{\sigma_t^2\}$  dada a informação até o tempo  $s < t$ ,  $I_s$ . Caso, para qualquer  $s$   $\{F_s(\sigma_t^2) - F(\sigma_t^2)\}$  convergir para 0 quase certamente para todo ponto de continuidade quando  $t \rightarrow \infty$ , isto é, a informação no tempo  $s$  não influencia a distribuição do processo quando  $t \rightarrow \infty$ , dizemos que os choques não persistem na distribuição de  $\{\sigma_t^2\}$ . Uma condição suficiente para que ocorra essa convergência é que  $\{\sigma_t^2\}$  seja ergódico e estritamente estacionário.

*ii)* Convergência nos momentos: a influência dos choques não persiste indefinidamente em  $\{\sigma_t^2\}$  se para quaisquer  $s$ ,  $E[\sigma_t^{2n} / I_s]$  converge a uma constante que independe de  $I_s$  quando  $t \rightarrow \infty$ .

Como no GARCH(1,1),  $E[\sigma_t^2 / I_s]$  converge à variância não condicional  $\frac{\omega}{1 - \alpha - \beta}$  quando  $t \rightarrow \infty$  e além disso o processo  $\{\sigma_t^2\}$  é estritamente estacionário e ergódico, consequentemente sob os critérios *i)* e *ii)* concluímos que no GARCH(1,1), o efeito dos choques não permanece indefinidamente na variância condicional. Em várias especificações GARCH(p,q), existe conflito entre os critérios *i)* e *ii)* pelo fato, já mencionado, de que os processos são estritamente estacionários e ergódicos, mas podem não possuir momentos finitos (para alguns valores de  $n$ ). Já nos modelos IGARCH, os choques não persistem segundo o critério de distribuição *i)*, mas persistem segundo o critério *ii)* pois  $E[\sigma_t^{2n} / I_s] \rightarrow \infty$  quando  $t \rightarrow \infty \forall n$ .

<sup>8</sup> Em Nelson (1990) são introduzidos os tipos de persistência.

Embora o efeito dos choques possa não influir indefinidamente, o decaimento da sua influência pode apresentar diferentes velocidades. Desta forma, além da escolha do tipo de persistência é necessário encontrar uma forma de medir essa velocidade de decaimento. Como anotam Bollerslev *et alii* (1994), na prática, a escolha do tipo de persistência depende dos objetivos do estudo (estimação ou previsão) e por outro lado, da facilidade para calculá-la. Assim, no modelo GARCH(1,1) a persistência ( $\rho$ ) é usualmente definida como  $\alpha + \beta$ . Nesta definição um valor igual a 1 indica não convergência em momentos.

Por outro lado, uma medida auxiliar que permite melhor interpretação da persistência  $\rho$  é a *meia-vida* (*half-life*);  $mv$ , a qual é definida como:

$$mv = \frac{-\log(2)}{\log(\rho)} \quad (3.3.3.1)$$

A meia-vida mede o período de tempo (por exemplo número de dias de funcionamento de mercado) necessário para que o efeito de um choque na variância condicional diminua até a metade.

### 3.3.4 EGARCH

Motivado pelos trabalhos de Pantula (1986) e Geweke (1986), Nelson (1988, 1991) propôs o modelo *GARCH Exponencial*; EGARCH, na tentativa de sanar as seguintes deficiências dos modelos GARCH:

- i) Apresentam dificuldades na estimação dos parâmetros com restrições.
- ii) Não incorporam Assimetria: diversos estudos como os de Black (1976) e Christie (1982), e as referências em Bollerslev *et alii* (1992, p.24) encontraram evidência que os ativos são correlacionados negativamente com mudanças na volatilidade, de forma que os retornos aumentam quando há um excesso de retornos menor que o esperado e tendem a cair quando há um excesso de retornos maior que o esperado. Conforme visto no Capítulo 2 isso é conhecido como *Leverage Effect*.

iii) Dificuldades em interpretar a Persistência já que alguns modelos (I)GARCH podem ser estritamente estacionários mas não covariante estacionários.

O modelo EGARCH é definido seguindo a forma usual de modelo produto como:

$$Y_t = \sigma_t \varepsilon_t \quad \text{com } \{\varepsilon_t\} \text{ i.i.d e } E[\varepsilon_t] = 0, \text{ Var}[\varepsilon_t] = 1,$$

no qual a variância condicional está definida por<sup>9</sup>:

$$\ln(\sigma_t^2) = \kappa_t + \sum_{k=1}^{\infty} \phi_k g(\varepsilon_{t-k}) \quad (3.3.4.1a)$$

com a função  $g(\cdot)$  especificada usualmente como:

$$g(\varepsilon_t) = \theta \varepsilon_t + |\varepsilon_t| - E|\varepsilon_t| \quad (3.3.4.1b)$$

onde  $\{\kappa_t\}_{t=-\infty}^{t=+\infty}$  e  $\{\phi_k\}_{k=1}^{\infty}$  são seqüências não estocásticas de números reais e  $\theta$  é uma constante real. O processo  $\{g(\varepsilon_t)\}_{t=-\infty}^{t=+\infty}$  é *i.i.d* com média zero e, portanto, constitui a inovação do processo da variância condicional em (3.3.4.1a). Por sua vez, supondo  $\{\kappa_t\}$  como uma seqüência de constantes o modelo é mais simples.<sup>10</sup>

Com o objetivo de incorporar *choques* com distribuições que apresentam cauda pesada, Nelson (1991) sugere o uso da Distribuição do Erro Generalizada (*Generalized Error Distribution, GED*), sendo a versão padronizada a seguinte:

$$f(\varepsilon_t) = \frac{v \exp\left[-0.5 \left|\frac{\varepsilon_t}{\lambda}\right|^v\right]}{\lambda 2^{\frac{v+1}{v}} \Gamma\left(\frac{1}{v}\right)} \quad \text{com} \quad \lambda = \sqrt{2^{\left(\frac{-2}{v}\right)} \Gamma\left(\frac{1}{v}\right) \Gamma\left(\frac{-3}{v}\right)} \quad (3.3.4.5)$$

onde

$$E[|\varepsilon_t|] = \frac{\lambda 2^{1/v} \Gamma(2/v)}{\Gamma(1/v)},$$

<sup>9</sup> Parametrização diferente do modelo proposto por Nelson (1991)

<sup>10</sup> Nelson(1991) numa aplicação empírica utiliza a seqüência  $\{k_t\}$  como não constantes para medir o efeito dos dias nos quais o mercado não funciona.

Nesta distribuição,  $\nu$  é o parâmetro responsável pelo peso das caudas. Assim, quando  $\nu > 2$  ( $\nu < 2$ ) a distribuição apresenta caudas mais (menos) pesadas que a normal. Se  $\nu = 2$  a distribuição GED torna-se normal padrão e, nesse caso,  $E[|\varepsilon_t|] = \sqrt{2/\pi}$ .

A discussão a seguir visa comentar as propriedades do modelo EGARCH que tratam as deficiências dos GARCH citadas anteriormente. Primeiro, com relação ao item *i*) fica claro pela definição (3.3.4.1) que na formulação do modelo EGARCH, não precisamos restrições nos parâmetros para garantir que a variância condicional seja positiva.

Por outro lado, o efeito assimétrico dos *choques* na variância condicional é introduzido nas inovações  $\{g(\varepsilon_t)\}$  através do parâmetro  $\theta$ . Com efeito, em (3.3.4.1b) temos que:

$$g(\varepsilon_t) = \begin{cases} (\theta + 1)\varepsilon_t & \varepsilon_t > 0 \\ (\theta - 1)\varepsilon_t & \varepsilon_t < 0 \end{cases}$$

Se  $\theta = 0$ , um *choque positivo* ( $\varepsilon_t > 0$ ) produz o mesmo efeito na volatilidade que um *choque negativo* da mesma magnitude. Por sua vez, se  $-1 < \theta < 0$ , um *choque positivo* aumenta a volatilidade menos que um *choque negativo*. Finalmente se  $\theta < -1$ , um choque positivo reduz a volatilidade enquanto que o choque negativo aumenta a volatilidade. Portanto, em geral se  $\theta < 0$  temos evidência de *Leverage Effect*.

Finalmente, Nelson (1991, p. 351-353) demonstra através de dois teoremas a existência de coerência nas definições de persistência *i*) e *ii*) da Seção 3.3.3. Apresentamos isso na forma de um Teorema:

*Teorema 1:*

Quando a distribuição de  $\{\varepsilon_t\}$  é GED com  $\nu > 1$  (isto é, com cauda menos pesada que a cauda da exponencial dupla) e se  $\sum_{k=1}^{\infty} \phi_k^2 < \infty$  então, o processo das observações  $\{Y_t\}$  e o processo das variâncias condicionais  $\{\sigma_t^2\}$  é ergódico, estritamente estacionário e possui todos os momentos finitos, o qual implica, por sua vez, ser covariante estacionários.

Por outro lado, na modelagem, precisamos expressões parcimoniosas de (3.3.4.1). Nesse sentido, é conveniente representar o modelo na forma ARMA. Assim, de (3.3.4.1):

$$\ln(\sigma_t^2) = \kappa_t + \pi(L)g(\varepsilon_{t-1}) \quad (3.3.4.7)$$

no qual :  $\pi(L) = \phi_1 + \phi_2 L + \phi_3 L^2 + \dots$ . Dessa forma, parametrizando esse polinômio como:

$$\pi(L) = \frac{\sum_{i=1}^{q+1} \psi_i L^{i-1}}{1 - \sum_{j=1}^p \Delta_j L^j} \quad (3.3.4.8)$$

e substituindo (3.3.4.8) em (3.3.4.7) obtemos:

$$\left[1 - \sum_{j=1}^p \Delta_j L^j\right] \ln(\sigma_t^2) = \kappa_t + \left[\sum_{i=1}^{q+1} \psi_i L^{i-1}\right] g(\varepsilon_{t-1}).$$

Portanto, podemos definir o modelo EGARCH (p,q) como:

$$\ln(\sigma_t^2) = \kappa_t + \sum_{j=1}^p \Delta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^{q+1} \psi_i g(\varepsilon_{t-i}) \quad (3.3.4.9)$$

Esse modelo é identificável quando os polinômios  $\left[\sum_{i=1}^{q+1} \psi_i L^{i-1}\right]$  e  $\left[1 - \sum_{j=1}^p \Delta_j L^j\right]$  possuem raízes diferentes. Dessa forma, pelo Teorema 1, os processos  $\{\sigma_t^2\}$  e  $\{Y_t\}$  são estritamente estacionários e ergódicos se e somente se as raízes de  $\left[1 - \sum_{j=1}^p \Delta_j L^j\right]$  estão fora do círculo unitário.

Por outro lado, o fato de trabalhar com o logaritmo da variância condicional e a forma de introduzir a assimetria fazem com que as expressões dos momentos para os modelos EGARCH sejam complexas. Essas expressões podem ser encontradas em Nelson (1991) supondo que a distribuição das inovações é GED.

### 3.4 Estimação

Entre os métodos paramétricos para estimar modelos ARCH, o mais estudado e ao mesmo tempo o mais utilizado nas aplicações empíricas é o Método de Máxima Verossimilhança (ML). Esse procedimento é desenvolvido sob a suposição de que é conhecida a distribuição das inovações; seja normal, *t-Student* ou do tipo GED. Outra alternativa é fazer uso da metodologia desenvolvida por White (1982) nomeada *Método de Quase-Máxima Verossimilhança (QML)* na qual assume-se que a distribuição das inovações pertence a uma família *próxima*<sup>11</sup> à normal. O objetivo da presente seção é apresentar de uma forma breve e sem rigor a estimação nos modelos ARCH através dos métodos ML e QML. Essas metodologias são desenvolvidas para o caso mais geral de um modelo de regressão com perturbações que evoluem como um modelo ARCH<sup>12</sup>. O conteúdo desta seção segue basicamente a abordagem de Bollerslev *et alii* (1994).

#### 3.4.1 Método de Máxima Verossimilhança

Seja o processo dos retornos  $\{Y_t\}$  e a amostra disponível:  $\{y_1, y_2, \dots, y_T\}$ . Trabalhamos com o modelo:

$$Y_t = \mu_t + Y_t^* \quad (3.4.1.1a)$$

$$Y_t^* = \sigma_t \varepsilon_t \quad (3.4.1.1b)$$

$$h_t = \sigma_t^2 \quad (3.4.1.1c)$$

onde  $\mu_t = E[Y_t / I_{t-1}]$  é a média condicional do processo, a qual pode admitir defasagens e  $\{\varepsilon_t\}$  é um processo *i.i.d* com  $E[\varepsilon_t] = 0$ ,  $\text{Var}[\varepsilon_t] = 1$  e densidade  $f(\varepsilon)$ . Como foi mencionado anteriormente, todo o processo de inferência está baseado nas variáveis  $\{\varepsilon_t\}$ .

O vetor de parâmetros do modelo (3.4.1.1),  $\Psi' \equiv (\Theta', \Gamma)$  possui ordem  $(m+k) \times 1$ , onde  $\Theta$  corresponde aos parâmetros da média e variância condicionais do ARCH, cuja ordem é  $(m \times 1)$  e  $\Gamma$  corresponde aos parâmetros de perturbação da densidade  $f(\varepsilon)$ .

<sup>11</sup>A distribuição está a uma distância específica seguindo a métrica de Kullback-Leibler.

<sup>12</sup>Os mesmos princípios podem ser utilizados na derivação dos resultados para um modelo ARMA com inovações ARCH. O primeiro tratamento desse modelo foi dado por Weiss (1984).

A função de distribuição conjunta das observações é dada por:

$$f(y_1, \dots, y_T) \equiv f(y_1)f(y_2 / y_1) \dots f(y_T / y_1, y_2, \dots, y_{T-1}),$$

e o Log-Verossimilhança:

$$\begin{aligned} L(\Psi) &\equiv \log f(y_1, y_2, \dots, y_T) \\ &= \log f(y_1) + \sum_{i=2}^T \log f(y_i / y_1, \dots, y_{i-1}) = \log f(y_1) + \sum_{i=2}^T \log f(y_i / I_{i-1}). \end{aligned} \quad (3.4.1.2)$$

Existem duas formas de tratar a expressão anterior. Frequentemente, a contribuição de  $\log f(y_1)$  na somatória em (3.4.1.2) é desprezível, e pode ser considerado como zero. No ARCH(1), experimentos de simulação com amostras pequenas (100 observações) realizados por Diebold e Schuermann (1992) mostram que a inclusão do termo  $\log f(y_1)$  na estimação, produz modificações insignificantes nas estimativas. Um segundo enfoque é considerar a fórmula (3.4.1.2) como a log-verossimilhança *condicional à observação*  $y_i$ . Como observam Bollerslev *et alii* (1994, p.21) essa abordagem não compromete as propriedades assintóticas dos estimadores. Em consequência, na parte restante do Capítulo consideramos que:

$$L(\Psi) = \sum_{i=2}^T \log f(y_i / I_{i-1}) \quad (3.4.1.3)$$

É conveniente expressar a relação anterior como função das inovações  $\{\varepsilon_i\}$ . Nesse sentido, em (3.4.1.1) considerando que  $y_i = \mu_i + \sigma_i \varepsilon_i = \varphi(\varepsilon_i)$ <sup>13</sup>, então  $\varepsilon_i = \varphi^{-1}(y_i) = \frac{y_i - \mu_i}{\sigma_i}$  e:

$$\frac{\partial \varphi^{-1}(y_i)}{\partial y_i} = \frac{1}{\sqrt{h_i}} = \left| \frac{1}{h_i} \right|^{1/2}, \text{ portanto:}$$

$$f_{y_i/I_{i-1}}(w) = f_{\varepsilon_i/I_{i-1}}(\varphi^{-1}(w)) \left| \frac{\partial}{\partial w} \varphi^{-1}(w) \right|,$$

$$f_{y_i/I_{i-1}}(w) = f_{\varepsilon_i/I_{i-1}}\left(\frac{w - \mu_i}{\sigma_i}\right) \frac{1}{\sqrt{h_i}},$$

<sup>13</sup> O símbolo  $\varphi(x)$  evidencia que a variável em questão é função de  $x$ .

$$\log f_{y_i/I_{i-1}}(w) = \log f_{\varepsilon_i/I_{i-1}}\left(\frac{w - \mu_i}{\sigma_i}\right) - \frac{1}{2} \log(h_i) ,$$

na discussão posterior não colocamos explicitamente a variável  $w$  a qual é suposta entendida. Assim:

$$\log f(y_i / I_{i-1}) = \log f(\varepsilon_i / I_{i-1}) - 0.5 \log(h_i) \quad (3.4.1.4)$$

e, substituindo (3.4.1.4) em (3.4.1.3) obtemos:

$$L(\Psi) = \sum_{i=1}^T \log f(\varepsilon_i / I_{i-1}) - 0.5 \sum_{i=1}^T \log(h_i) \quad (3.4.1.5)$$

Para um valor  $\Psi_0 = (\Gamma_0, \Theta_0)$  condicionada a algumas informações iniciais, a seqüência de variâncias condicionais  $h_i$  pode ser calculada e por sua vez, utilizadas no cálculo da log-verossimilhança em (3.4.1.5). Por exemplo, no GARCH (1,1) precisamos conhecer  $h_{-p+1}$ ,  $h_{-p+2}$ , ...,  $h_0$  e  $y_{-p+1}^*$ ,  $y_{-p+2}^*$ , ...,  $y_0^*$ . Nesse caso, Bollerslev (1986) sugere o uso de:

$$h_j = \hat{\sigma}^2 \quad j = -p+1, -p+2, \dots, 0 \quad , \quad (3.4.1.6)$$

onde  $\hat{\sigma}^2$  é a variância amostral de  $\{y_i^*\}$   $i=1,2,\dots,T$ , enquanto Shephard (1994b) sugere utilizar a variância amostral dos primeiros 20 valores.

Uma vez calculada a verossimilhança, essa pode ser maximizada utilizando algoritmos apropriados de otimização. Entre outros temos o *Método do Escore*, utilizado por Engle (1982, p.997) e o *Algoritmo BHHH* (Berndt, Hall, Hall e Hausman, 1974) utilizado por Bollerslev (1986, p.317).

A principal quantidade envolvida no processo de maximização é o *Gradiente*, o qual pode ser calculado numericamente através de diferenciação numérica em (3.4.1.4), ou de maneira analítica. Com efeito, assumindo que  $\log f(\varepsilon_i / I_{i-1})$ ,  $\mu_i$  e  $h_i = \sigma_i^2$  são diferenciáveis em  $\Psi$  obtemos que o gradiente da log-verossimilhança (VL) é igual a:

$$\nabla L(\Psi) \equiv \nabla L(y_1, \dots, y_T / \Psi) \equiv \sum_{i=1}^T s_i(y_i / \Psi), \quad (3.4.1.7a)$$

onde os *escores*  $s_i$ , de ordem  $(m+k) \times 1$  estão definidos como:

$$s_i(y_i / \Psi) \equiv \frac{\partial \log f(y_i / I_{i-1})}{\partial \Psi} \quad (3.4.1.7b)$$

Logo, o estimador ML é a solução das  $(m+k)$  equações não-lineares dadas por  $\nabla L(\Psi) = 0$ , e pode ser calculado utilizando, por exemplo, os algoritmos mencionados anteriormente. Explicitando a equação (3.4.1.7b) com ajuda de (3.4.1.4) temos:

$$s_i(y_i / \Psi) = \frac{\partial \log f(\varepsilon_i / I_{i-1})}{\partial \Psi} - \frac{1}{2} \frac{\partial \log(h_i)}{\partial \Psi}$$

Nas aplicações, a densidade das inovações  $\varepsilon_i$  é invariante no tempo. Na discussão a seguir consideramos  $f(\varepsilon_i / I_{i-1}) = f(\varepsilon_i)$ . Por sua vez, já que  $f(\varepsilon_i)$  depende só de  $\Gamma$  e  $h_i$  depende só de  $\Theta$ , então trabalhando na expressão anterior obtemos:

$$s_i(y_i / \Psi) = \frac{1}{f(\varepsilon_i)} \frac{\partial f(\varepsilon_i)}{\partial \Psi} - \frac{1}{2h_i} \frac{\partial h_i}{\partial \Psi} = \frac{f'(\varepsilon_i)}{f(\varepsilon_i)} \frac{\partial \varepsilon_i}{\partial \Theta} - \frac{1}{2h_i} \frac{\partial h_i}{\partial \Theta} \quad (3.4.1.8)$$

onde  $f'(\varepsilon_i) = \frac{\partial f(\varepsilon_i)}{\partial \Gamma}$ . Para avaliar efetivamente  $s_i$  resta calcular  $\frac{\partial \varepsilon_i}{\partial \Theta}$ , assim:

$$\frac{\partial \varepsilon_i}{\partial \Theta} = \frac{\partial}{\partial \Theta} \left( \frac{Y_i^*}{\sqrt{h_i}} \right) = Y_i^* \frac{\partial}{\partial \Theta} \left( \frac{1}{\sqrt{h_i}} \right) + \frac{1}{\sqrt{h_i}}$$

e como  $\frac{\partial Y_i^*}{\partial \Theta} = \frac{\partial (Y_i - \mu_i)}{\partial \Theta} = -\frac{\partial \mu_i}{\partial \Theta}$ ,  $\frac{\partial}{\partial \Theta} \left( \frac{1}{\sqrt{h_i}} \right) = -\frac{1}{2} h_i^{-3/2} \frac{\partial h_i}{\partial \Theta}$  então:

$$\frac{\partial \varepsilon_i}{\partial \Theta} = -\frac{1}{2} Y_i^* h_i^{-3/2} \frac{\partial h_i}{\partial \Theta} - \frac{1}{\sqrt{h_i}} \frac{\partial \mu_i}{\partial \Theta} \quad (3.4.1.9)$$

em consequência, substituindo (3.4.1.9) em (3.4.1.8):

$$s_i(y_i / \Psi) = \frac{f'(\varepsilon_i)}{f(\varepsilon_i)} \left[ -\frac{1}{2} Y_i^* h_i^{-3/2} \frac{\partial h_i}{\partial \Theta} - \frac{1}{\sqrt{h_i}} \frac{\partial \mu_i}{\partial \Theta} \right] - \frac{1}{2} \frac{1}{h_i} \frac{\partial h_i}{\partial \Theta}$$

$$s_i(y_i / \Psi) = \frac{1}{2h_i} \frac{\partial h_i}{\partial \Theta} \left[ \frac{-f'(\varepsilon_i)}{f(\varepsilon_i)} - 1 \right] - \frac{f'(\varepsilon_i)}{f(\varepsilon_i)} \frac{1}{\sqrt{h_i}} \frac{\partial \mu_i}{\partial \Theta}, \quad (3.4.1.10)$$

com  $\varepsilon_i = \frac{Y_i - \mu_i}{\sqrt{h_i}}$ .

Substituindo (3.4.1.10) em (3.4.1.7a) obtemos o Gradiente  $\nabla L(\Psi)$  e, dessa forma, o estimador de máxima verossimilhança é encontrado resolvendo as  $(m+k)$  equações não-lineares de  $\nabla L(\Psi) = 0$ . Finalmente, uma vez especificado  $f(\varepsilon_i)$ , encontramos  $\frac{f'(\varepsilon_i)}{f(\varepsilon_i)}$ . Na Tabela 3.1 apresentamos os resultados para as três densidades usualmente utilizadas como inovações nos modelos ARCH que foram apresentadas na seção anterior.

**Tabela 3.1**  
**Valores de  $\frac{f'(\varepsilon)}{f(\varepsilon)}$  para as densidades normal, *t-Student* e GED**

Densidade	Parâmetros de Perturbação	$f(\varepsilon)$	$\frac{f'(\varepsilon)}{f(\varepsilon)}$
Normal	Nenhum <sup>1</sup>	$f(\varepsilon) = \frac{1}{\sqrt{2\pi}} e^{-0.5\varepsilon^2}$	$-\varepsilon$
<i>t Student</i> <sub>(v)</sub>	$v^a$	$f(\varepsilon_i) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})} \frac{1}{\sqrt{v-2} \left[1 + \frac{\varepsilon_i^2}{v-2}\right]^{\frac{v+1}{2}}}$	$\frac{-\varepsilon(v+1)}{\varepsilon^2 + (v-2)}$
GED	$v^b$	$f(\varepsilon_i) = \frac{v \exp\left[-0.5 \left \frac{\varepsilon_i}{\lambda}\right ^v\right]}{\lambda 2^{\frac{v-1}{v}} \Gamma\left(\frac{1}{v}\right)}$  $\lambda^2 = 2^{-2/v} \Gamma(v^{-1}) \Gamma(3v^{-1})$	$-0,5 \frac{v}{\varepsilon} \left \frac{\varepsilon}{\lambda}\right ^{v-1}$

<sup>1</sup> A distribuição depende exclusivamente da média e variância condicionais.

<sup>a</sup>  $v > 4$  para garantir cauda pesada.

<sup>b</sup>  $v > 0$ .

A seguir apresentamos na forma de exemplos, as expressões dos escores para dois modelos supondo que as inovações  $\{\varepsilon_i\}$  possuem distribuição normal, em cujo caso

$$\frac{f'(\varepsilon)}{f(\varepsilon)} = -\varepsilon.$$

*Exemplo 3.1: Modelo de Regressão com perturbações AR(q)*

Seja o modelo:

$$Y_t = X_t\beta + Y_t^* ,$$

$$h_t = \omega + \sum_{j=1}^q \alpha_j Y_{t-j}^{*2} ,$$

onde  $X'=(X_1, X_2, \dots, X_r)$  são variáveis exógenas. Na notação de (3.4.1)  $\mu_i = X_i\beta$  e, como as inovações são normais então;  $\Psi = \Theta = (\tilde{\beta}, \omega, \tilde{\alpha})$  cuja ordem é  $(r+1+p)$  e

$$\frac{f'(\varepsilon)}{f(\varepsilon)} = -\varepsilon = \frac{Y_t - \mu_t}{\sqrt{h_t}} .$$

Substituindo esse valor em (3.4.10) obtemos:

$$s_t = \frac{1}{2h_t} \frac{\partial h_t}{\partial \Theta} \left[ \frac{(Y_t - X_t\tilde{\beta})^2}{h_t} - 1 \right] + \frac{(Y_t - X_t\tilde{\beta})}{h_t} \frac{\partial \mu_t}{\partial \Theta}$$

e dado que:

$$\frac{\partial \mu_t}{\partial \Theta} = (X_t, \tilde{0})' ,$$

$$\frac{\partial h_t}{\partial \Theta} = \frac{\partial}{\partial \Theta} \left[ \omega + \sum_{i=1}^q \alpha_i (Y_{t-i} - X_{t-i}\tilde{\beta})^2 \right] = \left( -2 \sum_{i=1}^q \alpha_i X_{t-i}' (Y_{t-i} - X_{t-i}\tilde{\beta}), \delta(\tilde{\beta}) \right)'$$

onde  $\delta(\tilde{\beta}) = (1, (Y_{t-1} - X_{t-1}\tilde{\beta})^2, (Y_{t-2} - X_{t-2}\tilde{\beta})^2, \dots, (Y_{t-q} - X_{t-q}\tilde{\beta})^2)$  é vetor  $q \times 1$  e

$\tilde{0}$  é vetor  $(q+1) \times 1$ . Portanto:

$$s_t = - \left\{ \frac{1}{2h_t} - \frac{(Y_{t-i} - X_{t-i}\tilde{\beta})^2}{2h_t^2} \right\} \left( -2 \sum_{i=1}^q \alpha_i X_{t-i}' (Y_{t-i} - X_{t-i}\tilde{\beta}), \delta(\tilde{\beta}) \right)' + \left( \frac{(Y_{t-i} - X_{t-i}\tilde{\beta})}{h_t}, \tilde{0} \right)$$

*Exemplo 3.2: Modelo AR(1) com inovações GARCH(1,1)*

O modelo é definido como:

$$Y_t = \rho Y_{t-1} + Y_t^* ,$$

$$h_t = \omega + \alpha Y_{t-1}^{*2} + \beta h_{t-1} ,$$

sendo  $\mu_t = \rho Y_{t-1}$ ,  $\Psi = \Theta = (\rho, \omega, \alpha, \beta)$  e  $\frac{f'(\varepsilon)}{f(\varepsilon)} = -\varepsilon = \frac{Y_i - \mu_i}{\sqrt{h_i}}$ . Logo, substituindo:

$$\frac{\partial \mu_t}{\partial \Theta} = (\rho Y_{t-1}, 0, 0, 0)'$$

$$\frac{\partial h_t}{\partial \Theta} = (2\alpha(Y_{t-1} - \rho Y_{t-2})Y_{t-2}, 1, (Y_{t-1} - \rho Y_{t-2})^2, h_{t-1})'$$

em (3.4.10) obtemos a expressão para o escore:

$$s_t = \frac{1}{2h_t^2} ((Y_t - \rho Y_{t-1})^2 - h_t) (2\alpha(Y_{t-1} - \rho Y_{t-2})Y_{t-2}, 1, (Y_{t-1} - \rho Y_{t-2})^2, h_{t-1})' \square$$

$$+ \frac{(Y_t - \rho Y_{t-1})}{h_t} (Y_{t-1}, 0, 0, 0)'$$

*Incerteza dos estimadores ML*

Para alguns modelos econométricos é possível mostrar que a distribuição assintótica dos estimadores de máxima verossimilhança  $\hat{\Psi}$ , satisfaz:

$$(\hat{\Psi} - \Psi_0) \xrightarrow{D} N(0, I_0^{-1}) \tag{3.4.11}$$

onde  $I$  é a matriz de informação<sup>14</sup>, a qual pode ser avaliada de duas formas, através das matrizes:

<sup>14</sup>O subíndice “0” indica que as quantidades estão avaliadas no verdadeiro parâmetro.

$$A_o = \frac{1}{T} \sum_{i=1}^T E \left[ \frac{\partial s_i(Y_i / \Psi_o)}{\partial \Psi} \right] \text{ ou } B_o = \frac{1}{T} \sum_{i=1}^T E \left[ s_i(Y_i / \Psi_o) s_i(Y_i / \Psi_o)' \right].$$

Estimadores consistentes dessas matrizes são:

$$\hat{A} = \frac{1}{T} \sum_{i=1}^T \frac{\partial s_i(Y_i / \hat{\Psi})}{\partial \Psi} \tag{3.4.12}$$

$$\hat{B} = \frac{1}{T} \sum_{i=1}^T \left[ s_i(Y_i / \hat{\Psi}) s_i(Y_i / \hat{\Psi})' \right] \tag{3.4.13}$$

Se o modelo está corretamente especificado então  $A_o=B_o$ , em cujo caso a eleição de  $\hat{A}$  ou  $\hat{B}$  está baseada na facilidade de cálculo dessas matrizes.

Em modelos ARCH, com exceção dos modelos ARCH(q), GARCH(1,1) e IGARCH(1,1), *não foi demonstrada ainda a veracidade da expressão (3.4.11).*

Existem poucos estudos relacionados com as propriedades de normalidade assintótica dos estimadores ML em amostras finitas. Entre eles, Engle *et alii* (1985) mostram que para tamanhos de amostra muito pequena (25,36,49,64,81) os estimadores do modelo AR(1) são viciados.

Por outro lado, experimentos de simulação com amostras de tamanho 500 realizados por Lumsdaine (1995) nos modelos GARCH(1,1) e IGARCH(1,1) mostram que as probabilidades de inclusão (*coverage probabilities*) de 90 e 95% na maioria das vezes subestima os níveis atuais e, no GARCH(1,1) essas probabilidades estão próximas dos níveis teóricos. Por sua vez, a estatística *t* correspondente a  $\alpha$  possui distribuição assintótica razoavelmente próxima da normal com assimetria à esquerda, especialmente para os IGARCH(1,1). As estatísticas *t* correspondentes aos parâmetros  $\omega$  e  $\beta$  possuem distribuições que estão muito afastadas da normal e apresentam forte assimetria à direita e à esquerda respectivamente. A assimetria é severa para o caso dos IGARCH(1,1). Esses resultados indicam que no caso dos modelos IGARCH(1,1) são necessários tamanhos de amostra muito grandes para realizar inferência; em particular Hong (1987) encontrou que são necessárias 5000 observações (aproximadamente 20 anos de retornos diários).

O cálculo das derivadas analíticas nos modelos ARCH é difícil para alguns dos modelos, sendo necessário recorrer ao cálculo numérico dessas expressões. Nesse sentido, do ponto de vista computacional é mais fácil trabalhar com (3.4.13). Experimentos de simulação realizados por Bollerslev e Wooldridge (1992) com amostras pequenas mostram que sob certas condições de regularidade as estimativas encontradas através do algoritmo *BHHH* são satisfatórias.

Por outro lado, se a matriz  $I$  for diagonal, o cálculo dessa fica simplificado e pode ser realizado em duas etapas. Isso é possível para alguns modelos ARCH, como por exemplo para o modelo de Regressão com perturbações ARCH( $q$ ) (Engle, 1982), mas não para outros como os EGARCH (Nelson, 1988, 1991) ou ARCH-M (Engle *et alii*, 1987). Assim, seja  $\theta = (\theta_1, \theta_2)$  onde  $\theta_1$  parametriza a média condicional  $\mu_t$  e  $\theta_2$  parametriza a variância condicional;  $h_t$ , considerando que  $\frac{\partial \mu_t}{\partial \theta_2} = 0$ , então a matriz  $I$  é diagonal sempre que as inovações sejam normais e sob certas condições de regularidade de tal forma que  $\frac{\partial h_t}{\partial \theta_1} = 0$ . A estimação em dois estágios consiste nos seguintes estágios:

- i)* Obter um estimador consistente dos parâmetros da média condicional ( $\hat{\theta}_1$ ) através do método de Quadrados Mínimos (OLS).
- ii)* Baseados nos resíduos de *i)* estimar  $\theta_2$  pelo método de máxima verossimilhança. Esses estimadores são assintoticamente eficientes.

O procedimento apresenta duas desvantagens; a primeira é que a perda de eficiência pode ser muito grande (Engle, 1982) e a segunda que os erros padrão das estimativas precisam ser corrigidas, por exemplo, pelo método de White (1982).

### 3.4.2 Método de Quase-Máxima Verossimilhança

Como mencionado na subseção 3.3.2, nas aplicações empíricas com modelos ARCH é freqüentemente observada a presença de cauda pesada na distribuição dos resíduos. Nesse sentido é desejável verificar se as propriedades de consistência e distribuição assintótica dos estimadores de máxima verossimilhança (ML) são robustos à normalidade nas inovações  $\{\varepsilon_t\}$ .

White (1982) demonstrou que em alguns modelos econométricos é possível definir uma família de distribuições para  $\{\varepsilon_t\}$  próximas da normal<sup>15</sup> de forma que os estimadores obtidos pela maximização da verossimilhança construída sob hipótese de normalidade ( $\hat{\Psi}_{QML}$ ) são consistentes e satisfazem (assintoticamente):

$$\left(\hat{\Psi}_{QML} - \Psi_0\right) \xrightarrow{D} N\left(0, \frac{1}{T} A_0^{-1} B_0 A_0^{-1}\right). \quad (3.4.14)$$

A abordagem descrita é conhecida como de *Quase-máxima Verossimilhança (QML)*. Se a distribuição de  $\{\varepsilon_t\}$  for normal, o modelo está corretamente especificado, então  $A_0=B_0$  e, em consequência obtemos a expressão (3.4.11).

Nos modelos ARCH, utilizando a metodologia proposta por White (1982) demonstra-se a validade da expressão (3.4.14) somente para os modelos ARCH(q), GARCH(1,1) e IGARCH(1,1). Para os modelos ARCH(q), Weiss (1986a) mostrou que os estimadores QML são consistentes e satisfazem (3.4.14), mas sob a hipótese muito restritiva de existência de quarto momento em  $\{Y_t\}$ . Bollerslev e Wooldridge (1992) relaxam as condições de Weiss (1986) e mostraram a validade de (3.4.14) para os modelos ARCH(q) supondo que o escore e a segunda derivada possuem distribuição normal e sob certas condições de regularidade. Essas condições são difíceis de verificar no caso dos modelos GARCH. Lumsdaine (1996) mostrou que o resultado (3.4.14) para os modelos ARCH(q), GARCH(1,1) e também para o IGARCH(1,1) é verdadeiro sob condições menos restritivas, exigindo-se apenas que a distribuição de  $\{\varepsilon_t\}$  possua momento de ordem 32 finito e que a densidade seja simétrica unimodal.

Nos três estudos anteriores é suposto que o processo  $\{\varepsilon_t\}$  é *i.i.d* e/ou normal. Lee e Hansen (1994) relaxam essa suposição mostrando que para condições bastante fracas sobre  $\{\varepsilon_t\}$ , é possível mostrar que os estimadores QML são assintoticamente normais.

Bollerslev e Wooldridge (1992) mostram que uma forma mais fácil de estimar A é através de:  $\hat{A} = \frac{1}{T} \sum_{i=1}^T \left[ -\frac{\partial \mu_t}{\partial \Psi} \frac{\partial \mu_t}{\partial \Psi} \cdot \frac{1}{h_t} - \frac{1}{2} \frac{\partial h_t}{\partial \Psi} \frac{\partial h_t}{\partial \Psi} \cdot \frac{1}{h_t^2} \right]$ , sob a suposição de que os dois primeiros momentos da distribuição estejam corretamente especificados. Adicionalmente, experimentos de simulação realizados por Engle e Gonzáles-Rivera(1991) mostram que as

<sup>15</sup>Considerando a distância de Kullbak-Leibler

estimativas ML e QML são próximas para inovações com densidades simétricas. Caso contrário, a perda de eficiência pode ser muito grande e é aconselhável utilizar densidades não paramétricas.

### 3.5 Testes de Hipóteses e Diagnóstico

Seja a seguinte hipótese de interesse:

$$H_0: r(\Psi_0) = 0 \tag{3.5.1}$$

na qual  $r: \Theta \times H \rightarrow R^g$  é uma função diferenciável no interior de  $\Psi$  e  $g < m+k$ . O Teste de Razão de Verossimilhança ( $T_{LR}$ ), o Testes Wald ( $T_W$ ) e o Teste de Multiplicadores de Lagrange ( $T_{LM}$ ) são definidos da seguinte forma:

$$T_{LR} = 2[L(Y; \hat{\Psi}_1) - L(Y; \hat{\Psi}_0)]$$

$$T_W = Tr(\hat{\Psi}_1)' \left[ \frac{\partial r(\hat{\Psi}_1)}{\partial \Psi} \hat{C}_1^{-1} \frac{\partial r(\hat{\Psi}_1)}{\partial \Psi} \right]^{-1} r(\hat{\Psi}_1)$$

$$T_{LM} = T \left[ \left\{ \frac{\partial r(\hat{\Psi}_0)}{\partial \Psi} \right\}' \hat{C}_0^{-1} \left\{ \frac{\partial r(\hat{\Psi}_0)}{\partial \Psi} \right\} \right]$$

onde os subíndices 0 e 1 correspondem as estimativas sob a hipótese nula e hipótese alternativa, respectivamente. A matriz C é uma estimativa consistente da matriz de covariância e quando é estimada dentro da abordagem de Quase-máxima verossimilhança, utilizando os resultados de White (1982), então os testes são ditos robustos. Normalmente, para vários modelos econométricos, se  $H_0$  é verdadeira e sob certas condições de regularidade, os três testes possuem distribuição assintótica  $\chi^2$  com  $(m+k)-g$  graus de liberdade.

Lumsdaine (1995) examina a validade dos testes apresentados através de experimentos de simulação (com séries de tamanho 500) com relação ao teste  $H_0: IGARCH(1,1)$  e  $H_1: GARCH(1,1)$ . (nesse caso  $r(\Psi) = \alpha + \beta - 1$ ). Os resultados encontrados por ele mostram que o teste Wald apresenta tamanho (percentagem de rejeição) próxima da normal, enquanto que para os outros testes é maior do que a normal. Adicionalmente, ao comparar os tamanhos do teste Wald e do teste Wald robusto, este último é preferível.

Os testes de Multiplicadores de Lagrange podem ser utilizados como testes de especificação. Os estudos sobre o desempenho desses testes (frequentemente utilizados na literatura) em amostras finitas são poucos (ver Bollerslev *et alii*, 1994). Por exemplo, quando testamos o modelo GARCH(1,1), Bollerslev e Wooldridge (1992) encontram que o Teste LM apresenta um comportamento satisfatório.

Outros Testes de Especificação avaliáveis são os baseados nos momentos, entre eles os propostos por Newey (1985) e Tauchen(1985), ver Bollerslev *et alii* (1994) e os testes BDS realizados nos resíduos, ver De Lima (1996) e Brock *et alii* (1996).

Por outro lado, uma forma simples de realizar o diagnóstico nos modelos ARCH através dos resíduos:  $\hat{\varepsilon}_t = \frac{Y_t - \hat{\mu}_t}{\hat{\sigma}_t}$ , é examinando se satisfazem os seguintes requisitos:

- i) possuam distribuição normal com média zero e variância (sempre que as inovações sejam supostas normais).
- ii) sejam não correlacionados.
- iii) os quadrados dos resíduos sejam não correlacionados.

Adicionalmente, é possível avaliar a adequabilidade do modelo através da previsão um-passo a frente *ex-ante*. Considerando que  $Y_t/I_{t-1} \sim N(\mu_t, \sigma_t^2)$ , se o modelo está corretamente especificado os intervalos de confiança  $100\alpha\%$  condicionados à informação anterior, devem incluir aproximadamente  $100\alpha\%$  dos retornos.

Finalmente, quando várias especificações ARCH são comparadas, dois critérios de escolha utilizados na literatura são o AIC (Akaike, 1973) e BIC (Schwarz, 1978), ver por exemplo Nelson (1991) e Bollerslev *et alii* (1994). Entretanto, as propriedades desses critérios ainda são matéria de estudo.

### 3.6 Previsão

Prever o comportamento futuro das séries financeiras é uma das principais preocupações dos investidores. A idéia de que os retornos são *choques* tem sido abandonada a

partir da descoberta dos *Fatos Estilizados* e com a utilização de metodologias não-lineares propostos basicamente na última década. Em consequência, o comportamento das séries financeiras pode ser modelado e, como consequência, fazer previsões.

Nesta Seção serão discutidos alguns aspectos da previsão da volatilidade dos retornos nos modelos ARCH apresentando fórmulas de previsão para os casos GARCH(1,1), IGARCH (1,1) e EGARCH (1,0). No Capítulo será assumido que a previsão é realizada usando a Esperança Condicional à informação disponível. Nos modelos ARCH além das propriedades de otimalidade existe a vantagem de simplicidade.

Tal como mostrado na Seção 3.2, nos modelos ARCH, os retornos são processos martingale diferença; portanto, *o valor predito dos retornos é zero para qualquer horizonte de previsão*, dessa forma o interesse está centrado no segundo momento, a volatilidade dos retornos. Como consequência da definição da volatilidade como função da informação disponível, *a previsão um-passo à frente da volatilidade é ótima*.

O caso que apresenta maior dificuldade é a previsão s-passos à frente. Embora a determinação da estimativa pontual constitua um exercício algébrico fácil ainda que tedioso, o cálculo dos erros médios quadráticos da previsão é uma tarefa complexa e pouco tratada na literatura. Nesta Seção será apresentada a metodologia para a previsão da volatilidade nos modelos GARCH, IGARCH e EGARCH. Como ilustrações obtemos as fórmulas de previsão para 3 casos particulares: GARCH(1,1), IGARCH e EGARCH (1,0). Adicionalmente é obtida a incerteza das estimativas da previsão do GARCH(1,1). As referências básicas para esta Seção são Baillie e Bollerslev (1992) e Nelson (1991).

### 3.6.1 Previsão nos Modelos GARCH e IGARCH

Seja o modelo GARCH (p,q) definido em (3.3.2.3) como:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (3.6.1.2)$$

É necessário enfatizar que conhecidos os parâmetros,  $\sigma_t^2$  é função de  $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \dots, Y_1\}$  para  $t$  suficientemente grande.

*Previsão Pontual da Volatilidade*

A previsão s-passos à frente no modelo GARCH (p,q) definido em (3.6.1.2) dada a informação  $I_T$  e com parâmetros conhecidos é:

$$\hat{\sigma}_{T+s}^2 = E[\sigma_{T+s}^2 / I_T] = \omega + \sum_{i=1}^q \alpha_i E[Y_{T+s-i}^2 / I_T] + \sum_{i=1}^p \beta_i E[\sigma_{T+s-i}^2 / I_T], \quad (3.6.1.3)$$

assim, essa expressão permite calcular de maneira recursiva a previsão da volatilidade, fazendo uso das seguintes igualdades, facilmente verificáveis:

$$E[Y_{T+i}^2 / I_T] = E[\sigma_{T+i}^2 / I_T] = \hat{\sigma}_{T+i}^2, \quad i > 0 \quad (3.6.1.4a)$$

$$E[\sigma_{T-i}^2 / I_T] = \sigma_{T-i}^2 \quad i > 0 \quad (3.6.1.4b)$$

$$E[Y_{T-i}^2 / I_T] = Y_{T-i}^2 \quad i > 0. \quad (3.6.1.4c)$$

O resultado em (3.6.1.4b) é uma consequência do fato de definir os modelos GARCH como processos baseados na previsão um passo à frente.

Na prática, os parâmetros são desconhecidos, em consequência, uma forma de obter as previsões é substituir na fórmula (3.6.1.3), as estimativas dos parâmetros. A seguir, como exemplo, apresentamos os preditores da volatilidade para os modelos GARCH(1,1) e IGARCH(1,1).

MODELO GARCH(1,1)

No modelo GARCH(1,1), a volatilidade está definida como:  $\sigma_{T+1}^2 = \omega + \alpha Y_T^2 + \beta \sigma_T^2$ , portanto utilizando (3.6.1.3) e (3.6.1.4) obtemos que a previsão um passo à frente é igual a:

$$\hat{\sigma}_{T+1}^2 = \omega + \alpha Y_T^2 + \beta \sigma_T^2,$$

isto é :

$$\hat{\sigma}_{T+1}^2 = \sigma_{T+1}^2. \quad (3.6.1.5)$$

Para o caso da previsão  $s$ -passos a frente ( $s \geq 2$ ), substituindo (3.6.1.4a) em (3.6.1.3) e operando repetidamente obtemos:

$$\hat{\sigma}_{T+s}^2 = \omega + (\alpha + \beta)\hat{\sigma}_{T+s-1}^2, \quad (3.6.1.6)$$

$$\hat{\sigma}_{T+s}^2 = \omega \sum_{i=0}^{s-2} (\alpha + \beta)^i + (\alpha + \beta)^{s-1} \sigma_{T+1}^2, \quad (3.6.1.7)$$

e, se o modelo é covariante estacionário, isto é, se  $\alpha + \beta < 1$  com  $\sigma^2 = \text{Var}[Y_t^2] = \frac{\omega}{1 - \alpha - \beta}$ , então, o preditor linear da volatilidade apresenta a seguinte forma:

$$\hat{\sigma}_{T+s}^2 = \omega \frac{1 - (\alpha + \beta)^{s-1}}{1 - (\alpha + \beta)} + (\alpha + \beta)^{s-1} \sigma_{T+1}^2,$$

$$\hat{\sigma}_{T+s}^2 = \sigma^2 \{1 - (\alpha + \beta)^{s-1}\} + (\alpha + \beta)^{s-1} \hat{\sigma}_{T+1}^2,$$

$$\hat{\sigma}_{T+s}^2 = \sigma^2 + (\alpha + \beta)^{s-1} \{\hat{\sigma}_{T+1}^2 - \sigma^2\}. \quad (3.6.1.8)$$

Ao analisar a expressão (3.6.1.7) observamos que no modelo GARCH(1,1), à medida que aumenta o horizonte da previsão, a informação disponível  $I_T$  (embutida em  $\hat{\sigma}_{T+1}^2$ ) torna-se cada vez menos importante no cálculo da previsão. Por sua vez, no caso estacionário, da fórmula (3.6.1.8) é fácil mostrar que a previsão da volatilidade converge a  $\sigma^2$ , isto é, à variância não condicional dos retornos. Isso acontece ainda no caso degenerado, quando  $\omega=0$ .

#### MODELO IGARCH(1,1)

O modelo IGARCH(1,1) é obtido impondo a restrição:  $\alpha + \beta = 1$  no modelo GARCH(1,1). Utilizando essa restrição em (3.6.1.5) e (3.6.1.6) obtemos as expressões para as previsões um passo à frente e  $s$  ( $s > 1$ ) passos à frente da volatilidade. Assim:

$$\hat{\sigma}_{T+1}^2 = \sigma_{T+1}^2 = \omega + \alpha Y_T^2 + (1 - \alpha)\sigma_T^2, \quad (3.6.1.9)$$

$$\hat{\sigma}_{T+s}^2 = \omega(s - 1) + \sigma_{T+1}^2. \quad (3.6.1.10)$$

<sup>16</sup>A expressão de  $\sigma_{T+s}^2$  em Baillie e Bollerslev (1992,p.98) apresenta um erro na somatoria.

Em (3.6.1.10), na medida que aumenta o horizonte da previsão, o preditor da volatilidade *crece linearmente*. Por sua vez, os choques na variância condicional são *Persistentes*, no sentido de Bollerslev e Engle (1986), já que:

$$\begin{aligned} P &= E[\sigma_{T+s}^2 / I_T] - E[\sigma_{T+s}^2 / I_{T-1}] \\ &= \{\omega(s-1) + \sigma_{T+1}^2\} - \{\omega(s) + \sigma_T^2\} = \sigma_{T+1}^2 - \omega - \sigma_T^2 \\ &= \alpha\{Y_T^2 - \sigma_T^2\}, \end{aligned}$$

é uma função não trivial da informação disponível  $I_T$ , para todos os horizontes  $s > 0$ . Para o caso degenerado, substituímos  $\omega=0$  em (3.6.1.10), dessa forma obtemos:  $\hat{\sigma}_{T+1}^2 = \sigma_{T+1}^2$ , isto é, *a volatilidade futura é constante*.

Nas aplicações, é freqüente encontrarmos modelos que exibem alta persistência. Como será mostrado na Seção 3.7, às vezes as estimativas dos modelos GARCH(1,1) e IGARCH(1,1) são muito próximas. Entretanto, os dois casos anteriores mostram que as conseqüências da escolha de um modelo particular (GARCH ou IGARCH) na previsão, são importantes. Desta forma, o comportamento das previsões pode constituir um importante critério para a escolha, ou especificação de um modelo. Nesse sentido, a discussão a seguir, visa explorar brevemente a natureza das diferenças entre o comportamento das previsões do modelo GARCH(1,1) estacionário e o modelo IGARCH.

O comportamento das previsões depende fortemente do tipo de modelo adotado, no sentido de considerar a variância dos retornos como finita ou infinita. Nos dois modelos, a previsão converge ao valor da variância não condicional dos retornos na medida que aumenta o horizonte de previsão. Em outras palavras, a volatilidade predita é o valor histórico da variância dos retornos.

Por outro lado, quando o horizonte de previsão é pequeno, podemos oferecer um argumento heurístico para a escolha de um modelo. Já que a volatilidade é em certa forma uma medida de *risco* ou *incerteza*, as pessoas que acreditam na profunda incerteza do mercado, escolhem o IGARCH, onde a volatilidade em geral aumenta. Pelo contrário, a pessoa que acredita que o risco não pode ser infinito, prefere escolher o GARCH estacionário, no qual as volatilidades decrescem até seu 'valor histórico'.

No entanto, quando trabalhamos com modelos de alta persistência e o interesse é a previsão um-passo à frente, as previsões obtidas pelos dois métodos são quase idênticas.

*Incerteza na Previsão da Volatilidade. Exemplo: GARCH (1,1)*

Como a volatilidade nos modelos ARCH é definida como a previsão um-passo a frente então o erro da previsão condicional é zero, quando os parâmetros são conhecidos. Em outras palavras, *o preditor linear da volatilidade um-passo à frente não só é ótimo, mas é medido sem incerteza.*

No entanto, no caso da previsão s-passos a frente, a situação não é simples. Nessa situação, para calcular o Erro Médio Quadrático, precisamos de expressões que envolvem os momentos da volatilidade, os quais são complexos. A seguir, como ilustração, apresentamos os resultados para o modelo GARCH (1,1). Esses resultados e os correspondentes para o caso geral GARCH (p,q) foram tomados de Baillie e Bollerslev (1992).

Seja o modelo GARCH (1,1) parametrizado, então é possível mostrar que o Erro Médio Quadrático Condicional é igual a:

$$EMQ \equiv E[V_{T,s}^2 / I_T] = (k_2 - 1) \sum_{i=1}^{s-1} \chi_{s-i}^2 E[\sigma_{T+i}^4 / I_T] \quad i = 1, 2, \dots, s-1 \quad (3.6.1.11)$$

onde :

$$V_{T,s} \equiv \sigma_{T+s}^2 - \hat{\sigma}_{T+s}^2 = \sum_{i=1}^{s-1} \chi_{s-i} v_{T+i},$$

$$v_i = Y_i^2 - \sigma_i^2$$

$$\chi_i = \alpha(\alpha + \beta)^{i-1}$$

$k_2 = 3$  quando o processo de inovações  $\{\epsilon_t\}$  é normal e

$$= (n-2)^2 \frac{\Gamma(3/2)\Gamma(0.5n-2)}{\Gamma(1/2)\Gamma(0.5n)} \text{ quando } \epsilon_t \text{ segue uma distribuição } t\text{-Student.}$$

Para avaliar (3.6.1.11) precisamos de  $E[\sigma_{T+i}^4 / I_T]$ :

$$E[\sigma_{T+s}^4 / I_T] = \omega^2 + 2(\alpha + \beta)\omega\hat{\sigma}_{T+s-1}^2 + (k_2\alpha^2 + \alpha\beta + \beta^2)E[\sigma_{T+s-1}^4 / I_T] \quad (3.6.1.12)$$

Quando  $s=2$  e as inovações são normais obtemos:

$$E[\sigma_{T+s}^4 / I_T] = \omega^2 + 2(\alpha + \beta)\omega\hat{\sigma}_{T+1}^2 + (k_2\alpha^2 + \alpha\beta + \beta^2)\hat{\sigma}_{T+1}^4. \quad (3.6.1.13)$$

Baillie e Bollerslev (1992) apresentam resultados que permitem calcular todos os momentos condicionais da volatilidade, assim como os momentos do processo de observações GARCH  $\{Y_t\}$ . Pela definição do modelo GARCH, quando as inovações  $\{\epsilon_t\}$  são normais, a previsão um-passo a frente dos retornos, isto é:  $Y_{T+1}$ , dada a informação  $I_T$ , tem distribuição normal com media 0 e variância  $\hat{\sigma}_{T+1}^2 = \sigma_{T+1}^2$ . Mas, como consequência de (3.6.1.13), a *distribuição da previsão s-passos a frente ( $s>1$ ) dos retornos não é normal !*. Entretanto, Baillie e Bollerslev (1992, p.105) observam que para valores pequenos de  $\alpha$  (o que usualmente acontece na prática) e horizontes de previsão pequenos, a aproximação da distribuição normal funciona bem; portanto, podemos estimar a distribuição de  $Y_{T+s/T}$  por  $N(0, \hat{\sigma}_{T+s}^2)$ .

### 3.6.2 Previsão nos Modelos EGARCH

Na literatura não foram encontrados estudos para medir a incerteza da previsão da volatilidade para os modelos EGARCH. A seguir apresentamos resultados para a previsão pontual do modelo EGARCH(1,0).

#### *Previsão Pontual da Volatilidade*

Seja o modelo EGARCH definido na seção 3.3.4, supondo que é estacionário e começa no tempo  $t = -\infty$ . Dessa forma, a previsão um-passo a frente da volatilidade está dada por:

$$\hat{\sigma}_{T+1}^2 \equiv E[\sigma_{T+1}^2 / I_T] = \sigma_{T+1}^2 = \exp\{\alpha_{T+1} + \sum_{i=1}^{\infty} \beta_i g(\epsilon_{T+1-i})\}, \quad (3.6.2.2)$$

e a fórmula, no caso de  $s$ -passos a frente, é:

$$\begin{aligned} \hat{\sigma}_{T+s}^2 &\equiv E[\sigma_{T+s}^2 / I_T] = E[\exp\{\alpha_{T+s} + \sum_{i=1}^{\infty} \beta_i g(\epsilon_{T+s-i})\} / I_T] \\ &= \exp\{\alpha_T + \sum_{i=s}^{\infty} \beta_i g(\epsilon_{T+s-i})\} E[\exp\{\sum_{i=1}^{s-1} \beta_i g(\epsilon_{T+s-i})\}] \end{aligned}$$

$$= \exp\{\alpha_T + \sum_{i=s}^{\infty} \beta_i g(\varepsilon_{T+s-i})\} \prod_{i=1}^{s-1} E[\exp\{\beta_i g(\varepsilon_{T+s-i})\}], \quad (3.6.2.3)$$

pois  $g(\varepsilon_{T+s-i})$  é função de  $I_T$  quando  $i \geq s$ , a distribuição de  $g(\varepsilon_{T+s-i})$  independe de  $I_T$  quando  $i < s$ , e as variáveis  $\varepsilon_t$  são mutuamente independentes. No entanto, para calcular efetivamente  $E[\exp\{\beta_i g(\varepsilon_{T+s-i})\}] = E[\exp\{\beta_i g(\varepsilon)\}]$  em (3.6.2.3), é necessário especificar a densidade  $f(\varepsilon)$ . Nelson (1991) deriva as expressões dos termos de (3.6.2.3) para densidades família GED. Em particular, no caso da normal temos:

$$E[\exp\{\beta_i g(\varepsilon)\}] = \exp[-\beta_i \phi (2/\pi)^{0.5}] \{ \Phi[\beta_i (\theta + \phi)] \exp[\frac{\beta_i^2}{2} (\theta + \phi)^2] + \Phi[\beta_i (\phi - \theta)] \exp[\frac{\beta_i^2}{2} (\phi - \theta)^2] \} \quad (3.6.2.4)$$

onde  $\Phi(\cdot)$  é a função de distribuição inversa da normal padrão. A seguir, ilustramos a previsão no modelo EGARCH (1,0).

#### Modelo EGARCH (1,0)

Este modelo é obtido fazendo  $\beta_i = \Delta^{i-1}$  em (3.6.2.1c), isto é:  $\ln(\sigma_t^2) = \alpha_t + \frac{1}{1-\Delta B} g(\varepsilon_{t-1})$ . Portanto, de (3.6.2.3) e (3.6.2.4) obtemos a fórmula para a previsão da volatilidade s-passos à frente :

$$\hat{\sigma}_{T+s}^2 = \exp\{\alpha_T + \sum_{i=s}^{\infty} \Delta^{i-1} g(\varepsilon_{T+s-i})\} \prod_{i=1}^{s-1} \exp[-\Delta^{i-1} \phi (2/\pi)^{0.5}] \{ \Phi[\Delta^{i-1} (\theta + \phi)] \exp[\frac{\Delta^{2(i-1)}}{2} (\theta + \phi)^2] + \Phi[\Delta^{i-1} (\phi - \theta)] \exp[\frac{\Delta^{2(i-1)}}{2} (\phi - \theta)^2] \}. \quad (3.6.2.5)$$

### 3.7 Aplicações nas *Séries Exemplo*

A metodologia desenvolvida nas seções anteriores será ilustrada através da modelagem das *Séries Exemplo*, considerando o seguinte modelo para os retornos:

$$Y_t = \mu_t + Y_t^* \quad (3.7.1a)$$

$$\mu_t = E[Y_t / I_{t-1}] \quad (3.7.1b)$$

$$Y_t^* = \sigma_t \varepsilon_t \quad (3.7.1c)$$

no qual a volatilidade é modelada através dos seguintes modelos:

$$\text{GARCH}(p,q): \quad \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i Y_{t-i}^{*2} + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (3.7.2)$$

$$\text{IGARCH}(1,1): \quad \sigma_t^2 = \omega + \alpha Y_{t-1}^{*2} + \beta \sigma_{t-1}^2 \quad \text{com} \quad \alpha + \beta = 1, \quad (3.7.3)$$

$$\text{EGARCH}(p,q) \quad \ln(\sigma_t^2) = \kappa + \sum_{i=1}^p \Delta_i \ln(\sigma_{t-i}^2) + \sum_{j=1}^{q+1} \psi_j g(\varepsilon_{t-j}) \quad (3.7.4a)$$

$$g(\varepsilon_t) = \theta \varepsilon_t + |\varepsilon_t| - \sqrt{2/\pi}, \quad (3.7.4b)$$

e supondo que as inovações  $\{\varepsilon_t\}$  possuem distribuição normal padrão.

Para cada série são ajustados vários modelos. No caso dos modelos GARCH consideramos que a soma de  $p+q$  é no máximo igual a 3 e, no caso dos modelos EGARCH a soma de  $p+q$  é no máximo igual a 4. Com o objetivo de comparar as características dos modelos GARCH e EGARCH é escolhido um representante de cada uma dessas classes utilizando os critérios AIC (Akaike, 1973) e BIC (Schwarz, 1978)<sup>17</sup>. Assim, dentro de cada classe é escolhido o modelo que apresente menores índices AIC e/ou BIC. A seguir, é realizada a análise de adequabilidade do modelo e, se o modelo satisfizer essa análise é considerado como representante. Adicionalmente, é realizado a análise da adequabilidade do GARCH(1,1) por dois motivos; primeiro por ser um dos mais utilizados na prática e, em segundo lugar, porque em alguns casos possui características semelhantes ao modelo de variância estocástica AR(1)-SV estacionário, que será estudado no capítulo seguinte.

O modelo GARCH(1,1) e os representantes GARCH e EGARCH obtidos pelos critérios AIC e BIC são comparados, discutindo as características das estimativas dos parâmetros, das volatilidades, a persistência e a influência de valores extremos dos quadrados dos retornos.

<sup>17</sup>Como mencionado no Seção 3.5, as propriedades dos índices AIC e BIC nos modelos ARCH ainda não são bem conhecidas.

Para a estimação dos modelos é utilizado o procedimento AUTOREG do pacote SAS versão 6.0, que utiliza o método de máxima verossimilhança<sup>18</sup>. Em cada série, a média e a variância condicional são estimadas conjuntamente.

### 3.7.1 Ações

Na análise dos *Fatos Estilizados* da série das *Ações* realizada no Capítulo 2, encontramos a presença de *Efeito Calendário* nas médias dos retornos e, por outro lado, a provável presença de correlação serial de primeiro ordem. Ambas características foram encontradas separadamente e com o objetivo de incorporá-las conjuntamente é proposto o seguinte modelo na média condicional:

$$\mu_t = E[Y_t / I_{t-1}] = \text{Seg}D_1 + \text{Ter}D_2 + \text{Quar}D_3 + \text{Quin}D_4 + \text{Sex}D_5 + \rho Y_{t-1}, \quad (3.7.1.1)$$

no qual  $D_i$  são variáveis *dummies* que assumem valor 1 no dia correspondente ao coeficiente em (3.7.1.1) e zero nos outros dias. A volatilidade está definida como (3.7.2)-(3.7.4). É necessário tomar cuidado ao interpretar os *Efeitos Calendário* na presença de autocorrelação significativa devido a que, nesse caso, existe um *efeito distributivo*<sup>19</sup>.

### Seleção do GARCH

Na Tabela 3.2 são apresentados os valores dos critérios AIC e BIC para algumas especificações de  $\sigma_t^2$ , inclusive quando essa é constante<sup>20</sup>. Segundo a análise dos resultados, os modelos com variância condicional heteroscedástica apresentam valores AIC e BIC notoriamente menores quando comparados com os apresentados pelo modelo com  $\sigma_t^2$  constante. Os modelos ARCH(1) e ARCH(2) apresentam valores de AIC e BIC muito maiores quando comparados com os GARCH estudados.

<sup>18</sup>No pacote, valores iniciais para as volatilidades são iguais à variância não condicional. Adicionalmente, os índices AIC e BIC estão definidos como:  $AIC = -2\ln(L) + 2k$ ,  $BIC = -2\ln(L) + \ln(N)k$ , onde  $L$  é a função de verossimilhança avaliada na estimativa dos parâmetros,  $N$  é o número de observações e  $k$  o número de parâmetros a serem estimados.

<sup>19</sup>(3.7.1.1) é um *autoregressive distributed-lag model*, ver por exemplo Hendry (1995, p.211).

<sup>20</sup>Como comentado na seção 3.3.2 não existe a versão MA da volatilidade GARCH(p,q), isto é, a menos que  $p=0$ ,  $q$  tem que ser maior que zero.

**Tabela 3.2**  
**Identificação na série Ações**

*I. Modelo com variância condicional constante*

AIC: -4427,57  
BIC: -4395,31

*II. Modelos GARCH(p,q)*

p	q	
	1	2
0	-4551,45 <sup>a</sup>	-4576,88
	-4508,44 <sup>b</sup>	-4528,49
1	-4771,34	-4772,47
	-4722,95	-4718,70
2	-4770,92	*
	-4717,15	

<sup>a</sup> Critério AIC

<sup>b</sup> Critério BIC

\* O modelo é estimado como GARCH(1,2)

*III. Modelo IGARCH(1,1)*

AIC: -4773,27  
BIC: -4730,26

*IV. Modelos EGARCH(p,q)*

p	q		
	0	1	2
0		-4543,84 <sup>a</sup>	-4653,79
		-4490,08 <sup>b</sup>	-4594,65
1	-4764,79	-4762,8	-4777,92
	-4711,02	-4703,66	-4713,40
2	-4779,11	-4779,23	-4790,45
	-4719,97	-4714,72	-4720,56

<sup>a</sup> Critério AIC

<sup>b</sup> Critério BIC

Tabela 3.3

Estimativas dos modelos GARCH, IGARCH e variância condicional constante na série *Ações*

	Constante	GARCH (1,1)	GARCH (1,2)	GARCH (2,1)	IGARCH (1,1)
$\rho$	<b>0,0865<sup>a</sup></b> 0,0249 <sup>b</sup> 0,0005 <sup>c</sup> *	<b>-0,0282</b> 0,0279 0,3117	<b>0,0327</b> 0,0278 0,2395	<b>0,0165</b> 0,0278 0,5530	<b>0,0283</b> 0,0283 0,3166
Seg	<b>-0,0077</b> 0,0034 0,0235 *	<b>-0,0059</b> 0,0025 0,0189 *	<b>-0,0062</b> 0,0026 0,0183 *	<b>-0,0048</b> 0,0026 0,0609 *	<b>-0,0059</b> 0,0025 0,018*
Ter	<b>0,0049</b> 0,0034 0,1494	<b>0,0066</b> 0,0029 0,0217 *	<b>0,0070</b> 0,0026 0,0074 *	<b>0,0055</b> 0,0026 0,0322 *	<b>0,0066</b> 0,0025 0,0090 *
Quar	<b>-0,0033</b> 0,0034 0,3304	<b>0,0050</b> 0,0026 0,0542 *	<b>0,0048</b> 0,0026 0,0664	<b>0,0052</b> 0,0025 0,0424 *	<b>0,0050</b> 0,0025 0,0471*
Quin	<b>0,0051</b> 0,0034 0,1298	<b>0,0022</b> 0,0026 0,3901	<b>0,0019</b> 0,0028 0,4981	<b>0,0028</b> 0,0026 0,2800	<b>0,0022</b> 0,0026 0,3890
Sex	<b>0,0013</b> 0,0034 0,7007	<b>0,0034</b> 0,0027 0,2091	<b>0,0040</b> 0,0026 0,1253	<b>0,0024</b> 0,0027 0,3596	<b>0,0034</b> 0,0026 0,1935
$\omega$		<b>0,0001</b> 0,00003 0,0004 *	<b>0,0001</b> 0,00003 0,0004 *	<b>0,0002</b> 0,00004 0,0002 *	<b>0,0001</b> 0,00002 0,0001 *
$\alpha_1$		<b>0,1872</b> 0,0230 0,0001 *	<b>0,1324</b> 0,0332 0,0001 *	<b>0,2574</b> 0,0285 0,0001 *	<b>0,1900</b> 0,0190 0,0001*
$\alpha_2$			<b>0,0821</b> 0,0452 0,0691		
$\beta_1$		<b>0,8095</b> 0,0202 0,0001 *	<b>0,7836</b> 0,0269 0,0001 *	<b>0,2148</b> 0,0808 0,0078 *	<b>0,8100</b> -
$\beta_2$				<b>0,5152</b> 0,0737 0,0001 *	
AIC	-4427,57	-4771,34	-4772,47	-4770,92	-4773,27
BIC	-4395,31	-4722,95	-4718,7	-4717,15	-4730,26

<sup>a</sup> estimativa<sup>b</sup> desvio padrão<sup>c</sup> valor-P associado

\* significativo 5%

Na análise das estimativas dos modelos (Tabela 3.3) vemos que no modelo com variância condicional constante somente a Segunda Feira e a autocorrelação são significativos ao nível de significância 5%. Dessa forma, o *Efeito Calendário* na Quarta Feira encontrado no Capítulo 2, já não é mais significativo na presença de correlação serial. A Segunda Feira é significativa (5%) em todos os modelos GARCH considerados na Tabela 3.3, com exceção do GARCH(2,1) no qual a estimativa possui valor-P=0,06. As estimativas da Terça Feira são significativas (5%) em todos os GARCH e, as correspondentes a Quarta Feira possuem valor-P inferior a 0,07.

Por outro lado, nos modelos com variância condicional heteroscedástica a autocorrelação não é significativa (valor-P>0,23). Isso demonstra (ver comentários da Seção 2.7.1), que séries catalogadas como AR(1) na base da função de autocorrelação (f.a.c), podem ser, na verdade, processos não correlacionados com variância condicional heteroscedástica. Em conseqüência, a leitura dos limites na f.a.c. dos retornos (Figura 2.7) deve ser realizada com cautela se é suspeitada a presença de  $\sigma_t^2$  heteroscedástica.

No processo de seleção do modelo<sup>21</sup>, quando utilizado o critério AIC, obtemos a seguinte ordenação em ordem decrescente de preferência, para a escolha: GARCH(1,2), GARCH(1,1), GARCH(2,1) e, segundo o critério BIC, GARCH(1,1), GARCH(1,2) e GARCH(2,1). Considerando que o modelo que apresenta o menor AIC, GARCH(1,2) é mais geral que o modelo GARCH(1,1) o qual apresenta o menor BIC é escolhido como candidato. A equação de volatilidade para o modelo GARCH(1,2) está dada por:

$$\sigma_t^2 = 1,12 \times 10^{-4} + 0,1324 Y_{t-1}^{*2} + 0,0821 Y_{t-2}^{*2} + 0,7836 \sigma_{t-1}^2 ,$$

(3,1x10<sup>-5</sup>)      (0,0332)      (0,0452)      (0,0269)

onde as quantidades entre parênteses são os valores dos desvios padrão das estimativas. Com exceção da estimativa de  $\alpha_2$ , que possui valor-P igual a 0,069, as outras estimativas são altamente significativas para qualquer valor razoável do valor-P. A soma  $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\beta}_1$  é igual a 0,9981, em conseqüência, pelos resultados da seção 3.3.2 o processo é covariante estacionário.

<sup>21</sup> Ao ajustar o modelo GARCH(2,2), o parâmetro  $\beta_2$  é estimado como igual a zero.

*Adequabilidade do GARCH(1,2)*

Como exposto na seção 3.5.2, a análise de adequabilidade é realizada utilizando os resíduos padronizados e considerando a previsão um-passo à frente ex-ante.

Ao analisar os valores-P do Teste Box-Ljung para os resíduos e seus quadrados (Tabela 3.4) observamos que os resultados apresentados são satisfatórios. Isso é confirmado ao observar as f.a.c nas Figuras 3.1a e 3.1b. Por sua vez, os resíduos apresentam valores da média; -0,034 e variância; 1,00 satisfatórios, enquanto que os valores do coeficiente de simetria (*s*): -0,58 e excesso de curtose (*ec*): 6,37, indicam que a distribuição dos resíduos é levemente assimétrica e possui cauda muito pesada à direita. Ao utilizar o teste de normalidade *W* (ver Seção 2.8) obtemos que a estimativa possui valor-P igual a 0,0001 indicando forte evidência contra a hipótese de normalidade.

**Tabela 3.4**  
**Análise dos resíduos e dos quadrados dos resíduos ARCH na série Ações**

	GARCH		GARCH		EGARCH	
	(1,1)		(1,2)		(2,2)	
A1	5		4		5	
A2	17		16		17	
B1	0		0		0	
B2	2		2		3	
Box-Ljung						
Lag 6	0,540 <sup>a</sup>	0,941 <sup>b</sup>	0,576	0,908	0,309	0,992
Lag 12	0,274	0,995	0,295	0,991	0,116	1,000
Lag 18	0,224	0,999	0,229	0,999	0,085	1,000
Lag 24	0,339	1,000	0,332	1,000	0,127	1,000
Lag 30	0,248	1,000	0,246	1,000	0,111	1,000
Lag 36	0,486	1,000	0,480	1,000	0,272	1,000
Lag 42	0,338	1,000	0,326	1,000	0,198	1,000

A1(A2) número de correlações fora de 1 desvio padrão nos 10 (50) primeiros lags, valores esperados: 3 (15).

B1(B2) número de correlações fora de 1,96 desvios padrão nos 10 (50) primeiros lags, valores esperados:0,5 (2,5).

<sup>a</sup> valores-P do Teste Box-Ljung dos resíduos

<sup>b</sup> valores-P do Teste Box-Ljung dos quadrados dos resíduos

É interessante analisar o que acontece com as estimativas do coeficiente de simetria, o excesso de curtose e o teste de normalidade  $W$  dos resíduos, ao remover o efeito das observações extremas dos retornos. Assim, retirando os 8 maiores resíduos (0,05%) em valor absoluto obtemos  $s=-0,08$  ;  $ec= 0,36$  e  $W=0,989$  (valor- $P=0,80$ ), conseqüentemente não é rejeitada a hipótese nula de distribuição normal nos resíduos. Adicionalmente, ao observar o Gráfico de Probabilidade Normal para os resíduos GARCH(1,2) na Figura 3.2 vemos que o comportamento é satisfatório com exceção de alguns pontos extremos, o que está de acordo com os comentários anteriores. Dentre esses pontos, alguns são possíveis *valores aberrantes (outliers)*<sup>22</sup>, como por exemplo (ver Figura 3.2), os correspondentes a 20/03/90, 4/02/91 e 13/01/92.

Na Tabela 3.5 são apresentados os valores dos dez maiores resíduos do modelo GARCH(1,2). Supondo que esse é o modelo verdadeiro, a existência de cada um dos resíduos (individualmente) na amostra é pouco provável e, a probabilidade de se obter 10 resíduos maiores (em valor absoluto) a 3,33 (décimo resíduo) é inferior a  $10^{-4}$  indicando que a presença desses (conjuntamente) na amostra é implausível<sup>23</sup>, ou o modelo é incapaz de acompanhar o aumento de volatilidade. Vários dos resíduos da Tabela 3.5 podem ser identificados com os acontecimentos descritos na Seção 1.5 que se relacionam com os altos valores observados nos retornos. Por exemplo, na Figura 3.3a é apresentado o período da *Mudança da Banda Cambial* (Março de 1995) no qual existe uma observação extrema em 10/03/95. Da mesma forma, no período da *Queda da Ministra Zélia Cardoso de Mello* (Janeiro 1992) apresentado na Figura 3.4a existe uma observação extrema em 13/01/92. Em ambas as situações, pelos valores dos retornos apresentados antes e depois da ocorrência dos valores extremos é muito provável que esses retornos sejam *valores aberrantes (outliers)* do tipo AO<sup>24</sup>. Os resíduos correspondentes a essas observações extremas foram identificadas na 3.7.1.2 que apresentam grandes desvios com relação à normal. Um outro gráfico com presença de 2 retornos extremos é o correspondente à implementação do *Plano Collor I* (Figura 3.5a).

<sup>22</sup> Existem vários tipos de valores aberrantes (outliers), entre eles os *aditivos* (AO) e de *inovação* (IO). Um valor aberrante AO (no instante  $t$ ) só afeta o nível da série na  $t$ -ésima observação enquanto que a principal característica de um IO (no instante  $t$ ) é que ele também tem efeito sobre as observações futuras.

<sup>23</sup> Para calcular essa probabilidade foi considerada a aproximação Poisson com média igual à probabilidade de ocorrência do evento  $[|\text{resíduo}| > 3,33]$  vezes 1599 (tamanho da amostra), o qual é inferior a  $10^{-4}$ .

<sup>24</sup> A rigor uma observação é tipificada de valor aberrante (*outlier*) através de um teste.

**Tabela 3.5**  
**Resíduos extremos GARCH(1,2) na série Ações**

Data	Retorno	Volatilidade (x100 )	Resíduo	Frequência Esperada
13/01/92	-0,6694	0,49	-9,47	≈0
20/03/90	-0,4209	0,57	-5,62	2x10 <sup>-5</sup>
4/02/91	0,4099	0,58	5,44	4x10 <sup>-5</sup>
18/01/89	-0,2811	0,42	-4,32	0,01
18/10/93	-0,1313	0,10	-4,01	0,05
10/03/95	0,3449	0,86	3,71	0,17
23/12/92	0,1757	0,22	3,66	0,20
17/01/94	0,1095	0,10	3,63	0,23
16/06/89	-0,3836	0,13	-3,60	0,25
27/08/90	0,2183	0,46	3,33	0,69

Uma outra forma de avaliar a adequabilidade do modelo é através da previsão um-passo a frente *ex-ante*. Considerando que  $Y_t / I_{t-1} \approx N(\mu_t, \sigma_t^2)$ , é possível construir intervalos de confiança para os retornos (i.c) condicionados à informação anterior. Ao utilizar a volatilidade GARCH(1,2), os i.c. 95% incluem 95,62% (70 retornos), mostrando que o desempenho do modelo é satisfatório. Na análise, podemos considerar também os níveis da volatilidade; baixo, médio e alto. Esses são definidos da seguinte forma: no nível baixo a volatilidade é menor que o quantil  $Q_{(1/3)}$ , no nível médio a volatilidade está compreendida entre  $Q_{(1/3)}$  e  $Q_{(2/3)}$  e, no nível alto as volatilidades são maiores que  $Q_{(2/3)}$ . Dessa forma, ao considerar os níveis na volatilidade (Tabela 3.6) comprovamos que com exceção do nível alto, a percentagem de retornos fora dos i.c. está dentro do esperado.

Do exposto, podemos concluir que de maneira geral o modelo GARCH(1,2) satisfaz os requisitos na análise de adequabilidade. Como ressalva, apenas a existência de alguns resíduos, cujos altos valores não podem ser explicados pelo modelo e, na análise de previsão *ex-ante* o número de retornos fora dos i.c. 95% quando a volatilidade é alta, é um pouco menor do que o esperado.

**Tabela 3.6**  
**Previsão um-passo à frente *ex-ante* considerando níveis na volatilidade estimada**  
**GARCH(1,2) na série Ações**

Nível	n	Volatilidade ( $\times 10^2$ )		numero de retornos fora	Porcentagem	Valor-P <sup>a</sup>
		b	c			
baixo	532	0,079	0,196	25	4,69	0,75
médio	533	0,196	0,373	28	5,25	0,79
alto	533	0,374	8,552	17	3,19	0,06

<sup>a</sup> Teste bilateral com aproximação Poisson supondo independência nas observações

<sup>b</sup> menor valor apresentado em cada nível.

<sup>c</sup> maior valor apresentado em cada nível.

### Seleção do EGARCH

Na Tabela 3.7 são apresentados os ajustes para os modelos EGARCH<sup>25</sup>, sendo observado que nos modelos EGARCH(2,2), EGARCH(2,1) e EGARCH(2,0), as estimativas da correlação serial não são significativas (valores-P maiores que 0,55). Com relação aos dias da semana, a Segunda (Quarta) Feira é significativo a 6% (4%) em todos os modelos. Por sua vez, a Terça Feira é significativa a 5% nos modelos EGARCH com  $p=1$ , EGARCH(2,2) e nos outros o valor-P é inferior a 0,09.

Dentro da classe dos modelos EGARCH, ao utilizar o critério BIC obtemos a seguinte ordenação decrescente, em ordem de preferência, para a escolha: EGARCH(2,2), EGARCH(2,0), EGARCH(2,1), EGARCH(1,2), EGARCH(1,0) e EGARCH(1,1). Já ao utilizar o AIC, essa ordenação apresenta uma única modificação com troca de posição entre o EGARCH(2,1) e EGARCH(2,0). Em consequência, ambos critérios conduzem à escolha do EGARCH(2,2) como candidato. Nesse modelo, todas as estimativas na equação da volatilidade são significativas com exceção da correspondente ao parâmetro  $\psi_2$  que possui valor-P igual a 0,96<sup>26</sup>. A expressão da volatilidade estimada é:

$$\ln(\sigma_t^2) = -0,5137 + 0,403 \ln(\sigma_{t-1}^2) + 0,5056 \ln(\sigma_{t-2}^2) + 0,2572g(\varepsilon_{t-1}) + 0,0030g(\varepsilon_{t-2}) + 0,2614g(\varepsilon_{t-3})$$

(0,1304) (0,0959) (0,0950) (0,0453) (0,0579) (0,0571)

$$g(\varepsilon_t) = -0,152 |\varepsilon_t| - \sqrt{2/\pi} \varepsilon_t,$$

(0,0562)

<sup>25</sup> não foi ajustada a versão com  $p=q=0$  dado que, nesse caso, a volatilidade independe das observações.

<sup>26</sup> Um estimador mais eficiente poderia ser conseguido ajustando novamente o modelo EGARCH(2,2) sem o parâmetro  $\psi_2$ , mas isso não é possível de ser realizado com o SAS.

Tabela 3.7  
Ajustes EGARCH na série Ações

	EGARCH (1,0)	EGARCH (1,1)	EGARCH (1,2)	EGARCH (2,0)	EGARCH (2,1)	EGARCH (2,2)
$\rho$	<b>0,0561</b> <i>0,0281</i> 0,0458 *	<b>0,0565</b> <i>0,0235</i> 0,0360 *	<b>0,0268</b> <i>0,0037</i> 0,0001 *	<b>0,0022</b> <i>0,0284</i> 0,9375	<b>0,0075</b> <i>0,0287</i> 0,7950	<b>0,0153</b> <i>0,0259</i> 0,5548
Seg	<b>-0,0076</b> <i>0,0029</i> 0,0076 *	<b>-0,0076</b> <i>0,0014</i> 0,0001 *	<b>-0,0076</b> <i>0,0010</i> 0,0001*	<b>-0,0052</b> <i>0,0027</i> 0,0583	<b>-0,0056</b> <i>0,0027</i> 0,0374 *	<b>-0,0062</b> <i>0,0026</i> 0,0165*
Terça	<b>0,0063</b> <i>0,0030</i> 0,0341*	<b>0,0063</b> <i>0,0027</i> 0,0185 *	<b>0,0064</b> <i>0,0015</i> 0,0001*	<b>0,0046</b> <i>0,0027</i> 0,0876	<b>0,0046</b> <i>0,0027</i> 0,0832	<b>0,0053</b> <i>0,0026</i> 0,0381*
Quarta	<b>0,0054</b> <i>0,0026</i> 0,0402 *	<b>0,0054</b> <i>0,0026</i> 0,0353	<b>0,0058</b> <i>0,0028</i> 0,0379 *	<b>0,0058</b> <i>0,0027</i> 0,0318 *	<b>0,0061</b> <i>0,0027</i> 0,0232 *	<b>0,0058</b> <i>0,0024</i> 0,0155 *
Quinta	<b>0,0020</b> <i>0,0027</i> 0,4435	<b>0,0020</b> <i>0,0026</i> 0,4311	<b>0,0028</b> <i>0,0016</i> 0,0711	<b>0,0032</b> <i>0,0027</i> 0,2327	<b>0,0028</b> <i>0,0025</i> 0,2675	<b>0,0027</b> <i>0,0025</i> 0,2696
Sexta	<b>0,0033</b> <i>0,0026</i> 0,2148	<b>0,0033</b> <i>0,0025</i> 0,1862	<b>0,0025</b> <i>0,0023</i> 0,2852	<b>0,0020</b> <i>0,0025</i> 0,4411	<b>0,0021</b> <i>0,0025</i> 0,4106	<b>0,0027</b> <i>0,0024</i> 0,2574
$\kappa$	<b>-0,2261</b> <i>0,0579</i> 0,0001*	<b>-0,2297</b> <i>0,0668</i> 0,0006 *	<b>-0,3635</b> <i>0,0863</i> 0,0001*	<b>-0,3104</b> <i>0,0839</i> 0,0008 *	<b>-0,3506</b> <i>0,0414</i> 0,0001*	<b>-0,5137</b> <i>0,1304</i> 0,0001*
$\Delta_1$	<b>0,9590</b> <i>0,0101</i> 0,0001*	<b>0,9584</b> <i>0,0115</i> 0,0001*	<b>0,9352</b> <i>0,0148</i> 0,0001*	<b>0,2804</b> <i>0,0406</i> 0,0001*	<b>0,1966</b> <i>0,0598</i> 0,0010 *	<b>0,4031</b> <i>0,0959</i> 0,0001*
$\Delta_2$				<b>0,6635</b> <i>0,0421</i> 0,0001*	<b>0,7401</b> <i>0,0568</i> 0,0001*	<b>0,5056</b> <i>0,0950</i> 0,0001*
$\Psi_1$	<b>0,2846</b> <i>0,0280</i> 0,0001*	<b>0,2790</b> <i>0,0534</i> 0,0001*	<b>0,2788</b> <i>0,0474</i> 0,0001*	<b>0,4102</b> <i>0,0380</i> 0,0001*	<b>0,3840</b> <i>0,0389</i> 0,0001*	<b>0,2572</b> <i>0,0453</i> 0,0001*
$\Psi_2$		<b>0,0076</b> <i>0,0582</i> 0,8961	<b>-0,1528</b> <i>0,0592</i> 0,0098*		<b>0,0823</b> <i>0,0518</i> 0,1119	<b>0,0030</b> <i>0,0579</i> 0,9591
$\Psi_3$			<b>0,2349</b> <i>0,0517</i> 0,0001*			<b>0,2614</b> <i>0,0571</i> 0,0001*
$\theta$	<b>-0,1420</b> <i>0,0682</i> 0,0371*	<b>-0,1446</b> <i>0,0710</i> 0,0417 *	<b>-0,1411</b> <i>0,0568</i> 0,0129 *	<b>-0,0993</b> <i>0,0596</i> 0,0959	<b>-0,1036</b> <i>0,0591</i> 0,0799	<b>-0,1521</b> <i>0,0562</i> 0,0068 *
AIC	-4764,79	-4762,8	-4777,92	-4779,11	-4779,23	-4790,45
BIC	-4711,02	-4703,66	-4713,4	-4719,97	-4714,72	-4720,56

<sup>a</sup> estimativa

<sup>b</sup> desvio padrão

<sup>c</sup> valor-P associado

\* significativo 5%

A estimativa correspondente a  $\theta$ : -0,1521 possui valor-P igual a 0,007, indicando evidência de presença de *Leverage Effect*, isto é, os choques positivos diminuem a volatilidade enquanto que os choques negativos a aumentam. Para encontrar a persistência do modelo calculamos as raízes da equação  $1-0,403103x-0,505624x^2$ , sendo iguais a 0,94064 e -0,53753, portanto, pelo Teorema 1 da seção 3.3.4 o efeito dos choques não persiste indefinidamente na volatilidade, pois o processo da volatilidade não é ergódico nem (estritamente) covariante estacionário.

#### *Adequabilidade do EGARCH(2,2)*

A análise da função de autocorrelação (f.a.c) dos resíduos (Tabela 3.4) revela que há evidência de ausência ou de baixa autocorrelação. Por sua vez, pelos resultados do Teste Box-Ljung nos quadrados dos resíduos, podemos observar que o modelo EGARCH(2,2) é capaz de remover a não-linearidade observada nos retornos. Nas Figuras 3.1c-d são mostradas as f.a.c dos resíduos e dos quadrados dos resíduos.

As estimativas da média ( $m$ ), variância ( $v$ ), coeficiente de simetria ( $s$ ) e excesso de curtose ( $ec$ ) dos resíduos são  $m=-0,024$ ,  $v=1,00$ ,  $s=-0,27$  e  $ec=3,71$  indicando que a média e variância são satisfatórios e, por sua vez, a distribuição dos resíduos apresenta leve assimetria e cauda pesada à direita. Embora os resíduos do modelo EGARCH(2,2) apresentem grande excesso de curtose, o valor da estatística  $W$  para os resíduos é igual a 0,9860 significando um valor-P=0,12, não dando grande evidência para rejeitar a hipótese de normalidade nos resíduos. Esta aparente contradição é explicada na análise do Gráfico de Probabilidade Normal dos resíduos (Figura 3.6), onde é observado que o comportamento é satisfatório para a quase totalidade dos valores, exceto para alguns valores extremos. De fato, retirando-se os 8 maiores resíduos temos que  $s=-0,044$  e  $ec=0,52$ .

Na tabela a seguir apresentamos uma lista dos 10 maiores resíduos para o ajuste do modelo EGARCH(2,2). Sob a suposição que esse é o verdadeiro modelo, a probabilidade de obter 10 resíduos maiores (em valor absoluto) ao valor do décimo maior resíduo, -3,29 é inferior a  $10^{-4}$ , isto é, a presença desses resíduos é implausível. Vários desses resíduos extremos coincidem nas datas obtidas para os resíduos extremos do GARCH(1,2). Ao observar a Figura 3.6 podemos identificar os resíduos extremos correspondentes a 13/01/92, 20/03/90 e 4/02/91, os quais se encontram muito afastados da reta de normalidade.

**Tabela 3.8**  
Resíduos extremos EGARCH (2,2) na série Ações

Data	Retorno	Volatilidade (x10 <sup>2</sup> )	Resíduo	Frequência Esperada
13/01/92	-0,6694	0,76	-7,58	≈0
04/02/91	0,4099	0,53	5,72	10 <sup>-5</sup>
20/03/90	-0,4209	0,58	-5,59	2x10 <sup>-5</sup>
18/01/89	-0,2811	0,51	-3,99	0,05
18/10/93	-0,1313	0,10	-3,92	0,07
23/12/92	0,1757	0,20	3,76	0,13
10/03/95	0,3449	0,90	3,63	0,22
6/10/89	0,1949	0,30	3,46	0,44
3/03/89	0,0953	0,07	3,44	0,46
16/06/89	-0,3836	1,37	-3,29	0,81

Na análise da previsão um-passo a frente *ex-ante* temos que a percentagem dos retornos incluídos dentro dos intervalos de confiança (95%) é igual a 95,68% (valor-P igual a 0,22), sendo esses resultados satisfatórios. Esses resultados mantêm-se ao considerar os níveis na volatilidade, ver Tabela 3.9, exceto para volatilidade alta, onde existe uma leve indicação de subestimação (valor-P igual a 0,09).

**Tabela 3.9**  
Previsão um-passo à frente *ex-ante* considerando níveis na volatilidade estimada EGARCH (2,2) na série Ações

Nível	n	Volatilidade (x10 <sup>2</sup> )		numero de retornos fora	Percentagem	Valor- P
		b	c			
baixo	533	0,072	0,203	27	5,07	0,95
médio	533	0,206	0,363	24	4,50	0,61
alto	533	0,364	13,522	18	3,38	0,09

<sup>a</sup> Teste bilateral com aproximação Poisson supondo independência nas observações

<sup>b</sup> menor valor apresentado em cada nível.

<sup>c</sup> maior valor apresentado em cada nível.

Em resumo, podemos dizer que de maneira geral o modelo EGARCH(2,2) apresenta bons resultados quanto à adequabilidade, mas existe falta de ajuste em poucos retornos que apresentam valores extremos.

*Modelo GARCH(1,1)*

Pelos motivos expostos anteriormente também será considerado o modelo GARCH(1,1). Esse modelo apresenta o menor índice BIC e o segundo AIC. A equação da volatilidade é a seguinte:

$$\sigma_t^2 = 10^{-5} + 0,1872Y_{t-1}^{*2} + 0,8095\sigma_{t-1}^2,$$

(3x10<sup>-5</sup>)    (0,0230)    (0,0202)

sendo que todas as estimativas são altamente (valores-P < 0,0001). A persistência do modelo GARCH(1,1) é altíssima, 0,9967<sup>27</sup>, significando uma meia-vida de 209,7 dias. Sendo que o valor da persistência é próximo de um, analisamos as estimativas da volatilidade IGARCH(1,1) na Tabela 3.3, observando-se que essas são quase idênticas às apresentadas pelo modelo GARCH(1,1).

Com relação à análise de adequabilidade do modelo, observamos que o comportamento apresentado pela f.a.c dos resíduos (Figura 3.1e) e pela f.a.c dos quadrados dos resíduos (Figura 3.1f) é bom. Os valores-P do teste Box-Ljung apresentados na Tabela 3.4 apoiam essa conclusão. Adicionalmente, os resíduos GARCH(1,1) apresentam valores da média; -0,034 e variância; 1,00 satisfatórios, enquanto que os valores do coeficiente de simetria (*s*): -0,61 e excesso de curtose (*ec*): 6,88, indicam que a distribuição dos resíduos possui cauda muito pesada à direita. A estatística *W* possui valor-P igual a 0,0001 rejeitando a hipótese de normalidade. No entanto, ao retirar os 8 maiores resíduos em valor absoluto obtemos: *s*=-0,03 e *ec*= 0,35; ainda, retirando somente o maior resíduo obtemos *s*=-0,06, *ec*=1,87, *W*=0,99, portanto, a hipótese de normalidade não é rejeitada. Embora não apresentado, o Gráfico de Probabilidade Normal para os resíduos apresenta um comportamento apenas insatisfatório nas caudas, observando-se poucos resíduos distantes da reta ideal de normalidade. Os 10 maiores resíduos são apresentados na Tabela 3.10 sendo observado que, se o verdadeiro modelo fosse o GARCH(1,1), a presença desses resíduos tanto individualmente, quanto conjuntamente (a probabilidade de um resíduo ser maior que 3,42 é inferior a 10<sup>-4</sup>), é implausível. Finalmente, enquanto à previsão um-passo à frente ex-ante, os intervalos de confiança 95% (i.c.) incluem 95,5% dos retornos mostrando que o desempenho do modelo é satisfatório e, ao considerar os

<sup>27</sup> Como observa Nelson (1991, p.360) quando as séries não são muito longas é necessário considerar as estimativas das raízes com cautela. Nesse sentido, um valor muito próximo de 1 pode ser uma indicação de que o modelo é não estacionário ou de longa memória (*long memory*).

níveis na volatilidade (Tabela 3.11) esses resultados mantêm-se com exceção no nível alto onde é observada uma proporção dos retornos menor do que a esperada. O modelo GARCH(1,1) satisfaz os requerimentos de adequabilidade no mesmo sentido que os modelos GARCH(1,2) e EGARCH(2,2).

**Tabela 3.10**  
Resíduos extremos GARCH(1,1) na série *Ações*

Data	Retorno	Volatilidade (x100)	Resíduo	Frequência Esperada
13/01/92	-0,6694	0,47	-9,70045	0
4/02/91	0,4099	0,57	5,56	$3,2 \times 10^{-5}$
20/03/90	-0,4209	0,59	-5,50	$6,4 \times 10^{-5}$
18/10/93	-0,1313	0,09	-4,05	0,08
18/01/89	-0,2811	0,49	-4,04	0,08
10/03/95	0,3449	0,87	3,69	0,36
23/12/92	0,1757	0,21	3,66	0,39
16/06/89	-0,3836	1,12	-3,61	0,48
17/01/94	0,1095	0,01	3,56	0,58
27/08/90	0,2183	0,43	3,42	1,00

**Tabela 3.11**  
Previsão um-passo à frente *ex-ante* considerando níveis na volatilidade estimada GARCH (1,1) na série *Ações*

Nível	n	Volatilidade ( $\times 10^2$ )		numero de retornos fora	Percentagem	Valor-P
		b	c			
baixo	532	0,080	0,198	26	4,88	0,90
médio	533	0,198	0,372	29	5,44	0,65
alto	533	0,372	8,668	17	3,19	0,06

<sup>a</sup> Teste bilateral com aproximação Poisson supondo independência nas observações

<sup>b</sup> menor valor apresentado em cada nível.

<sup>c</sup> maior valor apresentado em cada nível.

*Comparação dos Modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,2)*

Nos parágrafos a seguir será realizada uma comparação dos modelos GARCH(1,2) e EGARCH(2,2), os quais foram escolhidos na base dos critérios AIC e BIC. Mas, antes disso, são comparados os modelos GARCH(1,2) e GARCH(1,1). Na Figura 3.7 podemos observar que as volatilidades estimadas pelos modelos GARCH(1,2) e GARCH(1,1) são quase idênticas para a grande maioria dos pontos, mas existem alguns retornos para os quais as diferenças entre as estimativas são grandes. Essas grandes diferenças são identificadas com os pontos bem acima ou bem abaixo da linha reta. Em particular, existe uma observação que apresenta uma diferença extrema que corresponde ao período da implementação do *Plano Collor I*.

Por outro lado, como observado nas Figuras 3.8 e 3.9, as volatilidades estimadas pelos modelos GARCH(1,2) e EGARCH(2,2) são capazes de acompanhar os movimentos de aumentos e diminuições dos quadrados dos retornos<sup>28</sup>.

Na Figura 3.10 apresentamos a comparação das volatilidades estimadas GARCH(1,2) e EGARCH(2,2), sendo observado que para a grande maioria dos retornos, elas são bastante próximas. No entanto, existem vários retornos para os quais as diferenças são consideráveis e, como pode ser conferido na Figura 3.11 algumas dessas se produzem após a ocorrência de grandes valores dos retornos (em valor absoluto). Os altos valores ou *picos* nas volatilidades estimadas podem ser identificados em alguns casos com os acontecimentos descritos na Seção 1.5. Por exemplo, na Figura 3.3a, período correspondente à *Mudança da Banda Cambial*, devido ao *pico* em 10/03/95, as estimativas de volatilidade aumentam no dia seguinte, mas no caso do GARCH(1,2), o aumento é muito maior. Devido ao efeito dessa observação extrema e a alta persistência, a volatilidade estimada GARCH(1,2) apresenta grandes valores (comparados aos valores apresentados antes do 10/03/95) durante bastantes dias, decaindo lentamente. Por sua vez, a volatilidade estimada EGARCH(2,2) apresenta vários aumentos e diminuições bruscas após o 10/03/95 como consequência da forma de dependência (da volatilidade) as observações passadas. Com efeito, sendo  $\sigma_t^2 = \exp(\kappa)[\sigma_{t-1}^2]^{\Delta_1} [\sigma_{t-2}^2]^{\Delta_2} \exp[\Psi_1 g(\varepsilon_{t-1}) + \Psi_2 g(\varepsilon_{t-2}) + \Psi_3 g(\varepsilon_{t-3})]$  a volatilidade do modelo, onde  $\varepsilon_t = Y_t \sigma_t^{-1}$ , o maior efeito de uma observação extrema no instante  $T=t-3$ , se manifesta no tempo  $T=t$ .

<sup>28</sup> Assim como o GARCH(1,1) e outras especificações como o EGARCH(1,0) e GARCH(2,1).

Na Figura 3.4a podemos analisar o efeito do maior retorno da série, ocorrido no dia 13/01/92 correspondente à época da *Queda da Ministra Zélia Cardoso de Mello*. Neste caso, da mesma forma que no período da *Mudança da Banda Cambial*, a volatilidade estimada pelo modelo GARCH(1,2) é muito maior à volatilidade estimada EGARCH(2,2), mas no 16/01/92 acontece o contrário, sendo a volatilidade EGARCH(2,2) muito maior devido ao efeito atrasado explicado no parágrafo anterior. Observe-se que o efeito do retorno extremo permanece aproximadamente no mesmo período para os três modelos.

Diferentemente do que acontece nas duas situações anteriores, nos quais a volatilidade estimada GARCH(1,2) é geralmente maior que a volatilidade estimada EGARCH(2,2), no período seguinte à ocorrência da observação extrema, na época do *Plano Collor I* (Figura 3.5a) acontece o contrário. Nesse período ocorrem dois retornos extremos consecutivos; 20/03/90, 21/03/90 e, devido ao efeito desses, as volatilidades estimadas GARCH(1,2) e EGARCH(2,2) aumentam bastante no 22/03/90 e logo começam a diminuir até o 2/04/90 no qual a ocorrência de outro valor grande nos retornos a faz aumentar.

Uma forma adicional de comparar as volatilidades é através dos intervalos de confiança para os retornos (i.c). Assim, os i.c. 95% incluem 95,62% [95,68%] dos retornos para o modelo GARCH(1,2) [EGARCH(2,2)], isto é, nesta comparação, o desempenho dos modelos é praticamente o mesmo. Adicionalmente, ao observar os resultados das Tabelas 3.6 e 3.9 notamos que esse desempenho mantém-se ao considerar os níveis da volatilidade.

Para os três acontecimentos discutidos nos parágrafos anteriores apresentamos os gráficos dos i.c. 95% nas Figuras 3.3b, 3.4b, 3.5b. Observamos que, em nenhum dos casos, os i.c. 95% incluem os retornos extremos e, após a ocorrência desses, os i.c. são desnecessariamente largos. Isto é uma indicação de que temos um valor aberrante, que na literatura é conhecido como *outlier de observação* ou do tipo AO. Nesse sentido, no período da *Queda da Ministra Zélia Cardoso de Mello* (Figura 3.3b) preferimos o GARCH(1,2) ao EGARCH(2,2) por ser menos influenciado e por apresentar os i.c. menos largos. Por sua vez, no caso do *Plano Collor I* (Figura 3.5b) preferimos o EGARCH(2,2).

### Conclusões

As principais conclusões da análise realizada são:

- A volatilidade estimada pelos modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,2) são muito próximas para a grande maioria dos retornos, apresentando grandes diferenças só na presença de retornos extremos, os quais estão relacionados com os acontecimentos descritos na seção 1.5.
- Os modelos GARCH(1,2) e EGARCH(2,2) conseguem reproduzir os conglomerados de volatilidade observados na série das *Ações*.
- A análise dos resíduos dos modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,2) revela que esses modelos são capazes de remover a não-linearidade dos retornos. Por sua vez, o desempenho dos modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,2) na previsão *ex-ante* é similar e satisfatório, com exceção nos períodos de volatilidade alta. De maneira geral, os ajustes satisfazem os requerimentos da adequabilidade, mas existem vários resíduos cuja presença não pode ser explicada pelos modelos. Nesse sentido, modelos com intervenções poderiam melhorar os ajustes.
- Nos modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,2) a autocorrelação serial na média condicional não é significativa sendo os valores-P associados as estimativas iguais a 0,31; 0,24 e 0,55 respectivamente.
- No EGARCH(2,2) a estimativa correspondente à Segunda Feira é negativo na Terça e Quarta Feira positivos, sendo todos esses significativos a 5%. Da mesma forma, nos modelos GARCH(1,1) e GARCH(1,2) as estimativas correspondentes a Segunda e Terça Feira são significativas a 5% e da Quarta Feira também mas a 7%.
- Existe evidência de *Leveragge Effect* no modelo EGARCH (2,2).
- Os modelos GARCH(1,1) e GARCH(1,2) são covariante estacionários, mas o EGARCH(2,2) não. Nesse sentido, considerando o comportamento apresentado pelas variâncias recursivas (ver Capítulo 2), os modelos GARCH(1,1) e GARCH(1,2) são preferíveis.
- A persistência observada no modelo GARCH(1,1) é altíssima, 0,9967; sendo a meia-vida associada bastante grande.

- Os valores das volatilidades estimadas GARCH(1,2) e EGARCH(2,2) são muito diferentes logo após a ocorrência de valores extremos nos quadrados dos retornos. Valores extremos dos retornos (valor absoluto) produzem aumentos consideravelmente maiores na volatilidade GARCH(1,2) em comparação aos produzidos na volatilidade EGARCH(2,2), no instante seguinte à ocorrência da observação extrema. Também é observado que o efeito das observações extremas (entendido como grandes valores da volatilidade) permanece por vários dias.

- Como ilustrado em 3 situações (Figuras 3.3b, 3.4b e 3.5b) os limites dos intervalos de confiança dos modelos GARCH(1,2) e EGARCH(2,2) são muito próximos menos nos instantes logo após a ocorrência dos retornos extremos.

### 3.7.2 Taxa de Câmbio

Com o objetivo de incorporar o *Efeito Calendário*, observado no Capítulo 2, é ajustado o seguinte modelo:

$$Y_t = \mu_t + Y_t^* \quad (3.7.2.1a)$$

$$\mu_t = E[Y_t / I_{t-1}] = \text{Seg}D_1 + \text{Ter}D_2 + \text{Quar}D_3 + \text{Quin}D_4 + \text{Sex}D_5 \quad (3.7.2.1b)$$

$$Y_t^* = \varepsilon_t \sigma_t \quad (3.7.2.1c)$$

no qual  $D_i$  's são variáveis *dummy* que assumem valor 1 no dia correspondente ao coeficiente em (3.7.1.1b) e zero nos outros dias. A volatilidade ou variância condicional é da forma (3.7.1)-(3.7.3).

#### *Seleção do Modelo GARCH*

Nas Tabelas 3.12 e 3.13 são apresentados os valores AIC e BIC e as estimativas de algumas especificações GARCH incluindo o modelo com variância condicional constante.

**Tabela 3.12**  
**Identificação na série Taxa de Câmbio**

*I. Modelo com variância condicional constante*

AIC: -12248,6  
 BIC: -12221,2

*II. Modelos GARCH(p,q)*

p	q	
	1	2
0	-12257,9 <sup>a</sup>	-12267,4
	-12219,6 <sup>b</sup>	-12223,6
1	-12346,8	-12345,2
	-12303	-12295,9
2	-12348,8 <sup>a</sup>	-12350,8
	-12310,5	-12307,0

<sup>a</sup> Critério AIC  
<sup>b</sup> Critério BIC

*III. Modelo IGARCH(1,1)*

AIC: -12332,9  
 BIC: -12305,5

*IV. Modelos EGARCH(p,q)*

p	q		
	0	1	2
0		-12261,6 <sup>a</sup>	-12265,4
		-12212,4 <sup>b</sup>	-12210,6
1	-12345,7	-12344,8	-12345,8
	-12296,4	-12290,0	-12285,5
2	-12358,2	-12347,6	-12345,6
	-12303,5	-12287,4	-12279,9

<sup>a</sup> Critério AIC  
<sup>b</sup> Critério BIC

**Tabela 3.13**  
**Estimativas dos modelos GARCH, IGARCH(1,1) e variância condicional constante na série**  
**Taxa de Câmbio**

	CONSTAN TE	GARCH (1,1)	GARCH (1,2)	GARCH (2,1)	GARCH (2,2)	IGARCH (1,1)
Seg	<b>0,0005<sup>a</sup></b> 0,0004 <sup>b</sup> 0,2336 <sup>c</sup>	<b>0,0006</b> 0,0004 0,1674	<b>0,0006</b> 0,0004 0,1511	<b>0,0006</b> 0,0004 0,1744	<b>0,0006</b> 0,0004 0,1162	<b>0,0006</b> 0,0004 0,1920
Ter	<b>0,0001</b> 0,0004 0,7813	<b>-0,00002</b> 0,0004 0,9543	<b>0,00003</b> 0,0004 0,9363	<b>-0,00003</b> 0,0004 0,9454	<b>-0,00003</b> 0,0004 0,9356	<b>0,000044</b> 0,0004 0,9209
Quar	<b>-0,0001</b> 0,0004 0,8009	<b>-0,00001</b> 0,0004 0,9745	<b>-0,00002</b> 0,0004 0,9590	<b>-0,00001</b> 0,0004 0,9780	<b>-0,00007</b> 0,0004 0,8507	<b>-4,56x10<sup>-6</sup></b> 0,0004 0,9209
Quin	<b>0,0008</b> 0,0004 0,0481	<b>0,0007</b> 0,0004 0,0768	<b>0,0007</b> 0,0004 0,0719	<b>0,0007</b> 0,0004 0,0912	<b>0,0007</b> 0,0004 0,0597	<b>0,0006</b> 0,0004 0,1414
Sex	<b>-0,0001</b> 0,0004 0,8089	<b>-0,0003</b> 0,0004 0,5155	<b>-0,0003</b> 0,0004 0,4933	<b>-0,0003</b> 0,0004 0,5173	<b>-0,0002</b> 0,0004 0,5195	<b>-0,0003</b> 0,0004 0,5120
$\omega$		<b>2,48x10<sup>-6</sup></b> 7,68x10 <sup>-7</sup> 0,0012 *	<b>2,55x10<sup>-6</sup></b> 6,90x10 <sup>-7</sup> 0,0002 *	<b>2,50x10<sup>-6</sup></b> . .	<b>4,80x10<sup>-6</sup></b> . .	<b>7,06x10<sup>-7</sup></b> . .
$\alpha_1$		<b>0,0718</b> 0,0135 0,0001*	<b>0,0564</b> 0,0266 0,0337*	<b>0,0720</b> 0,0039 0,0001*	<b>0,0479</b> 0,0115 0,0001*	<b>0,0888</b> . .
$\alpha_2$			<b>0,0175</b> 0,0278 0,5303		<b>0,0882</b> 0,0140 0,0001*	
$\beta_1$		<b>0,8861</b> 0,0212 0,0001*	<b>0,8828</b> 0,0201 0,0001*	<b>0,8855</b> 0,0066 0,0001*	<b>0</b> . .	<b>0,9112</b> . .
$\beta_2$				<b>0</b> . .	<b>0,7820</b> 0,0161 0,0001*	
AIC	-12248,6	-12346,8	-12345,2	-12348,8	-12350,8	-12332,9
BIC	-12221,2	-12303,0	-12295,9	-12310,5	-12307	-12305,5

<sup>a</sup> estimativa pontual.

<sup>b</sup> desvio padrão estimado.

<sup>c</sup> valor-P.

Em três dos modelos, GARCH(2,1), GARCH(2,2) e IGARCH(1,1), o pacote SAS não fornece as estimativas dos desvios padrão de  $\alpha$ , da mesma forma, não temos a estimativa de  $\beta_2$  no

GARCH(2,1), de  $\beta_1$  no GARCH(2,2) e de  $\alpha_1$  no IGARCH(1,1)<sup>29</sup>. No processo de escolha, o modelo GARCH(2,2) apresenta o menor valor AIC e o segundo menor BIC, enquanto que o GARCH(2,1) apresenta o menor BIC e o segundo menor AIC. Considerando que o GARCH(2,2) é mais geral que o GARCH(2,1), o primeiro modelo é escolhido inicialmente como candidato.

No GARCH(2,2), a estimativa correspondente à Quinta Feira é significativa (6%) e o valor-P da estimativa correspondente à Segunda Feira é 0,12. Nos outros dias as estimativas não são significativas. A equação da volatilidade estimada é:

$$\sigma_t^2 = 4,80 \times 10^{-6} + 0,0479 Y_{t-1}^{*2} + 0,0882 Y_{t-2}^{*2} + 0,7820 \sigma_{t-2}^2 \quad (3.7.2.2)$$

(0,0115)            (0,0140)            (0,0160)

Como  $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\beta}_2$  é menor que um então o processo é covariante estacionário. Na Figura 3.12 apresentamos o gráfico da volatilidade estimada pelo modelo GARCH (2,2), sendo observado que ela é capaz de acompanhar o comportamento de aumentos e diminuições dos valores dos quadrados dos retornos.

#### *Adequabilidade do GARCH(2,2)*

Através da análise gráfica das f.a.c dos resíduos e seus quadrados apresentadas nas Figuras 3.13a e 3.13b respectivamente, podemos considerar que os resíduos do modelo GARCH(2,2) não apresentam correlação serial e que a não-linearidade foi removida. Esses resultados são confirmados pelos valores-P do Teste Box-Ljung apresentados na Tabela 3.14. Por outro lado, os resíduos do modelo GARCH(2,2) apresentam valores de média (0,003), variância (1,00) e coeficiente de simetria (0,16), satisfatórios. O valor do excesso de curtose é igual a 1,27 indicando que a distribuição dos resíduos apresenta cauda mais pesada que a normal. Mas, ao utilizar a estatística  $W$  para testar normalidade, encontramos que a hipótese de normalidade não é rejeitada, com valor-P igual a 0,49. Na Figura 3.14 apresentamos o Gráfico

<sup>29</sup> Na análise dos resultados apresentados pelo programa vimos que em alguns destes casos, o parâmetro é estimado na fronteira igual a zero e por isso não podemos utilizar o Hessiano para estimar a variância da estimativa.

de probabilidade normal dos resíduos GARCH(2,2) observando que o comportamento é apenas insatisfatório nas caudas.

**Tabela 3.14**  
**Análise dos resíduos e dos quadrados dos resíduos ARCH na série Taxa de Câmbio**

	GARCH (1,1)		GARCH (2,2)		EGARCH (2,0)	
A1	3		3		2	
A2	14		14		14	
B1	0		0		0	
B2	2		2		2	
Box-Ljung						
Lag 6	0,687 <sup>a</sup>	0,566 <sup>b</sup>	0,654	0,609	0,564	0,929
Lag 12	0,636	0,609	0,672	0,819	0,639	0,815
Lag 18	0,227	0,425	0,241	0,478	0,304	0,739
Lag 24	0,342	0,517	0,355	0,579	0,420	0,706
Lag 30	0,283	0,580	0,309	0,579	0,271	0,824
Lag 36	0,382	0,304	0,409	0,251	0,354	0,401
Lag 42	0,429	0,197	0,453	0,150	0,389	0,276

A1 (A2): número de autocorrelações maiores que um desvio padrão nos 10 (50) primeiros lags.

Valor esperado 3(15).

B1 (B2): número de autocorrelações maiores que 1,96 desvios padrões nos 10 (50) primeiros lags.

Valor esperado 0,5 (2,5).

<sup>a</sup> valores-P do teste Box-Ljung para os resíduos

<sup>b</sup> valores-P do teste Box-Ljung para quadrados dos resíduos

Com relação à previsão um-passo a frente *ex-ante*, os intervalos de confiança 95% construídos com a volatilidade GARCH(2,2) excluem 5,66% dos retornos (100 em 1767). O valor-P associado a essa estimativa é igual a 0,22, em consequência os resultados são satisfatórios. Por sua vez, como evidencia a Tabela 3.15, existem vários resíduos com valores muito grandes (em valor absoluto).

**Tabela 3.15**  
**Resíduos extremos GARCH(2,2) na Taxa de Câmbio**

Data	Retorno	Volatilidade ( $\times 10^5$ )	Resíduo	Frequência Esperada
23/09/85	0,04816	8,19	5,35	$8 \times 10^{-5}$
5/03/86	-0,03238	5,49	-4,45	$7,57 \times 10^{-3}$
25/09/89	0,02758	4,45	4,14	0,03
21/09/84	0,02998	6,42	3,66	0,23
15/01/88	-0,03255	8,14	-3,60	0,28
20/12/90	-0,02166	3,69	-3,53	0,37
4/01/90	0,02815	6,14	3,52	0,39
5/01/88	-0,02906	7,42	-3,45	0,49
20/10/87	-0,02205	4,36	-3,44	0,51
26/05/87	-0,01942	3,56	-3,37	0,66

Os resultados dessa tabela servem para mostrar que sob a suposição de que o modelo verdadeiro é o GARCH(2,2), a ocorrência desses resíduos é muito improvável, sendo a probabilidade de obter 10 resíduos maiores (em valor absoluto) a 3,37 inferior a  $10^{-4}$ .

Na Figura 3.15a mostramos os intervalos de confiança (95%) para o período da série em que ocorre o maior resíduo. Pelo comportamento dos quadrados dos retornos antes e depois de 23/09/85 podemos considerar esse retorno como sendo um *outlier* do tipo AO. O resíduo correspondente a essa observação encontra-se claramente afastado do resto dos resíduos (extremo superior a direita na Figura 3.14. O segundo maior resíduo (Figura 3.16a) ocorre num período no qual existem várias observações grandes nos retornos que não estão incluídas dentro dos intervalos de confiança (95%). Da análise realizada, podemos concluir que, embora o modelo possua um bom desempenho na previsão *ex-ante*, aparecem vários resíduos cuja existência é pouco provável se o modelo considerado for verdadeiro.

### *Seleção do Modelo EGARCH*

Considerando os modelos EGARCH, a Tabela 3.12 revela que o critério AIC conduz à eleição do modelo EGARCH(2,0) seguido do EGARCH(2,1), EGARCH(1,0) e EGARCH(2,2), em ordem de preferência. Por sua vez, ao considerar o BIC a ordenação para a escolha é: EGARCH(2,0), EGARCH(1,0), EGARCH(1,1) e EGARCH(2,1). Portanto, ambos os critérios conduzem à escolha do EGARCH(2,0). Na Tabela 3.16 observamos que no EGARCH(2,0) a estimativa correspondente à Quinta Feira é significativa 5% e a

correspondente à Segunda Feira possui valor-P igual a 0,10. As estimativas correspondentes aos outros dias são não significativas (valores-P > 0,25).

**Tabela 3.16**  
Estimativas dos modelos EGARCH na série *Taxa de Câmbio*

	EGARCH (1,0)	EGARCH (1,1)	EGARCH (1,2)	EGARCH (2,0)	EGARCH (2,1)	EGARCH (2,2)
Seg	<b>0,0006<sup>a</sup></b> <i>0,0004<sup>b</sup></i> 0,1183 <sup>c</sup>	<b>0,0006</b> <i>0,0004</i> 0,0802	<b>0,0007</b> <i>0,0004</i> 0,0741	<b>0,0006</b> <i>0,0004</i> 0,0989	<b>0,0006</b> <i>0,0004</i> 0,1308	<b>0,0006</b> <i>0,0004</i> 0,0922
Ter	<b>0,00004</b> <i>0,0004</i> 0,9132	<b>0,00001</b> <i>0,0004</i> 0,9725	<b>0,00003</b> <i>0,0004</i> 0,9363	<b>-8,456x10<sup>-9</sup></b> <i>9,115x10<sup>-6</sup></i> 0,9993	<b>8,973x10<sup>-6</sup></b> <i>0,0004</i> 0,9812	<b>4,662x10<sup>-6</sup></b> <i>0,0003</i> 0,9891
Quar	<b>-0,0002</b> <i>0,0004</i> 0,6798	<b>-0,0002</b> <i>0,0004</i> 0,6440	<b>-0,0001</b> <i>0,0004</i> 0,7045	<b>-0,0002</b> <i>0,0004</i> 0,6762	<b>-0,0002</b> <i>0,0004</i> 0,5333	<b>-0,0002</b> <i>0,0004</i> 0,5396
Quin	<b>0,0007</b> <i>0,0004</i> 0,0750	<b>0,0007</b> <i>0,0004</i> 0,0569	<b>0,0007</b> <i>0,0004</i> 0,0513	<b>0,0008</b> <i>0,0004</i> 0,0394 *	<b>0,0008</b> <i>0,0004</i> 0,0476 *	<b>0,0008</b> <i>0,0004</i> 0,0517
Sex	<b>-0,0002</b> <i>0,0004</i> 0,6437	<b>-0,0002</b> <i>0,0002</i> 0,2534	<b>-0,0002</b> <i>0,0004</i> 0,6128	<b>-0,0003</b> <i>0,0004</i> 0,4540	<b>-0,0001</b> <i>0,0004</i> 0,6955	<b>-0,0002</b> <i>0,0004</i> 0,6778
$\kappa$	<b>-0,3624</b> <i>0,1035</i> 0,0005*	<b>-0,3693</b> <i>0,1057</i> 0,0012*	<b>-0,4081</b> <i>0,1096</i> 0,0002*	<b>-0,1436</b> <i>0,0304</i> 0,0001*	<b>-0,7121</b> <i>0,1584</i> 0,0001*	<b>-0,7152</b> <i>0,1960</i> 0,0003
$\Delta_1$	<b>0,9625</b> <i>0,0105</i> 0,0001*	<b>0,9618</b> <i>0,0108</i> 0,0001*	<b>0,9578</b> <i>0,0112</i> 0,0001*	<b>1,8162</b> <i>0,0430</i> 0,0001*	<b>0,0294</b> <i>0,0376</i> 0,4335	<b>0,0277</b> <i>0,0394</i> 0,4816
$\Delta_2$				<b>-0,8310</b> <i>0,04334</i> 0,0001*	<b>0,8969</b> <i>0,0365</i> 0,0001*	<b>0,8983</b> <i>0,0379</i> 0,0001*
$\Psi_1$	<b>0,1477</b> <i>0,0231</i> 0,0001*	<b>0,0944</b> <i>0,0559</i> 0,0910	<b>0,0796</b> <i>0,0519</i> 0,1247	<b>0,0402</b> <i>0,0077</i> 0,0001*	<b>0,0993</b> <i>0,0295</i> 0,0008*	<b>0,0899</b> <i>0,0493</i> 0,0680
$\Psi_2$		<b>0,0573</b> <i>0,0539</i> 0,2877	<b>-0,0144</b> <i>0,0657</i> 0,8266		<b>0,1876</b> <i>0,0280</i> 0,0001*	<b>0,1875</b> <i>0,0286</i> 0,0001*
$\Psi_3$			<b>0,0950</b> <i>0,0542</i> 0,0797			<b>0,0114</b> <i>0,0389</i> 0,7704
$\theta$	<b>-0,0299</b> <i>0,0823</i> 0,7166	<b>-0,0366</b> <i>0,0856</i> 0,6694	<b>-0,0619</b> <i>0,0734</i> 0,3991	<b>-0,0783</b> <i>0,0770</i> 0,3087	<b>-0,0621</b> <i>0,0745</i> 0,4072	<b>-0,0217</b> <i>0,0762</i> 0,7756
AIC	-12345,7	-12344,8	-12345,8	-12358,2	-12347,6	-12345,6
BIC	-12296,4	-12290,0	-12285,5	-12303,5	-12287,4	-12279,9

<sup>a</sup> estimativa pontual.

<sup>b</sup> desvio padrão estimado.

<sup>c</sup> Valor-P.

\* significativo 5%

Na equação da volatilidade estimada:

$$\ln(\sigma_t^2) = -0,1436 + 1,8162 \ln(\sigma_{t-1}^2) - 0,8310 \ln(\sigma_{t-2}^2) + 0,0402g(\varepsilon_{t-1}) \quad (3.7.2.3)$$

(0,0304)    (0,0429)                    (0,0433)                    (0,0077)

$$g(\varepsilon_t) = -0,0783\varepsilon_t + |\varepsilon_t| - \sqrt{2/\pi} ,$$

(0,0769)

todas as estimativas são significativas (valores-P inferiores a 0,0001) com exceção do corresponde à assimetria que possui valor-P igual a 0,31, em consequência não há evidência de *Leverage Effect*. Por outro lado, as raízes de  $1 - 1,8162x + 0,8310x^2 = 0$  não são reais, mas as raízes de  $1 - 1,8x + 0,8x^2 = 0$  são iguais a 0,8 e 1. Portanto pelo Teorema 1 (seção 3.3.4) as volatilidades não são ergódicas nem estritamente (ou covariante) estacionárias. O gráfico da volatilidade estimada pelo modelo EGARCH (2,0) é apresentado na Figura 3.12 sendo observado que as variações nos quadrados dos retornos são acompanhados pelas variações da volatilidade estimada.

#### *Adequabilidade do modelo EGARCH(2,0)*

Ao analisar o comportamento da f.a.c. dos resíduos e seus quadrados (Figuras 3.13c-d), podemos observar que os resíduos EGARCH(2,0) não apresentam correlação serial e, além disso, a não linearidade é removida. Esses resultados são confirmados pelos dados da Tabela 3.14. Adicionalmente os resíduos apresentam estimativas da média, variância e coeficiente de simetria satisfatórios sendo iguais a 0,0006, 1,00 e 0,11 respectivamente. O excesso de curtose é igual a 1,23 e está afetado pelos valores extremos dos resíduos, sendo igual a 0,37 ao retirar os 9 maiores resíduos da amostra em valor absoluto. Isso indicaria que a distribuição dos resíduos é leptocúrtica, mas, através da estatística  $W$ , não rejeitamos a hipótese de normalidade com valor-P igual a 0,29. Adicionalmente, embora não apresentado, o Gráfico de Probabilidade Normal é satisfatório com exceção das caudas, onde é observada um pequeno número de pontos que apresentam afastamento com relação à reta normal.

Na Tabela 3.17 são apresentados os 10 maiores resíduos em valor absoluto, sendo observado que vários deles estão incluídos na Tabela 3.15 e, em particular os três maiores acontecem na mesma data.

**Tabela 3.17**  
**Resíduos extremos EGARCH(2,0) na série Taxa de Câmbio**

Data	Retorno	Volatilidade ( $\times 10^5$ )	Resíduo	Frequência Esperada
23/09/85	0,048161	9,29	5,03	$4,4 \times 10^{-4}$
5/03/86	-0,032379	5,52	-4,44	$7,9 \times 10^{-3}$
25/09/89	0,027580	4,34	4,22	0,022
5/01/88	-0,029055	5,62	-3,98	0,060
20/10/87	-0,022052	3,79	-3,71	0,18
21/09/84	0,029978	6,20	3,70	0,19
15/01/88	-0,032549	8,34	-3,56	0,32
4/01/90	0,028154	6,02	3,55	0,34
26/05/87	-0,019418	3,68	-3,33	0,76
5/07/89	0,025197	5,95	3,30	0,84

Supondo que o EGARCH(2,0) é o verdadeiro modelo, a probabilidade da existência de 10 resíduos maiores a 3,30 (décimo resíduo) na amostra é inferior a  $10^{-4}$ , o qual indica que a existência desses resíduos é implausível. Intervalos de confiança (95%) excluem 5,66 % dos retornos (valor-P para testar 5% é igual a 0,22) sendo esse resultado satisfatório.

#### *Modelo GARCH(1,1)*

Dos dados apresentados na Tabela 3.13, nenhuma das estimativas correspondentes aos dias da semana são significativas a 5% no GARCH(1,1). Entretanto, a estimativa correspondente à Quinta Feira possui um valor-P igual a 0,08. Por sua vez, todas as estimativas correspondentes à variância condicional são altamente significativas. A persistência estimada é bastante grande, 0,9579, significando uma média-vida igual a 16,12 (aproximadamente três semanas de funcionamento de mercado).

Quanto à adequabilidade do modelo GARCH(1,1), tanto pela análise gráfica das f.a.c dos resíduos e seus quadrados, quanto pelos valores-P do teste Box-Ljung apresentados na Tabela 3.14, podemos considerar que os resíduos não apresentam correlação serial e que a não-linearidade foi removida. Por outro lado, os resíduos do modelo GARCH(1,1) apresentam valores de média, variância e coeficiente de simetria iguais a 0,0004, 1,00 e 0,15 respectivamente, os quais são satisfatórios. Embora o valor do excesso de curtose, 1,3 indique leptocurtose na distribuição dos resíduos, o teste  $W$  possui valor-P igual a 0,33. Como nos casos anteriores, isto é explicado pelo fato da grande maioria dos dados ter comportamento

próximo da normal como foi verificado no Gráfico de Probabilidade Normal, que não é apresentado.

Com relação à previsão um-passo à frente *ex-ante*, os intervalos de confiança 95% construídos com a volatilidade GARCH(1,1) excluem 5,55% (98 em 1767) dos retornos, sendo esses resultados satisfatórios. Por sua vez, como mostra a Tabela 3.18, existem vários resíduos com valores muito grandes (em valor absoluto).

**Tabela 3.18**  
**Resíduos extremos GARCH(1,1) na série Taxa de Câmbio**

Data	Retorno	Volatilidade ( $\times 10^5$ )	Resíduo	Frequência Esperada
23/09/85	0,04816	8,56	5,23	$1,5 \times 10^{-4}$
5/03/86	-0,03238	5,42	-4,47	$6,8 \times 10^{-3}$
25/09/89	0,02758	4,27	4,22	0,02
15/01/88	-0,03255	6,99	-3,89	0,09
21/09/84	0,02998	6,20	3,72	0,18
4/01/90	0,02815	5,78	3,63	0,25
5/01/88	-0,02905	6,92	-3,58	0,31
26/05/87	-0,01942	3,43	-3,43	0,53
20/12/90	-0,02966	3,89	-3,43	0,54
20/10/87	-0,02205	4,58	-3,36	0,69

Os resultados dessa tabela servem para mostrar que sob a suposição que o modelo verdadeiro é o GARCH(1,1), a ocorrência desses resíduos é muito improvável, sendo a probabilidade de obter 10 resíduos maiores (em valor absoluto) a  $3,36 \times 10^{-4}$ .

Da análise podemos concluir que, embora o modelo possua um bom desempenho na previsão *ex-ante*, aparecem vários resíduos cuja existência é pouco provável se o modelo considerado for verdadeiro.

#### *Comparação dos Modelos GARCH(1,1), GARCH(2,2) e EGARCH(2,0)*

Os modelos GARCH(1,1) e GARCH(2,2) apresentam estimativas de volatilidade bastante próximas, como pode ser visto na Figura 3.17a. O mesmo é observado quando

comparamos os modelos GARCH (2,2) e EGARCH (2,0) (ver Figura 3.17b). Nos parágrafos a seguir é realizada uma comparação dos modelos escolhidos através dos critérios AIC e BIC.

Como é observado na Figura 3.12, as volatilidades estimadas pelos modelos GARCH(2,2) e EGARCH(2,0) são capazes de acompanhar as variações (aumentos e diminuições) dos quadrados dos retornos<sup>30</sup>, mas eles diferem em certas características. Primeiro, são observadas diferenças consideráveis entre as volatilidades para alguns retornos (Figura 3.17a), sendo em alguns casos bastante grandes. Isso é devido ao fato da volatilidade estimada pelo GARCH(2,2) reagir de forma diferente à presença de valores grandes (extremos) dos quadrados dos retornos quando comparada com a volatilidade EGARCH(2,0). Por exemplo, na Figura 3.15a observamos que no um dia após a observação extrema no dia 23/09/85, ambas volatilidades estimadas aumentam, mas a volatilidade estimada GARCH(2,2) o faz abruptamente. O mesmo acontece na Figura 3.16a nos dias 5/03/86 e 24/03/86. Nesse sentido, ao comparar as expressões (3.7.2.2) e (3.7.2.3), os dados evidenciam que o efeito aditivo de  $|Y_t|$  na volatilidade estimada GARCH(2,2) é maior que o efeito multiplicativo  $|Y_t|$  na volatilidade estimada EGARCH(2,0) no instante seguinte à ocorrência da observação extrema. Por sua vez, na Figura 3.17b observamos que a volatilidade EGARCH(2,0) também pode ser muito maior do que a volatilidade GARCH(2,2). Na Figura 3.18a é apresentado um período no qual isso acontece. Desde 27/02/85 é observado que a volatilidade EGARCH(2,0) é maior que a volatilidade estimada GARCH(2,2). Neste caso as diferenças começam no dia 1/03/85 em que ocorre uma observação muito pequena (em valor absoluto).

A segunda diferença, também relacionada com a primeira, deve-se ao fato de que logo após a ocorrência dos valores grandes dos retornos, em alguns períodos, como por exemplo nas apresentadas nas Figuras 3.15b e 3.16b, os intervalos de confiança 95% (i.c) para os retornos utilizando a volatilidade GARCH(2,2) apresentam de maneira geral, comprimento maior aos i.c. construídos com a volatilidade EGARCH(2,0).

Outra diferença consiste na forma da evolução das volatilidades sendo no EGARCH(2,0) mais *suave* quando comparada à volatilidade GARCH(2,2) (Figuras 3.12, 3.15a, 3.16a, 3.18a).

Por outro lado, com relação à previsão um-passo à frente *ex-ante*, o desempenho dos modelos GARCH(2,2) e EGARCH(2,0) é praticamente o mesmo. Assim, intervalos de

<sup>30</sup>Embora não apresentado, no gráfico da volatilidade GARCH(1,1) também é observado isso.

confiança 95% para os retornos (i.c.) incluem o mesmo número de retornos nos dois modelos e, por outro lado, as percentagens observadas nas Tabelas 3.19 e 3.20 são muito próximas.

**Tabela 3.19**  
Previsão um-passo à frente *ex-ante* considerando níveis na volatilidade estimada  
GARCH (2,2) na Taxa de Câmbio

Nível	n	Volatilidade ( $\times 10^5$ )		numero de retornos fora	Percentagem	Valor-P
		b	c			
baixo	589	2,842	4,492	42	7,13	0,02
médio	589	4,492	5,780	27	4,58	0,65
alto	589	5,784	27,6	31	5,26	0,78

<sup>a</sup> Teste bilateral com aproximação Poisson supondo independência nas observações

<sup>b</sup> menor valor apresentado em cada nível.

<sup>c</sup> maior valor apresentado em cada nível.

**Tabela 3.20**  
Previsão um-passo à frente *ex-ante* considerando níveis na volatilidade estimada  
EGARCH(2,0) na Taxa de Câmbio

Nível	n	Volatilidade ( $\times 10^5$ )		numero de retornos fora	Percentagem	Valor-P
		b	c			
baixo	589	2,151	4,478	41	6,96	0,03
médio	589	4,484	5,992	28	4,75	0,79
alto	589	5,992	16,2	31	5,26	0,78

<sup>a</sup> Teste bilateral com aproximação Poisson supondo independência nas observações

<sup>b</sup> menor valor apresentado em cada nível.

<sup>c</sup> maior valor apresentado em cada nível.

Nessas tabelas observamos que quando a volatilidade é baixa, os intervalos de confiança (i.c) em ambos os modelos excluem mais retornos do que os esperados.

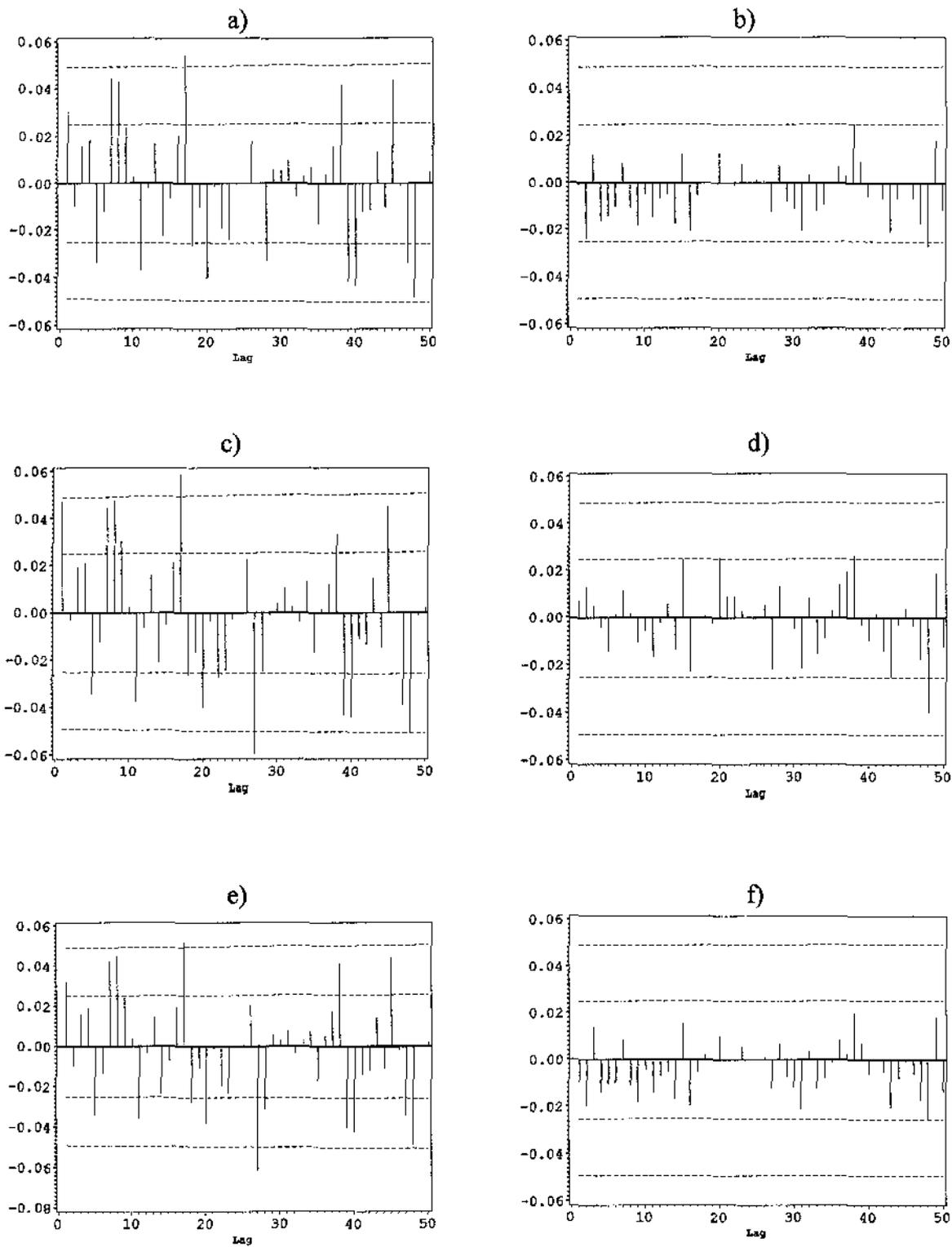
Nas Figuras 3.15b, 3.16b e 3.18b, correspondentes basicamente a períodos de volatilidades média e alta, é possível observar que os i.c construídos com ambas volatilidades sempre excluem os valores extremos dos retornos. Logo após a presença das observações extremas, os i.c obtidos utilizando o GARCH(2,2) são mais compridos do que os correspondentes aos i.c estimados pelo EGARCH(2,0).

### Conclusões

A discussão anterior permite estabelecer algumas conclusões:

- Os modelos GARCH(1,1), GARCH(2,2) e EGARCH(2,0) são capazes de remover a não-linearidade dos retornos.
- A volatilidade estimada EGARCH(2,0) evolui de forma mais suave quando comparada com a volatilidade estimada GARCH(2,2) que apresenta variações bruscas.
- Os modelos GARCH(1,1) e GARCH(2,2) apresentam estimativas da volatilidade muito próximas.
- As volatilidades estimadas pelos modelos GARCH(1,1), GARCH(2,2) e EGARCH(2,0) são capazes de acompanhar os movimentos de aumentos e diminuições dos quadrados dos retornos.
- Os modelos estimados GARCH(1,1) e GARCH(2,2) são covariante estacionários enquanto que o EGARCH(2,0) não é estacionário. Nesse sentido, considerando as evidências encontradas no Capítulo 2, os modelos GARCH são preferíveis.
- Os modelo ajustado GARCH(1,1) apresenta alta persistência e, por outro lado, no modelo ajustado EGARCH(2,0) não há evidência de *Leveragge Effect*.
- Em termos de previsão *ex-ante*, ambos os modelos GARCH(2,2) e EGARCH(2,0) exibem resultados satisfatórios e similares, resultados que se mantém quando consideramos níveis na volatilidade. Em particular, quando a volatilidade é baixa encontramos uma percentagem de retornos maior à esperada.
- Para grande quantidade de retornos, os valores apresentados pelas volatilidades estimadas GARCH(2,2) e EGARCH(2,0) são similares e as grandes diferenças entre elas ocorrem nos períodos posteriores à presença de observações extremas.
- Existe uma falta de ajuste nos modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,0) evidenciada pela existência de vários resíduos extremos que não podem ser explicados com os

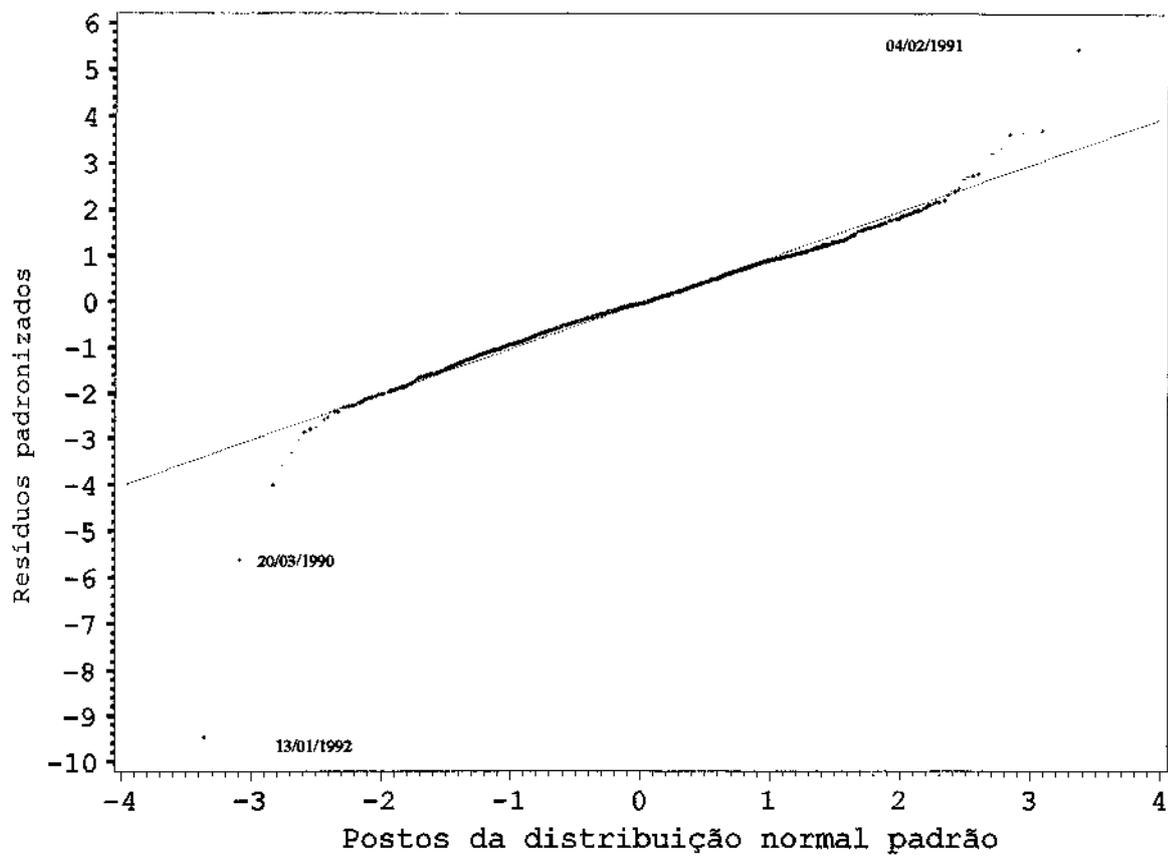
modelos estimados. Nesse sentido, é possível melhorar o ajuste considerando intervenções, e /ou, utilizando inovações com distribuições que apresentem cauda pesada, ou até procedimentos robustos.



**Figura 3.1: Funções de autocorrelação dos resíduos ARCH e dos quadrados dos resíduos nas Ações**

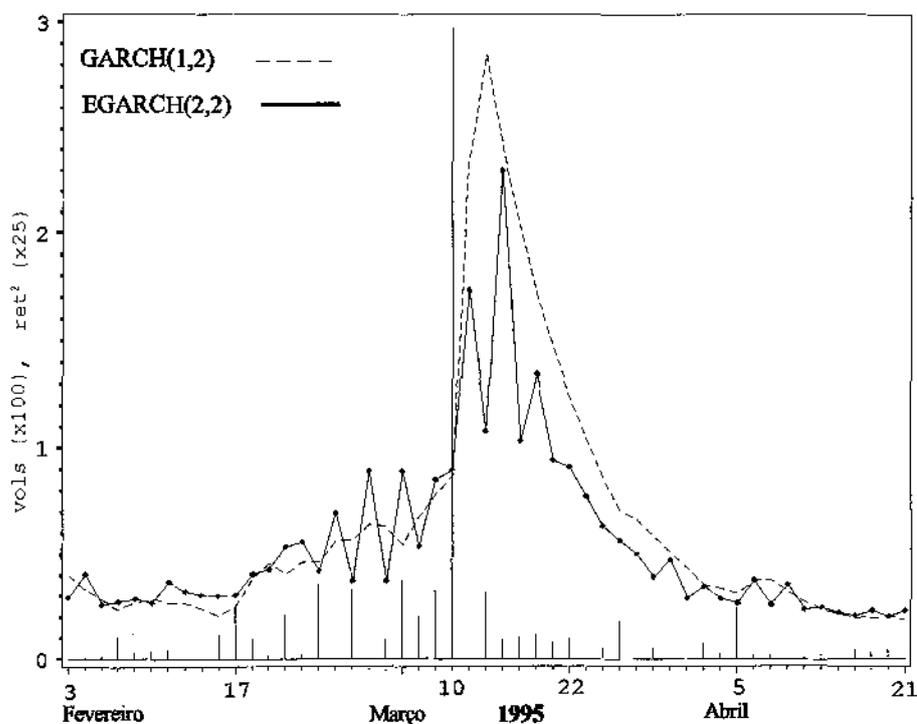
- a) f.a.c dos resíduos GARCH(1,2),    b) f.a.c dos quadrados dos resíduos GARCH(1,2)
- c) f.a.c dos resíduos EGARCH(2,2),    d) f.a.c dos quadrados dos resíduos EGARCH(2,2)
- e) f.a.c dos resíduos GARCH(1,1),    f) f.a.c dos quadrados dos resíduos GARCH(1,1)

**Legenda:** (---) limites usuais de Bartlett de 1 e 2 desvios padrão.

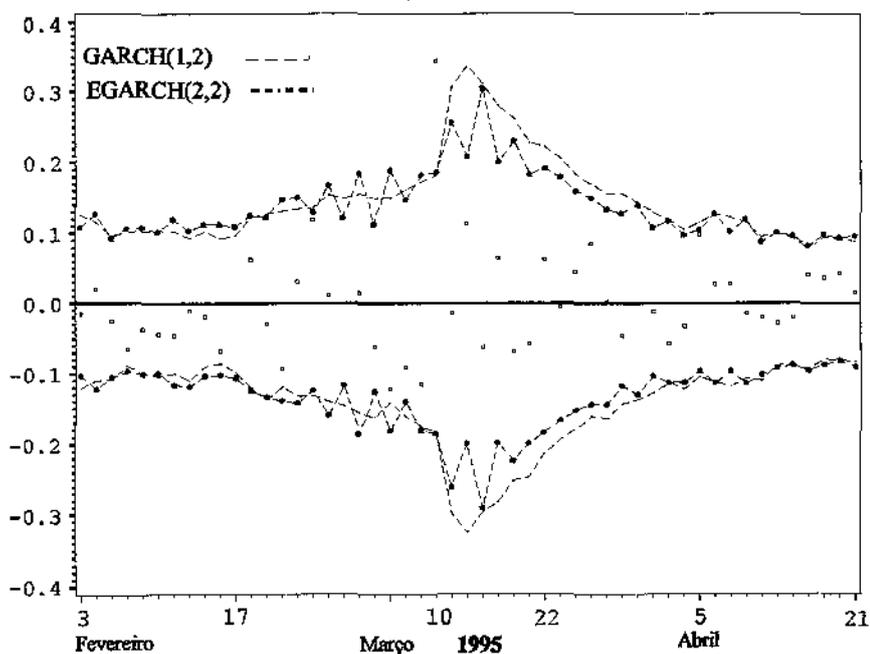


**Figura 3.2: Gráfico de Probabilidade Normal dos resíduos padronizados GARCH(1,2) nas Ações**

a) Volatilidades estimadas



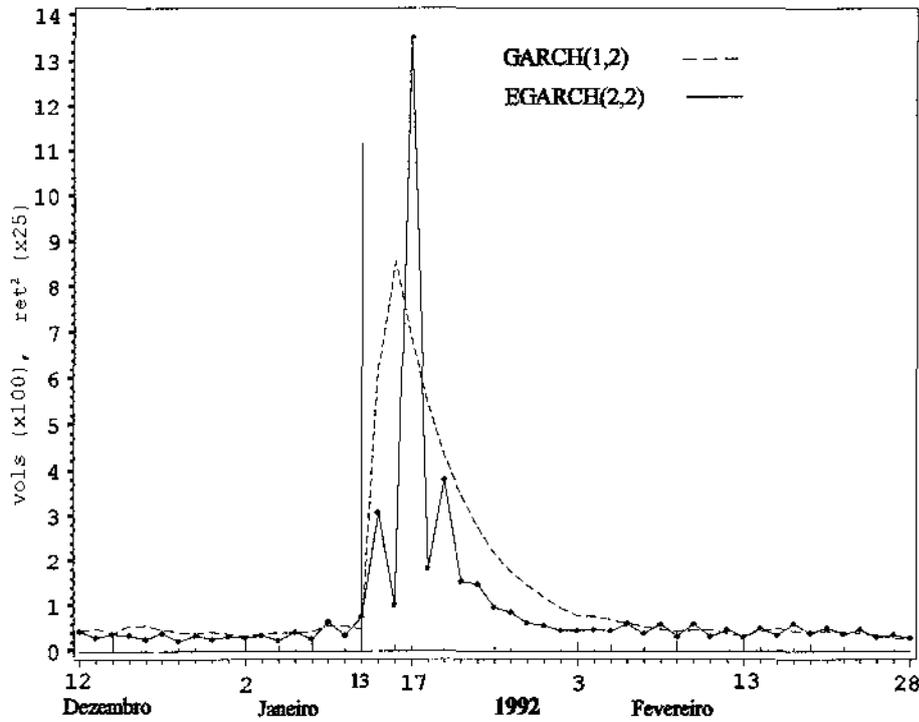
b) i.c 95% para os retornos



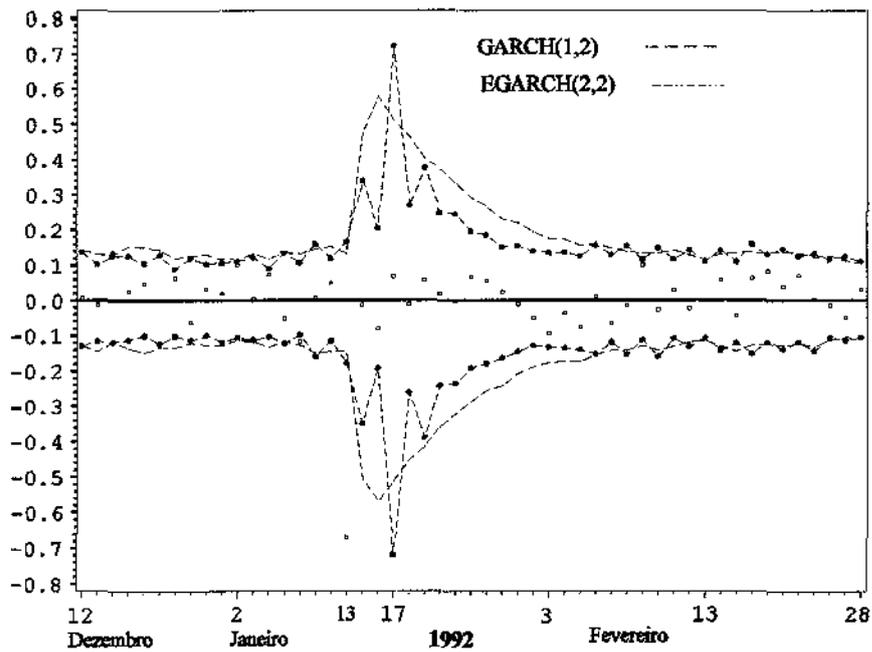
**Figura 3.3: Volatilidades estimadas ARCH e i.c 95% para os retornos no período da *Mudança da Banda Cambial***

**Legenda:** vol=volatilidades estimadas, ret<sup>2</sup>=quadrados dos retornos

a) Volatilidades estimadas



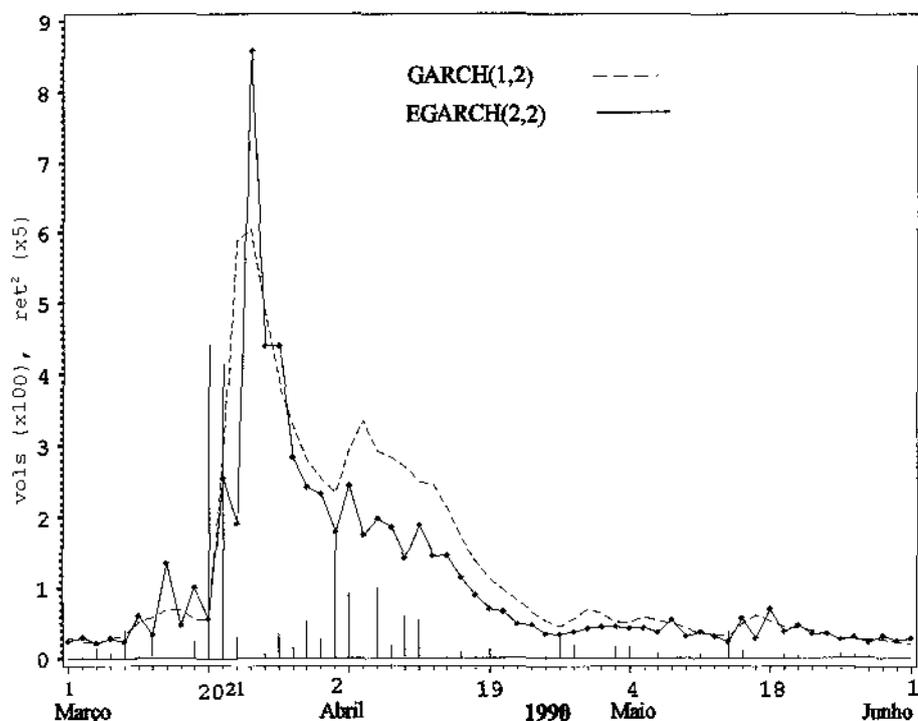
b) i.c. 95% para os retornos



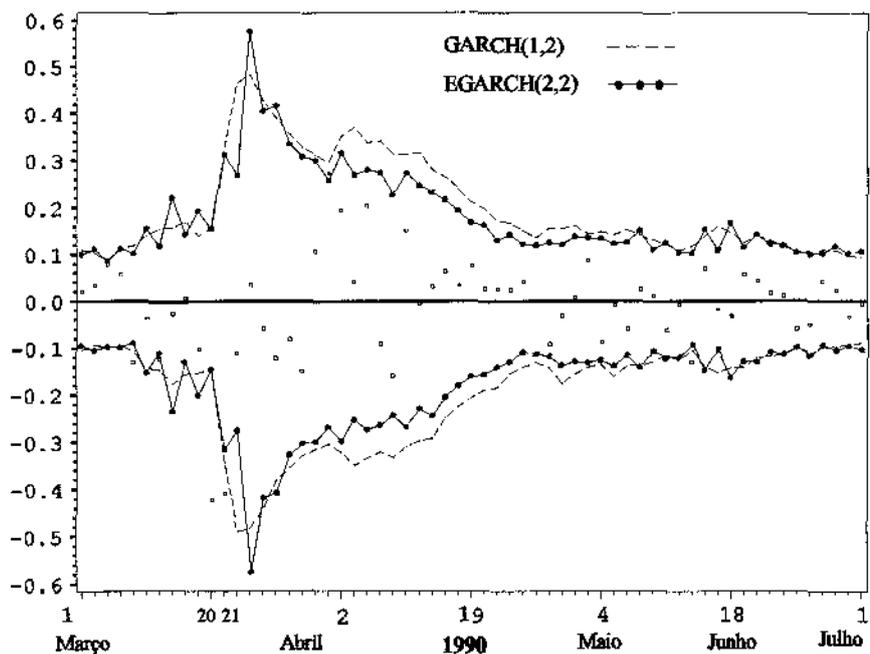
**Figura 3.4: Volatilidades estimadas ARCH e i.c 95% para os retornos no período da *Queda da Ministra Zélia Cardoso de Mello***

**Legenda:** vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos

a) Volatilidades estimadas

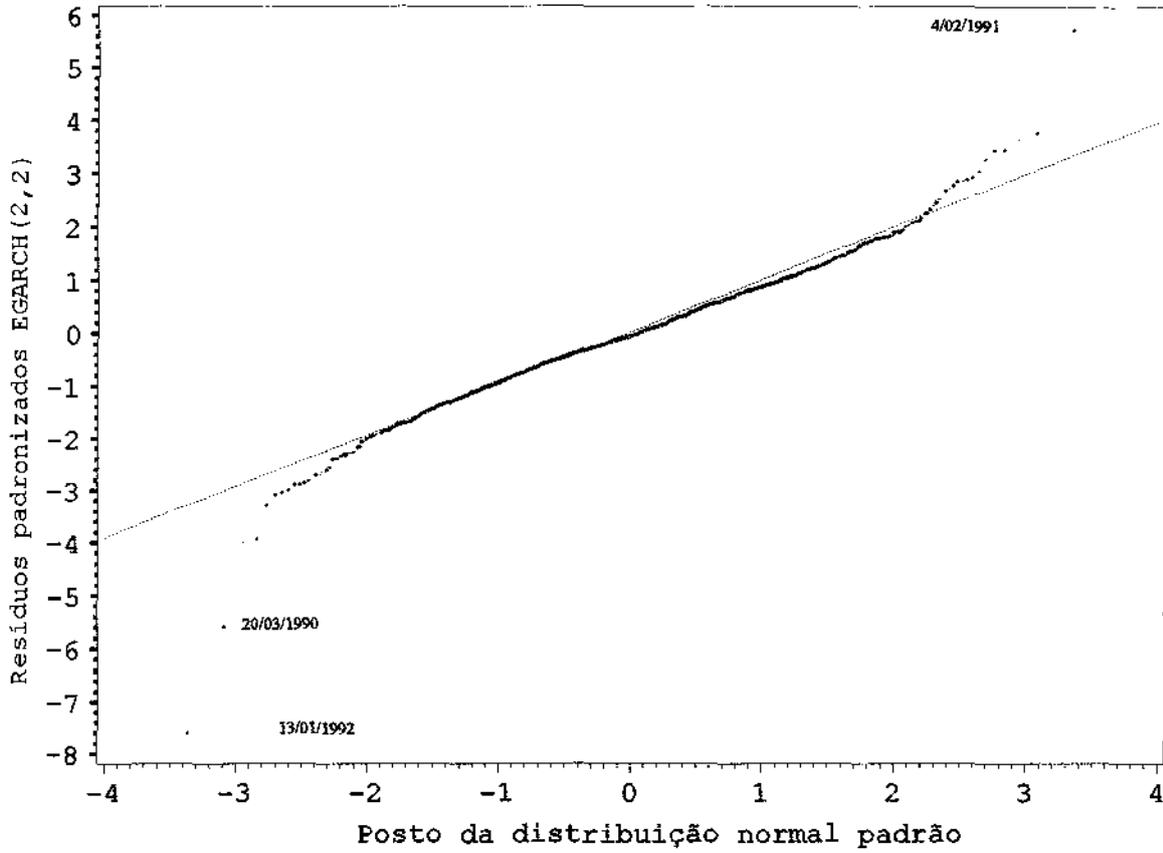


b) i.c 95% para os retornos

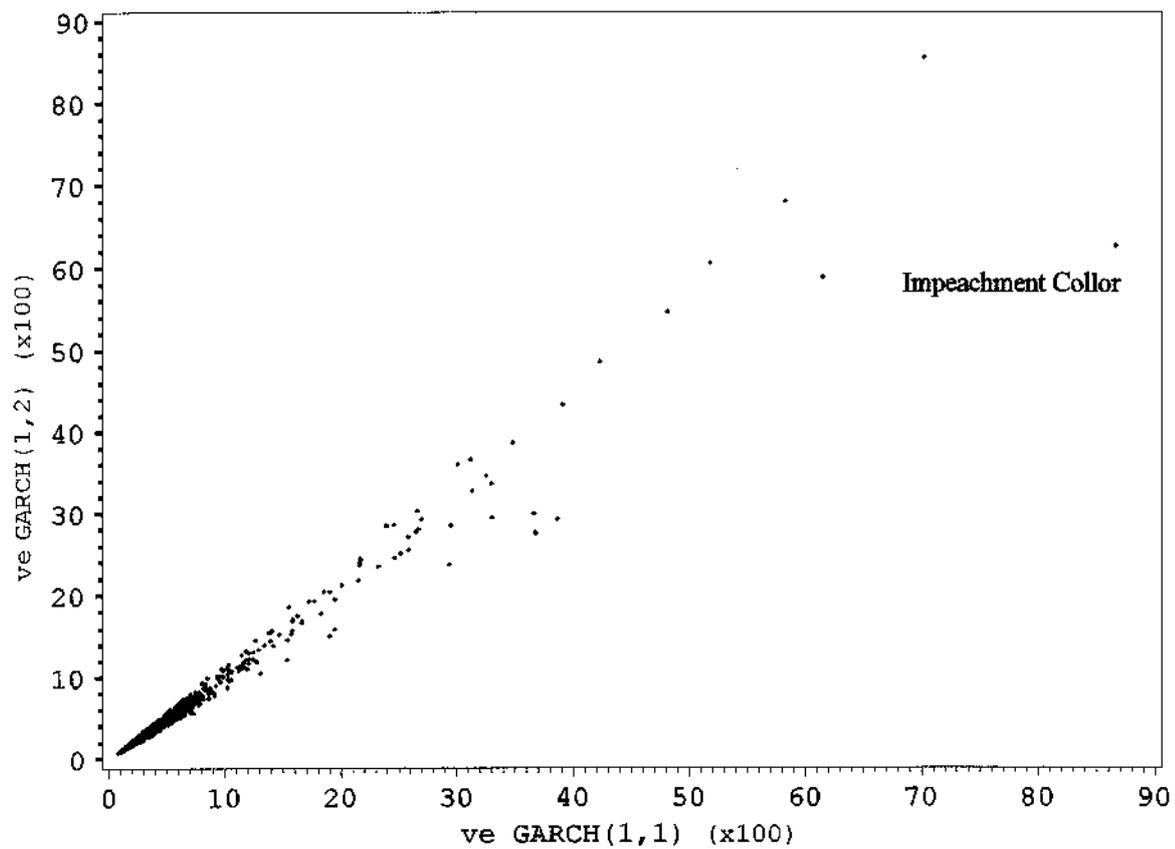


**Figura 3.5: Volatilidades estimadas e i.c 95% para os retornos no período da implementação do *Plano Collor I***

**Legenda:** vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos

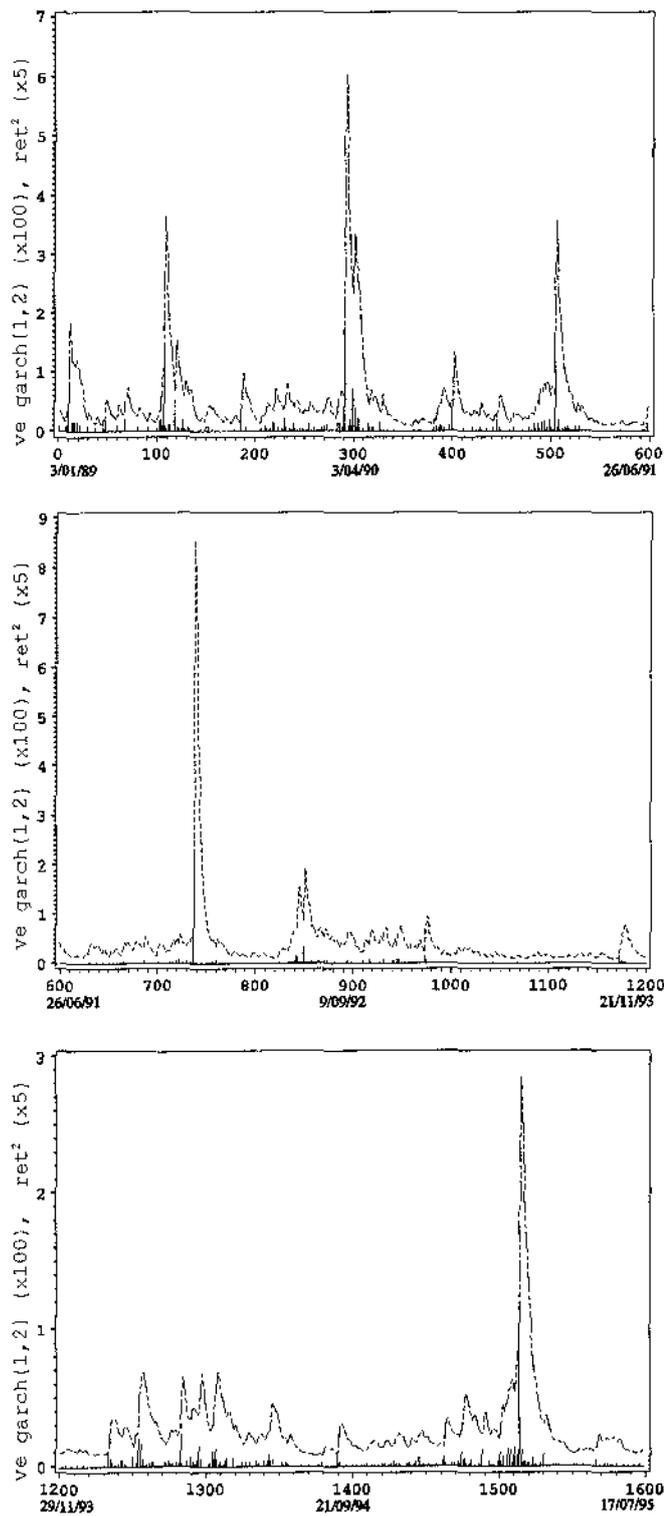


**Figura 3.6: Gráfico de Probabilidade Normal dos resíduos padronizados EGARCH(2,2) nas Ações**

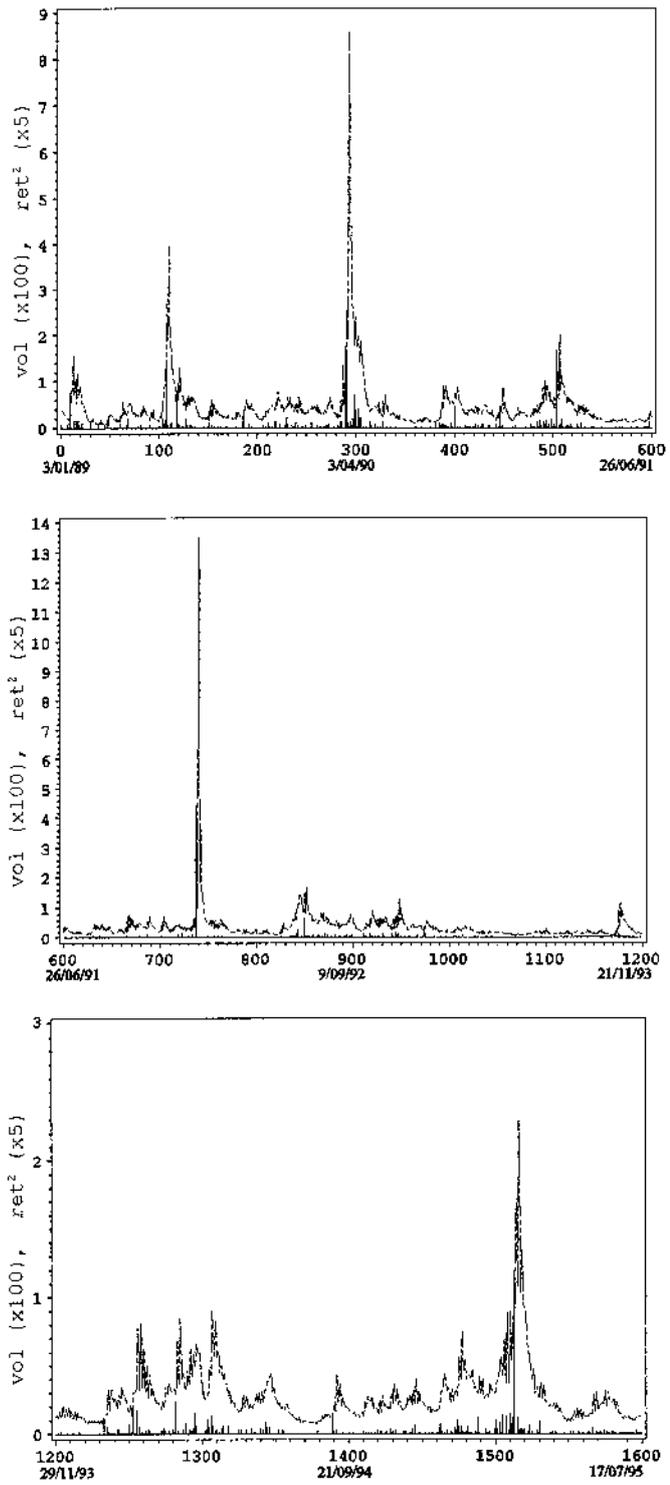


**Figura 3.7: Comparação das volatilidades estimadas GARCH(1,1) e GARCH(1,2) nas Ações**

**Legenda:** ve=volatilidade estimada



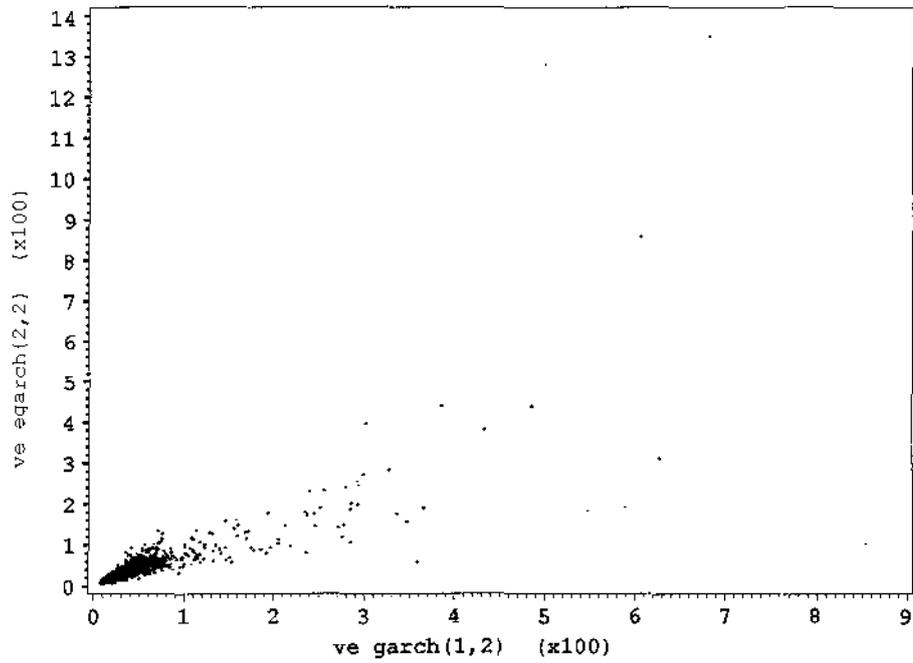
**Figura 3.8: Volatilidade estimada GARCH(1,2) nas Ações**  
**Legenda:** ve=volatilidade estimada, ret<sup>2</sup>=quadrados dos retornos



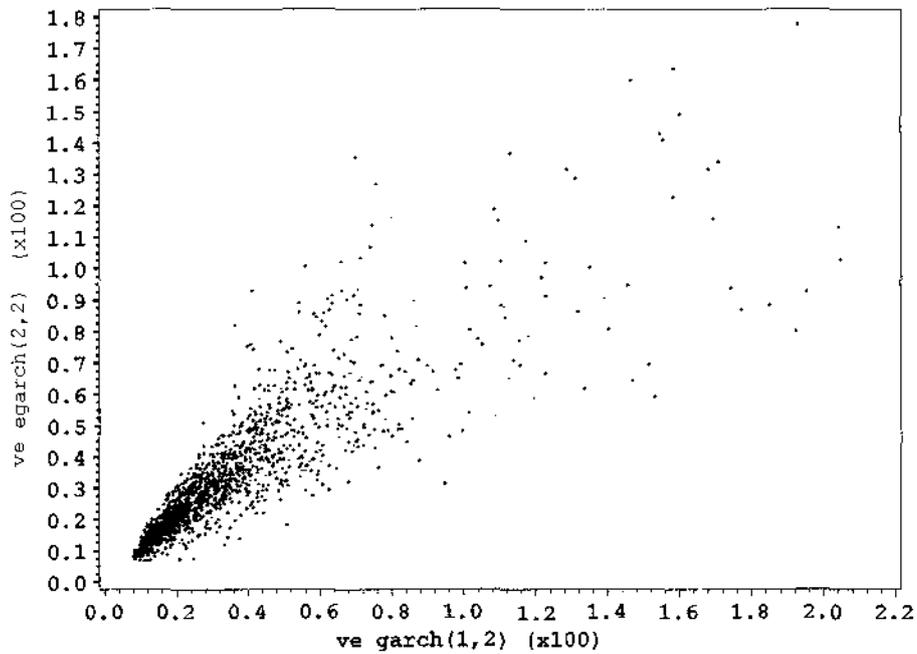
**Figura 3.9: Volatilidade estimada EGARCH(2,2) nas Ações**

**Legenda:** vol=volatilidade estimada, ret<sup>2</sup>=quadrados dos retornos

a) Todos os retornos

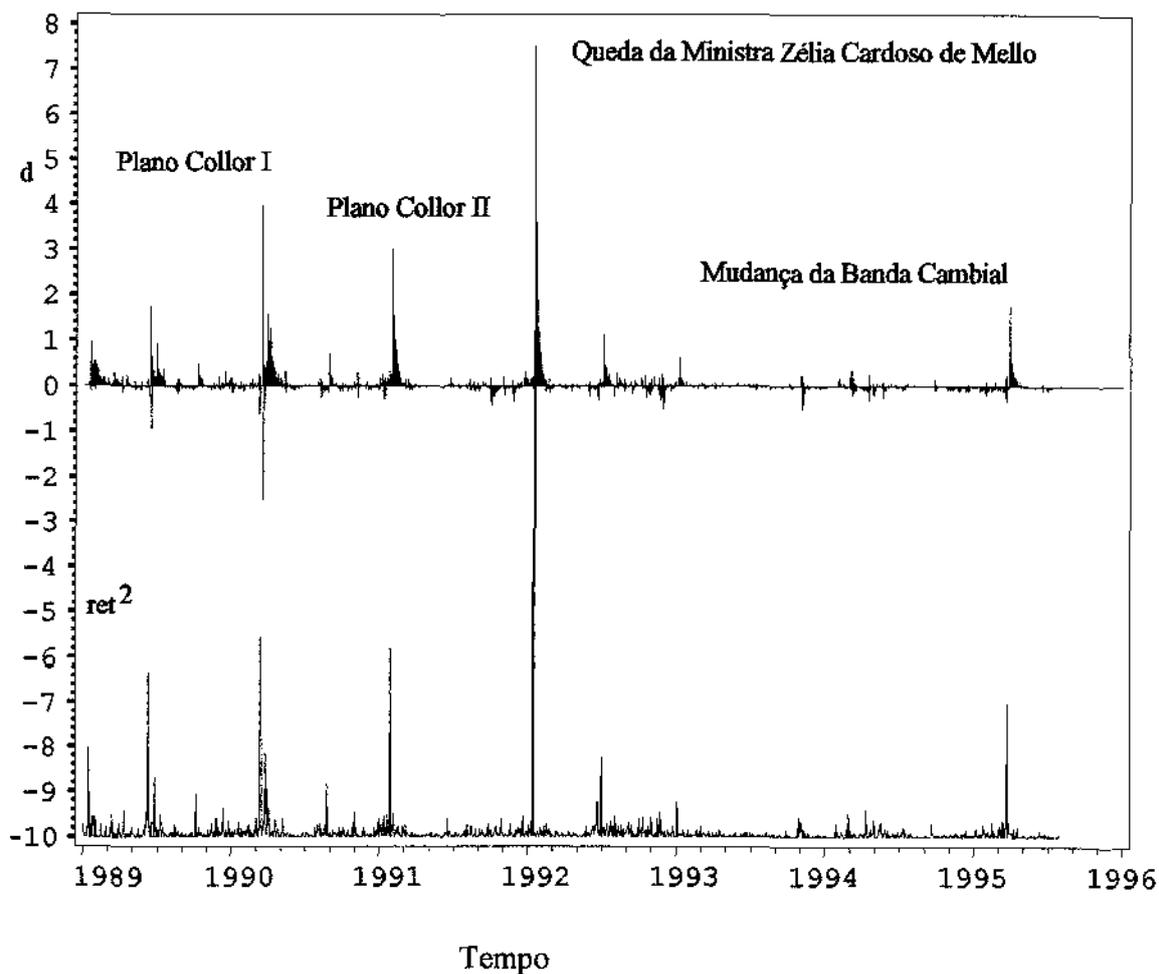


b) retornos que satisfazem:  $ve\ garch(1,2) < 2,2$



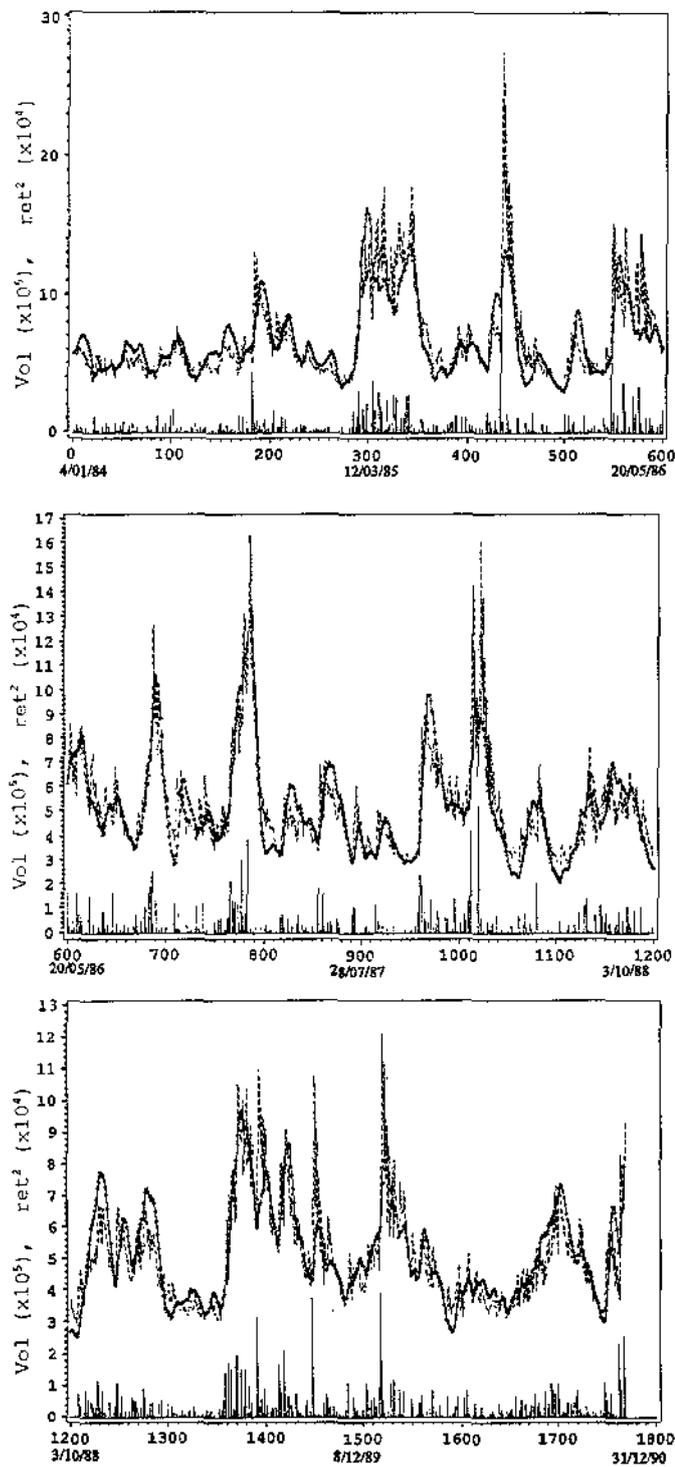
**Figura 3.10: Comparação das volatilidades estimadas GARCH(1,2) e EGARCH(2,2) nas Ações**

**Legenda:** ve=volatilidade estimada

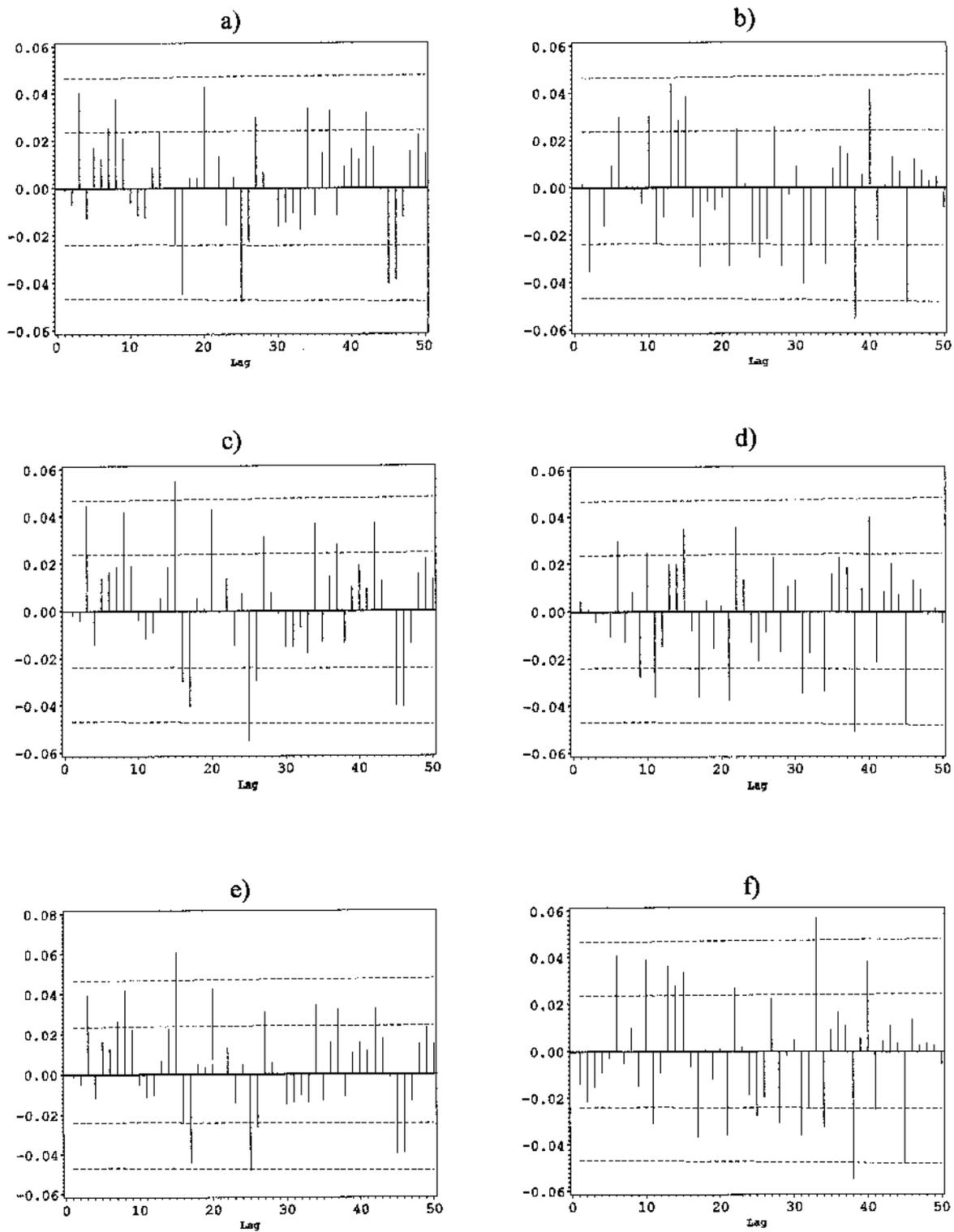


**Figura 3.11: Gráfico das diferenças entre as volatilidades estimadas GARCH(1,2) e EGARCH(2,2) nas Ações**

**Legenda:**  $d=100(\text{ve garch}(1,2)- \text{ve egarch}(2,2))$ ,  $\text{ret}^2=(25*\text{ret}*\text{ret})-10$   
ve=volatilidade estimada, ret=retorno



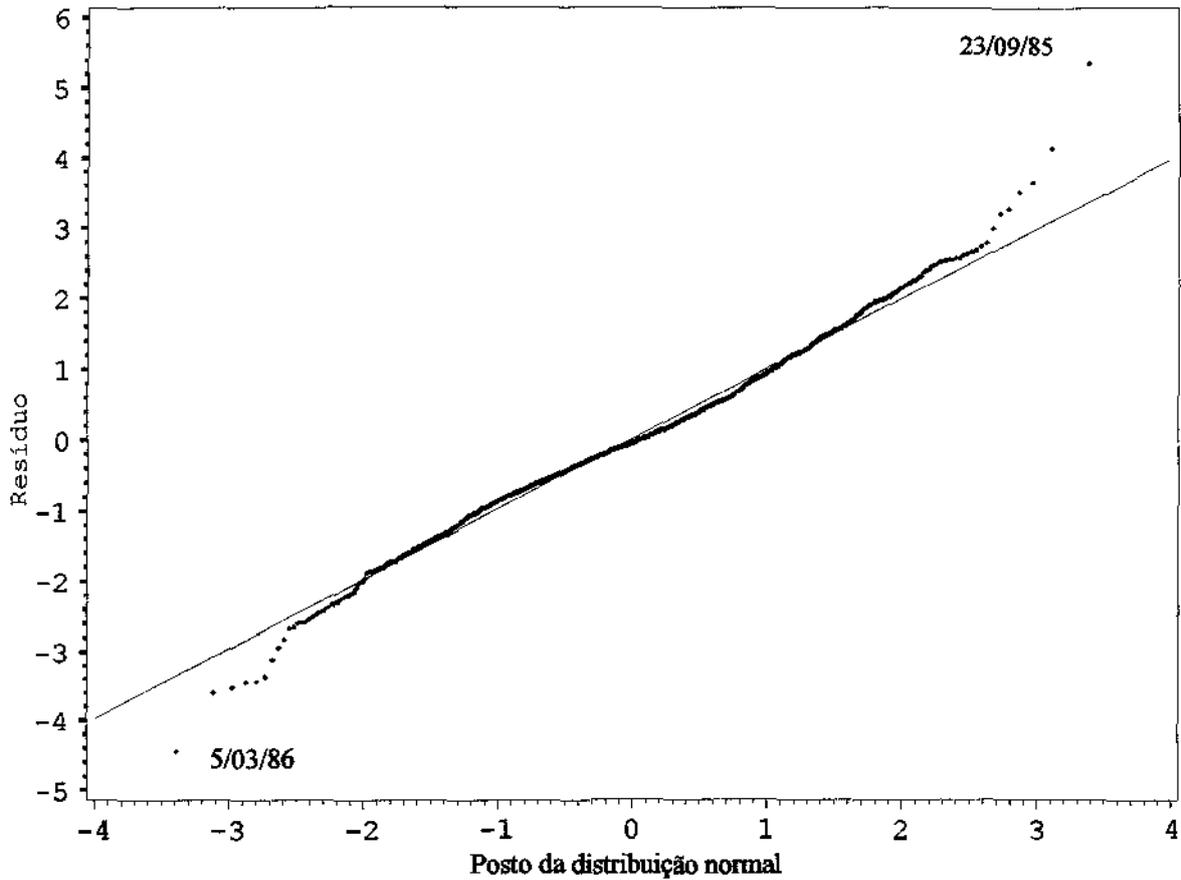
**Figura 3.12: Volatilidades estimadas GARCH(2,2) e EGARCH(2,0) na Taxa de Câmbio**  
**Legenda: GARCH(2,2) - - - , EGARCH(2,0) ——— , vol=volatilidades, ret²=quadrados dos retornos**



**Figura 3.13: Funções de autocorrelação dos resíduos ARCH e dos quadrados dos resíduos na Taxa de Câmbio**

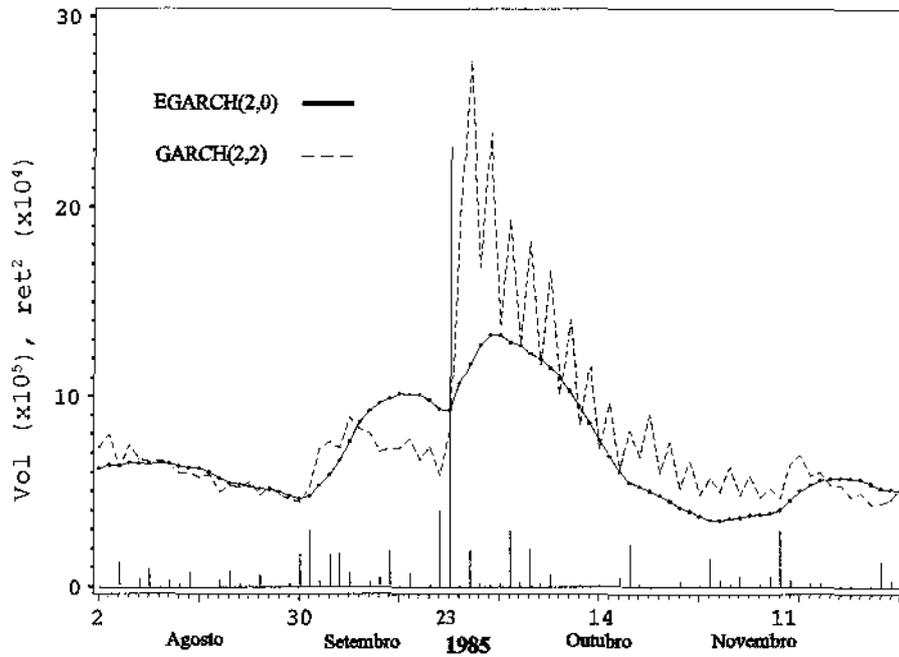
- a) f.a.c dos resíduos GARCH(2,2),    b) f.a.c dos quadrados dos resíduos GARCH(2,2)
- c) f.a.c dos resíduos EGARCH(2,0),    d) f.a.c dos quadrados dos resíduos EGARCH(2,0)
- e) f.a.c dos resíduos GARCH(1,1),    f) f.a.c dos quadrados dos resíduos GARCH(1,1)

**Legenda:** (---) limites usuais de Bartlett de 1 e 2 desvios padrão.

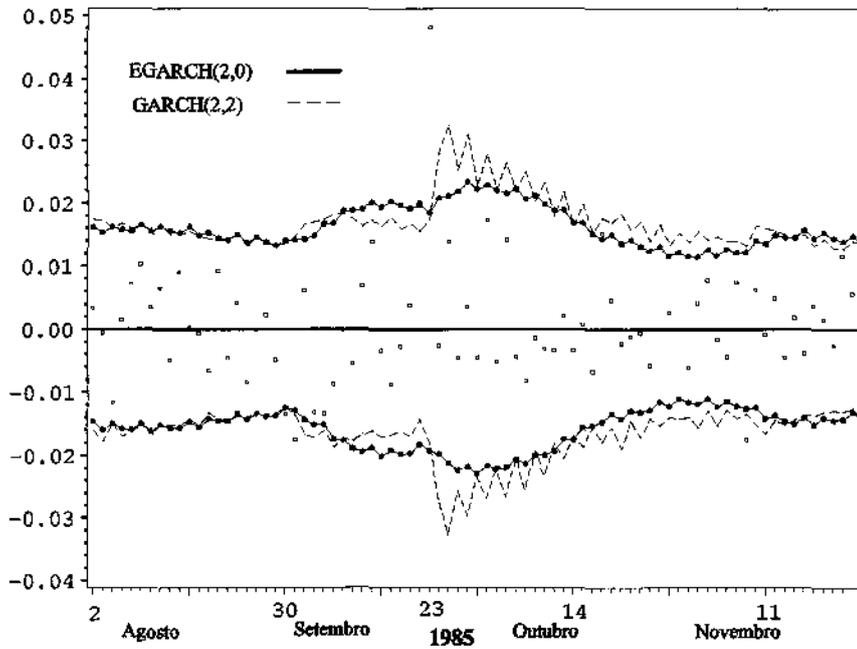


**Figura 3.14: Gráfico de Probabilidade Normal dos resíduos GARCH(2,2) na Taxa de Câmbio**

a) Volatilidades estimadas

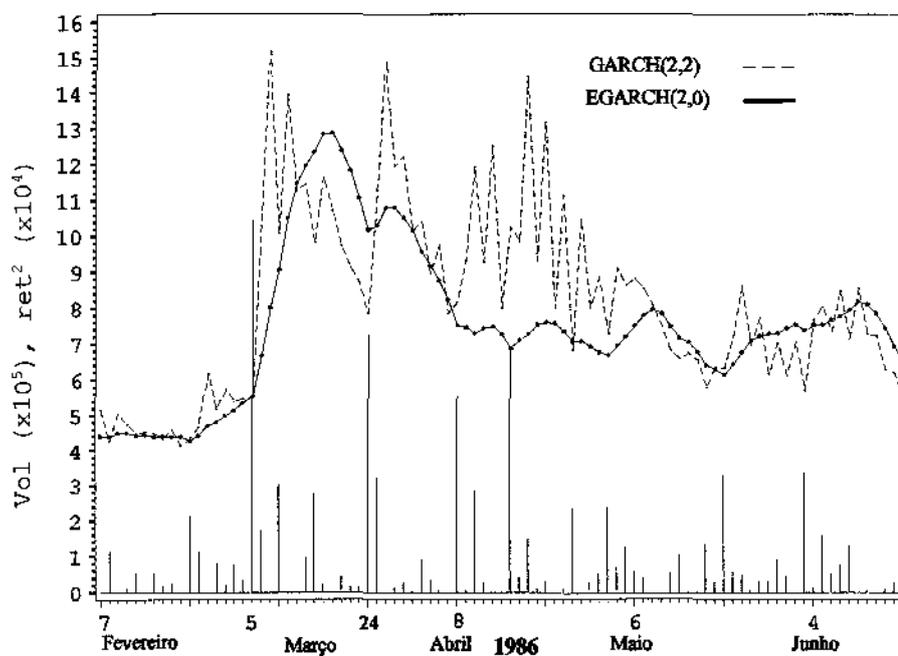


b) i.c 95%

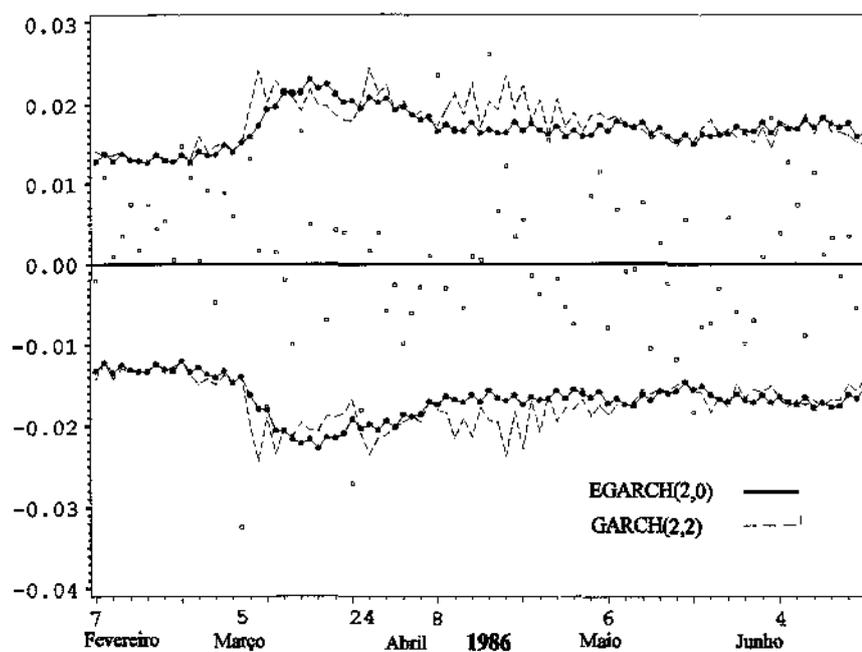


**Figura 3.15: Volatilidades estimadas ARCH e i.c 95% para os retornos no período da ocorrência do maior valor do retorno na Taxa de Câmbio**

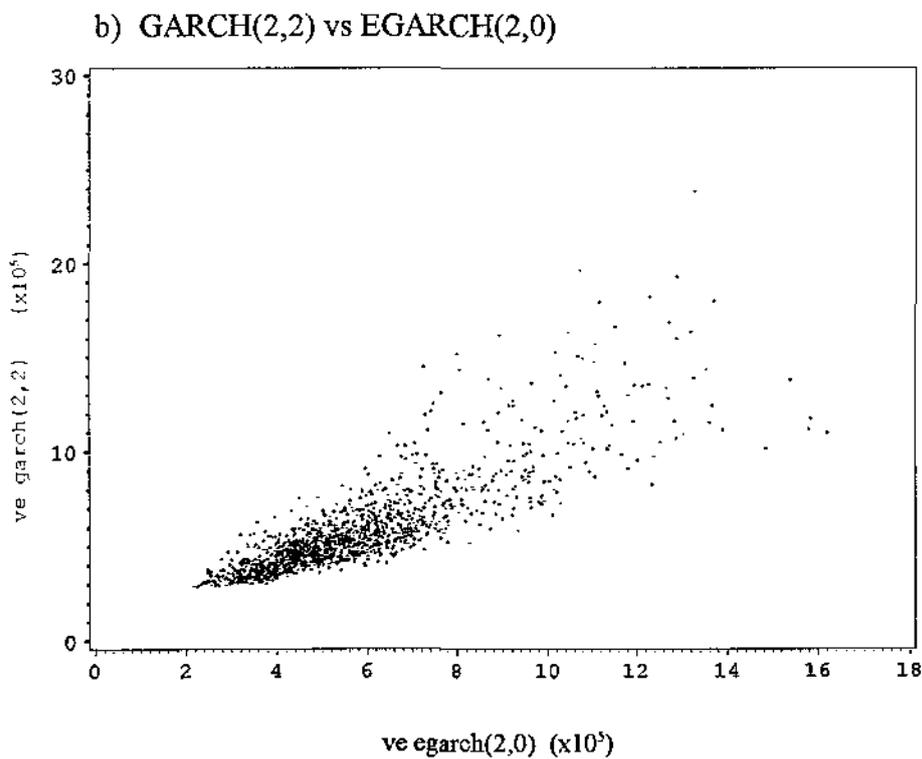
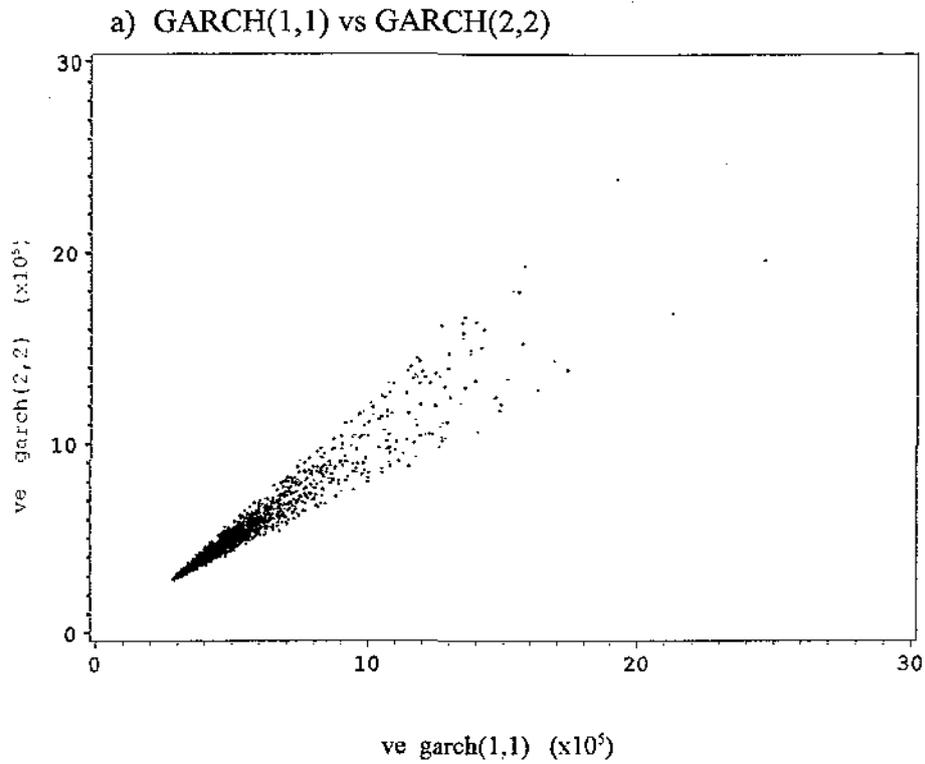
a) Volatilidades estimadas



b) i.c 95%



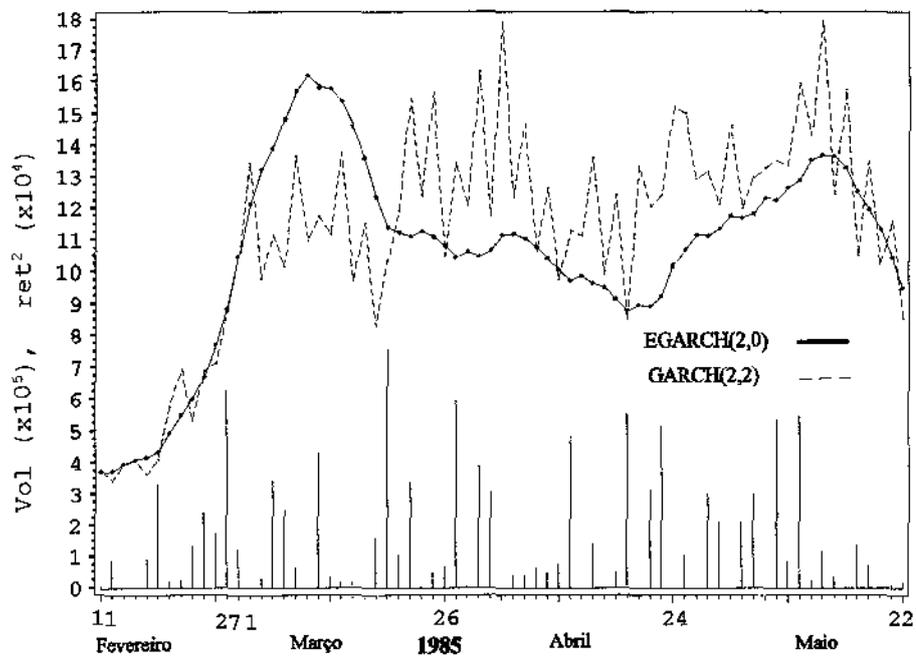
**Figura 3.16: Volatilidades estimadas ARCH e i.c. 95% para os retornos na Taxa de Câmbio no período de Fevereiro a Junho de 1986**



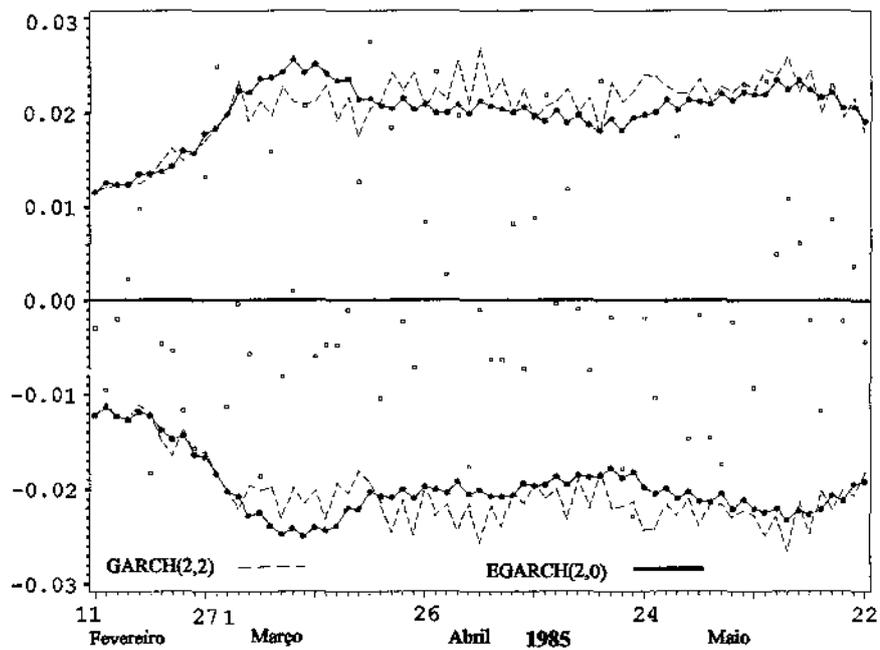
**Figura 3.17: Comparação das volatilidades estimadas GARCH(1,1), GARCH(2,2) e EGARCH(2,0) na Taxa de Câmbio**

**Legenda:** ve= volatilidade estimada

a) Volatilidades estimadas



b) i.c. 95%



**Figura 3.18: Volatilidades estimadas ARCH e i.c. 95% para os retornos da Taxa de Câmbio no período de Fevereiro a Maio de 1985**

# CAPÍTULO 4

## MODELOS DE VARIÂNCIA ESTOCÁSTICA

### 4.1 Introdução

Uma das principais características dos modelos ARCH, discutidos no Capítulo anterior, é o fato da volatilidade ser uma variável observável. Como consequência disso, a verossimilhança era obtida com facilidade o que, por sua vez, facilitava o processo de inferência. Este Capítulo está dedicado ao estudo de uma abordagem diferente, os *Modelos de Variância Estocástica (Modelos SV)* nos quais a volatilidade é uma variável *não observável*.

Entre os modelos que consideram a volatilidade como sendo não observável temos os propostos por Clark (1973), Tauchen e Pitts (1983), Hull e White (1987), Wiggins (1987), Melino e Turnbull (1990). Na formulação dos modelos *SV*, as propriedades assim como sua dinâmica eram facilmente derivadas. Entretanto as maiores dificuldades apareciam na estimação e no diagnóstico, por causa da ausência de procedimentos e ferramentas teóricas. Devido a isso, os modelos de variância estocástica eram pouco utilizados. Os progressos da última década relacionados com a análise de variáveis não observáveis, facilitaram o surgimento de diversas metodologias para calcular a volatilidade (ver Ghysels *et alii*, 1995 e Shephard, 1994b para uma revisão da bibliografia). Entre as abordagens mais promissoras está a utilização de procedimentos Bayesianos (Jacquier *et alii*, 1994), nos quais a distribuição da volatilidade é calculada fazendo uso de métodos computacionais intensivos como o *Amostrador de Gibbs* ou *Cadeias de Markov via Monte Carlo*. O principal problema nessas metodologias é a ausência de procedimentos de diagnóstico e, às vezes, o alto custo computacional.

Uma forma de tratar esses modelos, consiste em encontrar representações adequadas em espaço de estados. Nesse contexto, estimar a volatilidade constitui um problema de *filtragem* ou *extração de sinal*. O problema inerente a essa abordagem é a dificuldade em calcular a verossimilhança. Como uma alternativa para contorna essa dificuldade, Harvey *et alii* (1994) e Ruiz (1994) propõem o *Método de Quase Máxima Verossimilhança* que consiste em aproximar a verossimilhança pelos dois primeiros momentos e calcular a quase-verossimilhança

e as estimativas da volatilidade via Filtro de Kalman (Kalman, 1960). A representação em termos de espaço de estados permite estimar os resíduos recursivos, que podem ser utilizados para testes de diagnóstico do modelo e permite calcular a previsão de forma simples. Adicionalmente, permite a incorporação de diversos *Fatos Estilizados* como *Efeitos Calendário* e *Leverage Effect* de maneira simples. Finalmente, e quiçá seja a maior vantagem, a extensão para modelos multivariados é natural, sendo os modelos parcimoniosos (ver Harvey *et alii*, 1994). Apesar dessas vantagens, esse procedimento aproximado de estimação pode ser muito ineficiente.

Neste Capítulo será discutido um modelo de volatilidade estocástica específico; o modelo *Log-normal*, também conhecido como *AR(1)-SV*, que foi proposto pela primeira vez por Taylor (1980, 1986) e desenvolvido posteriormente por Nelson (1988) e Harvey *et alii* (1994), dentro de uma abordagem de espaço de estados. Inicialmente, na Seção 4.2 discutimos a representação em espaço de estados dos modelos de variância estocástica. A seguir, na Seção 4.3, é mostrado como o modelo *AR(1)-SV* é capaz de reproduzir teoricamente alguns *Fatos Estilizados*. A Seção 4.4 trata do processo de Estimação enquanto que, na Seção 4.5, discutimos algumas medidas para o diagnóstico e Bondade de Ajuste. Na Seção 4.6 são apresentadas algumas extensões do modelo *AR(1)-SV* para incorporar outros *Fatos Estilizados*. A previsão é o assunto da Seção 4.7 e, finalmente, na Seção 4.8 ilustramos a metodologia exposta através do ajuste das *Séries Exemplo*.

## 4.2 Abordagem em Espaço de Estados dos Modelos de Variância Estocástica.

Nesta seção será discutido, com algum detalhe, a construção de um modelo na representação de espaço de estados. Entre as referências para modelos em Espaço de Estados e Filtro de Kalman temos, Harvey (1989) e, para uma abordagem Bayesiana, West e Harrison (1989).

Nos modelos de Variância Estocástica a volatilidade é considerada como uma componente não observável que varia ao longo do tempo. Nesse sentido, dentro da especificação usual, é assumido que os retornos satisfazem:

$$Y_t / \sigma_t \approx N(0, \sigma_t^2), \quad (4.2.1)$$

ou, de outra forma, escrito como modelo produto:

$$Y_t = \varepsilon_t \sigma_t \quad \varepsilon_t \approx \text{NID}(0,1), \quad (4.2.2)$$

onde  $\varepsilon_t$  é independente do processo  $Y_t$ . Subsiste ainda o problema de modelar a variância condicional. Uma forma simples de modelar esse processo é através de um processo autoregressivo de ordem um e escrevê-lo como exponencial de um processo não observável, com o propósito de garantir que a volatilidade seja positiva:

$$\log(\sigma_t^2) = \delta + \gamma \log(\sigma_{t-1}^2) + \eta_t \quad \eta_t \approx \text{NID}(0, \sigma_\eta^2) \quad (4.2.3)$$

onde  $\delta, \gamma$  são parâmetros e  $|\gamma| < 1$ .

A formulação (4.2.2) e (4.2.3), denominado *Modelo AR(1)-SV Discreto Estacionário* conduz naturalmente a uma representação do modelo em Espaço de Estados. Com efeito, na equação (4.2.2), elevando ao quadrado cada termo e tomando o logaritmo obtemos a seguinte expressão:

$$\log(Y_t^2) = \log(\varepsilon_t^2) + \log(\sigma_t^2), \quad (4.2.4)$$

na qual, as variáveis  $\log(\varepsilon_t^2)$  seguem a distribuição log-normal com parâmetros (Abramovitz e Stegum, 1970):

$$\begin{aligned} E[\log(\varepsilon_t^2)] &= -1.27 \\ \text{Var}[\log(\varepsilon_t^2)] &= \pi^2/2 \cong 4.93 \end{aligned} \quad (4.2.5)$$

Por sua vez, podemos definir a variável:

$$\xi_t = \log(\varepsilon_t^2) - E[\log(\varepsilon_t^2)] = \log(\varepsilon_t^2) + 1.27, \quad (4.2.6)$$

e substituí-la em (4.2.4). Dessa forma obtemos a expressão:

$$\log(Y_t^2) = -1.27 + \log(\sigma_t^2) + \xi_t, \quad (4.2.7)$$

na qual  $E[\xi_t] = 0$ ,  $\text{Var}[\xi_t] = \pi^2/2$  e, dado que as variáveis aleatórias  $\varepsilon_t$  são mutuamente independentes, então as  $\xi_t$  também são.

Finalmente, as equações (4.2.3) e (4.2.7) constituem a representação em Espaço de Estados do modelo (4.2.2)-(4.2.3):

$$\log(Y_t^2) = -1.27 + \log(\sigma_t^2) + \xi_t, \quad \xi_t \approx \text{ID}(0, \pi^2 / 2) \quad (4.2.8a)$$

$$\log(\sigma_t^2) = \delta + \gamma \log(\sigma_{t-1}^2) + \eta_t, \quad \eta_t \approx \text{NID}(0, \sigma_\eta^2) \quad (4.2.8b)$$

$\xi_t, \eta_t$  independentes.

A inclusão de uma componente aleatória  $\eta_t$  na equação (4.2.8b), que envolve o processo da volatilidade, faz o modelo mais versátil no sentido que o torna suficientemente amplo, de forma a reproduzir os *Fatos Estilizados*<sup>1</sup>. Quando  $\gamma=1$ , definimos o *Modelo AR(1)-SV Discreto Não-Estacionário* como:

$$Y_t = \varepsilon_t \sigma_t, \quad \varepsilon_t \approx \text{NID}(0,1) \quad (4.2.9a)$$

$$\log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \eta_t, \quad \eta_t \approx \text{NID}(0, \sigma_\eta^2) \quad (4.2.9b)$$

$\xi_t, \eta_t$  independentes,

o qual possui uma representação em Espaço de Estados como (4.2.8) com  $\delta = 0$  e  $\gamma = 1$ , isto é, o logaritmo da variância se reduz a um passeio ao acaso sem *drift*.

O modelo (4.2.8) constitui uma aproximação discreta natural do modelo de difusão Orstein-Uhlenbeck utilizado, entre outros, por Hull e White (1987). Dassios (1992) mostrou que esse modelo constitui uma melhor aproximação discreta que o modelo EGARCH.

Na literatura existem várias metodologias para estimar a formulação (4.2.8), como por exemplo, entre outros, o Método dos Momentos Generalizado (GMM), a Moda, Amostragem por Importância (Importance Sampling), Cadeias de Markov via Monte Carlo, (para uma revisão atualizada ver Shephard, 1994b). A alternativa proposta por Harvey *et alii* (1994) e Ruiz (1994), consiste em assumir que a distribuição das perturbações  $\xi_t$  pode ser aproximada pela distribuição *Normal*. Dessa forma, (4.2.8) constitui a representação em Espaço de Estados com as perturbações consideradas *Gaussianas* e a verossimilhança calculada via Filtro de

<sup>1</sup>É possível generalizar essa equação de maneira que evolua como um ARMA.

Kalman. Como a verossimilhança é aproximada, temos uma estimativa pelo *Método de Quase-Maxima Verossimilhança*. Depois de estimados os parâmetros, o Filtro de Kalman pode ser utilizado para estimar a volatilidade. Uma característica importante deste método é que não precisamos conhecer a distribuição específica de  $\varepsilon_t$ , mas apenas os dois primeiros momentos.

### 4.3 Propriedades Básicas: Reprodução dos *Fatos Estilizados*

Na demonstração das propriedades básicas a seguir, a hipótese de normalidade nas perturbações é utilizada só quando necessária. Para facilitar o cálculo posterior da verossimilhança, substituímos o parâmetro de escala  $\delta$  pela expressão  $\log(\kappa)(1-\gamma)$ , onde a constante  $\kappa$  está fora da equação de estados. Dessa forma, o modelo AR(1)-SV discreto *Estacionário* é definido como:

$$Y_t = \varepsilon_t \sigma_t \tag{4.3.1a}$$

$$\sigma_t^2 = \kappa \exp(h_t) \tag{4.3.1b}$$

$$h_t = \gamma h_{t-1} + \eta_t \tag{4.3.1c}$$

onde  $|\gamma| < 1$ , e as perturbações  $\varepsilon_t$  e  $\eta_t$  satisfazem:

$$\varepsilon_t \sim \text{IID}(0, 1) \tag{4.3.1d}$$

$$\eta_t \sim \text{IID}(0, \sigma_\eta^2) \tag{4.3.1e}$$

$$E[\varepsilon_t, \eta_{t-s}] = 0 \quad \forall s \geq 0 \tag{4.3.1f}$$

Da mesma forma, a versão *Não-estacionária* é obtida substituindo  $\kappa = 1$  e  $\gamma = 1$ , em (4.3.1). As propriedades estatísticas do modelo de volatilidade estocástica definido em (4.3.1), serão apresentadas na forma de Teoremas, sendo o primeiro:

#### *Teorema 1*

No modelo (4.3.1), supondo que a variância da perturbação  $\eta_t$  é finita temos que:

i) Quando  $|\gamma| < 1$ , o processo  $\{h_t\}$  é estritamente estacionário, com média igual a:

$$\mu_h = 0, \text{ e variância } \sigma_h^2 = \frac{\sigma_\eta^2}{1-\gamma^2}.$$

ii) O processo  $\{Y_t\}$  é estritamente (fracamente) estacionário se o processo  $\{h_t\}$  é estritamente (fracamente) estacionário.

iii) O processo  $\{Y_t\}$  é martingale diferença com respeito a  $D_{t-1}$ ; a  $\sigma$ -álgebra gerada por  $I_t$ .

iv) O processo  $\{Y_t\}$  é ruído branco.

v) Se a distribuição do processo  $\{\varepsilon_t\}$  é simétrica, então a distribuição do processo  $\{Y_t\}$  é simétrica, portanto, seus momentos de ordem ímpar são nulos.

vi) Quando a distribuição do processo  $\{\varepsilon_t\}$  é normal com variância unitária, e a distribuição da perturbação  $\eta_t$  também é normal, então, a distribuição do processo  $\{Y_t\}$  apresenta cauda pesada, com variância e curtose dadas por :

$$\text{Var}[Y_t] = \kappa \exp(0.5\sigma_h^2), \tag{4.3.2}$$

$$\text{Curtose}[Y_t] = 3 \exp(\sigma_h^2) = 3 \exp\left(\frac{\sigma_\eta^2}{1-\gamma^2}\right), \tag{4.3.3}$$

se  $|\gamma| < 1$ . De maneira geral, os momentos de ordem par são obtidos através da expressão :

$$E[Y_t^{2m}] = \frac{(2m)!}{(2)^m m!} \exp(0.5m^2\sigma_h^2) \kappa^m. \tag{4.3.4}$$

*Prova:*

Para as provas são necessários resultados dos modelos autoregressivos, e da esperança condicional.

i) Resulta das propriedades dos modelos autoregressivos.

ii) Segue de (4.3.1a), (4.3.1b) e considerando que o produto de dois processos fracamente (estritamente) estacionários, é fracamente (estritamente) estacionário.

iii) Basta mostrar que  $E[Y_t / D_{t-1}] = 0$  e que  $E[|Y_t|]$  é finita.

a) Como  $\sigma_t$  é função de  $D_{t-1}$ , e as variáveis  $\varepsilon_t$  são mutuamente independentes, então:

$$E[Y_t / D_{t-1}] = E[\varepsilon_t \sigma_t / D_{t-1}] = \sigma_t E[\varepsilon_t / D_{t-1}] = 0, \text{ portanto } E[Y_t / D_{t-1}] = 0.$$

b) As perturbações  $\varepsilon_t$  são mutuamente independentes e possuem variância finita, em consequência  $E[|\varepsilon_t| / D_{t-1}] = E[|\varepsilon_t|]$  é finita. Por sua vez, considerando que  $\sigma_t^2$  é finito, então  $E[|\sigma_t|] < \infty$ . Usando ambos resultados obtemos que:

$$E[|Y_t|] = E[|\varepsilon_t \sigma_t|] = E[E[|\varepsilon_t \sigma_t| / D_{t-1}]] = E[|\sigma_t| E[|\varepsilon_t| / D_{t-1}]] \text{ é também finito.}$$

iv) Do item a) em iii) obtemos que  $E[Y_t] = E[E[Y_t / D_{t-1}]] = 0$ . Por outro lado, como as variáveis  $\varepsilon_t$  são mutuamente independentes, então para  $s > 0$ :  $E[\varepsilon_t \varepsilon_{t-s}] = 0$ . Em consequência,  $E[Y_t Y_{t-s}] = E[\varepsilon_t \sigma_t \varepsilon_{t-s} \sigma_{t-s}] = E[\sigma_t \sigma_{t-s} E[\varepsilon_t \varepsilon_{t-s} / D_{t-1}]] = 0$ . No item vi) será mostrado que o processo  $\{Y_t\}$  possui variância finita, com isso esse processo é ruído branco.

v) Como as perturbações  $\varepsilon_t$  são mutuamente independentes com distribuição simétrica, então  $E[\varepsilon_t^{2m+1} / D_{t-1}] = E[\varepsilon_t^{2m+1}] = 0$ , para  $m=0,1,2,\dots$ . Portanto:  $E[Y_t^{2m+1}] = E[\varepsilon_t^{2m+1} \sigma_t^{2m+1}] = E[\sigma_t^{2m+1} E[\varepsilon_t^{2m+1} / D_{t-1}]] = 0$ .

vi) Pelas propriedades da distribuição normal, e o fato da independência das perturbações  $\varepsilon_t$ , obtemos que:  $E[\varepsilon_t^{2m} / D_{t-1}] = E[\varepsilon_t^{2m}] = \frac{(2m)!}{(2)^m m!}$ . Por sua vez, a perturbação  $\eta_t$  possui distribuição normal, portanto o processo  $h_t = \log(\sigma_t^2 / \kappa)$  também tem distribuição normal com media e variância dadas em i). Dessa forma, é possível calcular os momentos de ordem par de  $\sigma_t^2$ , através da seguinte propriedade usada em Taylor (1986 p.74):

*Propriedade 1:*

Se  $\log(V_t)$  tem distribuição normal com media  $\alpha$  e variância  $\beta^2$  ( $\beta > 0$ ), então :

$$i) E[V_t] = \exp(\alpha + 0.5\beta^2), \quad (4.3.5.1)$$

$$ii) \text{ Como } \log[V_t^r] = r \log[V_t] \text{ em consequência, } E[V_t^r] = \exp(r\alpha + 0.5 r^2\beta^2). \quad (4.3.5.2)$$

Portanto, considerando  $V_t = \sigma_t^2 / \kappa$  obtemos que:  $E[(\sigma_t^2)^m / \kappa^m] = \exp(\frac{1}{2} m^2 \sigma_h^2)$  e  $E[(\sigma_t^2)^m] = \kappa^m \exp(\frac{1}{2} m^2 \sigma_h^2)$ . Logo, fazendo uso de  $E[Y_t^{2m}] = E[\varepsilon_t^{2m} \sigma_t^{2m}] = E[\sigma_t^{2m} E[\varepsilon_t^{2m} / D_{t-1}]]$ , obtemos a seguinte expressão:

$$E[Y_t^{2m}] = \frac{(2m)!}{(2)^m m!} \exp(0.5m^2 \sigma_h^2) \kappa^m,$$

em particular,  $E[Y_t^2] = \exp(0.5\sigma_h^2)\kappa$ , que é igual à variância de  $Y_t$ , e  $E[Y_t^4] = 3\exp(2\sigma_h^2)\kappa^2$ , em conseqüência a curtose é calculada facilmente.

□

Como acontece nos modelos ARCH, o processo dos retornos do modelo AR(1)-SV é não correlacionado, com distribuição simétrica em torno de zero. Se  $\sigma_\eta^2 > 0$ , do item vi), obtemos que a curtose é maior do que 3, em conseqüência, a distribuição dos retornos também é leptocúrtica. Entretanto, uma diferença importante é que no modelo (4.3.1), o quarto momento sempre existe quando o processo  $h_t$  é estacionário (mesmo quando  $\gamma$  é nulo<sup>2</sup>), ao contrário dos modelos ARCH, nos quais, a existência do quarto momento depende dos valores das raízes na equação da volatilidade (por exemplo no GARCH(1,1), quando  $\alpha+\beta<1$ ). A propriedade de cauda pesada vale também quando especificamos distribuições GED nas perturbações, ver Harvey e Shephard (1993b).

Com relação à Persistência, dado que a log-volatilidade é um processo autoregressivo linear (4.2.8b), o indicador natural para a persistência é o parâmetro  $\gamma$ .

Por outro lado, fazendo suposições sobre a distribuição conjunta das perturbações  $\varepsilon_t$  e  $\eta_t$ , é possível modificar o modelo (4.3.1) de forma a reproduzir outro *Fato Estilizado*: o *Leverage Effect*, essa modificação pode ser encontrada em Harvey e Shephard (1993c). Analogamente, incorporando o processo dos retornos como um processo de inovações dentro de um modelo de regressão, podemos imitar os *Efeitos Calendário*. Esse aspecto será discutido na Seção 4.7.

Na exposição das características das séries financeiras no Capítulo 2, encontramos que um dos *Fatos Estilizados* mais importantes é a não-linearidade dos retornos. Uma forma de confirmar essa característica é através da f.a.c de funções não-lineares dos retornos. No modelo AR(1)-SV estudado nesta dissertação, o interesse está centrado nos quadrados dos retornos e nos logaritmos dos quadrados dos retornos. O seguinte Teorema explicita sua dinâmica.

<sup>2</sup> Pois  $Y_t$  é mistura de distribuições, no qual o grau de mistura depende de  $\sigma_\eta^2$ .

*Teorema 2*

Seja o modelo de variância estocástica definido em (4.3.1). Supondo que as perturbações  $\epsilon_t$  e  $\eta_t$  são normais mutuamente independentes, então:

i) A f.a.c do processo dos quadrados dos retornos é igual a:

$$\rho_{Y_t^2}(s) = \frac{\exp[\sigma_h^2 \gamma^s] - 1}{3 \exp[\sigma_h^2] - 1} \quad (4.3.6)$$

ii) Entretanto, quando o parâmetro  $\gamma$  é próximo de um e ou  $\sigma_h^2$  é pequeno, podemos aproximar a f.a.c obtida em i) por:

$$\rho_{Y_t^2}(s) \cong \frac{\exp[\sigma_h^2] - 1}{3 \exp[\sigma_h^2] - 1} \gamma^s \quad (4.3.7)$$

iii) Finalmente, a f.a.c dos logaritmos dos quadrados dos retornos é:

$$\rho_{\log(Y_t^2)}(s) = \frac{\gamma^s}{1 + 4.93 / \sigma_h^2} \quad (4.3.8)$$

*Prova:*

i) Como as variáveis  $\epsilon_t$  são normais *i.i.d*, então:  $E[\epsilon_t^2 \epsilon_{t-s}^2 / D_{t-1}] = E[\epsilon_t^2 \epsilon_{t-s}^2] = 1$ . Dessa forma:

$$E[Y_t^2 Y_{t-s}^2] = E[\sigma_t^2 \sigma_{t-s}^2 E[\epsilon_t^2 \epsilon_{t-s}^2 / D_{t-1}]] = E[\sigma_t^2 \sigma_{t-s}^2] = E[\exp(h_t + h_{t-s})] \kappa^2 \quad (4.3.9)$$

Por sua vez, como  $h_t$  é um processo autoregressivo normal de ordem 1, então o processo:  $\log[\sigma_t^2 \sigma_{t-s}^2 / \kappa^2] = h_t + h_{t-s}$ , também tem distribuição normal com media igual a 0 e variância igual a  $2\sigma_h^2(1 + \gamma^s)$ . Logo, pela expressão (4.3.5.1) na Propriedade 1 obtemos:

$$E[\exp(h_t + h_{t-s})] = \exp[\sigma_h^2(1 + \gamma^s)], \quad (4.3.10)$$

e como por outro lado :

$$\text{Var}[Y_t] = E[Y_t^2] = \exp[0.5\sigma_h^2] \kappa \quad (4.3.11)$$

portanto, de (4.3.9), (4.3.10) e (4.3.11) obtemos:

$$\begin{aligned} \text{Cov}[Y_t^2 Y_{t-s}^2] &= \exp[\sigma_h^2(1 + \gamma^s)]\kappa^2 - (\exp[0.5\sigma_h^2]\kappa)^2 \\ &= \exp[\sigma_h^2]\{\exp[\sigma_h^2\gamma^s] - 1\}\kappa^2, \end{aligned} \quad (4.3.12)$$

$$\begin{aligned} \text{Var}[Y_t^2] &= E[Y_t^4] - (E[Y_t^2])^2 = 3\exp[2\sigma_h^2]\kappa^2 - (\exp[0.5\sigma_h^2]\kappa)^2 \\ &= \kappa^2 \exp[\sigma_h^2]\{3\exp[\sigma_h^2] - 1\}. \end{aligned} \quad (4.3.13)$$

Finalmente, como  $\text{Var}[Y_t^2] = \text{Var}[Y_{t-s}^2]$ , utilizando as expressões (4.12) e (4.13) é fácil obter:

$$\rho_{Y_t^2} = \text{Corr}[Y_t^2 Y_{t-s}^2] = \frac{\exp[\sigma_h^2\gamma^s] - 1}{3\exp[\sigma_h^2] - 1}.$$

ii) Expandindo  $\exp[\sigma_h^2\gamma^s]$  em série de McLaurin obtemos:  $\exp[\sigma_h^2\gamma^s] = \sum_{k=0}^{\infty} \frac{(\sigma_h^2\gamma^s)^k}{k!}$  e considerando que  $\{\gamma^s\}^k = 1$ , quando  $\gamma \cong 1$ , então:

$$\exp[\sigma_h^2\gamma^s] \cong 1 + \gamma^s \sum_{k=1}^{\infty} \frac{(\sigma_h^2)^k}{k!} = 1 + \gamma^s \{\exp[\sigma_h^2] - 1\},$$

logo, substituindo a última expressão em (4.3.6) obtemos:

$$\rho_{Y_t^2}(s) \cong \frac{\exp[\sigma_h^2] - 1}{3\exp[\sigma_h^2] - 1} \gamma^s.$$

iii) Em (4.2.4) temos que:  $\log(Y_t^2) = h_t + \log(\epsilon_t^2)$ . Por sua vez, como as variáveis  $\log(\epsilon_t^2)$  são mutuamente independentes e o processo  $\{h_t\}$  é AR(1), então o processo  $Z_t = \log(Y_t^2)$  é um ARMA (1,1); o qual possui uma função de autocorrelação da forma:

$$\rho_{Z_t}(s) = C\gamma^{s-1} \quad s = 1, 2, 3, \dots$$

onde C é o valor da primeira autocorrelação, que é calculada pela definição. Assim, considerando que os processos  $\{h_t\}$  e  $\{\log(\epsilon_{t-s}^2)\}$  são independentes ( $s \geq 0$ ), calculamos facilmente as seguintes quantidades:

$$E[Z_t] = E[h_t + \log(\varepsilon_t^2)] = E[\log(\varepsilon_t^2)],$$

$$\begin{aligned} E[Z_t Z_{t-1}] &= E[\log(Y_t^2) \log(Y_{t-1}^2)] = E[(h_t + \log(\varepsilon_t^2))(h_{t-1} + \log(\varepsilon_{t-1}^2))] \\ &= E[h_t h_{t-1}] + E[\log(\varepsilon_t^2)] E[\log(\varepsilon_{t-1}^2)], \end{aligned}$$

$$\begin{aligned} E[Z_t^2] &= E[\{(h_t + \log(\varepsilon_t^2))\}^2] = E[h_t^2] + E[\{\log(\varepsilon_{t-1}^2)\}^2] \\ &= E[h_t^2] + \{\text{Var}[\log(\varepsilon_t^2)] + (E[\log(\varepsilon_t^2)])^2\}, \end{aligned}$$

dessa forma obtemos:

$$\text{Var}[Z_t] = E[Z_t^2] - \{E[Z_t]\}^2 = E[h_t^2] + \text{Var}[\log(\varepsilon_t^2)] = \sigma_h^2 + 4.93,$$

$$\begin{aligned} \text{Cov}(Z_t, Z_{t-1}) &= E[Z_t Z_{t-1}] - E[Z_t] E[Z_{t-1}] = E[h_t h_{t-1}] \\ &= E[(\gamma h_{t-1} + \eta_t) h_{t-1}] = \gamma E[h_{t-1}^2] = \gamma \sigma_h^2, \end{aligned}$$

e, pela definição de correlação obtemos finalmente:

$$C = \frac{\gamma}{1 + 4.93/\sigma_h^2}.$$

□

Um aspecto interessante do resultado (4.3.7) é que essa função de autocorrelação é similar à obtida no caso de um modelo ARMA(1,1), a qual, como já vimos no Capítulo 2, é capaz de imitar a dinâmica que se observa frequentemente nas séries financeiras. Nas aplicações empíricas com modelos AR(1)-SV, usualmente encontramos estimativas de  $\gamma$  que permitem utilizar a expressão (4.3.7), como uma aproximação razoável. Da mesma forma, como visto na Seção 3.3.2.1 a f.a.c dos quadrados dos retornos do modelo GARCH(1,1) também é da forma ARMA(1,1). O fato dos modelos GARCH(1,1) e AR(1)-SV apresentarem a mesma f.a.c nos quadrados dos retornos será utilizado na Seção 5.4.1, para comparar os dois modelos.

Do exposto, vemos que modelo AR(1)-SV teoricamente é capaz de reproduzir os principais *Fatos Estilizados* inerentes à distribuição do processo dos retornos e da volatilidade. A seguir mostramos como pode ser realizada a estimação.

## 4.4 Estimação

Dentro da abordagem de Quase-Máxima Verossimilhança (QML), a estimação do modelo AR(1)-SV é simplificada pela utilização do Filtro de Kalman, que permite o cálculo da verossimilhança através da decomposição do erro de previsão. Uma vez obtidas as estimativas dos parâmetros através da maximização da verossimilhança, recalculamos os componentes via Filtro de Kalman. Na literatura, a verossimilhança é geralmente descrita como bem comportada não apresentando problemas numéricos, o que constitui uma diferença importante quando comparamos com a estimação nos modelos ARCH, nos quais existem dificuldades para maximizar com restrições, devido aos problemas de convergência, ou, como comentam Jacquier *et alii* (1994 p. 414) devido à presença de vários máximos locais. Entretanto, na aplicação com as *Séries Exemplo* encontramos problemas de convergência.

Nesta Seção será discutida a estimação do modelo AR(1)-SV dentro da abordagem QML. A seguir apresentamos uma breve discussão da eficiência desse método quando comparado com outros existentes na literatura.

### 4.4.1 Estimação Pontual

O procedimento de estimação será apresentado primeiro para o modelo AR(1)-SV estacionário, no qual o parâmetro do nível é estimado seguindo a abordagem proposta por Ruiz (1994)<sup>3</sup>.

Seja a representação em Espaço de Estados do modelo AR(1)-SV Estacionário:

$$X_t = h_t + \xi_t \quad \xi_t \approx \text{NID}(0, \pi^2 / 2) \quad (4.4.1.1a)$$

$$h_t = \gamma h_{t-1} + \eta_t \quad \eta_t \approx \text{NID}(0, \sigma_\eta^2) \quad (4.4.1.1b)$$

$$\xi_t, \eta_t \text{ não correlacionados,}$$

onde :

$$X_t = \log(Y_t^2) + 1.27 - \log(\kappa), \quad (4.4.1.1c)$$

e a volatilidade definida como:

<sup>3</sup>Em Ghysels *et alii* (1995, p. 13) discute-se uma alternativa não-paramétrica.

$$\sigma_t^2 = \kappa \exp(h_t). \tag{4.4.1.1d}$$

Uma reparametrização conveniente da representação (4.4.1.1), que facilita posteriormente o cálculo da quase-verossimilhança consiste em fazer  $\sigma_\eta^2 = \sigma_\xi^2 = \pi^2 / 2 = 4,93$  e  $q = \sigma_\eta^2 / \sigma_\xi^2$ . Dessa forma, os parâmetros a serem estimados são:  $\Psi = \{\kappa, \gamma, q\}$  e por sua vez, temos que estimar os componentes:  $\{h_t, \xi_t\}$ . A seguir apresentamos os passos para a estimação do modelo (4.4.1.1).

1) *Estimação consistente de  $\kappa$  (Ruiz, 1994).*

Aplicando o operador esperança em ambos termos de (4.4.1.1c) obtemos  $\log(\kappa) = 1.27 + E[\log(Y_t^2)]$  dado que  $E[X_t]$  é zero de (4.4.1.1a). Seguindo os resultados de Harvey (1989), estimamos  $E[\log(Y_t^2)]$  pela media amostral<sup>4</sup>, que é um estimador consistente. Logo,  $\log(\kappa)$  é estimado consistentemente por  $\log(\kappa) = 1.27 + \overline{\log(Y_t^2)}$ . Ao substituir essa expressão em (4.4.1.1c) obtemos que  $X_t = \log(Y_t^2) - \overline{\log(Y_t^2)}$ . No desenvolvimento posterior desta seção será assumido que os valores de  $X_t$  são conhecidos e, em consequência, só temos que estimar os parâmetros  $\Psi = \{\gamma, q\}$ .

2) *Filtro de Kalman*

O Filtro de Kalman é um algoritmo recursivo (em  $t=1,2,\dots,T$ ) dado pelas seguintes equações:

Equações de Previsão:  $m_{t|t-1} = \gamma m_{t-1} \tag{4.4.1.2a}$

$$P_{t|t-1} = \gamma^2 P_{t-1} + q \tag{4.4.1.2b}$$

Equações de Atualização:  $f_t = P_{t|t-1} + 1 \tag{4.4.1.2c}$

$$P_t = P_{t|t-1} - (1/f_t) P_{t|t-1}^2 \tag{4.4.1.2d}$$

$$m_t = m_{t|t-1} + k_t (X_t - m_{t|t-1}) \tag{4.4.1.2e}$$

<sup>4</sup>Esse estimador é QML e não correlacionado com os estimadores QML de  $\gamma$  e  $\sigma_\eta^2$

onde  $m_t$  ( $m_{t|t-1}$ ) é a estimativa de  $h_t$  ( $h_{t|t-1}$ ) baseada na informação  $I_t$ , ( $I_{t-1}$ ),  $P_t$  ( $P_{t|t-1}$ ) é a variância de  $m_t$  ( $m_{t|t-1}$ ),  $k_t = P_{t|t-1} / f_t$  é chamado Ganho de Kalman e  $f_t$  é a variância do erro da previsão um passo à frente  $v_t = X_t - m_{t|t-1}$ .

Para inicializar o filtro é necessário estabelecer as condições iniciais para o algoritmo:  $P_{1/0}$  e  $m_{1/0}$ . Uma opção é considerar os valores não condicionais da média e variância do processo, i.e.,  $P_{1/0} = \frac{\sigma_\eta^2}{1 - \gamma^2}$  e  $m_{1/0} = 0$ . Outra opção é considerar  $P_1 = 1$  e  $m_1 = Y_1$  que equivale a supor variância infinita em  $P_{1/0}$  (escolha equivalente a uma distribuição imprópria na abordagem Bayesiana).

### 3) Maximização da quase-verossimilhança

Na abordagem QML, a suposição básica é que as perturbações  $\xi_t$  em (4.4.1.1a) possuem distribuição normal, em conseqüência, utilizando os valores obtidos em 2) para  $v_t$  e  $f_t^* = \sigma_\eta^2 f_t$  é possível calcular a (log) quase-verossimilhança através da decomposição do erro de previsão. Essa expressão é:

$$\text{LogL}(\Psi) = \frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma_\eta^2) - \frac{1}{2} \sum_{t=1}^T \log(f_t) - \frac{1}{2\sigma_\eta^2} \sum_{t=1}^T \frac{v_t^2}{f_t} \quad (4.4.1.3a)$$

Já que  $\sigma_\eta^2 = 4,93$ , maximizar (4.4.1.3) equivale a minimizar:

$$\text{LogL}^*(\Psi) = \sum_{t=1}^T \log(f_t) + \frac{1}{\sigma_\eta^2} \sum_{t=1}^T \frac{v_t^2}{f_t} \quad (4.4.1.3b)$$

na qual  $f_t$  e  $v_t$  são funções dos parâmetros  $\Psi = \{\gamma, q\}$ . Seja qual for o método de estimação utilizado, *BHHH* ou *Score*, por exemplo, é necessário calcular o gradiente e, como resultado do processo de estimação obtemos  $\hat{\Psi} = \{\hat{\gamma}, \hat{q}\}$ . Valores iniciais para os parâmetros:  $\Psi_0 = \{\gamma_0, q_0\}$  são obtidos, por exemplo, através do *Algoritmo EM* (Dempster *et alii*, 1977), ou pelo método proposto por De Jong (1991).

#### 4) Estimação dos componentes

Estimativas dos componentes  $\{h_t, \xi_t\}$ , são obtidas através do Filtro de Kalman apresentado em 2) utilizando os valores de  $\hat{\Psi} = \{\hat{\gamma}, \hat{q}\}$ . Existem dois tipos de estimativas, *filtradas* e *suavizadas*. As estimativas *filtradas* podem ser calculadas fazendo uso da informação até o instante  $t$ ,  $I_t = \{Y_t, Y_{t-1}, \dots, Y_1\}$ , ou baseadas apenas na informação passada  $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \dots, Y_1\}$ . Essas são denotadas, respectivamente por  $\{\hat{h}_t, \hat{\xi}_t\}$  e  $\{\hat{h}_{t/t-1}, \hat{\xi}_{t/t-1}\}$ . Por outro lado, os estimadores que incorporam toda a informação disponível são conhecidos como *suavizados* por serem obtidos através de processos de suavização. Existem vários algoritmos de suavização, sendo o mais simples o resultante de inverter o processo do Filtro de Kalman começando pelo instante  $t=T$  até  $t=1$ . Um algoritmo mais eficiente é o proposto por Koopman (1993).

Os componentes estimados pelo Filtro de Kalman são ótimos no sentido de serem não-tendenciosos e apresentar variância mínima dentro da classe dos estimadores lineares.

#### 5) Estimadores da Volatilidade

Considerando a natureza das estimativas, como sendo filtradas ou suavizadas, distinguimos dois tipos de estimativas de volatilidade. Assim, a estimativa da *volatilidade filtrada* é definida como:

$$\hat{\sigma}_{f,t}^2 = \hat{\kappa} \exp(\hat{h}_t) \quad \text{baseada na informação } I_t \quad (4.4.1.4a)$$

$$\hat{\sigma}_{f,t/t-1}^2 = \hat{\kappa} \exp(\hat{h}_{t/t-1}) \quad \text{baseada na informação } I_{t-1} \quad (4.4.1.4b)$$

e, baseada na estimativa suavizada,  $\{\hat{h}_{t/T}, \hat{\xi}_{t/T}\}$ , definimos a *volatilidade suavizada* como:

$$\hat{\sigma}_{s,t}^2 = \hat{\kappa} \exp(\hat{h}_{t/T}). \quad (4.4.1.4c)$$

Por outro lado, no caso do modelo AR(1)-SV não-estacionário, a representação de espaço de estados é a mesma exceto pelo fato de que  $\kappa$  e  $\gamma$  são iguais a 1. Desta forma o espaço paramétrico reduz-se a  $\Psi = \{q\}$  e só temos que estimar as componentes  $\{h_t, \xi_t\}$ .

Nas aplicações, se o interesse é *entender o comportamento da volatilidade* no período de estudo, o estimador utilizado é o suavizado por dois motivos. Primeiro, porque incorpora toda a informação disponível na amostra e, em segundo lugar, porque é menos sensível aos possíveis *valores aberrantes (outliers)*. Já a volatilidade filtrada baseada na informação passada  $I_{t-1}$ , é apropriada para fazer *previsão* e na análise da *adequabilidade do modelo*. Ambos aspectos são tratados nas Seções 4.5 e 4.6.

Por outro lado, na prática, é possível que os retornos sejam nulos em algum instante de tempo (por exemplo quando os preços não mudam de um dia para o outro), invalidando a transformação logarítmica em (4.2.4). Duas formas são sugeridas com mais frequência na literatura. A primeira consiste em trabalhar com as observações ajustadas pela média amostral<sup>5</sup>, enquanto que o segundo método, sugerido por Fuller e citado em Ghysels *et alii* (1995), consiste em trabalhar com a seguinte transformação:

$$\log(\tilde{Y}_t^2) = \log(Y_t^2 + \sigma_{y^2}^2) - \frac{c\sigma_{y^2}^2}{Y_t^2 + c\sigma_{y^2}^2},$$

onde  $c$  é uma constante adequada (por exemplo 0,02) e  $\sigma_{y^2}^2$  é a estimativa da variância não condicional dos quadrados dos retornos. Observe-se que se  $c=0$  então  $\log(\tilde{Y}_t^2) = \log(Y_t^2)$ .

#### 4.4.2 Incerteza na Estimação dos Parâmetros e na Estimação da Volatilidade.

Expressões para a distribuição assintótica dos estimadores pelo método QML foram derivadas por Ruiz (1994) seguindo a teoria padrão de estimação de componentes não observáveis e baseada nos resultados de Dunsmuir (1979). Tanto para o caso do modelo AR(1)-SV estacionário quanto para o modelo não estacionário, esses resultados são aplicáveis, de forma que é possível mostrar que:

$$\sqrt{T}(\hat{\Psi} - \Psi) \approx N(0, C(\Psi)). \tag{4.4.2.1}$$

As expressões de  $C(\Psi)$  para os casos estacionário e não-estacionário podem encontrarse em Ruiz (1994).

<sup>5</sup>Nelson apresenta sérias críticas ao uso desses mecanismos de Trim, ver Jacquier *et alii* (1994, p.403-406).

Na literatura não encontramos expressões para a distribuição das estimativas de volatilidade  $\hat{\sigma}_{f,t}^2 = \hat{\kappa} \exp(\hat{h}_t)$ . Resultados assintóticos analíticos em modelos de tempo contínuo derivados por Nelson e Foster (1994), mostram que o estimador da volatilidade, obtido através da abordagem QML, é muito ineficiente e é possível que o mesmo ocorra no caso da aproximação discreta. Nesse sentido, uma alternativa é utilizar outros métodos de estimação como MCMC, ver Jacquier *et alii* (1994).

#### 4.4.3 Eficiência Comparativa dos Estimadores QML

Nesta subseção são discutidos dois assuntos. O primeiro, refere-se à influência da aproximação normal para a distribuição do processo  $\log(\varepsilon_t^2)$  (suposição básica) e, o segundo, à apresentação de alguns resultados sobre a eficiência do método QML, quando comparado com o *Método dos Momentos Generalizado* (GMM) e as *Cadeias de Markov via Monte Carlo* (MCMC).

##### *Eficiência dos Estimadores QML*

Em Ruiz (1994) foram realizados experimentos de simulação para medir a eficiência dos estimadores QML utilizando o critério do Erro Padrão Assintótico (Asymptotic Standard Error):  $ASES = \sqrt{C(\Psi) / T}$  e considerando três tamanhos de amostra: 500 (amostra pequena) 3000 e 6000 (amostras razoáveis).

No caso estacionário, o conjunto de parâmetros considerado nas simulações foi  $(\gamma, \sigma_{\eta}^2) = \{(0,7;0,09) (0,7;1) (0,9;0,09) (0,9;1)\}$ . Os resultados sugerem que em geral as estimativas QML são boas quando as amostras são razoáveis. Quando a amostra é pequena, tanto o vício de  $\hat{\sigma}_{\eta}^2$ , quanto os ASES podem ser bastante grandes, especialmente se o valor teórico de  $\sigma_{\eta}^2$  é pequeno (0,09). Por sua vez, de maneira geral, os vícios tendem a aumentar quando  $\gamma$  ou  $\sigma_{\eta}^2$  diminuem.

As simulações realizadas para o caso não-estacionário, considerando valores de:  $\sigma_{\eta}^2 = \{0,9; 0,01; 0,0009; 0,0001\}$ , mostram que as estimativas são boas quando trabalhamos com amostras grandes, mas, na medida que  $\sigma_{\eta}^2$  diminui, os vícios das estimativas tendem a

aumentar. Da mesma forma, valores pequenos de  $\sigma_\eta^2$  produzem o aumento dos ASES (chegando a ser grosseiros). Tudo isso pode ser ocasionado devida à pobre aproximação da normal para  $\hat{\sigma}_\eta^2$ . Por sua vez, para valores grandes de  $\sigma_\eta^2$ , os ASES estimados são razoavelmente corretos.

Experimentos de simulação realizados por Jacquier *et alii* (1994) confirmam os resultados anteriores. Utilizando como critério para a eficiência o MSE, eles encontraram que os estimadores QML não são satisfatórios quando a amostra é pequena e os valores de  $\sigma_\eta^2$  são baixos. Por sua vez, quando o tamanho da amostra é igual a 2000, os vícios diminuem bastante, mas ainda são grandes.

Existem duas razões para explicar os resultados apresentados. A primeira, apontada por Ruiz (1994), assinala que quando  $\sigma_\eta^2$  e  $\gamma$  são pequenos, a variância do processo da log-volatilidade;  $\sigma_h^2 = \frac{\sigma_\eta^2}{1-\gamma^2}$ , é muito pequena quando comparado com a variância de  $\xi_t$ . Consequentemente, a transformação  $\log(Y_t^2)$  é dominada por  $\xi_t$  e, portanto, a aproximação normal utilizada para  $\log(Y_t^2)$  é muito inadequada. Em segundo lugar, o pobre desempenho dos estimadores pode ser consequência da presença de observações muito pequenas ou *inliers*. Assim, na medida que os retornos estão próximos de zero (em valor absoluto), os valores de  $\log(Y_t^2)$  tendem a menos infinito, afetando (as vezes drasticamente) as estimativas obtidas via Filtro de Kalman. Por contornar esse problema, sugere-se utilizar as alternativas mencionadas na seção 4.4.1.

#### *Comparação dos Métodos QML, GMM e MCMC*

Na literatura tem sido utilizados outros métodos para estimar os modelos de variância estocástica, entre eles, o Método dos Momentos (Taylor, 1980,1986), o GMM em Melino e Turnbull (1990) e o MCMC em Jacquier *et alii* (1994).

Como apontado por Shephard (1994b), o GMM apresenta várias dificuldades. Uma delas é não fornecer estimadores suavizados ou filtrados para  $h_t$ , sendo necessária uma segunda forma de estimação (complementar). Por outro lado, o método Bayesiano MCMC sugerido recentemente por Jacquier *et alii* (1994) abre um futuro promissor na metodologia de

estimação dos modelos de variância estocástica. Mas, o custo computacional pode ser considerável e por outro lado, as propriedades dos estimadores são ainda matéria de estudo.

Numa aplicação empírica com taxa de câmbio Taylor (1994) encontra que as estimativas fornecidas pelo método dos momentos e pelo método QML são muito próximas. Em outro estudo, experimentos de simulação para comparar os métodos GMM e QML foram realizados por Ruiz (1994) utilizando como critério o ASE. Para modelos AR(1)-SV não estacionários ( $\sigma_\eta^2 = \{0,09; 0,01; 0,0009; 0,0001\}$ ) ela encontra que a eficiência relativa do GMM é menor quando comparada com a obtida pelo método QML, sendo que essa eficiência diminui na medida que diminui o valor de  $\sigma_\eta^2$ . Por outro lado, no caso estacionário o desempenho do QML é melhor quando  $\gamma$  é próximo de um e  $\sigma_\eta^2$  é relativamente grande. Pelo contrário, quando  $\gamma$  e  $\sigma_\eta^2$  são pequenos, o GMM é mais eficiente. A explicação desse comportamento é simples: quando os parâmetros  $\gamma$  e  $\sigma_\eta^2$  são pequenos o processo  $Y_t$  é aproximadamente normal (pouco excesso de curtose), em conseqüência, os momentos amostrais são bons estimadores dos momentos populacionais.

Por outro lado, utilizando como critério o MSE, Jacquier *et alii* (1994) apresentam comparações dos métodos QML GMM e Bayes MCMC através de simulações em séries de tamanhos 500 e 2000. Os resultados mostram que, na maioria das vezes, o QML é mais eficiente que o GMM e que a solução Bayesiana apresenta uma eficiência relativa de 1/3 quando comparada com os outros dois métodos, tendo, portanto, desempenho muito superior.

## 4.5 Diagnóstico e Bondade de Ajuste

### 4.5.1 Diagnóstico

Seguindo a metodologia proposta por Harvey (1989) e Harvey e Koopman (1992), o diagnóstico nos modelos de volatilidade estocástica é realizado examinando os *Quase-Resíduos Padronizados* (QRP),  $\tilde{v}_t = v_t / \sqrt{f_t}$  e *Quase-Resíduos Padronizados Auxiliares* (QRPA), os quais são estimativas suavizadas das inovações  $\xi_t$  e  $\eta_t$  definidas em (4.4.1.1) e denotados por  $\tilde{\xi}_t$  e  $\tilde{\eta}_t$ .

Quando o modelo é corretamente especificado e os parâmetros  $\psi$  são conhecidos, os QRP possuem média zero e variância unitária. Por sua vez, dado que a distribuição de  $\log(\varepsilon_t^2)$  é assimétrica e possui cauda pesada, então os QRP apresentam também essa característica. Adicionalmente, os QRP são não-correlacionados, embora dependentes Shephard (1994b, p.29), de forma que através da f.a.c. podem ser detectados comportamentos indesejáveis nas correlações.

Um recurso gráfico para avaliar o comportamento aleatório dos QRP é fazer o gráfico deles em relação ao tempo. Da mesma forma, através do gráfico do CUSUM dos QRP e CUSUM dos quadrados dos QRP, ver Brown *et alii* (1975) e Harvey (1993a, p.153-55), podemos detectar a presença de mudanças estruturais e /ou heteroscedasticidade.

Testes de correlação serial como o teste Portmanteau de Box-Ljung que tem distribuição  $\chi^2$  sob a hipótese de independência e o teste de Durbin-Watson (DW) que procura detectar a presença de autocorrelação, podem ser utilizados. De outro lado, um teste para detectar presença de heteroscedasticidade proposto por Harvey (1989, p.259) é:

$$H(h) = \frac{\sum_{t=T-h+1}^T \tilde{v}_t^2}{\sum_{t=d+1}^T \tilde{v}_t^2}, \quad (4.5.1.1)$$

onde  $h = \lfloor (T-d)/3 \rfloor$  e  $d$  é o número de observações utilizadas para inicializar o Filtro. Sob a hipótese nula de homoscedasticidade, a distribuição da estimativa  $H(h)$  é F-Snedecor com graus de liberdade iguais a  $h$ .

No caso dos QRPA sob a hipótese de ausência de autocorrelação, eles também possuem média zero e variância igual a um, mas, contrariamente aos QRP, apresentam correlação serial mesmo que o modelo esteja corretamente especificado (Harvey e Koopman, 1992). Os QRPA  $\tilde{\xi}_t$  apresentam distribuição assimétrica com cauda pesada e  $\tilde{\eta}_t$  possui distribuição normal. Através da análise deste último podemos detectar possíveis mudanças estruturais na volatilidade.

Por outro lado, seria interessante fazer o diagnóstico sobre as inovações originais  $\varepsilon_t$ , as quais são supostas normais padrão independentes. Nesse sentido, uma alternativa proposta por

Harvey e Shephard (1993b) consiste em trabalhar com as inovações corrigidas pela heteroscedasticidade, as quais são definidas como  $\tilde{\epsilon}_t = Y_t / \hat{\sigma}_{t/T}$ . Eles mostram que se os parâmetros fossem conhecidos, as inovações corrigidas são não-correlacionadas e possuem distribuição simétrica

#### 4.5.2 Bondade de Ajuste

Uma medida adequada para medir a bondade de ajuste nos modelos de séries temporais (ver Harvey ,1989 p. 264) é a Variância do Erro de Previsão um-passo à frente (*v.e.p*) após atingir o *regime estável* (Steady State), isto é:

$$\text{v. e. p} = \sigma^2 = \sigma^2 \bar{f} = \sigma^2 \lim_{t \rightarrow \infty} f_t.$$

Nas aplicações,  $f_t$  é substituído por  $f_T$ , o qual é praticamente constante logo das primeiras iterações do Filtro.

Quando duas especificações do modelo são comparados, preferimos aquele com menor *v.e.p*. Podemos também utilizar os critérios AIC (Akaike,1973) ou BIC (Schwarz ,1978) e escolher o modelo que apresente menores valores.

Outra estratégia para verificar a bondade do ajuste é mediante a previsão um-passo à frente *ex-ante*. Assim, os intervalos de confiança  $100\alpha\%$  (i.c) construídos com a *volatilidade filtrada* baseada na informação anterior, devem incluir aproximadamente  $100\alpha\%$  dos retornos.

#### 4.6 Previsão

O modelo de Variância Estocástica definido em (4.3.1) é martingale diferença. Consequentemente, da mesma forma que nos modelos ARCH, supondo que os parâmetros são conhecidos, o valor predito dos retornos é zero para qualquer horizonte de previsão.

Por outro lado, dentro da abordagem de Quase-Máxima Verossimilhança, é possível obter facilmente previsões para a volatilidade utilizando os componentes filtrados estimados

pelo Filtro de Kalman. Assim, no modelo AR(1)-SV estacionário (4.3.1) com a parametrização (4.4.1.1), a previsão da volatilidade  $s$ -passos à frente é obtida recursivamente:

$$\hat{\sigma}_{I,t+s/t}^2 = \hat{k} \exp(\hat{h}_{t+s-1/t}) \quad s \geq 2 \quad (4.6.1)$$

É interessante explicitar a expressão da estimativa da previsão como função das observações. A seguir mostramos a derivação dessa expressão supondo (para simplificar) que  $k=1$ .

Substituindo (4.4.1.2c) em (4.4.1.2d) e usando a definição de Ganho de Kalman obtemos que  $k_t = P_t$ . Em conseqüência, uma forma de re-escrever a equação (4.4.1.2e) é a seguinte:

$$m_t = [\gamma(1-P_t)]m_{t-1} + [P_t] X_t, \quad (4.6.2)$$

e se considerarmos (4.4.1.2a) obtemos:

$$m_{t/t-1} = [\gamma(1-P_{t-1})]m_{t-1/t-2} + \gamma P_{t-1} X_{t-1}. \quad (4.6.3)$$

Dessa forma, a partir das expressões (4.6.2) e (4.6.3) podemos calcular recursivamente tanto as estimativas do nível  $h_t$  dado  $I_t$  ( $m_t$ ), quanto as estimativa baseadas em  $I_{t-1}$ , ( $m_{t/t-1}$ ), de uma forma muito simples. Baseadas nessas expressões e supondo as condições iniciais  $P_{1/0}$  e  $m_{1/0}$ , não é difícil mostrar que a estimativa do nível  $h_{t+1}$  dada a informação passada  $I_t$  está dada por:

$$m_{t+1/t} = \sum_{i=1}^t C_i X_i + \psi \quad (4.6.4a)$$

onde

$$\psi = \gamma^t \prod_{i=1}^t (1-P_i) m_{1/0} \quad (4.6.4b)$$

$$C_i = \gamma^{t-i+1} P_i \prod_{j=i+1}^t (1-P_j) \quad i = 1, 2, \dots, t-1 \quad (4.6.4c)$$

$$C_t = \gamma P_t \quad (4.6.4d)$$

Tanto a expressão obtida em (4.6.4a) quanto a estimativa do nível  $h_t$  (dada por  $m_t = \gamma m_{t-1/t}$ ), constituem combinações lineares das observações  $X_i$ . Dessa forma, a importância de cada observação na estimativa é dada pelo coeficiente  $C_i$ . A discussão a seguir visa explorar a natureza desses coeficientes e da constante  $\psi$  na expressão das estimativas do nível  $h_t$ , na tentativa de simplificar o modelo.

Na modelagem de dados financeiros freqüentemente dispomos de quantidades razoavelmente grandes de observações (digamos 1000). Em conseqüência, a contribuição de  $\psi$  na estimação de  $m_t$ , quando  $t$  é grande, é praticamente nula. Isso pode ser observado a partir da expressão (4.6.4a), onde  $\psi$  é produto de quantidades menores que um e, além disso,  $\gamma^t$  converge a zero na medida que aumenta  $t$ .

Por outro lado, dentro dos modelos de nível local é conhecido que os valores de  $P_i$  convergem rapidamente aos regimes estáveis  $P^*$ . Assim, no limite :

$$P^* = \frac{1}{2\gamma^2} \left( (\gamma^2 - q - 1) + \sqrt{(q + 1 - \gamma^2) + 4q\gamma^2} \right). \quad (4.6.5)$$

Dessa forma, se  $i_0$  é o instante de tempo onde é alcançado o regime estável,  $P^*$ , isto é,  $P^* = P_{i_0+s}$ ,  $s=1,2,\dots$ , então os coeficientes  $C_i$  em (4.6.4c) e (4.6.4d) são:

$$C_i = \gamma^{t-i+1} P_i \prod_{j=i+1}^t (1 - P_j) \quad i=1,2,\dots,i_0-1 \quad (4.6.6a)$$

$$C_i = [\gamma P^*][\gamma(1 - P^*)]^{t-i} \quad i=i_0,\dots,t. \quad (4.6.6b)$$

Levando em consideração (4.6.6a), para valores de  $t$  grandes, a contribuição dos produtos  $C_i X_i$   $i=1,2,\dots,i_0-1$  na soma (4.6.4a) é desprezível, portanto, uma boa aproximação é:

$$m_{t+1/t} = \sum_{i=1}^t C_i X_i, \quad (4.6.7)$$

com coeficientes  $C_i$  dados por (4.6.7b). Supondo que o valor do parâmetro  $\gamma$  seja aproximadamente 1, o que acontece muitas vezes na prática, de (4.6.4) obtemos:

$$m_{t+1/t} = [P^*] X_t + [1-P^*] m_{t/t-1}, \quad (4.6.8)$$

sendo um Modelo Exponencial de Média Móvel (*Exponential Weighted Moving Average*, EWMA). Esse procedimento *ad-hoc* foi utilizado (com algumas modificações) por Taylor (1986) para estimar a volatilidade, sem a abordagem de espaço de estados utilizando argumentos de pesquisa operacional, sob a hipótese de alta persistência na volatilidade.

Finalmente, substituindo  $X_t = \log(Y_t^2) + 1.27$  (pois  $k=1$ ) na expressão 4.6.8 obtemos:

$$\begin{aligned} \sigma_{T+1/T}^2 &= \exp\left[\sum_{i=1}^T C_i (\log(Y_i^2) + 1.27)\right] \\ &= \exp\left(\sum_{i=1}^T \log[Y_i^2]^{C_i}\right) \exp\left(1.27 \sum_{i=1}^T C_i\right) \\ \sigma_{T+1/T}^2 &= \left(\prod_{i=1}^T (Y_i^2)^{C_i}\right) \exp\left(1.27 \sum_{i=1}^T C_i\right), \end{aligned} \quad (4.6.9)$$

portanto, nos modelos de variância estocástica, a previsão da volatilidade um-passo a frente é uma espécie de média geométrica dos quadrados dos retornos (pesos dados por  $C_i$ ) ajustada por um fator multiplicativo.

## 4.7 Extensões

Vários outros *Fatos Estilizados* podem ser modelados dentro da abordagem de espaço de estados fazendo leves modificações no modelo básico AR(1)-SV. Nesta Seção são apresentados, sem maiores detalhes, como podemos modelar algumas características empíricas observadas na literatura como: *Cauda Pesada*, *Efeitos Calendário*, *Correlação Serial* e *Intervenções*. Para um resumo atualizado dessas e outras regularidades ver Ghysels *et alii* (1995).

### 4.7.1. Cauda Pesada

Nas aplicações empíricas, é observado na literatura que os resíduos padronizados das perturbações  $\varepsilon_t$  apresentam cauda pesada. Nesse sentido, para contornar esse problema, uma proposta de Harvey *et alii* (1994) e Ruiz (1994), consiste em supor que as perturbações  $\varepsilon_t$  seguem uma distribuição *t-Student* com  $\nu$  graus de liberdade, ou considerar as distribuições GED. Em ambos os casos a representação em Espaço de Estados é obtida seguindo a metodologia da Seção 4.2. A seguir apresentamos os resultados quando a distribuição é *t-Student*.

Começamos com o caso estacionário. Definindo a inovação  $\xi_t = \log(\varepsilon_t^2) - E[\log(\varepsilon_t^2)]$ , pelas propriedades da distribuição log-normal obtemos (ver Abramovitz e Stegum, 1970):

$$\xi_t = \log(\varepsilon_t^2) + 1.27 + \psi(\nu/2) - \log(\nu/2), \quad (4.7.1.1a)$$

$$\sigma_\xi^2 = \text{Var}[\log(\varepsilon_t^2)] = \frac{\pi^2}{2} + \psi'(\nu/2), \quad (4.7.1.1b)$$

onde  $\psi'$  é a função trigamma. A representação em espaço de estados é (4.4.1.1) considerando  $X_t = \log(Y_t^2) - \log(\kappa) + 1.27 + \psi(\nu/2) - \log(\nu/2)$ . O processo de estimação é realizado da forma descrita na Seção 4.4.1, considerando que  $\log(\varepsilon_t^2)$  possui *distribuição normal*, sendo neste caso  $X_t^* = \log(Y_t^2) - \log(Y_t^2)$ . Utilizando o Filtro de Kalman calculamos a quase-verossimilhança e, ao maximizar essa sobre a restrição  $\sigma_\xi^2 \geq \pi^2/2$ , encontramos estimativas para os parâmetros  $\Psi = (\sigma_\xi^2, \gamma, \sigma_\eta^2)$ . O parâmetro  $\nu$  é obtido de (4.7.1.1b) e podemos calcular uma estimativa consistente de  $\log(\kappa)$  através da expressão:

$$\log(\kappa) = \log(Y_t^2) + 1.27 + \psi(\nu/2) - \log(\nu/2). \quad (4.7.1.2)$$

A distribuição assintótica de  $\hat{\Psi}$  é dada por:

$$\sqrt{T}(\hat{\Psi} - \Psi) \approx N(0, C(\Psi)),$$

onde  $C(\Psi)$  pode ser encontrada em Ruiz (1994).

Por outro lado, na representação em espaço de estados do modelo AR(1)-SV não-estacionário, a única modificação é considerar  $\kappa$  e  $\gamma$  iguais a 1.

É possível fazer inferência considerando  $v$  como conhecido, mas, segundo Ruiz (1994), as estimativas são menos eficientes que no caso quando  $v$  é desconhecido. Por sua vez, podemos testar a hipótese  $H_0: \sigma_{\xi}^2 = 4,93$ . Para isso, Ruiz (1994) propõe um teste de Razão de Verossimilhança que sob  $H_0$  possui distribuição aproximada  $\chi_{(1)}$ .

#### 4.7.2 Correlação Serial e Efeitos Calendário

A presença de autocorrelação de ordem um nos retornos, assim como o *Efeito Calendário* na média são incorporados facilmente no modelo AR(1)-SV através de uma extensão similar à realizada na Seção 3.3.5 para os ARCH. Dessa forma, definimos o modelo para os retornos  $\{Y_t\}$ :

$$Y_t = \rho Y_{t-1} + \sum_{i=1}^5 C_i D_i + Y_t^* \quad (4.7.2.1a)$$

$$Y_t^* = \varepsilon_t \sigma_t \quad (4.7.2.2a)$$

$$\sigma_t^2 = \kappa \exp(h_t) \quad (4.7.2.2b)$$

$$h_t = \gamma h_{t-1} + \eta_t \quad (4.7.2.2c)$$

onde  $D_i$  são variáveis *dummy* correspondentes aos dias da semana e assumem valores 0 ou 1. A estimação é realizada em dois estágios. Primeiro são estimados os parâmetros  $\rho$  e  $C_i$   $i=1...5$ , utilizando o método de Mínimos Quadrados Ordinários (OLS) e, por sua vez, obtidos os resíduos  $\hat{Y}_t^*$ . A seguir o modelo de variância estocástica (4.7.2.2) é ajustado aos resíduos  $\hat{Y}_t^*$  através do procedimento descrito em na subseção 4.3.1. Como mostram Harvey e Shephard (1993b) as estimativas fornecidas pelo método de dois estágios são satisfatórias.

#### 4.7.3 Intervenções

Mudanças estruturais abruptas na volatilidade podem ser capturadas especificando o modelo (4.4.1.1) como sendo:

$$\sigma_t^2 = \kappa \exp(h_t) \exp(\lambda D_t) \quad (4.7.3.1)$$

onde  $D_t$  é uma variável *dummy* que assume o valor 0 antes da ruptura e 1 depois da ruptura. Dessa forma, o efeito da mudança estrutural (no nível) é dado por  $\exp(\lambda)$ . Por exemplo, podemos modelar o *Efeito Final de Semana* na variância dos retornos definindo  $D$  como sendo 0 de terça a sexta-feira e 1 na segunda-feira. Em consequência, valores positivos de  $D$  implicam que a volatilidade aumenta nos finais de semana. O ajuste do modelo segue as mesmas bases da metodologia proposta na subseção 4.3.1.

Na modelagem ARCH, devido à realimentação entre as observações e a volatilidade, a incorporação destes efeitos é mais complicada. Por exemplo, no GARCH(1,1), uma mudança abrupta pode ser modelado como (ver Ghysels *et alii*, 1995):

$$\sigma_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2 + \lambda D_t - \lambda(\alpha + \beta)D_{t-1} \quad (4.7.3.2)$$

A implementação deste modelo é difícil devido à dificuldade de estimar os parâmetros, pois existem restrições para  $\lambda$  de tal forma que  $\sigma_t^2$  seja sempre positiva.

## 4.8 Aplicações nas *Séries Exemplo*

Para ilustrar a metodologia exposta nas seções anteriores, realizamos os ajustes das *Séries Exemplo* considerando o seguinte modelo para os retornos:

$$Y_t = \mu_t + Y_t^* \quad (4.8.1a)$$

$$\mu_t = E[Y_t / I_t] \quad (4.8.1b)$$

$$Y_t^* = \sigma_t \varepsilon_t \quad (4.8.1c)$$

Para especificar a volatilidade, distinguimos dois casos, as versões estacionária e não-estacionária do modelo AR(1)-SV.

### *Modelo AR(1)-SV Estacionário*

Neste caso utilizamos a especificação desenvolvida na Seção 4.6.1 no modelo de variância estocástica com inovações  $\varepsilon_t$  que possuem distribuição *t-Student* com  $\nu$  graus de liberdade. O modelo é o seguinte:

$$\sigma_t^2 = \kappa \exp(h_t), \quad h_t = \gamma h_{t-1} + \eta_t \quad (4.8.2a)$$

e a volatilidade:

$$\text{Vol} = \sigma_t^2 \frac{v}{v-2}. \quad (4.8.2b)$$

*Modelo AR(1)-SV Não-estacionário*

Neste caso, a volatilidade é definida simplesmente como  $\sigma_t^2$  onde:

$$\sigma_t^2 = \exp(h_t), \quad h_t = h_{t-1} + \eta_t, \quad (4.8.3)$$

sob a suposição que os choques  $\varepsilon_t$  em (4.7.1c) seguem distribuição normal padrão<sup>6</sup>.

Adicionalmente, com o objetivo de incorporar na volatilidade o efeito dos dias da semana, por exemplo, o *Efeito de Final de Semana* e o *Efeito de Feriados*, definidos como duas variáveis *dummy*:

$$\text{Efeito Final de Semana: } D_{s,t} = \begin{cases} 1 & Y_t, \text{ segunda } \quad Y_{t-1}, \text{ sexta} \\ 0 & \text{outro caso} \end{cases}, \quad (4.8.4a)$$

$$\text{Efeito de Feriados: } D_{f,t} = \begin{cases} 1 & \text{feriado ocorre entre os tempos } t \text{ e } t-1 \\ 0 & \text{outro caso} \end{cases}, \quad (4.8.4b)$$

podemos considerar quatro especificações do modelo AR(1)-SV estacionário:

Modelo I: volatilidade com *Efeito Final de Semana e Efeito Feriados*

Modelo II: volatilidade com *Efeito Final de Semana*.

Modelo III: volatilidade com *Efeito Feriados*.

Modelo IV: volatilidade sem nenhum dos efeitos.

(4.8.5)

---

<sup>6</sup>Não é considerada a distribuição *t-Student* para o modelo não-estacionário pois, neste caso, não é possível estimar o modelo com o pacote disponível.

Nesse contexto, a volatilidade no modelo AR(1)-SV estacionário é expressa como:

$$\text{Vol} = \kappa \exp(h_t) \exp(wD_s, t) \exp(fD_f, t) \frac{v}{v-2}, \quad (4.8.6)$$

e, para o caso não-estacionário substituímos na expressão anterior  $\kappa$  e  $\frac{v}{v-2}$  como sendo 1.

Para o caso estacionário, a representação de espaço de estados é:

$$X_t = T_t + h_t + wD_s + fD_f + \xi_t \quad (4.8.7a)$$

$$T_t = T_{t-1} = T \quad (4.8.7b)$$

$$h_t = \gamma h_{t-1} + \eta_t \quad (4.8.7c)$$

no qual:

$$X_t = \log(Y_t^{*2}) \quad (4.8.8a)$$

$$T_t = -1,27 - \psi(v/2) + \log(v/2) + \log(\kappa_t) \quad (4.8.8b)$$

$$E[\varepsilon_t \eta_{t-s}] = 0, \quad \xi_t \approx N(-1,27; 4,93), \quad \eta_t \approx N(0, \sigma_\eta^2) \quad (4.8.8c)$$

onde  $\psi(\cdot)$  é a função digamma já apresentada na seção 4.2.

Como anteriormente, no caso não-estacionário, a representação do modelo em espaço de estados é a mesma exceto pela inexistência da variável  $T$  e com  $\gamma$  igual a 1.

As estimativas dos componentes e dos parâmetros podem ser calculadas através do pacote *STAMP 5.0* (Koopman e Harvey, 1995). Esse pacote permite estimar o modelo AR(1)-SV estacionário na representação (4.8.7)<sup>7</sup>, fornecendo as estimativas de QML dos parâmetros;  $\gamma$ ,  $\sigma_\xi^2$  e  $\sigma_\eta^2$  e as estimativas suavizadas e filtradas dos componentes da tendência  $T_t$ , do componente autoregressivo  $h_t$  e do componente irregular  $\xi_t$ . Supondo que as inovações são  $t$ -

---

<sup>7</sup> Especificando nível fixo, sem inclinação, sem componente sazonal e presença de componentes irregular e autoregressiva.

*Student*, o parâmetro  $\nu$  pode ser estimado indiretamente através da relação  $\hat{\sigma}_\xi^2 = 4,93 + \psi'(\hat{\nu}/2)$ .<sup>8</sup> Logo, o nível e as volatilidades podem ser estimados da seguinte forma:

$$\text{nível: } \hat{\kappa}_t = \exp[\hat{T}_t + 1,27 + \psi(\hat{\nu}/2) - \log(\hat{\nu}/2)] \quad (4.8.10a)$$

$$\text{Vol\_f}_t = \exp[\hat{T}_f + 1,27 + \psi(\hat{\nu}/2) - \log(\hat{\nu}/2) + wD_{f,t} + fD_{f,t} + \hat{h}_{-f_t}] \frac{\hat{\nu}}{\hat{\nu}-2} \quad (4.8.10b)$$

$$\text{Vol\_s}_t = \exp[\hat{T}_s + 1,27 + \psi(\hat{\nu}/2) - \log(\hat{\nu}/2) + wD_{s,t} + fD_{s,t} + \hat{h}_{-s_t}] \frac{\hat{\nu}}{\hat{\nu}-2} \quad (4.8.10c)$$

sendo que a letra *f* corresponde à estimativa filtrada e a letra *s* à estimativa suavizada. Como discutido na seção 4.4.1, as estimativas filtradas no tempo  $t$  fazem uso da informação até esse instante ou até o instante anterior, enquanto que as estimativas suavizadas utilizam toda a informação da amostra.

Por sua vez, no caso do modelo não-estacionário, o pacote *STAMP 5.0*<sup>9</sup> estima os parâmetros  $\sigma_\xi^2$  e  $\sigma_\eta^2$  e as componentes (4.8.10). A rigor, a estimação teria que ser realizada com a restrição  $\sigma_\xi^2 = 4,93$ , portanto as estimativas serão analisadas com cautela.

A seguir apresentamos os resultados dos ajustes nas *Séries Exemplo*. Para completar o modelo (4.8.1) resta especificar a forma da média condicional. Nas duas séries essa média depende dos dias da semana através de dummies, nesse sentido é necessário ressaltar que os efeitos dos dias da semana na média e variância condicionais podem estar confundidos, de tal forma que as interpretações devem ser tomadas com cuidado. Esse confundimento também pode acontecer quando introduzimos na média condicional a correlação serial.

#### 4.8.1 Ações

Com o objetivo de incorporar o *Efeito Calendário* e a autocorrelação serial observadas no Capítulo 2, é proposto o seguinte modelo para a média condicional:

<sup>8</sup>Esse recurso é válido sempre que  $\hat{\sigma}_\xi^2 \geq 4,93$ , o que ocorre nas duas aplicações desta dissertação.

<sup>9</sup> Para estimar esse modelo especificamos como opções: nível estocástico, sem inclinação, sem componente sazonal e presença de componente irregular.

$$\mu_t = E[Y_t / I_t] = \rho Y_{t-1} + \text{Seg}D_1 + \text{Ter}D_2 + \text{Quar}D_3 + \text{Quin}D_4 + \text{Sex}D_5.$$

As estimativas são encontradas através do método de mínimos quadrados ordinários:

$$\hat{\mu}_t = 0,0865Y_{t-1} - 0,0077D_1 + 0,0049D_2 - 0,0033D_3 + 0,0051D_4 + 0,0013D_5$$

(0,0249)      (0,0034)      (0,0034)      (0,0034)      (0,0034)      (0,0034)

sendo as quantidades entre parêntesis os desvios padrão das estimativas. Segundo os resultados, podemos observar que a estimativa da autocorrelação de primeiro ordem (valor-P=0,005) e a estimativa da variável *dummy* correspondente à segunda feira (valor P=0,0235) são significativas. A seguir estimamos a volatilidade fazendo uso dos resíduos  $Y_t^*$ , considerando os modelos (4.8.5) para os casos estacionário e não-estacionário<sup>10</sup>.

#### *Estimação dos Modelos AR(1)-SV Estacionários*

Como nenhum dos resíduos é igual a zero, portanto, não há necessidade de aplicar uma transformação nos retornos. A seguir, utilizando o pacote *STAMP 5.0*<sup>11</sup> encontramos as estimativas que são mostradas na Tabela 4.1. Observa-se que os processos de estimação apresentam *convergência muito forte*<sup>12</sup> em 11 iterações para o Modelo IV, *convergência forte* em 12 iterações para os modelos I e III e, em 11 iterações para o Modelo II. Os resultados da Tabela 4.1 mostram que nos modelos I, II e III, nenhuma das variáveis *dummy* é significativa a 5%, portanto esses modelos são descartados. No entanto, ao comparar as estimativas de  $\hat{\sigma}_\xi^2$ ,  $\hat{\sigma}_\eta^2$ ,  $\hat{q}$ ,  $\hat{\gamma}$ ,  $\hat{v}$  encontramos que elas são muito próximas e, com relação aos valores das medidas de bondade de ajuste, o Modelo I apresenta menor p.e.v enquanto que o modelo IV (sem efeitos) apresenta menores valores AIC e BIC.

<sup>10</sup> Embora as estimativas das variáveis *dummy* não sejam significativas, elas são mantidas no modelo com o propósito de comparar este modelo com os modelos ARCH. Isso é realizado no capítulo seguinte.

<sup>11</sup> Valores iniciais para o Filtro e parâmetros são determinados automaticamente pelo pacote.

<sup>12</sup> Essa expressão indica que as estimativas são altamente confiáveis. No *STAMP 5.0* existem as seguintes classificações para a convergência em ordem de confiabilidade: muito forte, forte, fraca, muito fraca.

**Tabela 4.1**  
**Ações: estimativas dos modelos com Efeitos dos dias da semana na volatilidade**

	Modelo I	Modelo II	Modelo III	Modelo IV <sup>a</sup>
p.e.v	5,3764	5,3840	5,3800	5,3882
AIC	1,6883	1,6884	1,6877	1,6880
BIC	1,7051	1,7019	1,7012	1,6981
w	0,1468 (0,30) <sup>a</sup>	0,1585 (0,26)	-	-
f	0,4305 (0,13)	-	0,4468 (0,12)	-
$\hat{\sigma}_{\xi}^2$	5,0284	5,0326	5,0291	5,0341
$\hat{\sigma}_{\eta}^2$	0,0463	0,0463	0,0462	0,0461
$\hat{q}$	0,0092	0,0092	0,0092	0,0092
$\hat{\gamma}$	0,9656	0,9656	0,9657	0,9656
$\hat{\nu}$	22	20	22	20

<sup>a</sup> valores-P entre parêntesis

#### Estimação do Modelo AR(1)-SV Estacionário (Modelo IV)

Para este modelo temos que as estimativas são iguais a:

$$\hat{h}_t = 0,966\hat{h}_{t-1} + \hat{\eta}_t$$

com  $\hat{\sigma}_{\xi}^2 = 5,0341$ ,  $\hat{\sigma}_{\eta}^2 = 0,0461$ , conseqüentemente  $\hat{q} = 0,0092$ .

A persistência estimada é alta, 0,966, significando uma meia-vida de 20,04 dias, isto é, o tempo necessário para que o efeito de um choque na volatilidade estimada diminua até a metade é aproximadamente um mês de funcionamento de mercado. Por outro lado, o valor estimado do parâmetro  $\nu$  é encontrado através da relação  $\hat{\sigma}_{\xi}^2 = 4,93 + \psi'(\nu/2)$ , sendo aproximadamente igual a 20.

Na Figura 4.1a é apresentado o gráfico das estimativas suavizada e filtrada do componente  $T_t$  e na Figura 4.1b o gráfico das estimativas filtrada e suavizada do nível na

volatilidade  $\kappa$ . Ao analisar a variação da estimativa filtrada (tanto em  $T_t$  quanto em  $\kappa$ ) distinguimos três períodos (diferenciados no gráfico pelas linhas verticais). As grandes variações observadas no primeiro período são naturais na fase inicial do Filtro de Kalman. No período seguinte observamos que as variações ainda são bastante grandes colocando sob suspeito a suposição que o modelo possui nível constante  $T$ . Finalmente, no terceiro período a variação da estimativa filtrada é pequena e além disso possui valores próximos da estimativa suavizada.

Na Figura 4.2 é apresentado o gráfico dos componentes  $h_{f_t}$  e  $h_{s_t}$  os quais possuem variabilidade baixa (variâncias iguais a 0,26705 e 0,4260 respectivamente) quando comparados com o componente irregular filtrado (variância igual a 5,03) que é apresentado na Figura 4.3.

As volatilidades estimadas filtrada e suavizada são mostradas na Figura 4.4 e, como pode ser observado, ambas são capazes de acompanhar os movimentos de aumentos e diminuições nos quadrados dos retornos.

#### *Adequabilidade do Modelo AR(1)-SV Estacionário IV*

A avaliação da adequabilidade do modelo é realizada examinando os quase resíduos padronizados (QRP),  $\tilde{v}_t$ , definidos na seção 4.5.1. Os valores apresentados pela média, 0,00003 e pela variância, 1,00 são satisfatórios por serem próximos dos valores 0 e 1 respectivamente. Por sua vez, o coeficiente de simetria é igual a -1,21 e o excesso de curtose é igual a 2,63 indicando que a distribuição dos QRP é assimétrica e possui cauda pesada. Essas características podem ser observadas no gráfico do histograma e da densidade ajustada<sup>13</sup> na Figura 4.5a e, como comentado na Seção 4.5.1 tal comportamento é esperado. Adicionalmente, os gráficos do CUSUM e CUSUMSQ apresentados nas Figuras 4.5b e 4.5c mostram um bom comportamento. O valor da estatística  $H(532)$  (definida em 4.5.1.1) é igual a 0,968 (valor-P=0,65), em conseqüência, não existe evidência para rejeitar a hipótese nula de homoscedasticidade e/ou ausência de mudança estrutural nos QRP.

---

<sup>13</sup> Densidade não-paramétrica fornecida pelo STAMP 5.0

Analisando as autocorrelações dos QRP, a partir da função de autocorrelação (Figura 4.5d) e os valores do Teste Box-Ljung (Tabela 4.2), podemos concluir que não existe forte evidência de presença de correlação serial.

**Tabela 4.2**  
*Ações: valores P do Teste Box-Ljung nos QRP do AR(1)-SV Estacionário*

<i>Lag</i>	6	12	18	24	30	36	42
<i>Valor P</i>	0,127	0,209	0,283	0,318	0,281	0,218	0,210

Por outro lado, supondo que o modelo é o verdadeiro, os intervalos de confiança  $100(1-2\alpha)\%$  para os retornos (i.c) devem incluir aproximadamente  $100(1-2\alpha)\%$  dos retornos. Esses i.c. são calculados utilizando a estimativa da volatilidade suavizada:  $(\mu_t - t_{v,\alpha}\sqrt{\text{Vol}_{s_t}}, \mu_t + t_{v,\alpha}\sqrt{\text{Vol}_{s_t}})$ , onde  $t_{v,\alpha}$  corresponde ao  $(1-\alpha)$  percentil da distribuição *t-Student* padrão com  $v$  graus de liberdade. O fato de utilizar os mesmos retornos tanto para construir os intervalos quanto para avaliar a percentagem que está incluída dentro desses, justifica-se considerando que o valor de um retorno não influencia muito o valor da volatilidade suavizada na série. Dessa forma, considerando que  $t_{20,0975} = 2,086$ , os i.c. 95% excluem 3,44% dos retornos sendo esse resultado satisfatório.

Adicionalmente, podemos verificar o modelo através do critério de seu desempenho na previsão um-passo à frente *ex-ante*. Idealmente, para realizar isso seria necessário re-estimar o modelo cada vez que um retorno é incorporado na série, mas, levando em consideração o grande volume de trabalho computacional isso não é realizado. Mas, é possível utilizar uma alternativa supondo que as estimativas dos parâmetros no modelo, atualizadas cada vez que um retorno é incorporado na série, não apresentam grandes diferenças num período de, por exemplo, 50 retornos. Dessa forma, estimando os parâmetros até o instante  $t_0$  e supondo que esses são constantes nas próximas 50 observações, podemos calcular a componente filtrada via Filtro de Kalman e obter as previsões da volatilidade um-passo à frente através da expressão (4.6.1). A seguir, podemos calcular os i.c. para os retornos. Mas, através do pacote *STAMP 5.0* não é possível implementar este processo. Outra alternativa é modificar esse processo da seguinte forma:

- 1) Estimar os parâmetros do modelo até o instante  $t_0$ .
- 2) Baseados nas estimativas dos parâmetros no instante  $t_0$  e considerando que as estimativas dos parâmetros utilizados dados até  $t$  para  $t_0 - 49 \leq t \leq t_0$  são aproximadamente constantes, utilizamos as estimativas encontradas em 1) para calcular as estimativas filtradas da previsão um-passo à frente da volatilidade:  $Prev_{t/t-1}$  via Filtro de Kalman no período  $\{t_0 - 49, t_0 - 48, \dots, t_0\}$ .
- 3) Calculamos os intervalos de confiança 95% para os retornos. De forma geral os i.c.  $(1-2\alpha)$  são dados por:  $(\mu_t - t_{v,\alpha} \sqrt{Prev_{t/t-1}}, \mu_t + t_{v,\alpha} \sqrt{Prev_{t/t-1}})$ , onde  $t_{v,\alpha}$  corresponde à distribuição *t-Student* padrão (se  $\alpha = 0,025$  então  $t_{7,0,025} = 2,365$ ) sendo:  

$$Prev_{t/t-1} = \exp[\hat{\Gamma}_{f,t-1} + 1,27 + \psi(\hat{v}/2) - \log(\hat{v}/2) + \hat{\gamma} \hat{h}_{-f_{t-1}}] \frac{\hat{v}}{\hat{v} - 2}.$$

Já que as estimativas filtrada e suavizada do nível  $k$  apresentam valores similares a partir de  $t=1250$  aproximadamente (ver Figura 4.1), o modelo pode ser re-estimado considerando essa informação como informação básica. Dessa forma, temos os seguintes resultados:

**Tabela 4.3**  
*Ações: estabilidade das estimativas e previsão um-passo à frente ex-ante*

	Período	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_\eta^2$	$\rho$	$\gamma$	$v$	$nf$	$nf1$
1	1-1250	4,9718	0,0434	0,0087	0,9680	49	1	2
2	1-1300	5,0953	0,0464	0,0091	0,9667	13	3	3
3	1-1350	5,1022	0,0448	0,0088	0,9657	13	1	2
4	1-1400	5,1376	0,0430	0,0084	0,9688	11	1	1
5	1-1450	5,3647	0,0404	0,0075	0,9717	5,5	0	0
6	1-1500	5,2790	0,0434	0,0082	0,9680	7	0	0
7	1-1550	5,6687	0,0439	0,0077	0,9676	3,5	1	2
	1-1598	5,0341	0,0461	0,0092	0,9656	20	0	0

*nf*: número de retornos fora dos intervalos de confiança 95% utilizando toda a amostra  
*nf1*: número de retornos fora dos intervalos de confiança 95% utilizando parte da amostra

Em cada período da Tabela 4.3 ao comparar as duas últimas colunas, podemos constatar que não existe muita diferença entre os resultados, isto é, em termos de previsão um-passo à frente *ex-ante*, os intervalos de confiança (i.c.) construídos com as estimativas obtidas

utilizando a série completa são similares aos obtidos utilizando as estimativas de cada período. Por sua vez, considerando a informação completa da amostra (1598 retornos), o número de retornos fora dos i.c 95% a partir de  $t=1200$  é igual a 10, significando uma percentagem de 2,60%. Da mesma forma, considerando os dados da coluna  $\eta_f$ , a soma do número de retornos fora dos i.c. nos períodos considerados é igual a 7 (uma percentagem de 1,80%), consequentemente, ambas percentagens são menores do que a esperada (5%).

A análise dos resultados da Tabela 4.3 mostra que, tanto as estimativas da persistência  $\gamma$  quanto as de  $\sigma_{\eta}^2$ , apresentam estimativas estáveis e muito próximos enquanto que existem diferenças entre as estimativas de  $\sigma_{\xi}^2$ , que podem ser melhor observadas na variabilidade apresentada pelas estimativas de  $v$ .

Nas Figuras 4.6b, 4.7b e 4.8b apresentamos o gráfico dos intervalos de confiança 95% (i.c) em 3 períodos de volatilidades média e alta que correspondem a alguns dos acontecimentos importantes na história da série. Nos três gráficos é observado que os i.c não incluem os retornos extremos. Isso acontece no período da *Queda da Ministra Zélia Cardoso de Mello* (Figura 4.6b) no dia 13/01/92, no período da *Mudança da Banda Cambial* (4.7b) em 10/03/95 e para o *Impeachment Collor* (Figura 4.8b) em 29/06/92.

#### *Comparação das Volatilidades Suavizada e Filtrada*

As estimativas das volatilidades suavizada e filtrada apresentam diferenças que podem ser bastante grandes, tal como pode ser observado na Figura 4.9 (essas estão representadas pelos pontos bem acima ou bem abaixo da linha reta). Adicionalmente, a variabilidade da volatilidade suavizada é menor quando comparada à variabilidade apresentada pela volatilidade filtrada, sendo essa última mais sensível aos aumentos e diminuições dos quadrados dos retornos. Esse comportamento é constatado ao observar a Figuras 4.4 e, com maior detalhe, nas Figuras 4.5a, 4.6a, 4.7a, 4.8a.

Por outro lado, na Figura 4.4 vemos que as estimativas das volatilidades filtrada e suavizada são pouco influenciadas pelos valores extremos nos quadrados dos retornos. Por exemplo, na Figura 4.6a apresentamos o período onde acontece o maior retorno da série (pelo menos 4 vezes maior aos retornos nesse período), sendo observado que os valores de ambas as

volatilidades não são afetadas por essa observação extrema. O mesmo acontece em outro período que possui um valor extremo, ilustrado na Figura 4.7a. Esse comportamento é diferente ao apresentado pela volatilidade nos modelos ARCH, que são muito sensíveis as observações extremas. Finalmente, como é observado na Figura 4.4, as maiores estimativas da volatilidade ocorrem nos períodos dos maiores conglomerados de volatilidade, por exemplo, na vizinhança de  $t=300$ ,  $t=500$ ,  $t=850$  e  $t=1500$  e não nos períodos com presença de observações extremas, nesse sentido, as volatilidades são semelhantes ao EGARCH(2,0) analisado na seção 3.7.1.

*Ajuste dos Modelos AR(1)-SV Não-estacionários*

Considerando que nos modelos estacionários ajustados anteriormente (Tabela 4.1) a estimativa da persistência da volatilidade apresenta valores muito próximos de 1, são ajustados as versões não-estacionárias das especificações (4.8.5) supondo que as estimativas para a média condicional são as mesmas. A seguir apresentamos as outras estimativas<sup>14</sup>:

**Tabela 4.4**  
**Ações: estimativas dos modelos AR(1)-SV não-estacionários**

	Modelo I	Modelo II	Modelo III	Modelo IV
p.e.v	5,4373	5,4449	5,4408	5,4490
AIC	1,6983	1,6984	1,6977	1,6979
BIC	1,7117	1,7085	1,7078	1,7047
w	0,1444 (0,3112) <sup>a</sup>	0,1561 (0,2733)	-	-
f	0,4282 (0,1348)	-	0,4441 (0,1204)	-
$\hat{\sigma}_\xi^2$	5,1146	5,1196	5,1153	5,1210
$\hat{\sigma}_\eta^2$	0,0203	0,0202	0,0203	0,0201
$\hat{q}$	0,0040	0,0039	0,0040	0,0039

<sup>a</sup> valores P das estimativas

<sup>14</sup> Em todos os casos o processo converge *fortemente* em 3 iterações.

Seguindo os resultados da tabela acima, percebemos que nos 4 modelos os valores de  $\hat{\sigma}_\varepsilon^2$  são maiores a 4,93, o valor verdadeiro se a estimação fosse realizada com restrição. Por sua vez, os baixos valores de  $q$  indicam que a componente  $h$  e portanto, a volatilidade, apresenta muito pouca variabilidade, sendo mais suave quando comparada com o caso estacionário. Adicionalmente, os valores do p.e.v são menores no caso estacionário (sendo os valores dos AIC e BIC próximos em ambos modelos). Por esses motivos, o modelo estacionário é mais adequado que o modelo não-estacionário .

### *Conclusões*

Dos análises realizados podemos extrair as seguintes conclusões:

- 1) As estatísticas correspondentes à correlação serial e a segunda feira são significativas.
- 2) As volatilidades estimadas suavizada e filtrada do modelo AR(1)-SV estacionário são capazes de imitar os movimentos de aumentos e diminuições dos quadrados dos retornos.
- 3) A persistência do modelo AR(1)-SV estacionário é grande; 0,966 significando uma meia-vida de 20,04 dias.
- 4) Os intervalos de confiança 95% (i.c) para os retornos construídos com a volatilidade suavizada excluem 3,44% dos retornos sendo esse resultado satisfatório.
- 5) As estimativas dos parâmetros do modelo AR(1)-SV estacionário apresentam estabilidade ao serem estimados utilizando tamanhos de amostra diferentes.
- 6) Considerando as últimas 350 observações, os (i.c) 95% para a previsão um-passo a frente *ex-ante* excluem uma percentagem de retornos menor do que a esperada.
- 7) As volatilidades estimadas filtrada e suavizada do modelo AR(1)-SV estacionário são pouco influenciadas pelas observações extremas.
- 8) Em ambos os modelos, AR(1)-SV estacionário e não estacionário, as estimativas correspondentes aos parâmetros  $w$  (do *Final de Semana*) e  $f$  (dos *Feriados*) na volatilidade, não foram significativas, nos 4 modelos considerados.
- 9) O Modelo AR(1)-SV estacionário é preferível ao modelo não estacionário devido a que apresenta menor p.e.v., e a volatilidade é menos suave.

### 4.8.2 Taxa de Câmbio

Para esta série é proposto o modelo (4.8.1) considerando o *Efeito Calendário* na média condicional:

$$\mu_t = E[Y_t / I_t] = \text{Seg}D_1 + \text{Ter}D_2 + \text{Quar}D_3 + \text{Quin}D_4 + \text{Sex}D_5$$

Através do método de mínimos quadrados ordinários obtemos:

$$\hat{\mu}_t = 0,00048D_1 + 0,00011D_2 - 0,00010D_3 + 0,00079D_4 - 0,00010D_5 .$$

(0,00040)      (0,00040)      (0,00041)      (0,00040)      (0,00040)

Como comentado no Capítulo 2, existe *Efeito Calendário* na Quinta Feira (valor  $P=0,048$ ). A seguir, a volatilidade é estimada a partir dos resíduos  $\hat{Y}_t^*$ , considerando os casos estacionário e não-estacionário<sup>15</sup>.

#### *Estimação do Modelo AR(1)-SV Estacionário*

Utilizando o pacote *STAMP 5.0* ajustamos o Modelo IV<sup>16</sup> definido em (4.8.5). O processo de estimação realizado através do pacote *STAMP 5.0* converge em 32 iterações (convergência  *muito forte*), sendo estimado o seguinte modelo:

$$\hat{h}_t = 0,975\hat{h}_{t-1} + \hat{\eta}_t$$

com  $\sigma_\zeta^2 = 5,2687$ ,  $\sigma_\eta^2 = 0,0108$ , em conseqüência  $\hat{q} = 0,0020$ .

Como observado, esse modelo apresenta alta persistência (0,975) o qual equivale a uma meia-vida de 27,38. Por outro lado, através da relação  $\hat{\sigma}_\zeta^2 = 4,93 + \psi'(v/2)$  encontramos que  $\hat{v} = 7$ .

<sup>15</sup> Na análise do modelo seria conveniente considerar só uma variável *dummy* com valor um na quinta feira e zero nos outros dias, mas, com o objetivo de comparar as estimativas com os modelos ARCH, são consideradas as 5 variáveis *dummy*.

<sup>16</sup>As estimativas nos outros três modelos definidos em (4.8.5) são pouco confiáveis pois as convergências observadas no processo de estimação foram  *muito fracas*.

Na Figura 4.10 apresentamos o gráfico das estimativas filtrada e suavizada do componente  $T_t$ , as quais, apresentam valores similares no final da série (após a linha vertical pontilhada). Embora não mostrados os gráficos, as tendências  $h_f$  e  $h_s$ , possuem variabilidade baixa (variâncias iguais 0,058 e 0,11 respectivamente) quando comparadas com o componente irregular filtrado (variância igual a 5,42). O gráfico das volatilidades estimadas filtrada e suavizada é apresentado na Figura 4.11 Observe-se que ambas são capazes de imitar os movimentos de aumentos e diminuições nos quadrados dos retornos. Nas primeiras 20 observações aproximadamente, a volatilidade filtrada apresenta valores muito altos que são naturais na fase inicial do Filtro de Kalman.

*Adequabilidade do AR(1)-SV Estacionário*

Na análise dos quase resíduos padronizados (QRP),  $\tilde{v}_t$  obtemos que os valores da média, -0,035 e da variância, 0,999 são satisfatórios. Por sua vez, o coeficiente de simetria é igual a -1,16 e o excesso de curtose, 2,17; indicando que a distribuição dos QRP apresenta cauda à esquerda (ver Figura 4.12a). Os gráficos de CUSUM e CUSUMSQ nas Figuras 4.12b e 4.12c, apresentam comportamento satisfatório e a estimativa  $H(588)$  é igual 0,88536 com valor- $P=0,07$  mostrando que não existe grande evidência para rejeitar a hipótese de homoscedasticidade e/ou ausência de mudança estrutural nos QRP.

Ao examinar a função de autocorrelação dos QRP (Figura 4.12d) percebemos que não existe evidência forte de presença de correlação serial. Os valores P do Teste Box-Ljung, na tabela a seguir, apoiam essa conclusão:

**Tabela 4.5**  
**Taxa de Câmbio: Valores P do Teste Box-Ljung nos QRP do modelo AR(1)-SV Estacionário**

<i>Lag</i>	6	12	18	24	30	36	42
<i>Valor P</i>	0,346	0,067	0,314	0,276	0,130	0,102	0,025

Os baixos valores nos *lags* 12 e 42 são devidos às grandes autocorrelações nos *lags* 10 e 42 (Figura 4.8.2.5d). Embora não apresentado, o comportamento dos QRPA do componente irregular é satisfatório.

Por outro lado, os intervalos de confiança 95% para os retornos construídos utilizando a volatilidade suavizada excluem 48 em 1767 retornos (3,44%) sendo esse resultado bom.

Com relação à análise da estabilidade das estimativas dos parâmetros do modelo e a previsão um-passo à frente *ex-ante*, utilizamos a mesma metodologia desenvolvida na série das *Ações*. Para a série *Taxa de Câmbio*, é difícil apresentar resultados conclusivos já que o processo de estimação revela-se muito instável sendo necessário retirar observações iniciais em cada período considerado para obter estimativas que satisfizerem o critério de *convergência muito forte* no pacote *STAMP 5.0*. Todavia, em três períodos não alcançamos esse requerimento em várias tentativas retirando observações iniciais. Os resultados são apresentados na Tabela 4.6<sup>17</sup>.

Analisando os resultados da Tabela 4.6 podemos observar que, de maneira geral, as estimativas dos parâmetros nos períodos considerados são similares e, na parte final da série, a partir de 1600, as estimativas são muito próximas, mostrando a existência de estabilidade nas estimativas dos parâmetros no final da série. Por outro lado, o número de retornos fora dos intervalos de confiança 95% é o mesmo se trabalhamos com toda a série ou em cada período. Isso é devido a que as estimativas das componentes são similares quando as estimamos utilizando toda a série ou nos períodos considerados. Assim, se definimos  $d = \text{Prev}_p - \text{Prev}$ , onde  $\text{Prev}_p$  é a previsão considerando a fração da amostra e  $\text{Prev}$  a previsão com toda a amostra, a média e variância de  $d$  são respectivamente, inferiores a  $3 \times 10^{-6}$  e  $5 \times 10^{-12}$ . Considerando que a variância da componente filtrada na volatilidade é 0,058 podemos concluir que a diferença é desprezível. Portanto, o número de retornos fora dos intervalos de confiança 95% a partir de  $t=1200$  é menor ao esperado, sendo a percentagem igual a 2,29 % (13 em 568).

---

<sup>17</sup> Os resíduos  $Y_t^*$  em (4.8.1a) sob os quais é modelada a volatilidade são calculados estimando a média condicional utilizando todos os retornos (1767), de forma que a média condicional não é re-estimada em cada período.

**Tabela 4.6**  
**Taxa de Câmbio: estabilidade das estimativas e previsão um-passo á frente *ex-ante***

	Período	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_\eta^2$	q	$\gamma$	v	nf	nfl
1	2-1200	5,3845	0,0130	0,0024	0,9781	5,5	0	0
2	5-1250	5,3347	0,0157	0,0029	0,9728	6	0	0
3	2-1300	5,2674	0,0164	0,0031	0,9711	7	8	8
4	2-1350	5,1769	0,0152	0,0029	0,9725	9	0	0
5	1-1400	5,2053	0,0174	0,0033	0,9696	8	0	0
6	-1450*							
7	-1500*							
8	-1550*							
9	3-1600	5,2600	0,0126	0,0024	0,9745	7	0	0
10	2-1650	5,2799	0,0109	0,0020	0,9772	7	0	0
11	4-1700	5,2678	0,0112	0,0021	0,9758	7	0	0
12	7-1750	5,2574	0,0104	0,0020	0,9761	7	0	0
	Todo	5,2687	0,0108	0,0020	0,9753	7	13	

nf: número de retornos fora dos intervalos de confiança 95% utilizando toda a amostra

nfl: número de retornos fora dos intervalos de confiança 95% utilizando parte da amostra

\* não foi alcançada convergência

Nas Figuras 4.13b e 4.14b são apresentados os gráficos dos i.c. 95% para os retornos considerando dois períodos diferentes. No primeiro período (Figura 4.13b), que corresponde à época na qual acontece o maior retorno da série (no dia 23/09/85) e apresenta volatilidades média e alta, observamos que só um retorno (1,25%) está fora dos i.c. O segundo período (Figura 4.14b) é basicamente de volatilidade alta (volatilidade média até t=285) e corresponde à época onde ocorre um conglomerado de volatilidade com maiores valores dos retornos. Nesse período observamos que existe uma percentagem razoável de retornos fora dos i.c., 4%.

#### *Comparação das Volatilidades Suavizada e Filtrada*

Como pode ser observado na Figura 4.15, as estimativas das volatilidades suavizada e filtrada apresentam diferenças que são consideráveis para algumas observações. A volatilidade filtrada é mais sensível aos aumentos e diminuições dos quadrados dos retornos enquanto que a

volatilidade suavizada apresenta menor variabilidade. Esse comportamento pode ser observado na Figura 4.11 e com maior detalhe nas Figuras 4.13 e 4.14.

Por outro lado, ambas as volatilidades são pouco influenciadas pelas observações extremas nos quadrados dos retornos. Por exemplo, na Figura 4.13a observamos que as volatilidades não aumentam abruptamente nem bastante devido à observação extrema no 23/09/85, apesar de esse retorno apresentar um valor quatro vezes maior aos outros no mesmo período. Finalmente, é interessante notar que os maiores valores da volatilidade estimada estão no período 11/02/85 - 22/05/85 (Figura 4.16) no qual está o maior conglomerado de volatilidade com retornos grandes e não em 23/09/85 (Figura 4.13) no qual está o maior retorno da série.

#### *Ajuste do Modelo AR(1)-SV Não-estacionário*

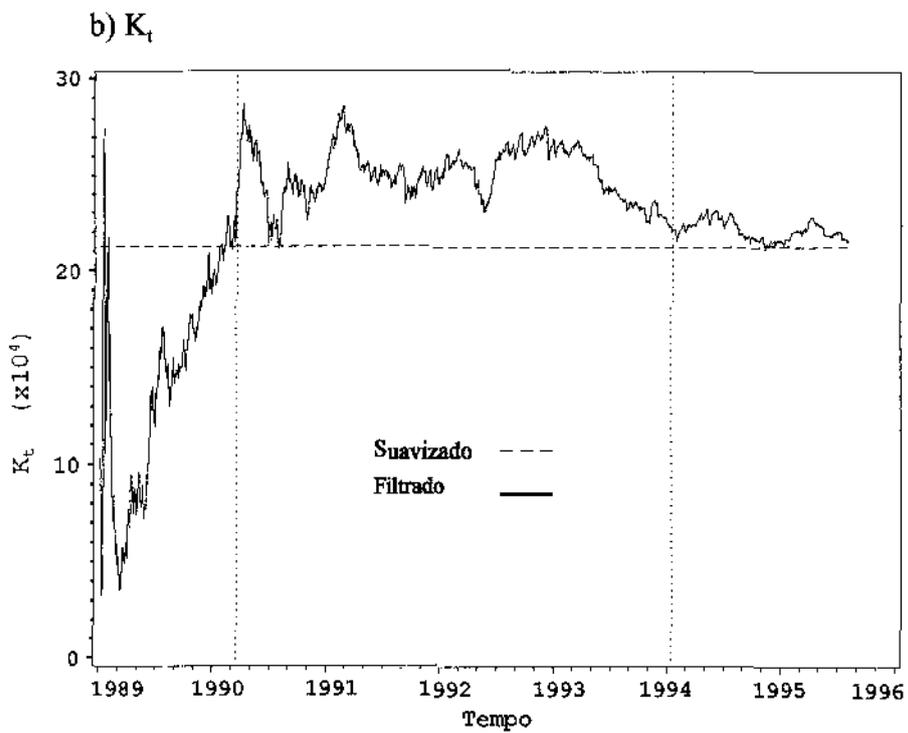
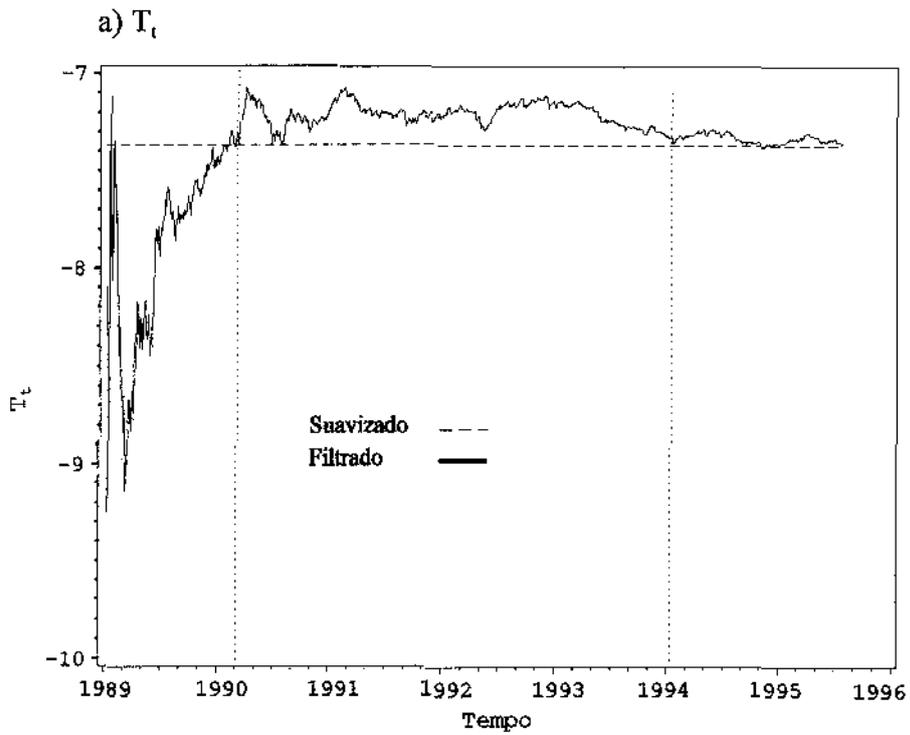
Já que a persistência da volatilidade no AR(1)-SV estacionário é bastante grande, tentamos ajustar a versão não-estacionária considerando as especificações (4.8.10). Em todos os modelos, o processo de estimação apresenta *convergência muito forte* em 5 iterações, mas analisando o gráfico da f.a.c. dos QRP (não mostrado) observamos presença de correlação serial e do Teste Box-Ljung temos evidência fraca a favor da hipótese de ruído branco. Portanto esses modelos não são adequados.

#### *Conclusões*

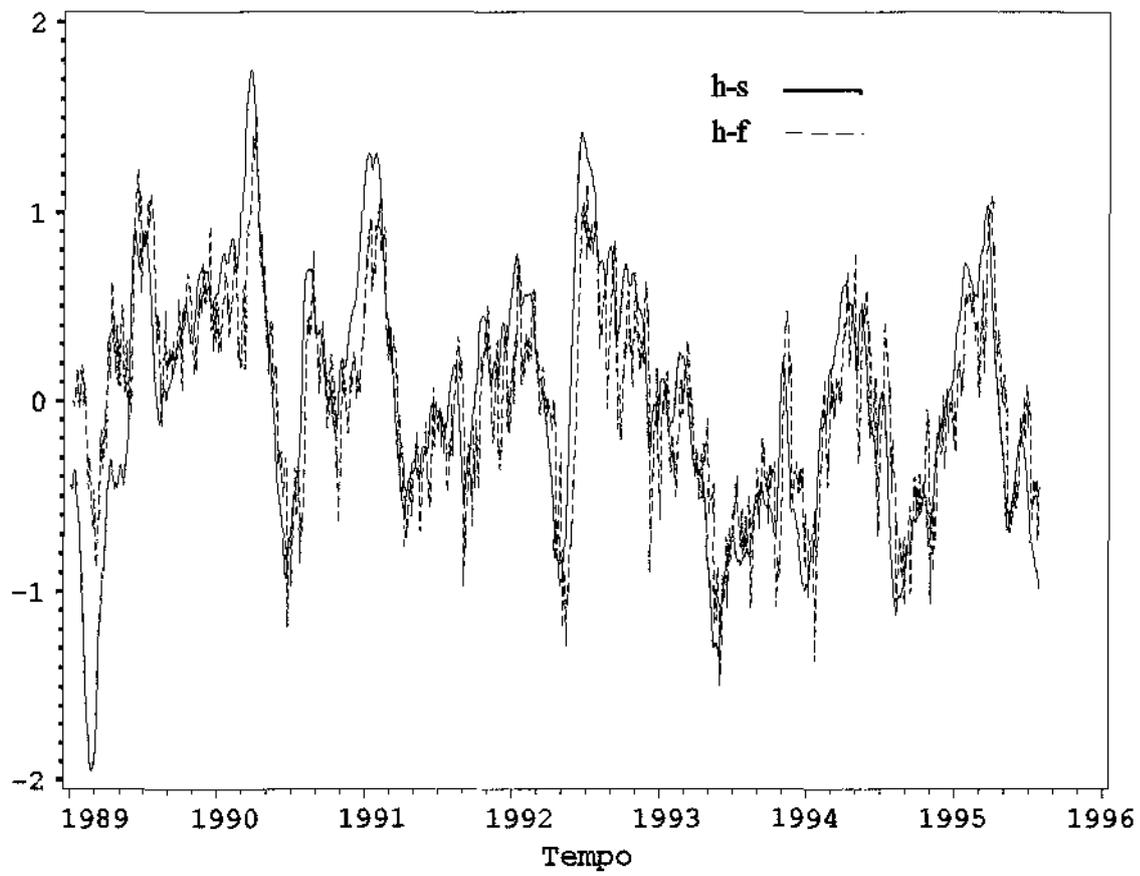
De todos os modelos ajustados, o único que satisfaz os requisitos de adequabilidade do modelo é o modelo estacionário, Modelo IV. Este apresenta as seguintes características:

- 1) As volatilidades estimadas suavizada e filtrada conseguem acompanhar os conglomerados de volatilidade dos retornos.
- 2) O modelo AR(1)-SV estacionário apresenta grande persistência; 0,975 significando uma meia-vida de 27,35 dias.
- 3) Os intervalos de confiança 95% para os retornos construídos com a volatilidade suavizada excluem 3,44% dos retornos sendo esse resultado satisfatório.

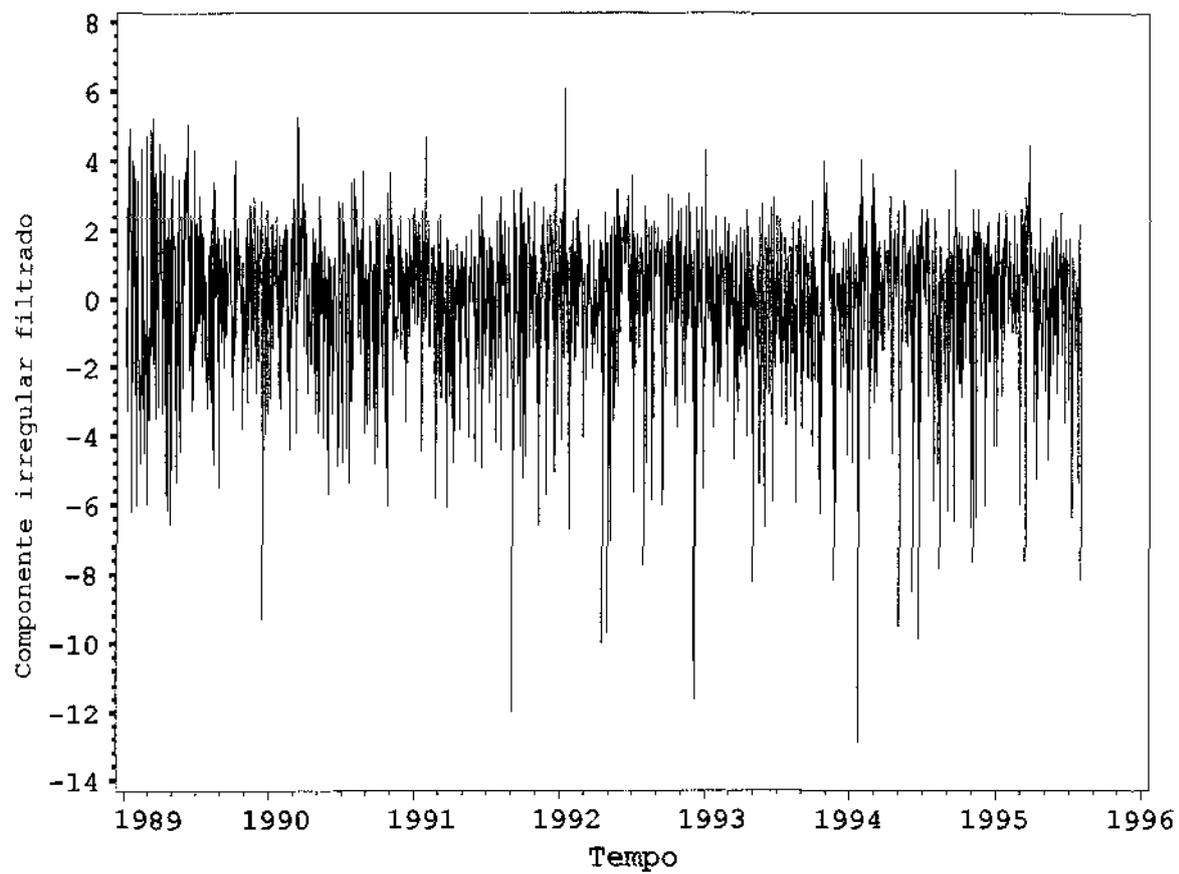
- 4) Para vários períodos da série, as estimativas apresentam estabilidade, especialmente na parte final da série. Por sua vez, há também períodos nos quais não é possível obter os requisitos de convergência para o modelo.
- 5) Considerando as últimas 567 observações, os i.c 95% para a previsão um-passo a frente *ex-ante* excluem uma percentagem de retornos menor do que a esperada.
- 6) As volatilidades filtrada e suavizada do modelo AR(1)-SV estacionário são pouco influenciadas pelas observações extremas, e por sua vez, os maiores valores de volatilidade ocorrem na presença de grandes conglomerados com valores altos dos retornos.



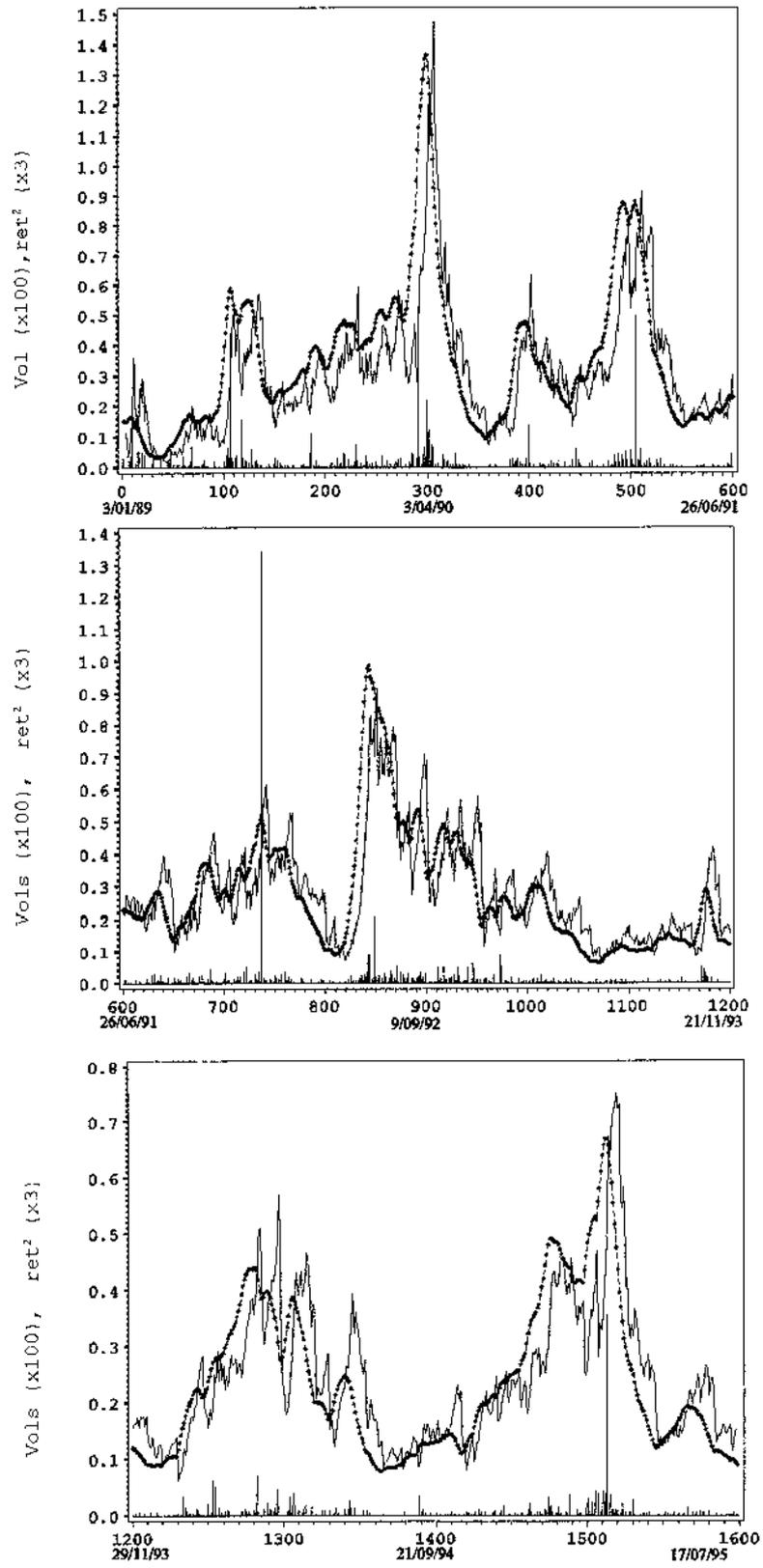
**Figura 4.1: Componentes estimados  $T_t$  e  $K_t$  nas *Ações***



**Figura 4.2: Componentes estimados  $h_{f_t}$   $h_{s_t}$  nas Ações**



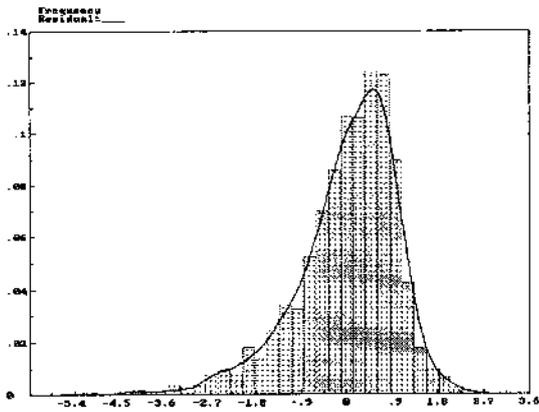
**Figura 4.3: Componente estimado irregular filtrado nas Ações**



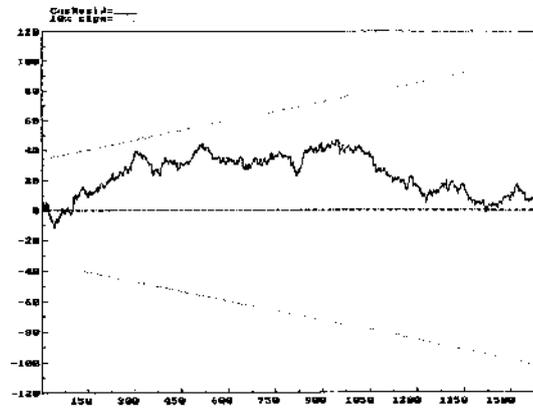
**Figura 4.4: Volatilidades estimadas filtrada e suavizada nas Ações**

**Legenda:** filtrada-----, suavizada ———  
 vols=volatilidades, ret²=quadrados dos retornos

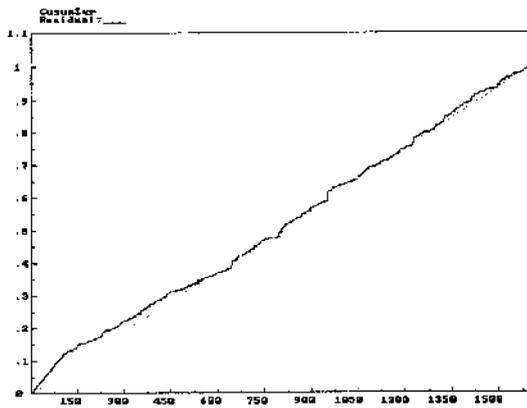
a)



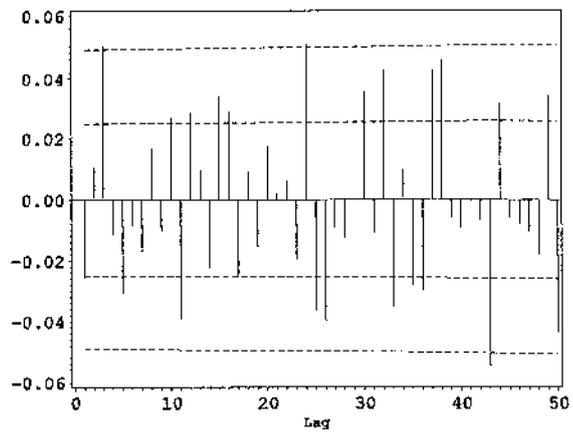
b)



c)



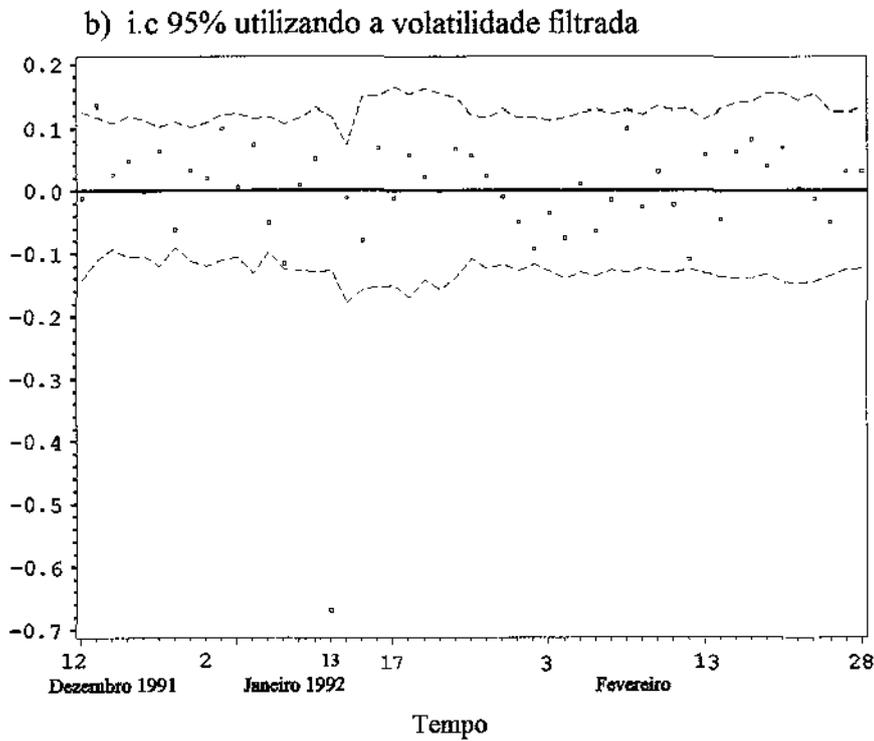
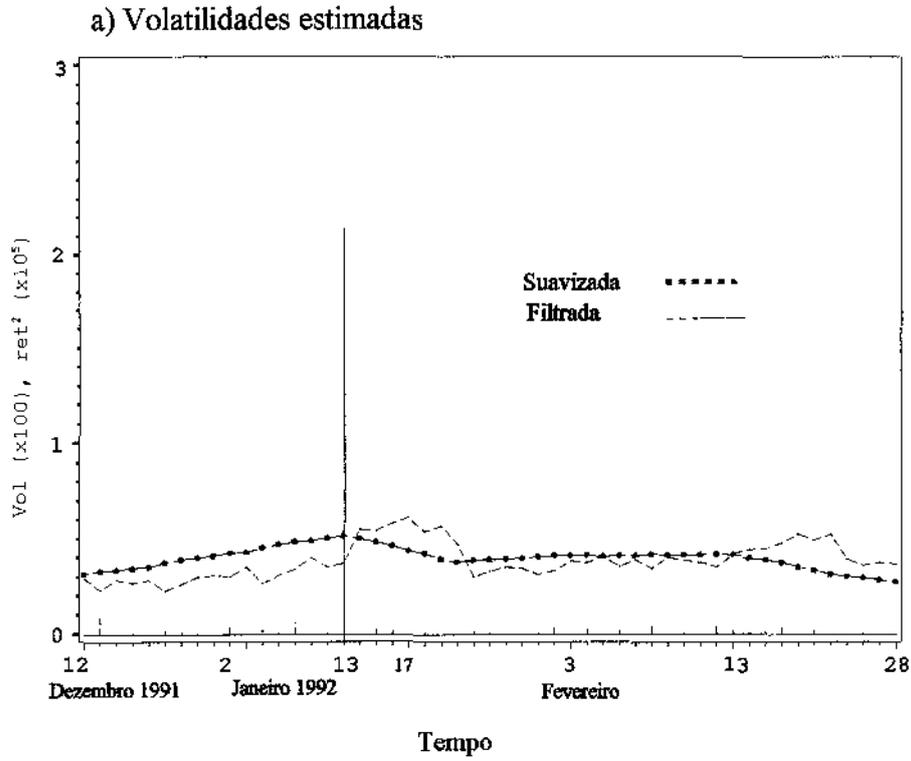
d)



**Figura 4.5: Análise dos Quase-resíduos padronizados nas Ações**

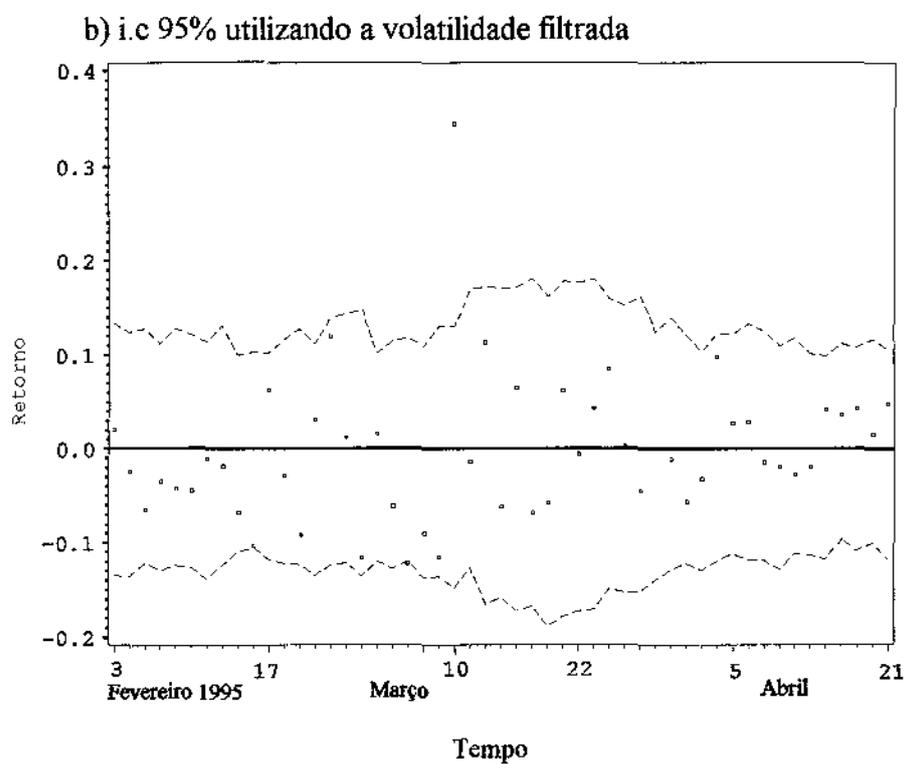
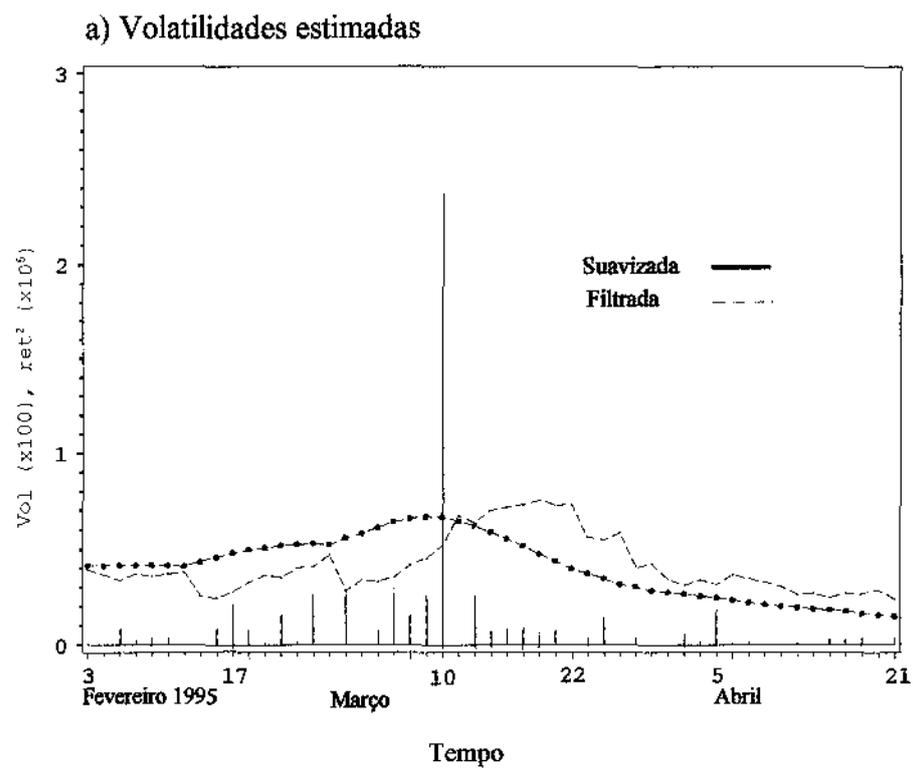
a) Histograma e densidade ajustada, b) CUSUM

c) CUSUMSq, d) f.a.c.



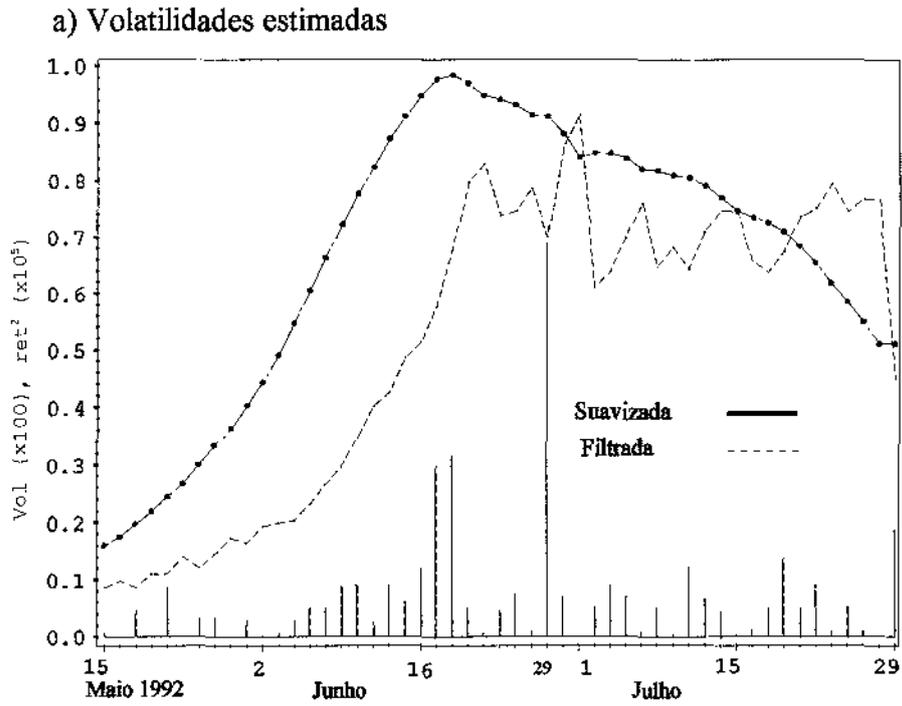
**Figura 4.6: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período da Queda da Ministra Zélia Cardoso de Mello**

**Legenda:** vol=volatilidade, ret<sup>2</sup>= quadrados dos retornos

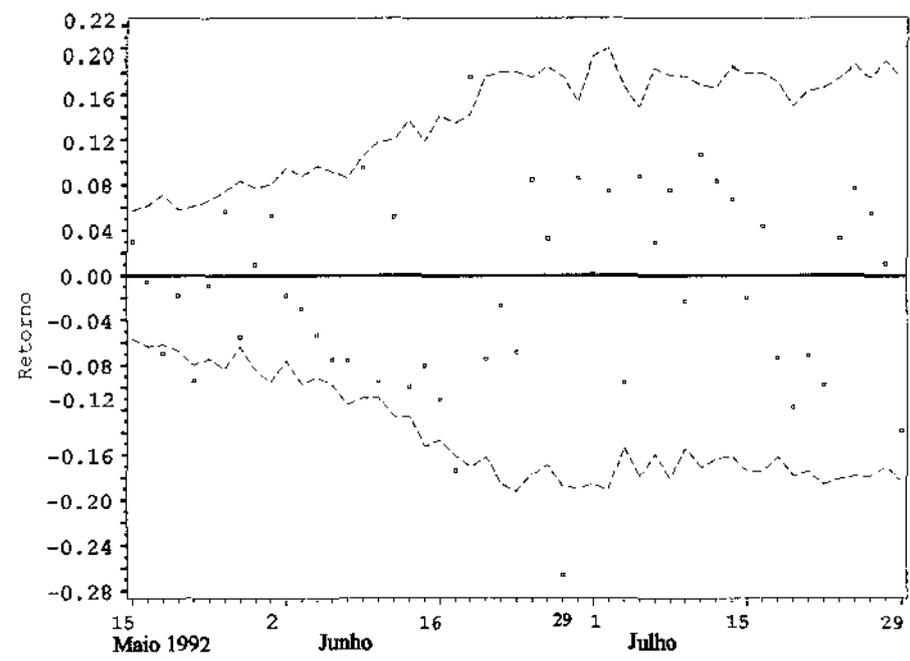


**Figura 4.7: Volatilidades estimadas e i.c 95% para os retornos no período da Mudança da Banda Cambial**

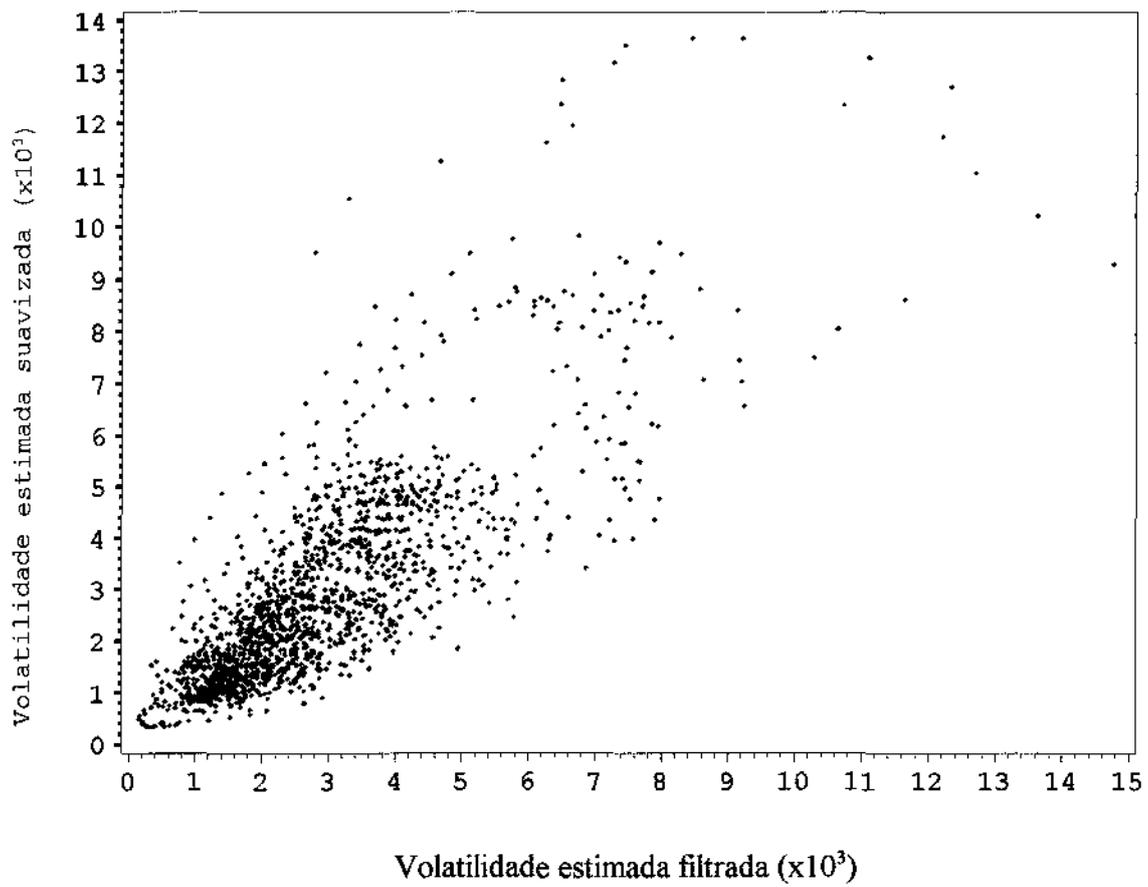
**Legenda:** vol=volatilidades,  $ret^2$ = quadrados dos retornos



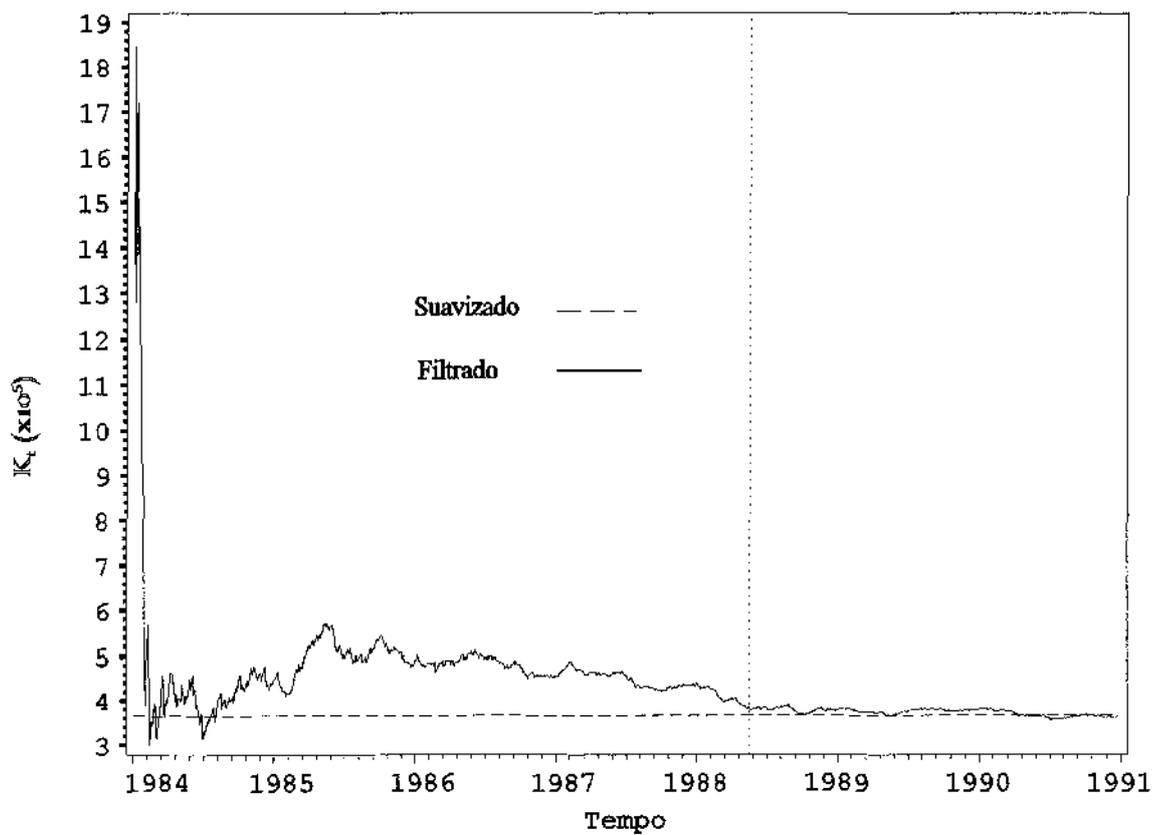
b) i.c 95% utilizando a volatilidade filtrada



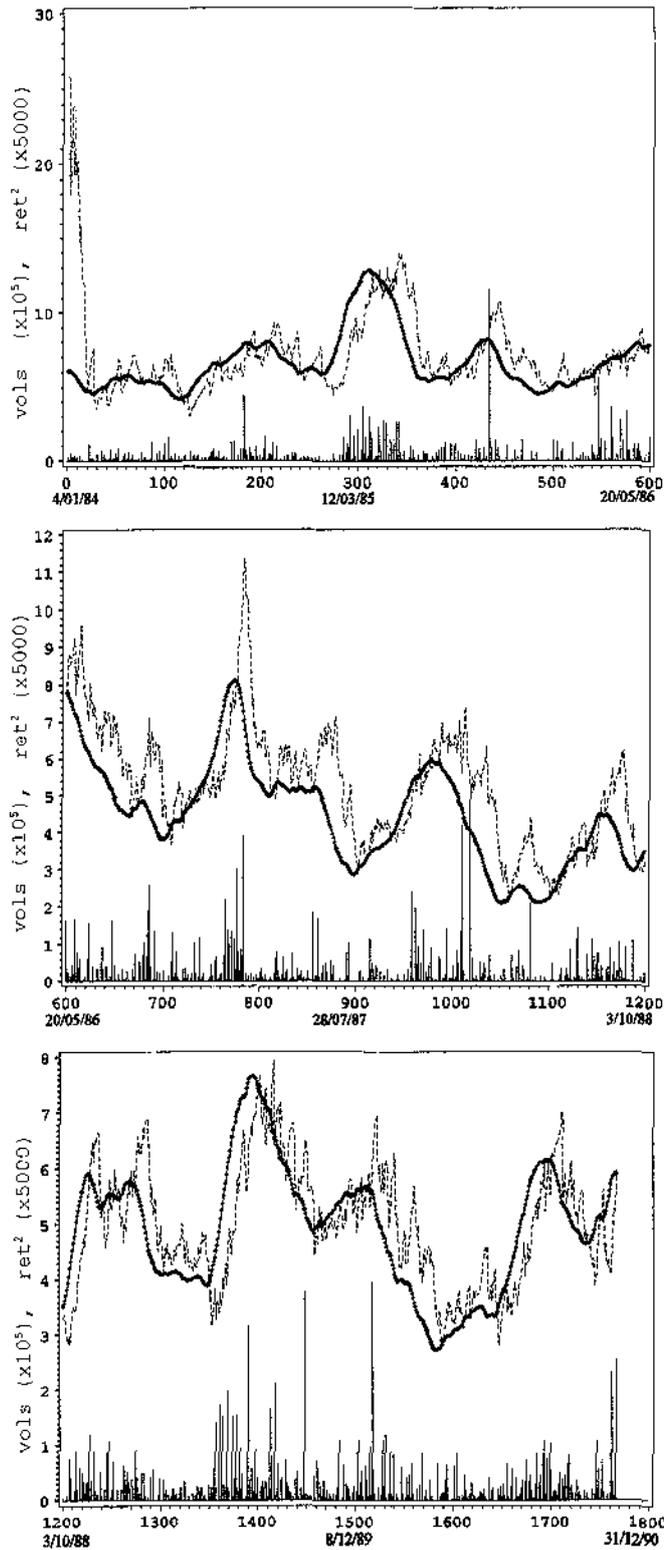
**Figura 4.8: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período do *Impeachment Collor***  
**Legenda:** vol=volatilidade,  $ret^2$ =quadrados dos retornos



**Figura 4.9: Comparação das volatilidades estimadas filtrada e suavizada nas *Ações***



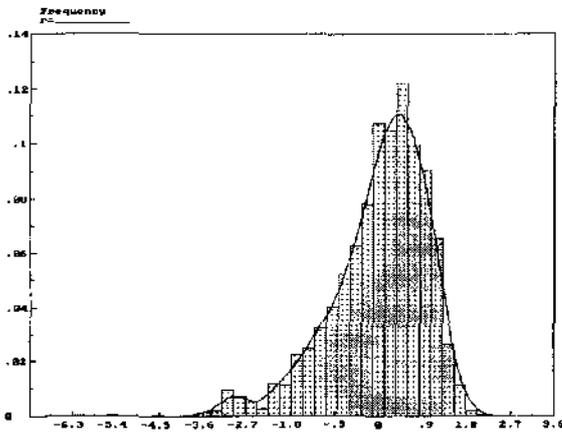
**Figura 4.10: Gráfico do componente  $K_t$  na *Taxa de Câmbio***



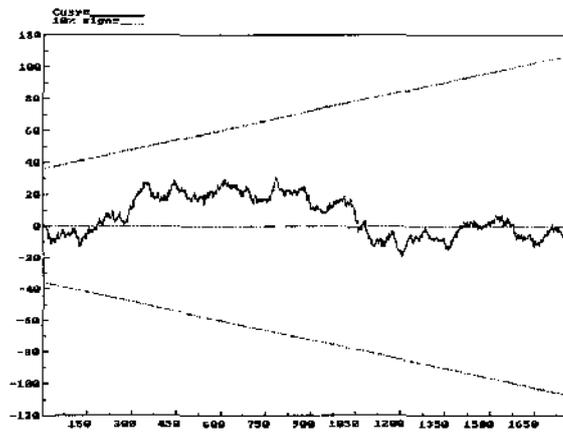
**Figura 4.11: Volatilidades estimadas suavizada e filtrada na Taxa de Câmbio**

**Legenda:** suavizada ————, filtrada - - - -  
 vols=volatilidades, ret<sup>2</sup>=quadrados dos retornos

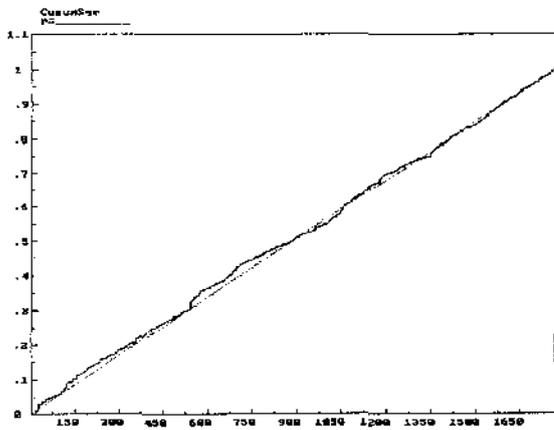
a)



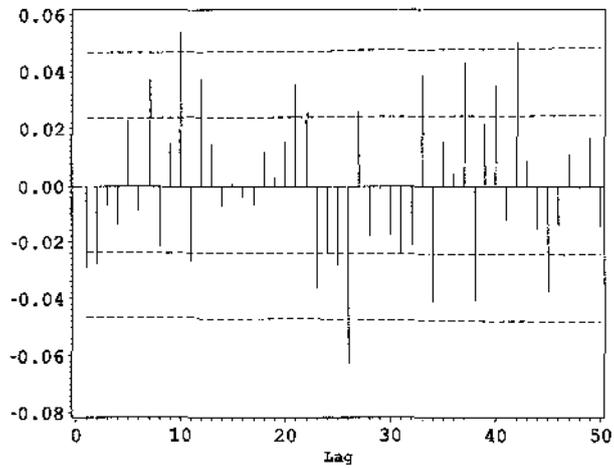
b)



c)



d)

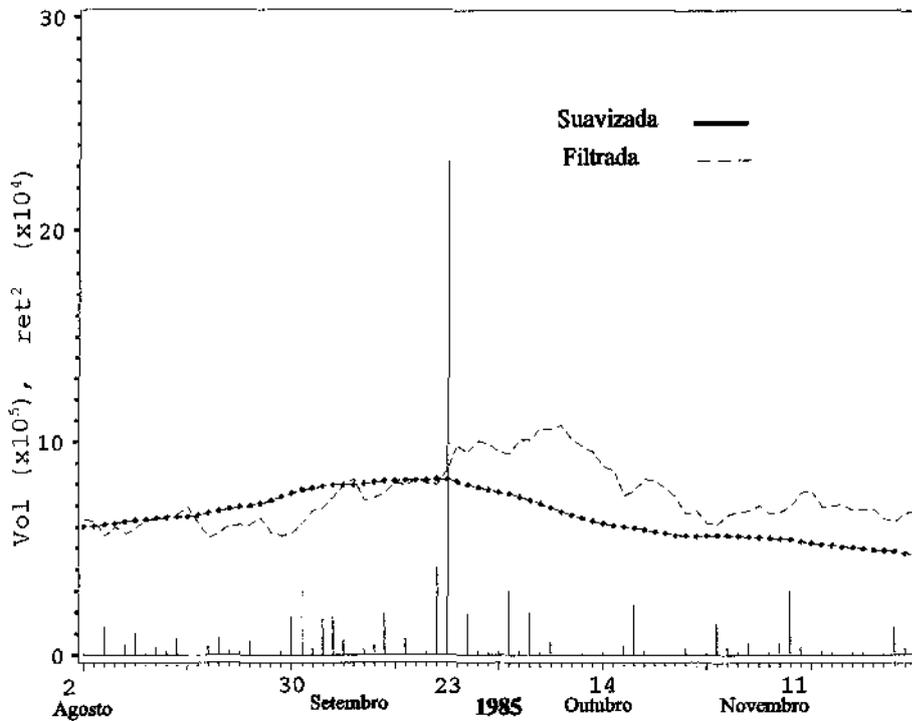


**Figura 4.12: Análise dos Quase-resíduos padronizados na Taxa de Câmbio**

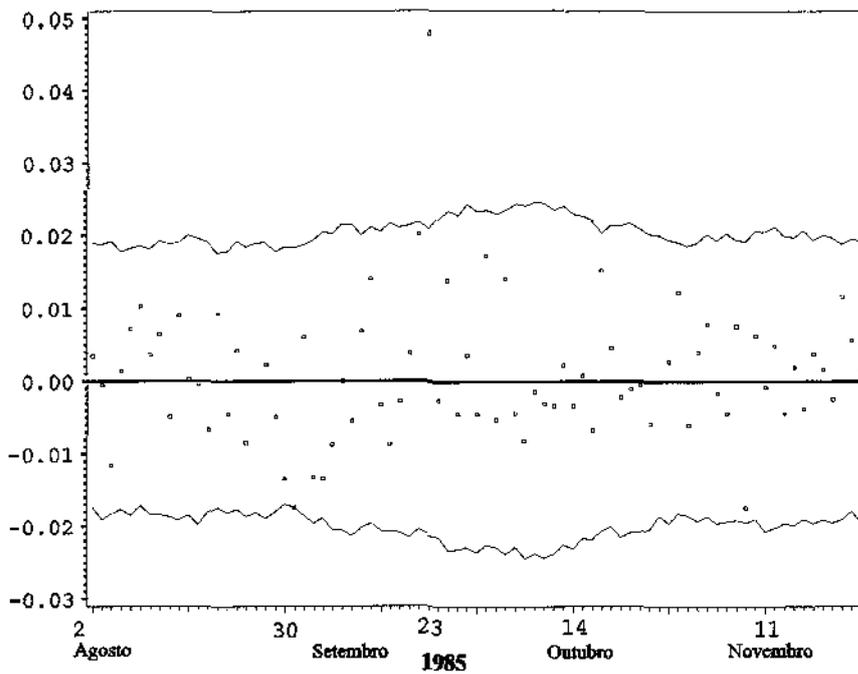
a) histograma e densidade ajustada, b) CUSUM,

c) CUSUMSq, d) f.a.c.

a) Volatilidades estimadas



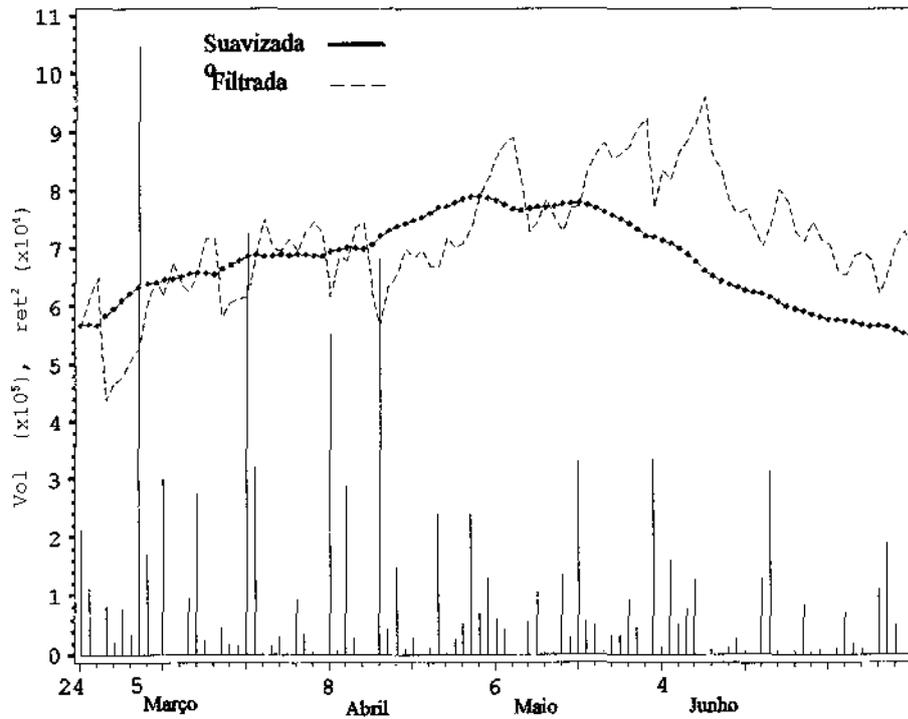
b) i.c 95% para os retornos



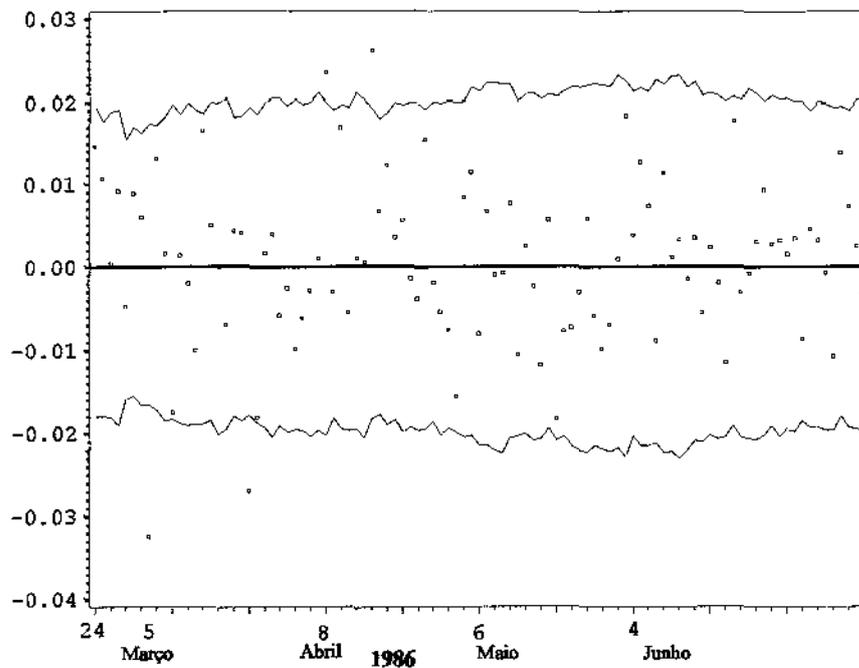
**Figura 4.13: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período da ocorrência do maior retorno na Taxa de Câmbio**

**Legenda:** vol=volatilidade,  $ret^2$ =quadrados dos retornos

a) Volatilidades estimadas

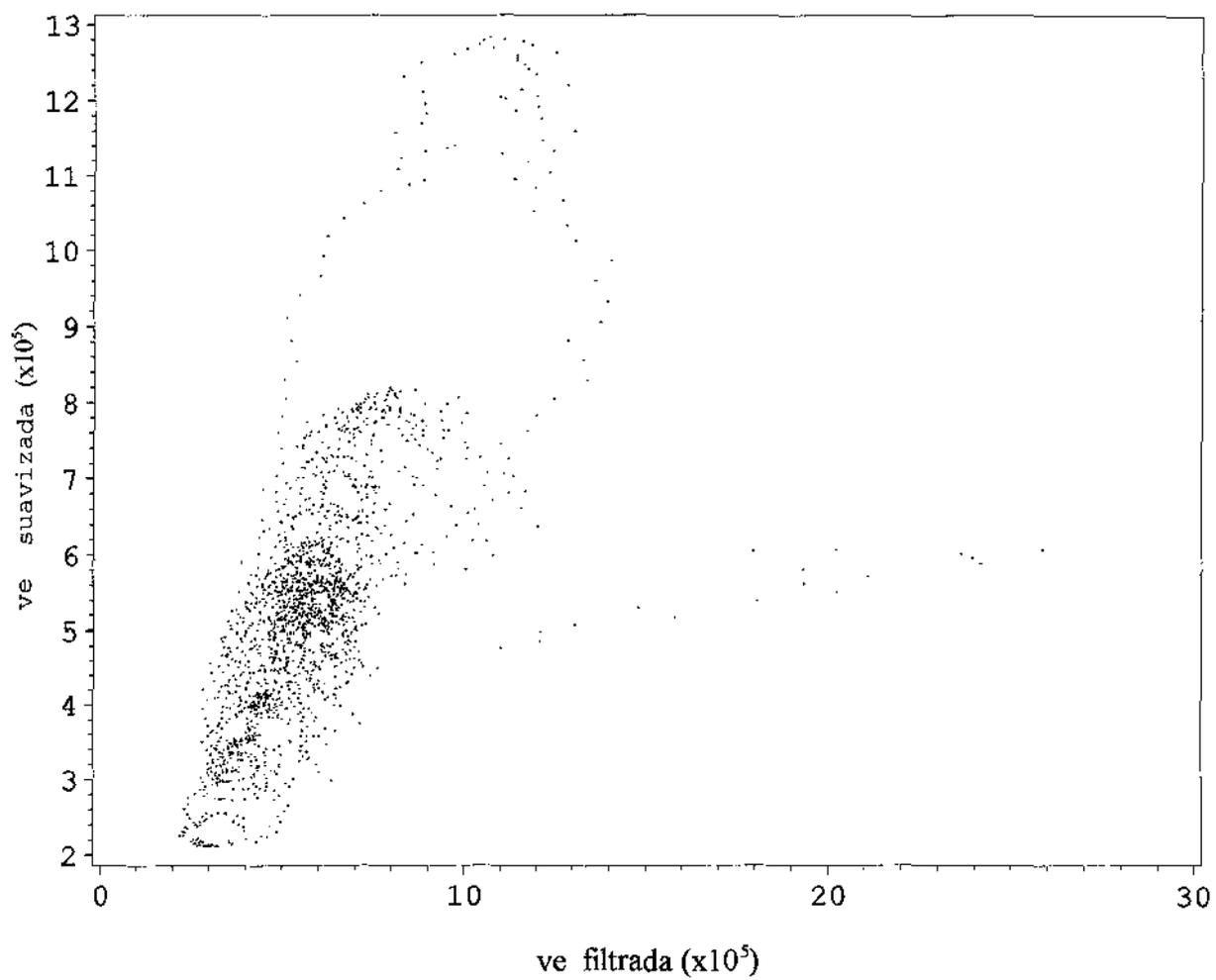


b) i.c 95% para os retornos



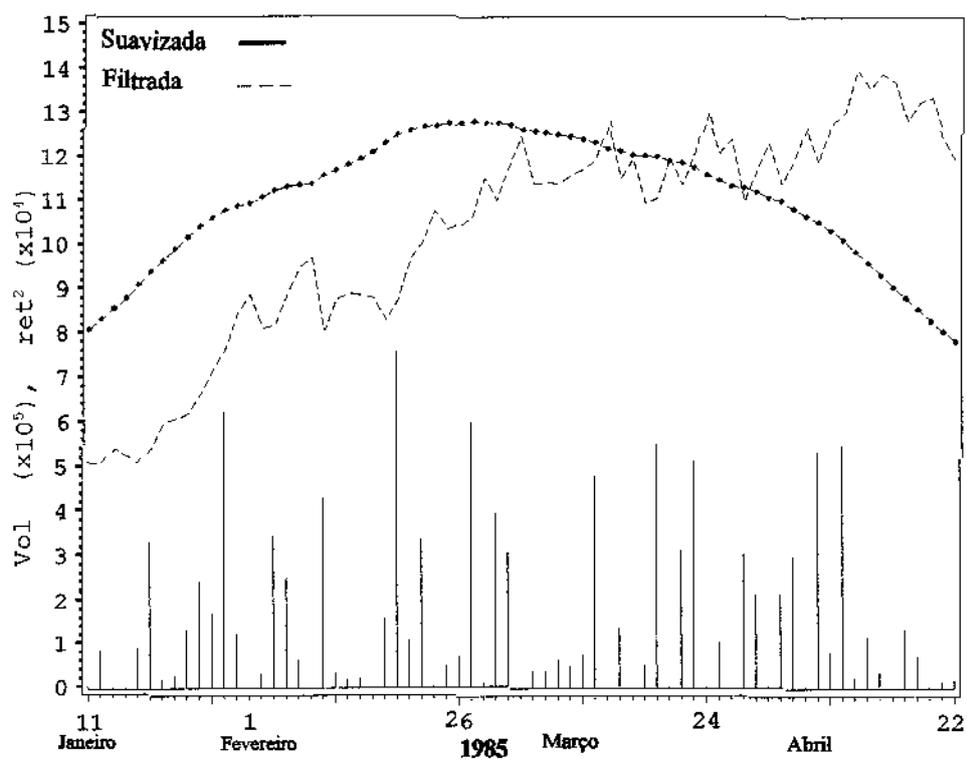
**Figura 4.14: Volatilidades estimadas AR(1)-SV e i.c 95% para os retornos no período de Março a Julho de 1986 na Taxa de Câmbio**

**Legenda:** vol=volatilidades,  $ret^2$ =quadrados dos retornos



**Figura 4.15: Comparação das volatilidades estimadas suavizada e filtrada na *Taxa de Câmbio***

**Legenda:** ve=volatilidade estimada



**Figura 4.16: Volatilidades estimadas AR(1)-SV no período do maior conglomerado de retornos grandes na *Taxa de Câmbio***

## CAPÍTULO 5

# COMPARAÇÃO DOS MODELOS ARCH E VARIÂNCIA ESTOCÁSTICA

### 5.1 Introdução

Na literatura financeira existem vários trabalhos que comparam empiricamente modelos de volatilidade para os retornos. Por exemplo, Pagan e Schwert (1990) comparam modelos ARCH e alternativas não-paramétricas; Cao e Tsay (1993) comparam modelos TAR e ARCH e Taylor (1994) compara modelos ARCH e de Variância Estocástica. Por sua vez, entre os estudos teóricos temos os de Nelson e Foster (1994) e Foster e Nelson (1996) que comparam modelos ARCH e de Variância Estocástica no tempo contínuo e Jacquier *et alii* (1994) que comparam as estimativas do modelo AR(1)-SV através das abordagens QML e MCMC. O grande problema na comparação de estimativas de volatilidade dos modelos reside no fato de que, na prática, *a volatilidade é uma variável não observável*. Para contornar esse problema, a solução frequentemente utilizada é adotar como referência (*benchmark*) para a volatilidade mensal a variância mensal não condicional dos retornos diários. Esta abordagem não será utilizada neste Capítulo. Neste Capítulo apresentamos as comparações das estimativas de volatilidade através dos modelos ARCH e Variância Estocástica utilizando as *Séries Exemplo* e simulações. Os aspectos considerados são basicamente dois: estimação pontual da volatilidade e a previsão um-passo a frente. A comparação das estimativas pontuais das volatilidades é realizada sob critérios basicamente qualitativos enquanto que na comparação das previsões um-passo a frente é utilizado um critério baseado nos retornos observados já utilizado em Herencia *et alii* (1995).

A comparação empírica dos modelos ARCH e de Variância Estocástica é realizada utilizando os resultados das seções 3.7 e 4.8 nos quais os dois modelos foram ajustados. Neste caso, o verdadeiro processo da volatilidade é desconhecido.

Por outro lado, com o objetivo de comparar as abordagens ARCH e de Variância Estocástica quando o processo gerador de dados e a volatilidade são conhecidos, recorreremos à

simulação. Primeiro é considerado um processo gerador de dados construído de forma que especificando os parâmetros ARCH, os parâmetros do modelo de Variância Estocástica sejam obtidos analiticamente, e vice-versa. Na seção 5.4.1 será mostrado que isso é possível. Em segundo lugar, os modelos ARCH e Variância Estocástica são comparados quando não é possível estabelecer relações analíticas entre os parâmetros dos modelos, sendo necessário estimar um dos modelos. Considerando que os resultados apresentados estão baseados em poucas simulações, esses resultados devem ser tomados como ilustrativos. A ênfase nas comparações será dada nos modelos GARCH(1,1) e AR(1)-SV estacionário.

Este Capítulo está organizado da seguinte forma. Nas Seções 5.2 e 5.3 é apresentada a comparação empírica utilizando as *Séries Exemplo*, dos modelos ARCH e de Variância Estocástica na estimativa da volatilidade e na previsão um-passo à frente. Na Seção 5.4 comparamos os modelos AR(1)-SV estacionário e GARCH(1,1) através de simulações e, finalmente, na Seção 5.5 apresentamos as conclusões.

## 5.2 Comparação Empírica dos Modelos

Nos Capítulos 3 e 4 foi demonstrado que os modelos ARCH e Variância Estocástica são capazes de reproduzir teoricamente vários *Fatos Estilizados* observados nas séries financeiras. Adicionalmente, a eficiência desses modelos foi ilustrada através da modelagem das *Séries Exemplo*.

Nesta seção será discutido outro aspecto na comparação das duas abordagens, as estimativas pontuais da volatilidade, para o qual utilizamos os resultados obtidos nas Seções 3.7 e 4.8 na modelagem das *Séries Exemplo*. Na ausência de um teste de hipóteses para comparar as volatilidades obtidas nas duas abordagens, nos parágrafos seguintes será realizada uma discussão basicamente qualitativa. Dentro da abordagem de Variância Estocástica o estimador utilizado é o *suavizado* e na abordagem ARCH o usual.

### *Conglomerados de Volatilidade e Suavidade*

Como pode ser constatado nas Figuras 5.1 a 5.4, tanto a volatilidade estimada pelos modelos ARCH quanto a volatilidade estimada suavizada do AR(1)-SV estacionário são capazes de imitar os movimentos de aumentos e diminuições dos quadrados dos retornos. É

possível observar também que a volatilidade do modelo de variância estocástica apresenta um comportamento bem mais suave quando comparado as volatilidades ARCH.

### *Estimativas pontuais da volatilidade*

Na série das *Ações*, ao comparar as volatilidades estimadas pelos modelos AR(1)-SV estacionário e GARCH(1,2) (Figura 5.5a) ou EGARCH(2,2) e AR(1)-SV (Figuras 5.5b), percebemos que as diferenças entre as estimativas podem ser bastante grandes<sup>1</sup>. Para ter uma idéia da magnitude das diferenças, calculamos a percentagem dos retornos cuja diferença entre as estimativas de volatilidade é maior que 20% e 45%, o qual é equivalente a calcular a percentagem dos retornos cujos desvios padrão da volatilidade se diferenciam em aproximadamente 10% e 20%, respectivamente. Nas Tabelas 5.1 I e II apresentamos os resultados quando a volatilidade estimada pelos modelos ARCH e AR(1)-SV estacionário, respectivamente, são tomados como referência.

Os dados da Tabela 5.1 mostram que para uma grande proporção dos retornos as diferenças entre as estimativas são consideráveis. Especificando a volatilidade ARCH como referência, para ambas as magnitudes de diferença (20% e 45%), a volatilidade AR(1)-SV subestima com maior frequência que superestima a volatilidade ARCH. Por sua vez, especificando a volatilidade AR(1)-SV como referência, a volatilidade ARCH superestima com mais frequência que subestima à volatilidade AR(1)-SV. Em consequência, a volatilidade estimada pelos modelos ARCH é maior que a volatilidade estimada pelo AR(1)-SV.

Analogamente, para a série *Taxa de Câmbio* encontramos que as diferenças entre as volatilidades estimadas, GARCH(2,2) e AR(1)-SV ou EGARCH(2,0) e AR(1)-SV podem ser bastante grandes para alguns retornos tal como pode ser observado nas Figuras 5.6a e 5.6b, respectivamente. Pelos resultados da Tabela 5.2 concluímos que as diferenças entre as estimativas são importantes para uma grande proporção dos retornos, sendo com maior frequência a volatilidade estimada ARCH maior que a volatilidade estimada pelo modelo AR(1)-SV.

---

<sup>1</sup>Os gráficos foram realizados sem considerar alguns retornos (7 em cada caso; a e b) nos quais as diferenças são bastante grandes. Esses pontos estão relacionados com valores extremos nos quadrados dos retornos, alguns dos quais são discutidos posteriormente.

**Tabela 5.1**  
**Diferenças entre as estimativas de volatilidade ARCH e AR(1)-SV estacionário nas Ações (%)**

*I. Quando a volatilidade ARCH é considerada como referência:*

	Diferença maior que 20%			Diferença maior que 45%		
	vs>1,2vg	vs<0,8vg	Total	vs>1,45vg	vs<0,55vg	Total
vs=AR(1)-SV vg=GARCH(1,1)	21,83	40,90	62,73	10,82	14,82	25,64
vs=AR(1)-SV vg=GARCH(1,2)	21,95	40,09	62,04	11,07	14,95	26,02
vs=AR(1)-SV vg=EGARCH(2,2)	21,64	40,59	62,23	10,76	14,20	24,96

*II Quando a volatilidade AR(1)-SV é considerada como referência*

	Diferença maior que 20%			Diferença maior que 45%		
	vg>1,2vs	vg<0,8vs	Total	vg>1,45vs	vg<0,55vs	Total
vs=AR(1)-SV vg=GARCH(1,1)	45,09	18,89	63,98	27,08	3,19	30,27
vs=AR(1)-SV vg=GARCH(1,2)	43,96	19,20	63,16	27,33	3,69	31,02
vs=AR(1)-SV vg=EGARCH(2,2)	44,78	18,32	63,10	28,08	2,75	30,83

**Tabela 5.2**  
**Diferenças entre as estimativas de volatilidade ARCH e AR(1)-SV na série Taxa de Câmbio (%)**

*I. Quando a volatilidade ARCH é considerada como referência:*

	Diferença maior que 20%			Diferença maior que 45%		
	vs>1,2vg	vs<0,8vg	Total	vs>1,45vg	vs<0,55vg	Total
vs=AR(1)-SV vg=GARCH(1,1)	19,07	24,22	43,29	4,07	4,81	8,88
vs=AR(1)-SV vg=GARCH(2,2)	20,77	24,45	45,22	4,30	4,41	8,71
vs=AR(1)-SV vg=EGARCH(1,2)	20,32	24,50	44,82	5,55	3,34	8,89

*II Quando a volatilidade AR(1)-SV é considerada como referência*

	Diferença maior que 20%			Diferença maior que 45%		
	vg>1,2vs	vg<0,8vs	Total	vg>1,45vs	vg<0,55vs	Total
vs=AR(1)-SV vg=GARCH(1,1)	28,81	14,20	43,01	13,24	0,79	14,03
vs=AR(1)-SV vg=GARCH(2,2)	28,64	15,11	43,75	13,98	0,90	14,98
vs=AR(1)-SV vg=EGARCH(1,2)	28,58	14,66	43,24	12,90	0,85	13,75

*Persistência*

Na tabela a seguir mostramos as estimativas da persistência nos ajustes considerados para cada uma das séries:

**Tabela 5.3**  
**Persistências estimadas nas Séries Exemplo**

	ARCH					VARIÂNCIA ESTOCÁSTICA
	GARCH (1,1)	GARCH (1,2)	GARCH (2,2)	EGARCH (2,0)	EGARCH (2,2)	AR(1)-SV
<i>Ações</i>	0,9967 209,7 <sup>a</sup>	*			**	0,9656 19,80
<i>Taxa de Câmbio</i>	0,9579		*	**		0,9753 27,72

<sup>a</sup> valores da meia-vida

\* não foi calculada a persistência

\*\* o modelo não é covariante estacionário.

Seguindo os resultados da Tabela 5.3, na série das *Ações* a persistência encontrada através do modelo GARCH(1,1) é muito maior que à correspondente ao AR(1)-SV estacionário. Pelo contrário, na série *Taxa de Câmbio* a persistência do AR(1)-SV é maior que a persistência do modelo GARCH(1,1). Neste caso a persistência do modelo AR(1)-SV é maior quando comparada com à das *Ações* e, utilizando o GARCH(1,1) acontece o contrário.

*Influência das observações extremas dos quadrados dos retornos*

Nos modelos ARCH ajustados para as *Séries Exemplo*; GARCH(1,1), EGARCH(1,0), GARCH(1,2) e EGARCH(2,2) constatamos que as observações extremas produzem um aumento bastante grande na estimativa da volatilidade nos instantes após a ocorrência dessa. Por exemplo, para a série das *Ações* no período da *Queda da Ministra Zélia Cardoso de Mello* (Figura 5.7a) observamos a presença de um retorno que, como comentado na Seção 3.7.1 pode ser considerada como um valor aberrante (VA) do tipo AO e produz um aumento brutal nas

estimativas das volatilidades futuras GARCH(1,2)<sup>2</sup> e EGARCH(2,2). O impacto é tão grande que somente após 20 dias (aproximadamente) da ocorrência do VA, a volatilidade estimada apresenta o mesmo valor do que o apresentado antes da ocorrência dessa observação extrema. Outro período no qual pode se observar a forte influência de um VA é nos dias seguintes à *Mudança da Banda Cambial* (Figura 5.7b). Por sua vez, na série *Taxa de Câmbio* durante o período de 02/08/85 a 25/11/85 (Figura 5.8a) podemos observar que as volatilidades estimadas pelos modelos GARCH(1,1) e EGARCH(2,0) explodem após a ocorrência dessa observação extrema em 23/09/85.

Pelo contrário, e como pode ser comprovado nas Figuras 5.7 e 5.8a, a volatilidade suavizada do modelo AR(1)-SV estacionário é menos influenciada pela presença de observações extremas. Essa é uma característica desejável para um modelo de volatilidade se o valor extremo é consequência da ocorrência de uma intervenção de tipo AO, enquanto o contrário ocorre quando usualmente existe um aumento real da volatilidade (por exemplo, devido à intervenção do tipo IO) como na Figura 5.8b. Nesse sentido, a eficiência relativa das estimativas de volatilidade pelo modelo ARCH ou de Variância Estocástica depende do tipo de intervenção.

#### *Influência de grandes Conglomerados de Volatilidade*

Na modelagem das *Séries Exemplo* pela abordagem de Variância Estocástica (Seção 4.8) foi observado que a volatilidade estimada do modelo AR(1)-SV estacionário é mais afetada pela presença de um conjunto de retornos cujos quadrados são grandes do que pela presença de valores aberrantes isolados (Figuras 5.1 e 5.3). Nos modelos ARCH ocorre o contrário, a volatilidade estimada é mais sensível à presença de valores extremos, embora certamente na presença de vários valores grandes próximos aumentem as estimativas. Na Figura 5.8b é mostrada um período onde existem conglomerados de valores com grande magnitude.

---

<sup>2</sup>Como apresentado na Seção 3.7.1, a volatilidade GARCH(1,1) é quase idêntica à volatilidade GARCH(1,2).

### 5.3 Comparação Empírica das Previsões um-passo à frente *ex-ante*

Como comentado anteriormente, de maneira geral, a dificuldade na comparação de modelos de volatilidade reside no fato de que a volatilidade não é observada na prática e, portanto, não é possível utilizar critérios como o Erro Médio Quadrático. Usualmente, esse problema tem sido contornado construindo uma série de referência (*benchmark*) para a volatilidade mensal que é estimada como a variância mensal não condicional dos retornos diários (ver, por exemplo, Cao & Tsay, 1993). Nesta dissertação propomos uma abordagem diferente comparando a qualidade da previsão dos modelos utilizando os retornos observados. A idéia, implementada em Herencia *et alii* (1995) consiste no seguinte: tanto para os modelos ARCH quanto os modelos de Variância Estocástica, conhecemos a distribuição condicional dos retornos (usualmente normal) dada a informação passada e o valor dos parâmetros. Supondo que o modelo é  $Y_t = \mu_t + Y_t^*$ , essa distribuição tem média igual à  $\mu_t$  e variância dada pela volatilidade estimada. Assim, podemos construir intervalos de confiança (i.c.) para as previsões um-passo à frente e comparar as abordagens ARCH e Variância Estocástica considerando dois critérios; *i*) o número de retornos incluídos nos i.c. e, *ii*) o comprimento dos i.c. O modelo que apresentar maior número de retornos dentro dos i.c. e menor comprimento nos i.c., será preferido, embora esses dois critérios sejam de certa forma contraditórios.

No cálculo das previsões para o instante  $t$ , idealmente, o modelo teria que ser estimado utilizando toda a informação até o instante  $t-1$ , mas por exigir muito trabalho computacional isso não será feito. Para contornar esse problema, utilizamos o esquema descrito na Seção 4.7.1<sup>3</sup> para cálculo dos i.c com volatilidade AR(1)-SV. Analogamente podemos fazer o mesmo com as volatilidades ARCH. No cálculo da previsão da volatilidade dentro da abordagem ARCH é utilizado o estimador usual e, dentro da abordagem de Variância Estocástica é utilizada a volatilidade filtrada. Os resultados são apresentados nas Tabelas 5.4 e 5.5.

Analisando os resultados para a série das Ações, dados na Tabela 5.4, constatamos que o número de retornos fora dos intervalos de confiança 95% (i.c.) é quase o mesmo nos modelos GARCH(1,1), GARCH(1,2) e EGARCH(2,2) e menores no modelo AR(1)-SV, embora a diferença não seja significativa estatisticamente. A não significância pode ser devida ao baixo número de observações. A mesma conclusão pode ser tomada a partir dos resultados da Tabela 5.5 na Taxa de Câmbio.

<sup>3</sup>Imediatamente antes da Tabela 4.3.

**Tabela 5.4**  
**Previsão um-passo à frente na série Ações**  
**Número de retornos fora dos intervalos de confiança 95% (# esperado = 4,75 )**

Período	ARCH						VARIÂNCIA ESTOCÁSTICA	
	GARCH(1,1)		GARCH(1,2)		EGARCH(2,2)		AR(1)-SV	
1451-1500	3	3 <sup>a</sup>	3	3	2	2	0	0
1501-1550	1	1	1	1	1	1	2	2
1551-1598	1	1	1	1	1	1	0	0

<sup>a</sup> Número de retornos fora dos i.c. 95% construídos com as estimativas dos parâmetros estimados utilizando toda a amostra.

**Tabela 5.5**  
**Previsão um-passo à frente na série Taxa de Câmbio**  
**Número de retornos fora dos intervalos de confiança 95% (# esperado=4,75 )**

Período	ARCH						VARIÂNCIA ESTOCÁSTICA	
	GARCH(1,1)		GARCH(2,2)		EGARCH(2,0)		AR(1)-SV	
1651-1700	0	3 <sup>a</sup>	3	3	0	0	0	0
1701-1750	1	1	1	1	1	1	0	0
1751-1767	2	2	2	2	2	2	2	2

<sup>a</sup> Número de retornos fora dos i.c. 95% construídos com as estimativas dos parâmetros estimados utilizando toda a amostra.

É importante observar que, como mostram as Tabelas 5.4 e 5.5, os resultados na análise da previsão um-passo à frente permanecem inalterados se consideramos as estimativas dos parâmetros utilizando toda a informação da amostra. A única diferença se dá no primeiro período da série *Taxa de Câmbio*. Outro aspecto importante é a percentagem de retornos fora dos i.c., que em todos os casos é inferior ao valor esperado 5%, existindo, portanto, uma indicação de superestimação da largura dos i.c. Assim, na série das *Ações*, utilizando os modelos GARCH(1,1), GARCH(1,2), EGARCH(2,2) e AR(1)-SV obtemos, respectivamente, 3,38%, 3,38%, 2,70% e 1,35%. de retornos fora do i.c. Por sua vez, na série *Taxa de Câmbio*, as percentagens são 2,56%, 5,13%, 2,56% e 1,71% para os modelos GARCH(1,1), GARCH(2,2), EGARCH(2,0) e AR(1)-SV, respectivamente.

Nas Figuras 5.9 e 5.10 apresentamos os i.c. 95% considerando os modelos re-estimados em cada período. Para a série das *Ações* (Figura 5.9) é possível observar que os i.c. construídos com a volatilidade AR(1)-SV estão muito próximos dos i.c. construídos com a volatilidade GARCH(1,1), sendo que os i.c. se cruzam e não existe um padrão como o observado para a série *Taxa de Câmbio* (Figura 5.10) no qual os i.c. do AR(1)-SV incluem os i.c. do GARCH(1,2).

A influência dos grandes valores nos quadrados dos retornos no comprimento dos i.c. pode ser observado na Figura 5.9 na *Mudança da Banda Cambial* da série das *Ações* (10/03/95). Como os i.c. são baseados na volatilidade, a presença de uma observação extrema, faz com que os i.c. dos modelos GARCH(1,2) e EGARCH(2,2) nos instantes posteriores à ocorrência do valor extremo sejam desnecessariamente longos, dado que o grande valor do retorno não influencia os retornos futuros (VA do tipo AO). Em consequência, considerando que os i.c. do modelo AR(1)-SV incluem o mesmo número de retornos, o modelo de Variância Estocástica é preferível nestas situações. Observe, no entanto, que após algumas observações os i.c. do modelo AR(1)-SV passam a ser desnecessariamente longos.

Portanto, em termos de previsão um-passo à frente, seguindo o critério adotado dos i.c. os resultados não mostram a dominância de um modelo sobre outro.

#### **5.4 Comparação dos Modelos GARCH(1,1) e AR(1)-SV estacionário através de simulações.**

Como comentado anteriormente uma das dificuldades de se comparar as estimativas da volatilidade através das séries exemplo era o fato das volatilidades serem desconhecidas. Uma forma de se evitar isto é através de simulação onde o processo gerador dos dados é conhecido (PGD). Conhecendo-se o PGD e a volatilidade é possível comparar as estimativas. Este procedimento, entretanto, adiciona um outro problema, o de escolher o PGD, já que um modelo pode ser adequado para um PGD e ruim para outro. Como estamos comparando estimativas de volatilidade através de modelos GARCH e modelos de Variância Estocástica é natural que estes dois modelos estejam entre os PGD escolhidos. Ao escolhermos estes modelos podemos verificar o quanto um modelo de uma família pode aproximar um modelo de outra família, ou o quanto um procedimento é robusto em termos de PGD.

Sabemos que tanto processos GARCH(1,1) como AR(1)-SV podem gerar processos com função de autocorrelação (f.a.c) dos quadrados do retorno da forma ARMA(1,1) com a primeira correlação pequena e decaimento lento. Quando as f.a.c. dos 2 modelos são próximos eles são ditos *equivalentes em termos de f.a.c. dos quadrados dos retornos*. Neste caso espera-se que não exista grande diferença nas estimativas de volatilidade, não perdendo, portanto, muita eficiência quando um modelo ARCH(1,1) é ajustado como AR(1)-SV e vice-versa. Porém, como a igualdade de f.a.c. não necessariamente implica em estimativas de volatilidades iguais é interessante que estes casos estejam incluídos entre os PGD escolhidos. Modelos equivalentes são apresentados na subseção 5.4.1 enquanto as comparações para estimativas de volatilidade destes modelos e a previsão um-passo à frente são apresentadas na subseção 5.4.2. Para evitar que problemas de estimação dificultem a comparação, vamos considerar que conhecemos os parâmetros dos modelos equivalentes. Outro motivo para se escolher modelos com este tipo de f.a.c., além da existência de modelos equivalentes, é o fato de várias séries financeiras empíricas possuírem f.a.c. dos quadrados dos retornos com essa característica, como ocorre com os casos das *Séries Exemplo* (ver Seção 2.7.2). Um segundo tipo de PGD a ser estudado, e apresentado na subseção 5.4.3, é quando não existem modelos equivalentes. Neste caso os parâmetros dos modelos ajustados têm que ser estimados.

#### 5.4.1 Modelos equivalentes GARCH(1,1) e AR(1)-SV estacionário

Sejam os modelos GARCH(1,1) definido em (3.3.2.1.1) e AR(1)-SV definido por (4.3.2.6) com a parametrização  $\sigma_{\varepsilon}^2 = \pi^2/2$  e  $q = \sigma_{\eta}^2 / \sigma_{\varepsilon}^2$ . Como foi mostrado em (3.3.2.4) e em (4.3.7), quando  $\gamma$  é próximo de um, as f.a.c. dos quadrados dos retornos apresentam forma ARMA(1,1) dadas por:

$$\text{GARCH (1,1)} \quad \rho(\tau) = \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} (\alpha + \beta)^{\tau-1} \quad \tau=1,2,\dots \quad (5.4.1.1)$$

$$\text{AR(1)-SV} \quad \rho^*(\tau) = \frac{\exp(\sigma_h^2) - 1}{3\exp(\sigma_h^2) - 1} \gamma^{\tau} \quad \tau=1,2,\dots \quad (5.4.1.2)$$

Tanto em (5.4.1.1) quanto em (5.4.1.2), todas as autocorrelações estão determinadas a partir das duas primeiras. Os parâmetros responsáveis pela dinâmica no modelo GARCH(1,1) são  $\{\alpha, \beta\}$  e no AR(1)-SV  $\{\gamma, \sigma_{\eta}^2\}$ . A seguir mostramos como é obtida a equivalência.

1) *Especificando o modelo GARCH(1,1):*

Calculamos  $\rho(1)$  e  $\rho(2)$  a partir de (5.4.1.1) e substituímos esses valores em (5.4.1.2) obtendo um sistema de duas equações com duas incógnitas. Assim:

$$\frac{\exp(\sigma_h^2) - 1}{3 \exp(\sigma_h^2) - 1} \gamma = \rho(1) \quad (5.4.1.3)$$

$$\rho(1) \gamma = \rho(2). \quad (5.4.1.4)$$

A solução do sistema é:

$$\gamma = \frac{\rho(2)}{\rho(1)} \quad (5.4.1.5)$$

$$\sigma_h^2 = \log \left( \frac{\rho(2) - \rho(1)^2}{\rho(2) - 3\rho(1)^2} \right). \quad (5.4.1.6)$$

Na expressão anterior observamos que é necessária a restrição:  $\rho(2) > 3\rho^2(1)$  e, por sua vez, devemos considerar as condições para a existência de segundo e quarto momento do GARCH(1,1), dadas em (3.3.2.1.3).

Por outro lado, substituindo (5.4.1.1) em (5.4.1.5), obtemos a seguinte relação importante:

$$\gamma = \alpha + \beta, \quad (5.4.1.7)$$

isto é, *as definições de persistência, quando grande (aproximadamente 1), levam ao mesmo valor de persistência nos dois modelos.*

Finalmente, podemos encontrar  $\sigma_{\eta}^2$  e  $q$  através das expressões:

$$\sigma_{\eta}^2 = \sigma_h^2(1 - \gamma^2) \quad (5.4.1.8)$$

$$q = \sigma_{\eta}^2 / 4.93. \quad (5.4.1.9)$$

Portanto, especificada a dinâmica do modelo GARCH(1,1) (através de  $\alpha$  e  $\beta$ ), obtemos o modelo equivalente (em dinâmica) AR(1)-SV determinado por (5.4.17) e (5.4.18) ou (5.4.19).

## 2) Especificando o modelo AR(1)-SV

Substituindo em (5.4.1.1)  $\rho(1)$  por  $\rho^*(1)$  calculada a partir de (5.4.1.2) e  $\alpha$  por  $\gamma - \beta$ , obtemos uma equação em  $\beta$  da seguinte forma:

$$\beta^2[\rho(1) - \gamma] + [\gamma^2 - 2\rho(1)\gamma + 1]\beta + [\rho(1) - \gamma] = 0$$

uma vez resolvida a equação anterior (isto é  $\beta$ ) obtemos  $\alpha$  através de  $\alpha = \gamma - \beta$ .

Para completar a equivalência dos modelos precisamos estabelecer o nível. Os parâmetros responsáveis desse valor são  $\omega$  no modelo GARCH (1,1) e  $\kappa$  no modelo AR(1)-SV. A relação entre os dois parâmetros é encontrada utilizando o valor da variância da série dos retornos. Assim:

$$\text{no GARCH(1,1):} \quad E[Y_t^2] = \frac{\omega}{1 - (\alpha + \beta)} \quad (5.4.1.10)$$

$$\text{no AR(1)-SV:} \quad E[Y_t^2] = \exp\left(\frac{\sigma_{\eta}^2}{2(1 - \gamma^2)}\right), \quad (5.4.1.11)$$

em conseqüência igualando essas expressões obtemos:

$$\omega = (1 - \gamma)\kappa \exp(0.5\sigma_h^2), \text{ ou}$$

$$\kappa = \omega(1 - \alpha - \beta)^{-1} \exp(-0.5\sigma_h^2).$$

Na tabela a seguir mostramos alguns modelos GARCH(1,1) e AR(1)-SV com grande persistência que apresentam igual dinâmica, sendo que o processo gerador é GARCH(1,1):

**Tabela 5.6**  
**Equivalência de alguns alguns modelos GARCH(1,1) e AR(1)-SV estacionário**

Modelo	$\alpha$	$\beta$	$\rho(1)$	$\rho(2)$	$\gamma$	$\sigma_n^2$	$q$
1	0,15	0,80	0,3	0,285	0,95	0,25	0,05
2	0,10	0,80	0,14	0,126	0,90	0,30	0,06
3	0,10	0,87	0,226	0,219	0,97	0,0552	0,011
4	0,05	0,94	0,155	0,153	0,99	0,0092	0,0019
5	0,18	0,80	0,54	0,529	0,98	ne	ne
6	0,14	0,85	0,562	0,556	0,99	ne	ne

ne: não existe equivalência

Da análise da Tabela 5.6, podemos observar que não existe equivalência do AR(1)-SV no quinto e sexto modelo, pois os parâmetros do GARCH(1,1) não satisfazem a restrição  $\rho(2) > 3\rho^2(1)$ . Comparando os modelos 1, 2 e 3 notamos que o tamanho de  $\alpha$  é determinante na equivalência dos modelos. Na Figura 5.11 é apresentada a região de equivalência do modelo AR(1)-SV estacionário supondo que o PGD é GARCH(1,1) para valores de  $\beta$  maiores que 0,8 e persistências maiores que 0,90.

#### 5.4.2 Comparação das estimativas de volatilidade e previsão um-passo à frente

Na simulação do PGD são escolhidos valores para os parâmetros dentro de um conjunto que ocorrem com frequência na literatura empírica. Assim, nos modelos ARCH muito frequentemente a persistência é maior do que 0,90 (Taylor, 1986) com valores de  $\beta$  maiores do que 0,80. Por sua vez, nos modelos AR(1)-SV, os valores de  $\gamma$  são usualmente maiores do que 0,90 e os valores de  $\sigma_n^2$  estão entre 0,01 e 2,77; ver Taylor (1986,1991) e Jacquier *et alii*

(1994). Na Tabela 5.7 são apresentadas as três especificações escolhidas para os modelos GARCH(1,1) e AR(1)-SV equivalentes:

**Tabela 5.7**  
**Modelos utilizados nas simulações**

	P	$\omega$	$\alpha$	$\beta$	$\kappa$	$\gamma$	$\sigma_{\eta}^2$	$\sigma_{\eta}^2$	q
Modelo I	0,95	$10^{-6}$	0,15	0,80	$5,547 \times 10^{-6}$	0,95	0,25008	2,56495	0,05073
Modelo II	0,97	$10^{-6}$	0,10	0,87	$2,09 \times 10^{-5}$	0,97	0,05522	0,93443	0,011
Modelo III	0,99	$10^{-6}$	0,05	0,94	$7,93 \times 10^{-5}$	0,99	0,00923	0,46363	0,00187

P é o valor da persistência

No Modelo I os parâmetros GARCH(1,1) são similares à estimativas da série *Ações* encontrada no Capítulo 3. Por sua vez, no Modelo II os parâmetros AR(1)-SV são iguais à estimativa da série DM/US\$ analisada por Shephard (1994b) e são similares também aos valores utilizados nas simulações realizadas por Ruiz (1994), no qual  $\gamma=0,97$  e  $\sigma_{\eta}^2=0,04$ . Finalmente, no Modelo III os parâmetros exibem valores similares aos encontrados por Harvey e Shephard (1993b) iguais a  $\gamma=0,975$  e  $\sigma_{\eta}^2=0,01$ .

As observações simuladas foram obtidas utilizando um programa desenvolvido na linguagem do Procedimento IML do pacote estatístico SAS 6.0. Para cada um dos modelos registrados na Tabela 5.6 são realizadas duas simulações, uma considerando como PGD o GARCH(1,1) e a outra considerando o AR(1)-SV. Na geração das séries foram inicialmente gerados 3100 observações e descartados os 100 primeiros valores gerados e, para inicializar a variância condicional, utilizamos a variância não condicional. Na Tabela 5.8 apresentamos os nomes dados para cada uma das simulações:

**Tabela 5.8**  
**Simulações realizadas**

Modelo	Persistência	Gerador	Nome
I	0,95	Garch(1,1)	Sim1
		AR(1)-SV	Sim1a
II	0,97	Garch(1,1)	Sim2
		AR(1)-SV	Sim2a
III	0,99	Garch(1,1)	Sim3
		AR(1)-SV	Sim3a

A seguir apresentamos os resultados para os dois casos considerados; quando simulamos o modelo GARCH(1,1) e quando simulamos o AR(1)-SV. Em cada um deles, a volatilidade verdadeira gerada pelo modelo é nomeada simplesmente de *volatilidade* enquanto que a volatilidade estimada a partir das observações geradas  $\{Y_t\}$  será chamada *volatilidade estimada*. Na discussão dos resultados a seguir, sempre será feita a distinção entre volatilidade estimada e volatilidade verdadeira.

*I. O processo gerador é GARCH(1,1)*

Neste caso, uma vez obtidos os parâmetros equivalentes do modelo AR(1)-SV, calculamos a volatilidade filtrada através do Filtro de Kalman (expressões 4.4.1.2) e encontramos a volatilidade suavizada pela inversão desse filtro começando em  $t=T$ . A volatilidade filtrada é utilizada para a previsão um-passo à frente, enquanto que a volatilidade suavizada é usada para compará-la com a volatilidade verdadeira.

Nas Figuras 5.12a-c são comparados os valores das volatilidades e suas estimativas sendo observado que existem vários pontos bem acima ou bem abaixo da reta  $x=y$  indicando que a volatilidade estimada subestima e ou superestima a volatilidade GARCH(1,1) de maneira considerável. Para obter uma idéia da superestimação, calculamos a percentagem de pontos alocados acima da reta  $x=y$  levando em consideração os níveis da volatilidade. Temos assim:

**Tabela 5.9**  
**Modelo simulado GARCH(1,1): percentagem de superestimação da volatilidade**

	baixo	médio	alto	Total
Sim 1	39*	42,4*	50,9	44,1
Sim 2	48,2	55,7*	50,8	51,57
Sim 3	48,8	50,2	51,2	50,07

\* rejeita hipótese de probabilidade 0,5 com nível de significância igual a 0,5% mas considerando independência.

A partir dos resultados na tabela vemos que em Sim3 e, em menor medida em Sim2, existe praticamente a mesma percentagem de pontos acima e abaixo da reta  $x=y$ . Por sua vez em Sim1 quando a volatilidade é baixa, existem mais pontos acima que abaixo da linha reta.

Para medir a magnitude da subestimação ou superestimação, utilizamos o mesmo procedimento da seção anterior e calculamos a percentagem dos retornos cuja estimativa da volatilidade AR(1)-SV é diferente em 20% e 45% da volatilidade verdadeira GARCH(1,1). Os resultados registrados na Tabela 5.10 mostram que para uma grande percentagem dos retornos, a subestimação e superestimação é considerável, sendo que essas percentagens diminuem na medida que a persistência aumenta, i.e, na medida que os modelos são mais próximos em dinâmica). Por sua vez, de maneira geral, as percentagens de superestimação e subestimação são similares. Para ilustrar a grande diferença observada nas volatilidades verdadeira e estimada apresentamos na Figura 5.13a o gráfico da variável  $q$ , definida como o quociente da volatilidade estimada e a volatilidade verdadeira ( $q = \text{vol. estimada} / \text{vol. verdadeira}$ ), com relação a volatilidade verdadeira no caso de Sim3. Embora não apresentado, para Sim1 e Sim2 obtemos figuras semelhantes.

**Tabela 5.10**  
**Magnitude da subestimação e super-estimação da volatilidade estimada (vs) AR(1)-SV com respeito à volatilidade verdadeira(vg) GARCH(1,1)**

	Diferença maior que 20%			Diferença maior que 45%		
	vs>1,2vg	vs<0,8vg	Total	vs>1,45vg	vs<0,55vg	Total
Sim 1	31,24	38,13	69,37%	18,67	16,80	35,47%
Sim 2	32,23	26,87	59,1%	14,06	4,07	18,13%
Sim 3	26,30	19,67	45,97%	8,87	3,16	12,03%

Com relação à previsão um-passo à frente e, como realizado anteriormente para as *Séries Exemplo*, calculamos a percentagem dos retornos fora dos intervalos de confiança 95% (i.c.) construídos a partir da previsão da volatilidade filtrada, considerando os níveis na volatilidade. Diferentemente do realizado para as *Séries Exemplo*, aqui os parâmetros do modelo são conhecidos e não há necessidade de estimar os parâmetros cada 50 observações. Neste caso a volatilidade filtrada é atualizada para cada retorno e a previsão é dada simplesmente pela expressão (4.6.2 ). Na Tabela 5.11 apresentamos os resultados e podemos constatar (ao observar a última linha) que para Sim1 a percentagem dos retornos fora dos i.c. construídos com a volatilidade estimada é muito maior do que quando os i.c. são construídos com a volatilidade verdadeira. Essa diferença de percentagem diminui na medida que a persistência aumenta, sendo bastante próximos para Sim3. Considerando os níveis de volatilidade, em Sim 1, os i.c. estimados excluem uma percentagem muito maior do que a

esperada enquanto que, a diferença de percentagem observada e esperada é menor nas três simulações quando a volatilidade é média.

**Tabela 5.11**  
**Modelo simulado GARCH(1,1): percentagem de retornos fora dos Intervalos de Confiança 95%**

	Sim 1 persistência=0,95		Sim 2 persistência=0,97		Sim 3 persistência=0,99	
Nível da Volatilidade						
baixo	12,4	3,5 <sup>a</sup>	6,3	5,2	6,7	6,2
médio	8,2	4,7	5,7	4,9	5,2	4,9
alto	8,3	5,3	6,5	4,6	5,5	5,0
% Total	9,63	4,47	6,17	4,9	5,80	5,37

<sup>a</sup> corresponde a percentagem observada com os i.c. construídos com a volatilidade verdadeira GARCH(1,1)

Uma das causas da diferença entre a percentagem estimada e a percentagem observada é a influência das *observações muito pequenas nos quadrados dos retornos*. Para ilustrar isso apresentamos na tabela a seguir os valores da volatilidade e a previsão para alguns instantes de Sim 1.

**Tabela 5.12**  
**Modelo simulado GARCH(1,1): efeito das observações muito pequenas nos quadrados dos retornos**

Tempo	$Y_t^2$	volatilidade ( $\times 10^5$ )	previsão ( $\times 10^5$ )
1978	$7 \times 10^{-7}$	0,8243	0,2233
1979	$8,2 \times 10^{-12}$	0,7706	0,2403
1980	$1,3 \times 10^{-6}$	0,7165	0,0401
1981	$1,2 \times 10^{-5}$	0,6932	0,0690

Os resultados da tabela mostram que o efeito do retorno T=1979 na previsão da volatilidade em T=1980 é bastante grande, produzindo uma subestimação considerável da volatilidade. Esse efeito, como será explicado a seguir, permanece nas previsões seguintes. Como produto da subestimação, os i.c. construídos apresentam largura muito menor ao verdadeiro e, como pode ser observado na Figura 5.14 esses i.c. estimados não conseguem incluir os retornos em T=1981 e T=1984.

Considerando que no instante  $T=t$ , tanto a volatilidade quanto a previsão um-passo a frente *dependem do mesmo tipo de informação*:  $\{Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_1\}$  é simples explicar a

diferença de comportamento observada ante a presença de observações muito pequenas. Com efeito, numa situação de alta persistência (no caso igual a 0.95) e  $\beta$  grande (igual a 0,80), a contribuição de um retorno pequeno (valor absoluto),  $Y_{t-1}$  na volatilidade GARCH(1,1) no instante  $t$ :  $\sigma_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2$  e menor quando comparada com a contribuição de  $\sigma_{t-1}^2$ . Por sua vez, para o modelo AR(1)-SV estacionário, temos pelas expressões derivadas na Seção 4.6 que a previsão um-passo à frente é dada por:  $\hat{\sigma}_{t/t-1}^2 = \kappa \exp[m_{t/t-1}]$  onde  $m_t = [\gamma(1-P)]m_{t-1} + [P]X_t$  e  $X_t = \log(Y_t^2) + 1,27 - \log(\kappa)$ . Neste caso, já que o efeito de  $Y_t^2$  é multiplicativo, uma quantidade muito pequena nos quadrados produz um decrescimento grande na previsão.

As observações com quadrados dos retornos muito grandes também produzem efeito na previsão um-passo à frente. Assim, quando no instante  $T=t-1$  a volatilidade e a volatilidade estimada apresentam valores próximos, uma observação muito grande de  $Y_t^2$  no instante  $T=t$  faz explodir a volatilidade, enquanto que a previsão pode subestimar a volatilidade. Na Figura 5.15 ilustramos esse efeito .

## II. O processo gerador é AR(1)-SV

Para as três simulações, uma vez geradas as observações  $Y_t$  e as volatilidades, calculamos as volatilidades estimadas a partir da relação  $\sigma_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2$ .

Os resultados da Tabela 5.13 mostram que para uma proporção bastante grande das observações, a subestimação e superestimação é considerável, sendo que essas proporções são menores na medida que a persistência aumenta. Claramente, existe uma frequência maior de superestimação, diferentemente do que acontece na Tabela 5.10 onde a superestimação é similar à subestimação. Na Figura 5.13b, correspondente a Sim 3a, podemos constatar as grandes diferenças encontradas entre a volatilidade estimada e a verdadeira, utilizando como medida o quociente ( $q$ ) entre a volatilidade estimada e a verdadeira volatilidade.

Com relação à previsão um-passo à frente, na Tabela 5.14 apresentamos os resultados da percentagem dos retornos fora dos intervalos de confiança 95% . Os dados dessa tabela mostram que considerando toda a informação (% total), a percentagem dos retornos fora dos i.c estimados é próxima à percentagem esperada, mas, ao considerar os níveis da volatilidade,

essas percentagens apresentam em alguns casos diferenças importantes. Assim, de maneira geral, quando a volatilidade verdadeira é baixa, a percentagem dos retornos fora dos i.c. é menor do que a esperada (i.e. existe superestimação da volatilidade) enquanto que quando a volatilidade é alta, as percentagens observada e esperada são próximas.

**Tabela 5.13**  
**Magnitude da subestimação e superestimação da volatilidade estimada (vg) GARCH(1,1) com respeito à volatilidade verdadeira(vs) AR(1)-SV**

	Diferença maior que 20%			Diferença maior que 45%		
	vg>1,2vs	vg<0,8vs	Total	vg>1,45vs	vg<0,55vs	Total
Sim 1a	71,80	16,70	88,50%	65,37	9,10	74,47%
Sim 2a	51,83	20,37	72,20%	37,30	6,47	43,77%
Sim 3a	44,93	18,87	63,80%	27,47	3,40	30,87%

**Tabela 5.14**  
**Modelo simulado AR(1)-SV : percentagem de retornos fora dos Intervalos de Confiança 95%**

Nível da Volatilidade	Sim 1a		Sim 2a		Sim 3a	
baixo	1,5	4,7 <sup>a</sup>	4,8	5,2	2,7	3,3
médio	3,7	4,4	7,4	5,4	5,4	5,9
alto	6,9	4,6	5,5	5,1	4,4	4,4
% Total	4,03	4,57	5,9	5,23	4,17	4,53

<sup>a</sup> corresponde a percentagem observada com os i.c. construídos com a volatilidade verdadeira: AR(1)-SV

A diferença do caso anterior, quando o processo gerador é GARCH(1,1), neste caso os valores da volatilidade e da previsão um-passo (igual à volatilidade) dependem de informações diferentes. Com efeito, a volatilidade AR(1)-SV é um processo que evolui independentemente dos retornos passados e depende dos choques  $\eta_t$  segundo a regra:  $\log(\sigma_t^2) = \gamma \log(\sigma_{t-1}^2) + \eta_t$ , sendo os retornos definidos a partir da volatilidade e dos choques  $\varepsilon_t$  como  $Y_t = \sigma_t \varepsilon_t$ . Por sua vez a previsão um passo a frente (ou volatilidade estimada) depende do valor da volatilidade e dos retornos no instante anterior. Portanto, a volatilidade no instante  $T=t$  depende da informação  $\{\sigma_{t-1}^2, \eta_t, \sigma_{t-2}^2, \eta_{t-1}, \dots\}$  enquanto que a previsão um-passo a frente (ou volatilidade estimada) depende da informação  $\{\sigma_{t-1}^2, Y_{t-1}^2, \sigma_{t-2}^2, Y_{t-2}^2, \dots\}$ .

A utilização de informações diferentes levam a reações diferentes diante de observações extremas (tanto grandes quanto pequenas), que por sua vez, produzem grandes diferenças entre a volatilidade e a previsão um-passo a frente (ou volatilidade estimada). Para ilustrar isso apresentamos a seguir duas situações de Sim 2a. Na Figura 5.16 é observado que em  $T=828$  temos uma observação muito grande como produto de um choque  $\varepsilon_t$  igual a  $-2,88$ . O valor muito grande nos quadrados dos retornos faz com que o valor da previsão GARCH(1,1) aumente de forma abrupta superestimando a volatilidade (Figuras 5.16, 5.17) e, já que a persistência é grande, os efeitos do retorno em  $T=828$  permanecem significativos por vários instantes de tempo. Como resultado da superestimação da volatilidade após a ocorrência da observação extrema, os intervalos de confiança 95% estimados são muito maiores aos verdadeiros, mas, ambos os i.c incluem o mesmo número de retornos até  $T=845$  (instante no qual a volatilidade e a volatilidade estimada são próximas). Por sua vez, em  $T=862$  (Figura 5.16) ocorre uma observação muito pequena nos quadrados dos retornos que produz a diminuição da volatilidade e, em menor medida, a diminuição da volatilidade (previsão) estimada no instante seguinte. Neste caso observa-se uma subestimação da volatilidade. Na Figura 5.17 podemos observar que em  $T=864$  os i.c. estimados não conseguem incluir o retorno observado.

#### **5.4.3 Comparação dos modelos GARCH(1,1) e AR(1)-SV estacionário quando não existem modelos equivalentes**

Com o objetivo de ilustrar a comparação dos modelos GARCH(1,1) e AR(1)-SV quando não existe equivalência na f.a.c. da série dos quadrados do processo gerador de dados, são realizadas 4 simulações. Primeiro são simulados dois processos GARCH(1,1) e estimados os parâmetros e a volatilidade AR(1)-SV, depois simulamos o AR(1)-SV e estimamos os parâmetros e a volatilidade GARCH(1,1).

##### *1) Simulando GARCH(1,1)*

Na Tabela 5.15 são apresentados os modelos gerados e as estimativas dos parâmetros correspondentes. Para os dois ajustes do modelo AR(1)-SV temos convergência *muito forte*, e na análise de adequabilidade o modelo é satisfatório. Como realizado na Seção 5.4.2, para a previsão um-passo à frente utilizamos a volatilidade filtrada e para a comparação com a volatilidade verdadeira utilizamos a volatilidade suavizada.

**Tabela 5.15**  
**Parâmetros do modelo simulado GARCH(1,1) e estimativas do modelo AR(1)-SV**  
**no caso quando não existe equivalência**

	1ra autocorrelação	Parâmetros GARCH(1,1)				Estimativas AR(1)-SV <sup>a</sup>			
		$\omega$	$\alpha$	$\beta$	$\alpha+\beta$	$\hat{\kappa}$	$\hat{\gamma}$	$\sigma_{\eta}^2$	$\hat{q}$
Sim 4	0,1793	$10^{-6}$	0,15	0,60	0,75	$3,5 \times 10^{-6}$	0,8548	0,1028	0,0215
Sim 5	0,1189	$10^{-6}$	0,10	0,70	0,80	$4,9 \times 10^{-6}$	0,8212	0,0990	0,0213

<sup>a</sup> O modelo foi estimado sem a restrição  $\sigma_{\xi}^2 = 4,93$ , e as estimativas são muito próximas desse valor.

Os resultados da Tabela 5.16 mostram que para uma proporção bastante grande de observações a diferença entre a volatilidade verdadeira e a estimada é considerável, sendo que as percentagens de subestimação e superestimação são similares. Ao comparar esses dados com os resultados da Tabela 5.10 vemos que neste caso a frequência de subestimação e superestimação é menor. Para Sim 4, é apresentado na Figura 5.18a, o gráfico de  $q$  (vol. estimada/vol.verdadeira) com relação à volatilidade verdadeira. Comparando esse gráfico com a Figura 5.13a observamos que os padrões exibidos pelos pontos são semelhantes.

**Tabela 5.16**  
**Magnitude da subestimação ou superestimação da volatilidade estimada (vs) AR(1)-SV com**  
**respeito à volatilidade verdadeira(vg) GARCH(1,1) no caso de modelos não equivalentes**

	Diferença maior que 20%			Diferença maior que 45%		
	vs>1,2vg	vs<0,8vg	Total	vs>1,45vg	vs<0,55vg	Total
Sim4	19,17	25,10	44,27	4,57	4,76	9,33
Sim 5	23,67	17,00	40,67	4,00	1,47	5,47

Com relação à previsão um-passo à frente, os resultados estão registrados na Tabela 5.17. Esses dados mostram que para Sim 5 (maior persistência) os resultados são melhores quando comparados com Sim 4, no sentido de que a percentagem das observações fora dos i.c 95% é mais próxima da observada (e esperada). Analisando por níveis, obtemos os piores resultados nos níveis de volatilidade alta.

**Tabela 5.17**  
**Percentagem de retornos fora dos Intervalos de Confiança 95% se o modelo simulado é GARCH(1,1) e não existe equivalência**

	Sim 4		Sim 5	
Nível da volatilidade				
baixo	3,60	3,50 <sup>a</sup>	3,90	5,50
médio	4,4	4,00	4,00	4,10
alto	10,10	5,8	7,70	5,80
% Total	6,03	4,43	5,20	5,13

<sup>a</sup> corresponde a percentagem observada com os i.c. construídos com a volatilidade verdadeira GARCH(1,1)

Os resultados obtidos na estimação da volatilidade mostram que as percentagens de subestimação e superestimação (diferenças em 20 e 45%) são menores quando comparadas com os resultados da Tabela 5.10. Por sua vez, com relação à previsão um-passo à frente, temos que as percentagens totais da Tabela 5.17 são similares as percentagens totais encontradas para Sim2 e Sim3. Isto indica que o modelo AR(1)-SV equivalente não necessariamente leva a uma melhor performance.

*II) Simulando AR(1)-SV*

Os modelos AR(1)-SV gerados e as estimativas do modelo GARCH(1,1) são mostrados na Tabela 5.18.

**Tabela 5.18**  
**Modelos AR(1)-SV simulados e estimativas GARCH(1,1) quando não existe equivalência**

	1ra autocorrelação	Parâmetros AR(1)-SV				Estimativas GARCH(1,1) <sup>a</sup>			
		$\hat{\kappa}$	$\hat{\gamma}$	$\sigma_{\eta}^2$	$\hat{q}$	$\omega$	$\alpha$	$\beta$	$\alpha+\beta$
Sim 4a	0,11517	$4 \times 10^{-6}$	0,75	0,1972	0,04	$1,3 \times 10^{-6}$ ( $3,3 \times 10^{-7}$ ) <sup>b</sup>	0,149177 (0,02695)	0,586696 (0,08018)	0,7359
Sim 5a	0,11533	$4 \times 10^{-6}$	0,80	0,1479	0,03	$9,4 \times 10^{-7}$ ( $2,4 \times 10^{-7}$ )	0,141677 (0,02426)	0,667058 (0,0677)	0,8087

<sup>a</sup> O modelo foi estimado sem a restrição  $\sigma_{\xi}^2 = 4,93$ , e as estimativas são muito próximas desse valor.

<sup>b</sup> Desvios padrão das estimativas

Seguindo os resultados da Tabela 5.19 constatamos que a frequência de subestimação e superestimação é bastante grande. Considerando as diferenças de 20%, neste caso, as

percentagens totais de superestimação e subestimação são similares as encontradas nos casos Sim 1a, Sim 2a, Sim 3a da seção 5.4.2, quando existem modelos equivalentes. Já ao considerar as diferenças de 45%, neste caso, as percentagens totais são inferiores as apresentadas quando existem modelos equivalentes. Na Figura 5.18b ilustramos o que acontece para Sim 4a sendo observado que o padrão das observações é similar ao padrão da Figura 5.13b. Além disso, a variável  $q$  (definida como o quociente entre a volatilidade estimada GARCH(1,1) e a volatilidade verdadeira AR(1)-SV) apresenta valores muito altos em alguns casos indicando grande superestimação.

**Tabela 5.19**  
**Magnitude da subestimação e superestimação da volatilidade estimada (vg) GARCH(1,1) com respeito à volatilidade verdadeira(vs) AR(1)-SV quando não existe equivalência**

	Diferença maior que 20%			Diferença maior que 45%		
	vg>1,2vs	vg<0,8vs	Total	vg>1,45vs	vg<0,55vs	Total
Sim4a	49,87	26,26	76,13	37,97	9,76	47,73
Sim 5a	48,30	25,63	73,93	35,53	9,17	44,7

Quanto à previsão um-passo à frente, os resultados da Tabela 5.20 mostram que o desempenho dos modelos GARCH(1,1) e AR(1)-SV são similares ao considerar a percentagem dos retornos fora dos i.c 95%. Mas, quando comparamos os resultados considerando os níveis da volatilidade as diferenças entre essas abordagens são grandes, sendo que no nível alto a percentagem dos retornos fora dos i.c é muito maior que a esperada.

**Tabela 5.20**  
**Percentagem de retornos fora dos Intervalos de Confiança 95% quando o modelo simulado é AR(1)-SV e não existe equivalência**

Nível da volatilidade	Sim 4a		Sim 5a	
	baixo	0,9	4,8 <sup>a</sup>	4,1
médio	4,1	4,3	1,0	4,7
alto	12,4	4,7	11,7	4,6
% Total	5,8	4,6	5,6	4,6

<sup>a</sup> corresponde a percentagem observada com os i.c construídos com a volatilidade verdadeira AR(1)-SV

Finalmente, ao observar as percentagens totais da Tabela 5.20 e da Tabela 5.14 podemos dizer que os resultados são comparáveis, mas, ao considerar os níveis da volatilidade, não.

## 5.5 Conclusões

As volatilidades estimadas assim como as previsões um passo à frente fornecidas pelas abordagens ARCH e Variância Estocástica foram comparadas sob dois aspectos, empiricamente mediante o ajuste das *Séries Exemplo* e através de simulações. A seguir apresentamos as principais conclusões:

### *Ajuste das Séries Exemplo:*

- Na comparação das volatilidades estimadas nas *Séries Exemplo* encontramos que para uma proporção muito grande de retornos as diferenças entre as volatilidades estimadas ARCH e Variância Estocástica são consideráveis, sendo que, na maioria das vezes a volatilidade estimada ARCH é maior que a volatilidade estimada pelo modelo AR(1)-SV estacionário.
- Com relação à previsão um-passo à frente e utilizando o critério dos retornos observados, encontramos que o desempenho dos modelos ARCH e Variância Estocástica são similares. Dado ao exposto anteriormente, principalmente em períodos posteriores à ocorrência de observações extremas nos quadrados dos retornos, pelo menos para as *Séries Exemplo*, o modelo AR(1)-SV é mais adequado.
- Na série das *Ações* a persistência estimada pelo modelo GARCH(1,1) é muito maior à apresentada pelo modelo AR(1)-SV estacionário. Já na *Taxa de Câmbio* a persistência estimada pelo modelo AR(1)-SV estacionário é maior que à correspondente ao modelo GARCH(1,1).

### *Simulações quando existem modelos equivalentes:*

- Os resultados das simulações evidenciam que, quando o modelo simulado é GARCH(1,1), para uma grande proporção dos retornos a volatilidade estimada pelo modelo AR(1)-SV

estacionário superestima ou subestima a volatilidade GARCH(1,1) consideravelmente. Por sua vez, quando simulamos o modelo AR(1)-SV encontramos que também há superestimação e subestimação considerável para bastantes retornos, sendo que, nesse caso, a superestimação é mais freqüente quando comparada com o caso de gerar o GARCH(1,1). Quando o modelo é GARCH(1,1) o ajuste do modelo AR(1)-SV superestima ( $vs > 1,2vg$ ) com menos freqüência do que quando o modelo é AR(1)-SV e a volatilidade é estimada através do modelo GARCH(1,1). Em relação à freqüência de subestimação, quando consideramos a freqüência em que existe subestimação (volatilidade estimada menor que volatilidade verdadeira) ocorre o contrário.

- Quanto à previsão um-passo à frente temos que se o modelo simulado é GARCH(1,1), os desempenhos dos modelos AR(1)-SV e GARCH(1,1) não são muito diferentes (em termos da porcentagem de retornos observada dentro dos i.c.) sempre que a persistência seja maior ou igual a 0,97. Quando a persistência é igual a 0,95, devido à subestimação comentada anteriormente, a freqüência de valores de retornos fora dos i.c. construídos através do modelo AR(1)-SV é maior. Por sua vez, quando o modelo simulado é AR(1)-SV, o desempenho desse modelo é similar ao apresentado pelo GARCH(1,1).

Do exposto acima não podemos concluir se um modelo é superior ao outro.

*Simulações quando não existem modelos equivalentes:*

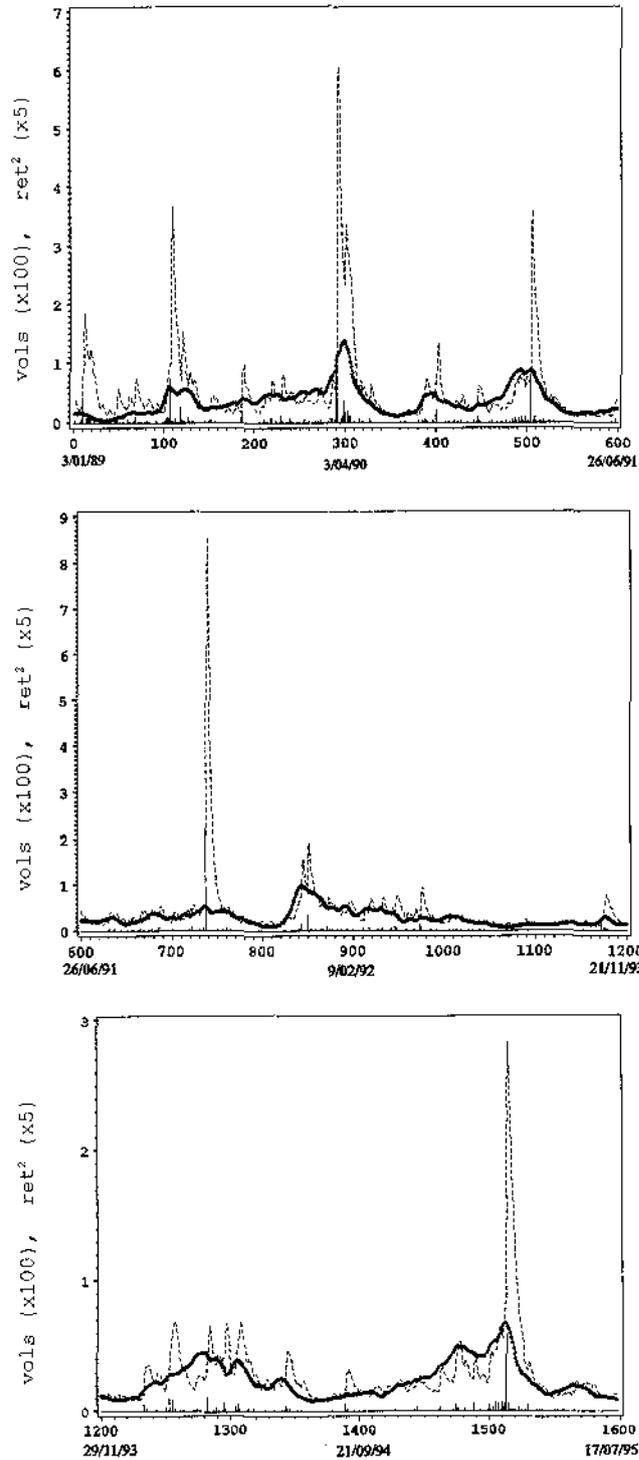
- Ao simular o modelo GARCH(1,1) observamos que a volatilidade estimada pelo modelo AR(1)-SV superestima e subestima a verdadeira volatilidade para uma grande proporção dos retornos, sendo que essas porcentagens são menores às obtidas quando existem modelos equivalentes (comparando os modelos com igual persistência). Essa mesma situação acontece quando o modelo simulado é AR(1)-SV.

- Tanto ao simular o GARCH(1,1) como quando simulamos o AR(1)-SV, observamos que para uma quantidade bastante grande de retornos existe superestimação e subestimação de forma considerável. A superestimação é mais freqüente quando o modelo simulado é o GARCH(1,1) e ajustado o AR(1)-SV do que quando o modelo simulado é AR(1)-SV. Em relação à subestimação, as freqüências observadas ao simular o GARCH(1,1) são similares às encontradas ao simular o AR(1)-SV.

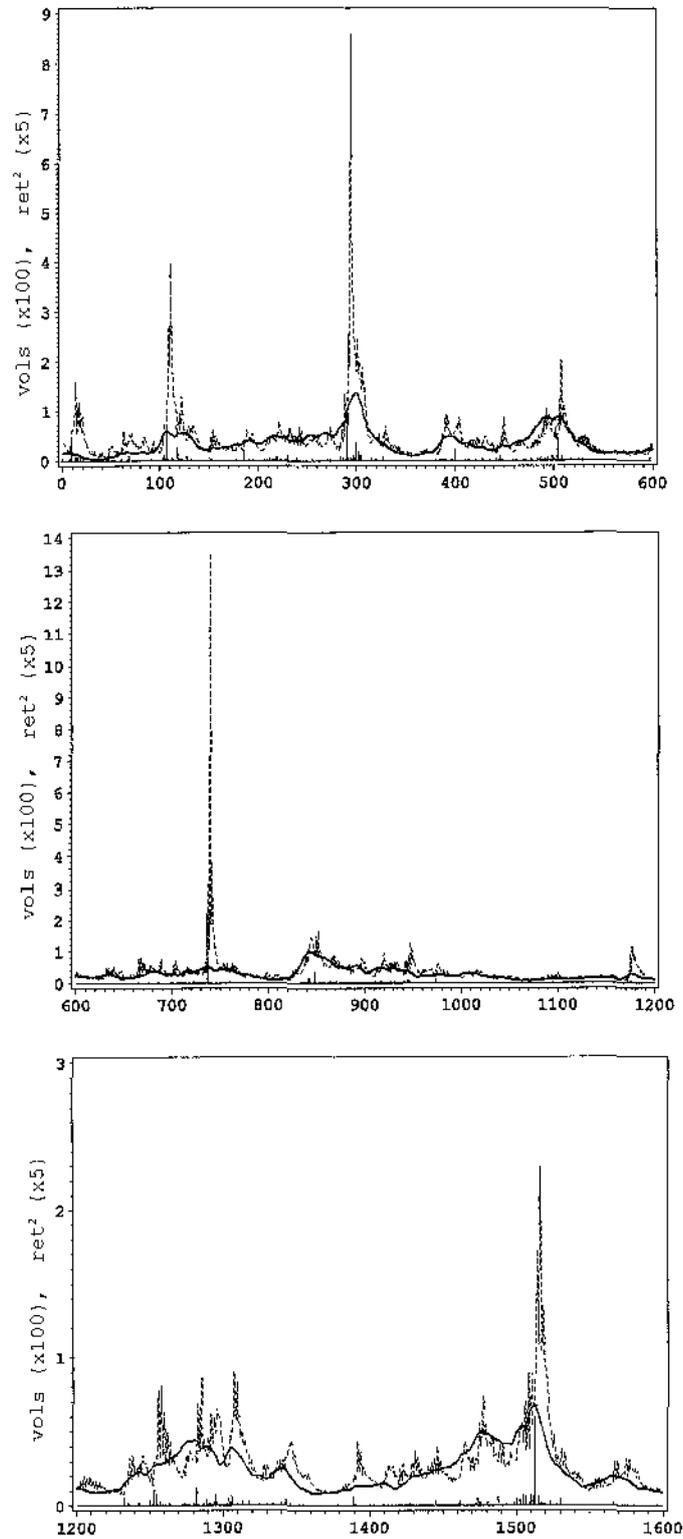
- Quando o modelo GARCH(1,1) é simulado, as percentagens observadas de superestimação e subestimação são menores do que quando existe modelo equivalente AR(1)-SV. Esta aparente contradição pode ser explicada pelo fato de que no caso de não existir modelos equivalentes, os parâmetros são estimados, ou seja, é escolhido o valor do parâmetro de forma que seja escolhido o modelo AR(1)-SV que melhor ajuste os dados.

- Na previsão um-passo à frente, ao simular o GARCH(1,1) encontramos que os desempenhos dos modelos GARCH(1,1) e AR(1)-SV são similares (em termos da percentagem de retornos observada dentro dos i.c). O mesmo ocorre quando o modelo gerado é AR(1)-SV e estimamos o GARCH(1,1). Esses desempenhos são similares aos encontrados no caso quando há equivalência entre os modelos.

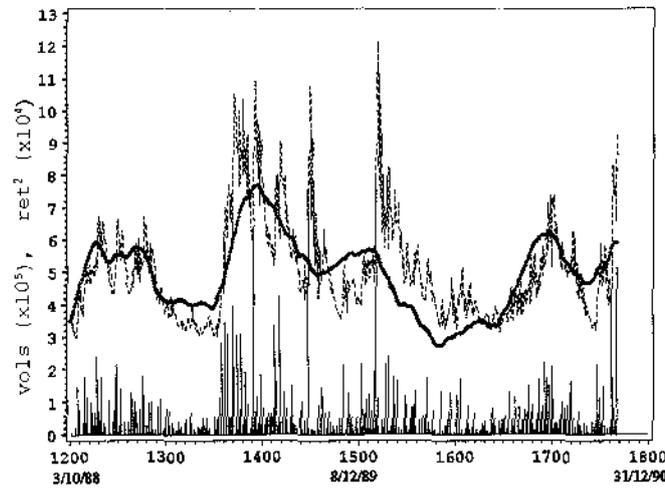
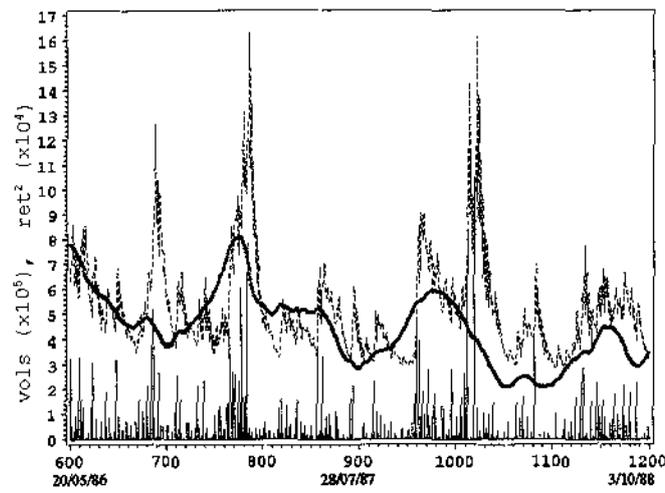
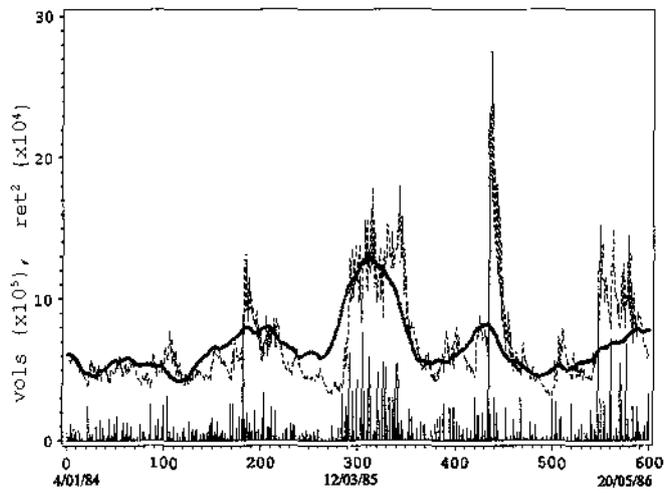
-Resumindo, neste caso existe uma leve superioridade do modelo GARCH(1,1) dada a maior frequência de superestimação no ajuste de modelos AR(1)-SV quando o modelo verdadeiro é GARCH(1,1).



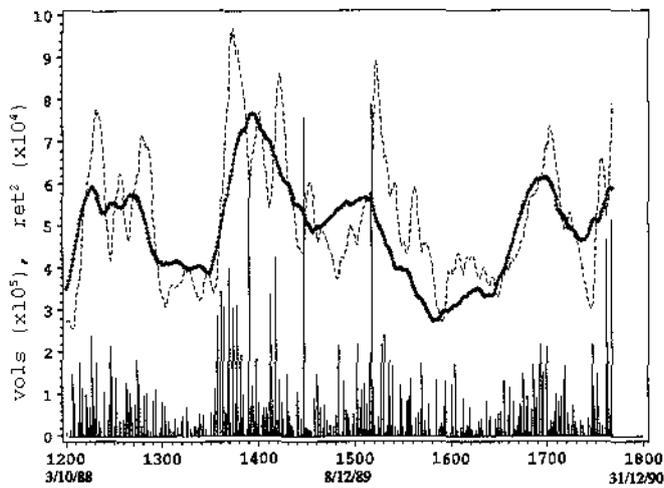
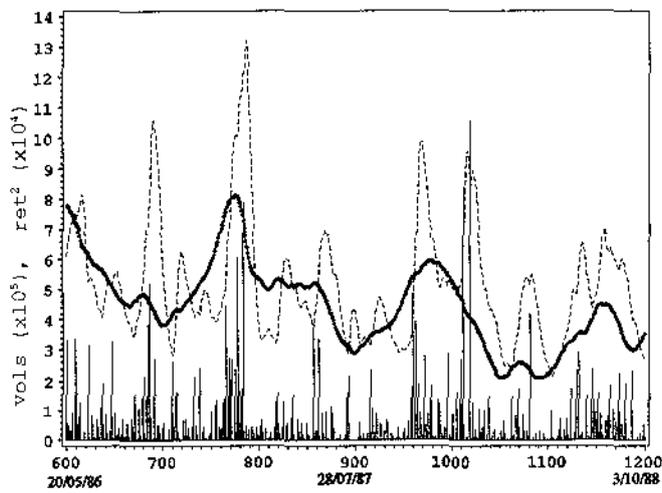
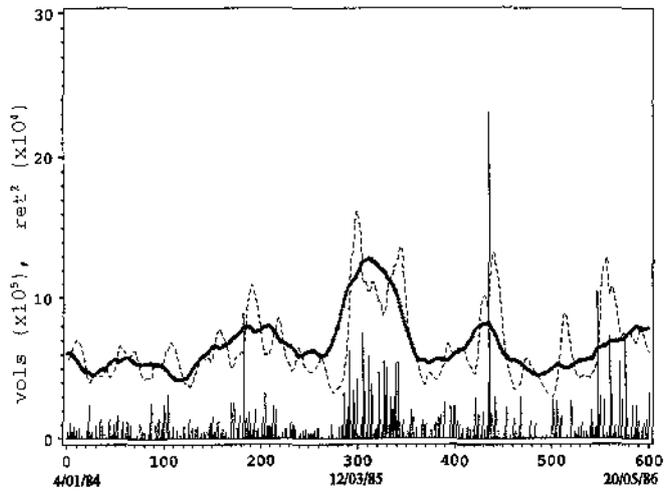
**Figura 5.1: Volatilidades estimadas GARCH(1,2) e AR(1)-SV nas Ações**  
**Legenda:** GARCH(1,2) - - - - , AR(1)-SV - - - -  
 vols=volatilidades,  $ret^2$ =quadrados dos retornos



**Figura 5.2: Volatilidades estimadas EGARCH(2,2) e AR(1)-SV nas Ações**  
**Legenda: EGARCH(2,2) - - -, AR(1)-SV ———**  
 vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos



**Figura 5.3: Volatilidades estimadas GARCH(2,2) e AR(1)-SV na Taxa de Câmbio**  
**Legenda: GARCH(2,2) - - - , AR(1)-SV ———**  
 vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos

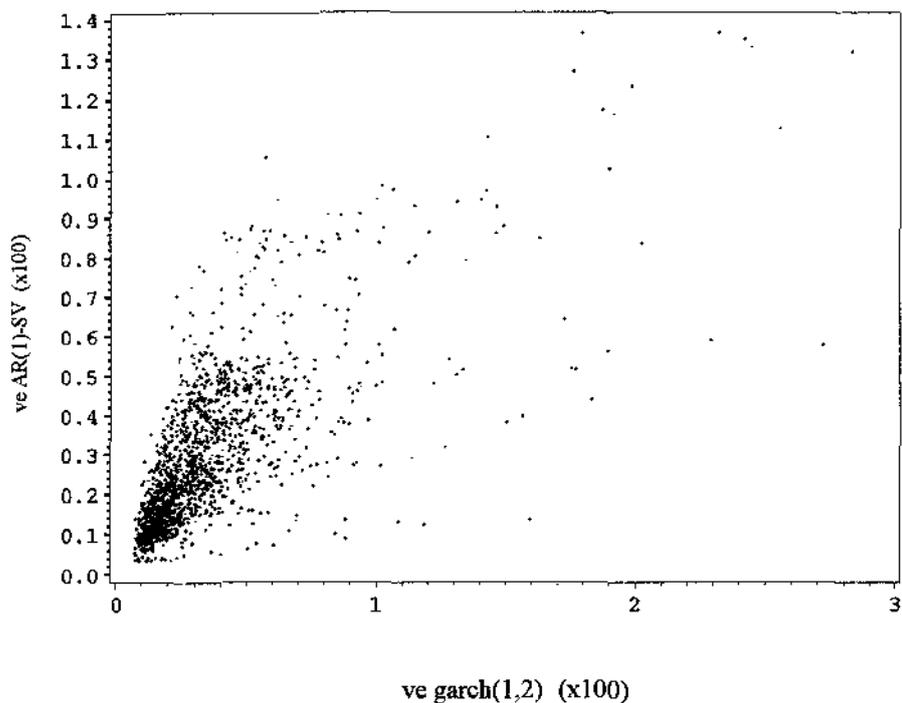


**Figura 5.4: Volatilidades estimadas EGARCH(2,0) e AR(1)-SV na Taxa de Câmbio**

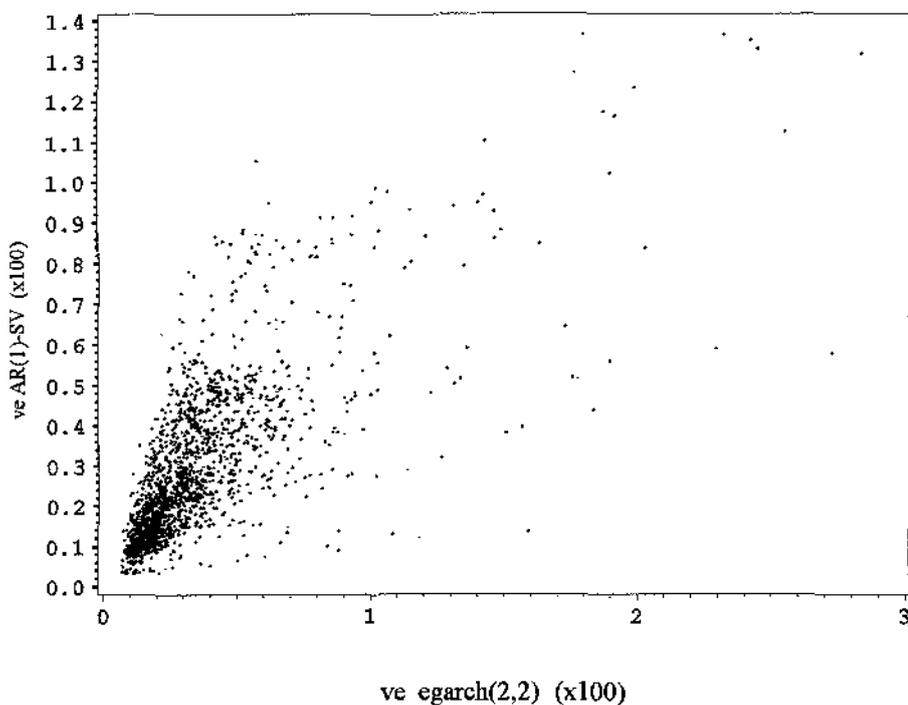
**Legenda:** EGARCH(2,0) - - - , AR(1)-SV ———

vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos

a) GARCH(1,1) e AR(1)-SV



b) EGARCH(2,2) e AR(1)-SV



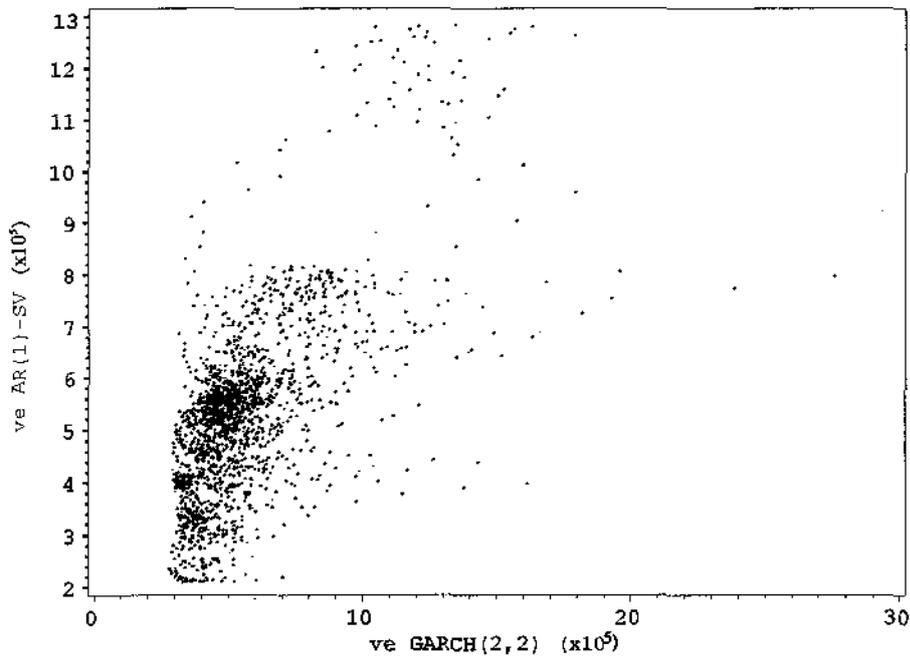
**Figura 5.5: Comparação das volatilidades estimadas ARCH e Variância Estocástica nas Ações<sup>a</sup>**

**Legenda:** ve=volatilidade estimada.

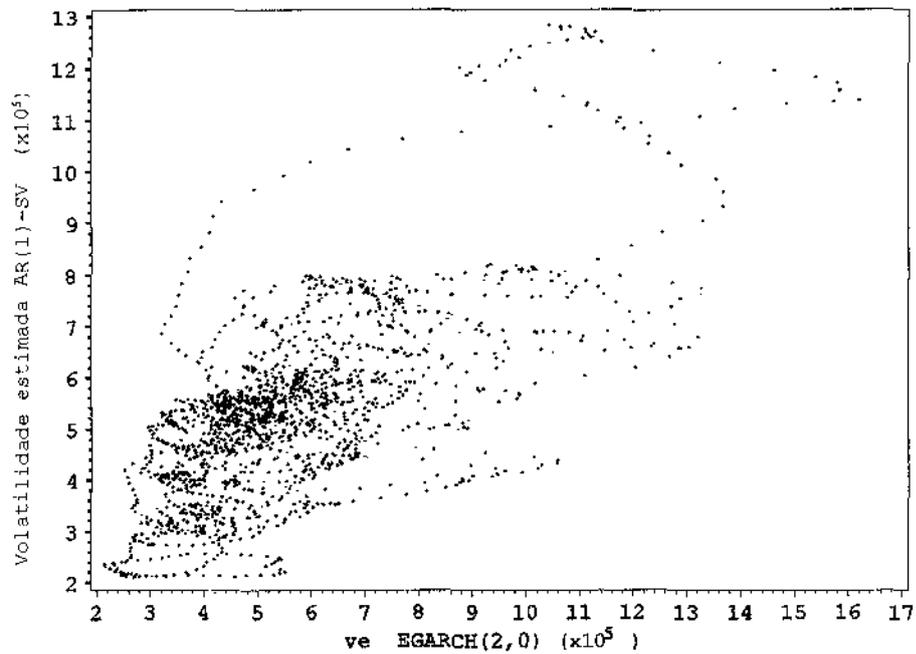
<sup>a</sup> Não foram considerados os retornos que: a)  $100 \cdot ve \text{ garch}(1,2) > 3$  b)  $100 \cdot ve \text{ egarch}(1,2) > 3$ .

Nos dois casos o número total de pontos é 1592 (99,56%)

a) GARCH(2,2) e AR(1)-SV



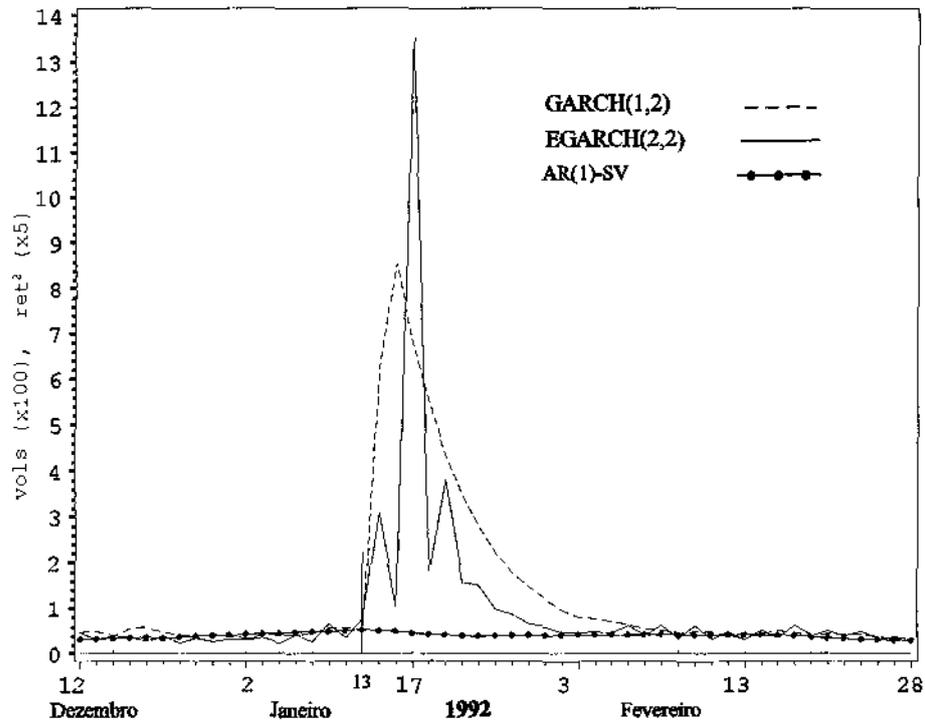
b) EGARCH(2,0) e AR(1)-SV



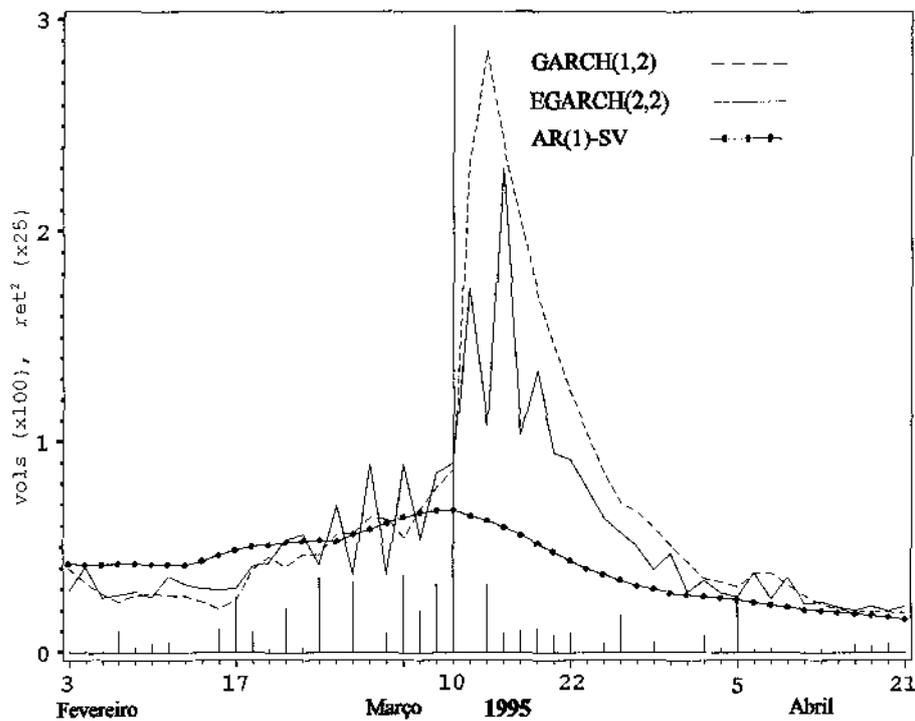
**Figura 5.6: Comparação das volatilidades estimadas ARCH e AR(1)-SV na Taxa de Câmbio**

**Legenda:** ve= volatilidade estimada

a) *Queda da Ministra Zélia Cardoso de Mello*

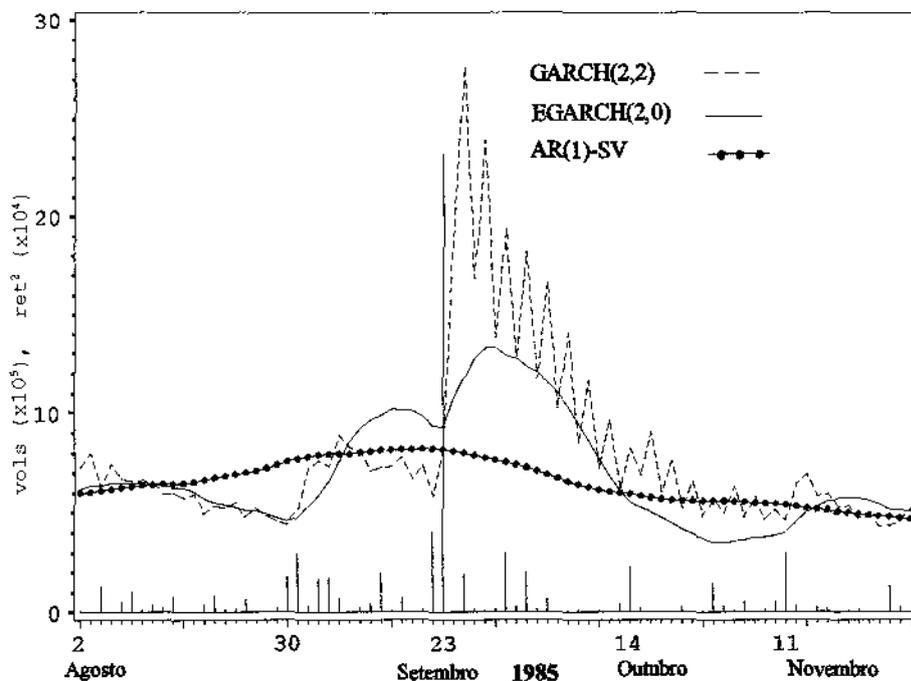


b) *Mudança da Banda Cambial*

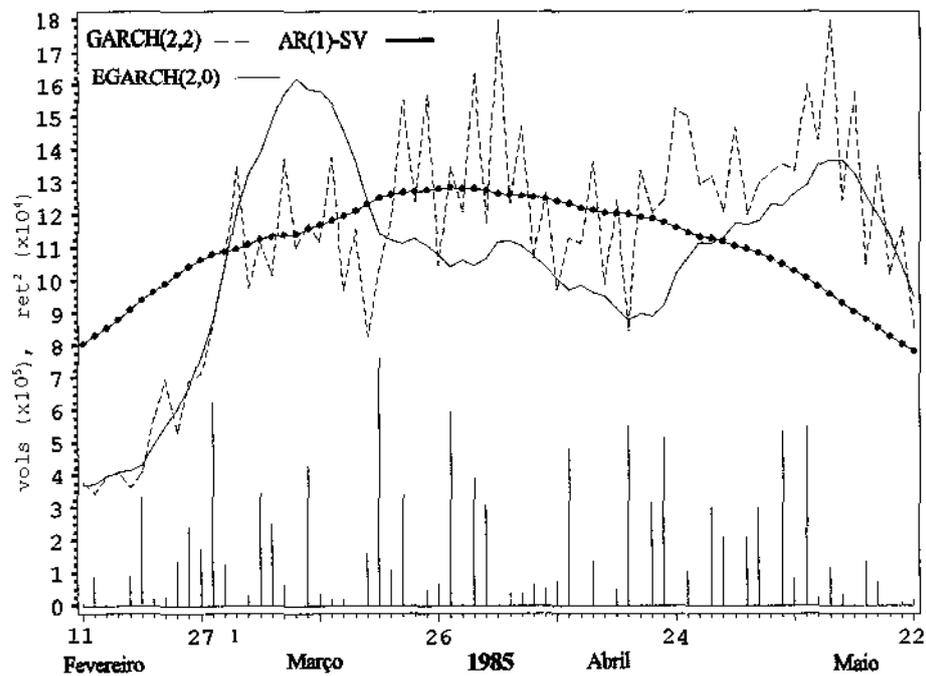


**Figura 5.7: Influência das observações extremas dos retornos nas volatilidades estimadas ARCH e Variância Estocástica**  
**Legenda:** vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos

a) Período com uma observação extrema

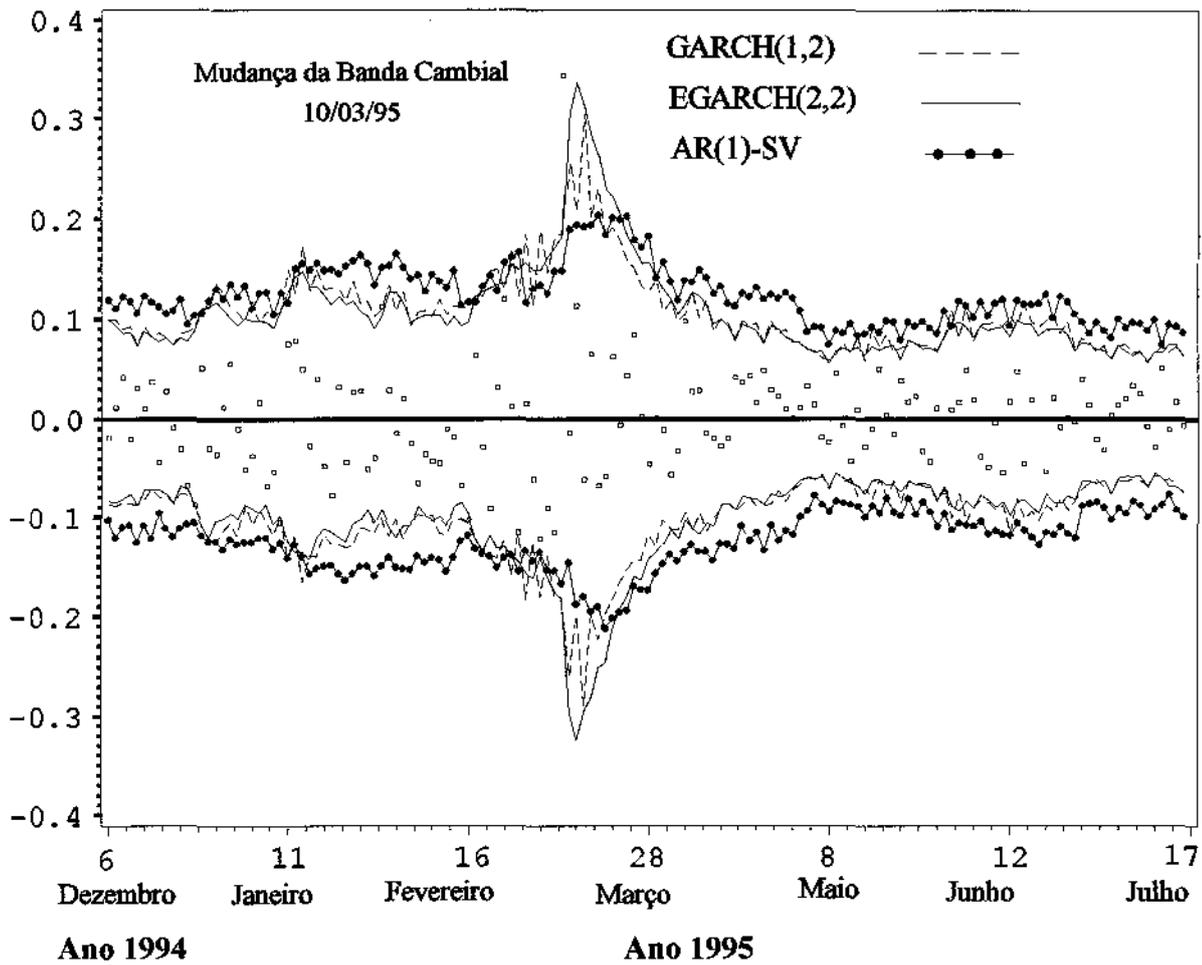


b) Período com um conjunto de grandes valores de retornos

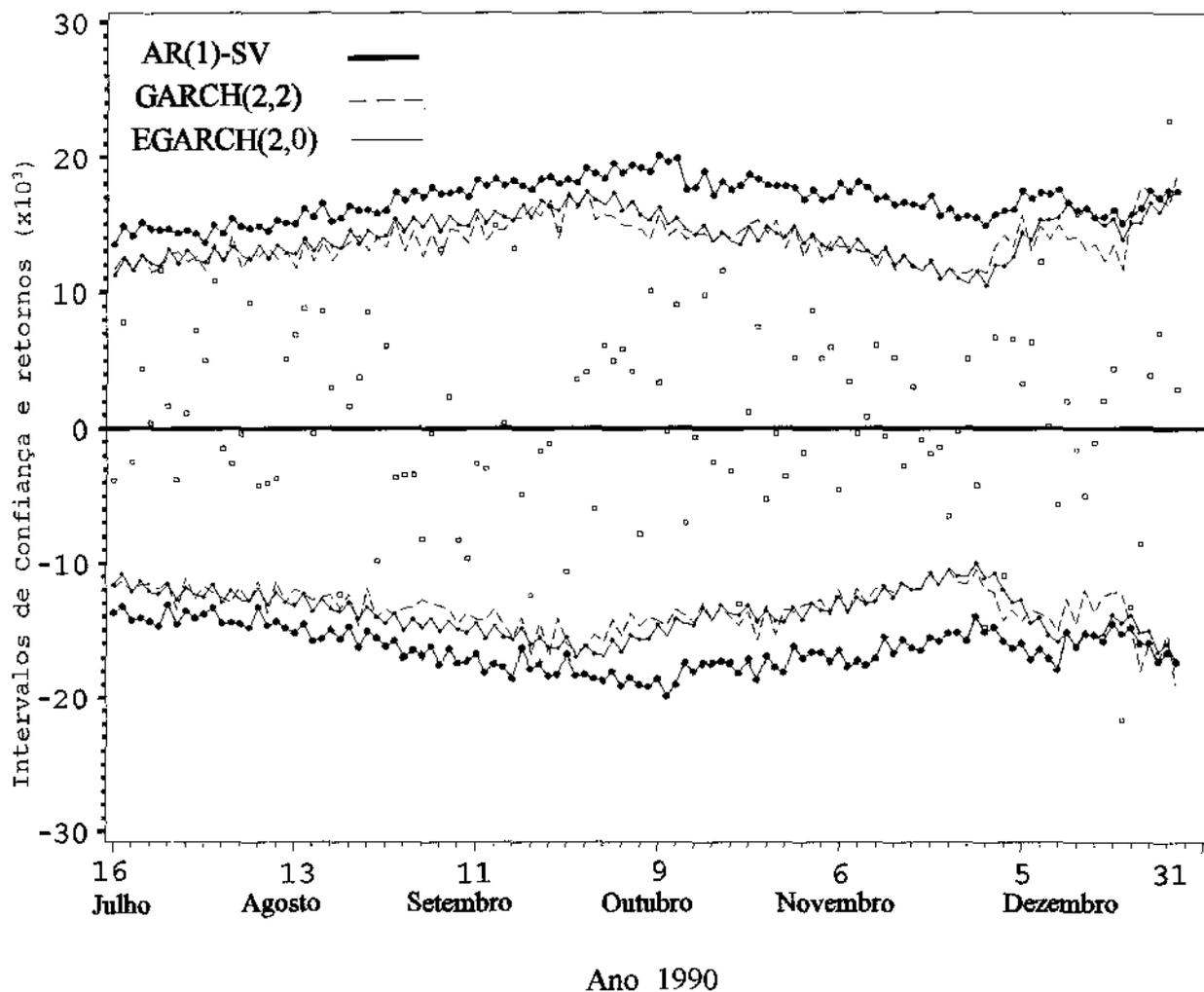


**Figura 5.8: Volatilidades estimadas ARCH e Variância Estocástica em dois períodos da Taxa de Câmbio**

**Legenda:** vols=volatilidades estimadas,  $ret^2$ =quadrados dos retornos

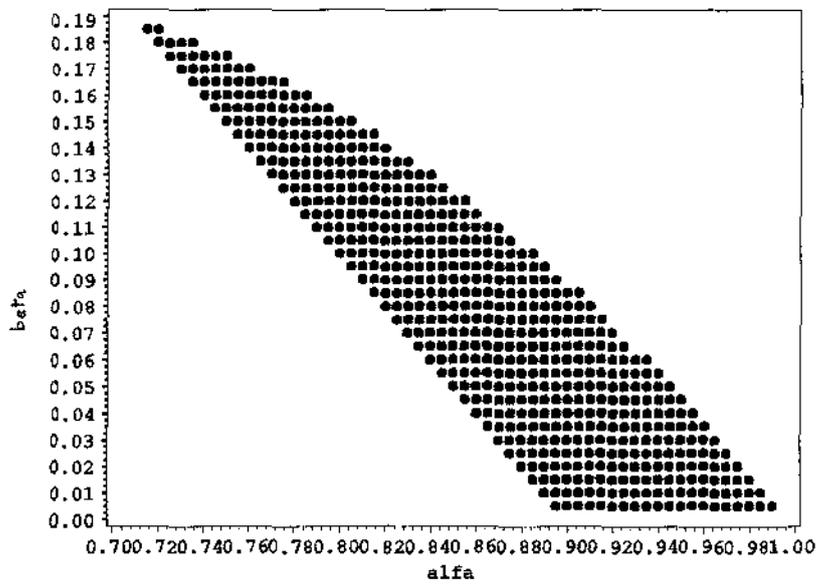


**Figura 5.9: Comparação dos intervalos de confiança 95% para os retornos obtidos pelos modelos ARCH e Variância Estocástica nas Ações**

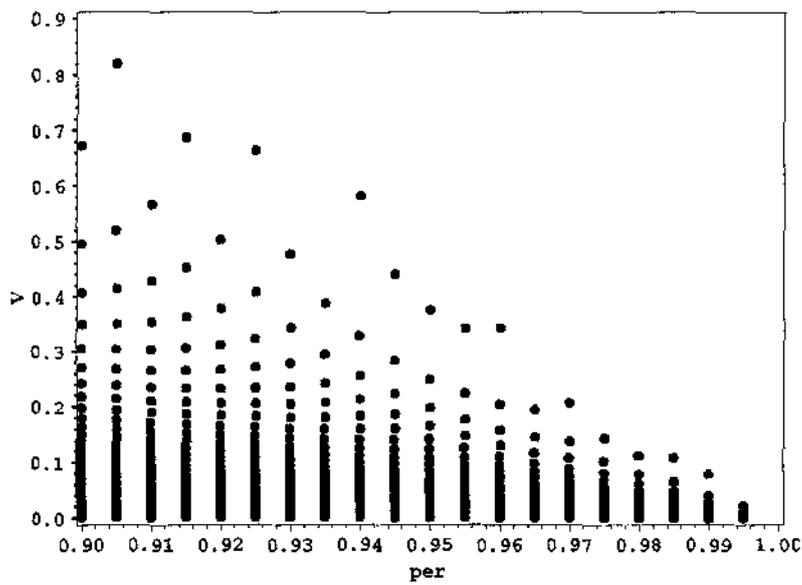


**Figura 5.10: Comparação dos intervalos de confiança 95% para os retornos obtidos pelos modelos ARCH e Variância Estocástica na Taxa de Câmbio**

a) Região GARCH(1,1)



b) Região AR(1)-SV



**Figura 5.11: Regiões de equivalência dos modelos GARCH(1,1) e AR(1)-SV estacionário sob alta persistência**

**Legenda:** alfa= $\alpha$ , beta= $\beta$ , per= $\gamma$ , v= $\sigma^2_{\eta}$

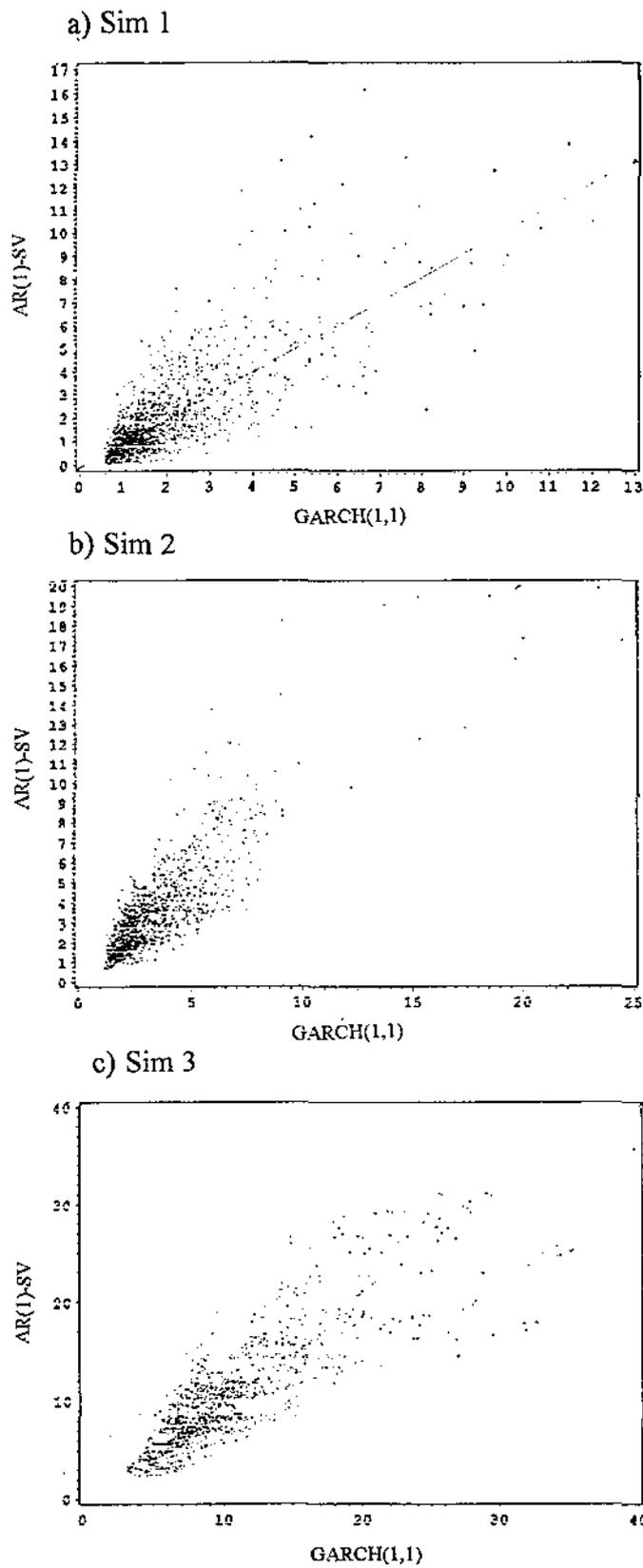
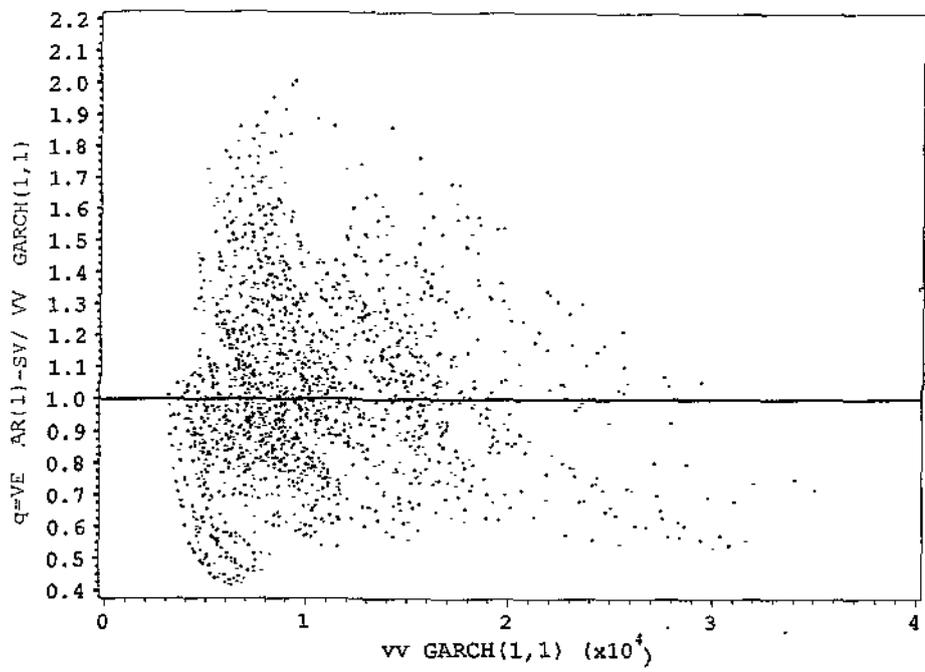


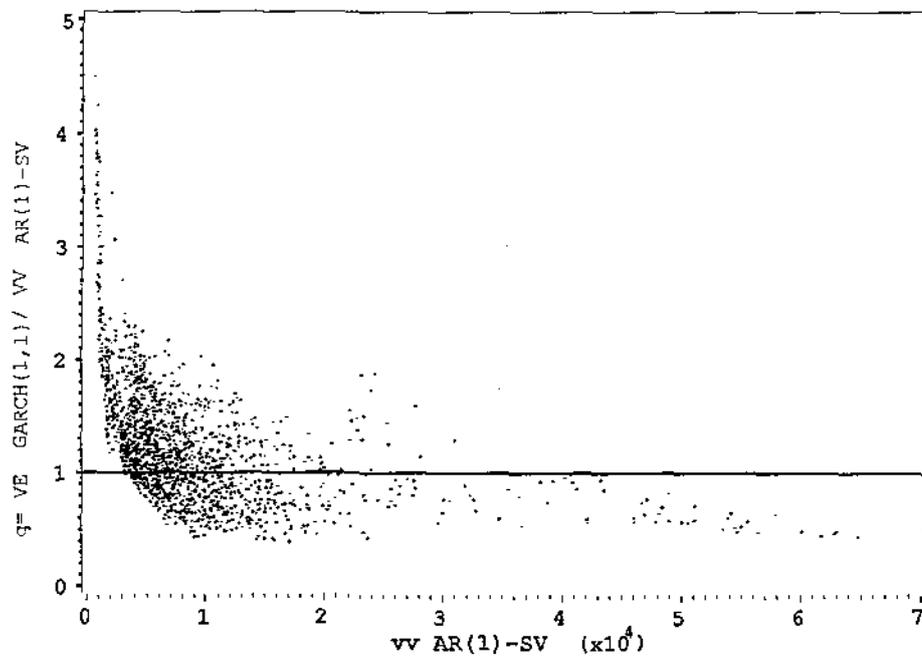
Figura 5.12: Comparação das volatilidades geradas GARCH(1,1) e estimadas AR(1)-SV

Legenda: volatilidades ( $\times 10^5$ )

a) Sim 3

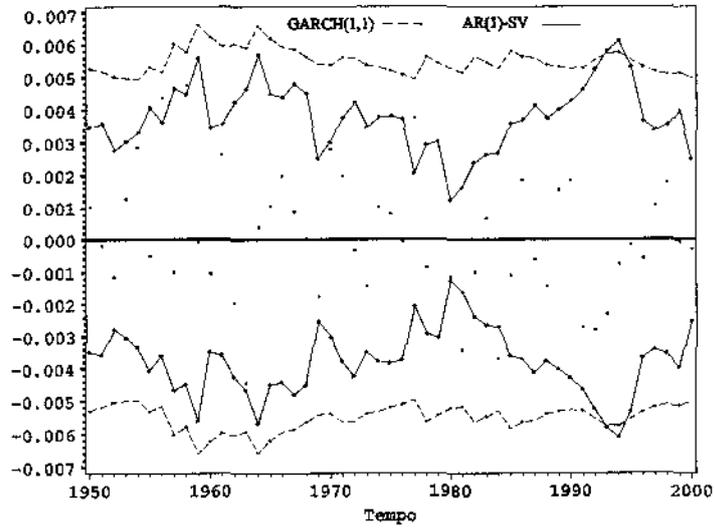


b) Sim 3a

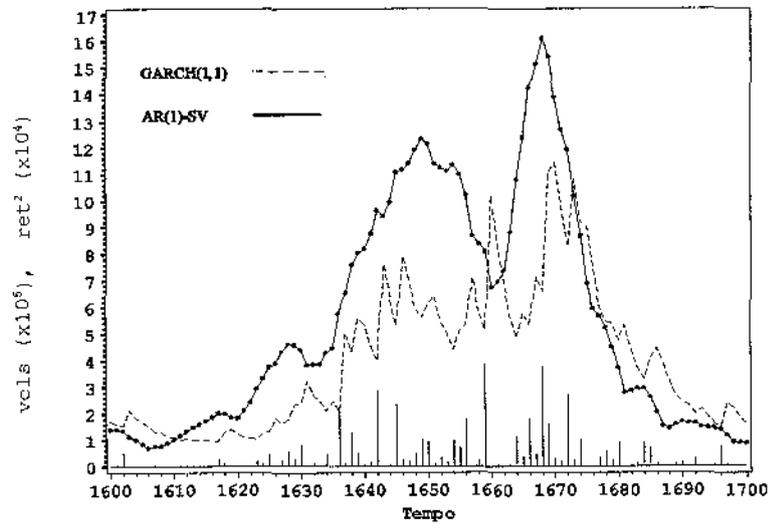


**Figura 5.13: Comparação das volatilidades geradas e estimadas quando existem modelos equivalentes**

**Legenda:** VV= volatilidade verdadeira, VE=volatilidade estimada

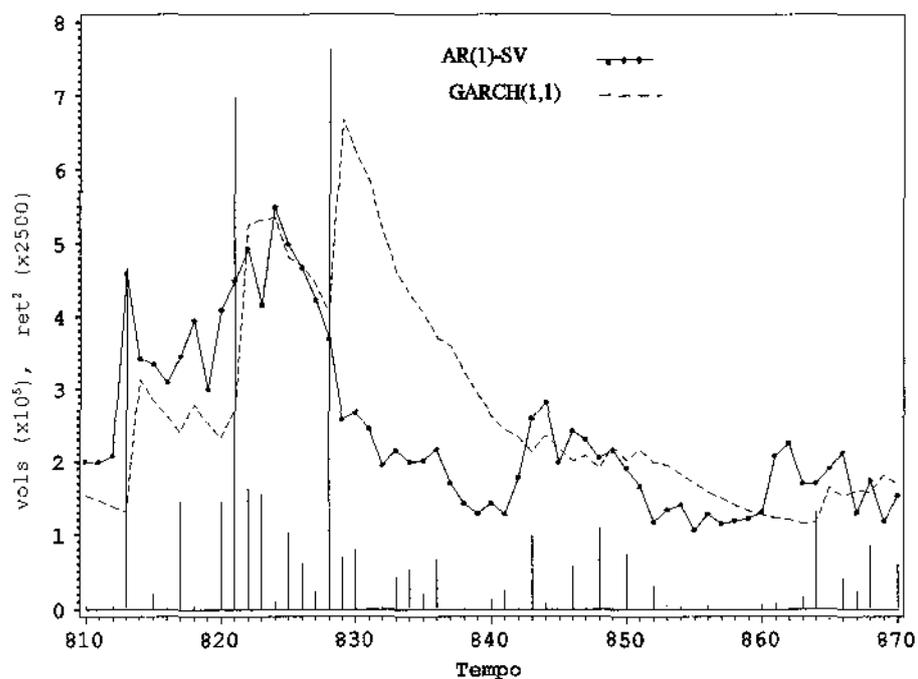


**Figura 5.14: i.c 95% para os retornos em um período de Sim 1**

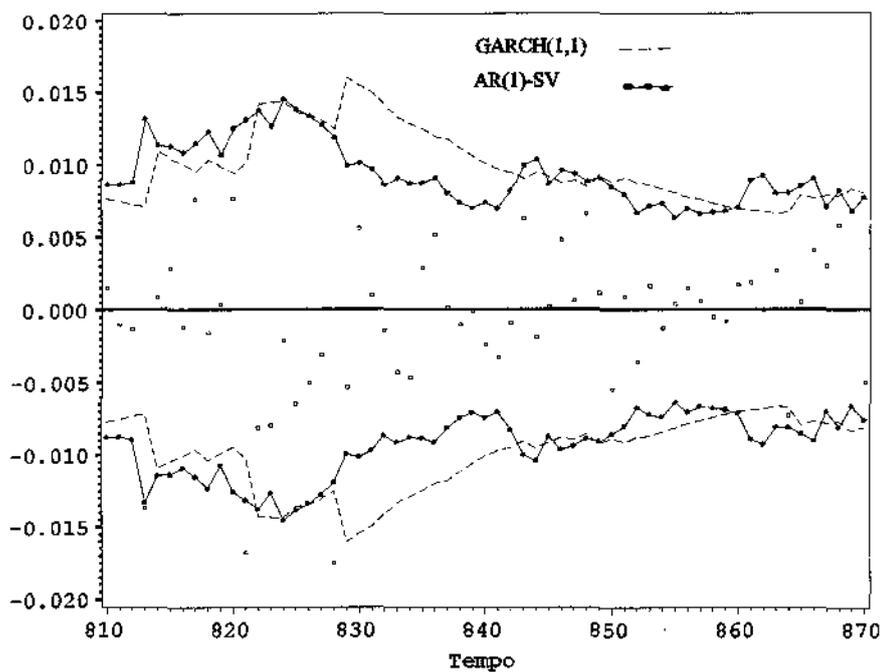


**Figura 5.15: volatilidades GARCH(1,1) e AR(1)-SV em um período de Sim 1**

**Legenda:** vols=volatilidades,  $ret^2$ =quadrados dos retornos

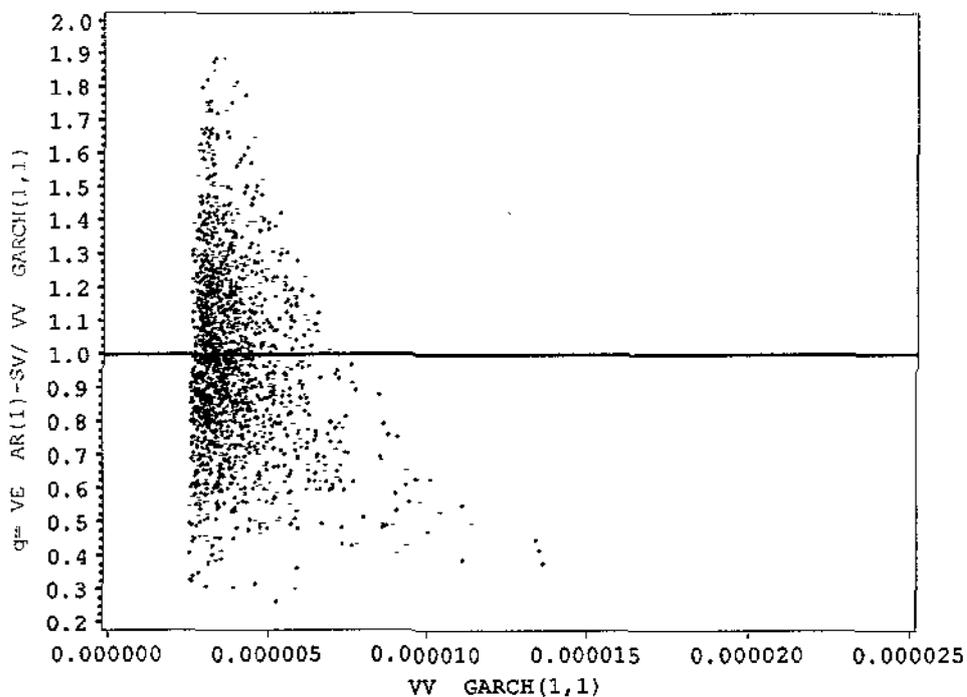


**Figura 5.16: Volatilidades GARCH(1,1) e AR(1)-SV em um período de Sim2a**  
**Legenda:** vol=volatilidades, ret<sup>2</sup>=quadrados dos retornos

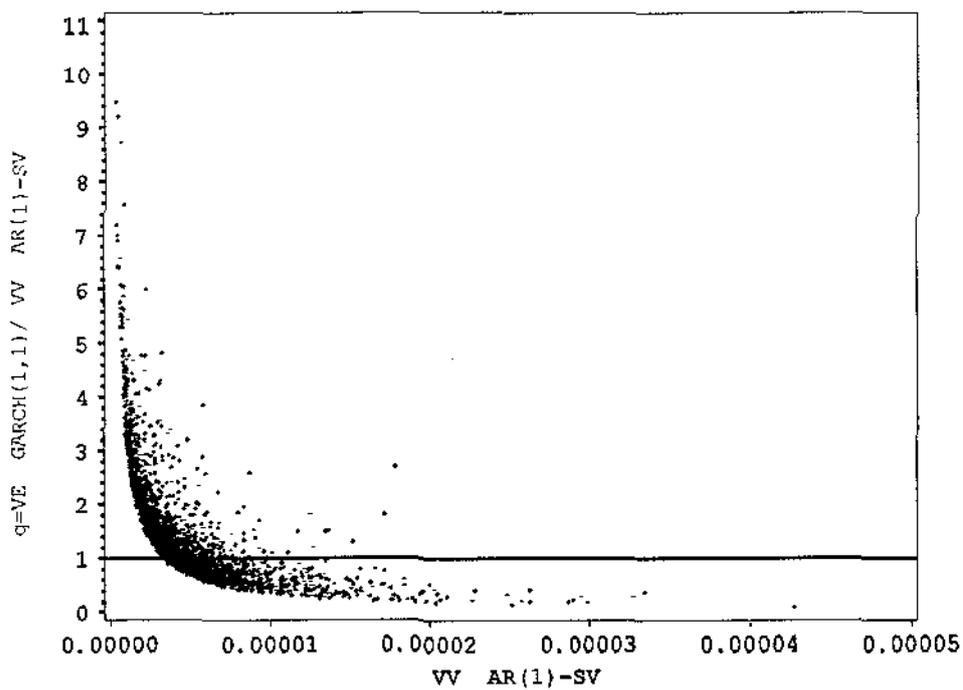


**Figura 5.17: i.c 95% para os retornos em um período de Sim2a**

a) Sim 4



b) Sim 4a



**Figura 5.18: Comparação das volatilidades geradas e estimadas quando não existem modelos equivalentes**

**Legenda:** VV=volatilidade verdadeira, VE=volatilidade estimada

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAMOVITZ, M. & STEGUM, N.C. (1970). *Handbook of Mathematical Functions*. New York : Dover Publications Inc.
- AIT-SAHALIA, Y. (1992). "The delta and bootstrap methods for nonparametric kernel functionals". Mimeo, MIT.
- \_\_\_\_\_ (1996). "Nonparametric pricing of interest rate derivative securities". *Econometrica*. **64**: 527-560.
- AKAIKE, H. (1973). "Information theory and extension of the maximum likelihood principle". In: *Second International Symposium on Information Theory*. 267-281. B.N. Petrov & F. Csáki (eds). Budapest: Akadémiai Kiadó.
- ANDERSEN, T.G. (1994). "Stochastic autoregressive volatility: a framework for volatility modelling". *Mathematical Finance*. **4**: 75-102.
- BAILLIE, R.T. & BOLLERSLEV, T. (1991). "Intra day and inter day volatility in foreign exchange rates". *Review of Economic Studies*. **58**: 565-585.
- \_\_\_\_\_ (1992). "Prediction in dynamic model with time-dependent conditional variances". *Journal of Econometrics*. **52**: 91-113.
- BACHELIER, L. (1900). *Théorie de la Spéculation*. Annales de l'Ecole Normale Supérieure, 3rd series, 17, 21-86- Trans.by A.J. BONESS in *The Random Character of Stock Market Prices*, ed. P.H. Cootner, Cambridge, Mass: MIT Press, 1967.
- BERNDT, E.R.; HALL, B.M.; HALL, R.E.S & HAUSMAN, J.A. (1974). "Estimation and inference in nonlinear structure models". *Annals of Economic and Social Measurement*. **4**: 653-665.
- BLACK, F. (1976). "Studies in stock price volatility changes". *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statistical Association*. 177-181.
- BLACK, F. & SCHOLES, M. (1973). "The pricing of options and corporate liabilities". *Journal of Political Economy*. **81**: 637-654.

- BOLLERSLEV, T. (1986). "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics* .31: 307-327.
- \_\_\_\_\_ (1987). "A conditionally heteroskedastic time series model for speculative prices and rate of return". *Review of Economics and Statistics* .69: 542-547.
- \_\_\_\_\_ (1988). "On the correlation structure for the generalized autoregressive conditional heteroskedastic process". *Journal of Time Series Analysis* .9: 121-131.
- BOLLERSLEV, T., CHOU, Y.C. & KRONER, K. (1992). "ARCH modelling in finance: a selective review of the theory and empirical evidence". *Journal of Econometrics* .52: 201-224.
- BOLLERSLEV, T. & ENGLE, R.F. (1993). "Common persistence in conditional variances: definition and representation". *Econometrica* .61: 166-187.
- BOLLERSLEV, T., ENGLE, R.F. & NELSON, D. (1994). "ARCH Models". *Handbook of Econometrics* . Volume IV, Cap. 49: 2959-3038. Engle, R.F. & McFadden, D. (eds). Amsterdam: North Holland.
- BOLLERSLEV, T. & GHYSELS, E. (1996). "Periodic autoregressive conditional heteroscedasticity". *Journal of Business & Economic Statistics* .14: 139-151.
- BOLLERSLEV, T. & WOOLDRIDGE, J.M. (1992) "Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances". *Econometric Reviews* .11: 143-172.
- BOUGEROLL, P. & PICARD, N. (1992). "Stationary of GARCH processes and of some non-negative time series". *Journal of Econometrics* .52: 115-127.
- BOWMAN, K.O. & SHENTON, L.R. (1975). "Omnibus contours for departures from normality based on  $\sqrt{b_1}$  and  $b_2$ ". *Biometrika* .62: 243-250.
- BROCK, W.A., DECHERT, W.D. & SCHEINKMAN, J.A. (1987). "A test for independence based on the correlation dimension". Mimeo, University of Wisconsin.
- BROCK, W.A., DECHERT, W.D., SCHEINKMAN, J.A. & LE BARON, B. (1996). "A test for independence based on the correlation dimension". *Econometric Reviews* .15: 197-235.
- BROCKWELL, P.J. & DAVIS, R.A. (1991). *Time Series: Theory and Methods*. New York: Second Edition Springer Verlag.
- BROWN, R.L., DURBIN, J. & EVANS, J.M. (1975). "Techniques for testing the constancy of regressions relationships over time (with discussion)". *Journal of the Royal Statistical Society B* .37: 149-192.

- CAO, C.Q. & TSAY, R.S. (1993). "Nonlinear time-series analysis of stock volatilities". In: *Nonlinear Dynamics, Chaos and Econometrics*, Pesaran, M.H. e Potter, M. (eds) John Wiley & Sons.
- CHESNEY, M. & SCOTT, L. (1989). "Pricing european currency options: a comparison of the modified Black-Scholes model and a random variance model". *Journal of Financial and Quantitative Analysis*. **24**: 267-284.
- CHRISTIE, A.A. (1982). "The stochastic behavior of common stock variances: value, leverage and interest rate effects". *Journal of Financial Economics*. **10**: 407-432.
- CLARK, P.K. (1973). "A subordinated stochastic process model with finite variance for speculative prices". *Econometrica*. **41**: 135-156.
- DANIELSSON, J. (1994). "Stochastic volatility in asset prices: estimation with simulated maximum likelihood". *Journal of Econometrics*. **61**: 375-400.
- DASSIOS, A. (1992). "Asymptotic expressions for approximations to stochastic variance models". Mimeo. London School of Economics.
- DeJONG, P. (1991). "The diffuse Kalman filter". *Annals of Statistics*. **19**:1073-1083.
- DE LIMA, P.J.F. (1996). "Nuisance parameter free properties of correlation integral based statistics". *Econometric Reviews*. **15**: 237-259.
- DELGADO, S. (1995). "A distribuição de probabilidade dos retornos das ações no Brasil: uma abordagem não paramétrica". *Revista de Econometria*. **15**: 77-100.
- DEMPSTER, A.P., LAIRD, N.M. & RUBIN, D.B. (1977). "Maximum likelihood from incomplete data via the EM algorithm". *Journal of the Royal Statistical Society Series B*. **39**: 1-38.
- DIEBOLD, F.X & SCHUERMAN, T (1992) "Exact maximum likelihood estimation of ARCH models". Mimeo. University of Pennsylvania.
- DUNSMUIR, W. (1979). "A central limit theorem for parameter estimation in stationary time series and its applications to models for a signal observed white noise". *Annals of Statistics*. **7**: 490-506.
- DUAN, J.C. (1995). "The GARCH option pricing model". *Mathematical Finance*. **5**: 13-32.
- ENGLE, R. F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation". *Econometrica*. **50**: 987-1007.
- ENGLE, R. F. & BOLLERSLEV, T. (1986). "Modelling the persistence of conditional variances" (with discussion). *Econometrics Reviews*. **5**: 1-50, 81-87.

- ENGLE, R.F. & GONZÁLEZ-RIVERA, G. (1991). "Semiparametric ARCH models." *Journal of Business and Economic Statistics*. 9:345-359.
- ENGLE, R.F., HENDRY, D. F. & TRUMBLE, D. (1985). "Small sample properties of ARCH estimators and tests". *Canadian Journal of Economics*. 18: 66-93.
- ENGLE, R.F.; KANE, A. & NOH, J. (1993a). "Option index pricing with stochastic volatility and the value of accurate variance forecasts". Manuscript, UCSD Economic Dept.
- ENGLE, R.F, HONG, C-H, KANE, A. & NOH, J (1993b). "Arbitrage valuation of variance forecasts with simulated options". *Advances in Futures and Options Research*. 6: 393-415.
- ENGLE, R.F; LILIEN, D.M. & ROBINS, R.P (1987). "Estimative time varying risk premia in the term structure: The ARCH-M Model". *Econometrica*. 55: 391-407.
- ENGLE, R.F. & MUSTAFA. (1992). "Implied ARCH Models for options prices". *Journal of Econometrics*. 52: 289-311.
- FAMA, E.F. (1963). "Mandelbrot and the stable Paretian distribution". *Journal of Business*. 36: 420-429.
- \_\_\_\_\_ (1965). "Behavior of stock market prices". *Journal of Business*. 38: 34-105.
- \_\_\_\_\_ (1970). "Efficient capital market: a review of theory and empirical work". *Journal of Finance*. 25: 383-423.
- FRENCH, K.R. & ROLL, R (1986). "Stock return variances: the arrival of information and the reaction of traders". *Journal of Financial Economics*. 17: 5-26.
- FOSTER, D. & NELSON, D.B. (1996). "Continuous record asymptotics for rolling sample variance estimators". *Econometrica*. 64: 139-174.
- GEWEKE, (1986). "Modelling the persistence of conditional variance: a comment". *Econometric Reviews*. 5: 71-73.
- GHYSELS, E., HARVEY, A.C. & RENAULT, E. (1995). "Stochastic Volatility". *Handbook of Statistics*. 14 (forthcoming).
- GLOSTEN, L.R., JAGANNATHAN, R. & RUNKLE, D. (1993). "Relationship between the expected value and the volatility of the nominal excess return on stocks". *Journal of Finance*. 48: 1779-1802.
- GRANGER, C.W.J. & ANDERSEN, A.P. (1978). "Non-linear Time Series Modelling". In: Applied Time Series Analysis, D.F. Findley (ed) New York: Academic Press.
- HARVEY, A.C. (1981). *Time Series Models*. Oxford : Phillip Allan.

- \_\_\_\_\_ (1989). *Forecasting, Estructural Time Series Models and the Kalman Filter*. Cambridge : Cambridge University Press.
- \_\_\_\_\_ (1993). "Long Memory in Stochastic Volatility". Mimeo, LSE Statistical Dept.
- \_\_\_\_\_ (1993a). *The Econometric Analysis of Time Series*. Cambridge: MIT Press.
- HARVEY, A.C. & KOOPMAN, S.J. (1992). "Diagnostic checking of unobserved-components time series models". *Journal of Business & Economic Statistics*. **10**: 377-389.
- HARVEY, A.C., RUIZ, E. & SENTANA, (1992). "Unobserved component time series models with ARCH disturbances". *Journal of Econometrics*. **52**: 129-158.
- HARVEY, A.C., RUIZ E. & SHEPHARD,N. (1994). "Multivariate stochastic variance models". *Review of Economic Studies*. **61**: 247-264.
- HARVEY, A.C. & SHEPHARD, N. (1993a). "The econometrics of stochastic volatility". Discussion Paper 166, LSE Statistical Dept Financial Market Group.
- \_\_\_\_\_ (1993b). "Estimation and testing of stochastic variance models". Discussion Paper EM/93/268, LSE Statistical Dept. STICERD.
- \_\_\_\_\_ (1993c). "Estimation of an asymmetric stochastic volatility model for asset returns". Working paper, LSE Statistical Dept.
- HENDRY, D.F.(1995). *Dynamic Econometrics*. New York: Oxford University Press.
- HERENCIA, M.Z., HOTTA, L.K., PEREIRA, P.L.V., FERREIRA, J. & MECCHI, M. (1995). "Volatilidade nos retornos de TELEBRÁS: uma comparação entre abordagem ARCH e modelo de volatilidade estocástica". *Anais do XVII Encontro Brasileiro de Econometria*, Vol 1: 587-621.
- HESTON, S.L. (1993). "A closed-form solution for options with stochastic volatility, with applications to bond and currency options". *Review of Financial Studies* **6**: 327-343.
- HIGGINS, M.L. & BERA, A.K.(1992). "A class of nonlinear ARCH models". *International Economic Review*. **33**: 137-158.
- HSIEH, D.A. (1989). "Testing for nonlinear dependence in daily foreign exchange rates". *Journal of Business* **62**: 339-368.
- \_\_\_\_\_ (1991). "Modelling heteroskedasticity in foreign exchange rates". *Journal of Business & Economic Statistics*. **7**: 307-317.
- HOAGLIN, D.C., MOSTELLER, F. & TUKEY, J.W. (1983). *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley.

- HONG, C.H. (1987). "The integrated generalized autoregressive conditional heteroskedastic model: the process, estimation and some Monte Carlo experiments". Unpublished manuscript, UCSD Economics Dept.
- HULL, J. & WHITE, A. (1987). "The pricing of options on assets with stochastic volatilities". *Journal of Finance*. **42**: 281-300.
- JACQUIER, E., POLSON, N.G. & ROSSI, P.E. (1994). "Bayesian analysis of stochastic volatility models" (with discussion). *Journal of Business & Economic Statistics*. **12**: 371-417.
- KALMAN, R.E. (1960). "A new approach to linear filtering and prediction problems". *Journal of Basic Engineering, Transactions ASMA, Series D*. **82**: 35-45.
- KEARNS (1990). "Non-linearities in the term structure". Mimeo, University of Rochester.
- KIM, C.M. (1989). "Nonlinear dependence of exchange rates changes". Unpublished Ph.D dissertation Graduate School of Business, University of Chicago.
- KING, M., SENTANA, E. & WADHWANI, S. (1994). "Volatility and links between national stock markets". *Econometrica*. **62**: 901-934.
- KOOPMAN, S.J. (1993). "Disturbance smoother for state space models". *Biometrika*. **80**: 117-126.
- KOOPMAN, S.J. & HARVEY, A.C. (1995). *Stamp 5.00*. New York: Chapman and Hall.
- LAMOREAUX, C.G. & LASTRAPES, W.D. (1990). "Persistence in variance, structural change and the GARCH model". *Journal of Business & Economic Statistics*. **8**: 225-233.
- LEE, S.W. & HANSEN, B.E. (1994). "Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator". *Econometric Theory*. **10**: 29-52.
- LIU, S.M. & BRORSEN, B.W. (1995). "Maximum likelihood estimation of a GARCH-Stable model". *Journal of Applied Econometrics*. **10**: 273-285.
- LO, A. & MACKINLAY, C. (1988). "Stock market prices do not follow random walks: evidence from a simple specification test". *Review of Financial Studies*. **1**: 41-66.
- LUMSDAINE, R.L. (1995). "Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: a Monte Carlo investigation". *Journal of Business & Economic Statistics*. **13**: 1-10.
- \_\_\_\_\_ (1996). "Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models". *Econometrica*. **64**: 575-596.

- MANDELBROT, B.B. (1963). "The variation of certain speculative prices". *Journal of Business*. **36**: 394-416.
- MELINO, A. & TURNBULL, S.M (1990). "Pricing foreign currency options with stochastic volatility". *Journal of Econometrics*. **45**: 239-265.
- MILLOJ, A.(1987). "A multiplicative parametrization of ARCH models". Mimeo. University of Copenhagen.
- MILLS (1993). *The Econometric Modelling of Financial Time Series*. Cambridge: University Press.
- McLEOD, A.I. & LI, W.K. (1983). "Diagnostic checking ARMA time series models using squared residual autocorrelations". *Journal of Time Series Analysis*. **4**: 269-273.
- NELSON, D.B. (1988). "The time series behavior of stock market volatility and returns". Unpublished Doctoral Dissertation, MIT Economics Dept.
- \_\_\_\_\_ (1990). "Stationary and persistence in the GARCH (1,1) model". *Econometric Theory*. **6**: 318-334.
- \_\_\_\_\_ (1991). "Conditional heteroskedasticity in asset returns: a new approach". *Econometrica*. **59**: 347-370.
- NELSON, D.B. & CAO, C.Q. (1992). "Inequality constraints in the univariate GARCH model". *Journal of Business & Economic Statistics*. **10**: 229-235.
- NELSON, D.B. & FOSTER, D.P. (1994). "Asymptotic filtering theory for univariate ARCH models". *Econometrica*. **62**: 1-41.
- NEWBY, W.F. (1985). "Maximum likelihood specification testing and conditional moment tests". *Econometrica*. **53**: 1047-1070.
- NOH, J.; ENGLE, R.F & KANE, A. (1994). "Forecasting volatility and option prices of the S&P 500 index". *Journal of Derivatives*. : 17-30.
- PAGAN, A. (1994). "The econometrics of financial markets". Mimeo, University of Rochester e Australian National University.
- PAGAN, A. & SCHWERT, G.W.(1990). "Alternative models for conditional stock volatility". *Journal of Econometrics*. **45**: 267-290.
- PANTULA, S. (1986). "Modelling the persistence of conditional variance: a comment". *Econometric Reviews*. **5**: 71-73.

PATTEL, J.M. & WOLFSON, M.A. (1979). "Anticipated information releases reflected in call option prices". *Journal of Accounting and Economics*. 1: 117-140.

\_\_\_\_\_ (1981). "The ex-ante and ex-post effects of quarterly earnings announcement reflected in option and stock prices". *Journal of Accounting Research*. 19: 434-458.

PHILLIPS, P.C.B. & LORETAN, M. (1990). "Testing covariance stationarity under moment condition failure with an application to common stock returns". Discussion Paper 947, Cowles Foundation.

POTERBA, J. & SUMMERS, L. (1986). "The persistence of volatility and stock market fluctuations". *American Economic Review*. 76: 1124-1141.

\_\_\_\_\_ (1988). "Mean reversion in stock prices: evidence and implications". *Journal of Financial Economics*. 22: 27-59.

RUIZ, E. (1994). "Quase maximum likelihood estimation of stochastic variance models". *Journal of Econometrics*. 63: 284-306.

SAS 6.0 (1990). SAS Institute Inc.

SCOTT, L. (1987). "Options pricing when the variance changes randomly: theory estimation and an application". *Journal of Financial and Quantitative Analysis*. 22: 419-438.

SCHWARZ, G. (1978). "Estimating the dimension of a model". *The Annals of Statistics*. 6: 461-464.

SCHEINKMAN, J.A. & LE BARON (1989). "Non-linear dynamics and stock returns". *Journal of Business*. 62: 311-337.

SCHOLES, M. & WILLIAMS, J. (1977). "Estimating beta from nonsynchronous data". *Journal of Financial Economics*. 5: 309-327.

SCHWERT, G.W. (1989a). "Business cycles, financial crises and stock volatility". Carnegie-Rochester Conference Series on Public Policy. 39: 83-126.

\_\_\_\_\_ (1989b). "Why does stock market volatility change over time". *Journal of Finance*. 44: 1115-1153.

SHAPIRO, S.S. & WILK, M.B. (1965). "An analysis of variance test for normality (complete samples)". *Biometrika*. 52: 591-611.

SHEPARD, N. (1994a). "Local scale models: state space alternative to integrated GARCH processes". *Journal of Econometrics*. 60: 181-202.

- \_\_\_\_\_ (1994b). "Statistical aspects of ARCH and stochastic volatility". Mimeo, Nuffiel College.
- SILVERMAN, B. W (1986). *Density Estimation for Statistical and Data Analysis*. London: Chapman-Hall.
- STEIN, E.M. & STEIN, J. (1991). "Stock price distributions with stochastic volatility: an analytic approach". *Review of Financial Studies*. **4**: 727-752.
- TAUCHEN, G. (1985). "Diagnostic testing and evaluation of maximum likelihood models". *Journal of Econometrics*. **30**: 415-443.
- TAUCHEN, G. & PITTS, M. (1983). "The price variability-volume relationship on speculative markets". *Econometrica*. **51**: 485:505.
- TAYLOR, S.J. (1980). "Conjetured models for trends in financial prices, tests and forecasts". *Journal of the Royal Statistical Society A*. **143**: 338-362.
- TAYLOR, S.J. (1986). *Modelling Financial Time Series*. Chitester: John Wiley.
- TAYLOR, S. J. (1994). "Modelling stochastic volatility: a review and comparative study". *Mathematical Finance*. **4**: 183-204.
- TONG,H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. New York: Oxford University Press.
- TSAY,R.S. (1986). "Nonlinearity tests for time series". *Biometrika*. **73**: 461-466.
- WEISS,A.A. (1984) . "ARMA models with ARCH errors". *Journal of Time Series Analysis*. **5**: 129-143.
- \_\_\_\_\_ (1986). "Asymptotic theory for ARCH Models: estimation and testing". *Econometric Theory*. **2**: 107-131.
- WEST, M. & HARRISON, J. (1989). *Bayesian Forecasting and Dynamic Models*. New York: Springer Verlag.
- WHITE,H.(1982) . "Maximum likelihood estimation of misspecified models". *Econometrica*. **50**: 1-25.
- WIGGINS,J.B. (1987). "Options values under stochastic volatility. theory and empirical estimates". *Journal of Financial Economics*. **19**: 351-372.
- WOOLDRIDGE,J.M. (1994) "Estimation and Inference for Dependent Processes". In: *Handbook of Econometrics*, Vol IV. Engle,R.F & Mc Fadden,D (eds). Amsterdam, North-Holland.

ZIEGELMANN, F.A. (1996). Modelos de Variabilidade Estocástica e Deformação Temporal. Dissertação de Mestrado, Universidade de São Paulo Instituto de Matemática e Estatística.