

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE CIÊNCIAS ECONÔMICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**DOUGLAS GOMES DOS SANTOS**

**ESTIMAÇÃO DE VOLATILIDADE EM SÉRIES FINANCEIRAS:  
MODELOS ADITIVOS SEMI-PARAMÉTRICOS E GARCH**

**Porto Alegre**

**2008**

**DOUGLAS GOMES DOS SANTOS**

**ESTIMAÇÃO DE VOLATILIDADE EM SÉRIES FINANCEIRAS:  
MODELOS ADITIVOS SEMI-PARAMÉTRICOS E GARCH**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. Flávio A. Ziegelmann

**Porto Alegre  
2008**

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
Responsável: Biblioteca Gládis W. do Amaral, Faculdade de Ciências Econômicas da  
UFRGS

S237e

Santos, Douglas Gomes dos

Estimação de volatilidade em séries financeiras : Modelos aditivos semi-paramétricos e Garch / Douglas Gomes dos Santos. – Porto Alegre, 2008.

147 f. : il.

Orientador: Flávio A. Ziegelmann.

Ênfase em Economia Aplicada.

Dissertação (Mestrado em Economia) - Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, 2008.

1. Volatilidade : Índice Bovespa : Previsão. 2. Volatilidade : Índice Standard and Poor's 500 : Previsão. 3. Mercado financeiro : Volatilidade. I. Ziegelmann, Flávio A. II. Universidade Federal do Rio Grande do Sul. Faculdade de Ciências Econômicas. Programa de Pós-Graduação em Economia. III. Título.

CDU 330.115  
336.761

Aos meus pais

## **AGRADECIMENTOS**

Gostaria de agradecer às pessoas e às instituições que, direta ou indiretamente, contribuíram para este trabalho:

À Universidade Federal do Rio Grande do Sul, por ter continuado a me proporcionar ensino gratuito e de ótima qualidade após a graduação, tendo consciência de que essa circunstância é restrita a apenas uma pequena parcela da sociedade;

À CAPES, a quem sou grato por ter me proporcionado uma bolsa de estudos;

Ao programa de Pós-Graduação em Economia;

Ao professor Flávio Augusto Ziegelmann, por todo seu apoio e primordial orientação neste trabalho;

Aos demais professores do PPGE, em especial a Jorge Paulo de Araújo e Marcelo Savino Portugal;

À minha querida e paciente irmã e à minha namorada por todo seu amor, confiança e presença constante;

E, finalmente, aos meus excelentes amigos.

## RESUMO

A estimação e previsão da volatilidade de ativos são de suma importância para os mercados financeiros. Temas como risco e incerteza na teoria econômica moderna incentivaram a procura por métodos capazes de modelar uma variância condicional que evolui ao longo do tempo. O objetivo principal desta dissertação é comparar alguns métodos de regressão global e local quanto à extração da volatilidade dos índices Ibovespa e Standard and Poor's 500. Para isto, são realizadas estimações e previsões com os modelos GARCH paramétricos e com os modelos aditivos semi-paramétricos. Os primeiros, tradicionalmente utilizados na estimação de segundos momentos condicionais, têm sua capacidade sugerida em diversos artigos. Os segundos provêm alta flexibilidade e descrições visualmente informativas das relações entre as variáveis, tais como assimetrias e não linearidades. Sendo assim, testar o desempenho dos últimos frente às estruturas paramétricas consagradas apresenta-se como uma investigação apropriada. A realização das comparações ocorre em períodos selecionados de alta volatilidade no mercado financeiro internacional (crises), sendo a performance dos modelos medida dentro e fora da amostra. Os resultados encontrados sugerem a capacidade dos modelos semi-paramétricos em estimar e prever a volatilidade dos retornos dos índices nos momentos analisados.

Palavras chave: Volatilidade; modelos aditivos; regressão polinomial local; modelos GARCH; crise.

## ABSTRACT

Volatility estimation and forecasting are very important matters for the financial markets. Themes like risk and uncertainty in modern economic theory have encouraged the search for methods that allow for the modeling of time varying variances. The main objective of this dissertation is to compare global and local regressions in terms of their capacity to extract the volatility of Ibovespa and Standard and Poor 500 indexes. To achieve this aim, parametric GARCH and semiparametric additive models estimation and forecasting are performed. The first ones, traditionally applied in the estimation of conditional second moments, have their capacity suggested in many papers. The second ones provide high flexibility and visually informative descriptions of the relationships between the variables, like asymmetries and nonlinearities. Therefore, testing the last ones' performance against the acknowledged parametric structures is an appropriate investigation. Comparisons are made in selected periods of high volatility in the international financial market (crisis), measuring the models' performance inside and outside sample. The results that were found suggest the capacity of semiparametric models to estimate and forecast the Indexes returns' volatility at the analyzed moments.

Keywords: Volatility; additive models; local polynomial regression; GARCH models; crisis.

## LISTA DE ILUSTRAÇÕES

Gráfico 1 - Evolução do Ibovespa e seus retornos.....	72
Gráfico 2 – Evolução do S&P500 e seus retornos.....	73
Gráfico 3 – Fatos estilizados dos retornos do Ibovespa.....	74
Gráfico 4 – Fatos estilizados dos retornos do S&P500.....	75
Gráfico 5 – Funções de autocorrelação.....	79
Gráfico 6 – Funções univariadas do Modelo aditivo (3).....	82
Gráfico 7 – Funções univariadas do Modelo aditivo (4).....	84
Gráfico 8 – Funções univariadas do Modelo aditivo (5).....	85
Gráfico 9 – Funções de autocorrelação dos quadrados dos resíduos padronizados.....	86
Gráfico 10 – Funções univariadas do Modelo aditivo (4) – $span=0,7$ .....	88
Gráfico 11 – Medidas de ajuste do Modelo aditivo (4) – $span=0,7$ .....	89
Gráfico 12 – Resíduos ao quadrado e volatilidade estimada do Ibovespa.....	90
Gráfico 13 – Funções univariadas do Modelo aditivo (3).....	93
Gráfico 14 – Funções univariadas do Modelo aditivo (4).....	94
Gráfico 15 – Funções univariadas do Modelo aditivo (5).....	95
Gráfico 16 – Funções de autocorrelação dos quadrados dos retornos padronizados.....	96
Gráfico 17 – Funções univariadas do Modelo aditivo (4) – $span=0,8$ .....	98
Gráfico 18 – Medidas de ajuste do Modelo aditivo (4) – $span=0,8$ .....	99
Gráfico 19 – Retornos ao quadrado e volatilidade estimada do S&P500.....	100
Gráfico 20 – Medidas de ajuste do Modelo $ARCH(4)$ .....	103



Gráfico 21 – Medidas de ajuste do Modelo $TARCH(3,2)$ .....	105
Gráfico 22 – Resíduos ao quadrado e volatilidade estimada do Ibovespa.....	106
Gráfico 23 – Medidas de ajuste do Modelo $ARCH(4)$ .....	108
Gráfico 24 – Medidas de ajuste do Modelo $TARCH(1,2)$ .....	110
Gráfico 25 – Retornos ao quadrado e volatilidade estimada do S&P500.....	111
Gráfico 26 – Intervalos de confiança um passo à frente em 2001 – Ibovespa.....	114
Gráfico 27 – Intervalos de confiança um passo à frente em 2002 – Ibovespa.....	139
Gráfico 28 – Intervalos de confiança um passo à frente em 2007 – Ibovespa.....	140
Gráfico 29 – Intervalos de confiança um passo à frente em 2001 – Ibovespa.....	141
Gráfico 30 – Intervalos de confiança um passo à frente em 2002 – Ibovespa.....	116
Gráfico 31 – Intervalos de confiança um passo à frente em 2007 – Ibovespa.....	142
Gráfico 32 – Intervalos de confiança um passo à frente em 2001 - S&P500.....	118
Gráfico 33 – Intervalos de confiança um passo à frente em 2002 - S&P500.....	143
Gráfico 34 – Intervalos de confiança um passo à frente em 2007 - S&P500.....	144
Gráfico 35 – Intervalos de confiança um passo à frente em 2001 - S&P500.....	145
Gráfico 36 – Intervalos de confiança um passo à frente em 2002 - S&P500.....	120
Gráfico 37 – Intervalos de confiança um passo à frente em 2007 - S&P500.....	146

## LISTA DE TABELAS

1 – Estatísticas descritivas dos retornos diários do Ibovespa.....	134
2 – Estatísticas descritivas dos retornos diários do S&P500.....	135
3 – Observações extremas excluídas da série completa dos retornos do Ibovespa.....	136
4 – Estimativas do modelo para a média condicional do Ibovespa.....	137
5 – Seleção da defasagem nos modelos aditivos para o Ibovespa.....	81
6 – Seleção do <i>span</i> no Modelo aditivo (4).....	87
7 – Observações extremas excluídas da série completa dos retornos do S&P500.....	138
8 – Seleção da defasagem nos modelos aditivos para o S&P500.....	92
9 – Seleção do <i>span</i> no Modelo aditivo (4).....	97
10 – Seleção dos modelos GARCH para o Ibovespa.....	101
11 – Estimativas dos modelos GARCH para o Ibovespa.....	102
12 – Seleção dos modelos GARCH para o S&P500.....	107
13 – Estimativas dos modelos GARCH para o S&P500.....	107
14 – Desempenho dos modelos (dentro da amostra) para o Ibovespa.....	113
15 – Desempenho dos modelos (fora da amostra) para o Ibovespa.....	115
16 – Desempenho dos modelos (dentro da amostra) para o S&P500.....	117
17 – Desempenho dos modelos (fora da amostra) para o S&P500.....	119

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>11</b>
<b>2</b>	<b>MODELOS GARCH PARAMÉTRICOS.....</b>	<b>15</b>
2.1	DEFINIÇÕES INICIAIS.....	16
2.2	REGULARIDADES EMPÍRICAS EM SÉRIES DE TEMPO FINANCEIRAS.....	18
2.2.1	<b>Caudas Pesadas.....</b>	<b>18</b>
2.2.2	<b>Agrupamentos de Volatilidade.....</b>	<b>19</b>
2.2.3	<b>Assimetria na Volatilidade.....</b>	<b>19</b>
2.2.4	<b>Períodos sem Negociação.....</b>	<b>20</b>
2.2.5	<b>Outros Fatos Estilizados.....</b>	<b>20</b>
2.3	MODELOS GARCH UNIVARIADOS.....	21
2.3.1	<b>O Modelo ARCH(q) Linear.....</b>	<b>21</b>
2.3.2	<b>O Modelo GARCH(p,q) Linear.....</b>	<b>23</b>
2.3.3	<b>O Modelo EGARCH(p,q).....</b>	<b>25</b>
2.3.4	<b>O Modelo TARARCH(p,q).....</b>	<b>26</b>
2.3.5	<b>Parametrizações Alternativas.....</b>	<b>26</b>
2.3.6	<b>Estimação.....</b>	<b>28</b>
2.3.7	<b>Tópicos Adicionais em Inferência.....</b>	<b>30</b>
2.3.8	<b>Previsão.....</b>	<b>32</b>
<b>3</b>	<b>MÉTODOS NÃO PARAMÉTRICOS.....</b>	<b>35</b>
3.1	REGRESSÃO UNIVARIADA.....	36
3.1.1	<b>Suavização Kernel Polinomial Local.....</b>	<b>36</b>
3.2	ESPECIFICAÇÕES EM REGRESSÃO POLINOMIAL LOCAL.....	42
3.2.1	<b>A Função Peso (Kernel).....</b>	<b>42</b>
3.2.2	<b>A Escolha da Janela.....</b>	<b>44</b>
3.2.2.1	Seletores Anexos ( <i>Plug-in</i> ) e a Janela Ótima.....	44
3.2.2.2	Validação Cruzada.....	45
3.3	REGRESSÃO MÚLTIPLA.....	46
3.4	ESTIMAÇÃO DE VOLATILIDADE.....	48
3.4.1	<b>Estimação de Volatilidade em Regressão Múltipla.....</b>	<b>53</b>
3.4.1.1	O Modelo Saturado e o Estimador Residual Multivariado.....	54
3.5	MODELOS ADITIVOS.....	54
3.5.1	<b>Inferência.....</b>	<b>58</b>
<b>4</b>	<b>APLICAÇÃO NUMÉRICA: MODELOS ADITIVOS SEMI- PARAMÉTRICOS E GARCH.....</b>	<b>63</b>
4.1	PERÍODOS SELECIONADOS.....	63
4.1.1	<b>A Crise de 2001.....</b>	<b>64</b>
4.1.2	<b>A Crise de 2002.....</b>	<b>66</b>

4.1.3	<b>A Crise de 2007</b> .....	<b>68</b>
4.2	CARACTERÍSTICAS DOS DADOS.....	71
4.3	ESTIMAÇÃO.....	77
4.3.1	<b>Modelos Aditivos</b> .....	<b>77</b>
4.3.1.1	Modelos Aditivos para o Ibovespa.....	77
4.3.1.2	Modelos Aditivos para o S&P500.....	91
4.3.2	<b>Modelos GARCH</b> .....	<b>101</b>
4.3.2.1	Modelos GARCH para o Ibovespa.....	101
4.3.2.2	Modelos GARCH para o S&P500.....	106
4.4	COMPARAÇÃO ENTRE OS MODELOS ADITIVOS E GARCH.....	111
4.4.1	<b>Comparação na Série do Ibovespa</b> .....	<b>112</b>
4.4.2	<b>Comparação na Série do S&amp;P500</b> .....	<b>116</b>
5	<b>CONSIDERAÇÕES FINAIS</b> .....	<b>122</b>
	<b>REFERÊNCIAS</b> .....	<b>126</b>
	<b>APÊNDICE A – Características dos dados</b> .....	<b>134</b>
	<b>APÊNDICE B – Intervenção em observações extremas do Ibovespa</b> .....	<b>136</b>
	<b>APÊNDICE C – Equação da esperança condicional do Ibovespa</b> ....	<b>137</b>
	<b>APÊNDICE D – Intervenção em observações extremas do S&amp;P500</b> .....	<b>138</b>
	<b>APÊNDICE E – Intervalos de confiança com base nas estimativas um passo à frente da volatilidade do Ibovespa</b> .....	<b>139</b>
	<b>APÊNDICE F – Intervalos de confiança com base nas previsões um passo à frente da volatilidade do Ibovespa</b> .....	<b>141</b>
	<b>APÊNDICE G – Intervalos de confiança com base nas estimativas um passo à frente da volatilidade do S&amp;P500</b> .....	<b>143</b>
	<b>APÊNDICE H – Intervalos de confiança com base nas previsões um passo à frente da volatilidade do S&amp;P500</b> .....	<b>145</b>

## 1 INTRODUÇÃO

Em razão da relevância de temas como risco e incerteza na teoria econômica moderna, foram propostos na literatura métodos capazes de modelar uma variância condicional que evolui ao longo do tempo. Neste contexto, o modelo ARCH (*Autoregressive Conditional Heteroskedastic*) difundido a partir do artigo seminal de Engle (1982), e sua versão generalizada (GARCH) proposta por Bollerslev (1986), destacam-se por sua ampla utilização.

A principal característica apresentada pelos modelos GARCH se encontra na diferenciação entre os momentos de segunda ordem, não condicionais e condicionais. A volatilidade condicional é definida pela variância condicionada ao conjunto de informações disponíveis. Desta maneira, assume-se a variância não condicional como invariante (constante), ao passo que a volatilidade condicional (estocástica) é modelada como uma variável aleatória. Salienta-se que uma forma de obter medidas para o risco e a incerteza, inerentes aos mercados financeiros, costuma se dar a partir da volatilidade dos retornos. Estes são definidos neste trabalho como  $y_t = \ln X_t - \ln X_{t-1}$  (onde  $X_t$  é o preço do ativo no tempo  $t$ ).

Inúmeras pesquisas foram realizadas com os modelos GARCH nos últimos vinte e seis anos, levando a diversas inovações nos mesmos. Estas inovações visam (em sua maioria) a descrever de forma mais apropriada as regularidades empíricas em séries de tempo financeiras.

Em relação aos modelos não paramétricos, a principal motivação para se trabalhar com os mesmos ao invés dos paramétricos, apresenta-se quando o pesquisador não possui conhecimento suficiente sobre o processo gerador explicativo, permitindo desta forma que os dados “falem por si mesmos”. Assim, parte-se para a análise dos dados supondo-se modelos menos rígidos. A única condição que deve ser especificada para os estimadores não paramétricos está relacionada à definição de um certo grau de suavidade para as funções estimadas. Portanto, pode-se dizer que a principal vantagem dos modelos não paramétricos está relacionada a sua flexibilidade e à conseqüente diminuição do risco de errar na

especificação do modelo. Contudo, os custos desta maior flexibilidade se apresentam em taxas de convergência mais lentas, perda do poder de extrapolação e o problema da dimensionalidade (ZIEGELMANN, 2002a).

Sob o enfoque multivariado, onde há a presença de mais de uma variável defasada na regressão, o modelo saturado (ou totalmente não paramétrico) destaca-se como a opção mais flexível na redução do viés de especificação. Todavia, ressalta-se que problemas relacionados a espaços de alta dimensão são freqüentes e, como citado anteriormente, recebem o nome de problema da dimensionalidade (*curse of dimensionality*). Este leva à perda da noção de vizinhança local, essencial em modelagem não paramétrica, o que abre espaço para o uso de modelos semi-paramétricos. Estes, por sua vez, impõem algum tipo de restrição nas funções estimadas. Por essa razão, são mais rígidos que o modelo não paramétrico saturado, mas substancialmente mais flexíveis que os paramétricos.

O modelo aditivo semi-paramétrico, amplamente discutido em Hastie e Tibshirani (1990), é apresentado como alternativa ao problema da dimensionalidade. Neste, a idéia é modelar funções univariadas de forma aditiva, abandonando-se o pressuposto da linearidade definido no modelo de regressão linear paramétrico. A motivação principal de um modelo aditivo é a de permitir que componentes do modelo de regressão assumam formas não paramétricas. Estes superam o problema em espaços de alta dimensão devido ao sistema de ajuste ser construído a partir de suavizadores univariados. Nota-se ainda que as suposições sobre o erro do modelo são substancialmente relaxadas.

Os modelos aditivos podem ser definidos de forma a representar estruturas similares a um modelo ARCH. Pode-se especificar a volatilidade condicional como uma soma de funções univariadas, uma para cada coordenada direcional. As mesmas são estimadas por algum suavizador não paramétrico, sendo que neste trabalho optou-se pelo loess, desenvolvido por Cleveland (1979). Cabe destacar que este procedimento de estimação é distinto da realização de  $n$  regressões univariadas na medida em que as demais dimensões são levadas em conta no momento do ajuste em uma direção específica. Além do suavizador univariado, utiliza-se o estimador residual da variância que, juntos ao algoritmo iterativo *backfitting*, estimam cada função do modelo através da suavização dos resíduos parciais.

O objetivo principal desta dissertação é comparar técnicas distintas de regressão global e local quanto à extração da volatilidade dos índices selecionados. Para tanto, são realizadas estimações e previsões com os modelos GARCH paramétricos e com os modelos aditivos semi-paramétricos. Os primeiros são tradicionalmente utilizados na estimação de segundos momentos condicionais, tendo sua capacidade sugerida em diversos estudos (ISSLER, 1999; MORAIS; PORTUGAL, 1999; PEREIRA *et al.*, 1999). Já os modelos aditivos provêm alta flexibilidade e descrições visualmente informativas das relações entre as variáveis envolvidas, tais como assimetrias, não linearidades etc. Dentre as referências que utilizam os modelos aditivos (ou sua versão generalizada) na estimação de volatilidade, pode-se citar Kim e Linton (2004), Levine e Li (2007), e Ziegelmann (2002a). Neste sentido, testar o desempenho dos últimos frente às estruturas paramétricas consagradas apresenta-se como uma investigação apropriada.

Com este propósito, são utilizadas séries de retornos diários, representativas dos índices Ibovespa e Standard and Poor's 500. A realização das comparações ocorre em períodos selecionados de alta volatilidade no mercado financeiro internacional (crises), onde a performance dos modelos é medida dentro e fora da amostra. As aplicações numéricas são realizadas no *software* livre R para os modelos aditivos, ao passo que, para os modelos GARCH, utiliza-se o *software* EViews 5.1. O restante da dissertação é estruturado da forma que segue.

O capítulo 2 reúne as técnicas paramétricas de extração da volatilidade condicional a partir dos modelos da família GARCH. No mesmo são discutidas as principais regularidades empíricas encontradas em séries de tempo financeiras (fatos estilizados). Adicionalmente, realiza-se a apresentação das parametrizações mais utilizadas na literatura, aliadas ao processo de estimação, seleção, diagnóstico e previsão nos modelos GARCH.

O capítulo 3 engloba o método não paramétrico de estimação da volatilidade via regressão polinomial local. Neste sentido, apresenta-se inicialmente a técnica de regressão para a média condicional, tanto no contexto univariado, como no contexto multivariado. Em seguida, o método é estendido para a estimação da variância condicional (volatilidade). Procura-se descrever ao longo das seções o ferramental básico para a utilização posterior dos modelos semi-paramétricos, tal como o suavizador loess, o estimador residual da variância, entre outros. Além disso, como tema principal do capítulo, são apresentados os modelos aditivos. Para

estes, são definidas a forma de estimação por meio do algoritmo iterativo *backfitting* e algumas maneiras disponíveis de inferência.

O capítulo 4 contempla a aplicação numérica de comparação entre os modelos paramétricos e semi-paramétricos. Como eles são comparados em períodos de crise, apresentam-se inicialmente os momentos selecionados somados a uma breve contextualização dos principais fatos ocorridos nestes períodos. A partir disto, após a descrição dos dados utilizados na dissertação, parte-se para a seleção das estruturas mais adequadas entre os modelos aditivos e GARCH em cada série completa dos retornos. Posteriormente, os modelos selecionados na etapa anterior têm suas estimativas e previsões um passo à frente comparadas nos períodos de alta volatilidade no mercado financeiro.

Por fim, as considerações finais são reunidas no capítulo 5. Neste, são avaliados os resultados encontrados entre as técnicas distintas de regressão e, de forma complementar, são realizadas algumas proposições de trabalhos futuros sobre o tema.



## 2 MODELOS GARCH PARAMÉTRICOS

A partir da década de 1980 houve uma mudança de foco na análise de séries temporais financeiras. Anteriormente, as modelagens macroeconômicas e financeiras se concentravam no estudo dos primeiros momentos condicionais, tratando qualquer dependência temporal nos momentos superiores como um incômodo. Devido à importância de temas como risco e incerteza na teoria econômica moderna, novos métodos capazes de modelar uma variância condicional que evolui ao longo do tempo surgiram (BOLLERSLEV; ENGLE; NELSON, 1994). Neste contexto, foram propostos os modelos ARCH (*Autoregressive Conditional Heteroskedastic*) por Engle (1982) e sua versão generalizada (GARCH) por Bollerslev (1986).

A principal característica apresentada pelos modelos da classe GARCH se encontra na diferenciação entre os momentos de segunda ordem, não condicionais e condicionais. A volatilidade condicional é definida pela variância condicionada ao conjunto de informações disponíveis. Desta maneira, assume-se a variância não condicional como invariante (constante), ao passo que a volatilidade condicional (estocástica) é modelada como uma variável aleatória. Desta forma, a distribuição condicional dos retornos muda a cada ponto no tempo (ZIEGELMANN, 2002a).

Uma característica especial da volatilidade<sup>1</sup> de ativos financeiros é que a mesma não é diretamente observável. Segundo Tsay (2002), considerando-se os retornos de um ativo, definidos como  $y_t = \ln X_t - \ln X_{t-1}$  (onde  $X_t$  é o preço do ativo no tempo  $t$ ), a volatilidade diária não é diretamente observada a partir dos retornos diários, pois neste caso existe apenas uma observação para cada dia de negociação. Teoricamente, se dados intradiários fossem observados em tempo contínuo, poder-se-ia observar a volatilidade diária. Entretanto, este processo é observado na prática em tempo discreto, levando à utilização de dados em tempo discreto para a inferência da volatilidade. Neste contexto, existem diversos estudos

---

<sup>1</sup> Ao longo deste capítulo, poderá fazer-se referência tanto a  $\sigma_t$  como a  $\sigma_t^2$  como volatilidade.

com a utilização de dados em alta freqüência para a estimação e previsão de volatilidade, onde maiores detalhes podem ser encontrados em Andersen e Bollerslev (1997, 1998). Cabe destacar, que mesmo não sendo diretamente observável, a volatilidade apresenta algumas características comuns nas séries dos retornos. Estas características peculiares das séries de tempo financeiras recebem o nome de fatos estilizados.

Neste capítulo, busca-se apresentar técnicas paramétricas de extração da volatilidade condicional a partir dos modelos da família GARCH. Para isto, são discutidas as principais características encontradas em séries de tempo financeiras (fatos estilizados) na seção 2.2. Além disso, de forma complementar, a seção 2.3 é constituída das parametrizações mais utilizadas na literatura, bem como do processo de estimação, seleção, diagnóstico e previsão nos modelos GARCH.

## 2.1 DEFINIÇÕES INICIAIS

Para iniciar, pode-se definir um processo estocástico discreto  $\{\varepsilon_t(\boldsymbol{\theta})\}$ , assumindo-se que  $\varepsilon_t(\boldsymbol{\theta})$  é um escalar no contexto univariado, com esperança e variância condicionais parametrizadas pelo vetor de dimensão finita  $\boldsymbol{\theta} \in \Theta \subseteq R^m$ , tal que  $\boldsymbol{\theta}_0$  representa o valor verdadeiro dos parâmetros. Define-se a esperança condicional, em relação ao conjunto de informação disponível até o tempo  $t-1$ , como  $E_{t-1}(\cdot)$ . Pode-se dizer que, se a esperança condicional do processo  $\{\varepsilon_t(\boldsymbol{\theta}_0)\}$  for igual à zero,

$$E_{t-1}(\varepsilon_t(\boldsymbol{\theta}_0)) = 0 \quad t = 1, 2, \dots, \quad (2.1)$$

e a variância condicional,

$$\sigma_t^2(\boldsymbol{\theta}_0) \equiv \text{Var}_{t-1}(\varepsilon_t(\boldsymbol{\theta}_0)) = E_{t-1}(\varepsilon_t^2(\boldsymbol{\theta}_0)) \quad t = 1, 2, \dots, \quad (2.2)$$

dependem de forma não trivial da sigma-álgebra gerada pelas observações passadas,  $\{\varepsilon_{t-1}(\boldsymbol{\theta}_0), \varepsilon_{t-2}(\boldsymbol{\theta}_0), \dots\}$ , então este processo segue um modelo ARCH (BOLLERSLEV; ENGLE; NELSON, 1994).

Pode-se pensar no processo  $\{\varepsilon_t\}$ , como inovações, choques ou retornos corrigidos pela média. Neste caso, pode-se definir o processo estocástico  $\{y_t(\boldsymbol{\theta}_0)\}$  como representativo dos retornos de um ativo. Desta forma, a esperança condicional de  $\{y_t(\boldsymbol{\theta}_0)\}^2$ ,

$$\mu_t(\boldsymbol{\theta}_0) \equiv E_{t-1}(y_t) \quad t = 1, 2, \dots, \quad (2.3)$$

é definida. Ressalta-se que ambas estatísticas,  $\mu_t(\boldsymbol{\theta}_0)$  e  $\sigma_t^2(\boldsymbol{\theta}_0)$ , são medidas em relação ao conjunto de informação em  $t-1$ . Assim, o processo  $\{\varepsilon_t(\boldsymbol{\theta}_0)\}$  pode ser representado como:

$$\varepsilon_t(\boldsymbol{\theta}_0) \equiv y_t - \mu_t(\boldsymbol{\theta}_0) \quad t = 1, 2, \dots \quad (2.4)$$

A variância condicional, por sua vez, será igual para os processos  $\{\varepsilon_t\}$  e  $\{y_t\}$ .

A partir das expressões (2.1) e (2.2), pode-se representar o processo padronizado,

$$z_t(\boldsymbol{\theta}_0) \equiv \varepsilon_t(\boldsymbol{\theta}_0) \sigma_t^2(\boldsymbol{\theta}_0)^{-1/2} \quad t = 1, 2, \dots, \quad (2.5)$$

tal que o mesmo possui esperança condicional igual a zero e, variância condicional unitária ( $z_t \sim IID(0,1)$ ). No caso da distribuição condicional de  $z_t$  ser assumida constante no tempo e, com o quarto momento finito, a partir da desigualdade de Jensen segue que:

$$E(\varepsilon_t^4) = E(z_t^4) E(\sigma_t^4) \geq E(z_t^4) E(\sigma_t^2)^2 = E(z_t^4) E(\varepsilon_t^2)^2, \quad (2.6)$$

---

<sup>2</sup> Quando o contexto deixar clara a dependência explícita em relação aos parâmetros,  $\boldsymbol{\theta}$  será suprimido.

onde a igualdade se mantém verdadeira apenas se a variância condicional for constante. Assumindo-se uma distribuição normal para  $z_t$ , a distribuição não condicional de  $\varepsilon_t$  será leptocúrtica (BOLLERSLEV; ENGLE; NELSON, 1994).

## 2.2 REGULARIDADES EMPÍRICAS EM SÉRIES DE TEMPO FINANCEIRAS

Embora a volatilidade condicional não seja diretamente observável, algumas características peculiares nas séries de tempo financeiras, ditas fatos estilizados, foram documentadas em diversos trabalhos empíricos. Pode-se dizer que a vasta quantidade de modelos da classe GARCH foi surgindo ao longo do tempo da necessidade de se modelar tais fatos estilizados. Neste sentido, para selecionar o modelo GARCH adequado, deve-se antes de tudo saber qual regularidade empírica busca-se capturar. Esta seção baseia-se primordialmente em Bollerslev, Chou e Kroner (1992), Bollerslev, Engle e Nelson (1994) e Ziegelmann (2002a).

### 2.2.1 Caudas Pesadas

Este fato estilizado está ligado à tendência existente da distribuição não condicional dos retornos de ativos ser leptocúrtica. Esta regularidade empírica, documentada primeiramente por Mandelbrot (1963) e Fama (1965), pode ocasionar um efeito significativo nas propriedades dos estimadores de volatilidade (HALL; YAO, 2003). Pode-se dizer que as caudas pesadas (*heavy tails*) estão relacionadas à existência de autocorrelação entre os quadrados dos retornos, embora os retornos propriamente ditos, não sejam costumeiramente autocorrelacionados. Desta forma, sempre que se observa um retorno de elevado valor absoluto, não importando seu sinal, existe uma grande possibilidade de se obter o próximo retorno com elevado valor absoluto, levando a existência de um excesso de valores extremos.

### 2.2.2 Agrupamentos de Volatilidade

O agrupamento de volatilidade (*volatility clustering*) está ligado a autocorrelação nos quadrados dos retornos de ativos financeiros. Desta forma, segundo Mandelbrot (1963), variações elevadas tendem a ser seguidas por variações elevadas, não importando o sinal, ao passo que pequenas variações tendem a ser seguidas por pequenas variações. Assim, como retornos próximos no tempo  $t$  irão influenciar a volatilidade de  $n$  períodos à frente, explicitando a persistência na volatilidade. Por outro lado, no longo prazo, assumindo-se que a variância não condicional é finita, períodos de alta volatilidade tendem a ser seguidos por períodos de menor volatilidade, e vice versa. Esta característica da volatilidade é conhecida como reversão a média.

### 2.2.3 Assimetria na Volatilidade

Primeiramente descrito por Black (1976), e posteriormente investigado por Christie (1982), o efeito assimetria se refere à tendência de as mudanças nos preços dos ativos serem negativamente correlacionadas com as mudanças na volatilidade dos mesmos. O efeito assimetria ocorre quando uma queda não esperada no preço de um ativo (más notícias) aumenta a volatilidade futura do mesmo mais do que um aumento não esperado no preço (boas notícias) de similar magnitude (ENGLE; NG, 1993). Este efeito é frequentemente observado em retornos de índices de ações (BOLLERSLEV, 2008). Segundo Ziegelmann (2002a), a assimetria na volatilidade costuma ser atribuída em algumas ocasiões ao efeito alavancagem (*leverage effect*), e em outras, a um prêmio de risco. O efeito alavancagem está ligado à razão *debt-to-equity* de uma firma, sendo sua capacidade isolada de influência sobre a assimetria na volatilidade questionada na literatura (BOLLERSLEV, 2008). Como explicitado por Nelson (1991), Grigoletto e Lisi (2008), as razões econômicas por trás do efeito assimetria ainda não são claras.

#### 2.2.4 Períodos sem Negociação

Os mercados financeiros costumam comportar-se de forma distinta após feriados e fins de semana, sendo observado que a volatilidade apresenta-se mais elevada na reabertura dos mesmos. Porém, de acordo com Fama (1965), French e Roll (1986), as informações tendem a se acumular em menor ritmo quando o mercado está fechado, do que quando está em funcionamento.

#### 2.2.5 Outros Fatos Estilizados

**Eventos previsíveis:** a divulgação de informações relevantes costuma levar ao aumento da volatilidade *ex-ante*. Como exemplo, Cornell (1978), Patell e Wolfson (1979, 1981) descrevem o comportamento de elevação da volatilidade dos retornos individuais das firmas, em decorrência da proximidade da divulgação de seus resultados financeiros.

**Co-movimentos na volatilidade:** a volatilidade de diferentes ativos mostra-se intimamente ligada dentro do mesmo mercado, bem como em mercados distintos. Engle, Ito e Lin (1992) descrevem a íntima ligação na volatilidade dos títulos americanos de maturidades distintas. Em outro exemplo, Schwert (1989) encontra indícios de que as volatilidades das ações e títulos americanos movimentam-se conjuntamente, ao passo que Engle e Susmel (1993) descrevem ligações nas volatilidades de mercados financeiros internacionais. Sobre o último tema, pode-se citar Baur e Jung (2006) e Buseti e Manera (2008).

**Variáveis exógenas e volatilidade:** esta regularidade está relacionada ao fato do comportamento dos preços dos ativos financeiros ser influenciado por outras variáveis do mercado. Como exemplo, medidas de incerteza macroeconômica, expressadas através das variâncias condicionais de taxas de juro, taxas de câmbio, produção industrial entre outras, ajudam a explicar a volatilidade

nos mercados financeiros. Estudos sobre o tema podem ser encontrados em Glosten *et al.* (1993), Bollerslev e Melvin (1994) e, Engle e Mezrich (1996).

## 2.3 MODELOS GARCH UNIVARIADOS

Existem diversos modelos paramétricos para se estimar a volatilidade condicional. Entre os mesmos, os modelos da família GARCH são amplamente utilizados. Embora a primeira aplicação dos últimos estivesse relacionada à modelagem de incerteza inflacionária (ENGLE, 1982), o método passou ao longo do tempo a ser substancialmente utilizado para capturar a dependência temporal existente nas séries dos retornos de ativos financeiros. Busca-se apresentar nesta seção, de maneira geral, as principais parametrizações da classe de modelos GARCH, bem como a forma de estimação dos mesmos. Contudo, para uma descrição mais detalhada, ver Bollerslev, Chou e Kroner (1992) e Bollerslev, Engle e Nelson (1994).

### 2.3.1 O Modelo ARCH( $q$ ) Linear

No modelo  $ARCH(q)$  linear de Engle (1982), a variância condicional é definida como uma função linear do quadrado dos  $q$  choques passados,

$$\varepsilon_t = \sigma_t z_t \quad (2.7)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \equiv \omega + \alpha(L) \varepsilon_{t-1}^2, \quad (2.8)$$

tal que  $L$  denota o operador de defasagem,  $L^i y_t \equiv y_{t-i}$ . Algumas restrições devem ser impostas aos parâmetros para se obter uma variância condicional não negativa, tal como  $\omega > 0$  e  $\alpha_1 \geq 0, \dots, \alpha_q \geq 0$ .

De acordo com Fan e Yao (2003), a condição necessária e suficiente para (2.7) e (2.8), definindo um processo estritamente estacionário único  $\{\varepsilon_t\}$ , com  $E\varepsilon_t^2 < \infty$  é,

$$\sum_{i=1}^q \alpha_i < 1. \quad (2.9)$$

Além disso,  $E\varepsilon_t = 0$ ,  $E\varepsilon_t^2 = \omega / (1 - \sum_{i=1}^q \alpha_i)$  e  $\varepsilon_t \equiv 0$  para todo  $t$  se  $\omega = 0$ . Acrescenta-se que, a solução estritamente estacionária de (2.7) e (2.8) terá um quarto momento finito,  $E\varepsilon_t^4 < \infty$ , desde que

$$\max \left\{ 1, (Ez_t^4)^{1/2} \right\} \sum_{i=1}^q \alpha_i < 1. \quad (2.10)$$

A partir da relação,  $\nu_t \equiv \varepsilon_t^2 - \sigma_t^2$ , o modelo definido na expressão (2.8) pode ser reescrito como segue:

$$\varepsilon_t^2 = \omega + \alpha(L)\varepsilon_{t-1}^2 + \nu_t. \quad (2.11)$$

Assim, como  $E_{t-1}(\nu_t) = 0$ , o modelo (2.11) representa um processo  $AR(q)$  para os choques ao quadrado,  $\varepsilon_t^2$ . Os choques  $\varepsilon_t$ , embora não correlacionados, são dependentes no tempo. Além disso, a função de autocorrelação do quadrado dos choques,  $\{\varepsilon_t^2\}$ , é estritamente positiva devido às restrições impostas nos parâmetros. Desta forma, o modelo  $ARCH(q)$  apresenta agrupamento de volatilidade (ZIEGELMANN, 2002a).

Como citado anteriormente, no caso da distribuição condicional dos choques padronizados ( $z_t$  da expressão (2.5)) ser assumida constante no tempo, a distribuição não condicional de  $\varepsilon_t$  apresentará caudas mais pesadas que a distribuição de  $z_t$ . Neste sentido, para um modelo  $ARCH(1)$  com distribuição condicional gaussiana para  $z_t$ ,



$$\frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} = 3 \frac{(1-\alpha_1^2)}{(1-3\alpha_1^2)}, \text{ se } \alpha_1^2 < \frac{1}{3};$$

e (2.12)

$$\frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} = \infty \quad \text{caso contrário.}$$

Assim, observa-se que em ambas as situações o valor obtido é superior a três, sendo a distribuição de  $\varepsilon_t$  leptocúrtica (BOLLERSLEV; ENGLE; NELSON, 1994).

### 2.3.2 O Modelo GARCH(p,q) Linear

Embora o modelo ARCH seja relativamente simples, freqüentemente deve-se utilizar diversos parâmetros para descrever adequadamente o processo da volatilidade dos retornos. Com a intenção de resolver este problema, Bollerlev (1986) propôs o modelo ARCH generalizado, ou  $GARCH(p, q)$ ,

$$\varepsilon_t = \sigma_t z_t \tag{2.13}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \equiv \omega + \alpha(L) \varepsilon_{t-1}^2 + \beta(L) \sigma_{t-1}^2. \tag{2.14}$$

Na expressão (2.14), define-se  $\omega \geq 0$ ,  $\alpha_i, \beta_j \geq 0$ , para  $i=1,2,\dots,q$  e  $j=1,2,\dots,p$ . A variância condicional, definida em (2.14), é modelada como uma função linear do quadrado das  $q$  observações passadas do processo  $\{\varepsilon_t\}$ , mais as  $p$  variâncias condicionais passadas do processo endógeno  $\{\sigma_t^2\}$ . As restrições impostas aos parâmetros são necessárias para a obtenção de uma variância condicional não negativa (ZIEGELMANN, 2002a).

Segundo Fan e Yao (2003), a condição necessária e suficiente para (2.13) e (2.14), definindo um processo estritamente estacionário único  $\{\varepsilon_t\}$ , com  $E\varepsilon_t^2 < \infty$  é,

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (2.15)$$

Além disso,  $E\varepsilon_t = 0$ ,  $E\varepsilon_t^2 = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}$  e a  $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$  para qualquer  $k \neq 0$ . Acrescenta-se que, a solução estritamente estacionária de (2.13) e (2.14) terá um quarto momento finito,  $E\varepsilon_t^4 < \infty$ , desde que

$$\max \left\{ 1, E(z_t^4)^{1/2} \right\} \frac{\sum_{i=1}^q \alpha_i}{1 - \sum_{j=1}^p \beta_j} < 1. \quad (2.16)$$

Reescrevendo o modelo  $GARCH(p, q)$  como em (2.11), segue que

$$\varepsilon_t^2 = \omega + [\alpha(L) + \beta(L)]\varepsilon_{t-1}^2 + v_t - \beta(L)v_{t-1}, \quad (2.17)$$

que representa um modelo  $ARMA[\max(p, q), p]$  para  $\varepsilon_t^2$ . O modelo é dito estacionário em covariância se e somente se todas as raízes de  $\alpha(x) + \beta(x) = 1$  estão fora do círculo unitário (BOLLERSLEV; ENGLE; NELSON, 1994). De forma distinta ao modelo ARCH, mesmo um modelo  $GARCH(1,1)$  permite que a função de autocorrelação do processo  $\{\varepsilon_t^2\}$  seja pequena na primeira defasagem e apresente um lento decaimento posterior. Como o processo  $\{\varepsilon_t^2\}$  é inversível, ou seja,  $\{\varepsilon_t^2\}$  pode ser escrito como um  $AR(\infty)$ , de modo similar um  $GARCH(1,1)$  é equivalente a um  $ARCH(\infty)$ . Este fato explica o porquê de um simples  $GARCH(1,1)$  ser mais parcimonioso que um  $ARCH(q)$  e, ainda assim, permitir que se lide com uma estrutura mais complexa para a autocorrelação de  $\{\varepsilon_t^2\}$  (ZIEGELMANN, 2002a).

Em diversas aplicações com dados financeiros é comum encontrar a estimativa para  $\alpha(1) + \beta(1)$  muito próxima da unidade. Isto motivou Engle e Bollerslev (1986) para a introdução de um modelo  $GARCH(p, q)$  integrado, ou  $IGARCH(p, d, q)$ .

### 2.3.3 O Modelo EGARCH(p,q)

O modelo  $GARCH(p,q)$  consegue de forma adequada captar retornos com caudas pesadas, agrupamento de volatilidade e, pode ser modificado para captar outras regularidades empíricas das séries financeiras. Porém é incapaz, por exemplo, de captar o efeito *leverage*. Isto se deve a variância condicional, definida na expressão (2.14), ser modelada somente através das magnitudes dos choques defasados, não levando em consideração seus sinais.

No modelo GARCH exponencial, ou  $EGARCH(p,q)$  de Nelson (1991),  $\sigma_t^2$  depende tanto da magnitude quanto do sinal dos choques passados. O modelo assume a forma

$$\varepsilon_t = \sigma_t z_t \quad (2.18)$$

$$\ln(\sigma_t^2) = \omega + \left(1 + \sum_{i=1}^q \alpha_i L^i\right) \left(1 - \sum_{j=1}^p \beta_j L^j\right)^{-1} \left\{ \theta z_{t-1} + \gamma \left[ |z_{t-1}| - E|z_{t-1}| \right] \right\}. \quad (2.19)$$

Assim,  $\{\ln(\sigma_t^2)\}$  segue um processo  $ARMA(p,q)$ . Segundo Nelson (1991), o processo  $\{\ln(\sigma_t^2) - \omega\}$  da expressão (2.19) é estritamente estacionário, ergódico e estacionário em covariância se e somente se todas as raízes de  $\left(1 - \sum_{j=1}^p \beta_j L^j\right)$  se encontram fora do círculo unitário.

Pode-se notar que o termo  $\gamma$  da expressão (2.19) representa o efeito da magnitude do retorno na volatilidade, desde que, para  $\gamma > 0$ , quanto maior for o valor absoluto de  $z_{t-1}$ , maior será o valor de  $\gamma$ . Por outro lado,  $\theta$  determina o efeito do sinal dos retornos, se negativos ou positivos, no comportamento da variância condicional. No caso de  $\theta < 0$ , retornos negativos contribuem com uma quantidade positiva em  $\left\{ \theta z_{t-1} + \gamma \left[ |z_{t-1}| + E|z_{t-1}| \right] \right\}$ , ao passo que na presença de retornos positivos, vale o contrário.

### 2.3.4 O Modelo TAR<sub>CH</sub>(p,q)

O modelo  $TARCH(p,q)$  de Zakoian (1994) apresenta-se como uma parametrização alternativa, que é capaz de captar assimetrias na volatilidade (*leverage effect*). O modelo é definido especificando-se  $\gamma = 1$  na expressão que segue:

$$\varepsilon_t = \sigma_t z_t \quad (2.20)$$

$$\sigma_t^\gamma = \omega + \sum_{i=1}^q \left[ \alpha_i^+ I(\varepsilon_{t-i} > 0) |\varepsilon_{t-i}|^\gamma + \alpha_i^- I(\varepsilon_{t-i} \leq 0) |\varepsilon_{t-i}|^\gamma \right] + \sum_{j=1}^p \beta_j \sigma_{t-j}^\gamma, \quad (2.21)$$

tal que  $I(\cdot)$  representa uma função indicadora.

Modificando-se apenas um parâmetro na expressão (2.21), obtém-se o modelo estimado por Glosten, Jagannathan e Runkle (1993). O modelo GJR é definido quando  $\gamma = 2$  e permite uma resposta quadrática da variância condicional em relação aos choques, com coeficientes distintos para choques positivos e negativos. A volatilidade será mínima quando não houver choques (BOLLERSLEV; ENGLE; NELSON, 1994).

### 2.3.5 Parametrizações Alternativas

Existem diversos modelos da família GARCH, sendo que em cada parametrização alternativa, busca-se captar alguma regularidade empírica (fato estilizado). Após apresentar os modelos mais utilizados para a estimação de volatilidade nas seções anteriores, algumas estruturas alternativas seguem nesta seção. Para uma visão geral atualizada dos modelos existentes da família GARCH, ver Bollerslev (2008).

Para captar a aglomeração na variância, o modelo GARCH assume que a volatilidade condicional,  $\sigma_t^2$ , se iguala a defasagens distribuídas dos choques (resíduos) ao quadrado. Neste sentido, Taylor (1986) e Schwert (1989) modelam o desvio padrão condicional como defasagens distribuídas dos choques absolutos,

$$\sigma_t = \omega + \sum_{i=1}^q \alpha_i |\varepsilon_{t-i}| + \sum_{j=1}^p \beta_j \sigma_{t-j}. \quad (2.22)$$

Pode-se apresentar mais uma modelagem alternativa, capaz de captar o efeito *leverage* em retornos de ativos. A mesma é definida como um caso especial do modelo ARCH quadrático (QARCH) de Sentana (1995). Neste modelo, para um  $\kappa \neq 0$ , a volatilidade dependerá tanto da magnitude como do sinal dos choques passados,

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i (\varepsilon_{t-i} - \kappa)^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2. \quad (2.23)$$

Engle (1990) também desenvolve um modelo próximo, cujo nome é ARCH assimétrico (AARCH). Para um *AARCH*(1,1), o modelo assume a forma:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \delta \varepsilon_{t-1} + \beta \sigma_{t-1}^2, \quad (2.24)$$

onde para  $\delta < 0$ , retornos negativos aumentam a volatilidade de maneira mais expressiva do que retornos positivos.

Existem modelos ARCH que incorporam mudança de regime na análise da volatilidade. Propostos independentemente por Hamilton e Susmel (1994) e Cai (1994), os mesmos são conhecidos como *Regime Switching* ARCH, ou SWARCH. Em linhas gerais, o modelo SWARCH permite que o intercepto  $\omega_{s(t)}$  e (ou) a magnitude dos choques ao quadrado,  $\varepsilon_{t-i}^2 |s(t-i)$ , entrem na equação da volatilidade condicional dependendo de uma variável latente de estado,  $s(t)$ , com a transição entre os diferentes estados governada por uma cadeia de Markov (BOLLERSLEV, 2008).

Outro exemplo, ligado a Engle, Lilien e Robins (1987), chamado ARCH-M (*ARCH-in-Mean*), especifica a média condicional como uma função explícita da

variância condicional dos retornos. Neste modelo, um aumento da volatilidade condicional será associado a um crescimento ou decréscimo da esperança condicional dos retornos, dependendo da relação existente com  $\sigma_t^2$ . O modelo busca lidar com questões teóricas em finanças onde existe um *tradeoff* explícito entre risco e retorno esperado (BOLLERSLEV; CHOU; KRONER, 1992).

Para finalizar esta breve seção, pode-se citar o modelo proposto por Baillie, Bollerslev e Mikkelsen (1996), conhecido como FIGARCH (*Fractionally Integrated GARCH*). O mesmo se baseia em uma especificação do tipo ARFIMA para captar a dinâmica da dependência de longo prazo da volatilidade. Além disso, o FIGARCH pode ser visto como uma extensão do modelo IGARCH, permitindo uma ordem de integração fracionada ( $0 < d < 1$ ).

### 2.3.6 Estimação

A maneira mais difundida para se estimar a volatilidade nos modelos da família GARCH está ligada à maximização de funções de verossimilhança condicionais. Duas funções de verossimilhança são amplamente utilizadas na estimação desses modelos. Sob a hipótese que  $z_t$  segue uma distribuição normal, a função de verossimilhança de um modelo *ARCH*( $q$ ) pode ser definida como

$$\begin{aligned} f(\varepsilon_1, \dots, \varepsilon_T | \boldsymbol{\theta}) &= f(\varepsilon_T | F_{T-1}) f(\varepsilon_{T-1} | F_{T-2}) \dots f(\varepsilon_{q+1} | F_q) f(\varepsilon_1, \dots, \varepsilon_q | \boldsymbol{\theta}) \\ &= \prod_{t=q+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[-\frac{\varepsilon_t^2}{2\sigma_t^2}\right] \times f(\varepsilon_1, \dots, \varepsilon_q | \boldsymbol{\theta}), \end{aligned} \quad (2.25)$$

onde  $\boldsymbol{\theta} = (\omega, \alpha_1, \dots, \alpha_q)'$ ,  $T$  é o tamanho da amostra,  $F_{T-1}$  denota o conjunto de informação disponível até  $T-1$  e  $f(\varepsilon_1, \dots, \varepsilon_q | \boldsymbol{\theta})$  é a função densidade de probabilidade conjunta de  $\varepsilon_1, \dots, \varepsilon_q$ . Devido à forma exata da função  $f(\varepsilon_1, \dots, \varepsilon_q | \boldsymbol{\theta})$  ser de difícil acesso, a mesma é comumente extraída da função de verossimilhança

(2.25), especialmente quando se trabalha com amostras grandes. Isto leva a utilização da função de verossimilhança condicional

$$f(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \varepsilon_1, \dots, \varepsilon_q) = \prod_{t=q+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[-\frac{\varepsilon_t^2}{2\sigma_t^2}\right], \quad (2.26)$$

tal que  $\sigma_t^2$  pode ser avaliada recursivamente.

Costuma-se maximizar o logaritmo da função de verossimilhança condicional, neste caso

$$l(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \varepsilon_1, \dots, \varepsilon_q) = \sum_{t=q+1}^T \left[ -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{1}{2} \frac{\varepsilon_t^2}{\sigma_t^2} \right]. \quad (2.27)$$

Na medida em que o termo  $\ln(2\pi)$  não envolve nenhum parâmetro, a função log-verossimilhança pode ser escrita como

$$l(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \varepsilon_1, \dots, \varepsilon_q) = - \sum_{t=q+1}^T \left[ \frac{1}{2} \ln(\sigma_t^2) + \frac{1}{2} \frac{\varepsilon_t^2}{\sigma_t^2} \right], \quad (2.28)$$

onde  $\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$  pode ser avaliada recursivamente (TSAY, 2002).

Embora os modelos ARCH sejam tipicamente estimados supondo-se uma distribuição condicional normal para  $z_t$ , em diversas aplicações empíricas os choques (resíduos) padronizados,  $\hat{\varepsilon}_t \hat{\sigma}_t^{-1}$ , apresentam caudas mais pesadas em relação à distribuição normal. Pode-se desta forma, relaxar o pressuposto da normalidade condicional de  $z_t$ , como em Bollerslev (1987), assumindo-se que os choques padronizados seguem uma distribuição t-Student padronizada. Para isto, deixe  $x_\nu$  ser uma distribuição t-Student, com  $\nu$  graus de liberdade, tal que a  $\text{Var}(x_\nu) = \nu/(\nu-2)$  para  $\nu > 2$  e  $z_t = x_\nu / \sqrt{\nu/(\nu-2)}$ . Desta maneira, pode-se apresentar a função densidade de probabilidade de  $z_t$ ,

$$f(z_t | \nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2) \sqrt{(\nu-2)\pi}} \left(1 + \frac{z_t^2}{\nu-2}\right)^{-(\nu+1)/2}, \quad \nu > 2, \quad (2.29)$$

tal que  $\Gamma(x)$  representa a função Gama tradicional, ou seja,  $\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} dy$ . Para a definição,  $\varepsilon_t = \sigma_t z_t$ , obtém-se a função de verossimilhança condicional

$$f(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \mathbf{C}_q) = \prod_{t=q+1}^T \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2) \sqrt{(\nu-2)\pi}} \frac{1}{\sigma_t} \left[ 1 + \frac{\varepsilon_t^2}{(\nu-2)\sigma_t^2} \right]^{-(\nu+1)/2}, \quad (2.30)$$

onde  $\nu > 2$  e  $\mathbf{C}_q = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q)$ .

Salienta-se que os graus de liberdade ( $\nu$ ) da distribuição t podem ser definidos previamente, ou estimados conjuntamente aos demais parâmetros. Quando os últimos são predefinidos, costuma-se utilizar valores entre 3 e 6 (TSAY, 2002). Neste caso, a função log-verossimilhança condicional segue a forma

$$l(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \mathbf{C}_q) = - \sum_{t=q+1}^T \left[ \frac{\nu+1}{2} \ln \left( 1 + \frac{\varepsilon_t^2}{(\nu-2)\sigma_t^2} \right) + \frac{1}{2} \ln(\sigma_t^2) \right]. \quad (2.31)$$

Se por outro lado, deseja-se estimar os graus de liberdade de forma conjunta aos parâmetros, a função log-verossimilhança é definida como

$$l(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \nu, \mathbf{C}_q) = (T-q) \left[ \ln(\Gamma((\nu+1)/2)) - \ln(\Gamma(\nu/2)) - \frac{1}{2} \ln((\nu-2)\pi) \right] + l(\varepsilon_{q+1}, \dots, \varepsilon_T | \boldsymbol{\theta}, \mathbf{C}_q), \quad (2.32)$$

sendo o último termo de (2.32) definido na expressão (2.31).

Salienta-se que o método de estimação discutido nesta seção para o modelo  $ARCH(q)$  é válido em parametrizações mais complexas da família GARCH, necessitando apenas de algumas especificações extras.

### 2.3.7 Tópicos Adicionais em Inferência

Antes de se estimar um modelo GARCH em uma série de tempo financeira, deve-se analisar a presença, ou não, do efeito ARCH nos resíduos



(retorno corrigido pela média),  $\hat{\varepsilon}_t$ . Normalmente, a média condicional dos retornos,  $E_{t-1}(y_t)$ , segue uma estrutura simples ARMA.

Na medida em que um modelo ARCH pode ser escrito como um modelo AR em relação aos choques (resíduos) ao quadrado, como na expressão (2.11), um teste do Multiplicador de Lagrange (ML) para efeitos ARCH pode ser definido com base na regressão auxiliar (2.11)<sup>3</sup>. Assim, sob a hipótese nula que não há presença de efeitos ARCH, ou seja,  $\alpha_1 = \alpha_2 = \dots = \alpha_q = 0$ , o teste assume a forma

$$ML = T \cdot R^2 \stackrel{A}{\sim} \chi^2(q), \quad (2.33)$$

onde  $T$  representa o tamanho da amostra, e o  $R^2$  é calculado a partir da regressão (2.11) com os resíduos estimados (ZIVOT; WANG, 2002). Além do último, pode-se utilizar outro teste assintoticamente equivalente de correlação serial para  $\hat{\varepsilon}_t^2$ , como o de Ljung e Box (1978).

Através dos últimos testes, na medida em que há indícios da presença de efeitos ARCH, pode-se junto à análise da estrutura de autocorrelação e autocorrelação parcial dos  $\hat{\varepsilon}_t^2$ , identificar a ordem adequada dos modelos ARCH. Feito isto, parte-se para o processo de estimação como apresentado na seção (2.3.6). A partir dos modelos estimados, o processo de seleção costuma estar ligado as parametrizações que minimizam algum critério de informação, como o de Akaike (AIC) ou o de Schwarz (BIC). Além disso, outros fatores são avaliados, como a significância dos termos e o valor do logaritmo da verossimilhança.

Após ser estimada a variância condicional, com base na estimativa do desvio padrão condicional, pode-se analisar o comportamento dos resíduos padronizados,  $\hat{\varepsilon}_t \hat{\sigma}_t^{-1}$ . Aplicando-se tanto o teste do multiplicador de Lagrange (ARCH) como o de Ljung e Box aos resíduos padronizados, pode-se acessar se o modelo está bem ajustado. Em caso de sucesso na modelagem, tanto da esperança quanto da variância condicionais, não deverá restar autocorrelação nos resíduos padronizados tampouco em seus quadrados.

Outra forma de avaliar a qualidade do ajuste está ligada ao exame da série dos resíduos padronizados,  $\hat{\varepsilon}_t \hat{\sigma}_t^{-1}$ , em relação à distribuição teórica assumida.

<sup>3</sup> Para maiores detalhes, ver Engle (1982).

Como a distribuição de  $z_t = \varepsilon_t \sigma_t^{-1}$  é predefinida, seguindo costumeiramente uma normal padrão ou t-Student padronizada, pode-se analisar a assimetria e a curtose de  $\{\hat{\varepsilon}_t \hat{\sigma}_t^{-1}\}$ , bem como os quantis teóricos contra os quantis da série  $\{\hat{\varepsilon}_t \hat{\sigma}_t^{-1}\}$  através de um gráfico quantil-quantil.

### 2.3.8 Previsão

A partir do modelo  $ARCH(q)$ , definido em (2.7) e (2.8), podem-se obter previsões para a variância condicional de maneira recursiva. Para isto, considere o modelo  $ARCH(q)$ , tal que a origem de previsão é definida como  $h$ . Assim, a previsão um passo à frente de  $\sigma_{h+1}^2$  é dada por

$$\sigma_h^2(1) = \omega + \alpha_1 \varepsilon_h^2 + \dots + \alpha_q \varepsilon_{h+1-q}^2, \quad (2.34)$$

ao passo que a previsão dois passos à frente é

$$\sigma_h^2(2) = \omega + \alpha_1 \sigma_h^2(1) + \alpha_2 \varepsilon_h^2 + \dots + \alpha_q \varepsilon_{h+2-q}^2. \quad (2.35)$$

Generalizando, pode-se expressar a previsão  $\ell$  passos à frente para  $\sigma_{h+\ell}^2$  como segue:

$$\sigma_h^2(\ell) = \omega + \sum_{i=1}^q \alpha_i \sigma_h^2(\ell-i), \quad (2.36)$$

onde  $\sigma_h^2(\ell-i) = \varepsilon_{h+\ell-i}^2$  se  $(\ell-i) \leq 0$ .

De forma complementar ao modelo ARCH, previsões da volatilidade condicional podem ser obtidas para o modelo  $GARCH(p,q)$ , definido em (2.13) e (2.14), em referência ao método utilizado nos modelos ARMA (TSAY, 2002).

Considerando-se o modelo  $GARCH(1,1)$ , define-se a previsão um passo à frente como,

$$\sigma_{h+1}^2 = \omega + \alpha_1 \varepsilon_h^2 + \beta_1 \sigma_h^2, \quad (2.37)$$

tal que  $\varepsilon_h$  e  $\sigma_h^2$  são conhecidos na origem de previsão  $h$ . Desta forma, pode-se apresentar a previsão um passo à frente como

$$\sigma_h^2(1) = \omega + \alpha_1 \varepsilon_h^2 + \beta_1 \sigma_h^2. \quad (2.38)$$

Utilizando-se  $\varepsilon_t^2 = \sigma_t^2 z_t^2$ , pode-se reescrever o modelo  $GARCH(1,1)$  para apresentar a previsão da variância condicional em múltiplos passos à frente,

$$\sigma_{t+1}^2 = \omega + (\alpha_1 + \beta_1) \sigma_t^2 + \alpha_1 \sigma_t^2 (z_t^2 - 1). \quad (2.39)$$

Além disso, pode-se definir a expressão (2.39) quando  $t = h+1$ , da seguinte forma:

$$\sigma_{h+2}^2 = \omega + (\alpha_1 + \beta_1) \sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2 (z_{h+1}^2 - 1). \quad (2.40)$$

Como em (2.40) a  $E(z_{h+1}^2 - 1 | F_h) = 0$ , a previsão da variância condicional dois passos à frente passa a ser escrita como

$$\sigma_h^2(2) = \omega + (\alpha_1 + \beta_1) \sigma_h^2(1). \quad (2.41)$$

Generalizando a análise, a previsão  $\ell$  passos à frente pode ser escrita como,

$$\sigma_h^2(\ell) = \omega + (\alpha_1 + \beta_1) \sigma_h^2(\ell - 1), \quad (2.42)$$

para  $\ell > 1$ . Através de substituições sucessivas em (2.42), pode-se representar a previsão  $\ell$  passos à frente da seguinte forma:

$$\sigma_h^2(\ell) = \frac{\omega [1 - (\alpha_1 + \beta_1)^{\ell-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{\ell-1} \sigma_h^2(1). \quad (2.43)$$

Deve-se notar que quando  $\ell \rightarrow \infty$ , a previsão da volatilidade em (2.43) converge para a variância não condicional  $\omega / (1 - \alpha_1 - \beta_1)$ , desde que o modelo  $GARCH(1,1)$  seja estacionário, ou seja,  $(\alpha_1 + \beta_1 < 1)$ .

A avaliação das previsões da variância condicional apresenta-se como uma tarefa delicada na medida em que a volatilidade não é diretamente observável, dificultando o cálculo de estatísticas convencionais como a do erro quadrático médio (EQM). Desta forma, a abordagem tradicional quando se deseja comparar a capacidade preditiva de um número de modelos distintos está ligada à definição da variância condicional  $E_{t-1}(\varepsilon_t^2) = \sigma_t^2$ . Neste caso, estando o modelo corretamente especificado para  $\sigma_t^2$ , então  $E_{t-1}(\varepsilon_t^2) = E_{t-1}(\sigma_t^2 \cdot z_t^2) = \sigma_t^2$ , o que justifica o uso dos quadrados dos retornos corrigidos pela média  $\{\varepsilon_t^2\}$  sobre o horizonte de previsão, como uma medida aproximada da volatilidade *ex-post*. Entretanto, embora  $\{\varepsilon_t^2\}$  gere estimativas não viesadas da volatilidade latente, as mesmas são altamente voláteis e inconsistentes (ANDERSEN; BOLLERSLEV, 1998). Neste sentido, Andersen e Bollerslev (1998) propõem uma base de comparação alternativa, dada por um estimador assintoticamente livre de erro obtido a partir de dados em alta frequência, conhecido como estimador da volatilidade realizada.

Contudo, salienta-se que a abordagem a ser utilizada no capítulo 4 será a tradicional, ou seja, comparando-se o erro de previsão da volatilidade extraída dos modelos em relação aos quadrados dos retornos corrigidos pela média  $\{\varepsilon_t^2\}$ .

Neste capítulo, buscou-se discutir algumas técnicas paramétricas para a estimação dos segundos momentos condicionais. Partindo-se de uma breve revisão dos fatos estilizados em séries financeiras e dos modelos mais utilizados na literatura, espera-se poder facilitar a compreensão das técnicas não paramétricas de estimação da volatilidade discutidas no capítulo 3.

### 3 MÉTODOS NÃO PARAMÉTRICOS

A principal motivação para se trabalhar com modelos não paramétricos ao invés dos paramétricos se apresenta quando o pesquisador não possui conhecimento suficiente sobre o processo gerador explicativo, permitindo desta forma que os dados “falem por si mesmos”. Parte-se, desta maneira, para a análise dos dados sem levar em consideração opiniões prévias. A única condição que deve ser especificada para os estimadores não paramétricos está ligada a definição de um certo grau de suavidade para as funções que serão estimadas.

Segundo Ziegelmann (2002a), os modelos não paramétricos são aqueles que contêm uma ou mais funções desconhecidas e não apresentam espaço paramétrico de dimensão finita. Os semi-paramétricos, por sua vez, são compostos por uma ou mais funções desconhecidas e, por um espaço paramétrico desconhecido, de dimensão finita. Por fim, os modelos paramétricos possuem como única estrutura desconhecida o espaço paramétrico de dimensão finita. Assim, pode-se dizer que a principal vantagem dos modelos não paramétricos está ligada a sua flexibilidade e a conseqüente diminuição do risco de errar na especificação do modelo. Por outro lado, os custos desta maior flexibilidade se apresentam em taxas de convergência mais lentas, perda do poder de extrapolação e o problema da dimensionalidade (para o último item, ver seção 3.3). Distintos métodos não paramétricos podem ser utilizados para se modelar as estruturas explicativas, dentre eles, podem ser citados a suavização *kernel* e *splines*. As técnicas de suavização *kernel* podem ser encontradas em referências como Silverman (1986), Wand e Jones (1995), Fan e Gijbels (1996), Bowman e Azzalini (1997), Ziegelmann (2002a) e em Fan e Yao (2003).

Neste capítulo, busca-se apresentar o método de estimação da volatilidade via regressão polinomial local. Para isto, apresenta-se inicialmente a técnica de regressão para a média condicional tanto no contexto univariado (seção 3.1), como no contexto multivariado (seção 3.3). O método é estendido para a

estimação da variância condicional (volatilidade) na seção 3.4 e, por fim, os Modelos aditivos são discutidos na seção 3.5.

### 3.1 REGRESSÃO UNIVARIADA

O método de regressão linear paramétrica se apresenta como uma das técnicas mais utilizadas, onde para os pares de dados  $(X_i, Y_i)$ ,  $i = 1, \dots, n$ , procura-se ajustar globalmente uma reta ao longo das observações. A técnica de regressão linear é bastante útil se a esperança condicional aos dados é linear. Contudo, se a relação entre as variáveis for não linear, pode-se dizer que a técnica irá criar um viés elevado na modelagem. A forma habitual de lidar com a última situação se dá através do aumento do número de parâmetros, utilizando regressão polinomial. Embora esta prática seja amplamente utilizada, ela apresenta algumas desvantagens. Uma delas está relacionada as funções polinomiais não serem muito flexíveis na modelagem de algumas estruturas. Isto se deve ao fato das mesmas apresentarem todas as ordens de derivadas em toda a parte. Outra desvantagem é que observações individuais podem apresentar grande influência em partes remotas da curva estimada (FAN; GIJBELS, 1996).

Dentro do arcabouço de regressão não paramétrica, existem diversas formas de superar as deficiências do ajuste polinomial. Uma das possibilidades, via regressão local, encontra-se no uso da suavização *kernel*. A descrição dessa técnica será efetuada nesta seção, sendo que a mesma será utilizada ao longo de toda a dissertação.

#### 3.1.1 Suavização *Kernel* Polinomial Local

Apresenta-se aqui uma sucinta descrição de regressão de núcleo (*kernel*) através do ajuste polinomial local e, posteriormente, destacam-se os estimadores Nadaraya-Watson, linear local (*local linear kernel*) e loess.

O estimador polinomial local de suavização *kernel* estima a função de regressão em um ponto particular, ajustando localmente um polinômio de ordem  $p$  via mínimos quadrados ponderados. Na apresentação da técnica, considere os dados bivariados  $(Y_t, X_t)$ , onde no contexto de séries temporais  $X_t$  representa  $Y_{t-1}$  e  $T$  denota o número disponível de pares na amostra. A esperança condicional e a variância condicional são definidas respectivamente como  $m(x) = E(Y|X = x)$  e  $\sigma^2(x) = \text{Var}(Y|X = x) > 0$ . Supõe-se um modelo que relaciona as variáveis com a seguinte forma:

$$Y_t = m(X_t) + \sigma(X_t)\varepsilon_t, \quad (3.1)$$

onde  $E(\varepsilon_t|X_t) = 0$  e  $\text{Var}(\varepsilon_t|X_t) = 1$ . Adicionalmente,  $m'(z), m''(z), \dots, m^{(p)}(z)$  são as derivadas da função de regressão  $m(z)$ . Por meio de uma expansão de Taylor, para  $x$  na vizinhança local de  $x_0$ , assumindo que a derivada de  $m(x_0)$  de ordem  $(p+1)$  existe,  $m(x)$  pode ser escrita como:

$$m(x) = m(x_0) + m'(x_0)(x - x_0) + \frac{m''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{m^{(p)}(x_0)}{p!}(x - x_0)^p + O\left\{(x - x_0)^{p+1}\right\}, \quad (3.2)$$

ou seja, pode-se aproximar a função de regressão desconhecida  $m(X_t)$  da equação (3.1) por um polinômio de ordem  $p$ . Considerando-se um ponto de interesse  $x$  no domínio da variável aleatória  $X$ , então pode-se definir o estimador polinomial local de  $m(x)$  como  $\hat{m}_p(x) = \hat{\beta}_0$ , com  $\hat{\beta}_0$  dado pela solução do problema de mínimos quadrados ponderados a seguir:

$$\left(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p\right) = \arg \min_{\beta_0, \beta_1, \dots, \beta_p} \sum_{t=1}^T \left\{ Y_t - \sum_{i=0}^p \beta_i (X_t - x)^i \right\}^2 K_h(X_t - x), \quad (3.3)$$

onde  $K(\cdot)$  é uma função densidade de probabilidade simétrica em  $\mathfrak{R}$ ,  $K_h(u) = (1/h)K(u/h)$ , e  $h > 0$  é o parâmetro de suavização ou janela (*bandwidth*). Os

parâmetros  $\beta_i$  dependem de  $x$  e são chamados de parâmetros locais. Da teoria tradicional de mínimos quadrados, sabe-se que o estimador  $\hat{\beta}_0$  tem forma fechada e, que em notação matricial é o primeiro elemento de

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} = (\mathbf{X}'\mathbf{K}\mathbf{X})^{-1} \mathbf{X}'\mathbf{K}\mathbf{y} , \quad (3.4)$$

onde

$$\mathbf{X} = \begin{pmatrix} 1 & (X_1 - x) & \cdots & (X_1 - x)^p \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_T - x) & \cdots & (X_T - x)^p \end{pmatrix} , \quad \mathbf{y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_T \end{pmatrix} ,$$

e  $\mathbf{K}$  é a matriz diagonal de pesos  $T \times T$

$$\mathbf{K} = \text{diag} \{K_h(X_t - x)\} .$$

Olhando-se para as expressões (3.2) e (3.3), conclui-se que  $\beta_\nu = m^{(\nu)}(x)/\nu!$ . Isto significa que estimadores das derivadas de  $m(x)$  podem ser obtidos através de  $\hat{m}^{(\nu)}(x) = \hat{\beta}_\nu \nu!$ .

A idéia principal da suavização de núcleo local (*local kernel smoothing*) é que, sem assumir uma forma específica para a função de regressão  $m(\cdot)$ , um determinado ponto do conjunto de dados distante do ponto de interesse  $x$  deveria carregar pouca informação sobre o valor de  $m(x)$  (ZIEGELMANN, 2002a).

Dos suavizadores discutidos nesta seção, Nadaraya-Watson, linear local e loess, todos são casos especiais do estimador polinomial local. O primeiro dos três é o que apresenta a estrutura mais simples, sendo uma média local

$$\hat{m}_{nw}(x) = \frac{\sum_{t=1}^T K_h(X_t - x) Y_t}{\sum_{t=1}^T K_h(X_t - x)} , \quad (3.5)$$



e proposto por Nadaraya (1964) e Watson (1964). O suavizador (3.5) é definido, no ajuste polinomial local, com a ordem do polinômio igual a zero ( $p=0$ ). O mesmo efetua aproximações da função através de constantes locais. Sua simplicidade leva a um número de desvantagens frente a outros métodos, sendo a principal delas o viés nas regiões de fronteira (LOADER, 2004).

O estimador linear local, por sua vez, é utilizado ajustando-se localmente aos dados um polinômio, com grau  $p=1$ , ou seja, incluindo o termo linear da expansão de Taylor. A forma explícita não-matricial para este suavizador é apresentada abaixo:

$$\hat{m}_t(x) = T^{-1} \sum_{t=1}^T \frac{[\hat{s}_2(x) - \hat{s}_1(x)(X_t - x)] K_h(X_t - x) Y_t}{\hat{s}_2(x) \hat{s}_0(x) - \hat{s}_1(x)^2}, \quad (3.6)$$

onde

$$\hat{s}_r(x) = T^{-1} \sum_{t=1}^T (X_t - x)^r K_h(X_t - x).$$

A vantagem principal do estimador linear local diz respeito a sua capacidade de adaptação às regiões de fronteira (ZIEGELMANN, 2002a).

O suavizador ponderado localmente (*locally weighted smoother*) de Cleveland (1979), conhecido como loess na linguagem S, é mais um caso especial de regressão polinomial local. Para apresentá-lo, deve-se primeiramente definir uma função peso  $W$ , com as características que seguem:

- a)  $W(x) > 0$  para  $|x| < 1$ ;
- b)  $W(-x) = W(x)$ ;
- c)  $W(x)$  é uma função não-crescente para  $x \geq 0$ ;
- d)  $W(x) = 0$  para  $|x| \geq 1$ .

(3.7)

As vizinhanças locais são determinadas por uma janela com base em vizinhos próximos (*nearest neighbour bandwidth*). De forma complementar,  $0 < f \leq 1$  exprime uma fração (*span*) dos dados ( $n$ ) a serem incluídos em cada vizinhança, ao

passo que,  $fn$  representa uma quantidade de observações. Deve-se arredondar  $fn$  ao número inteiro mais próximo, denotado como  $r$ . O termo  $f$  representa o parâmetro de suavização e deve ser definido pelo pesquisador.

O procedimento de regressão pode ser explicado em linhas gerais, onde designa-se para cada  $x_i$ , pesos  $w_k(x_i)$ , que são definidos para todos  $x_k$ ,  $k=1, \dots, n$ , utilizando a função peso  $W$ . Isto é efetuado centrado a função  $W$  em  $x_i$  e, definindo a escala, tal que o ponto onde  $W$  primeiro se torna igual a zero é no  $r$ -ésimo vizinho próximo de  $x_i$ . Em loess a função peso normalmente usada é a tricúbica (*tricube kernel*), que segue a forma:

$$\begin{aligned} W(x) &= (1 - |x|^3)^3, \text{ para } |x| < 1 \\ &= 0, \quad \text{para } |x| \geq 1. \end{aligned} \quad (3.8)$$

Assim, o valor ajustado em cada  $x_i$ ,  $\hat{y}_i$ , é o valor estimado por um polinômio de ordem  $d$ , utilizando-se mínimos quadrados ponderados com os pesos  $w_k(x_i)$ .

Supondo-se que os dados são gerados por  $y_i = g(x_i) + \varepsilon_i$ , onde  $g$  é uma função suave, como nas estruturas mais comuns utilizadas em regressão supõe-se que  $\varepsilon_i$  é um erro aleatório independente, com média igual a zero e variância  $\sigma^2$ . Neste contexto,  $\hat{y}_i$  é uma estimativa de  $g(x_i)$  onde o pressuposto de suavidade permite que os pontos em uma vizinhança de  $(x_i, y_i)$  sejam utilizados na formação de  $\hat{y}_i$ . A função peso  $W(x)$  decresce para valores crescentes (não negativos) de  $x$  e, os pesos  $w_k(x_i)$  decrescem à medida que a distância de  $x_k$  em relação à  $x_i$  aumenta. Desta forma, os pontos próximos a  $x_i$  exercem grande influência na determinação de  $\hat{y}_i$ , ao contrário dos pontos distantes. O ato de elevar  $f$  aumenta a vizinhança em torno do ponto alvo e, assim, tende a exercer um maior grau de suavização (CLEVELAND, 1979).

No procedimento de estimação via loess, para cada  $i$ , define-se  $h_i$  (a janela) como a distância de  $x_i$  em relação ao seu  $r$ -ésimo vizinho próximo. Neste caso,  $h_i$  é o  $r$ -ésimo menor número entre  $|x_i - x_j|$ , para  $j=1, \dots, n$ . Para  $k=1, \dots, n$ ,

define-se  $w_k(x_i) = W(h_i^{-1}(x_k - x_i))$ . Assim, para cada  $i$ , calcula-se as estimativas  $\hat{\beta}_j(x_i)$ ,  $j = 0, \dots, d$  dos parâmetros, em uma regressão polinomial de grau  $d$ , de  $y_k$  em  $x_k$ . O ajuste se dá via mínimos quadrados ponderados, com pesos  $w_k(x_i)$  para  $(x_k, y_k)$ . Desta maneira,  $\hat{\beta}_j(x_i)$  são obtidos através dos valores de  $\beta_j$  que minimizam:

$$\sum_{k=1}^n w_k(x_i) (y_k - \beta_0 - \beta_1 x_k - \dots - \beta_d x_k^d)^2. \quad (3.9)$$

O ponto suavizado em  $x_i$ , é  $(x_i, \hat{y}_i)$ , onde  $\hat{y}_i$  é o valor estimado da regressão em  $x_i$ . Assim,

$$\hat{y}_i = \sum_{j=0}^d \hat{\beta}_j(x_i) x_i^j = \sum_{k=1}^n r_k(x_i) y_k, \quad (3.10)$$

onde  $r_k(x_i)$  não depende de  $y_j$ , para  $j = 1, \dots, n$ .

A principal diferença entre o estimador loess, com o grau da função polinomial ( $p = 1$ ), e o estimador linear local, dá-se na definição da janela. Para o primeiro, a janela é variável (*nearest neighbour bandwidth*), sendo definida pelo pesquisador a fração (*span*) de observações incluídas em cada vizinhança local (BOWMAN; AZZALINI, 1997). Para o segundo método, a janela é constante, ou seja, usa-se o mesmo parâmetro de suavização nos pesos definidos para cada ponto  $(x_i, y_i)$ .

Ressalta-se ainda a diferença de loess em relação à versão robusta de regressão ponderada localmente, chamada *Locally weighted scatterplot smoothing* (Lowess), introduzida por Cleveland (1979). O último utiliza os resíduos da regressão polinomial local para ponderar novamente o problema de minimização original. O mesmo é repetido para um número de estágios, e tem como idéia básica diminuir o peso, em cada estágio, dos pontos que apresentarem grandes resíduos.

## 3.2 ESPECIFICAÇÕES EM REGRESSÃO POLINOMIAL LOCAL

Utilizando-se o estimador polinomial local, destaca-se que algumas escolhas devem ser realizadas pelo pesquisador, tais como: a ordem  $p$  do polinômio, a função peso (*kernel*)  $K$  e a janela  $h$ .

Em relação à janela, caso a mesma seja muito grande ( $h = \infty$ ), o ajuste polinomial local se torna um ajuste polinomial global, legando a complexidade do modelo à especificação da ordem  $p$  do polinômio. No caso de uma janela igual a zero ( $h = 0$ ), a estimativa obtida será uma interpolação dos dados. Desta forma, com o parâmetro de suavização partindo de 0 a  $+\infty$ , a função estimada varia do modelo mais complexo (interpolação) ao modelo mais simples (modelo paramétrico polinomial). Assim, de forma distinta aos modelos paramétricos, a complexidade do ajuste polinomial local é primordialmente controlada pelo parâmetro de suavização. Por isso, a ordem  $p$  do polinômio é usualmente pequena e sua escolha se torna secundária nas aplicações (FAN; GIJBELS, 1996). A estimação de polinômios de ordem elevada pode levar a uma redução no viés, mas em contrapartida leva a um aumento da variância (devido à inclusão de um maior número de parâmetros locais).

Nas seções abaixo, apresentam-se alguns tipos de função peso (*kernel*), bem como métodos de escolha da janela, parâmetro de suma importância em regressão local.

### 3.2.1 A Função Peso (*Kernel*)

Na escolha do *kernel*, o qual é responsável por ponderar as observações na vizinhança de cada ponto de interesse  $x_0$ , salienta-se que algumas características devem fazer parte desta função. De acordo com Cleveland e Loader (1996), a função peso deve ser contínua, simétrica, com maior peso em torno de  $x_0$  e decrescente na medida em que  $x$  se distancia de  $x_0$ .

Entre as possíveis escolhas, apresentam-se as funções retangular, tricúbica, de Epanechnikov e a gaussiana. As funções peso  $K$  são definidas em torno

da variável transformada  $u$ , ou seja,  $K(u) = K[(x_i - x_0)/h]$ . Inicia-se com a apresentação da função retangular, sendo que esta pondera as observações na vizinhança de interesse de  $x_0$ , da seguinte forma:

$$\begin{aligned} K(u) &= 1 \text{ se } |u| < 1, \\ &= 0 \text{ se } |u| \geq 1. \end{aligned} \quad (3.11)$$

A definição acima não é utilizada com freqüência pois apresenta estimativas com descontinuidades.

Outra escolha disponível, definida anteriormente na expressão (3.8), é a função tri-cúbica (*tricube kernel*). Esta é utilizada em loess por apresentar a característica de decair suavemente a zero (FAN; GIJBELS, 1996). Como terceira opção, pode-se citar o *kernel* de Epanechnikov (*optimal kernel*) o qual minimiza o critério *AMISE* (erro em média quadrática integrado assintótico). Para maiores detalhes sobre os critérios de escolha da função *kernel*, ver Ziegelmann (2002a), Fan e Yao (2003). A definição da função peso segue:

$$\begin{aligned} K(u) &= \frac{3}{4}(1-u^2) \text{ se } |u| < 1, \\ &= 0 \quad \text{se } |u| \geq 1. \end{aligned} \quad (3.12)$$

Uma desvantagem existente no último *kernel*, é que o mesmo não é diferenciável em todos os pontos.

Como último exemplo, tem-se a função peso gaussiana. Devido à função densidade normal estar centrada em  $x_0$ , a janela  $h$  é representada pelo desvio padrão (BOWMAN; AZZALINI, 1997). Desta forma, valores que se encontrem a mais de dois desvios ( $2h$ ) receberão um peso negligenciável,

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} = \frac{1}{\sqrt{2\pi}} e^{-(x_i - x_0)^2/2h^2} \text{ se } -\infty < u < +\infty. \quad (3.13)$$

A busca por uma função peso ótima não é uma questão crítica no contexto de regressão polinomial local, sendo que diversos formatos de *kernel* possuem pequenas diferenças em termos de eficiência assintótica.

### 3.2.2 A Escolha da Janela

Os métodos para a escolha da janela podem ser separados em dois grupos. O primeiro é formado por métodos baseados em anexos (*plug-ins*), onde em linhas gerais, o viés de uma estimativa  $\hat{m}$  é escrito em termos da função  $m$  desconhecida, sendo geralmente aproximada por meio de uma expansão de Taylor. Uma estimativa de  $m$  é então anexada (*plugged-in*) para derivar uma estimativa do viés e, uma medida de ajuste, tal como o erro em média quadrática integrado (*MISE*). Desta maneira, escolhe-se uma janela ótima que minimize a medida de ajuste definida (LOADER, 1995). O segundo grupo, constitui-se de métodos clássicos, derivados em certa forma dos procedimentos utilizados em regressão paramétrica, onde dentre os mesmos, tem-se: validação cruzada (*cross-validation*) e o critério de informação de Akaike (*AIC*).

#### 3.2.2.1 Seletores Anexos (*Plug-in*) e a Janela Ótima

Para iniciar, apresenta-se o erro em média quadrática ponderado integrado, de  $\hat{m}_h(\cdot)$ , o estimador da média condicional  $m(\cdot)$ :

$$MISE(h) = \int E \{ \hat{m}_h(x) - m(x) \}^2 f(x) dx, \quad (3.14)$$

onde  $f(x)$ , a densidade de probabilidade de  $X_i$ , é a função peso. Minimizando a aproximação assintótica da última expressão em relação a  $h$  (*bandwidth*), no ajuste polinomial local, obtém-se a janela ótima abaixo:

$$h_{opt} = C_p(K) \left\{ \frac{\int \sigma^2(x) dx}{T \int [m^{(p+1)}(x)]^2 f(x) dx} \right\}^{1/(2p+3)}, \quad (3.15)$$

onde

$$C_p(K) = \left\{ \frac{(p+1)p! \int K^{*2}(t) dt}{2 \left[ \int t^{p+1} K^*(t) dt \right]^2} \right\}^{1/(2p+3)},$$

$$\text{e } K^*(t) = \mathbf{e}_1' \mathbf{S}^{-1} (1, t, \dots, t^p)' K(t) \quad (3.16)$$

é denominado *kernel* equivalente (expressa como o método de aproximação polinomial local atribui os pesos para cada ponto da amostra). O termo  $\mathbf{e}_1 = (1, 0, \dots, 0)$  representa um vetor unitário  $(p+1)$ , preenchido com o valor 1 na primeira posição. Define-se  $\mathbf{S}$  como uma matriz  $(p+1) \times (p+1)$  tal que o elemento  $(i, j)$  é  $\mu_{i+j-2}(K) = \int_{-\infty}^{\infty} u^{i+j-2} K(u) du$ . A intenção do método é a de usar a expressão (3.15) como um seletor de janela, mas como existem alguns termos desconhecidos na última expressão, os mesmos necessitam ser estimados. O método *direct plug-in* consiste em estimar os termos desconhecidos da expressão (3.15) e substituir os mesmos por suas estimativas (ZIEGELMANN, 2002a).

### 3.2.2.2 Validação Cruzada

O método de validação cruzada se apresenta como uma alternativa para a seleção de parâmetros de suavização através da estimação do erro de previsão de um estimador. O princípio básico do método consiste em prever cada valor da resposta  $Y_t$ , do restante dos dados, isto é, deixando-se um dos pontos da amostra de lado para a validação do modelo e, utilizando-se as observações restantes para a estimação. A função de validação cruzada é definida como:

$$CV(h) = T^{-1} \sum_{t=1}^T \{Y_t - \hat{m}_{h,-t}(X_t)\}^2, \quad (3.17)$$

onde  $\hat{m}_{h,-t}(X_t)$  é um estimador da média condicional  $m(X_t)$ , com a observação  $(X_t, Y_t)$  excluída. Pode-se mostrar que

$$E\{CV(h)\} = T^{-1} \sum E\{\hat{m}_{h,-t}(X_t) - m(X_t)\}^2 + T^{-1} \sum \sigma^2(X_t). \quad (3.18)$$

Partindo-se do pressuposto que o segundo termo  $T^{-1} \sum \sigma^2(X_t)$  não depende de  $h$ , a validação cruzada (3.17) apresenta-se como um estimador da estatística  $MISE(h)$ , definida em (3.14). Assim, este estimador da janela, representado por  $\hat{h}_{CV}$ , é definido como o  $h$  que minimiza a expressão (3.17). Uma das vantagens de se utilizar o método está relacionada a generalidade de sua definição, que permite sua aplicação a uma gama de estruturas distintas. Neste sentido, salienta-se que a função  $CV(h)$  pode ser representada apropriadamente quando o estimador  $\hat{m}(\cdot)$  é definido em termos de uma janela variável (usando distâncias de vizinhos próximos), como em loess (BOWMAN; AZZALINI, 1997).

Como método alternativo à validação cruzada, pode-se utilizar um critério de informação como o de Akaike ( $AIC$ ). Apresenta-se a última estatística, em um contexto não paramétrico, na seção 3.5.1.

### 3.3 REGRESSÃO MÚLTIPLA

Para trabalhar com regressão múltipla, deve-se focar a atenção na relação estrutural entre a variável resposta  $Y$  e o vetor de covariáveis  $(X_1, X_2, \dots, X_d)'$ . Desta maneira, pode-se estender o modelo (3.1) ao caso multivariado, onde  $\mathbf{X}$  passa a ser um vetor de dimensão- $d$  em  $\Re^d$  (em séries de tempo  $\mathbf{X}_t = (Y_{t-1}, Y_{t-2}, \dots, Y_{t-d})'$ ) e  $Y$  um escalar em  $\Re$ . Também, pode-se representar



a esperança e a variância condicionais como  $m(\mathbf{x}) = E(Y|\mathbf{X} = \mathbf{x})$  e  $\sigma^2(\mathbf{x}) = \text{Var}(Y|\mathbf{X} = \mathbf{x}) > 0$ , respectivamente. O modelo (3.1) assume a forma:

$$Y_t = m(\mathbf{X}_t) + \sigma(\mathbf{X}_t)\varepsilon_t, \quad (3.19)$$

onde  $E(\varepsilon_t|\mathbf{X}_t) = 0$  e  $\text{Var}(\varepsilon_t|\mathbf{X}_t) = 1$ .

Nesta seção, estende-se a análise para regressão múltipla através do ajuste linear local, ou seja,  $p = 1$  no contexto de regressão polinomial local. Para isto, define-se  $K$  como uma função densidade de probabilidade simétrica  $d$ -variada e, além disso,

$$K_{\mathbf{B}}(\mathbf{u}) = \frac{1}{|\mathbf{B}|} K(\mathbf{B}^{-1}\mathbf{u}), \quad (3.20)$$

onde  $\mathbf{B}$  é uma matriz invertível  $d \times d$ , que representa a matriz de suavização (*bandwidth matrix*). Assim, o estimador de regressão multivariada de  $m(x)$  é o primeiro elemento de:

$$\begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_d \end{pmatrix} = (\mathbf{X}'_{\mathbf{D}} \mathbf{W} \mathbf{X}_{\mathbf{D}})^{-1} \mathbf{X}'_{\mathbf{D}} \mathbf{W} \mathbf{y}, \quad (3.21)$$

onde

$$\mathbf{X}_{\mathbf{D}} = \begin{pmatrix} 1 & (X_{11} - x_1) & \cdots & (X_{1d} - x_d) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_{T1} - x_1) & \cdots & (X_{Td} - x_d) \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_T \end{pmatrix},$$

e  $\mathbf{W}$  é a matriz diagonal de pesos  $T \times T$ ,

$$\mathbf{W} = \text{diag} \{ K_B(\mathbf{X}_t - \mathbf{x}) \}.$$

O suavizador loess, discutido na seção 3.1.1, também apresenta sua versão multivariada, desenvolvida em Cleveland e Devlin (1988). Porém, como nas aplicações numéricas do capítulo 4 trabalha-se somente com o suavizador univariado, a versão estendida dos autores não é apresentada nesta dissertação.

Como principal desvantagem da regressão múltipla não paramétrica, pode-se citar o problema da dimensionalidade (*curse of dimensionality*). O problema está relacionado à escassez de dados em espaços de alta dimensão (BELLMAN, 1961), refletindo o fato de que uma vizinhança local em alta dimensão deixa de ser local, necessitando-se de uma vizinhança muito maior para conter um certo número suficiente de dados. Como resultado, necessita-se de muitas observações para somente um número moderado de dimensões (ZIEGELMANN, 2002a). As soluções para este problema costumam estar ligadas a algum mecanismo de redução de dimensão, como será apresentado posteriormente com os modelos aditivos semi-paramétricos.

### 3.4 ESTIMAÇÃO DE VOLATILIDADE

Nesta etapa da dissertação, busca-se apresentar alguns estimadores univariados de volatilidade. A partir de uma breve discussão dos últimos, pode-se partir para a próxima seção, onde estende-se a utilização de um dos estimadores apresentados ao caso multivariado (modelo (3.19)), ou seja, permitindo-se que a função de autoregressão  $m(\cdot)$  e a função de volatilidade condicional  $\sigma(\cdot)$  dependam de variáveis com distintas defasagens.

Para iniciar a apresentação, pode-se partir do modelo univariado definido em (3.1):

$$Y_t = m(X_t) + \sigma(X_t)\varepsilon_t, \quad (3.22)$$

onde anteriormente discutiu-se como estimar a função  $m(\cdot)$  e, neste momento, busca-se apresentar métodos para se estimar a função de volatilidade condicional  $\sigma(\cdot)$ . Uma das formas de se estimar a última função, proposta por Müller e Stadtmüller (1987), é através da expressão (3.22) no caso de um modelo de delineamento fixo (*fixed design model*). Os autores propõem estimar  $\sigma^2(\cdot)$  através de uma seqüência diferencial, assumindo a forma:

$$\hat{\sigma}_{ds}^2(x_t) = \tilde{\sigma}^2(x_t) + \tilde{\varepsilon}_t, \quad (3.23)$$

onde

$$\tilde{\sigma}^2(x_t) = \left( \sum_{u=-m}^m \omega_u Y_{t+u} \right)^2, \quad (3.24)$$

para  $m > 0$ . A seqüência diferencial  $\{\omega_u\}$  satisfaz as condições abaixo,

$$\sum_{u=-m}^m \omega_u = 0, \quad \sum_{u=-m}^m \omega_u^2 = 1. \quad (3.25)$$

Um segundo método é motivado pela conhecida decomposição da variância condicional na seguinte forma:  $Var(Y|X=x) = E(Y^2|X=x) - E(Y|X=x)^2$ .

Isto leva ao estimador direto de  $\sigma^2(x)$ :

$$\hat{\sigma}_{dir}^2(x) = \hat{v}(x) - \hat{m}(x)^2, \quad (3.26)$$

onde  $\hat{m}(x)$  e  $\hat{v}(x)$  são os respectivos estimadores de  $m(x)$  e de  $v(x) = E(Y^2|X=x)$ . Em outra abordagem, Yang, Härdle e Nielsen (1999) derivam resultados para um modelo não paramétrico, com média aditiva, mas com a volatilidade multiplicativa. A estimação das funções se dá através do método de integração marginal (ZIEGELMANN, 2002a).

O método que será utilizado nas aplicações numéricas desta dissertação é baseado nos resíduos, cujo nome é estimador residual da variância. Alguns autores como Ruppert, Wand, Holst e Hössjer (1997), Fan e Yao (1998) e Ziegelmann (2002b, 2008) apresentam diversas implementações sobre a técnica. Resumidamente, pode-se dizer que o método consiste em reescrever o modelo<sup>4</sup> (3.22) como

$$Y_t - m(X_t) = \sigma(X_t)\varepsilon_t, \quad (3.27)$$

tal que tomando-se a esperança condicional dos resíduos ao quadrado,  $r_t^2 = [Y_t - m(X_t)]^2 = \sigma^2(X_t)\varepsilon_t^2$ , obtém-se  $E(r_t^2|X_t) = \sigma^2(X_t)$ . Assim, pode-se estimar  $\sigma^2(X_t)$  a partir de uma regressão de  $\hat{r}_t^2$  contra  $X_t$ , onde  $\hat{r}_t^2 = Y_t - \hat{m}(X_t)$ . A prova da normalidade assintótica deste estimador é feita por Fan e Yao (1998), quando o estimador linear local é utilizado para o ajuste tanto de  $m(x)$  quanto de  $\sigma^2(x)$ . Além disso, os autores provam que o estimador desenvolvido para a variância condicional é completamente adaptativo (assintoticamente) a média condicional desconhecida. Com o intuito de fazer uma revisão dos estimadores univariados da volatilidade condicional, baseados nos resíduos, busca-se seguir Ziegelmann (2002a, b, 2008). Contudo, cabe destacar que as provas teóricas são deixadas de lado, podendo ser encontradas nas três referências.

Para iniciar, pode-se apresentar o estimador residual linear local da variância condicional, denotado como  $\hat{\sigma}_t^2(x) = \hat{\alpha}$ , obtido através do problema de mínimos quadrados ponderados em relação a  $\alpha$ :

---

<sup>4</sup> Como anteriormente,  $X_t$  representa  $Y_{t-1}$  em séries de tempo.

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{\alpha, \beta} \sum_{t=1}^T \left\{ \hat{r}_t^2 - \alpha - \beta(X_t - x) \right\}^2 K_{h_2}(X_t - x), \quad (3.28)$$

onde  $K$  é uma função densidade de probabilidade simétrica e  $h_2 > 0$  representa uma segunda janela ( $h_1$  é definida na estimação de  $m(x)$ ). Com o intuito de obter somente estimativas não negativas para a volatilidade condicional, pode-se impor uma restrição que atue no estimador linear local com este propósito. Este procedimento leva ao estimador residual linear local restrito,  $\hat{\sigma}_c^2(x) = \exp \hat{\alpha}$ , onde pode-se obter  $\hat{\alpha}$  da seguinte forma:

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{\alpha, \beta} \sum_{t=1}^T \left\{ \hat{r}_t^2 - \exp\{\alpha\} - \beta(X_t - x) \right\}^2 K_{h_2}(X_t - x). \quad (3.29)$$

Um método alternativo ao último, desenvolvido por Ziegelmann (2002b), recebe o nome de estimador residual local exponencial da variância condicional. A idéia do método é assegurar a obtenção de valores não negativos para as estimativas de volatilidade. Quando compara-se o estimador linear local restrito ao local exponencial, a diferença entre os dois se encontra na medida que o primeiro efetua o ajuste de forma linear e o segundo de forma exponencial. Além disso, o autor prova teoricamente a completa adaptação assintótica do estimador exponencial a esperança condicional desconhecida. O estimador local exponencial para a volatilidade condicional é obtido através da solução da minimização que segue:

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{\alpha, \beta} \sum_{t=1}^T \left\{ \hat{r}_t^2 - \psi\{\alpha + \beta(X_t - x)\} \right\}^2 K_{h_2}(X_t - x), \quad (3.30)$$

onde, como anteriormente,  $K$  é uma função densidade de probabilidade simétrica e  $h_2 > 0$  é a janela. Destaca-se que quando  $\psi(x) \equiv x$ , obtém-se o estimador linear local  $\hat{\sigma}_t^2(x) = \hat{\alpha}$ . A forma funcional do estimador local exponencial é definida como

$\psi(x) \equiv \exp(x)$ , e o estimador da volatilidade condicional assume a forma  $\hat{\sigma}_e^2(x) = \exp \hat{\alpha}$ .

Uma versão robusta do estimador residual da variância condicional é apresentada em Ziegelmann (2008). Esta versão se baseia na evidência empírica que os retornos de séries de tempo financeiras apresentam distribuições marginais com caudas pesadas. Desta forma, segundo o autor, estes retornos podem não apresentar momentos superiores finitos, o que afeta o uso de estimadores baseados em mínimos quadrados (*least-square-deviations* (LSD)). Assim, Ziegelmann (2008) desenvolve o estimador residual linear local baseado em mínimos desvios absolutos (*least-absolute-deviations* (LAD)) para a variância condicional. Para isto, um modelo um pouco distinto ao (3.27) é definido, onde neste, a função de volatilidade se torna a mediana condicional do quadrado dos resíduos.

Considerando-se os dados bivariados  $(Y_t, X_t)$ , define-se respectivamente a esperança condicional e a volatilidade condicional como  $m(x) = E(Y|X = x)$  e  $\tau(x) = M\left\{\left[Y - m(x)\right]^2 | X = x\right\} > 0$ , onde  $M\{Y|X\}$  representa a mediana condicional de  $Y$  dado  $X$ . Assim, o modelo assume a forma:

$$Y_t = m(X_t) + \tau^{1/2}(X_t)\varepsilon_t, \quad (3.31)$$

onde  $E(\varepsilon_t|X_t) = 0$  e  $M(\varepsilon_t^2|X_t) = 1$ . Neste método, a função  $m(x)$  continua a ser estimada pelo estimador linear local (LSD), definido em (3.6). Todavia, para estimar a volatilidade condicional,  $\tau(\cdot)$ , em uma segunda etapa, deve-se reescrever o modelo (3.31) como:

$$r_t^2 = (Y_t - m(X_t))^2 = \tau(X_t)\varepsilon_t^2, \quad (3.32)$$

onde tomando-se a mediana condicional em (3.32), obtém-se  $M(r_t^2|X_t) = \tau(X_t)$ . Assim, o estimador da volatilidade condicional é dado por  $\hat{\tau}(x) = \hat{\gamma}_1$ , que é a solução

para  $\gamma_1$  no problema de minimização dos desvios absolutos ponderados (*weighted least-absolute-deviations*) que segue:

$$(\hat{\gamma}_1, \hat{\gamma}_2) = \arg \min_{\gamma_1, \gamma_2} \sum_{t=1}^T \left| \hat{r}_t^2 - \gamma_1 - \gamma_2 (X_t - x) \right| K_{h_2} (X_t - x), \quad (3.33)$$

onde  $K$  é uma função densidade de probabilidade simétrica e  $h_2 > 0$  é a janela.

### 3.4.1 Estimação de Volatilidade em Regressão Múltipla

Após a apresentação de métodos para se obter estimativas da volatilidade condicional, no contexto univariado, busca-se nesta seção estender a discussão ao caso multivariado. Neste sentido, supõe-se que o comportamento da variável dependente, o quadrado dos resíduos, seja melhor explicado quando utiliza-se mais de uma variável defasada no conjunto de informação.

No contexto multivariado, o modelo saturado (ou totalmente não paramétrico) destaca-se como a opção mais flexível na redução do viés de especificação. Entretanto, ressalta-se que problemas relacionados a espaços de alta dimensão são freqüentes, e como citado anteriormente, recebem o nome de problema da dimensionalidade (*curse of dimensionality*). Este leva a perda da noção de vizinhança local, o que abre espaço para o uso de diversos modelos como os não paramétricos não-saturados ou semi-paramétricos. Estes, por sua vez, impõem algum tipo de restrição nas funções de volatilidade, por isso, são menos livres que o modelo não paramétrico saturado (ZIEGELMANN, 2002a).

O modelo aditivo, amplamente discutido em Hastie e Tibshirani (1990), é apresentado como alternativa ao problema da dimensionalidade na seção 3.5. Neste método, modelam-se funções de volatilidade univariadas de maneira aditiva, abandonando-se o pressuposto da linearidade definido no modelo de regressão paramétrico.

### 3.4.1.1 O Modelo Saturado e o Estimador Residual Multivariado

O método de estimação da variância condicional, a partir do estimador residual multivariado, é uma generalização do caso univariado apresentado na seção anterior. A partir do modelo de regressão múltipla definido em (3.19),

$$Y_t = m(\mathbf{X}_t) + \sigma(\mathbf{X}_t)\varepsilon_t,$$

onde  $\mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-d+1})'$  é um vetor  $d$ -dimensional e,  $E(\varepsilon_t | \mathbf{X}_t) = 0$  e  $\text{Var}(\varepsilon_t | \mathbf{X}_t) = 1$ , pode-se definir o estimador residual linear local da volatilidade. O último é obtido através da regressão de  $\hat{r}_t^2 = (Y_t - \hat{m}(\mathbf{X}_t))^2$  contra o vetor  $\mathbf{X}_t$ . Assim, define-se a versão multivariada como  $\hat{\sigma}_{lm}^2(\mathbf{x}) = \hat{\beta}_0$ , que representa o primeiro elemento de:

$$\begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_d \end{pmatrix} = (\mathbf{X}'_D \mathbf{W} \mathbf{X}_D)^{-1} \mathbf{X}'_D \mathbf{W} \hat{\mathbf{r}}, \quad (3.34)$$

$$\mathbf{X}_D = \begin{pmatrix} 1 & (X_d - x_0) & \cdots & (X_1 - x_{d-1}) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_T - x_0) & \cdots & (X_{T-d+1} - x_{d-1}) \end{pmatrix}, \quad \hat{\mathbf{r}} = \begin{pmatrix} \hat{r}_d^2 \\ \vdots \\ \hat{r}_T^2 \end{pmatrix}.$$

Destaca-se que  $\mathbf{W}$  é a matriz diagonal de pesos  $(T-d+1) \times (T-d+1)$ ,  $\mathbf{W} = \text{diag}\{K_B(\mathbf{X}_t - \mathbf{x})\}$ , e  $K$ ,  $\mathbf{B}$  e  $K_B(\mathbf{u})$  são definidas como na seção 3.3.

## 3.5 MODELOS ADITIVOS

Para apresentar o modelo aditivo, considere primeiramente o modelo geral de regressão múltipla discutido em Fan e Gijbels (1996),



$$Y = m(\mathbf{X}) + \varepsilon, \quad (3.35)$$

onde  $E(\varepsilon) = 0$ ,  $Var(\varepsilon) = \sigma^2$  e  $\varepsilon$  é independente do vetor de covariáveis  $\mathbf{X}$ . No modelo de regressão linear múltipla a função de regressão  $m(\cdot)$  é assumida como linear, e portanto, aditiva nas variáveis explicativas. Nos modelos aditivos, o pressuposto de linearidade é abandonado, mas a forma aditiva é mantida. A idéia de um modelo aditivo, como sugerido por Friedman e Stuetzle (1981) e substancialmente desenvolvido por Hastie e Tibshirani (1990), é a de permitir que componentes do modelo de regressão linear assumam formas não paramétricas (BOWMAN; AZZALINI, 1997), o que leva ao modelo:

$$Y = \alpha + \sum_{j=1}^d g_j(X_j) + \varepsilon, \quad (3.36)$$

onde  $g_1, \dots, g_d$  são funções univariadas desconhecidas. No sentido de evitar constantes livres nas funções e, para garantir a identificabilidade do modelo, é requerido que

$$E\{g_j(X_j)\} = 0, \quad j = 1, \dots, d. \quad (3.37)$$

Isto implica que  $E(Y) = \alpha$  e, este modelo pode ser enquadrado como uma ferramenta de redução de dimensão. Salienta-se que o modelo aditivo (3.36) é um caso especial do modelo de procura projetiva (*projection pursuit model*), que apresenta o seguinte formato,

$$Y = \sum_{l=1}^L g_l(\boldsymbol{\alpha}_l' \mathbf{X}) + \varepsilon, \quad (3.38)$$

onde  $\boldsymbol{\alpha}_l$ ,  $l = 1, \dots, L$ , denota os vetores direcionais onde as observações são projetadas. As contribuições das projeções unidimensionais  $\boldsymbol{\alpha}_l' \mathbf{X}$ ,  $l = 1, \dots, L$ , são modeladas aditivamente através de funções univariadas  $g_l$ ,  $l = 1, \dots, L$  (FAN; GIJBELS, 1996).

Quando o modelo aditivo é válido, tem-se que

$$E \left\{ Y - \alpha - \sum_{j \neq k} g_j(X_j) \mid X_k \right\} = g_k(X_k), \quad k = 1, \dots, d. \quad (3.39)$$

Surge então um algoritmo iterativo que permite o cálculo de cada função univariada  $g_1, \dots, g_d$ . Desta forma, para um dado  $\alpha$  e para funções dadas  $g_j, j \neq k$ , a função  $g_k$  pode ser obtida através de uma simples regressão univariada baseada nas observações  $\{(X_{ik}, Y_i) : i = 1, \dots, n\}$ . O suavizador univariado de  $g_k$  é definido por  $S_k$ , onde nesta situação, qualquer técnica de suavização univariada pode ser utilizada, tal como *splines*, regressão polinomial local, entre outras. No sentido de respeitar a condição (3.37), a estimativa resultante  $\hat{g}_k(\cdot)$  de  $g_k(\cdot)$  obtida através do suavizador  $S_k$ , é substituída por sua versão centralizada:

$$\hat{g}_k^*(\cdot) = \hat{g}_k(\cdot) - \frac{1}{n} \sum_{j=1}^n \hat{g}_k(X_{jk}). \quad (3.40)$$

Uma escolha inicial das funções univariadas,  $g_k^0$ , torna-se necessária, tanto quanto o esquema de iteração. Este procedimento leva ao chamado *backfitting algorithm*, que segue os próximos passos:

- **Passo 1-** Inicialização:  $\hat{\alpha} = n^{-1} \sum_{i=1}^n Y_i$ ,  $\hat{g}_k = g_k^0, k = 1, \dots, d$ .
- **Passo 2-** Para cada  $k = 1, \dots, d$ , obter  $\hat{g}_k = S_k \left\{ Y - \hat{\alpha} - \sum_{j \neq k} \hat{g}_j(X_j) \mid X_k \right\}$  e obter  $\hat{g}_k^*(\cdot)$  como em (3.40).
- **Passo 3-** Manter rodando o passo 2 até convergir.

A idéia básica do algoritmo é a de através da regressão inicial, calcular os resíduos parciais desta, e regredir novamente. O termo *backfitting* se refere a esta ação e foi primeiramente utilizado por Friedman e Stuetzle (1981). O estimador  $n^{-1} \sum_{i=1}^n Y_i$  de  $\alpha$ , no algoritmo acima, segue naturalmente o fato que  $E(Y) = \alpha$ . Pode-se utilizar como funções iniciais  $g_1^0, \dots, g_d^0$ , por exemplo, os resultados obtidos através de uma regressão linear local de  $Y$  contra os  $X_k$  individualmente.

Ressalta-se que os modelos aditivos superam o “problema da dimensionalidade” (*curse of dimensionality*) devido ao sistema de ajuste ser construído a partir de suavizadores univariados. O preço pago por isto se mostra em uma perda de flexibilidade do modelo em relação ao modelo não paramétrico saturado (ZIEGELMANN, 2002a). A partir do modelo (3.36), espera-se que o algoritmo *backfitting* gere estimativas que sejam as melhores aproximações aditivas à superfície de regressão (FAN; GIJBELS, 1996).

Para estimar a volatilidade, o modelo aditivo é construído da seguinte forma:

$$\sigma^2(\mathbf{X}_t) = \sum_{i=0}^{d-1} f_i(X_{t-i}), \quad (3.41)$$

onde  $\mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-d+1})'$  é o vetor de covariáveis. Destaca-se que a última representação do modelo aditivo pode ser utilizada para ajustar de forma não paramétrica uma estrutura ARCH. A equação (3.41) apresenta a volatilidade condicional como uma soma de funções univariadas, uma para cada coordenada direcional. Salieta-se, como anteriormente, que qualquer suavizador não paramétrico, como o linear local, loess ou outros, podem ser usados para estimar cada função  $f_i(\cdot)$ ,  $i = 0, \dots, d-1$ . Contudo, este procedimento de estimação é distinto da simples realização de  $d$  regressões univariadas ordinárias, isto porque as outras dimensões devem ser levadas em conta no momento do ajuste em uma direção (ZIEGELMANN, 2002a).

Como antes, para prevenir constantes livres nas funções  $f_i(\cdot)$ , e desta forma assegurando-se a identificação do modelo (3.41), é imposto que  $Ef_i(X_{t-i}) = 0$ ,  $i = 0, \dots, d-1$ , o que leva a um termo adicional no modelo (3.41),

$$\sigma^2(\mathbf{X}_t) = \alpha + \sum_{i=0}^{d-1} f_i(X_{t-i}). \quad (3.42)$$

Desta forma, quando (3.42) é bem especificado, tem-se:

$$E \left\{ r_t^2 - \alpha - \sum_{i \neq k} f_i(X_{t-i}) \mid X_{t-k} \right\} = f_k(X_{t-k}), \quad k = 0, \dots, d-1,$$

tal que o quadrado dos resíduos,  $r_t^2 = \{Y_t - m(\mathbf{X}_t)\}^2$ , é utilizado como variável resposta. Assim, para um dado  $\alpha$  e para as funções dadas  $f_i$ ,  $i \neq k$ , a função  $f_k$  pode ser obtida através de uma simples regressão univariada. Denota-se o suavizador univariado de  $f_k$  como  $S_k$ , onde neste trabalho utiliza-se loess. As estimativas iniciais de  $f_k$  são definidas como  $f_k^0$ ,  $k = 0, \dots, d-1$ , e desta forma, o algoritmo *backfitting* segue os seguintes passos:

- **Passo 1-** Inicialização:  $\hat{\alpha} = (T-d+1)^{-1} \sum_{t=d}^T \hat{r}_t^2$ ,  $\hat{f}_k = f_k^0$ ,  $k = 0, \dots, d-1$ .
- **Passo 2-** Rodar  $k = 0, \dots, d-1$  e calcular  $\hat{f}_k^*(\cdot) = S_k \left\{ \hat{r}_t^2 - \hat{\alpha} - \sum_{i \neq k} \hat{f}_i(\cdot) \mid X_{t-k} \right\}$ . Centrar o estimador para obter,  $\hat{f}_k(\cdot) = \hat{f}_k^*(\cdot) - \frac{1}{T-d+1} \sum_{t=d}^T \hat{f}_k^*(X_{t-k})$ .
- **Passo 3-** Repetir o passo 2 até a convergência.

### 3.5.1 Inferência

Os modelos aditivos provêm alta flexibilidade e descrições visualmente informativas dos dados. Porém, métodos que permitam a inferência dos mesmos devem ser especificados, com o intuito de guiar o processo de estimação. Para isto, Hastie e Tibshirani (1990) recomendam o uso da soma dos quadrados dos resíduos e os seus graus de liberdade aproximados como mecanismo de auxílio na escolha dos modelos. Salienta-se que estas quantidades podem ser interpretadas em analogia aos modelos lineares paramétricos, embora a natureza não paramétrica dos modelos aditivos não permita o uso das distribuições  $\chi^2$  e  $F$  tradicionais, desta forma, trata-se de um procedimento heurístico de inferência.

A partir do modelo aditivo  $y_i = \alpha + \sum_{j=1}^d g_j(x_{ji}) + \varepsilon_i$ , a soma dos quadrados dos resíduos pode ser definida como segue,

$$SQR = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (3.43)$$

onde  $\hat{y}_i$  denota o valor ajustado através da avaliação do modelo na observação  $x_i$ . Como cada passo do algoritmo *backfitting* envolve uma matriz de suavização  $S_j$  (cujas linhas consistem dos pesos apropriados à estimação em cada observação  $x_i$ ), cada componente ( $\hat{g}_j$ ) ajustado pelo modelo aditivo pode ser representado na forma  $S_j y$ , onde  $y$  denota o vetor de respostas observadas. Por meio da última matriz os graus de liberdade aproximados (que assumem a função de número equivalente de parâmetros) podem ser calculados para cada componente estimado, bem como para o modelo completo (BOWMAN; AZZALINI, 1997).

Ainda em relação aos graus de liberdade,  $df$ , destaca-se que os mesmos são calculados através do traço da matriz de suavização,  $tr(S)$ , onde  $S$  é definida como a matriz que produz o vetor de estimativas  $\hat{g} = Sy$ . Segundo Bowman e Azzalini (1997), a matriz  $S$  é análoga a matriz de projeção (*hat matrix*) dos modelos de regressão linear. Adicionalmente, os graus de liberdade para cada elemento estimado ( $\hat{g}_j$ ) podem ser definidos como  $df_j = tr(S_j)$ .

A comparação entre modelos alternativos pode ser efetuada, por exemplo, a partir de um teste similar ao ANOVA, dito ANODEV (*Analysis of Deviance*), através de um teste  $F$  aproximado ou a partir do critério de informação aproximado de Akaike ( $AIC$ ).

O desvio (*deviance*) é uma função do logaritmo da verossimilhança, sendo freqüentemente utilizado como um critério de qualidade do ajuste. O desvio,  $D(y; \mu)$ , é definido por

$$\frac{D(y; \mu)}{\phi} = 2l(\mu^*; y) - 2l(\mu; y), \quad (3.44)$$

onde  $\mu^*$  maximiza o logaritmo da verossimilhança sobre  $\mu$  (não restrito) e  $\phi$  representa um parâmetro de dispersão. Na medida em que se utiliza uma variável resposta contínua com erros normais, o parâmetro de dispersão é dado pela

variância,  $\sigma^2$ , e o desvio reduz-se à soma dos quadrados dos resíduos  $\sum_i (y_i - \mu_i)^2$  (HASTIE, 1993).

O desvio pode ser utilizado na comparação de modelos concorrentes e, de forma similar à soma dos quadrados dos resíduos, pode ser tornado muito pequeno na medida em que se estabeleça um modelo que interpole os dados. Todavia, nos modelos aditivos, assim como nos modelos lineares paramétricos, procura-se manter as estruturas as mais simples possíveis. No caso linear, por exemplo, tenta-se manter mínimo o número de parâmetros no modelo. Os modelos aditivos, por sua vez, são geralmente não paramétricos, no entanto pode-se definir para os mesmos “um número equivalente de parâmetros”. Em analogia ao caso linear, busca-se minimizar os últimos. Como citado anteriormente, este número equivalente de parâmetros é definido em termos dos graus de liberdade ( $df$ ). Desta forma, como nos modelos paramétricos, os graus de liberdade dos resíduos são obtidos a partir da diferença do número de observações ( $n$ ) em relação ao número equivalente de parâmetros ( $df$ ) (HASTIE, 1993). Sendo assim, como  $df = tr(\mathbf{S})$ , logo  $df^{res} = n - tr(\mathbf{S})$ .

Para os modelos lineares paramétricos, a teoria distributiva assintótica é amplamente conhecida. Neste caso, uma forma de selecionar modelos alternativos se dá através da Análise dos Desvios (*Analysis of Deviance*). Em linhas gerais, supondo-se que  $\mu_1$  e  $\mu_2$  sejam dois modelos lineares, com  $\mu_1$  aninhado em  $\mu_2$ , então sob condições regulares e assumindo que  $\mu_1$  está especificado corretamente,  $D(\hat{\mu}_2; \hat{\mu}_1) = D(\mathbf{y}; \hat{\mu}_1) - D(\mathbf{y}; \hat{\mu}_2)$  tem distribuição assintótica  $\chi^2$  com graus de liberdade iguais a diferença nas dimensões dos dois modelos. Ressalta-se que para os modelos aditivos, o teste de desvio segue fazendo sentido como uma maneira de captar diferenças entre os modelos. Entretanto, como a teoria distributiva ainda não está desenvolvida, os testes de desvio são efetuados de forma heurística, onde a distribuição  $\chi^2$  é utilizada como uma aproximação razoável (HASTIE; TIBSHIRANI, 1990).

Outro método disponível para inferência nos modelos aditivos é o teste  $F$  aproximado. Neste teste, para fazer a comparação entre dois modelos, pode-se especificar o primeiro modelo, dito completo, e o segundo, definido a partir da

eliminação de uma das variáveis explicativas do primeiro (modelo aninhado ao primeiro). Desta forma, o teste  $F$  aproximado pode ser especificado como segue:

$$F = \frac{(SQR_2 - SQR_1) / (df_2^{res} - df_1^{res})}{SQR_1 / df_1^{res}}, \quad (3.45)$$

tal que sua estrutura é derivada do teste  $F$  tradicional, utilizado na comparação de modelos lineares paramétricos. Salienta-se que a analogia não se estende aos cálculos da distribuição, e tampouco existe expressão geral para a distribuição em (3.45). Apesar disso, Hastie e Tibshirani (1990) sugerem que ao menos alguma aproximação pode ser feita para auxiliar na tomada de decisão sobre os efeitos das variáveis na superfície de regressão. Os autores referem a estatística obtida por meio do teste  $F$  não paramétrico a uma distribuição  $F$  com  $(df_2^{res} - df_1^{res})$  e  $df_1^{res}$  graus de liberdade (BOWMAN; AZZALINI, 1997).

Como terceira técnica disponível para inferência nos modelos aditivos, o critério de informação aproximado de Akaike ( $AIC$ ) pode ser utilizado, tanto para selecionar modelos alternativos, como na escolha do parâmetro de suavização (*bandwidth ou span*). O cálculo do  $AIC$  pode ser definido da seguinte forma:

$$AIC = D(\mathbf{y}; \hat{\boldsymbol{\mu}}) / n + 2df\phi / n, \quad (3.46)$$

onde os graus de liberdade,  $df = tr(\mathbf{S})$ , tornam o  $AIC$  assintoticamente não viesado para o erro de previsão. O termo  $\phi$  da expressão (3.46) representa o parâmetro de dispersão.

O erro de previsão pode ser medido utilizando-se a distância de Kullback-Leibler, como segue:

$$PE = E \left\{ \frac{1}{n} \sum_{i=1}^n D(Y_i^0; \hat{\mu}_i) \right\}, \quad (3.47)$$

onde  $Y_i^0$  possui a mesma distribuição da realização  $y_i$  e,  $\hat{\mu}_i = \hat{\mu}(x^i)$  representa o ajuste aditivo baseado em  $\mathbf{y}$ . Utilizando a especificação (3.46) do  $AIC$ , pode-se descrever a aproximação do mesmo em relação à validação cruzada. Em

comparação a última estatística, o *AIC* apresenta-se como método menos custoso computacionalmente para a escolha da janela (HASTIE; TIBSHIRANI, 1990). Sua definição segue abaixo:

$$AIC = \frac{1}{n} \sum_{i=1}^n D(y_i; \hat{\mu}_i) + 2tr(\mathbf{S})\phi / n . \quad (3.48)$$

Por fim, assim como procedeu-se no capítulo 2 para os modelos paramétricos, no capítulo 3 buscou-se apresentar algumas propriedades básicas para a utilização de métodos não paramétricos. Concluída a discussão sobre as técnicas, parte-se para o capítulo 4 da dissertação, onde a implementação numérica será realizada confrontando as técnicas alternativas de estimação da volatilidade em um contexto multivariado.



## 4 APLICAÇÃO NUMÉRICA: MODELOS ADITIVOS SEMI-PARAMÉTRICOS E GARCH

Este capítulo irá confrontar as técnicas apresentadas ao longo da dissertação. Para isto, são selecionadas duas séries de tempo compostas dos preços de fechamento diários, de dois índices amplamente representativos do mercado financeiro brasileiro e norte-americano, o Índice Bovespa e o Standard and Poor's 500, respectivamente. A modelagem é realizada a partir da série dos retornos diários dos índices, definidos como  $y_t = \ln X_t - \ln X_{t-1}$  (onde  $X_t$  é o preço do ativo no tempo  $t$ ).

Baseando-se na amostra completa dos retornos, serão selecionados os modelos mais adequados para a estimação da volatilidade dos índices mencionados. Após a escolha das estruturas mais apropriadas, os modelos terão suas estimativas e previsões um passo à frente comparadas em períodos denominados de crise (alta volatilidade) no mercado financeiro internacional.

A estruturação do capítulo destaca na seção 4.1 os momentos selecionados de crise, com uma breve contextualização dos principais fatos ocorridos no ano. A partir disto, os dados utilizados são apresentados na seção 4.2, ao passo que as estimações nas amostras completas são realizadas na seção 4.3. Por fim, a seção 4.4 reúne as comparações dos modelos paramétricos e semi-paramétricos, dentro e fora da amostra, nos momentos de alta volatilidade de cada ano escolhido.

### 4.1 PERÍODOS SELECIONADOS

Para comparar a capacidade de extração da volatilidade dos modelos aditivos semi-paramétricos com os tradicionalmente utilizados na literatura, modelos GARCH paramétricos, foram selecionados três anos com crises no mercado

financeiro internacional. De forma mais específica, a análise foi delimitada para os momentos de alta volatilidade (crises) mais recentes em ambas as séries. Haja vista os modelos serem selecionados a partir das amostras completas dos retornos diários, decidiu-se dispor de metade das observações de cada série antes da primeira crise analisada. Este procedimento visa à obtenção de estimativas mais estáveis para os termos incluídos nos modelos. Sendo assim, a partir da avaliação do comportamento dos últimos quanto à estimação e previsão da variância condicional nos períodos de elevada volatilidade, conclusões acerca dos distintos métodos poderão ser obtidas.

Fundamentando-se nestes objetivos, será apresentada uma breve descrição dos principais fatos ocorridos nos anos que compreendem as crises, concentrando a atenção nos fatores que majoritariamente influenciaram o mercado financeiro. Salieta-se que o objetivo da revisão em torno das turbulências anuais é o de contextualizar meramente os períodos em que a volatilidade será avaliada através dos modelos. Além disso, como os índices selecionados são provenientes dos mercados brasileiro e norte-americano, o foco das revisões se dá em torno dos dois países.

#### **4.1.1 A Crise de 2001**

Em 2001, a economia mundial foi marcada pela recessão nos Estados Unidos (EUA), que influenciou negativamente a dinâmica produtiva nos demais países. Neste cenário, a evolução do produto mundial reduziu-se acentuadamente. Após crescer 4,7% em 2000 (preços constantes), registrou-se um aumento de apenas 2,3% no produto mundial em 2001 (INTERNATIONAL MONETARY FUND, 2003).

A economia norte-americana interrompeu em 2001 uma década de crescimento contínuo, entrando em recessão a partir de março, segundo os critérios técnicos do National Bureau of Economic Research (NBER). Devido à acentuada desaceleração (retração) dos principais indicadores de oferta e demanda agregados, o Federal Reserve decidiu, na ausência de pressões inflacionárias, reduzir a meta para os juros básicos (*fed funds*) em 475 pontos base (p.b.) ao longo do ano. Em

dezembro, após onze cortes sistemáticos nos juros básicos, a meta dos *fed funds* foi fixada em 1,75% a.a.. Somadas à política monetária expansionista, outras medidas foram tomadas no âmbito fiscal, como a redução de impostos e o aumento de gastos federais.

O governo Republicano, em início de mandato, anunciou um programa de cortes de impostos e aumentos de gastos em estímulo à economia em cumprimento às promessas de campanha. Sob este cenário recessivo, novas propostas de gastos públicos foram enviadas ao Legislativo em decorrência dos atentados de 11 de setembro. Devido ao ataque terrorista, o mercado financeiro ficou fechado por quatro dias. Em sua reabertura, no dia 17 de setembro, os preços das ações caíram abruptamente nas principais bolsas de valores. Nas semanas posteriores ao atentado, observou-se muita volatilidade no mercado financeiro. Em setembro, o Nasdaq apresentou variação negativa de 16,9%, seguido pelo Dow Jones e pelo Standard and Poor's 500 (S&P500) que caíram 11,1% e 8,2%, respectivamente. A partir de outubro, os três índices apresentaram recuperação. Porém, o resultado anual do Dow Jones, do Nasdaq e do S&P500 foi negativo em 7,2%, 21,0% e 11,9%, respectivamente (BANCO CENTRAL DO BRASIL, 2002).

Ao final do ano, os efeitos das políticas monetária e fiscal expansionistas passaram a ser notados na economia. O Produto Interno Bruto (PIB) real dos EUA apresentou decréscimo, em taxas trimestrais anualizadas, para o primeiro (0,5%) e o terceiro (1,4%) trimestres do ano. Contudo, nos últimos três meses, a economia norte-americana surpreendeu positivamente, crescendo 1,6%. Apesar da recuperação apresentada nos últimos três meses do ano, o crescimento real do PIB do País em 2001 foi de 0,8% (BUREAU OF ECONOMIC ANALYSIS, 2008).

No Brasil, o cenário positivo para o mercado de ações entre o final de 2000 e início de 2001, relacionado com as expectativas otimistas quanto à evolução da economia brasileira, foi revertido a partir de fevereiro de 2001 em virtude dos fracos indicadores apresentados pela economia norte-americana. No primeiro trimestre de 2001, a variação do Índice da Bolsa de Valores do Estado de São Paulo (Ibovespa) acumulou queda de 5,4%.

Alguns fatores negativos à economia, como o racionamento de energia iniciado no final do primeiro semestre, a crise econômica e política na Argentina, além do atentado aos EUA em 11 setembro, aumentaram a incerteza no mercado brasileiro. Neste contexto, o câmbio foi impactado, apresentando elevada

depreciação no segundo semestre do ano. Partindo-se de uma taxa de R\$2,16/US\$ em março, atingiu-se a maior cotação do ano em outubro R\$2,84/US\$, até alcançar a cotação do final de dezembro de R\$2,32/US\$. Em relação ao nível de preços, pode-se dizer que as decisões em 2001 quanto à taxa básica de juros da economia refletiram majoritariamente as preocupações com o câmbio e com o racionamento de energia. Desta forma, a taxa Selic, fixada em janeiro em 15,25% a.a., atingiu o patamar de 19% a.a. ao final de dezembro. Quanto ao crescimento real do PIB de 1,3% em 2001, pode-se dizer que as previsões otimistas do início do ano foram desfavorecidas pelos desdobramentos internos e externos que ocorreram ao longo dos meses.

Na Bolsa de Valores de São Paulo, refletindo primordialmente os efeitos do atentado terrorista nos EUA, o índice Bovespa atingiu 10.034 pontos em 14 de setembro. Este foi o menor nível registrado desde agosto de 1999. Neste cenário, o valor de mercado das companhias listadas na Bovespa atingiu US\$142,8 bilhões em setembro, representando uma queda de 36,7% no ano, resultado da combinação de queda nos preços dos papéis e depreciação cambial. Alinhado com o mercado financeiro internacional, o Índice Bovespa registrou variação negativa de 17,2% em setembro. Nos meses seguintes, com o baixo valor das empresas em dólares e a gradual mudança nas expectativas de mercado, os investimentos tornaram-se mais atraentes. A maior procura por papéis nacionais confirmou-se pelo aumento no volume médio negociado nos dois últimos meses do ano, aproximadamente R\$700 milhões. Contudo, mesmo com a recuperação do Ibovespa ao final de 2001, o índice terminou o ano com variação negativa de 11,5%, no patamar de 13.509 pontos (BANCO CENTRAL DO BRASIL, 2002).

#### **4.1.2 A Crise de 2002**

No ano de 2002 o desempenho da economia mundial esteve condicionado à evolução da economia norte-americana. A recuperação da economia dos Estados Unidos revelou-se menos acentuada do que sugeriam os resultados obtidos no início do ano. O crescimento trimestral anualizado do PIB de 2,7% em termos reais, obtido no início do ano não foi mantido ao final do período

subseqüente. No segundo semestre, a diminuição do ritmo de expansão da economia foi sinalizada pela taxa de crescimento do produto (anualizada) que, após crescer 2,4% no terceiro trimestre, expandiu-se apenas 0,2% nos três últimos meses do ano. O crescimento real do PIB norte-americano em 2002 foi de 1,6% (BUREAU OF ECONOMIC ANALYSIS, 2008).

Um fator relevante na análise de 2002 diz respeito às denúncias de fraudes contábeis em grandes corporações norte-americanas e européias (atreladas majoritariamente ao setor de alta tecnologia), que conduziram o mercado financeiro a um clima de elevada incerteza. Neste cenário, a trajetória descendente das principais bolsas de valores internacionais foi acentuada. As mesmas apresentaram elevadas quedas entre abril e setembro, com as bolsas norte-americanas registrando perdas de 27% no Dow Jones Industrial Average (DJIA) e 36,5% no National Association of Securities Dealers Automated Quotations (Nasdaq). No Brasil, seguindo a tendência mundial, o Ibovespa registrou queda de 34,9% (BANCO CENTRAL DO BRASIL, 2003). Sob este contexto, o movimento de queda prosseguiu na primeira quinzena de outubro nos Estados Unidos, onde os principais índices negociados (Standard and Poor's 500 (S&P500), DJIA e o Nasdaq) registraram os níveis mais baixos desde o final da década de 1990. No ano, o DJIA registrou variação negativa de 17,7%, o Nasdaq decresceu 32,8% e o S&P500 registrou queda de 22,1% (INTERNATIONAL MONETARY FUND, 2003).

Adicionalmente às más expectativas e ao aumento da aversão ao risco no mercado de capitais, pode-se citar a perspectiva de guerra entre os Estados Unidos e o Iraque como fator prejudicial às decisões de investimento e produção no segundo semestre do ano.

No Brasil, a taxa de câmbio mostrou-se relativamente estável nos quatro primeiros meses de 2002 (em torno de R\$2,35/US\$). A partir do final do mês de abril, novos condicionantes afetaram a economia nacional além da crise no mercado acionário norte-americano. Como fatores adicionais à instabilidade já presente, pode-se mencionar incertezas quanto ao processo eleitoral brasileiro, o aprofundamento da crise argentina e o rebaixamento da classificação de risco país do Brasil. Nestas condições, a taxa de câmbio ultrapassou a cotação de R\$3/US\$ ao final de julho, quando o risco Brasil, medido pelo Emerging Market Bond Index (Embi-Brasil) atingiu 2390 pontos, comparativamente a 857 pontos ao final de abril. Além dos últimos acontecimentos, no segundo semestre as taxas de inflação foram

afetadas pelo comportamento do câmbio. Assim, foi necessário que o Banco Central efetuasse três elevações na meta para a taxa Selic a partir de outubro, terminando o ano com a definição de 25% a.a. (BANCO CENTRAL DO BRASIL, 2003).

Com a conclusão do processo eleitoral e a mudança de expectativas dos agentes quanto ao compromisso do novo governo com a manutenção da austeridade fiscal e do regime de metas para a inflação, o câmbio, que havia se aproximado de R\$4/US\$, iniciou uma trajetória declinante encerrando o ano em R\$3,53/US\$. O crescimento real do PIB brasileiro foi de 2,7% em 2002.

Na Bolsa de Valores de São Paulo, o Ibovespa atingiu o maior valor do ano no mês de fevereiro (14.033 pontos) finalizando o movimento de alta com início em outubro de 2001. Nos meses subseqüentes, alguns fatores como a depreciação cambial, as perspectivas de aceleração inflacionária, além do rebaixamento da classificação dos papéis brasileiros efetuado por agências de risco, foram fundamentais para o comportamento declinante deste índice. A partir da segunda metade do mês de outubro, a definição do processo eleitoral brasileiro influenciou de forma positiva o desempenho do Ibovespa, cuja valorização atingiu 30,7% no último trimestre de 2002. Ao final do ano, o valor nominal de mercado das companhias listadas na Bovespa manteve-se relativamente estável (em R\$438,3 bilhões) com expansão de 1,9%. Entretanto, o índice Bovespa registrou variação negativa de 16,6% no ano, tendo atingido em dezembro 11.268 pontos (BANCO CENTRAL DO BRASIL, 2003).

#### **4.1.3 A Crise de 2007**

O ano de 2007 foi marcado pela crise do setor de crédito hipotecário *subprime* (crédito de alto risco, ligado a clientes com histórico de inadimplência) nos Estados Unidos. O aumento do índice de inadimplência no mercado de crédito *subprime*, observado na última semana de julho, repercutiu em alguns fundos de aplicação que fornecem recursos a bancos operadores neste mercado, provocando restrições à liquidez bancária. Embora a intervenção do Federal Reserve (seguida por bancos centrais europeus) tenha sido imediata, através do aporte de recursos para o sistema, os principais fundos de aplicação financeira sofreram perdas. As

perdas se concretizaram especialmente devido à queda nos índices das principais bolsas de valores. No mês de julho, o Dow Jones, o Nasdaq e o S&P500 recuaram 1,5%, 2,2% e 3,2%, respectivamente.

O mercado financeiro internacional continuou com elevada volatilidade nas primeiras semanas de agosto, refletindo as incertezas geradas pela crise no mercado imobiliário norte-americano. A inadimplência nos financiamentos *subprime* exerceu um “efeito cascata” sobre o mercado de crédito mundial, levando a graves problemas de liquidez e aumento do custo do crédito. Devido à alta volatilidade e ao aumento substancial da percepção de risco que se seguiu em agosto, as autoridades monetárias dos Estados Unidos, Canadá e Europa anunciaram a criação de linhas de redesconto de curtíssimo prazo para gerenciar a liquidez bancária e evitar o risco sistêmico. Esta medida foi fundamental para a gradual normalização dos mercados, com efeitos sobre as taxas de juros dos títulos norte-americanos, bem como sobre os índices das bolsas de valores. Sob estas condições, o Dow Jones fechou agosto em alta de 1,1%, o Nasdaq subiu 2,0% e o S&P500 apresentou variação positiva de 1,3% (BANCO CENTRAL DO BRASIL, 2007). Nos meses posteriores, o mercado financeiro nos EUA mostrou-se bastante instável, com altas em setembro e outubro e baixas em novembro e dezembro.

Pode-se dizer que os efeitos negativos da crise no mercado financeiro afetaram a economia real dos Estados Unidos somente no último trimestre do ano. Partindo-se de uma taxa de crescimento real do PIB (taxas anualizadas) de 0,6% no primeiro trimestre de 2007, observou-se no segundo e terceiro trimestres taxas de crescimento de 3,8% e 4,9%, respectivamente. No entanto, o produto gerado nos três últimos meses do ano expandiu-se apenas 0,6% (BUREAU OF ECONOMIC ANALYSIS, 2008). Ao longo de 2007, com o intuito de aumentar a liquidez na economia, o Comitê de Mercado Aberto do Federal Reserve (FOMC) cortou a taxa básica de juros em 100 pontos base, fixando-a em 4,25% a.a. em dezembro.

Embora as principais bolsas mundiais tenham sido afetadas a partir do mês de julho, os efeitos na economia real ficaram restritos ao fim de 2007, especialmente nos Estados Unidos e em menor magnitude na Europa Ocidental. As economias emergentes e em desenvolvimento, por sua vez, foram pouco atingidas no ano pelos desdobramentos do mercado financeiro, sendo que o grupo, liderado por China e Índia, apresentou elevado crescimento do produto em 2007. Para exemplificar, as taxas de crescimento anual de EUA, Europa (Zona do Euro) e Japão

foram de 2,2%, 2,6% e 2,1%, respectivamente. Por outro lado, nos países emergentes como Brasil, Rússia, Índia e China (grupo conhecido como BRIC), as mesmas foram de 5,4%, 8,1%, 9,2% e 11,4%, respectivamente (INTERNATIONAL MONETARY FUND, 2008).

Para o Brasil, o ano de 2007 mostrou-se extremamente favorável, mostrando que o País encontra-se mais resistente a turbulências internacionais em comparação ao passado. O crescimento do produto nacional, de 5,4%, foi robusto e representou a dinâmica da demanda interna ao longo do ano. Observou-se que o consumo foi o componente com maior contribuição para a expansão do PIB em 2007. Neste resultado, a maior estabilidade do emprego, com um maior número de trabalhadores formais, impulsionou o consumo de bens a partir da obtenção de crédito. Destaca-se que o aumento do acesso ao crédito foi fundamental para a participação do consumo no PIB. Além disso, o crescimento da demanda interna também refletiu o aumento dos investimentos público e privado (sustentado por expectativas positivas) a despeito da crise no mercado financeiro internacional.

Ao longo do ano, o Comitê de Política Monetária (Copom) cortou a taxa básica de juros (Selic) em 175 pontos base. Após a meta ser fixada em 11,25% a.a. na reunião de setembro, a taxa não foi mais reduzida até dezembro. A taxa de câmbio, por sua vez, obteve forte apreciação em 2007, com o Real valorizando-se 17% frente ao Dólar e fechando o ano em R\$1,77/US\$. Ao mesmo tempo, devido aos bons resultados acumulados ao longo do ano no balanço de pagamentos, o Banco Central acumulou um montante de US\$ 180 bilhões em reservas internacionais ao final de dezembro (BANCO CENTRAL DO BRASIL, 2008).

Em relação ao mercado financeiro, salienta-se que as cotações dos títulos de países emergentes sofreram redução ao final de julho face à crise hipotecária nos EUA. Nessas condições, o Emerging Markets Bond Index Plus (Embi+)<sup>5</sup> do Brasil passou de 160 pontos no final de junho para 213 pontos no final de julho. Sob este contexto, o câmbio não repercutiu a instabilidade financeira internacional em julho e a taxa de câmbio apresentou apreciação de 2,52%, situando-se em R\$1,87/US\$ ao final do mês. Na Bovespa, o volume médio diário negociado situou-se em R\$4,8 bilhões, valor 8,5% inferior ao de junho. Alinhado com o cenário internacional, o Ibovespa registrou recuo mensal de 0,4%, atingindo 54.182 pontos.

---

<sup>5</sup> No (Embi+) quanto maior a pontuação menor a cotação dos títulos.



O mês de agosto refletiu primordialmente o recrudescimento das incertezas quanto à crise imobiliária norte-americana. Porém, após a ação dos principais bancos centrais, o mercado reagiu de forma positiva: os títulos soberanos brasileiros apresentaram valorização, com o Emerging Markets Bond Index Plus (Embi+) do País situando-se em 195 pontos ao final de agosto. Ainda, o Dólar apreciou-se frente ao Real, chegando a atingir a casa de R\$2,11, mas fechando o mês em R\$1,96. A Bovespa, por sua vez, apresentou resultado positivo em agosto, expresso no aumento de 5,1% no volume médio diário transacionado no mês. O Ibovespa, após registrar queda superior a 10% até a terceira semana do mês, atingiu 54.637 pontos no final de agosto, elevando-se 0,8%. Nos meses seguintes o Índice Bovespa apresentou um comportamento volátil. Contudo, valorizou-se 43,6% no ano, concluindo o ano de 2007 em 63.886 pontos (BANCO CENTRAL DO BRASIL, 2007, 2008).

#### 4.2 CARACTERÍSTICAS DOS DADOS

Busca-se apresentar nesta seção os dados utilizados na comparação dos modelos GARCH com os modelos aditivos quanto à extração da volatilidade. As estimações e previsões serão efetuadas sobre duas séries de retornos diários, ou seja, tomando-se a primeira diferença dos logaritmos de cada índice analisado. A primeira série dos retornos é construída a partir dos preços de fechamento diários do Índice da Bolsa de Valores do Estado de São Paulo (Ibovespa). A segunda série dos retornos, por sua vez, é obtida a partir dos preços de fechamento do Índice Standard & Poor's 500 (S&P500). O Ibovespa é o mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro, e retrata o comportamento dos principais papéis negociados na Bolsa de Valores de São Paulo. O S&P500 é um índice ponderado, calculado pela Standard & Poor's, sobre 500 ações negociadas em bolsas norte-americanas, como New York Stock Exchange (NYSE) e American Stock Exchange (AMEX). O S&P500 é considerado um dos melhores parâmetros para se avaliar o desempenho do mercado de ações dos Estados Unidos.

A amostra completa dos retornos diários do Ibovespa contém 3.212 observações, iniciando em 03/01/1995 e estendendo-se até 28/12/2007. A evolução do índice e seus retornos seguem no GRAF. 1.

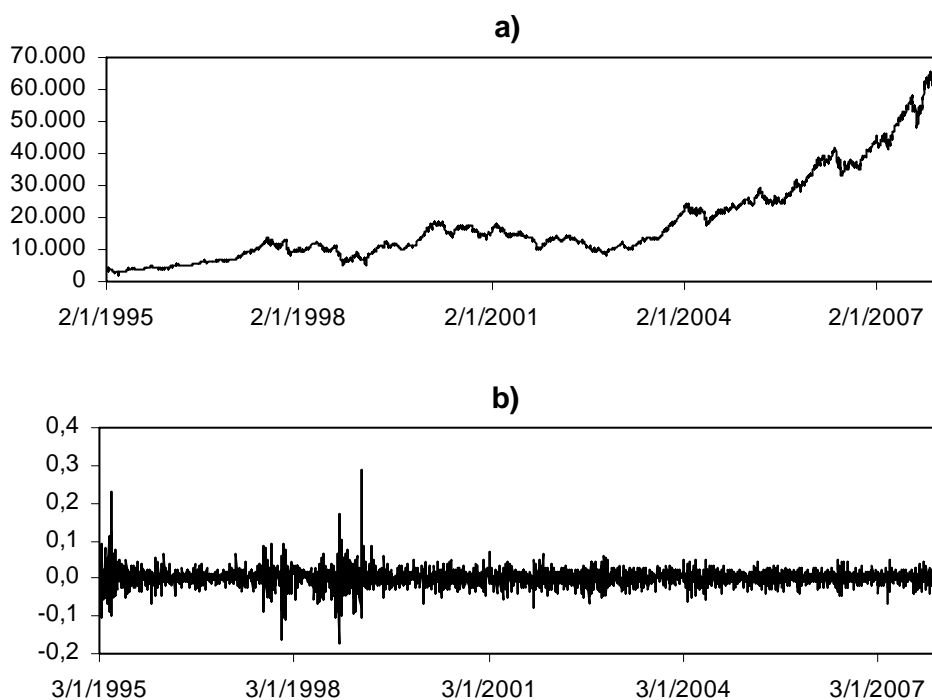


GRÁFICO 1 - Evolução do Ibovespa e seus retornos

a) Índice Bovespa

b) Retornos do Ibovespa

Fonte: Dados da Economatica.

Adicionalmente, a amostra completa dos retornos diários do S&P500 contém 3.272 observações, iniciando em 04/01/1995 e estendendo-se até 31/12/2007. Pode-se visualizar a evolução do índice e seus respectivos retornos no GRAF. 2.

Ambas as séries compreendem um período posterior à crise do México (fim de 1994) e atravessam momentos de alta volatilidade, como as crises da Ásia (1997) e da Rússia (1998). A partir de 2001, notam-se turbulências decorrentes do atentado de 11 de setembro nos Estados Unidos, e mais à frente em 2002, os meses de setembro e outubro refletem os desdobramentos da bolha das empresas de alta tecnologia. Para completar o período amostral, nos meses de julho e agosto de 2007, os índices sofrem influência da crise no mercado hipotecário *subprime* norte-americano.

Ao visualizar o item (b) dos GRAF. 1 e GRAF. 2, pode-se perceber a elevada influência das crises anteriores a de 2001 nas séries dos retornos diários. Todavia, devido aos motivos anteriormente salientados, a análise comparativa dos modelos não será realizada nestes períodos.

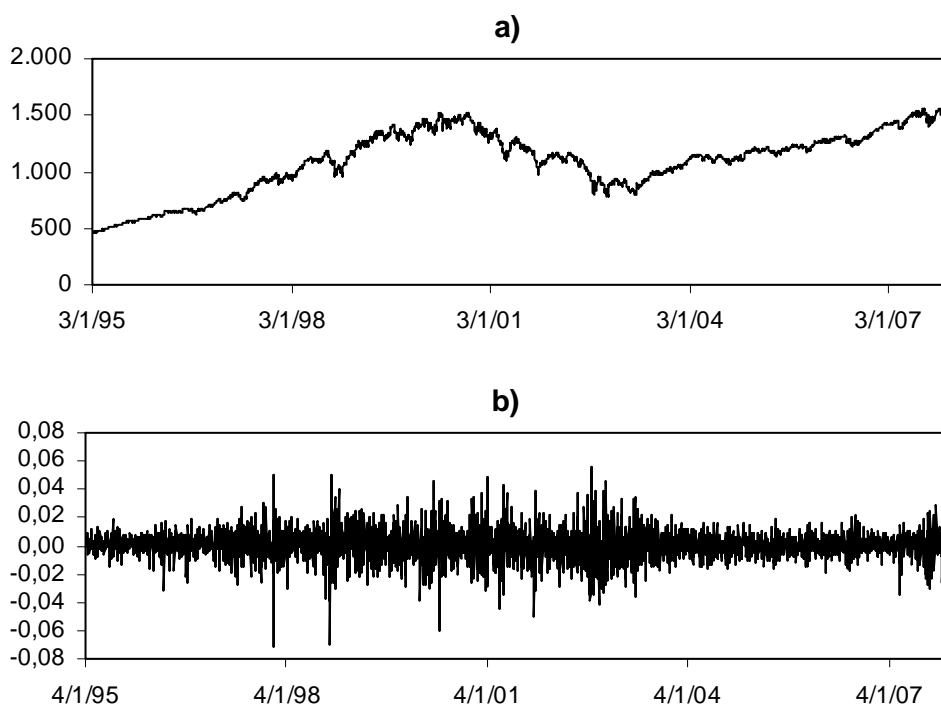


GRÁFICO 2 - Evolução do S&P500 e seus retornos  
a) Índice S&P500  
b) Retornos do S&P500

Fonte: Dados da Economatica.

Quanto às regularidades empíricas (fatos estilizados) das séries em estudo, pode-se visualizar evidências de caudas pesadas em relação à distribuição normal, nos itens (a) e (b) do GRAF. 3 e do GRAF. 4. Como discutido no capítulo 2, as caudas pesadas estão intimamente relacionadas a autocorrelação nos quadrados dos retornos, o que acaba levando a um excesso de valores elevados concentrados nas mesmas. Adicionalmente, devido à existência de autocorrelação nos quadrados dos retornos, pode-se notar agrupamentos de volatilidade nas séries dos retornos diários, tanto no GRAF. 1 como no GRAF. 2. Outra característica em séries financeiras diz respeito aos retornos propriamente ditos, na medida em que estes costumemente não são autocorrelacionados. A partir do item (d), do GRAF. 3 e do GRAF. 4, observa-se que ambas as funções de autocorrelação dos quadrados dos retornos são significativas, embora a estrutura de autocorrelação do S&P500 se

mostre mais persistente que a do Ibovespa. Por outro lado, no item (c) do GRAF. 3 e 4 aparentemente não há autocorrelação nos retornos, o que satisfaz os fatos estilizados.

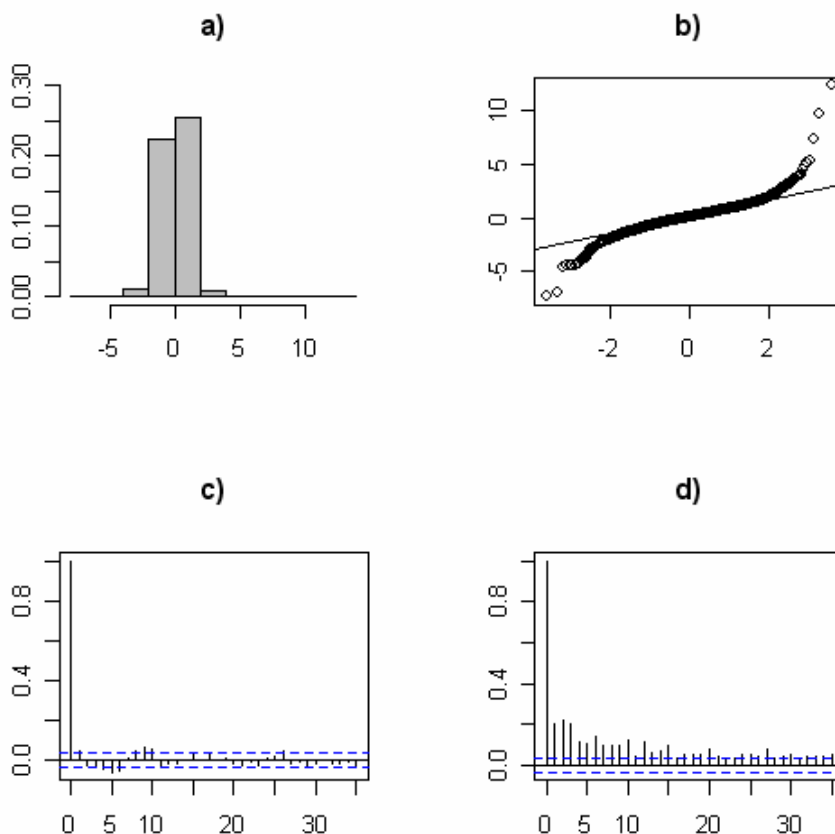


GRÁFICO 3 – Fatos estilizados dos retornos do Ibovespa

- a) Histograma dos retornos padronizados
- b) Gráfico QQ-normal dos retornos padronizados
- c) Função de autocorrelação dos retornos
- d) Função de autocorrelação dos quadrados dos retornos

Fonte: Dados da Economatica.

A escolha dos modelos aditivos bem como os da família GARCH se dará sobre as duas amostras completas dos retornos. Após a escolha das estruturas que melhor se ajustam aos dados, serão efetuadas as estimações dos parâmetros em três sub-amostras. As estimativas e previsões um passo à frente atualizadas serão calculadas sobre períodos de cem dias de negociação após a última observação de cada sub-amostra. A primeira partição dos dados do Ibovespa compreende os retornos diários de 03/01/1995 até 31/08/2001, ou seja, até o último dia do mês anterior ao ataque terrorista nos Estados Unidos. A segunda sub-amostra engloba

os dados de 03/01/1995 até 30/08/2002. Nesta situação o objetivo é mensurar o comportamento dos modelos a partir do mês de setembro de 2002, cuja volatilidade nos mercados financeiros é elevada. Por fim, a terceira partição dos dados inicia em 03/01/1995 e se estende até 29/06/2007. Este período antecede os desdobramentos da crise do mercado hipotecário *subprime* nos EUA, que se intensificam ao final de julho e durante as primeiras semanas de agosto.

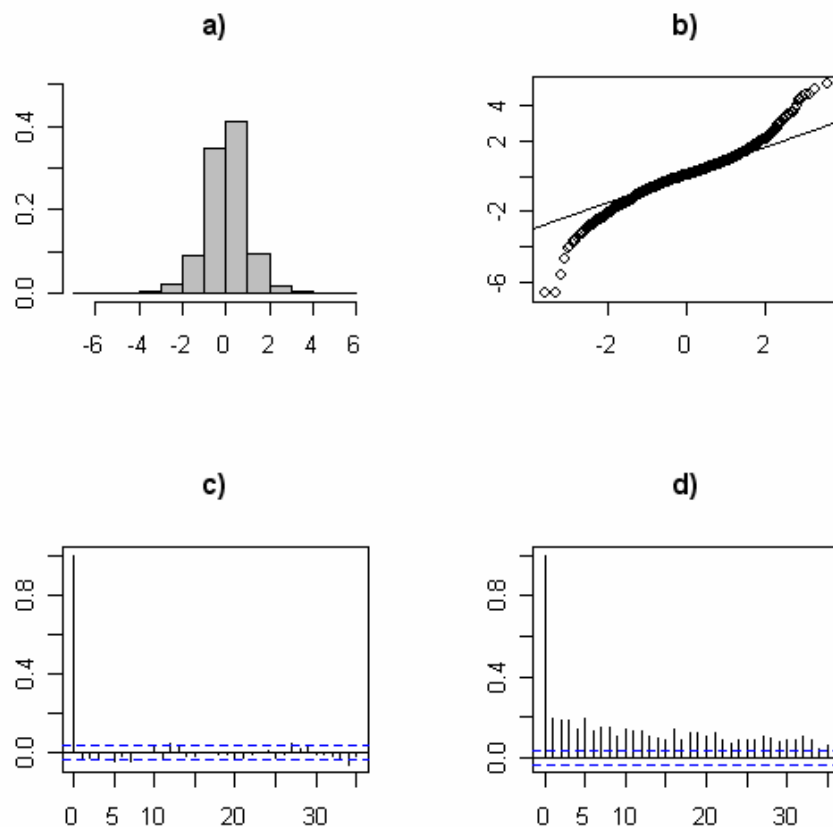


GRÁFICO 4 – Fatos estilizados dos retornos do S&P500

- a) Histograma dos retornos padronizados
- b) Gráfico QQ-normal dos retornos padronizados
- c) Função de autocorrelação dos retornos
- d) Função de autocorrelação dos quadrados dos retornos

Fonte: Dados da Economatica.

O mesmo método será utilizado para o índice S&P500, cuja primeira subamostra inicia em 04/01/1995 e termina em 31/08/2001. A segunda partição tem início em 04/01/1995 e se estende até 30/08/2002, ao passo que a última inicia novamente em 04/01/1995 e termina em 29/06/2007. Adicionalmente, cabe

apresentar no APÊNDICE A as principais estatísticas descritivas, de ambos os índices, para as três partições definidas e também para a amostra completa.

A partir da TAB. 1 do APÊNDICE A observa-se que tanto na amostra completa do Ibovespa quanto nas três partições definidas, pode-se rejeitar a hipótese nula de normalidade da série dos retornos diários através das estatísticas de Jarque e Bera (1980). Ademais, os coeficientes de curtose encontrados revelam que as distribuições não condicionais em cada período devem ser leptocúrticas.

Quanto à autocorrelação nos retornos, não claramente perceptível no item (c) do GRAF. 3, pode-se constatá-la para a amostra completa (nas trinta e seis defasagens avaliadas) através da estatística Q de Ljung e Box (1978). Além disso, as três partições restantes também apresentam autocorrelação significativa nos retornos, sendo esta observada da primeira defasagem até a trigésima sexta em duas delas. Em relação aos quadrados dos retornos, pode-se dizer que todos os períodos revelam a existência de autocorrelação através de estatísticas Q altamente significativas. Por fim, pode-se rejeitar a hipótese da presença de raiz unitária nas quatro séries do Ibovespa baseando-se nos testes Augmented Dickey-Fuller – ADF (1979) e Phillips-Perron – PP (1988). Este resultado é visualmente sugerido pelo GRAF. 1, item (b).

Em relação ao índice S&P500, a TAB. 2 do APÊNDICE A revela que a distribuição não condicional dos retornos possui caudas pesadas em relação à distribuição normal em todos os períodos, fato anteriormente notado nos itens (a) e (b) do GRAF. 4. Os retornos do S&P500 apresentam uma fraca autocorrelação, sendo perceptível sua significância na trigésima sexta defasagem de cada partição. Por outro lado, os quadrados dos retornos mostram-se autocorrelacionados da primeira até a trigésima sexta defasagem. Para completar a análise das estatísticas descritivas, pode-se dizer que todas as sub-amostras são estacionárias de acordo com os testes ADF e PP. O último resultado também é sugerido pelo GRAF. 2, item (b).

### 4.3 ESTIMAÇÃO

Nesta seção busca-se selecionar os modelos mais indicados para a amostra completa de ambos os índices. Inicialmente, efetua-se a estimação dos modelos aditivos, e posteriormente a dos GARCH paramétricos. Salienta-se uma maior descrição visual no processo de inferência dos modelos semi-paramétricos, haja vista a importância de se observar o comportamento dos componentes funcionais incluídos nos mesmos.

#### 4.3.1 Modelos Aditivos

Iniciar o processo de estimação a partir dos modelos aditivos se deve a necessidade de excluir algumas observações extremas das séries analisadas. Salienta-se que, ao efetuar intervenções nas amostras completas dos retornos dos índices, estimou-se tanto os modelos aditivos como os GARCH paramétricos com os mesmos dados. A eliminação de observações extremas faz-se necessária em decorrência da estrutura bastante flexível dos modelos aditivos que, ajustados localmente, acabam tendo suas estimativas prejudicadas por estas observações isoladas (não freqüentes).

##### 4.3.1.1 Modelos Aditivos para o Ibovespa

Ao iniciar a modelagem pela série completa do Ibovespa, deve-se ressaltar a exclusão de doze observações, dentre elas: cinco mínimas e sete máximas. A série do Ibovespa visualizada no GRAF. 1 deixa de apresentar valores exageradamente extremos a partir do ano de 1999. Os dados que foram retirados da estimação dos modelos são observados entre os anos de 1995 e 1999. Mais especificamente, são dados influenciados por desdobramentos das crises do México (fim de 1994), da Ásia (1997) e da Rússia (1998). Os valores extremos de janeiro de

1999 estão relacionados aos efeitos da desvalorização cambial no Brasil. As datas específicas e os valores das observações podem ser encontrados na TAB. 3 do APÊNDICE B.

Observando uma autocorrelação significativa nos retornos do Índice Bovespa, ajustou-se uma estrutura para a esperança condicional dos mesmos. Devido ao método definido para comparação das estimativas e previsões entre os modelos aditivos e os da família GARCH (ver seção 2.3.8, capítulo 2), optou-se por utilizar uma estrutura ARMA na equação da média condicional para ambos os modelos. Desta maneira, na medida em que a volatilidade extraída dos últimos é comparada com os quadrados dos resíduos estimados, deve-se utilizar a mesma técnica de ajuste na média para preservar a base de comparação.

Através dos procedimentos tradicionais de identificação em séries temporais, selecionou-se o modelo  $AR(2)$  como o mais adequado para filtrar a série dos retornos do Ibovespa. Esta decisão foi tomada após a comparação de algumas estruturas alternativas, sendo que a seleção final ocorreu entre um modelo  $AR(1)$  e um  $AR(2)$ , ambos sem intercepto. O segundo modelo apresentou os melhores resultados (reunidos na TAB. 4 do APÊNDICE C) tanto no critério de informação de Akaike ( $AIC$ ) quanto na redução da autocorrelação dos resíduos medida por meio das estatísticas Q de Ljung-Box. Adicionalmente, o GRAF. 5 reúne as funções de autocorrelação e autocorrelação parcial dos retornos (sem *outliers*), bem como a função de autocorrelação dos resíduos estimados do modelo  $AR(2)$  e seus quadrados.



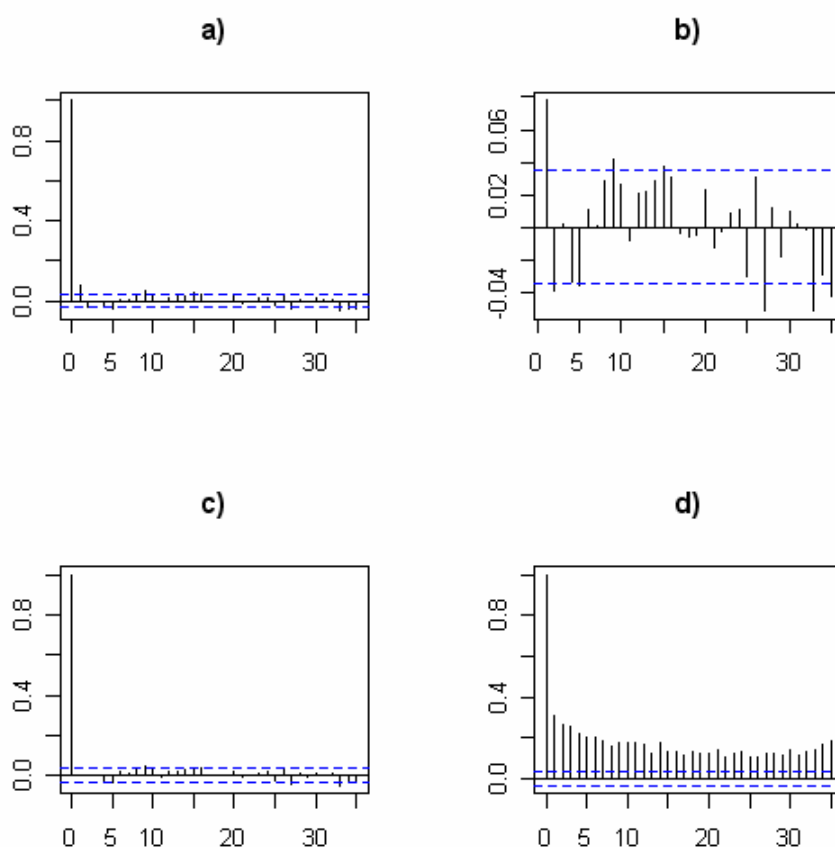


GRÁFICO 5 – Funções de autocorrelação

a) Autocorrelação dos retornos (sem *outliers*)

b) Autocorrelação parcial dos retornos (sem *outliers*)

c) Autocorrelação dos resíduos do  $AR(2)$

d) Autocorrelação dos quadrados dos resíduos do  $AR(2)$

Fonte: Dados da Economatica.

Pode-se perceber no item (c) do GRAF. 5 que a estrutura de autocorrelação dos resíduos é inferior à ilustrada no item (a) para a série dos retornos. Porém, como pode ser constatado no item (d), os quadrados dos resíduos permanecem altamente correlacionados. Sendo assim, os resíduos do modelo  $AR(2)$  serão utilizados para se obter as estimativas da volatilidade do Ibovespa, a partir dos modelos aditivos e posteriormente dos GARCH paramétricos.

Na estimação da variância condicional através dos modelos aditivos, cabe ressaltar que na ausência de estatísticas não paramétricas com distribuições teóricas desenvolvidas, faz-se necessário levar em consideração diversos resultados, sejam quantitativos ou visuais, ao longo do processo de inferência.

Embora o teste  $F$  aproximado (ver seção 3.5.1, capítulo 3) sirva de guia na escolha entre modelos alternativos, salienta-se que outras formas de inferência são possíveis devendo ser avaliadas em conjunto (HASTIE; TIBSHIRANI, 1990). Nesta dissertação, por exemplo, a obtenção de estimativas negativas para a volatilidade pode ser considerada como um indicativo da má especificação do modelo. Outra maneira importante de analisar o comportamento dos modelos dá-se a partir da visualização da contribuição individual das variáveis explicativas na construção aditiva da superfície de regressão. Neste caso, descrições visuais das funções univariadas muito distintas do que se espera teoricamente mostram-se como uma indicação da falta (ou excesso) de suavização nas mesmas, ou má especificação. Além das últimas considerações, a análise tradicional dos resíduos padronizados se mostra fundamental.

Explicada de forma geral a maneira como se conduz a seleção dos modelos aditivos nesta dissertação, parte-se primeiramente ao encontro da defasagem adequada (número de funções univariadas) para o modelo aditivo estruturado como um  $ARCH(q)$ , mas estimado não parametricamente. Em um segundo momento, após o número de funções aditivas ter sido definido, seleciona-se o parâmetro de suavização ( $span$ ) mais indicado. Deve-se ressaltar que este procedimento em duas etapas foi utilizado após a realização de diversos testes simultâneos. Nestes constatou-se que a decisão quanto à defasagem adequada não era influenciada pelas modificações no parâmetro de suavização.

Neste caso, buscou-se a defasagem apropriada no modelo aditivo com o uso do teste  $F$  aproximado, mantendo-se o  $span$  constante. O último foi mantido fixo no valor padrão do software<sup>6</sup> ( $span=0,5$ ), o que indica a utilização de 50% das observações (com pesos distintos) em cada vizinhança local em torno do ponto alvo estimado (ver seção 3.1.1, capítulo 3). Resumidamente, o teste  $F$  leva em consideração a soma dos quadrados dos resíduos (*deviance*) e os seus graus de liberdade aproximados, sendo utilizado para avaliar a contribuição não linear de cada componente não paramétrico (funções univariadas) incluído no modelo (HASTIE, 1993).

Desta forma, efetuou-se a comparação de um modelo com apenas uma função em termos dos resíduos em  $(t-1)$  (similar a um  $ARCH(1)$ ) na hipótese nula,

---

<sup>6</sup> Software livre R, Pacote gam.

contra um modelo com duas funções aditivas na hipótese alternativa: a primeira em termos dos resíduos em  $(t-1)$  e a segunda agregando em  $(t-2)$ . No próximo passo, comparou-se o modelo com duas defasagens (como um  $ARCH(2)$ ) contra o de três, e assim sucessivamente até a nona. A inclusão dos componentes funcionais mostrou-se significativa até a nona defasagem. Todavia, a partir da quinta, os modelos passaram a não se comportar adequadamente, levando a escolha de estruturas com um menor número de funções.

Os modelos com mais de cinco defasagens não serão apresentados, pois não foram aprovados em dois dos critérios de inferência utilizados, quais sejam: a apresentação de funções univariadas com apelo teórico, e a não obtenção de estimativas negativas para a variância condicional (volatilidade). A TAB. 5 contempla os resultados obtidos a partir do teste  $F$  bem como os critérios de informação aproximados de Akaike ( $AIC$ ) de cada modelo.

TABELA 5  
Seleção da defasagem nos modelos aditivos para o Ibovespa

	$SQR$	$df^{res}$	Teste $F$	$AIC$
Modelo aditivo (1)	0,0023673	3.186,80		-35.992,57
Modelo aditivo (2)	0,0022609	3.181,59	28,75 (0,000)	-36.129,03
Modelo aditivo (3)	0,0021702	3.176,39	25,49 (0,000)	-36.249,36
Modelo aditivo (4)	0,0021350	3.171,18	10,04 (0,000)	-36.291,14
Modelo aditivo (5)	0,0021103	3.165,97	7,12 (0,000)	-36.317,89

Fonte: Dados da Economatica.

Nota (1): Nos modelos aditivos o número entre parênteses indica a ordem de defasagens.

Nota (2): No teste  $F$  os números em parênteses indicam os p-valores.

A partir da TAB. 5 percebe-se que a inclusão sucessiva de componentes funcionais se mostra altamente significativa até a quinta defasagem. Os critérios de informação ( $AIC$ ) reforçam os resultados das estatísticas do teste  $F$  que indicam o modelo de ordem cinco como o mais adequado.

Como salientou-se anteriormente, as decisões em relação aos modelos não serão tomadas baseando-se somente nos testes apresentados na TAB. 5. Neste sentido, dois critérios serão avaliados para a escolha da defasagem adequada no modelo aditivo. O primeiro será em relação ao comportamento dos três últimos modelos quanto à descrição visual das funções univariadas (aplicadas nos resíduos defasados) na explicação aditiva da variável resposta (volatilidade). O segundo irá

levar em conta a capacidade (dos três últimos modelos) de redução da autocorrelação nos quadrados dos resíduos padronizados. Destaca-se que os modelos com menos de três funções univariadas reduzem muito pouco a autocorrelação nos quadrados dos resíduos padronizados, desta forma, são excluídos das análises que seguem.

Iniciando com as descrições visuais, o GRAF. 6 apresenta as funções componentes do modelo aditivo com três defasagens (como um  $ARCH(3)$ ). Pode-se visualizar abaixo da estimativa de cada função unidimensional (que leva em consideração as demais dimensões) um *rug plot* que indica a região onde ocorre a maior concentração de observações.

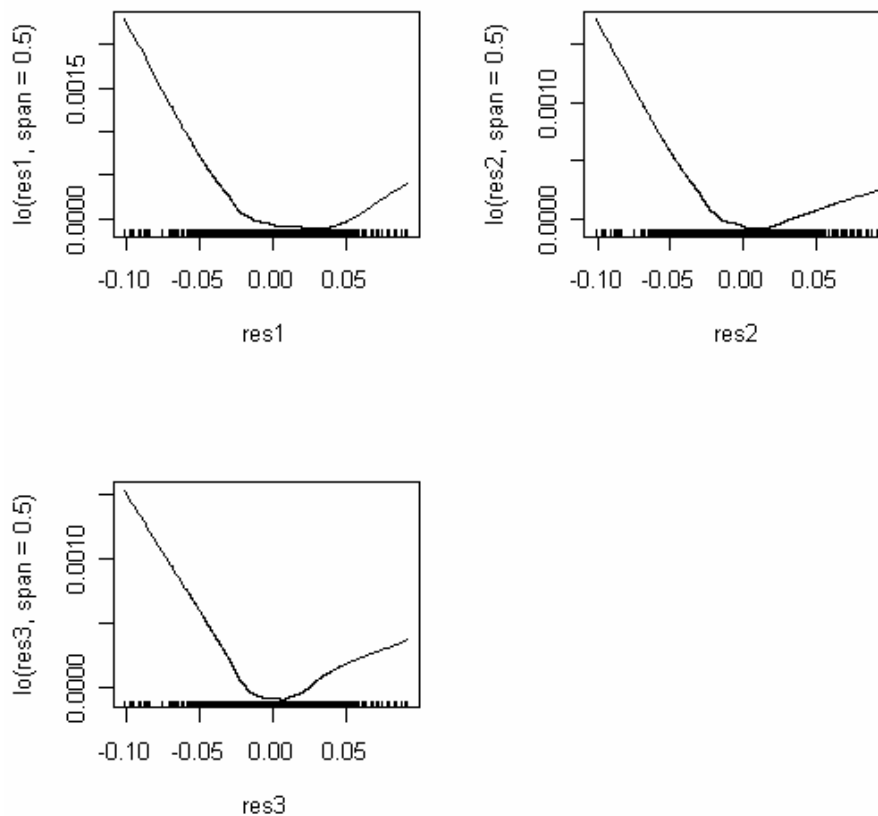


GRÁFICO 6 – Funções univariadas do Modelo aditivo (3)

Fonte: Dados da Economatica.

Nota (1): Os termos res1, res2 e res3 indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo lo indica o uso do suavizador loess na função.

A partir do GRAF. 6 pode-se fazer alguma relação com a curva proposta por Engle e Ng (1993) (*News Impact Curve*) que mede a resposta da volatilidade frente à ocorrência de choques positivos (boas notícias) ou negativos (más notícias)

no mercado. Neste sentido, as funções nas duas primeiras defasagens indicam um deslocamento das curvas em relação a zero, sendo centradas (ponto de mínimo) em um valor positivo dos retornos (neste caso, retornos corrigidos pela média). Dentre os modelos paramétricos pode-se citar um exemplo de parametrização que capta este tipo de assimetria, o modelo *AARCH*(1,1) de Engle (1990), representado na forma  $\sigma_t^2 = \omega + \alpha(\varepsilon_{t-1} + \gamma)^2 + \beta\sigma_{t-1}^2$ .

Contudo, as três funções preditoras indicam uma assimetria em relação à ocorrência de retornos negativos ou positivos de mesma magnitude. Esta pode ser notada pela inclinação mais acentuada das funções quando ocorrem choques negativos. A última característica pode ser captada por modelos paramétricos usualmente utilizados, tais como o *EGARCH*(*p,q*) de Nelson (1991) e o *TARCH*(*p,q*) proposto independentemente por Zakoian (1994), Glosten, Jagannathan e Runkle (1993).

O modelo aditivo (3) oferece descrições visuais condizentes com a teoria dos fatos estilizados. Deve-se, portanto, avaliar o comportamento dos demais modelos frente ao mesmo critério.

Para o modelo aditivo (4), pode-se acompanhar o GRAF. 7 que relaciona as funções univariadas aos quatro resíduos defasados até (*t* - 4). Neste modelo, o comportamento dos quatro elementos funcionais não é tão adequado quanto o apresentado no GRAF. 6, mas devido às funções serem extremamente flexíveis, as assimetrias podem ser acentuadas na medida em que novos termos são incluídos.

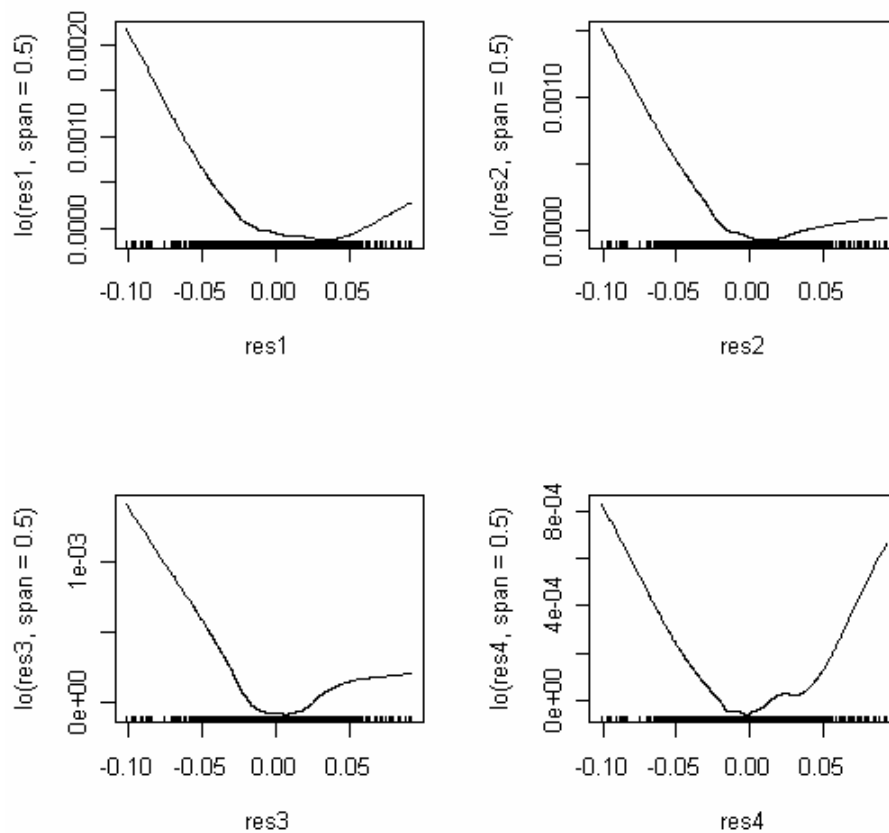


GRÁFICO 7 – Funções univariadas do Modelo aditivo (4)

Fonte: Dados da Economatica.

Nota (1): Os termos res1, res2, res3 e res4 indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo lo indica o uso do suavizador loess na função.

As duas primeiras funções univariadas mostram seus pontos de mínimo deslocados de zero e, os três primeiros termos funcionais apresentam uma resposta mais acentuada na presença de retornos negativos (más notícias). A segunda função descreve uma resposta muito suave na presença de retornos positivos, sendo um comportamento similar ao apresentado por modelos *EGARCH* paramétricos em curvas de impacto (PATTERSON, 2000).

No GRAF. 8 surge a primeira restrição quanto a escolha do modelo aditivo (5). Pode-se notar na segunda e terceira defasagens um comportamento pouco provável para a volatilidade frente aos retornos (corrigidos pela média) positivos. Não é razoável supor que na ocorrência de retornos positivos (boas notícias) a resposta na volatilidade seja constante ou até mesmo negativa.

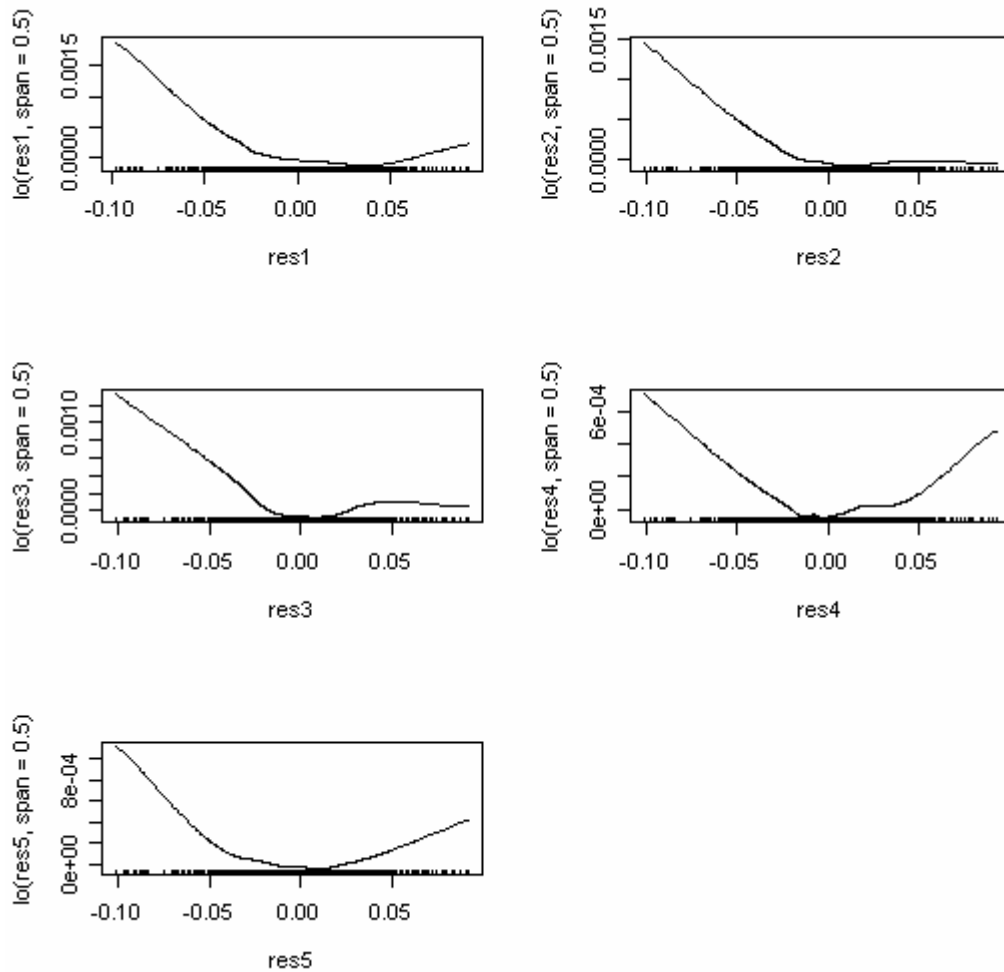


GRÁFICO 8 – Funções univariadas do Modelo aditivo (5)

Fonte: Dados da Economatica.

Nota (1): Termos res1, res2, res3, res4 e res5 indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo lo indica o uso do suavizador loess na função.

Para completar a comparação dos três modelos, o GRAF. 9 destaca as funções de autocorrelação dos quadrados dos resíduos padronizados para as três estruturas com distintas defasagens. Pode-se perceber no item (a) uma autocorrelação nos quadrados mais significativa do que a apresentada pelos modelos concorrentes nos itens (b) e (c).

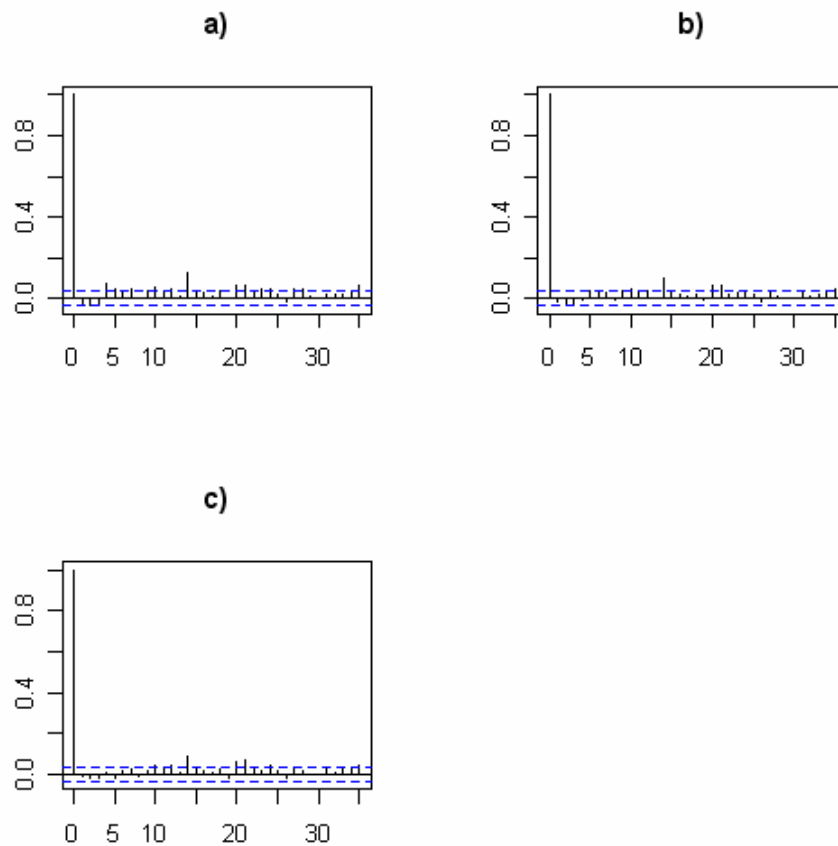


GRÁFICO 9 – Funções de autocorrelação dos quadrados dos resíduos padronizados

- a) Modelo aditivo (3)
- b) Modelo aditivo (4)
- c) Modelo aditivo (5)

Fonte: Dados da Economatica.

Analisando-se de maneira geral os resultados obtidos até o momento, pode-se dizer que o modelo aditivo (3) apresentou adequadas descrições visuais da relação funcional dos resíduos (retornos corrigidos pela média) com a variância condicional.

Contudo, o teste  $F$  aproximado e o  $AIC$  indicam que a inclusão de mais componentes funcionais é significativa. Outro critério que leva a exclusão do modelo aditivo (3) como o mais apto está vinculado a manutenção de uma estrutura de autocorrelação nos quadrados superior aos demais modelos.

O modelo aditivo (5), por sua vez, seria o escolhido levando-se em consideração os resultados da TAB. 5, mas o comportamento visual dos seus componentes aditivos não se mostrou plausível em determinadas defasagens. Além do último critério, salienta-se que a autocorrelação nos quadrados dos resíduos padronizados do GRAF. 9, item (c), é muito similar à apresentada no item (b). Neste



caso, elegeu-se o modelo aditivo com quatro funções univariadas (modelo aditivo (4)) como o mais adequado para a série do Ibovespa, na medida em que o mesmo apresentou resultados satisfatórios em todos os critérios de escolha.

A próxima etapa do processo de estimação diz respeito à escolha do *span* no modelo aditivo (4). Para isto, as análises foram efetuadas com o parâmetro de suavização variando de 0,3 até 0,8, com incrementos de 0,1.

A escolha do *span* no suavizador loess é extremamente intuitiva, podendo ser efetuada através dos mesmos critérios utilizados até o momento, dispensando o uso de métodos custosos computacionalmente (HASTIE, 1993). Destaca-se que, a complexidade dos modelos aumenta na medida em que o *span* é reduzido de 0,8 em direção a 0,3. Neste sentido, efetuou-se um teste *F* aproximado, partindo-se de um modelo aditivo (4) com *span*=0,8 sob a hipótese nula, comparando o mesmo com um modelo aditivo (4) com *span*=0,7 sob a hipótese alternativa. Este processo foi repetido até se atingir o *span*=0,3. Definiu-se esta amplitude para o *span* com a intenção de não suavizar demasiadamente as funções (aumentando o viés do estimador), tampouco torná-las exageradamente erráticas (aumentando a variância do estimador). Foram definidos incrementos de 0,1 para o *span*, em cada teste, devido aos resultados terem se mostrado muito similares para incrementos inferiores a 0,1. A TAB. 6 reúne as estatísticas do teste *F* e os valores dos critérios de informação (*AIC*) dos modelos.

TABELA 6  
Seleção do *span* no Modelo aditivo (4)

	<i>SQR</i>	$df^{res}$	Teste <i>F</i>	<i>AIC</i>
<i>span</i> =0,8	0,0021559	3.180,18		-36.288,41
<i>span</i> =0,7	0,0021513	3.178,18	3,36 (0,034)	-36.298,15
<i>span</i> =0,6	0,0021409	3.175,12	5,07 (0,002)	-36.300,59
<i>span</i> =0,5	0,0021373	3.172,17	1,81 (0,144)	-36.300,06
<i>span</i> =0,4	0,0021332	3.168,13	1,50 (0,198)	-36.298,09
<i>span</i> =0,3	0,0021251	3.161,13	1,73 (0,098)	-36.296,30

Fonte: Dados da Economatica.

Nota (1): No teste *F* os números em parênteses indicam os p-valores.

Em face dos resultados da TAB. 6, nota-se que o parâmetro de suavização deve ser definido entre 0,7 e 0,6. Na escolha definitiva do *span*, levou-se em consideração a indicação do teste *F* aproximado, o comportamento conjunto das funções univariadas e a capacidade de redução da autocorrelação nos

quadrados dos resíduos padronizados. Embora o  $AIC$  do modelo com  $span=0,7$  não tenha sido o menor entre as duas opções, decidiu-se a favor de um pouco mais de suavização nas funções, ou seja, optou-se pelo estimador com menor variância e maior viés. Sendo assim, pode-se apresentar no GRAF. 10 os termos funcionais univariados do modelo aditivo (4) com  $span=0,7$ .

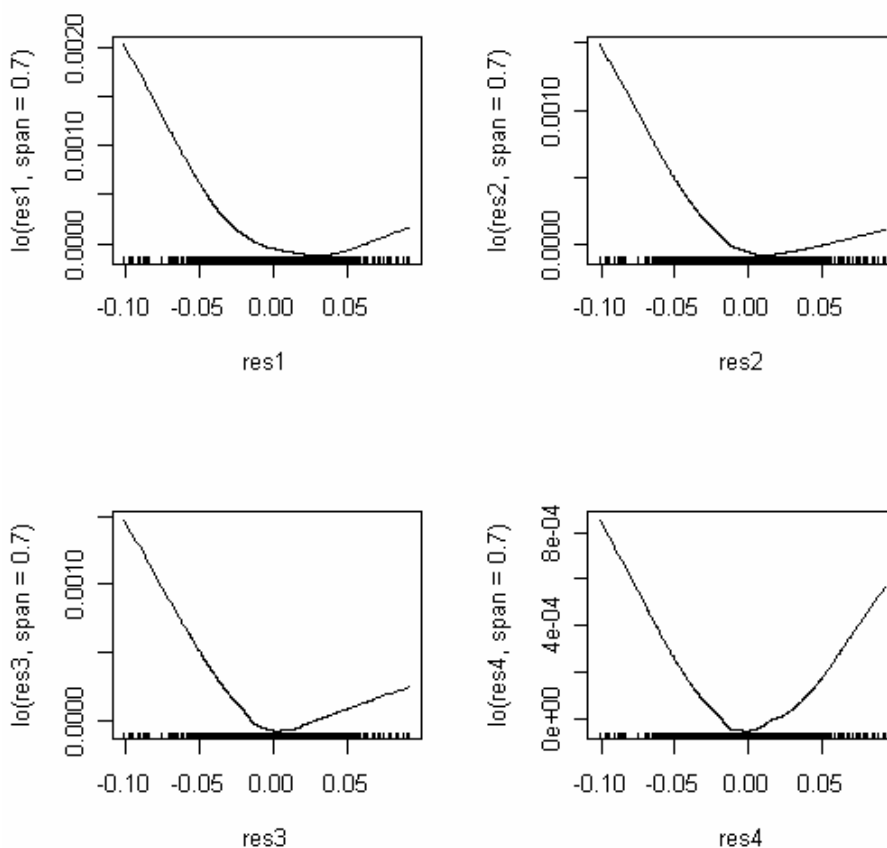


GRÁFICO 10 – Funções univariadas do Modelo aditivo (4) –  $span=0,7$

Fonte: Dados da Economatica.

Nota (1): Os termos res1, res2, res3 e res4 indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo lo indica o uso do suavizador loess na função.

Fazendo uma comparação do GRAF. 7 com o GRAF. 10 pode-se perceber que as funções do segundo são mais suaves que as do primeiro, apresentando comportamento satisfatório. Em relação às medidas de ajuste, o GRAF. 11 reúne o histograma dos resíduos padronizados, sobreposto por uma densidade não paramétrica, um gráfico quantil-quantil normal, além das funções de autocorrelação dos resíduos padronizados e seus quadrados.

Pode-se observar no item (d) do GRAF. 11 que as autocorrelações dos quadrados dos resíduos corrigidos por heterocedasticidade, pelo modelo aditivo (4),

apresentam uma substancial redução quando comparadas aos quadrados dos resíduos do  $AR(2)$ , contidos no item (d) do GRAF. 5. Isto se mostra como uma indicação de que o modelo foi capaz de capturar boa parte da variação da volatilidade.

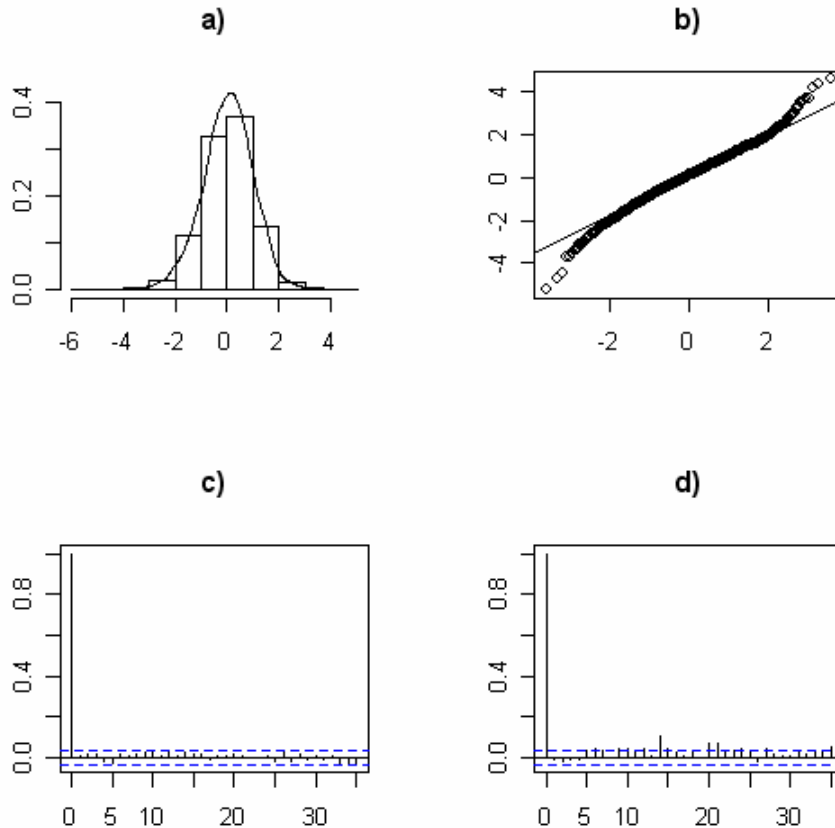


GRÁFICO 11 – Medidas de ajuste do Modelo aditivo (4) –  $span=0,7$

- a) Histograma dos resíduos padronizados
- b) Gráfico QQ-normal dos resíduos padronizados
- c) Função de autocorrelação dos resíduos padronizados
- d) Função de autocorrelação dos quadrados dos resíduos pad.

Fonte: Dados da Economatica.

Entretanto, percebe-se uma estrutura de autocorrelação remanescente nos quadrados, indicando que nem toda a não-linearidade foi captada pelo modelo. Este fato é confirmado pelo histograma contido no item (a), bem como pelo gráfico quantil-quantil normal do item (b), que mostram a manutenção de caudas pesadas.

Após a seleção do modelo aditivo com quatro termos funcionais (similar a um  $ARCH(4)$ ) e com um parâmetro de suavização que indica a inclusão de 70%

das observações da amostra dentro de cada vizinhança local, pode-se visualizar as estimativas da volatilidade contidas no GRAF. 12.

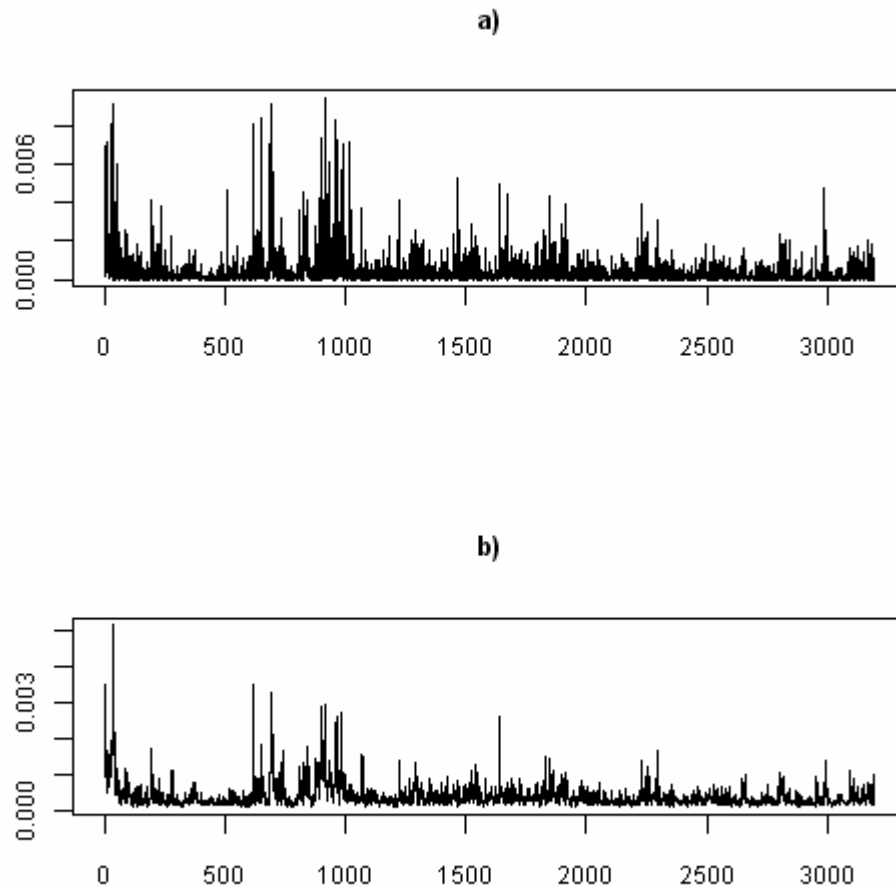


GRÁFICO 12 – Resíduos ao quadrado e volatilidade estimada do Ibovespa

a) Resíduos estimados do modelo  $AR(2)$  ao quadrado

b) Estimativas da variância condicional do Modelo aditivo (4)- $span=0,7$

Fonte: Dados da Economatica.

O GRAF. 12 reúne os resíduos do modelo  $AR(2)$  ao quadrado, item (a), que serão posteriormente utilizados como base de comparação entre as estimativas dos modelos aditivos e GARCH paramétricos nos períodos de crise. Adicionalmente, o item (b) destaca a evolução da variância condicional estimada pelo modelo aditivo (4). Embora a comparação visual das estimativas na amostra completa não seja ideal, pode-se perceber o acompanhamento das mesmas frente à variabilidade do estimador adotado para a variância condicional verdadeira.

#### 4.3.1.2 Modelos Aditivos para o S&P500

Utilizando o mesmo método de inferência da série do Ibovespa, partiu-se para a análise da série completa do Standard and Poor's 500 (S&P500). Cabe destacar que no processo de inferência do S&P500, foi novamente necessário intervir na série dos retornos. Desta vez menos observações extremas foram excluídas, sendo oito no total. Foram extraídos da amostra completa dos retornos quatro mínimos e quatro máximos, sendo a intervenção fundamental na adequação dos modelos aditivos. Esses dados são referentes a períodos de crise, tais como a da Ásia (1997), da Rússia (1998) e das empresas “ponto.com” (2000). Ademais, o retorno registrado na reabertura do mercado norte-americano em 17 de setembro de 2001 (pós-ataque terrorista) foi descartado, assim como dois retornos referentes à crise de 2002 (desdobramentos da bolha tecnológica). Os valores das observações e suas respectivas datas de ocorrência estão contidos na TAB. 7 do APÊNDICE D.

Analisando os resultados contidos na TAB. 7 do APÊNDICE D quanto à autocorrelação nos retornos (sem os *outliers*), pode-se notar sua presença por meio das estatísticas Q de Ljung-Box somente na trigésima sexta defasagem. Diante disto, ainda testou-se a significância estatística de alguns termos para a equação da média condicional. Porém, nenhuma evidência foi encontrada. Sendo assim, partiu-se diretamente em busca de um modelo aditivo para a estimação da volatilidade. Salienta-se que, com a esperança condicional igual a zero, as comparações das estimativas da volatilidade serão efetuadas contra a série dos retornos ao quadrado.

Iniciando a procura pela estrutura mais adequada à série, procedeu-se inicialmente a escolha da defasagem (número de funções) ideal para o modelo aditivo. Para isto, o parâmetro de suavização foi mantido constante no valor padrão, ou seja  $span=0,5$ , e utilizou-se o teste  $F$  aproximado para capturar a significância das funções incluídas no modelo. As estruturas com seis ou mais elementos funcionais geraram estimativas negativas para a variância condicional, mesmo após a exclusão dos *outliers*, ficando assim de fora da análise.

O teste  $F$  foi realizado inicialmente, comparando-se um modelo com uma função univariada dos retornos em  $(t-1)$  na hipótese nula (de maneira similar a um

$ARCH(1)$ ), contra um modelo contendo dois elementos funcionais, o primeiro em  $(t-1)$  e o segundo em  $(t-2)$  (como um  $ARCH(2)$ ). Testou-se a significância, a cada inclusão de um novo componente funcional, até atingir um modelo com cinco funções aditivas (modelo aditivo (5)). Os resultados do teste estão reunidos na TAB. 8.

TABELA 8  
Seleção da defasagem nos modelos aditivos para o S&P500

	$SQR$	$df^{res}$	Teste $F$	$AIC$
Modelo aditivo (1)	0,0001381	3.252,84		-46.062,96
Modelo aditivo (2)	0,0001307	3.247,69	35,83 (0,000)	-46.232,98
Modelo aditivo (3)	0,0001284	3.242,53	11,37 (0,000)	-46.281,08
Modelo aditivo (4)	0,0001261	3.237,37	11,35 (0,000)	-46.329,18
Modelo aditivo (5)	0,0001240	3.232,21	10,94 (0,000)	-46.375,28

Fonte: Dados da Economatica.

Nota (1): Nos modelos aditivos o número entre parênteses indica a ordem de defasagens.

Nota (2): No teste  $F$  os números em parênteses indicam os p-valores.

Tomando os resultados da TAB. 8, percebe-se que o modelo aditivo (5) se mostra o mais indicado, tanto pelo teste  $F$  como pelo  $AIC$  aproximado. Para não realizar a escolha definitiva levando somente em consideração os testes acima, buscou-se averiguar o comportamento das funções preditoras dos modelos aditivos (3), (4) e (5). Além disso, a capacidade dos últimos frente à redução da autocorrelação dos quadrados dos retornos padronizados também serviu para a seleção adequada da ordem do modelo.

Iniciando com as descrições visuais, o GRAF. 13 apresenta as funções componentes do modelo aditivo com três defasagens (como um  $ARCH(3)$ ). Ressalta-se que, embora cada função seja univariada (devido à forma como são estimadas pelo algoritmo *backfitting*<sup>7</sup>), cada termo funcional sofre influência dos demais.

<sup>7</sup> Ver seção 3.5 do capítulo 3.

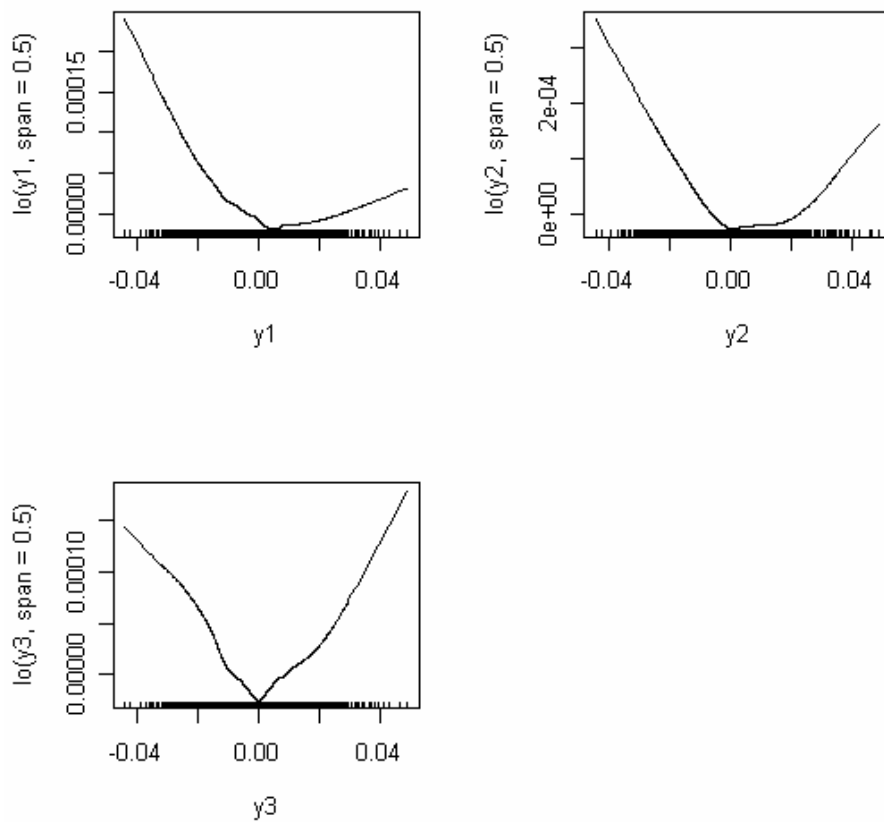


GRÁFICO 13 – Funções univariadas do Modelo aditivo (3)

Fonte: Dados da Economatica.

Nota (1): Os termos  $y_1$ ,  $y_2$  e  $y_3$  indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo lo indica o uso do suavizador loess na função.

A partir do GRAF. 13 pode-se observar que a primeira função apresenta um pequeno deslocamento em relação a zero de seu ponto de mínimo. Além disso, as duas primeiras defasagens indicam uma assimetria em resposta à ocorrência de retornos negativos ou positivos de mesma magnitude. Isto pode ser notado pela inclinação mais acentuada das funções quando ocorrem retornos negativos (más notícias no mercado). A função em  $(t-3)$  apresenta uma elevada resposta na volatilidade frente a grandes retornos (independentes do sinal).

Para o modelo aditivo (4), pode-se acompanhar o GRAF. 14 que relaciona as funções univariadas aos retornos defasados até  $(t-4)$ . Neste modelo, o comportamento das quatro funções não se modifica substancialmente em relação ao apresentado no GRAF. 13. Pode-se notar uma resposta mais suave na presença de retornos positivos na primeira defasagem, sendo que um comportamento inverso é percebido na quarta defasagem. Contudo, devido às funções serem extremamente

flexíveis, as assimetrias e não linearidades podem ser acentuadas na medida em que novos componentes são incluídos.

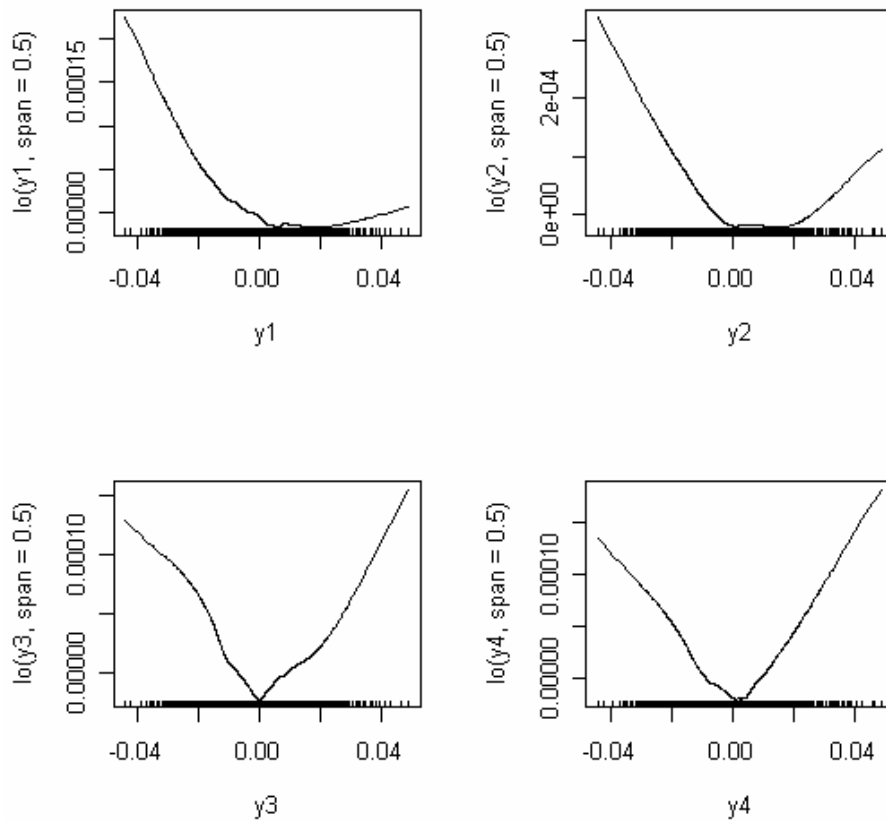


GRÁFICO 14 – Funções univariadas do Modelo aditivo (4)

Fonte: Dados da Economatica.

Nota (1): Os termos  $y_1$ ,  $y_2$ ,  $y_3$  e  $y_4$  indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo  $lo$  indica o uso do suavizador loess na função.

No GRAF. 15 surge uma restrição quanto à escolha do modelo aditivo (5). Pode-se visualizar na primeira defasagem um comportamento pouco provável para a volatilidade frente aos retornos positivos. O decaimento contínuo da função não é aceitável na presença de retornos positivos (boas notícias). Nas demais defasagens o modelo se comporta razoavelmente, apresentando no segundo elemento funcional um deslocamento do ponto de mínimo (a direita de zero), e uma resposta mais acentuada para valores positivos dos retornos nas funções em  $(t-4)$  e  $(t-5)$ .



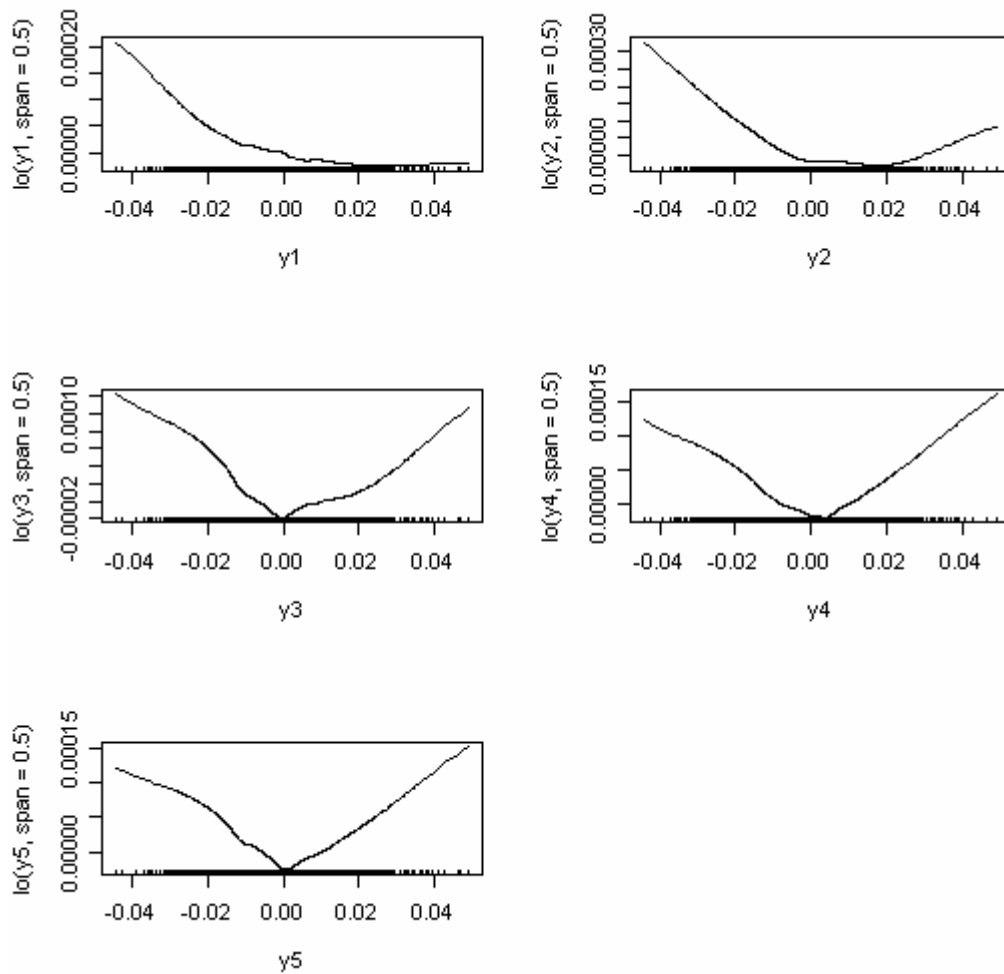


GRÁFICO 15 – Funções univariadas do Modelo aditivo (5)

Fonte: Dados da Economatica.

Nota (1): Termos  $y_1$ ,  $y_2$ ,  $y_3$ ,  $y_4$  e  $y_5$  indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo  $lo$  indica o uso do suavizador loess na função.

Levando-se adiante a comparação dos três modelos, o GRAF. 16 destaca as funções de autocorrelação dos quadrados dos retornos padronizados. Neste, pode-se perceber no item (a), referente ao modelo aditivo (3), uma autocorrelação nos quadrados mais significativa do que a apresentada pelos modelos alternativos nos itens (b) e (c).

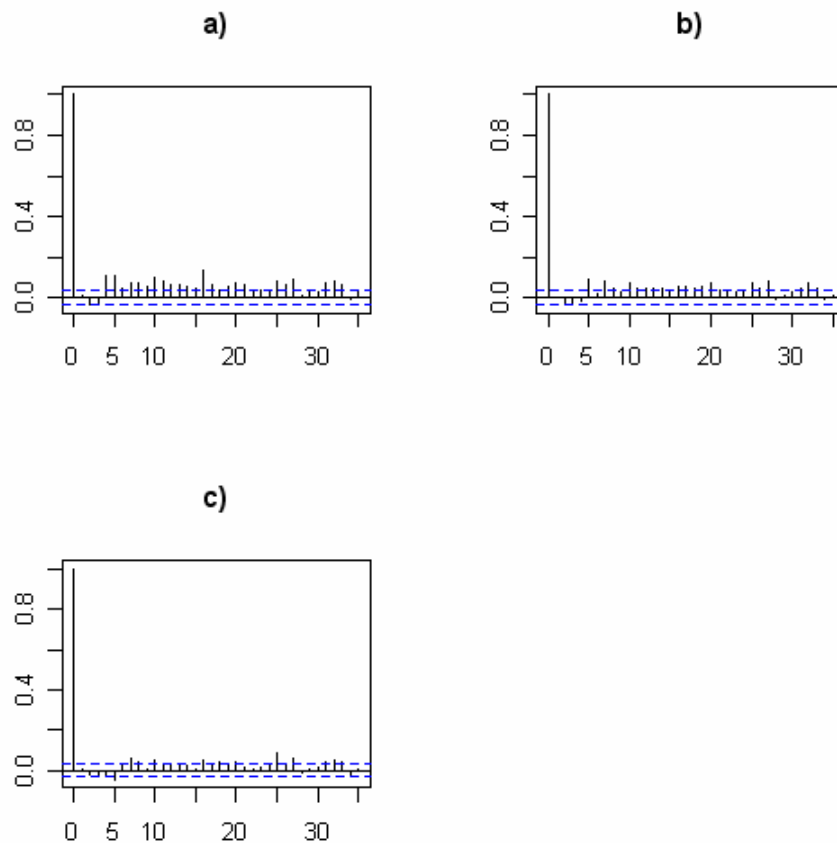


GRÁFICO 16– Funções de autocorrelação dos quadrados dos retornos padronizados

- a) Modelo aditivo (3)
- b) Modelo aditivo (4)
- c) Modelo aditivo (5)

Fonte: Dados da Economatica.

Avaliando os resultados obtidos até o momento, pode-se dizer que o modelo aditivo (3) apresentou descrições visuais satisfatórias da relação funcional dos retornos com a variância condicional. Entretanto, o teste  $F$  aproximado e o  $AIC$  indicam que a inclusão de mais elementos funcionais é significativa. Além disso, outro fator que leva a eliminação do modelo aditivo (3) está relacionado à manutenção de uma estrutura de autocorrelação, nos quadrados dos retornos padronizados, superior aos demais modelos.

Quanto ao modelo aditivo (5), cabe dizer que este seria o selecionado levando-se em consideração os resultados da TAB. 8. Contudo, o comportamento visual da função univariada na primeira defasagem não se mostrou aceitável frente aos retornos positivos.

Embora se possa perceber que a autocorrelação nos quadrados dos retornos padronizados do GRAF. 16, item (c), é ligeiramente inferior à apresentada

no item (b), a descrição visual inadequada descaracteriza o modelo aditivo (5) como o mais indicado. Frente aos fatos apresentados, elegeu-se o modelo aditivo com quatro funções univariadas (modelo aditivo (4)) como o mais apropriado para a série dos retornos do S&P 500 na medida em que o mesmo apresentou resultados satisfatórios em todos os critérios de escolha.

Como próximo passo no processo de estimação, buscou-se à escolha do *span* no modelo aditivo (4). Para isto, foram efetuados testes com o parâmetro de suavização variando de 0,3 até 0,8, com incrementos de 0,1.

Destaca-se novamente que a complexidade dos modelos aumenta na medida em que o *span* é reduzido de 0,8 em direção a 0,3. Neste caso, efetuou-se um teste *F* aproximado partindo-se de um modelo aditivo (4) com *span*=0,8 sob a hipótese nula, comparando o mesmo com um modelo aditivo (4) com *span*=0,7 sob a hipótese alternativa. Este processo foi repetido até se atingir o *span*=0,3.

Foram definidos incrementos de 0,1 para o *span*, em cada teste, devido aos resultados terem se mostrado muito similares para incrementos inferiores a 0,1. A TAB. 9 contempla as estatísticas do teste *F* e os valores dos critérios de informação aproximados (*AIC*) dos modelos.

TABELA 9  
Seleção do *span* no Modelo aditivo (4)

	<i>SQR</i>	$df^{res}$	Teste <i>F</i>	<i>AIC</i>
<i>span</i> =0,8	0,0001266	3.246,44		-46.346,75
<i>span</i> =0,7	0,0001265	3.244,30	1,61 (0,198)	-46.345,91
<i>span</i> =0,6	0,0001263	3.241,28	2,26 (0,078)	-46.346,74
<i>span</i> =0,5	0,0001261	3.238,36	1,18 (0,314)	-46.344,38
<i>span</i> =0,4	0,0001259	3.234,36	1,20 (0,305)	-46.341,25
<i>span</i> =0,3	0,0001253	3.227,84	2,57 (0,015)	-46.345,07

Fonte: Dados da Economatica.

Nota (1): No teste *F* os números em parênteses indicam os p-valores.

Em face dos resultados da TAB. 9 decidiu-se a favor de uma maior suavização no modelo, seguindo a indicação do critério de informação aproximado (*AIC*). A partir da última estatística, selecionou-se o *span*=0,8, que reflete uma menor complexidade na estrutura do modelo aditivo.

Ademais, na escolha do *span* levou-se em consideração o comportamento conjunto das funções univariadas, e a capacidade de redução da autocorrelação nos quadrados dos retornos padronizados. Desta forma, pode-se

apresentar no GRAF. 17 os termos funcionais univariados do modelo aditivo (4) com  $span=0,8$ .

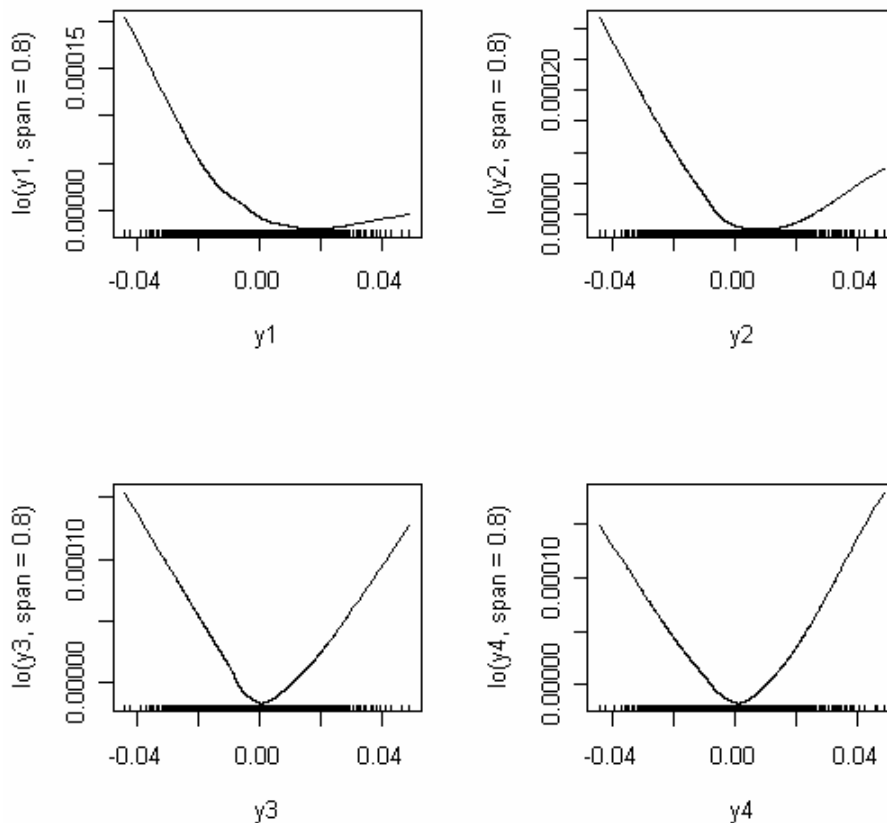


GRÁFICO 17 – Funções univariadas do Modelo aditivo (4) –  $span=0,8$

Fonte: Dados da Economatica.

Nota (1): Os termos  $y_1$ ,  $y_2$ ,  $y_3$  e  $y_4$  indicam as defasagens da variável explicativa.

Nota (2): No eixo das ordenadas o termo  $lo$  indica o uso do suavizador loess na função.

Comparando o GRAF. 14, do modelo aditivo (4) com  $span=0,5$ , com o GRAF. 17, pode-se perceber que as funções do segundo são mais suaves que as do primeiro. Isto indica a escolha de um estimador com menor variância, mas com um maior viés.

Em se tratando das medidas de ajuste, o GRAF. 18 reúne o histograma dos retornos padronizados, sobreposto por uma densidade não paramétrica, um gráfico quantil-quantil normal, além das funções de autocorrelação dos retornos padronizados e seus quadrados.

Percebe-se ao observar o item (d), do GRAF. 18, que as autocorrelações dos quadrados dos retornos corrigidos por heterocedasticidade apresentam uma redução expressiva quando comparadas aos quadrados dos retornos contidos no

item (d), do GRAF. 4. Isto indica que o modelo foi capaz de capturar boa parte da variação da volatilidade.

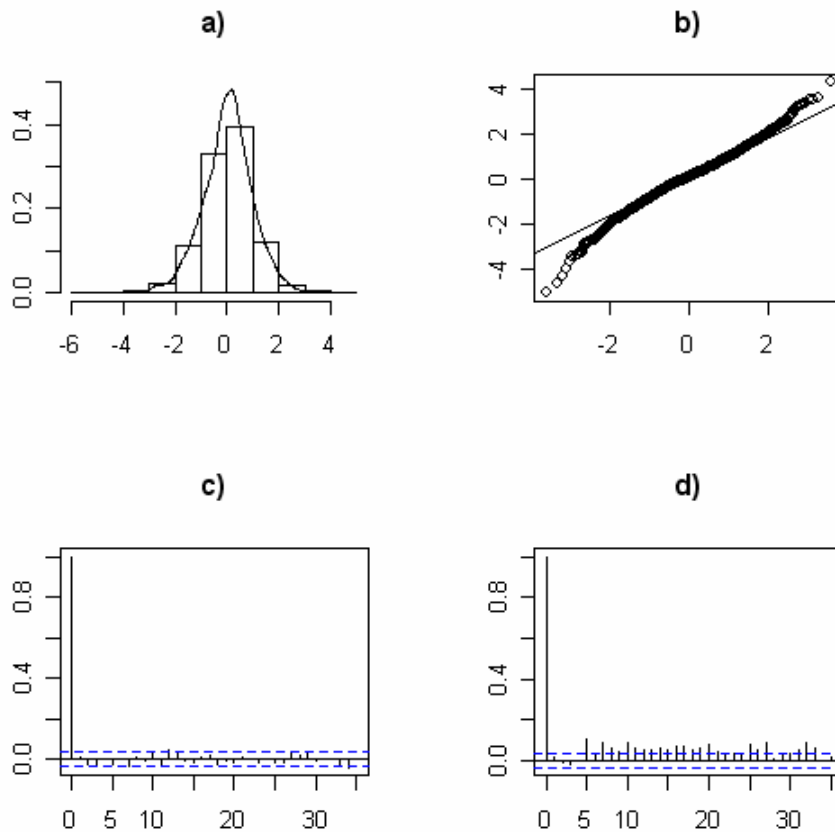


GRÁFICO 18 – Medidas de ajuste do Modelo aditivo (4) –  $span=0,8$

- a) Histograma dos retornos padronizados
- b) Gráfico QQ-normal dos retornos padronizados
- c) Função de autocorrelação dos retornos padronizados
- d) Função de autocorrelação dos quadrados dos retornos pad.

Fonte: Dados da Economatica.

Todavia, nota-se um resquício de autocorrelação nos quadrados, indicando que nem toda a não-linearidade foi captada pelo modelo. Isto é confirmado pelo histograma contido no item (a) do GRAF. 18, bem como pelo gráfico quantil-quantil normal do item (b). Estes refletem a manutenção de caudas pesadas, com a ocorrência de algumas observações extremas na cauda inferior.

A partir do modelo aditivo composto por quatro funções univariadas (similar a um  $ARCH(4)$ ) e, com um parâmetro de suavização que indica a inclusão de 80% das observações da amostra dentro de cada vizinhança local, pode-se visualizar as estimativas da volatilidade e os retornos ao quadrado no GRAF. 19.

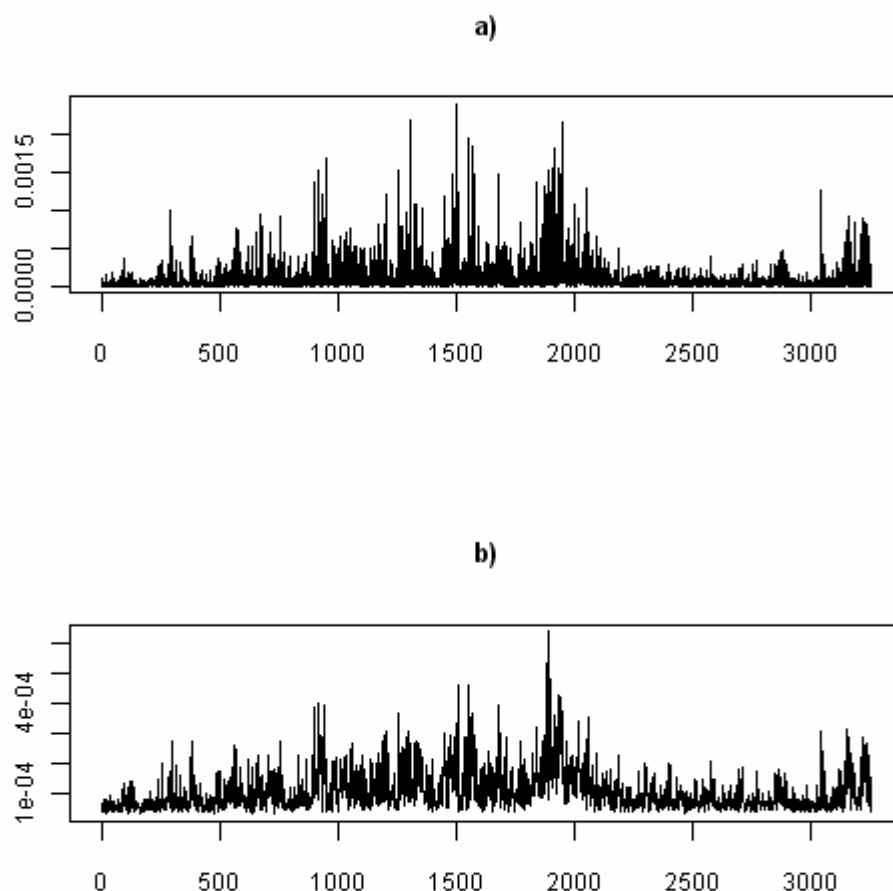


GRÁFICO 19 – Retornos ao quadrado e volatilidade estimada do S&P500

a) Retornos do S&P500 ao quadrado

b) Estimativas da variância condicional do Modelo aditivo (4)- $span=0,8$

Fonte: Dados da Economatica.

O GRAF. 19 destaca a evolução dos retornos ao quadrado do S&P500, item (a), que serão utilizados à frente como base de comparação entre as estimativas dos modelos aditivos e GARCH paramétricos. Conjuntamente, o item (b) reflete a evolução da variância condicional estimada pelo modelo aditivo (4). Pode-se notar na evolução das estimativas do modelo aditivo que as mesmas refletem de maneira geral o comportamento do estimador da volatilidade não observável.

Por fim, de posse dos modelos semi-paramétricos selecionados, busca-se na próxima seção escolher os modelos da família GARCH mais indicados para a estimação da volatilidade dos retornos dos dois índices.

### 4.3.2 Modelos GARCH

Utilizando as mesmas séries de retornos (sem *outliers*) adotadas nos modelos aditivos, procurou-se avaliar as parametrizações mais adequadas dentro da classe GARCH. Como o processo de estimação dos modelos aditivos se fez em duas etapas, ou seja, primeiro a equação da média e posteriormente a da variância condicional, deve-se lembrar que as estruturas definidas na seção anterior para a média são as mesmas na estimação dos modelos GARCH paramétricos.

#### 4.3.2.1 Modelos GARCH para o Ibovespa

Iniciando a análise pela série dos retornos do Ibovespa, buscou-se selecionar os modelos mais apropriados através dos critérios de informação de Akaike (AIC) e de Schwarz (BIC) e a partir do valor máximo do logaritmo da verossimilhança. Na estimação via função de verossimilhança condicional foram supostas duas distribuições condicionais para  $z_t = \varepsilon_t \sigma_t^{-1}$ , a normal e a t-Student. A TAB. 10 reúne os modelos comparados.

TABELA 10  
Seleção dos modelos GARCH para o Ibovespa

Modelo	Normal			t-Student		
	AIC	BIC	Max [Log L]	AIC	BIC	Max [Log L]
ARCH(4)	-5,014547	-5,001260	8.025,26	<b>-5,032479</b>	<b>-5,017294</b>	<b>8.054,94</b>
GARCH(1,1)	-5,053130	-5,043639	8.084,95	-5,066498	-5,055109	8.107,33
GARCH(3,2)	-5,052914	-5,037729	8.087,61	-5,066170	-5,049087	8.109,81
GARCH(3,3)	-5,055011	-5,037927	8.091,96	-5,067778	-5,048797	8.113,38
EGARCH(1,1)	-5,071864	-5,060475	8.115,91	-5,083033	-5,069745	8.134,77
TARCH(1,2)	-5,077988	-5,064701	8.126,70	-5,087981	-5,072796	8.143,68
TARCH(3,2)	-5,078771	-5,061688	8.129,96	<b>-5,088500</b>	-5,069518	<b>8.146,51</b>

Fonte: Dados da Economatica.

Nota (1): Dados em negrito indicam o melhor modelo.

Procurou-se apresentar na TAB 10 somente os modelos com os parâmetros da equação da variância significativos. Além disso, as estruturas cujas estimativas não convergiram foram eliminadas do processo seletivo. Devido à

intenção de comparar os modelos GARCH com os aditivos na próxima seção, optou-se por selecionar um modelo  $ARCH(4)$  e outra parametrização entre os GARCH com o melhor resultado nos critérios anteriormente definidos.

A partir dos resultados da TAB.10 foram selecionados os modelos  $ARCH(4)$  e  $TARCH(3,2)$ , ambos baseados na função de verossimilhança derivada da distribuição t-Student. Salienta-se que o modelo  $TARCH(3,2)$  segue a especificação de Glosten, Jagannathan e Runkle (1993). A TAB.11 apresenta as estimativas dos parâmetros dos modelos escolhidos.

TABELA 11  
Estimativas dos modelos GARCH para o Ibovespa

	ARCH(4)	TARCH(3,2)
Média condicional		
$\phi_1$	0,0584 (0,018371)	0,0707 (0,017856)
$\phi_2$	-0,0278 (0,017880)	-0,0199 (0,018570)
Variância condicional		
$\omega$	0,0002 (0,000013)	1,58E-05 (0,000003)
$\alpha_1$	0,1023 (0,023606)	-0,0335 (0,011272)
$\alpha_2$	0,1274 (0,028218)	0,0501 (0,015488)
$\alpha_3$	0,1951 (0,030340)	..
$\alpha_4$	0,1446 (0,030644)	..
$\beta_1$	..	1,2452 (0,215016)
$\beta_2$	..	-0,7251 (0,326537)
$\beta_3$	..	0,3579 (0,154501)
$\gamma_1$	..	0,1356 (0,022988)
Graus de liberdade	9,9532 (1,564467)	15,2440 (2,837413)

Fonte: Dados da Economatica.

Nota (1): Erros padrão entre parênteses.

Nota (2): O modelo  $AR(2)$  segue a forma,  $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$ .

Nota (3): O modelo  $ARCH(4)$  segue a forma,  $\sigma_t^2 = \omega + \sum_{i=1}^4 \alpha_i \varepsilon_{t-i}^2$ .

Nota (4): O modelo  $TARCH(3,2)$  segue a forma,  $\sigma_t^2 = \omega + \sum_{i=1}^2 \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^3 \beta_j \sigma_{t-j}^2 + \gamma_1 \varepsilon_{t-1}^2 Y_{t-1}$ , onde  $Y_t = 1$  se  $(\varepsilon_t < 0)$  e  $(0)$  caso contrário.

O modelo  $ARCH(4)$  apresentou todos os termos da equação da variância condicional significativos. Entretanto, o segundo termo autoregressivo da equação da média condicional, quando estimado conjuntamente aos demais parâmetros,



mostrou-se não significativo. Decidiu-se por sua manutenção no modelo para garantir a equidade nas comparações posteriores.

Quanto às medidas de ajuste do modelo  $ARCH(4)$ , o GRAF. 20 destaca o histograma dos resíduos padronizados, sobreposto por uma densidade não paramétrica, item (a), bem como as funções de autocorrelação dos resíduos padronizados e seus quadrados, nos itens (b) e (c), respectivamente.

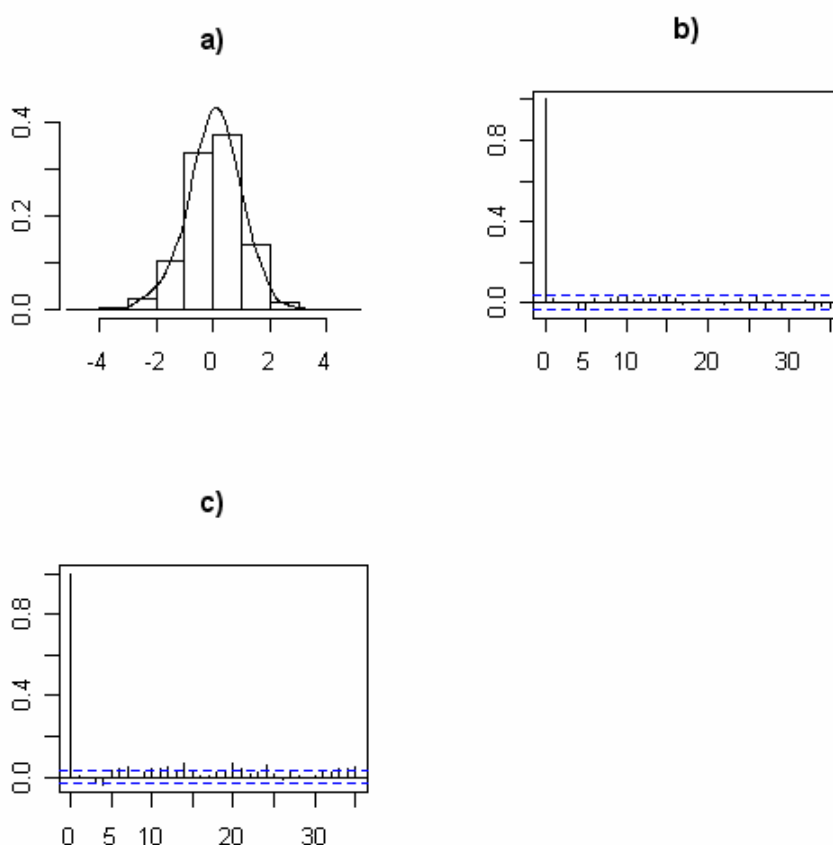


GRÁFICO 20 – Medidas de ajuste do Modelo  $ARCH(4)$

- a) Histograma dos resíduos padronizados
- b) Função de autocorrelação dos resíduos padronizados
- c) Função de autocorrelação dos quadrados dos resíduos pad.

Fonte: Dados da Economatica.

A partir do GRAF. 20, item (c), pode-se perceber a manutenção de uma estrutura de autocorrelação nos quadrados dos resíduos padronizados. Esta percepção é reforçada pela apresentação do histograma no item (a), cujas caudas permanecem pesadas.

Todavia, quando se compara o item (c) do GRAF. 20 com a função de autocorrelação dos quadrados dos resíduos do modelo  $AR(2)$ , presente no item (d) do GRAF. 5, nota-se uma substancial redução.

O modelo  $TARCH(3,2)$ , por sua vez, também apresentou todos os parâmetros da equação da variância significativos. De modo similar ao  $ARCH(4)$ , no momento da estimação conjunta dos termos da equação da média com os da variância, o segundo coeficiente do  $AR(2)$  não se mostrou significativo. Mais uma vez, optou-se por sua manutenção.

A assimetria na volatilidade foi captada pelo modelo, onde o parâmetro ( $\gamma_1 > 0$ ) indica a presença do efeito *leverage*. Cabe destacar que a evidência de assimetria já havia sido notada no comportamento visual das funções univariadas do modelo aditivo (4), presente no GRAF. 10.

Analisando as medidas de ajuste do modelo  $TARCH(3,2)$ , pode-se perceber no GRAF. 21, item (c), uma redução da autocorrelação nos quadrados dos resíduos padronizados superior à apresentada pelo modelo  $ARCH(4)$ , no item (c) do GRAF. 20. Porém, o comportamento descrito pelo histograma no item (a) do GRAF. 21 indica a existência de algumas observações extremas remanescentes nas caudas.

Pode-se dizer que a escolha de um modelo  $AR(2)$  para a equação da esperança condicional foi adequada, na medida em que não é possível visualizar no item (b), dos GRAF. 20 e 21, um comportamento significativo para a autocorrelação nos resíduos padronizados.

Para completar a descrição dos modelos GARCH selecionados para a estimação da volatilidade dos retornos do Ibovespa, destacam-se no GRAF. 22 suas estimativas da variância condicional para a amostra completa. O item (a) descreve a evolução dos quadrados dos resíduos do  $AR(2)$ , sendo que os itens (b) e (c) apresentam as estimativas da volatilidade dos modelos  $ARCH(4)$  e  $TARCH(3,2)$ , respectivamente.

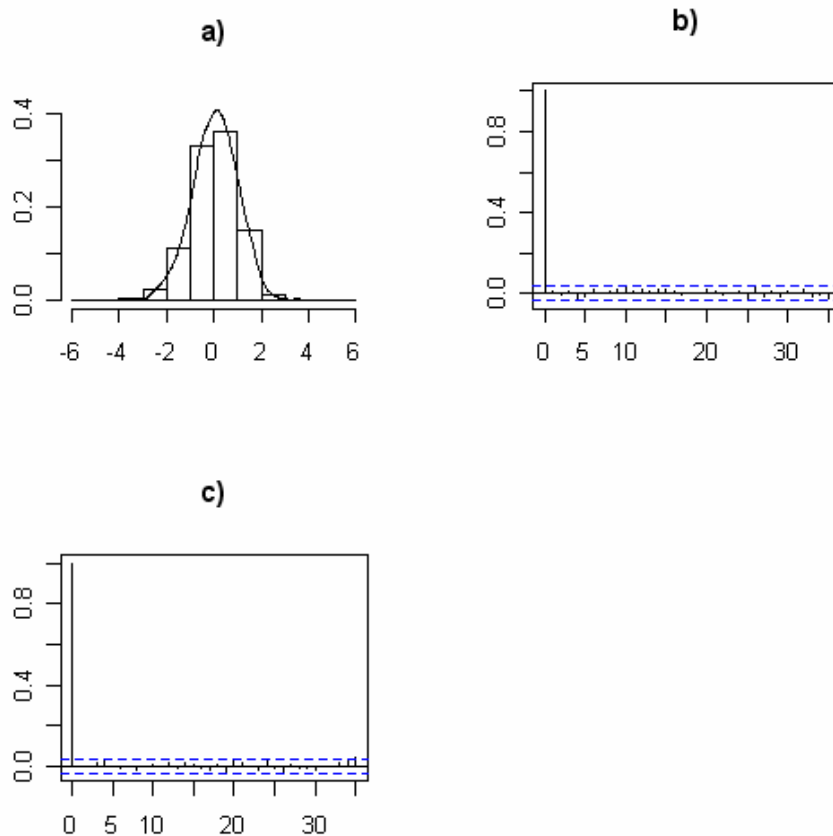


GRÁFICO 21 – Medidas de ajuste do Modelo  $TARCH(3,2)$

- a) Histograma dos resíduos padronizados
- b) Função de autocorrelação dos resíduos padronizados
- c) Função de autocorrelação dos quadrados dos resíduos pad.

Fonte: Dados da Economatica.

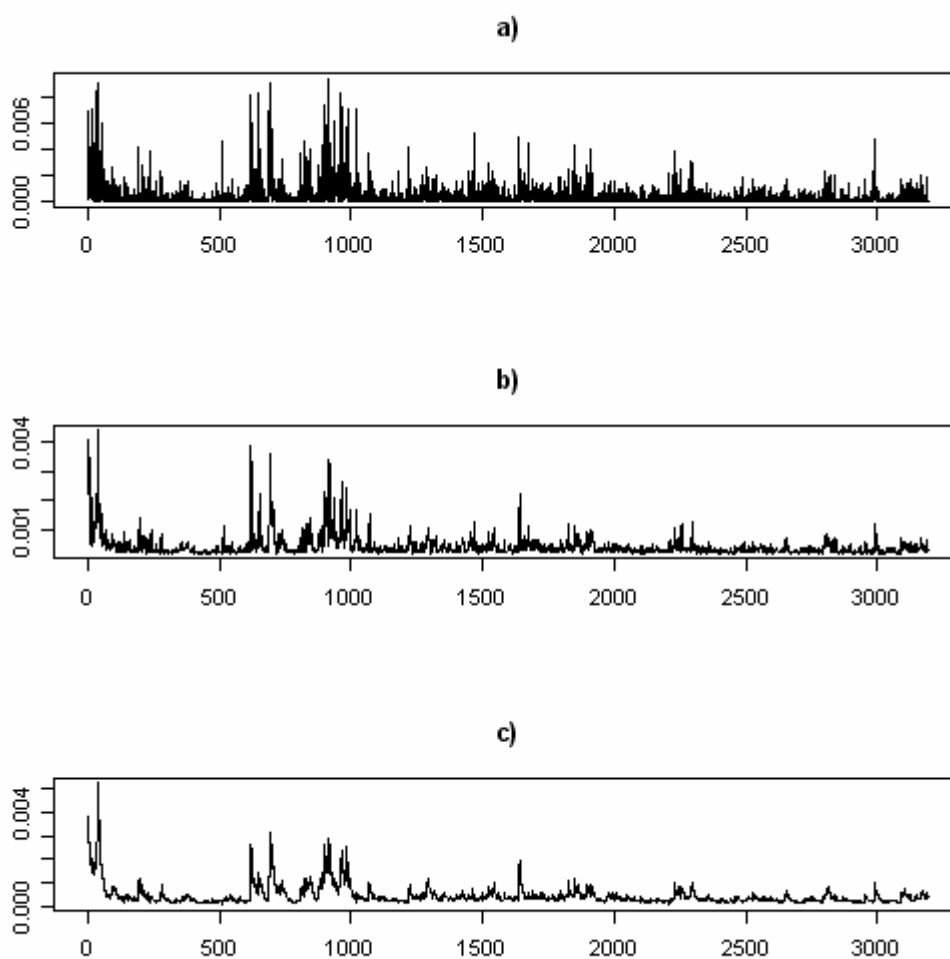


GRÁFICO 22 – Resíduos ao quadrado e volatilidade estimada do Ibovespa  
 a) Resíduos estimados do modelo  $AR(2)$  ao quadrado  
 b) Estimativas da variância condicional do modelo  $ARCH(4)$   
 c) Estimativas da variância condicional do modelo  $TAR(3,2)$

Fonte: Dados da Economatica.

#### 4.3.2.2 Modelos GARCH para o S&P500

Utilizando a mesma forma de inferência da série do Ibovespa, buscou-se selecionar os modelos mais indicados para os retornos do S&P500 através dos critérios de informação de Akaike (AIC) e de Schwarz (BIC) e a partir do valor máximo do logaritmo da verossimilhança. Novamente foram definidas na estimação duas distribuições condicionais para os erros padronizados ( $z_t = \varepsilon_t \sigma_t^{-1}$ ), a normal e a t-Student. Os melhores resultados foram reunidos na TAB. 12, sendo que esta exclui

os modelos cujas estimativas não convergiram, bem como os que não apresentaram parâmetros estatisticamente significativos ao nível de 10%.

TABELA 12  
Seleção dos modelos GARCH para o S&P500

Modelo	Normal			t-Student		
	AIC	BIC	Max [Log L]	AIC	BIC	Max [Log L]
ARCH(4)	<b>-6,432259</b>	<b>-6,422928</b>	<b>10.502,45</b>	..	..	..
GARCH(1,1)	-6,523916	-6,518318	10.650,03	-6,540567	-6,533103	10.678,21
GARCH(2,1)	..	..	..	-6,541639	-6,532309	10.680,95
EGARCH(1,1)	-6,557413	-6,549949	10.705,70	-6,570409	-6,561079	10.727,91
EGARCH(1,2)	-6,563358	-6,554028	10.716,40	..	..	..
TARCH(1,1)	-6,560185	-6,552721	10.710,22	-6,571902	-6,562571	10.730,34
TARCH(1,2)	-6,564905	-6,555575	10.718,93	<b>-6,577243</b>	<b>-6,566047</b>	<b>10.740,06</b>

Fonte: Dados da Economatica.

Nota (1): Dados em negrito indicam o melhor modelo.

Nota (2): [...] indica o modelo cujas estimativas não convergiram.

Dentre os resultados apresentados na TAB.12, foram selecionados os modelos  $ARCH(4)$  (gaussiano) e  $TARCH(1,2)$  com distribuição t-Student. Destaca-se que ambos os modelos possuem a média condicional definida como zero (assim como nos modelos aditivos). Sendo assim, a TAB.13 reúne as estimativas dos parâmetros contidos na equação da variância.

TABELA 13  
Estimativas dos modelos GARCH para o S&P500

	ARCH(4)	TARCH(1,2)
Média condicional		
..	..	..
Variância condicional		
$\omega$	4,35E-05 (0,000002)	1,22E-06 (0,0000002)
$\alpha_1$	0,1038 (0,014061)	-0,0833 (0,012207)
$\alpha_2$	0,1847 (0,022041)	0,0731 (0,013307)
$\alpha_3$	0,1555 (0,019481)	..
$\alpha_4$	0,1860 (0,022085)	..
$\beta_1$	..	0,9310 (0,008548)
$\gamma_1$	..	0,1430 (0,014448)
Graus de liberdade	..	12,5735 (2,058625)

Fonte: Dados da Economatica.

Nota (1): Erros padrão entre parênteses.

Nota (2): O modelo  $ARCH(4)$  segue a forma,  $\sigma_t^2 = \omega + \sum_{i=1}^4 \alpha_i \varepsilon_{t-i}^2$ .

Nota (3): O modelo  $TARCH(1,2)$  segue a forma,  $\sigma_t^2 = \omega + \sum_{i=1}^2 \alpha_i \varepsilon_{t-i}^2 + \beta_1 \sigma_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 Y_{t-1}$ , onde  $Y_t = 1$  se  $(\varepsilon_t < 0)$  e  $(0)$  caso contrário.

Todos os parâmetros do modelo  $ARCH(4)$  gaussiano se apresentaram altamente significativos. Sua escolha se deu na medida em que a versão estimada com distribuição t-Student não convergiu. As medidas de ajuste do modelo  $ARCH(4)$  são apresentadas no GRAF. 23. Este reúne o histograma dos retornos padronizados, sobreposto por uma densidade não paramétrica, no item (a), um gráfico quantil-quantil normal no item (b), além das funções de autocorrelação dos retornos padronizados e seus quadrados, nos itens (c) e (d), respectivamente.

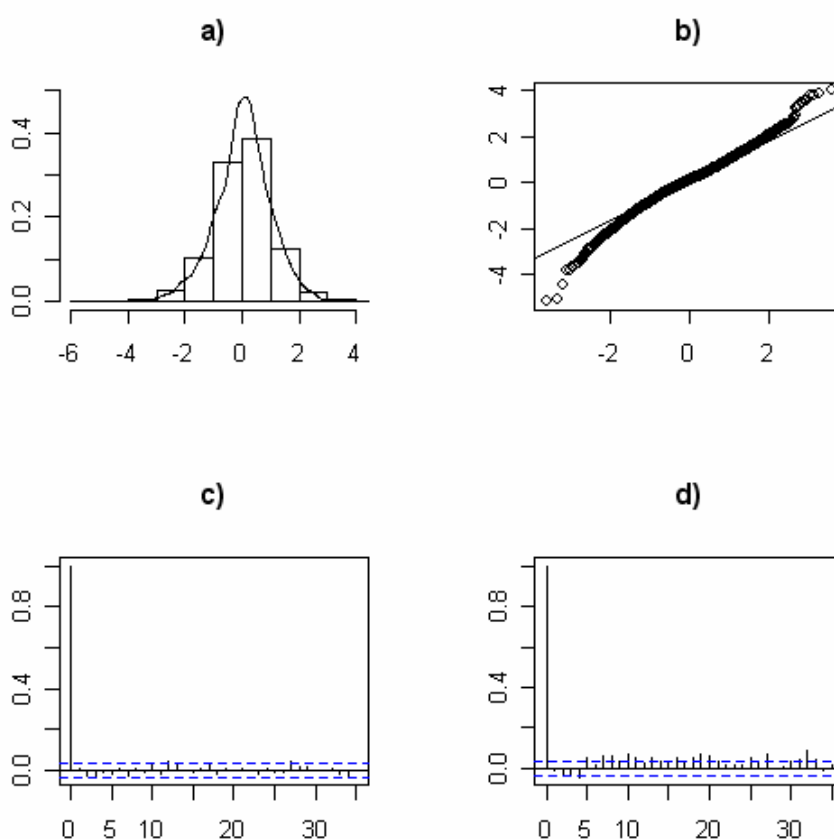


GRÁFICO 23 – Medidas de ajuste do Modelo  $ARCH(4)$

- a) Histograma dos retornos padronizados
- b) Gráfico QQ-normal dos retornos padronizados
- c) Função de autocorrelação dos retornos padronizados
- d) Função de autocorrelação dos quadrados dos retornos pad.

Fonte: Dados da Economatica.

Junto ao GRAF. 23, item (d), pode-se observar um resquício de autocorrelação nos quadrados dos retornos padronizados. Adicionalmente, pode-se perceber que nem toda a não-linearidade foi captada pelo modelo a partir do

histograma (item (a)) e do gráfico QQ-normal (item (b)). Ambos os gráficos demonstram a manutenção de caudas pesadas. Todavia, o modelo reduziu sensivelmente a autocorrelação nos quadrados quando em comparação com o item (d) do GRAF. 4. Através do item (c) pode-se concluir que a especificação da média condicional (igual a zero) não é imprópria.

O modelo  $TARCH(1,2)$  também apresentou todos os parâmetros altamente significativos. O mesmo captou assimetria na volatilidade dos retornos do S&P500, onde o parâmetro ( $\gamma_1 > 0$ ) indicou a presença do efeito *leverage*. Novamente, pode-se dizer que este comportamento assimétrico da variância condicional já havia sido notado nas descrições visuais das funções do modelo aditivo (4), presente no GRAF. 17.

Em relação às medidas de ajuste do modelo  $TARCH(1,2)$ , pode-se observar no item (c), do GRAF. 24, uma redução substancial da autocorrelação nos quadrados dos retornos padronizados em relação à realizada pelo modelo  $ARCH(4)$ , presente no item (d) do GRAF. 23. Através da comparação com o GRAF. 4, item (d), pode-se dizer que o modelo foi capaz de captar boa parte da variação da volatilidade. Contudo, mesmo utilizando uma distribuição com caudas mais pesadas que a normal ainda se percebe a manutenção das mesmas no histograma do item (a).

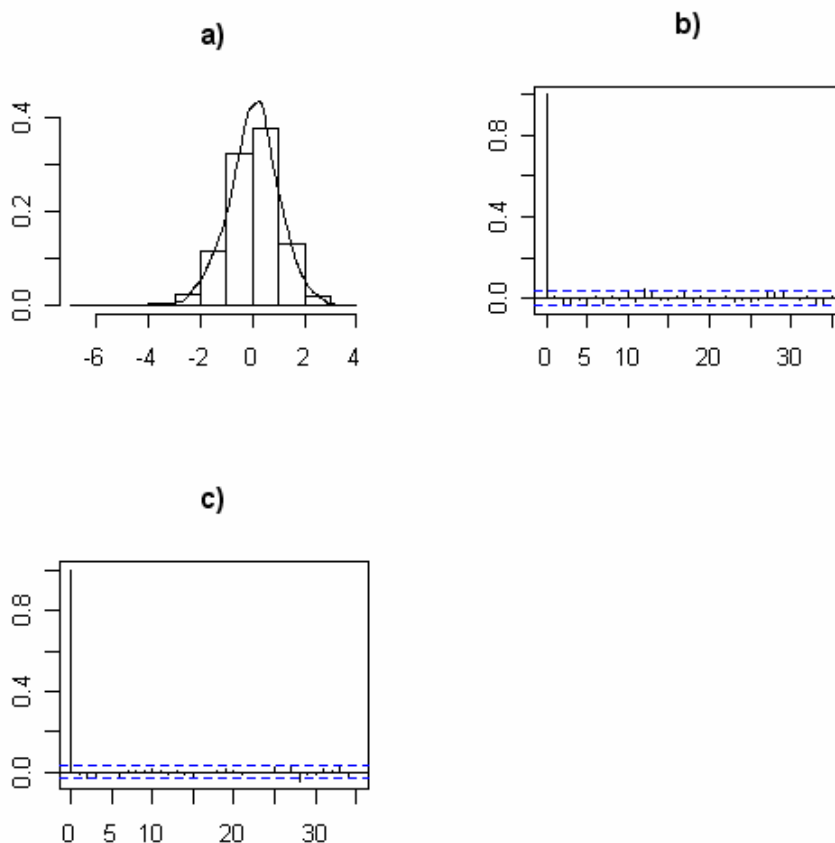


GRÁFICO 24 – Medidas de ajuste do Modelo  $TARCH(1,2)$

- a) Histograma dos retornos padronizados
- b) Função de autocorrelação dos retornos padronizados
- c) Função de autocorrelação dos quadrados dos retornos pad.

Fonte: Dados da Economatica.

Após o processo de inferência dos modelos GARCH na série dos retornos do S&P500, cabe descrever graficamente as estimativas da volatilidade obtidas a partir dos mesmos para a amostra completa. O GRAF. 25 destaca no item (a) os retornos ao quadrado, seguido das estimativas da variância condicional dos modelos  $ARCH(4)$  gaussiano e  $TARCH(1,2)$ , nos itens (b) e (c), respectivamente.

Finalizada a escolha dos modelos paramétricos, nas duas séries de retornos, pode-se partir para a realização das comparações entre os modelos aditivos e GARCH na próxima seção.



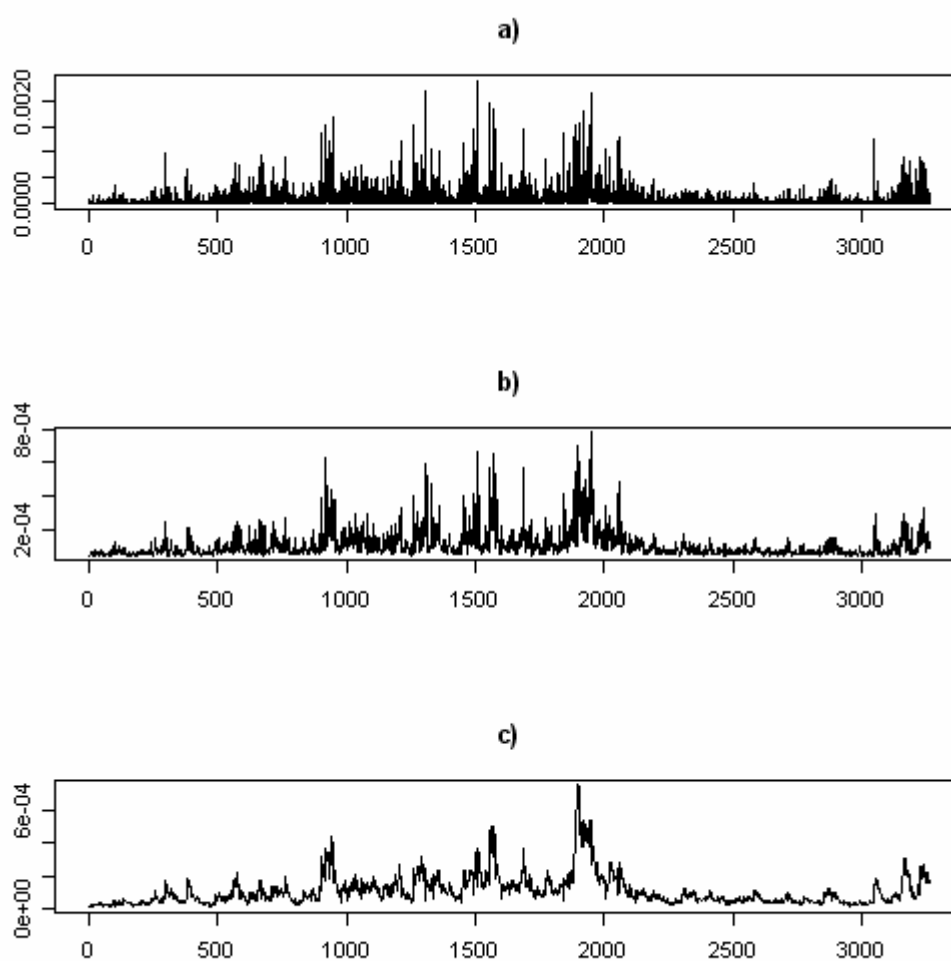


GRÁFICO 25 – Retornos ao quadrado e volatilidade estimada do S&P500  
a) Retornos ao quadrado  
b) Estimativas da variância condicional do modelo  $ARCH(4)$   
c) Estimativas da variância condicional do modelo  $TAR(1,2)$

Fonte: Dados da Economatica.

#### 4.4 COMPARAÇÃO ENTRE OS MODELOS ADITIVOS E GARCH

As comparações entre os modelos paramétricos e semi-paramétricos são realizadas a partir das três primeiras partições da série dos retornos do Ibovespa e do S&P500, apresentadas respectivamente nas TAB. 1 e 2 do APÊNDICE A. Para comparar o desempenho dos modelos dentro da amostra, os mesmos são estimados incorporando as observações contidas nos momentos de crise. Para o Ibovespa, este período é definido em 2001 de (03.09.2001-04.02.2002), em 2002 de

(02.09.2002-24.01.2003) e em 2007 de (02.07.2007-26.11.2007). Da mesma maneira para o S&P500, este período é definido em 2001 de (04.09.2001-01.02.2002), em 2002 de (03.09.2002-24.01.2003) e em 2007 de (02.07.2007-20.11.2007).

As medidas de erro são calculadas somente sobre as estimativas do intervalo de cem dias de negociação. Para a análise fora da amostra, as previsões um passo à frente atualizadas são efetuadas nos mesmos momentos referenciados de cem dias. Sendo assim, os modelos são estimados até o período  $t$ , e então são obtidas as previsões para  $t+1$ . Em  $t+1$ , os modelos são reestimados e as previsões são obtidas para  $t+2$ , e assim sucessivamente.

As medidas de erro utilizadas são duas: a raiz do erro quadrático médio (REQM) e o erro absoluto médio (EAM). Estas estatísticas avaliam as estimativas e previsões dos modelos em comparação aos quadrados dos retornos corrigidos pela média (estimador da volatilidade latente). As mesmas oferecem um indicativo do quão distantes estão as estimativas e previsões em relação aos quadrados dos resíduos (ou quadrados dos retornos para o S&P500):

$$REQM(x) = \left[ \frac{1}{T} \sum_{t=1}^T (x_t - \hat{\varepsilon}_t^2)^2 \right]^{1/2}, \quad (4.1)$$

$$EAM(x) = \frac{1}{T} \sum_{t=1}^T |x_t - \hat{\varepsilon}_t^2|, \quad (4.2)$$

onde  $\{\hat{\varepsilon}_t^2\}_{t=1}^T$  representa a série dos quadrados dos retornos corrigidos pela média,  $T$  é o número de dias estimados ou previstos e  $\{x_t\}_{t=1}^T$  denota a série de estimativas ou previsões da variância condicional de um determinado modelo.

#### 4.4.1 Comparação na Série do Ibovespa

Iniciando a análise pela série do Ibovespa, pode-se apresentar na TAB. 14 o desempenho dentro da amostra dos modelos selecionados (nas seções 4.3.1.1 e 4.3.2.1) para os três momentos de alta volatilidade.

TABELA 14  
Desempenho dos modelos (dentro da amostra) para o Ibovespa

	M. Aditivo (4)- <i>span</i> =0,7	ARCH (4) <i>t</i> (·)	TARCH (3,2) <i>t</i> (·)
Crise de 2001			
REQM	<b>0,0008412376</b>	0,0008990026	0,0008741969
EAM	<b>0,0005206397</b>	0,0005725927	0,0005428123
Crise de 2002			
REQM	0,0007112360	0,0006957946	<b>0,0006629027</b>
EAM	0,0004999630	0,0004883705	<b>0,0004564408</b>
Crise de 2007			
REQM	0,0005062302	0,0004984613	<b>0,0004836501</b>
EAM	0,0003785241	0,0003933001	<b>0,0003607623</b>

Fonte: Dados da Economatica.

Nota (1): Valor em negrito denota a melhor medida.

Os resultados da TAB. 14 evidenciam a superioridade do modelo aditivo (4) em estimar a variância condicional no período mais conturbado do ano de 2001. Ambas as medidas indicam o modelo semi-paramétrico como o mais adequado neste cenário, tendo um comportamento substancialmente superior a estrutura ARCH paramétrica com o mesmo número de defasagens. Para a crise de 2002 nota-se um melhor desempenho nas medidas, REQM e EAM, em favorecimento do modelo  $TARCH(3,2)$ . O resultado deste é sensivelmente superior ao apresentado pelos modelos concorrentes, sendo que o modelo aditivo (4) registra os valores mais elevados nas estatísticas de erro. Para completar os anos de crise analisados, em 2007 o modelo  $TARCH(3,2)$  obtém as menores medidas novamente. O modelo aditivo (4), por sua vez, supera o  $ARCH(4)$  com distribuição t-Student somente na medida EAM. Além disso, pode-se perceber que ao se incorporar mais observações da amostra em direção a 2007, o modelo paramétrico e assimétrico respondeu com resultados mais apropriados que o semi-paramétrico.

Uma forma complementar de visualizar o ajuste, nos períodos selecionados de crise, dá-se através dos gráficos que reúnem os retornos corrigidos pela média junto aos intervalos de confiança. Os últimos são baseados nas estimativas um passo à frente para o período de cem dias. São construídos a partir da duplicidade das estimativas um passo à frente dos desvios padrão condicionais. Ao mesmo tempo, este tipo de gráfico baseado nas informações passadas pode ser utilizado em análise de risco, municiando os profissionais do mercado financeiro na elaboração de estratégias de compra e venda de ativos com base na amplitude de

variação das bandas. Neste caso, menores amplitudes nos intervalos diminuem o tempo de exposição ao risco (ZIEGELMANN, 2002a).

Para a crise de 2001, pode-se apresentar no GRAF. 26 os resíduos do modelo  $AR(2)$  aliados aos intervalos de confiança construídos a partir das estimativas dos modelos reunidos na TAB. 14. Pode-se perceber no GRAF. 26, item (a), que as bandas do modelo aditivo (4) possuem menor amplitude por um maior período de tempo. O intervalo construído com as estimativas do modelo  $TARCH(3,2)$ , item (c), mostra-se mais adequado na medida em que se afasta das primeiras observações dos resíduos. Contudo, sua amplitude inicial é superior a apresentada pelo item (a). Nesta situação, o resultado corrobora as medidas de erro para o período, registradas na TAB. 14.

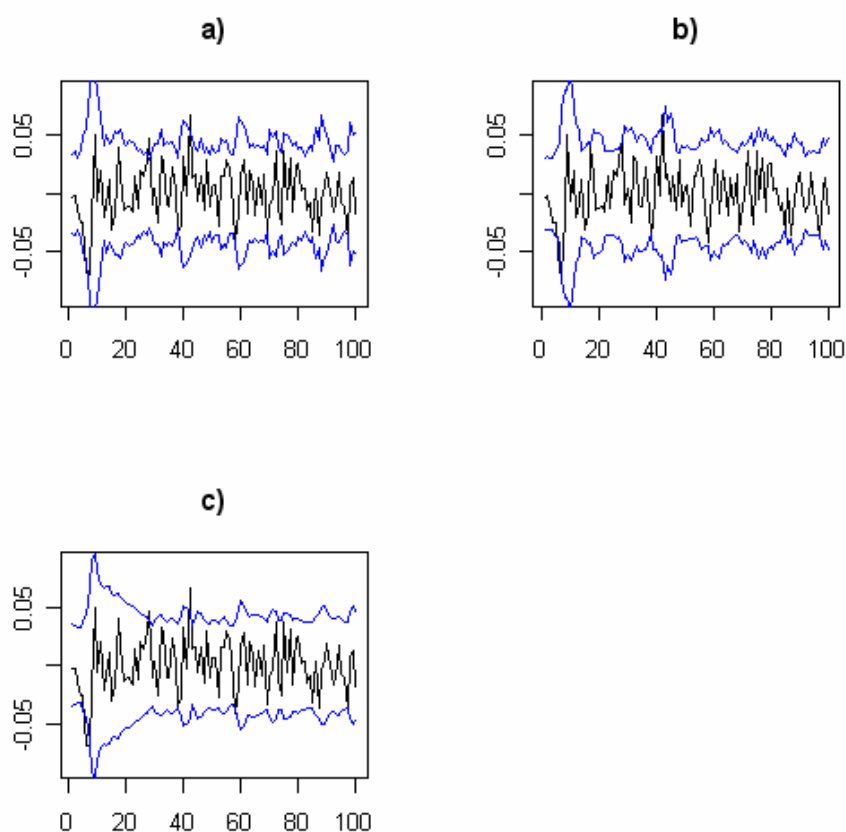


GRÁFICO 26 – Intervalos de confiança um passo à frente em 2001 - Ibovespa

a) Estimativas do Modelo aditivo (4)- $span=0,7$

b) Estimativas do Modelo  $ARCH(4)$   $t(\cdot)$

c) Estimativas do Modelo  $TARCH(3,2)$   $t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Resíduos do  $AR(2)$  e estimativas de 03/09/2001 até 04/02/2002.

Nas amostras referentes às crises de 2002 e 2007, o mesmo comportamento é notado. O modelo  $TARCH(3,2)$ , que obteve as melhores medidas de erro em 2002 e 2007, apresenta os intervalos de confiança com menor amplitude. Os GRAF. 27 e 28 que englobam os ajustes dentro da amostra em 2002 e 2007, respectivamente, podem ser avaliados no APÊNDICE E.

Na avaliação dos modelos quanto à capacidade preditiva, salienta-se que os resultados são semelhantes aos obtidos dentro da amostra. A TAB. 15 registra os erros de previsão um passo à frente, por modelo, em cada período selecionado.

TABELA 15  
Desempenho dos modelos (fora da amostra) para o Ibovespa

	M. Aditivo (4)- $span=0,7$	$ARCH(4) t(\cdot)$	$TARCH(3,2) t(\cdot)$
Crise de 2001			
REQM	<b>0,0008473683</b>	0,0009058024	0,0008761446
EAM	<b>0,0005238287</b>	0,0005774203	0,0005435897
Crise de 2002			
REQM	0,0007181177	0,0007046350	<b>0,0006892591</b>
EAM	0,0005060001	0,0004948852	<b>0,0004756111</b>
Crise de 2007			
REQM	0,0005101338	0,0004997983	<b>0,0004839464</b>
EAM	0,0003817302	0,0003941795	<b>0,0003608995</b>

Fonte: Dados da Economatica.

Nota (1): Valor em negrito denota a melhor medida.

Novamente, nota-se a superioridade do modelo aditivo (4) para o ano de 2001. Esta é evidenciada pelas estatísticas de erro inferiores as registradas pelos modelos concorrentes, especialmente em relação ao ARCH paramétrico de mesma ordem. Para o período seguinte (crise de 2002) o modelo  $TARCH(3,2)$  com distribuição t-Student destaca-se como o mais indicado. Este apresenta ambas as medidas sensivelmente inferiores em relação as demais estruturas. Adicionalmente, o modelo semi-paramétrico obteve um desempenho pior que o  $ARCH(4)$  fora da amostra. Por fim, em 2007 o resultado se repete, com o modelo assimétrico  $TARCH(3,2)$  sendo o escolhido a partir das duas estatísticas utilizadas.

Avaliando os gráficos que contém os intervalos de confiança (junto aos resíduos) construídos através das previsões um passo à frente dos desvios padrão condicionais, pode-se apontar que os mesmos ratificam os números da TAB. 15.

Como a apresentação dos mesmos não se diferencia substancialmente, em comparação aos gráficos construídos com as estimativas dos modelos, destaca-se o GRAF. 30 referente à crise de 2002. Os demais, GRAF. 29 e 31 podem ser observados no APÊNDICE F.

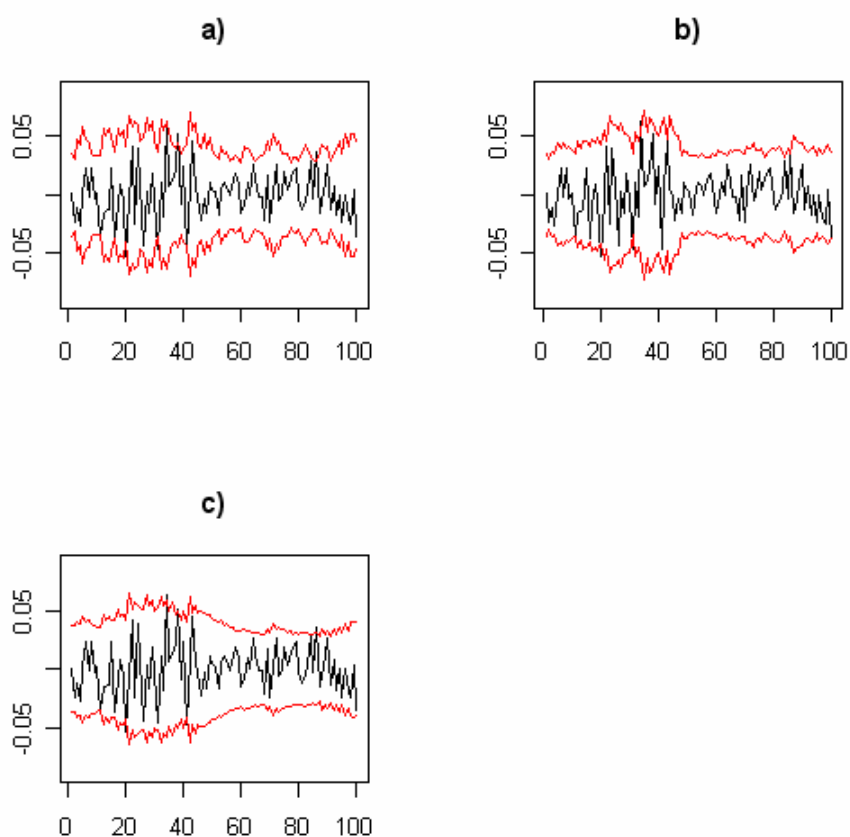


GRÁFICO 30 – Intervalos de confiança um passo à frente em 2002 - Ibovespa

a) Previsões do Modelo aditivo (4)- $span=0,7$

b) Previsões do Modelo  $ARCH(4)$   $t(\cdot)$

c) Previsões do Modelo  $TARCH(3,2)$   $t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Resíduos do  $AR(2)$  e previsões de 02/09/2002 até 24/01/2003.

#### 4.4.2 Comparação na Série do S&P500

O mesmo método é utilizado para o S&P500, onde as comparações entre os modelos da classe GARCH (seção 4.3.2.2) e os aditivos semi-paramétricos

(seção 4.3.1.2) são efetuadas a partir das três primeiras sub-amostras dos retornos do índice. Desta forma, os resultados obtidos nas medidas de erro, REQM e EAM, entre os retornos ao quadrado e as estimativas da volatilidade podem ser visualizados na TAB. 16.

TABELA 16  
Desempenho dos modelos (dentro da amostra) para o S&P500

	M. Aditivo (4)- <i>span</i> =0,8	<i>ARCH</i> (4) <i>N</i> (·)	<i>TARCH</i> (1,2) <i>t</i> (·)
Crise de 2001			
REQM	0,0002131765	0,0002267755	<b>0,0002080361</b>
EAM	<b>0,0001383292</b>	0,0001456224	0,0001393204
Crise de 2002			
REQM	0,0003956438	0,0003903236	<b>0,0003755018</b>
EAM	<b>0,0002415433</b>	0,0002599567	0,0002672503
Crise de 2007			
REQM	0,0002235213	0,0002254274	<b>0,0002220203</b>
EAM	<b>0,0001472079</b>	0,0001549095	0,0001547281

Fonte: Dados da Economatica.

Nota (1): Valor em negrito denota a melhor medida.

Observa-se na TAB. 16 que segundo o critério (REQM), o modelo assimétrico *TARCH*(1,2) é escolhido como o mais adequado para o período de 2001. Todavia, o critério EAM que leva em conta as diferenças absolutas seleciona o modelo aditivo (4) com *span*=0,8. Ambas as medidas de erro do modelo semi-paramétrico são inferiores as registradas para o modelo paramétrico com o mesmo número de defasagens. Em relação à crise de 2002, o modelo *TARCH*(1,2) com distribuição t-Student continua apresentando o menor valor para a REQM, ao passo que o modelo aditivo (4) se destaca na medida EAM. Neste período, o modelo semi-paramétrico erra mais (segundo a REQM) ao estimar a volatilidade que o *ARCH*(4) gaussiano. Para o último momento avaliado (2007), a situação permanece a mesma, com o modelo aditivo (4) sendo selecionado pelo EAM e o *TARCH*(1,2) pela REQM.

Junto ao GRAF. 32, item (a), pode-se perceber que o intervalo de confiança gerado, a partir das estimativas do modelo aditivo (4), apresenta uma amplitude ligeiramente inferior a dos demais modelos nas primeiras observações dos retornos. Para o ano de 2002, representado no GRAF. 33 do APÊNDICE G, o intervalo de confiança visualizado no item (a), construído através das estimativas do

modelo semi-paramétrico, destaca-se nitidamente como o de menor amplitude. Contudo, em 2007 as diferenças nas dispersões das bandas dos itens (a), (b) e (c) não são tão perceptíveis, ficando difícil discernir sobre o melhor ajuste (GRAF. 34, APÊNDICE G).

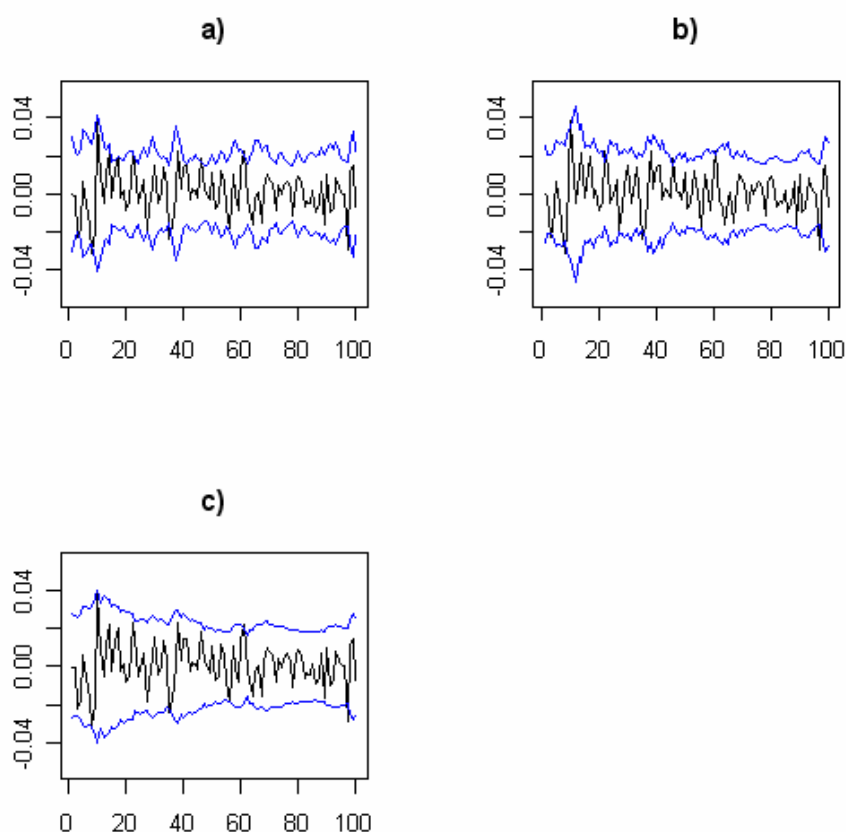


GRÁFICO 32 – Intervalos de confiança um passo à frente em 2001 - S&P500

a) Estimativas do Modelo aditivo (4)- $span=0,8$

b) Estimativas do Modelo  $ARCH(4) N(\cdot)$

c) Estimativas do Modelo  $TARCH(1,2) t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Retornos do S&P500 e estimativas de 04/09/2001 até 01/02/2002.

O desempenho entre os modelos alternativos, quanto à capacidade de previsão um passo à frente atualizada, também foi medido através das estatísticas de erro REQM e EAM. Neste processo, os resultados obtidos na avaliação do desempenho dentro da amostra se repetiram. As medidas utilizadas não selecionaram as mesmas estruturas em cada período de crise analisado. A TAB. 17 reúne os valores obtidos para os erros de previsão dos modelos em cada partição da série dos retornos do S&P500.



TABELA 17  
Desempenho dos modelos (fora da amostra) para o S&P500

	M. Aditivo (4)- $span=0,8$	$ARCH(4) N(\cdot)$	$TARCH(1,2) t(\cdot)$
Crise de 2001			
REQM	0,0002147466	0,0002276209	<b>0,0002082377</b>
EAM	<b>0,0001392696</b>	0,0001466709	0,0001395953
Crise de 2002			
REQM	0,0004049819	0,0003918914	<b>0,0003755383</b>
EAM	<b>0,0002428046</b>	0,0002599835	0,0002672952
Crise de 2007			
REQM	0,0002252650	0,0002260639	<b>0,0002220382</b>
EAM	<b>0,0001477665</b>	0,0001553677	0,0001547516

Fonte: Dados da Economatica.

Nota (1): Valor em negrito denota a melhor medida.

A TAB. 17 evidencia que para a crise de 2001 o modelo mais indicado, segundo o critério REQM, é o  $TARCH(1,2)$  com distribuição t-Student. Todavia, o valor registrado pelo erro absoluto médio é inferior para o modelo aditivo(4). Neste período a superioridade do modelo semi-paramétrico, em relação ao  $ARCH(4)$ , é observada através de ambas as medidas. No ano de 2002 o impasse na seleção de um modelo único segue, haja vista cada estatística de erro ter selecionado um modelo distinto. O modelo assimétrico  $TARCH(1,2)$  obtém o menor valor para a REQM e o modelo aditivo (4) registra o menor valor para o EAM. De acordo com o critério REQM o modelo semi-paramétrico realiza as piores previsões em 2002. Contudo, em 2007 o modelo aditivo (4) registra ambas as medidas inferiores ao  $ARCH(4)$ , ao passo que o EAM o indica como a estrutura mais apropriada para o período de crise. De modo similar a 2001 e 2002, o modelo  $TARCH(1,2)$  é selecionado pela estatística REQM.

Devido à similaridade dos gráficos, compostos pelos intervalos de confiança das estimativas e previsões, selecionou-se o GRAF. 36 que aponta a superioridade do modelo aditivo (4) em 2002, item (a), frente às estruturas alternativas. Pode-se notar a partir do mesmo, que a amplitude das bandas em torno dos retornos iniciais é substancialmente inferior a descrita nos itens (b) e (c). De maneira complementar, os GRAF. 35 e 37 que destacam o ajuste das bandas para os anos de 2001 e 2007, seguem no APÊNDICE H.

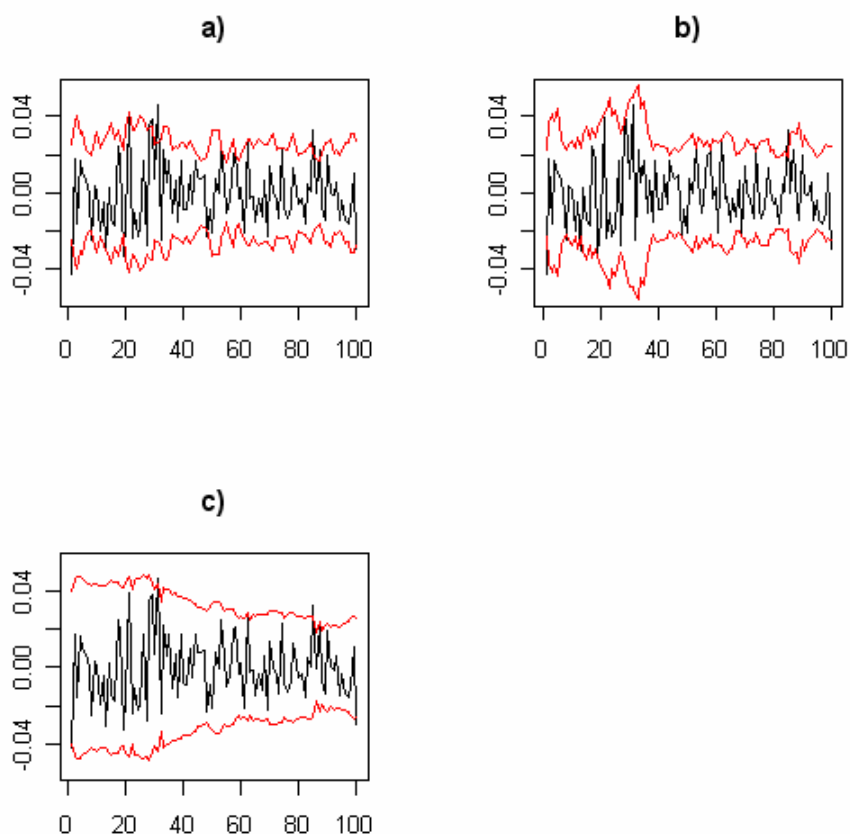


GRÁFICO 36 – Intervalos de confiança um passo à frente em 2002 - S&P500

a) Previsões do Modelo aditivo (4)- $span=0,8$

b) Previsões do Modelo  $ARCH(4)$   $N(\cdot)$

c) Previsões do Modelo  $TARCH(1,2)$   $t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Retornos do S&P500 e previsões de 03/09/2002 até 24/01/2003.

Ao final deste capítulo, pode-se considerar que as comparações efetuadas entre os modelos GARCH paramétricos e os aditivos semi-paramétricos destacam a capacidade dos últimos em estimar e prever adequadamente a volatilidade dos retornos de ambos os índices.

Ao avaliar o comportamento das estruturas estimadas de forma distinta, ou seja, regressão global para os GARCH e local para os aditivos, obteve-se resultados satisfatórios na modelagem semi-paramétrica. Os modelos aditivos, em ambas as séries estudadas, foram superiores em alguns períodos de crise em relação às estruturas tradicionalmente utilizadas da classe GARCH. Estes resultados foram corroborados dentro e fora da amostra, tanto pelas medidas de ajuste adotadas, quanto pelos gráficos compostos dos intervalos de confiança.

Para a série dos retornos do Ibovespa, que apresenta uma maior volatilidade em relação à série do S&P500, notou-se uma superioridade do modelo paramétrico, especificamente o  $TARCH(3,2)$  com distribuição t-Student, frente ao modelo aditivo composto de quatro elementos funcionais.

Porém, para a série do S&P500 (a de comportamento menos volátil), os resultados avaliados nos períodos de crise foram dúbios, sendo que cada medida de erro selecionou um estimador distinto. O modelo aditivo (4) com  $span=0,8$  foi indicado pelo EAM, talvez a estatística mais apropriada para a situação por ser mais robusta a observações extremas, e o  $TARCH(1,2)$  assimétrico foi apontado pela REQM.

## 5 CONSIDERAÇÕES FINAIS

Algumas considerações podem ser realizadas sobre o objetivo proposto de comparar os modelos aditivos semi-paramétricos com os amplamente empregados na literatura de finanças, modelos GARCH paramétricos.

A eliminação de algumas observações extremas pode ser entendida como um ponto negativo na estimação dos modelos aditivos. Devido às funções univariadas serem bastante flexíveis, a exclusão dos *outliers*, principalmente na série do Ibovespa (mais volátil), fez-se necessária para a obtenção de estimativas apropriadas para a variância condicional dos retornos, visto que estimativas de volatilidade para estes valores extremos seriam prejudicadas pela ausência de vizinhança local. Nota-se que distintos parâmetros de suavização poderiam ser uma alternativa nesta situação. Entretanto, tal abordagem estaria fora do escopo deste trabalho. Outra característica encontrada para ambas as séries foi a inadequação de modelos com mais de cinco variáveis defasadas. Deve-se ressaltar que em diversas aplicações com modelos paramétricos também se faz uso da redução de peso de valores extremos (*outliers*). De qualquer forma, as séries utilizadas em ambos os modelos foram as mesmas, não prejudicando o exercício de comparação entre os mesmos.

Os modelos aditivos descrevem comportamentos visuais informativos da dependência (cuja forma não é pré-determinada) da variável resposta frente às variáveis explicativas. Esta propriedade se apresenta como um ponto extremamente positivo no processo de inferência. A partir da visualização do ajuste em cada componente funcional pode-se notar a presença de não linearidades, assimetrias etc.

Avaliando os resultados obtidos na série completa dos retornos do Ibovespa quanto ao ajuste, pôde-se constatar uma superioridade do modelo  $TARCH(3,2)$  com distribuição t-Student em relação ao modelo aditivo com quatro elementos funcionais. A capacidade de redução da autocorrelação nos quadrados dos resíduos (do  $AR(2)$ ) padronizados, apresentada pelo primeiro modelo, mostrou-

se substancialmente maior que a apresentada pelo segundo. Todavia, o modelo aditivo quando comparado ao ARCH paramétrico de mesma ordem mostrou-se equivalente sob o mesmo critério. Por meio da visualização dos elementos funcionais do modelo aditivo, puderam-se observar indícios de assimetria na volatilidade dos retornos, fato posteriormente corroborado pela escolha de um modelo da classe GARCH assimétrico. Talvez nesta situação a assimetria seja um fator demasiadamente importante, fazendo com que o  $TARCH(3,2)$ , por apresentar esta particular estrutura, seja superior.

Para a amostra completa do S&P500, os resultados do ajuste evidenciaram novamente a superioridade de um modelo assimétrico da família GARCH, o  $TARCH(1,2)$  com distribuição t-Student. Além disso, o modelo semi-paramétrico com quatro funções univariadas apresentou uma redução ligeiramente inferior da autocorrelação nos quadrados dos retornos padronizados frente ao  $ARCH(4)$ . Mais uma vez o modelo aditivo indicou uma assimetria na volatilidade dos retornos, também captada pelo modelo  $TARCH(1,2)$ .

Com a proposição de comparar os modelos em períodos de alta volatilidade, procedeu-se à escolha de três anos com crises no mercado financeiro internacional. Os anos selecionados foram os de 2001 (atentado de 11 de setembro), 2002 (desdobramentos da bolha tecnológica) e 2007 (mercado de crédito *subprime*). Foram utilizadas duas estatísticas, a REQM e o EAM, para medir o desempenho das estimativas e previsões da volatilidade em relação aos quadrados dos retornos (corrigidos pela média). Salieta-se que a segunda medida (EAM) é mais robusta frente a valores extremos causados por grandes movimentos no mercado, sendo possivelmente a mais indicada nos períodos de crise.

Para o período de cem dias de negociação, definido como representativo da crise de 2001, observou-se para a série do Ibovespa que ambas as medidas de erro apontaram (dentro da amostra) o modelo aditivo com quatro funções univariadas como o mais apropriado.

Entretanto, o modelo  $TARCH(3,2)$  registrou estatísticas de erro inferiores aos demais modelos para os anos de 2002 e 2007. Estes mesmos resultados se repetiram nas previsões um passo à frente, onde o modelo aditivo foi selecionado em 2001, ao passo que o modelo  $TARCH(3,2)$  foi indicado como superior em 2002

e 2007. Ressalta-se que somente na crise de 2002 o modelo semi-paramétrico apresentou desempenho inferior nas duas medidas (dentro e fora da amostra) em comparação ao  $ARCH(4)$ .

Para a série dos retornos do S&P500, em cada crise as estatísticas indicaram um modelo entre o aditivo com quatro funções univariadas e o  $TARCH(1,2)$ . O primeiro modelo foi apontado pelo EAM em todos os três anos como o de melhor performance, ao passo que a REQM selecionou para os mesmos anos o modelo assimétrico. Este resultado foi encontrado tanto para as estimativas da volatilidade quanto para as previsões um passo à frente.

Os resultados comparativos, de maneira geral, revelaram a capacidade dos modelos aditivos em estimar adequadamente a volatilidade dos retornos dos índices escolhidos. Adicionalmente, o desempenho dos mesmos em relação aos períodos de crise mostrou-se satisfatório, indicando uma superioridade em relação a especificações ARCH de mesma ordem. Entretanto, na série mais volátil (Ibovespa), o modelo paramétrico  $TARCH(3,2)$  mostrou-se o mais indicado, ao passo que na série menos volátil (S&P500) os resultados foram inconclusivos.

Tendo utilizado o estimador tradicional da volatilidade latente (ou seja, o quadrado do retorno corrigido pela média  $\{\varepsilon_t^2\}$ ), pôde-se comparar de forma adequada os modelos propostos. Segundo Andersen e Bollerslev (1998) este estimador é capaz de gerar estimativas não viesadas da volatilidade não observável, embora as mesmas sejam altamente voláteis e inconsistentes. Desta forma, a substituição da base de comparação utilizada neste trabalho pelo estimador da volatilidade realizada (desenvolvido pelos autores) apresenta-se como uma alternativa a ser apreciada. A última medida é obtida a partir de dados em alta frequência (retornos intradiários) e, segundo Andersen e Bollerslev (1998), é assintoticamente livre de erro.

Sugere-se como trabalho futuro a utilização do estimador da variância realizada, sendo esta proposta válida para estender os resultados obtidos nesta dissertação. Adicionalmente, há ainda a possibilidade de comparação com outros modelos paramétricos e semi-paramétricos, além da estimação de modelos aditivos aplicados a outros problemas em economia.

Concluindo, os modelos aditivos não devem necessariamente ser entendidos como concorrentes aos paramétricos, mas sim complementares. Por exemplo, suas propriedades visualmente informativas podem ser utilizadas em auxílio a parametrizações mais adequadas entre as variáveis envolvidas em determinado problema. Assim, destacam-se como mais uma ferramenta disponível ao pesquisador.

## REFERÊNCIAS

ANDERSEN, T. G.; BOLLERSLEV, T. Intraday Periodicity and Volatility Persistence in Financial Markets. *Journal of Empirical Finance*, v. 4, p. 115-158, 1997.

ANDERSEN, T. G.; BOLLERSLEV, T. Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, v. 39, p. 885-905, 1998.

BAILLIE, R. T.; BOLLERSLEV, T.; MIKKELSEN, H. O. Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, v. 74, p. 3-30, 1996.

BANCO CENTRAL DO BRASIL. *Boletim do Banco Central do Brasil*. Relatório Anual 2001, v. 37, set. 2002. Disponível em: <<http://www.bcb.gov.br/?BOLETIM2001>>. Acesso em: 03 abr. 2008.

BANCO CENTRAL DO BRASIL. *Boletim do Banco Central do Brasil*. Relatório Anual 2002, v. 38, Jul. 2003. Disponível em: <<http://www.bcb.gov.br/?BOLETIM2002>>. Acesso em: 05 abr. 2008.

BANCO CENTRAL DO BRASIL. *Boletim do Banco Central do Brasil*. Relatório Mensal, v. 43, n. 10. Out. 2007. Disponível em: <<http://www.bcb.gov.br/ftp/histbole/Bol200710P.pdf>>. Acesso em: 07 abr. 2008.

BANCO CENTRAL DO BRASIL. *Boletim do Banco Central do Brasil*. Relatório Mensal, v. 44, n. 02. Fev. 2008. Disponível em: <<http://www.bcb.gov.br/ftp/histbole/Bol200802P.pdf>>. Acesso em: 07 abr. 2008.

BAUR, D.; JUNG, R. Return and Volatility Linkages Between the US and the German Stock Market. *Journal of International Money and Finance*, v. 25, p. 598-613, 2006.

BELLMAN, R. E. *Adaptive Control Processes*. New Jersey: Princeton University Press, 1961.



BLACK, F. Studies of Stock Price Volatility Changes. *Proceedings from the American Statistical Association, Business and Economic Statistics Section*, p. 177-181, 1976.

BOLLERSLEV, T. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, v. 31, p. 307-27, 1986.

BOLLERSLEV, T. A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, v. 69, p. 542-47, 1987.

BOLLERSLEV, T. *Glossary to ARCH (GARCH)*. Disponível em: <<http://faculty.chicagogsb.edu/jeffrey.russell/teaching/finecon/readings/glossary.pdf>>. Acesso em: 20 fev. 2008.

BOLLERSLEV, T.; CHOU, R. Y.; KRONER, K. F. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, v. 52, p. 5-59, 1992.

BOLLERSLEV, T.; ENGLE, R. F.; NELSON, D. B. ARCH Models. In: ENGLE, R. F.; MCFADDEN, D. L. *Handbook of Econometrics*. Amsterdam: Elsevier Sciences B. V., v. IV, cap. 49, 1994.

BOLLERSLEV, T; MELVIN, M. Bid-Ask Spreads in the Foreign Exchange Market: An Empirical Analysis. *Journal of International Economics*, v. 36, p. 355-372, 1994.

BOWMAN, A.; AZZALINI, A. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford: Oxford University Press, 1997.

BUREAU OF ECONOMIC ANALYSIS (BEA). *National Economic Accounts - Gross Domestic Product (GDP)*. Disponível em: <<http://www.bea.gov/national/xls/gdpchg.xls>>. Acesso em: 28 mar. 2008.

BUSETTI, G.; MANERA, M. *STAR-GARCH Models for Stock Market Interactions in the Pacific Basin Region, Japan and US*. Disponível em: <[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=419081](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=419081)>. Acesso em: 04 mar. 2008.

CAI, J. A Markov Model of Switching-Regime ARCH. *Journal of Business and Economic Statistics*, v. 12, p. 309-316, 1994.

CHRISTIE, A. A. The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. *Journal of Financial Economics*, v. 10, p. 407-432, 1982.

CLEVELAND, W. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, v. 74, p. 829-836, 1979.

CLEVELAND, W.; Devlin, S. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* v. 83, p. 596-610, 1988.

CLEVELAND, W.; LOADER, Clive. Smoothing by Local Regression: Principles and Methods. In: HÄRDLE, Wolfgang; SCHIMEK, M. G. (Eds.). *Statistical Theory and Computational Aspects of Smoothing*. Heidelberg: Physica-Verlag, p. 10-49, 1996.

CORNELL, B. Using the Options Pricing Model to Measure the Uncertainty Producing Effect of Major Announcements. *Financial Management*, v. 7, p. 54-59, 1978.

DICKEY, D. A.; FULLER, W. A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, v. 74, p. 427-431, 1979.

ENGLE, R. F. Autoregressive Conditional Heteroskedasticity with Estimates of the Variances of U.K. Inflation. *Econometrica*, v. 50, n. 4, p. 987-1008, 1982.

ENGLE, R. F. Discussion: Stock Market Volatility and the Crash of 87. *Review of Financial Studies*, v. 3, p. 103-106, 1990.

ENGLE, R. F.; BOLLERSLEV, T. Modelling the Persistence of Conditional Variances. *Econometric Reviews*, v. 5, p. 1-50, 1986.

ENGLE, R. F.; ITO, T.; LIN, W. L. Where Does the Meteor Shower Come From? The Role of Stochastic Policy. *Journal of International Economics*, v. 32, p. 221-240, 1992.

ENGLE, R. F.; LILIEN, D. M.; ROBINS, R. P. Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica*, v. 55, p. 391-407, 1987.

ENGLE, R. F.; MEZRICH, J. GARCH for Groups. *Risk*, v. 9, p. 36-40, 1996.

ENGLE, R. F.; NG, V. K. Measuring and Testing the Impact of News on Volatility. *The Journal of Finance*, v. 48, p. 1749-1778, 1993.

ENGLE, R. F.; SUSMEL, R. Common Volatility in International Equity Markets. *Journal of Business and Economic Statistics*, v. 11, p. 167-176, 1993.

FAMA, E. F. The Behavior of Stock Market Prices. *Journal of Business*, v. 38, p. 34-105, 1965.

FAN, J.; GIJBELS, I. *Local Polynomial Modeling and Its Applications*. London: Chapman and Hall, 1996.

FAN, J.; YAO, Q. Efficient Estimation of Conditional Variance Functions in Stochastic Regression. *Biometrika*, v. 85, p. 645-660, 1998.

FAN, J.; YAO, Q. *Nonlinear Time Series: Nonparametric and Parametric Methods*. New York: Springer-Verlag, 2003.

FRENCH, K. R.; ROLL, R. Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics*, v. 17, p. 5-26, 1986.

FRIEDMAN, J. H.; STUETZLE, W. Projection Pursuit Regression. *Journal of the American Statistical Association*, v. 76, p. 817-823, 1981.

GLOSTEN, L. R.; JAGANNATHAN, R.; RUNKLE, D. On the Relation Between the Expected Value and the Volatility of the Normal Excess Return on Stocks. *Journal of Finance*, v. 48, p. 1779-1801, 1993.

GRIGOLETTO, M.; LISI, F. *Looking for Skewness in Financial Time Series*. Disponível em: <<http://homes.stat.unipd.it/lisif/garchdsk.pdf>>. Acesso em: 06 mar. 2008.

HALL, P.; YAO, Q. Inference in ARCH and GARCH Models with Heavy-Tailed Errors. *Econometrica*, v. 71, p. 285-317, 2003.

HAMILTON, J. D.; SUSMEL, R. Autoregressive Conditional Heteroskedasticity and Changes in Regimes. *Journal of Econometrics*, v. 64, p. 307-333, 1994.

HASTIE, T. J. Generalized additive models. In: CHAMBERS, J. M.; HASTIE, T. J. *Statistical Models in S*. London: Chapman and Hall, cap. 7, 1993.

HASTIE, T.; TIBSHIRANI, R. *Generalized Additive Models*. London: Chapman and Hall, 1990.

INTERNATIONAL MONETARY FUND (IMF). *World Economic Outlook (WEO)*. Abr. 2003. Disponível em: <<http://www.imf.org/external/pubs/ft/weo/2003/01/index.htm>>. Acesso em: 25 mar. 2008.

INTERNATIONAL MONETARY FUND (IMF). *World Economic Outlook (WEO)*. Abr. 2008. Disponível em: <<http://www.imf.org/external/pubs/ft/weo/2008/01/index.htm>>. Acesso em: 27 mar. 2008.

ISSLER, J. V. Estimating and Forecasting the Volatility of Brazilian Finance Series Using ARCH Models. *Brazilian Review of Econometrics*, v. 19, p. 5-56, 1999.

JARQUE, C. M.; BERA, A. K. Efficient Tests for Normality, Heteroskedasticity and Serial Independence of Regression Residuals. *Economics Letters*, v. 6, p. 255-259, 1980.

KIM, W.; LINTON, O.B. The Live Method for Generalized Additive Volatility Models. *Econometric Theory*, v. 20, n. 6, p. 1094-1139, 2004.

LEVINE, M.; LI, J. Local Instrumental Variable (LIVE) Method for the Generalized Additive-Interactive Nonlinear Volatility Model. Disponível em: <<http://www.stat.purdue.edu/~mlevins/Articles/ch1234ref.pdf>>. Technical Report, 2007. Acesso em: 15 abr. 2008.

LJUNG, G. M.; BOX, G. E. P. On a Measure of Lag of Fit in Time Series Models. *Biometrika*, v. 67, p. 297-303, 1978.

LOADER, Catherine. Smoothing: Local Regression Techniques. In: GENTLE, James; HÄRDLE, Wolfgang; MORI, Yoichi (Eds.). *Handbook of Computational Statistics*. Heidelberg: Springer-Verlag, 2004.

LOADER, Clive. *Old Faithful Erupts: Bandwidth Selection Reviewed*. Working paper, AT&T Bell Laboratory, 1995. Disponível em: <<http://citeseer.ist.psu.edu/cache/papers/cs/96/http:zSzzSzcm.bell-labs.comzSzcmzSzmszSzdepartmentszSzsiatzSzdoczSz95.9.pdf/loader95old.pdf>>. Acesso em: 18 mar. 2008.

MANDELBROT, B. The Variation of Certain Speculative Prices. *Journal of Business*, v. 36, p. 394-419, 1963.

MORAIS, I. A. C.; PORTUGAL, M. S. Modelagem e Previsão de Volatilidade Determinística e Estocástica para a Série do Ibovespa. *Estudos Econômicos*, v. 29, n. 3, p. 303-341, 1999.

MÜLLER, H. G.; STADMÜLLER, U. Estimation of Heteroscedasticity in Regression Analysis. *Annals of Statistics*, v. 15, p. 610-625, 1987.

NADARAYA, E. A. On Estimating Regression. *Theory of Probability and Its Applications*, v. 10, p. 186-190, 1964.

NELSON, D. B. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, v. 59, p. 347-370, 1991.

PAGAN, A. R.; SCHWERT, G. W. Alternative Models for Conditional Stock Volatility. *Journal of Econometrics*, v.45, p. 267-290, 1990.

PATELL, J. M.; WOLFSON, M. A. Anticipated Information Releases Reflected in Call Option Prices. *Journal of Accounting and Economics*, v. 1, p. 117-140, 1979.

PATELL, J. M.; WOLFSON, M. A. The Ex-Ante and Ex-Post Price Effects of Quarterly Earnings Announcement Reflected in Option and Stock Price. *Journal of Accounting Research*, v. 19, p. 434-458, 1981.

PATTERSON, K. *An Introduction to Applied Econometrics*. London: Palgrave Publishers, 2000.

PEREIRA, P. L. V.; HOTTA, L. K.; SOUZA, L. A. R.; ALMEIDA, N. M. C. G. Alternative Models to Extract Asset Volatility: A Comparative Study. *Brazilian Review of Econometrics*, v. 19, p. 57-109, 1999.

PHILLIPS, P. C. B.; PERRON, P. Testing for a Unit Root in Time Series Regression. *Biometrika*, v. 75, p. 335–346, 1988.

RUPPERT, D.; WAND, M. P. Multivariate Locally Weighted Least Squares Regression. *Annals of Statistics*, v. 22, p. 1346-1370, 1994.

RUPPERT, D.; WAND, M. P., HOLST, U.; HÖSSJER, O. Local Polynomial Variance Function Estimation. *Technometrics*, v. 39, p. 262-273, 1997.

SCHWERT, G. W. Why Does Stock Market Volatility Change Over Time. *Journal of Finance*, v. 44, p. 1115-1153, 1989.

SENTANA, E. Quadratic ARCH Models. *Review of Economic Studies*, v. 62, p. 639-661, 1995.

SILVERMAN, B. W. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: London, 1986.

TAYLOR, S. *Modeling Financial Time Series*. New York: John Wiley and Sons, 1986.

TSAY, R. S. *Analysis of Financial Time Series*. New York: John Wiley and Sons, 2002.

WAND, M. P.; JONES, M. C. *Kernel Smoothing*. Chapman and Hall: London, 1995.

WATSON, G. S. Smooth Regression Analysis. *Sankhya Series A*, v. 26, p. 101-116, 1964.

YANG, L.; HÄRDLE, W.; NIELSEN, J. Nonparametric Autoregression with Multiplicative Volatility and Additive Mean. *Journal of Time Series Analysis*, v. 20, p. 579-604, 1999.

ZAKOIAN, J. M. Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, v. 18, p. 931-955, 1994.

ZIEGELMANN, F. A. *Estimation of Volatility Functions: Nonparametric and Semi-Parametric Methods*. PhD thesis - Department of Statistics, University of Kent at Canterbury (UK), 2002a.

ZIEGELMANN, F. A. Estimation of Volatility Functions: The Local Exponential Estimator. *Econometric Theory*, v. 18, n. 4, p. 985-992, 2002b.

ZIEGELMANN, F. A. A Local Linear Least-Absolute-Deviations Estimator of Volatility. *Communications in Statistics-Simulation and Computation*, to appear, 2008.

ZIVOT, E.; WANG, J. *Modelling Financial Time Series with S-PLUS*. New York: Springer Science, 2002.

## APÊNDICE A – Características dos dados

TABELA 1  
Estatísticas descritivas dos retornos diários do Ibovespa

	03.01.1995 – 31.08.2001	03.01.1995 – 30.08.2002
Média	0,000663	0,000466
Mediana	0,001342	0,001070
Desvio padrão	0,027950	0,027288
Assimetria	0,745545	0,690969
Curtose	16,02324	15,58598
Máximo	0,288325	0,288325
Mínimo	-0,172082	-0,172082
Observações	1650	1893
Jarque-Bera	11.813,18 (0,000)	12.644,97 (0,000)
$Q_1(1)$	3,98 (0,046)	5,87 (0,015)
$Q_1(6)$	31,64 (0,000)	32,46 (0,000)
$Q_1(36)$	86,81 (0,000)	84,94 (0,000)
$Q_2(1)$	66,18 (0,000)	76,49 (0,000)
$Q_2(6)$	262,57 (0,000)	301,34 (0,000)
$Q_2(36)$	409,12 (0,000)	475,26 (0,000)
ADF	-38,66 (0,000)	-41,15 (0,000)
Phillips-Perron	-38,62 (0,000)	-41,09 (0,000)
	03.01.1995 – 29.06.2007	03.01.1995 – 28.12.2007
Média	0,000821	0,000840
Mediana	0,001406	0,001438
Desvio padrão	0,023675	0,023520
Assimetria	0,593462	0,572547
Curtose	17,14970	16,98146
Máximo	0,288325	0,288325
Mínimo	-0,172082	-0,172082
Observações	3090	3212
Jarque-Bera	25.958,95 (0,000)	26.337,37 (0,000)
$Q_1(1)$	7,67 (0,006)	7,42 (0,006)
$Q_1(6)$	39,85 (0,000)	39,55 (0,000)
$Q_1(36)$	104,20 (0,000)	103,73 (0,000)
$Q_2(1)$	131,75 (0,000)	136,32 (0,000)
$Q_2(6)$	534,47 (0,000)	555,06 (0,000)
$Q_2(36)$	905,52 (0,000)	944,52 (0,000)
ADF	-52,84 (0,000)	-53,97 (0,000)
Phillips-Perron	-52,79 (0,000)	-53,93 (0,000)

Fonte: Dados da Economatica.

Nota (1): Os testes ADF e PP também rejeitam a hipótese nula na presença de um termo constante e tendência linear.

Nota (2): Números em parênteses indicam p-valores.  $Q_1(k)$  e  $Q_2(k)$  referem-se às estatísticas de Ljung-Box nas séries dos retornos e retornos ao quadrado com  $k$  defasagens, respectivamente.



TABELA 2  
Estatísticas descritivas dos retornos diários do S&P500

	04.01.1995 – 31.08.2001	04.01.1995 – 30.08.2002
Média	0,000537	0,000358
Mediana	0,000683	0,000445
Desvio padrão	0,011114	0,011698
Assimetria	-0,281816	-0,167485
Curtose	6,836844	6,346719
Máximo	0,049887	0,055744
Mínimo	-0,071127	-0,071127
Observações	1683	1930
Jarque-Bera	1.054,61 (0,000)	909,73 (0,000)
$Q_1(1)$	0,04 (0,840)	0,50 (0,479)
$Q_1(6)$	12,22 (0,057)	12,73 (0,048)
$Q_1(36)$	73,14 (0,000)	59,59 (0,008)
$Q_2(1)$	68,68 (0,000)	70,17 (0,000)
$Q_2(6)$	177,15 (0,000)	273,38 (0,000)
$Q_2(36)$	404,77 (0,000)	541,29 (0,000)
ADF	-41,12 (0,000)	-43,17 (0,000)
Phillips-Perron	-41,37 (0,000)	-43,30 (0,000)
	04.01.1995 – 29.06.2007	04.01.1995 – 31.12.2007
Média	0,000377	0,000355
Mediana	0,000654	0,000654
Desvio padrão	0,010671	0,010739
Assimetria	-0,126805	-0,135562
Curtose	6,650646	6,443363
Máximo	0,055744	0,055744
Mínimo	-0,071127	-0,071127
Observações	3145	3272
Jarque-Bera	1.754,85 (0,000)	1.626,49 (0,000)
$Q_1(1)$	0,75 (0,387)	1,99 (0,158)
$Q_1(6)$	13,12 (0,041)	12,33 (0,055)
$Q_1(36)$	70,47 (0,001)	69,90 (0,001)
$Q_2(1)$	126,24 (0,000)	123,51 (0,000)
$Q_2(6)$	595,19 (0,000)	593,66 (0,000)
$Q_2(36)$	1.651,7 (0,000)	1.646,10 (0,000)
ADF	-56,86 (0,000)	-58,55 (0,000)
Phillips-Perron	-57,17 (0,000)	-58,95 (0,000)

Fonte: Dados da Economatica.

Nota (1): Os testes ADF e PP também rejeitam a hipótese nula na presença de um termo constante e tendência linear.

Nota (2): Números em parênteses indicam p-valores.  $Q_1(k)$  e  $Q_2(k)$  referem-se às estatísticas de Ljung-Box nas séries dos retornos e retornos ao quadrado com  $k$  defasagens, respectivamente.

## APÊNDICE B – Intervenção em observações extremas do Ibovespa

TABELA 3  
Observações extremas excluídas da série completa dos retornos do Ibovespa

Data do retorno	Mínimas	Máximas
23/02/1995	..	0,113426
10/03/1995	..	0,228116
14/03/1995	..	0,121814
27/10/1997	-0,162137	..
12/11/1997	-0,107646	..
27/08/1998	-0,104786	..
10/09/1998	-0,172082	..
11/09/1998	..	0,125571
15/09/1998	..	0,171289
23/09/1998	..	0,104140
14/01/1999	-0,105024	..
15/01/1999	..	0,288325

Fonte: Dados da Economatica.

### APÊNDICE C – Equação da esperança condicional do Ibovespa

TABELA 4  
Estimativas do modelo para a média condicional do Ibovespa

	AR(1)	AR(2)
$\phi_1$	0,0787 (0,000)	0,0807 (0,000)
$\phi_2$	..	-0,0379 (0,032)
<i>AIC</i>	-15.528,21	-15.530,80
<i>Q</i> (2)	4,85 (0,089)	..
<i>Q</i> (3)	4,86 (0,183)	0,03 (0,998)
<i>Q</i> (4)	7,61 (0,107)	3,72 (0,446)
<i>Q</i> (5)	12,38 (0,030)	9,00 (0,109)
<i>Q</i> (6)	12,73 (0,048)	9,68 (0,134)
<i>Q</i> (7)	12,74 (0,079)	9,80 (0,201)
<i>Q</i> (8)	14,70 (0,065)	12,28 (0,139)
<i>Q</i> (9)	20,88 (0,013)	18,51 (0,030)
<i>Q</i> (10)	23,53 (0,009)	21,59 (0,018)
<i>Q</i> (11)	23,92 (0,013)	21,99 (0,025)
<i>Q</i> (12)	24,64 (0,016)	22,74 (0,030)
<i>Q</i> (25)	40,55 (0,026)	39,51 (0,035)

Fonte: Dados da Economatica.

Nota (1): O modelo selecionado  $AR(2)$  segue a forma,  $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$ .

Nota (2): Números em parênteses indicam p-valores.  $Q(k)$  refere-se às estatísticas de Ljung-Box na série dos resíduos, com  $k$  defasagens.

## APÊNDICE D – Intervenção em observações extremas do S&P500

TABELA 7  
Observações extremas excluídas da série completa dos retornos do S&P500

Data do retorno	Mínimas	Máximas
27/10/1997	-0,071127	..
28/10/1997	..	0,049887
31/08/1998	-0,070438	..
08/09/1998	..	0,049646
14/04/2000	-0,060045	..
17/09/2001	-0,050468	..
24/07/2002	..	0,055744
29/07/2002	..	0,052667
$Q_1(1)$		0,19 (0,666)
$Q_1(6)$		11,54 (0,073)
$Q_1(36)$		62,65 (0,004)
$Q_2(1)$		76,46 (0,000)
$Q_2(6)$		552,36 (0,000)
$Q_2(36)$		2.363,69 (0,000)

Fonte: Dados da Economatica.

Nota (1): Números em parênteses indicam p-valores.  $Q_1(k)$  e  $Q_2(k)$  referem-se às estatísticas de Ljung-Box nas séries dos retornos e retornos ao quadrado (sem *outliers*) com  $k$  defasagens, respectivamente.

**APÊNDICE E – Intervalos de confiança com base nas estimativas um passo à frente da volatilidade do Ibovespa**

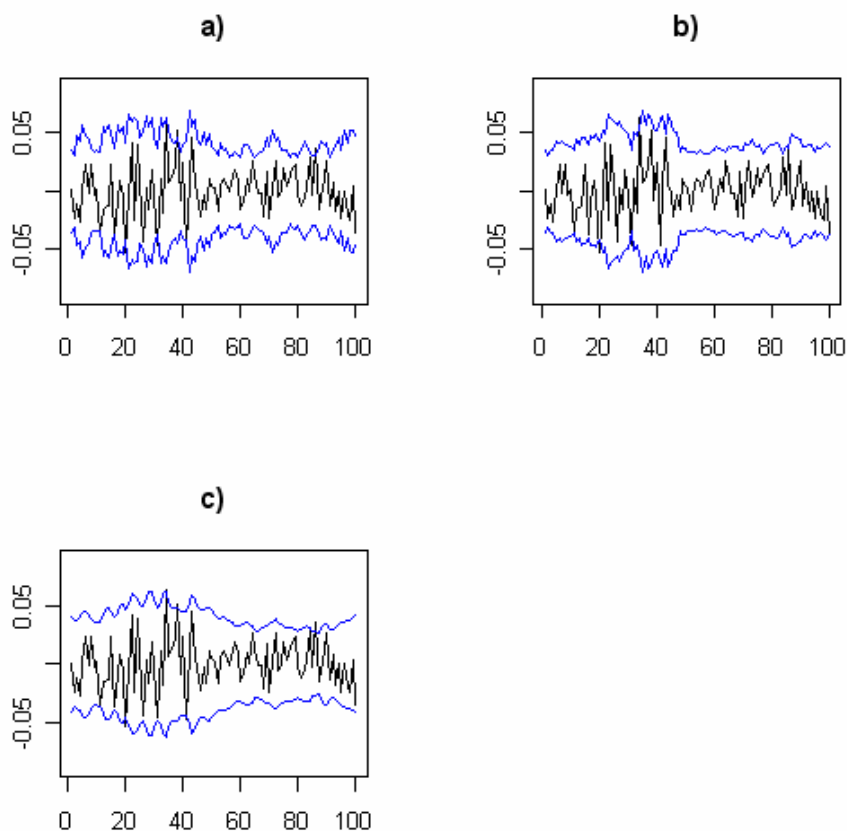


GRÁFICO 27 – Intervalos de confiança um passo à frente em 2002 - Ibovespa

a) Estimativas do Modelo aditivo (4)- $span=0,7$

b) Estimativas do Modelo  $ARCH(4) t(\cdot)$

c) Estimativas do Modelo  $TARCH(3,2) t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Resíduos do  $AR(2)$  e estimativas de 02/09/2002 até 24/01/2003.

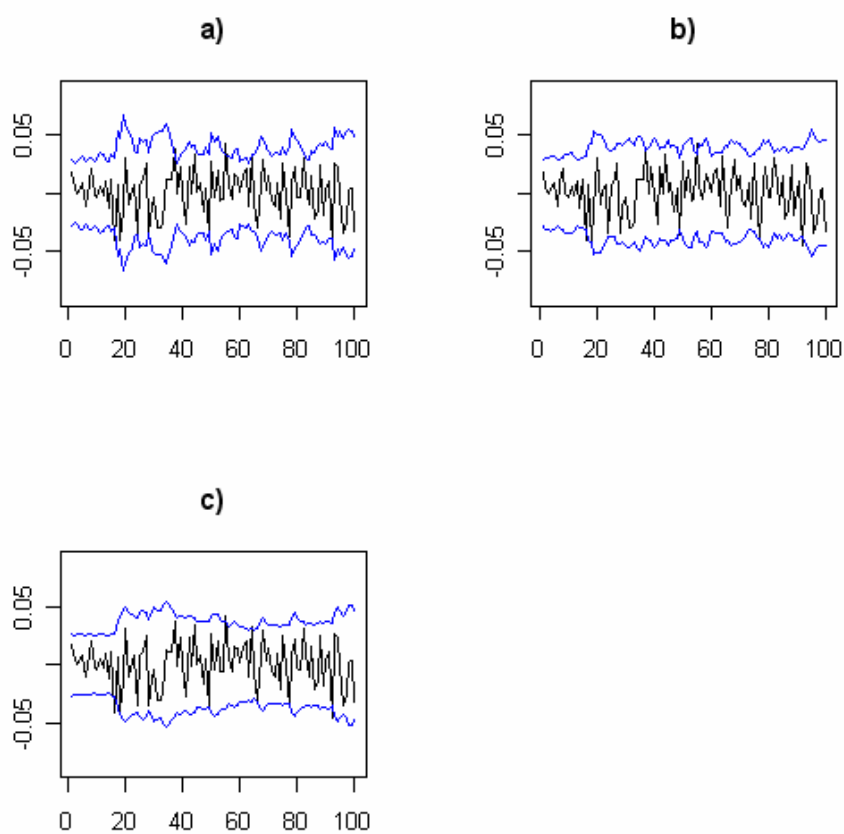


GRÁFICO 28 – Intervalos de confiança um passo à frente em 2007 - Ibovespa

a) Estimativas do Modelo aditivo (4)- $span=0,7$

b) Estimativas do Modelo  $ARCH(4) t(\cdot)$

c) Estimativas do Modelo  $TARCH(3,2) t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Resíduos do  $AR(2)$  e estimativas de 02/07/2007 até 26/11/2007.

**APÊNDICE F – Intervalos de confiança com base nas previsões um passo à frente da volatilidade do Ibovespa**

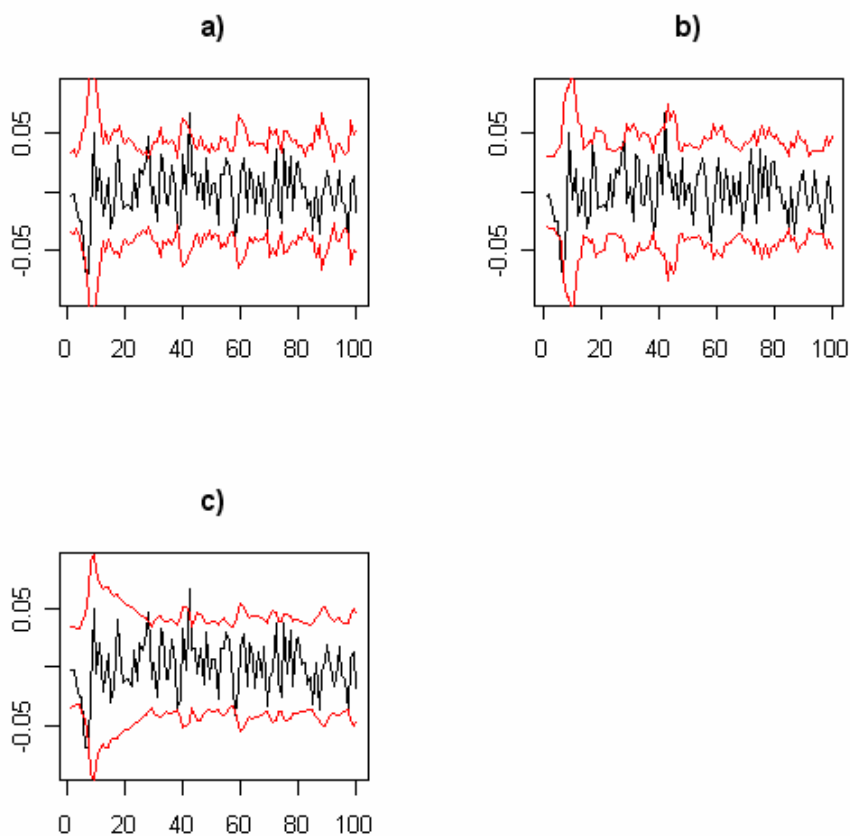


GRÁFICO 29 – Intervalos de confiança um passo à frente em 2001 - Ibovespa

a) Previsões do Modelo aditivo (4)- $span=0,7$

b) Previsões do Modelo  $ARCH(4) t(\cdot)$

c) Previsões do Modelo  $TARCH(3,2) t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Resíduos do  $AR(2)$  e previsões de 03/09/2001 até 04/02/2002.

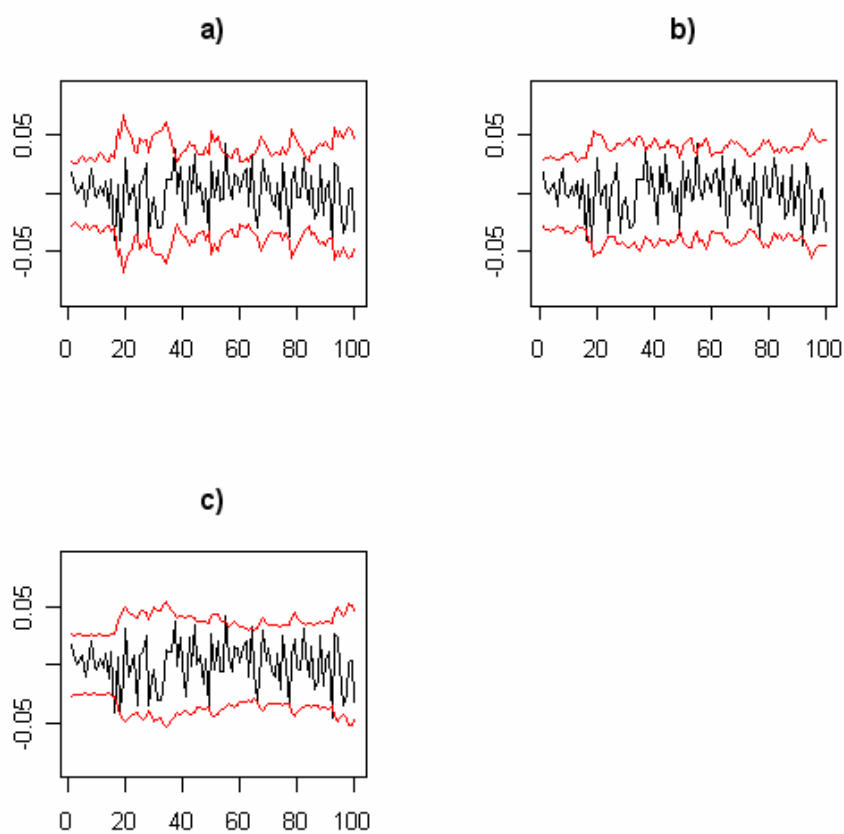


GRÁFICO 31 – Intervalos de confiança um passo à frente em 2007 - Ibovespa

a) Previsões do Modelo aditivo (4)- $span=0,7$

b) Previsões do Modelo  $ARCH(4)$   $t(\cdot)$

c) Previsões do Modelo  $TARCH(3,2)$   $t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Resíduos do  $AR(2)$  e previsões de 02/07/2007 até 26/11/2007.



**APÊNDICE G – Intervalos de confiança com base nas estimativas um passo à frente da volatilidade do S&P500**

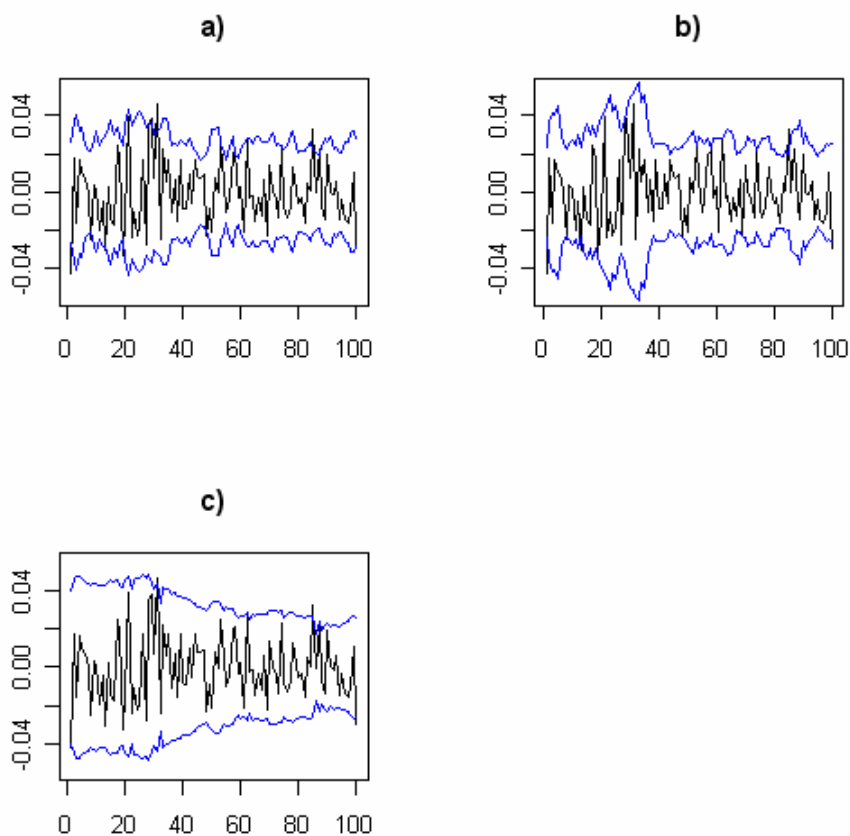


GRÁFICO 33 – Intervalos de confiança um passo à frente em 2002 - S&P500

a) Estimativas do Modelo aditivo (4)- $span=0,8$

b) Estimativas do Modelo  $ARCH(4) N(\cdot)$

c) Estimativas do Modelo  $TARCH(1,2) t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Retornos do S&P500 e estimativas de 03/09/2002 até 24/01/2003.

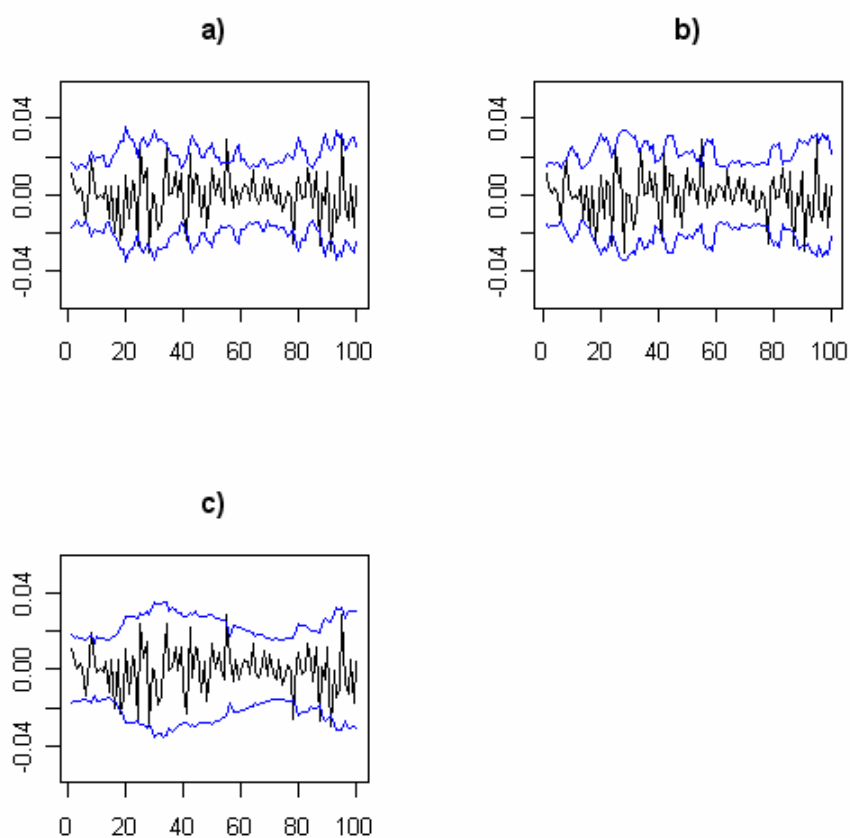


GRÁFICO 34 – Intervalos de confiança um passo à frente em 2007 - S&P500

a) Estimativas do Modelo aditivo (4)- $span=0,8$

b) Estimativas do Modelo  $ARCH(4)$   $N(\cdot)$

c) Estimativas do Modelo  $TARCH(1,2)$   $t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Retornos do S&P500 e estimativas de 02/07/2007 até 20/11/2007.

**APÊNDICE H – Intervalos de confiança com base nas previsões um passo à frente da volatilidade do S&P500**

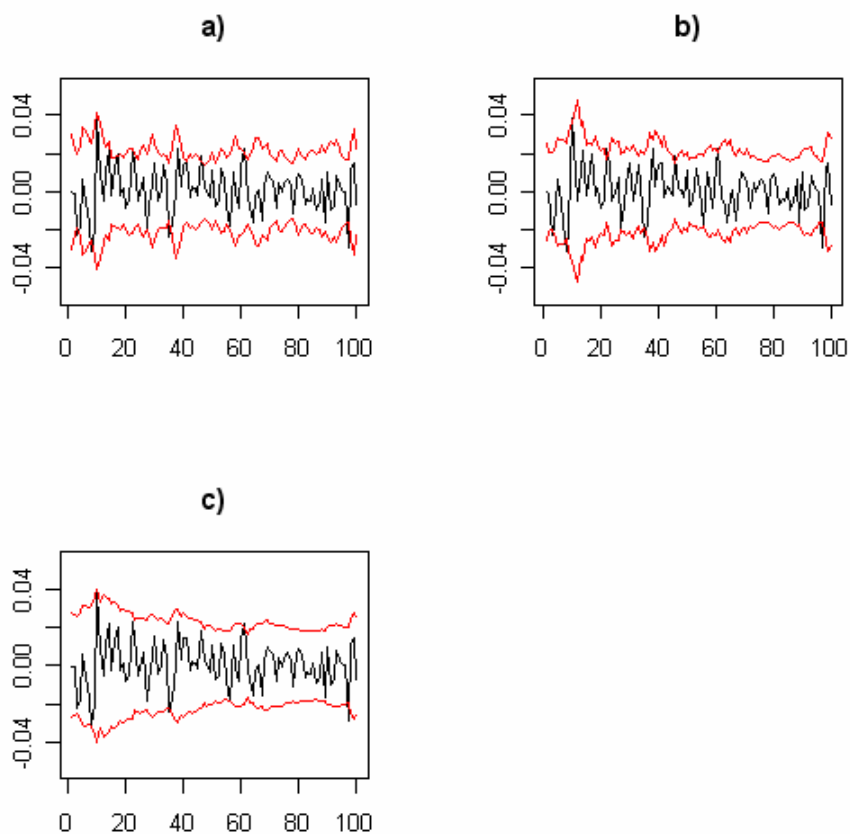


GRÁFICO 35 – Intervalos de confiança um passo à frente em 2001 - S&P500

a) Previsões do Modelo aditivo (4)- $span=0,8$

b) Previsões do Modelo  $ARCH(4) N(\cdot)$

c) Previsões do Modelo  $TARCH(1,2) t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Retornos do S&P500 e previsões de 04/09/2001 até 01/02/2002.

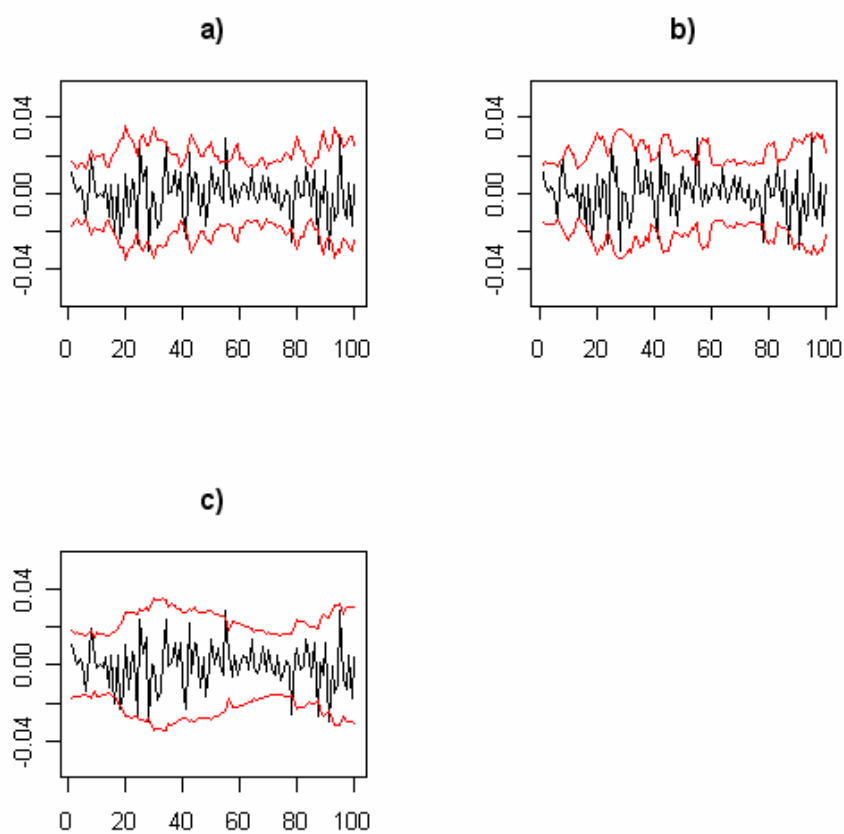


GRÁFICO 37 – Intervalos de confiança um passo à frente em 2007 - S&P500

a) Previsões do Modelo aditivo (4)- $span=0,8$

b) Previsões do Modelo  $ARCH(4)$   $N(\cdot)$

c) Previsões do Modelo  $TARCH(1,2)$   $t(\cdot)$

Fonte: Dados da Economatica.

Nota(1): Retornos do S&P500 e previsões de 02/07/2007 até 20/11/2007.