

3 Estimações

O objetivo deste capítulo é analisar como os modelos com mudanças estruturais considerados descrevem as séries de volatilidade realizada das ações do Dow Jones. Existem quebras estruturais e/ou múltiplos regimes estatisticamente significantes? Quais são as magnitudes e durações das possíveis alterações? As mudanças de níveis são relevantes economicamente? Diante da possibilidade de estimação de quebras estruturais espúrias, os resultados fazem sentido quando lidas em um contexto mais amplo? O que as estimações de modelos com mudanças estruturais dizem sobre o mercado de ações no período analisado? Quais são as virtudes de ajuste desses modelos em relação à alternativa de integração fracional? Até que ponto modelos *ARFIMA* são consistentes com virtuais inflexões na dinâmica da volatilidade?

3.1 Dados

A análise empírica é concentrada na volatilidade realizada de dezesseis ações atualmente pertencentes ao índice de ações DJIA (Dow Jones Industrial Average): Alcoa, American International Group, Boeing, Caterpillar, General Electric, General Motors, Hewlett Packard, IBM, Intel, Johnson and Johnson, Coca-Cola, Microsoft, Merck, Pfizer, Wal-Mart e Exxon. Os dados intradiários originais são compostos de cotações extraídas da base de dados *Trade and Quote* (TAQ) da bolsa de Nova York para o período entre 3 de janeiro de 1994 e 31 de dezembro de 2003. São excluídos dias de negócios com volumes extraordinariamente baixos causados pela proximidade de feriados (e.g., *Good Friday*), restando 2476 dias. Tendo em vista a concisão na exposição dos resultados no texto principal, as seções seguintes têm como enfoque as ações da IBM, enquanto os demais papéis são contemplados detalhadamente pelo apêndice D.

Inicialmente, são removidas cotações fora do padrão, computados preços através da média para compra e venda, filtrados possíveis erros e obtidos retornos segundo a segundo para o intervalo entre 9:30 e 16:05, diariamente. Observando as considerações de consistência discutidas por Hansen e Lunde

(2006b), o método de última cotação é adotado para a determinação de preços nos segundos exatos. Com base nos resultados de Hasbrouck (1995), que reporta uma parcela mediana de informação de 92,7% para a NYSE em termos de ações do Dow Jones, e Blume e Goldstein (1997), que concluem que as cotações da NYSE correspondem à melhor cotação nacional na maior parte do tempo, foram privilegiadas cotações da bolsa de Nova York (ou NASDAQ, para a Microsoft e a Intel) quando as mesmas se localizassem suficientemente perto da hora desejada em relação às demais atualizações.

A busca por métodos não-viesados, consistentes e eficientes para estimação de volatilidade realizada na presença de ruídos de microestrutura tem sido um dos tópicos mais intensos em econometria financeira nos últimos anos. Embora referências de vanguarda tais como Andersen et al. (2001) advogassem a simples seleção de uma frequência arbitrária (tipicamente 5-15 minutos) com base numa intuição de balanceamento entre precisão e dissipação do viés de microestrutura, procedimento denominado *amostragem espaçada*, diversos artigos recentes desenvolveram estimadores que dominam essa estratégia. As contribuições podem ser agrupadas em diversas categorias: seleção da frequência ótima para amostragem espaçada (e.g., Bandi e Russell, 2005), subamostragem (Zhang, Mykland e Ait-Sahälia, 2005; Ait-Sahälia, Mykland e Zhang, 2005), estimações baseadas em kernel (Zhou, 1996; Barndorff-Nielsen et al., 2005; Hansen e Lunde, 2006b) e filtragem MA (e.g., Hansen, Large e Lunde, 2006), dentre outros. Uma revisão da literatura e comparação dos métodos está disponível em McAleer e Medeiros (2006).

Exemplos de estimadores consistentes disponíveis atualmente são o estimador de kernel modificado (Barndorff-Nielsen et al., 2005), o filtro MA modificado de Hansen, Large e Lunde (2006) e o estimador em duas escalas de tempo de Zhang, Mykland e Ait-Sahälia (2005), o escolhido para este trabalho. Ait-Sahälia, Mykland e Zhang (2005) mostram que o estimador funciona adequadamente quando a hipótese de ruídos de microestrutura independentes é violada, é estável na escolha das escalas de tempo (neste estudo são aplicados *grids* de cinco minutos) e produz estimativas próximas ao método mais eficiente –porém mais exigente computacionalmente– de múltiplas escalas (Zhang, 2005). A variável dependente final, como usual, é dada pelo logaritmo da volatilidade realizada diária.¹

3.2

¹Assim dispensando restrições nas estimações de séries de tempo que eliminam a possibilidade de volatilidades negativas.

Especificação

O ciclo de modelagem é baseado no princípio do específico-ao-geral, sendo iniciado a partir do nó original (nível 0). A equação geral é dada por:

$$\begin{aligned} \log(RV_t) = & \varpi + \alpha_1 \log(RV_{t-1}) + \dots + \alpha_k \log(RV_{t-k}) + \delta_1 I[Seg] \\ & + \delta_2 I[Ter] + \delta_3 I[Qua] + \delta_4 I[Qui] + \delta_5 I[FOMC] + \delta_6 I[EMP] \\ & + \delta_7 I[CPI] + \delta_8 I[PPI] + \varepsilon_t \end{aligned} \quad (3-1)$$

onde $\log(RV_t)$ é o logaritmo de volatilidade realizada diária, $I[Seg]$, $I[Ter]$, $I[Qua]$ e $I[Qui]$ são *dummies* incluídas para capturar possíveis estruturas de autocorrelação remanescentes causadas por efeitos de dias da semana (ver Martens, van Dijk e de Pooter, 2004, e outras referências no próprio artigo) e $I[FOMC]$, $I[EMP]$, $I[CPI]$ e $I[PPI]$ são *dummies* que indicam datas para os seguintes anúncios macroeconômicos: taxa de juros (reuniões do Federal Open Market Committee), desemprego (*Employment Situation*, relatório mensal do Bureau of Labor Statistics) e índices de preço aos consumidores (CPI) e produtores (PPI). Diversos autores discutem a relação entre a divulgação dessas informações e possíveis mudanças e saltos na volatilidade dos mercados financeiros americanos (e.g., Barndorff-Nielsen e Shephard, 2006; Huang, 2006).

O primeiro passo no ciclo de modelagem consiste no uso dessa equação para a seleção do número de defasagens na estrutura autoregressiva e dos efeitos de dias da semana e anúncios relevantes, levando a uma especificação primária a ser contraposta à hipótese de não-linearidade/múltiplos regimes. Os coeficientes autoregressivos examinados vão até a 15^a defasagem. Em busca de uma destinação parcimoniosa, a escolha é feita pelo Critério de Informação de Schwarz (SBIC), que seleciona inicialmente as defasagens 1-3, 5, 7,10, as variáveis indicadoras para segundas e sextas-feiras e os anúncios de taxas de juros e relatório da situação de desemprego² para todos os papéis. Apesar da possibilidade de abreviação do número de coeficientes AR via inclusão de um termo MA, a especificação AR com mais variáveis é favorecida por duas razões: o alto custo computacional da alternativa sob a delineação não-linear e disponibilidade de suficientes graus de liberdade. A presença de um coeficiente MA poderia ser justificada por erros de medida ou componentes de saltos³ na

²Possíveis alterações na volatilidade nos períodos pré-anúncios também foram testadas, mas foram irrelevantes. Por outro lado, os coeficientes para os dois dias seguintes às divulgações de taxas de juros e desemprego tenderam a ser levemente negativo e estatisticamente significantes para algumas ações

³Ver Andersen, Bollerslev e Diebold (2005a), Tauchen e Zhou (2005) e Barndorff-Nielsen e Shephard (2006).

volatilidade.

A etapa subsequente consiste na definição do conjunto de variáveis de transição para cada nó e de quais variáveis explicativas da equação inicial farão parte da estrutura da árvore (i.e., são afetadas por múltiplos regimes). Nas próximas seções, as variáveis de transição candidatas caem em duas aplicações: quebras estruturais (o tempo é a única variável de transição) e efeitos de assimetria (defasagens de retornos acumulados). Como o foco está na avaliação das qualidades de modelagem de diferentes modelos capazes de gerar autocorrelações persistentes, não serão analisadas combinações de quebras estruturais e efeitos de assimetria. Uma terceira possibilidade explorada por Martens, van Dijk e de Pooter (2004) seria o uso da própria volatilidade realizada como variável de transição, alternativa que não se mostrou significativa em nenhum caso; o uso de volatilidades acumuladas, por outro lado, é significativo mas se assemelha ao modelo HAR-RV, examinado posteriormente.

Se mais de uma variável de transição candidata é significativa em um algum nó, aquela com o menor p-valor no teste LM é selecionada, a não ser quando especificado contrariamente. Um nó é estimado se alguma variável de transição é significativa a $(1/2^{\text{nível}})\%$. Após uma análise inicial abrangente, uma opção mais conservadora envolvendo uma estrutura em árvore somente para a constante foi adotada no caso do modelo com assimetria; efeitos não-lineares nos termos AR, apesar de estatisticamente significantes, não trouxeram ganhos fora da amostra, além de aumentarem excessivamente a quantidade de parâmetros estimados e ocasionalmente causarem problemas numéricos nas estimações. Por razões semelhantes, os modelos com quebras estruturais são estimados com não-linearidades na constante e nos coeficientes autoregressivos de primeira e segunda ordem.

Se a hipótese de linearidade é preterida, retorna-se à expressão inicial para a estimação de modelos lineares aumentados pela expansão de Taylor, com o propósito de checar se a estrutura não-linear afeta a especificação definida anteriormente pelo critério de Schwarz. No caso da IBM, a especificação básica permanece inalterada. O ciclo de modelagem pode então ser resumido da seguinte maneira:

1. Seleção do conjunto relevante de variáveis de transição para cada nível da árvore.
2. Estimações de modelos AR aumentados por *dummies* de dias da semana para determinação do número de defasagens e outros efeitos relevantes pelo SBIC.

3. Testes LM para não-linearidade em todas as variáveis de transição relevantes no primeiro nó.

Caso significativo:

4. Seleção da variável candidata como o menor p-valor e revisão do número de defasagens AR e outros efeitos relevantes através de estimações de modelos lineares aumentados pela expansão de Taylor do teste.
5. Especificação seqüencial do modelo através dos testes LM e seleção de variáveis de transição pelo critério do menor p-valor.
6. Estimação do modelo final.
7. Diagnósticos.

3.3

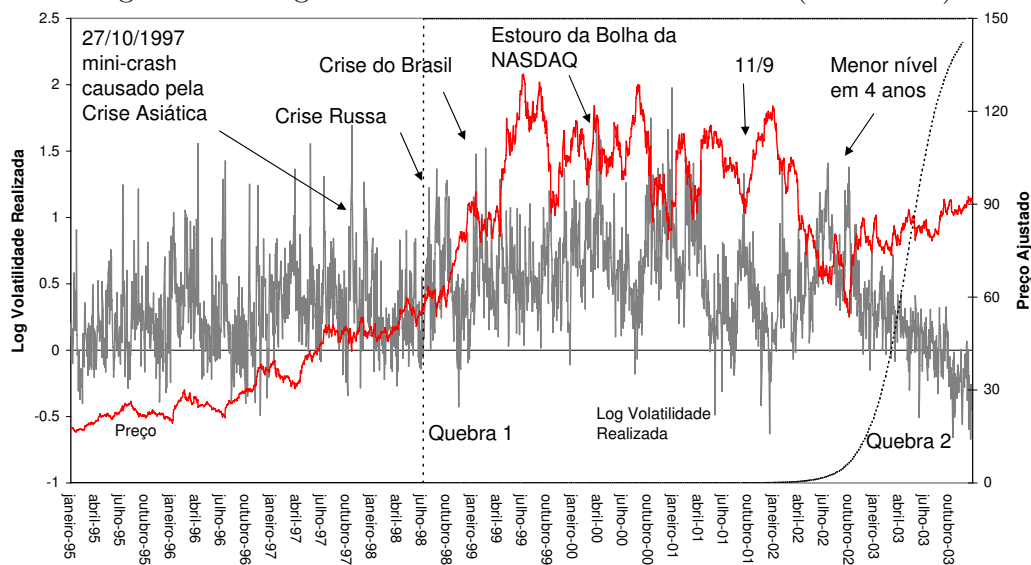
Quebras-estruturais

Seguindo a literatura recente (e.g, Granger e Hyung, 2004; Martens, van Dijk e de Pooter, 2004; Morana e Beltratti, 2004), esta seção investiga os efeitos de possíveis quebras estruturais sobre os níveis de volatilidade. Um modelo em árvore é estimado tendo o tempo como variável de transição. Os resultados são apresentados na Tabela 3.1.

Duas quebras são estimadas: uma em agosto de 1998 (volatilidade e persistência aumentam; média da volatilidade realizada diária vai de 1,50% a 2,10%, alta de 40%) e outra em abril de 2003 (volatilidade declina acentuadamente; média cai de 2,10% para 1,15%, queda de 45%). Embora os erros dos parâmetros de inclinação pareçam indicar que as quebras não são significantes, o problema de identificação discutido na seção 2.3 tem como consequência o fato que sob a hipótese nula $H_0 : \gamma = 0$ a estatística-t usual não segue a distribuição padrão –os p-valores dos testes LM são as estatísticas adequadas para avaliar a importância das mudanças estruturais. Em geral, é muito difícil obter estimativas precisas para essas variáveis; uma razão para isso é que altas gamas implicam que as derivadas da função de transição se aproximam de funções degeneradas, encolhendo as estatísticas-t dos respectivos parâmetros. Ver a discussão em Eitrheim e Terasvirta (1996).

A Figura 3.1 contextualiza as quebras através do tempo, plotando as duas funções de transição $G(t; \gamma, c)$, o logaritmo da volatilidade realizada e a evolução do preço da ação (ajustado para proventos) no período 1995-2003. A primeira mudança estrutural coincide precisamente com a crise russa de 1998, enquanto a posterior limita duas trajetórias distintas para o índice Dow Jones: enquanto o índice viria a alcançar o menor nível em quatro anos no

Figura 3.1: Log da Volatilidade Realizada da IBM (1995-2003)



mês de outubro de 2002, o ano seguinte mostrou-se altamente positivo, com alta de cerca de 25%. A figura é sugestiva de outras relações semelhantes: há diversos agregados de alta volatilidade associados com diminuições expressivas no preço da ação, seguidos de recuos acentuados na volatilidade realizada uma vez que os preços retomaram patamares superiores; alguns exemplos são os períodos do mini-crash de outubro de 1997, a crise da Rússia em 1998, o rompimento da bolha da NASDAQ, os aglomerados no fim de 2000 e no começo de 2001, o período do 11/9 e o mercado deprimido de 2002. A subamostra entre a primeira quebra estrutural e a segunda – o período de alta volatilidade histórica – é marcada pela incidência mais intensa de episódios de queda de preços. A próxima seção investiga relações dessa natureza.

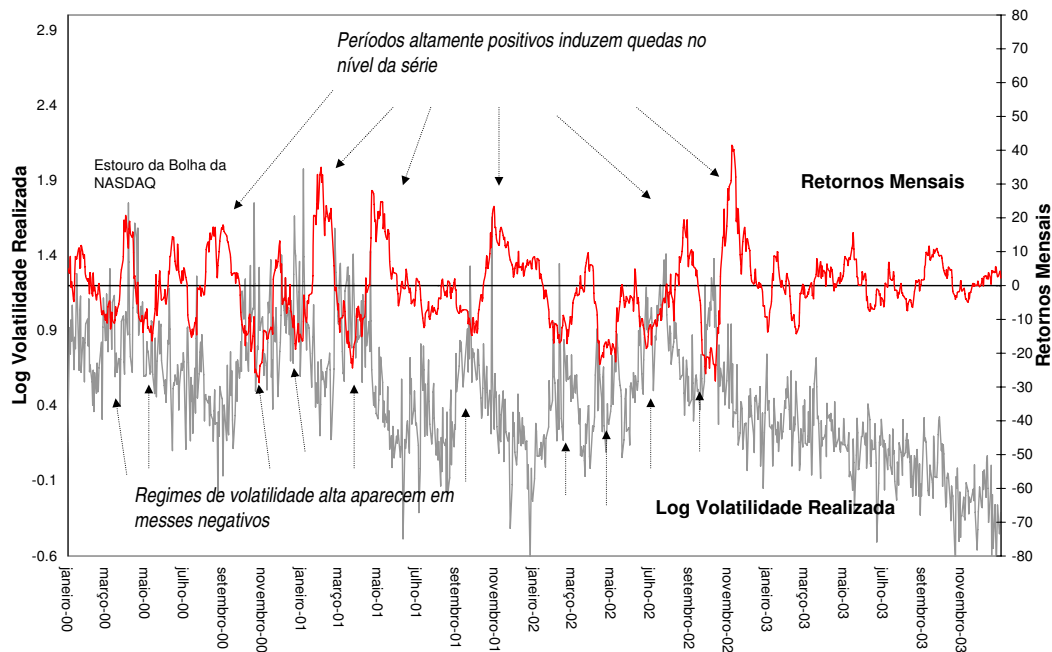
3.4

Efeitos de Assimetria

A motivação para estimação do modelo em árvores com retornos acumulados defasados como fonte de múltiplos regimes na volatilidade é ilustrada pela Figura 3.2, que exhibe a trajetória da volatilidade realizada e os retornos mensais da IBM para o período entre 2000 e 2003. As setas indicam a relação a ser estimada: parece haver um padrão recorrente de transições para níveis superiores de volatilidade relacionadas com quedas nos preços da ação – a reversão a patamares mais baixos ocorre em meses com retornos positivos. A única exceção expressiva ocorre no período imediatamente anterior ao estouro da bolha da Nasdaq.

Para conferir maior interpretabilidade ao modelo de regressão em árvore

Figura 3.2: Log da Volatilidade Realizada da IBM e Retornos Mensais (2000-2003)



estimado na presença de vários horizontes de retornos altamente significantes, a estimação é organizada de forma que os retornos acumulados testados nos dois primeiros nós correspondem a períodos superiores a um mês, enquanto em níveis superiores todas as possibilidades são consideradas. Entretanto, se o retorno do dia anterior for imposto como variável de transição no primeiro nó –o efeito de alavancagem usual– e então efeitos não-lineares adicionais de retornos acumulados forem testados em nós superiores, os resultados permanecem inalterados, com efeitos de assimetria prolongados surgindo em outros nós.

Os coeficientes estimados para o modelo em árvore com efeitos de assimetria (doravante também abreviado como AE) estão detalhados na Tabela 3.2. O modelo é descrito por cinco regimes altamente significantes estatisticamente, introduzindo quatro níveis de efeitos de assimetria. O primeiro nó aponta para um regime de baixa volatilidade e ausência de efeitos de alavancagem de curto prazo, associado com um mercado em alta num horizonte de quatro meses; no sentido contrário, quedas de 12% em aproximadamente dois meses introduzem o regime de maior volatilidade, enquanto retornos maiores nesse intervalo proporcionam um regime de volatilidade intermediária e efeitos de alavancagem diante de retornos negativos em poucos dias. Embora as constantes nos regimes dois e três sejam pequenas, parecendo indicar uma volatilidade menor do que nos intervalos de alta sugeridos pelo primeiro nó, os baixos gamas,

em conjunto com os valores dos parâmetros de locação, implicam num efeito mais abrangente que envolve uma interação entre as transições III-IV: há um aumento de volatilidade para retornos levemente negativos que se torna mais expressivo quando estes últimos se tornam ainda mais negativos, elevando a média. A distinção entre regimes nesses nós, portanto, não é tão clara. As correlações negativas entre as funções de transição e o logaritmo da volatilidade realizada sublinham a importância de componentes linear, indicando que o uso de termos desse tipo seria preferível à inclusão de *dummies*, procedimento mais comum na literatura.

Com base nos regimes estimados da Tabela 3.2 e nos gráficos das funções de transição apresentados no Apêndice A, as observações são divididas de acordo com a seguinte classificação em grupos: o primeiro reúne os dias tais que os retornos nos últimos noventa dias de negócios foram superiores a 9% (isto é, a função de transição da primeira divisão esteve acima 0,5), o segundo representa o regime de volatilidade intermediária sem efeitos de alavancagem (dias mais próximos dos regimes dois e três), o terceiro representa um regime de volatilidade intermediária com efeitos de alavancagem no curto prazo (mais próximo do regime quatro) e o último corresponde a regimes de grandes quedas de preços em dois meses.

A Tabela 3.3 abaixo reporta o número de observações classificadas em cada grupo e as respectivas médias e desvios-padrões. A correlação entre a classificação e a volatilidade realizada é 0,37. Os grupos dois, três e quatro representam sucessivamente níveis de volatilidade 11%, 20% e 57% maiores em relação ao regime de volatilidade baixa. A tabela também sugere que a dispersão é positivamente relacionada com o próprio nível de volatilidade.

3.5

Comparação Dentro da Amostra

3.5.1

Modelos Alternativos

ARFIMA O modelo $ARFIMA(p, d, q)$ padrão é definido pela seguinte equação

$$\phi(L)(1 - L)^d(\log(RV_t) - \mu) = \theta(L)\varepsilon_t, \quad (3-2)$$

onde d denota o parâmetro de integração fracional, L o operador de defasagem, ε_t é um ruído branco, p e q são as ordens dos polinômios $\phi(L)$ e $\theta(L)$ respectivamente, cujas raízes estão fora do círculo unitário. Para cada ação,

a combinação mais apropriada de parâmetros p e q é obtida através de estimações de diversas especificações $ARFIMA(p, d, q)$ através do método de Whittle (1951) (máxima verossimilhança no domínio da frequência)⁴ e comparações pelo critério de informação de Schwarz. A abordagem conduz à escolha de uma especificação $ARFIMA(0, d, 0)$ (daqui para frente também abreviada como FI) em todas as séries, resultado minuciosamente confirmado pela análise das propriedades das séries filtradas e de performances preditivas. As previsões para o modelo $ARFIMA(0, d, 0)$ são computadas através da truncagem da representação autoregressiva infinita na 150^a defasagem. O modelo FI definitivo é dado por:

$$(1 - L)^d(\log(RV_t) - \mu - \delta_1 I[Seg] - \delta_2 I[Sex] - \delta_3 I[FOMC] - \delta_4 I[EMP]) = \varepsilon_t \quad (3-3)$$

Na discussão que se segue, os modelos $ARFIMA$ são considerados os padrões para séries de volatilidade, tendo sido estimados no contexto de volatilidade realizada em inúmeros artigos tais como Andersen et al. (2003), Areal e Taylor (2002), Beltratti e Morana (2005), Deo, Hurvich e Lu (2006), Martens, van Dijk e de Pooter (2004), Thomakos e Wang (2003), dentre outros.

HAR-RV Os modelos autoregressivos heterogêneos de volatilidade realizada (HAR-RV), propostos por Corsi (2003), são fundamentados no modelo ARCH Heterogêneo (HARCH) desenvolvido por Müller et al. (1997), sendo especificados como modelos de volatilidade de múltiplos componentes em estruturas hierárquicas aditivas. Em outras palavras, a classe HAR-RV modela a volatilidade como uma soma de diversos componentes da própria série em horizontes distintos. O modelo vem ganhando importância na literatura pela facilidade e flexibilidade de estimação (em contraste com a classe $ARFIMA$) e pela capacidade de reproduzir satisfatoriamente os padrões de autocorrelações de modelos com memória longa em horizontes não muito longos (e.g., Andersen, Bollerslev e Diebold, 2005a). Defina a volatilidade realizada normalizada no horizonte h como

$$\log(RV_t)_{t+h} = \frac{\log(RV_{t+1}) + \log(RV_{t+2}) + \dots + \log(RV_{t+h})}{h} \quad (3-4)$$

A especificação HAR-RV é dada por:

⁴Ver também Baillie (1996).

$$\begin{aligned} \log(RV_t) = & \varpi + \alpha_1 \log RV_{t-1} + \alpha_2 \log RV_{t-2} + \alpha_3 \log(RV_t)_{t-5} \\ & + \alpha_4 \log(RV_t)_{t-22} + \delta_1 I[Seg] + \delta_2 I[Sex] \\ & + \delta_3 I[FOMC] + \delta_4 I[EMP] + \varepsilon_t \end{aligned} \quad (3-5)$$

Um termo autoregressivo de segunda ordem é acrescentado na formulação típica com o objetivo de eliminar autocorrelações restantes.

3.5.2 Resultados

A Tabela 3.4 exibe as estatísticas descritivas para os resíduos dos quadro modelos, onde JB representa o teste de normalidade de Jarque-Bera (p-valor), $Q(k)$ indica o p-valor dos testes apropriados para correlação serial até a defasagem k (teste de Ljung-Box para os modelos FI e HAR e um teste do tipo LM para os modelos em árvores; ver a descrição deste último em Medeiros e Veiga, 2003), $Q(k)^2$ fornece os p-valores dos mesmos testes para os quadrados dos resíduos e SBIC é o critério de informação de Schwarz. Os coeficientes de determinação são ajustados de acordo com o procedimento descrito em Andersen, Bollerslev e Meddahi (2005b).⁵

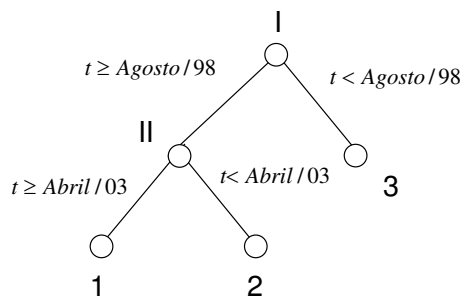
A tabela mostra que o modelo com efeitos de assimetria possui melhor ajuste dentro da amostra como medido pelo R^2 , enquanto o modelo HAR –mais parcimonioso– é superior pelo critério de Schwarz. O modelo com integração fracional exibe ajuste marcadamente inferior aos demais, mesmo pelo critério de Schwarz; em contraste, o modelo com efeitos de assimetria envolve a estimação de mais de dez parâmetros em excesso. Todos os modelos geram resíduos assimétricos e leptocúrticos, o que é explicado por um conjunto de cerca de quarenta *outliers* à direita, provavelmente correspondendo à influência de saltos.

As estatísticas de teste para autocorrelação dos resíduos indicam que os modelos HAR, com efeitos de assimetria e com quebras estruturais não implicam em nenhuma estrutura de correlação serial significativa até a 20^a defasagem (a 5%); por outro lado, o modelo com integração fracional retorna resíduos altamente autocorrelacionados mesmo em defasagens pequenas. Como o resultado poderia ser consequência de termos AR ou MA ignorados, tal hipótese foi examinada através de diversos testes com especificações ARFIMA adicionais; nenhuma delas foi capaz de reverter as autocorrelações significantes (ou superar o modelo $ARFIMA(0, d, 0)$ em previsões dentro e fora da amostra). Razões para essa dissonância são consideradas no capítulo 4. Por fim, há fortes

⁵O estimador de quarticidade utilizado é o mesmo proposto em Andersen, Bollerslev e Diebold (2005a).

evidências de dependência nos quadrados dos resíduos. Entretanto, ao contrário dos resultados de Beltratti e Morana (2005) para séries de volatilidade de taxas de câmbio, não existe memória longa na variância condicional da volatilidade.

Tabela 3.1: Estimação com Quebras Estruturais



Nó	p-valor	c		γ	
I	0,000	1,744	(0,136)	13,359	(6,154)
II	0,000	3,273	(0,101)	7,003	(12,716)

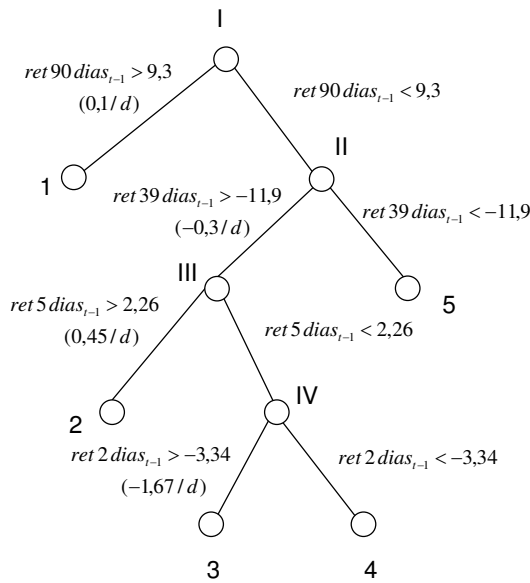
Regime	1	2	3
Const	0,005 (0,048)	0,140 (0,021)	0,118 (0,014)
AR(1)	0,261 (0,164)	0,449 (0,036)	0,409 (0,033)
AR(2)	0,224 (0,078)	0,156 (0,037)	0,080 (0,033)

Outros Coeficientes					
AR(3)	0,084	(0,021)	SEG	-0,064	(0,013)
AR(5)	0,074	(0,020)	SEX	-0,063	(0,014)
AR(7)	0,044	(0,019)	FOMC	0,067	(0,032)
AR(10)	0,047	(0,018)	EMP	0,094	(0,023)

R^2	0,619
-------	-------

A tabela apresenta a estimação do modelo em árvore para o logaritmo da volatilidade realizada dos retornos das ações da IBM entre 3 de janeiro de 1994 e 31 de dezembro de 2003 (2476 observações diárias), tendo o tempo como única variável de transição e permitindo que a mesma afete a constante e os coeficientes autoregressivos de primeira e segunda ordem; a estrutura em árvore é definida por uma seqüência de testes LM discutidos nas seções 2.3 e 3.2. Os erros padrões estão em parênteses. γ e c são reportados para a variável de transição padronizada. A coluna p-valor reporta a estatística LM robusta para a divisão na posição indicada na primeira coluna. SEG, SEX, FOMC e EMP são coeficientes para a *dummies* para segundas-feiras, sextas-feiras, anúncios de taxas de juros e divulgações dos relatórios de desemprego. O R^2 é corrigido segundo Andersen, Bollerslev e Meddahi (2005).

Tabela 3.2: Estimação de Efeitos de Assimetria



Nó	p-valor	c	γ	Correl
I	0,000	0,541 (0,344)	2,000 (1,082)	-0,310
II	0,004	-0,955 (0,319)	2,000 (1,018)	-0,370
III	0,000	0,479 (0,469)	2,000 (1,794)	-0,103
IV	0,000	-1,091 (0,284)	2,423 (1,211)	-0,325

Regime	Const	Outros Coeficientes
1	0,081 (0,013)	AR(1) 0,386 (0,022)
2	-0,004 (0,046)	AR(2) 0,118 (0,023)
3	0,069 (0,044)	AR(3) 0,107 (0,021)
4	0,447 (0,127)	AR(5) 0,091 (0,020)
5	0,184 (0,030)	AR(7) 0,065 (0,019)
		AR(10) 0,078 (0,018)
R^2	0,631	SEG -0,068 (0,012)
		SEX -0,064 (0,014)
		FOMC 0,068 (0,032)
		EMP 0,092 (0,023)

A tabela apresenta a estimação do modelo de regressão em árvores com transições suaves para o logaritmo da volatilidade realizada diária dos retornos das ações da IBM entre 3 de janeiro de 1994 e 31 de dezembro de 2003 (2476 observações), onde os retornos defasados e retornos acumulados defasados são possíveis variáveis de transição determinantes de múltiplos regimes para a constante. A estrutura da árvore é definida por uma seqüência de testes LM discutidos nas seções 2.3 e 3.2; nos dois primeiros nós, a estimação seleciona os vetores de retornos acumulados em períodos superiores a um mês com menores p-valores (caso significantes), enquanto em níveis maiores a regra se estende a todas as variáveis. Os erros padrões estão em parênteses. γ e c são reportados para a variável de transição padronizada. A coluna p-valor reporta a estatística LM robusta para a divisão na posição indicada na primeira coluna. SEG, SEX, FOMC e EMP são coeficientes para a *dummies* para segundas-feiras, sextas-feiras, anúncios de taxas de juros e divulgações dos relatórios de desemprego. *Correl* indica a correlação entre a função de transição $G(x; \gamma, c)$ e a variável dependente. O R^2 é corrigido segundo Andersen, Bollerslev e Meddahi (2005).

Tabela 3.3: Regimes de Volatilidade das Ações da IBM

Grupo	Média	Desv. Pad.	Obs.
1	1,57	0,54	1264
2	1,75	0,71	601
3	1,88	0,81	357
4	2,46	0,82	254
Todos	1,75	0,71	2476

Médias e desvios-padrões das medidas diárias de volatilidade realizada para grupos de observações divididos pela classificação baseada na da Tabela 3.2.

Tabela 3.4: Estatísticas dos Resíduos

	AE	FI	SB	HAR
R^2	0,631	0,505	0,619	0,610
DP	0,223	0,255	0,226	0,229
Assimetria	0,697	0,336	0,725	0,707
Curtose	4,703	4,166	4,535	4,503
JB	0,000	0,000	0,000	0,000
$Q(5)$	0,367	0,000	0,4319	0,637
$Q(10)$	0,115	0,000	0,3077	0,275
$Q(20)$	0,399	0,000	0,4219	0,530
$Q(50)$	0,003	0,000	0,0177	0,027
$Q^2(10)$	0,012	0,000	0,001	0,001
$Q^2(20)$	0,032	0,000	0,008	0,008
SBIC	-2,905	-2,699	-2,889	-2,918

A tabela mostra estatísticas descritivas para os resíduos de quatro modelos estimados para o log da volatilidade realizada da IBM: o modelo em árvore com retornos acumulados passados como variáveis de transição (AE, Tabela 3.2), o modelo em árvore com quebras estruturais (SB, Tabela 3.1), o $ARFIMA(0, d, 0)$ estimado pelo método de Geweke-Porter-Hudak (1983) e o modelo autoregressivo heterogêneo (HAR-RV) de Corsi (2003). JB é o teste de normalidade de Jarque-Bera (p-valor), $Q(k)$ indica o p-valor dos testes apropriados para correlação serial até a defasagem k , $Q^2(k)$ fornece os p-valores dos mesmos testes para os quadrados dos resíduos e SBIC é o critério de informação de Schwarz. O R^2 é corrigido segundo Andersen, Bollerslev e Meddahi (2005).