

## 4

### Memória Longa

Para avaliar as características de memória longa de cada modelo, as estimações do capítulo anterior são tomadas como base para a simulação de mil séries de cada um deles com tamanho idêntico à amostra original, permitindo assim uma análise do vetor de estimativas para os parâmetros de integração fracional ( $d$ ). As condições iniciais são dadas pela média amostral das primeiras observações para o modelo com quebras estruturais e pela média amostral geral no caso do modelo em árvore com efeitos de assimetria. São incluídas também simulações de modelos AR com os parâmetros lineares do modelo em árvore, tornando possível enfatizar a habilidade dos efeitos não-lineares em descrever decrescimentos hiperbólicos nas autocorrelações diante da influência possivelmente enganosa das estruturas autoregressivas persistentes.

Dois métodos semi-paramétricos são aplicados para as estimações dos parâmetros de memória longa: o estimador de Geweke e Porter-Hudak (1983), o mais comum na literatura, e o estimador com viés reduzido de Andrews e Guggenberger (2003). São empregadas duas alternativas para o valor do número de ordenadas utilizadas em cada regressão ( $l$ ), onde  $n$  é o tamanho da amostra:  $n^{1/2}$ , o valor da regra de bolso sugerido por Geweke e Porter-Hudak (1983) (fundamentado em resultados de simulações) e o valor selecionado pelo método de Hurvich e Deo (1999), que aponta um valor por volta de  $n^{0.65}$  para todas as séries.

Para cada conjunto de simulações, também é avaliado o poder do teste de Ohanissian, Russell e Tsay (2004) para a hipótese nula de um processo gerador de dados com memória longa verdadeira, baseado na propriedade de invariância do parâmetro de memória longa em agregações temporais sob a hipótese nula. As especificações do teste podem ser consultados no apêndice C. Andersen et al. (2001) examinam essa propriedade para ações do Dow Jones como evidência de memória longa, embora sem implementar um teste formal.

A Tabela 4.1 reporta as médias e o desvios-padrões (em parênteses) do parâmetro de integração fracional para os conjuntos de simulações, além do valor estimado para a série original do logaritmo da volatilidade realizada das ações da IBM. A primeira linha da tabela indica que o modelo com múltiplos

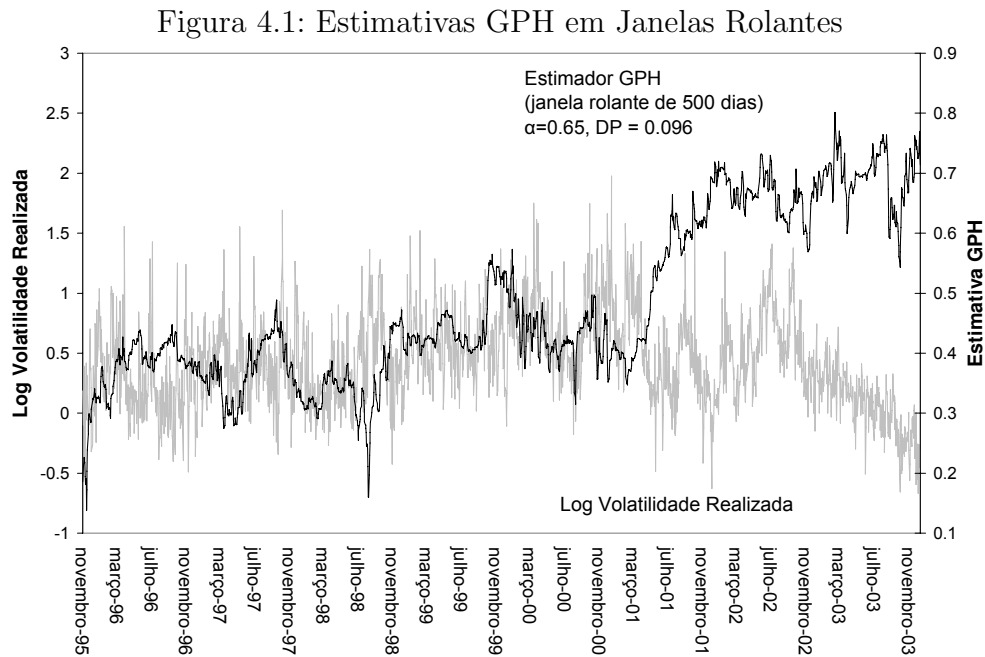
regimes associados com efeitos de retornos acumulados é capaz de gerar altas estimativas para o parâmetro, mesmo em amostras extensas. Em concordância com a literatura, o mesmo é válido para o modelo com quebras estruturais. A tabela também sublinha que o teste de Ohanissian, Russell e Tsay (2004) exibe baixa potência contra essas alternativas. O teste não rejeita a hipótese nula na série original, embora com alta sensibilidade à especificação e ao próprio período analisado – se a primeira semana é removida da amostra, por exemplo, o teste rejeita a hipótese nula a 5%.

Tabela 4.1: Memória Longa - Simulações e Logaritmo da Volatilidade Realizada

	$l = n^{0,5}$		$l = n^{0,65}$		$l = n^{0,7}$
	GPH	AG	GPH	AG	LMT
Árvore (AE)	0,48 (0,15)	0,30 (0,25)	0,60 (0,08)	0,44 (0,17)	33,8%
Árvore (SB)	0,42 (0,08)	0,51 (0,12)	0,50 (0,04)	0,42 (0,09)	25,5%
AR	0,14 (0,11)	0,02 (0,16)	0,38 (0,05)	0,11 (0,11)	94,5%
Log Vol Realizada	0,6 (0,10)	0,35 (0,17)	0,46 (0,05)	0,59 (0,10)	0,556*

A tabela reporta as médias e os desvios-padrões dos vetores de parâmetros de integração fracional estimados para mil simulações de três modelos (efeitos de assimetria, quebras estruturais e autoregressivo), cujos parâmetros são dados pelas estimações apresentadas nas seções anteriores. GPH e AG indicam os estimadores de Geweke e Porter-Hudak (1983) e Andrews e Guggenberger (2003) respectivamente. O número de ordenadas utilizadas em cada regressão está indicada na primeira linha. O número de ordenadas usadas em cada conjunto de regressões é dado por  $n^\alpha$ , onde  $n$  é o tamanho da amostra e  $\alpha$  é o parâmetro indicado na primeira linha. A última coluna exibe os resultados para o teste de Ohanissian, Russell e Tsay (2004) para a hipótese nula de um processo com memória longa verdadeira; os três primeiros valores indicam a porcentagem de simulações tais que a hipótese nula é rejeitada e a última linha indica p-valor do teste para a série original da IBM.

Inicialmente documentada por Granger e Ding (1996), uma importante questão empírica acerca da abordagem *ARFIMA* para a modelagem de memória longa em dados financeiros é a excessiva variação observada ao longo tempo das estimativas do parâmetro de integração fracional –possivelmente incluindo longos períodos em regiões não estacionárias. Esse problema é ilustrado no presente contexto pela Figura 4.1, que plota a evolução de estimativas do parâmetro de integração do logaritmo da volatilidade realizada da IBM pelo método GPH ( $l = n^{0.65}$ ) em janelas rolantes de quinhentos dias de negócios. O gráfico mostra que os estimadores GPH variam entre cerca de 0,2 e 0,8 através da amostra –apesar do desvio padrão de 0,096–, além



de apresentarem uma elevação consistente para níveis não-estacionários (i.e., acima de 0,5) a partir de 2001.

Uma flexibilidade interessante do modelo de regressão em árvore com efeitos de assimetria é a possibilidade de reproduzir esse fato. Esse ponto é exemplificado abaixo através de simulações parciais do modelo a partir da série observada de retornos, analisadas em duas subamostras determinadas pela primeira quebra estrutural do modelo em árvore correspondente. Embora trate-se de um procedimento *ad hoc* que tende a subestimar a capacidade do modelo de gerar autocorrelações persistentes, o exercício configura uma forma útil de obter uma indicação acerca das características do modelo.

A Tabela 4.2 mostra os resultados em contraste com as estimativas para a série original. Como a Figura 4.1 sugere, todas as estimativas apontam para uma persistência significativamente menor na primeira parte da amostra. Com efeito, esta é a fonte da performance irregular do modelo com integração fracional em termos de ajuste, reportada na seção 3.5.2 – a estimativa elevada para o  $d$  na série como um todo produz grandes erros na primeira parte da amostra e autocorrelações residuais (também induzidas pela queda contínua da volatilidade em 2003). A tabela revela ainda que o modelo em árvore com efeitos de assimetria de fato é capaz de reproduzir essas variações temporais em estimativas dos parâmetros de integração fracional, não obstante os valores médios inferiores às estimativas não-estacionárias observadas na série original.

Tabela 4.2: Estimativas para o Parâmetros de Memória Longa: Simulações Parciais

Jan/1994 a Ago/1998	GPH (0,5)	GPH (0,65)	AG (0,65)
Simulação Parcial	0,33 (0,13)	0,52 (0,07)	0,29 (0,14)
Log Vol Realizada	0,34 (0,13)	0,29 (0,07)	0,36 (0,14)
Set/1998 a Dez/2003	GPH	GPH	AG
Simulação Parcial	0,46 (0,11)	0,6 (0,06)	0,43 (0,12)
Log Vol Realizada	0,65 (0,12)	0,66 (0,07)	0,74 ( 0,14)

A tabela reporta a média e o desvio padrão (em parênteses) das estimativas para o parâmetro de integração fracional ( $d$ ) em duas subamostras da série do log da volatilidade realizada da IBM e de mil simulações parciais do modelo em árvore com efeitos de assimetria (as funções de transição são calculadas a partir da série real de retornos). GPH e AG indicam os estimadores de Geweke e Porter-Hudak (1983) e Andrews e Guggenberger (2003) respectivamente. O número de ordenadas usadas em cada conjunto de regressões é dado por  $n^\alpha$ , onde  $n$  é o tamanho da amostra e  $\alpha$  é o parâmetro indicado na primeira linha.