

# Modelos para a Persistência na Volatilidade da Taxa de Câmbio R\$/US\$ - Análise Comparativa entre GARCH e Mudança Markoviana

Márcio Poletti Laurini<sup>1</sup>  
Marcelo Savino Portugal<sup>2</sup>

## Resumo

Neste artigo mostramos que a persistência dos choques estimada pelo modelo GARCH para a volatilidade condicional para a série de taxas de câmbio R\$/US\$ pós Real é uma medida inadequada. Para mostrar tal inadequação, realizamos estimações com janelas móveis, com amostra crescente e, por fim, um experimento de Monte Carlo onde estudamos a persistência estimada pelo modelo GARCH quando o processo gerador dos dados é sujeito a mudanças de regime. Os resultados destes procedimentos mostram que a persistência estimada pelo modelo GARCH é instável, e sofre severamente de um problema de viés relacionado ao tamanho da amostra quando o processo gerador dos dados é sujeito a mudanças de regimes nas estruturas da média condicional e da variância incondicional. Em substituição ao modelo GARCH, utilizamos um modelo de Mudança Markoviana (Markov Switching) para prever a volatilidade desta série. As principais vantagens deste método são duas: a persistência dos choques na volatilidade condicional é menor para intervalos pequenos de previsão; as previsões são condicionais às probabilidades de vigência de cada regime, incorporando mudanças de parâmetros, e o processo de volatilidade neste modelo é muito mais suave do que a do modelo GARCH. Mostramos também, que o decaimento dos choques no modelo markoviano é hiperbólico, e não exponencial, como é o caso no modelo GARCH, sendo consistente com modelos de Memória Longa na Variância da classe FIGARCH/HYGARCH.

Palavras Chave: Volatilidade, GARCH, Mudanças de Regime, Memória Longa, Mudança Markoviana.

JEL CODE: C22, C15, C53.

## 1 Introdução

Uma característica fundamental de séries financeiras é a existência de heterocedasticidade: o fato de que a volatilidade dos retornos não é constante no tempo. Esta heterocedasticidade também é de um tipo especial, já que depende das realizações passadas da própria série e da volatilidade nos momentos anteriores. O modelo mais utilizado para modelar a heterocedasticidade presente em séries financeiras é a classe de modelos GARCH (*Generalised Autoregressive Conditional Heterocedasticity*) de Engle [16] e Bollerslev [6], onde a volatilidade é modelada como um processo determinístico em função do passado dos erros ao quadrado da estrutura da média condicional e da própria volatilidade no passado. Entretanto modelos da classe GARCH sofrem de vários problemas, sendo os mais conhecidos os problemas de persistência espúria devido a presença de quebras estruturais e o fato de que o estimador de máxima verossimilhança para um processo GARCH em amostras pequenas é negativamente viesado de forma muito significativa.

Em comparação com o modelo GARCH, iremos mostrar que um modelo autoregressivo sujeito a mudanças controladas por uma cadeia de Markov para os parâmetros autoregressivos e a variância

---

<sup>1</sup> Ibmec Business School - Rua Maestro Cardim 1170, 11 andar, tel (55) 11 3175 2311, fax (55) 11 3175 2311, email - marciopl@ibmec.br

<sup>2</sup> Professor de Economia na Universidade Federal do Rio Grande do Sul (UFRGS) e pesquisador associado do CNPq - email - msp@ufrgs.br

incondicional é capaz de capturar uma estrutura de dependência condicional para a variância semelhante a uma estrutura GARCH, mas aonde o decaimento temporal da persistência é de uma forma hiperbólica, em substituição ao decaimento exponencial imposto pelos modelos GARCH. Por último mostraremos que o decaimento dos choques neste modelo de Mudança Markoviana é compatível com os modelos de Memória Longa para a variância do tipo FIGARCH/HYGARCH.

A justificativa para o uso de um modelo de mudança de regimes é dada pelo fato de que a presença de heterocedasticidade condicional é diretamente relacionada à existência de mudanças de parâmetros no processo gerador dos dados. Um resultado conhecido na literatura bayesiana é que um modelo de regressão com coeficientes aleatórios e variância constante pode ser transformado em uma regressão heterocedástica com coeficientes fixos (Bauwens, Lubrano e Richard [5]), e esta relação entre heterocedasticidade condicional e mudança de parâmetros é também explorada nos modelos CHARMA de Tsay [42] e no modelo RCA de Nicholls e Quinn [36], que usam coeficientes aleatórios para produzir heterocedasticidade condicional.

A estrutura do artigo é a seguinte: na Seção 2 revisamos as propriedades do modelo GARCH para chegar a definição de persistência que iremos utilizar no artigo. Na Seção 3 estimamos os primeiros modelos do tipo GARCH para esta série. Nas Seções 4 e 5 revisamos alguns problemas apontados pela literatura para modelos do tipo GARCH e realizamos alguns experimentos com a série de câmbio para mostrar a efetividade destes problemas. O modelo de Mudança Markoviana é apresentado e estimado na Seção 6, e a Seção 7 contém um experimento de Monte Carlo que mostra as propriedades do modelo GARCH quando o processo gerador dos dados é um modelo de Mudança Markoviana. Mostramos que um modelo de Memória Longa na Variância é consistente com um processo de Mudança Markoviana na Seção 9 e as conclusões estão na Seção 10.

## 2 Modelos ARCH-GARCH

### 2.1 Modelo ARCH

A modelagem de volatilidade condicional baseada na classe de modelos ARCH teve sua criação através do trabalho de Engle [16] para o estudo da variância da taxa de inflação na Inglaterra. A literatura básica de estudo do comportamento empírico da volatilidade da taxa de câmbio é baseada principalmente em formulações ARCH.

O modelo ARCH de Engle [16] define o comportamento da volatilidade condicional de uma série como uma função linear dos choques ao quadrado ocorridos no passado desta série. Definindo  $y_t$  como a série em interesse, podemos representar o comportamento da média da série como:

$$y_t = X\beta + \varepsilon_t \quad (1)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (2)$$

onde  $X$  é um conjunto de variáveis explicativas, podendo incluir defasagens de  $y_t$ . O processo  $z_t$  é um IID com média zero e variância unitária. Um processo ARCH(1) é representado como:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 \quad (3)$$

onde  $h_t$  representa a variância condicional da série no momento  $t$ . Uma condição necessária é que todas as realizações de  $h_t$  não sejam negativas, o que requer que  $w$  seja maior que 0 e que  $\alpha_1$  seja  $\geq 0$ . Se  $\alpha_1$  é igual a zero o modelo é condicionalmente homocedástico.

A representação 3 permite ver que o modelo ARCH captura agrupamentos de volatilidade devido ao fato de que o choque  $\varepsilon_t$  é função crescente do choque passado  $\varepsilon_{t-1}^2$ . Desta forma se o choque no momento passado foi grande, o choque no presente também tende a ser grande. Uma vantagem adicional do modelo ARCH é que ele permite capturar o excesso de kurtose existente em séries financeiras. Conforme Engle [16], a kurtose de  $\varepsilon_t$  sempre supera a kurtose de  $z_t$  já que:

$$E[\varepsilon_t^4] = E[z_t^4]E[h_t^2] \geq E[z_t^4]E[\varepsilon_t^2]^2 \quad (4)$$

o que é valido pela desigualdade de Jensen. Se  $z_t$  possui distribuição normal, a kurtose de  $\varepsilon_t$  é igual a

$$K_\varepsilon = \frac{E[\varepsilon_t^4]}{E[\varepsilon_t^2]^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2} \quad (5)$$

que é finita se  $3\alpha_1^2 < 1$  e sempre maior que 3.

O modelo ARCH(1) pode ser generalizado para uma forma geral ARCH( $q$ ) representada como:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (6)$$

onde a variância incondicional  $\sigma^2$  é definida pela relação:

$$\sigma^2 = \frac{w}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_q} \quad (7)$$

O modelo será covariância-estacionário se todas as raízes do polinômio  $1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_q L^q$  estiverem fora do círculo unitário, o que equivale a dizer que os choques não têm efeitos permanentes, existindo uma reversão à volatilidade incondicional.

## 2.2 ARCH Generalizado - GARCH

É comum a necessidade de uma representação com um valor elevado de  $q$  no modelo ARCH para representar adequadamente o comportamento da volatilidade existente em séries financeiras. Bollerslev [6] sugeriu uma formulação geral, conhecida como GARCH (*Generalized ARCH*) que é uma generalização do modelo ARCH para incluir valores defasados de  $h_t$  em 6. O modelo GARCH(1,1) é representado como:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (8)$$

As condições de positividade são que  $w$  e  $\alpha_1 > 0$  e  $\beta_1 \geq 0$ . A representação 8 evita a necessidade de colocarmos muitos lags de  $\varepsilon_t$ , o que notamos se 8 for reescrita como:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 (w + \alpha_1 \varepsilon_{t-2}^2 + \beta_1 h_{t-2}) \quad (9)$$

Continuando esta recursão chegamos a:

$$h_t = \sum_{i=1}^{\infty} \beta^i w + \alpha_1 \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^2 \quad (10)$$

A representação 10 corresponde a um modelo *ARCH*( $\infty$ ) para esta estrutura de  $\varepsilon_t^2$ . Adicionando  $\varepsilon_t^2$  em ambos os lados de 8 e movendo o termo  $h_t$  para o lado direito da equação, o modelo GARCH(1,1) pode ser escrito como um modelo ARMA(1,1) para  $\varepsilon_t^2$ :

$$\varepsilon_t^2 = w + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1} \quad (11)$$

onde  $v_t = \varepsilon_t^2 - h_t$ . O processo definido em 11 para  $\varepsilon_t^2$  será covariância estacionário se  $\alpha_1 + \beta_1 < 1$ . Definimos a variância incondicional de  $\varepsilon_t$  como:

$$\sigma^2 = \frac{w}{1 - \alpha_1 - \beta_1} \quad (12)$$

A kurtose de  $\varepsilon_t$ , novamente assumindo que  $z_t$  seja normalmente distribuído, é representada como:

$$K_\varepsilon = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \quad (13)$$

que é também sempre maior do que 3. Uma característica importante derivada por Bollerslev [6] é que as autocorrelações de  $\varepsilon_t$  são dadas por:

$$\rho_\varepsilon = (\alpha_1 + \beta_1)^{k-1} \rho_1 \quad (14)$$

o que faz com que as correlações decaiam exponencialmente mas no fator de  $\alpha_1 + \beta_1$ . Quando  $\alpha_1 + \beta_1$  está próxima de um, o decaimento é bastante lento, como ocorre no modelo autoregressivo com raízes unitárias.

Da mesma forma que o modelo ARCH(1) pode ser generalizado para um processo ARCH(q), podemos generalizar um processo GARCH(1,1) para um processo GARCH(p,q) dado por:

$$h_t = w + \sum_{i=1}^q \varepsilon_{t-i}^2 + \sum_{i=1}^p h_{t-i} \quad (15)$$

Embora a estimação de modelos GARCH de ordem maior que (1,1) seja trivial, a especificação mais empregada é a de um GARCH(1,1), e esta parece ser a forma mais adequada na prática, conforme Bollerslev, Chou e Kroner [7].

A generalização do modelo ARCH permite modelar a volatilidade através de modelos ARMA(p,q) (GARCH), formas exponenciais (EGARCH) e ARIMA(p,I,q) (IGARCH) para o processo da volatilidade. Afora estes modelos, uma variada gama de modelos derivados do modelo ARCH original de Engle [16] tem sido proposta, incluindo modelos de integração fracionária, modelos quadráticos, modelos de fatores ARCH entre muitas outras formas. Uma especificação que parecia adequada aos nossos dados, o modelo EGARCH (Exponential GARCH) de Nelson [34] que permite uma resposta assimétrica da variância condicional aos choques ocorridos, não foi estudada por problemas de convergência na maximização da verossimilhança. Em nenhuma das várias especificações testadas para o modelo EGARCH na amostra utilizada ocorreu convergência forte, o que nos levou a abandonar o uso deste modelo.

### 2.3 GARCH Integrado

Uma especificação relevante em modelos da classe é o modelo *Integrated GARCH* (IGARCH). O modelo IGARCH força a restrição de que a soma de parâmetros  $\alpha_1 + \beta_1$  seja igual a um, uma restrição válida para a maioria das séries financeiras. Essa restrição impõe uma raiz unitária na representação ARMA(1,1) para  $\varepsilon_t^2$ , o que origina o termo integrado para essa série. Com essa restrição podemos escrever:

$$(1 - L)\varepsilon_t^2 = w - v_t - \beta_1 v_{t-1} \quad (16)$$

Como colocam Franses e Van Dijk [18], a analogia com processos de raízes unitárias deve ser considerada com cuidado. Embora a variância incondicional não exista, conforme mostra a equação 12, e o modelo não seja covariância-estacionário, o modelo IGARCH pode ainda ser estritamente estacionário se  $E[\ln(\alpha_1 z_{t-1}^2 + \beta_1)] < 0$ , como mostrado por Nelson [35].

As autocorrelações para  $\varepsilon_t^2$  neste modelo não estão definidas, mas segundo Ding e Granger [13] podem ser aproximadas por:

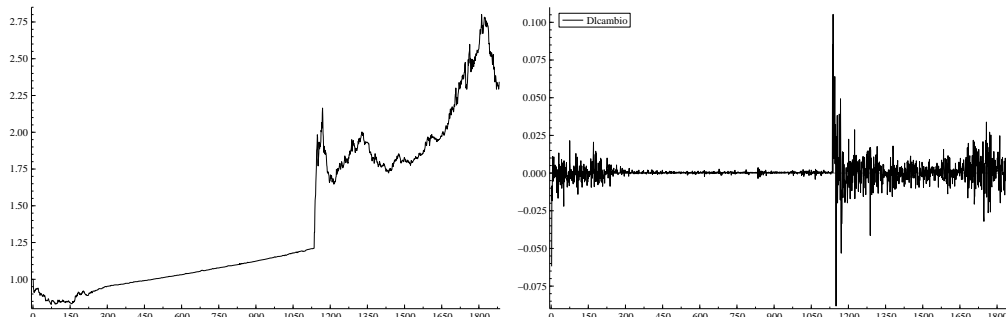
$$\rho_k = \frac{1}{3}(1 + 2\alpha_1)(1 + 2\alpha_1)^{-k/2} \quad (17)$$

e assim decaem exponencialmente, ao contrário das autocorrelações de um modelo com raiz unitária onde as autocorrelações devem se manter teoricamente no valor um.

### 2.4 Persistência

Uma questão fundamental é verificar se os choques têm efeitos transitórios ou permanentes na estrutura da variância condicional da série. Desta forma necessitamos encontrar uma expressão que determine o efeito final do choque  $\varepsilon_0^2$  na variância condicional  $h_t$ , o que pode ser representado como:

Figura 1: Taxas de Câmbio R\$/US\$ e Log-Retornos



$$\lim_{t \rightarrow \infty} E \left( \frac{\partial h_t}{\partial e_0^2} \right) \quad (18)$$

Nelson [35] mostra que, assumindo que  $h_0$  é independente da seqüência  $\varepsilon_t$ , ela mesma uma seqüência também independente, o valor de 18 é dado por:

$$\begin{aligned} E \left( \frac{\partial h_t}{\partial e_0^2} \right) &= E(h_0) E \left( \frac{\partial \prod_{i=1}^t (\alpha_1 \varepsilon_{i-1}^2 + \beta_1)}{\partial e_0^2} \right) \\ &= E(h_0) E \left( \prod_{i=1}^t (\alpha_1 \varepsilon_{i-1}^2 + \beta_1) \right) E \left( \frac{\partial (\alpha_1 \varepsilon_0^2 + \beta_1)}{\partial e_0^2} \right) \\ &= E(h_0) (\alpha_1 E(\varepsilon^2) + \beta_1)^{t-1} \alpha_1 \end{aligned} \quad (19)$$

Existe persistência se o valor de 19 converge para um número finito e diferente de zero quando  $t \rightarrow \infty$ , o que ocorre quando  $\alpha_1 E(\varepsilon^2) + \beta_1 = 1$ , e o efeito permanente é dado por  $\alpha_1$ .

Os choques terão efeito permanente na variância da série quando a soma dos parâmetros  $\alpha_1$  e  $\beta_1$  em um processo GARCH(1,1) for igual a 1, o que corresponde ao modelo IGARCH de Engle e Bollerslev [17]. Nosso estudo da persistência dos choques na variância condicional da série de log-retornos do câmbio concentra-se na análise da soma dos parâmetros ARCH(1) e GARCH(1) estimados.

### 3 Estimação Modelo GARCH - Série R\$/US\$

Os dados utilizados no estudo compreendem os log-retornos da série de taxas de câmbio Real/Dólar no período de 01 de julho de 1994 até 4 de janeiro de 2002, o que leva a um número de 1881 observações<sup>3</sup>. Essa série é caracterizada pela presença de 3 regimes exógenos de câmbio, Pode-se observar no início da série um período de flutuação com tendência de valorização do Real frente ao Dólar, correspondendo a introdução do Real. Seguindo esse período, temos uma tendência linear de desvalorização do Real, correspondendo ao período de adoção do regime de bandas cambiais, e após a crise cambial de 13 de janeiro de 1999 temos um regime de livre flutuação do câmbio. A figura 1 mostra a série de câmbio R\$/US\$ em nível e os log-retornos utilizados no estudo.

Iniciamos a estimação de um modelo GARCH geral com uma especificação AR(6) para a média do processo e uma especificação GARCH(3,3) para a variância incondicional. O modelo ARMA(0,0)-GARCH(1,1) foi a especificação escolhida para a média e variâncias condicionais do processo, utilizando o critério de informação de Akaike

<sup>3</sup>É importante notar que estas observações não tiveram nenhum ajuste para os dados faltantes, correspondentes aos dias em que não houveram transações no mercado de câmbio. Tratamos a amostra como se fosse contínua, indo de 1 a 1882 sem dados faltantes. Seria necessário um estudo para verificar se a utilização de métodos para tratamento de dados faltantes alteraria os resultados principais de nosso estudo

Tabela 1: ARMA(0,0) - GARCH(1,1)

ARMA(0,0)	GARCH(1,1)		amostra	8-1882	
parâmetro	valor est.	desvio-padrão	dp-robusto	<i>t-value</i>	<i>t-prob</i>
Constante	0.000333736	8.970e-005	8.971e-005	3.72	0.000
Intercepto-Garch	1.56549e-006	2.244e-007	1.715e-006	0.913	0.361
ARCH(1)	0.148378	0.02360	0.04071	3.64	0.000
GARCH(1)	0.824514	0.02312	0.03913	21.1	0.000
log-lik	7353.55396		AIC	-7.84952423	
média(h_t)	5.81099e-005		var(h_t)	3.04398e-008	
alpha(1)+beta(1)	0.972892				

Tabela 2: Testes de Especificação

Resíduos				Resíduos ao quadrado			
ARCH 1-2 teste	F(2,1867)	0.18546	[0.8307]	ARCH 1-2 teste	F(2,1867)	0.18546	[0.8307]
Normalidade	Chi <sup>2</sup> (2)	8112.7	[0.0000]**	Normalidade	Chi <sup>2</sup> (2)	8112.7	[0.0000]**
Portmanteau(43)	Chi <sup>2</sup> (43)	52.563	[0.1505]	Portmanteau(43)	Chi <sup>2</sup> (41)	0.76315	[1.0000]

Os teste para efeitos ARCH remanescentes e o teste *Portmanteau* (Tabela 2) indicam que o modelo estimado controla a heterocedasticidade condicional e autocorrelação existentes na série, tanto para os resíduos normais quanto para os resíduos ao quadrado. Entretanto a normalidade no resíduo é rejeitada a qualquer nível de significância, o que é devido ao excesso de kurtose remanescente nos resíduos do modelo.

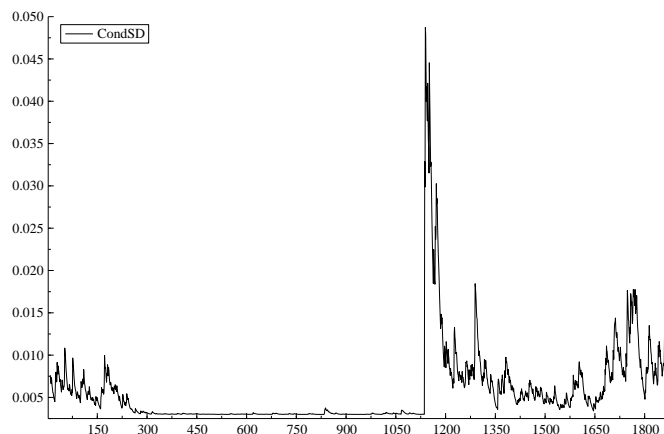
Como o único parâmetro presente na especificação da média é uma constante, o valor previsto para a média da série é sempre 0.000333736. Os valores previstos para a variância condicional da série estão colocados na Figura 2. Este gráfico mostra que o comportamento da volatilidade da série é compatível com a classificação dos 3 subperíodos observáveis neste período para a série de câmbio. O comportamento nas observações 6-300 e 1138-1881 mostra a volatilidade decorrente da liberdade de flutuação do câmbio, enquanto que no período de bandas o valor previsto para a variância condicional é extremamente baixo.

A persistência dos choques nessa série, medida pela soma dos parâmetros  $\alpha_1$  (ARCH(1)) e  $\beta_1$  (GARCH(1)) é igual a 0.972892. Como os parâmetros  $\alpha_1$  e  $\beta_1$  têm desvios padrões 0.02360 e 0.02312, realizamos um teste com um modelo restrito IGARCH (forçando a restrição de  $\alpha_1 + \beta_1 = 1$ ) e comparamos com o modelo irrestrito. Por este teste não podemos rejeitar que a soma dos parâmetros seja diferente de um. Desta forma, a volatilidade condicional pode ser considerada um processo não-estacionário, e os choques na variância possuem uma persistência muito longa por este modelo estimado.

## 4 Problemas com a Modelagem ARCH

A modelagem ARCH e suas derivações têm se mostrado limitadas no seu sucesso em modelar a volatilidade condicional existente em algumas séries, já que a literatura aponta que estes modelos levam a alguns resultados implausíveis. Uma evidência colocada por Gallant, Hsieh e Tauchen [19] no estudo da taxa de câmbio Libra/Dólar diária de 1974 a 1983 é o fato de que os resíduos dos modelos GARCH apresentam magnitudes de simetria e kurtose muito maiores dos que os dados

Figura 2: Variância Condicional Modelo GARCH(1,1)



originais:

"... the Studentized residuals from the non-normal GARCH-type models display substantially more skewness and kurtosis than the original data. The large skewness is inconsistent with the assumption that Studentized residuals have a symmetric, homogeneous distribution. The magnitude of the kurtosis is rather dramatic in some cases: The kurtosis of the GARCH(1,1)-t residuals is 357.351 versus 8.995 for the raw data. ... While the large kurtosis cannot, strictly speaking, be construed as evidence against the GARCH-t fit, it is disconcerting because it is accustomed to having residuals that are better behaved than the original data."

Mas como colocam Gallant, Hsieh e Tauchen [19], a não-normalidade dos retornos das séries financeiras é inesperada. Segundo eles:

"The lack of Gaussianity of daily price changes is a paradox. Financial markets are active and the price often moves many times within the day. The within-day movements are triggered by pieces of new information about market fundamentals and are therefore independent, or nearly so. The daily price change is thus the sum of many within-day independent price movements. By the central theorem one would expect gaussianity to obtain, but it does not."

Uma tentativa de resolução deste paradoxo veio do trabalho de Clark [10], que mostrou que embora os choques à estas séries sejam independentes, a frequência de ocorrência destes choques varia no tempo. Em determinados períodos, os mercados estão agitados e a frequência de novas informações e choques é muito maior. Em outros períodos, no entanto, novas informações e choques são escassos, e a mudança de preços ocorre de forma bem mais lenta. A abordagem de Clark [10] é pioneira porque aponta para o fato de que este fenômeno tem uma importância estatística fundamental, ao mostrar que o modelo mais realista para estes preços é um modelo de misturas (*mixture model*), onde os choques vem de distribuições diferentes. Comentando o trabalho de Clark [10], Gallant, Hsieh e Tauchen [19] sintetizam:

"As is well known, a mixture of random variables, even Gaussian random variables, is non-Gaussian and has a leptokurtic density. Clark estimated a mixture of normal models with an i.i.d log-normal mixing process and found that the implied distribution fits better than a straight Gaussian distribution or the stable distribution."

Os resultados dos trabalhos de Clark [10] e de Gallant, Hsieh e Tauchen [19] apontam para o fato de que embora existam poucas evidências de dependências nas médias (retornos) para essas séries, existem fortes evidências de dependências condicionais nos demais momentos destas séries, e que não podem ser explicados corretamente pela classe de modelos ARCH usualmente empregados nestas séries. A abordagem de Gallant, Hsieh e Tauchen [19] para trabalhar com estas séries é utilizar uma classe de modelos semi-paramétricos de mistura conhecidos como modelos SNP.

Outro problema existente na modelagem das taxas de câmbio por modelos da classe ARCH é que os parâmetros estimados para estes processos geralmente implicam variâncias explosivas e persistências excessivamente longas no tempo. Como coloca Gray [21]:

"One problem with fitting Garch models ... is that the parameter estimates often imply that the conditional variance processes is not covariance-stationary. Bollerslev (1986) demonstrates that this is the case when  $a+b$  is greater than one. Nelson (1990) shows that in such cases shocks accumulate in the sense that  $E[h(t+m)|h(t)] \rightarrow \infty$  as  $m \rightarrow \infty$ . Note, however, that this does not imply that  $h(t+m) \rightarrow \infty$  as  $m \rightarrow \infty$ "

O problema com a modelagem GARCH é que esta persistência pode ser espúria. Lamoureux e Lastrapes[32] demonstram que qualquer mudança estrutural nos parâmetros pode levar a estimação incorreta dos parâmetros do GARCH de forma que eles indicam persistência excessiva na volatilidade. O modelo selecionado é normalmente um modelo IGARCH, enquanto que o processo gerador da volatilidade é estacionário.

Hamilton e Susmel [25] apontam que a persistência calculada pelos modelos da classe ARCH contradiz a baixa capacidade preditiva apresentada por estes modelos. Se a persistência estimada fosse válida, isto facilitaria o cálculo de previsões para fora da amostra. Mas, como colocam Hamilton e Susmel, [25] acontece exatamente o contrário:

"Such persistence is difficult to reconcile with the poor forecasting performance - if the variance were this persistent, the model should be a much better job at forecasting. However, the models forecasts continue to worsen relative to a constant-variance specification as the forecast horizon increases"

A presença de *outliers* também pode afetar o resultado das estimações de modelos GARCH, indicando persistência excessiva de forma incorreta. O problema se torna ainda mais grave por outra razão: Doornik e Ooms [15] apontam que o uso de *dummies* para o controle de *outliers* e quebras na variância pode induzir a multi-modalidade na função de verossimilhança do modelo ARCH, tornando as inferências estatísticas inválidas. O estudo de influência de *outliers* em modelos ARCH pode ser vista nos artigos de Hotta e Tsay [28], que derivam testes para presença de *outliers* em modelos GARCH, e em Sakata e White [41], discutem métodos robustos de estimação na presença de *outliers*.

Um problema adicional ocorre devido ao problema do viés existente em pequenas amostras quando estimando processos GARCH pelo método de máxima verossimilhança. Segundo Hwang e Valls Pereira [29], o estimador de máxima verossimilhança irrestrito é severamente viesado para baixo em amostras pequenas para processos GARCH, e quando são impostas as condições de não negatividade de Bollerslev [6] existem problemas de não convergência e viés positivo, e que os termos de viés diminuem muito lentamente com o tamanho da amostra.

Como substituição ao modelo GARCH, as vantagens do modelo de Mudança Markoviana para as séries de câmbio são que, ao permitir que a volatilidade seja o resultado de distintos regimes (distribuições), os parâmetros estimados não indicam as persistências explosivas capturadas pelos modelos ARCH, e que são contraditórias aos fenômenos observados nas séries financeiras. Como as previsões para dentro e fora da amostra são realizadas utilizando-se as probabilidades suavizadas e filtradas de cada regime, a persistência dos choques é dependente do regime vigente, o que permite uma maior flexibilidade na realização das previsões, contornando o problema da persistência espúria existente nos modelos GARCH. A modelagem de mudança markoviana permite que tenhamos regimes de alta volatilidade, que podem ser associados a momentos de crise, e regimes de baixa volatilidade, condizentes com momentos de calma nos mercados. Como as previsões são realizadas utilizando-se as probabilidades filtradas de cada regime



Tabela 3: Estimação GARCH subamostras

Estimação em sub-amostras	parâmetros	modelo GARCH		
Amostra 1-300	ARMA(0,1)	GARCH(1,1)	valor t	<i>p-value</i>
ARCH(1)	0.083804	0.022397	3.742	0.0002
GARCH(1)	0.916496	0.019368	47.32	0.0000
Alpha[1]+Beta[1]	1.00029			
Amostra 301-1137	ARMA(0,0)	GARCH(1,1)	valor t	<i>p-value</i>
ARCH(1)	-0.234306	0.01526	0.2352	0.320
GARCH(1)	0.938565	0.01023	150	0.000
Alpha[1]+Beta[1]	0.704259			
Amostra 1138-1881	ARMA(1,1)	GARCH(1,1)	valor t	<i>p-value</i>
ARCH(1)	0.169422	0.037638	4.501	0.0000
GARCH(1)	0.817106	0.036699	22.27	0.0000
Alpha[1]+Beta[1]	0.98652			

## 5 Problemas do modelo GARCH para a série R\$/US\$

Na literatura de modelos GARCH são apresentadas uma série de procedimentos para contornar os problemas apontados na Seção 4, principalmente a existência de quebras estruturais no processo gerador da série. Os principais procedimentos sugeridos são a estimação em subamostras e o uso de janelas móveis para as observações. Mostramos que estes procedimentos são incapazes de efetivamente resolver o problema de persistência elevada causada por quebras estruturais.

### 5.1 Estimação em subamostras

Para tentar verificar qual é o efeito das mudanças estruturais na série de log-retornos sobre a estimação de parâmetros no modelo GARCH, utilizamos o procedimento de reestimar os parâmetros nas 3 subamostras conhecidas exogenamente - período de introdução do Real (observações 1-300), período da Banda Cambial (observações 301-1137) e livre flutuação (1138-1881).

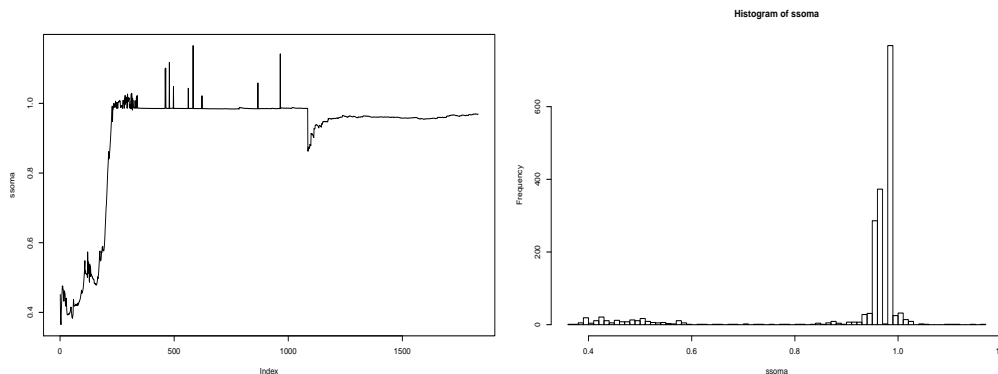
A persistência estimada na primeira e na última subamostra continuam estatisticamente não diferentes de um, enquanto que no período de vigência da banda cambial o valor se reduziu para 0.70. É interessante notar que no período de bandas cambiais o valor do termo ARCH(1) não é estatisticamente significativo, (Tabela 3), o que indica que os choques na volatilidade não possuem efeitos permanentes. Desta forma, embora a soma seja igual a 0.70 devemos considerar que, como o valor do ARCH(1) não é diferente de zero, existe uma persistência elevada já que o termo GARCH(1) tem valor 0.93.

A estimação dos modelos GARCH para esses subperíodos mostra que existe uma mudança significativa no comportamento da variância condicional. Mas embora este procedimento capture esta mudança, ele sofre do problema de que utilizamos apenas informação externa ao modelo sobre pontos de quebra na série, e assim esta divisão em 3 subamostras não captura mudanças endógenas existentes na série. Outro problema fundamental com este procedimento é o problema do viés em pequenas amostras estudado por Hwang e Valls Pereira [29].

### 5.2 Teste Para a Constância de Parâmetros

Como a estimação em subamostras nos mostrou que houve alterações significativas nos parâmetros estimados e na persistência associada, seria importante algum tipo de teste de constância de parâmetros para validar estatisticamente estes resultados. Como a validade da persistência

Figura 3: Estimação com Amostra Crescente - Persistência e Histograma



estimada depende fundamentalmente da constância dos parâmetros, ou seja, ausência de quebras estruturais, precisamos de algum teste de constância de parâmetros no modelo GARCH.

Lundberg e Teräsvirta [33] desenvolveram um teste com este propósito. A hipótese nula neste teste é que os parâmetros são constantes durante toda a amostra, contra hipótese alternativa de que os parâmetros estão mudando de forma suave. O teste pode ser realizado através de uma regressão auxiliar:

$$h_t = w^* + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \zeta t + \gamma_1 \varepsilon_{t-1}^2 t + \delta_1 h_{t-1} t \quad (20)$$

A hipótese nula corresponde à restrição  $\zeta^* = \gamma_1^* = \delta_1^* = 0$ . Usando os parâmetros estimados anteriormente, testamos esta restrição através de um teste Wald, e obtivemos as estatísticas  $F=12907.87$  com p-valor de 0.0000 e uma estatística  $\chi^2=38723.60$  com p-valor de 0.0000. A hipótese nula de que os parâmetros do GARCH são constantes é rejeitada em qualquer nível de significância, indicando que o modelo GARCH é inadequado para modelar a variância condicional da série de log-retornos do câmbio no período em estudo. A não constância dos parâmetros é uma evidência de que a elevada persistência pode ser espúria.

### 5.3 Estimação com Amostra Crescente

Para analisar o comportamento do modelo GARCH na presença das quebras estruturais existentes em nossos dados partimos de um número reduzido de observações e progressivamente foi se aumentando o tamanho da amostra utilizada na estimação. Estimamos inicialmente com um número de observações igual a 50 e, a cada estimação seguinte, aumentamos em uma observação o tamanho da amostra.

Como esperado, o valor final utilizando a amostra total de log-retornos do câmbio é igual ao da estimação realizada anteriormente. O valor médio dos parâmetros de persistência estimada nesse procedimento é igual a 0.9171421 com variância 0.02507348. A Figura 3 mostra que o comportamento da persistência estimada para o modelo GARCH é instável, mostrando no início da amostra uma relação crescente entre valor da persistência estimada e o tamanho da amostra, confirmando o viés negativo do estimador de máxima verossimilhança para pequenas amostras apontado por Hwang e Valls Pereira[29]. Quando a amostra em questão começa a incluir os dados pertencentes ao regime de bandas cambiais, o valor da persistência se estabiliza em valores muito próximos de um, sendo que este comportamento persiste até o fim do regime de bandas. Após o início do câmbio flexível, novamente se reproduz o comportamento de aumento da persistência com o aumento da amostra, embora esse efeito seja suavizado pelo peso das observações anteriores na estimação.

Tabela 4: Estimação Janela Móvel

Tamanho da Janela	Valor médio da persistência	Variância
250	0.8727	0.0504
500	0.9117	0.0252
1000	0.9498	0.0053

Este procedimento permite visualizar alguns problemas existentes com a modelagem GARCH. Comparando o valor encontrado na estimação usando apenas as observações pertencentes ao período da banda cambial encontramos um valor de persistência de 0.704259 (Tabela 3), mas o valor da persistência estimada usando os dois períodos é 0.988201, o que claramente superestima a persistência de choques na volatilidade durante o período de bandas cambiais.

#### 5.4 Estimação com Janelas Móveis

Como forma de contornar os problemas apontados acima na estimação do modelo GARCH, alguns autores sugerem que o procedimento indicado é estimar o processo GARCH usando janelas móveis para os dados. Isto equivale a reestimar o GARCH utilizando os valores mais recentes dos dados, o que podemos considerar um refinamento do procedimento de estimar o modelo em subamostras realizado anteriormente.

Este procedimento é baseado no fato de que o uso de uma janela móvel faz com que, utilizando a informação mais nova e desprezando a informação mais antiga, os valores previstos para fora da amostra sejam teoricamente mais confiáveis, incorporando as transições suaves nos parâmetros. Para ilustrar esse procedimento estimamos um AR(0)-GARCH(1,1) para a série de log-retornos do câmbio com janelas móveis de tamanho 250, 500 e 1000 e calculamos os valores médios e a variância da persistência estimada para cada janela.

Olhando inicialmente apenas os resultados médios obtidos por esse procedimento (Tabela 4) notamos que a persistênciaa relação entre o valor da persistência estimada e o tamanho da amostra, neste caso, a janela de estimação utilizada. O uso de janelas móveis é desaconselhado pelo fato do viés negativo do estimador de máxima verossimilhança ser muito significativo para os valores usuais usados em janelas móveis.

Uma observação relevante neste procedimento é que ele não consegue reduzir a variabilidade existente nos valores estimados para a persistência na série. Olhando para o histograma da distribuição dos estimadores da persistência na janela de 250 observações (Figura 4) notamos que a distribuição empírica é multimodal e com uma grande variabilidade nos valores estimados, o que fica evidente no gráfico com os valores estimados. Este gráfico também mostra um grande número de valores extremos, que correspondem aos pontos onde a função de verossimilhança não converge, o que se mostrou mais freqüente para os valores mais baixos (250 e 500 observações) da janela de estimação.

Quando utilizamos uma janela de 500 observações os problemas encontrados na estimação com a janela de 250 observações se reduzem, já que a distribuição fica mais centrada em torno de .9 e o número de processos estimados cuja verossimilhança não converge diminui de forma significativa, mas a variabilidade nos valores estimados para a persistência continua elevada, conforme pode ser visto nas Figura 5. É relevante notar que o intervalo de maior variabilidade ocorre no período de vigência das bandas cambiais, onde justamente a volatilidade é restrita. Esse fenômeno ocorre pelo fato da janela com 500 observações conter dados dos períodos pré e pós bandas cambiais que contaminam a estimação da persistência no período da banda cambial.

Na estimação utilizando a janela de 1000 observações notamos que a distribuição dos valores estimados para a persistência tem menor variância, mas da mesma forma a distribuição não é unimodal. O número de processos sem convergência se reduz de forma drástica (Figura 6), mas

Figura 4: Valores Estimados e Histograma - Janela Móvel de 250 observações

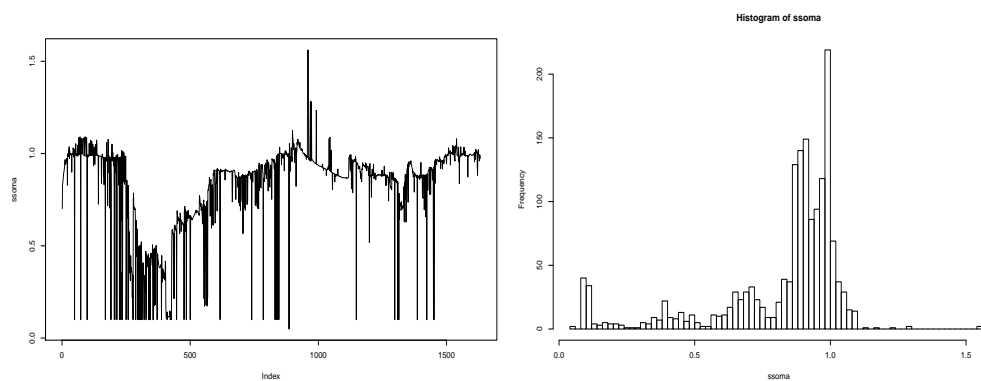


Figura 5: Valores Estimados e Histograma - Janela Móvel de 500 observações

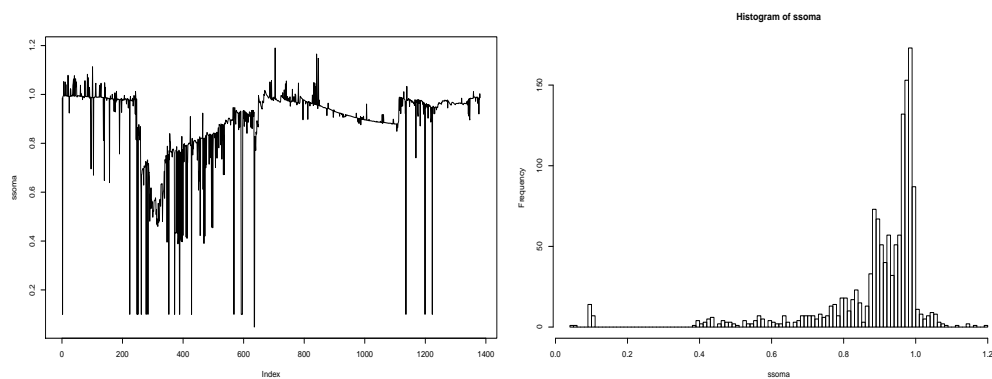
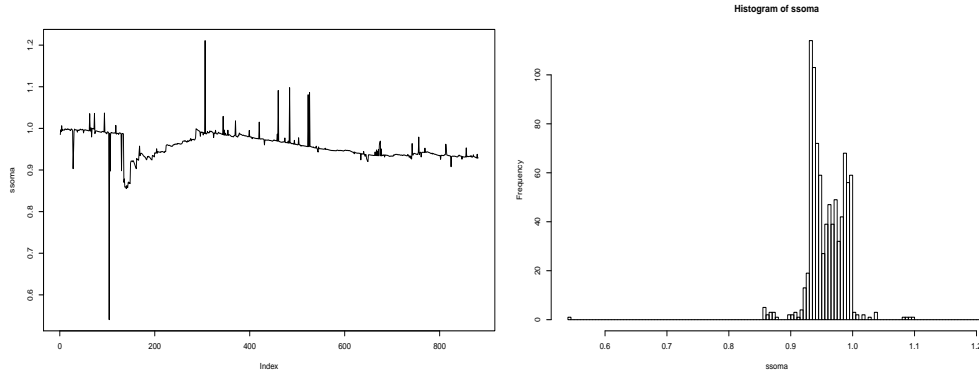


Figura 6: Valores Estimados e Histograma - Janela Móvel de 1000 observações



o valor médio da persistência calculado de 0.94 se mantém dentro do intervalo de confiança em torno do valor 1, o que define um processo de choques com efeitos permanentes na volatilidade.

O procedimento de utilizar janelas móveis na estimação do GARCH não parece ser suficiente para contornar os problemas existentes na presença de quebras estruturais no processo gerador da série. Além da dificuldade em se estabelecer o tamanho ótimo da janela, verificamos que uma janela mais curta, que potencialmente teria a vantagem de dar um peso menor às quebras anteriores existentes na série, sofre de uma instabilidade numérica muito grande na estimação e também do viés negativo do estimador. Janelas mais longas têm valores numericamente mais estáveis, mas não resolvem o problema de persistência excessiva causada pelas quebras estruturais.

## 6 Modelo de Mudança Markoviana

Quando estimamos modelos da classe GARCH para as 3 subamostras na Seção 5.1, esta divisão baseada nos únicos eventos que eram explicitamente delimitados: a introdução do Real, seguido pelo regime de bandas cambiais, e após a crise cambial 13 de janeiro de 1999 o regime de câmbio flutuante. Mas uma delimitação com base nesse evento é claramente insatisfatória como forma de modelar todas as mudanças no processo gerador da série nesse período, já que existe a possibilidade de mudanças endógenas de regime.

Uma técnica estatística reconhecidamente capaz de capturar esta mudança de parâmetros é a classe de modelos conhecidas como modelos de *Markov Switching*, ou como são conhecidos em português, modelos de Mudança Markoviana. Os modelos de mudança markoviana fazem parte da classe de modelos conhecidos como *lineares por partes (piece-wise linear)*, já que o processo gerador dos dados é linear dentro de cada regime.

### 6.1 Descrição do Modelo de Mudança Markoviana

A idéia básica do modelo de mudança markoviana é decompor uma série de dados em uma seqüência finita de processos estocásticos distintos, ou como ficou mais conhecido na literatura, regimes. O processo vigente em cada regime é linear, mas a combinação dos processos gera um regime não-linear. Um exemplo simples seria um modelo autoregressivo de primeira ordem sujeito a mudanças no parâmetro autoregressivo, representado pelo seguinte sistema :

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \epsilon_{1t} | se s_t = 1 \\ Y_t &= \phi_2 Y_{t-1} + \epsilon_{2t} | se s_t = 2 \end{aligned} \quad (21)$$

onde  $r$  representa o regime vigente. Desta forma o parâmetro  $\phi_1$  descreve o comportamento da série caso o regime vigente seja o regime 1, enquanto que se o regime vigente for o 2, o parâmetro que descreve o comportamento da série é  $\phi_2$ .

Como não observamos o processo estocástico que determina qual o regime vigente, necessitamos de algum mecanismo para inferir probabilidades sobre qual regime é vigente no período  $t$ . A idéia básica do modelo de mudança markoviana é descrever o processo estocástico que determina a mudança de um regime para outro através de uma cadeia de Markov. A cadeia de Markov é utilizada como forma de modelar o comportamento de uma variável (ou combinação de variáveis) de estado que determina qual o regime vigente, já que esta variável não é diretamente observável. Podemos representar uma cadeia de Markov da seguinte forma: suponha que a probabilidade da variável  $s_t$  assumir algum valor particular  $j$  depende somente do valor passado  $s_{t-1}$  é dada pela seguinte equação

$$P \{s_t = j | s_{t-1} = k, \dots\} = P \{s_t = j | s_{t-1} = i\} = P_{ij} \quad (22)$$

Este processo é descrito como uma cadeia de Markov com  $n$ -estados, onde  $P_{ij}$  indica a probabilidade de que o estado  $i$  será seguido pelo estado  $j$ . Notando que

$$P_{i1} + P_{i2} + P_{in} = 1 \quad (23)$$

podemos construir o que é chamado de *matriz de transição*, onde a linha  $i$  e a coluna  $j$ , são a probabilidade de que o estado  $i$  será sucedido pelo estado  $j$ .

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix} \quad (24)$$

A característica fundamental do uso da matriz de transição markoviana de primeira ordem é o fato de que a probabilidade de transição para o próximo regime só depende do regime corrente, o que simplifica a modelagem e principalmente os métodos de estimação.

Uma vantagem do modelo de mudança markoviana, conforme colocado por Campbell, Lo e MacKinley [9], é que ele não sofre dos vieses de alguns modelos tradicionais de quebra estrutural, já que a determinação da probabilidade associada a cada regime é determinada por interações entre os dados e a cadeia de Markov, e não por mecanismos *ad hoc* baseados em análises preliminares dos dados, como ocorre por exemplo nos modelos de *Threshold*.

## 6.2 Estimação e Testes de Especificação

A estratégia que utilizamos na busca da especificação correta do modelo de mudança markoviana adequado aos nossos dados foi uma mistura de um procedimento geral para o específico com análises de especificação. Este procedimento é devido a intensa carga computacional presente na estimação desses modelos e ao fato de que o teste do número de regimes não tem distribuição padrão, devido a presença de parâmetros de incômodo.

Partimos de um modelo geral com 10 regimes e 12 defasagens e mudanças de regime no intercepto, nos parâmetros auto-regressivos e na variância, o que, usando a notação de Krolzig [30] é o um modelo MSIAH(10)-ARX(12) (*Markov Switching Intercept Autoregressive Heterocedasticity*). Também incluímos na especificação do modelo uma variável de tendência que se mostrou necessária para a identificação do modelo e estatisticamente significativa. O modelos foram estimados utilizando-se o algoritmo EM de Dempster, Laird e Rubin [12] na forma do filtro BHLK (Baum-Lindgren-Hamilton-Kim) usando o software MSVAR de Krolzig [30]. Detalhes deste método podem ser encontrados na monografia de Krolzig [30](Capítulos 5, 6 e 8).

Tabela 5: Log-Verossimilhança e Critérios de Informação

Regimes	Log-Lik	Parâmetros	AIC	BIC	<i>Nuisance</i>	Restrições
10	8944.3934	170	-9.3794	-8.8765	90	72
9	8924.5196	144	-9.3859	-8.9600	72	64
8	8908.4118	120	-9.3943	-9.0394	56	56
7	8883.6830	98	-9.3974*	-9.1016	42	48
6	8858.2377	78	-9.3856	-9.1549	30	40
5	8832.4047	60	-9.3772	-9.1998*	20	32
4	8769.3628	44	-9.3270	-9.1968	12	24
3	8611.4266	30	-9.1731	-9.0844	6	16
2	8309.3226	18	-8.8630	-8.8097	2	8
1	6516.8861	8	-6.9577	-6.9340	0	0

### 6.3 Determinação do Número de Regimes

O principal problema na determinação da especificação adequada de um modelo de mudança markoviana é determinar o número de regimes. Testes para determinar se a hipótese nula de  $n - 1$  regimes contra a hipótese alternativa de  $n$  regimes não têm distribuição padrão, já que a hipótese nula é não identificada devido à presença dos parâmetros de incômodo (*nuisance parameters*). O procedimento usual de testar esta hipótese através de um teste de razão de verossimilhança não é válido pelo fato de que, na hipótese nula as probabilidades, associadas ao regime adicional não são identificadas, o que viola as condições normais de regularidade para esse teste.

Embora alguns procedimentos para a derivação da distribuição assintótica tenham sido propostos por Hansen [26] e Garcia e Perron [20], eles não são válidos para nosso modelo geral e exigem a simulação dos dados condicionados a um *grid* de valores para os parâmetros de incômodo, o que exigiria uma demorada simulação para cada especificação testada. Na determinação do número de regimes utilizaremos três métodos. O primeiro é o uso dos critérios de informação, já que é demonstrado que os critérios de informação de Akaike e Schwartz nunca subestimam o número mínimo de regimes, conforme Rydén [39]. O segundo é usar uma aproximação da distribuição assintótica do teste, devida a Ang e Bekaert [2]. A terceira é realizar testes de especificação para checar a necessidade de um regime adicional.

Na Tabela 5 vemos a log-verossimilhança, os critérios de informação AIC e BIC, e o número de parâmetros, restrições e parâmetros de incômodo associados a cada regime. Estimamos modelos MSIAH com número de regimes de 10 a 2, e o modelo com um regime que corresponde ao modelo linear. O número de defasagens para a parte autoregressiva foi selecionado através dos critérios de informação de BIC, escolhendo um número de 5 defasagens como o mais adequado. Pela Tabela 5 vemos que o número de regimes selecionado pelo critério de informação de Akaike (AIC) corresponde ao modelo com 7 regimes, enquanto que o critério de informação de Bayes (BIC) seleciona 5 regimes.

Ang e Bekaert [2] mostram que a distribuição assintótica do teste de razão de verossimilhança entre  $n - 1$  e  $n$  regimes pode ser aproximada por uma distribuição qui-quadrada, onde o número de graus de liberdade é dado pelo número de parâmetros de incômodo do modelo com  $n$  regimes mais o número de restrições impostas do regime  $n$  para o regime  $n - 1$ . A estatística de teste é calculada da forma usual em testes de razão de verossimilhança,  $LR=2(\log\text{-verossimilhança}(n)-\log\text{-verossimilhança}(n-1))$ , onde  $n$  e  $n-1$  são os modelos com  $n$  e  $n-1$  regimes.

Pela Tabela 6 este procedimento de teste indica a um nível de significância de 1% a necessidade de um modelo com 5 regimes, enquanto que a um nível de significância de 5% ele aponta para a necessidade de um modelo com 6 regimes. Este teste, no entanto, é baseado em uma aproximação dos valores críticos corretos, e por isso necessitamos de mais suporte para a decisão quanto ao número ótimo de regimes.

Tabela 6: Teste LR Número de Regimes

Teste	Estat.	Dist	<i>p-value</i>
1 contra 2	3854.873	Chi <sup>2</sup> (10)	0.0000 **
2 contra 3	604.208	Chi <sup>2</sup> (14)	0.0000 **
3 contra 4	315.8724	Chi <sup>2</sup> (20)	0.0000 **
4 contra 5	126.0838	Chi <sup>2</sup> (28)	0.0000 **
5 contra 6	51.666	Chi <sup>2</sup> (38)	0.0487*
6 contra 7	50.8956	Chi <sup>2</sup> (50)	0.4381
7 contra 8	49.4576	Chi <sup>2</sup> (64)	0.9095
8 contra 9	32.2156	Chi <sup>2</sup> (80)	1.0000
9 contra 10	39.7476	Chi <sup>2</sup> (98)	1.0000

Tabela 7: Teste BDS Especificação

Teste BDS Modelo 6 Regimes				Teste BDS Modelo 7 Regimes			
Dimensão	BDS Estat.	D. Padrão	<i>p-valor</i>	Dimensão	BDS Esta.	D. Padrão	<i>p-valor</i>
2	-0.002602	0.001363	0.0461	2	-0.001400	0.001307	0.2840
3	-0.004220	0.002159	0.0406	3	-0.000443	0.002068	0.8303
4	-0.003577	0.002562	0.1628	4	0.001066	0.002451	0.6638
5	-0.003316	0.002661	0.2128	5	0.002315	0.002543	0.3626
6	-0.002880	0.002557	0.2600	6	0.002775	0.002441	0.2555

O critério que utilizamos como forma de decidir qual o número de regimes necessário vem da idéia de verificar por meio de testes de especificação a necessidade de um regime adicional. Um teste que se mostra adequado neste contexto é a estatística BDS, um teste robusto para determinar a presença de estruturas remanescentes tanto na média quanto na variância do processo, e assim é eficaz para verificar se a especificação proposta consegue capturar toda a estrutura existente com  $n - 1$  regimes.

A Tabela 7 mostra o resultado do teste BDS aplicado aos resíduos dos modelos com 6 e 7 regimes. Nos resíduos do modelo com 6 regimes, a estatística BDS rejeita a 5% que eles sejam IID nas dimensões 2 e 3, mostrando que, com 6 regimes, há evidência de que existe alguma estrutura na média e/ou variância ainda não capturada. Para verificar se 7 regimes são suficientes para capturar toda a estrutura existente na média e variância aplicamos a estatística BDS aos resíduos do modelo com 7 regimes. Os resultados do teste mostram que não conseguimos rejeitar que os resíduos do modelo com MSIAH(7)-ARX(5) sejam IID, o que indica que devemos trabalhar com um número de regimes igual a 7, capturando toda a estrutura de dependência existente na série de log-retornos de câmbio.

## 6.4 Determinação da Forma Funcional

Havíamos partido de um modelo geral MSIAH com mudança de regimes em todos os parâmetros, e a partir desta especificação, determinamos o número adequado de regimes. É obrigatório no entanto verificar se é necessário sujeitar todos os parâmetros do modelo à mudança de regime, já que um princípio de modelagem econométrica é sempre buscar o modelo mais parcimonioso possível.

Ao contrário da determinação do número de regimes, o teste da especificação da forma funcional mais adequada, em relação aos parâmetros sujeitos à mudança markoviana, tem distribuição



Tabela 8: Testes LR Forma Funcional

Modelo Irrestrito	Modelo Restrito	Restrições	Valor Teste	Valor Crítico	<i>p-value</i>
MSIAH(7) 8883.683	MSIA(7) 7667.708	6	2341.94	14.0671	[0.0000] **
MSIAH(7) 8883.683	MSIH(7) 8743.989	35	279.38	49.8118	[0.0000] **
MSIAH(7) 8883.683	MSAH(7) 8866.492	6	34.38	14.0671	[0.0000] **
MSIAH(7) 8883.683	MSI(7) 7134.9259	42	3497.51	58.1240	[0.0000] **

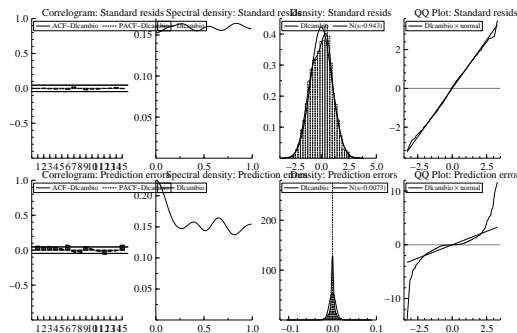
Tabela 9: Testes Especificação Resíduos

Resíduos Padronizados	dist.	Estat.	<i>p-value</i>
portmanteau(31)	Chi(26)	22.4019	[0.6665]
normalidade	Chi(2)	10.7503	[0.004] **
assint.norm.	Chi(2)	9.7392	[0.0077] **
heterocedasticidade	Chi(12)	12.6262	[0.3968]
hetero-X test	Chi(27)	39.6994	[0.0546]
hetero Quadrado:	Chi(12)	12.6440	[0.3954]
hetero-X Quadrado	Chi(27)	39.7478	[0.0541]
ARCH(5)	Chi(5)	5.44	[0.3536]

Qui-Quadrado padrão. Usando o princípio de razão de verossimilhança, testamos o modelo geral MSIAH (mudanças no intercepto, parâmetros auto-regressivos e variância) contra as demais especificações possíveis. Conforme a Tabela 8, vemos que pelo teste LR (razão de verossimilhança), a especificação adequada é a do modelo MSIAH, com mudanças em todos os parâmetros (intercepto, parâmetros autoregressivos e variância incondicional).

A análise dos resíduos, presente na Tabela 9 e na Figura 7 mostra que os os resíduos padronizados do modelo de Mudança Markoviana são não correlacionados, homocedásticos e não têm estrutura de variância condicional do tipo ARCH. Os resíduos padronizados ainda rejeitam a normalidade, mas na Figura 7 vemos, pela densidade empírica e pelo gráfico Q-Q, que os desvios em relação a normalidade são aparentemente pouco significantes em contraste com os resíduos do modelo GARCH (Tabela 2).

Figura 7: Análise dos Resíduos



## 6.5 Modelo Estimado

O modelo estimado MSIAH(7)-ARX(5) corresponde à seguinte especificação:

$$\left\{ \begin{array}{l} y_{1t} = c_{1t} + \phi_{11}y_{1t-1} + \phi_{12}y_{1t-2} + \phi_{13}y_{1t-2} + \phi_{14}y_{1t-4} + \phi_{15}y_{1t-5} + \beta_1t + \varepsilon_{1t} \quad \varepsilon_{1t} \sim N(0, \sigma_1^2) \\ y_{2t} = c_{2t} + \phi_{21}y_{2t-1} + \phi_{22}y_{2t-2} + \phi_{23}y_{2t-2} + \phi_{24}y_{2t-4} + \phi_{25}y_{2t-5} + \beta_2t + \varepsilon_{2t} \quad \varepsilon_{2t} \sim N(0, \sigma_2^2) \\ y_{3t} = c_{3t} + \phi_{31}y_{3t-1} + \phi_{32}y_{3t-2} + \phi_{33}y_{3t-2} + \phi_{34}y_{3t-4} + \phi_{35}y_{3t-5} + \beta_3t + \varepsilon_{3t} \quad \varepsilon_{3t} \sim N(0, \sigma_3^2) \\ y_{4t} = c_{4t} + \phi_{41}y_{4t-1} + \phi_{42}y_{4t-2} + \phi_{43}y_{4t-2} + \phi_{44}y_{4t-4} + \phi_{45}y_{4t-5} + \beta_4t + \varepsilon_{4t} \quad \varepsilon_{4t} \sim N(0, \sigma_4^2) \\ y_{5t} = c_{5t} + \phi_{51}y_{5t-1} + \phi_{52}y_{5t-2} + \phi_{53}y_{5t-2} + \phi_{54}y_{5t-4} + \phi_{55}y_{5t-5} + \beta_5t + \varepsilon_{5t} \quad \varepsilon_{5t} \sim N(0, \sigma_5^2) \\ y_{6t} = c_{6t} + \phi_{61}y_{6t-1} + \phi_{62}y_{6t-2} + \phi_{63}y_{6t-2} + \phi_{64}y_{6t-4} + \phi_{65}y_{6t-5} + \beta_6t + \varepsilon_{6t} \quad \varepsilon_{6t} \sim N(0, \sigma_6^2) \\ y_{7t} = c_{7t} + \phi_{71}y_{7t-1} + \phi_{72}y_{7t-2} + \phi_{73}y_{7t-2} + \phi_{74}y_{7t-4} + \phi_{75}y_{7t-5} + \beta_7t + \varepsilon_{7t} \quad \varepsilon_{7t} \sim N(0, \sigma_7^2) \end{array} \right. \quad (25)$$

Os parâmetros estimados estão na Tabela 10, em conjunto com os seus desvios-padrões e estatísticas t. Os parâmetros estimados são muito interessantes. Enquanto existem regimes com todos os parâmetros significativos, como o regime 1, no regime 2 nenhum parâmetro é estatisticamente significativo. Outro fato interessante é notar que a tendência só é diferente de zero no regime 7, mas a necessidade da tendência era fosse confirmada por testes de especificação. A matriz de transição estimada se encontra na Tabela 11.

A Figura 8 contém as probabilidades estimadas de cada regime para cada observação na amostra. O gráfico mostra as probabilidades previstas, filtradas (usando a informação até o período t) e as probabilidades suavizadas (usando a informação da amostra toda para inferir as probabilidades no momento t). Este gráfico nos mostra que o modelo associa ao regime de bandas cambiais três regimes exclusivos (regimes 3, 4 e 5), enquanto que os demais regimes são presentes nos períodos restantes da amostra. Os regimes 2 e 6 são identificados com os períodos de variação livre do câmbio, enquanto que os regimes 1 e 7 podem ser identificados, conforme a Tabela 10, com os períodos de crise cambial, combinando elevada variância na taxa de câmbio com tendências a forte desvalorização.

Esta associação entre regimes também pode ser notada na matriz de transição, já que as probabilidades de transição dos regimes 3, 4 e 5 para os demais regimes são quase nulas, e entre eles são bastante significantes. Podemos interpretar a existência de 3 regimes dentro do período de bandas cambiais como a presença de 2 regimes agindo como barreiras superiores e inferiores na banda, e um terceiro regime como valor do câmbio dentro dos limites da banda. Pelos parâmetros estimados para estes regimes podemos identificar o regime 3 como sendo o de valores dentro dos limites normais da banda cambial, enquanto que o regime 4 teria valores próximos a barreira inferior e, o regime 5, os retornos que levariam a proximidade do limite superior da banda cambial.

Os autovalores da matriz de transição estão colocados na Tabela 14. Como o primeiro autovalor é igual a um e os demais autovalores estão dentro do círculo unitário, a matriz de transição é ergódica, já que é também irredutível. Desta forma, o autovetor associado ao autovalor unitário presente em 14 representa as probabilidades ergódicas do processo. Esse vetor também indica a probabilidade incondicional de cada regime, e desta forma construímos a Tabela 13. O fato de a matriz de transição ser ergódica confirma que o processo é estacionário, já que, conforme Hamilton [24] (pgs 681 e 682), demonstra-se que um processo de mudança markoviana com uma matriz de transição ergódica é sempre covariância-estacionário.

## 6.6 Falta de Parcimônia no Modelo Final Estimado

O modelo MSIAH(7)-ARX(5) contém um grande número de parâmetros estimados, e assim não pode ser considerado um modelo parcimonioso pelos padrões usuais em econometria. Entretanto a existência de quebras estruturais e não linearidades tanto na média quanto na variância, quando se estimando modelos lineares por partes, como o caso do modelo MSIAH, levam a necessidade um grande número de parâmetros.

O número de regimes escolhido, que gera de forma multiplicativa o número de parâmetros no modelo final, foi justificado pelos critérios de Informação de Akaike e Schwarz, este último um critério com propriedades de super consistência para apontar o número de regimes e parâmetros

Tabela 10: Parâmetros Estimados

Regime 1	Coef.	D.Padrão	<i>t-value</i>	Regime 2	Coef.	D.Padrão	<i>t-value</i>
Const	-0.0209	0.0019	-11.1947	Const	-0.0004	0.0002	-1.9478
Dlcambio_1	-0.1225	0.0940	-1.3037	Dlcambio_1	-0.0007	0.0635	-0.0114
Dlcambio_2	-0.3699	0.0892	-4.1483	Dlcambio_2	-0.0083	0.0539	-0.1547
Dlcambio_3	-0.5672	0.0996	-5.6919	Dlcambio_3	-0.0427	0.0527	-0.8097
Dlcambio_4	-0.2190	0.0906	-2.4176	Dlcambio_4	0.0837	0.0484	1.7289
Dlcambio_5	-0.7564	0.0791	-9.5628	Dlcambio_5	-0.0021	0.0497	-0.0414
Trend	0.0000	0.0000	0.4548	Trend	0.0000	0.0000	0.0422
D.Padrão	0.0095764			D.Padrão	0.0042088		

Regime 3	Coef.	D.Padrão	<i>t-value</i>	Regime 4	Coef.	D.Padrão	<i>t-value</i>
Const	0.0002	0.0001	2.0211	Const	-0.0000	0.0000	-1.9315
Dlcambio_1	-0.0146	0.0663	-0.2201	Dlcambio_1	-0.0150	0.0266	-0.5664
Dlcambio_2	-0.1437	0.0622	-2.3103	Dlcambio_2	-0.0873	0.0180	-4.8547
Dlcambio_3	-0.0203	0.0607	-0.3343	Dlcambio_3	0.0309	0.0154	2.0102
Dlcambio_4	-0.1258	0.0526	-2.3922	Dlcambio_4	-0.0646	0.0136	-4.7414
Dlcambio_5	0.0795	0.0674	1.1790	Dlcambio_5	-0.0241	0.0151	-1.5928
Trend	0.0000	0.0000	0.0345	Trend	0.0000	0.0000	0.0333
D.Padrão	0.0004381			D.Padrão	0.0001678		

Regime 5	Coef.	D.Padrão	<i>t-value</i>	Regime 6	Coef.	D.Padrão	<i>t-value</i>
Const	0.0021	0.0001	15.8368	Const	-0.0000	0.0004	-0.1106
Dlcambio_1	-0.1979	0.0713	-2.7755	Dlcambio_1	0.1077	0.0486	2.2137
Dlcambio_2	-0.1540	0.0745	-2.0686	Dlcambio_2	-0.1328	0.0453	-2.9307
Dlcambio_3	-0.1840	0.0735	-2.5049	Dlcambio_3	0.1168	0.0458	2.5513
Dlcambio_4	-0.1806	0.0752	-2.4018	Dlcambio_4	-0.0053	0.0436	-0.1227
Dlcambio_5	-0.2012	0.0758	-2.6533	Dlcambio_5	0.0421	0.0420	1.0007
Trend	-0.0000	0.0000	-0.2335	Trend	0.0000	0.0000	0.0860
D.Padrão	0.0013201			D.Padrão	0.0084794		

Regime 7	Coef.	D.Padrão	<i>t-value</i>
Const	0.0924	0.0062	14.9538
Dlcambio_1	-0.2350	0.1496	-1.5707
Dlcambio_2	0.1601	0.1539	1.0405
Dlcambio_3	0.2282	0.1587	1.4377
Dlcambio_4	-0.1669	0.1628	-1.0256
Dlcambio_5	-0.2677	0.1648	-1.6237
Trend	-0.0001	0.0000	-3.6738
D.Padrão	0.0227350		

Tabela 11: Matriz de Transição

	regime1	regime2	regime3	regime4	regime5	regime6	regime7
regime1	0.7122	0.0009091	2.585e-008	6.118e-009	1.005e-007	0.2869	1.217e-011
regime2	8.171e-007	0.9744	8.937e-008	2.735e-008	0.004562	0.02103	3.011e-011
regime3	2.909e-011	9.485e-008	0.8714	0.06670	0.06194	9.380e-009	1.885e-015
regime4	0.003840	1.068e-007	0.1716	0.7283	0.09189	1.926e-008	0.004342
regime5	0.004077	6.174e-007	0.06717	0.1963	0.7324	7.645e-008	7.992e-015
regime6	0.01168	0.02171	1.034e-008	0.002224	4.566e-008	0.9558	0.008568
regime7	0.1250	8.225e-007	3.491e-011	8.977e-012	1.277e-010	1.128e-007	0.8750

Note que  $p[i|j]=\Pr\{s(t+1)=j|s(t)=i\}$

Figura 8: Probabilidades Estimadas

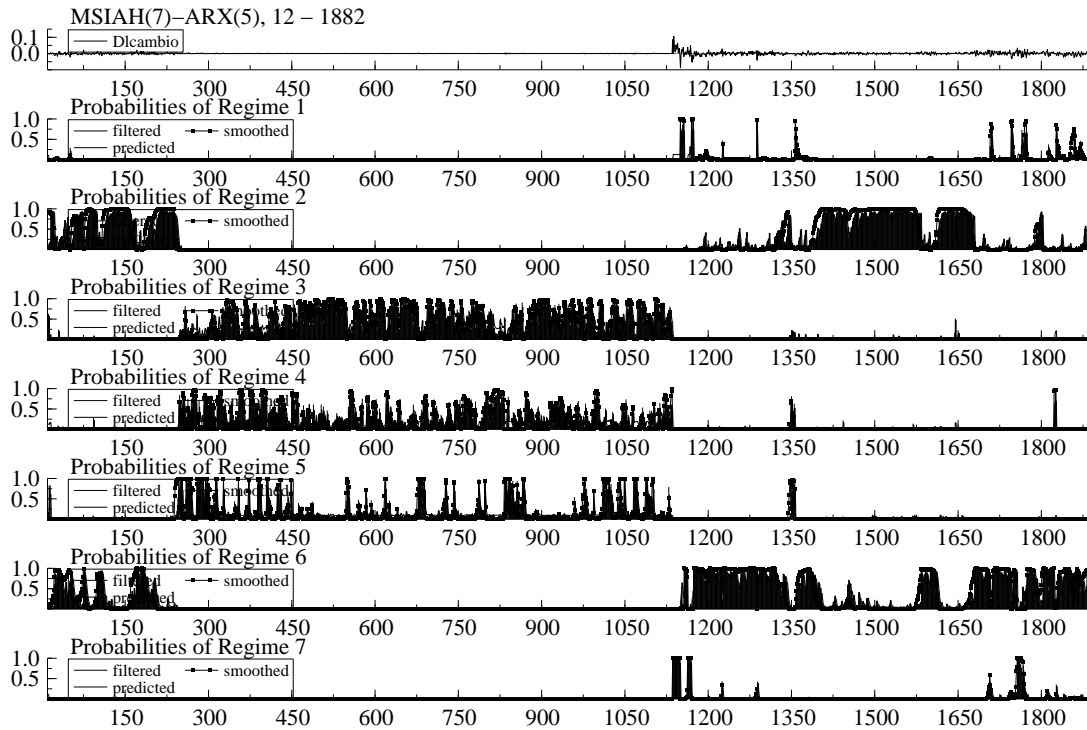


Tabela 13: Regimes e Duração

	número de Observações	Probabilidade Incondicional	Duração
Regime 1	42.3	0.0233	3.47
Regime 2	420.7	0.2171	39.08
Regime 3	452.9	0.2388	7.77
Regime 4	256.4	0.1374	3.68
Regime 5	201.1	0.1062	3.74
Regime 6	457.3	0.2549	22.63
Regime 7	40.2	0.0222	8.00

Tabela 14: Autovalores da Matriz de Transição

real	1.0000	0.99353	0.95765	0.85702	0.73215	0.70606	0.60312
------	--------	---------	---------	---------	---------	---------	---------

Figura 9: Valores Observados e Previstos

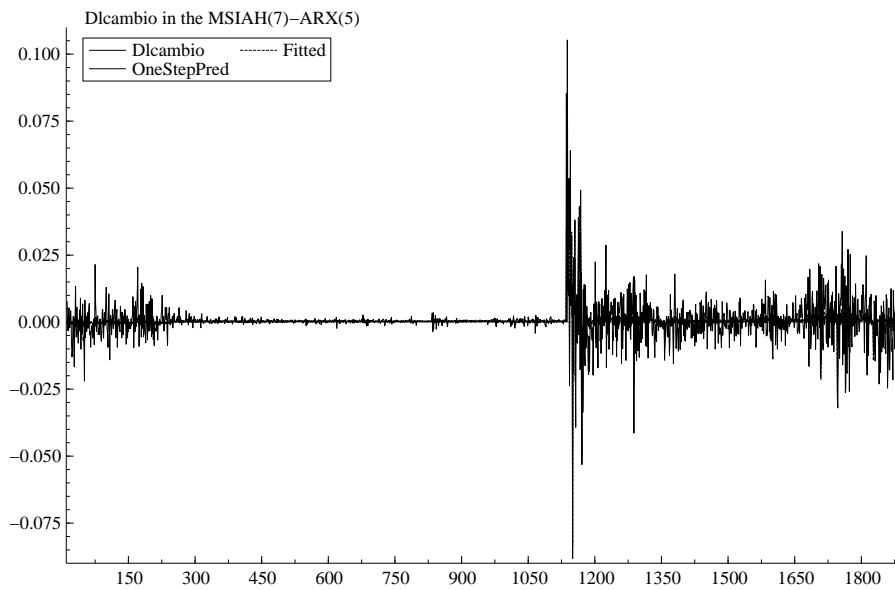
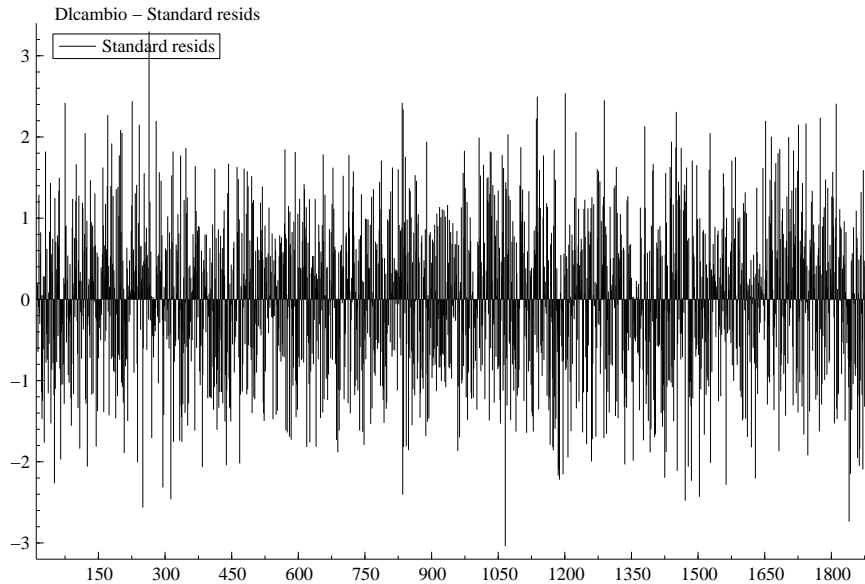


Figura 10: Resíduos Padronizados



necessários, conforme demonstrado por Ryden [39]. O critério de Schwarz, que foi o que indicava o menor número de regimes, apontava à necessidade de 5 regimes e 60 parâmetros estimados, o que leva à um modelo pouco parcimonioso, mas corretamente especificado. Os testes para o número de regimes baseados na aproximação de Ang e Bekaert e as análises dos resíduos continuavam a indicar a necessidade de um grande número de regimes e consequentemente parâmetros. O único modelo que obtinha resíduos I.I.D. era o modelo com 7 regimes.

Conforme Poskitt e Chung [38], a função de autocorrelação de um modelo com mudanças markovianas no intercepto e variância pode ser aproximado por um modelo linear  $ARMA(p+k-1, p+k-1)$ , aonde  $p$  é o número de parâmetros autoregressivos e  $k$  o número de regimes. O modelo linear que estimamos para nossos dados era um modelo  $AR(12)$ , o que seria aproximadamente consistente com um modelo  $MSIH$  com cinco regimes e cinco termos autoregressivos, já que estamos comparando com um modelo  $AR$  puro sem termos  $MA$  e nosso modelo contém mudanças nos parâmetros autoregressivos, mas serve como evidência da necessidade de um modelo pouco parcimonioso conforme os padrões usuais.

Outra forma de verificar a necessidade de 7 regimes é notar que temos 3 regimes que vigoram no período da banda cambial e 4 fora da banda, sendo que desses últimos 2 estão identificados com crises. Se segmentássemos a amostra e estimássemos modelos para cada período, o número de regimes necessário seria bem inferior ao do modelo para toda a amostra, já que os regimes identificados para o regime cambial e os dados com livre flutuação são quase que excludentes. Desta forma um modelo para a amostra toda com 7 regimes captura estas quebras estruturais existentes, sem nenhuma necessidade de informação ad hoc, como a introdução de *dummies* ou procedimentos de identificação de outliers.

## 6.7 Modelo SWARCH

Uma forma de se buscar um modelo mais parcimonioso e um critério de comparação para o modelo  $MSIAH(7)-ARX(5)$ , para modelar a estrutura de dependência nos retornos incluindo às mudanças de regimes, seria o modelo  $SWARCH$  de Hamilton e Susmel [25]. Entretanto ocorreram vários problemas na aplicação do modelo  $SWARCH$  para esta série. O principal problema era

um problema de convergência na estimação. Foram testadas diversas especificações possíveis, e para a maioria delas os resultados eram de falhas na maximização e quando a maximização era possível ocorria a não invertibilidade do Hessiano, assim impossibilitando o cálculo da variância dos estimadores.

Este problema é causado pelo fato de uma especificação parcimoniosa com 2 ou 3 regimes para o modelo SWARCH não captura toda a dinâmica existente nesta série, e quando da aplicação deste modelo, os testes de especificação nos resíduos apontavam para a necessidade de mais termos para a média e variância condicional. Nenhuma das especificações testadas resultava em um ruído branco nos resíduos, o que acontece no modelo MSIAH. A necessidade de inclusão de mais parâmetros, ou seja, uma especificação menos parcimoniosa, levava ao problema da não invertibilidade do Hessiano, o que está relacionado ao fato de que a matriz de Informação singular, um sintoma de que o modelo estimado é não identificado. Como o modelo SWARCH com mais parâmetros era não identificado, e em adição não eliminava toda a estrutura de dependência existente na série e tinha problema graves de convergência, este modelo se mostrou inadequado para modelar a série de log-retornos do câmbio no período estudado. Possivelmente os problemas de não identificação do modelo SWARCH se devem ao fato de que existe mudança na estrutura dos parâmetros tanto da média, quanto da variância condicional, e o modelo SWARCH captura apenas mudanças de regime na escala da volatilidade de um modelo GARCH.

## 7 Análise de Monte Carlo - Propriedades dos estimadores GARCH na presença de múltiplos regimes

Para verificar o comportamento do modelo GARCH com nossos dados, aonde assumimos que existam mudanças de regimes modeladas por uma cadeia de Markov, realizamos um experimento de Monte Carlo para verificar o comportamento do modelo GARCH(1,1) estimado sob condições de mudança estrutural no processo estatístico gerador da série.

O experimento foi realizado gerando 2000 replicações do processo definido pelos parâmetros da Tabela 10 e a matriz de transição disposta na Tabela 11. Para cada replicação estimamos um modelo ARMA(0,0)-GARCH(1,1) e coletamos os valores da soma dos parâmetros ARCH(1) e GARCH(1) ( $\alpha_1 + \beta_1$ ) que medem a persistência dos choques na variância. A especificação fixa ARMA(0,0) é baseada em duas razões. A primeira é um problema de implementação, já que, se para cada processo estimado fosse realizada uma busca pela melhor especificação da média, a complexidade e principalmente o tempo necessário à realização do experimento seriam muito maiores. A segunda razão é o fato de que mesmo sendo o processo gerador dos dados um AR(5) para cada regime, a transição de regimes faz com que a ordem do processo AR selecionada para o modelo linear seja um AR(0), conforme descrito na Seção 3. Desta forma, o experimento fica mais consistente com o fato de que geralmente não conhecemos o processo gerador dos dados quando trabalhamos com séries reais. Seria entretanto necessário um estudo adicional para verificar se a especificação do termo AR no modelo GARCH influencia os resultados do estudo de Monte Carlo realizado.

Geramos séries com tamanhos de amostra 50, 100, 250, 500, 1000, 2500 e 5000 observações. Para o conjunto de 2000 replicações de cada tamanho de amostra calculamos a média e a variância da soma dos termos ARCH(1) e GARCH(1), e, para ilustrar a distribuição empírica da persistência estimada, geramos histogramas para os tamanhos de amostra 100, 250, 500 e 1000. Os resultados estão na Tabela 15.

Os resultados do experimento mostram que quando o processo gerador dos dados é sujeito a mudanças nos parâmetros da média, variância e parâmetros autoregressivos, a estimação do modelo GARCH(1,1) se torna extremamente instável e dependente do tamanho da amostra. Era esperado que, nas diferentes amostras, um mesmo valor médio para a persistência na variância fosse encontrado, e que com o aumento do tamanho da amostra a variância do valor estimado para a persistência diminuísse. Mesmo esperando que o coeficiente de persistência fosse negativamente viesado devido aos problemas de amostras pequenas reportados por Hwang e Valls Pereira [29], os

Tabela 15: Monte Carlo - GARCH

Tamanho Amostra	Média	Variância
50	0.3742	0.2793
100	0.5431	0.3395
250	0.7722	0.2546
500	0.8563	0.1513
1000	0.8947	0.0681
2500	0.9112	0.0205
5000	0.9194	0.009

Figura 11: Histograma Persistência Estimada (Amostras Tamanho 100 e 250)

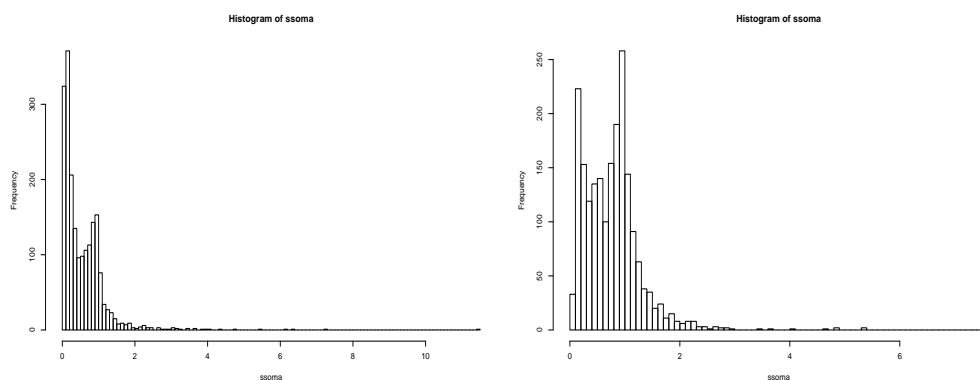
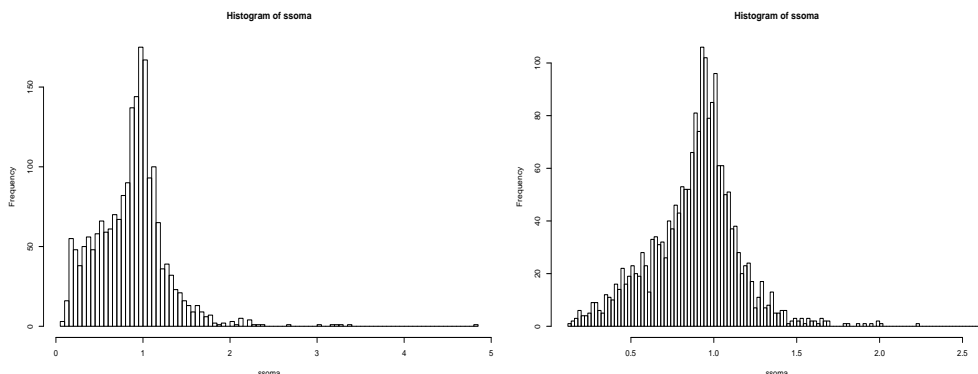




Figura 12: Histograma Persistência Estimada (Amostras Tamanho 500 e 1000)



valores encontrados no procedimento de Monte Carlo mostram que o modelo GARCH é inadequado na presença de mudanças de parâmetros (regimes).

O resultado obtido mostra que, com o aumento do tamanho da amostra o valor médio estimado para a persistência aumenta, partindo de um valor médio de 0.3742 para as 2000 replicações com amostra de tamanho 50, passando por 0.3742 para uma amostra de tamanho 500 e atingindo um valor de 0.9194 para a amostra de tamanho 5000.

Observando os histogramas (Figuras 11 e 12) notamos que a distribuição empírica dos estimadores é muito assimétrica para tamanhos menores de amostra, como mostram os histogramas para os tamanhos 100 e 250, e, com o aumento do tamanho da amostra, eles se tornam mais simétricos e aparentemente convergem para uma distribuição normal, como indica o histograma associado às amostras de tamanho 1000.

O experimento de Monte Carlo nos permite concluir que o valor da persistência estimada é diretamente relacionada ao tamanho da amostra. Em um estimador consistente é esperado que a variância do estimador seja elevada para um número reduzido de observações, e que este valor tenda ao verdadeiro valor do parâmetro com o aumento da amostra.

A simulação de Monte Carlo mostra claramente que existe algum tipo de viés relacionado ao tamanho da amostra na estimação da persistência no modelo GARCH(1,1) além do problema do tamanho de amostra apontado por Hwang e Valls Pereira [29], gerado pela existência de múltiplos regimes no processo gerador da série, invalidando as estimativas da persistência do modelo GARCH(1,1). Como o processo estimado não é o modelo gerador dos dados, não é trivial discutir se o estimador é assintoticamente consistente ou não, já que necessitaríamos discutir uma forma de comparação entre os parâmetros dos dois modelos. Na Seção 8.1 mostramos que a variância no modelo de mudança markoviana utilizado no estudo é dependente do regime vigente determinado pela matriz de transição modelada por uma cadeia de Markov, e assim não existe um valor de persistência único para a amostra, como é estimado pelo GARCH(1,1).

## 8 Comparação com a Variância Prevista no Modelo de Mudança Markoviana

Para verificar o comportamento das previsões dinâmicas para a variância condicional do modelo GARCH(1,1) em comparação com as do modelo de mudança markoviana (MSIAH) estudado na Seção 6. Para esta análise verificamos de forma visual o comportamento da previsão da variância realizada por cada modelo em horizontes de 30 e 100 observações, primeiro utilizando o final da amostra (observações de 1850 a 1880) e uma previsão para as observações entre 500 e 600. Os

critérios usuais de comparação de previsões para a volatilidade são baseados em comparar o erro quadrático médio entre as previsões para a variância condicional e os resíduos ao quadrado do modelo não são diretamente aplicáveis aqui, já que embora o modelo GARCH seja uma representação para os resíduos ao quadrado, no modelo de Mudança Markoviana a variância em cada regime é a variância incondicional, sendo dois critérios não diretamente comparáveis, já que a variância não é diretamente observada.

A realização das previsões dinâmicas no modelo GARCH(1,1) é trivial, e pode ser realizada com a representação do processo (8) utilizando-se os parâmetros estimados, o que equivale a uma previsão a  $n$  passos à frente de um processo ARMA(1,1), conferindo o decaimento exponencial observado.

## 8.1 Previsão da Variância no Modelo MSH (Mudança Markoviana na Variância)

A previsão da variância no modelo de mudança markoviana é baseada na previsão do regime vigente  $n$  passos à frente condicionada às probabilidades vigentes para cada regime no período  $t$ . Desta forma a previsão da variância é a esperança dos valores da variância previstos  $n$  passos a frente, ou seja, a soma dos valores estimados em cada regime para a variância incondicional, multiplicados pela probabilidade do regime no período. Diferente do modelo GARCH, aonde a variância está condicionada aos erros da estrutura da média ao quadrado e a própria variância condicional no período passado, neste modelo de mudança markoviana a previsão condicional para a variância é condicionada às probabilidades estimadas para cada regime no período imediatamente anterior, o que permite incorporar na previsão às probabilidades de mudanças de regimes, assim evitando o problema de persistência espúria que acontece no modelo GARCH quando existem mudanças no processo gerador da série.

A questão fundamental para a previsão no modelo de mudança Markoviana é determinar a probabilidade de vigência de cada regime  $n$  períodos à frente. Dado que a estrutura de transição dos regimes é determinada por uma cadeia de Markov, necessitamos de uma equação que determine previsões para a cadeia de Markov.

Seguindo Hamilton [24] (pgs 679-680) vemos que para realizar previsões para uma cadeia de Markov, uma representação útil é reescreve-la como uma autoregressão vetorial (VAR). Definindo um vetor aleatório ( $N \times 1$ )  $\xi_t$ , cujo elemento  $j$  é igual a 1 se  $s_t = j$  e zero se  $s_t \neq j$ , onde  $s_t$  define o valor da variável aleatória  $s$  no período. A probabilidade de que  $s_t$  seja igual a  $j$  é determinada por uma cadeia de Markov, conforme a representação usada na Seção 6. O vetor  $\xi_t$  pode ser representado como:

$$\xi_t = \begin{cases} (1, 0, 0, \dots, 0), & \text{se } s_t = 1 \\ (0, 1, 0, \dots, 0), & \text{se } s_t = 2 \\ (0, 0, 1, \dots, 0), & \text{se } s_t = 3 \\ (0, 0, 0, \dots, 1), & \text{se } s_t = N \end{cases} \quad (26)$$

Assim, se  $s_t = 1$  o vetor  $\xi_t$  é igual à primeira coluna da matriz identidade, quando  $s_t = 2$ , é igual à segunda coluna da matriz Identidade, e assim por diante. Dado a estrutura markoviana, se  $s_t = i$  o  $j$ -ésimo elemento de  $\xi_{t+1}$  assume valor 1 com probabilidade  $p_{ij}$  e zero nos demais casos. Conforme Hamilton [24], essa matriz tem esperança  $p_{ij}$ , e a esperança condicional  $\xi_{t+1} | s_t = i$  é dada por:

$$E(\xi_{t+1} | s_t = i) = \begin{bmatrix} p_{i1} \\ p_{i2} \\ \vdots \\ p_{iN} \end{bmatrix} \quad (27)$$

O vetor representado por 27 é simplesmente a  $i$ -ésima coluna da matriz de probabilidades de transição definida na Tabela 24. Quando  $s_t = i$ , o vetor  $\xi_t$  corresponde a  $i$ -ésima coluna da matriz

identidade de ordem  $N$ , e assim a esperança condicional é dada por  $P\xi_t$ . Desta forma, a equação 27 implica que

$$E(\xi_{t+1}|\xi_t) = P\xi_t \quad (28)$$

e segue da propriedade markoviana (equação 26) que:

$$E(\xi_{t+1}|\xi_t, \xi_{t-1}, \xi_{t-2}, \dots) = P\xi_t \quad (29)$$

Assim podemos representar uma cadeia de Markov na forma de uma autoregressão vetorial para  $\xi_t$  dada por:

$$\begin{aligned} \xi_{t+1} &= P\xi_t + v_t \\ v_t &\equiv \xi_{t+1} - E(\xi_{t+1}|\xi_t, \xi_{t-1}, \dots) \end{aligned} \quad (30)$$

Desta forma,  $v_t$  é um processo de inovação definido por uma seqüência *martingale* à diferenças com média zero, impossível de ser prevista com base nas realizações anteriores do processo.

A equação 30 implica que:

$$\xi_{t+n} = v_{t+n} + Pv_{t+n-1} + P^2v_{t+n-2} + \dots + P^{n-1}v_{t+1} + P^n\xi_t \quad (31)$$

onde  $P^n$  indica a matriz de transição multiplicada por ela mesma  $n$  vezes. Como o vetor  $v_t$  tem esperança igual a zero, previsões  $n$  períodos a frente para uma cadeia de Markov são dadas por:

$$E(\xi_{t+n}|\xi_t, \xi_{t-1}, \dots) = P^n\xi_t \quad (32)$$

Utilizando as definições anteriores podemos ver que o  $j$ -ésimo elemento do vetor  $\xi_t$  indica a probabilidade de que  $s_{t+n}$  assumo o valor  $j$  condicional aos valores dados no tempo  $t$ . Supondo que o valor de  $s_t$  no tempo  $t$  seja  $i$  a equação 32 estabelece que:

$$\begin{bmatrix} P\{s_{t+n} = 1|s_t = i\} \\ P\{s_{t+n} = 2|s_t = i\} \\ \vdots \\ P\{s_{t+n} = N|s_t = i\} \end{bmatrix} = P^n \cdot e_i \quad (33)$$

onde  $e_i$  é a  $i$ -ésima coluna da matriz identidade de dimensão  $N$ . A expressão 33 estabelece que a probabilidade de que o regime  $i$  será seguido  $n$  períodos a frente pelo regime  $j$  ( $P\{s_{t+n} = j|s_t = i\}$ ), é dada pelo elemento da linha  $i$ , coluna  $j$ , da matriz de transição de probabilidade multiplicada  $n$  vezes por ela mesma.

Para encontrar o valor previsto para a variância  $n$  passos à frente substituímos a matriz identidade por uma matriz  $V_m$  contendo as variâncias associadas a cada regime na diagonal principal e com os demais elementos iguais a zero, na forma:

$$V_m = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \sigma_N^2 \end{bmatrix} \quad (34)$$

Definindo um vetor  $p_e$  contendo as probabilidades estimadas para cada regime no período  $t$ , a esperança da previsão da variância  $n$  passos a frente, condicionada as probabilidades de cada regime em  $t$ , é dada por :

$$E(\sigma_{t+n}^2|p_e) = P^n \cdot V_m \cdot p_e \cdot i \quad (35)$$

Figura 13: Previsões do Modelo de mudança markoviana para a Variância

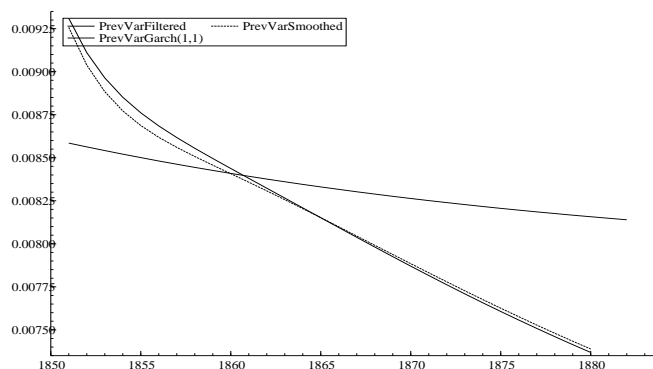
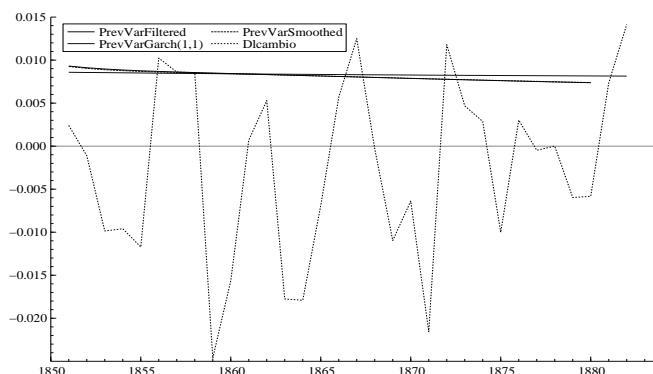


Figura 14: Previsões para a Variância e Valores Observados



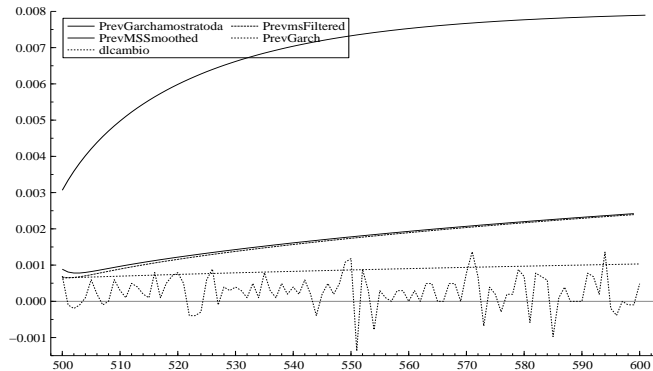
onde  $i$  é um vetor ( $1 \times N$ ) contendo valores iguais a 1. A equação 35 é a esperança das variâncias associadas a cada regime previstas  $n$  períodos à frente condicionada às probabilidades de cada regime no tempo  $t$ .

Comparamos as previsões dinâmicas para a variância condicional em duas situações. A primeira comparação é feita utilizando as observações 1850 a 1880, que correspondem ao final da amostra de log-retornos de câmbio. Utilizamos a amostra toda na estimação dos modelo AR(0)-GARCH (1,1) (Tabela 1) e MSIAH (Tabelas 10, 11 e 8), e calculamos as previsões para esse intervalo utilizando as equações 8 e 35. Calculamos as previsões da variância do modelo de mudança markoviana usando as probabilidades filtradas e suavizadas.

A Figura 13 mostra que as previsões do modelo de mudança markoviana indicam um decaimento muito mais rápido do que o do modelo GARCH(1,1), o que parece mais consistente com o comportamento verificado nos dados. Como os parâmetros estimados no modelo GARCH(1,1) apontam uma persistência de 0.972892 a previsão para a volatilidade condicional ainda tem decaimento exponencial, mas é importante notar que este parâmetro não é estatisticamente diferente de 1, o que indica que a hipótese de que os choques na variância apresentam efeitos permanentes não pode ser rejeitada no modelo GARCH(1,1), como observamos anteriormente.

A comparação dos valores previstos para a variância com os valores observados na série de log-retornos mostra que mesmo com a grande diferença na persistência dos choques calculada pelos dois modelos, para o horizonte de 30 observações, em comparação com os dados observados, os

Figura 15: Previsão - Amostra 500-600 e Valores Observados



valores previstos para a variância são muito próximos, como mostra a Figura 14. Para esta parte da amostra os dois modelos aparentemente subestimam o valor da variância.

Realizamos previsões para o intervalo de observações 500-600, que correspondem na amostra a parte do período de bandas cambiais. Nesse período a volatilidade da série era controlada por intervenções no mercado de câmbio. Para este intervalo realizamos previsões dinâmicas para a variância condicional da série de 4 formas. Duas usando o modelo de mudança markoviana, através do uso das probabilidades filtradas e suavizadas, e usando o modelo GARCH(1,1) prevemos a variância usando os parâmetros estimados usando a amostra toda e usando parâmetros estimados até a observação 500.

Os valores previstos pelo modelo GARCH(1,1) usando toda a informação da amostra são muito maiores do que as demais previsões, embora todos os modelos apontem para uma variância crescente neste horizonte de previsão de 100 passos à frente. As previsões do modelo GARCH(1,1) estimado até a observação 500 no entanto, são bem menores do que os valores previstos pelos outros modelos. As previsões do modelo de mudança markoviana são intermediárias em relação às do modelo GARCH, como mostra a Figura 15.

Em comparação com os valores realizados (Figura 15) vemos que nesse período as previsões para a variância condicional dadas pelo modelo GARCH(1,1) estimado até a observação 500 subestimam a variância, enquanto que o o modelo estimado GARCH(1,1) usando toda a amostra superestima a volatilidade. Quando usamos a amostra toda na estimação do GARCH os valores previstos são muito maiores do que os observados, o que é mais uma evidência da quebra estrutural na variância existente nessa série. As previsões para a variância do modelo de Mudança Markoviana parecem ser as mais adequadas nesse subperíodo. O mais importante é notar que as previsões desse modelo se adaptam às mudanças existentes nos parâmetros dentro da amostra, e que levam a uma superestimação da persistência no modelo GARCH (1,1). Uma vantagem do modelo de Mudança Markoviana é que a variância do modelo é muito mais suave (Figura 18) do que a do modelo GARCH, o que é desejável em contraste com a estrutura de variância extremamente errática que é estimada nos modelos da classe GARCH.

É interessante notar que as previsões do modelo de mudança markoviana, independente das probabilidades no período  $t$ , sempre irão convergir para um mesmo ponto. Este ponto equivale à uma média das variâncias associadas aos 7 regimes, ponderada pelas probabilidades incondicionais de cada regime, e que é aproximadamente igual a 0.003951565. A Figura 16 ilustra essa propriedade.

Uma questão adicional é verificar como podemos aproximar os resultados estimados para a persistência nos dois modelos. Embora a estrutura da variância nos modelos de mudança markoviana (MSH) e no modelo GARCH sejam muito diferentes, para horizontes curtos as previsões do modelo de mudança markoviana aparentemente podem ser aproximadas por uma função de

Figura 16: Previsão 2000 passos e Variância modelo MSIAH

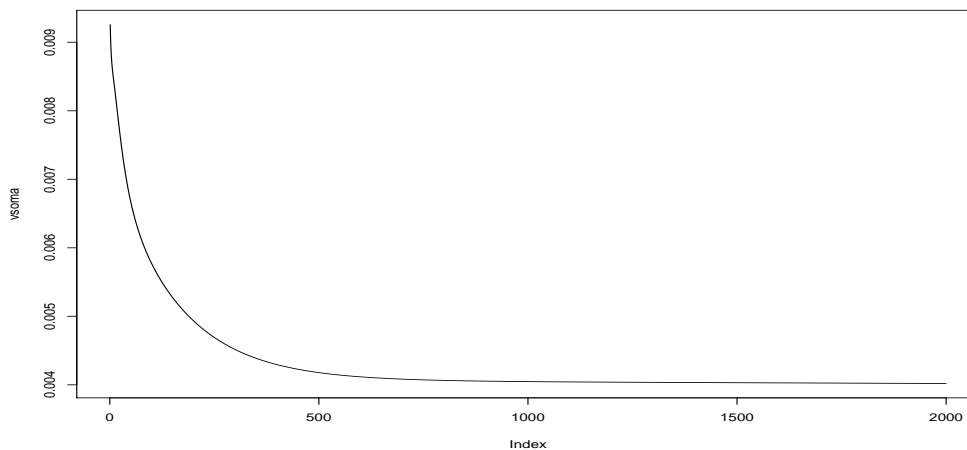


Tabela 16: Aproximação Persistência

Horizonte de previsões	Valor aproximado AR(1)	Horizonte de previsões	Valor aproximado AR(1)
10	0.6356	250	0.9774
15	0.7268	500	0.9840
20	0.7899	1000	0.9866
30	0.8656	2000	0.9875
50	0.9246	5000	0.9881
100	0.9595	10000	0.9889

decaimento exponencial, e assim torna possível comparar diretamente a persistência calculada por esses modelos.

Através das Figuras 17 e 17 que mostram o ACF e o PACF das variâncias previstas para os últimos 30 valores da série de log-retornos de câmbio vemos que estas previsões podem ser aproximadas por um modelo AR(1).

Estimando um modelo AR(1) para os valores previstos pelo modelo de mudança markoviana, utilizando as probabilidades filtradas e suavizadas, obtivemos respectivamente os valores de 0.8656 e 0.8703. Neste procedimento simples de aproximação por um processo autoregressivo, as persistências aproximadas para o modelo de mudança markoviana serão inferiores as do modelo GARCH(1,1) estimado para a série toda, indicando uma persistência menor dos choques, nesse horizonte de 30 passos à frente, uma característica que é desejável em vista do problema de persistência excessiva existente no modelo GARCH(1,1).

É importante notar que, com o aumento do horizonte de previsão, o valor da persistência estimado pelos dois modelos se aproxima. A Tabela 16 mostra que para as primeiras observações a persistência é muito baixa, e com o aumento do intervalo de previsão a persistência aumenta.

Fixando o horizonte de previsões em 2000, adotamos um procedimento oposto ao realizado acima, estimando um modelo AR(1) para as últimas n observações previstas. O resultado desse procedimento, exposto na Tabela 17, mostra que nas últimas n observações ocorre um resultado análogo ao da Tabela 16, onde as últimas observações tem decaimento mais rápido do que as

Figura 17: Autocorrelações e Variâncias - Filtradas e Suavizadas

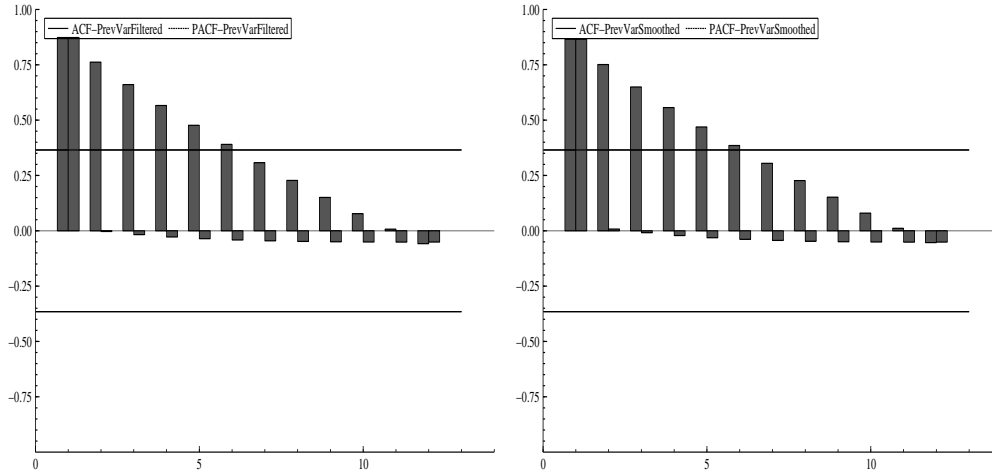


Tabela 17: Aproximação Persistência (n últimas observações)

Horizonte de Previsões	Valor aproximado AR(1)	Horizonte de Previsões	Valor aproximado AR(1)
1900-2000	0.9703	1970-2000	0.9032
1925-2000	0.9605	1980-2000	0.8571
1950-2000	0.9412	1990-2000	0.7273

observações iniciais. Como exemplo as observações 1980-2000 tem persistência aproximada de 0.85 enquanto que aumentando o intervalo de previsão para 1950-2000 a persistência aproximada se eleva para 0.94. Este resultado, quando confrontado com a Figura 16, que ilustra a previsão para 2000 passos à frente do modelo MSIAH estimado, mostra que a função de previsão desse modelo é dada por uma função hiperbólica em vez de uma função exponencial como no modelo GARCH(1,1).

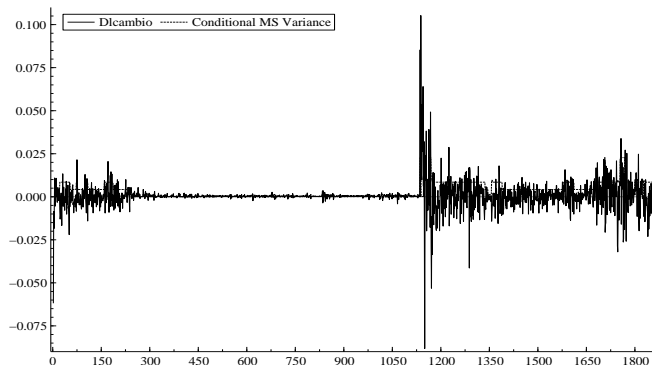
Este padrão de decaimento hiperbólico é consistente com o previsto pela classe de modelos de memória longa na variância do tipo FIGARCH-HYGARCH, que estudaremos a seguir. O resultado obtido nas Tabelas 16 e 17 mostra que a existência de quebras na estrutura da variância pode induzir o comportamento de memória longa na variância que é capturado pelos modelos que colocam uma ordem de integração fracionária na estrutura autoregressiva da variância condicional.

Este resultado seria uma transposição para a estrutura da variância dos resultados obtidos no artigo Laurini e Portugal, onde demonstra-se que um modelo que coloca mudanças markovianas para os todos os parâmetros de um modelo autoregressivo induz a um comportamento de memória longa para a média condicional da série, mesmo que em cada regime a persistência do processo seja curta.

## 9 Modelos de Memória Longa na Variância

Uma generalização do modelo GARCH que vem sendo crescentemente utilizada na literatura sobre persistência na volatilidade condicional é considerar modelos de memória longa na variância, uma adaptação dos modelos de ordem de integração fracionária ARFIMA para a variância condicional presente nos modelos da classe GARCH. Da mesma forma que os modelos ARFIMA têm o objetivo

Figura 18: Variância Condicional



de flexibilizar a estrutura de autocorrelações existentes nos modelos ARMA, os modelos FIGARCH e HYGARCH permitem a captura de uma ordem de longas autocorrelações que porventura existam na variância condicional

## 9.1 Modelo FIGARCH

O modelo FIGARCH foi proposto por Baillie, Bollerslev e Mikkelsen [3] como forma de superar o problema da persistência excessiva existente no modelo GARCH, já que a soma dos parâmetros  $\alpha_1 + \beta_1$  é sempre próxima de um nesses modelos. Para definir o modelo FIGARCH começamos definindo o modelo IGARCH usando uma representação para os quadrados dos resíduos. A representação ARMA-nos quadrados pode ser escrita em termos de um polinômio *lag* como:

$$(1 - L)^d \varepsilon_t^2 = w - v_t - \beta_1 v_{t-1} \quad (36)$$

que é uma generalização do modelo IGARCH para incluir a ordem de integração fracionária  $d$ , que define o modelo FIGARCH.

O modelo com ordem de integração fracionária impõe uma estrutura de decaimento diferente do modelo GARCH. Nos modelos GARCH, o impacto de um choque  $\varepsilon_t$  na variância condicional  $h_{t+k}$  é dado por  $\alpha_1 \beta_1^{k-1}$ , e, desta forma, com a soma de  $\alpha_1 + \beta_1$  próxima de um o decaimento dos efeitos do choque é muito lento, o que equivale a uma persistência muito longa. Embora o decaimento no modelo GARCH seja exponencial, ele poderia ser ainda muito rápido para capturar a estrutura de dependência longa existente em séries financeiras, por mais próximo de um que a soma de  $\alpha_1 + \beta_1$  esteja. Ding e Granger [13] observam que a taxa de decaimento dos retornos ao quadrado e, principalmente retornos absolutos, é melhor aproximada por uma hipérbole. Este decaimento hiperbólico é obtido inserindo uma ordem de integração fracionária de acordo com a representação 36.

Ding e Granger [13] mostram que as autocorrelações amostrais inicialmente decaem a uma taxa mais rápida do que um decaimento exponencial, mas para as autocorrelações mais distantes, o decaimento é mais lento, o que é equivalente a uma curva hiperbólica, uma evidência a favor do modelo FIGARCH. Uma fato interessante a ser observado é que observando o gráfico 16 vemos que os valores previstos para a variância no modelo de mudança markoviana (MSIAH) são consistentes com um função hiperbólica. Seria análoga à relação entre modelos de memória longa (ARFIMA) e modelos de mudança markoviana para a média.

Uma observação importante, feita por Davidson [11] é que as estruturas de modelos autoregressivos para a média, como raízes unitárias, são inadequadas para estruturas na variância da série. Um primeiro exemplo dessa inadequação é o fato de que embora o modelo IGARCH não seja covariância-estacionário ele pode ser estritamente estacionário sob determinadas condições. Uma



segunda observação importante é que a memória nos processos FIGARCH varia inversamente com o parâmetro  $d$ , ao contrário do que acontece nos modelos de memória longa na média da classe ARFIMA.

O artigo de Davidson [11] apresenta uma detalhada análise da estrutura de memória no processo imposta pela soma dos parâmetros autoregressivos na representação dos quadrados da série. Um dos resultados importantes derivados por Davidson é que a taxa de convergência das defasagens é irrelevante para a propriedade da estacionaridade da série. Davidson prova que o modelo IGARCH possui memória curta, já que os parâmetros correspondentes à estrutura de defasagens são absolutamente somáveis.

O modelo IGARCH é um caso particular do modelo FIGARCH, obtido quando o termo  $d$  é igual a zero, mas como aponta Davidson, o fato interessante é que a memória do processo aumenta quando  $d$  se aproxima de zero. No caso extremo de  $d$  ser igual a um, a memória do processo é descontínua, saltando imediatamente para infinito. Quando  $d$  se aproxima de zero, os pesos dos lags vão se aproximando da não somabilidade absoluta, mas o caso extremo  $d$  igual a zero apresenta memória curta, sendo na verdade um processo GARCH estável. O modelo FIGARCH, com valores diferentes de zero e um na verdade é um modelo cuja memória é muito maior do que a dos modelos GARCH e IGARCH, ao contrário do que esperavam Baillie, Bollerslev e Mikkelsen [3].

Na Tabela 18 temos o resultado da estimação do melhor modelo FIGARCH estimado para nossos dados. O modelo estimado contém uma estrutura ARMA(0,0) na média, ou seja somente uma constante foi representativa nessa estimação. A equação da variância era dada por um processo FIGARCH (1,  $d$ , 1) estimado pelo método de Chung. Note que temos um problema nessa estimação já que as condições de positividade para os valores da volatilidade condicional não são respeitadas, e desta forma as estimativas para a persistência longa não são confiáveis. O resultado do modelo FIGARCH aplicado aos nossos dados mostra um parâmetro  $d$  com valor igual a 0.135576 e estatisticamente significativo. Como já observado anteriormente, a memória nessa classe de modelos varia inversamente com o valor do parâmetro  $d$ , e dessa forma o valor do parâmetro  $d$  estimado para os dados em estudo apontaria para uma persistência muito longa para na estrutura da variância.

Esse resultado deve ser visto com cuidado no entanto, já que a condição de positividade não é respeitada e os valores para os parâmetros ARCH e GARCH estimados são negativos e muito distantes de todos os valores que estimamos anteriormente. O melhor modelo FIGARCH estimado para nossos dados impõe uma estrutura inconsistente para a variância da série.

## 9.2 Modelo HYGARCH

Davidson [11] aponta que alguns dos problemas existentes no modelo FIGARCH fazem com que a estrutura imposta por este modelo seja inadequada. Em substituição ao modelo FIGARCH, Davidson propõe um modelo de memória longa cuja estrutura das defasagens na representação dos quadrados decai hiperbolicamente. Usando a notação de Davidson [11], o modelo GARCH hiperbólico (HYGARCH) têm a seguinte estrutura:

$$\theta(L) = 1 - \frac{\delta(L)}{\beta(L)}(1 + \alpha((1-l)^d - 1)), \quad \alpha > 0 \quad (37)$$

A presença de um componente GARCH modifica a estrutura dos coeficientes dos lags mais baixos, e assim o modelo pode ser aproximado por:

$$\theta(L) = 1 - \frac{\delta(L)}{\beta(L)}(1 - \alpha\phi(L)) \quad (38)$$

onde a estrutura das defasagens corresponde a :

$$\theta(L) = \zeta(1+d)^{-1} \sum_{j=1}^{\infty} j^{-1-d} L^j, \quad d > 0 \quad (39)$$

Tabela 18: ARMA(0, 0)-FIGARCH(1, d, 1) e ARMA(0, 0)-HYGARCH (1, d, 0)

ARMA (0, 0)	FIGARCH	(1, d, 1)			ARMA (0, 0)	HYGARCH	(1, d, 0)		
	Coefficiente	D. Pad.	<i>t-valor</i>	<i>t-prob</i>		Coefficiente	D. Pad.	<i>t-value</i>	<i>t-prob</i>
Cst(M)	0.033782	0.010037	3.366	0.0008	Cst(M)	0.030016	0.009128	3.288	0.0010
Cst(V)	0.239301	0.023384	10.23	0.0000	Cst(V)	0.032689	0.008857	3.691	0.0002
d-Figarch	0.135576	0.011915	11.380	0.0000	d-Figarch	0.855270	0.082473	11.80	0.0000
ARCH	-0.906189	0.055692	-16.27	0.0000	GARCH(1)	0.636272	0.095971	6.630	0.0000
GARCH	-0.897911	0.059411	-15.11	0.0000	Log Alpha (HY)	-0.066920	0.029649	-2.2570	0.0241
LogLik.	-1369.475				Log Lik.	-1296.912			

onde  $\zeta(\cdot)$  é a função Zeta de Riemman, que leva a estrutura hiperbólica nos pesos das defasagens nesse polinômio. A melhor especificação de modelo HYGARCH encontrada para nossos dados é a estrutura ARMA(0,0)-HYGARCH(1,d,0) da Tabela 18.

O resultado interessante é que, embora o valor  $d$  estimado igual a 0.855270 seja estatisticamente significante, ele é muito menor do que o valor  $d$  encontrado no modelo FIGARCH, indicando uma estrutura de memória muito mais curta do que a apontada pelo modelo FIGARCH. Entretanto um intervalo de confiança com dois desvios padrões do valor estimado para  $d$  inclui o valor um<sup>4</sup>, o que nos indica que não podemos rejeitar que este modelo seja estatisticamente diferente de um modelo IGARCH.

Embora o modelo não possa ser considerado estatisticamente diferente de um modelo IGARCH, temos que notar que o parâmetro da ordem de integração fracionária é significativo. Este resultado, no entanto, tem que ser confrontado com a evidência encontrada na Seção 8.1 que mostra que um modelo de mudança markoviana também é consistente com um decaimento hiperbólico na estrutura da variância condicional. Desta forma, a memória longa encontrada na variância pode ser causada pelas mudanças abruptas na variância condicional, e, conseqüentemente, os modelos de memória longa na variância podem ser inadequados para capturar a persistência existente na variância condicional.

## 10 Conclusões

Nosso estudo da persistência a choques na volatilidade da série de log-retornos do câmbio é baseado na comparação da persistência estimada nos modelos GARCH(1,1) e o modelo de mudança markoviana. Complementando o estudo empírico, realizamos uma análise de Monte Carlo para verificar a adequação do modelo GARCH nas condições de quebras estruturais que possivelmente existem em nossos dados.

Os resultados obtidos mostram que o modelo GARCH(1,1) apresenta sérios problemas quando aplicado nestas situações, já que o estimador da persistência do modelo GARCH (a soma dos parâmetros  $\alpha_1$  e  $\beta_1$ ), é diretamente relacionado ao tamanho da amostra em estudo; amostras maiores levam a maiores valores estimados para a persistência. Este viés faz com que conclusões sobre a persistência dadas pelo modelo GARCH(1,1) sejam frágeis. Os procedimentos tentados para contornar o problema causado por este viés (estimação em subamostras e janelas móveis) não se mostraram adequados para superar este problema.

O modelo de Mudança Markoviana que permite mudanças na variância incondicional se mostra mais robusto no estudo da persistência a choques na série de log-retornos do câmbio, já que aponta que os choques têm efeitos transitórios sobre a variância condicional da série, o que parece mais adequado aos dados. Uma evidência adicional é que, observando o gráfico das previsões dinâmicas

<sup>4</sup>É necessário olhar com cuidado essa inferência, já que as distribuições desses parâmetros ainda não foram estudadas de forma completa.

para a variância no modelo MSIAH elas são consistentes com o decaimento hiperbólico previsto pelos modelos de memória longa na variância.

A literatura sobre persistência em modelos GARCH aponta que a presença de quebras na média do processo de volatilidade condicional pode levar a uma persistência elevada espúria. Em nosso estudo, esta possibilidade se mostra efetiva, já que a persistência estimada sempre é próxima de um e a persistência observada na série não reflete este valor. Os diversos procedimentos de estimação da volatilidade baseados no modelo GARCH(1,1) não conseguem eliminar os efeitos das quebras estruturais.

Utilizando o modelo de mudança markoviana para realizar previsões da volatilidade obtemos alguns resultados interessantes. As previsões deste modelo para horizontes curtos de tempo indicam uma persistência significativamente menor do que a do modelo GARCH(1,1), e que a variância condicional estimada por este modelo é bem mais suave do que a de um GARCH.

Um resultado obtido no estudo de persistência na variância, relacionada a existência de quebras estruturais, foi demonstrar que um processo estocástico com mudanças de parâmetros, principalmente mudanças na variância condicional, controlados por uma cadeia de Markov, pode induzir uma relação positiva entre a persistência estimada no modelo GARCH(1,1) e o tamanho da amostra.

## Referências

- [1] Andersen, T. G., Bollerslev, T. Diebold, F. X. e Labys, P. . The Distribution of Exchange Rate Volatility. **NBER Working Paper** 6961. 1999
- [2] Ang, A. e Bekaert, G. (). Regime Switches in Interest Rates. **Research Paper** 1486, Stanford University. 1998
- [3] Baillie, R.T, Bollerslev, T. e Mikkelsen, H.-O. (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. **Journal of Econometrics** 74:3-30.
- [4] Barnett, W., Powell, J. e Tauchen, G. (1992). **Nonparametric and Semiparametric Methods in Econometrics and Statistics**. Cambridge University Press.
- [5] Bauwens, L., Lubrano, M. e Richard, J-F. (1999). Bayesian Inference in Dynamic Econometric Models. Cambridge University Press.
- [6] Bollerslev, T. (1986). Generalised Autoregressive Conditional Heteroscedasticity. **Journal of Econometrics** 31:307-327
- [7] Bollerslev, T, Chou, R. e Kroner, K. (1992). ARCH Modelling in Finance:A Review of Theory and Empirical Evidence. **Journal of Econometrics** 53:5-60
- [8] Bollerslev, T. e Mikkelson, H.O. (1996). Modelling and Pricing Long Memory in Stock Market Volatility. **Journal of Econometrics** 73:151-184
- [9] Campbell, John Y., Lo, Andrew W. e MacKinlay, Archie G. (1997). **The Econometrics of Financial Markets**. Princeton University Press
- [10] Clark, P. B. (1973). A Subordinated Stochastic Process Model With Finite Variance for Security Prices. **Econometrica** 68:153-55.
- [11] Davidson, J. (2002). Moment and Memory Properties of Linear Conditional Heteroscedasticity models. manuscrito em processo - <http://www.cf.ac.uk/carbs/econ/davidsonje/hygarch4.pdf>
- [12] Dempster, A.P. , Laird, N.M. e Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM algorithm. **Journal of The Royal Statistical Society, Series B**, 39:1-38
- [13] Ding, Z. e Granger, C.W.J (1996). Modeling Volatility Persistence of Speculative Returns: a New Approach. **Journal of Econometrics** 73:185-215.

- [14] Doornik, J.A. (2001). **Ox: An Object-Oriented Matrix Language** (4th edition), London: Timberlake Consultants Press
- [15] Doornik, J.A. and Ooms, Marium (2000). Multimodality and the GARCH Likelihood. manuscripto
- [16] Engle, Robert. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. **Econometrica** 46:987-1007
- [17] Engle, R. e Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. **Econometric Reviews** 5:1-50.
- [18] Franses, Philip Hans e Dijk, Dick van. (2000) . **Non-linear Time Series Models in Empirical Finance**. Cambridge University Press
- [19] Gallant, A. R, Hsieh, D. e Tauchen, G. (1992) . On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974-1983. in **Nonparametric and semiparametric methods in econometrics and statistics** (1992), eds Barnett, W., Powell, J. e Tauchen, G. Cambridge University Press.
- [20] Garcia, R. e Perron, P. (1990). An Analysis of the Real Interest Rate under Regime Shifts. **Université of Montreal working paper**.
- [21] Gray, Stephen F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime Switching Process. **Journal of Financial Economics** 42:27-62
- [22] Hamilton, James D. (1989) A New Approach to the Economic Analysis to Nonstationary Time Series and the Business Cycle. **Econometrica** 57:357-384
- [23] Hamilton, James D. (1990). Analysis of Time Series Subject to Changes In Regime. **Journal Of Econometrics** 45:39-70
- [24] Hamilton, James D. (1994). **Time Series Analysis**. Princeton University Press
- [25] Hamilton, James D. e Susmel, Raul (1994). Autoregressive Heterokedasticity and Changes in Regime. **Journal of Econometrics** 64:307-333
- [26] Hansen, Bruce E. (1992). The Likelihood Ratio Test under Non-Standard Condition: Testing the Markov Switching Model of GNP. **Journal of Applied Econometrics** 7:S61-82
- [27] Herencia, M. Z, Hotta, L K. e Valls Pereira, P. L. (1998). Filtragem e Previsão com Modelos de Volatilidade: Volatilidade Estocástica versus GARCH. **Revista Brasileira de Economia** 52(2):241-278
- [28] Hotta. L. K. e Tsay. R. S. ( 1998). Outliers in GARCH Process. Graduate Scholl of Business, University of Chicago, manuscripto não publicado.
- [29] Hwang, S. e Valls Pereira, P.L. (2002). Small Sample Properties of Garch Estimates and Persistence of Volatility. Finance Lab Working Papers 48, Ibmecc Business School, São Paulo.
- [30] Krolzig, Hans-Martin (1997). **Markov-Switching Vector Autoregressions: Modeling, Statistical Inference and Application to Business Cycle Analysis**. Lecture Notes in Economics and Mathematical Systems . Springer
- [31] Laurini, M.P e Portugal, M.S (2002). Long Memory in R\$/US\$ Exchange Rates: A Robust Analysis. Anais do XXIV Encontro da Sociedade Brasileira de Econometria.
- [32] Lamoureux, C.G e Lastrapes, W. (1990). Persistence in Variance, Structural Change, and the GARCH Model. *Journal of Business & Economics Statistics* 8:2, 224-234.

- [33] Lundberg, S. e Teräsvirta, T. (1998). Evaluating GARCH Models. **Working Papers in Economics and Finance** 291. Stockholm School of Economics.
- [34] Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. **Econometrica** 59:347-370
- [35] Nelson, D. (1990). Stationary and Persistence in a GARCH(1,1) Model. in **ARCH - Selected Readings** (1995), editor Robert Engle, Oxford University Press 176-192
- [36] Nicholls, D.F. e Quinn, B.G. (1982). **Random Coefficient Autoregressive Models: An Introduction**. Lecture Notes in Statistics 11. Springer-Verlag.
- [37] Pagan, A. R. & Schwert, G. W. (1990). Alternative Models for Conditional Stock Market Volatility. **Journal of Econometrics** 45:267-29
- [38] Poskitt, D.S e Chung, S. H. (1996). Markov Chain Models, Time Series Analysis and Extreme Value Theory. **Adv. Applied Probability** 28 405-425.
- [39] Rydén, T. (1995). Estimating the Order of Hidden Markov Models. **Statistics** 26 345-354.
- [40] Rydén, T, Teräsvirta, T e Asbrink, S. (1998). Stylized Facts of Daily Return Series and the Hidden Markov Switching model. **Journal Of Applied Econometrics** 13:217-244
- [41] Sakata, S. e White. H. (1998). High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility. **Econometrica** 66:529-567.
- [42] Tsay, Ruey S. (1987). Conditional Heteroscedastic Time Series Models. **Journal of the American Statistical Association - Theory and Methods** (82)(398) 590:604
- [43] Valls Pereira, P. L., Hotta, L. K., Souza, L. A. R e Almeida, N. M. C. G. (1999). Alternative Models to Extract Asset Volatility: A Comparative Study. **Brazilian Review of Econometrics** 19:1, 57-109.