

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

**Medidas de Persistência a Choques e Eficiência de
Mercado para a Taxa de Câmbio R\$/US\$**

Márcio Poletti Laurini
Orientador - Prof. Dr. Marcelo Savino Portugal

PORTO ALEGRE

2002

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

**Medidas de Persistência a Choques e Eficiência de
Mercado para a Taxa de Câmbio R\$/US\$**

Márcio Poletti Laurini
Orientador - Prof. Dr. Marcelo Savino Portugal

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da Ufrgs, como quesito parcial para obtenção do Grau de Mestre em Economia.

PORTO ALEGRE

2002

Agradecimentos

Acredito que o principal mérito de uma dissertação de mestrado seja demonstrar os resultados de um período de aprendizado. Desta forma tenho de agradecer a uma série de pessoas que fortemente influenciaram esse processo, não somente no período de elaboração da dissertação, mas em toda minha vida.

Sem o incessante apoio de meus pais Amadeu e Elisabeth nada disto teria sido possível. Os sacrifícios realizados por eles foram fundamentais para que eu pudesse alcançar meus objetivos.

Ao Prof. Pedro Valls Pereira pela grande ajuda no desenvolvimento deste trabalho, pelas críticas e mais do que tudo, por ser o responsável pelo meu interesse em econometria de séries temporais. O exemplo de dedicação e rigor acadêmico são inestimáveis.

A orientação do Prof. Marcelo Portugal representou a experiência e o equilíbrio que tornaram esta dissertação possível. Sem o direcionamento e os ajustes de foco realizados pelo meu orientador este trabalho nunca poderia ser concluído.

Aos professores e mais que isso, amigos, Otaviano Canuto, José Maria Silveira e Domingos Alves pelos conselhos e pelos exemplos de competência e amor a ciência.

Aos meus grandes amigos Marcelo Gouvea e Guilherme Aires devo agradecer pela amizade e pela companhia. A todos meus colegas no programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Sul - uma turma cujo grau de amizade e companheirismo não tem equivalentes. Cristiano, Alexandre, Rogério, Maurício e Filipe por serem mais do que colegas, grandes amigos.

Também tenho uma dívida com os desenvolvedores dos softwares estatísticos utilizados neste programa, especialmente ao R Core Team (Projeto R), Jurgen Doornik (Linguagem Ox e pacote Arfima e PCGIVE), e também as equipes dos melhores editores científicos existentes - Lyx e Texmacs - onde esta dissertação foi escrita.

Os comentários, correções e sugestões colocados pelos pareceristas da dissertação: Sérgio da Silva (PPGE-UFRGS), Pedro Valls Pereira (IBMEC-SP) e João Victor Issler (EPGE-FGV), levaram a uma melhora substancial na qualidade desta dissertação. Não tenho palavras para agradecer a valiosa contribuição que foi dada por eles.

Finalmente quero agradecer o apoio financeiro fundamental dado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

Sumário

Agradecimentos	iii
Prefácio	xiii
1 O Modelo Martingale, Persistência e Eficiência de Mercado	1
1.1 Introdução	1
1.2 O Modelo <i>Martingale</i>	1
1.2.1 Definição	1
1.2.2 Modelo de HANSEN e SINGLETON	3
1.3 Passeio Aleatório (<i>Random Walk</i>)	5
1.3.1 Passeio Aleatório I - Incrementos IID	5
1.3.2 Passeio Aleatório II - Incrementos Independentes	5
1.3.3 Passeio Aleatório III - Incrementos Não Correlacionados	5
1.4 Eficiência de Mercado	6
1.5 Persistência em modelos de séries temporais	7
1.6 Testes de Raiz Unitária e Eficiência de Mercado para a série de câmbio	8
1.6.1 Análise preliminar dos dados	8
1.6.2 Testes de Raiz Unitária	15
1.7 Testes de Eficiência de Mercado	18
1.7.1 Testes de <i>Random Walk I</i> e <i>Random Walk II</i>	18
1.7.2 Testes de <i>Random Walk III</i>	20
1.8 Conclusões	23
2 Modelo de Mudança Markoviana	25
2.1 Introdução	25
2.2 Descrição do Modelo de Mudança Markoviana	26
2.2.1 Estimação dos parâmetros desconhecidos	27
2.3 Revisão das Aplicações	27
2.4 Estimação e Testes de Especificação	28
2.4.1 Determinação do Número de Regimes	29
2.4.2 Determinação da Forma Funcional	31
2.5 Modelo Estimado	31
2.5.1 Falta de Parcimônia no Modelo Final Estimado	38
2.5.2 Modelo SWARCH	41
2.6 Análises de Significância Conjunta - Teste de Eficiência de Mercado	42
2.7 Variância na Média	43
2.8 Razão de Variância no Modelo de Mudança Markoviana	47
2.9 Conclusões	48

3	Modelos de Memória Longa	51
3.1	Introdução	51
3.2	Propriedades do Modelo ARFIMA	51
3.3	Estimação	53
3.4	Modelos ARFIMA-GARCH	56
3.4.1	ARFIMA-GARCH - Amostra Completa	56
3.4.2	Modelo ARFIMA-GARCH(1,1) - Primeira subamostra (1-301) (Estabilização Inflacionária)	57
3.4.3	Modelo ARFIMA-GARCH(1,1) - Segunda Subamostra (301- 1137) (Regime de Bandas Cambiais)	58
3.4.4	Modelo ARFIMA-GARCH(1,1) - Terceira Subamostra (301- 1137) - (Livre Flutuação da Taxa de Câmbio)	59
3.5	Modelo ARFIMA Aplicado aos Resíduos do Modelo MSH	60
3.6	Estudo de Monte Carlo	63
3.6.1	Descrição do Experimento de Monte Carlo	63
3.6.2	Simulação Utilizando as Probabilidades Incondicionais e Vari- ância Unitária	63
3.6.3	Simulação Utilizando as Probabilidades Incondicionais e Vari- âncias do Modelo MSIAH	64
3.6.4	Simulação Usando a Estrutura Markoviana nos Regimes - Va- riância Fixa	65
3.6.5	Simulação Usando a Estrutura Markoviana Total nos Regimes	65
3.7	Conclusões	68
4	Persistência na Volatilidade	70
4.1	Introdução	70
4.2	Modelos ARCH-GARCH	71
4.2.1	Preliminares Sobre Modelagem ARCH	71
4.2.2	ARCH Generalizado - GARCH	72
4.2.3	GARCH Integrado	73
4.2.4	Persistência	74
4.3	Estimação	74
4.4	Problemas com a Modelagem ARCH	79
4.5	Estimação em subamostras	81
4.6	Teste Para a Constância de Parâmetros	82
4.7	Análise de Monte Carlo	85
4.8	Estimação com Amostra Crescente	89
4.9	Estimação em Subamostras - Dados Simulados	90
4.10	Estimação com Janelas Móveis	91
4.11	Comparação com a Variância Prevista no Modelo de Mudança Marko- viana	94
4.11.1	Previsão da Variância no Modelo MSH (Mudança Markoviana na Variância)	96
4.12	Modelos de Memória Longa na Variância	105
4.12.1	Modelo FIGARCH	105
4.12.2	Modelo HYGARCH	108
4.13	Conclusões	109
5	Conclusões	114

Lista de Tabelas

1.1	Estatísticas Descritivas	11
1.2	Teste ADF Câmbio	16
1.3	Teste KPSS	16
1.4	Teste Phillips-Perron	17
1.5	Teste ADF Sub-amostras	17
1.6	Teste BDS Independência	18
1.7	Testes de Linearidade na média	19
1.8	BDS Resíduos do Modelo Linear	19
1.9	Teste <i>Runs</i>	21
1.10	Variance Ratio Amostra completa	22
1.11	Variance Ratio Pré Banda Cambial	22
1.12	Variance Ratio Banda Cambial	23
1.13	Variance Ratio Pós Banda Cambial	23
2.1	Log-Verossimilhança e Critérios de Informação	29
2.2	Teste LR Número de Regimes	30
2.3	Teste BDS Especificação	30
2.4	Testes LR Forma Funcional	31
2.5	Testes Especificação Resíduos	33
2.6	Parâmetros Estimados	34
2.7	Matriz de Transição	35
2.9	Regimes e Duração	36
2.10	Regimes e Autovalores	37
2.11	Autovalores da Matriz de Transição	37
2.12	Testes Wald de Exclusão	42
2.13	GARCH na Média	44
2.14	Relação Média-Variância	46
2.15	Variance Ratios - Retornos Padronizados	47
2.16	Regimes e Observações	50
3.1	Medidas de Persistência	54
3.2	ARFIMA Amostra 12-1881	54
3.3	ARFIMA(2,d,0)	54
3.4	ARFIMA-GARCH(1,1) Amostra Completa	56
3.5	ARFIMA(2,d,0)-GARCH(1,1)	57
3.6	ARFIMA-GARCH(1,1) Amostra 1-299	58
3.7	ARFIMA(2,d,0)-GARCH(1,1) Amostra 1-299	58
3.8	ARFIMA-GARCH(1,1) Amostra 301-1137	59
3.9	ARFIMA(1,d,2)-GARCH(1,1)	59

3.10	ARFIMA-GARCH(1,1) Amostra 1138-1881	60
3.11	ARFIMA(2,d,0)-GARCH(1,1) Amostra 1137-1881	60
3.12	ARFIMA Resíduos Modelo MSH	62
3.13	ARFIMA(1,d,0) Resíduos Modelo MSH	62
3.14	DGP Probabilidades Incondicionais, Variância Unitária	64
3.15	DGP Probabilidades Incondicionais, Variância markoviana	64
3.16	DGP Probabilidades Condicionais (<i>Markov Switching</i>) - Variância Unitária	65
3.17	DGP Probabilidades Condicionais (<i>Markov Switching</i>) - mudanças na Variância	66
4.1	ARMA(1,0) - GARCH(1,1)	74
4.2	ARMA(0,0) - GARCH(1,1)	75
4.3	ARMA(0,0) - GARCH(2,2)	75
4.4	Testes de Especificação	76
4.5	Resíduos ao Quadrado Estatísticas Descritivas	76
4.6	Estimação GARCH subamostras	81
4.7	Monte Carlo - GARCH	86
4.8	Estimação em subamostras	90
4.9	Estimação Janela Móvel	92
4.10	Aproximação Persistência	103
4.11	Aproximação Persistência (n últimas observações)	105
4.12	ARMA(0, 0)-FIGARCH(1, d, 1)	108
4.13	ARMA(0, 0)-HYGARCH (1, d, 0)	109

Lista de Figuras

1.1	Taxa de Câmbio R\$/US\$	9
1.2	Log-Retornos do Câmbio	11
1.3	Histogramas	12
1.4	Densidades - Câmbio e Log-Retornos do Câmbio	13
1.5	ACF - Câmbio	14
1.6	ACF - Log-Retornos do Câmbio	15
2.1	Análise dos Resíduos	32
2.2	Probabilidades Estimadas	35
2.3	Valores Observados e Previstos	39
2.4	Resíduos Padronizados	40
2.5	Densidades - Regimes	40
2.6	Variância Condicional	45
3.1	Correlograma dos Log-Retornos do Câmbio	52
3.2	Valores Previstos e Observados Modelo ARFIMA(2,d,0)	55
3.3	Resíduos Modelo MSH	61
3.4	Convergência	67
4.1	Densidade Resíduos	77
4.2	Variância Condicional Modelo GARCH(1,1)	78
4.3	Garch(1,1) Pré Banda	82
4.4	Garch(1,1) Banda Cambial	83
4.5	Garch(1,1) Pós Banda Cambial	84
4.6	Histograma Persistência Estimada (Amostra Tamanho 100)	86
4.7	Histograma Persistência Estimada (Amostra Tamanho 250)	87
4.8	Histograma Persistência Estimada (Amostra Tamanho 500)	88
4.9	Histograma Persistência Estimada (Amostra Tamanho 1000)	88
4.10	Estimação com Amostra Crescente - Densidade Empírica	89
4.11	Persistência estimada Amostra Crescente	90
4.12	Série Gerada Amostra Tamanho 2500	91
4.13	Histograma - Janela Móvel de 250 observações	92
4.14	Valores Estimados - Janela de 250 Observações	93
4.15	Histograma - Janela Móvel de 500 observações	93
4.16	Valores Estimados - Janela de 500 Observações	94
4.17	Histograma - Janela Móvel de 1000 observações	95
4.18	Valores Estimados - Janela de 1000 Observações	95
4.19	Previsões do Modelo de mudança markoviana para a Variância	98
4.20	Previsões para a Variância e Valores Observados	99

4.21	Previsão Amostra 500-600	100
4.22	Previsão - Amostra 500-600 e Valores Observados	101
4.23	Previsão 2000 passos e Variância modelo MSIAH	102
4.24	autocorrelação e Variância Filtrada	103
4.25	autocorrelação e Variância Suavizada	104
4.26	Variância Condicional	106
4.27	Previsões - Modelo HYGARCH	110
4.28	Intervalo de Confiança - Modelo HYGARCH	111

Resumo

Neste estudo analisamos persistência a choques na média e na volatilidade e a validade da hipótese de Eficiência Fraca de Mercado para a série de câmbio nominal diária R\$/US\$, através de Modelos de mudança markoviana, Memória Longa e GARCH. Mostramos que o modelo de mudança markoviana é adequado para capturar a estrutura de dependência existente nesta série, tanto na média quanto na variância, enquanto que os modelos de Memória Longa e GARCH não são robustos na presença das quebras estruturais existentes neste período. Conduzimos uma série de procedimentos de análise de especificação e experimentos de Monte Carlo para mostrar que mudanças na estrutura da variância podem induzir padrões espúrios de memória longa e persistência.

Abstract

In this study we analyze the persistence of shocks in the mean and volatility and the validity of the hypothesis of Weak Market Efficiency for the series of daily R\$/US\$ nominal exchange rates, through models of Markov Switching model, Long Memory and GARCH. We show that the Markov switching model is adequate to capture the structure of dependence in mean and variance of this series, while that the models of Long Memory and GARCH are not robust in presence of the existing structural breaks in this period. We lead a series of procedures of analysis of specification and Monte Carlo experiments to show that changes in the structure of the variance can induce spurious standards of long memory and persistence.

Prefácio

O comportamento das taxas de câmbio em economias modernas sempre foi um tema de grande interesse teórico e prático. O processo de crescentes aberturas comerciais e financeiras, em conjunto a um processo de abandono de políticas de fixação total ou parcial das taxas de câmbio, tornam cada vez maior a importância de se conhecer e prever o comportamento do câmbio.

Embora seja um tema de estudo extremamente enfatizado na teoria econômica, a criação de modelos analíticos para a determinação do câmbio, principalmente do comportamento das taxas de câmbio de curto prazo, tem obtido um sucesso bem restrito. Em um artigo seminal, MEESE e ROGOFF (1983) mostram que a performance preditiva de um modelo simples de Random Walk supera os valores previstos por uma série de modelos monetários de determinação do câmbio (incluindo o modelo de DORN-BUSH (1976) e modelos de preços flexíveis) em termos de previsão de um até doze meses a frente. MEESE e ROGOFF (1983) consideram uma série de explicações para esse fenômeno, incluindo a quebra de funções de demanda por moeda e desvios prolongados da paridade do poder de compra de longo prazo, sem achar uma conclusão definitiva. A conclusão destes autores é que os modelos estruturais de determinação do câmbio, no entanto, tem performance preditiva superior para horizontes de dois a três anos. Mas o comportamento do câmbio, tanto em seu preço como sua volatilidade, em dados de alta frequência ainda é um tema que não encontrou uma resposta definitiva, ou ao menos satisfatória na teoria econômica.

Em razão da performance insatisfatória dos modelos estruturais, tem se tentado uma série de modelos empíricos para o estudo das taxas de câmbio. Estes modelos tentam capturar alguns fatos estilizados presentes nas séries de câmbio.

Estes comportamentos das séries temporais de câmbio são compartilhados pelas séries de ativos financeiros em geral; entre essas características podemos incluir o fato de que as mudanças de preços são (ao menos linearmente) não-previsíveis, indicando que estes processos podem ser descritos como *martingales*, o que constitui uma forte evidência em favor da teoria de mercados eficientes. Outras características destas séries financeiras são que as mudanças de preços não podem ser consideradas Gaussianas, já que são leptocurticas (excessivamente centradas em torno do zero e possuindo caudas mais pesadas do que as de uma distribuição gaussiana equivalente) e apresentam assimetria.

Objetivos

O objetivo desta dissertação é elaborar um extensivo estudo sobre a dinâmica empírica da taxa de câmbio nominal no Brasil no período pós Plano Real de uma perspectiva univariada, ou seja usando apenas o próprio passado desta série. Os focos principais

são estudar o poder explicativo dos diversos modelos considerados no estudo e a caracterização da persistência a choques que cada modelo impõe à série de câmbio R\$/US\$.

A determinação correta do modelo estatístico gerador dos dados e principalmente sobre de que forma os choques externos se propagam na série de câmbio tem implicações críticas para a formulação de estratégias de intervenção no câmbio por parte do Banco Central, e pelo lado dos agentes privados, na formulação de estratégias de captação externa, exportações e *hedge*.

O período de estudo inicia-se com a fase do plano de estabilização inflacionária, que utilizou-se intensamente das taxa de câmbio como instrumento de política antiinflacionária. A dinâmica do câmbio pós-1994 é interessante por ser uma seqüência de mecanismos de intervenção governamental sobre a taxa de câmbio, iniciando com uma taxa de câmbio valorizada, seguindo-se de fases de depreciação controlada e bandas cambiais (*target zones*) e, após seguidas crises externas e um ataque especulativo, mudança para um regime de câmbio totalmente flutuante, em conjunto com um regime de metas de inflação.

Essa seqüência de mudanças nos processos econômicos que determinam o valor da taxa R\$/US\$ tem profundas conseqüências na modelagem econométrica. Uma característica marcante da taxa dessa série é que o mecanismo estatístico gerador dessa série não pode, se considerarmos modelos lineares do tipo autoregressivo média móvel (ARMA), ser considerado estável.

Essa série é marcada por mudança de parâmetros tanto no processo gerador da média quanto no processo gerador da variância. Essa instabilidade estrutural leva a diversas conseqüências. A primeira é que modelos lineares tradicionais não são capazes de modelar adequadamente o comportamento desta série, já que justamente não capturam os fatos mais notáveis presentes no seu comportamento. Além disso tentaremos mostrar que as mudanças de parâmetros induzem alguns fenômenos estatísticos como dependência longa e persistência elevada na volatilidade, que podem ser espúrios.

Desta forma esta dissertação também é um estudo sobre os problemas decorrentes da aplicação de modelos econométricos em séries cujo DGP é marcado por mudanças estruturais. Um dos resultados dessa dissertação é mostrar que os modelos da classe ARFIMA, que capturam dependência longa, e os tradicionais modelos GARCH, a ferramenta mais usada na modelagem de variância condicional, são inadequados quando o processo gerador da série possui quebras.

Esta dissertação possui a seguinte estrutura: O Capítulo 1 revisa o modelo mais importante de precificação de ativos conhecido como modelo *Martingale* e avalia suas conseqüências sobre a persistência de choques e mostra sua relação com o conceito de Eficiência de Mercado. Nesse capítulo também estimamos algumas medidas tradicionais relacionadas ao modelo *Martingale* para a série de câmbio em estudo.

O Capítulo 2 propõe a aplicação de um modelo de mudança markoviana para a série de log-retornos de câmbio, analisando as características e vantagens desse modelo. Também discutimos quais as conseqüências do uso desse modelo para a persistência a choques e a Eficiência de Mercado para essa série. O modelo MSIAH (*Markov Switching Intercept Autoregressive Heterocedasticity*) usado é o padrão de comparação com os demais modelos estimados no presente trabalho.

No Capítulo 3 analisaremos a aplicação de modelos de memória longa do tipo ARFIMA para a série de log-retornos de câmbio. A existência de memória longa está bastante ligada a séries com quebras estruturais. Dessa forma, um estudo aprofundado desse modelo é realizado, e através de um estudo de Monte Carlo mostramos que a existência de memória longa não é sustentável, e sim gerada pelo processo de mudança na estrutura de auto-correlação e de variância-covariância desta série.

O Capítulo 4 tem o objetivo de estudar a persistência de choques sobre a volatilidade da série de log-retornos do câmbio. Mostramos através de estimação com janela móvel e de um estudo de Monte Carlo que a aplicação de modelos GARCH nessa série pode levar a persistência excessivamente elevada de choques na variância dessa série. Mostramos também que o modelo de mudança markoviana usado no capítulo 2 captura de forma adequada as mudanças na volatilidade.

O Capítulo 5 discute os resultados encontrados nesse estudo e mostra como esses resultados são consistentes com o modelo teórico para precificação de ativos de KIRMAN e TEYSSIERRE (2001).

Capítulo 1

O Modelo Martingale, Persistência e Eficiência de Mercado

1.1 Introdução

O primeiro passo em nosso estudo da taxa de câmbio diária Real/US\$ é definir os conceitos fundamentais relacionados aos modelos e testes econométricos desenvolvidos. Partiremos do modelo mais importante de precificação de ativos, o modelo *martingale*, e a partir deste modelo iremos definir os conceitos de Eficiência de Mercado e persistência a choques.

A taxa de câmbio reflete o preço em moeda estrangeira de um ativo, sendo esse ativo a moeda nacional. Para realizar um estudo empírico sobre as características estatísticas desse ativo, necessitamos de teorias econômicas sobre o processo estatístico gerador da taxa de câmbio.

A teoria mais antiga e importante sobre precificação de ativos é a teoria conhecida como *martingale model*, e suas origens podem ser traçadas aos manuscritos de Cardano de 1565, mas cuja formulação moderna foi estabelecida por BACHELIER (1900) e SAMUELSON (1965). De forma compacta, esta teoria estabelece que mudanças de preços nos ativos (retornos) não podem ser previstas de forma sistemática. Isso equivale a dizer que estatisticamente os retornos de qualquer ativo devem ser um processo aleatório não correlacionado (ruído branco fraco). Se em adição a isso os retornos tiverem distribuição Normal eles isso equivale a que eficiência de mercado implica que os retornos devam ser um processo independente e identicamente distribuído (I.I.D). Segundo este modelo quaisquer tentativas de previsão dos preços futuros de um ativo não podem ter poder explicativo estatisticamente significante.

1.2 O Modelo *Martingale*

1.2.1 Definição

Um *martingale* é um processo estocástico P_t que satisfaz, conforme CAMPBELL, LO & MACKINLAY (1997), a seguinte condição:

$$E [P_{t+1}|P_t, P_{t-1}, P_{t-2}, \dots] = P_t \quad (1.1)$$

ou de forma equivalente :

$$E [P_{t+1} - P_t|P_t, P_{t-1}, P_{t-2}, \dots] = 0 \quad (1.2)$$

O modelo *Martingale* estabelece que se P_t é o preço do ativo, o melhor predictor para o preço desse ativo no próximo período é o preço do ativo no momento presente. A representação 1.2 nos diz que mudanças no preço do ativo (retornos) devem ter média zero e assim não exibir nenhuma tendência ou padrão sistemático. A condição de que P_t é o melhor predictor *linear* está relacionada a dois conceitos importantes: a Lei das Expectativas Iteradas e a propriedade markoviana de preços de ativos.

Para estabelecer a relação entre a aleatoriedade nos retornos e a Lei das Expectativas Iteradas, usaremos a formulação proposta por SAMUELSON (1965). A Lei das Expectativas Iteradas estabelece que $E [X|I_t] = E [E [X|J_t] |I_t]$ onde I_t e J_t são dois conjuntos de informação e $I_t \subset J_t$, isto é, o conjunto de informação J_t é superior porque contém toda a informação contida em I_t além de possuir informação extra. Conforme CAMPBELL, LO & MACKINLAY (1997) reescrever esta equação como $E [X - E [X|J_t] |I_t] = 0$ tem a intuitiva interpretação de que ninguém com um conjunto inferior de informação I_t pode prever o erro de previsão condicionado em um conjunto de informação superior J_t .

A relevância da Lei das Expectativas Iteradas na precificação de ativos e sua relação com a natureza aleatória dos retornos, de acordo com a formulação pioneira de SAMUELSON (1965), pode ser colocada na seguinte formulação de LEROY (1973): suponha que o ativo P possa ser escrito no tempo t como a expectativa racional de algum "valor fundamental" V^* , condicionado em um conjunto de informação I_t . Dessa forma, podemos escrever:

$$P_t = E [V^*|I_t] = E_t [V^*] \quad (1.3)$$

a mesma relação vale um período a frente:

$$P_{t+1} = E [V^*|I_{t+1}] = E_{t+1} [V^*] \quad (1.4)$$

A expectativa da mudança no preço para próximo período é então:

$$E_t [P_{t+1} - P_t] = E_t [E_{t+1} [V^*] - E_t [V^*] = 0] \quad (1.5)$$

isto é válido pois $I_t \subset I_{t+1}$ e pela Lei das Expectativas Iteradas, se os preços são processos estocásticos do tipo *Martingale*, mudanças de preços devem ser aleatórias. Esta formulação tem o mérito de esclarecer a percepção errônea de que, se os preços de um ativo são determinados por algum fundamento, eles poderiam ser previsíveis e não aleatórios. A equação 1.5 mostra que aleatoriedade nos retornos e determinação de preços baseada em fundamentos são consistentes.

Uma segunda noção relevante é o que chamamos de propriedade markoviana dos preços. Um processo estocástico possui a propriedade markoviana de primeira ordem se a seguinte relação for válida:

$$Prob(X_{t+1}|X_t, X_{t-1}, X_{t-2}, \dots) = Prob(X_{t+1}|X_t) \quad (1.6)$$

Essa propriedade estabelece que somente o passado mais próximo ($t-1$) é relevante na determinação do preço no período subsequente. Um modelo de precificação de

ativos que justifica a propriedade Markov bem como o modelo *martingale* baseada na exposição de LJUNQGVIST e SARGENT (2000) sobre o modelo de dois ativos de HANSEN e SINGLETON (1983) é colocada a seguir. Embora seja um modelo geral de precificação de opções e não diretamente um modelo estrutural para a determinação de câmbio, ele ajuda a compreender as características existentes na taxa de câmbio que são geralmente compartilhadas por todos os ativos financeiros.

1.2.2 Modelo de HANSEN e SINGLETON

O modelo de HANSEN e SINGLETON (1983) coloca o problema de um indivíduo representativo que tem a oportunidade de trocar dois ativos. A troca de ativos, bem como toda estrutura resultante de preços é baseada em restrições sobre o consumo. Estas restrições permanecem válidas mesmo se mais ativos são oferecidos ao agente. O agente possui uma riqueza $A_t > 0$ e quer maximizar sua utilidade esperada em todo seu período de vida:

$$E_t \sum_{j=0}^{\infty} \beta^j u(c_{t+j}) \quad (1.7)$$

onde u é uma função utilidade côncava, estritamente crescente e duas vezes continuamente diferenciável. O parâmetro β é um fator de desconto subjetivo e c_{t+j} é o consumo no período $t+j$. Para financiar o consumo, o agente pode transferir intertemporalmente riqueza contida em títulos e ações. Títulos de um período são remunerados com a taxa de juros sem risco R_t medida em unidades de consumo. Seja L_t o resultado bruto dos ganhos de capital do agente sobre os títulos entre os períodos t e $t+1$ com um valor presente $R_t^{-1} L_t$ no tempo t . O valor de L_t é negativo se o agente lança títulos e desta forma toma emprestado recursos. A participação do agente em equidades entre os períodos t e $t+1$ são denotados por s_t . As restrições impostas de endividamento são de que $L_t > -b_b$ e $s_t > -b_s$ onde b_b e $b_s > 0$.

A participação em ações rende um fluxo de rendimentos estocástico y_t , e o preço de cada ação no período t é denotado por p_t . Desta forma, podemos formular a restrição orçamentária

$$c_t + R_t^{-1} L_t + p_t s_t \leq A_t \quad (1.8)$$

e a riqueza do próximo período:

$$A_{t+1} = L_t + (p_{t+1} + y_{t+1}) s_t \quad (1.9)$$

O dividendo y_t é a única fonte de incerteza exógena. A resolução desse problema é através de programação dinâmica considerando A_t e valores presentes e passados de y_t como variáveis de estado no tempo t e como variáveis de controle L_t e s_t . As equações de Euler associadas com os controles L_t e s_t que representam soluções interiores do problema são, conforme LJUNQGVIST e SARGENT (2000):

$$w'(c_t) R_t^{-1} = E_t \beta w'(c_{t+1}) \quad (1.10)$$

$$w'(c_t) p_t = E_t \beta (y_{t+1} + p_{t+1}) w'(c_{t+1}) \quad (1.11)$$

Essas equações de Euler nos permitirão encontrar propriedades fundamentais sobre precificação de ativos. A maximização também deve respeitar as seguintes condições de transversalidade:

$$\lim_{k \rightarrow \infty} E_t \beta^k u'(c_{t+k}) R_{t+k}^{-1} L_{t+k} = 0 \quad (1.12)$$

$$\lim_{k \rightarrow \infty} E_t \beta^k u'(c_{t+k}) p_{t+k} s_{t+k} = 0 \quad (1.13)$$

Uma intuição sobre essas condições é que elas impedem que o agente sobre-acumule ou sub-acumule ativos. Em um modelo de horizontes finitos, a violação de 1.12 e 1.13 faria com que o agente morresse com ativos ainda em carteira, o que representaria uma sub-otimalidade no consumo.

A partir desse modelo básico podemos derivar o modelo *martingale* de consumo e preços de ativos. O modelo de preços de ativos é obtido fazendo uma série de restrições sobre 1.10 e 1.11. Assumindo que uma taxa de juros constante $R_t = R > 1$ a equação 1.10 implica em:

$$E_t w'(c_{t+1}) = (\beta R)^{-1} w'(c_t) \quad (1.14)$$

Reescrevendo a equação de Euler 1.11 como :

$$E_t \beta (y_{t+1} + p_{t+1}) \frac{w'(c_{t+1})}{w'(c_t)} = p_t \quad (1.15)$$

e notando que para quaisquer duas variáveis aleatórias temos que $E_t xy = E_t x E_t y + cov_t(x, y)$. Aplicando este resultado em 1.15 chegamos a

$$\beta E_t (y_{t+1} + p_{t+1}) E_t \frac{w'(c_{t+1})}{w'(c_t)} + \beta cov_t \left[(y_{t+1} + p_{t+1}), \frac{w'(c_{t+1})}{w'(c_t)} \right] = p_t \quad (1.16)$$

Para obter o modelo *martingale* é necessário supor que $E_t [w'(c_{t+1}) / w'(c_t)]$ é constante e que

$$cov_t \left[(y_{t+1} + p_{t+1}), \frac{w'(c_{t+1})}{w'(c_t)} \right] = 0 \quad (1.17)$$

o que são condições bastante restritivas. Conforme LJUNQGVIST e SARGENT (2000) uma condição suficiente é que os agentes são neutros ao risco e que $u(c_t)$ é linear em c_t , e dessa forma a derivada $w'(c_t)$ é independente de c_t . Nesse caso a equação 1.16 se reduz a

$$E_t (y_{t+1} - p_{t+1}) = p_t \quad (1.18)$$

A equação 1.18 nos permite chegar às duas propriedades que havíamos definido. A propriedade Markov é alcançada notando que, ajustando para dividendos e desconto inter-temporal, os preços devem seguir um processo de Markov de primeira ordem e que nenhuma outra variável deve causar o preço no sentido de Granger.

A propriedade de que preços de ativos devem seguir um processo estocástico do tipo *martingale* pode ser visualizada notando que a equação estocástica a diferenças 1.18, conforme LJUNQGVIST e SARGENT (2000), tem soluções da classe:

$$p_t = E_t \sum_{j=1}^{\infty} \beta^j y_{t+j} + \xi_t \left(\frac{1}{\beta} \right)^t \quad (1.19)$$

onde ξ_t é um processo estocástico que respeite a relação $E_t \xi_{t+1} = \xi_t$, o que corresponde a definição de um *martingale*. Esta equação tem o significado de que o preço do ativo deve corresponder ao valor descontado dos dividendos mais um termo de "bolha", que é geralmente igualado a zero em modelos de equilíbrio geral ou evitado por condições de transversalidade.

1.3 Passeio Aleatório (*Random Walk*)

Um modelo relacionado ao processo *martingale* e que é a representação mais utilizada nos testes de previsibilidade de retornos é o modelo de passeio aleatório - (*random walk*). A representação de um passeio aleatório é dada por:

$$P_t = \mu + P_{t-1} + \varepsilon_t \quad (1.20)$$

Este modelo diz que o preço do ativo no tempo t é dado pelo preço no momento imediatamente anterior ($t - 1$), um termo de mudança esperada conhecido como *drift* mais um componente de erro não previsível ε_t . O modelo de passeio aleatório pode ser obtido do processo *martingale* ao modelarmos o termo de erro ε_t . O comportamento do termo de erro ε_t é extremamente importante, e restrições sobre o comportamento deste termo levam a três versões do modelo de Passeio Aleatório, conforme colocado por CAMPBELL, LO & MACKINLAY (1997), que discutiremos a seguir.

1.3.1 Passeio Aleatório I - Incrementos IID

A versão mais forte do modelo de passeio aleatório é aquela em que os incrementos ao preço P_t dados pelo termo de erro ε_t pertencem a uma mesma distribuição (identicamente distribuídos) e são independentes. Adicionalmente pode se colocar a distribuição de origem, o que no caso mais comum corresponde a assumir que ε_t pertence a uma distribuição normal com média 0 e variância constante σ^2 . O Passeio Aleatório I, também conhecido como RW1 é ainda mais restritivo do que o modelo *martingale*, já que nesse modelo além de os incrementos serem não linearmente não-correlacionados, qualquer combinação não-linear dos incrementos deve ser não correlacionada também. Em tempo contínuo isso corresponde ao modelo de movimento browniano aritmético utilizado como base dos modelos de precificação de opções de Black-Scholes.

1.3.2 Passeio Aleatório II - Incrementos Independentes

O modelo RW1 é extremamente restritivo para ser considerado em séries financeiras reais, já que ele exclui a possibilidade de mudanças estruturais no processo gerador dos dados, como por exemplo mudanças de parâmetros, onde a mais relevante é a mudança na volatilidade. Uma versão mais adequada é conhecida como passeio aleatório do tipo II (RW2), que somente impõe que os incrementos devem ser independentes, mas não necessariamente devem vir da mesma distribuição. Isso mantém a característica de não-previsibilidade linear mas permite comportamentos de mudança na volatilidade incondicional como iremos testar no capítulo 2.

1.3.3 Passeio Aleatório III - Incrementos Não Correlacionados

A forma mais geral do modelo de passeio aleatório é aquela em que apenas restringimos o processo ε_t a ser não correlacionado, também conhecido como RW3. Como exemplo,

séries financeiras com efeitos do tipo ARCH podem respeitar o comportamento de RW3 já que os retornos em nível podem ser não correlacionados, mas os retornos ao quadrado possuem autocorrelação, o que faz que o processo não seja independente. Por ser a forma de passeio aleatório menos restritiva é a forma mais consistente com o comportamento observado em séries financeiras reais. O RW3 é usualmente a forma mais testada de passeio aleatório.

Na Seção 1.5 discutiremos formas de aplicar estes testes, mas antes discutiremos dois conceitos importante que são o de eficiência de mercado e o de persistência de choques.

1.4 Eficiência de Mercado

Um conceito ligado às propriedades de precificação de ativos obtidas nos modelos acima é a *eficiência de mercado*. A definição de eficiência de mercado é relacionada ao uso racional da informação disponível publicamente. Um mercado financeiro é considerado Eficiente (informacionalmente eficiente) quando o preço do ativo reflete toda a informação disponível.

Essa definição faz com que o preço corrente seja o melhor previsor para o preço a $t+h$ períodos a frente, já que ele incorpora toda a informação existente até esse momento. Mudanças de preço não podem ser sistematicamente previstas se o mercado for informacionalmente eficiente. Isto corresponde à definição de que mudanças nos preços devem ser um processo estocástico *martingale*, como definido anteriormente.

Como a definição de eficiência de mercado é baseada no uso do conjunto de informações disponível, a seguinte taxonomia proposta por ROBERTS (1967) propõe três definições de mercados eficientes, relacionando cada uma delas a um conjunto de informação relevante. Essas definições são:

1. Eficiência fraca - Preços refletem a informação disponível no histórico da série . Testes de Eficiência de Mercado usando modelos univariados de séries temporais são testes de eficiência fraca.
2. Eficiência semi-forte - Preços refletem toda a informação publicamente disponível.
3. Eficiência forte - Preços refletem toda a informação existente, tanto pública quanto privada. Aqui fica contemplada a possibilidade de *insider trading*.

A teoria de mercados eficientes, em formulações mais recentes, leva em conta o fato de que existe um custo na coleta de informações. Uma formulação mais adequada é que o conjunto informacional relevante é aquele formado até o ponto onde o ganho marginal do uso de informação iguala o custo marginal de coletar informação. A teoria, no entanto, não contempla dois problemas. O primeiro é a incerteza que temos sobre a qualidade da informação; o segundo problema é que temos capacidade limitada de processamento de informações, conforme foi analisado por GROSSMAN e STIGLITZ (1980).

Se um mercado financeiro não é eficiente, então existem possibilidades de ganhos excessivos usando os padrões encontrados nas séries através de arbitragem. Os testes de eficiência e racionalidade de mercado usualmente tentam encontrar correlações ou padrões estatisticamente significantes existentes nas séries.

A hipótese de mercados eficientes pode ser testada de diversas formas, mas todas consistem em verificar se existe algum padrão nos dados. Alguns dos testes existentes

na literatura buscam verificar a presença fenômenos como efeitos *intraday*, efeitos dia da semana, sazonalidades nos retornos, testes de auto-correlação, teste *Runs*, teste de razão da variância (*variance ratio*) ou a lucratividade do uso de regras de filtro e de análise técnica.

Efeitos *intraday* e dia da semana são um aspecto bem estudado em séries financeiras. Alguns fatos observados são que retornos na segunda-feira tendem a ser menores. No Brasil existia o famoso efeito quinta-feira, em que a propagação de boatos levava a significativas oscilações nos preços dos ativos mais importantes. Um efeito *intraday* relevante é que nos últimos 30 minutos de transação se observam os maiores aumentos nos preços ocorridos no dia. Dentro do ano também foram estudados alguns efeitos sazonais, entre eles a existência de retornos maiores em janeiro, principalmente para empresas menores, devidos a divulgação de balanços e efeitos de manipulação fiscal.

1.5 Persistência em modelos de séries temporais

Conforme visto o modelo *martingale* e o modelo de eficiência de mercado tem são representados pelo modelo de séries temporais em tempo discreto conhecido como *passeio aleatório*, que é a versão em tempo discreto do movimento browniano. O passeio aleatório é uma versão do processo autoregressivo de primeira ordem, em que o parâmetro autoregressivo é igual a um. Esse processo é denotado como:

$$y_t = \phi y_{t-1} + \varepsilon_t \quad (1.21)$$

Para entender a persistência dos choques ε_t na série, é ilustrativo determinar a solução desta equação à diferenças de forma recursiva:

$$\begin{aligned} y_1 &= \phi y_0 + \varepsilon_1 \\ y_2 &= \phi y_1 + \varepsilon_2 = \phi(\phi y_0 + \varepsilon_1) + \varepsilon_2 = \phi^2 y_0 + \phi \varepsilon_1 + \varepsilon_2 \\ y_3 &= \phi y_2 + \varepsilon_3 = \phi(\phi^2 y_0 + \phi \varepsilon_1 + \varepsilon_2) + \varepsilon_3 \\ &= \phi^3 y_0 + \phi^2 \varepsilon_1 + \phi \varepsilon_2 + \varepsilon_3 \end{aligned} \quad (1.22)$$

$$y_t = \phi^t y_0 + \phi^{t-1} \varepsilon_1 + \phi^{t-2} \varepsilon_2 + \dots + \phi \varepsilon_{t-1} + \varepsilon_t$$

Essa recursão mostra que escrevendo uma processo autoregressivo de primeira ordem AR(1) como um processo de média móvel de ordem infinita MA(∞), o valor de y no período t é determinado pelo valor inicial y_0 multiplicado por ϕ^t mais os choques passados ponderados por ϕ^{k-1} , onde k é distância do tempo t em relação a data do choque. Se impusermos a restrição $|\phi| < 1$, é fácil notar que o efeito dos choques na série decaim exponencialmente. Uma forma de se verificar isso é através dos multiplicadores dinâmicos, que medem o efeito do choque no tempo $t - s$ sobre o valor da variável no tempo t . Escrevendo y_{t+s} como uma função de y_{t-1} e os choques $\varepsilon_t, \varepsilon_{t+1}, \dots, \varepsilon_{t+s}$:

$$y_{t+s} = \phi^{s+1} y_{t-1} + \phi^s \varepsilon_t + \phi^{s-1} \varepsilon_{t+1} + \phi^{s-2} \varepsilon_{t+2} + \dots + \phi \varepsilon_{t+s-1} + \varepsilon_{t+s} \quad (1.23)$$

o efeito de ε_t em y_{t+s} , conhecido como multiplicador dinâmico, é dado por:

$$\frac{\partial y_{t+s}}{\partial \varepsilon_t} = \phi^s \quad (1.24)$$

Se a restrição $|\phi| < 1$ for válida, o efeito de ε_t em y_{t+s} tende a diminuir com o aumento do intervalo de tempo s . Os modelos de precificação de ativos ligados ao

modelos *martingale* e à teoria de eficiência de mercado estabelecem que o parâmetro ϕ deva ser igual a um, o que corresponde a dizer que o preço do período imediatamente anterior é o melhor previsor para o preço no período presente, dado o conjunto de informações disponível. Essa restrição no entanto torna o comportamento da série totalmente diferente. Para visualizar essa mudança, vamos usar a equação 1.22 que define o processo AR(1) como um processo MA(∞). Quando ϕ é igual a um, vemos que y_t é escrito como o valor inicial mais uma soma não descontada de todos os choques ε , o que equivale a dizer que o efeito de cada choque não é dissipado com o passar do tempo. Isto fica evidente quando olhamos o multiplicador dinâmico definido pela equação 1.24, que mostra que, quando $\phi = 1$ o efeito de ε_t para em y_{t+s} é o mesmo em todos os períodos seguintes. No caso de um processo com $\phi = 1$, também conhecido como processo de Raíz Unitária, o efeito de um choque só é anulado por um choque de mesma magnitude e de sinal contrário.

Séries temporais com raízes unitárias têm sido objeto de estudo intenso desde a elaboração do trabalho de DICKEY e FULLER (1979), que determinaram a distribuição correta para o teste de hipóteses $\phi = 1$ contra a hipótese $\phi < 1$, já que este teste não tem distribuição padrão Student t . Para uma revisão de processos auto-regressivos com raízes unitárias ver HAMILTON (1994), capítulos 15, 17 e 18.

Nosso objetivo inicial neste estudo é determinar se nossa série de taxa de câmbio segue um processo auto-regressivo com $\phi = 1$, o que representa testar pelos testes de raízes unitárias se, em nossa série, os choques se dissipam ou se acumulam de forma permanente.

Uma ressalva extremamente importante é que o teste de raiz unitária não pode ser considerado adequadamente um teste de Eficiência de Mercado, já que um processo de raiz unitária pode ser mais geral do que o processo de passeio aleatório que utilizamos para definir a persistência a choques no caso mais simples. Processos com raízes unitárias podem contar com autocorrelações significativas na série em diferenças, o que é uma violação da hipótese de eficiência de mercado, já que se trabalharmos com a série em logaritmo, a primeira diferença dessa série corresponde a série de log-retornos.

Desta forma o teste de raiz unitária serve como um requisito inicial para a análise de eficiência de mercado, testando de forma robusta se o parâmetro que liga o preço em t com o preço em $t-1$ é igual a um, mas a não rejeição dessa hipótese não significa necessariamente evidência a favor da hipótese de eficiência de mercado. A interpretação correta de testes de raízes unitárias é verificar qual procedimento é necessário para induzir estacionariedade em um dada série não estacionária. Se a série possui uma raiz unitária o procedimento correto é tomar uma diferença, e se a série é estacionária em torno de uma tendência, o procedimento indicado é ajustar uma tendência linear aos dados.

1.6 Testes de Raiz Unitária e Eficiência de Mercado para a série de câmbio

1.6.1 Análise preliminar dos dados

Antes de iniciar uma seqüência de testes de raiz unitária e testes de eficiência de mercado, devemos mostrar quais os comportamentos estatísticos apresentados pela série de câmbio nominal pós-real, características que nos servirão para determinar o *design* e o poder dos testes que iremos realizar.



Figura 1.1: Taxa de Câmbio R\$/US\$

A taxa de câmbio Real/Dólar apresenta características interessantes. A amostra utilizada em nosso estudo consiste de 1882 observações diárias, de 01 de julho de 1994 até 4 de janeiro de 2002¹. Como mostra a Figura 1.1, a taxa de câmbio possui notavelmente três períodos com comportamentos distintos. 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.

O regime de bandas cambiais introduzido pelo ministro Gustavo Franco não se mostrou sustentável frente as crises externas ocorridas nesses períodos, sendo substituído por um regime de livre flutuação da taxa de câmbio, que corresponde em nossa amostra a observação 1337, data em que ocorreu a crise cambial de 13 de janeiro de 1999. Após essa mudança, fica claro pela figura 1.1 alguns períodos de valorização e desvalorização da taxa de câmbio.

Essa característica de distintos regimes será explicitamente usada em nosso estudo, já que basicamente podemos classificar o comportamento do câmbio no Brasil pós-Real em três períodos com características distintas. O primeiro período após a implantação

¹É 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

do Real corresponde a uma época de grande intervenção no câmbio, já que a taxa de câmbio era o principal instrumento de controle da inflação via valorização da taxa de câmbio e redução do valor em moeda interna dos preços externos. O segundo período, correspondendo ao regime de banda cambial (*target zone*), tem uma característica estatística interessante por ser um período de volatilidade restrita e lenta desvalorização do câmbio a uma taxa constante. O terceiro período, que corresponde ao final da amostra, é um período de flutuação livre e volatilidades maiores, mas com movimentos de valorização e desvalorização (*swings* curtos na nomenclatura de ENGEL e HAMILTON (1990)).

Essa divisão em três principais regimes corresponde aos regimes conhecidos de forma exógena, onde podemos determinar com precisão o início e o final de cada regime, já que cada um deles é correspondente a uma política cambial explicitamente divulgada pelo Banco Central do Brasil. No Capítulo 2, onde iremos trabalhar com um modelo de mudança markoviana, determinaremos estatisticamente a existência de 7 regimes, mas estes são determinados de forma endógena e a determinação da probabilidade associada a cada regime é estimada através dos dados da série, sem utilizar-se de informação externa.

Utilizaremos nesse trabalho quase que exclusivamente log-retornos do câmbio, que correspondem a uma capitalização contínua dos retornos de câmbio. Essa preferência deve-se o fato de que log-retornos possuem características estatísticas atrativas. A principal motivação para trabalhar com log-retornos é que estes são geralmente estacionários (covariância-estacionários). Uma segunda vantagem de se trabalhar com log-retornos em vez da série em nível é que estes apresentam, de forma mais intuitiva, o comportamento da volatilidade condicional da série, conforme aparece na Figura 1.2.

As estatísticas descritivas estão colocadas na Tabela 1.1 para as séries de câmbio em nível, retornos $(P_t - P_{t-1})/P_{t-1}$, e log-retornos. O fato mais notável é que os retornos não possuem distribuição normal ou log-normal no caso de log-retornos, já que os coeficientes de simetria estão distantes dos valores associados à distribuição normal. A simetria positiva indica a tendência de desvalorização do câmbio dado pelo maior número de retornos positivos do que negativos, enquanto que os valores do coeficiente de Kurtose indicam caudas muito mais pesadas do que seria correspondente a uma distribuição normal com mesma média e variância. A existência de excesso de kurtose aponta que valores extremos (ganhos e perdas excessivas) são mais frequentes nessa série do que seria esperado em uma distribuição normal equivalente. O gráficos da Figura 1.4 mostram o comportamento das caudas da distribuição de câmbio e log-retornos de câmbio em comparação a uma distribuição normal através do uso de um QQ-Plot (*quantile-quantile Plot*), mostrando grandes desvios nas caudas em relação aos valores esperados em uma distribuição normal com mesma média e variância dos dados, reforçando a evidência de não normalidade nas séries.

Como veremos mais profundamente no Capítulo 2 e no Capítulo 4, essas características são evidências de que o modelo gerador desses dados não pode ser considerado estável, indicando que o modelo mais adequado seria um modelo do tipo de mistura de distribuições como utilizado no capítulo 2 ou então modelos de variância condicional do tipo ARCH como estudado no capítulo 4.

O histograma das séries de câmbio e log-retornos de câmbio na Figura 1.3 aponta para algumas características interessantes. A primeira é que a multi-modalidade na densidade empírica da série de câmbio (primeiro gráfico na Figura 1.3) mostra que a série não pode ser considerada estacionária, e por isso medidas de média e variância incondicional não são representativas dessa série, já que nos processos não-estacionários do tipo diferença estacionária a média e a variância dependem do tempo. A densidade

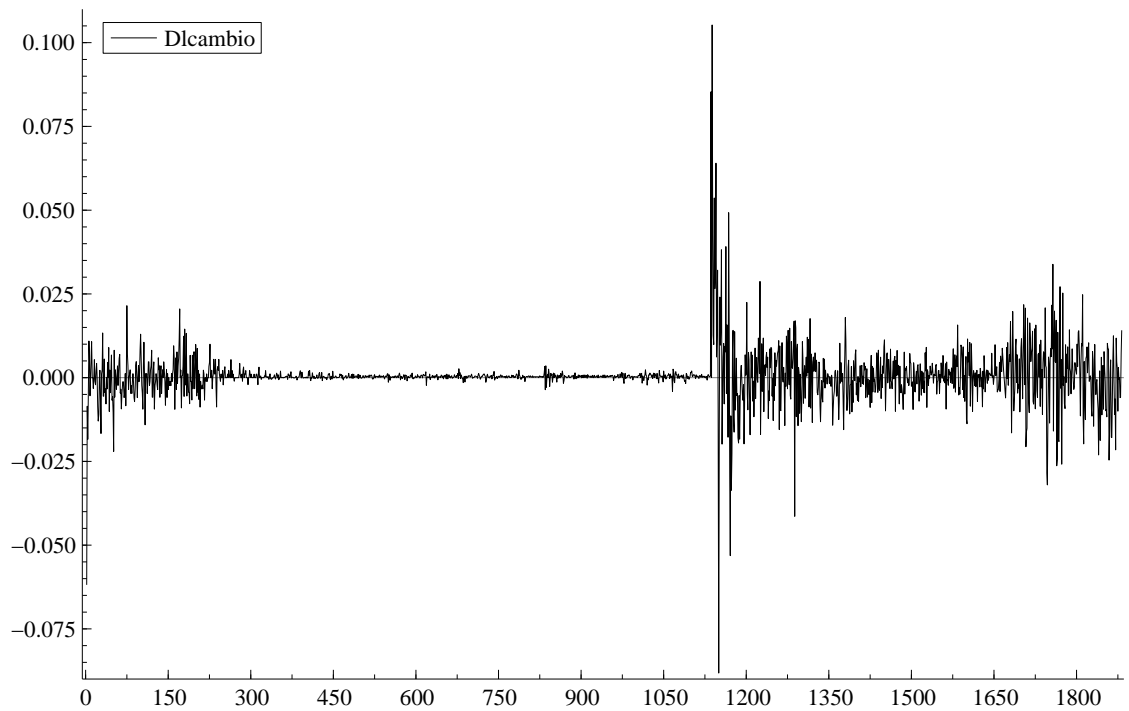


Figura 1.2: Log-Retornos do Câmbio

Tabela 1.1: Estatísticas Descritivas

	câmbio	retornos	log-retornos
média	1.417051	0.000753	0.000491
desvio padrão	0.5252	0.0141	0.007634
simetria	0.7734	0.2451	1.8322043
excesso de kurtose	2.3660	30.44971	48.22841
Jarque-Bera	211789	58759.63	160508.6
p-value- JB	0.0000	0.000	0.000

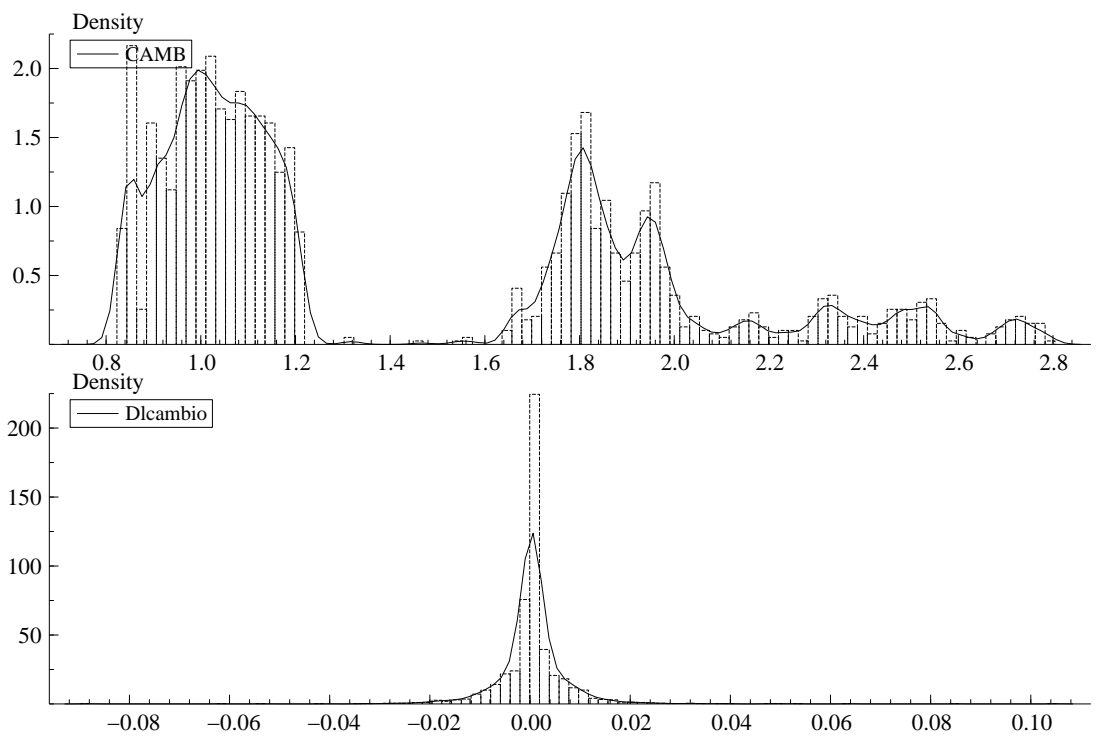


Figura 1.3: Histogramas

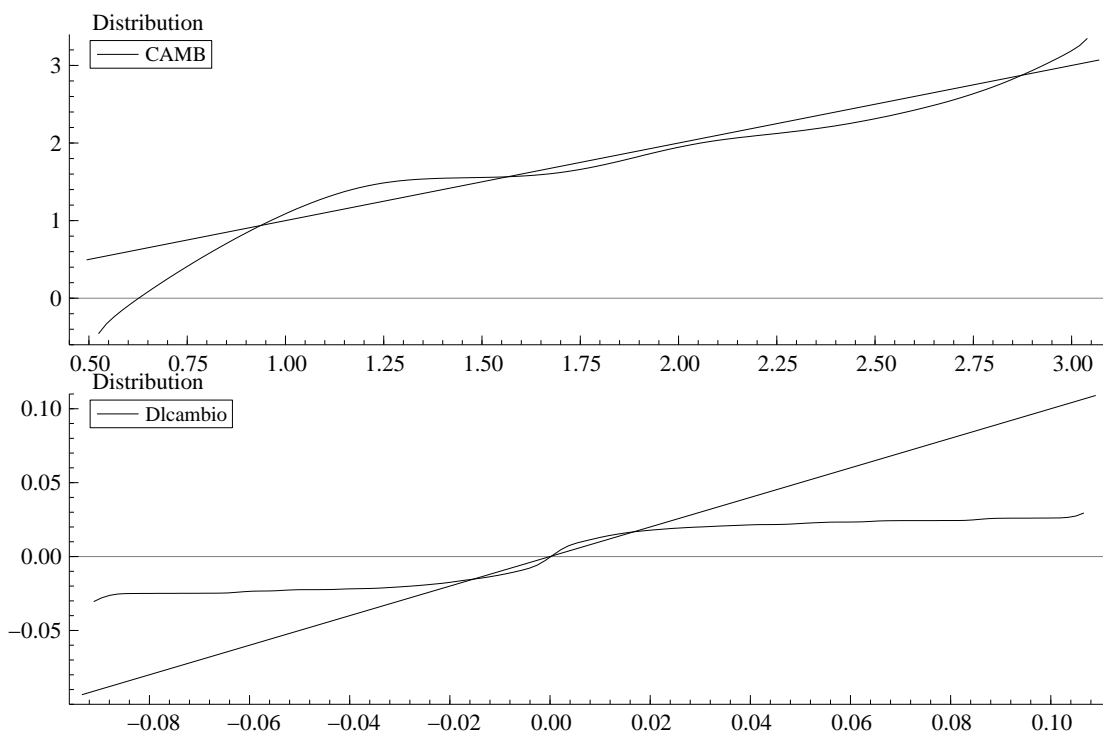


Figura 1.4: Densidades - Câmbio e Log-Retornos do Câmbio

empírica da série de log-retornos (segundo gráfico na Figura 1.3) confirma a evidência de caudas pesadas medidas pelo coeficiente de kurtose.

Em nossa análise da previsibilidade e da eficiência do mercado de câmbio duas referências importantes serão as funções de auto-correlação (ACF) e a função de auto-correlação parcial (PACF) das séries de câmbio e log-retornos de câmbio (Figuras 1.5 e 1.6). O ACF e o PACF da série de câmbio, expostos na figura 1.5, mostram que a primeira autocorrelação é próxima de um, com um decaimento muito lento, enquanto que só a primeira auto-correlação parcial é significativa. Isso evidencia que o processo que gera o câmbio em nível é um processo autoregressivo de primeira ordem AR(1) com um parâmetro autoregressivo próximo de um, o que leva a necessidade de realizarmos testes de raiz unitária para verificar a hipótese de que este parâmetro seja estatisticamente igual a um.

As funções de autocorrelação e autocorrelação parcial (ACF E PACF) da série de log-retornos de câmbio são bastante interessantes. Notamos que existem diversas auto-correlações estatisticamente significativas, e que o ACF e o PACF não permitem identificar com clareza a ordem de um modelo ARMA(p,q) para essa série. É interessante perceber também que embora as autocorrelações decaiam com o aumento do intervalo entre períodos, elas não decaem de forma exponencial como esperado em um modelo do tipo AR. Aparentemente, elas decaem de forma polinomial e, mesmo após um significativo número de períodos, ainda existem autocorrelações significativas. Este comportamento é compatível com a existência de memória longa nessa série, uma ca-

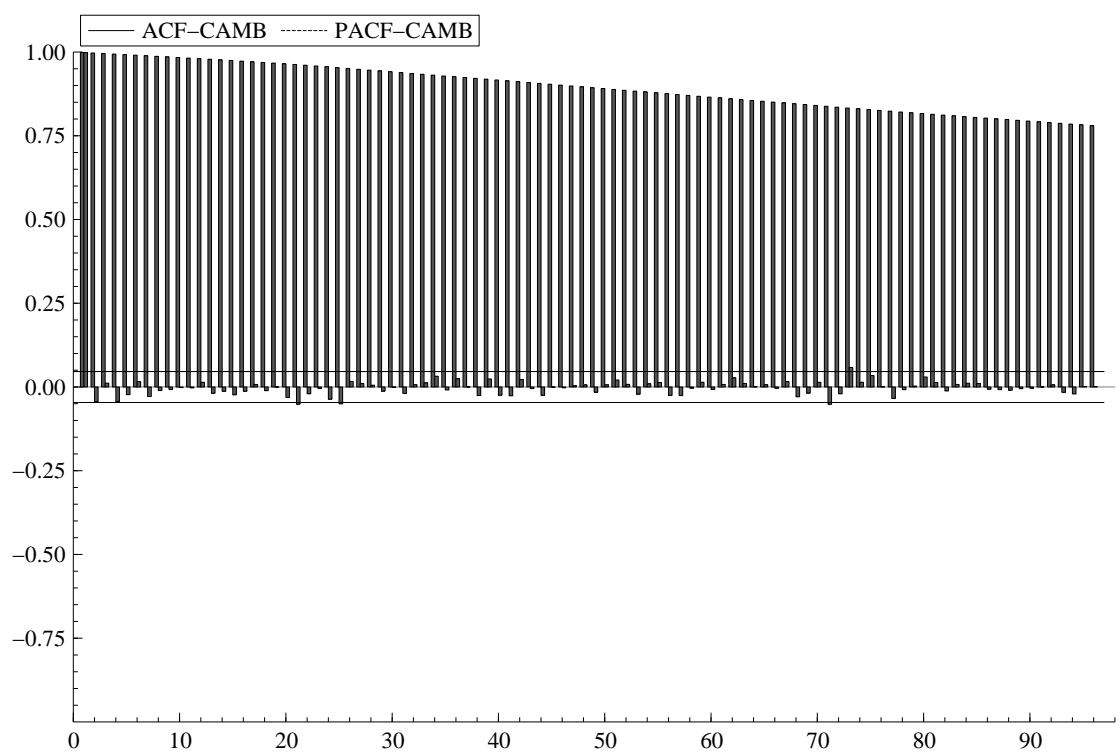


Figura 1.5: ACF - Câmbio

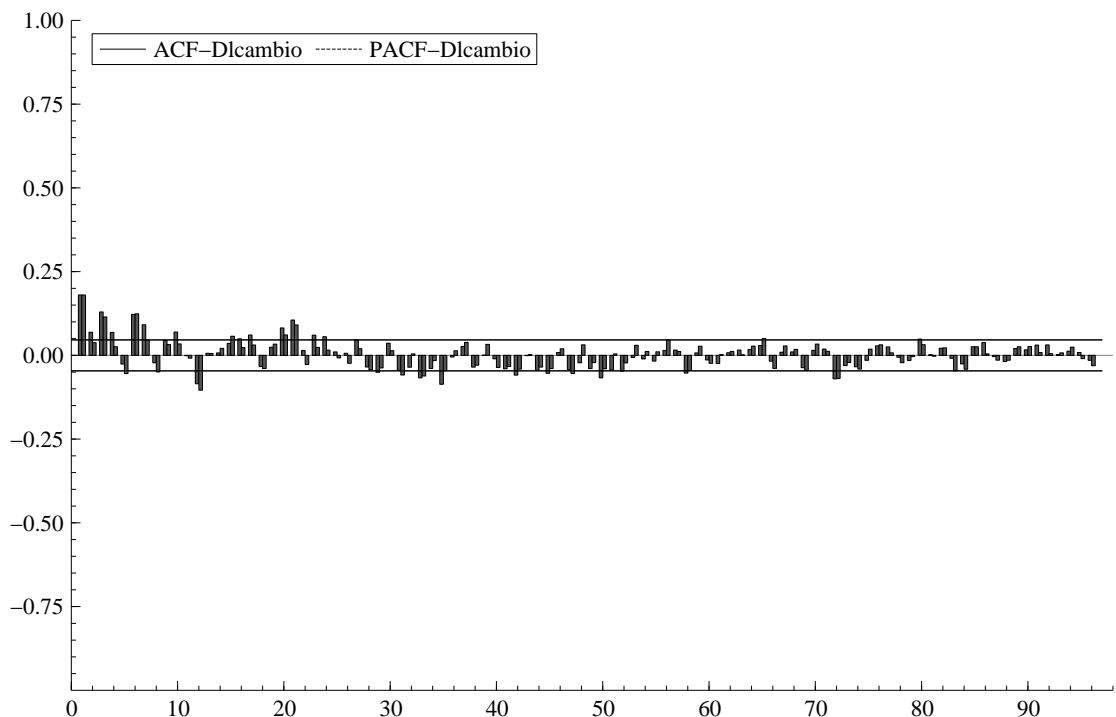


Figura 1.6: ACF - Log-Retornos do Câmbio

racterística que será estudada com profundidade no Capítulo 3.

1.6.2 Testes de Raiz Unitária

Dada a forte evidência de um parâmetro autoregressivo próximo a um mostrada pela Figura 1.5, temos de verificar a evidência de não estacionaridade causada por uma raiz unitária para a série de Câmbio em nível. A importância do teste de raiz unitária para nosso estudo é verificar se a relação entre a taxa de câmbio (P_t) e a taxa de câmbio no período imediatamente anterior (P_{t-1}) é dada por um parâmetro estatisticamente igual a um, o que seria uma evidência a favor do modelo de passeio aleatório².

Usamos três metodologias nesse teste: o teste ADF (DICKEY e FULLER (1979)), o teste KPSS (KWIATKOWSKI, PHILLIPS, SCHIMIDT e SHIN (1992)) e o teste PP de PHILLIPS E PERRON (1998). O teste ADF e o teste PHILLIPS-PERRON tem como hipótese nula a presença de raiz unitária, enquanto que no teste KPSS a hipótese nula é a de estacionaridade. Estes três testes diferem na forma de se construir a estatística de teste, mas são as metodologias mais utilizadas em testes de estacionaridade. Embora a construção de testes de raiz unitária seja um tópico em econometria extremamente fér-

²Conforme já colocamos, a raiz unitária é necessária mas não suficiente para a hipótese de Eficiência de Mercado. Um processo de raiz unitária pode ser mais geral do que o modelo de passeio aleatório correspondente a eficiência de mercado, já que na série em primeira-diferença podem existir auto-correlações significativas, o que seria um fator contra a hipótese de EM.

Tabela 1.2: Teste ADF Câmbio

modelo	Método Seleção	Lags	Est. de Teste	Valor Crítico 5%
cst, tendência	BIC	6	-2.455614	-3.4148
cst, tendência	AIC	21	-3.084253	-3.4148
cst,tendência	Wald	21	-3.084253	-3.4148
cst,tend., dummie*	BIC	6	-2.466431	-3.4148
cst,tend., dummie*	AIC	21	-5.024730	-3.4148*

Tabela 1.3: Teste KPSS

		Teste	KPSS		
variável - câmbio		estat	5%	10%	m
nível	teste com drift	0.8395*/**	0.146	0.119	32
diferença	teste sem drift	0.1172	0.463	0.347	32
diferença	teste com drift	0.0326	0.146	0.146	32
		*rejeição à 5%	**rejeição a 10%		

til, iremos apenas relatar os resultados do teste, remetendo o leitor a literatura específica do tema sobre os problemas de poder e tamanho existentes nesse teste e as metodologias mais recentes tornando o teste mais poderoso na presença de quebras estruturais e heterocedasticidade. Uma referência completa sobre esse tema pode ser encontrada em HAMILTON (1994), e sobre as abordagens mais recentes KIM e MADDALA (1998).

O teste ADF (*Augmented Dickey-Fuller*) foi construído inicialmente colocando-se tendência e intercepto na equação de teste, especificação que se mostrou significativa em todas as equações testadas. O número de defasagens colocadas para controle de auto-correlação, procedimento que visa aumentar o poder do teste, foi determinado utilizando-se os critérios de Schwartz, Akaike e um teste Wald de redução da ordem de defasagens. O número de defasagens selecionado em cada critério se encontra na Tabela 1.2. Não podemos rejeitar a hipótese de raiz unitária a nenhum nível crítico utilizando-se as especificações com tendência e intercepto na equação de teste.

Um procedimento posterior foi incluir na equação de teste um *dummie* de inclinação correspondendo ao período de bandas cambiais, para tentar obter evidências de que a presença de uma quebra estrutural poderia influenciar o teste de raiz unitária. Selecionando o número de intervalos pelo BIC, vemos que a conclusão não se altera, mas quando selecionamos o número de defasagens pelo AIC obtemos um resultado contraditório aos demais, já que a estatística de teste está acima do valor crítico de rejeição da hipótese nula. Uma ressalva importante é que isto é apenas uma evidência, já que os valores críticos utilizados no teste não são adequados a um teste onde a quebra estrutural é controlada usando uma *dummie* de inclinação da forma como foi construída (zeros antes e depois do regime de bandas, um durante o regime). Os valores críticos definidos por PERRON (1989), a referência mais importante no teste de raiz unitária na presença de quebras estruturais, *não* são adequados a este teste, e o procedimento fica apenas como uma evidência não conclusiva.

O teste KPSS tem como hipótese nula estacionaridade contra a hipótese alternativa de não-estacionaridade. Desta forma, a rejeição da hipótese nula seria consistente com

Tabela 1.4: Teste Phillips-Perron

Teste Phillips-Perron					
Série	Estatística- PP	1%	5%	10%	m
Câmbio	-2.460201	-3.9683	-3.4148	-3.1292	7
D(Câmbio)	-36.95256	-3.9683*	-3.4148*	-3.1292*	7

Tabela 1.5: Teste ADF Sub-amostras

	Modelo	Seleção	Lags	Observações	Est. de Teste	Valor Crítico
Pré-Banda	cst, tendência	BIC	11	13-300	-2.716233	-3.4268
Banda	cst, tendência	BIC	1	301-1137	2.338878	3.4175
Pós-Banda	cst, tendência	BIC	6	1138-1882	-2.605007	-3.4181

a presença de uma raiz unitária na série de câmbio em nível. Conforme a Figura 1.3 vemos que a hipótese nula de estacionaridade é rejeitada para a série em nível. Para confirmar este resultado, realizamos o teste de estacionaridade para a série em diferenças, já que se a série em nível é não estacionária devido a uma raiz unitária, a primeira diferença da série deve torná-la estacionária. Nos testes com a série em diferenças não podemos rejeitar a hipótese nula de estacionaridade, o que confirma nossos resultados.

O teste de PHILLIPS-PERRON obtém os mesmos resultados do teste ADF e do teste KPSS. O teste não rejeita a hipótese nula de raiz unitária para a série em nível, e rejeita a um elevado nível de significância a raiz unitária para a série em diferenças, o que é uma evidência de que a não-estacionaridade existente na série de câmbio em nível é devido ao parâmetro autoregressivo ser estatisticamente igual a um.

A presença de uma mudança no processo gerador dos dados, correspondendo a quebras estruturais, pode prejudicar o resultado do teste de raiz unitária. Para verificar se o regime de banda cambial poderia afetar o resultado do teste, realizamos o teste nas três sub-amostras representativas, conforma a Tabela 1.5. Os resultados do teste não se alteraram, ou seja, não conseguimos rejeitar a hipótese nula de raiz unitária em cada sub-amostra. Duas ressalvas são necessárias: a primeira é que a redução do tamanho da amostra reduz o poder do teste, principalmente na primeira sub-amostra que conta apenas com 287 observações após o ajuste de defasagens. A segunda é uma questão conceitual. Processos com raízes unitárias tem como característica o fato de que a variância deve explodir para infinito com o aumento do tamanho da amostra. Em um regime cambial, a variação em torno da tendência de desvalorização é controlada por intervenções do Banco Central. Desta forma a variância é restrita, o que não é consistente com um processo de raiz unitária. Outro fato relevante é notar que para o período de bandas cambiais foi encontrada uma estatística de teste positiva, indicando um processo explosivo, o que é inconsistente com a restrição de flutuação dada pela banda cambial. Este resultado nos coloca mais uma evidência de que o uso de testes de raiz unitária para períodos de bandas cambiais é inadequado.

Os resultados dos testes de raízes unitárias apontam para a não rejeição da hipótese nula da presença de uma raiz unitária nessa série. Desta forma, a taxa de câmbio no período em estudo é um processo não estacionário do tipo diferença estacionária, o que

Tabela 1.6: Teste BDS Independência

Teste BDS					
Amostra: 1 1882					
Dimensão	BDS Estat.	D. Padrão	Estat. z	Prob.Normal	Prob.Btrap
2	0.059795	0.003039	19.67719	0.0000	0.0000
3	0.118890	0.004852	24.50301	0.0000	0.0000
4	0.169718	0.005809	29.21424	0.0000	0.0000
5	0.203105	0.006090	33.34803	0.0000	0.0000
6	0.226233	0.005909	38.28321	0.0000	0.0000

implica que não existe dissipação de choques nesta série. O valor da taxa de câmbio é dado pela acumulação dos choques, sem tendências de reversão à média.

1.7 Testes de Eficiência de Mercado

1.7.1 Testes de *Random Walk I* e *Random Walk II*

Como já comentamos na Seção 1.3, os modelos de passeio aleatório tipo I e II são restritivos demais para serem válidos em séries financeiras reais. O pressuposto de que os incrementos na série de preços sejam de uma mesma distribuição e independentes não é sustentável nas situações de mudança estrutural e aprendizado dos agentes, conforme pode ser visto em KIRMAN e TEYSSIERRE (2001). A violação do pressuposto IID para o componente de inovações ε_t devido a mudanças no processo gerador da série de câmbio é a característica mais importante que os estudos realizados nos Capítulos 2, 3 e 4 tentam capturar.

A simples visualização da Figura 1.3 mostra que a densidade empírica da taxa de câmbio e dos log-retornos não pode ser originária de uma distribuição IID, já que além de mostrar que a série de câmbio em nível é não estacionária, ilustra o comportamento dos diversos regimes existentes na taxa de câmbio pós- Real. Para confirmar estatisticamente a violação do passeio aleatório dos tipos I e II usamos a estatística BDS de BROCK, DECHERT e SCHEINKMAN (1987). A estatística BDS adapta o conceito de integral de correlação da literatura de sistemas caóticos proposta por GRASSBERGER e PROCACCIA (1983) para estabelecer um teste de independência adequado à quantidade finita de dados existentes em séries econômicas.

A estatística BDS é calculada usando a seguinte formulação:

$$J_{n,T}(k) = \sqrt{T} \frac{C_{n,T}(k) - C_{1,T}(k)^n}{\hat{\sigma}_{n,T}(k)} \quad (1.25)$$

onde $C_{n,T}(k)$ e $C_{1,T}(k)^n$ são as integrais de correlação definidas em GRASSBERGER e PROCACCIA (1983). Sob a hipótese nula de IID para os incrementos a estatística tem distribuição assintótica Normal. O poder e o tamanho desse teste para amostras finitas são estabelecidos por HSIEH (1988). Como esperado, a estatística BDS rejeita a qualquer nível de significância que os log-retornos de câmbio sejam uma seqüência IID. Desta forma existe alguma estrutura, linear ou não-linear, nos log-retornos.

Tabela 1.7: Testes de Linearidade na média

Teste <i>Neural Network</i> - Teräsvirta		
χ Quadrado = 6.4885	df = 50	p-value = < 2.2e-16
Teste <i>Neural Network</i> - White		
χ Quadrado = 36.2891	df=2	p-value=1.318e-08
Teste White - Regressão 12 lags		
χ Quadrado = 29.2501	df = 2	p-value = 4.451e-07

Tabela 1.8: BDS Resíduos do Modelo Linear

Teste BDS	RESID01			
Amostra: 13 1882				
Dimensão	BDS Estat.	D. Padrão	Estat. z	Prob.Normal
2	0.054853	0.003034	18.07861	0.0000
3	0.115550	0.004841	23.86796	0.0000
4	0.167904	0.005792	28.98713	0.0000
5	0.203718	0.006068	33.57100	0.0000
6	0.226750	0.005884	38.53876	0.0000

Como primeiro passo na determinação da forma funcional da estrutura existente testamos a linearidade do processo gerador da média dos dados. Utilizamos os testes de TERASVIRTA, LIN e GRANGER (1993) e LEE, WHITE e GRANGER (1993) para não linearidade na média, conhecido na literatura como teste de WHITE para não linearidade.

O procedimento de Teräsvirta tem como hipótese nula a linearidade na média. O teste usa uma expansão de séries de Taylor da função de ativação de uma rede neural como forma de obter a estatística de teste. O resultado deste teste leva à rejeição da hipótese nula de que a média do processo seja gerada por um mecanismo gerador linear. O teste de WHITE usa também uma rede neural, testando para a existência de não-linearidade na série, podendo ser utilizado na própria série ou em resíduos de uma regressão. Na Tabela 1.7 vemos que a hipótese de linearidade na média é rejeitada pelos testes de Teräsvirta e White. Nessa Tabela também vemos que um modelo autoregressivo com 12 lags não consegue ajustar toda a estrutura existente na série, e o teste aponta que um modelo não-linear é necessário.

Para confirmar a necessidade de um modelo não-linear na modelagem dessa série realizamos um teste BDS nos resíduos de um modelo AR(12) para log-retornos. Nesse caso, a nossa hipótese nula é de que um modelo linear capture toda a estrutura de dependência na média, enquanto que a hipótese alternativa é que existe alguma estrutura não-linear restante nos resíduos do modelo linear. Conforme a Tabela 1.8, rejeitamos que um modelo linear autoregressivo seja capaz de capturar a estrutura de dependência existente na média, assim ficando a necessidade de buscar alguma forma de modelo não-linear para a correta modelagem da estrutura existente na série de log-retornos de câmbio.

Para testar se os resíduos ao quadrado da série possuem alguma estrutura de dependência, o que também levaria a rejeição da hipótese nula de IID no teste BDS,

realizamos um teste ARCH nos resíduos de um modelo AR(12) para os log-retornos do câmbio. O teste ARCH é baseado no princípio LM (*Lagrange Multiplier*), foi realizado estimando-se um modelo autoregressivo tendo como variável dependente os resíduos ao quadrado contra as n primeiras defasagens do resíduo ao quadrado. Sob a hipótese nula de não existência de auto-regressividade nos resíduos ao quadrado, o tamanho da amostra vezes o R^2 dessa regressão tem distribuição $\chi^2(n)$, onde n é a ordem do processo ARCH em teste. Realizando esta regressão obtemos um R^2 de 0.06401, o que leva a uma estatística de teste $1876 \cdot 0.06401 = 120.028$. O valor crítico de 95% de uma distribuição $\chi^2(12)$ é igual a 21.02607. Rejeitamos a hipótese nula de que não existe uma estrutura ARCH nos resíduos.

Pelos testes BDS rejeitamos a hipótese de que os log-retornos sejam independentes e provenientes de uma mesma distribuição. Os testes de não-linearidade na média e pelo teste ARCH, mostram que possivelmente a violação da independência devem-se a existência de estruturas não-lineares na média e também na variância.

1.7.2 Testes de *Random Walk* III

A forma mais comum de se realizar testes de eficiência de mercado³ é através do uso do modelo de passeio aleatório do tipo III, que pressupõe apenas ausência de correlação nos incrementos. Esta forma é, conforme visto na Seção 1.3 bem menos restritiva e assim mais factível para séries reais.

Uma forma bem simples de se testar eficiência fraca usando o modelo de passeio aleatório é através de uma regressão da forma:

$$R_t = \alpha + \sum_{i=1}^n \beta_n R_{t-i} + e_t \quad (1.26)$$

A hipótese nula de mercado eficiente corresponde a um teste de que todos os parâmetros β_n são iguais a zero contra a hipótese alternativa de que pelo menos um dos β_n é estatisticamente diferente de zero, o que corresponde a violação da eficiência de mercado. Esse teste conjunto pode ser realizado utilizando-se um teste F tradicional utilizado para a verificar a existência de uma regressão.

Realizamos uma primeira regressão com 30 lags de log-retornos de câmbio como variáveis explicativas. O resultado da estatística F dessa regressão foi $F(30,1820) = 7.438$ com um p-valor de 0.000, o que rejeita a qualquer nível de significância que os 30 primeiros lags sejam não significantes. Para aumentar o poder do teste, limitamos o número de lags a 12, que corresponderam aos parâmetros significativos pelo teste t. No teste com 12 lags a estatística foi $F(12,1838) = 14.22$ com novamente um p-valor de 0.000, rejeitando também que os 12 primeiros lags sejam não significantes. Uma forma alternativa de se realizar este procedimento é através de um teste *Portmanteau* para as n auto-correlações em teste. Realizando o teste *Portmanteau* obtivemos uma estatística de 191.146, o que correspondeu a um p-valor de 0.000, resultado consistente com o obtido através das regressões.

Um teste alternativo de aleatoriedade nos retornos é conhecido como teste *runs*. Se os incrementos não têm estrutura de dependência e correlação, espera-se que o nú-

³Uma ressalva importante a se fazer sobre toda a discussão de testes de Eficiência de Mercado se refere a que na verdade este teste é um teste conjunto da eficiência de mercado com a validade do processo gerador que estamos estabelecendo para os ativos. A hipótese nula é a de uma hipótese conjunta de equilíbrio e eficiência de mercado, conforme discutido em Campbell, Lo e Mackinlay, pgs. 24 e 25. O mais relevante é analisar como desvios da eficiência de mercado poderiam ser utilizados na busca de lucros excessivos através de previsibilidade e arbitragem.

Tabela 1.9: Teste *Runs*

	1-1882	1-300	301-1137	1137-1882
<i>Runs</i>	606	95	231	282
dist.	N(870.50, 20.84)	N(151.00, 20.84)	N(419.50, 20.84)	N(302.00, 20.844)
<i>z-score</i>	12.689099	2.686539	9.043082	0.959478
<i>p-value</i>	0.000000	0.007219	0.000000	0.337318

mero de retornos positivos e negativos seja estatisticamente igual. Quando ocorre uma mudança no sinal dos retornos, dizemos que houve um *run*. O teste *runs* é realizado contando-se o número de mudanças de sinal existentes na série. Na hipótese nula de aleatoriedade nos incrementos, esta estatística tem distribuição assintoticamente Normal. A Tabela 1.9 contém o resultado do teste para a amostra toda e para as sub-amostras que utilizamos anteriormente. Vemos que o único período em que não rejeitamos a hipótese nula é o período de livre flutuação do câmbio, enquanto que para a série toda e os demais sub-períodos a hipótese nula de RW III é rejeitada.

O mais utilizado teste do modelo RW III no entanto é o teste de razão de variância (*variance ratio*). O teste de razão de variância (VR) foi proposto originalmente por COCHRANE (1984) para verificar o tamanho do componente de *Random Walk* no PNB. Na hipótese nula de que o processo gerador da série é um passeio aleatório, a variância dos retornos de n -períodos deve ser proporcional ao número de períodos.

Definindo o retorno de n períodos como

$$R_t^n = P_t - P_{t-k} \quad (1.27)$$

se o processo gerador da série é um passeio aleatório, os retornos de um período $P_t - P_{t-1} = \mu_t$ devem ter uma distribuição IID com médias μ e variância constante σ^2 . Como os retornos de n períodos são a acumulação de n sucessivos μ_t a variância de R_t^n deve ser igual a $n \times \sigma^2$. A estatística VR (*variance ratio*) é definida como:

$$VR(n) = \frac{Var(R^n)}{Var(R^1)} \frac{1}{n}$$

que deve ser igual a um na hipótese nula de *random walk*. LO e MACKINLAY (1989) derivaram a distribuição amostral desse teste em amostras finitas e também um estimador para essa estatística que é consistente na existência de heterocedasticidade. Uma derivação completa do teste e das distribuições assintóticas pode ser encontrada em CAMPBELL, LO e MACKINLAY (1997). Utilizamos o estimador consistente para heterocedasticidade bem como as distribuições de teste de LO e MACKINLAY (1989) para construir os testes VR. Calculamos a razão de variância para retornos de 5, 10, 20, 50, 75, 100 e 300 períodos, para analisar se de alguma forma existiria um poder de previsão tanto para períodos curtos quanto para períodos mais longos de tempo. Seguindo nossa metodologia, estimamos esses valores tanto para a série toda quanto para os sub-períodos relevantes.

Para a amostra toda, a hipótese nula de um passeio aleatório é rejeitada para retornos de até 50 períodos, enquanto que para períodos mais longos a hipótese nula não é rejeitada. Isso seria uma evidência de que teríamos poder de previsão para horizontes curtos, mas em horizontes mais longos, acima de 3 meses, seria válida a hipótese

Tabela 1.10: Variance Ratio Amostra completa

Variance Ratio			
q	VR	psi	<i>p</i> -valor
5	1.4935	2.9559	0.0031
10	1.8392	3.2939	9.8817e-004
20	2.1866	3.3171	9.0973e-004
50	2.4678	2.8140	0.0049
75	2.1229	1.9574	0.0503
100	2.0267	1.5660	0.1173
300	1.8184	1.0049	0.3149

Tabela 1.11: Variance Ratio Pré Banda Cambial

Variance Ratio			
q	VR	psi	<i>p</i> -value
5	0.8781	-0.7202	0.4714
10	0.8860	-0.4712	0.6375
20	0.9162	-0.2641	0.7917
50	1.4609	0.9909	0.3218
75	1.6463	1.1576	0.2470
100	2.0382	1.6396	0.1011
300	-	-	-

de eficiência fraca de mercado representada pelo modelo de passeio aleatório. Entretanto é importante notar que, para horizontes mais longos, o tamanho da amostra se reduz consideravelmente, possivelmente reduzindo o poder do teste e introduzindo a possibilidade de viés a favor da não rejeição da hipótese nula.

Esse é um fenômeno interessante, já que as evidências encontradas na literatura são que o modelo de passeio aleatório seria válido para períodos curtos de tempo, enquanto que no longo prazo teríamos poder de previsão, o que contraria os resultados encontrados.

Uma questão fundamental é verificar se de alguma forma a presença de quebras estruturais, ou como colocado por KIM, NELSON e STARTZ (1998), a presença de padrões de heterocedasticidade, poderiam afetar o resultado do nosso teste. A forma mais simples de verificar esta possibilidade foi então analisar o comportamento da estatística VR nos sub-períodos em que conhecíamos com certeza a existência de mudanças estruturais na série, dadas pelas mudanças no regime cambial.

No sub-período correspondente à introdução do Real até à adoção do regime de bandas cambiais não conseguimos rejeitar a hipótese nula de que o processo gerador de câmbio seja um passeio aleatório para todos os intervalos de retornos testados, conforme Tabela 1.11. Para o sub-período posterior (bandas cambiais), o teste rejeita o modelo de passeio aleatório para retornos de até 100 períodos (Tabela 1.12), o que seria uma evidência de violação do modelo de passeio aleatório.

Uma ressalva é necessária no entanto. A estatística VR assume que a acumulação das variâncias seja um processo irrestrito, mas o que acontece nesse período é que, dada

Tabela 1.12: Variance Ratio Banda Cambial

Variance Ratio			
q	VR	psi	<i>p-value</i>
5	0.2276	-108.0825	0
10	0.4580	-3.2386	0.0012
20	0.2460	-3.4694	5.2162e-004
50	0.1424	-2.8472	0.0044
75	0.1225	-2.5382	0.0111
100	0.1005	-2.3291	0.0199
300	0.0614	-1.4656	0.1428

Tabela 1.13: Variance Ratio Pós Banda Cambial

Variance Ratio			
q	VR	psi	<i>p-value</i>
5	1.2984	1.9315	0.0534
10	1.2384	0.9751	0.3295
20	1.0595	0.1677	0.8669
50	0.8974	-0.1908	0.8487
75	0.9367	-0.1019	0.9188
100	0.9474	-0.0798	0.9364
300	3.4112	2.5488	0.0108

a existência das bandas, existiam limites de flutuação para a taxa de câmbio e, desta forma, o processo da variância era restrito. Assim a evidência de previsibilidade apontada pela estatística VR para a série de câmbio nesse período deve ser analisada com cuidado, já que a estatística não é definida nesse contexto de volatilidade controlada.

No sub-período final da série (Tabela 1.13), correspondendo ao período de livre flutuação do valor da taxa de câmbio, vemos que a hipótese nula de passeio aleatório não é rejeitada para nenhum dos intervalos de retornos, e é interessante notar que com o aumento do horizonte de retornos a estatística VR se torna cada vez mais próxima de um, o que é contrário à visão de que o aumento do horizonte levaria a um maior poder de previsibilidade.

O uso da estatística de razão de variância (VR) encontra evidências favoráveis à eficiência fraca de mercado, já que o modelo de passeio aleatório se mostra consistente para os períodos iniciais e finais da amostra e, no período de vigência das bandas cambiais, a estatística VR, que rejeita a hipótese nula, não pode ser considerada robusta porque o pressuposto de variância irrestrita não é respeitado.

1.8 Conclusões

Nesse capítulo apresentamos os modelos de precificação de ativos que iremos utilizar como forma de testar persistência e eficiência fraca de mercado. Os resultados empíricos dos testes realizados, no entanto, não são consistentes.

Verificamos que a presença de uma raiz unitária na série faz com que as inovações

tenham efeitos permanentes e, dessa forma a taxa de câmbio em nível é um processo não-estacionário em que o efeito dos choques não se dissipa no tempo, sendo necessária a presença de um choque de sinal e magnitude opostas para compensar o efeito de um choque passado.

Utilizando uma versão mais restrita do que a de um processo de raiz unitária, através de um modelo de passeio aleatório, encontramos evidências de que o processo gerador de log-retornos de câmbio não é um processo IID (independente e identicamente distribuído) compatível com modelos de passeio aleatório do tipo II e III. Nesta mesma análise encontramos fortes evidências de que o processo gerador desta série é não-linear na média e na variância.

Os resultados dos testes para o modelo de passeio aleatório III no entanto não são conclusivos. Embora pela regressão linear e pelo teste *Portmanteau* rejeitamos a não existência de correlação na série de log-retornos, pelos testes *runs* e de razão de variância vimos que as mudanças estruturais existentes na série influenciam o resultado sobre a existência de ineficiências de mercado, já que por estes testes o modelo de passeio aleatório não é rejeitado pelos dados para o último período da série e, pelo teste, VR o único período que rejeita o modelo de passeio aleatório é o de bandas cambiais, mas com as restrições que colocamos quanto a robustez do teste nestas situações.

O mais importante porém é analisar que, se de alguma forma encontramos evidências de violação da eficiência de mercado para a taxa de câmbio, é necessário verificar se estas evidências são realmente significantes. Isto leva à a necessidade de se encontrar o modelo mais adequado para modelar a estrutura existente na série, e temos evidências da necessidade de um processo não linear. É obrigatório verificar se o poder de previsão encontrado é economicamente significativo, no sentido de representar uma possibilidade real de arbitragem com a taxa de câmbio utilizando apenas a informação contida no passado da série.

O passo posterior em nossa série é a busca de um modelo econométrico não-linear adequado ao comportamento de nossa série. O primeiro modelo que iremos testar é o modelo conhecido como mudança markoviana, que vem atraindo grande atenção na literatura econômica nos últimos anos.

Notas adicionais

Os testes ADF, PHILLIPS-PERRON e BDS foram realizados utilizando se o software Eviews versão 4.0. A estatística KPSS foi calculada usando o software EasyReg de Herman J. Bierens. O teste *Runs* foi calculado com o software GRET (Gnu Regression, Econometrics and Time Series Library de Allin Cottrell). As estatísticas de Razão de Variância consistentes sob heterocedasticidade foram calculadas usando o módulo VARRATIO para Matlab de Ludwig Kansler. Os testes de não-linearidade de Teräsvirta e WHITE foram calculados usando o pacote TSeries para a linguagem R. Os procedimentos restantes utilizaram o software PcGive.

Capítulo 2

Modelo de Mudança Markoviana

2.1 Introdução

Neste capítulo discutiremos a aplicação de um modelo de mudança markoviana para os log-retornos de câmbio. A aplicação deste modelo é derivada do reconhecimento de duas características presentes na série de log-retornos de câmbio. A primeira delas é o reconhecimento da necessidade de um modelo não-linear para explicar o comportamento da série de log-retornos, o que foi identificado pelos testes realizados no Capítulo 1.

A segunda característica é a existência de quebras estruturais no mecanismo estatístico gerador da série de câmbio, o que nos leva em busca de modelos estatísticos que incorporem a mudança nos parâmetros ocorrida nesta série. A metodologia que utilizamos no Capítulo 1, dividindo a série em três sub-períodos, era baseada em particionar a série em períodos em que a existência de uma quebra era conhecida com certeza, ou seja, quando sabíamos com exatidão como demarcar o início e o fim de cada regime.

A classificação que empregamos era 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. Desta forma existia a necessidade de se buscar um modelo econométrico que identificasse as quebras estruturais através da informação existente na amostra, ou seja, um método estatístico para estimar as datas onde aconteceram quebras estruturais.

Uma técnica estatística reconhecidamente capaz de capturar esta mudança de parâmetros é a classe de modelos conhecidos 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. Uma grande vantagem do modelo de mudança markoviana é que sua interpretação econômica é mais trivial do que a grande maioria dos modelos não-lineares, já que, condicionado ao conhecimento do regime vigente, estaremos utilizando um modelo linear autoregressivo, cujo comportamento já foi exaustivamente analisado na literatura.

2.2 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.

O comportamento da série é governado pela equação correspondente ao regime que vigora. Um exemplo simples seria um modelo autoregressivo de primeira ordem sujeito a mudanças no parâmetro autoregressivo, representado pelo seguinte sistema :

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \epsilon_{1t} | se\ s_t = 1 \\ Y_t &= \phi_2 Y_{t-1} + \epsilon_{2t} | se\ s_t = 2 \end{aligned} \quad (2.1)$$

onde r representa o regime vigente. Desta forma o parâmetro ϕ_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 ϕ_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

$$P \{s_t = j | s_{t-1} = k, \dots\} = P \{s_t = j | s_{t-1} = i\} = P_{ij} \quad (2.2)$$

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

$$P_{i1} + P_{i2} + \dots + P_{in} = 1 \quad (2.3)$$

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 .

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix} \quad (2.4)$$

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 MACKINLAY (1997), 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.

2.2.1 Estimação dos parâmetros desconhecidos

A estimação destes modelos é realizada em nosso estudo através de métodos de máxima verossimilhança. As principais formas constituem aproximações da função de verossimilhança, como o algoritmo utilizado por HAMILTON (1989), ou então através do algoritmo EM (*Expectation-Maximization*) de DEMPSTEIR, LAIRD e RUBIN (1977). KIM (1989) desenvolve um filtro baseado na estimação por máxima verossimilhança que permite uma estimação computacionalmente mais eficiente de modelos auto-regressivos com mudança markoviana com formulações mais complexas e mais estados

KIM e NELSON (1993) e (1999), KIM e YOO (1995), desenvolvem novas formas de se estimar os parâmetros destes modelo, com bases em formulações em espaço de estado e alterações no filtro de Kalman que permitem a estimação dessas formas não lineares. O passo seguinte é evitar as aproximações utilizadas no algoritmos de estimação utilizando-se métodos de *Monte Carlo Markov Chain* (MCMC) para a estimação Bayesiana destes parâmetros, mais especificamente através de amostragem de Gibbs e amostragem de Gibbs com rejeição (*rejection sampling*). Estes procedimentos de estimação Bayesiana usando MCMC foram introduzidos de forma independente por ALBERT e CHIB (1993) e MCCULLOCH e TSAY (1994).

2.3 Revisão das Aplicações

A aplicação dos modelos de mudança markoviana em econometria remonta ao modelo de *switching regression* de GOLDFELD e QUANDT (1973), e, em estatística, ao modelos de *Hidden Markov Chain* e de misturas de distribuições. Mas a grande influência destes modelos em econometria decorre dos trabalhos de HAMILTON (1989), que generaliza o modelo de mudança markoviana para autoregressões. HAMILTON e SUSMEL (1994), DUEKER (1997), CAI (1994) e GRAY (1996) fazem generalizações do modelo de mudança markoviana para processos onde os parâmetros de modelos ARCH e GARCH são sujeitos a regimes determinados por uma cadeia de Markov.

A utilização empírica destes modelos é bastante ampla hoje. Os trabalhos clássicos dos modelos de mudança markoviana referem-se ao estudo de ciclos econômicos de HAMILTON, e o estudo do fenômeno conhecido como *long swing* do Dólar no estudo de ENGEL e HAMILTON (1990). Outras aplicações do modelo de mudança markoviana, afora as já citadas generalizações do modelo de mudança markoviana para processo ARCH/GARCH incluem modelos de múltiplos regimes de GARCIA e PERRON (1990), onde a taxa de juros é modelada através de três regimes de média e variância distintas, onde se testa a hipótese de que a taxa de juros é constante dentro de cada regime; GRAY (1996) e DAHQUIST e GRAY (2000) aplicam modelos de mudança markoviana generalizada para as séries de juros, e mostram que o modelo de mudança markoviana generalizado domina diversas formulações existentes, como os modelos de difusão, modelos quadráticos e especificações ARCH-GARCH.

O estudo do mercado acionário através de modelos de mudança markoviana foi feito nos trabalhos de CAI (1994) e DUEKER (1997). No Brasil, os estudo de VALLS PEREIRA, HOTTA, SOUZA e ALMEIDA (1999), RABI e VALLS PEREIRA (2001) comparam o ajuste dos modelos de mudança de regimes e SWARCH para o índice BOVESPA, retornos de Telebrás e retorno de títulos de 30 anos do Tesouro Americano contra especificações de passeio aleatório e GARCH, encontrando evidências a favor dos modelos de mudança markoviana.

A análise das taxas de inflação por esta classe de modelos é realizada nos estudos de DUEKER e FISHER (1996) e ERLANDSSON (2000). KIM (1993) aplica um modelo de mudança markoviana onde uma das especificações testadas é a que as inovações na série de inflação provêm de regimes distintos, e com isso faz uma análise entre inflação esperada e incerteza nas taxas de inflação.

Uma área em que os modelos de mudança markoviana tem sido especialmente produtivos é na análise de ciclos econômicos. Os trabalhos de DIEBOLD e RUDEBUSH (1996), KAUFMANN (2000), KIM e NELSON (1998) e KROLZIG (1997), entre outros, mostram que o modelo de mudança markoviana consegue capturar um dos fatos estilizados presentes no estudo de ciclos econômicos, que é a existência de não-linearidades dadas por quebras no processo. Os trabalhos de DIEBOLD e RUDEBUSH (1996) e de KIM e NELSON (1998) generalizam o modelo de fatores comuns de STOCK e WATSON (1989) para incluir a possibilidade de mudança de regimes, e com isso estes modelos conseguem capturar os dois fatos estilizados existentes na literatura de ciclos econômicos, que é a existência de co-movimentos entre as variáveis e as quebras estruturais existentes nestas séries.

Em relação ao tema de nosso estudo, a modelagem da taxa de câmbio, os já citados trabalhos de ENGEL (1992) e ENGEL e HAMILTON (1990) representam os marcos iniciais dessa abordagem. HAMILTON (1996) utiliza o estudo do câmbio para também desenvolver uma série de testes de especificação para os modelos de mudança markoviana, já que dado a presença de *nuisance parameters* (parâmetros de "incômodo", alguns dos testes habituais de especificação não são válidos. Exemplos mais recentes do estudo de câmbio incluem o trabalho de BERGMAN e HANSSON (2000) no estudo de taxas reais de câmbio, BOLLEN, GRAY e WHALEY (2000), que utilizam a modelagem de regimes para preços de opções. MASSON (2000) testa um modelo de *Hidden Markov Chains* para capturar a probabilidade de transição de regimes de câmbio, e ROSSI (2000) trabalha as questões de poder de previsão e persistência nos modelos de mudança markoviana.

2.4 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 adequação das especificações. 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.

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 (1997) 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 e significativa.

O modelos foram estimados utilizando-se o algoritmo EM de Dempster, LAIRD e RUBIN (1997) na forma do filtro BHLK (Baum-Lindgren-Hamilton-Kim) usando o software MSVAR de KROLZIG (1997). Para discutir de forma adequada este método de estimação seria necessária uma apresentação que foge do escopo desta dissertação. A monografia de KROLZIG (1997) (Capítulos 5, 6 e 8) apresenta uma discussão completa sobre o assunto.

Tabela 2.1: 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

2.4.1 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 (1992) e GARCIA e PERRON (1990), 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. O segundo é usar uma aproximação da distribuição assintótica do teste, devida a ANG e BEKAERT (1998). A terceira é realizar testes de especificação para checar a necessidade de um regime adicional.

Na Tabela 2.1 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 um modelo com um regime que corresponde ao modelo linear. O número de defasagens foi selecionado através dos critérios de informação de AIC e BIC, escolhendo um número de 5 defasagens como o mais adequado.

Pela Tabela 2.1 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 (1998) 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 parâme-

Tabela 2.2: Teste LR Número de Regimes

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

Tabela 2.3: Teste BDS Especificação

Teste BDS		BDSSTR6			Teste BDS		BDSSTR7		
Dimensão	BDS Estat.	D. Padrão	Estatística z	<i>p-value</i>	Dimensão	BDS Esta.	D. Padrão	Estatística z	<i>p-value</i>
2	-0.002602	0.001363	-1.909946	0.0461	2	-0.001400	0.001307	-1.071312	0.2840
3	-0.004220	0.002159	-1.954688	0.0406	3	-0.000443	0.002068	-0.214278	0.8303
4	-0.003577	0.002562	-1.395865	0.1628	4	0.001066	0.002451	0.434681	0.6638
5	-0.003316	0.002661	-1.245971	0.2128	5	0.002315	0.002543	0.910344	0.3626
6	-0.002880	0.002557	-1.126457	0.2600	6	0.002775	0.002441	1.137098	0.2555

tros 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 2.2 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.

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 2.3 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.

Tabela 2.4: 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] **

2.4.2 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 padrão. Usando o princípio de razão de verossimilhança, a estatística de teste é $2(\log\text{-verossimilhança}(\text{modelo irrestrito}) - \log\text{-verossimilhança}(\text{modelo restrito}))$, e a distribuição de teste é uma qui-quadrada com o número de graus de liberdade correspondente ao número de restrições impostas. Testamos o modelo geral MSIAH (mudanças no intercepto, parâmetros auto-regressivos e variância) contra as demais especificações possíveis. Conforme a Tabela 2.4, vemos que pelo teste LR (razão de verossimilhança), a especificação adequada é a do modelo MSIAH, com mudanças em todos os parâmetros, já que as demais especificações que restringem que alguns dos parâmetros sejam constantes através dos regimes são inadequadas.

É interessante notar pelas estatísticas LR calculadas que o fator que causa mais perda no valor da log-verossimilhança é a restrição de o modelo ter uma única variância, sem mudança markoviana nos valores da variância incondicional. Os testes do modelo MSIAH contra os modelos MSIA e MSI mostram que uma característica fundamental do modelo de mudança markoviana utilizado é capturar de forma adequada às mudanças na variância incondicional da série.

A análise dos resíduos, presente na Tabela 2.5 e na Figura 2.1 mostra que os erros de predição do modelo sofrem de problemas de autocorrelação, ARCH e heterocedasticidade. Mas os resíduos padronizados, ou seja, os resíduos divididos pela variância correspondente ao regime que tem maior probabilidade de pertencerem, são não correlacionados, homocedásticos e não têm estrutura de variância condicional do tipo ARCH. A correção da estrutura de mudanças na variância incondicional pela padronização dos resíduos mostra que, para a série de log-retornos do câmbio, grande parte da estrutura existente é gerada pelas mudanças na variância incondicional. Os resíduos padronizados ainda rejeitam a normalidade, mas na Figura 2.1 vemos, pela densidade empírica e pelo gráfico Q-Q, que os desvios em relação a normalidade são aparentemente pouco significantes.

2.5 Modelo Estimado

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

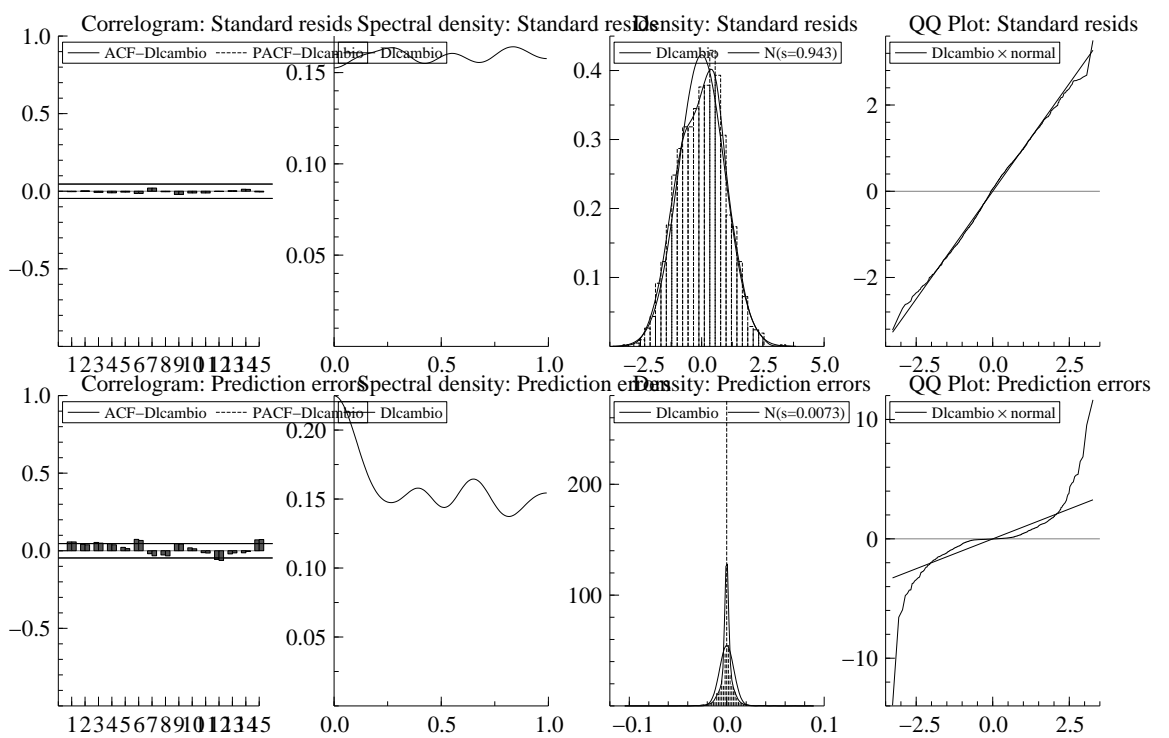


Figura 2.1: Análise dos Resíduos

Tabela 2.5: 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]

Erro de Predição	dist.	Estat.	<i>p-value</i>
portmanteau(31)	Chi(26)	136.0621	[0.0000] **
normalidade	Chi(2)	9054.0432	[0.0000] **
assint.norm.	Chi(2)	13043.066	[0.0000] **
heterocedasticidade	Chi(12)	326.3924	[0.0000] **
hetero-X	Chi(27)	512.9613	[0.0000] **
hetero Quadrado	Chi(12)	326.3890	[0.0000] **
hetero-X Quadrado	Chi(27)	512.9570	[0.0000] **
ARCH(5)	Chi(5)	148.08437	[0.0000] **

$$\begin{cases}
y_{1t} = c_{1t} + \phi_{11}y_{1t-1} + \phi_{12}y_{1t-2} + \phi_{13}y_{1t-2} + \phi_{14}y_{1t-4} + \phi_{15}y_{1t-5} + \beta_1t + \varepsilon_{1t} & \varepsilon_{1t} \sim N(0, \sigma_1^2) \\
y_{2t} = c_{2t} + \phi_{21}y_{2t-1} + \phi_{22}y_{2t-2} + \phi_{23}y_{2t-2} + \phi_{24}y_{2t-4} + \phi_{25}y_{2t-5} + \beta_2t + \varepsilon_{2t} & \varepsilon_{2t} \sim N(0, \sigma_2^2) \\
y_{3t} = c_{3t} + \phi_{31}y_{3t-1} + \phi_{32}y_{3t-2} + \phi_{33}y_{3t-2} + \phi_{34}y_{3t-4} + \phi_{35}y_{3t-5} + \beta_3t + \varepsilon_{3t} & \varepsilon_{3t} \sim N(0, \sigma_3^2) \\
y_{4t} = c_{4t} + \phi_{41}y_{4t-1} + \phi_{42}y_{4t-2} + \phi_{43}y_{4t-2} + \phi_{44}y_{4t-4} + \phi_{45}y_{4t-5} + \beta_4t + \varepsilon_{4t} & \varepsilon_{4t} \sim N(0, \sigma_4^2) \\
y_{5t} = c_{5t} + \phi_{51}y_{5t-1} + \phi_{52}y_{5t-2} + \phi_{53}y_{5t-2} + \phi_{54}y_{5t-4} + \phi_{55}y_{5t-5} + \beta_5t + \varepsilon_{5t} & \varepsilon_{5t} \sim N(0, \sigma_5^2) \\
y_{6t} = c_{6t} + \phi_{61}y_{6t-1} + \phi_{62}y_{6t-2} + \phi_{63}y_{6t-2} + \phi_{64}y_{6t-4} + \phi_{65}y_{6t-5} + \beta_6t + \varepsilon_{6t} & \varepsilon_{6t} \sim N(0, \sigma_6^2) \\
y_{7t} = c_{7t} + \phi_{71}y_{7t-1} + \phi_{72}y_{7t-2} + \phi_{73}y_{7t-2} + \phi_{74}y_{7t-4} + \phi_{75}y_{7t-5} + \beta_7t + \varepsilon_{7t} & \varepsilon_{7t} \sim N(0, \sigma_7^2)
\end{cases} \quad (2.5)$$

O modelo foi estimado usando o pacote MSVAR de KROLZIG (1997) para a linguagem Ox. O modelo obteve convergência forte após 123 interações, e o método de estimação através de máxima verossimilhança usando o algoritmo EM (*Expectation-Maximization*).

Os parâmetros estimados estão na Tabela 2.6, 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 inclusão da tendência também teve um papel fundamental na estimação do modelo, já que sem inclusão da tendência não era possível calcular os desvios padrões, o que ocorria pelo fato da matriz de informação tornar-se singular, um sintoma de possível não identificação do modelo na ausência da tendência.

A matriz de transição estimada se encontra na Tabela 2.7.

A Figura 2.2 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 pode ser identificado, conforme a Tabela 2.6, com os períodos de crise cambial, combinando elevada variância na taxa de câmbio com tendências a forte desvalorização.

A Tabela 2.9 mostra o número de observações pertencentes a cada regime, e com isso associa a cada um dos regimes uma probabilidade incondicional. Os regimes 3,

Tabela 2.6: 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 2.7: 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 $p[i|j]=\Pr\{s(t+1)=j|s(t)=i\}$

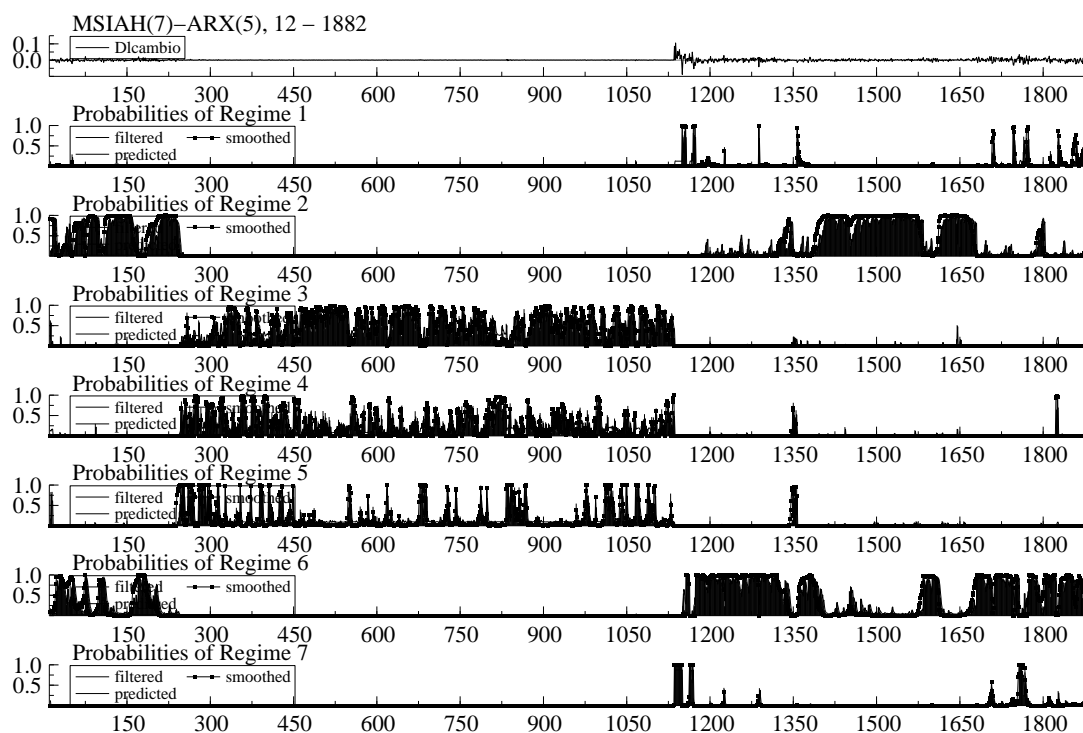


Figura 2.2: Probabilidades Estimadas

Tabela 2.9: 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

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.

Esta conclusão é consistente com os valores médios de duração desses regimes, já que a duração média dos regimes 4 e 5 é de 3,5 dias em média, enquanto que a duração média do regime 3 é de 7,8 dias. O fato de as durações destes regimes que vigoram dentro da banda cambial serem bem curtas mostra o fato de que eram necessárias freqüentes intervenções no mercado de câmbio para manter o valor dentro dos intervalos fixados pelo Banco Central.

A duração média de cada regime também é apresentada na mesma Tabela 2.9. Vemos que o regime mais persistente é o regime 2, com duração média de 38.08 dias, seguido pelo regime 6 com uma duração média de 22.63 dias. Estes dois regimes são caracterizados pelas mais baixas correlações entre os regimes estimados. Os demais regimes são bem menos persistentes, ocorrendo trocas freqüentes.

Uma característica interessante é que o regime 7, associado aos momentos de crise no mercado de câmbio, mesmo tendo a menor probabilidade incondicional entre todos os regimes, possui a terceira maior persistência média, durando em média 8 dias. Isso é devido a matriz de transição associar uma probabilidade de uma vez estando no regime 7, a probabilidade de permanecer neste mesmo regime 7 é de 87%. Também é relevante notar que o regime que tem maior probabilidade de preceder o regime 7 é o regime 1, ou seja aqueles associado a períodos de desvalorização da taxa de câmbio. Esta característica é consistente com um comportamento de existência de valores extremos e agrupamentos de alta volatilidade na série, o que é freqüentemente modelado através de modelos do tipo GARCH.

Desta forma, utilizando apenas a informação contida na amostra, o modelo de mudança markoviana nos parâmetros auto-regressivos, intercepto e variância consegue inferir estatisticamente uma série de informações economicamente relevantes.

Os autovalores que descrevem o comportamento dentro de cada regime estão colocados na Tabela 2.10. Como todos os autovalores calculados são em valor absoluto menores que um, o processo é sempre estacionário. Como a soma de todos os autovalores complexos para cada regime é sempre igual a zero, cada processo tem a característica de apresentar um decaimento exponencial não cíclico, caracterizado pelos autovalores reais existentes. Os valores baixos dos parâmetros auto-regressivos estimados em todos os regimes também caracterizam que a persistência de choques é muito

Tabela 2.10: Regimes e Autovalores

Regime 1			Regime 2			Regime 3		
real	complexo	módulo	real	complexo	módulo	real	complexo	módulo
0.65000	-0.75086	0.99312	-0.57489	0.00000	0.57489	-0.49429	-0.46305	0.67730
0.65000	0.75086	0.99312	0.030743	0.54203	0.54290	-0.49429	0.46305	0.67730
-0.96752	0.00000	0.96752	0.030743	-0.54203	0.54290	0.27847	0.58137	0.64462
-0.22749	-0.86078	0.89033	0.48777	0.00000	0.48777	0.27847	-0.58137	0.64462
-0.22749	0.86078	0.89033	0.024909	0.00000	0.024909	0.41707	0.00000	0.41707

Regime 4			Regime 5			Regime 6		
real	complexo	módulo	real	complexo	módulo	real	complexo	módulo
0.39886	0.36779	0.54255	0.52725	-0.57722	0.78178	0.57771	0.00000	0.57771
0.39886	-0.36779	0.54255	0.52725	0.57722	0.78178	-0.33604	0.41083	0.53076
-0.26628	-0.47026	0.54042	-0.69199	0.00000	0.69199	-0.33604	-0.41083	0.53076
-0.26628	0.47026	0.54042	-0.28020	-0.63025	0.68973	0.10102	-0.49819	0.50833
-0.28023	0.00000	0.28023	-0.28020	0.63025	0.68973	0.10102	0.49819	0.50833

Regime 7		
real	complexo	módulo
-0.42538	-0.68356	0.80511
-0.42538	0.68356	0.80511
0.64972	0.42643	0.77716
0.64972	-0.42643	0.77716
-0.68367	0.00000	0.68367

Tabela 2.11: Autovalores da Matriz de Transição

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

baixa nessa série, já que em esta série tem decaimento exponencial controlado pelos parâmetros auto-regressivos.

Os autovalores da matriz de transição estão colocados na Tabela 2.11. 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 2.11 representa as probabilidades ergódicas do processo. Esse vetor também indica a probabilidade incondicional de cada regime, e desta forma construímos a Tabela 2.9. O fato de a matriz de transição ser ergódica confirma que nosso regime é estacionário, já que, conforme HAMILTON (1994), demonstra-se que um processo de mudança markoviana com uma matriz de transição ergódica é sempre covariância-estacionário.

Uma discussão fundamental é medir o poder de ajuste do modelo estimado em relação ao modelo linear. As Figuras 2.3 e 2.4 mostram os gráficos dos valores previstos e observados e os resíduos padronizados desse modelo. Graficamente notamos que o poder de ajuste do modelo de mudança markoviana é muito superior ao dos modelos lineares auto-regressivos. O coeficiente de determinação do modelo MSIAH(7)-ARX(5)

é de 0.43462, contra 0.08422 do modelo linear AR(12) que é selecionado pelo critério de informação de BIC. Embora o coeficiente de determinação não seja um instrumento recomendado de análise, já que o número de parâmetros do modelo MSIAH (98 parâmetros) é muito superior ao do modelo AR(12), que contém 14 parâmetros estimados, e assim o coeficiente de determinação do modelo MSIAH nunca poderá ser inferior ao do modelo linear. Porém, isto serve como ilustração do poder superior de ajuste do modelo de mudança markoviana, confirmado de forma robusta pelos critérios de informação.

Uma forma de corrigir o coeficiente de determinação para levar em conta o número de parâmetros estimado é a correção para graus de liberdade, como proposto por KROLZIG (1997):

$$\overline{R^2} = 1 - \frac{T - 1}{T - M(M - 1 + K + K(K + 1)/2) - Kp - 1} (1 - R^2) \quad (2.6)$$

onde T é o número de observações, M, o número de regimes e Kp é o número de regressores em cada equação. O $\overline{R^2}$ de nosso modelo é igual a 0.3011, o que é bastante superior ao valor de 0.08422 do modelo linear.

O uso de critérios de informação, mesmo que os modelos não sejam diretamente comparáveis (os modelos não podem ser considerados diretamente aninhados (*nested*) pela presença dos parâmetros de incômodo, i.e as probabilidades de transição), mostra que o ajuste do modelo MSIAH é muito superior ao do modelo linear, mesmo penalizando o elevado número de parâmetros existente nesse modelo, confirmando a conclusão encontrada pelo uso do $\overline{R^2}$.

A Figura 2.5 mostra a densidade estimada pela parte auto-regressiva do modelo para cada um dos regimes pelo modelo de mudança markoviana. Cada regime provém de uma distribuição Normal com média μ_s e variância σ_s^2 conforme a representação dada pela equação 2.5 e os valores dados pela Tabela 2.6 . Este gráfico nos permite notar que os regimes cuja médias estão mais próximas de zero são aqueles com a menor variância, enquanto que o regime de maior variância é o regime 7, com também a maior média incondicional.

2.5.1 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 necessários, conforme demonstrado por RYDEN (1995). 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.

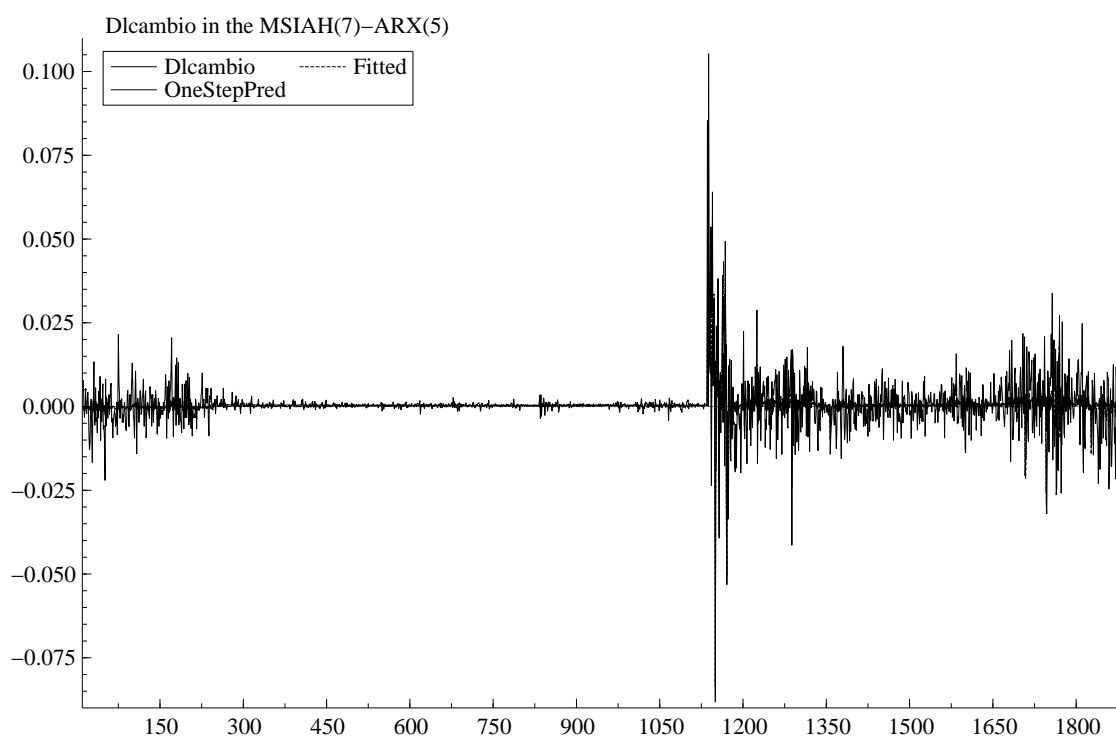


Figura 2.3: Valores Observados e Previstos

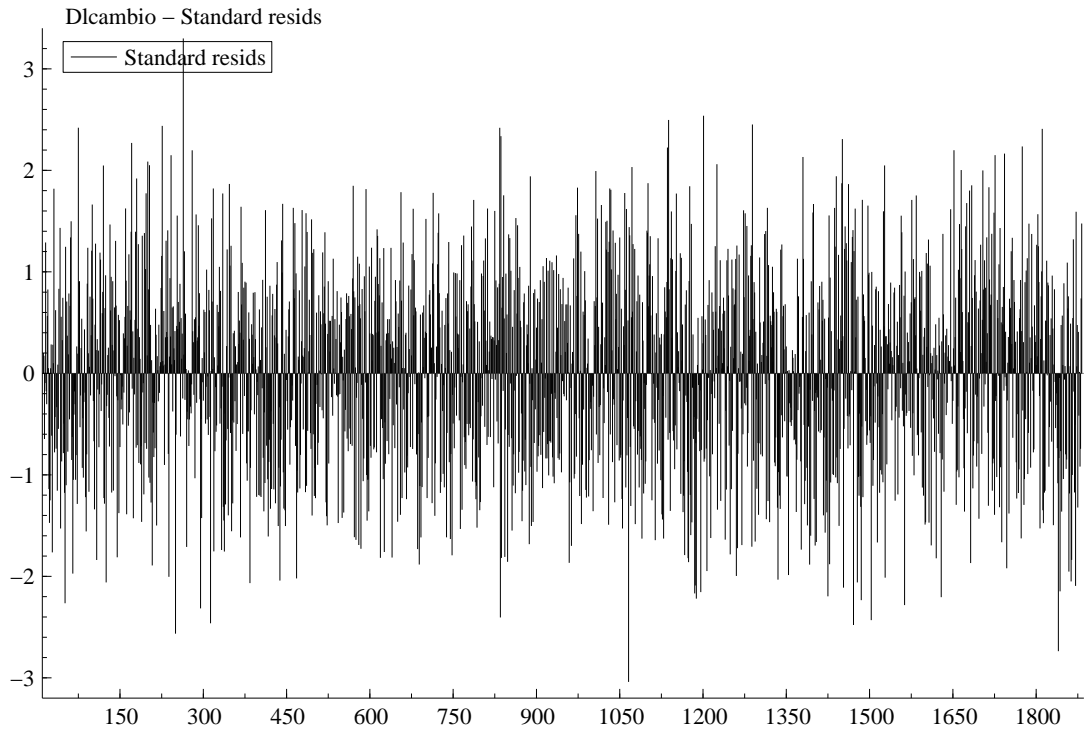


Figura 2.4: Resíduos Padronizados

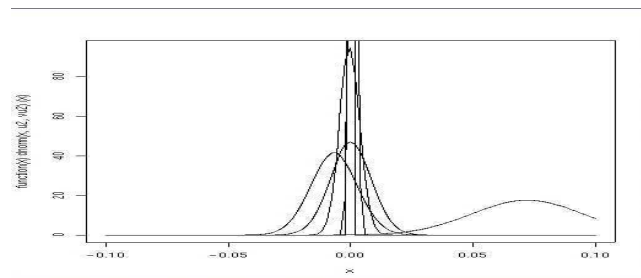


Figura 2.5: Densidades - Regimes

Conforme POSKITT e CHUNG (1996), 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 de necessitamos 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.

2.5.2 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 (1994). 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 se tornava 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 convergencia, 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.

Tabela 2.12: Testes Wald de Exclusão

Regime	Dist.	Valor de teste	<i>p-value</i>
Regime1	$\chi^2(7)$	270.004	[0.0000] **
Regime2	$\chi^2(7)$	6.98191	[0.4308]
Regime3	$\chi^2(7)$	11.3812	0.1228
Regime4	$\chi^2(7)$	162.51	[0.0000] **
Regime5	$\chi^2(7)$	288.391	[0.0000] **
Regime6	$\chi^2(7)$	17.2896	[0.0120]*
Regime7	$\chi^2(7)$	477.018	[0.0000] **

$h_0: \mu_{1s} = \phi_{1s} = \phi_{2s} = \phi_{3s} = \phi_{4s} = \phi_{5s} = \beta_{trend} = 0 | s = Regime$

2.6 Análises de Significância Conjunta - Teste de Eficiência de Mercado

A forma proposta para testar a eficiência de mercado, análoga a da Seção 1.6, é testar se os coeficientes autoregressivos de cada regime são em conjunto estatisticamente diferentes de zero. Para realizar este teste, utilizamos o princípio do teste Wald, em que apenas necessitamos estimar o modelo restrito, já que a estimação de um modelo restrito seria computacionalmente complicada, e a distribuição assintótica do teste Wald é válida para esse teste.

O teste de exclusão realizado correspondeu a testar que todos os parâmetros (exceto a variância) estimados eram iguais a zero, contra a hipótese alternativa de que pelo menos um deles era diferente de zero. O objetivo desse teste é verificar em quais regimes a hipótese de eficiência de mercado é válida, o que testamos notando que na hipótese nula o logaritmo dos preços é um passeio aleatório, e desta forma os log-retornos são não correlacionados.

Como visto na Seção 1.6.2, isto pode ser testado através de um teste conjunto para verificar se todos os parâmetros auto-regressivos são iguais a zero contra a hipótese nula de que pelo menos um deles é diferente de zero. Aqui testamos a hipótese mais forte de que o intercepto e o componente de tendência sejam iguais a zero também, já de que no modelo de mudança markoviana a necessidade de se fazer inferências sobre estes parâmetros que são condicionados ao regime vigente também pode representar possibilidades de arbitragem.

A Tabela 2.12 nos mostra que podemos rejeitar a hipótese nula de que todos os parâmetros são iguais a zero para os regimes 1,4,5,6 e 7. A hipótese nula não é rejeitada para os regimes 2 e 6. Para estes regimes não apenas é válido o modelo de passeio aleatório do tipo 3, mas também log-retornos têm média zero nesses dois regimes. É interessante notar que esta evidência contraria o resultado obtido nos testes de razão de variância que havíamos feito para o período de bandas cambiais, que apontavam a rejeição do modelo de passeio aleatório nesse sub-período.

Os testes Wald de exclusão mostram que existe poder de previsão nos regimes onde se rejeita a hipótese nula. Isto contraria alguns dos resultados do Capítulo 1, e é devido ao fato que não se corrigiu a estrutura existente na variância incondicional da série. Quando trabalhamos com os log-retornos ajustados para risco, ou seja, padronizados pela variância vigente em cada regime, as evidências de poder de previsão se tornam frágeis, como será visto na Seção 2.8.

O valor das estatísticas Wald calculadas apontam que a evidência de desvios em relação a eficiência fraca de mercado são obtidas justamente nos momentos de contínua desvalorização do câmbio, que identificamos respectivamente com os regimes 1 e 7. Principalmente na vigência do regime 7, a aposta na desvalorização da taxa de câmbio parece levar a lucros excessivos, já que o componente de intercepto nesse regime é muito significativo, o que mostra uma tendência linear de desvalorização enquanto o regime é válido contudo, o risco dessa aposta é extremamente alto devido à variância nesse regime ser muito superior a dos demais regimes.

A evidência coletada pelo modelo de mudança markoviana mostra que, enquanto nos regimes 2, 3, 4, 5 e 6 não existem tendências significativas de mudança no valor do câmbio (componentes de intercepto e tendência pouco significativos), o regime 1 pode ser identificado com um período de elevada e relativamente tranqüila (em relação ao regime 7) desvalorização da taxa de câmbio. O regime 7, no entanto, é um período de forte desvalorização da taxa de câmbio, mas acompanhado de uma volatilidade extremamente elevada.

2.7 Variância na Média

A relação encontrada no regime 7 de associação positiva entre desvalorização do câmbio e aumento da volatilidade leva-nos a indagar se existe algum tipo de relação risco-retorno do tipo previsto pelo modelo CAPM (*Capital Asset Pricing Model*) de SHARP (1964) e LINTNER (1965). Essa relação tradicional na literatura de finanças afirma que, se um ativo é mais arriscado (possue uma variância superior aos dos demais ativos), para que ele seja aceito na carteira de investimentos de um investidor sua remuneração média deve ser superior.

A modelagem tradicional desta relação é através do modelo ARCH na média (*ARCH-in-mean*) introduzido por ENGLE, LILIEN e ROBINS (1987). A especificação deste modelo é dada por

$$y_t = X_t \beta + \delta h_t + u_t \quad (2.7)$$

$$u_t = \sqrt{h_t} \cdot v_t \quad (2.8)$$

$$h_t = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_m u_{t-m}^2 \quad (2.9)$$

onde v_t é um processo IID com média zero e variância unitária. O processo de variância condicional existente na série é representado pelo componente h_t . Na equação 2.7 que representa a média do processo, notamos que o componente da variância condicional entra diretamente na especificação da média do processo ponderado pelo parâmetro δ . Como já havíamos detectado a existência de um processo ARCH na série de log-retornos de câmbio, estimamos a forma mais geral deste processo, equivalente a colocar um processo GARCH(1,1) na especificação da média do processo.

A Tabela 2.13 mostra os resultados da estimação deste modelo. Os componentes ARCH(1) e GARCH(1) são estatisticamente significativos. O componente GARCH na média também é significativo para a equação da média, o que confirma que na série de log-retornos de câmbio existe uma relação média-variância.

No capítulo 4 iremos mostrar que modelos do tipo GARCH podem ser inadequados se o processo gerador da série for um processo MSIAH. Desta forma, em vez de

Tabela 2.13: GARCH na Média

Variável Dependente	DLCAMB			
	Coef.	D.Padrão	Estat. z	<i>p-value</i>
equação da média				
GARCH	8.870188	3.018029	2.939067	0.0033
C	0.000189	0.000220	0.855994	0.3920
equação var				
C	2.03E-06	1.98E-07	10.25363	0.0000
ARCH(1)	0.144671	0.016236	8.910448	0.0000
GARCH(1)	0.816476	0.017936	45.52286	0.0000
R-Quadrado	0.029491	Média Var. Dep	0.000484	
R-Quadrado Ajustado	0.027409	D.P. Var. Dep.	0.007630	
D.P. da Regressão	0.007524	AIC	-7.825715	
Soma Quad. Resíduo	0.105590	BIC	-7.810919	
Log- Verossimilhança	7322.044	Estatística F	14.16800	
Durbin-Watson	1.710888	Prob(F)	0.000000	

trabalharmos com a especificação GARCH na média, realizaremos um procedimento em dois passos para estimar um modelo de heterocedasticidade markoviana na média.

O primeiro passo é construir um vetor contendo a variância incondicional estimada pelo modelo de mudança markoviana MSIAH para cada regime vigente em cada período da amostra. A Figura 2.6 mostra o valor da variância condicional estimada pelo modelo de mudança markoviana em relação aos valores observados de log-retornos do câmbio.

A partir desse vetor estimamos um modelo de defasagens distribuídas como forma de ilustrar a relação existente entre os valores de log-retornos do câmbio e a variância determinada pelo modelo de mudança markoviana. A regressão estimada se encontra na Tabela 2.14. Os resultados desse modelo simples mostram que a relação de risco-retorno é significativa, confirmando a evidência do modelo GARCH na média.

É necessário entretanto notar que este procedimento em dois passos é um procedimento ineficiente de estimação, já que o indicado é estimar tanto de forma simultânea tanto a equação contida na Tabela 2.14 como o próprio vetor de variâncias sujeitas à mudança markoviana. Como o processo gerador da variância é claramente dependente do processo gerador da média (não somente por a variância entrar diretamente na especificação da média, mas também pelo fato de que tanto a média quanto a variância do processo são determinadas pela cadeia de transição de Markov), este procedimento é ineficiente, e para verificar a validade deste procedimento seria necessário um estudo específico. Desta forma, este procedimento fica apenas como ilustração da relação entre média e variância.

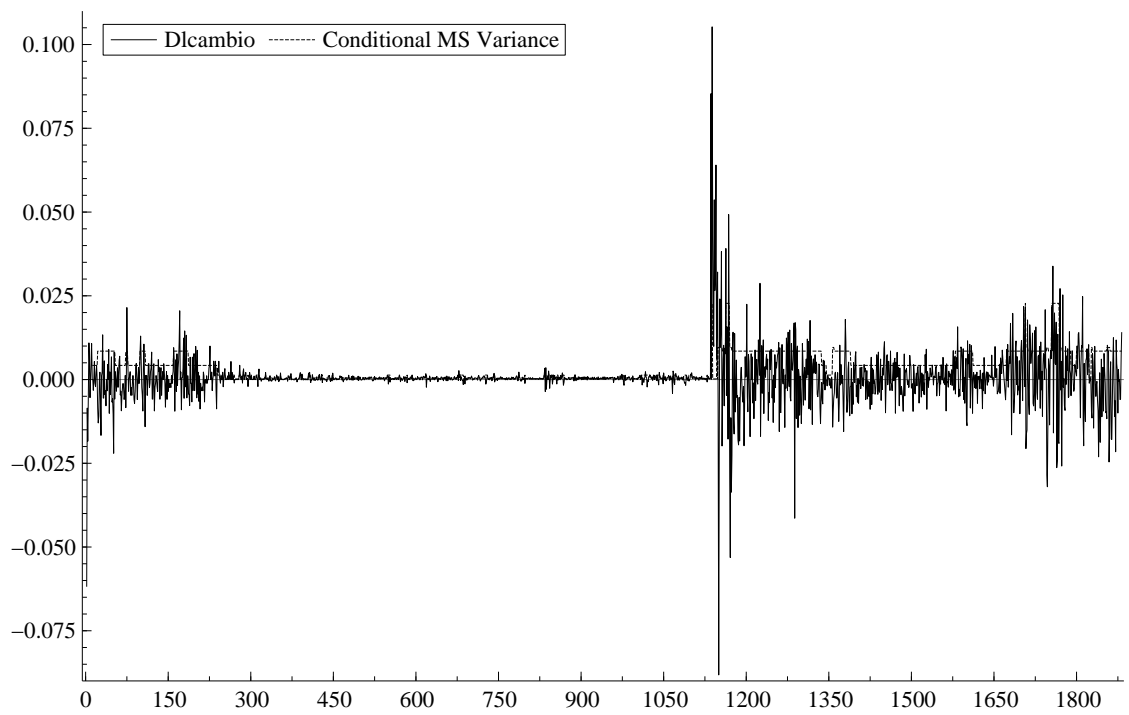


Figura 2.6: Variância Condicional

Tabela 2.14: Relação Média-Variância

	Coef.	D.Padrão	<i>t-value</i>	<i>t-prob</i>
Constant	0.00054761	0.0002487	2.2	0.028
ConditionalMSVariance	0.804718	0.1398	5.76	0.000
ConditionalMSVariance_1	-0.543989	0.1803	-3.02	0.003
ConditionalMSVariance_2	0.412934	0.1498	2.76	0.006
ConditionalMSVariance_5	-0.307771	0.1080	-2.85	0.004
ConditionalMSVariance_8	0.223009	0.1503	1.48	0.138
ConditionalMSVariance_9	-0.307690	0.1653	-1.86	0.063
ConditionalMSVariance_11	0.638238	0.1664	3.84	0.000
ConditionalMSVariance_12	-0.711456	0.1656	-4.30	0.000
ConditionalMSVariance_14	-0.581895	0.1643	-3.54	0.000
ConditionalMSVariance_15	0.295212	0.1820	1.62	0.105
ConditionalMSVariance_16	0.429232	0.1643	2.61	0.009
ConditionalMSVariance_18	-0.273024	0.1140	-2.40	0.017
ConditionalMSVariance_24	-0.615421	0.1454	-4.23	0.000
ConditionalMSVariance_25	0.532603	0.1404	3.79	0.000
sigma	0.00735829	RSS	0.0988675985	
R^2	0.0825883	F(14,1826)	11.74 [0.000]**	
Log-Verossimilhança	6438.13	DW	1.65	
no. de observações	1841	no. de parâmetros	15	
média(Dlcambio)	0.000527005	var(Dlcambio)	5.85377e-005	
Amostra	42 to 1882			

Tabela 2.15: Variance Ratios - Retornos Padronizados

q	VR	psi	<i>p-value</i>
5	1.0686	0.3733	0.373
10	1.1498	0.5393	0.5897
20	0.9579	-0.1071	0.9147
50	0.9504	-0.0821	0.9346
75	0.9537	-0.0640	0.9490
100	1.0179	0.0219	0.9825
150	1.0295	0.0309	0.9754
200	1.1194	0.1144	0.9089
300	1.1877	0.1635	0.8701

2.8 Razão de Variância no Modelo de Mudança Markoviana

Os testes de razão de variância que utilizados no Capítulo 1 tinham valores críticos derivados de distribuições amostrais simuladas via procedimentos de Monte Carlo, na hipótese nula de que os retornos são não serialmente correlacionados. Para analisar o efeito de reconhecida heterocedasticidade existente nos retornos de ativos, POTERBA e SUMMERS (1988) estudam o comportamento da estatística VR sob condições homocedásticas e heterocedásticas, e não encontram diferenças significativas nos resultados. O estudo posterior de KIM, NELSON e STARTZ (1998), no entanto, recorre a uma estratégia diferente para definir o teste VR em condições de heterocedasticidade. Eles utilizam uma estratégia que permite preservar o padrão histórico da volatilidade existente nas séries. A descrição desse procedimento segue a apresentação de KIM e NELSON (1999), que adaptam a estratificação na aleatoriedade de retornos propostas por KIM, NELSON e STARTZ (1998) para o padrão de volatilidade apresentado pelo modelo de mudança markoviana na variância incondicional.

O procedimento de KIM e NELSON (1999) utiliza os retornos padronizados pelas variâncias estimadas para cada observação da amostra pelo modelo de mudança markoviana para calcular as estatísticas VR. Desta forma, os retornos supostamente preservariam a informação contida na heterocedasticidade, e assim o teste teria mais poder por permitir que os padrões específicos da volatilidade apresentados na série em estudo sejam levados em conta na especificação do modelo. Como este procedimento é baseado nos padrões da variância apresentados por cada série, é necessário derivar, via métodos de Monte Carlo, os valores críticos para cada amostra.

O estudo de KIM e NELSON e STARTZ (1998) é baseado na estimação do modelo de mudança markoviana via amostragem de Gibbs, o que permite diretamente obter a estratificação aleatória utilizada. Recorreremos aqui a um procedimento mais simples, apenas testar o valor das Razões de Variância para os retornos ajustados para risco, ou seja padronizados pelas variâncias do modelo de *markov switching*.

As razões de variância calculadas para os retornos padronizados estão na Tabela 2.15. Os valores estimados mostram que não podemos rejeitar que os log-retornos padronizados provenham de um modelo de passeio aleatório, indicando que ao corrigirmos os retornos pelo sua volatilidade, o modelo de passeio aleatório, que era rejeitado nos regimes 1, 4, 5, 6 e 7, não pode ser rejeitado após ajustes para o risco.

Este resultado é consistente com os resultados obtidos pela literatura de *equivalent martingale measure*, que propõem que após ajustes para risco e dividendos, os preços de ativos devem ser um processo *martingale*. Uma apresentação sucinta sobre *equivalent martingale measure* pode ser encontrada em LJUNGQVIST e SARGENT (2000).

Uma conclusão importante desse resultado é mostrar que a violação do modelo de passeio aleatório do tipo III, ou seja a existência de correlação entre os log-retornos de câmbio, é causada por mudanças na estrutura da variância da série. Este resultado é muito importante. No Capítulo 3 mostramos que as medidas de memória longa encontradas nessa série são espúrias e causadas por essas mudanças na estrutura da variância, enquanto que no capítulo 4 mostramos que a existência um processo markoviano na variância incondicional faz com que o uso de modelos do tipo GARCH sejam inadequados como medidas de persistência na variância.

2.9 Conclusões

O modelo estimado nesse capítulo, um modelo de mudança markoviana para todos os parâmetros de uma autoregressão (intercepto, coeficientes auto-regressivos e variância) se mostrou adequado para a modelagem da estrutura existente na série de log-retornos de câmbio.

Uma medida de adequação do modelo é o fato de que por não podermos rejeitar pela estatística BDS que os resíduos do modelo MSIAH(7)-ARX(5) sejam um processo IID, o modelo consegue capturar toda a estrutura existente na média e na variância do processo. O modelo também se mostra capaz de capturar, usando apenas a informação presente na amostra, tanto as quebras estruturais conhecidas com precisão (início e fim do regime de bandas cambiais) como também identifica de forma robusta as quebras endógenas existentes na série de log-retornos de câmbio.

A discussão sobre eficiência fraca de mercado obtém dois resultados interessantes. O primeiro é mostrar que existem na série regimes onde a eficiência de mercado é válida (regimes 2 e 3), enquanto que nos demais regimes rejeitamos a eficiência de mercado através de um teste Wald, assim existindo a possibilidade de arbitragem e lucros excessivos através de transações com a taxa de câmbio. O segundo resultado é confirmar que, após ajustes para risco, a série de log-retornos de câmbio pode ser considerada um processo estocástico consistente com o modelo *martingale* de precificação de ativos apresentado no Capítulo 1.

Embora o modelo de mudança markoviana identifique a possibilidade de usar a informação disponível na amostra como um mecanismo de exploração das ineficiências no uso da informação disponível, algumas considerações são necessárias. Os coeficientes de determinação do modelo (0.43462 não ajustado e 0.3011 ajustado) são muito superiores ao do modelo linear, mas mesmo assim são apenas razoáveis tendo em vista a elaboração de previsões. Um segundo problema é de ordem técnica. A realização de previsões para fora da amostra em um processo MSIAH é um procedimento complicado, já que a equação de previsões nesse modelo é um processo não-linear em função dos parâmetros estimados e dos dados. A terceira restrição na construção de previsões é que nesse modelo temos um elemento adicional de incerteza que são as probabilidades do valor observado em um momento específico pertencer a cada regime. Também é relevante notar que os padrões mais significativos nos dados são encontrados no regime 7, que aponta para forte tendência de desvalorização, mas ao mesmo tempo este é o regime de maior variância, e assim maior risco.

Desta forma a utilização do modelo de mudança markoviana como mecanismo de arbitragem tem de ser vista com cuidado. Embora a eficiência fraca de mercado seja rejeitada pelos dados, fica o questionamento se é possível mesmo explorar de forma efetiva as ineficiências capturadas pelo modelo. Um conceito mais amplo de eficiência de mercado é necessário, que leve em conta a efetividade de estratégias de aproveitamento de ineficiências de mercado. Uma literatura que aborda este problema é analisar se o uso de regras de Filtro e análise técnica levam a lucros relevantes. Uma extensão da análise apresentada nesse capítulo seria verificar se são significativos os lucros obtidos pelas previsões do modelo MSIAH levando em conta custos de transação, o que fica para um trabalho posterior.

Notas Adicionais

Os modelos de mudança markoviana foram estimados usando a classe MSVAR de Hans-Martin Krolzig para a linguagem Ox de Jurgen Doornik. et all. As estatísticas BDS foram calculadas pelo Eviews versão 4.0 e as estatísticas de razão de variância consistentes sob heterocedasticidade foram calculadas usando o módulo VAR-RATIO para Matlab de Ludwig Kansler.

Tabela 2.16: Regimes e Observações

Regime1	Regime 3	Regime 4	Regime 5	Regime 6
1150-1156 [0.98]	257-259 [0.53]	247-248 [0.67]	239-246 [0.87]	22-53 [0.85]
1170-1173 [0.98]	306-306 [0.69]	254-256 [0.83]	249-253 [0.94]	73-76 [0.81]
1288-1288 [0.98]	325-326 [0.62]	262-263 [0.50]	260-261 [0.98]	99-108 [0.84]
1357-1359 [0.73]	329-339 [0.83]	271-278 [0.93]	264-270 [0.87]	160-186[0.86]
1708-1711 [0.73]	342-352 [0.77]	292-294 [0.79]	279-291 [0.91]	1157-1162 [0.94]
1746-1748 [0.90]	364-371 [0.76]	296-297 [0.72]	295-295 [0.99]	1174-1287 [0.96]
1769-1773 [0.84]	382-385 [0.73]	302-305 [0.71]	298-301 [0.76]	1289-1336 [0.86]
1827-1828 [0.80]	409-423 [0.63]	307-309 [0.54]	310-315 [0.76]	1360-1389 [0.87]
1855-1860 [0.61]	437-447 [0.70]	316-324 [0.79]	327-327 [0.98]	1580-1611 [0.89]
	454-455 [0.65]	328-328 [0.52]	354-354 [0.99]	1673-1706 [0.90]
Regime 2	462-547 [0.83]	340-341 [0.52]	372-373 [0.83]	1712-1745 [0.96]
12-21 [0.82]	562-573 [0.76]	353-353 [0.39]	389-394 [0.85]	1749-1754 [0.84]
54-72 [0.75]	577-583 [0.83]	374-381 [0.88]	405-408 [0.86]	1774-1789 [0.91]
77-98 [0.89]	589-595 [0.72]	386-388 [0.57]	424-425 [0.78]	1800-1822 [0.91]
109-159 [0.96]	602-616 [0.87]	395-404 [0.91]	429-430 [0.96]	1829-1854 [0.87]
187-238 [0.90]	624-641 [0.80]	426-428 [0.69]	448-449 [0.78]	1861-1882 [0.83]
1337-1345 [0.74]	645-675 [0.87]	431-436 [0.71]	548-553 [0.81]	
1390-1579 [0.95]	694-703 [0.87]	450-453 [0.79]	584-584 [0.72]	Regime 7
1612-1672 [0.95]	707-725 [0.77]	456-461 [0.77]	617-619 [0.85]	1136-1149 [1.00]
1790-1799 [0.58]	734-741 [0.78]	554-561 [0.83]	676-686 [0.84]	1163-1169 [0.99]
	748-748 [0.58]	574-576 [0.52]	689-689 [0.98]	1707-1707 [0.58]
	751-751 [0.49]	585-588 [0.73]	726-730 [0.67]	1755-1768 [0.91]
	753-757 [0.77]	596-601 [0.61]	742-743 [0.90]	
	773-786 [0.72]	620-623 [0.86]	787-787 [0.91]	
	788-795 [0.73]	642-644 [0.63]	798-798 [0.94]	
	805-805 [0.54]	687-688 [0.65]	833-838 [0.89]	
	849-855 [0.65]	690-693 [0.71]	841-848 [0.83]	
	860-864 [0.59]	704-706 [0.64]	856-858 [0.66]	
	877-914 [0.83]	731-733 [0.59]	865-869 [0.89]	
	920-924 [0.62]	744-747 [0.60]	974-978 [0.78]	
	931-932 [0.61]	749-750 [0.56]	994-994 [0.70]	
	939-949 [0.71]	752-752 [0.50]	1010-1025 [0.85]	
	951-961 [0.76]	758-772 [0.68]	1037-1043 [0.77]	
	963-963 [0.47]	796-797 [0.50]	1049-1050 [0.99]	
	966-973 [0.81]	799-804 [0.76]	1066-1072 [0.94]	
	980-993 [0.78]	806-832 [0.84]	1087-1091 [0.83]	
	1004-1009 [0.86]	839-840 [0.66]	1096-1101 [0.67]	
	1026-1036 [0.86]	859-859 [0.44]	1129-1129 [0.51]	
	1044-1044 [0.50]	870-876 [0.70]	1346-1348 [0.80]	
	1053-1063 [0.88]	915-919 [0.59]	1351-1356 [0.83]	
	1073-1086 [0.81]	925-929 [0.57]		
	1092-1095 [0.57]	933-938 [0.69]		
	1103-1119 [0.80]	950-950 [0.52]		
	1127-1128 [0.49]	962-962 [0.49]		
	1130-1131 [0.64]	964-965 [0.53]		
		979-979 [0.41]		
		995-1003 [0.82]		
		1045-1048 [0.53]		
		1051-1052 [0.56]		
		1064-1065 [0.52]		
		1102-1102 [0.61]		
		1120-1126 [0.71]		
		1132-1135 [0.82]		
		1349-1350 [0.66]		
		1823-1826 [0.95]		

Capítulo 3

Modelos de Memória Longa

3.1 Introdução

A classe de modelos conhecida como memória longa tem recentemente recebido grande atenção na modelagem de séries temporais. A especificação mais utilizada destas equações é a de modelos auto-regressivos com ordem de integração fracionária, conhecidos como ARFIMA(p,d,q).

O modelo ARFIMA, introduzido por GRANGER e JOYEUX (1980) e HOSKING (1981), é a generalização do conceito de ordem de integração utilizado nos modelos ARIMA, permitindo que a ordem de integração seja um número fracionário. Podemos representar esse processo como:

$$(1 - L)^d y_t = e_t \quad (3.1)$$

onde e_t é um processo de ruído branco. Ao contrário dos modelos em que a função de autocorrelação decai em uma taxa exponencial, nos modelos com ordem de integração fracionária a função de auto-correlação decai a uma taxa polinomial, o que corresponde à idéia de memória longa capturada nesses modelos. Visualizando o correlograma da série de log-retornos do câmbio (Figura 3.1) notamos essa propriedade.

3.2 Propriedades do Modelo ARFIMA

As propriedades do modelo ARFIMA têm sido bastante estudadas na literatura e as mais importantes serão colocadas a seguir. Quando a ordem d de integração da série é menor que 0.5, a série y_t é covariância-estacionária e possui uma representação de média móvel infinita:

$$y_t = e_t + \sum_{i=1}^{\infty} \psi_i e_{t-i} \quad (3.2)$$

onde

$$\psi_i = \frac{d(1+d)\dots(i-1-d)}{i!} = \frac{(i+d-1)!}{i!(d-1)!} \quad (3.3)$$

Quando d é maior do que -0.5 y_t é invertível e possui uma representação auto-regressiva infinita

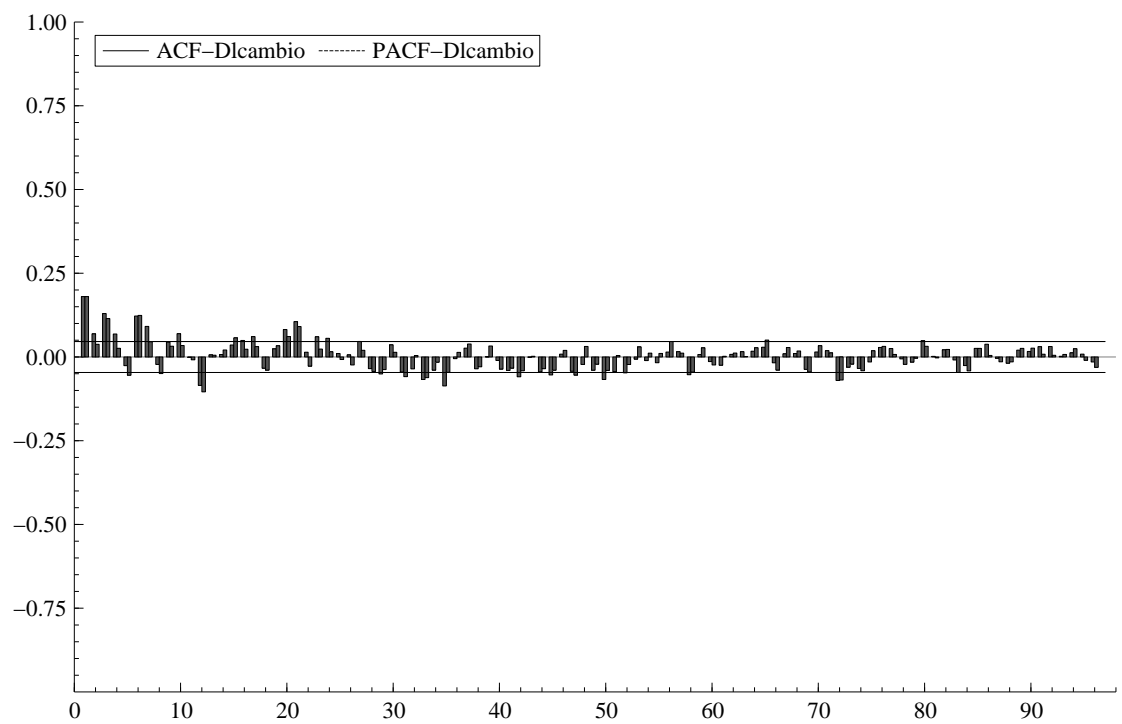


Figura 3.1: Correlograma dos Log-Retornos do Câmbio

$$y_t = \sum_{i=1}^{\infty} \pi_i y_{t-i} + e_t \quad (3.4)$$

com

$$\pi_i = \frac{d(1+d)\dots(i-1-d)}{i!} = \frac{(i+d-1)!}{i!(d-1)!} \quad (3.5)$$

Para $-0.5 < d < 0.5$ a ACF de y_t é dado por

$$\rho_i = \frac{d(1+d)\dots(i-1+d)}{(1-d)(2-d)\dots(k-d)} \quad (3.6)$$

nesse mesmo caso a PACF é dada por

$$\phi_{i,i} = \frac{d}{(i-d)}$$

Se a série fracionalmente diferenciada $(1-L)^d y_t$ segue um processo ARMA(p,q) dizemos que esta série é um processo ARFIMA (p,d,q). Para encontrar a representação de $(1-L)^d$ pode se usar uma expansão binomial para potências fracionárias:

$$(1-L)^d = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} L^k \text{ onde } \binom{d}{k} = \frac{d(d-1)\dots(d-k+1)}{k!} \quad (3.7)$$

A característica mais importante é que, embora a ACF de uma série ARFIMA não seja de magnitude grande, ela decai lentamente, o que corresponde à noção de memória longa, e as auto-correlações são significativas mesmo em intervalos muito longos, como fica claro nos gráficos de ACF e PACF da série de log-retornos do câmbio. Essa característica significa que choques na série de câmbio teriam uma dissipação extremamente lenta, ou seja, elevada persistência.

3.3 Estimação

Dada a evidência gráfica de memória longa mostrada pelos gráficos da função de autocorrelação e da função de autocorrelação parcial, em conjunto com medidas de memória longa como a estatística R/S de Lo, estatística KPSS e a estimação não-paramétrica da ordem de integração de Robinson (Tabela 3.1), estimamos uma série de especificações de um modelo ARFIMA (p,d,q) usando o método de Whittle para a série de log-retornos do câmbio. Como ordens superiores a 3 para os parâmetros AR e MA não eram significativas nesse modelo, testamos as combinações possíveis com parâmetros AR e MA inferiores a 4, conforme Tabela 3.2.

Testando todas as combinações possíveis para a parte ARMA(p,q) com p e q máximos de 2, selecionamos pelo critério de informação de Akaike o modelo ARFIMA(2,d,0) colocado na Tabela 3.3. O modelo foi estimado usando o método de máxima verossimilhança total de Whittle, com o módulo Arfima no programa PcGive.

Todos os parâmetros estimados são significativos, incluindo o parâmetro correspondente à ordem de integração fracionária d de 0.201205. Este resultado seria consistente com as evidências mostradas pelo gráfico da autocorrelação e as medidas de persistência da Tabela 3.1. O valor d de 0.201205 estimado define um processo covariância-estacionário, conforme as propriedades vistas na Seção 3.2. Os valores previstos pelo

Tabela 3.1: Medidas de Persistência

Medidas de Persistência	amostra 1-1181
RS	1.19579
KPSS	0.173743
d de Robinson	0.112466

Tabela 3.2: ARFIMA Amostra 12-1881

modelo	d	d. padrão(d)	<i>t-value</i> (d)	<i>t-prob</i> (d)	amostra 12-1881 AIC-T
ARFIMA(2,d,2)	0.149541	0.04600	3.25	0.001	-5.66584723
ARFIMA(1,d,2)	0.159742	0.04760	3.36	0.001	-5.66654016
ARFIMA(0,d,2)	0.18872	0.04231	4.32	0.000	-6.917660
ARFIMA(2,d,1)	0.191031	0.03804	5.02	0.000	-6.9181742
ARFIMA(2,d,0)	0.201205	0.04139	4.86	0.000	-6.92795259
ARFIMA(1,d,1)	0.131943	0.02338	5.64	0.000	-6.91704088
ARFIMA(0,d,1)	0.132439	0.0311	4.26	0.000	-6.91651336
ARFIMA(1,d,0)	0.136198	0.03097	4.40	0.000	-6.91642264
ARFIMA(0,d,0)	0.156432	0.01896	8.25	0.000	-6.91712088

Tabela 3.3: ARFIMA(2,d,0)

ARFIMA(2,d,0)	amostra 12-1881	
parâmetro	valor estimado	desvio-padrão
Intercepto	0.000722	7.49074e-005
ARFIMA d	0.201205	0.0413921
AR(1)	-0.004658249	0.0007196
AR(2)	-0.054229896	0.0009225

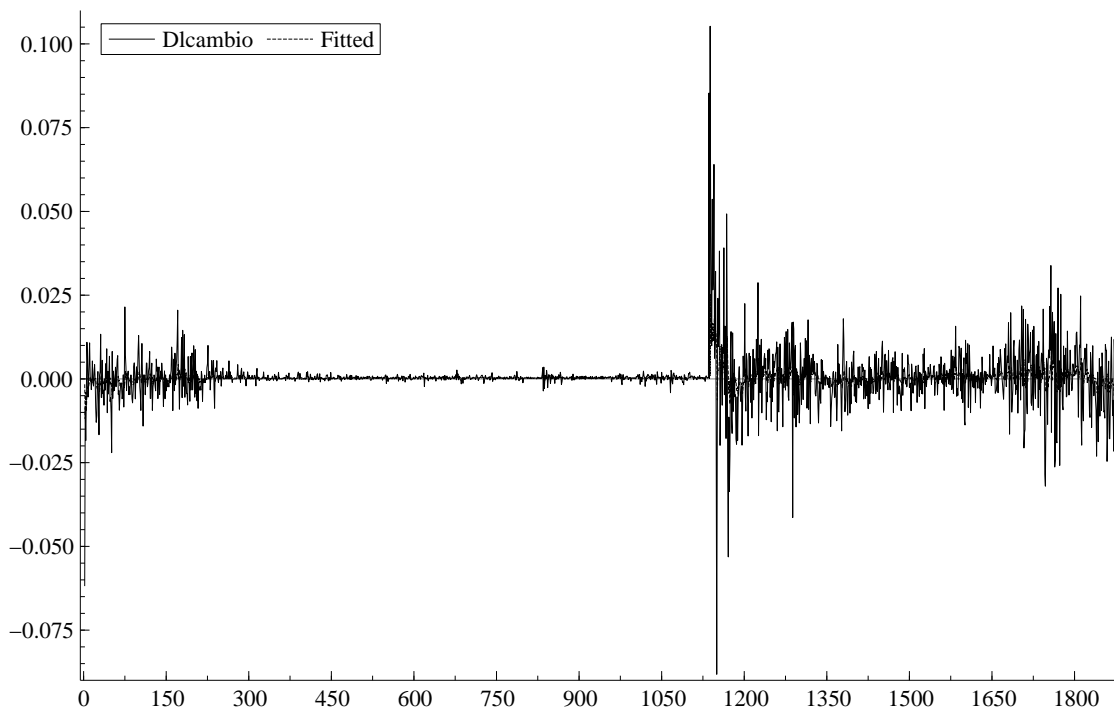


Figura 3.2: Valores Previstos e Observados Modelo ARFIMA(2,d,0)

modelo ARFIMA(2,d,0) para a série de log-retornos de câmbio podem ser visualizados na Figura 3.2.

Mas um parâmetro significativo de integração fracionária contradiz os resultados obtidos no Capítulo 2, onde os parâmetros autoregressivos estimados definiam um processo de memória curta (decaimento exponencial rápido).

Esse resultado no entanto deve ser visto com cautela. Existe uma ampla literatura em econometria que mostra que quebras estruturais e existência de heterocedasticidade podem influir no resultado de procedimentos econométricos. Talvez o exemplo mais representativo dessa literatura sejam os trabalhos sobre a influência de quebras estruturais e heterocedasticidade no resultado dos testes de raízes unitárias (PERRON (1989)). Os testes de raízes unitárias têm como hipótese nula um processo $I(0)$ contra hipótese alternativa de um processo $I(1)$, enquanto que em nosso processo estamos interessados em determinar o valor da ordem de integração I .

Para verificar se o procedimento de estimação de modelos de memória longa é robusto na presença de quebras estruturais iremos realizar uma série de procedimentos para verificar o comportamento do teste nas situações de mudanças na heterocedasticidade e nos parâmetros autoregressivos.

Tabela 3.4: ARFIMA-GARCH(1,1) Amostra Completa

modelo	d	D. Padrão d	AIC-T
ARFIMA(0,d,0)-GARCH(1,1)	0.0176186	(0.0356433)	5632
ARFIMA(1,d,0)-GARCH(1,1)	0.0154518	(0.0591366)	5631.93
ARFIMA(2,d,0)-GARCH(1,1)	-0.00900782	(0.0738184)	5628.18
ARFIMA(3,d,0)-GARCH(1,1)	-0.0160677	(0.0591613)	5643.33
ARFIMA(0,d,1)-GARCH(1,1)	0.883722	(0.0433847)	6077.77
ARFIMA(1,d,1)-GARCH(1,1)	0.999795	(0.000242347)	5678.42
ARFIMA(2,d,1)-GARCH(1,1)	-0.00169456	(0.12095)	5632.05
ARFIMA(3,d,1)-GARCH(1,1)	-0.0156386	(0.0586161)	5642.41
ARFIMA(1,d,2)-GARCH(1,1)	0.999996	(8.88598e-005)	5655.39
ARFIMA(2,d,2)-GARCH(1,1)	não convergiu	*	*
ARFIMA(1,d,3)-GARCH(1,1)	não convergiu	*	*
ARFIMA(3,d,2)-GARCH(1,1)	-0.00624241	(0.0518974)	5649.5
ARFIMA(2,d,3)-GARCH(1,1)	-0.008264	(0.0689502)	5637.77
ARFIMA(3,d,3)- GARCH(1,1)	-0.00748883	(0.0553174)	5648.52

3.4 Modelos ARFIMA-GARCH

Uma característica marcante na série de log-retornos do câmbio é o fato de que a variância dessa série não é constante, exibindo heterocedasticidade condicional. Nos Capítulos 2 e 4 trabalhamos com formas de capturar o comportamento da variância através de modelos paramétricos, condicionando o comportamento da variância da série usando uma cadeia de Markov através do modelo de mudança markoviana na variância (*Markov Switching Heterocedasticity- MSH*) no Capítulo 2, e usando modelos da classe GARCH (Capítulo 4). Uma questão relevante é verificar se a existência de heterocedasticidade condicional pode induzir fenômenos de persistência longa.

Para verificar esse efeito estimaremos inicialmente um modelo que une ordem de integração fracionária com uma estrutura GARCH(1,1) para a variância condicional. A estimação conjunta nos permite testar de forma eficiente se, controlando o efeito da heterocedasticidade condicional a ordem de integração ainda é significativa.

3.4.1 ARFIMA-GARCH - Amostra Completa

Estimamos uma série de modelos ARFIMA mantendo uma estrutura GARCH(1,1) fixa utilizando a amostra completa¹. Os resultados da estimação se encontram na Tabela 3.4. Reportamos o valor do critério AIC multiplicado pelo tamanho da amostra, que é um procedimento comum na modelagem ARFIMA.

O modelo selecionado pelo critério AIC-T é o modelo ARFIMA(2,d,0)-GARCH(1,1). Dois fenômenos que ocorrem na estimação destes modelos são que, para alguns deles, não foi alcançada convergência na maximização da função de verossimilhança, e o fato de que em alguns modelos com componente Média Móvel o valor estimado para a ordem de integração fracionária foi muito significativamente diferente dos demais, e muito próximo de um. Esta diferença acontece pela presença de raízes comuns para os termos autoregressivos e de média móvel no polinômio definidor do processo ARMA

¹ A amostra é fixada no tamanho da amostra disponível na especificação com a maior defasagem

Tabela 3.5: ARFIMA(2,d,0)-GARCH(1,1)

ARFIMA(2,d,0)	GARCH(1,1)	amostra 1-1881
parâmetro	valor estimado	desvio padrão
Intercepto	0.000320622	7.49074e-005
ARFIMA d	-0.00900782	0.0738184
AR(1)	0.078832	0.0779679
AR(2)	0.0512	0.0992569
GARCH Intercepto ^(1/2)	0.00302543	0.00130698
ARCH(1)	0.1447	0.0505813
GARCH(1)	0.82767	0.0653265

estimado. Este problema torna a estimação do parâmetro da ordem de integração não confiável, e por isso desconsideramos os modelos onde ocorriam raízes comuns.

No modelo ARFIMA(2,d,0)-GARCH(1,1) selecionado (Tabela 3.5) notamos que o parâmetro da ordem de integração é significativamente menor do que o estimado no modelo ARFIMA(2,d,0) sem o componente GARCH. Além disso, dado o valor do desvio-padrão deste parâmetro também não podemos rejeitar a hipótese nula de que é estatisticamente diferente de zero.

Quando realizamos a estimação conjunta do coeficiente de integração fracionária com os parâmetros GARCH, controlando a existência de heterocedasticidade condicional, a existência de memória longa não é mais constatada, e assim temos evidências de que o fenômeno de persistência longa pode ser gerado pelas mudanças na variância da série.

No entanto a não significância do parâmetro d pode ser causada por alguma quebra estrutural no processo da média ou então por algum problema de amostra. Para tentar verificar se estes efeitos estão presentes, estimamos modelos ARFIMA-GARCH para os 3 subperíodos em que sabemos com certeza que existe uma quebra estrutural, definida pela introdução e fim do regime de bandas cambiais. As subamostras correspondem ao período do início da série até a introdução do regime de bandas (observações 1-300), o período de vigência do regime de bandas (observações 301-1137) e o período final da livre flutuação da taxa do câmbio (observações 1138-1881).

3.4.2 Modelo ARFIMA-GARCH(1,1) - Primeira subamostra (1-301) (Estabilização Inflacionária)

O período que compreende a primeira parte da amostra é um mercado por valorização da taxa de câmbio como um instrumento de estabilização inflacionária e pesadas intervenções no mercado de câmbio. Os modelos ARFIMA-GARCH(1,1) estimados para essa subamostra estão na Tabela 3.6.

O modelo selecionado pelo critério AIC nessa subamostra foi também um modelo ARFIMA(2,d,0)-GARCH(1,1). Embora a variância dos valores estimados, compreendendo valores de -0.04 a 0.22, em nenhum dos modelos o parâmetro d foi estatisticamente diferente de zero. No modelo selecionado pelo critério AIC-T (Tabela 3.7) estimamos um parâmetro d igual a -0.026, mas com desvio-padrão de 0.118633. Embora o reduzido tamanho da amostra diminua de forma significativa o poder deste teste, continuamos com a evidência de que a memória longa é artificialmente induzida pela

Tabela 3.6: ARFIMA-GARCH(1,1) Amostra 1-299

modelo	d	D. Padrão	AIC -T
ARFIMA(0,d,0)-GARCH(1,1)	0.138724	(0.087866)	892.111
ARFIMA(1,d,0)-GARCH(1,1)	0.127289	(0.149866)	894.843
ARFIMA(2,d,0)-GARCH(1,1)	0.0267857	(0.118633)	890.674
ARFIMA(3,d,0)-GARCH(1,1)	-0.0499879	(0.0687544)	905.201
ARFIMA(4,d,0)-GARCH(1,1)	0.069302	(0.0842878)	908.011
ARFIMA(1,d,1)-GARCH(1,1)	0.103195	(0.202888)	893.916
ARFIMA(2,d,1)-GARCH(1,1)	0.018632	(0.0991393)	894.046
ARFIMA(3,d,1)-GARCH(1,1)	-0.0390525	(0.0640671)	904.84
ARFIMA(0,d,2)-GARCH(1,1)	0.0717403	(0.129686)	891.442
ARFIMA(1,d,2)-GARCH(1,1)	0.179319	(0.344876)	894.259
ARFIMA(2,d,2)-GARCH(1,1)	0.154775	(0.18559)	894.416
ARFIMA(0,d,3)-GARCH(1,1)	0.126345	(0.142234)	891.167
ARFIMA(1,d,3)-GARCH(1,1)	0.227192	(0.130004)	895.608

Tabela 3.7: ARFIMA(2,d,0)-GARCH(1,1) Amostra 1-299

ARFIMA(2,d,0)	GARCH(1,1)	amostra 1-299
parâmetro	valor estimado	desvio padrão
Intercepto	0.000275717	0.000235758
ARFIMA d	-0.0267857	0.118633
AR(1)	0.17185	0.109255
AR(2)	0.16203	0.0883352
GARCH Intercepto ^{1/2}	0.00217053	0.000493419
ARCH(1)	0.40324	0.193478
GARCH(1)	0.61216	0.137393

heterocedasticidade condicional.

3.4.3 Modelo ARFIMA-GARCH(1,1) - Segunda Subamostra (301-1137) (Regime de Bandas Cambiais)

Os modelos da classe ARFIMA-GARCH estimados para o período de vigência do regime de bandas cambiais foram muito instáveis. Os parâmetros para a ordem de integração fracionária flutuaram de -0.25 a .59, o que, no segundo caso, define um processo não-estacionário. No caso do modelo ARFIMA(1,d,1)-GARCH(1,1), a presença de raízes comuns no polinômio ARMA induz um parâmetro próximo de um.

Para o período de bandas cambiais o modelo selecionado pelo critério de AIC-T foi o ARFIMA(1,d,2)-GARCH(1,1) da Tabela 3.9. Nesse modelo o parâmetro estimado está muito próximo de zero, mas pelo valor muito baixo do desvio-padrão associado ele se mostrou significativo (*t-value* de 44.22). É importante notar que este modelo sofre de raízes comuns nos termos AR e MA. Veja que o valor estimado para o AR(1) foi de 1 e o de MA(1) foi de -0.99. Desta forma, o modelo estimado não é numericamente estável. Na maioria dos demais modelos estimados no entanto o parâmetro *d* não era

Tabela 3.8: ARFIMA-GARCH(1,1) Amostra 301-1137

modelo	d	D. Padrão	AIC-T
ARFIMA(0,d,0)-GARCH(1,1)	-0.213025	(0.103438)	4260.16
ARFIMA (0,d,1)-GARCH(1,1)	0.812579	(0.0919414)	4260.22
ARFIMA(1,d,0)-GARCH(1,1)	não convergiu	*	*
ARFIMA(1,d,1)-GARCH(1,1)	0.999958	(3.44731e-005)	4247.16
ARFIMA(0,d,2)-GARCH(1,1)	0.615022	(0.0786477)	4270.27
ARFIMA(1,d,2)-GARCH(1,1)	0.00139078	(3.14503e-005)	4239.23
ARFIMA(2,d,0)-GARCH(1,1)	não convergiu	*	*
ARFIMA(2,d,1)-GARCH(1,1)	-0.0771185	(0.000160507)	4247.87
ARFIMA (2,d,2)-GARCH(1,1)	não convergiu	*	*
ARFIMA (0,d,3)-GARCH(1,1)	0.598528	(0.0988306)	4269.39
ARFIMA(1,d,3)-GARCH(1,1)	-0.391281	(0.111832)	4269.79
ARFIMA(2,d,3)-GARCH(1,1)	0.208043	(0.166738)	4276.92
ARFIMA(3,d,0)-GARCH(1,1)	-1.8292	(0.08706)	4284.58
ARFIMA(3,d,1)-GARCH(1,1)	não convergiu	*	*
ARFIMA(3,d,2)-GARCH(1,1)	0.0750084	(0.167758)	4274.68
ARFIMA(3,d,3) GARCH(1,1)	-0.252808	(0.0201196)	4269.41

significativo.

3.4.4 Modelo ARFIMA-GARCH(1,1) - Terceira Subamostra (301-1137) - (Livre Flutuação da Taxa de Câmbio)

Os resultados obtidos com os modelos ARFIMA-GARCH(1,1) para o período final da amostra são bem semelhantes aos obtidos no início da amostra. Para as estimações onde não ocorre o problema de raízes comuns no polinômio ARMA, os valores estimados foram próximos de zero e estatisticamente não significativos, conforme pode ser visualizado na Tabela 3.10.

Embora o critério AIC-T selecionasse o modelo ARFIMA(1,d,1)-GARCH(1,1), este modelo tinha o mesmo problema de raízes comuns encontrado anteriormente. Se-

Tabela 3.9: ARFIMA(1,d,2)-GARCH(1,1)

ARFIMA(1,d,2)	GARCH(1,1)	amostra 301-1137
parâmetro	valor estimado	desvio padrão
Intercept	0.11038	0.00381037
ARFIMA d	0.00139078	3.14503e-005
AR(1)	1	2.16481e-005
MA(1)	-0.99352	0.0665635
MA(2)	-0.0095972	0.067653
GARCH Intercept ^(1/2)	0.000393278	6.49607e-005
ARCH(1)	0.25077	0.0721675
GARCH(1)	0.65098	0.11242

Tabela 3.10: ARFIMA-GARCH(1,1) Amostra 1138-1881

modelo	d	D. Padrão	AIC-T
ARFIMA(0,d,0)-GARCH(1,1)	0.0158103	(0.0390674)	1860.11
ARFIMA(1,d,0)-GARCH(1,1)	0.0427113	(0.136904)	1878.6
ARFIMA(2,d,0)-GARCH(1,1)	0.0168324	(0.263252)	1841.04
ARFIMA(3,d,0)-GARCH(1,1)	não convergiu	*	*
ARFIMA(0,d,1)-GARCH(1,1)	1.00427	(0.0424434)	1864.85
ARFIMA(1,d,1)-GARCH(1,1)	0.997465	(0.00631198)	1816.51
ARFIMA(2,d,1)-GARCH(1,1)	0.999929	(0.000317033)	1828.96
ARFIMA(3,d,1)-GARCH(1,1)	0.0342652	(0.25382)	1855.18
ARFIMA(0,d,2)-GARCH(1,1)	0.967159	(0.0388119)	1864.51
ARFIMA(1,d,2)-GARCH(1,1)	0.911526	(0.0480465)	1832.59
ARFIMA(2,d,2)-GARCH(1,1)	1.01056	(0.0100994)	1837.42
ARFIMA(0,d,3)-GARCH(1,1)	não convergiu	*	*
ARFIMA(0,d,3)-GARCH(1,1)	não convergiu	*	*

Tabela 3.11: ARFIMA(2,d,0)-GARCH(1,1) Amostra 1137-1881

ARFIMA(2,d,0)	GARCH(1,1)	amostra 301-1137
parâmetro	valor estimado	desvio padrão
Intercept	0.00013915	0.000512075
ARFIMA d	0.0168324	0.263252
AR(1)	0.14366	0.278143
AR(2)	-0.13158	0.097985
GARCH Intercept ^{^(1/2)}	0.00386452	0.000463498
ARCH(1)	0.20392	(0.0643499)
GARCH(1)	0.7724	(0.0529712)

leccionamos então o modelo ARFIMA(2,d,0)-GARCH(1,1) (Tabela 3.11) que era consistente com a representação usada na primeira parte da amostra, e assim servia para comparação direta com os resultados obtidos. Neste modelo o parâmetro d estimado da ordem de integração era também muito próximo de zero e estatisticamente não significativo, o que é consistente com os resultados anteriores.

3.5 Modelo ARFIMA Aplicado aos Resíduos do Modelo MSH

Para verificar se a estrutura da variância determinada pelo modelo de mudança markoviana apresentado no Capítulo 2 também poderia induzir a presença de memória longa, realizamos um procedimento em dois passos para testar essa hipótese.

O primeiro passo foi estimar um modelo de mudança markoviana para a série de log-retornos do câmbio. Para não afetar a estrutura de autocorrelações existentes na série, o único parâmetro sujeito à mudança markoviana foi a variância incondicional. Utilizando o mesmo número de regimes determinados para o modelo do capítulo

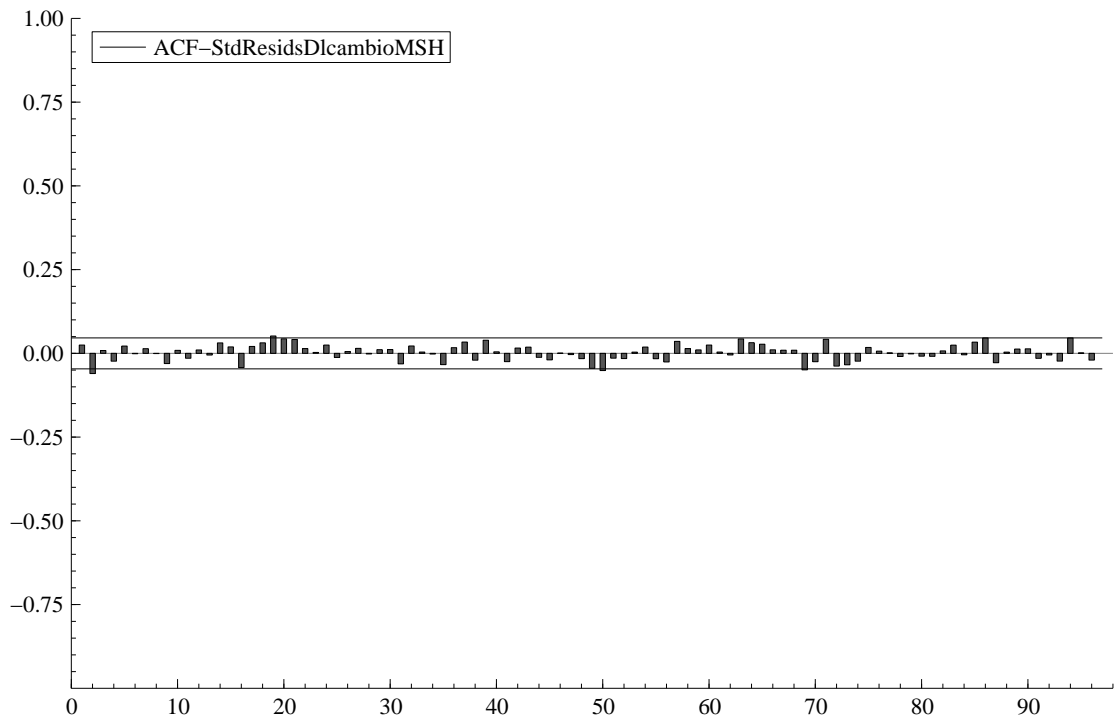


Figura 3.3: Resíduos Modelo MSH

2, estimamos então um modelo MSH (*Markov Switching Heterocedasticity*) com 7 regimes.

O segundo passo do procedimento foi estimar uma seqüência de modelos ARFIMA (p,d,q) para a série de resíduos padronizados do modelo MSH(7)-AR(0) que estimamos. A Figura 3.3 mostra a gráfico da função de autocorrelação para esta série estimada. É interessante notar que, embora o padrão da autocorrelação seja o mesmo disposto pela série original de log-retornos, na padronização utilizando as 7 variâncias estimadas pelo modelo MSH, todas as autocorrelações estão dentro da banda de significância, fazendo com que após a correção pela variância as autocorrelações apresentadas se tornarem estatisticamente não-significantes.

Os modelos estimados para a série de resíduos padronizados do modelo MSH estão dispostos na Tabela 3.12. Como os resíduos possuem média zero por definição, na estimação destes modelos ARFIMA forçamos a restrição de que a média fosse zero não incluindo uma constante na estimação.

Embora pelo gráfico da função de autocorrelação parcial nenhuma das correlações estivesse fora da banda, o modelo escolhido pelo critério de informação AIC-T foi um modelo ARFIMA(1,d,0), contendo um termo autoregressivo de primeira ordem na especificação da média. Os parâmetros estimados se encontram na Tabela 3.13. Nenhum dos parâmetros estimados foi estatisticamente diferente de zero.

O resultado deste procedimento de dois passos também é consistente com a hipótese de que a persistência longa encontrada pelos modelos ARFIMA sem ajuste para a

Tabela 3.12: ARFIMA Resíduos Modelo MSH

modelo	d	D. Padrão	AIC-T
ARFIMA(0,d,0)	0.00654766	(0.0189905)	847.429
ARFIMA(1,d,0)	-0.0184685	(0.0285412)	847.205
ARFIMA(2,d,0)	0.0328249	(0.0326322)	849.906
ARFIMA(3,d,0)	0.0335371	(0.0374943)	848.907
ARFIMA(0,d,1)	0.053764	(0.0399949)	847.425
ARFIMA(1,d,1)	-0.01643	(0.02187)	849.546
ARFIMA(2,d,1)	0.00869183	(0.0311261)	849.242
ARFIMA(3,d,1)	0.022684	(0.0337418)	848.571
ARFIMA(0,d,2)	0.0469924	(0.0429947)	850.372
ARFIMA(1,d,2)	0.0719771	(0.210842)	849.416
ARFIMA(2,d,2)	0.142631	(0.0913865)	850.248
ARFIMA(0,d,3)	0.0484714	(0.0548457)	849.374
ARFIMA(1,d,3)	0.254321	(0.119458)	850.176

Tabela 3.13: ARFIMA(1,d,0) Resíduos Modelo MSH

ARFIMA(1,d,0) amostra 1-1871		
parâmetro	valor estimado	desvio padrão
ARFIMA-d	-0.0184685	0.0285412
AR(1)	0.043508	0.033922

estrutura da variância é um fenômeno espúrio induzido pela existência de heterocedasticidade condicional.

Duas ressalvas quanto a validade do procedimento de estimação em dois passos que realizamos devem ser feitas. A primeira é o fato de que, como não estimamos conjuntamente o parâmetro d e as variâncias utilizadas na padronização dos dados, o procedimento é estatisticamente ineficiente. A segunda é que já que estamos trabalhando com uma variável dependente construída usando um procedimento de estimação o teste t da hipótese de que o parâmetro é diferente de zero utilizando uma distribuição t (ou Normal, pelo número de graus de liberdade existente) só é válido assintoticamente.

3.6 Estudo de Monte Carlo

A possibilidade da memória longa detectada em nossa série de log-retornos do câmbio ser espúria já havia sido estudada na literatura. Em um estudo sobre a possibilidade de memória longa na volatilidade em índices de mercados de ações, BOLLERSLEV and MIKKELSON (1996) sugerem que a memória longa detectada nesses índices poderia ser uma consequência da agregação, já que os retornos individuais não tinham componentes significativos de memória longa na volatilidade.

Já GRANGER (1980) mostra que a soma de processos autoregressivos de primeira ordem AR(1), retirados aleatoriamente de alguma distribuição adequada (nesse estudo FOI utilizada uma distribuição Beta), gera processos com memória longa na média do processo, quando o número de termos na soma cresce.

Com base nestes resultados, elaboramos um estudo de Monte Carlo para identificar se a memória longa existente em nossa série de log-retornos do câmbio poderia ser gerada pela soma de componentes autoregressivos ou por mudanças na estrutura da variância condicional.

3.6.1 Descrição do Experimento de Monte Carlo

O objetivo do nosso procedimento é testar, via simulação, se um processo gerador consistente com o modelo que estimamos no Capítulo 2 poderia induzir a presença de memória longa em uma série, mesmo que na realidade a memória deste processo seja curta.

O mecanismo estatístico gerador de nossos dados (*Data Generating Process* - DGP) escolhido foi um modelo de mudança markoviana para todos os parâmetros (intercepto, parâmetros autoregressivos e variância incondicional) de um processo autoregressivo de ordem 5, conforme a representação 2.5.

Os parâmetros escolhidos para representar o DGP de nossos dados foram os parâmetros estimados pelo modelo MSIAH(7)-AR(5) para os dados de log-retornos de câmbio. Retiramos o componente de tendência que era apenas significativo no regime 7 como forma de tornar os resultados obtidos mais facilmente comparáveis. Utilizando este DGP, testamos quatro especificações para o processo.

3.6.2 Simulação Utilizando as Probabilidades Incondicionais e Variância Unitária

A primeira especificação do processo gerador foi gerar 2000 replicações de um processo com os parâmetros definidos na Tabela 2.6, sendo o regime vigente determinado

Tabela 3.14: DGP Probabilidades Incondicionais, Variância Unitária

T	sign.	>0.05	>0.10	>.25	média	var.
100	99 (4.95%)	145 (7.25%)	96 (4.80%)	25 (1.25%)	0.0185	0.0024
500	203 (10.15%)	298 (14.90%)	136 (6.80%)	5 (0.25%)	0.0238	0.00158
2000	249 (12.45%)	191 (9.55%)	17 (0.35%)	0 (0.0%)	0.0189	0.0004

Tabela 3.15: DGP Probabilidades Incondicionais, Variância markoviana

T	sign.	0.05	0.10	.25	média	var.
100	149 (7.45%)	239 (11.95%)	158 (7.90%)	35 (1.75%)	0.0247	0.0030
500	225 (11.25%)	335 (16.75%)	136 (6.80%)	0 (0.0%)	0.0242	0.0014
2000	610 (30.50%)	152 (7.60%)	28 (1.40%)	0 (0.0%)	0.0115	0.0003

pela probabilidade incondicional dos dados. Nesta primeira especificação do experimento de Monte Carlo os regimes eram retirados de uma distribuição uniforme com valores determinados pela Tabela 2.6, e assim não tínhamos a estrutura de uma cadeia de Markov na determinação do regime vigente. Neste primeiro experimento a variância do componente de erro de cada replicação tinha variância unitária. Este procedimento resumidamente testa se a soma de processos AR(5) retirados de uma distribuição uniforme pode induzir memória longa. Foram realizadas 2000 replicações para amostras de tamanho 100, 500 e 2000.

Os resultados deste experimento estão apresentados na Tabela 3.14. Esta Tabela apresenta os valores estatisticamente diferentes de zero encontrados (coluna sign) e a percentagem em relação ao total de replicações. As colunas > 0.05, > 0.10 e > .025 mostram o total de coeficientes de integração fracionária maiores que o valor indicado na coluna, bem como a proporção em relação ao total. As duas colunas restantes mostram a média e a variância do vetor de valores de integração fracionária estimados para as 2000 replicações do experimento.

A média estimada para cada tamanho de amostra é próxima de zero, mas o número de d estimados significativos é superior a 10% para as amostras de tamanho 500 e 2000, embora seja difícil estabelecer algum padrão de comportamento nesse experimento. Mas comprova que a soma de componentes AR(5) selecionados por uma distribuição uniforme com as probabilidades utilizadas pode induzir a presença de memória longa, para o conjunto de parâmetros utilizado no DGP do experimento.

3.6.3 Simulação Utilizando as Probabilidades Incondicionais e Variâncias do Modelo MSIAH

O segundo experimento utiliza a mesma estrutura do experimento anterior (transição de regimes determinada por uma distribuição uniforme, sem utilizar a matriz de transição de Markov), mas agora o componente de erro de cada autoregressão tem variância diferente. Para o experimento associamos as variâncias estimadas na Tabela 2.6 com cada autoregressão utilizada.

Tabela 3.16: DGP Probabilidades Condicionais (*Markov Switching*) - Variância Unitária

T	sign.	0.05	0.10	.25	média	var.
100	161 (8.05%)	269 (13.45%)	192 (9.60%)	37 (1.85%)	0.0204	0.0036
500	246 (12.30%)	357 (17.85%)	149 (7.45%)	2 (0.01%)	0.0960	0.0178
2000	279 (13.95%)	234 (11.70%)	19 (0.90%)	0 (0.0%)	0.1695	0.0252

Nesse experimento conseguimos comprovar o fato de que mudanças na variância do processo induzem memória longa no comportamento na série. O número de parâmetros de ordem de integração fracionária significativos aumenta significativamente com o aumento do tamanho da amostra, conforme visto na segunda coluna da Tabela 3.15.

3.6.4 Simulação Usando a Estrutura Markoviana nos Regimes - Variância Fixa

Neste experimento o mecanismo gerador dos dados é um processo de mudança markoviana nos parâmetros autoregressivos e no intercepto, e a transição entre os regimes é determinado por uma matriz de transição conforme descrito na Tabela 2.4 . Para definir o DGP utilizado no experimento utilizamos os parâmetros estimados no Capítulo 2 (Tabela 2.6) e também a matriz de transição (Tabela 2.7) estimada por esse mesmo modelo. Para isolar o efeito da estrutura markoviana na variância nesse experimento o componente de erro têm variância unitária. Da mesma forma que nos experimentos anteriores geramos 2000 replicações para amostras de tamanho 100, 500 e 2000.

Os resultados desse experimento (Tabela 3.16) mostram uma proporção relevante de parâmetros da ordem de integração fracionário que são estatisticamente significativos (8.05% na amostra de tamanho 100 e 13.95% na amostra de tamanho 2000). O número de parâmetros d estimados maiores que 0.05 e 0.10 também aumenta para as amostras de tamanho 100 e 500, embora uma parte deles não seja estatisticamente significativa (a proporção de valores na coluna 3 é maior que a proporção de valores na coluna 2). Um fenômeno interessante capturado nesse experimento é o fato de o aumento do tamanho da amostra elevar o número de processos com memória longa. A média do vetor de parâmetros d estimados também aumento de forma muito significativa (médias de 0.02, 0.09 e 0.16 para as amostras de tamanho 100, 500 e 2000).

3.6.5 Simulação Usando a Estrutura Markoviana Total nos Regimes

O DGP do último experimento realizado permite que todos os parâmetros de um processo autoregressivo mudem segundo uma cadeia de transição de Markov, já que nesse experimento permitimos também que a variância incondicional seja sujeita a mudanças. O processo gerador dos dados é um modelo MSIAH(7)-AR(5) com parâmetros definidos pela Tabela 2.6 e a matriz de transição disposta na Tabela 2.7. Geramos novamente 2000 replicações para cada tamanho de amostra, mas nesse experimento utilizamos os tamanhos de amostra 50,100, 250,500,750,1000,1500,2000, 5000 e 10000.

Os resultados do experimento de Monte Carlo para o DGP que permite que as variâncias mudem segunda uma cadeia de Markov são muito interessantes. A Tabela 3.17

Tabela 3.17: DGP Probabilidades Condicionais (*Markov Switching*) - mudanças na Variância

T	sign.	0.05	0.10	0.25	média	var.
50	52(2.60%)	64(3.20%)	50(2.50%)	19(0.90%)	0.0128	0.0015
100	96(4.80%)	136(6.80%)	105(5.25%)	43(2.15%)	0.0204	0.0036
250	224(11.20%)	554(27.70%)	430(21.50%)	199(9.95%)	0.0635	0.0122
500	264(13.20%)	818(40.90%)	663(33.15%)	287(14.35%)	0.0960	0.0178
750	360(18.00%)	936(46.80%)	808(40.40%)	340(17%)	0.1141	0.0203
1000	378(18.90%)	984(49.2%)	840(42.00%)	400(20%)	0.1251	0.0223
1500	400(20%)	1280(64.00%)	1104(55.20%)	509(25.45%)	0.1618	0.0238
2000	449(22.45%)	1288(64.4%)	1097(54.85%)	626(31.30%)	0.1695	0.0252
5000	736(36.80%)	1776(88.80%)	1616(80.80%)	768(38.40%)	0.2205	0.0117
10000	1264(63.20%)	1952(97.6%)	1824(91.20%)	936(46.80%)	0.2454	0.0112

que contém os resultados deste experimento mostra que, com o aumento do tamanho da amostra, o número de processos com memória longa espúria cresce de forma hiperbólica. O número de processos com valores estatisticamente significantes para a ordem de integração fracionária é 2.60% para a amostra de tamanho 50, chega a 18.90% na amostra de tamanho 1000 e alcança 63.20% para a amostra de tamanho 10000.

Quando analisamos a coluna que contém a proporção de parâmetros estimados de ordem de integração fracionária que foram maiores que 0.05, os resultados são ainda mais impressionantes. A proporção de processos com d maior que 0.05 é de 3.20% com a amostra de tamanho 50, alcança 49.2% com uma amostra de tamanho 1000 e para a amostra de tamanho 10000 o número de parâmetros d chega a 97.6%, quase a totalidade dos dados gerados. Os resultados para parâmetros d maiores que 0.10 é parecido (91,20% dos processos tem d maior que 0.10) e quando analisamos os parâmetros d maiores que 0.25 a proporção diminui (46.80% na amostra de tamanho 10000) mas ainda assim é extremamente elevada.

A média dos valores estimados para o parâmetro d nas 2000 replicações de cada processo também cresce com o aumento do tamanho da amostra, o que reflete os resultados do parágrafo anterior. A média para amostra de tamanho 50 é 0.01, alcança 0.12 na amostra de tamanho 1000 e chega a 0.2454 com a amostra de tamanho 10000.

Um fato interessante a ser notado é que embora a média dos valores d estimados aumente com o tamanho da amostra, em nenhum dos experimentos realizados com essa especificação do DGP o valor d foi superior a 0.5, que é a região limite da estacionaridade para processos de memória longa. Deste modo podemos inferir que com o aumento do tamanho da amostra para infinito, a média dos parâmetros d estimados deve convergir para um valor próximo mas inferior a 0.5. Isto é derivado do fato de que o DGP gerador dos dados era sempre um processo estacionário, como visto pelos autovalores de cada autoregressão. Um gráfico mostrando a tendência de convergência dos parâmetros está colocado na Figura 3.4.

O resultado dos experimentos de Monte Carlo aponta que, ao menos para o DGP utilizado, correspondente ao processo estimado para a série de log-retornos de câmbio pelo modelo MSIAH, a probabilidade de que a memória longa estimada na seção 1.3 seja espúria é elevada.

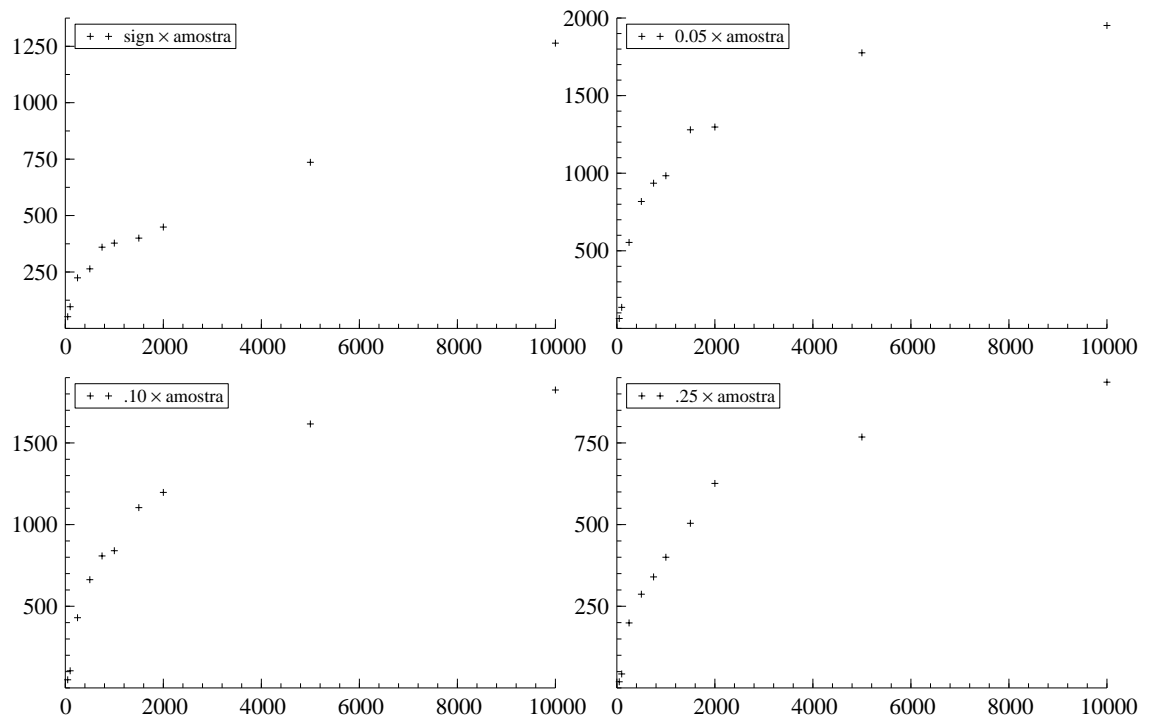


Figura 3.4: Convergência

3.7 Conclusões

Nesse Capítulo estudamos se as evidências de persistência longa (memória longa) encontradas na série de log-retornos do câmbio são robustas. A estimação do modelo mais utilizado para modelar memória longa, o modelo ARFIMA de GRANGER e JOYEUX (1980), encontrou um parâmetro significativo para a ordem de integração fracionária, o que seria consistente com as longas autocorrelações significativas encontradas para a série de log-retornos do câmbio.

Devido a existência de heterocedasticidade condicional na série, realizamos uma série de procedimentos para verificar se os procedimentos de estimação de medidas de persistência longa são robustos na presença de heterocedasticidade condicional.

O primeiro procedimento realizado foi estimar um modelo ARFIMA(p,d,q)-GARCH(1,1) onde modelamos de forma conjunta a ordem de integração fracionária com um modelo GARCH. O resultado deste procedimento mostrou que a evidência de memória longa é frágil quando controlamos a heterocedasticidade existente na amostra.

Para verificar se este resultado se mantém na presença de quebras estruturais, estimamos os modelos ARFIMA(p,d,q)-GARCH(1,1) para as 3 subamostras importantes da série de log-retornos do câmbio. O resultado nas 3 amostras, embora prejudicado pela instabilidade do método de estimação na presença de raízes comuns no polinômio ARMA, confirmou que o parâmetro d estimado na Seção 1.3 não é sustentável quando controlamos a heterocedasticidade condicional usando um componente GARCH.

A etapa seguinte foi tentar controlar a heterocedasticidade através de um modelo MSH (*Markov Switching Heterocedasticity*), realizando um procedimento de dois passos. Primeiro estimando um modelo MSH para a série de log-retornos do câmbio e depois, estimando um modelo ARFIMA para os resíduos padronizados desse modelo, que corrigem a existência de mudanças na variância incondicional da série. Os resultados desse procedimento mostraram que a evidência de memória longa também não era sustentável quando controlávamos a heterocedasticidade existente por um modelo de mudança markoviana.

O procedimento final foi realizar uma série de estudos de Monte Carlo para confirmar que, em processos onde existisse mudança na variância incondicional controlado por uma cadeia de Markov (modelos MSH), existe uma grande probabilidade de existência de memória longa espúria. A probabilidade de um modelo de mudança markoviana induzir memória longa na série cresce de forma hiperbólica com o tamanho da amostra. Com isto confirmamos que, além da conclusão de GRANGER (1980) que a soma de processos AR(1) geram processos com memória longa pode ser generalizada para processos AR(5), o componente que induz a presença de memória longa é a soma de processos autoregressivos com diferentes variâncias incondicionais.

Este estudo nos permite inferir com um razoável grau de certeza que a persistência de choques na série de log-retornos de câmbio é muito baixa, de acordo com os resultados do Capítulo 2. A hipótese de que choques na série de log-retornos do câmbio apresentam efeitos após períodos relativamente longos não é sustentável pelos resultados dos procedimentos realizados nesse estudo.

A conclusão final é a que modelos de memória longa da classe ARFIMA não são robustos na presença de heterocedasticidade condicional, e que os resultados encontrados por esses modelos nesta situação devem ser vistos com cautela. Fica a necessidade de desenvolvimento de correções para a existência de heterocedasticidade nos estimadores de modelos do tipo ARFIMA.

Notas Adicionais

Os modelos ARFIMA foram estimados usando o pacote ARFIMA de Marium Ooms e Jurgen DOORNIK para PcGive. Usamos a classe LMMod para Ox de James DAVIDSON para estimar os modelos ARFIMA-GARCH. Estimamos o modelo MSH (*Markov Switching Heterocedasticity*) usando a classe MSVAR 1.3.0 de Hans-Martin KROLZIG para a linguagem Ox. O estudo de Monte Carlo foi programado usando a linguagem R usando o pacote Fracdiff para estimar os parâmetros do modelo ARFIMA.

Capítulo 4

Persistência na Volatilidade

4.1 Introdução

Uma das características mais marcantes de séries financeiras em geral é a existência de heterocedasticidade: o fato de que a volatilidade dos retornos não é constante no tempo. Conforme vimos no Capítulo 2, temos fortes evidências de que a variância da série de log-retornos do câmbio não é constante. Neste capítulo iremos explorar duas classes de modelos para capturar a volatilidade condicional existente na série de log-retornos do câmbio. A primeira abordagem é a classe de modelos ARCH (*Autoregressive Conditional Heterocedasticity*) de ENGLE (1982) e BOLLERSLEV (1986) (incluindo modelos de memória longa), e a segunda é um estudo mais detalhado das previsões para a volatilidade dadas pelo modelo MSH (*Markov Switching Heterocedasticity*) que utilizamos no Capítulo 2.

A modelagem da variância da série de câmbio é muito importante porque medidas da volatilidade dessa série medem o risco associado a este ativo, sendo fundamentais em situações como o controle de risco de portfólios com moedas estrangeiras como um de seus ativos e a precificação de derivativos de câmbio.

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 (1999)). A relação entre heterocedasticidade condicional e mudança de parâmetros é também explorada nos modelos CHARMA de TSAY (1987) e no modelo RCA de NICHOLLS e QUINN (1982), que usam coeficientes aleatórios para produzir heterocedasticidade condicional.

Como já identificamos através do modelo de mudança markoviana estudado no Capítulo 2 a existência de mudanças abruptas nos parâmetros para a série de log-retornos de câmbio, é natural continuar nosso estudo verificando se os modelos da classe ARCH e o modelo de mudança markoviana na variância (MSH) são adequados para capturar o comportamento da variância condicional presente em nossos dados.

4.2 Modelos ARCH-GARCH

4.2.1 Preliminares Sobre Modelagem ARCH

A modelagem baseada na classe de modelos ARCH e suas derivações teve sua criação através do trabalho de ENGLE (1982) 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. Como exemplos dessa literatura aplicada a estudo da séries de taxas de câmbio temos os trabalhos de ANDERSEN, BOLLERSLEV, DIEBOLD e LABYS, P. (1999) e ENGLE e MUSTAFA (1995).

Os trabalhos de HERENCIA, HOTTA e VALLS PEREIRA (1998), NELSON (1990) e PAGAN e SCHWERT (1990) estudam o tópico da persistência de choques na volatilidade de séries financeiras em perspectiva teórica. Também são importantes os trabalhos comparando a metodologia ARCH a outros modelos de volatilidade condicional, como o trabalho de HERENCIA, HOTTA e VALLS PEREIRA (1998), que compara os modelos ARCH com modelos de volatilidade estocástica.

O modelo ARCH de ENGLE (1982) 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:

$$y_t = X\beta + \varepsilon_t \quad (4.1)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (4.2)$$

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:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 \quad (4.3)$$

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 α_1 seja ≥ 0 . Se α_1 é igual a zero o modelo é condicionalmente homocedástico.

A representação 4.3 permite ver que o modelo ARCH captura agrupamentos de volatilidade devido ao fato de que o choque ε_t é função crescente do choque passado ε_{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 (1982), a kurtose de ε_t sempre supera a kurtose de z_t já que:

$$E[\varepsilon_t^4] = E[z_t^4]E[h_t^2] \geq E[z_t^4]E[\varepsilon_t^2]^2 \quad (4.4)$$

o que é válido pela desigualdade de Jensen. Se z_t possui distribuição normal, a kurtose de ε_t é igual a

$$K_\varepsilon = \frac{E[\varepsilon_t^4]}{E[\varepsilon_t^2]^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2} \quad (4.5)$$

que é finito 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:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 \quad (4.6)$$

onde a variância incondicional σ^2 é definida pela relação:

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

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 média na volatilidade.

4.2.2 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 (1986) sugeriu uma formulação geral, conhecida como GARCH (Generalized ARCH) que é uma generalização do modelo ARCH para incluir valores defasados de h_t em 4.6. O modelo GARCH(1,1) é representado como:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (4.7)$$

As condições de positividade são que w e $\alpha_1 > 0$ e $\beta_1 \geq 0$. A representação 4.7 evita a necessidade de colocarmos muitos lags de ε_t , o que notamos se 4.7 for reescrita como:

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 (w + \alpha_1 \varepsilon_{t-2}^2 + \beta_1 h_{t-2}) \quad (4.8)$$

Continuando esta recursão chegamos a:

$$h_t = \sum_{i=1}^{\infty} \beta^i w + \alpha_1 \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^2 \quad (4.9)$$

A representação 4.9 corresponde a um modelo $ARCH(\infty)$ para esta estrutura de ε_t^2 . Adicionando ε_t^2 em ambos os lados de 4.7 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 ε_t^2 :

$$\varepsilon_t^2 = w + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1} \quad (4.10)$$

onde $v_t = \varepsilon_t^2 - h_t$. O processo definido em 4.10 para ε_t^2 será covariância estacionário se $\alpha_1 + \beta_1 < 1$. Definimos a variância incondicional de ε_t como:

$$\sigma^2 = \frac{w}{1 - \alpha_1 - \beta_1} \quad (4.11)$$

A kurtose de ε_t , novamente assumindo que z_t seja normalmente distribuído, é representada como:

$$K_\varepsilon = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \quad (4.12)$$

que é também sempre maior do que 3. Uma característica importante derivada por BOLLERSLEV (1986) é que as autocorrelações de ε_t são dadas por:

$$\rho_\varepsilon = (\alpha_1 + \beta_1)^{k-1} \rho_1 \quad (4.13)$$

ou seja, as correlações decaem 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:

$$h_t = w + \sum_{i=1}^q \varepsilon_{t-i}^2 + \sum_{i=1}^p h_{t-i} \quad (4.14)$$

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), esta parece ser a forma mais adequada na prática, conforme BOLLERSLEV, CHOU e KRONER (1992).

A generalização dos modelos 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 (1982) 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 (1991) 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.

4.2.3 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 ε_t^2 , o que origina o termo integrado para essa série. Com essa restrição podemos escrever:

$$(1 - L)\varepsilon_t^2 = w - v_t - \beta_1 v_{t-1} \quad (4.15)$$

Como colocam FRANSES e VAN DIJK (2000), a analogia com processos de raízes unitárias deve ser considerada com cuidado. Embora a variância incondicional não exista, conforme mostra a equação 4.11, 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 (1990).

As autocorrelações para ε_t^2 neste modelo não estão definidas, mas segundo DING e GRANGER (1996) podem ser aproximadas por:

$$\rho_k = \frac{1}{3}(1 + 2\alpha_1)(1 + 2\alpha_1)^{-k/2} \quad (4.16)$$

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

Tabela 4.1: ARMA(1,0) - GARCH(1,1)

ARMA(1,0)	GARCH(1,1)		amostra	8-1882	
parâmetro	valor est.	desvio-padrão	dp-robusto	<i>t-value</i>	<i>t-prob</i>
Constante	0.000315297	9.040e-005	9.013e-005	3.50	0.000
Dlcambio_1	0.0579005	0.03289	0.03449	1.68	0.093
Intercepto-Garch	1.50810e-006	0.02281	0.03968	3.59	0.000
ARCH(1)	0.142515	0.02281	0.03968	3.59	0.000
GARCH(1)	0.830830	0.02249	0.03836	21.7	0.000
log-lik	7355.04897		AIC	-7.84005223	
média(h_t)	5.70609e-005		var(h_t)	2.78904e-008	
alpha(1)+beta(1)	0.973345				

4.2.4 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 ε_0^2 na variância condicional h_t , o que pode ser representado como:

$$\lim_{t \rightarrow \infty} E \left(\frac{\partial h_t}{\partial \varepsilon_0^2} \right) \quad (4.17)$$

NELSON (1990) mostra que, assumindo que h_0 é independente da seqüência ε_t , ela mesma uma seqüência também independente, o valor de 4.17 é dado por:

$$\begin{aligned} E \left(\frac{\partial h_t}{\partial \varepsilon_0^2} \right) &= E(h_0) E \left(\frac{\partial \prod_{i=1}^t (\alpha_1 \varepsilon_{t-i}^2 + \beta_1)}{\partial \varepsilon_0^2} \right) \\ &= E(h_0) E \left(\prod_{i=1}^t (\alpha_1 \varepsilon_{t-i}^2 + \beta_1) \right) E \left(\frac{\partial (\alpha_1 \varepsilon_0^2 + \beta_1)}{\partial \varepsilon_0^2} \right) \\ &= E(h_0) (\alpha_1 E(\varepsilon^2) + \beta_1)^{t-1} \alpha_1 \end{aligned} \quad (4.18)$$

Existe persistência se o valor de 4.18 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 α_1 .

Os choques terão efeito permanente na variância da série quando a soma dos parâmetros α_1 e β_1 em um processo GARCH(1,1) for igual a 1, o que corresponde ao modelo IGARCH de ENGLE e BOLLERSLEV (1986). 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.

4.3 Estimação

Iniciamos a estimação de um modelo GARCH com uma especificação AR(6) para a média do processo e uma especificação GARCH(1,1) para a variância incondicional. Este modelo têm os parâmetros autoregressivos de ordem maior que um não significantes, e por isso reduzimos o modelo usando o teste t para o modelo ARMA(1,0)-GARCH(1,1) disposto na Tabela 4.1.

Tabela 4.2: 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 4.3: ARMA(0,0) - GARCH(2,2)

ARMA(0,0)	GARCH(2,2)		amostra		
parâmetro	valor est.	desvio-padrão	dp-robusto	<i>t-value</i>	<i>t-prob</i>
Constante	0.000310972	9.036e-005	8.759e-00	3.55	0.000
Intercepto-Garch	2.00722e-006	1.860e-006	3.647e-006	0.550	0.582
ARCH(1)	0.121130	0.03849	0.06221	1.95	0.052
ARCH(2)	0.0688515	0.1717	0.3074	0.224	0.823
GARCH(1)	0.573146	1.207	2.108	0.272	0.786
GARCH(2)	0.202003	1.001	1.741	0.116	0.908
log-lik			AIC	-7.83828564	
média(h_t)	5.71898e-005		var(h_t)	2.84053e-008	
alpha(1)+beta(1)	0.9751469				

A redução foi válida pelo critério de Akaike. Nesta nova especificação, no entanto, o parâmetro autoregressivo de log-retornos do câmbio não era mais significativo a 5%, e por isso tentamos uma nova redução deixando apenas uma constante na média, conforme a Tabela 4.2.

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. É necessário verificar se a especificação da variância condicional é satisfatória. Testamos então o modelo ARMA(0,0)-GARCH(2,2) colocado na Tabela 4.3.

O critério de informação de Akaike mostra que o modelo mais recomendado é o ARMA(0,0)-GARCH(1,1), e também podemos verificar que os parâmetros ARCH(2) e GARCH(2) não são significativos pelo teste t. É escolhida a especificação da Tabela 4.2.

Os teste para efeitos ARCH remanescentes e o teste *Portmanteau* (Tabela 4.4) 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.

Podemos ver na Tabela 4.5 que os resíduos ao quadrado não são simétricos e que o excesso de kurtose é extremamente elevado, o que causa a rejeição da normalidade. O

Tabela 4.4: Testes de Especificação

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

Tabela 4.5: Resíduos ao Quadrado Estatísticas Descritivas

Resíduos ao Quadrado	Estatísticas Descritivas
Observações	1875
Média	0.015633
Desvio Padrão	0.99938
Simetria	12.078
Excesso de Kurtose	341.39
Mínimo	-4.2422
Máximo	28.296
Normalidade($\chi^2(2)$)	9.1510e+006 [0.0000]**
Normalidade Assintótica($\chi^2(2)$)	8112.7 [0.0000]**

excesso de kurtose muito acima do valor presente nos dados modelados indica que o modelo pode ser inadequado, já que o peso elevado nas caudas da distribuição dos resíduos mostra que os valores extremos têm baixa probabilidade de serem capturados por esse modelo. A Figura 4.1 mostra a densidade estimada dos resíduos em comparação com os valores de uma distribuição normal com mesma média e variância.

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 4.2. Este gráfico mostra que o comportamento da volatilidade da série é compatível com a classificação dos 3 subperíodos que temos utilizado. 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 α_1 (ARCH(1)) e β_1 (GARCH(1)) é igual a 0.972892. Como os parâmetros α_1 e β_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.

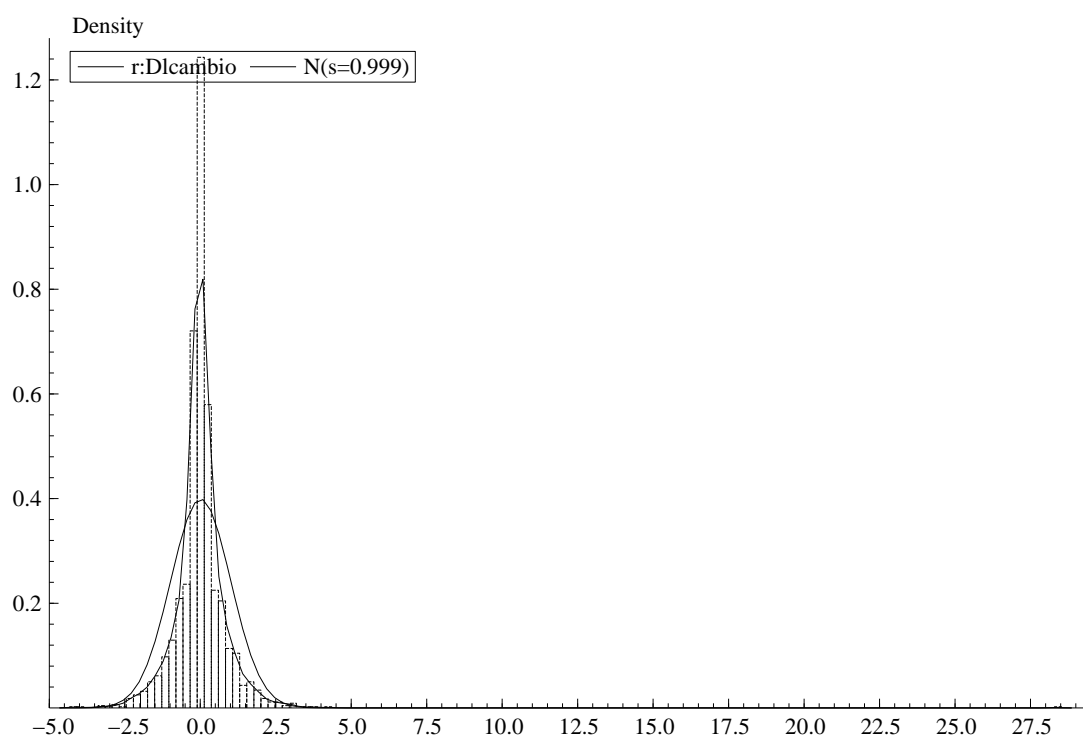


Figura 4.1: Densidade Resíduos

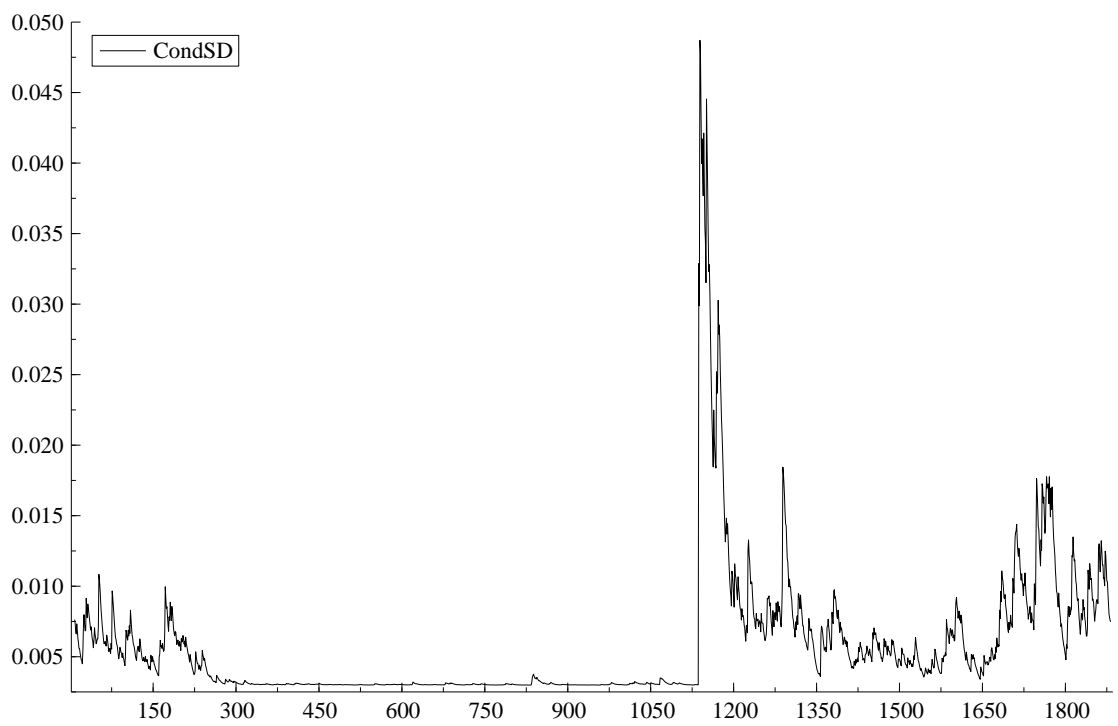


Figura 4.2: Variância Condicional Modelo GARCH(1,1)

4.4 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 (1992) 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 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 is accustomed to having residuals that are better behaved than the original data."

Mas como colocam GALLANT, HSIEH e TAUCHEN (1992), 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 (1973), 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 (1973) 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 (1973), GALLANT, HSIEH e TAUCHEN (1992) 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 (1973) e de GALLANT, HSIEH e TAUCHEN (1992) 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 (1992) 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 (1996):

"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 it 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. LA-MOUREAUX e LASTRAPES (1990) 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 (1994) 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 (1994) 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 (2000) apontam que o uso de *dummies* para o controle de *outliers* e quebras na variância pode induzir a multimodalidade 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 (1998), que derivam testes para presença de *outliers* em modelos GARCH, e em SAKATA e WHITE (1998), discutem métodos robustos de estimação na presença de *outliers*.

A hipótese alternativa à modelagem ARCH que pode ser explorada, apontada no trabalho de GALLANT, HSIEH e TAUCHEN (1992), é modelar a volatilidade dessas séries como proveniente de uma mistura de distribuições, ao invés de um processo único como nos modelos ARCH. Isto permite que tenhamos regimes de variâncias provenientes de distintas distribuições.

Nossa abordagem para a modelagem desse processo de mistura no entanto foi diferente das abordagens de CLARK (1973) e de GALLANT, HSIEH e TAUCHEN (1992) de modelos semi-paramétricos (SNP). O modelagem do câmbio como resultado de inovações originárias de diferentes distribuições realizada no Capítulo 2 foi baseada

Tabela 4.6: Estimação GARCH subamostras

Estimação em sub-amostras	parâmetros	modelo GARCH	valor t	<i>p-value</i>
Amostra 1-301	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			

nos modelos de mudança markoviana. Nos modelos de mudança markoviana, os parâmetros são determinados pela probabilidade de se pertencer a uma série de regimes, e a mudança de um regime para outro seria determinada por uma cadeia de Markov de primeira ordem.

A vantagem do modelo de mudança markoviana para as séries de câmbio é 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. 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 calmaria nos mercados.

A analogia possível é com o modelo de raízes unitárias com quebras estruturais de PERRON (1989). Nestes modelos, a não incorporação das quebras estruturais leva à possível não rejeição da hipótese de raízes unitárias, que impõem uma persistência infinita aos choques. A idéia da modelagem de mudança markoviana é que, ao modelarmos a variância como proveniente de diversos regimes, evitamos aceitar uma hipótese incorreta de persistência excessiva de choques à volatilidade .

4.5 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 definidas nos capítulos anteriores.

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 4.6), 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

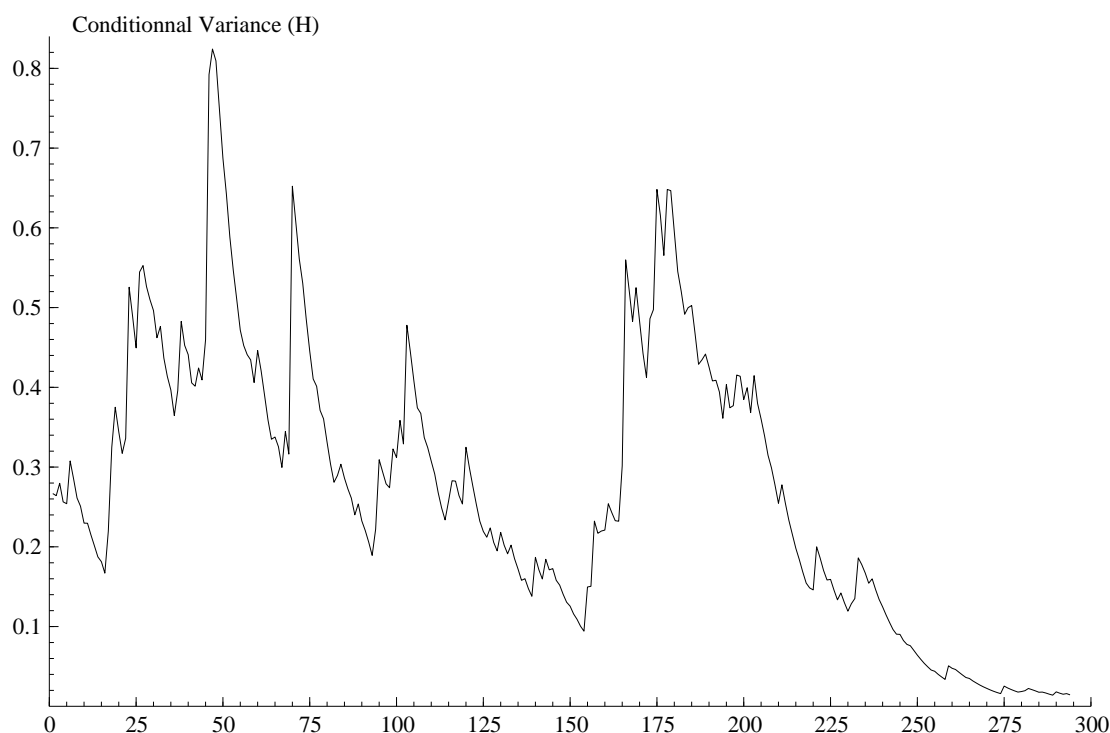


Figura 4.3: Garch(1,1) Pré Banda

valor 0.93. As variâncias condicionais estimadas para os 3 períodos estão colocadas nas Figuras 4.3, 4.4 e 4.5.

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.

Como vimos no Capítulo 2, o número de regimes determinado pelo modelo de mudança markoviana era maior do que os regimes determinados pelas políticas cambiais do Banco Central. Se existirem mudanças de regime endógenas no comportamento da volatilidade da série de log-retornos do câmbio, o modelo GARCH estimado pode superestimar valor previsto para a persistência da série, como apontado por HAMILTON e SUSMEL (1994).

4.6 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 estimada depende fundamentalmente da constância dos parâme-

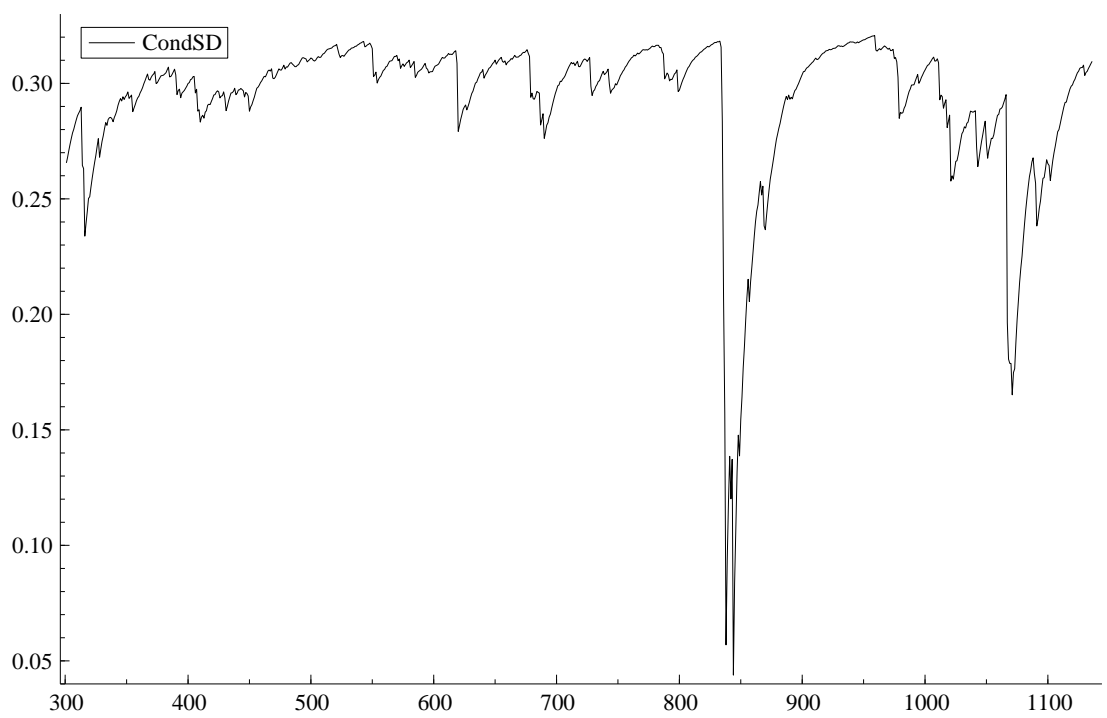


Figura 4.4: Garch(1,1) Banda Cambial

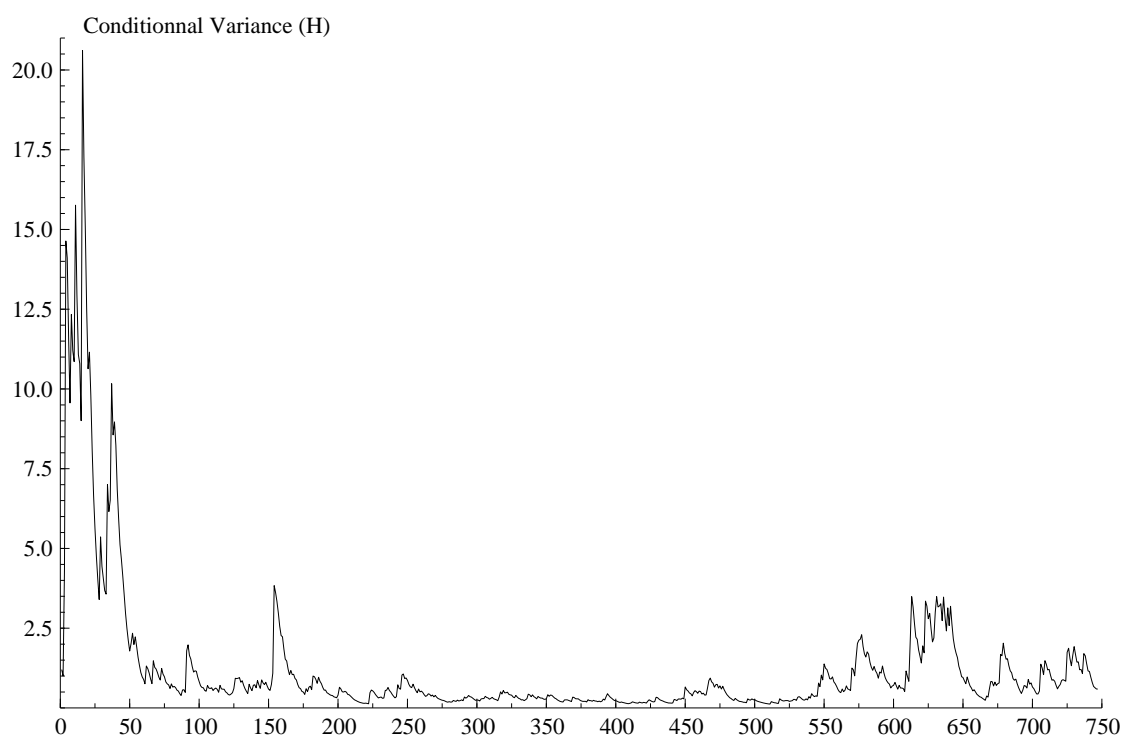


Figura 4.5: Garch(1,1) Pós Banda Cambial

tros, ou seja, ausência de quebras estruturais, precisamos de algum teste de constância de parâmetros no modelo GARCH.

LUNDBERG e TERASVIRTA (1998) 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:

$$h_t = w^* + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \zeta t + \gamma_1 \varepsilon_{t-1}^2 t + \delta_1 h_{t-1} t \quad (4.19)$$

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.

4.7 Análise de Monte Carlo

Para verificar o comportamento do modelo GARCH com nossos dados, 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. A literatura sobre estimação de modelos GARCH em condições de quebra estrutural aponta o fato de que a persistência nesse modelo é superestimada na presença de quebras estruturais no processo gerador da série, conforme HAMILTON e SUSMEL (1994).

Partindo do modelo estimado no Capítulo 2, escolhemos como mecanismo gerador dos dados em nosso estudo de Monte Carlo um modelo de mudança markoviana nos parâmetros do intercepto, parâmetros autoregressivos e variância incondicional (modelo MSIAH). Os parâmetros utilizados no estudo são os parâmetros estimados para os dados de log-retornos do câmbio pelo modelo MSIAH, conforme Tabela 2.6.

O objetivo do estudo é verificar se as estimativas da persistência dadas pelo modelo GARCH(1,1) são robustas na presença de mudanças no processo gerador dos dados. Assumindo o modelo e os parâmetros estimados no Capítulo 2 estamos sujeitos à incerteza quanto aos parâmetros estimados e principalmente quanto ao modelo escolhido, mas este estudo ajuda a esclarecer o comportamento da persistência do modelo GARCH quando o processo gerador contém quebras endógenas.

O experimento foi realizado gerando 2000 replicações da série definida pelos parâmetros da Tabela 2.6 e a matriz de transição disposta na Tabela 2.7. 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 4.3. Desta forma, o experimento fica mais consistente com o fato de que geralmente não conhecemos o processo

Tabela 4.7: 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

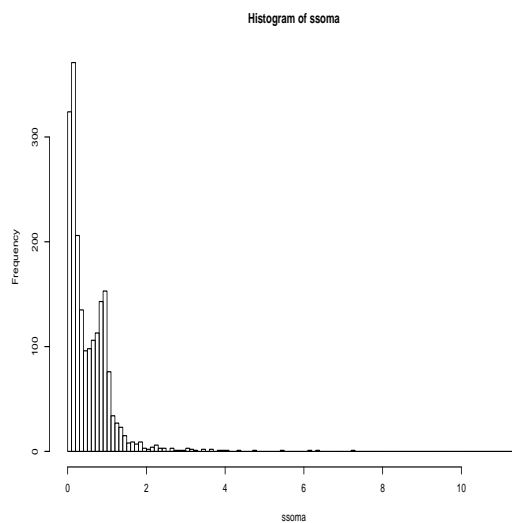


Figura 4.6: Histograma Persistência Estimada (Amostra Tamanho 100)

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 4.7.

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.

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

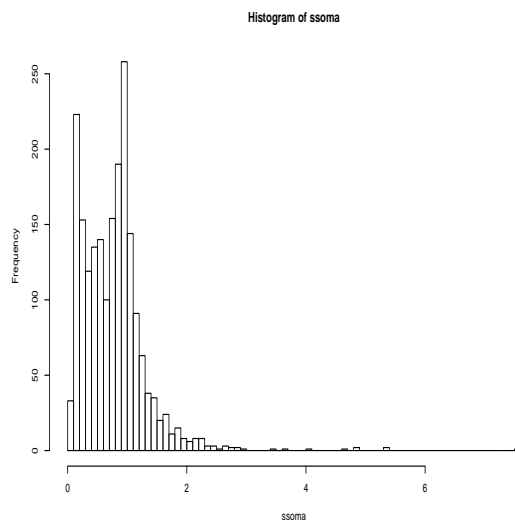


Figura 4.7: Histograma Persistência Estimada (Amostra Tamanho 250)

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 4.6, 4.7, 4.8 e 4.9) 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). 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 4.11.1 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).

Desta forma, a idéia de viés e principalmente o conceito de consistência assintótica não ficam claros já que o modelo estimado não corresponde ao modelo gerador dos dados. A questão relevante é determinar se um processo GARCH pode aproximar as previsões para a variância condicional caso o processo gerador dos dados fosse um MSIAH como o utilizado em nosso estudo.

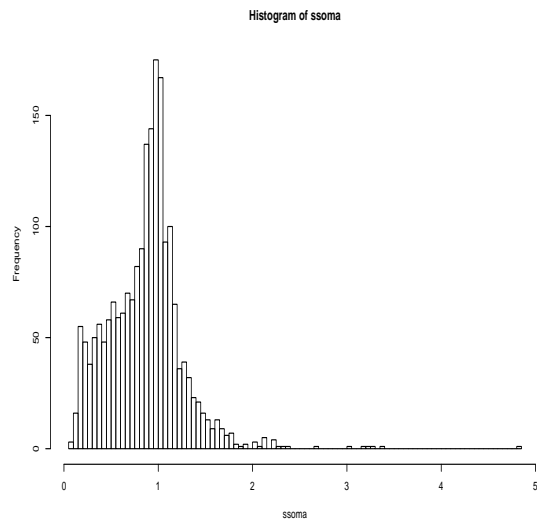


Figura 4.8: Histograma Persistência Estimada (Amostra Tamanho 500)

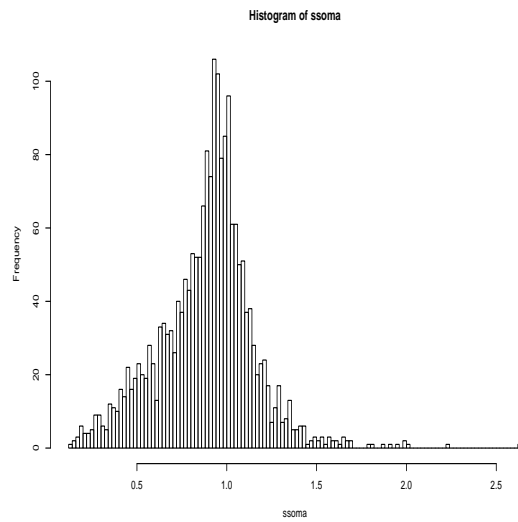


Figura 4.9: Histograma Persistência Estimada (Amostra Tamanho 1000)

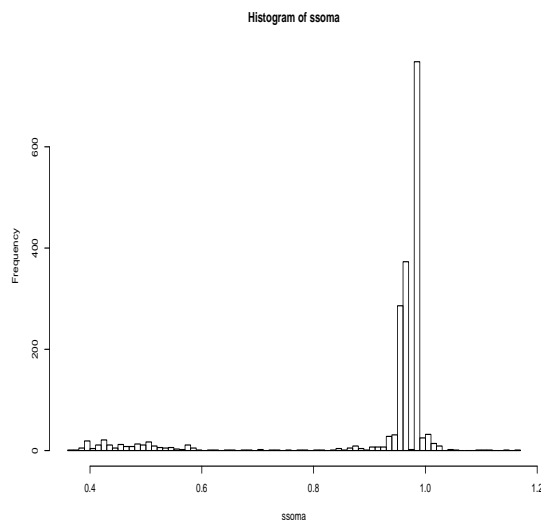


Figura 4.10: Estimação com Amostra Crescente - Densidade Empírica

4.8 Estimação com Amostra Crescente

Para analisar o comportamento do modelo GARCH na presença das quebras estruturais existentes em nossos dados adaptamos o procedimento de Monte Carlo utilizado acima, partindo de um número reduzido de observações e progressivamente 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 4.10 mostra que o comportamento da persistência estimada para o modelo GARCH é instável, sendo que, no início da amostra, é reproduzida a relação encontrada no experimento de Monte Carlo, relacionando o valor crescente da persistência estimada com o tamanho da amostra. 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.

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 4.6), 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.

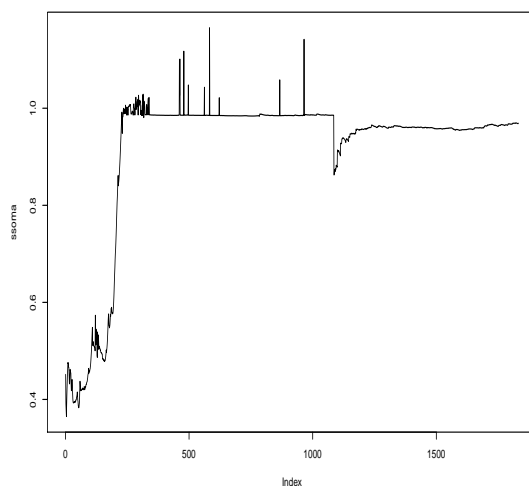


Figura 4.11: Persistência estimada Amostra Crescente

Tabela 4.8: Estimação em subamostras

amostra	persistência	desvio-padrão
1:2500	0.9063	0.0820
1:500	0.4182	0.093895
500:1000	1.1004	0.003231
1000:1500	1.0543	0.000221
1500:2000	0.9947	0.00054
2000:2500	0.9703	1.9323e-05

4.9 Estimação em Subamostras - Dados Simulados

Embora a estimação em subamostras como realizado na Seção 4.5 nos permita analisar com mais cuidado a consistência do valor estimado para a volatilidade, o fato de o valor estimado para a persistência depender do tamanho da amostra pode inviabilizar o uso do modelo GARCH como estimador da variância condicional, já que fica a questão prática de como determinar o tamanho ideal de amostra.

Estimar o modelo em subamostras, da forma como fizemos na Seção 4.5, parece adequado, mas este procedimento possui alguns problemas, além do fato de termos amostras reduzidas. O primeiro problema desse procedimento é como determinar as subamostras, o que equivale a como determinar os pontos de quebra estrutural relevantes. Esta questão é fundamental, já que a determinação dos pontos de quebra só seria adequada com o uso de informações extra-amostra. O uso da informação presente na amostra só é possível através de algum mecanismos estatístico como o utilizado na estimação do modelo de mudança markoviana, para evitar os problemas associados a *data mining* e *data snopping*.

Para ilustrar esse procedimento utilizamos uma das replicações (Figura 4.12) do es-

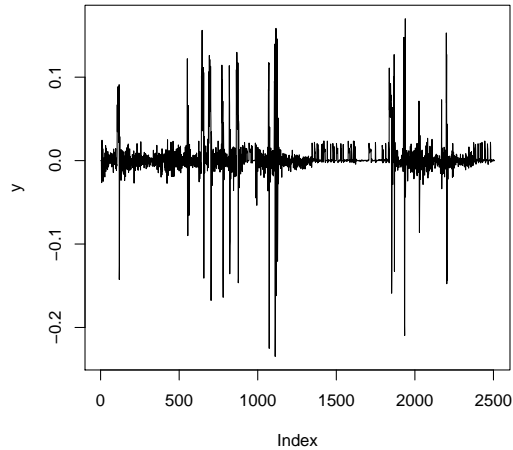


Figura 4.12: Série Gerada Amostra Tamanho 2500

tudo de Monte Carlo da amostra de tamanho 2500 e estimamos um processo $ARMA(0,0)$ - $GARCH(1,1)$ para as subamostras de tamanho 500. Os resultados desse procedimento mostram que os valores se mantêm próximos de 1, mas com uma variação relativamente grande. Embora a divisão das subamostras seja arbitrária em nosso exemplo, a validade deste procedimento é condicionada pela seleção correta das subamostras, e isso pode representar um problema caso a informação de fora da amostra seja insuficiente ou incorreta.

4.10 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, o que é somente válido caso os choques nessa série não tenham efeitos permanentes, o que levaria a mudanças abruptas 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 4.9) notamos que a persistência reproduz o padrão encontrado no estudo de Monte Carlo, de uma relação entre o valor da persistência estimada e o tamanho da amostra, neste caso, a janela de estimação utilizada.

Tabela 4.9: 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

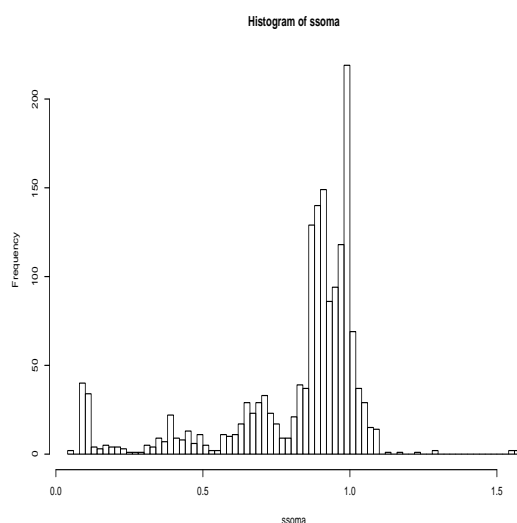


Figura 4.13: Histograma - Janela Móvel de 250 observações

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 4.13) 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 (Figura 4.14). 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 Figuras 4.15 e 4.16. 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.

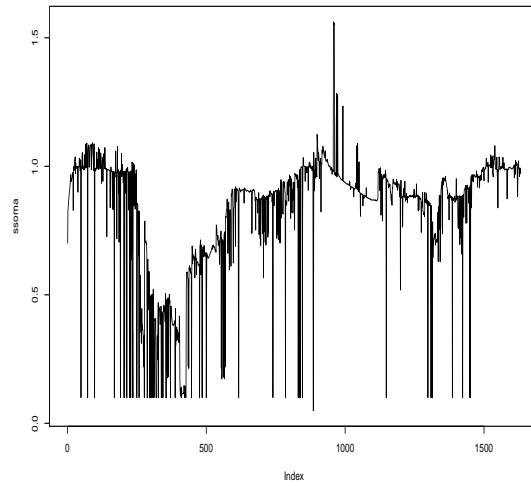


Figura 4.14: Valores Estimados - Janela de 250 Observações

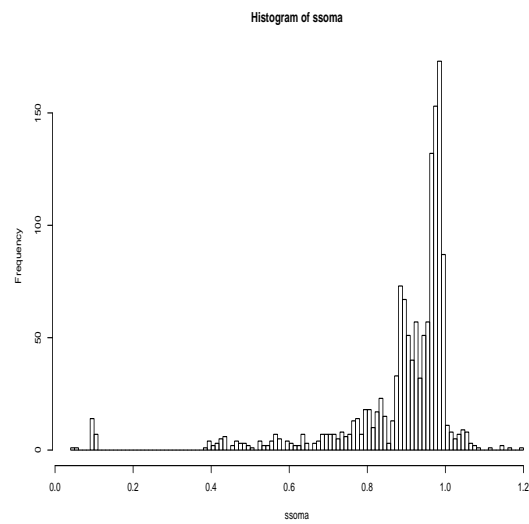


Figura 4.15: Histograma - Janela Móvel de 500 observações

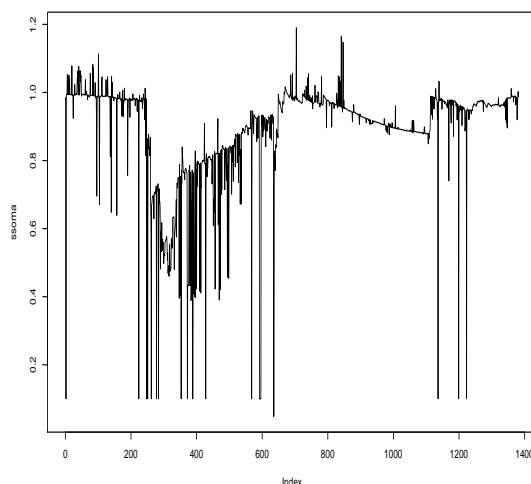


Figura 4.16: Valores Estimados - Janela de 500 Observações

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 (Figuras 4.17 e 4.18), mas 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. Janelas mais longas têm valores numericamente mais estáveis, mas não resolvem o problema de persistência excessiva causada pelas quebras estruturais.

4.11 Comparação com a Variância Prevista no Modelo de Mudança Markoviana

Uma análise adicional realizada foi 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 no Capítulo 2. 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.

A realização das previsões dinâmicas no modelo GARCH(1,1) é trivial, e pode ser realizada com a representação do processo (4.7) utilizando-se os parâmetros estima-

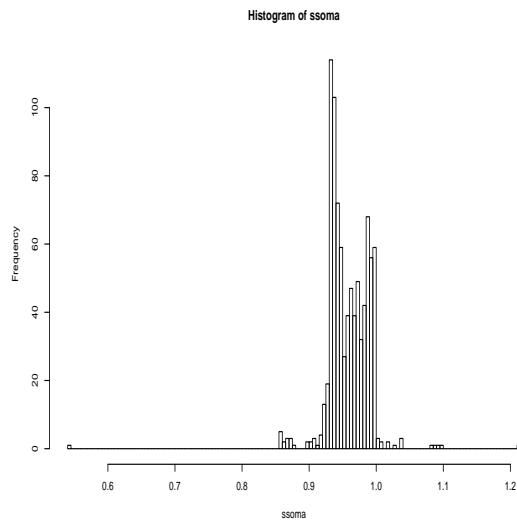


Figura 4.17: Histograma - Janela Móvel de 1000 observações

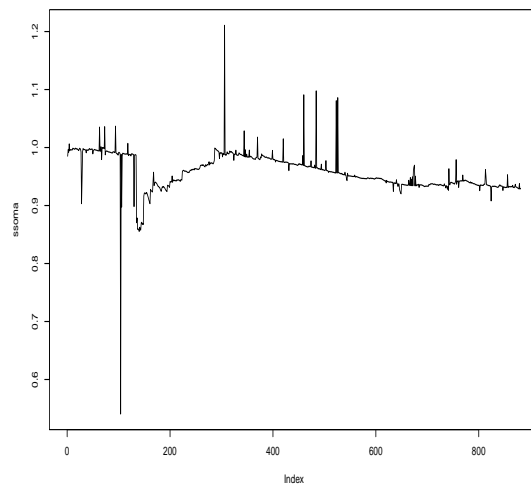


Figura 4.18: Valores Estimados - Janela de 1000 Observações

dos, o que equivale a uma previsão a n passos à frente de um processo ARMA(1,1), conferindo o decaimento exponencial observado.

4.11.1 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.

A questão fundamental é então 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 (1994) 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 ($N \times 1$) ξ_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 no Capítulo 2. O vetor ξ_t pode ser representado como:

$$\xi_t = \begin{cases} (1, 0, 0, \dots, 0), & \text{se } s_t = 1 \\ (0, 1, 0, \dots, 0), & \text{se } s_t = 2 \\ (0, 0, 1, \dots, 0), & \text{se } s_t = 3 \\ (0, 0, 0, \dots, 1), & \text{se } s_t = N \end{cases} \quad (4.20)$$

Assim, se $s_t = 1$ o vetor ξ_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 ξ_{t+1} assume valor 1 com probabilidade p_{ij} e zero nos demais casos. Conforme HAMILTON (1994), essa matriz tem esperança p_{ij} , e a esperança condicional $\xi_{t+1}|s_t = i$ é dada por:

$$E(\xi_{t+1}|s_t = i) = \begin{bmatrix} p_{i1} \\ p_{i2} \\ \cdot \\ \cdot \\ p_{iN} \end{bmatrix} \quad (4.21)$$

O vetor representado por 4.21 é simplesmente a i -ésima coluna da matriz de probabilidades de transição definida na Tabela 2.4. Quando $s_t = i$, o vetor ξ_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 4.21 implica que

$$E(\xi_{t+1}|\xi_t) = P\xi_t \quad (4.22)$$

e segue da propriedade markoviana (equação 4.20) que:

$$E(\xi_{t+1}|\xi_t, \xi_{t-1}, \xi_{t-2}, \dots) = P\xi_t \quad (4.23)$$

Assim podemos representar uma cadeia de Markov na forma de uma autoregressão vetorial para ξ_t dada por:

$$\begin{aligned}\xi_{t+1} &= P\xi_t + v_t \\ v_t &\equiv \xi_{t+1} - E(\xi_{t+1}|\xi_t, \xi_{t-1}, \dots)\end{aligned}\quad (4.24)$$

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

A equação 4.24 implica que:

$$\xi_{t+n} = v_{t+n} + Pv_{t+n-1} + P^2v_{t+n-2} + \dots + P^{n-1}v_{t+1} + P^n\xi_t \quad (4.25)$$

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:

$$E(\xi_{t+n}|\xi_t, \xi_{t-1}, \dots) = P^n\xi_t \quad (4.26)$$

Utilizando as definições anteriores podemos ver que o j -ésimo elemento do vetor ξ_t indica a probabilidade de que s_{t+n} assuma o valor j condicional aos valores dados no tempo t . Supondo que o valor de s_t no tempo t seja i a equação 4.26 estabelece que:

$$\begin{bmatrix} P\{s_{t+n} = 1|s_t = i\} \\ P\{s_{t+n} = 2|s_t = i\} \\ \vdots \\ P\{s_{t+n} = N|s_t = i\} \end{bmatrix} = P^n \cdot e_i \quad (4.27)$$

onde e_i é a i -ésima coluna da matriz identidade de dimensão N . A expressão 4.27 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 i , 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:

$$V_m = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^2 & \dots & 0 \\ & & \dots & \dots & \\ 0 & 0 & 0 & 0 & \sigma_N^2 \end{bmatrix} \quad (4.28)$$

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 :

$$E(\sigma_{t+n}^2|p_e) = P^n \cdot V_m \cdot p_e \cdot i \quad (4.29)$$

onde i é um vetor (1xN) contendo valores iguais a 1. A equação 4.29 é a esperança das variâncias associadas a cada regime previstas n períodos à frente condicionada às probabilidades de cada regime no tempo t .

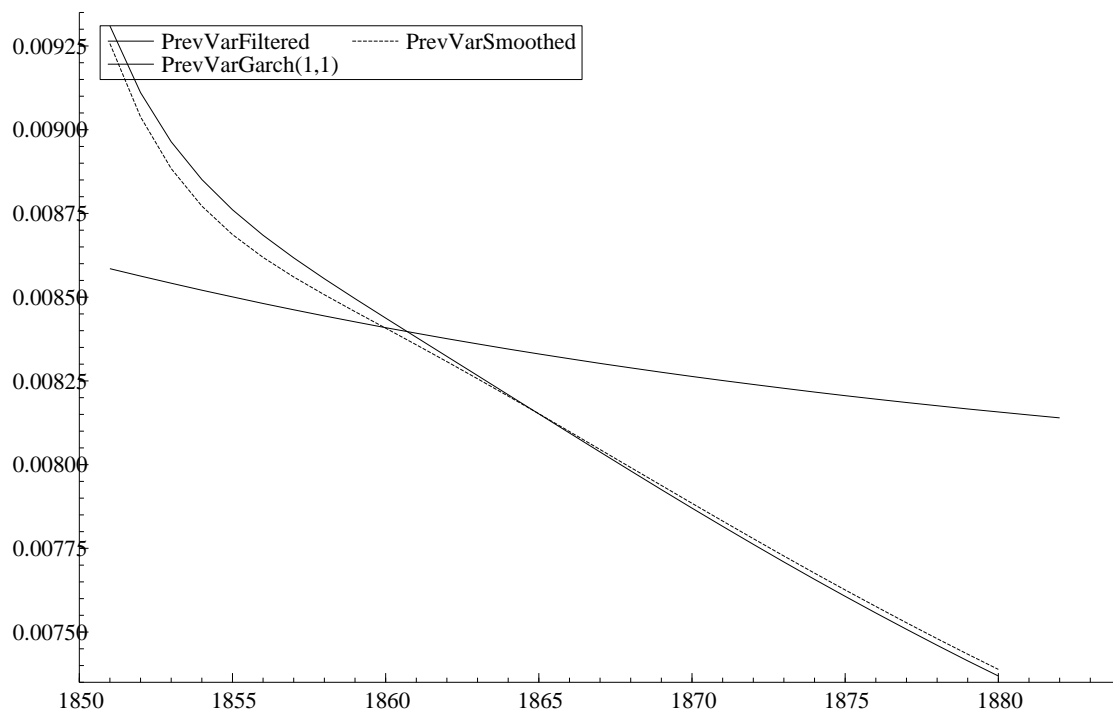


Figura 4.19: Previsões do Modelo de mudança markoviana para a Variância

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 4.2) e MSIAH (Tabelas 2.6, 2.7 e 2.2), e calculamos as previsões para esse intervalo utilizando as equações 4.7 e 4.29. Calculamos as previsões da variância do modelo de mudança markoviana usando as probabilidades filtradas e suavizadas.

A Figura 4.19 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 valores previstos para a variância são muito próximos, como mostra a Figura 4.20. Para esta parte da amostra os dois modelos aparentemente subestimam o valor da variância.

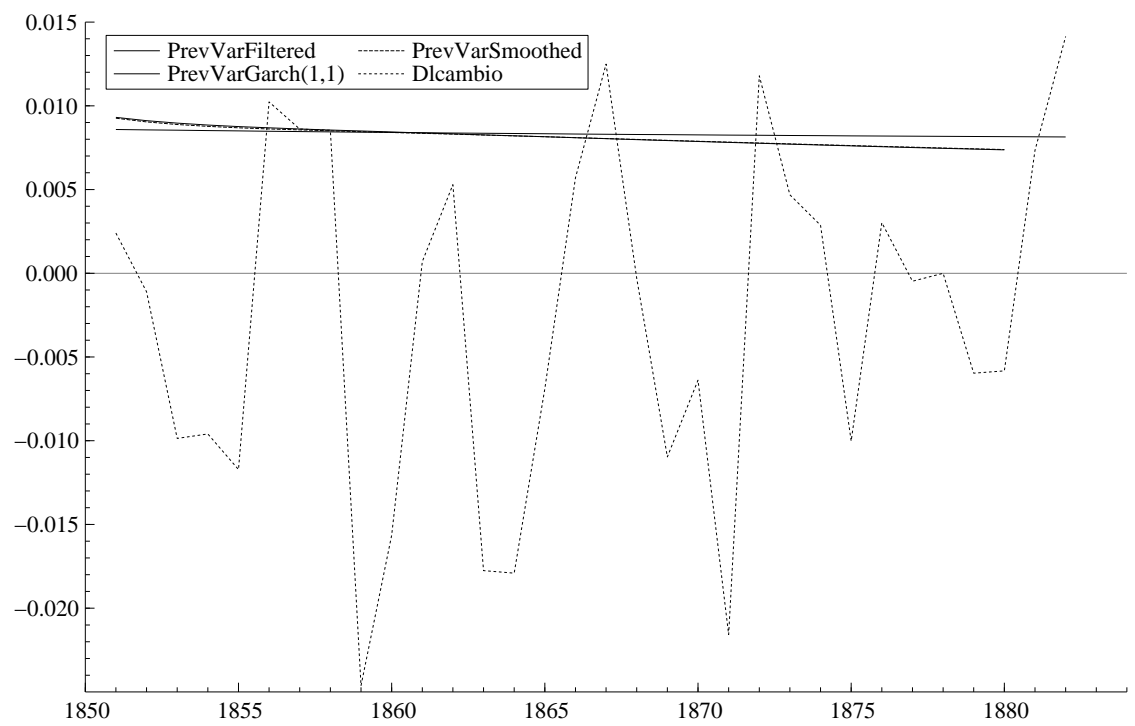


Figura 4.20: Previsões para a Variância e Valores Observados

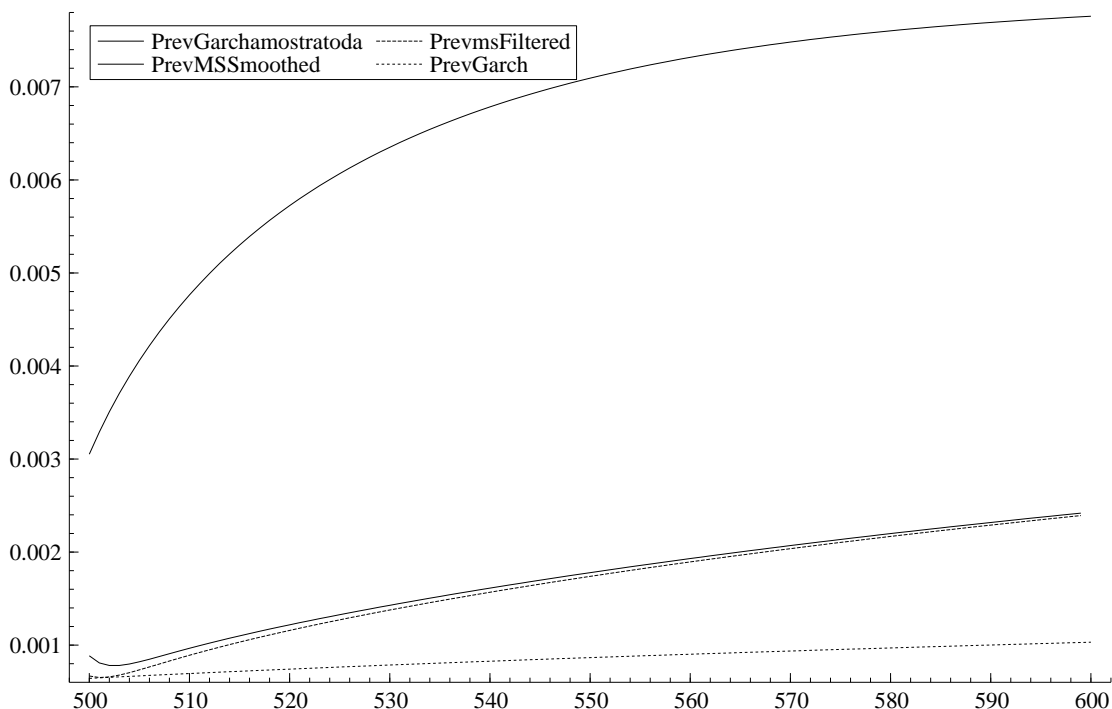


Figura 4.21: Previsão Amostra 500-600

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 4.21.

Em comparação com os valores realizados (Figura 4.22) vemos que nesse período as melhores previsões para a variância condicional são dadas pelo modelo GARCH(1,1) estimado até a observação 500. 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.

É interessante notar que as previsões do modelo de mudança markoviana, independente das probabilidades no período t , sempre irão convergir para um mesmo ponto.

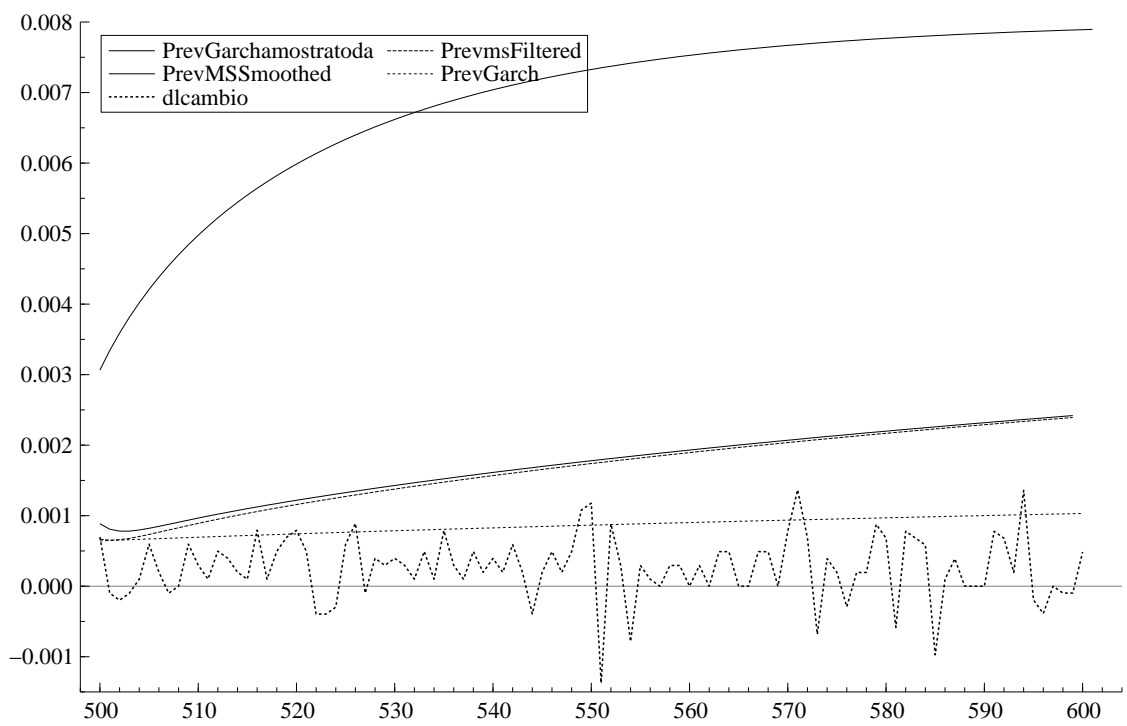


Figura 4.22: Previsão - Amostra 500-600 e Valores Observados

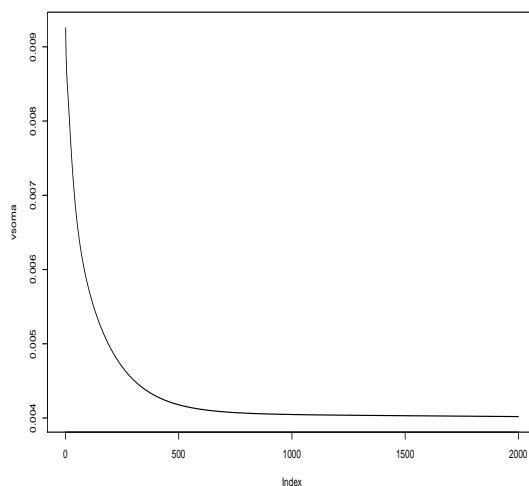


Figura 4.23: Previsão 2000 passos e Variância modelo MSIAH

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 4.23 ilustra essa propriedade.

Uma questão adicional é verificar se 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 decaimento exponencial, e assim torna possível comparar diretamente a persistência calculada por esses modelos.

Através das Figuras 4.24 e 4.25 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.

É importante notar que, com o aumento do horizonte de previsão, o valor da persistência estimado pelos dois modelos se aproxima. A Tabela 4.10 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 4.11, mostra que nas últimas n observações ocorre um resultado análogo ao da Tabela 4.10, onde as últimas observações tem decaimento mais rápido do que as observações iniciais. Como exemplo as observações 1980-2000 tem persistência aproximada de 0.85 enquanto que aumentando o

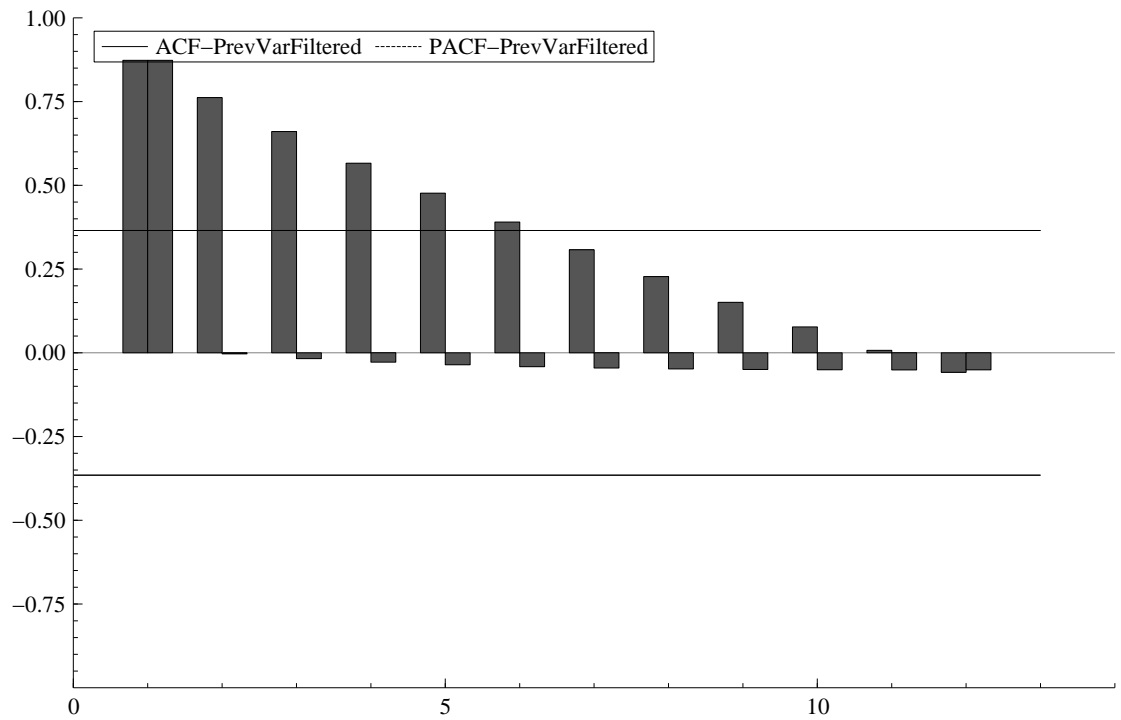


Figura 4.24: autocorrelação e Variância Filtrada

Tabela 4.10: Aproximação Persistência

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

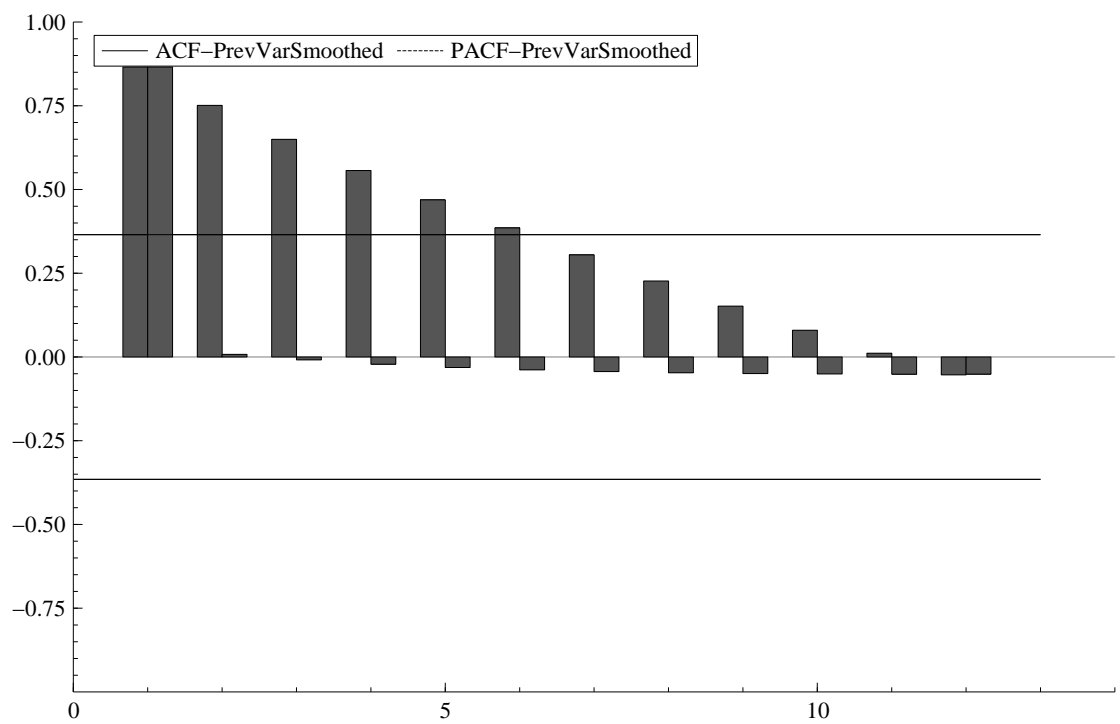


Figura 4.25: autocorrelação e Variância Suavizada

Tabela 4.11: Aproximação Persistência (n últimas observações)

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

intervalo de previsão para 1950-2000 a persistência aproximada se eleva para 0.94.

Este resultado, quando confrontado com a Figura 4.23, 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 4.10 e 4.11 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 que obtivemos no Capítulo 3, onde demonstramos 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 série em nível, mesmo que em cada regime a persistência do processo seja curta.

4.12 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 que estudamos no Capítulo 3, para modelar o comportamento da estrutura ARMA para a variância condicional presente nos modelos da classe GARCH. Da mesma forma que os modelos ARFIMA têm o objetivo de flexibilizar a estrutura de autocorrelações existentes nos modelos ARMA, os modelos FIGARCH permitem a captura de uma ordem de longas autocorrelações que porventura existam na variância condicional. Estudaremos dois modelos de memória longa na variância, os modelos FIGARCH e HYGARCH.

4.12.1 Modelo FIGARCH

O modelo FIGARCH foi proposto por BAILLIE, BOLLERSLEV e MIKKELSEN (1996) 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

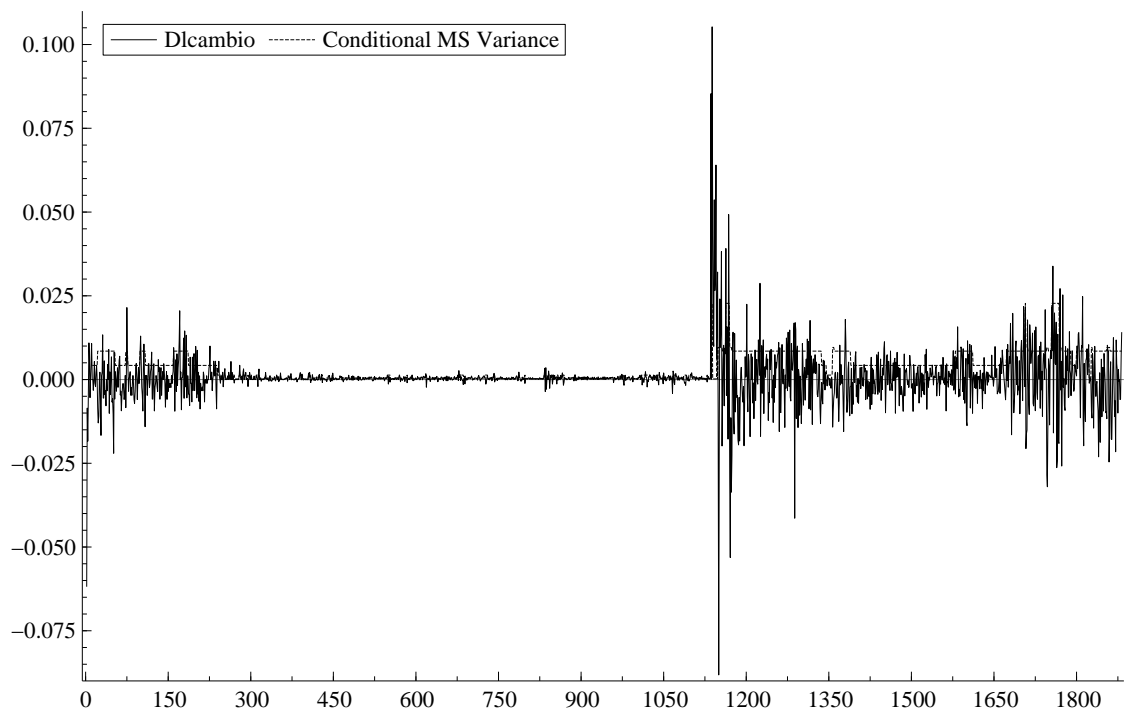


Figura 4.26: Variância Condicional

usando uma representação para os quadrados dos resíduos. A representação ARMA-nos quadrados pode ser escrita em termos de um polinômio *lag* como:

$$(1 - L)^d \varepsilon_t^2 = w - v_t - \beta_1 v_{t-1} \quad (4.30)$$

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 ε_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 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 (1996) 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 4.30.

DING e GRANGER (1996) 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 4.23 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 (2002) 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 (2002) 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 estacionariedade 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,

Tabela 4.12: ARMA(0, 0)-FIGARCH(1, d, 1)

ARMA (0, 0)	FIGARCH (1, d, 1)			
	Coefficiente	D. Padrão	<i>t-value</i>	<i>t-prob</i>
Cst(M)	0.033782	0.010037	3.366	0.0008
Cst(V)	0.239301	0.023384	10.23	0.0000
d-Figarch	0.135576	0.011915	11.380	0.0000
ARCH	-0.906189	0.055692	-16.27	0.0000
GARCH	-0.897911	0.059411	-15.11	0.0000
Log Likelihood	-1369.475			

BOLLERSLEV e MIKKELSEN (1996).

Na Tabela 4.12 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.

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.

4.12.2 Modelo HYGARCH

DAVIDSON (2002) 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 (2002), o modelo GARCH hiperbólico (HYGARCH) têm a seguinte estrutura:

$$\theta(L) = 1 - \frac{\delta(L)}{\beta(L)}(1 + \alpha((1-l)^d - 1)), \quad \alpha > 0 \quad (4.31)$$

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

$$\theta(L) = 1 - \frac{\delta(L)}{\beta(L)}(1 - \alpha\phi(L)) \quad (4.32)$$

onde a estrutura das defasagens corresponde a :

Tabela 4.13: ARMA(0, 0)-HYGARCH (1, d, 0)

ARMA (0, 0)	HYGARCH (1, d, 0)			
	Coefficiente	D. Padrão	<i>t-value</i>	<i>t-prob</i>
Cst(M)	0.030016	0.009128	3.288	0.0010
Cst(V)	0.032689	0.008857	3.691	0.0002
d-Figarch	0.855270	0.082473	11.80	0.0000
GARCH(1)	0.636272	0.095971	6.630	0.0000
Log Alpha (HY)	-0.066920	0.029649	-2.2570	0.0241
Log Likelihood	-1296.912			

$$\theta(L) = \zeta(1+d)^{-1} \sum_{j=1}^{\infty} j^{-1-d} L^j, \quad d > 0 \quad (4.33)$$

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 4.13.

O resultado interessante é que, embora o valor d estimado igual a 0.855270 seja estatisticamente significativo, 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 um¹, o que nos indica que não podemos rejeitar que este modelo seja estatisticamente diferente de um modelo IGARCH. Como a estrutura de modelo de memória longa HYGARCH na variância não pode ser considerada estatisticamente diferente de um IGARCH no modelo estimado, achamos mais adequado trabalhar com modelos de memória curta do tipo GARCH-IGARCH. As Figuras 4.27 e 4.28 mostram os valores previstos para a variância condicional e a média pelo modelo HYGARCH estimado.

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 4.11.1 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.

4.13 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

¹É necessário olhar com cuidado essa inferência, já que as distribuições desses parâmetros ainda não foram estudadas de forma completa.

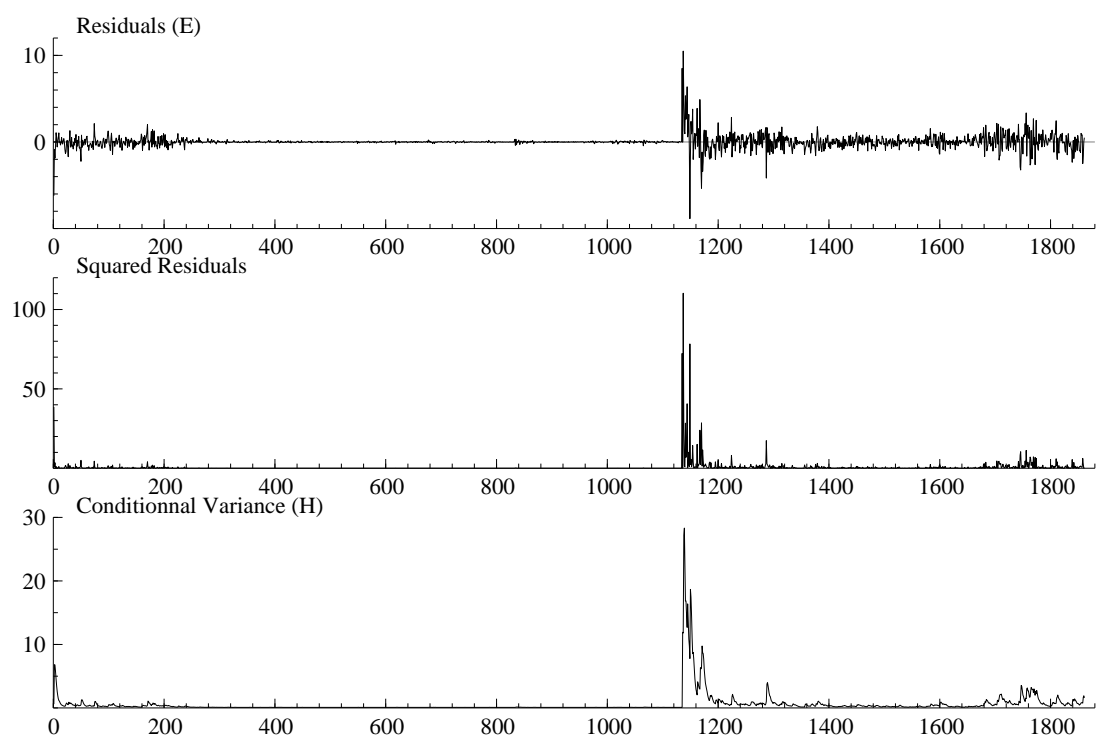


Figura 4.27: Previsões - Modelo HYGARCH

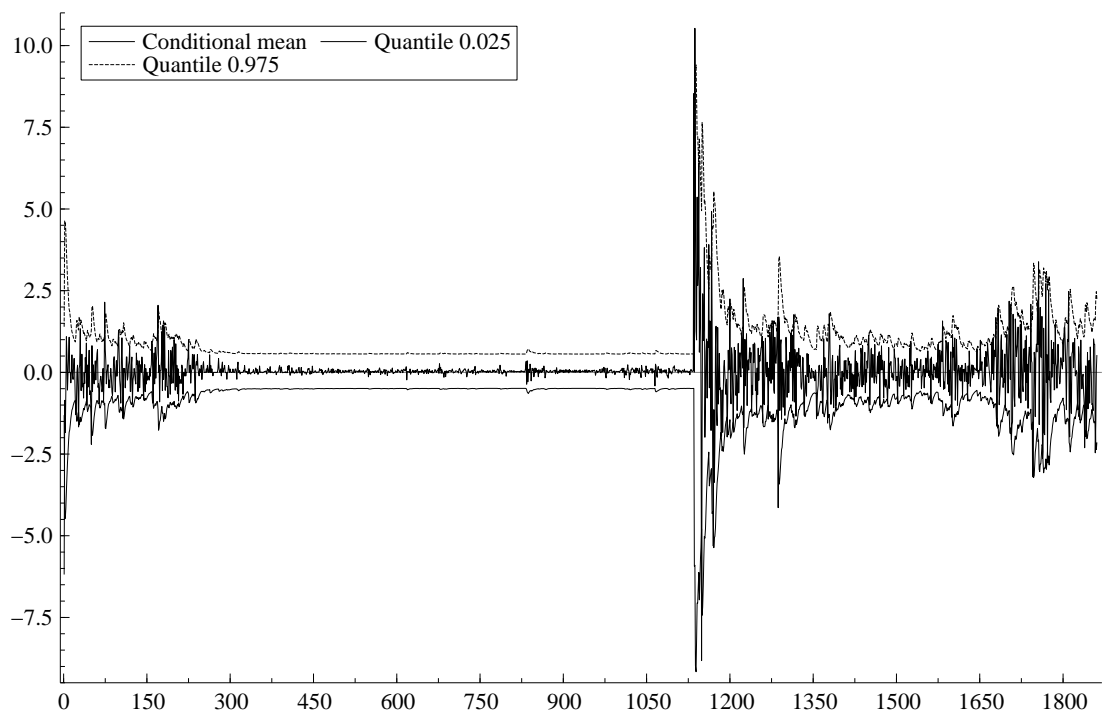


Figura 4.28: Intervalo de Confiança - Modelo HYGARCH

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 α_1 e β_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. Outra evidência a favor do modelo MSIAH utilizado é que, como colocado na Seção 2.4.1, este modelo consegue capturar toda a estrutura existente nesta série, tanto na média quanto na variância, como mostra a estatística BDS aplicada aos resíduos deste modelo.

Uma evidência adicional é que, observando o gráfico das previsões dinâmicas para a variância no modelo MSIAH elas são consistentes com o decaimento hiperbólico previsto pelos modelos de memória longa na variância, o que representa um tema adicional de estudo, análogo ao que realizamos no Capítulo 3, comparando o modelo de mudança markoviana com o modelo de memória longa ARFIMA, na modelagem da média da série.

Como colocamos na Introdução, a existência de variância condicional é causada por mudanças de parâmetros no mecanismo estatístico da série. Nosso estudo indica que quando as mudanças de parâmetros são abruptas, uma característica presente do modelo de mudança markoviana, o modelo GARCH é inadequado, já que este é construído com o pressuposto de que as mudanças nos parâmetros da média são suaves.

Outro resultado importante obtido nesse capítulo é mostrar que a existência de quebras na estrutura da variância gera um decaimento hiperbólico na estrutura autoregressiva da variância. Isto pode ser entendido como um indício de memória longa nesta série. Embora tenhamos fortes indícios de que esta memória longa seja espúria, induzida pelas quebras na variância condicional da série, é necessário um estudo de Monte Carlo (como o realizado no Capítulo 3) para validar essa conclusão e verificar se este resultado é possível de ser generalizado.

Como o modelo de mudança markoviana prevê mudanças abruptas para os parâmetros da média e da variância, ele se mostrou adequado para modelar a estrutura da variância condicional existente na série de log-retornos do câmbio. O mercado de câmbio no Brasil parece não evoluir de forma suave, e sim de forma mais instável, onde períodos de relativa calma são seguidos por momentos de extremo nervosismo, o que faz com que, segundo nossa análise, o modelo de mudança markoviana seja mais robusto do que o modelo GARCH(1,1).

Notas Adicionais

Os modelos GARCH foram estimados com o pacote de modelos para a volatilidade (*Volatility Models*) presente no programa PcGive. Para estimar os modelos FIGARCH e HYGARCH utilizamos a classe LMod de James Davidson. O modelo MSIAH estimado corresponde ao mesmo modelo do Capítulo 2 estimado usando MSVAR para

Ox. O experimento de Monte Carlo foi programado na linguagem R usando o pacote *tseries* para a estimação dos processos GARCH(1,1). As previsões para a variância no modelo MSIAH também foram programadas na linguagem R.

Capítulo 5

Conclusões

Nesse estudo realizamos uma análise dos resultados sobre as questões de persistência a choques e eficiência de mercado na média e na variância da série de log-retornos da taxa R\$/US\$. Verificamos estas questões através de procedimentos relacionados a modelos lineares no Capítulo 1, modelos de quebra estrutural usando mudança markoviana no Capítulo 2, modelos de memória longa no Capítulo 3 e no Capítulo 4, analisamos a questão da persistência a choques na volatilidade condicional desta série.

A primeira evidência encontrada é a que as inovações na taxa de câmbio em nível têm efeitos permanentes, como mostram todos os testes para raízes unitárias que foram realizados. A série do câmbio, inclusive no subperíodo de vigência das bandas cambiais, acumula todos os choques ocorridos no nível da série sem mecanismos robustos de dissipação de choques. Nenhuma das análises realizadas conseguiu estabelecer alguma tendência de reversão à média para a série do câmbio em nível.

Este resultado, como mostrado no Capítulo 1, indica que o valor da taxa de câmbio R\$/US\$ pode ser considerado como a acumulação dos choques existentes em todo o histórico da série. Este comportamento, no entanto, é consistente com a determinação da taxa de câmbio através de fundamentos econômicos, como fica demonstrado nas Seções 1.2 e 1.3. A incorporação dos fundamentos econômicos no valor do câmbio é realizada através do componente de inovação dado pelos choques diários na série. Como estes choques não se dissipam, cada inovação se incorpora ao conjunto informacional existente no passado da série.

Mas se a presença de uma raiz unitária no série de câmbio leva a efeitos permanentes dos choques na série de câmbio, foi necessário analisar se existe alguma estrutura de correlação na primeira diferença da série, ou de forma economicamente mais informativa, o log da primeira diferença dos preços, chamado de log-retornos de câmbio. A presença de algum padrão nos log-retornos significaria a existência de informação que possibilitaria poder de previsão sobre as mudanças no valor da taxa de câmbio.

Se os agentes usam de forma racional toda a informação disponível, o preço desse ativo no período presente já incorpora toda a informação econômica relevante, e assim não é possível usar este mesmo conjunto informacional para prever de forma significativa as mudanças futuras no nível da série do câmbio R\$/US\$. Esta formulação, baseada na lei das expectativas iteradas, é a base da hipótese de eficiência de mercado que testamos durante este estudo. A presença de alguma estrutura de correlação nos log-retornos indicaria que ainda existe a possibilidade de ganho especulativo usando o padrão de correlação, o que seria uma violação da eficiência de mercado, já que ainda existiria informação não explorada no mercado.

Os procedimentos para analisar a hipótese de eficiência de mercado realizados no Capítulo 1 obtinham resultados conflitantes, principalmente os resultados obtidos na estimação em subamostras para o modelo de passeio aleatório do tipo III. Estes resultados, no entanto, eram baseados no pressuposto de que a estrutura de dependência existente nos dados é uma estrutura linear, um pressuposto que se mostra inadequado em vista das mudanças de parâmetros causadas pelas quebras estruturais na série.

Os resultados dos testes realizados sobre a validade da hipótese de eficiência fraca de mercado, para o mercado de câmbio no período em estudo nos permitem algumas conclusões relacionadas à forma funcional das estruturas de dependência existentes nesta série e ao poder de previsão possível. Os testes de independência realizados no Capítulo 1 mostram que existe alguma estrutura de dependência tanto na média quanto na variância da série, e temos fortes evidências de que esta estrutura é possivelmente não-linear, como mostram os testes BDS e de não-linearidade de White e Teräsvirta.

Tendo em vista que a estrutura de dependência é não-linear, formulamos dois modelos para tentar capturá-la: um modelo de mudança markoviana (MSIAH) em todos os parâmetros e um modelo de memória longa com ordem de integração fracionária (ARFIMA). O modelo de mudança markoviana se mostrou adequado para capturar as características de dependência existentes na série de log-retornos do câmbio, tanto na média quanto na variância. Uma evidência da adequação deste modelo é o fato de que usando a estatística BDS não podemos rejeitar a hipótese nula de que os resíduos padronizados deste modelo sejam independentes e identicamente distribuídos, ou seja, não resta estrutura nenhuma nos resíduos, mostrando que o modelo MSIAH consegue capturar toda a estrutura de dependência existente na série.

Um resultado importante foi mostrar que a dependência não linear existente nos dois primeiros momentos desta série é adequadamente modelada por um processo linear em partes composto de 7 regimes, onde cada regime é um processo autoregressivo com parâmetros distintos. Embora a soma dos regimes gere um processo não linear, cada processo em separado é linear, o que torna a análise dos resultados mais fácil e comparável às análises realizadas nos modelos lineares tradicionais.

A possível crítica de que este ajuste é devido ao elevado número de parâmetros estimados no modelo MSIAH, caracterizando a possibilidade de *overfitting*, pode ser respondida notando que, mesmo com o elevado número de parâmetros, este modelo é escolhido pelos critérios de informação em relação aos demais modelos, e o número de 7 regimes que leva ao grande número de parâmetros estimados é justificado por uma série de testes de especificação.

A questão fundamental é se esta estrutura não linear presente no modelo de mudança markoviana é realmente uma violação do hipótese de Eficiência Fraca de Mercado, no sentido do potencial de exploração destes padrões para a obtenção de lucros significantes. Os resultados obtidos pelo modelo de mudança markoviana explorado no Capítulo 2 mostram que, mesmo após a identificação e estimação dos regimes associados às quebras estruturais ocorridas nesta série, podem ser identificados alguns padrões de persistência na média.

Estes padrões remanescentes correspondem aos processos autoregressivos que são estatisticamente significativos nos regimes 1, 4, 5, 6 e 7. Para verificar se estes padrões são uma efetiva violação da hipótese de eficiência fraca de mercado é necessário analisar se estas estruturas autoregressivas são suficientemente fortes para a possibilidade de arbitragem. Sob este aspecto, vemos que os regimes 4, 5 e 6 possuem uma estrutura autoregressiva que embora seja estatisticamente significativa, é muito fraca, e nas condições existentes de custos de transação não são economicamente viáveis. Os regimes 1 e 7, entretanto representam possibilidades reais de ganho especulativo, já

que estão associados aos períodos de grande valorização do Dólar em relação ao Real, e correspondem aos períodos onde esta especulação efetivamente aconteceu.

A hipótese de eficiência fraca de mercado, no entanto, continua válida se consideramos a série de log-retornos do câmbio após o ajuste para o risco. Este ajuste, que equivale a padronizar os retornos utilizando a variância associada a cada retorno estimada pelo modelo de mudança markoviana, faz com que não se possa rejeitar a hipótese de que os retornos ajustados sejam provenientes de um modelo de passeio aleatório, modelo que está associado à não previsibilidade das inovações que definem a hipótese de Eficiência de Mercado. Esta evidência era obtida através do uso da estatística de Razão de Variância aplicada aos retornos padronizados.

A hipótese de eficiência fraca de mercado também pode ser justificada no contexto do modelo de mudança markoviana pela complexidade computacional de se realizar previsões quando o conjunto de parâmetros sujeitos à troca de regimes inclui os parâmetros autoregressivos. Nesta situação, a previsão para fora da amostra se torna um procedimento não linear relativamente complexo, e o número de possíveis previsões é condicionada ao número de regimes existentes, crescendo exponencialmente com o número de períodos à frente. A isso se soma o fato de que a estrutura autoregressiva é muito fraca, exceto nos regimes 1 e 7. O grau de incerteza associada a estas previsões também se torna um problema complicado, já que a construção de intervalos de confiança para as previsões futuras também se torna um procedimento não linear, como demonstramos no Capítulo 4.

As conclusões gerais sobre o modelo de mudança markoviana nos parâmetros de intercepto, coeficientes autoregressivos e variância nos 7 regimes são que este modelo obteve grande sucesso em determinar a estrutura existente nos de log-retornos do câmbio, determinando pontos de quebra estrutural endógena e exógena consistentes com os fatos econômicos relevantes que influenciaram o comportamento desta série. O ajuste do modelo também foi muito superior aos do modelo linear e do modelo de memória longa. O poder de previsão deste modelo e a conseqüente análise sobre eficiência fraca de mercado não são conclusivas, contudo, devido às dificuldades em se calcular e estabelecer intervalos de confiança para as previsões devido a incerteza sobre qual é o regime vigente.

No segundo modelo estudado, o modelo ARFIMA, a existência de uma estrutura de memória longa nesta série, capturada pelo parâmetro da ordem de integração fracionária, não se mostrou robusta. Embora os modelos de memória longa estimados sem controle para a heterocedasticidade condicional se mostrem significativos, quando estimamos modelos ARFIMA onde se inclui a estrutura autoregressiva existente na variância condicional da série (através de um componente GARCH), o parâmetro de integração fracionária deixa de ser estatisticamente significativo.

Para confirmar que a existência de memória longa na média desta série é gerada por mudanças na estrutura da variância, construímos um experimento de Monte Carlo que gera mudanças nos valores da variância condicional através de uma cadeia de Markov, com os parâmetros estimados pelo modelo MSIAH do Capítulo 2. Este experimento mostra que, com o aumento do tamanho da amostra, o número de processos que apresentam memória longa tende à quase totalidade das séries geradas, enquanto que o verdadeiro mecanismo gerador dos dados é um processo de memória curta.

A existência de heterocedasticidade faz com que a análise da estrutura de dependência não seja estatisticamente confiável. Podemos notar isso pelo fato de que a definição das autocorrelações significativas (através de procedimentos como o teste de Box-Pierce) são explicitamente baseadas no pressuposto de que a variância é constante. O fato de a variância em nossa série não ser constante torna significativas estruturas de

dependências que, na realidade, são estatisticamente fracas.

Este problema se torna ainda mais grave pelo fato de que os testes conjuntos para n autocorrelações são baseados no fato de que a soma dos quadrados da raiz quadrada das autocorrelações, multiplicadas pelo tamanho da amostra, tenham distribuição qui-quadrada. Como nossa série é composta pela soma de componentes provenientes de distintas distribuições, o pressuposto de normalidade é violado, e assim também se viola o pressuposto de distribuição qui-quadrada (soma de distribuições Normal padrão elevadas ao quadrado) utilizado como distribuição no teste das autocorrelações.

Aplicando este mesmo raciocínio para o modelo ARFIMA, vemos que a estimação do parâmetro da ordem de integração fracionária não é robusta na presença de heterocedasticidade. O procedimento de Monte Carlo realizado no Capítulo 3 mostra que se geramos uma série com memória curta mas com mudanças markovianas na heterocedasticidade incondicional, induzimos um processo de memória longa espúria. Como o processo de memória longa é definido como a existência de autocorrelações significativas mesmo em intervalos muito distantes, o problema da heterocedasticidade e não normalidade nos retornos se torna presente no teste da significância do parâmetro da ordem de integração fracionária que corresponde ao processo de memória longa.

Estas duas evidências mostram que a estrutura de longa dependência entre as autocorrelações pode ser gerada por um processo de memória curta com quebras estruturais nos parâmetros da variância incondicional, e, deste modo, a existência de memória longa nesta série deve ser vista com muita cautela. Embora exista uma grande literatura sobre a equivalência observacional entre modelos de mudança markoviana e memória longa, nosso estudo incorpora às explicações existentes que o principal componente responsável por esse comportamento são quebras estruturais criadas por mudanças nos parâmetros da variância incondicional da série.

A existência de mudanças abruptas de parâmetros e os efeitos destas mudanças sobre a modelagem empírica da série do câmbio no período em estudo é a característica central em estudo nesta dissertação. A existência destas quebras estruturais também afeta toda a análise sobre a persistência de choques na volatilidade da série, estudadas no Capítulo 4.

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), mas, com o aumento do horizonte de previsões, a persistência apontada era equivalente a do modelo GARCH.

O relacionamento da persistência estimada com o horizonte de previsões mostra um comportamento análogo ao observado entre o parâmetro de integração fracionária no modelo ARFIMA e o tamanho da amostra. Observando que a função de previsão da variância no modelo de mudança markoviana é melhor aproximado por uma função hiperbólica, também notamos que a equivalência entre modelos de mudança estrutural markoviana e modelos de memória longa para a média da série é possivelmente válida para a variância do processo.

Os modelos estimados de memória longa na volatilidade não se mostram apropriados para os nossos dados, mas a possibilidade de que a memória longa na variância

seja devida à quebras estruturais no processo da variância condicional se torna real. A verificação desta possibilidade fica para estudos futuros.

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, como estudado no Capítulo 4, e também pode causar a falsa percepção de estruturas de memória longa tanto na média quanto na volatilidade, como demonstramos nos Capítulos 3 e 4. Estes resultados mostram a necessidade de se desenvolver métodos de estimação de processos de memória longa robustos na presença de heterocedasticidade.

Os resultados obtidos nesta dissertação são consistentes com novas classes de modelos de precificação de ativos propostos por KIRMAN e TEYSSIERRE (2001), baseados em fundamentos microeconômicos comportamentais com interações entre os agentes. Nestes modelos, as previsões são influenciadas pela expectativa de comportamento entre os agentes, e a formação de expectativas é baseada em processos de aprendizagem dinâmica, com o uso e abandono de regras de previsão.

Os resultados das simulações realizadas nos modelos de KIRMAN e TEYSSIERRE (2001) mostram que, quando existe um processo de aprendizagem com troca de regras, são observados efeitos análogos aos relatados nos mercados de ativos reais, como comportamentos de manada (*herding*) e bolhas nos preços dos ativos. Estes comportamentos parecem ter tido influência significativa da determinação da taxa de câmbio R\$/US\$ no período estudado.

As características estatísticas das séries de preços de ativos geradas por estes modelos são totalmente consistentes com as características encontradas na série de taxa do câmbio que estudamos. As características comuns são a rejeição da hipótese de reversão à média, a presença de estruturas não lineares de dependência nos retornos e nos retornos ao quadrado, existência de heterocedasticidade condicional e também a indução de memória longa.

Nesse estudo conseguimos identificar que o responsável pelas estruturas de memória longa na média da série e da alta persistência nas volatilidades estimadas pelo modelo GARCH são devidas a mudanças na variância incondicional controladas por uma cadeia de Markov. Uma possível inferência é que as mudanças nas regras de comportamento e previsão dos agentes econômicos são responsáveis pelas características de dependência não-linear nos dois primeiros momentos das séries de preços de ativos, causadas por mudanças na estrutura da variância incondicional dos preços.

Os novos modelos de precificação de ativos propostos por KIRMAN e TEYSSIERRE (2001) estão inseridos na classe de modelos econômicos com aprendizagem e incerteza sobre a estrutura da economia. Estas formulações são uma tentativa de capturar comportamentos importantes como a presença de bolhas, irracionalidade e comportamento de manada, que embora não sejam discutidos nessa dissertação dada a orientação econométrica do trabalho, são fundamentais na determinação da taxa de câmbio e preços de ativos em geral no curto prazo.

Referências Bibliográficas

- ALBERT, J. e CHIB, S. (1993). Bayes Inference Via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. **Journal of Business & Economic Statistics** v. 11, p. 1-16.
- ANDERSEN, T. G., BOLLERSLEV, T. DIEBOLD, F. X. e LABYS, P. (1999). The Distribution of exchange Rate Volatility. **NBER Working Paper** 6961.
- ANG, A. e BEKAERT, G. (1998). Regime Switches in Interest Rates". **Research Paper** 1486, Stanford University.
- BACHELIER, L. (1900). Theory of Speculation, reimpressa em COOTNER, P (ed.) **The Random Character of STOCK Market Prices**, (1964) . MIT Press.
- BAILLIE, R.T, BOLLERSLEV, T. e MIKKELSEN, H.-O. (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. **Journal of Econometrics** v. 74, p. 3-30.
- BARNETT, W., POWELL, J. e TAUCHEN, G. (1992). **Nonparametric and Semiparametric Methods in Econometrics and Statistics**. Cambridge University Press.
- BAUWENS, L., LUBRANO, M. e RICHARD, J-F. (1999). Bayesian Inference in Dynamic Econometric Models. Cambridge University Press.
- BERGMAN, M. e HANSSON (2000). Real exchange Rate and Switching Regimes, manuscrito, Lund University
- BLACK, F. e SCHOLES. M. (1973). The Pricing of Options and Corporate Liabilities. **Journal of Political Economy** v. 81,p. 637-654.
- BOLLEN, N. P. B., GRAY, S. e WHALEY, R. E. (2000). Regime Switching in Foreign exchange Rates: Evidence from Currency Options Prices. **Journal of Econometrics** v. 94, p. 239-276.
- BOLLERSLEV, T. (1986). Generalised Autoregressive Conditional Heteroscedasticity. **Journal of Econometrics** v. 31, p. 307-327.
- BOLLERSLEV, T, CHOU, R. e KRONER, K. (1992). ARCH Modelling in Finance:A Review of Theory and Empirical Evidence. **Journal of Econometrics** v. 53, p. 5-60.

- BOLLERSLEV, T. e MIKKELSON, H.O. (1996). Modelling and Pricing Long Memory in Stock Market Volatility. **Journal of Econometrics** v. 73, p. 151-184
- BROCK, W., DECHERT, W. e SCHEINKMAN, J. (1987). A Test for Independence Based on the Correlation Dimension. working paper não publicado, University of Wisconsin at Madison, University of Houston e University of Chicago.
- CAI, J. (1994). A Markov Model of Switching Regime ARCH. **Journal of Business & Economic Statistics** 12, p. 309-316.
- CAMPBELL, J. Y., LO, A. W. e MACKINLAY, A. G. (1997). **The Econometrics of Financial Markets**. Princeton University Press.
- CLARK, P. B. (1973). A Subordinated Stochastic Process Model With Finite Variance for Security Prices. **Econometrica** v. 68, p. 153-55.
- COCHRANE, J. How Big is the Random Walk in the GNP? . **Journal of Political Economy** v. 96, p. 893-920.
- DAHLQUIST, M. e GRAY, S. (2000). Regime-Switching and Interest Rates in the European Monetary System. **Journal of International Economics** v. 50, p. 399-419.
- 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>.
- 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**, v. 39, p. 1-38.
- DICKEY, D. A. e FULLER, W. (1979). Distribution of Estimators for Autoregressive Time Series with a Unit Root . **Journal of the American Statistical Association** v. 74, p. 427-431.
- DIEBOLD, F. X. e RUDEBUSH, G. D. (1996). Measuring business cycles: a Modern Perspective. **Review of Economic Studies** v. 78, p. 67-77.
- DING, Z. e GRANGER, C.W.J (1996). Modeling Volatility Persistence of Speculative Returns: a New Approach. **Journal of Econometrics** v. 73, p. 185-215.
- DOORNIK, J.A. (2001). **Ox: An Object-Oriented Matrix Language** (4th edition), London: Timberlake Consultants Press.
- DOORNIK, J.A. e OOMS, M. (2000). Multimodality and the GARCH Likelihood. manuscrito.
- DORNBUSH, R. (1976). Expectations and exchange Rate Dynamics. **Journal of Political Economy**, v. 84 1161-1176.
- DURKER, M. J. e FISHER, A. (1996). Inflation Targeting in a Small Open Economy: Results from for Switzerland. **Journal of Monetary Economics** v. 37, p. 89-107.

- DUEKER, M. J. (1997). Markov Switching in GARCH Processes and Mean-Reverting Stock Market Volatility. **Journal of Business & Economic Statistics** v. 15, p. 26-34.
- ENGEL, C. M. (1992). Can the Markov Switching Model Forecast Exchange Rates ?. **NBER Working Paper** 4210.
- ENGEL, C. e HAMILTON, J. D. (1990). Long Swings in the Dollar: Are They in the Data and Do Markets Know It? . **American Economic Review** v. 80, p. 689-713.
- ENGLE, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. **Econometrica** v. 46, p. 987-1007.
- ENGLE, R. e BOLLERSLEV, T. (1986). Modelling the Persistence of Conditional Variances. **Econometric Reviews** v. 5, p. 1-50.
- ENGLE, R. F , LILIEN D. e ROBBINS, R. (1987). Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model. **Econometrica** v. 55, p. 391-407.
- ENGLE, R. F e MUSTAFA, C. (1995). Implied ARCH Models from Options Prices. in **ARCH - Selected Readings** (1995), editor ENGLE,R. Oxford University Press.
- ERLANDSSON, U. (2000). Forecasting Swedish Interest Rate Volatility- A Regime Switching Approach . **Lund Macroeconomic Studies** 2000-5 . Lund University.
- FRANSES, P. H. e DIJK, D. (2000) . **Non-linear Time Series Models in Empirical finance**. Cambridge University Press.
- 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.
- GARCIA, R. e PERRON, P. (1990). An Analysis of the Real Interest Rate under Regime Shifts. **Université of Montreal working paper**.
- GOLDFELD, S. M. e QUANDT, R. E. (1973). A Markov Model for Switching Regressions. **Journal of Econometrics** v. 1, p. 3-16.
- GRANGER, C. (1980) Long-memory Relationships and the Aggregation of Dynamic Models **Journal of Econometrics**, v. 14, p. 227-238.
- GRANGER, C. e JOYEUX (1980). An Introduction to Long Memory Time Series Models and Fractional Differencing. **Journal of Time Series Analysis** v. 1, p. 15-19.
- GRASSBERGER, P. e PROCACCIA, I. (1983). Measuring the Strangeness of Strange Attractors. **Physica**, v. 9D, p. 189:208.
- GRAY, S. F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime Switching Process. **Journal of Financial Economics** v. 42, p. 27-62.

GROSSMAN, S. e STIGLITZ, J. (1980). On the Impossibility of Informationally Efficient Markets. **American Economic Review** v. 70, p. 393-408.

HAMILTON, J. D. (1989.) A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. **Econometrica** v. 57, p. 357-384.

HAMILTON, J. D. (1990). Analysis of Time Series Subject to Changes in Regime. **Journal of Econometrics** v. 45, p. 39-70.

HAMILTON, J. D. (1994). **Time Series Analysis**. Princeton University Press.

HAMILTON, J.D. (1996). Specification Testing in Markov Switching Time Series Models. **Journal of Econometrics** v. 70, p. 127-157.

HAMILTON, J. D. e SUSMEL, R. (1994). Autoregressive Heteroskedasticity and Changes in Regime. **Journal of Econometrics** v. 64, p. 307-333.

HANSEN, B. E. (1992). The Likelihood Ratio Test under Non-Standard Conditions: Testing the Markov Switching Model of GNP. **Journal of Applied Econometrics** v. 7, p. 61-82.

HANSEN, L. P. e SINGLETON, K. (1983). Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns. **Journal of Political Economy** v. 91, p. 249-265.

HARVEY, A., RUIZ, E. e SHEPARD, N. (1995). Multivariate Stochastic Variance Models. in **ARCH - Selected Readings** (1995), editor ENGLE, R. Oxford University Press.

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** v. 52(2), p. 241-278.

HOSKING, J.R.M (1981). Fractional Differencing. **Biometrika** v. 68, p. 165-178.

HOTTA, L. K. e TSAY, R. S. (1998). Outliers in GARCH Process. Graduate School of Business, University of Chicago, manuscrito não publicado.

HSIEH, D. Testing for Nonlinear Dependence in Daily Foreign Exchange Rates. **Journal of Business** v. 62, p. 339-368.

KAUFMANN, Sylvia.(2000). Measuring Business Cycle Models with a Dynamic Markov Switching Factor Model: an Assessment Using Bayesian Simulation Methods. **Econometrics Journal** v. 3, p. 39-65.

KIM, C. (1993). Dynamic Linear Models with Markov Switching. **Journal of Econometrics** v. 60, p. 1-22.

KIM, C. (1993b). Unobserved-Component Time Series Models with Markov-Switching

Heteroskedasticity: Changes in Regime and the Link between Inflation Rates and Inflation Uncertainty. **Journal of Business and Economic Statistics** v. 11 341:349.

KIM, C. e KIM, M. (1996). Transient Fads and the Crash of '87. **Journal of Applied Econometrics** v. 11, p. 41-58.

KIM, C. e NELSON, C. R. (1998). Business Cycle Turning Points , New Coincident Index and Tests of Duration Dependence based on a Dynamic Factor Model with Regime Switching. **The Review of Economics and Statistics** v. 82 , p. 188-201.

KIM, C. e NELSON, C. R. (1999). **State Space Models With Regime Switching : Classical and Gibbs-Sampling Approaches with Applications** . MIT Press .

KIM, C. e NELSON, C. R. e STARTZ. R. (1998). Testing for Mean Reversion in Heteroskedastic Data Based on Gibbs-Sampling-Augmented-Randomization. **Journal of Empirical Finance** v. 5, p. 131-154.

KIM, M. e Maddala, G.S (1998). **Unit Roots, Cointegration, and Structural Change**. Cambridge University Press.

KIM, M. e Yoo, J (1995). New index of coincident indicators A multivariate Markov switching factor model approach. **Journal of Monetary Economics** v. 36, p. 607-630.

KIRMAN, A. e TEYSSIERRE, G. (2001). Bubbles and Long RANGE Dependence in Asset Price Volatilities. A sair em **Equilibrium, Markets and Dynamics** editado por HOMMES, C., RAMER, R. e WITHAGEN, C.

KWIATKOWSKI, D., PHILLIPS, P., SCHIMDT, P., e Shin, Y. (1992). "Testing the Null of Stationarity Against the Alternative of a Unit Root. **Journal of Econometrics** v. 54, p. 159-178.

KROLZIG, H. (1997). **Markov-Switching Vector Autoregressions: Modeling, Statistical Inference and Application to Business Cycle Analysis**. Lecture Notes in Economics and Mathematical Systems . Springer.

LAM, P. (1990). The Hamilton Model with a General Autoregressive Component: Estimation and Comparison with Other Models of Economic Time Series. **Journal of Monetary Economics** v. 26, p. 409-432.

LAMOUREAUX, C.G e LASTRAPES, W. (1990). Persistence in Variance, Structural Change, and the GARCH Model. **Journal of Business & Economics Statistics** v. 8:2, p. 224-234.

LEE, T. H. WHITE, H. and GRANGER, C. W. J. (1993): Testing for Neglected Non-linearity in Time Series Models. **Journal of Econometrics** v. 56, p. 269-290.

LEROY, S (1973). Risk Aversion and the Martingale Property of Stock Prices. **International Economic Review** v. 14, p. 436-446.

LINTNER, J. 1965. Security Prices, Risk and the Gains from Diversification. **Journal**

of **Finance** v. 20, p. 587-615.

LJUNQGVIST, L e SARGENT, T. (2000). **Recursive Macroeconomic Theory**. MIT Press.

LO, A. e MACKINLAY, A.C. (1989). The Size and the Power of Variance Ratio Test in Finite Samples: A Monte Carlo Investigation. **Journal of Econometrics** v. 40, p. 203-238.

LUNDBERG, S. e TERASVIRTA, T. (1998). Evaluating GARCH Models. **Working Papers in Economics and Finance** 291. Stockholm School of Economics.

LUTKPOHL, H. (1991). **Introduction to Multiple Series Analysis**. Springer.

MASSON, P. (2000). Exchange Rate Regime Transitions. **IMF Working Paper** 00-134.

MCCULLOCH, R. E e TSAY, R.S. (1994). Statistical Analysis of Economic Time Series via Markov Switching Models, **Journal of Time Series Analysis** v. 15, p. 521-539.

MEESE, R. A. e ROGOFF, K. (1983). Empirical exchange Rate Models of the Seventies: Do They Fit Out of Sample ?. **Journal of International Economics** v. 14, p. 3-24.

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

NELSON, D. (1990). Stationary and Persistence in a GARCH(1,1) Model. in **ARCH - Selected Readings** (1995), editor ENGLE, R. Oxford University Press p. 176-192.

NICHOLLS, D.F. e QUINN, B.G. (1982). **Random Coefficient Autoregressive Models: An Introduction**. Lecture Notes in Statistics 11. Springer-Verlag.

PAGAN, A. R. & SCHWERT, G. W. (1990). Alternative Models for Conditional STOCK Market Volatility. **Journal of Econometrics** v. 45, p. 267-29.

PERRON, P. (1989). The great crash, the oil shock, and the unity root hypotheses. **Econometrica** v. 33, p. 1361-1401.

PHILLIPS, P. e PERRON, P. (1988). Testing for a Unit Root in Time Series Regression. **Biometrika** v. 75, p. 335-346.

POSKITT, D.S e CHUNG, S. H. (1996). Markov Chain Models, Time Series Analysis and Extreme Value Theory. **Adv. Applied Probability** v. 28, p. 405-425.

POTERBA, J. e SUMMERS, L. (1988). Mean Reversion in STOCK Prices: Evidence and Implications. **Journal of Financial Economics** v. 22, p. 27-59.

RABI, L. e VALLS PEREIRA, P. L. (2001). Markovian Switch Models - Applicati-

ons to Financial Time Series - Trabalho em progresso disponível em [http://pvalls. ib-mec.br/artigos/ifj20001.pdf](http://pvalls.ib-mec.br/artigos/ifj20001.pdf).

ROBERTS, H. (1967). **Statistical versus Clinical Prediction of the STOCK Market**. Manuscrito não publicado, Center for Research in Security Prices, University of Chicago.

ROSSI, B. (2000). Testing Out-Of-Sample Predictive Ability with High Persistence. An Application to Models of Nominal exchange Rate Determination. manuscrito.

RYDEN, T. (1995). Estimating the Order of Hidden Markov Models. **Statistics** v. 26, p. 345-354.

RYDEN, T., TERASVIRTA, T. e ASBRINK, S. (1998). Stylized Facts of Daily Return Series and the Hidden Markov switching model. **Journal Of Applied Econometrics** v. 13, p. 217-244.

SAMUELSON, P. (1965). Proof that Properly Anticipated Prices Fluctuate Randomly. **Industrial Management Review** v. 6, p. 41-49.

SAKATA, S. e WHITE. H. (1998). High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility. **Econometrica** v. 66, p. 529-567.

SHARPE. W. Capital Asset Prices. A Theory of Market Equilibrium under Conditions of Risk. **Journal of Finance** v. 19, p. 425-442.

STOCK, J. H. e WATSON, M. (1989). New Indexes of Coincident and Leading Economic Indicators. **NBER Macroeconomic Annual**, eds Blanchard, Oliver e Fisher, Stanley . MIT Press.

TERASVIRTA, T. Lin, C.F e GRANGER, C.W.J (1993): Power of the Neural Network Linearity Test. **Journal of Time Series Analysis** v. 14, p. 209-220.

TSAY, R. S. (1987). Conditional Heteroscedastic Time Series Models. **Journal of the American Statistical Association - Theory and Methods** v. 82, p. 590-604.

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** v. 19:1,p. 57-109.

ZHANG, J. e STINE, R. A. (1997). Autocovariance Structure of Markov Regime Switching Models and Model Selection. **Working Paper, Dept. of Statistics, University of Pennsylvania**.