

**FUNDAÇÃO GETÚLIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO**

Alessandra Gazzoli dos Santos

**Estrutura Fractal em Séries Temporais: uma
Investigação quanto à Hipótese de Passeio Aleatório no
Mercado à Vista de *Commodities* Agrícolas Brasileiro.**

**São Paulo
2013**

ALESSANDRA GAZZOLI DOS SANTOS

**Estrutura Fractal em Séries Temporais: uma
Investigação quanto à Hipótese de Passeio Aleatório no
Mercado à Vista de *Commodities* Agrícolas Brasileiro.**

Dissertação apresentada à Escola de
Economia de São Paulo da Fundação
Getúlio Vargas, como requisito para
obtenção do título de Mestre em
Ciências Econômicas.

Campo de conhecimento:
Finanças Quantitativas

Orientador: Prof. Dr. Juan Carlos
Ruilova Teran.

**São Paulo
2013**

SANTOS, Alessandra Gazzoli.

Estrutura fractal em séries temporais: uma investigação quanto à hipótese de passeio aleatório no mercado à vista de *commodities* agrícolas brasileiro. / Alessandra Gazzoli dos Santos. – 2013.

151 f.

Orientador: Juan Carlos Juilova Teran

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Economia agrícola - Brasil. 2. Análise de séries temporais. 3. Passeio aleatório (Matemática). 4. Fractais. I. Teran, Juan Carlos Juilova. II. Dissertação (MPFE) - Escola de Economia de São Paulo. III. Título.

CDU 631.116

ALESSANDRA GAZZOLI DOS SANTOS

**Estrutura Fractal em Séries Temporais: uma
Investigação quanto à Hipótese de Passeio Aleatório no
Mercado à Vista de *Commodities* Agrícolas Brasileiro.**

Dissertação apresentada à Escola de
Economia de São Paulo da Fundação
Getúlio Vargas, como requisito para
obtenção do título de Mestre em
Ciências Econômicas.

Campo de conhecimento:
Finanças Quantitativas

Data de Aprovação:

____/____/____

Banca Examinadora:

Prof. Dr. Luis Carlos Ruilova Teran
(orientador)
FGV - EESP

Prof. Dr. Alessandro Martim Marques
FGV-EESP

Prof. Dr. Pedro Alberto Morettin
USP-IME

Agradecimentos

Agradeço primeiramente a Deus, por me oferecer tranquilidade e força para enfrentar os desafios com perseverança e coragem.

Um agradecimento especial aos professores Dr. Alessandro Martim Marques e Dr. Juan Carlos Ruilova Teran pelos ensinamentos, pelas claras explicações e especialmente pela paciência e confiança depositadas em mim. A orientação e a amizade de vocês foram essenciais para minha formação como *Quant!*

Agradeço aos meus pais, Maria e Valdemir, pelo apoio incondicional, pela atenção, o carinho e pela motivação incessante que sempre me deram. Serei eternamente grata a vocês por todo o esforço, dedicação e amor para me fazer feliz. Dedico principalmente a vocês o resultado de mais esta conquista.

Ao Luis Antonio, pelo apoio e carinho, pela companhia em tantas idas e vindas para a FGV desde a prova de admissão, pelas conversas e pelos apontamentos sempre muito pertinentes e atenciosos sobre o este trabalho.

Agradeço também aos meus colegas da turma deste Mestrado e do Itaú, que em tantas conversas acabaram contribuindo de forma significativa para meu aprendizado e crescimento.

E por fim, um agradecimento especial também a minha avó, Zulmira, meus tios Beto, Dalva e Luci, minha prima Luciana e as minhas amigas, pela compreensão diante de minhas ausências. Agora tudo volta ao normal!

“The classical theorists resemble Euclidean geometers in a non-Euclidean world who, discovering that in experience straight lines apparently parallel often meet, rebuke the lines for not keeping straight – as the only remedy for the unfortunate collisions which are occurring. Yet in truth, there is no remedy except to throw over the axiom of parallels and to work out a non-Euclidean geometry. Something similar is required today in economics.”

- John Maynard Keynes

RESUMO

As variáveis econômicas são frequentemente governadas por processos dinâmicos e não-lineares que podem gerar relações de dependência de longo prazo e padrões cíclicos não-periódicos com mudanças abruptas de tendências. Para o caso dos preços agrícolas este comportamento não é diferente e as peculiaridades destes mercados podem gerar séries temporais fracionalmente integradas, cujas singularidades não seriam adequadamente capturadas pelos tradicionais modelos analíticos fundamentados na hipótese dos mercados eficientes e de passeio aleatório. Sendo assim, o presente estudo buscou investigar a presença de estruturas fractais no mercado à vista de algumas das principais *commodities* agrícolas brasileiras: café, boi gordo, açúcar, milho, soja e bezerro. Foram empregadas técnicas tradicionais e específicas para a análise de séries temporais fractais como a análise de R/S e a aplicação de modelos das famílias *ARFIMA* e *FIGARCH*. Os resultados indicaram que, com exceção do bezerro, o componente de *drift* destas séries não apresentou comportamento fractal, ao contrário do observado para o componente da volatilidade, que apresentou aspecto de estrutura fractal para todas as *commodities* analisadas.

PALAVRAS-CHAVE: hipótese de mercados eficientes; passeio aleatório; mercados à vista; *commodities* agrícolas; fractais; leptocurtose; memória longa; análise R/S; autossimilaridade; modelos fracionalmente integrados; *ARFIMA*; *FIGARCH*.

ABSTRACT

Economic variables are often governed by dynamic and non-linear processes that can originate long-term relationship and non-periodic and non-cyclical patterns with abrupt trend changes. Commodity prices exhibit this type of behavior and the peculiarities of those markets could generate fractionally integrated time series, whose singularities could not be properly captured by the traditional analytic models based on the efficient market hypothesis and random walk processes. Therefore, this study has investigated the presence of fractal structures on some very important Brazilian commodity spot markets such as coffee, cattle, sugar, soybean and calf. Some traditional techniques were used as well as other specific for fractal time series analysis, such as rescaled range (R/S) analysis, different fractal hypothesis tests and *ARFIMA* and *FIGARCH* models. The results showed that the drift component has not shown fractal behavior, except for the calf series, however, volatility has demonstrated fractal behavior for all the commodities that were analyzed.

KEY WORDS: Efficient market hypothesis; random walk; spot markets; commodities; fractals; leptokurtosis; long-term memory; R/S analysis; self-similarity; fractionally integrated models; ARFIMA; FIGARCH.

Lista de Ilustrações

QUADRO 1 – EXPOENTE DE HURST VS CORRELAÇÃO.....	41
GRÁFICO 1 – EXPOENTE DE HURST VS CORRELAÇÃO	41
GRÁFICO 2 – EXPOENTE DE HURST VS CORRELAÇÃO VS DIMENSÃO FRACTAL.....	45
GRÁFICO 3 – EXPOENTE DE HURST VS CORRELAÇÃO VS DIMENSÃO FRACTAL VS ORDEM DE INTEGRAÇÃO.....	50
QUADRO 2 – VERSÕES DO MODELO DE RANDOM WALK.	54
QUADRO 3 – COEFICIENTE DE CURTOSE	62
QUADRO 4 – COMMODITIES EM ANÁLISE	74
GRÁFICO 4 – ORDEM DE INTEGRAÇÃO ESTIMADA PARA OS LOG-RETORNOS.	80
GRÁFICO 5 – ORDEM DE INTEGRAÇÃO ESTIMADA PARA OS RETORNOS ABSOLUTOS.....	82
GRÁFICO 6 – ORDEM DE INTEGRAÇÃO ESTIMADA PARA OS RETORNOS AO QUADRADO.....	82

Lista de Tabelas

TABELA 1 – RESULTADOS DOS TESTES DE HIPÓTESE PARA OS LOG-RETORNOS.....	76
TABELA 2 – RESULTADOS DOS TESTES DE HIPÓTESE PARA OS RETORNOS ABSOLUTOS.....	78
TABELA 3 – RESULTADOS DOS TESTES DE HIPÓTESE PARA OS RETORNOS AO QUADRADO.	78
TABELA 4 – RESULTADOS DOS TESTES DE HIPÓTESE FRACTAIS PARA OS LOG-RETORNOS.	78
TABELA 5 – RESULTADOS DOS TESTES DE HIPÓTESE FRACTAIS PARA OS RETORNOS ABSOLUTOS	81
TABELA 6 – RESULTADOS DOS TESTES DE HIPÓTESE FRACTAIS PARA OS RETORNOS AO QUADRADO.....	81
TABELA 7 – RESULTADOS DA ESTIMAÇÃO DO COMPONENTE DE DRIFT DOS PROCESSOS	83
TABELA 8 – RESULTADOS DA ESTIMAÇÃO CONJUNTA DO DRIFT E DA VARIÂNCIA DOS PROCESSOS.....	84

Lista de Abreviaturas e Siglas

HME – Hipótese dos Mercados Eficientes

HMF – Hipótese dos Mercados Fractais

CEPEA-ESALQ/USP – Centro de Estudos Avançados em Economia Aplicada da Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo.

ACF – Função de Autocorrelação

PACF – Função de Autocorrelação Parcial

AR – Processo Autorregressivo

ARMA – Processo Autorregressivo com Médias Móveis

ARIMA – Processo Autorregressivo Integrado com Médias Móveis

ARFIMA – Processo Autorregressivo Fracionalmente Integrado com Médias Móveis

GARCH – Autorregressivos com Heterocedasticidade Condicional Generalizado

EGARCH – GARCH Exponencial

FIGARCH – GARCH Fracionalmente Integrado

FIEGARCH - EGARCH Fracionalmente Integrado

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	OBJETIVOS	17
1.2	ESTRUTURA DA DISSERTAÇÃO.....	18
2	A EFICIÊNCIA DOS MERCADOS E HIPÓTESE DE PASSEIO ALEATÓRIO.....	19
2.1	O MODELO DE <i>RANDOM WALK</i>	20
2.1.1	O MOVIMENTO BROWNIANO	23
2.2	A PESQUISA EMPÍRICA SOBRE O COMPORTAMENTO DOS PREÇOS DOS ATIVOS.....	25
2.3	A HIPÓTESE DOS MERCADOS EFICIENTES	27
2.4	ALGUNS ESTUDOS NACIONAIS SOBRE A HIPÓTESE DO PASSEIO ALEATÓRIO	35
3	A MEMÓRIA LONGA E A HIPÓTESE DOS MERCADOS FRACTAIS.....	38
3.1	ESTRUTURAS FRACTAIS.....	38
3.2	SÉRIES TEMPORAIS FRACIONÁRIAS.....	46
4	METODOLOGIA	53
4.1	A VERIFICAÇÃO DA HIPÓTESE DE PASSEIO ALEATÓRIO	53
4.1.1	A ESTACIONARIEDADE DA SÉRIE	54
4.1.1.1	ANÁLISE DA FUNÇÃO DE AUTOCORRELAÇÃO	55
4.1.1.2	ANÁLISE DA INDEPENDÊNCIA	58
4.1.1.3	ANÁLISE DE RAIZ UNITÁRIA	59
4.1.2	A NORMALIDADE	61
4.1.3	A LINEARIDADE DA SÉRIE.....	63
4.2	A IDENTIFICAÇÃO DE UMA SÉRIE TEMPORAL FRACIONALMENTE INTEGRADA.....	65
4.2.1	A ESTATÍSTICA R/S E A DETERMINAÇÃO DO EXPOENTE DE HURST.....	65
4.2.2	A MENSURAÇÃO DA DIMENSÃO FRACTAL.....	68
4.2.3	MÉTODO GPH.....	68
4.2.4	MÉTODO DO PERIODOGRAMA	68
4.2.5	MÉTODO DE WHITTLE.....	69
4.3	ESTIMAÇÃO DE MODELOS COM INTEGRAÇÃO FRACIONÁRIA.....	70
5	RESULTADOS E CONCLUSÕES	74
5.1	A VERIFICAÇÃO DA HIPÓTESE DE PASSEIO ALEATÓRIO.....	74
5.2	A VERIFICAÇÃO DA HIPÓTESE DE SÉRIES TEMPORAIS FRACIONALMENTE INTEGRADAS	78
5.3	ESTIMAÇÃO DOS MODELOS	83
5.4	OBSERVAÇÕES FINAIS E SUGESTÕES PARA ESTUDOS FUTUROS	85
6	REFERÊNCIAS BIBLIOGRÁFICAS.....	88
7	ANEXOS.....	94
7.1	ESTATÍSTICAS DESCRITIVAS	94
7.2	ESTIMATIVAS DA ORDEM DE INTEGRAÇÃO	95
7.3	PREÇOS (SÉRIES EM NÍVEL)	96
7.4	LOG-RETORNOS (SÉRIES EM PRIMEIRA DIFERENÇA).....	96
7.5	HISTOGRAMA DOS LOG-RETORNOS (SÉRIES EM PRIMEIRA DIFERENÇA)	97
7.6	GRÁFICOS QXQ DOS LOG-RETORNOS (SÉRIES EM PRIMEIRA DIFERENÇA)	97
7.7	RETORNOS ABSOLUTOS	98
7.8	RETORNOS AO QUADRADO	98
7.9	FUNÇÕES DE AUTOCORRELAÇÃO (ACF) DAS SÉRIES DE PREÇOS	99
7.10	FUNÇÕES DE AUTOCORRELAÇÃO (ACF) DAS SÉRIES DE LOG-RETORNOS	99

7.11	FUNÇÕES DE AUTOCORRELAÇÃO (ACF) DOS RETORNOS ABSOLUTOS.....	100
7.12	FUNÇÕES DE AUTOCORRELAÇÃO (ACF) DOS RETORNOS AO QUADRADO	100
7.13	FUNÇÕES DE AUTOCORRELAÇÃO PARCIAL (PACF) DAS SÉRIES DE PREÇOS	101
7.14	FUNÇÕES DE AUTOCORRELAÇÃO PARCIAL (PACF) DAS SÉRIES DE LOG-RETORNOS	101
7.15	FUNÇÕES DE AUTOCORRELAÇÃO PARCIAL (PACF) DOS RETORNOS ABSOLUTOS.....	102
7.16	FUNÇÕES DE AUTOCORRELAÇÃO PARCIAL (PACF) DOS RETORNOS AO QUADRADO	102
7.17	RESULTADOS DOS TESTES DE RAZÃO DE VARIÂNCIA	103
7.18	OUTPUT DA ESTIMAÇÃO DE UM MODELO ARFIMA (0,d,0).....	109
7.18.1	SÉRIES DE LOG-PREÇOS	109
7.18.2	SÉRIES DE LOG-RETORNOS	111
7.18.3	SÉRIES DE RETORNOS ABSOLUTOS.....	113
7.18.4	SÉRIES DE RETORNOS AO QUADRADO	116
7.19	RESULTADOS DA ESTIMAÇÃO DO COMPONENTE DE <i>DRIFT</i> (LOG-RETORNOS)	119
7.20	OUTPUT DA ESTIMAÇÃO CONJUNTA DO COMPONENTE DE <i>DRIFT</i> E DA VARIÂNCIA	122
7.20.1	BOI GORDO	122
7.20.2	BEZERRO.....	127
7.20.3	CAFÉ	132
7.20.4	AÇÚCAR	137
7.20.5	MILHO.....	142
7.20.6	SOJA (REGIÃO DO PARANÁ).....	147

1 Introdução

Os esforços iniciais de pesquisa no âmbito da teoria econômica visando sua inserção no hall das Ciências (como a Física e a Matemática) estiveram fundamentados nos procedimentos analíticos desenvolvidos e consagrados por Descartes, Galileu e Newton, que compõem o que se convencionou denominar *método reducionista clássico ou mecanicismo*.

Segundo este referencial, o *todo* representativo do objeto de estudo seria formado por uma somatória de suas partes constituintes, de modo que a análise destas permitiria a compreensão do objeto em sua totalidade. Neste sentido, a análise científica deveria ser realizada em duas etapas: a primeira envolveria um exercício de decomposição deste *todo* em suas partes menores tal que estas pudessem ser individual e profundamente analisadas, e a segunda, por sua vez, envolveria nova agregação destes elementos previamente decompostos de modo a permitir o entendimento do objeto de estudo por completo.

A aplicação do método reducionista clássico nas análises do sistema econômico resultava assim, na necessidade de que o este fosse didaticamente subdividido em termos de suas partes menores, tais como os agentes econômicos (indivíduos, firmas, governo) institucionais ou privados e também o ambiente em que estes agentes se relacionam entre si (ambiente no qual estes agentes assumem papéis bem definidos de ofertantes e/ou demandantes). A partir disto procedia-se a compreensão detalhada do comportamento de cada agente individualmente, assim como do ambiente de mercado em que estas inter-relações entre os agentes ocorrem, tal que a agregação posterior destes elementos já previamente analisados pudesse propiciar uma compreensão aprofundada do sistema em sua totalidade.

Observa-se assim, que na perspectiva da abordagem reducionista clássica, pouca relevância era concedida ao entendimento da dinâmica própria destas inter-relações, na medida em que se supunha que as relações de causa e efeito que perpassam das partes componentes para o *todo* sistêmico ocorrem de maneira aproximadamente linear e que por isso podem ser expressas através de uma agregação simples. Neste sentido, as mudanças de comportamento das partes causariam efeitos linearmente proporcionais no *todo* sistêmico que as engloba (posto que do contrário, a totalidade do objeto de estudo não seria resultante da mera somatória de suas partes constituintes).

Foi assim, sob esta perspectiva, que surgiram as bases da Economia como ciência, visando o enquadramento do sistema econômico no contexto dos demais sistemas naturais (objeto de estudo da Matemática e da Física, por exemplo). Diante disto, a apreensão do conceito de sistema em equilíbrio destas ciências permitiu que fosse construído um arcabouço teórico que ainda serve de base para as principais escolas de pensamento econômico mundiais e que se fundamenta no entendimento da Economia como um sistema que tende constantemente para uma situação de equilíbrio. Neste sentido, a interação dos agentes nos mercados apesar de gerar desequilíbrios momentâneos não seria capaz de alterar esta tendência de que as forças de mercado se reajustassem em direção a um novo equilíbrio, uma vez que esta característica seria intrínseca ao próprio sistema.

No entanto, conforme ressalta Bertalanffy (1969), alguns sistemas, como o econômico, caracterizam objetos de investigação científica tais que o entendimento de suas partes isoladamente não permite a compreensão do todo sistêmico.

Sendo assim, conforme ressalta Beinhocker (2006) esta compreensão advinda do reducionismo clássico, que fundamentou o que chamamos de ciência moderna e embasa o pensamento de grande parte dos estudiosos do atual *mainstream economics*, ao colaborar para que se explicasse cada fenômeno por meio da redução em seu sentido estrito¹ teria criado um desafio para a ciência contemporânea. Este desafio consistiria em passar a entender determinados *todos sistêmicos* tanto por meio da compreensão de suas partes componentes quanto, e principalmente, a partir das características próprias de cada sistema que emergem justamente das relações entre suas partes componentes. Desta maneira, o desafio está no entendimento do sistema em sua totalidade considerando essencialmente a forma como suas partes constituintes estão organizadas e interagem entre si no ambiente sistêmico, levando em conta, portanto, sua dependência e principalmente as conseqüências destas inter-relações.

Em especial a partir dos anos 1980, muitos autores² passaram a se dedicar a estas questões, e seus estudos indicam que uma forma mais adequada de caracterização do sistema econômico seria não como um sistema que se direciona ao equilíbrio, mas, sobretudo, como um sistema complexo, marcado por desequilíbrios e choques e que por

¹ Indicando cada objeto e seus respectivos elementos componentes como unidades separadas e independentes.

² Tais como BAK, (1988), ANDERSON P., K. J. e K. ARROW, D. PINES, Eds. (1988), JOHNSON, N. F. (2009), entre outros.

isso talvez exigiriam, um instrumental de análise mais adequado a suas características peculiares do que os tradicionalmente utilizados.

Segundo Mandelbrot e Hudson (2004), as teorias tradicionais de finanças (pensamento ortodoxo que remonta à tese de Bachelier de 1900 e fundamenta modelos como o de Black & Scholes, o CAPM e suas derivações) falham ao capturar as singularidades (fatos estilizados³) presentes em séries temporais representativas deste sistema. Além disto, acabam produzindo modelos cujos resultados se mostram muitas vezes inadequados para descrever a realidade deste sistema complexo marcado pela influência de diversas forças e agentes heterogêneos.

Neste sentido, modelos que consideram esta complexidade dos agentes e estruturas dos mercados componentes do sistema econômico, assim como as conexões não lineares entre os acontecimentos que permeiam este ambiente seriam capazes de capturar as influências dos ciclos de oferta e demanda sobre os preços, que implicam em situações de dependências e reversões constantes nos preços.

Muitos dos estudos nesta linha de pesquisa tem se concentrado na avaliação do comportamento dos ativos financeiros mais tradicionais, como ações negociadas em bolsa, índices de ações, taxas de inflação, de câmbio, entre outros, enquanto que os estudos direcionados ao comportamento dos mercados agrícolas como sistemas complexos são ainda restritos, em especial se considerados somente os estudos no mercado brasileiro. Todavia, a análise e previsão do comportamento dos preços nos mercados de *commodities* agrícolas são relevantes não somente para os agentes que se utilizam destes mercados de maneira direta como os agentes da cadeia produtiva do agronegócio (produtores, cooperativas, indústrias entre outros), como também no que se refere a sua representatividade nos mercados financeiros mundiais.

Conforme destacam Jin e Frechette (2002), as variáveis econômicas são frequentemente governadas por processos dinâmicos e não-lineares e que podem gerar relações de dependência de longo prazo e padrões cíclicos não-periódicos com mudanças abruptas de tendências. Para o caso dos preços agrícolas este comportamento não é diferente, na medida em que estes são determinados por fatores naturais e socioeconômicos que são conhecidos pela sua autossimilaridade em diferentes escalas e por seguirem movimentos cíclicos não-periódicos.

³ Tais como a não-normalidade dos retornos, a existência de caudas largas, as respostas assimétricas dos retornos diante de retornos negativos e positivos, a existência de heterocedasticidade, entre outros.

Segundo os autores, a presença destas características acaba gerando séries temporais com singularidades como a existência de estrutura fractal e a observação de uma ordem de integração fracionária, características que podem ser modeladas por técnicas, como a análise de R/S, os modelos autorregressivos fracionalmente integrados e os modelos fracionalmente integrados da família *GARCH*.

No entanto, a maior parte dos estudos nacionais sobre *commodities* agrícolas utilizou de técnicas de modelagem fundamentadas na hipótese de que estes processos são regidos por um passeio aleatório. Além disso, outros modelos com tendências e ciclos também foram propostos. Todavia, Jin e Frechette (2002) destacam que os modelos que permitem a modelagem destas séries como processos fracionalmente integrados podem ser mais adequados para capturar estas características particulares.

Conforme ressalta Barkoulas, Labys e Onochie (1997), a existência de estrutura fractal nos preços dos ativos levanta questões relacionadas aos testes estatísticos tradicionais baseados em métodos padronizados, modelagem econométrica teórica e linear dos preços dos ativos, assim como a precificação da eficiência dos mercados e da racionalidade.

Os trabalhos mais recentes sobre a investigação da dinâmica fractal estão ligados diretamente à avaliação da existência de memória de longo prazo⁴ nos retornos de ações e outros ativos e derivativos financeiros, sendo pouca atenção fornecida ao estudo do comportamento dos preços e retornos das *commodities*. Barkoulas, Labys e Onochie (1997) indicam que a depender do valor que se observa em termos da ordem de integração, d , do processo, este pode ser caracterizado como um processo estacionário fracionalmente integrado com memória longa, com memória curta, ou até mesmo não estacionário. A detecção de memória de longo prazo indica a existência de uma estrutura de dependência temporal entre as observações na série, que podem configurar padrões cíclicos não periódicos que caracterizam uma forma especial de dinâmica não linear e que conforme Lo (1991) pode ser problemática na medida em que é inconsistente com os métodos que tratam de processos estocásticos contínuos aplicados aos métodos martingales de precificação de contratos futuros e de opções, o que ressalta a necessidade de análise cuidadosa para identificação desta característica.

⁴ Segundo Jin e Frechette (2002), a memória longa em séries temporais caracteriza-se como uma forma especial de dinâmica não linear e a existência de processos fracionalmente integrados pode dar origem a memória longa em séries temporais.

Os elementos mais característicos das séries de *commodities* agrícolas seriam, segundo Jin e Frechette (2002), a existência de autocorrelação positiva (períodos com altos preços tendem a ser seguidos por períodos com altos preços, assim como períodos de baixos preços tendem a ser seguidos por períodos com preços baixos); a presença de picos (períodos com fortes saltos de preços bem distantes da média histórica, seja para cima ou para baixo). Segundo os autores, este comportamento sinalizaria a existência de heterogeneidade em termos dos horizontes de investimentos dos agentes, que são considerados pelos modelos específicos de séries temporais fracionárias e que não são refletidos quando se assume que o processo que rege estas séries é um passeio aleatório.

Além disto, também os choques fractais na oferta e demanda das *commodities* agrícolas podem induzir estruturas fractais nos preços que não seriam capturadas pelo modelo de passeio aleatório tradicional (Jin e Frechette, 2002).

Neste sentido, o presente estudo busca colaborar para o preenchimento da lacuna existente na literatura nacional direcionada à avaliação dos mercados de *commodities* agrícolas nacional, visando colaborar para os avanços na identificação de modelos cada vez mais adequados para o tratamento de peculiaridades de cada mercado.

1.1 Objetivos

O presente estudo tem como objetivo investigar a presença de estruturas fractais no mercado à vista de algumas das principais *commodities* agrícolas brasileiras: café, boi gordo, açúcar, milho, soja (região do Paraná) e bezerro. Neste sentido, buscou-se avaliar a hipótese de que o mercado de *commodities* agrícolas pode ser modelado por um processo do tipo passeio aleatório (*random walk*) em comparação à hipótese da existência de comportamento fractal.

Para tanto foram utilizados os indicadores diários de preços à vista calculados pelo Centro de Estudos em Economia Aplicada (CEPEA), da Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP).

Esta análise procura detectar se as séries temporais em estudo podem ser caracterizadas como séries temporais fracionalmente integradas, o que se acredita, pode exigir um tratamento diferenciado em sua modelagem, assim como impactar os tradicionais modelos analíticos fundamentados na hipótese dos mercados eficientes.

Por meio de testes de correlação serial, de raiz unitária, de testes portmanteau⁵, de razão de variância, entre outras, o presente estudo buscou avaliar a hipótese de passeio aleatório (muitas vezes confundida como a Hipótese dos Mercados Eficientes) em séries temporais de mercado à vista das *commodities* agrícolas supracitadas.

Já a análise da estrutura fractal destes mercados será realizada por meio da utilização de técnicas mais adequadas para séries fracionalmente integradas, como a análise de R/S de Hurst, os métodos GPH e do periodograma e de Whittle. Além disto, serão brevemente analisados alguns modelos fracionalmente integrados, como os autorregressivos fracionalmente integrados (ARFIMA) e os modelos GARCH fracionalmente integrados. Para tanto foram utilizadas series log-retornos, retornos absolutos e retornos ao quadrado, sendo estas duas últimas analisadas como *proxies* do comportamento da variância dos retornos do mercado à vista.

1.2 Estrutura da Dissertação

O primeiro capítulo buscou introduzir brevemente o ambiente que envolve as pesquisas em séries temporais fracionárias e ressaltar a relevância do tratamento adequado de séries como estas. Já o segundo capítulo busca descrever a Hipótese dos Mercados Eficientes (HME) e a coincidência que há na literatura em termos da admissão desta como sinônimo da premissa de que os ativos podem ser regidos por processos do tipo passeio aleatório. No capítulo três serão discutidos alguns aspectos característicos de estruturas fractais e de séries temporais fracionalmente integradas. Na sequência, no capítulo quatro, serão descritas as metodologias de análise normalmente empregadas em análises tradicionais dos mercados e algumas voltadas para a análise de séries temporais fracionalmente integradas e que serão empregadas neste trabalho.

No capítulo cinco seguem então os resultados, conclusões e sugestões para estudos futuros.

⁵ Um teste de hipóteses Portmanteau é aquele em que enquanto a hipótese nula é bem especificada, a hipótese alternativa é determinada de maneira mais frouxa. Em análises de séries temporais, testes como o de Box e Pierce (1970) e Ljung e Box (1978), que serão abordados no capítulo quatro, podem ser caracterizados como tal.

2 A eficiência dos mercados e hipótese de passeio aleatório

O conceito de eficiência é central para os estudos em finanças e pode descrever dois aspectos distintos analisados em economia: eficiência em termos operacionais, cuja ênfase está na análise da melhor alocação dos recursos escassos (eficiência de Pareto) e, eficiência em termos informacionais. Esta última abordagem serve para caracterizar a Hipótese dos Mercados Eficientes (HME) e é utilizada para descrever um mercado em que informações relevantes são corretamente apreendidas pelos preços.

Segundo Lo e MacKinlay (1997), as origens da Hipótese dos Mercados Eficientes podem ser rastreadas desde a contribuição pioneira de Bachelier (1900) e os estudos empíricos de Cowles (1933).

De maneira inovadora para sua época, Bachelier (1900) buscou avaliar o mercado financeiro baseando-se em uma analogia entre o movimento de uma partícula em um fluido e o comportamento dos preços dos títulos públicos na bolsa de Paris.

Esta analogia se deveu à recuperação de um estudo de 1827 do biólogo escocês Robert Brown, no qual ele analisou ao microscópio a difusão de grânulos de pólen suspensos em um fluido e identificou que estas finas partículas se moviam rapidamente com incrementos altamente irregulares que eram estimulados tanto pelo calor quanto pela diminuição da viscosidade do fluido⁶ (WIERSEMA, U. F., 2008). Posteriormente aos relatos de Brown, outros autores identificaram que os movimentos irregulares observados pelo autor seriam justificados por um altíssimo número de colisões entre as moléculas de pólen e as do fluido (menores que as do pólen) que aconteceriam de maneira irregular, em intervalos muito curtos de tempo e de maneira independente umas às outras.

Bachelier, ao analisar os movimentos dos preços dos títulos públicos na bolsa de Paris identificou também incrementos altamente irregulares como os observados por Brown. A partir disto, Bachelier desenvolveu então a primeira especificação matemática para o movimento das partículas relatado por Brown e utilizou-a como um modelo para descrever também o movimento dos preços dos ativos analisados (WIERSEMA, U. F., 2008).

⁶ A influência da viscosidade e do calor foi um tópico estudado de forma matematicamente mais precisa por Albert Einstein em seu artigo de 1905.

Partindo da ideia central de que as variações dos preços perfariam movimentos que podem ser descritos pelas leis da probabilidade e poderiam, portanto, ser avaliadas por meio de equações e cálculos probabilísticos, Bachelier considerou os mercados financeiros como um ambiente de “jogo honesto” em que poderiam ocorrer ganhos ou perdas diante de movimentos de subida ou descida dos preços. Cada ocorrência de um movimento de subida e/ou descida seria equiprovável, assim como as chances de serem obtidos ganhos ou perdas no jogo. Cada intervalo de tempo em que ocorreria um movimento representaria uma jogada e cada jogada ocorreria de maneira independente em relação às demais (anteriores ou posteriores) não havendo, portanto, memória entre as jogadas sucessivas.

Bachelier admitiu ainda que estas variações dos preços poderiam ser grandes ou pequenas, mas que na ausência de informações externas que promovessem alterações acentuadas, estas flutuações tenderiam a ocorrer em torno de seu valor inicial.

O autor admitiu ainda que eventos passados, presentes e até mesmo futuros descontados são refletidos nos preços de mercado, mas em geral não apresentam relação aparente com mudanças de preços. Sendo assim, afirmou que apesar de o mercado não prever de fato suas flutuações ele ao menos pode indicar se determinadas flutuações são ou muito ou pouco prováveis e esta probabilidade pode ser calculada em termos matemáticos.

Como ressaltam Dimson e Mussavian (2000) esta percepção levou Bachelier a uma análise que antecipou em grande medida a derivação da versão Einstein-Wiener do movimento Browniano assim como diversos resultados analíticos redescobertos pelos estudiosos em finanças somente na segunda metade do século XX.

2.1 O modelo de *random walk*

Cinco anos após a tese de Bachelier, Karl Pearson (1905) introduziu o termo *random walk* (ou passeio aleatório). Em seu artigo, Pearson analisava o problema do procedimento ótimo de busca para um bêbado abandonado no meio de um campo. Sua observação foi de que, assumindo-se que o bêbado cambaleie de modo totalmente aleatório e imprevisível, é provável que ele termine em um local mais próximo ao qual ele começou a andar do que em qualquer outro ponto.

Para o caso do andar de um bêbado (muitas vezes citado como um caminhante aleatório), a descrição mais simples de um passeio aleatório em uma dimensão segue três regras: trata-se de um caminho percorrido no qual há um ponto inicial, a distância

entre um passo e outro percorrido possui um comprimento l constante e a cada passo é possível estabelecer uma nova direção para o próximo passo (direita ou esquerda) de forma arbitrária.

Neste contexto conforme Silva e Lima (2006), admite-se n_1 como o total de passos para a direita que o caminhante pode realizar ao longo de um intervalo de tempo e n_2 o total de passos para a esquerda. Define-se também p como a probabilidade de que, a cada novo instante, a direção do passo seja a direita, tal que será dada por p e conseqüentemente a probabilidade de um passo para a esquerda será dada por

$$q = 1 - p \quad (2.1)$$

indique a probabilidade de que a direção do passo seja à esquerda.

Ao passar de um tempo, o caminhante terá realizado um total de N passos, sendo n_1 para direita⁷, e n_2 para a esquerda, tal que a distância percorrida pelo caminhante pode ser dada por:

$$m = l \times (n_1 - n_2). \quad (2.2)$$

Admitindo-se que os passos são estatisticamente independentes a determinação da distância percorrida implica na determinação de quantas trajetórias são possíveis de ocorrer. Isto é o mesmo de definir quais as maneiras possíveis de arranjar os n_1 passos para a direita e n_2 passos para a esquerda de tal forma que as permutações entre eles seja irrelevante. Isto implica que o número de possibilidades distintas pode ser dado por:

$$\frac{N!}{n_1!n_2!} \quad (2.3)$$

Sendo assim, como todas as sequências são independentes, a probabilidade de que o caminhante dê n_1 passos para a direita e n_2 passos para a esquerda em um total de N passos em qualquer ordem é dada por:

$$P_N(n_1) = \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} \quad (2.4)$$

Nota-se assim que $P_N(n_1)$ segue uma distribuição binomial, cuja expansão pode ser representada por:

⁷ Considerando-se a perspectiva da representação da trajetória que o caminhante deseja realizar a partir das coordenadas cartesianas, admite-se que a linha reta possa ser descrita pelo eixo x , de tal forma que os passos dados para a direita possam ser considerados como movimentos acima do eixo x , ou seja, com valores positivos. Desta forma, os passos realizados para a esquerda são considerados para efeitos de cálculo como movimentos abaixo do eixo x , apresentando assim, valores negativos.

$$(p + q)^N = \sum_{n_1=0}^N \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} \quad (2.5)$$

$$(p + q)^N = 1 = \sum_{n_1=0}^N \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} \quad (2.6)$$

$$\sum_{n_1=0}^N P_N(n_1) = 1. \quad (2.7)$$

Desta forma é possível dizer que a distribuição de $P_N(n_1)$ neste molde é dita normalizada.

A partir da equação (2.4) acima, para a determinação da probabilidade de que o caminhante se encontre na posição m :

$$P_N(m) = \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} p^{\frac{N+m}{2}} q^{\frac{N-m}{2}} \quad (2.8)$$

A percepção de que o próximo passo de um caminhante aleatório apresenta a mesma probabilidade de que seja para a direita ou para a esquerda é análogo ao entendimento de Bachelier de que o preço dos ativos no mercado financeiro pode, a cada instante de tempo, realizar de forma equiprovável um movimento de subida ou descida. Desta forma, ambos os eventos poderiam ser modelados como um processo que descrevesse uma sequência de eventos de um jogo justo.

Esta analogia entre o mercado financeiro como um ambiente de jogo justo permitiu ainda o desenvolvimento de teorias sobre os processos aleatórios que regem os movimentos de preços de mercado baseadas em uma semelhança com o modelo do tipo *random walk* e os fenômenos de difusão de partículas.

Em termos matemáticos, admitir que o movimento dos preços dos ativos segue um processo estocástico do tipo *random walk* implica afirmar que, em tempo discreto, o processo pode ser descrito, segundo Shreve (2004) a partir da definição de uma variável aleatória X , baseada no lançamento de uma moeda, tal que:

$$M_n = \sum_{j=1}^n X_j \quad n \geq 1 \quad (2.9)$$

$$\text{com } M_0 = 0$$

Onde:

$$X_j = \begin{cases} 1 & \text{se } \omega_j = \textit{cara} \\ -1 & \text{se } \omega_j = \textit{coroa} \end{cases}, \quad (2.10)$$

sendo que ω representa uma série de lançamentos de uma moeda tal que a probabilidade de sair qualquer uma das faces, cara ou coroa, é a mesma em todos os lançamentos, e igual a $\frac{1}{2}$, portanto:

$$\begin{aligned} & (\omega \in \Omega \mid \omega = \omega_1 \omega_2 \omega_3 \dots) \\ & \begin{cases} p = P(\omega_j = \textit{cara}) = \frac{1}{2} \\ q = P(\omega_j = \textit{coroa}) = \frac{1}{2} \end{cases} \end{aligned} \quad (2.11)$$

Sendo assim, a cada lançamento da moeda, a variável aleatória X_j pode indicar que o processo estocástico M_n realize um movimento de subida, quando o lançamento da moeda resultar em cara e logo $X_j = 1$, ou descida, quando o lançamento da moeda resultar em coroa e assim $X_j = -1$.

Sob estas condições, o processo estocástico M_n é denominado passeio aleatório simétrico (já que as probabilidades de subida e descida são iguais) e em tempo discreto.

Até este ponto, a descrição do movimento dos preços dos ativos por meio da analogia com o jogo justo se assemelha em muito à descrição por meio da interpretação sobre o andar do caminhante aleatório. Todavia, como ressaltam Silva e Lima (2006), uma conexão com o fenômeno de difusão de uma partícula em um fluido tal como descrita por Einstein (1905) e Pearson (1905) exige a descrição do problema do caminhante aleatório considerando variáveis aleatórias contínuas.

2.1.1 O movimento Browniano

Em 1920, Norbert Wiener, físico-matemático do MIT, apresentou em um congresso de matemática em Strasbourg um artigo em que desenvolveu o arcabouço probabilístico rigoroso por trás do modelo de Bachelier, denominado atualmente de Movimento Browniano ou Processo de Wiener.

Wiener utilizou o ferramental da física estatística para analisar o processo de difusão das partículas em um fluido, com ênfase na avaliação da posição provável de uma partícula num dado instante de tempo. Em seu estudo Wiener observou que os

incrementos nos movimentos das partículas possuíam valores contínuos, eram independentes uns dos outros (se considerados em momentos disjuntos de tempo) e não apresentavam uma direção definida, de modo que a distribuição dos incrementos apresentava média zero. Admitiu ainda que o deslocamento das partículas era provocado por um somatório de choques aleatórios independentes e que a dispersão no valor da posição final da partícula na superfície analisada seria proporcional ao tempo decorrido.

Diante disto, considerou que a trajetória de uma partícula em um fluido, formada pela sequência dos incrementos, se configuraria como uma trajetória definida por um conjunto de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) que, sendo geradas por um mesmo processo de ruído aleatório em intervalos de tempo contínuo (e não mais discretos), poderia ser descrita pelo modelo que recebeu o nome de movimento Browniano.

Desta forma, a transposição da análise em tempo discreto do jogo justo para o tempo contínuo característico do movimento Browniano envolve um artifício de considerar a partição da variável indicativa do tempo em intervalos cada vez menores de modo que estes tendessem a zero. Isto garante que particionando cada vez mais a dimensão temporal a dimensão espacial dos choques também tenda a zero, de tal maneira que um exemplo em tempo discreto se aproxime da realidade de uma análise em tempo contínuo.

Considerando as equações (2.9), (2.10) e (2.11) acima e admitindo-se uma partição do intervalo temporal dada por $0 = t_0 < t_1 < \dots < t_n$ tal que cada nt_j seja um inteiro, define-se um passeio aleatório escalado como:

$$\omega^n(t) = \frac{1}{\sqrt{n}} M_t, \quad n > 0 \quad e \quad t \geq 0. \quad (2.12)$$

Shreve (2004) demonstra que o cálculo da função geradora dos momentos deste processo estocástico resulta, quando $n \rightarrow \infty$ (caso limite do particionamento do intervalo temporal), na função geradora dos momentos de uma distribuição normal, dada por:

$$\varphi(u) = e^{u^2/2}. \quad (2.13)$$

Sendo assim, para o caso limite do particionamento do intervalo temporal este processo estocástico converge em distribuição para uma normal, de modo que o passeio aleatório escalado torna-se o movimento Browniano.

2.2 A pesquisa empírica sobre o comportamento dos preços dos ativos

Na década de 1950, a utilização dos computadores passou a estar mais disponível para fins de pesquisa empírica, o que colaborou para os avanços nos campos de finanças e nos estudos do comportamento das séries de mercado.

Por volta desta época, em especial a partir da publicação do trabalho seminal de Markowitz (1952), parte dos pesquisadores acreditava que as análises de séries temporais apresentavam não somente componentes aleatórios (porção residual das oscilações de curto prazo) como também tendências. Para estes autores era possível extrair os movimentos de longo prazo (tendências) para uma análise individual e então, em separado, examinar a porção residual das oscilações de curto prazo e suas variações aleatórias.

Em 1953, todavia, Kendall analisou 22 séries de preços (de ações e de *commodities*) do mercado britânico em intervalos semanais e descobriu que suas variações eram essencialmente aleatórias, de modo que a metodologia de quebra da série temporal entre componentes aleatórios e não aleatórios não seria condizente com a característica dos dados. Além disso, Kendall (1953) notou também que a variância dos ativos era não-estacionária (já que era observada dependência na variância empírica). Sua conclusão foi de que em séries de preços analisadas em intervalos curtos de tempo, as mudanças aleatórias entre um período e outro são tão grandes que se sobrepõem a quaisquer efeitos sistemáticos que podem estar presentes, de tal forma que os dados se comportam quase como séries errantes.

Entretanto, a correlação serial das mudanças de preços próximas a zero (característica de processos aleatórios) era uma observação que parecia inconsistente com as observações empíricas. Todavia, adotou-se o processo de *random walk* para a descrição destas mesmas observações empíricas.

Seguindo os caminhos trilhados pelas análises de Working (1934) e Kendall (1953), Roberts (1959) demonstrou que uma série temporal gerada por uma sequência de números aleatórios era muito semelhante a uma série temporal de preços de ações norte-americanas, o que indicava a seu ver, que a série temporal das ações norte-

americanas poderia também ser expressa por um processo do tipo *random walk*, o que implicava que a previsão de valores futuros dos ativos não seria uma atividade possível.

Simultaneamente, Maury Osborne (1959) encontrou evidências de que os preços das ações não seguiam uma distribuição normal, ao contrário da sugestão advinda dos estudos de Bachelier de que estes seguiriam um Movimento Browniano. O autor identificou que os preços apresentavam propriedades próximas ao movimento das moléculas, mas que era a taxa de retorno de um ativo (o percentual médio de variação dos preços a cada instante) que seguiam uma distribuição normal. Como os preços e a taxa de retorno das mudanças de preços são relacionados por uma relação logarítmica, aplicando os métodos da mecânica estatística aos mercados de ações, as observações de Osborne indicavam que os preços de uma ação deveriam apresentar uma distribuição log-normal.

Além disto, Osborne encontrou evidências sobre a chamada regra da raiz quadrada do tempo, que indica que, para este modelo, a dispersão dos retornos aumenta segundo a proporção da raiz quadrada do tempo decorrido, assim como sinais de uma grande “dispersão tangencial” (presença de caudas longas) nos limites das distribuições dos retornos. Na mesma linha, Larson (1960) apresentando os resultados de um novo método de análise de séries temporais, observou que a distribuição das variações de preços se aproximava bastante de uma distribuição normal para os 80% centrais dos dados, todavia apresentavam um excesso de retornos extremos. Ambas as análises de Osborne (1959) e Larson (1960) encontraram uma característica que futuramente se convencionou denominar *fat tails*, ou caudas pesadas, indicando que as distribuições empíricas eram leptocúrticas e, portanto, diferentes da gaussiana.

Além destes estudos, como destacam Dimson e Mussavian (2000), apesar da emergência de evidências sobre o caráter aleatório dos preços, outros como de Working (1934), Cowles e Jones (1937) e Kendall (1953) haviam encontrado subconjuntos de informações em que os preços pareciam se comportar de maneira previsível e não imprevisível como em um *random walk*.

Em detrimento destas análises divergentes, em 1964 Cootner publicou uma coleção de artigos referentes às pesquisas relacionadas ao caráter aleatório dos preços dos ativos e no ano seguinte, a dissertação de doutorado de Fama foi reproduzida por completo na revista *Journal of Business*. Neste trabalho Fama havia revisado a literatura existente referente ao comportamento dos preços de mercado, examinado as distribuições de diferentes ativos e avaliado a dependência serial dos retornos. Sua

conclusão foi de que havia evidências suficientes a favor da hipótese de que os ativos seguiam um *random walk* (Fama, 1965).

2.3 A Hipótese dos Mercados Eficientes

Conforme ressaltam Lo e MacKinlay, a Hipótese dos Mercados Eficientes, por si só, não é uma hipótese bem definida e empiricamente refutável. Para torná-la operacional é necessário que se especifique elementos adicionais, como as preferências dos investidores, a estrutura informacional, as condições de mercado, entre outros. Desta forma, um teste sobre a HME torna-se um teste conjunto sobre uma série de hipóteses auxiliares e um teste como tal não permite que se determine o fator relevante exato para certa característica observada no mercado, de modo que acaba por dizer pouco sobre os dados (Lo e MacKinlay, 1999).

Entretanto, a partir dos anos 1960, em especial devido às publicações de Cootner (1964), Fama (1965) e Samuelson (1965) a linha de pesquisa em Economia sobre a Hipótese dos Mercados Eficientes, passou a adotar o pressuposto de que os ativos seguem um movimento do tipo *random walk* (para análises em tempo discreto, ou Movimento Browniano para as em tempo contínuo) como sinônimo para mercados eficientes.

Sendo assim, a hipótese de *random walk* ou do movimento Browniano são, neste sentido, somente um dentre vários aspectos da Hipótese dos Mercados Eficientes, mas que todavia, acaba sendo interpretada erroneamente como seu equivalente. (Lo e MacKinlay, 1999).

Ao longo dos anos, entretanto, o conceito de eficiência dos mercados acabou recebendo diferentes definições por diversos autores, no entanto, a categorização clássica relacionada às informações disponíveis introduzida por Roberts (1959) e adotada posteriormente por Fama (1970) tornou-se a definição mais difundida e utilizada.

Roberts (1959) definiu três níveis de eficiência dos mercados: a eficiência na forma fraca, semi-forte ou forte, de acordo com a ausência de possibilidades de obtenção de lucros (livres de custos de transação) em função do grau de ajustamento dos preços a diferentes conjuntos de informação: preços históricos (ou retornos passados), outras informações públicas e informações privadas.

Desenvolvendo a abordagem de Roberts, Fama (1970) estabeleceu a definição de mercado eficiente segundo a qual este seria aquele em que os preços sempre

“refletem totalmente” todas as informações disponíveis. Diante de uma revisão bibliográfica teórica e empírica da literatura de seu tempo e fundamentando-se nos diferentes testes realizados pelos autores analisados, Fama listou, para os três níveis de eficiência supracitados, três graus de testes de eficiência dos mercados.

Sob a hipótese nula de que os mercados refletem totalmente toda a informação disponível, os testes de eficiência na forma fraca seriam aqueles em que o conjunto de informações relevante seria somente os preços históricos. Já na forma semi-forte, o conjunto relevante seria além dos preços históricos o ajuste a outras informações públicas disponíveis (como anúncios de lucros de empresas, *splits* de ações, etc.). Por fim, os testes na forma forte se preocuparam também com o acesso monopolista de determinados investidores a informações privadas privilegiadas.

Além destes, no chamado teste de previsibilidade dos retornos (Fama, 1991), sob a hipótese nula de que os retornos são imprevisíveis, o conjunto disponível de informações pode incluir, adicionalmente às informações históricas, as características passadas específicas das firmas (como tamanho, razão preço/lucro, a taxa de dividendos), variáveis macroeconômicas (relacionadas à estrutura a termo das taxas de juros e inflação) ou mesmo efeitos de calendário (Lo, 2004). Desta maneira, diante de mercados eficientes, a hipótese nula de imprevisibilidade dos retornos (portanto de que os retornos dos ativos seguem um processo do tipo *random walk*) não deveria ser rejeitada.

A pesquisa teórica realizada por Fama (1970) indicou que a maioria dos estudos à época baseava-se unicamente na suposição de que as condições de equilíbrio de mercado podem ser expressas por meio da noção de retornos esperados (condicionais). Segundo este conceito, os retornos esperados de equilíbrio de um determinado ativo, condicionados a um determinado conjunto de informações, é uma função do risco do ativo. Sendo assim, em termos matemáticos, é possível descrever de maneira geral que:

$$E(p_{j,t+1} | \Phi_t) = [1 + E(r_{j,t+1} | \Phi_t)] \times p_{j,t} \quad (2.14)$$

na qual

$p_{j,t}$: preço de um ativo j no período t ,

$r_{j,t+1}$: retorno percentual do ativo j entre o período t e $t + 1$,

Φ_t : conjunto total de informações disponíveis no período t .

Fama ressalta que Φ_t indica o conjunto total de informações que é inteiramente utilizada na determinação dos retornos esperados de equilíbrio e neste sentido Φ_t é “totalmente” refletido pelos preços $p_{j,t}$.

Fama observa ainda que o operador esperança (ou o valor esperado) é somente uma dentre diversas medidas de resumo de uma distribuição de retornos e a eficiência dos mercados por si só não impõe tanta importância a este operador. Diante disto, os resultados dos testes baseados nesta suposição dependem tanto da validade da própria suposição (descrita na equação 2.14) quanto da eficiência do mercado em análise.

Desta definição deriva a implicação de que não é possível obter retornos extraordinários aos retornos esperados de mercado nos possíveis negócios fundamentados somente nas informações disponíveis, ou seja, Φ_t .

Sendo $x_{j,t+1}$ a diferença entre os preços observados e o valor esperado dos preços projetados no período t com base nas informações disponíveis no mesmo período Φ_t , ou seja, o retorno extraordinário de mercado de um ativo j no período $t + 1$, é possível escrever:

$$x_{j,t+1} = p_{j,t+1} - E(p_{j,t+1} | \Phi_t) \quad \text{tal que} \quad E(x_{j,t+1} | \Phi_t) = 0 \quad (2.15)$$

Ou ainda, de maneira equivalente,

$$z_{j,t+1} = r_{j,t+1} - E(r_{j,t+1} | \Phi_t) \quad \text{tal que} \quad E(z_{j,t+1} | \Phi_t) = 0 \quad (2.16)$$

Onde $z_{j,t+1}$ seria o retorno extraordinário em $t + 1$, em relação aos retornos esperados de equilíbrio projetados em t .

Desta forma, Fama ressalta que por definição, tanto a sequência de variáveis aleatórias $\{x_{j,t}\}$ quanto $\{z_{j,t}\}$ seriam “jogos honestos” com relação à sequência de informações Φ_t .

Conforme Milionis (2006), esta última observação em Fama (1970) foi a que gerou maiores dúvidas e controvérsias quanto à noção de eficiência dos mercados e suas implicações testáveis. Segundo o autor, estas controvérsias talvez pudessem ter sido desfeitas mediante uma definição mais precisa em termos matemáticos e estatísticos de Fama sobre o tema da eficiência informacional dos mercados. Em seu artigo Milionis sugere uma abordagem fundamentada em argumentos econométricos em oposição aos

exclusivamente financeiros de Fama, sem contudo, alterar o *framework* desenvolvido pelo autor.

Segundo esta abordagem Milionis sugere um modelo descrito por:

$$E(\tilde{R}_{j,t+1} | \Phi_t) = \sum_{i=1}^N p_i R_{j,t+1} \quad (2.17)$$

na qual :

$\tilde{R}_{j,t+1}$: estimador que o mercado usa para prever $R_{j,t+1}$,

$R_{j,t+1}$: retorno percentual do ativo j entre o período t e $t + 1$,

Φ_t : conjunto total de informações disponíveis no período t ,

p_i : probabilidade de todos os possíveis resultados de $R_{j,t+i}$.

Neste caso, dado o conjunto total de informações Φ_t , todos os possíveis resultados de $\tilde{R}_{j,t+1}$ são $R_{j,t+1}$ com probabilidade p_1 , $R_{j,t+2}$ com probabilidade p_2 , e assim por diante até $R_{j,t+N}$ com probabilidade p_N .

Desta forma, o erro de previsão $U_{j,t+1}$ poderia ser descrito pelo desvio entre o valor realizado dos retornos no período $t + 1$ e o valor esperado pelo mercado para os retornos no período $t + 1$, portanto:

$$U_{j,t+1} = R_{j,t+1} - E(\tilde{R}_{j,t+1} | \Phi_t) \quad (2.18)$$

O erro de previsão apresentaria então as seguintes propriedades:

1. O valor esperado dos erros seria igual a zero:

$$E(U_{j,t+1}) = 0. \quad (2.19)$$

2. O valor esperado dos erros condicional ao conjunto total de informações disponíveis seria também igual a zero:

$$E(U_{j,t} | \Phi_t) = 0. \quad (2.20)$$

Diante das propriedades (2.19) e (2.20) podemos afirmar que o processo estocástico dos erros de previsão é um martingale com respeito ao conjunto de informações analisado (o que, na terminologia de Fama, implicaria na afirmação de que se trata de um jogo justo).

De maneira geral, um processo X_t é martingale se $E(X_{j,t} | \Phi_t) = X_t$.

Segundo esta perspectiva, Milionis sugere ainda a definição de um mercado eficiente como aquele em que o estimador que o mercado utiliza para precificar os retornos futuros esperados de um ativo é o valor esperado condicional diante de toda a informação disponível até o momento presente.

É nesta linha de abordagem mais estatística e econométrica que se inserem a maior parte dos estudos produzidos a partir da década de 1960 sobre a eficiência dos mercados, impulsionados pelo artigo de Paul Samuelson (1965). Neste sentido, Milionis afirma que a adoção da linha de pensamento desenvolvida por Samuelson em seu artigo de 1965 seria mais adequada para evitar o aparecimento de controvérsias. Neste sentido, Milionis afirma que, apesar de o artigo de Fama (1970) e suas posteriores revisões terem se tornado as referências mais lembradas sobre os textos clássicos que versam sobre a eficiência dos mercados, é possível dizer que a literatura moderna sobre os estudos de mercados eficientes começa com o artigo de Samuelson em 1965 (MILIONIS, A. 2006).

Com base em uma análise referente à precificação temporal de *commodities* estocáveis, Samuelson demonstrou que, em mercados eficientes em termos informacionais, uma vez que todas as expectativas dos agentes e informações disponíveis são incorporadas de maneira completa, as variações de preços observadas ontem não exerceriam influência sobre as variações a serem observadas hoje, sendo portanto, cada variação independente de qualquer anterior ou futura, de maneira que as mudanças nos preços devem ser imprevisíveis.

Diante disto, a série de retornos observada no mercado acabaria por apresentar a propriedade martingale e os preços futuros seguiriam um movimento de passeio aleatório (*random walk*). Sendo assim, frente ao conjunto de informações disponíveis até o momento, o preço de mercado corrente constitui-se a cada momento, a melhor estimativa para o preço futuro do ativo, caracterizando assim um mercado futuro perfeito (WORKING, H. 1962).

Desta forma, a partir da perspectiva aberta pelo artigo de Samuelson, tornou-se corriqueira a adoção do pressuposto de que os ativos seguem um passeio aleatório como sinônimo para mercados eficientes. Deste momento em diante diversos artigos com base em diferentes metodologias foram escritos visando identificar mercados eficientes a partir da existência de processos de passeio aleatório nas séries temporais analisadas.

Segundo Lo e MacKinlay (1997) os primeiros estudos que avaliavam a hipótese de passeio aleatório neste contexto definiam que as séries temporais financeiras poderiam ser modeladas a partir da seguinte equação:

$$X_t = \mu + X_{t-1} + \varepsilon_t \quad , \quad (2.21)$$

onde

X_t : logaritmo dos preços,

μ : termo de *drift* arbitrário,

ε_t : termo de erro aleatório.

Esta forma convencional, a qual os autores denominam *random walk* I ou RWI, admitia ainda que os incrementos ε_t eram independentes e identicamente distribuídos (i.i.d.), com média igual a zero e variância constante no tempo, ou seja:

$$\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

Esta hipótese implica que os incrementos são não somente um ruído branco, mas um processo puramente aleatório (ou ruído branco forte). Conforme Morettin (2011), uma sequência de variáveis aleatórias $\{\varepsilon_t, t \in Z\}$ é um ruído branco fraco em tempo discreto se:

$$\begin{cases} E[\varepsilon_t] = 0 \\ E[\varepsilon_t^2] = \sigma^2 \\ E[\varepsilon_t, \varepsilon_s] = 0 \quad , \quad t \neq s \end{cases} \quad . \quad (2.22)$$

Na medida em que além das três características acima é incluída a hipótese de que os incrementos são independentes e identicamente distribuídos obtém-se um processo puramente aleatório.

Lo e MacKinlay (1997) ressaltam que a adoção da hipótese de que os incrementos são independentes implica que o processo de passeio aleatório é também um jogo justo assim como o que se observa a partir da perspectiva dos martingales de Samuelson, todavia, em um sentido mais forte. Isto se deve ao fato de que a independência implica que não somente que os incrementos são descorrelacionados como também que quaisquer funções não lineares dos incrementos também são descorrelacionadas. Além disto uma hipótese muito comum adotada na literatura sobre o tema é a admissão de que os incrementos seguem uma distribuição gaussiana, o que configura o termos de erro como um ruído branco gaussiano.

Segundo Lo e MaCkinlay (1997) também é possível identificar na literatura outra definição para o passeio aleatório, denominada por eles *random walk II* (RWII), segundo a qual a hipótese de que a distribuição se mantém a mesma é relaxada tal que se assume que os incrementos são independentes, mas não identicamente distribuídos (*INID, independent not identically distributed*), logo:

$$\varepsilon_t \sim INID(0, \sigma_\varepsilon^2).$$

Esta segunda versão do passeio aleatório passa, portanto a conter a primeira versão como um caso especial e inclui processos mais gerais, como é o caso de incrementos com heterocedasticidade incondicional, aspecto relevante se considerada a existência de variância não constante no tempo (Lo e Mackinlay, 1997)

Conforme Eitelman (2008) as duas versões de passeio aleatório acima, RWI e RWII, adotam hipóteses um pouco simplistas e já refutadas por diversos autores, em especial quando se trata de séries temporais financeiras, já que a heterocedasticidade e a não normalidade são fatos estilizados em finanças sobretudo se considerados horizontes temporais mais distantes. Todavia Lo e Mackinlay (1997) destacam que muitos testes econométricos iniciais sobre a hipótese de passeio aleatório em séries temporais se sustentavam fortemente nestas premissas. Apesar disto Eitelman (2008) ressalta que uma rejeição da hipótese de que uma série segue um passeio aleatório tal como definido pelos modelos RWI e RWII acima com base somente nestes últimos argumentos seria de pouco interesse atualmente.

Já a terceira versão para o passeio aleatório que pode ser encontrada na literatura sobre o assunto, a qual os autores denominam *random walk III* ou RWIII, relaxa então a hipótese de que os retornos são independentes e assim passa a incluir processos com incrementos decorrelacionados, mas não independentes, e não identicamente distribuídos (*NIID, not independent and not identically distributed*), de tal forma que:

$$\varepsilon_t \sim NIID(0, \sigma^2).$$

Destaca-se ainda, que para quaisquer das três versões supracitadas a esperança e a variância condicionais a um determinado valor inicial X_0 podem ser escritas por:

$$E[X_t | X_0] = X_0 + \mu t, \quad (2.23)$$

$$Var[X_t | X_0] = \sigma^2 t, \quad (2.24)$$

o que implica que o modelo de passeio aleatório é um processo não estacionário⁸ na medida em que tanto a esperança quanto a variância condicionais são variáveis no tempo (Lo e Mackinlay, 1997).

Além destas, segundo Fava e Alves (1998) outra representação que poderia modelar uma série temporal financeira é um $ARIMA(p, d, q)$ tal que:

$$\phi(B)(1-B)^d(Y_t - \mu) = \theta(B)\varepsilon_t, \quad (2.25)$$

onde:

$(1-B)$: operador de *lag* ou de *backshift*

d : parâmetro de diferenciação que assume valores inteiros

$\phi(B)$: polinômio de grau p

$$\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p), \quad (2.26)$$

$\theta(B)$: polinômio de grau q onde

$$\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q), \quad (2.27)$$

ε_t : ruído branco com média zero e variância σ_ε^2 .

A utilização desta equação geral para a modelagem de um passeio aleatório implicaria assumir que $d = 1$, e portanto, que todas as raízes do polinômio $\phi(B)$ estão sobre do círculo unitário (ou seja, $\phi = 1$). Isto significa considerar que, para d assumindo valores inteiros, o processo seria integrado de ordem um, ou $I(d = 1)$, na medida em que o processo estocástico apresentaria raiz unitária. Sendo assim, o processo a ser analisado poderia ser descrito como um $ARIMA(0,1,0)$.

Neste sentido, um dos procedimentos tradicionais de avaliação de que uma série poderia ser representada por um passeio aleatório envolve uma análise quanto a presença de raiz unitária por meio de testes convencionais como o de Dickey e Fuller (1979), Dickey-Fuller aumentado (1981), Phillips e Perron (1988). Estes testes adotam como hipótese nula a existência de raiz unitária e como hipótese alternativa, que a série é estacionária podendo assim apresentar um processo gerador do tipo $ARIMA(p,0,q)$, que é o mesmo que um $ARMA(p,q)$, com $p \geq 0$ e $q \geq 0$ (Fava e Alves, 1998).

⁸ Um processo é dito fracamente estacionário quando, entre outras coisas, o primeiro e o segundo momento não são constantes ao longo do tempo. Já os processos fortemente estacionários são aqueles em que a distribuição se mantém constante ao longo do tempo (MORETTIN, P. 2011).

No entanto, os testes de raiz unitária apresentam vários problemas dentre eles o baixo poder para distinguir processos não estacionários como um passeio aleatório (em que $\phi = 1$) de estacionários persistentes (em que $|\phi|$ está próximo de um), também chamados de processos de memória longa (Morettin, 2011).

As séries que apresentam um processo de memória longa são aquelas em que existe uma estrutura de dependência de longo prazo entre as observações, de tal forma que a função de autocorrelação pode apresentar comportamento muito semelhante ao de uma série não estacionária, $I(d = 1)$. Diante desta semelhança é possível que, frente a não rejeição da hipótese nula, os testes de raiz unitária indiquem que uma série com memória longa (processo estacionário persistente) é na realidade uma série não estacionária (Fava e Alves, 1998).

Nesse sentido, uma melhor representação para um processo estacionário persistente poderia ser a partir da consideração de um parâmetro de diferenciação que apresentasse valores não inteiros, mas sim fracionários. Para estes casos, a série temporal é dita fracionalmente integrada, serie fractal, ou de memória longa.

Diante disto, é necessário cuidado especial para avaliar a possibilidade de que a série temporal em análise apresente memória de longo prazo, o que demanda tanto uma especificação de modelo mais adequado quanto a aplicação de métodos distintos dos convencionais testes de raiz unitária, que serão discutidos com maior detalhe nos capítulos que se seguem.

2.4 Alguns estudos nacionais sobre a hipótese do passeio aleatório

Na modelagem de séries temporais de *commodities* é preciso levar em conta uma série de fatores, como a sazonalidade nos preços devido às oscilações na oferta e demanda em decorrência dos períodos de safra e entressafra.

Em termos dos estudos práticos voltados para a avaliação da hipótese do passeio aleatório no mercado brasileiro, o foco de grande parte dos autores é em grande medida voltado para os ativos financeiros. Diante disto, conforme destaca Bitencourt (2007), os estudos direcionados aos mercados de *commodities* agrícolas no Brasil são incipientes e o aprofundamento metodológico é ainda pouco explorado.

Ademais, é comum que estes estudos avaliem a hipótese de passeio aleatório como um sinônimo de mercados eficientes, ou fundamentando-se na Hipótese das

Expectativas Racionais⁹, de tal forma que os resultados encontrados pelos diversos autores dirijam-se para a determinação do grau de eficiência destes mercados: fraca, semi-forte e forte.

Bressan e Leite (2001) avaliaram a hipótese de passeio aleatório em séries de contratos futuros de café negociados entre março de 1992 e março de 1998, para testar o desempenho destes contratos como indicador do comportamento futuro dos preços à vista, por meio da aplicação da análise de correlação serial com até cinco defasagens. Além disto, os autores testaram a hipótese de que todas as correlações das séries avaliadas eram iguais a zero, ou seja, de que a série era um ruído branco, por meio da avaliação da estatística Q de Ljung e Box (1978).

Em termos da análise autocorrelação, em um teste inicial que não considerou horizontes fixos¹⁰, 27 do total de 155 calculadas (aproximadamente 17,5%) estavam fora dos intervalos de confiança indicando a presença de autocorrelação serial. Já no que diz respeito à análise da estatística Q de Ljung e Box, 13 dos 31 contratos avaliados não seguiam um passeio aleatório. Diante de suas hipóteses de trabalho, Bressan e Leite concluíram que em 42% dos contratos analisados foi identificada ineficiência nos mercados, no sentido de que os agentes não teriam utilizado toda a informação que lhes era disponível para tomada de decisões, de tal forma que os preços não teriam incorporado de maneira instantânea as expectativas de mercado.

Em um segundo momento, porém, a consideração de horizontes fixos¹¹ levou os autores à conclusão de que os coeficientes de autocorrelação eram estatisticamente iguais a zero a um nível de 5%, indicando que não era observado um padrão de comportamento nos erros de previsão. Sendo assim, concluíram que, para o período avaliado, os preços futuros de café podiam ser descritos por um passeio aleatório (Bressan e Leite, 2001).

Outra metodologia de avaliação da hipótese da caminhada aleatória nos mercados futuros agropecuários brasileiros foi adotada por Cruz Junior e Silveira (2007)

⁹ Segundo a teoria das expectativas racionais, os agentes econômicos são indivíduos que tomam suas decisões de compra e venda nos mercados de forma racional, com base em cálculos fundamentados em expectativas, de tal maneira que estas decisões acabam por refletir toda a informação disponível no mercado.

¹⁰ Inicialmente foram analisados os diferentes contratos futuros sem que se considerassem os impactos que poderiam ser gerados por períodos em que a vigência de ambos era coincidentes (ou seja, períodos em que o contrato mais recente já estava sendo negociado enquanto as negociações com o contrato mais antigo ainda não haviam sido encerradas). Esta sobreposição, todavia, pode desencadear problemas de autocorrelação serial devido aos erros das diferentes amostras de cada contrato.

¹¹ Os horizontes fixos eram caracterizados por períodos cinco dias antes do vencimento dos contratos até noventa dias, com janelas de cinco em cinco dias.

e diz respeito à análise de raiz unitária. A presença de raiz unitária em uma série temporal indica que a série é não-estacionária. Todavia, muitos autores utilizam os testes referentes à presença de raiz unitária nas séries como uma maneira de avaliar a hipótese da eficiência dos mercados em sua forma fraca, no sentido de que os preços em análise conteriam todas as informações relevantes disponíveis até o momento.

Cruz Junior e Silveira avaliaram séries diárias dos logaritmos dos preços dos contratos futuros (de primeiro vencimento) de café arábica cotados nas bolsas BM&FBOVESPA (Brasil), ICE Futures (Nova Iorque) e LIFFEE (Londres) , para o período de janeiro de 2002 a dezembro de 2006. A análise da estacionariedade destas séries em nível foi feito por meio dos testes de raiz unitária de Dickey-Pantula e Dickey e Fuller Aumentado (ADF). Os resultados indicaram que todas as séries em nível apresentaram raiz unitária, de modo que eram não estacionárias. Os testes indicaram ainda que as séries, para o período analisado, eram integradas de ordem igual a um, sendo estacionárias somente após uma diferenciação. Diante disto, os autores afirmam que os mercados seguem um *random walk* e por isso são eficientes (Cruz Junior e Silveira, 2007).

Os autores avaliaram ainda, por meio de testes de cointegração que os mercados da ICE Futures e da BM&FBOVESPA apresentam relação de longo prazo, o que foi justificado por eles diante do fato de que os preços futuros de ambas as bolsas serem referenciados em um mesmo tipo de *commodity*, o café arábica. Ademais, Cruz Junior e Silveira avaliaram a causalidade no sentido de Granger entre os mercados e obtiveram resultados que indicaram que os movimentos dos preços futuros na bolsa de Nova Iorque causam os retornos observados tanto na BM&FBOVESPA quanto na LIFFEE, o que pode ser explicado por esta ser a praça com maior volume de negociações de contratos futuros de café arábica do mundo.

3 A memória longa e a hipótese dos mercados fractais

3.1 Estruturas fractais

A existência de formas assimétricas e de alta complexidade como um litoral recortado, flocos de neve, uma bola de papel amassado, uma nuvem ou um pinheiro, é bastante comum na realidade da vida cotidiana, entretanto, até pouco antes das décadas de 1960 e 1970 estas estruturas eram um desafio à matemática euclidiana. Sua descrição somente era possível por meio de um grande número de equações que acabavam resultando em meras simplificações das formas observáveis. Sendo assim, a definição de comprimento, área e volume de objetos e fenômenos como estes era uma tarefa de grande esforço e que gerava informações pouco precisas. Todavia, a partir das décadas de 1960 e 1970, com o desenvolvimento da geometria fractal por Benoit Mandelbrot estes desafios puderam então ser superados.

Não somente objetos e fenômenos naturais podem ser descritos e tratados como fractais, mas também séries temporais econômicas. A semelhança entre gráficos de séries temporais financeiras e o perfil dos litorais recortados que formam linhas serrilhadas (*jagged lines*), assim como outras estruturas que puderam ser descritas somente após os avanços da geometria fractal de Mandelbrot, são exemplos de *insights* que conduziram o avanço de estudos em finanças com base interpretação das séries temporais financeiras como estruturas fractais.

Além deste aspecto visual, também o avanço da perspectiva de que o sistema econômico, ao contrário da interpretação do *mainstream*, é caracterizado como um sistema não em equilíbrio, mas regido por processos dinâmicos, complexos e não lineares, colaborou para o avanço da aplicação de técnicas de processos fracionários às análises de séries temporais econômicas. Na medida em que estas técnicas exigem em geral, hipóteses menos restritivas do que os modelos mais tradicionais, elas podem ser mais adequadas para capturar algumas peculiaridades observadas em séries financeiras, como a não normalidade dos retornos.

Segundo Peters (1996), apesar do desenvolvimento de extensas pesquisas sobre estas estruturas em diferentes áreas do conhecimento, Mandelbrot não estabeleceu em seus trabalhos uma definição formal e precisa sobre uma estrutura fractal. Jin e Frechette, todavia, utilizam uma definição interessante, segundo a qual a fractalidade de um objeto ou fenômeno pode ser entendida como um conjunto de regras sistemáticas

que governam as propriedades deste objeto ou fenômeno ao longo do tempo e espaço (JIN, H. e FRECHETTE, D.L., 2002).

Diante disto, a determinação de uma estrutura fractal implica na identificação deste conjunto de regras que pode ser percebido, como ressalta Peters (1994), justamente por meio da investigação quanto às características comuns e mensuráveis a estas, tais como a autossimilaridade (*self similarity*) e a dimensão fractal.

A autossimilaridade ou auto-semelhança, em matemática, indica que parte do objeto ou estrutura em análise é semelhante e está de alguma forma relacionada com o todo a qual ela pertence. Isso implica que as partes de um sistema, em diferentes escalas espaciais ou temporais, são semelhantes ao sistema como um todo (LORENZ, E. D 1996).

No exemplo da estrutura de um pinheiro, esta característica pode ser percebida na medida em que se detecta que o pinheiro é formado por uma rede de galhos que são em geral qualitativamente similares ao pinheiro como um todo, mas que individualmente são de fato diferentes. A partir de cada galho surgem outros novos que se tornam cada vez menores, mas que permanecem com a característica similar ao pinheiro em sua totalidade. Em termos da análise das séries temporais financeiras, a autossimilaridade pode ser percebida visualmente, como destaca Mandelobrot e Hudson (2004), por meio da análise gráfica de séries de retornos de ativos: mesmo se plotadas em gráficos referentes a diferentes horizontes temporais, se retiradas as legendas, é muito difícil identificar o horizonte temporal a que se refere cada gráfico.

A autossimilaridade pode ser de três tipos: exata, quase-autossimilaridade (caso dos fractais auto-afins) ou ainda autossimilaridade estatística.

A autossimilaridade exata é aquela que revela outro importante elemento de estruturas fractais: a invariância da forma em relação à escala (*scale invariance*). Isto implica que independente da variação do comprimento da escala segundo a qual a estrutura fractal é analisada, sua forma é conservada idêntica à original. Este tipo de autossimilaridade é em geral associado aos fractais determinísticos gerados por sistemas de funções iterativas (como algoritmos recursivos, por exemplo) e por isso não é normalmente observada nos elementos reais da natureza. Exemplos de estruturas fractais que apresentam autossimilaridade exata citam-se o conjunto de Cantor, o triângulo de Spierpinski, o floco de neve de Koch, entre outros (MANDELBROT, B. e HUDSON, R., 2004).

Existem, todavia, estruturas fractais formadas por cópias menores da estrutura como um todo, mas que as proporções originais não são mantidas fixas em todas as direções do espaço frente uma mudança de escala, o que caracteriza os fractais auto-afins. Estruturas fractais deste tipo podem ser encontradas em análises de séries de precipitações (MUÑOZ, P. C. et al, 2008)

Já a terceira forma de autossimilaridade, a estatística, envolve a busca de propriedades estatísticas que se mantêm inalteradas diante de mudanças de escala, ou seja, agregação temporal (TARGINO, R. S., 2010). Desta maneira é forma aquela menos evidente através da observação visual, dado que as medidas preservadas em diferentes escalas são aspectos numéricos ou estatísticos, tais como o expoente de Hurst e a dimensão fractal.

De acordo com Targino (2010), um processo Z definido abaixo, pode ser dito auto-similar se existe um expoente de autossimilaridade H , também chamado de expoente de Hurst, tal que, para qualquer fator de escala c tanto $(Z_{ct})_{t \in R}$ e $(c^H Z_t)_{t \in R}$ possuem a mesma lei, ou seja:

$$(Z_{ct})_{t \in R} \stackrel{d}{=} (c^H Z_t)_{t \in R}, \quad (3.1)$$

onde:

$$H > 0,$$

$$c > 0.$$

Segundo Peters (1996) o expoente de Hurst é uma estatística robusta, na medida em que necessita de poucas hipóteses sobre o sistema a ser estudado, que permite a classificação de séries temporais em processos aleatórios e não aleatórios mesmo diante de séries aleatórias não normalmente distribuídas.

Segundo Peters, a relação entre o expoente de Hurst e uma medida de correlação dada pela equação (3.2) abaixo permite três tipos de classificação para o expoente de Hurst, dadas pelo quadro abaixo, que implicam na observação de diferentes comportamentos nas séries analisadas:

$$C = 2^{(H-1)} - 1, \quad (3.2)$$

onde:

C : medida de correlação entre as observações,

H : expoente de Hurst.

Analisando-se a equação obtém-se a seguinte classificação:

Tipo	Expoente (H)	Correlação (C)	Característica da Série
(1)	$0 \leq H < 0,5$	$-0,5 \leq C < 0$	- Antipersistente ou ergódica - Com reversão à media - Volatilidade elevada
(2)	$H=0,5$	$C = 0$	- Eventos aleatórios e descorrelacionados - distribuição pode ou não ser Normal
(3)	$0,5 < H < 1,0$	$0 < C \leq 1,0$	- Persistente - Tendências aparentes (<i>trend reinforcing</i>) - Movimento Browniano fracionário ou <i>random walk</i> viesado

Fonte: Peters, E. E. (1996), elaboração própria

Quadro 1 – Expoente de Hurst vs Correlação.

O comportamento do Expoente de Hurst reflete, portanto, a seguinte estrutura de correlações na série:

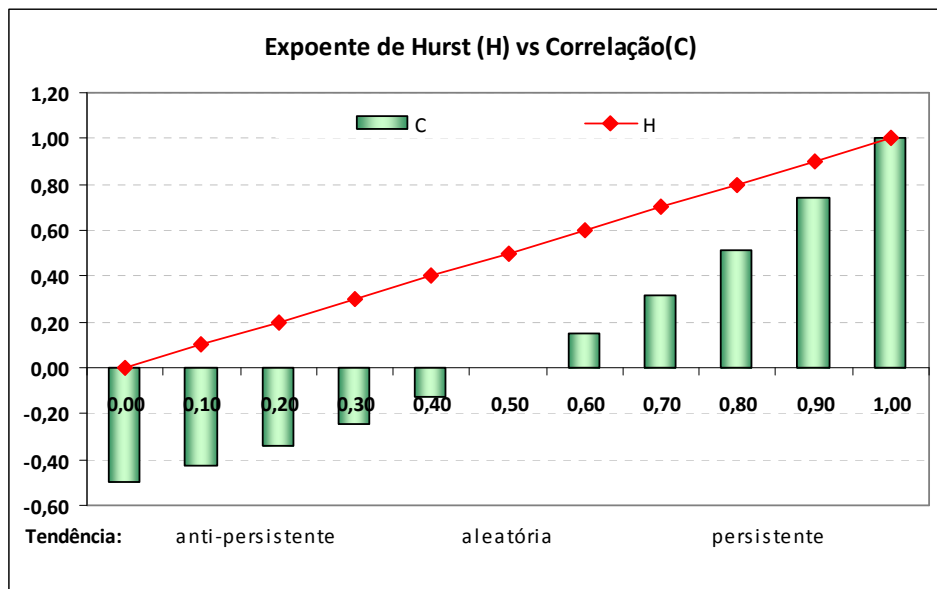


Gráfico 1 – Expoente de Hurst vs Correlação.

Nota-se, portanto, que o tipo dois de expoente de Hurst, onde $H = 0,5$, é encontrado para os casos em que os eventos que compõem o processo são aleatórios e descorrelacionados, de tal forma que a correlação entre as observações é nula. Esta situação é em geral observada para o caso de processos aleatórios, cuja distribuição pode ser gaussiana ou não.

Diante disto observa-se que, para $H \neq 0,5$ as observações não são independentes e carregam uma espécie de memória de todos os eventos que a precederam, de tal forma que um processo que apresenta expoente de Hurst diferente do

tipo dois é caracterizado como um processo não-martingale. Trata-se de um processo que apresenta memória longa que teoricamente, poderia durar para sempre (Peters, 1996).

Conforme Zivot e Wang (2005), um processo estacionário com memória longa é aquele que apresenta uma função de autocorrelação

$$\rho(k) \rightarrow Ck^{-\alpha} \quad \text{quando } k \rightarrow \infty, \quad (3.3)$$

onde

C : medida de correlação entre as observações, tal que $C > 0$,

α : número real tal que $0 < \alpha < 1$.

Em função disto a função de autocorrelação de um processo com memória longa decai para zero hiperbolicamente (ao contrário de um processo de memória curta, cujo decaimento é exponencial) tal que suas autocorrelações não são somáveis:

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty, \quad (3.4)$$

Sabe-se ainda que:

$$H = 1 - \frac{\alpha}{2} \in (1/2, 1), \quad (3.5)$$

Isto permite que se reescreva que os processos com memória longa são aqueles que apresentam $0,5 < H < 1$, de tal forma que quanto maior H , maior a memória de longo prazo do processo. Sendo assim, um sistema que apresenta estatística de Hurst nestes termos é resultado de uma longa série de eventos interconectados tal que um acontecimento no momento presente é resultado de tudo o que ocorreu no passado e continuará influenciando as situações futuras, de modo que a ordem das observações tal como elas ocorrem é relevante para a análise dos dados.

Expoentes de Hurst do tipo um são em geral encontrados em análises de processos que apresentam reversão à média, ou seja, se o sistema estava com tendência de alta em um determinado período, no seguinte é mais provável que ele apresente uma tendência de queda (e vice versa), o que implica em processos que apresentam maior volatilidade se comparados aos aleatórios. Apesar de esta ser uma hipótese teórica muito comum em estudos fundamentados na abordagem tradicional da economia como um sistema em equilíbrio, poucas séries anti-persistentes foram encontradas em análises práticas ((PETERS, E. E. 1996).

Já expoentes de Hurst do tipo três, são normalmente encontrados em processos com comportamento persistente, ou seja, se o sistema estava com tendência de alta em

determinado período, no seguinte é mais provável que se mantenha uma tendência de alta (e vice versa), tal que a tendência é mais aparente em análises gráficas.

Conforme H se aproxima mais de um, indicando uma correlação de 100% entre as observações, o comportamento persistente se torna mais acentuado. De maneira oposta, quanto mais o expoente de Hurst se aproximar de 0,5, mais a série será ruidosa e menor a tendência aparente.

Segundo Peters, uma série que apresenta comportamento persistente, portanto com expoente de Hurst do tipo 3 variando entre 0,5 e 1,0, é caracterizada como um movimento browniano fracionário ou passeio aleatório viesado (*biased random walk*). Nesse contexto, a força do viés dependerá de quão acima de 0,5 está o expoente de Hurst (PETERS, E. E. 1996). Neste sentido, um movimento browniano fracionário pode ser definido como um processo gaussiano de média zero, auto-similar e com incrementos estacionários (TARGINO, R. S, 2010).

A estimação do expoente de Hurst pode ser feita pelo método de *rescaled range* que envolve a estimação da estatística R/S e que será discutida com mais detalhes na seção (4.2.1).

A dimensão fractal ou fracionária, segunda propriedade marcante de uma estrutura fractal, é um elemento que permite a quantificação do grau de irregularidade ou fragmentação de um objeto natural, um conjunto geométrico, uma figura (como um gráfico ou uma fotografia) assim como de uma série temporal.

De acordo com a matemática euclidiana, a variável indicativa da dimensão topológica D_t de um elemento somente assume valores inteiros: pontos desconectados apresentam dimensão igual a zero; linhas apresentam dimensão igual a um, superfícies dimensão igual a dois e volumes (entendidos como objetos tridimensionais passíveis de diferenciação em toda sua superfície) apresentariam dimensão igual a três. Segundo este referencial, uma bola oca e/ou perfurada (como uma folha de papel amassada em formato de esfera) não pode ser considerada um objeto tridimensional a rigor, uma vez que ao apresentar protuberâncias e buracos sua superfície não seria diferenciável pelos instrumentos de cálculo tradicionais. Nota-se assim, que um valor adequado para a dimensão de uma bola oca e/ou perfurada se situaria entre dois e três, tal como um processo de caminhada aleatória, que não sendo propriamente uma reta nem um plano, teria sua dimensão refletida de forma mais adequada por um valor entre um e dois, portanto, um valor fracionário (PETERS, E. E. 1996).

Nesse sentido, na medida em que a dimensão fracionária permitiria uma análise da irregularidade de uma estrutura fractal na natureza, como o litoral de um país, também a avaliação deste coeficiente permitiria uma análise do grau de irregularidade que pode ser observada de um gráfico referente a uma série temporal financeira, reflexo da volatilidade (ou desvio padrão) do ativo em análise.

Em termos estatísticos, o desvio padrão busca medir quão distante uma observação estará em relação ao valor médio observado para a série, de tal forma que a volatilidade (dada pelo quadrado do desvio padrão) seria uma medida da dispersão dos retornos dos ativos, muitas vezes interpretada como o risco do mesmo. Nesse sentido, quanto mais irregular for a série maior seria a volatilidade ou o risco do ativo por ela representado. Ressalta-se, porém, que esta relação somente é válida quando o sistema analisado é aleatório, pois do contrário, o desvio e a volatilidade não são medidas de dispersão adequadas.

Segundo Peters (1996), assim como o formato observado para o perfil dos litorais recortados é decorrente das forças naturais que atuam sobre esta estrutura, também a distribuição de uma série financeira no tempo é produto de todos os fatores e que são evidenciados através das correlações e/ou dependência entre as observações.

Para o caso de uma série temporal aleatória (como um passeio aleatório) o que se observa é a influência de um grande número de eventos igualmente prováveis de ocorrer, configurando um sistema com um número muito elevado de graus de liberdade. Devido à complexidade extrema do sistema a previsão é praticamente impossível e as hipóteses da geometria clássica e da abordagem tradicional dos mercados financeiros (aquelas fundamentadas na hipótese dos mercados eficientes) poderiam ser utilizadas com pouco prejuízo. Entretanto uma série temporal não aleatória reflete justamente a falta de aleatoriedade na natureza dos fatores que a influenciam, de tal forma que os dados vão se agrupar para refletir essa correlação inerente às influências, fazendo com que a série seja uma estrutura fractal.

Sendo assim, para o caso da análise das séries temporais, exceto se forem representadas por um processo *random walk* a geometria clássica e a abordagem tradicional dos mercados financeiros não seria muito adequada para o entendimento dos elementos causais por trás desta estrutura (PETERS, E. E. 1996).

Mandelbrot ressalta ainda que esta inadequação se deve principalmente ao fato de que a abordagem tradicional dos mercados financeiros tem como pressuposto que as variações dos preços de mercado apresentam processos aleatórios do tipo brando, como

se cada movimento aleatório de subida ou descida dos preços fosse determinado por um jogo honesto de cara ou coroa (seguindo portanto, um processo *random walk*) e onde grandes flutuações deveriam ser consideradas *outliers*. No entanto, a realidade observável indica que grandes flutuações, ou processos aleatórios mais turbulentos, apresentam maior probabilidade de ocorrerem do que aquela considerada pela teoria por trás desta abordagem. De maneira oposta, a análise fractal seria capaz de explicitar outras formas de aleatoriedade além da branda: a lenta e a turbulenta, característica dos mercados financeiros reais (MANDELBROT, B. e HUDSON, R., 2004).

Sendo assim para uma análise melhor de processos não aleatórios, comumente observados no mercado, a dimensão fractal poderia ser uma medida mais apropriada do grau de irregularidade destas séries, em especial por não exigir qualquer suposição sobre a distribuição das observações.

Greene e Fielitz (1977) desenvolveram uma fórmula que relacionava o expoente de Hurst, H , e a dimensão fractal, D , como uma medida desta irregularidade (ruído) da série, dada por:

$$D = 2 - H \quad (3.6)$$

Diante disto, é possível avaliar a relação que existe entre a dimensão fractal, dada pela equação acima, com o expoente de Hurst e a medida de correlação da equação (2.13), melhor expressa pelo gráfico abaixo.

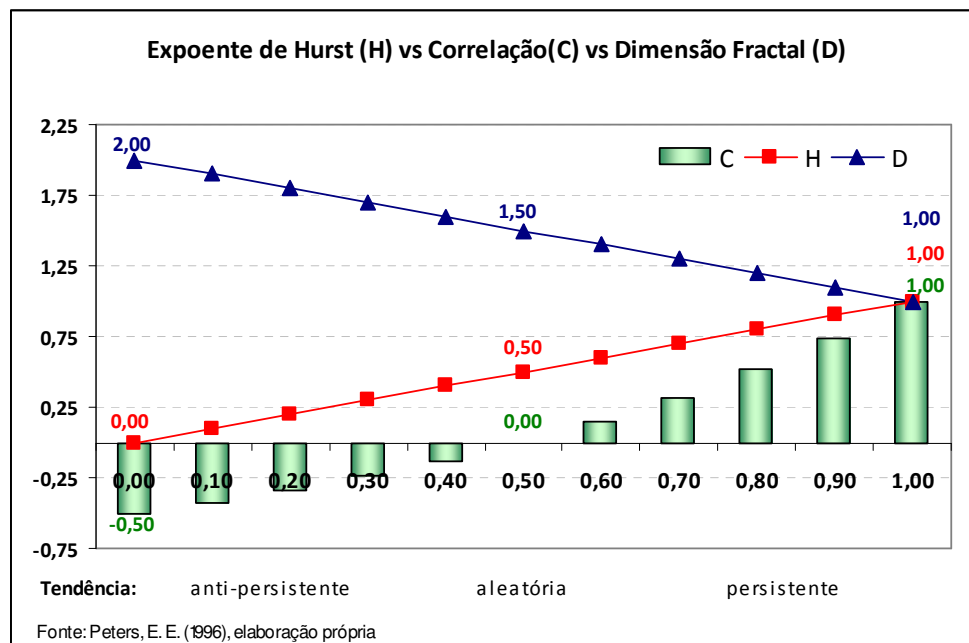


Gráfico 2 – Expoente de Hurst vs Correlação vs Dimensão Fractal.

Nota-se assim que, para processos aleatórios como um *random walk*, que apresentam $H = 0,5$ e $C = 0$ obtém-se $D = 1,5$, se situando portanto entre a dimensão fractal de uma reta (igual a um) e a dimensão fractal de um plano geométrico (igual a dois). Já para processos que apresentam expoente de Hurst no intervalo $0 \leq H < 0,5$, a dimensão fractal apresenta valores entre $2 \leq D < 1,5$ enquanto que para os processos em que $0,5 < H \leq 1,0$, a dimensão fractal assume valores entre $1,5 < D \leq 1,0$. Quanto mais altos forem os valores da dimensão fractal, maior será a irregularidade (ou o ruído) da série analisada.

3.2 Séries temporais fracionárias

Da literatura que trata de séries temporais, sabe-se que séries econômicas e financeiras apresentam características bastante peculiares, tais como a presença de tendências; sazonalidade; pontos atípicos; heterocedasticidade condicional e ausência de linearidade. No que diz respeito a séries de retornos, outros fatos estilizados são observados, tais como a existência autocorrelação entre as observações assim como no quadrado dos retornos, agrupamentos de volatilidade ao longo do tempo, a observação de distribuições com caudas pesadas (*fat tails*) e não normais (portanto leptocúrticas), assim como a não linearidade dos retornos (MORETTIN, P. A. 2011).

Diversos estudos econômicos identificaram além destas características comuns presentes em séries temporais de diferentes mercados, também outras que se caracterizam como indícios de que estas séries podem ser estruturas fractais, tais como a observação de dimensão fractal e de padrões periódicos e não cíclicos com mudanças abruptas de tendência (JIN, H. e FRECHETTE, D.L., 2002).

Apesar da observação destes elementos em diversos mercados, em geral os autores continuam a adotar a hipótese de que os preços e também os retornos seguem alguma variante de passeio aleatório.

Nota-se, porém, que a presença destes fatos estilizados em séries temporais configura-se como uma oposição aos requisitos necessários para a caracterização da mesma como um processo de caminhada aleatória no seu sentido estrito (seja em tempo discreto ou contínuo). Isto se deve ao fato de que a configuração como tal depende da identificação de elementos como por exemplo, a normalidade, a linearidade dos retornos e a independência temporal entre as observações. Diante disto, a hipótese de que os mercados financeiros são fractais (no sentido de que os processos são estruturas fractais)

é concorrente à hipótese de que os mercados são eficientes (no sentido de que os processos são aleatórios), como ressaltam Mandelbrot e Hudson (2004), Jin e Frechette (2002), Peters (1996) entre outros.

Sendo assim, é interessante que um estudo mais completo dos mercados envolva além das análises tradicionais para passeios aleatórios, também a avaliação das características fractais destes mercados.

Segundo Jin e Frechette (2002), existem duas formas consagradas de análise de séries temporais fracionárias que envolvem técnicas referentes ao estudo do movimento browniano fracionário (Fractal Brownian Motion - FBM) ou a partir da utilização de distribuições estáveis¹².

Apesar de inicialmente estudado por Andrei Nikolaevich Kolmogorov, com o nome de "Espiras de Wiener" (*Wiener Spirals*), o movimento browniano fracionário foi assim introduzido na literatura de finanças por Mandelbrot e Van Ness (1968).

Segundo Jin e Frechette (2002), o estudo do movimento browniano fracionário pode ser feito principalmente por duas técnicas: análise da estatística R/S (*rescaled range*) que envolve a determinação do expoente de Hurst, e a utilização de modelos de integração fracionária (acompanhados da estimação da ordem de integração fracionária do processo), como os das famílias $ARFIMA(p, d, q)$ e $FIGARCH(m, d, q)$.

Segundo os autores, a análise de R/S pode detectar de maneira indireta a memória de longo prazo em séries temporais assim como a autossimilaridade a partir do uso de diferentes escalas temporais. Os procedimentos específicos destas análises serão melhor explicados no capítulo 4.

Já a segunda técnica, referente aos modelos de integração fracionária e a estimação da ordem de integração fracionária dos processos, por sua vez, ganharam força a partir dos artigos de Granger e Joyeux (1980) e Hosking (1981).

Conforme Zivot e Wang (2005), um processo estacionário com memória longa pode ser representado a partir da função de autocorrelação dada pela equação (3.3). Para o caso de um processo estacionário, os autores observam que a função de autocorrelação contém a mesma informação que a densidade espectral definida por:

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ik\omega}, \quad (3.7)$$

¹² Entretanto, distribuições estáveis não serão abordadas neste estudo.

onde $f(\omega)$ é a frequência de Fourier, tal que é possível demonstrar que

$$f(\omega) \rightarrow C_f \omega^{\alpha-1} \quad \text{quando } \omega \rightarrow 0, \quad \text{com } C_f > 0. \quad (3.8)$$

Com base na propriedade de escala em (3.3) e a propriedade do domínio da frequência em (3.8) Granger e Joyeux (1980) e Hosking (1981) demonstraram de maneira independente que processos de memória longa podem também ser modelados parametricamente a partir da extensão de um processo integrado para um processo *fracionalmente* integrado (Zivot e Wang, 2005). Conforme Marques (2009) os autores recuperaram o conceito matemático de integração/diferenciação fracionária, $I(d)$ com $d \in R$, e desenvolveram, de maneira independente, o modelo autorregressivo de médias móveis fracionalmente integrado ou $ARFIMA(p, d, q)$. A recuperação deste conceito se refletiu, no caso do estudo de séries temporais econômicas, na possibilidade de que a ordem de integração, d , das séries pudesse admitir valores fracionários e não somente inteiros, o que permite sua caracterização como estruturas fractais (ou séries temporais fracionalmente integradas).

Nesse sentido, esta abordagem permitiu a modelagem de um processo fracionalmente integrado por:

$$(1-B)^d (X_t - \mu) = u_t, \quad (3.9)$$

onde

u_t : é um processo estacionário com memória curta, com média zero e função densidade espectral igual a $f_u(\lambda)$,

d : é a ordem de integração do processo, tal que $d \in R$ e $d > -1$,

B : operador do *lag*, da defasagem do processo,

$(1-B)^d$: operador de diferença fracionária.

Conforme ressaltam Zivot e Wang (2005), quando uma série temporal é fortemente persistente e não apresenta estacionariedade, uma forma de se obter estacionariedade é considerar $d=1$ e diferenciar a série uma vez para obter estacionariedade. Todavia, para algumas séries com memória particularmente longa, assumir uma ordem de integração dada por um número inteiro pode ser muito devido à propriedade de que a densidade espectral desaparece para a série diferenciada. Logo, para permitir a memória longa evitando uma ordem de diferenciação inteira para X_t ,

assume-se $d \in R$ tal que $d > -1$ (portanto fracionário) e procede-se um filtro de diferenças fracionário definido por:

$$(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k (B)^k, \quad (3.10)$$

$$\sum_{k=0}^{\infty} \binom{d}{k} (-1)^k (B)^k = 1 - dB + \frac{1}{2!} d(d-1)B^2 - \frac{1}{3!} d(d-1)(d-2)B^3 + \dots, \quad (3.11)$$

na qual,

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}. \quad (3.12)$$

Segundo Marques (2009), os modelos de integração fracionária podem descrever de maneira direta tanto a memória de longo prazo quanto de curto prazo, sendo mais utilizados quando o objetivo da análise está relacionada à previsão dos movimentos futuros da série.

A relação entre a ordem de integração d do processo e o expoente H de Hurst é dada por:

$$d = H - 0,5. \quad (3.13)$$

Da equação (3.6) é possível escrever que:

$$d = 1,5 - D. \quad (3.14)$$

Fava e Alves (1998) ressaltam que para o intervalo em que $-0,5 < d < 0,5$ o processo é estacionário e invertível. Para o caso de processos aleatórios como um *random walk*, que apresentam $H = 0,5$, a ordem de integração é dada por $d = 0$. Já para processos que apresentam expoente de Hurst no intervalo $0 \leq H < 0,5$, a ordem de integração se situa no intervalo $-0,5 < d < 0$, o processo X_t é estacionário e possui memória curta sendo dito anti-persistente, com maior irregularidade. Os autores destacam que para este intervalo, todas as autocorrelações são negativas e tendem exponencialmente para zero, tal que a memória do processo é intermediária. Já nos casos em que $0,5 < H \leq 1,0$, a ordem de integração do processo pode assumir valores entre $0 < d < 0,5$, o que se reflete em todas as correlações positivas e declinantes hiperbolicamente, tal que o processo X_t é estacionário e com memória longa, sendo dito persistente.

Fora destes intervalos, ou seja, quando $|d| > 0,5$, em especial se $0,5 \leq d < 1,0$, Fava e Alves (1998) destacam que o processo apresenta reversão à média, mas deixa de ser estacionário na medida em que sua variância não é finita.

Uma melhor visualização entre o comportamento do expoente de Hurst H , a correlação C_ρ , a dimensão fractal D e a ordem de integração d , em séries temporais temporais fracionárias pode ser obtida no gráfico a seguir.

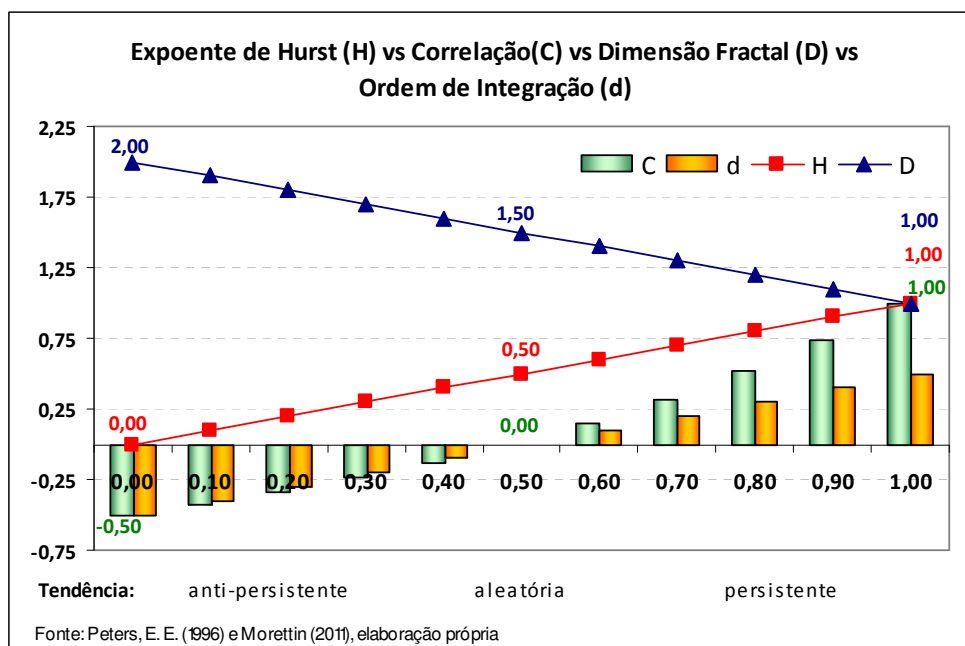


Gráfico 3 – Expoente de Hurst vs Correlação vs Dimensão Fractal vs Ordem de Integração.

O enfoque nas análises referentes à determinação da ordem integração fracionária do processo em análise requer sua estimação. Para tanto existem diversas técnicas, como o método da regressão do periodograma (que é um estimador do espectro $f_u(\lambda)$ do processo estacionário), o método GPH¹³, e o método de Whittle, descritos com mais detalhes no capítulo 4 (MORETTIN, P. A., 2011. e ZIVOT, E. e WANG, J., 2005).

No Brasil, em especial a partir da década de 1980, alguns estudos foram publicados utilizando-se este referencial, sobretudo no que diz respeito às avaliações de processos de memória longa em variáveis como taxas de juros e inflação, assim como na modelagem da volatilidade de ativos financeiros (MORETTIN, P. A. 2011).

¹³ Em referência a Geweke e Porter-Hudak (1983).

Esta característica de memória longa pode, todavia, ser observada tanto nos retornos das séries quanto na volatilidade. Para os processos em que a memória longa pode ser observada na série de retornos sua modelagem pode ser realizada especialmente a partir de modelos ARFIMA, enquanto que para o caso de a memória longa e comportamento persistente ser observada em séries de volatilidade dos retornos¹⁴, os modelos FIGARCH são mais utilizados para modelagem do processo.

Fava e Alves (1998) estudaram as séries temporais das taxas de inflação brasileira no período de 1974 a 1994 com o objetivo de identificar processos fracionalmente integrados utilizando modelos do tipo ARFIMA estimados a partir dos métodos de regressão do periodograma e de máxima verossimilhança. Para ambos os métodos concluíram que a ordem de integração dos processos eram inferiores a um, indicando que o processo gerador das taxas de inflação brasileira no período em análise apresentavam comportamento de memória longa e seriam mais bem estimados considerando seu caráter fracionário.

Marques (2009) analisou a dinâmica de longo prazo entre o consumo das famílias e a renda disponível no Brasil no período de 1947 a 2005. Por meio de testes de co-integração e integração fracionária, o autor encontrou resultados que apoiaram a existência de uma relação estocástica estável entre as variáveis. Para o autor, o avanço dos estudos fundamentados em modelos de integração fracionária mostrou que eles são mais aptos a captar comportamento de baixa frequência das séries temporais, tal que se tornam mais adequados para a representação de processos que exibem memória de longo prazo decorrente da existência de uma estrutura de longa dependência entre as observações.

No âmbito dos estudos dos mercados de *commodities*, Barkoulas, Labys, e Onochie. (1997) testaram a dependência de longo prazo para séries históricas de 21 *commodities* no período de 1960 a 1994 por meio de um processo de integração fracionária, ARFIMA. Das vinte e uma *commodities* analisadas, seis apresentaram indícios de que as séries temporais analisadas eram estacionárias, sendo elas cobre, ouro, soja em grão, chá, lã e o índice UCAM¹⁵.

¹⁴ Zivot e Wang (2005) indicam que, para o caso em que se deseje modelar de maneira conjunta a dinâmica dos retornos e a volatilidade do processo, a extensão dos modelos da família GARCH são mais adequados do que apenas a extensão dos modelos ARIMA.

¹⁵ Trata-se de uma média ponderada dos preços de exportação de *commodities* primárias de países em desenvolvimento.

Já Jin e Frechette (2002) analisaram o comportamento dos preços diários à vista listados na bolsa de Chicago (CBOT), em Kansas City (KC) e na bolsa de Nova Iorque (ICE Futures) para diversas *commodities*¹⁶ principalmente entre os períodos de 1992 e 2000¹⁷, utilizando tanto as análises R/S quanto aproximação de distribuições estáveis. Seus resultados indicaram que a maioria das séries não era estacionária em nível, mas em log-retornos. Com base nelas as análises de R/S resultaram em expoentes de Hurst estimados superiores a 0,5, sugerindo que a maioria das séries apresentava caráter de memória de longo prazo¹⁸, com dimensão fractal inferior a 1,5. O ajuste das distribuições empíricas às distribuições estáveis indicou que elas não eram normalmente distribuídas, tal que de forma geral seus resultados indicaram que as séries analisadas são fracionárias.

Matias, Silva e Vieira (2007) estudaram o comportamento fracionário da série histórica de preços diários do minério cobre para o período de 1974 a 2005. Seus resultados indicaram que os preços não eram estatisticamente independentes e identicamente distribuídos, assim como não apresentavam distribuição normal. Os testes de raiz unitária não foram rejeitados, indicando comportamento não-estacionário. A avaliação da estatística R/S indicou que as séries apresentaram memória curta, com tendência de reversão. Segundo os autores, apesar de a não-linearidade e a não-normalidade serem condições recorrentes em padrões fractais, mas não suficientes para afirmar a fractalidade das séries, o conjunto dos resultados encontrados os levou a afirmar que, para o período considerado, a série apresentou comportamento fractal.

¹⁶ As *commodities* analisadas foram: frango de corte; cacau; café; ovos brancos grandes; milho de Kansas City; aveia; soja; farelo de soja; óleo de soja; trigo de primavera; açúcar #11, trigo e preços a vista de trigo #1. Para maiores detalhes

¹⁷ Segundo os autores, o tamanho da amostra considerado por eles foi bastante superior ao de outros autores como é o caso de Barkoulas, J., W.C. Labys, and J. Onochie. (1997), por exemplo, que utilizaram uma amostra com aproximadamente 408 observações.

¹⁸ Embora tenham sido superiores a 0,5, os autores destacaram que elas não se mostraram estatisticamente diferentes de 0,5.

4 Metodologia

Diante da inadequação da geometria clássica e da abordagem tradicional dos mercados financeiros para a análise de séries que não seguem um processo tipo *random walk*, o primeiro passo em um estudo que envolve análises de séries temporais financeiras deve ser a verificação da própria hipótese de que a série apresenta um processo do tipo *random walk* para somente depois verificar sua fractalidade.

A identificação de um processo do tipo *random walk* ou movimento browniano em séries temporais se dá pela verificação de algumas características, tais como a linearidade da série temporal e a assunção de que os retornos em relação à média e os incrementos seguem algum tipo de distribuição específica (sendo a mais utilizada a gaussiana). Já os indícios de que a série apresenta um movimento fractal derivam da falta dos elementos supracitados, assim como da identificação de memória de longo prazo (*long-term dependency* ou *long term correlation*).

Entretanto, conforme ressalta Bueno (2011), o aspecto que deve preceder qualquer análise e do qual depende a estimação correta de uma série temporal é a avaliação adequada conforme sua estacionariedade, elemento fundamental também para o estudo do comportamento de séries temporais fractais (MANDELBROT, B. et al, 1997). Segundo Brooks (2002), um processo é considerado estacionário se apresenta média e variância finita e se estas medidas, assim como a autocovariância, se mantêm constantes ao longo do tempo.

Neste contexto, a modelagem correta de séries temporais consiste em avaliar todos estes elementos e realizar os seguintes passos: identificar as ordens p e/ou q dos modelos em análise; estimar o modelo econométrico representativo da série; verificar se os resíduos estimados seguem um processo tipo *random walk* (portanto se não rejeitam a hipótese nula de que sejam um ruído branco) e realizar a previsão (caso este seja o intuito do pesquisador).

4.1 A verificação da hipótese de passeio aleatório

De acordo com Charles e Darné (2009), os testes que envolvem a hipótese de que as séries temporais se comportam como um passeio aleatório se caracterizam como uma forma de avaliar se os retornos são imprevisíveis, o que é muitas vezes interpretado como indicativo de mercados fracamente eficientes.

Conforme cada uma destas três versões de passeio aleatório citadas no capítulo 2 é possível categorizar algumas das metodologias de teste mais comumente utilizadas na literatura de finanças, sumarizadas no quadro abaixo.

Versão	Incrementos	Testes
Random Walk I	iid	- Estatísticos Tradicionais - Sequências e reversões e <i>runs</i>
Random Walk II	independentes	- Regras de Filtragem - Análise Técnica
Random Walk III	descorrelacionados	- Coeficientes de Autocorrelação - Estatísticas de Portmanteau - Razão de Variâncias

Fonte: Lo e Mackinlay (1997), elaboração própria

Quadro 2 – Versões do modelo de Random Walk.

Os tipos mais antigos de testes de *random walk* são aqueles voltados para as hipóteses características dos processos RWI e RWII e na atualidade, servem em sua maioria para interesses históricos (CAMPBELL, J. Y., LO, A. W. e MACKINLAY, A. C., 1997).

Os primeiros métodos para testar a hipótese dos mercados eficientes do ponto de vista aqui definido (ou seja, entendida a partir do sinônimo com a hipótese de que a série segue um passeio aleatório) foram os testes de correlação serial, por Fama (1965) e de análise espectral, em Granger e Morgensten (1963). Ao longo dos anos, muitos outros testes foram desenvolvidos e as metodologias comumente utilizadas para verificar a hipótese de *random walk* na atualidade são: análise de correlação serial; de raiz unitária e de razão de variância.

4.1.1 A estacionariedade da série

Um processo estocástico estacionário é, de maneira geral, aquele apresenta primeiro e segundo momentos (média e variância) constantes ao longo do tempo e cuja covariância entre dois períodos de tempo depende somente da defasagem (do intervalo) entre os mesmos e não do próprio tempo em que a covariância é calculada (Gujarati, 2004).¹⁹ Em uma série temporal que possui estas características é possível observar

¹⁹ A rigor processos assim descritos, com os dois primeiros momentos de sua distribuição de probabilidades (a média e a variância) constantes ao longo do tempo são disto *fracamente estacionários*. Os processos estacionários no sentido estrito são aqueles em que todos os momentos da distribuição de probabilidades são constantes ao longo do tempo. Gujarati ressalta ainda que, para o caso de o processo estacionário apresentar distribuição normal, o processo fracamente estacionário é também estritamente estacionário uma vez que um processo normal é plenamente definido por seus dois primeiros momentos. (Gujarati, 2004).

tendência de reversão à média assim como flutuações em torno da média (mensuradas pela variância) com amplitude mais ou menos constante.

A relevância da verificação da estacionariedade da série diz respeito ao fato de que em processos não estacionários (aqueles em que a média e/ou a variância variam com o tempo) somente é possível avaliar o comportamento da série para o período considerado já que esta varia ao longo de diferentes intervalos temporais. Neste sentido, cada conjunto de dados torna-se um evento específico de tal forma que a generalização das características observadas para demais períodos e, portanto, a realização de análises com a finalidade de previsões não é confiável.

Além de processos estacionários e não estacionários, existe ainda um tipo especial de processo estocástico, que é puramente aleatório, também chamado de ruído branco, ou processo de Wiener (para o caso de o ruído branco apresentar distribuição normal). Em geral, em modelos econométricos de séries temporais, é comum admitir que o termo de erro é aleatório, ou seja, segue um ruído branco, hipótese que a rigor, deve também ser avaliada.

4.1.1.1 Análise da função de autocorrelação

Considerando-se o caso em que os dados analisados sigam um processo *random walk* descrito pela versão RWIII acima, isto implica que os incrementos (ou as primeiras diferenças da variável em nível) são descorrelacionados para quaisquer intervalos passados de tempo (*lags*). Sendo assim, pode-se testar se a série analisada segue um processo tipo RWIII por meio da hipótese nula de que os coeficientes de autocorrelação das primeiras diferenças em vários *lags* são iguais a zero.

Esta ideia aparentemente simples é a base para uma variedade muito extensa e testes sobre processos tipo *random walk*. Como destacam Campbell, Lo e Mackinlay (1997) alguns testes podem ser baseados nos próprios coeficientes de correlação, enquanto outros podem estar fundamentados na soma dos quadrados das autocorrelações e mesmo na combinação linear das autocorrelações (que podem ser adequadas para identificação de aspectos particulares).

Quando se trabalha com uma série estacionária univariada, a análise dos coeficientes de correlação ou da função de autocorrelação (que relaciona a autocorrelação com a defasagem do modelo analisado) oferece indicativos sobre o tipo de processo gerador da série temporal em questão.

A função de autocorrelação populacional (FAC) com defasagem k é denotada por ρ_k :

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad (4.1)$$

onde

γ_k : covariância com defasagem k ,

γ_0 : variância.

Observa-se assim que a função de autocorrelação oferece um número puro sem unidades dado que a variância e a autocovariância apresentam a mesma unidade de medida, podendo variar no intervalo entre -1 e 1. Na prática, no entanto, estão em geral disponíveis somente dados amostrais, de tal forma que a função de autocorrelação amostral $\hat{\rho}_k$ é definida de forma análoga à populacional, a partir da razão entre a covariância amostral com defasagem k , dada por $\hat{\gamma}_k$ e a variância amostral $\hat{\gamma}_0$.

Após o cálculo da FAC para diferentes defasagens do modelo é possível se obter um gráfico da autocorrelação amostral $\hat{\rho}_k$ para cada defasagem e desta forma, como ressalta Bueno (2011), a análise gráfica da autocorrelação pode permitir não somente uma melhor caracterização do processo gerador dos dados avaliados como também em alguns casos a identificação das ordens dos mesmos. Isso se deve ao fato de que, a depender da série temporal a que o correlograma se refere, são observados diferentes padrões para o coeficiente de autocorrelação.

Em séries temporais que seguem um processo do tipo ruído branco, observa-se que os coeficientes de autocorrelação para diferentes defasagens variam em torno de zero. Já para séries temporais não-estacionárias, observam-se coeficientes de autocorrelação com valores muito altos para as defasagens iniciais tendendo a se reduzir para zero conforme o número de defasagens vai aumentando.

Esse decaimento a zero pode ser brusco, como observado em processos do tipo médias móveis (*MA*) a partir de certa defasagem q ; exponencial, como é o caso de processos Autorregressivos (*AR*) ou apresentar diversas configurações, caso típico de processos Autorregressivos com médias móveis *ARMA*.

Este decaimento para zero dos coeficientes de autocorrelação em processos autoregressivos está relacionada à correlação implícita entre as variáveis de preço no período corrente e nos períodos anteriores presentes no modelo como, por exemplo, X_t ,

e X_{t-1} em (2.21). Todavia, Bueno (2011) ressalta a possibilidade de que estas correlações implícitas sejam eliminadas por meio do processo de filtragem que envolve a geração da função de autocorrelação parcial (FACP), segundo a qual seria possível avaliar somente as correlações puras entre duas observações. Desta maneira, a análise da FACP truncada a partir de certa defasagem p seria mais indicada para avaliar a defasagem de um processo, enquanto que a análise da FAC permitiria a determinação da defasagem de um processo modelado por uma equação de médias móveis.

Estas análises, entretanto, devem ser utilizadas como uma regra de bolso, já que em alguns casos, não oferecem a determinação exata da defasagem de truncagem. Para tanto seria necessária além da análise gráfica, uma avaliação do coeficiente de autocorrelação $\hat{\rho}_k$ em termos de sua significância para cada defasagem, o que poderia ser realizado a partir de uma análise do erro-padrão. Outra possibilidade além da análise individual de cada coeficiente de correlação seria a avaliação da hipótese conjunta de que todos os coeficientes $\hat{\rho}_k$ até certa defasagem são iguais a zero, o que implica em testar se os resíduos do modelo estimado para a série temporal seguem um ruído branco. Este teste pode ser realizado por meio da análise da estatística Q desenvolvida por Box e Pierce, ou uma variante desta denominada estatística de Ljung e Box (ambas conhecidas como estatísticas *Portmanteau*).

Diante das dificuldades em termos da exatidão dos resultados obtidos a partir da análise da FAC e FACP, a identificação do modelo na prática deve ser realizada acompanhada da estatística de Ljung e Box (ou estatísticas *Portmanteau*) sobre os resíduos do modelo estimado para confirmação dos resultados.

Conforme ressalta Bueno, a não rejeição da hipótese nula de ausência de autocorrelação dos resíduos por meio da FAC, FACP e Ljung e Box, indica que os resíduos são um ruído branco.

Todavia, como indicado no capítulo dois, Fava e Alves (1998) ressaltam que, para os casos de processos que apresentam memória longa, a análise da função de autocorrelação pode gerar conclusões equivocadas, na medida em que a série original (em nível) aparenta ser não estacionária, $I(d \neq 0)$, e a série em diferenças pode apresentar ordem de integração “superdiferenciada” (MORETTIN, P. A. 2011).

4.1.1.2 Análise da independência

As estatísticas *Portmanteau*, avaliadas no âmbito do tempo²⁰, caracterizam-se como testes de hipóteses que buscam identificar se as observações em análise (seja a série em nível ou defasada), assim como os resíduos de um modelo estimado, como por exemplo, os incrementos ε_t em (2.21), podem ser considerados estatisticamente independentes.

A estatística Q desenvolvida inicialmente por Box e Pierce (1970) e depois modificada por Ljung e Box (1978) para Q^2 são baseadas na função de autocorrelação parcial (FACP) dos resíduos do modelo estimado. A estatística Q pode ser definida como:

$$Q_m = T \sum_{k=1}^m \rho^2(k). \quad (4.2)$$

na qual,

T : número de observações

ρ^2 : coeficiente de autocorrelação,

k : número de defasagens,

m : número de autocorrelações.

Nesse sentido, sob a hipótese de que de que o processo segue um RWI dado que $\varepsilon_t \sim IID(0, \sigma^2)$, tem-se que o estimador $\hat{Q} = T \sum_{k=1}^m \hat{\rho}^2(k)$ em amostras grandes segue assintoticamente uma distribuição de probabilidades χ_m^2 , ou seja, qui-quarado com m graus de liberdade.

Na prática, sob a hipótese nula de que o processo analisado segue um ruído branco, se a estatística \hat{Q} calculada for superior ao valor de Q na distribuição qui-quadrado ao nível de significância escolhido, isto implica que ao menos um dos coeficientes de autocorrelação é diferente de zero e que os resíduos do modelo estimado não seguem um ruído branco.

Ljung e Box (1978) desenvolveram uma correção para amostras finitas que levam a um ajuste melhor à distribuição χ_m^2 para amostras pequenas, tal que:

²⁰ As estatísticas Portmanteau podem ser avaliadas não somente no domínio o tempo, segundo a qual considera-se a função de autocorrelação dos resíduos, como também no domínio da frequência. Neste caso, as estatísticas baseiam-se na função densidade espectral (PRIESTLEY, M. B. 1981).

$$LB = Q_m^* = T(T+2) \sum_{k=1}^m \frac{\rho^2(k)}{T-k}, \quad (4.3)$$

onde:

T : T : número de defasagens,

ρ^2 : coeficiente de autocorrelação,

k : número de defasagens,

m : número de autocorrelações.

Ao realizar a soma dos quadrados das autocorrelações a estatística de Box-Pierce é capaz de detectar desvios de autocorrelações não nulas para quaisquer direções (positiva ou negativa) e para qualquer *lag* temporal. Sendo assim, é adequada para uma série de suposições alternativas à hipótese de *random walk*, desde que com o cuidado da escolha do número de autocorrelações avaliadas (CAMPBELL, J. Y., LO, A. W. e MACKINLAY, A. C., 1997).

4.1.1.3 Análise de raiz unitária

Na análise de séries temporais, a presença de uma ou mais raízes unitárias é um indicativo da ausência de estacionariedade. Desta forma, um teste de raiz unitária sobre a série temporal pode auxiliar na determinação se a série é estacionária em nível ou em diferenças.

Considerando um processo estocástico dado por:

$$Y_t = \phi Y_{t-1} + \varepsilon_t, \quad (4.4)$$

$$\phi(B)(1-B)^d (Y_t - \mu) = \varepsilon_t, \quad (4.5)$$

onde

$(1-B)$: operador de *lag* (tal que B indica a defasagem do processo),

d : parâmetro de diferenciação que assume valores inteiros,

$\phi(B)$: polinômio de grau p onde

$$\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p), \quad (4.6)$$

ε_t : ruído branco com média zero e variância σ_ε^2 .

Para os processos autorregressivos, a condição para que Y_t seja estacionário é que todas as raízes de $\phi(B)=0$ estejam fora do círculo unitário, portanto, que $|\phi| < 1$ (MORETTIN, P. A., 2011).

Sendo assim, em um sentido amplo, a hipótese geral de testes por traz de uma análise de raiz unitária é de que $\phi = 1$, o que indica que o processo apresenta raiz unitária contra a hipótese alternativa de que $|\phi| < 1$, tal que o processo é estacionário, logo²¹:

$$\begin{cases} H_0 : \phi = 1 \Rightarrow Y_t \sim I(1) \\ H_1 : |\phi| < 1 \Rightarrow Y_t \sim I(0) \end{cases} \quad (4.7)$$

Os testes mais comumente utilizados para avaliar a presença de raiz unitária são DF e ADF de Dickey e Fuller (1979, 1981), com a ressalva de que admitem que há no máximo uma raiz unitária, de Dickey e Pantula, que sugere procedimentos para o caso de a série apresentar no máximo duas raízes unitárias, e de Phillips e Perron (1988), que supõe erros correlacionados e possivelmente heterocedásticos.

Morettin (2011) ressalta que os testes de raiz unitária, todavia, apresentam uma série de problemas, dentre eles a baixa potência²² em especial quando se procede a avaliação de processos com memória longa (BUENO, R. L S., 2011).

Segundo Fava e Alves (1998) os testes de raiz unitária, ao apresentarem baixa potência na identificação desta característica, podem levar a conclusões errôneas na medida em que não conseguem distinguir os processos integrados de ordem um, $I(1)$, dos processos com ordem de integração fracionária. De acordo com os autores diversos estudos analisaram a questão da potência dos testes frente a integração fracionária e demonstraram que o teste KPSS (Kwiatkowski, Phillips, Schmidt e Shin, 1992) são mais consistentes (para amostras grandes, em torno de 1000 observações) em relação à hipótese de que a série segue um processo fracionalmente integrado ou seja, que a série é $I(d)$, com $-0,5 < d < 0,5$.

Conforme Fava e Alves (1998), o poder mais elevado do teste KPSS para a identificação de raiz unitária diante de séries fracionalmente integradas pode estar relacionado ao fato de que sua hipótese nula é de que a série é estacionária²³.

²¹ A partir desta configuração percebe-se a razão para a nomenclatura para estes testes, na medida em que são assim denominados devido ao fato de que, sob a hipótese nula, o polinômio autorregressivo possui uma raiz unitária igual à unidade.

²² Segundo o autor, o baixo poder do teste está relacionado ao fato de que com bastante frequência o teste induz ao erro tipo II, que implica em não rejeitar a hipótese nula quando ela é falsa.

Nesse sentido, em oposição aos testes supracitados, o teste KPSS, desenvolvido por Kwiatkowski, Phillips, Schmidt e Shin (1992) é um teste de estacionariedade cuja hipótese nula é de que o processo analisado é estacionário, contra a hipótese alternativa de que o processo é integrado de ordem um, ou seja:

$$\begin{cases} H_0 : Y_t \sim I(0) \\ H_1 : Y_t \sim I(1) \end{cases} \quad (4.8)$$

Segundo Bueno (2011), diante da baixa potência dos tradicionais testes de raiz unitária, muitos autores utilizam os resultados do teste KPSS como uma forma de complementar os testes de raiz unitária tradicionais.

No entanto, diante da possibilidade de que a série apresente memória de longo prazo, é interessante a aplicação de métodos próprios para a avaliação de séries fracionalmente integradas, conforme ressaltado no capítulo 2.

4.1.2 A normalidade

Um aspecto comumente analisado na literatura no que diz respeito aos modelos do tipo *random walk* é a assunção que os incrementos seguem algum tipo de distribuição específica, sendo a mais utilizada, a gaussiana, tal que $\varepsilon_t \sim RBN(0, \sigma^2)$.

Desta forma, quando se assume além da normalidade dos incrementos, que o processo em análise segue um RWI, obtém-se um modelo tipo movimento browniano aritmético (em intervalos regulares de tempo), dado por:

$$P_t = \mu + P_{t-1} + \varepsilon_t \quad \varepsilon_t \sim IID \quad N(0, \sigma^2). \quad (4.9)$$

A adoção da hipótese da normalidade dos incrementos, todavia, gera uma consequência importante para as análises de finanças e diz respeito à mesma questão levantada por Bachelier (1900) referente à hipótese da normalidade dos retornos: a probabilidade positiva de existência de preços negativos, ou seja, $P(P_t < 0) \geq 0$.

Diante disto, o ideal seria uma análise a partir do logaritmo natural dos preços (como levantado também por Osborne, 1959), tal que a hipótese adotada seria:

$$p_t = \mu + p_{t-1} + \varepsilon_t \quad \varepsilon_t \sim IID \quad N(0, \sigma^2). \quad (4.10)$$

onde:

$$p_t = \ln P_t$$

²³ Em detrimento da hipótese nula dos testes DF, ADF, de Dickey e Pantula, e de PP, para os quais a hipótese nula é de que a série apresenta raiz unitária.

Sendo assim, os retornos continuamente compostos são independentes e identicamente distribuídos (conforme uma distribuição normal), com média μ e variância σ^2 , o que leva aos modelos adotados tanto por Bachelier (1900) e Einstein (1905).

Para a verificação da normalidade nas séries avaliadas foram analisados os coeficientes de assimetria e curtose, assim como o teste de Jarque e Bera (1987).

Este último trata-se de um teste assintótico (de grandes amostras), que parte da análise dos resíduos estimados pelo método dos mínimos quadrados ordinários. Por meio de uma análise comparativa da assimetria e da curtose da distribuição dos resíduos do modelo estimado em relação a uma distribuição normal, o teste de Jarque e Bera busca verificar se os momentos estimados são iguais aos da distribuição gaussiana.

O grau de assimetria S indica a posição da distribuição em relação à média e é dada pela razão entre o terceiro momento da distribuição e o desvio padrão elevado ao cubo:

$$S = \frac{E(X - \mu)^3}{\sigma^3}. \quad (4.11)$$

Sabe-se que, para o caso de uma variável normalmente distribuída, todos os momentos em torno do valor médio de ordem ímpar, superiores a dois, são iguais a zero. Sendo assim, espera-se que para uma variável que segue uma distribuição gaussiana, o valor do coeficiente de assimetria deve ser nulo (GUJARATI, D. 2004).

Já o coeficiente de curtose K indicativo do grau de achatamento da distribuição, é dado pela razão entre o quarto momento em torno da média e o quadrado do segundo momento, ou seja:

$$K = \frac{E(X - \mu)^4}{\sigma^4}. \quad (4.12)$$

Sendo assim:

Curtose (K)	Tipo da distribuição	Característica da Distribuição
K < 3	Platicúrticas	- Caudas curtas; - Distribuição mais achatada que a gaussiana.
K = 3	Mesocúrticas	- Distribuição gaussiana.
K > 3	Leptocúrticas	- Caudas longas (<i>fat tails</i>); - Distribuição mais magras que a gaussiana.

Fonte: Gujarati, D. (2004). Elaboração própria.

Quadro 3 – Coeficiente de Curtose.

Conforme o quadro acima nota-se que para o caso de uma variável que segue uma distribuição gaussiana, espera-se que sua curtose seja igual a três.

Diante disto, a lógica por trás do teste de normalidade de Jarque-Bera (JB) é justamente a verificação da hipótese conjunta de que os coeficientes de assimetria e de curtose são iguais às suas normas zero e três, respectivamente.

Testa-se, então a estatística abaixo, sob a hipótese nula da normalidade, onde n é o tamanho da amostra.

$$JB = n \left[\frac{S^2}{6} + \frac{(K - 3)^2}{24} \right]. \quad (4.13)$$

Sob a hipótese nula de que a série em análise é normalmente distribuída, Jarque e Bera demonstraram que a estatística (JB) segue, em grandes amostras, a distribuição χ^2 com dois graus de liberdade.

4.1.3 A linearidade da série

Conforme Zivot e Wang (2005) o sistema econômico e financeiro é em geral marcado por mudanças estruturais e comportamentais significativas que se refletem no comportamento não linear das séries temporais resultante em geral do aparecimento de diferentes dinâmicas e mudanças de regimes nas variáveis que compõem estes sistemas.

Segundo Morettin (2011) a ausência de linearidade na série temporal indica que a resposta a choques grandes ou pequenos, positivos ou negativos, se dá de maneira diferente. Tradicionalmente na análise de séries temporais financeiras, nota-se que a variância de um processo aumenta diante de choques negativos nos retornos, enquanto que este aumento não é tão pronunciado frente a choques positivos.

No entanto grande parte dos modelos mais tradicionais utilizados para a modelagem de séries temporais financeiras é linear e não captam este comportamento não linear de resposta das variáveis.

Como exemplo, conforme Campbell, Lo e MacKinlay (1997) ressaltam, as três diferentes versões de *random walk* citadas no capítulo 2 são modelos lineares, na medida em que apresentam uma importante propriedade que indica que a variância dos incrementos de um processo *random walk* deve ser uma função linear dos intervalos de tempo, ou seja, a variância amostral dos retornos em um período k de uma série Y_t é igual a m vezes a variância amostral do retorno de um período. Nesse sentido, é

possível afirmar que a soma das variâncias dos incrementos deve ser igual à variância da soma (CAMPBELL, J. Y., LO, A. W. e MACKINLAY, A. C., 1997).

Com base nesta propriedade de linearidade da variância é que se fundamentam os testes de razão de variância (ou *variance ratio tests*).

Considerando uma série temporal dos retornos Y_t de um determinado ativo em um intervalo de tempo $t = 1, 2, \dots, T$, a razão de variância da k-ésima diferença em relação à primeira diferença pode ser representada por:

$$VR(k) = \frac{\sigma^2(k)}{\sigma^2(1)} \quad (4.14)$$

na qual:

$\sigma^2(k)$: é igual a $\frac{1}{k}$ da variância da k-ésima diferença,

$\sigma^2(1)$: variância da primeira diferença.

Sendo assim, sob a hipótese nula de que a série em questão segue um processo tipo *random walk*, com variância constante, a razão $VR(k)$ deve ser aproximar de um. Caso a razão de variância for menor do que um para horizontes temporais mais longos, há indícios de correlação serial negativa o que indica reversão à média, e no caso de razões superiores a um há indícios de correlação serial positiva indicando persistência ou aversão a média (CHARLES, A. e DARNÉ, O. 2009).

Zivot e Wang (2005) ressaltam porém que, apesar de grande parte dos modelos tradicionais ser linear, existe uma gama muito grande de modelos voltados para o tratamento de séries não lineares. Segundo os autores estes modelos são bastante flexíveis e por isso é possível que se cometam erros no tratamento das séries derivados da possibilidade de se obter um bom ajuste aos modelos não lineares mesmo diante de séries lineares.

Além dos testes de razão de variância, Zivot e Wang (2005) indicam o teste BDS como uma das técnicas mais populares para avaliar a linearidade versus a não linearidade da série temporal em função de possuir alto poder frente a uma diversidade de modelos não lineares e da ausência da necessidade de uma hipótese prévia de que a série apresenta uma determinada distribuição característica.

O teste BDS (Brock, Dechert e Scheinkman, 1987) foi desenvolvido inicialmente para a análise da hipótese nula de que a série temporal apresenta variáveis

i.i.d. contra a hipótese alternativa de que existe alguma estrutura de dependência nas variáveis, seja ela linear ou não, com o propósito de avaliar dinâmicas caóticas não aleatórias.²⁴ No entanto, conforme Matias, Silva e Vieira (2007), o teste demonstrou ser capaz de detectar diversas situações em que as variáveis não são i.i.d. além do caos determinístico, tais como a ausência de estacionariedade e a não-linearidade.

O teste baseia-se na análise da integral de correlação²⁵, que é uma medida da frequência com que padrões temporais se repetem nos dados. Em condições moderadas de regularidade, a estatística do teste converge em distribuição a uma gaussiana com média zero e variância igual a um, tal que a hipótese nula de i.i.d. pode ser rejeitada com 5% de significância quando o módulo da estatística é maior do que 1,96.

Zivot e Wang (2005) destacam ainda que é relevante avaliar se existe um comportamento não linear da série econômica em análise de modo a que se alcance uma modelagem mais adequada a esta dinâmica particular caso necessário.

4.2 A identificação de uma série temporal fracionalmente integrada

Segundo Matias et al (2007), a fractalidade em séries temporais pode ser identificada diante da observação de não-linearidade; distribuição não-gaussiana dos retornos em relação à média; a persistência de memória de longo prazo, e a existência de ordem de integração fracionária, o que exige a aplicação de testes estatísticos específicos. Os testes de linearidade e de normalidade mais utilizados foram descritos nas seções anteriores deste capítulo entretanto, conforme citado no capítulo 3, existem duas técnicas consagradas para avaliar se uma série temporal apresenta a memória longa sendo elas a análise da estatística R/S (*rescaled range*) acompanhada da determinação do expoente de Hurst H , e a utilização de modelos de integração fracionária que exigem a estimação da ordem de integração d fracionária do processo (chamada às vezes de parâmetro de longo prazo).

4.2.1 A estatística R/S e a determinação do expoente de Hurst

Desde o trabalho de Peters (1989) uma das mais difundidas técnicas para a detecção da fractalidade nas séries temporais tem sido a adaptação do método

²⁴ Segundo Zivot e Wang (2005), uma série temporal é dita caótica se segue um processo determinístico não linear, mas que aparenta ser aleatório.

²⁵ Mais detalhes podem ser obtidos em Zivot e Wang (2005).

desenvolvido pelo hidrólogo Harold Edwin Hurts chamado *rescaled range analysis*, ou R/S de Hurst e que envolve a determinação do expoente H .

A pesquisa de Hurst em hidrologia²⁶, que envolvia a previsão das cheias do rio Nilo e também o nível das mesmas, avaliou a relação entre o número de medições anuais e a amplitude das máximas e mínimas de cada registro e constatou que a amplitude das cheias se dava, conforme uma lei de potência (potência de 0,73 do desvio padrão da distribuição empírica). Estes resultados indicavam que o processo analisado não seria adequadamente representado por um modelo de *random walk*, posto que, neste caso, o crescimento da amplitude das cheias deveria ocorrer conforme uma lei multiplicativa segundo a raiz quadrada do desvio padrão da distribuição (MANDELROT, B. e HUDSON, R., 2004). Hurst elaborou então uma fórmula baseada em uma lei de potência que dava conta do entendimento deste processo e, em estudos posteriores, notou que a mesma era capaz não somente de explicar as cheias do Nilo, mas também outros processos diversos, como o padrão anual das chuvas em Nova Iorque, o crescimento dos anéis de árvores, entre outros (MANDELROT, B. e HUDSON, R., 2004).

Em 1963, sob uma analogia ao estudo de Hurst, segundo a qual as cheias do Nilo seriam como grandes saltos de preços e as grandes secas seriam como desastrosos *crashes* de mercados, Mandelbrot (1963) avaliou o comportamento dos preços do algodão. A partir deste artigo, Mandelbrot colaborou fortemente para o avanço das pesquisas sobre fractalidade e a aplicação da análise de R/S de Hurst nos estudos dos mercados financeiros visando a identificação de processos de memória longa.

Peters (1989) adaptou então a fórmula para a análise dos retornos mensais das séries do índice S&P500 e de títulos do Tesouro americano e também do retorno relativo de ambas as séries no período que vai de 1950 a 1988. Em seu artigo concluiu que para as séries analisadas foi observada a persistência de memória de longo-prazo entre as observações o que indicaria dependência entre elas, tal que as informações passadas são relevantes para a explicação das observações presentes, uma característica de que as séries temporais avaliadas apresentavam memória longa, podendo ser caracterizadas como fracionárias.

²⁶ Realizada na década de 1950, tinha o objetivo de conceber um processo chamado de “armazenagem centenária” que visava à construção de um reservatório adequado para o acúmulo da água necessária para prevenir o Egito dos problemas causados pelas possíveis secas.

Alguns anos depois Andrew W. Lo (1991), identificou que, na presença de heterogeneidade estatística na série analisada, a estatística R/S apresentaria viés, de tal forma que os testes poderiam deturpar a memória longa com os efeitos da memória mais recente, não sendo, portanto, robusta à memória de curto prazo. Lo sugeriu um refinamento na técnica (Mandelbrot, 2004). Conforme Morettin (2011), a estatística R/S pode ser definida por:

$$\tilde{Q}_T = \frac{1}{\hat{\sigma}_T(q)} \left[\max_{1 \leq k \leq T} \sum_{j=1}^k (X_j - \bar{X}) - \min_{1 \leq k \leq T} \sum_{j=1}^k (X_j - \bar{X}) \right], \quad (4.15)$$

onde:

$$X = \{X_1, \dots, X_T\},$$

\bar{X} : é a média amostral,

$\hat{\sigma}_T^2(q)$: é o estimador da variância de longo prazo de Newey-West com largura de faixa (*bandwidth*) q descrito por:

$$\hat{\sigma}_T^2(q) = S_T^2 \left(1 + \frac{2}{T} \sum_{j=1}^q w_{qj} r_j \right), \text{ tal que} \quad (4.16)$$

$$w_{qj} = 1 - \frac{j}{(q+1)}, \quad q < T, \text{ e} \quad (4.17)$$

$$q = \left[4 \left(\frac{T}{100} \right)^{2/9} \right]^{27}, \quad (4.18)$$

onde:

S_T^2 : variância amostral,

q e r_j são as autocorrelações amostrais usuais de X_t .

Segundo Morettin (2011), na ausência de memória longa a estatística R/S dada por (4.15) deve convergir para a distribuição limite segundo uma taxa de $T^{1/2}$. Todavia, na presença de memória longa a taxa de convergência será de T^H onde H é o expoente de Hurst.

Tradicionalmente procede-se a construção de um gráfico na escala log-log de R/S versus o tamanho da amostra, no qual se espera que, para séries de memória curta

²⁷ Segundo sugestão de Newey West (1987). Alguns pacotes de softwares estatísticos utilizam outros expoentes, como é o caso do S+FinMetrics em que o expoente utilizado na função padrão é igual a 0,25 (MORETTIN, P. A. 2011).

os pontos devem se situar ao longo da reta com inclinação igual a 0,5, enquanto que para as séries de memória longa a inclinação da reta deve ser igual a $H > 0,5$ para grandes amostras (MORETTIN, P. A., 2011).

4.2.2 A mensuração da dimensão fractal

Conforme descrito no capítulo 3 a dimensão fractal D pode ser calculada a partir da equação (3.14) após a determinação do expoente de Hurst H .

4.2.3 Método GPH

O método GPH²⁸, desenvolvido por Geweke e Porter-Hudak (1983) visa a estimação do parâmetro de memória longa d , ou seja, a ordem de integração fracionária de um processo dado pela equação (3.9).

Baseado na representação de um processo fracionalmente integrado com memória longa dado pela equação (3.9), os autores propuseram uma abordagem semi-não paramétrica para testar memória longa fundamentado em uma análise no domínio da frequência²⁹.

Sob a hipótese nula de que a ordem de integração do processo é igual a zero, sua rejeição implica afirmar que a distribuição não apresenta memória longa, podendo ser assim fracionalmente integrada.

Segundo Morettin (2011) é possível também estimar diretamente a ordem de integração do processo pelo método da máxima verossimilhança a partir da estimação de um processo $ARFIMA(p, d, q)$ que é na realidade um $ARIMA(p, d, q)$ ajustado para uma série temporal com memória longa, a ser abordado na seção 4.3.

4.2.4 Método do Periodograma

Conforme abordado no capítulo 3, é possível descrever a função de autocorrelação de um processo com memória longa a partir de sua densidade espectral, segundo a equação (3.8). Levando-se em conta (3.7), Zivot e Wang (2005)

²⁸ O método GPH é também conhecido como método de regressão do periodograma, sendo este um estimador do espectro $f_u(\lambda)$ do processo.

²⁹ Mais detalhes podem ser obtidos em Morettin (2011) e Zivot e Wang (2005).

demonstram que para processos de memória longa a densidade espectral $f(\omega)$ tende para $C_f \omega^{1-2H}$ quando a frequência tende para zero, logo:

$$f(\omega) \rightarrow C_f \omega^{1-2H} \text{ quando } \omega \rightarrow 0, \text{ com } C_f > 0. \quad (4.19)$$

Como a densidade espectral pode ser estimada a partir do periodograma³⁰, a análise do gráfico log-log do periodograma versus a frequência deve revelar, para processos com memória longa, maior concentração dos dados com frequência próxima de zero em torno da reta com inclinação $1 - 2H$.

Sendo assim, o método pode também ser utilizado para a estimação do coeficiente de Hurst H , a partir do qual se obtém os valores do parâmetro d .

4.2.5 Método de Whittle

Outra técnica de estimação da ordem de integração fracionária do processo no domínio da frequência é o método de Whittle, cuja derivação pode ser obtida em Beran (1994).

A partir de (3.9) Zivot e Wang (2005) destacam que é possível estimar, por meio de máxima verossimilhança, o parâmetro d via minimização de uma versão discreta de:

$$Q(\theta) = \int_{-\pi}^{\pi} \frac{I(\omega)}{f(\theta; \omega)} d\omega, \quad (4.20)$$

onde:

θ é o vetor de parâmetros desconhecidos a serem estimados (incluindo a ordem de integração d ,

$I(\omega)$ é o periodograma de X_t ,

$f(\theta; \omega)$ é a densidade espectral teórica de X_t .

Zivot e Wang (2005) destacam que apesar de o método ter sido definido para um processo fracionalmente integrado geral, como em (3.9), sua implementação em softwares estatísticos (como o Splus utilizado neste estudo) assume que os resíduos do modelo estimado apresentam distribuição normal. Isto implica que o processo X_t em (3.9) com resíduos assim considerados, segue um processo $ARFIMA(0, d, 0)$, ou seja, um ruído branco fracionário.

³⁰ Conforme Morettin (2011), o periodograma é um estimador do espectro de um processo.

4.3 Estimação de modelos com integração fracionária

As seções anteriores abordaram alguns métodos de teste para detecção de memória longa e estimação dos parâmetros H e d indicativos da fracionalidade (ou fractalidade) da série temporal.

Nesta seção, busca-se introduzir brevemente a estimação de alguns modelos representativos de séries temporais fracionárias, que são capazes de refletir tanto o comportamento de memória de curto prazo quanto de longo prazo quando presentes em séries temporais estacionárias.

Conforme ressaltam Zivot e Wang (2005), a técnica tradicionalmente utilizada para a modelagem de séries temporais estacionárias integradas de ordem zero parte inicialmente da consideração de um modelo geral do tipo ARIMA dado pela equação (3.9), tal que o parâmetro d assume somente valores inteiros maiores ou iguais a zero.

Ao permitir que o parâmetro d assuma valores fracionários, o tradicional modelo ARIMA torna-se um modelo ARFIMA ou FARIMA, portanto, autorregressivo fracionalmente integrado com médias móveis. Para diferentes valores do parâmetro d é possível que o modelo estimado apresente um determinado comportamento característico, sendo estacionário somente no intervalo em que $-0,5 < d < 0,5$.

Conforme Zivot e Wang (2005), Beran (1995) estendeu a estimação dos modelos $ARFIMA(p, d, q)$ para qualquer $-0,5 < d$ admitindo a seguinte variante do processo $ARFIMA(p, d, q)$:

$$\phi(B)(1-B)^\delta[(1-B)^m Y_t - \mu] = \theta(B)\varepsilon_t, \quad (4.21)$$

onde $-0,5 < \delta < 0,5$, m é o número de diferenciações necessárias para que se obtenha uma série estacionária e as demais variáveis são definidas conforme a equação (3.9).

Sendo assim, a ordem de integração é dada por

$$d = \delta + m, \quad (4.22)$$

tal que, quando $m = 0$, μ é a esperança Y_t , enquanto que para $m = 1$, μ é o coeficiente angular do componente de tendência linear em Y_t ((ZIVOT, E. e WANG, J., 2005).

Segundo Morettin (2011), conforme aumentam os *lags* temporais considerados o efeito do parâmetro d sobre estas observações mais distantes decai de maneira hiperbólica, enquanto que os efeitos dos parâmetros ϕ e θ decaem exponencialmente.

Sendo assim, a partir desta resposta diferenciada dos parâmetros é possível que modelos do tipo ARFIMA captem tanto as estruturas de correlação de ordens altas da série (os efeitos de memória longa) quanto as estruturas de correlação de ordens baixas (os efeitos de memória curta) que podem existir.

No presente estudo foi utilizada a função FARIMA do S+Finmetrics para ajustar um modelo ARFIMA sobre os log-retornos das diferentes *commodities* e a identificação das ordens p e q que compõem o modelo se deu em função da minimização do critério BIC³¹.

Além da estimação de um modelo fracionalmente integrado no âmbito dos retornos é possível que seja encontrada memória longa também na volatilidade de séries temporais³². Diante deste comportamento, Zivot e Wang (2005) destacam que é possível aplicar os modelos ARFIMA para a modelagem da variância de um processo de forma independente.

Em geral as séries de retornos ao quadrado ou de retornos absolutos são utilizadas como estimativas para a volatilidade. Todavia, conforme ressaltam Zivot e Wang (2005), no caso da dificuldade de obtenção de uma estimativa confiável para a volatilidade ou diante da necessidade da estimação desta em conjunto com a dinâmica dos retornos os modelos da família ARCH são mais adequados.

O modelo ARCH (autorregressivos com heterocedasticidade condicional), introduzidos por Engle (1982) é dado por

$$\begin{aligned} r_t &= \sqrt{h_t} \varepsilon_t \\ h_t &= \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_m r_{t-m}^2 \end{aligned} \quad (4.23)$$

tal que $\varepsilon_t \sim i.i.d(0, \sigma^2)$ ³³, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \dots, m-1$, e $\alpha_m > 0$. Nota-se assim, que o modelo é não linear na variância, no sentido de que a volatilidade (variância condicional) depende dos retornos passados por meio de uma função quadrática (Morettin, 2011). Este modelo, todavia, não captura um aspecto característico da volatilidade de séries financeiras: a resposta diferenciada da volatilidade diante de retornos positivos e negativos (conhecido como *leverage*).

³¹ BIC trata-se de um critério de informação bayesiano que auxilia na identificação de um modelo mais adequado, favorecendo a parcimônia. Neste estudo escolheu-se portanto, o modelo que minimizasse o BIC.

³² A existência de memória longa na volatilidade indica que não deve ser possível assumir a hipótese de que a variância é homocedástica.

³³ Morettin (2011) ressalta que em geral assume-se que ε_t segue uma distribuição normal padrão ou ainda, pode se considerar que segue uma distribuição t-Student ou outra que admita caudas mais pesadas.

Bollerslev (1986) desenvolveu então o modelo $GARCH(m,n)$, uma generalização do modelo $ARCH$ que passou a considerar que a volatilidade é não somente uma função dos retornos passados ao quadrado como também das volatilidades passadas (Zivot e Wang, 2005). No entanto, sendo ainda uma função quadrática dos retornos, esta generalização não foi capaz de capturar a resposta diferente da volatilidade, de modo que trata simetricamente os retornos (MORETTIN, P. A., 2011).

Uma generalização do modelo $GARCH$ foi proposta por Nelson (1991) e trata-se do $EGARCH$, ou $GARCH$ exponencial. Seu objetivo era de inserir nos modelos de volatilidade a resposta assimétrica aos retornos positivos e negativos, tal que os retornos negativos tenham maior impacto na volatilidade, na medida em que a resposta da volatilidade aos retornos negativos é mais rápida. Isto foi alcançado a partir da admissão a volatilidade como uma função dos retornos absolutos, inserida em um modelo exponencial de modo que não existissem volatilidades negativas.

Além destes, outras generalizações possíveis são os modelos da família $GARCH$ fracionalmente integrados, derivados dos modelos $IGARCH$, caracterizados por serem modelos integrados, no sentido de que apresentam raiz unitária e uma tendência persistente.

Conforme destacam Zivot e Wang (2005) os modelos $FIEGARCH(q,d,s)$, ou seja, um $EGARCH$ fracionalmente integrado, desenvolvido por Bollerslev e Mikkenlesen (1996) e ainda o $FIGARCH(m,d,q)$, desenvolvido por Baillie, Bollerslev e Mikkenlesen (1996), partem de uma representação do modelo $IGARCH(r,s)$ como um $ARFIMA(m,d,q)$, tal que:

$$\phi(B)(1-B)^d \varepsilon_t^2 = a + b(B)u_t, \quad (4.24)$$

onde

$\phi(B)$: polinômio de grau $m = \max(p,q)$ dado por:

$$\phi(B) = (1 - \phi_1 B - \dots - \phi_m B^m), \quad (4.25)$$

$b(B)$: polinômio de grau q dado por:

$$b(B) = (1 - \theta_1 B - \dots - \theta_q B^q), \quad (4.26)$$

em que todas as raízes dos polinômios $\phi(z)=0$ e $b(z)=0$ se situam fora do círculo unitário.

Quando $d = 0$ o modelo torna-se um GARCH usual, quando $d = 1$ o modelo torna-se um IGARCH e quando $0 < d < 1$ os resíduos ao quadrado fracionalmente integrados, $(1 - B)^d \varepsilon_t^2$, seguem um processo $ARMA(m, q)$.³⁴ Zivot e Wang (2005) destacam que, reescrevendo a equação acima em termos da variância condicional σ_t^2 obtém-se o modelo $FIGARCH(m, d, q)$:

$$b(B)\sigma_t^2 = a + [b(B) - \phi(B)(1 - B)^d] \varepsilon_t^2. \quad (4.27)$$

Já o modelo $FIEGARCH(q, d, s)$ pode ser representado pela equação

$$\phi(B)(1 - B)^d \ln \sigma_t^2 = a + \sum_{j=1}^q (b_j |x_{t-j}| + \gamma_j x_{t-j}), \quad (4.28)$$

onde x_t são os resíduos padronizados, dados por

$$x_t = \frac{\varepsilon_t}{\sigma_t}. \quad (4.29)$$

Para $0 < d < 1$ o modelo FIEGARCH é estacionário e para $\gamma_j \neq 0$ o modelo admite a existência dos efeitos de *leverage* (ZIVOT, E. e WANG, J., 2005).

³⁴ Maiores detalhes podem ser obtidos em Baillie, Bollerslev e Mikkelsen (1996).

5 Resultados e Conclusões

No presente estudo foram utilizadas séries diárias do indicador de preços do mercado à vista calculado pelo CEPEA-ESALQ/USP³⁵ para as *commodities* café, boi gordo, açúcar, milho, soja (região do Paraná) e bezerro brasileiros. Para cada uma destas *commodities* foram avaliados diferentes períodos de dados disponíveis, que compreenderam um número variável de observações, descritos no quadro abaixo.

<i>Commodity</i>	Café	Boi Gordo	Açúcar	Milho	SojaPR	Bezerro
Início	2/9/1996	23/7/1997	20/5/2003	30/7/2004	29/7/1997	21/2/2000
Fim	4/7/2013	4/7/2013	4/7/2013	4/7/2013	4/7/2013	4/7/2013
nº Observações	4187	3965	2508	2219	3964	3327

Fonte: CEPEA/ESALQ, elaboração própria.

Quadro 4 – Commodities em Análise.

5.1 A verificação da hipótese de passeio aleatório

Foi analisado inicialmente o comportamento das séries em nível (preços), por meio das funções de autocorrelação e autocorrelação parcial, que demonstraram que todas as séries avaliadas eram não-estacionárias em nível, conforme gráficos disponíveis nos Anexos.

O mesmo procedimento foi realizado para a avaliação dos log-retornos, para as quais também foi identificada a ausência de estacionariedade em primeira diferença.

Realizou-se então o teste de Phillips e Perron para verificar a hipótese nula de que as séries apresentavam raiz unitária, tal que seriam $I(1)$. Para os log-retornos do café, soja e bezerro os resultados indicaram a não rejeição da hipótese nula, oferecendo mais indícios de que estas séries não eram estacionárias em primeira diferença. Já para os log-retornos do boi gordo, açúcar e milho, a hipótese nula foi rejeitada, tal que estas séries não seriam $I(1)$, sendo necessário avaliar se elas seriam integradas de ordem zero $I(0)$ ou fracionalmente integradas.

Diante da possibilidade de existência de memória longa nestas séries e da baixa potência dos testes de raiz unitária diante de séries com esta característica (conforme abordado no capítulo 4), foi realizado então o teste KPSS para avaliar a hipótese nula de

³⁵ Mais detalhes sobre a forma de composição dos indicadores diários de preços de commodities podem ser obtidos no site do Centro de Estudos Avançados em Economia Aplicada (www.cepea.esalq.usp.br).

que a série é estacionária³⁶ considerando a possibilidade de estacionariedade em torno de uma constante e em torno de uma tendência. Para ambos os casos de teste, a maior parte das *commodities* avaliadas, exceto o bezerro, não foi rejeitada a hipótese nula, havendo indícios de que as séries seriam estacionárias. Já para o caso do bezerro, quando considerada a hipótese nula de que os dados eram estacionários em torno de uma tendência linear, a hipótese nula foi rejeitada ao nível de confiança de 5%.

Notou-se, assim, que os resultados dos testes de raiz unitária e KPSS ofereceram respostas não muito conclusivas quanto à classificação das séries como integradas de ordem zero (estacionárias), integradas de ordem um (não estacionárias), ou fracionalmente integradas (podendo ser estacionárias ou não).

Foi então realizado o teste de Ljung e Box modificado para verificação da ausência de autocorrelação nas observações componentes de cada série. Para todas as *commodities* analisadas, os resultados implicaram na rejeição da hipótese nula de que não há autocorrelação, oferecendo indícios de que os dados componentes de cada série apresentam alguma relação de dependência, de tal forma que não seria possível afirmar que estas séries seguem um processo aleatório com observações independentes.

Realizou-se também o teste BDS, para verificar a hipótese de que os dados componentes das séries das *commodities* analisadas eram independentes e identicamente distribuídos, o que foi rejeitado para todas as séries de log-retornos, indicando também a existência de alguma relação de dependência entre as observações.

Além disto, avaliou-se a hipótese nula de que os dados são normalmente distribuídos por meio do teste de Jarque e Bera e da avaliação do coeficiente de assimetria e curtose das diferentes séries. Para todas as *commodities* a hipótese nula foi rejeitada, e os coeficientes indicaram um comportamento bem distante de uma distribuição gaussiana.

Um resumo dos resultados dos testes supracitados para os log-retornos segue no quadro abaixo.

³⁶ De acordo com a configuração do teste, o sentido de estacionariedade indicaria que a série seria integradas de ordem zero.

Série: Log-Retornos							
Testes de Hipótese	H0	Café	Boi Gordo	Açúcar	Milho	Soja (Paraná)	Bezerro
Jarque Bera	Dados são normalmente distribuídos	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
Ljung-Box	Ausência de Autocorrelação	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
BDS	Dados são i.i.d.	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
Phillips-Perron	Existe raiz unitária [dados são I(1)]	NÃO REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0
KPSS Test (tendência = "c")	Dados são estacionários I(0) em torno de uma constante	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0
KPSS Test (tendência = "ct")	Dados são estacionários I(0) em torno de uma tendência linear	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0 A 5%

Fonte: Resultados da pesquisa. Elaboração própria.

Tabela 1 – Resultados dos Testes de Hipótese para os Log-Retornos.

Adicionalmente foram realizados os testes de razão de variância para a verificação da hipótese nula de que as séries analisadas seguem um passeio aleatório com erros homocedásticos e também com erros heterocedásticos. As *commodities* boi gordo, açúcar, milho e soja, apresentaram para ambos os casos de teste, comportamento distante da hipótese de passeio aleatório, uma vez que é esperado uma razão de variância próxima de um, conforme pode ser observado nos gráficos que constam nos Anexos.

Para o caso do bezerro, observou-se um comportamento próximo de um passeio aleatório para *lags* temporais inferiores a 50 dias e um distanciamento do comportamento de um *random walk* para períodos temporais mais longos. Já para o café, para *lags* temporais mais curtos (inferiores a 20 dias), o gráfico indica um comportamento mais próximo de um *random walk*, em especial se considerada a existência de homocedasticidade. No entanto, diante da consideração da presença de heterocedasticidade, o comportamento do gráfico indica um ajuste mais próximo do comportamento de um passeio aleatório.

Nestes mesmos moldes foram estudadas também as séries de retornos absolutos e retornos ao quadrado, visando analisar melhor o comportamento da variância do processo.

A função de autocorrelação e a de autocorrelação parcial dos retornos absolutos demonstraram a ausência de estacionariedade, indicando que a variância do processo poderia ser variável no tempo (presença de heterocedasticidade).

Realizou-se então o teste de Phillips e Perron para verificar a hipótese nula de que as séries apresentavam raiz unitária. Para os retornos absolutos do café, boi gordo, soja e bezerro os resultados indicaram a não rejeição da hipótese nula, oferecendo mais indícios de que estas séries não eram estacionárias. Já para os retornos absolutos do

açúcar e milho, a hipótese nula foi rejeitada, tal que estas séries não seriam $I(1)$, tal que elas poderiam ser integradas de ordem zero $I(0)$ ou fracionalmente integradas.

Diante da possibilidade de existência de memória longa nestas séries e da baixa potência dos testes de raiz unitária diante de séries com esta característica (conforme abordado no capítulo 4), foi realizado então o teste KPSS para avaliar a hipótese nula de que a série é estacionária considerando a possibilidade de estacionariedade em torno de uma constante e em torno de uma tendência.

Na análise dos retornos absolutos, a hipótese nula foi rejeitada para todas as *commodities*, tal que não seriam integradas de ordem zero. Considerando o caso da estacionariedade em torno de uma constante a hipótese nula foi rejeitada ao nível de 1% de significância para o café, a soja e o bezerro, enquanto que para o milho, o açúcar e o boi gordo, rejeitou-se a hipótese nula ao nível de significância de 5%. Considerando o caso da estacionariedade em torno de uma tendência linear, a hipótese nula foi rejeitada ao nível de 1% de significância para o boi gordo, açúcar, milho, soja e bezerro, enquanto que para o café, rejeitou-se a hipótese nula ao nível de significância de 5%.

Na análise dos retornos ao quadrado, considerando o caso da estacionariedade em torno de uma constante a hipótese nula foi rejeitada ao nível de 1% de significância para o café, enquanto que para o boi gordo e o milho rejeitou-se a hipótese nula ao nível de 5%. No entanto, para o açúcar, a soja e o bezerro, a hipótese nula de estacionariedade em torno de uma constante não pôde ser rejeitada. Já no caso da estacionariedade em torno de uma tendência linear, rejeitou-se H_0 ao nível de 1% para o açúcar, milho, soja e bezerro, enquanto que para o boi gordo e o café a hipótese nula não foi rejeitada.

Da mesma forma que o caso dos log-retornos, notou-se que os resultados dos testes de raiz unitária e KPSS ofereceram respostas não muito conclusivas, quanto à classificação das séries como integradas de ordem zero (estacionárias), integradas de ordem um (não estacionárias), ou fracionalmente integradas (podendo ser estacionárias ou não), o que destacou a importância dos testes de memória longa para uma análise mais apurada das séries.

Foi então realizado o teste de Ljung e Box modificado para verificação da ausência de autocorrelação nas observações componentes de cada série (de retornos ao quadrado e de retornos absolutos). Para todas as *commodities* analisadas, tanto para o caso dos retornos ao quadrado quanto para o caso dos retornos absolutos, os resultados implicaram na rejeição da hipótese nula de que não há autocorrelação, oferecendo

indícios de que os dados componentes de cada série apresentam alguma relação de dependência, de tal forma que não seria possível afirmar que estas séries seguem um processo aleatório com observações independentes.

Realizou-se também o teste BDS³⁷, para verificar a hipótese de que as observações eram independentes e identicamente distribuídas, o que foi rejeitado para todas as séries de retornos ao quadrado e de retornos absolutos, indicando também a existência de alguma relação de dependência entre as observações.

Um resumo dos testes realizados para as séries de retornos absolutos e retornos ao quadrado segue nos quadros abaixo.

Série: Retornos Absolutos							
Testes de Hipótese	H0	Café	Boi Gordo	Açúcar	Milho	Soja (Paraná)	Bezerro
Jarque Bera	Dados são normalmente distribuídos	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
Ljung-Box	Ausência de Autocorrelação	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
BDS	Dados são i.i.d.	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
Phillips-Perron	Existe raiz unitária	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0	REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0
KPSS Test (tendência = "c")	Dados são estacionários [I(0)] em torno de uma constante	REJEITA H0 A 1%	REJEITA H0 A 5%	REJEITA H0 A 5%	REJEITA H0 A 5%	REJEITA H0 A 1%	REJEITA H0 A 1%
KPSS Test (tendência = "ct")	Dados são estacionários [I(0)] em torno de uma tendência linear	REJEITA H0 A 5%	REJEITA H0 A 1%	REJEITA H0 A 1%	REJEITA H0 A 1%	REJEITA H0 A 1%	REJEITA H0 A 1%

Fonte: Resultados da pesquisa. Elaboração própria.

Tabela 2 – Resultados dos Testes de Hipótese para os Retornos Absolutos.

Série: Retornos ²							
Testes de Hipótese	H0	Café	Boi Gordo	Açúcar	Milho	Soja (Paraná)	Bezerro
Jarque Bera	Dados são normalmente distribuídos	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
Ljung-Box	Ausência de Autocorrelação	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
BDS	Dados são i.i.d.	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0	REJEITA H0
Phillips-Perron	Existe raiz unitária = não é estacionário	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0	REJEITA H0	NÃO REJEITA H0	REJEITA H0
KPSS Test (tendência = "c")	Dados são estacionários [I(0)] em torno de uma constante	REJEITA H0 A 1%	REJEITA H0 A 5%	NÃO REJEITA H0	REJEITA H0 A 5%	NÃO REJEITA H0	NÃO REJEITA H0
KPSS Test (tendência = "ct")	Dados são estacionários [I(0)] em torno de uma tendência linear	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0 A 1%	REJEITA H0 A 1%	REJEITA H0 A 1%	REJEITA H0 A 1%

Fonte: Resultados da pesquisa. Elaboração própria.

Tabela 3 – Resultados dos Testes de Hipótese para os Retornos ao quadrado.

5.2 A verificação da hipótese de séries temporais fracionalmente integradas

Conforme abordado no capítulo 4, a presença de certos elementos como a não-linearidade, a distribuição não-gaussiana dos retornos e principalmente a existência de ordem de integração fracionária são indicativos de que as séries em análise apresentam

³⁷ É importante ressaltar neste ponto, que sendo a série de retornos ao quadrado e de retornos absolutos transformações não lineares das séries de retornos das commodities analisadas, a análise por meio do teste BDS faz sentido se considerarmos estas séries como *proxies* da variância.

características de estruturas fractais. Observa-se assim, que os resultados dos testes tradicionais citados na seção anterior, oferecem indícios de que as séries de retornos, retornos absolutos e retornos ao quadrado das *commodities* aqui analisadas sejam séries temporais fracionalmente integradas. Neste sentido, a aplicação de testes específicos foi de suma importância para uma identificação mais adequada das mesmas.

Inicialmente foram realizados os testes de R/S e de GPH para verificar a presença de memória longa nas séries de log-retornos, cujos resultados são apresentados na tabela abaixo.

Série: Log-Retornos				
Teste de Hipótese	R/S (Hurst, 1951)	R/S (Lo, 1991)	GPH	ARFIMA (0,d,0)
Café	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0
Boi Gordo	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0
Açúcar	REJEITA H0 A 1%	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0
Milho	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0
Soja (Paraná)	NÃO REJEITA H0	NÃO REJEITA H0	NÃO REJEITA H0	REJEITA H0
Bezerro	NÃO REJEITA H0	REJEITA H0 A 5%	REJEITA H0 A 5%	NÃO REJEITA H0

Fonte: Resultados da pesquisa. Elaboração própria.

Tabela 4 – Resultados dos Testes de Hipótese Fractais para os Log-Retornos.

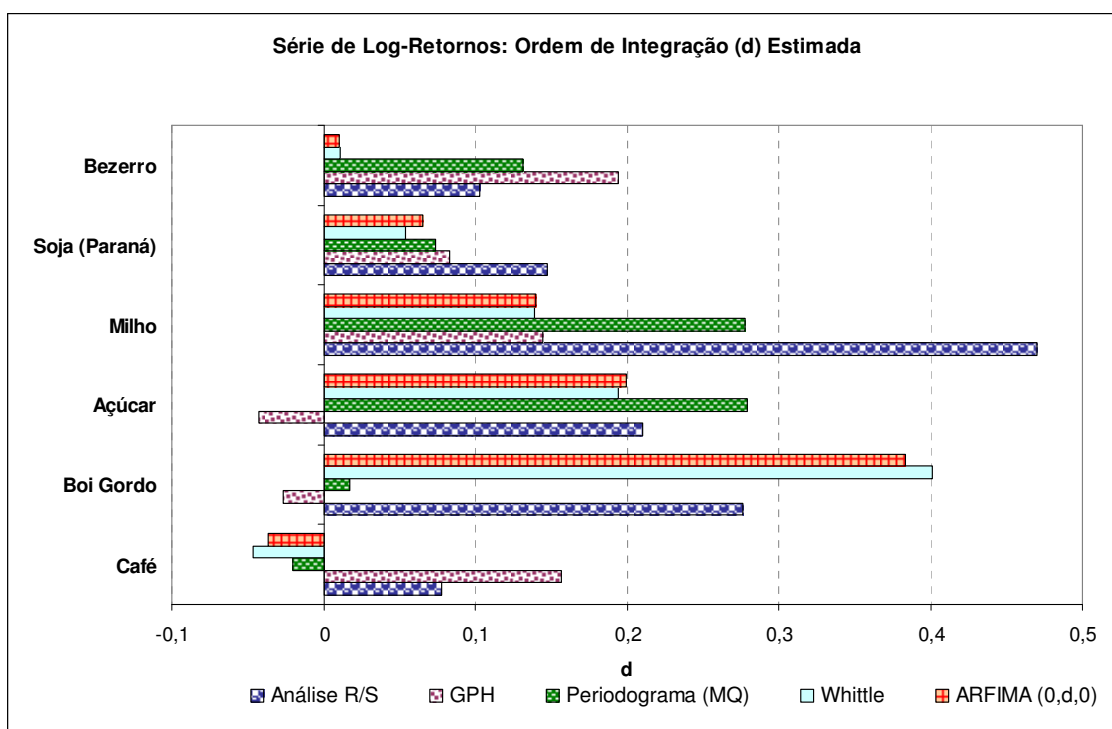
Sob a hipótese nula da ausência de memória longa, os testes de R/S na versão de Hurst (com bandwidth igual a zero) rejeitaram H_0 ao nível de 1% somente para a série do açúcar, enquanto que na versão de Lo (com bandwidth igual a 10), a hipótese nula foi rejeitada ao nível de 5% para a série do bezerro, tal que poderia ser possível que os log-retornos destas séries apresentassem memória longa. Todavia, esta discrepância no resultado dos testes conforme a versão adotada poderia indicar a presença de heterogeneidade estatística nas séries de açúcar e do bezerro. De maneira geral, porém os testes de R/S em ambas as versões, para o café, boi gordo, milho e soja não foram capazes de rejeitar a hipótese nula da ausência de memória longa. Isto pode se dever ao fato de que o teste apresenta baixa potência ou de que da estatística de teste se concentrou próximo dos limites do intervalo de confiança.

Sendo assim, procedeu-se o teste GPH sobre os log-retornos, sob a hipótese nula de que a ordem de integração d dos processos é igual a zero, ou seja, não há memória longa. Com exceção da série de log-retornos do bezerro, não foi possível rejeitar H_0 para as demais *commodities*, tal que para estas não houve indícios da existência de memória longa. No entanto, a estimação da ordem de integração destas séries poderia indicar o comportamento do componente da média destes processos de tal forma a oferecer melhores subsídios para a determinação de um modelo representativo mais adequado.

Foram então estimadas as ordens de integração segundo os diferentes métodos abordados na seção 4.2, cujos resultados seguem nos Anexos.

Segundo os resultados dos testes de R/S e GPH, era de se esperar que a ordem de integração estimada referente às séries em que a hipótese nula da ausência de memória longa não foi rejeitada se mostrasse fora ou na fronteira do intervalo em que $0 < d < 0,5$ (onde fica caracterizada a presença de memória longa e cujo resultado esperado do teste seria de rejeição da H_0).

Conforme o gráfico abaixo e os resultados da tabela que consta no item 6.2 dos Anexos é possível observar que os diferentes métodos indicaram que o parâmetro d estimado se localiza para a maioria das séries, dentro do intervalo característico de séries com memória longa, no entanto, mais próximos ao limite inferior deste intervalo. Para as séries do açúcar e do boi gordo, o método GPH indicou valores no intervalo característico de séries estacionárias com memória curta, assim como o método de Whittle e o método do periodograma para o caso dos log-retornos do café.



Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 4 – Ordem de Integração Estimada para os Log-Retornos.

Diante disto, nota-se que as divergências encontradas nos testes de raiz unitária de Phillips e Perron e o KPSS podem ser devidas à existência de uma ordem de integração fracionária para as séries analisadas neste estudo.

Sendo assim, em especial para o caso da série dos log-retornos do bezerro, cujos resultados do teste indicaram a existência de memória longa, seria interessante considerar uma modelagem baseada em modelos fracionalmente integrados com os *ARFIMA* para descrição do comportamento da média.

Foram também avaliadas as séries de retornos absolutos e de retornos ao quadrado como *proxies* para a volatilidade destas *commodities*.

Para os retornos absolutos, os resultados dos testes R/S e GPH indicaram a rejeição da hipótese nula de ausência de memória longa para todas as séries, conforme resultados da tabela abaixo.

Série: Retornos Absolutos				
Teste de Hipótese	R/S (Hurst, 1951)	R/S (Lo, 1991)	GPH	ARFIMA (0,d,0)
Café	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 5%	REJEITA HO
Boi Gordo	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO
Açúcar	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	NÃO REJEITA HO
Milho	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO
Soja (Paraná)	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO
Bezerro	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO

Fonte: Resultados da pesquisa. Elaboração própria.

Tabela 5 – Resultados dos Testes de Hipótese Fractais para os Retornos Absolutos.

Já a análise da série de retornos ao quadrado, os resultados dos testes R/S e GPH indicaram a rejeição da hipótese nula de ausência de memória longa para quase todas as séries, com exceção da série de boi gordo, conforme resultados da tabela abaixo.

Série: Retornos²				
Teste de Hipótese	R/S (Hurst, 1951)	R/S (Lo, 1991)	GPH	ARFIMA (0,d,0)
Café	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO
Boi Gordo	REJEITA HO A 1%	NÃO REJEITA HO	NÃO REJEITA HO	REJEITA HO
Açúcar	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	NÃO REJEITA HO
Milho	REJEITA HO A 1%	REJEITA HO A 5%	REJEITA HO A 1%	REJEITA HO
Soja (Paraná)	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO
Bezerro	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO A 1%	REJEITA HO

Fonte: Resultados da pesquisa. Elaboração própria.

Tabela 6 – Resultados dos Testes de Hipótese Fractais para os Retornos ao quadrado.

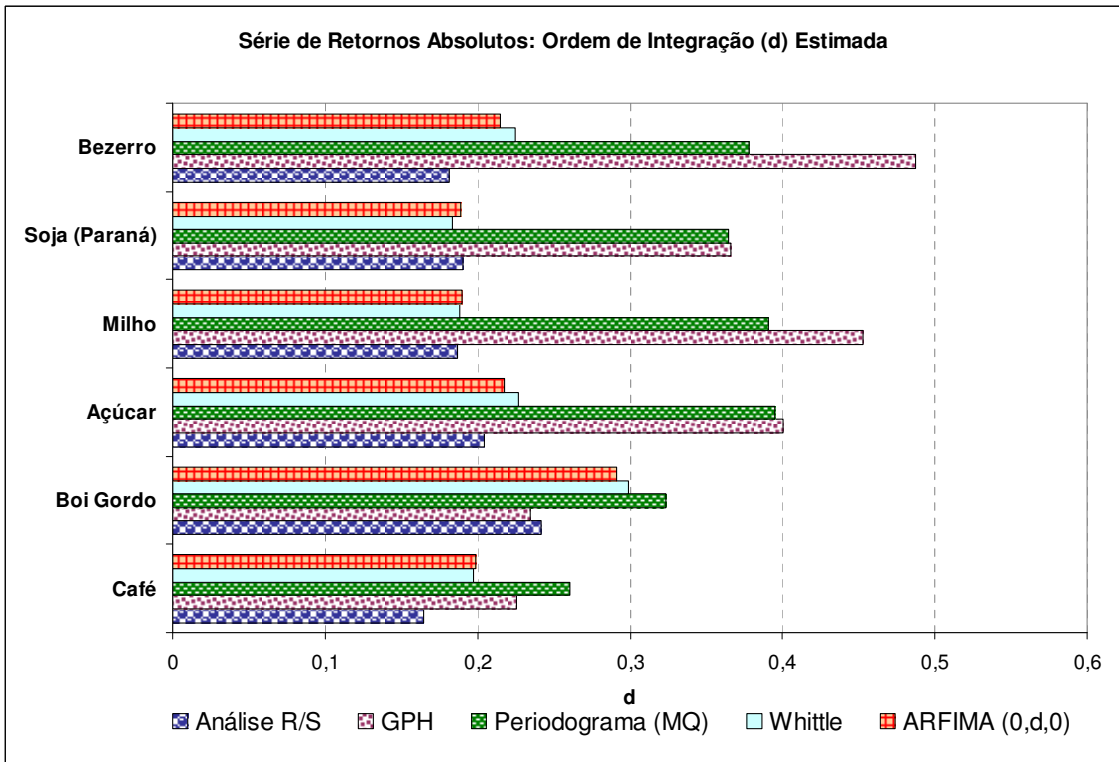
A estimação dos parâmetros de memória longa para os retornos absolutos e os retornos ao quadrado indicou uma concentração maior das estimativas entre 0 e 0,2, como pode ser analisado nos gráficos abaixo e nas tabelas em anexo.

Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 4 – Ordem de Integração Estimada para os Log-Retornos.

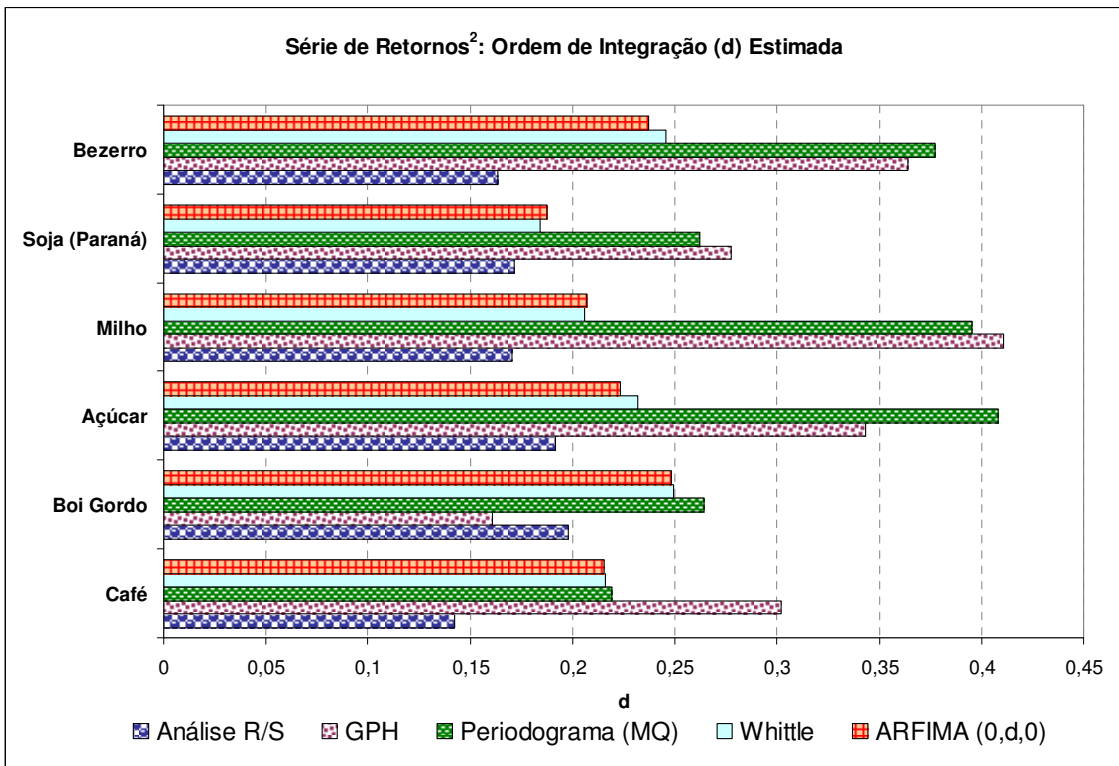
Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 4 – Ordem de Integração Estimada para os Log-Retornos.



Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 5 – Ordem de Integração Estimada para os Retornos Absolutos.



Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6 – Ordem de Integração Estimada para os Retornos ao quadrado.

5.3 Estimação dos modelos

Diante das evidências obtidas, buscou-se estimar, como um elemento adicional às análises que se pretendia este estudo, os modelos para a parte do *drift* dos processos das commodities que fossem mais adequados às características sinalizadas pelos testes. Para tanto, foi utilizada a função FARIMA do software S+Finmetrics para ajustar um modelo mais geral do tipo $ARFIMA(p,d,q)$, além dos modelos $ARIMA(p,d,q)$, $ARMA(p,q)$ e $AR(p)$ sobre os log-retornos das diferentes *commodities* e a identificação do modelo mais adequado se deu em função da minimização do critério BIC³⁸.

Estes procedimentos resultaram nos seguintes ajustes para modelagem do componente de *drift* para os processos das respectivas commodities:

Log-Retornos: coponente de <i>drift</i>	
Teste de Hipótese	Modelo Estimado
Café	AR (2)
Boi Gordo	AR (4)
Açúcar	ARMA (1,1)
Milho	AR (3)
Soja (Paraná)	AR (2)
Bezerro	AR (2)

Fonte: Resultados do Modelo. Elaboração Própria.

Tabela 7 – Resultados da estimação do componente de *drift* dos processos.

A partir destes resultados, foram avaliados então diferentes modelos visando o ajuste do componente da variância. Considerou-se que esta poderia ser modelada a partir de distintas formas funcionais da família GARCH, tais como o EGARCH, FIGARCH, FIEGARCH.

Procedeu-se então a modelagem conjunta do *drift* com os modelos supracitados para as seis *commodities* deste estudo³⁹ e segundo a minimização do critério BIC, foi escolhido o modelo de combinação mais ajustada, que pode ser visualizada na tabela abaixo e também nos gráficos QxQ que constam nos Anexos.

³⁸ BIC trata-se de um critério de informação bayesiano que auxilia na identificação de um modelo mais adequado, favorecendo a parcimônia. Neste estudo escolheu-se portanto, o modelo que minimizasse o BIC.

³⁹ Para tanto, foram consideradas as funções “garch” e “fgarch” no S+Finmetrics.

Ajuste conjunto do componente de <i>drift</i> e da variância		
Teste de Hipótese	Componente de <i>drift</i>	Componente da variância
Café	AR (2)	FIEGARCH
Boi Gordo	AR (4)	FIEGARCH
Açúcar	ARMA (1,1)	GARCH
Milho	AR (3)	GARCH
Soja (Paraná)	AR (2)	EGARCH
Bezerro	AR (2)	FIEGARCH

Fonte: Resultados do Modelo. Elaboração Própria.

Tabela 8 – Resultados da estimação conjunta do *drift* e da variância dos processos.

Para as *commodities* açúcar e milho, o melhor ajuste foi observado diante de modelos GARCH, para a soja, o modelo EGARCH (que admite resposta assimétrica diante de retornos positivos e negativos) foi o mais adequado. Já para as *commodities* café, boi gordo e bezerro, o melhor ajuste foi observado diante de modelos FIEGARCH, que admitem resposta assimétrica diante de retornos positivos e negativos e incluem a característica da fractalidade das séries.

Isto pode estar relacionado ao maior volume de negócios nos mercados de café, boi gordo e bezerro, assim como ao maior desenvolvimento dos mercados futuros destas *commodities* no Brasil se comparados às demais aqui analisadas, o que pode fazer com que a resposta dos agentes em relação aos retornos positivos e negativos seja mais acentuada.

Ressalta-se, todavia, que as funções padronizadas de ajuste às distribuições fracionalmente integradas presentes no S+Finmetrics não disponibilizam a opção de alteração da distribuição dos resíduos do componente da variância. Sendo assim, não foi considerado na modelagem o ajuste à distribuição t-Student, tal que o componente de caudas mais pesadas observado nos testes não foi avaliado. Da mesma forma, as funções padronizadas disponíveis no S+Finmetrics não oferecem a opção de ajuste conjunto do componente de *drift* fracionário com o componente da variância, de modo que este elemento também não foi avaliado. Isto seria interessante em especial para o caso do bezerro, cuja hipótese nula da ausência de memória longa nos testes R/S (versão de Lo, 1991) e GPH foi rejeitada e o ajuste individual do componente de *drift* a um modelo autorregressivo fracionalmente integrado (um $AR(2)$ fracionário) indicou um critério BIC superior ao ajuste de um modelo $AR(2)$ não fracionário.

Seria interessante, neste sentido, o desenvolvimento de um algoritmo próprio que fosse capaz de admitir estes elementos relevantes para a análise, tal que uma nova

estimação e análise do ajustamento destes modelos pudessem ser realizadas considerando estes aspectos.

5.4 Observações finais e sugestões para estudos futuros

Os resultados deste estudo indicaram que as séries representativas do mercado à vista brasileiro de café, boi gordo, açúcar, milho, soja e bezerro apresentam características de estruturas fractais, em especial em termos da variância destes processos, e que o uso de técnicas apropriadas para o tratamento de séries temporais fracionárias pode ser relevante para o melhor entendimento do seu comportamento.

Entretanto, a maior parte dos estudos disponíveis ainda se concentra na análise da hipótese de que estes mercados são regidos por processo do tipo passeio aleatório, de tal forma que as diversas análises que são realizadas sobre estes mercados estão fundamentadas na admissão desta premissa.

Todavia, a detecção de características fractais nestes mercados levanta a questão da possibilidade de que as outras técnicas de estudo, como as análises de co-integração entre os mercados à vista e futuros e do *hedge*, amplamente difundidas no caso dos mercados de *commodities* agrícolas, possam conduzir o pesquisador a obtenção de resultados viesados diante de um tratamento inadequado do elemento fractal presente nestas séries.

Destaca-se, em especial, o caso dos mercados de bezerro e boi gordo, uma vez que estas *commodities* fazem parte da mesma cadeia produtiva e houve a detecção de fractalidade na série de log-retornos do bezerro. Seria interessante, neste sentido, uma análise mais aprofundada destes mercados visando identificar as razões desta característica e os impactos que isto pode gerar em termos da análise e modelagem dos mesmos. Além disto, recomenda-se também a realização de estudos mais aprofundados sobre o impacto do tratamento de séries fracionalmente integradas por técnicas tradicionais de cointegração (que não a cointegração fracionária).

Ressalta-se, ainda que a detecção de fracionalidade em todas as séries *proxies* da volatilidade das *commodities* aqui analisadas sinaliza a necessidade de estudos mais aprofundados sobre esta característica, de modo a identificar os impactos que a não admissão da fractalidade na volatilidade destas *commodities* pode gerar em sua modelagem.

Outra questão a ser mais bem analisada seria a avaliação destas séries descontando-se as influências da inflação. É possível que ao se proceder desta forma, as características de memória longa também nos log-retornos possam ser observadas, na medida em que refletiriam a estrutura existente nos mercados físicos destas *commodities*. Sendo assim, este procedimento permanece também como sugestão para estudos posteriores.

Além destas questões mais técnicas, o exercício da pesquisa para a concretização deste estudo revelou que não há no mercado brasileiro um referencial teórico extenso para o estudo das características de séries temporais fracionalmente integradas, seja em termos de artigos que avaliem se estas estruturas são encontradas em outras séries temporais ou mesmo em termos de livros de referência. Neste sentido, a investigação quanto a existência de séries temporais fracionalmente integradas no mercado brasileiro pode representar ainda um campo fértil de pesquisa e publicações já que aparentemente ainda pouco explorado.

Ademais, foi observado que parte significativa dos softwares estatísticos disponíveis para análise econométricas ainda não dispõe de ferramental extenso que permita análises mais abrangentes das características particulares de séries fracionalmente integradas, tal como a questão da distribuição dos resíduos abordada ao final da seção 5.3. Neste sentido, como sinalizado na seção anterior, seria interessante, o desenvolvimento de algoritmos próprio que permitissem análises mais variadas destas séries.

Desta forma, os avanços nas pesquisas quanto a características fractais de séries temporais fracionalmente integradas é de suma importância senão para testar a adequação da aplicação dos modelos tradicionais diante da presença de séries temporais fracionalmente integradas, ao menos para identificar a presença destas estruturas em diferentes mercados brasileiros de modo a ter uma compreensão mais aprofundada do seu funcionamento.

Ressalta-se, por fim, a relevância deste estudo no sentido de colaborar para o avanço da teoria econômica e dos avanços das pesquisas no âmbito da Economia não como um sistema em equilíbrio simples, mas marcado por inúmeras relações complexas entre seus agentes e que se traduzem em dinâmicas sistêmicas não-lineares. Sendo assim, este estudo soma-se aos esforços de outros autores para a renovação da pesquisa como ciência.

6 Referências Bibliográficas

- ANDERSON P., K. J. e K. ARROW, D. PINES, Eds. *The Economy as an Evolving Complex System*, Reading, Mass.: Addison-Wesley, 1988. 317 p.
- ARTHUR, W. B., S. N. DURLAUF, D. A. LANE, Eds., *The Economy as an Evolving Complex System II*, Reading, Mass.: Addison-Wesley , 1997. 583 p.
- ARTHUR, W.B., Complexity and the Economy, *Science*, v. 284, p.107-109, 1999.
- BAK, P., C. Tang, and K. Wiesenfeld, Self-Organized Criticality, *Physical Review A*, v.38, p.364, 1988.
- BACHELIER, L. *Théorie de La Speculation*, 1900. Tese de doutorado traduzida e reimpressa em Cootner, *op. cit.*
- BAILLIE, R. T. Long memory processes and fractional integration in econometrics. *Journal of Econometric.*, v. 73, p .5-59, 1996.
- BAILLIE, R. T. BOLLERSLEV, T., e MIKKENLESEN, H. O. Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometric.* v. 74, p. 3-30, 1996.
- BARKOULAS, J., LABYS, W. C. e ONOCHIE, J. Fractional dynamics in international commodity prices. *Journal of Futures Markets.* v. 17, p. 161–189, 1997.
- BARKOULAS, J. T., LABYS, W. C. e ONOCHIE, J. I. Long Memory In Futures Prices. *Financial Review*, v. 34, p. 91–100, 1999.
- BEINHOCKER, E. D. *Origin of Wealth: Evolution, Complexity and the radical remaking of Economics*. Harvard Business Press, 2006. 527 p.
- BERA, A. K. e JARQUE, C. M., Efficient tests for normality, homoscedasticity and serial dependence of regression residuals: Monte Carlo Evidence. *Economics Letters*, v. 7, p. 313-318, 1981.
- BERAN, J. (1995) *Statistics for Long Memory Processes*. New York: Chapman e Hall. 1994, 315 p.
- BERTALANFFY, L. V. *General System Theory: Foundations, Development, Applications*. New York: George Braziller, 1969. 296 p.
- BITENCOURT, W. A. *Ensaio empírico sobre a eficiência do mercado futuro de café*. Dissertação (Mestrado em Administração) – Departamento de Administração e Economia, Universidade Federal de Lavras, Lavras, 2007. Disponível em: < http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=180787>. Acesso em: 01.nov.2012.

- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*. v. 31, p. 301-327, 1986.
- BOLLERSLEV, T., e MIKKENLESEN, H.O. Modeling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics*. v. 73, p. 151-184, 1996.
- BOUCHAUD, J.-P. Economics needs a scientific Revolution. *Nature*, v. 455, p. 1181, 2008.
- BOX, G. E. P.; PIERCE D. A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, v. 65, n. 332, p.1509–1526, 1970.
- BRESSAN, A. e LEITE, C. A. M., Eficiência do mercado futuro de café no Brasil. *Reúna*, v.6, n.1, p. 11-32, 2001.
- BROOKS, C. *Introductory Econometrics for Finance*. 1. ed. Cambridge University Press, 2002. 728 p.
- BROCK, W. A., DECHERT, W. D., SCHEINKMAN, J. A., A test for Independence based on the correlation dimension. *Econometric Reviews*, 15, 197-235, 1987.
- BUENO, R. L. S., *Econometria de séries temporais*. 2. ed. São Paulo: Cengage, 2011. v. 1. 341 p.
- CAMPBELL, J. Y., LO, A. W. e MACKINLAY, A. C. *The Econometrics of Financial Markets*. Princeton University Press. 1997. 611 p.
- CEPEA – Centro de Estudos Avançados em Economia Aplicada. Disponível em: <<http://www.cepea.esalq.usp.br>>.
- CHARLES, A. e DARNÉ, O. Variance-Ratio Tests of Random Walk: An Overview. *Journal of Economic Surveys*, v. 23, n.3, p. 503-527, 2009.
- CONT, R. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, v. 1, p. 223-236, 2001.
- COOTNER, P. H. *The Random Character of Stock Market Prices*. Cambridge, Mass.: The M.I.T. Press., 1964, 552p.
- COWLES, 3rd, A. Can stock market forecasters forecast? *Econometrica*, v. 1, n.3, 309-324, 1933.
- CRUZ JÚNIOR, J. C.; SILVEIRA, R. L. F. Análise de eficiência, co-integração e exogeneidade nos mercados futuros de café na BM&F, NYBoT e LIFFE. In: CONGRESSO BRASILEIRO DE ECONOMIA E SOCIOLOGIA RURAL, XLV., 2007, Rio Branco. *Anais...* Londrina: Sociedade Brasileira de Economia, administração e Sociologia Rural, 2007. p. 21.

DICKEY, D. A. e FULLER, W. A. Distribution of the estimator for autoregressive time series with a unit root. *Journal of the American Statistical Association*, v.74, p. 427-431, 1979.

DICKEY, D. A. e FULLER, W. A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, v. 49, p. 1052-1072, 1981.

DIMSON, E. e MUSSAVIAN, M. The Efficient Markets Hypothesis. In: DAHIYA, Shri Bhagwan *The current state of business disciplines*, Spellbound Publications, v.3, p. 959-970, 2000.

EINSTEIN, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, *Annalen der Physik* v. 332, n. 8, p. 549-560, 1905.

EITELMAN, P. S. e VITANZA, J. T. A Non-Random Walk Revisited: Short- and Long-Term Memory in Asset Prices. FRB International Finance Discussions Paper. N° 956.

ENGLE, R. F. Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation. *Econometrica*, v. 50, p. 987-1008, 1982.

FAMA, E. F. Efficient Capital Markets: A review of Theory and Empirical Work, *Journal of Finance*, v. 25, n. 2, p. 383-417, 1970.

FAVA, V. L. e ALVES, D. C. O. Longa persistência nas taxas de inflação, *Revista de Econometria*, v. 18, n. 2, p. 245-264, 1998.

GEWEKE, J. e PORTER-HUDAK, S. The estimation and publication of long memory time series models. *Journal of Time Series Analysis*, v. 4, p. 221-238, 1983.

GLERIA, I.; MATSUSHITA R.; Da SILVA, S. Sistemas complexos, criticalidade e leis de potência. *Revista Brasileira do Ensino de Física*, v.26, n.2, 2004.

GRANGER, C. M. G. e JOYEUX, R., An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, v.1, p.15-29, 1980.

GREENE, M. T. e FIELITZ, B. D., Long Term dependence in common stock returns, *Journal of Financial Economics*, v.5, p.339-380,1977.

GUJARATI, D. *Econometria Básica*. São Paulo:Makron Books, 2004. 812 p.

HOSKING, J. R. M. Fractional Differencing, *Biometrika*, v. 68, p. 165-176, 1981.

JIN, H. e FRECHETTE, D.L. Fractal Geometry in Agricultural Cash Price Dynamics. *Annual Meeting of the American Agricultural Economics Association*, Long Beach, 2002.

JOHNSON, N. F. *Simply Complexity: A Clear Guide to Complexity Theory*. Oneworld Publications, 2009. 236p.

- KENDALL, M. G. The analysis of economic time-series – Part I: Prices, *Journal of the Royal Statistical Society Series A (General)* v.116, n. 1, p. 11-25, 1953.
- KWIATKOWSKI, D., PHILLIPS, P. C. B. e SCHMIDT, P. e SHIN, Y., Testing the null Hypothesis of Stationary Against the Alternative of a Unit Root, *Journal of Econometrics*, v. 54, p. 159-178, 1992.
- LARSON, A. B. Measurement of a random process in future prices, *Food Research Institute Studies* v.1, n. 3, p. 313-24, 1960.
- LEROY, S. F. Efficient Capital Markets: Comment, *The Journal of Finance*, v. XXXI, n.1 , p. 139-141, 1976.
- LEROY, S. F. Efficient Capital Markets and Martingales, *Journal of Economic Literature*, v. XXVII, p. 1583-1621, 1989.
- LJUNG, G. e Box, G. E. P. On a measure of lack off fit in time series models. *Biometrika*, v. 65, p. 297-303, 1978.
- LO, A. W. Long term memory in stochastic market prices. *Econometrica*, v. 59, p. 1279-1313, 1991.
- LO, A. W. e MACKINLAY, A. C. *A Non-Random Walk Down Wall Street*. Princeton University Press. 1999. 448 p.
- LORENZ, E. D., Predictability: A problem partly solved, *Proceeding of the ECMWF Seminar on Predicatability (4-9 September 1995, READING, UK)*, ECMWF, p. 1-18, 1996.
- LAUX, P. R. e AGUIRRE, A. J. Estimativa da dimensão fractal de nuvens em imagem de satélites meteorológicos. In: X SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, X, 2001, Foz do Iguaçu. *Anais...* São José dos Campos: INPE, 2001, p. 811-818.
- MANDELBROT, B; VAN NESS, J.W., Fractional Brownian motions, fractional noises and applications, *SIAM Review*, v. 10, n. 4, p. 422–437, 1968.
- MANDELBROT, B. e HUDSON, R. *The (mis)behaviour of markets: a fractal view of risk, ruin, and reward*. New York: Basic Books, 2004. 328 p.
- MARKOWITZ, H. Portfolio Selection. *The Journal of Finance*. v. 7, n. 1, p. 77– 91, 1952.
- MARQUES, G. O. L. C. Equilíbrio de Longo Prazo entre o consumo das famílias e renda disponível: aspectos da co-integração I(1)-I(0) e fracionária. *Revista de Economia e Administração*, v.8, n.4, p. 371-390, out./dez. 2009.
- MATIAS, Márcia Athayde. Análise do comportamento de preços da Commodity cobre: uma abordagem sob a ótica da teoria dos fractais. 2006. Dissertação (Mestrado em Ciências Contábeis) - Universidade Federal do Rio Grande do Norte, Natal, 2006.

MATIAS, M. A., SILVA, C. A. T. e VIEIRA, L. Fractalidade em séries temporais: uma aplicação para uma série histórica de preços de minério de cobre, 2007. Disponível em: < <http://www.congressosp.fipecafi.org/artigos72007/279.pdf> >. Acesso: 05/12/2012.

MILIONIS, A. E. An alternative definition of market efficiency and some comments on its empirical testing. Bank Of Greece Economic Research Department – Special Studies Division, 2006.

MORETTIN, P. A. *Econometria Financeira: um curso em séries temporais financeiras*. 2ª Ed. São Paulo: Bluchner, 2011. 383 p.

MUÑOZ, P.C, da Silva, MIRANDA, M.F., CLARO, J.G.V. F. e DINIZ, R. G., Physica A Geometria fractal: propriedades e características de fractais ideais. *Revista Brasileira de Ensino de Física*, v. 30, n. 2, 2304, 2008.

NELSON, D. B. Conditional heteroskedasticity in asset returns. *Econometrica*, v. 59, p. 347-370, 1991.

OSBORNE, M. F. M. Brownian motion in the stock market, *Operations Research*, v. 7, n. 2, p. 145-73, 1959.

PEARSON, K., The problem of the random walk, *Nature*, v.72, p.294-294, 1905.

PETERS, E. E. *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*. John Wiley & Sons, Inc, 1994. 336 p.

PETERS, E. E. *Chaos and Order in the Capital Markets*. 2ª ed. New view of Cycles, Prices, and Market Volatility, 1996. 336 p.

PETERS, E. E. Fractal Structure in the Capital Markets, *Financial Analysts Journal* July/August, v. 45, n.4, p. 32-37, 1989.

PARK, C.-H. e IRWIN, S. H. *The profitability of Technical Analysis: a review*. Urbana: University of Illinois at Urbana-Champaign, Department of Agricultural and Consumer Economics, 2004. 102 p. (AgMAS Project Research reports, n.04)

PRIESTLEY, M. B. *Spectral Analysis and Time Series: Multivariate series, prediction and control*, Nova Iorque: Academic Press, 1981, 890 p.

ROBERTS, H. Stock-market patterns and financial analysis: methodological suggestions, *Journal of Finance* 14, 1-10, 1959.

RODRIGUES, M. A., MARTINES FILHO, J. G. Eficiência nos mercados futuros agropecuários brasileiros. 2ª Conferência em Gestão de Risco e Comercialização de *Commodities*. Instituto Educacional BM&FBOVESPA.

SALINAS, S. R. A.. Einstein e a teoria do movimento browniano. *Revista Brasileira de Ensino de Física*, v. 27, n. 2, p. 263 - 269, 2005.

SAMUELSON, P. A. Proof that properly anticipated prices fluctuate randomly, *Industrial Management Review* 6:2, 41-49, 1965.

SHREVE, S. E., *Stochastic Calculus for Finance*. Nova Iorque: Springer-Verlag, v.1, 2004. 187 p.

SILVA, J.M. e LIMA, J. A. S. Quatro abordagens para o movimento browniano. *Revista Brasileira de Ensino de Física*, v. 29, n. 1, p. 25-35, 2007.

TARGINO, R. S. Teoremas de Não-Arbitragem em Mercados Regidos pelo Movimento Browniano Fracionário. (Mestrado Estatística, Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

ZIVOT, E. e WANG, J., *Modelling Financial Time Series With SPLUS*. First Edition. Nova Iorque: Springer, 2005.

WIERSEMA, U. F. *Brownian Motion Calculus*, John Wiley & Sons, Inc., 2008. 330 p.

WORKING, H. New concepts concerning futures markets and prices. *The American Economic Review*, v. 52, n. 3, p. 431-459, 1962.

7 Anexos

7.1 Estatísticas Descritivas

Série: Log-Retornos						
Estatísticas Descritivas	Café	Boi Gordo	Açúcar	Milho	SojaPR	Bezerro
nº de observações	4186	3964	2507	2218	3963	3326
Mínimo	-0.1665	-0.0357180826	-0.0748648790	-0.0768540950	-0.0896546959	-0.0731922502
1º Qu.	-0.01127	-0.0015727933	-0.0066378639	-0.0065253274	-0.0064855078	-0.0061014726
Mediana	0.0001313	0.0002477134	0.0004054328	0.0000000000	0.0000000000	0.0005497263
3º Qu.	0.01148	0.0022812508	0.0073200071	0.0071578282	0.0075080798	0.0069292767
Máximo	0.2521	0.0345327002	0.0791857362	0.0940693072	0.1009866378	0.0953287933
Momentos Amostrais:						
Média	0.00001006	0.0003435475	0.0002873805	0.0002928289	0.0001435791	0.0002668077
Desvio Padrão	0.02301	0.004758	0.01287	0.0129	0,01377	0.01256
Skewness	0.07804	-0.01134	-0.1908	0.1878	-0,2092	-0.1003
Kurtosis	11.12	9.086	6.797	7.921	7,294	7.034

Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6.1.a – Estatísticas descritivas e momentos das séries de log-retornos.

Série: Retornos Absolutos						
Estatísticas Descritivas	Café	Boi Gordo	Açúcar	Milho	SojaPR	Bezerro
nº de observações	4186	3964	2507	2218	3963	3326
Mínimo	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
1º Qu.	0.005151419	0.0007428731	0.003181033	0.003030766	0.003201284	0.002999543
Mediana	0.011388500	0.0019638656	0.006936444	0.006804173	0.007009702	0.006480823
3º Qu.	0.022106723	0.0043440920	0.012544903	0.012554941	0.013111063	0.011881788
Máximo	0.252059161	0.0357180826	0.079185736	0.094069307	0.100986638	0.095328793
Momentos Amostrais:						
Média	0.016113085	0.0031534474	0.009315927	0.009221547	0.009775795	0.008942923
Desvio Padrão	0.01642	0.003579	0.008881	0.00903	0.009695	0.008824
Skewness	3.108	2.588	2.345	2.619	2.404	2.349
Kurtosis	23.93	14.08	12.25	15.08	12.8	12.12

Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6.1.b – Estatísticas descritivas e momentos das séries de retornos absolutos.

Série: Retornos ²						
Estatísticas Descritivas	Café	Boi Gordo	Açúcar	Milho	SojaPR	Bezerro
nº de observações	4186	3964	2507	2218	3963	3326
Mínimo	0.0000000000	0.000000e+000	0.0000000000	0.000000e+000	0.0000000000	0.000000e+000
1º Qu.	0.00002653712	5,51861E-07	0.00001011897	9,19E+00	0.00001024822	8,997264
Mediana	0.00012969792	3,85677E-06	0.00004811426	4,63E+01	0.00004913592	42,001070
3º Qu.	0.00048870735	1,88711E-05	0.00015737507	1,58E+02	0.00017190021	1,41E+02
Máximo	0.06353382079	1,27578E-03	0.00627038081	8,85E+03	0.01019830100	9,09E+03
Momentos Amostrais:						
Média	0.00052929934	2,27508E-05	0.00016562387	1,67E+02	0.00018952905	1,58E+02
Desvio Padrão	0.001684	0.00006444	0.0003981	0.0004386	0.0004752	0.0003873
Skewness	17.69	9.254	7.542	8.895	8.392	8.457
Kurtosis	527.9	133.7	84.61	119	113.9	129.2

Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6.1.c – Estatísticas descritivas e momentos das séries de retornos ao quadrado.

7.2 Estimativas da ordem de integração

Série: Log-Retornos					
Método de Estimação	Análise R/S	GPH	Periodograma (MQ)	Whittle	ARFIMA (0,d,0)
Café	0,0778068	0,1565	-0,02032401	-0,04653242	-0,0364
Boi Gordo	0,2764278	-0,0267	0,01723713	0,4013715	0,3838
Açúcar	0,2103852	-0,0426	0,2796232	0,1945017	0,1997
Milho	0,4697858	0,1442	0,2777283	0,1390243	0,1399
Soja (Paraná)	0,1476535	0,0833	0,07419926	0,05439037	0,0652
Bezerro	0,1034167	0,1939	0,1316716	0,01095701	0,0103

Fonte: Resultados da pesquisa. Elaboração própria.

Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6.2.a – Estimativas da ordem de integração das séries de retornos ao quadrado.

Série: Retornos Absolutos					
Método de Estimação	Análise R/S	GPH	Periodograma (MQ)	Whittle	ARFIMA (0,d,0)
Café	0,1643518	0,2252	0,2606086	0,1972009	0,2026
Boi Gordo	0,2413918	0,2345	0,3231792	0,2991319	0,2915
Açúcar	0,2043252	0,4007	0,3952086	0,2266753	0,2179
Milho	0,1872492	0,4533	0,3908639	0,1884979	0,1898
Soja (Paraná)	0,1905760	0,3663	0,3649658	0,1832069	0,1888
Bezerro	0,1812055	0,4871	0,3777266	0,2249452	0,2146

Fonte: Resultados da pesquisa. Elaboração própria.

Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6.2.b – Estimativas da ordem de integração das séries de retornos absolutos.

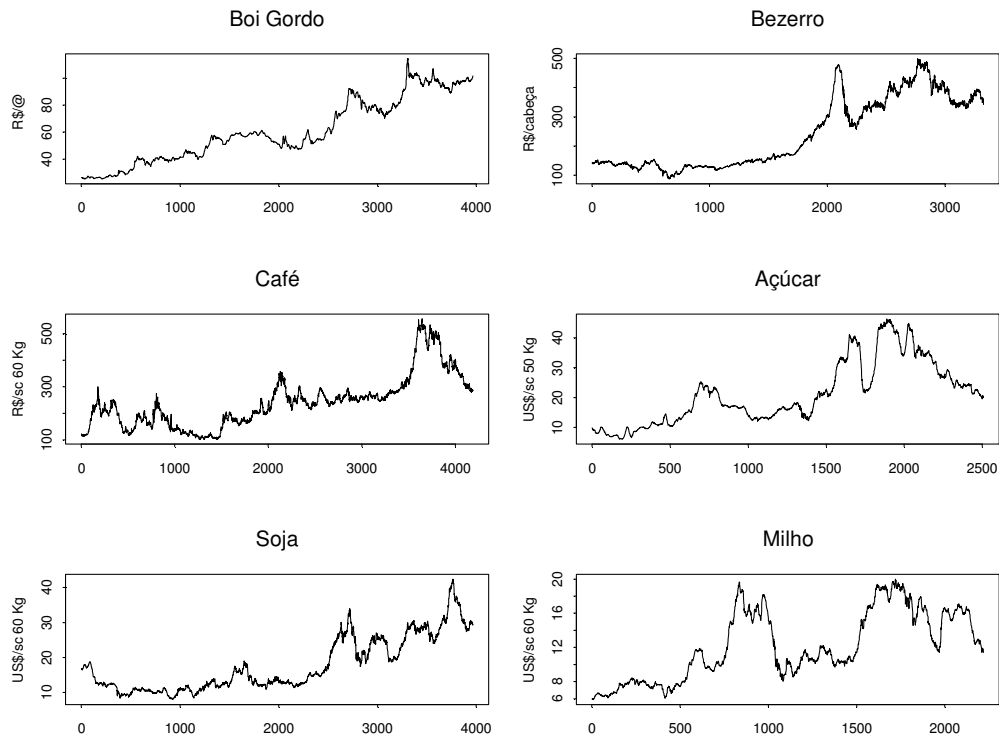
Série: Retornos ²					
Método de Estimação	Análise R/S	GPH	Periodograma (MQ)	Whittle	ARFIMA (0,d,0)
Café	0,1421595	0,3018	0,2193576	0,2159172	0,2183
Boi Gordo	0,1976743	0,1607	0,2644121	0,2495366	0,2481
Açúcar	0,1913236	0,3432	0,4081863	0,2317625	0,2233
Milho	0,1702437	0,4108	0,3955521	0,2059817	0,207
Soja (Paraná)	0,1712247	0,2775	0,2622902	0,1842896	0,1871
Bezerro	0,1633317	0,3639	0,3775007	0,2456745	0,2374

Fonte: Resultados da pesquisa. Elaboração própria.

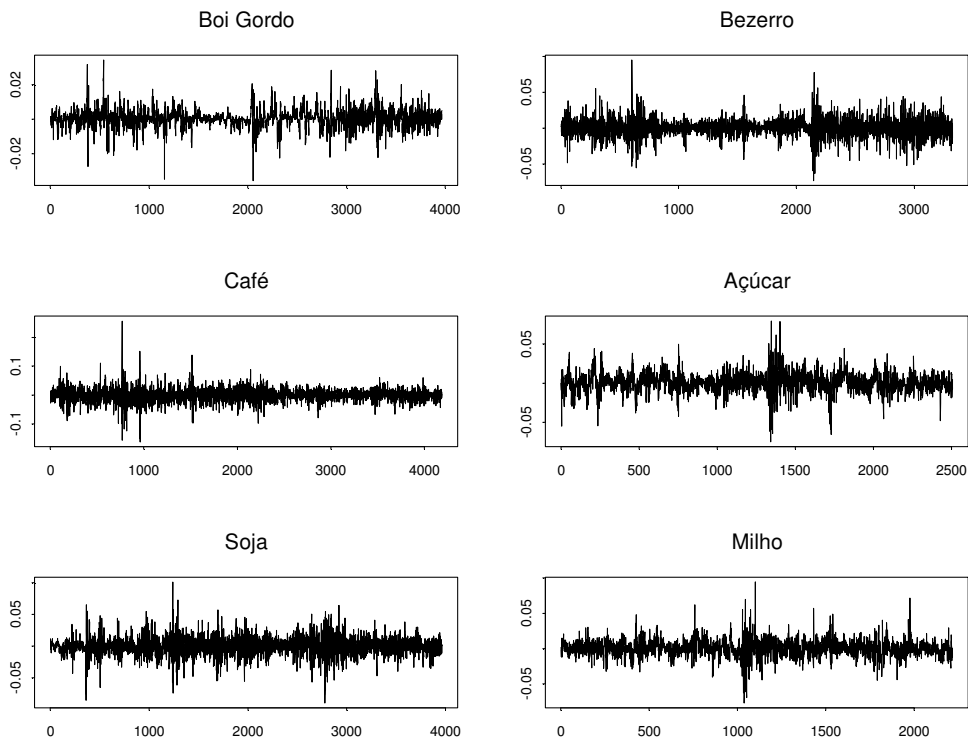
Fonte: Resultados da pesquisa. Elaboração própria.

Gráfico 6.2.c – Estimativas da ordem de integração das séries de retornos ao quadrado.

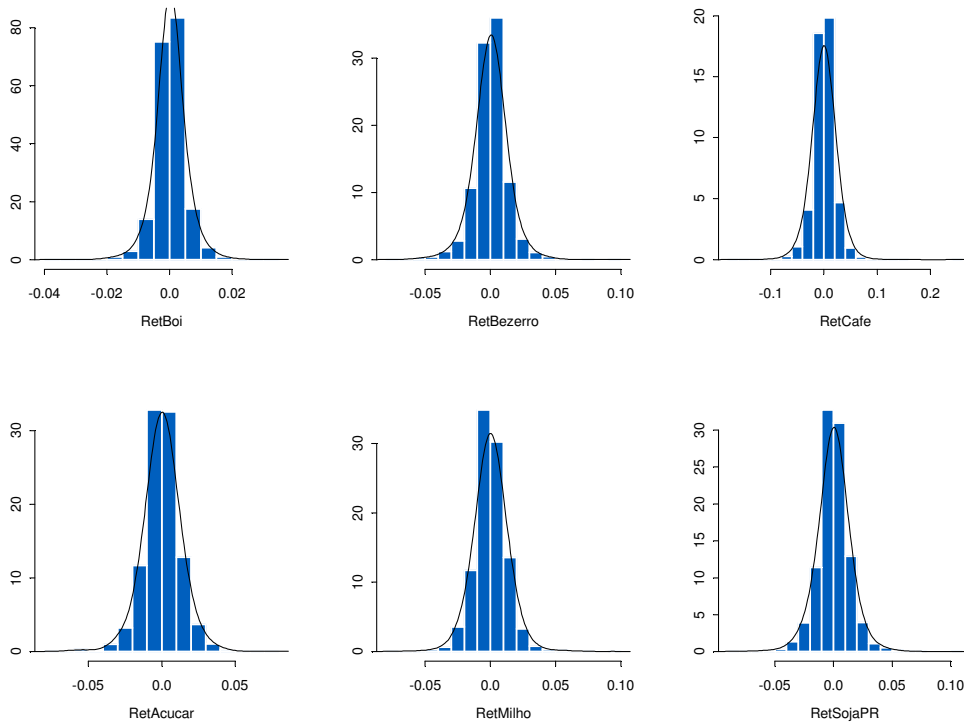
7.3 Preços (séries em nível)



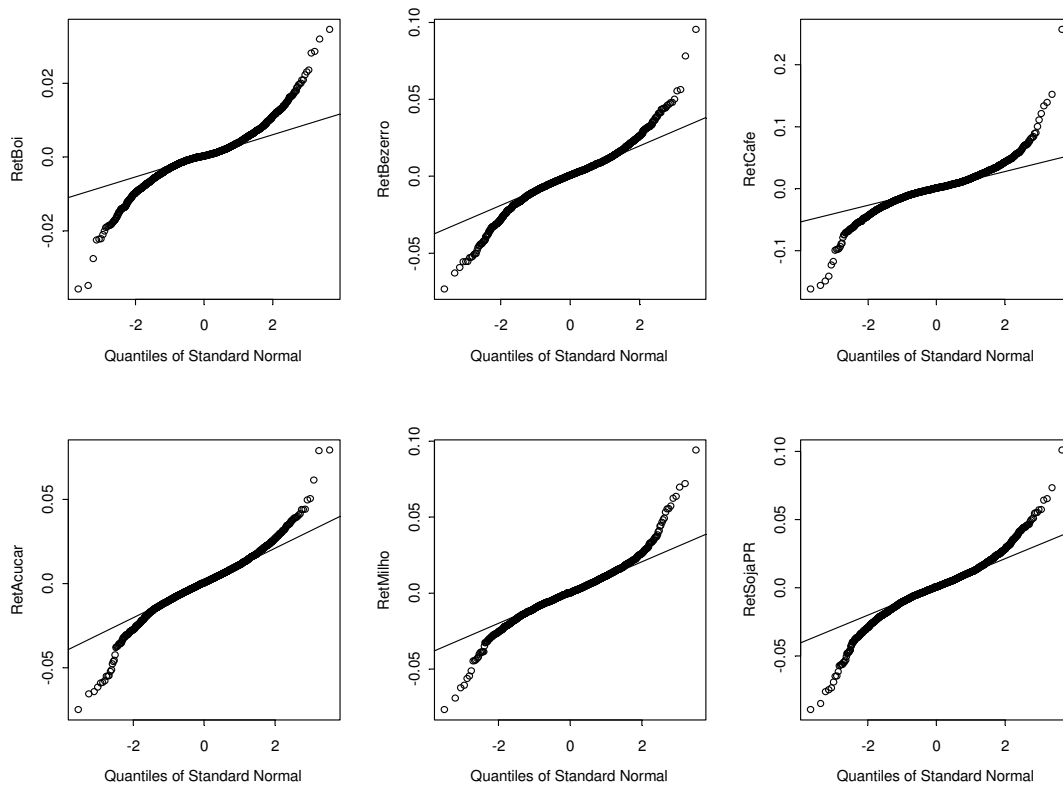
7.4 Log-Retornos (séries em primeira diferença)



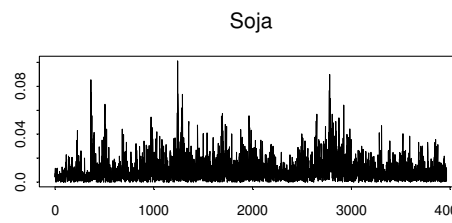
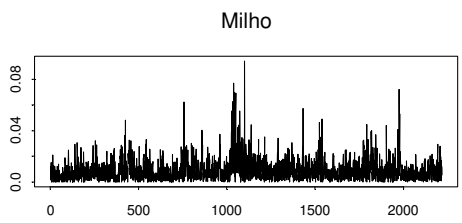
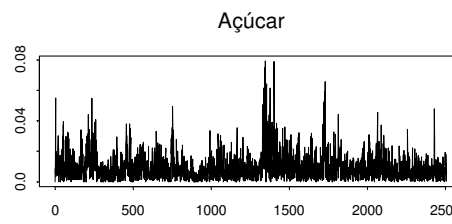
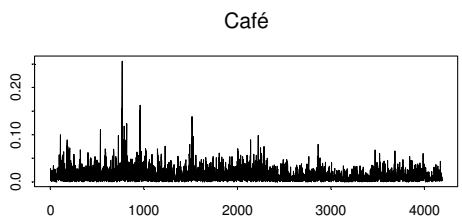
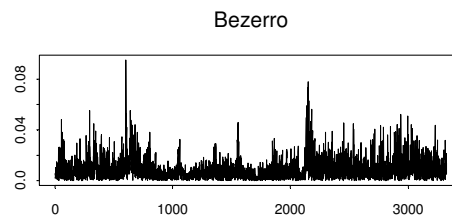
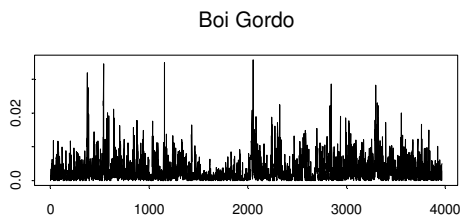
7.5 Histograma dos Log-Retornos (séries em primeira diferença)



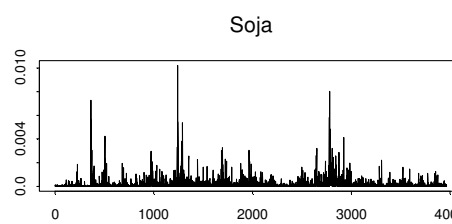
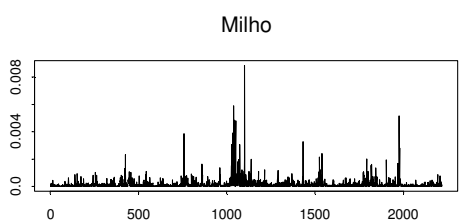
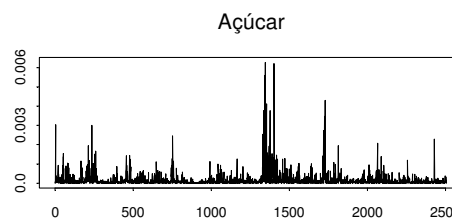
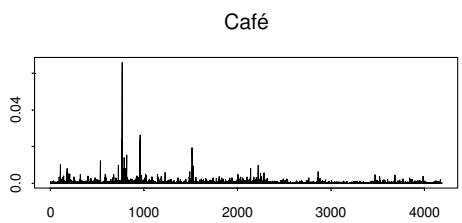
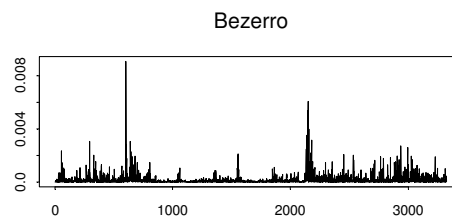
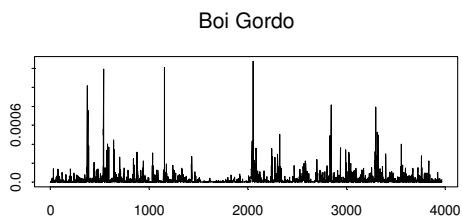
7.6 Gráficos QxQ dos Log-Retornos (séries em primeira diferença)



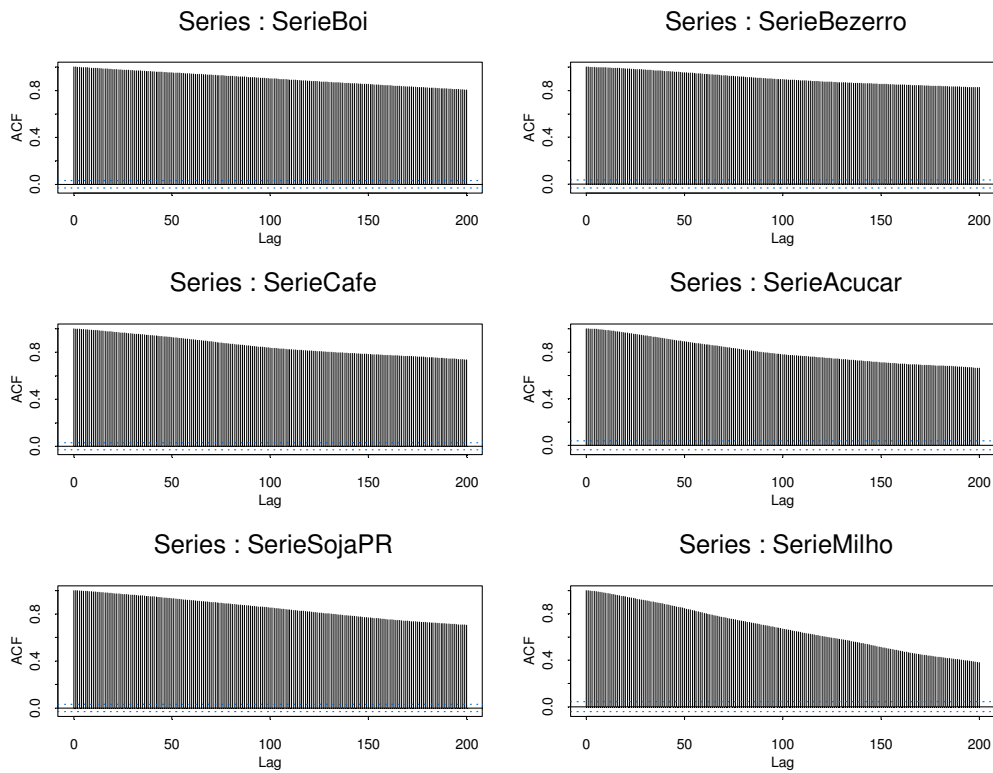
7.7 Retornos Absolutos



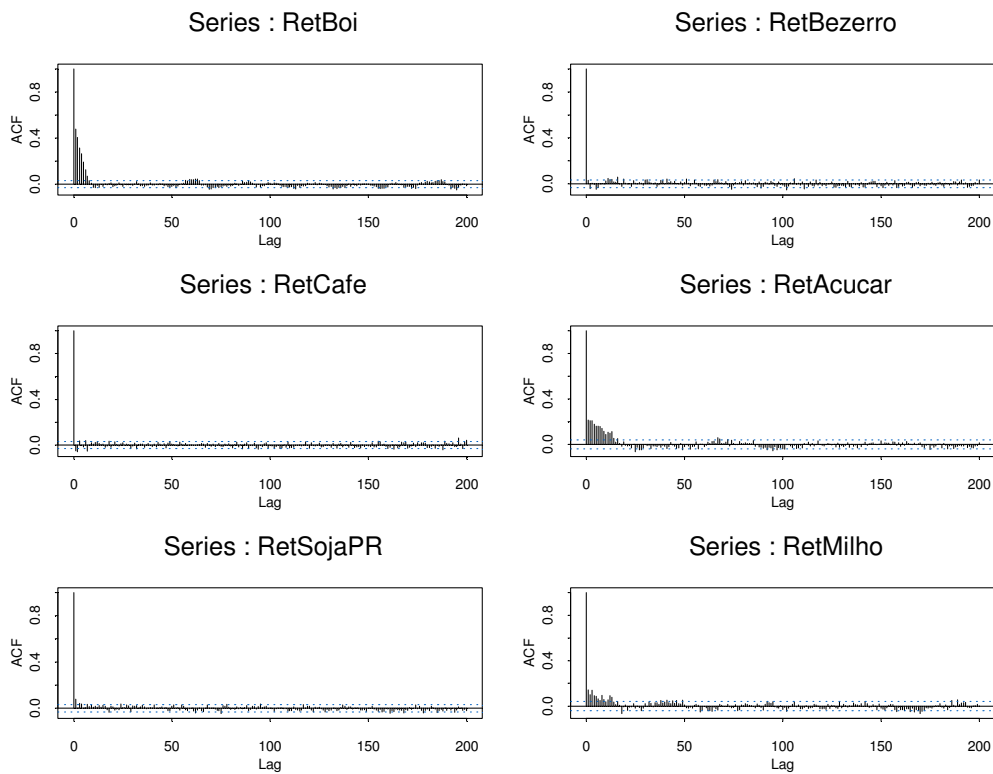
7.8 Retornos ao quadrado



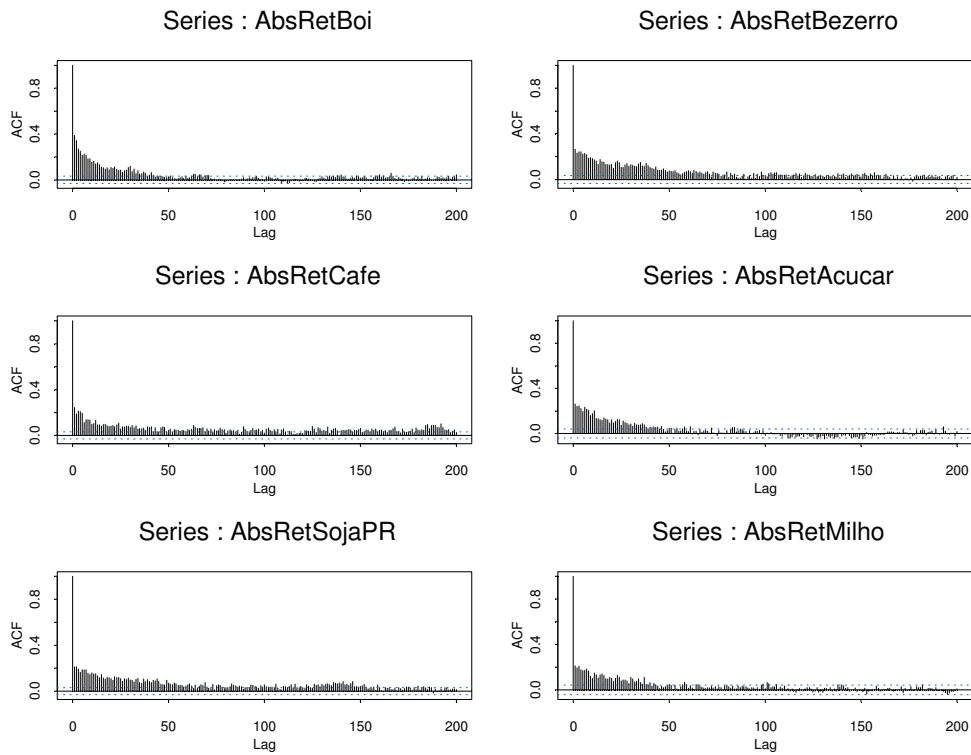
7.9 Funções de Autocorrelação (ACF) das séries de Preços



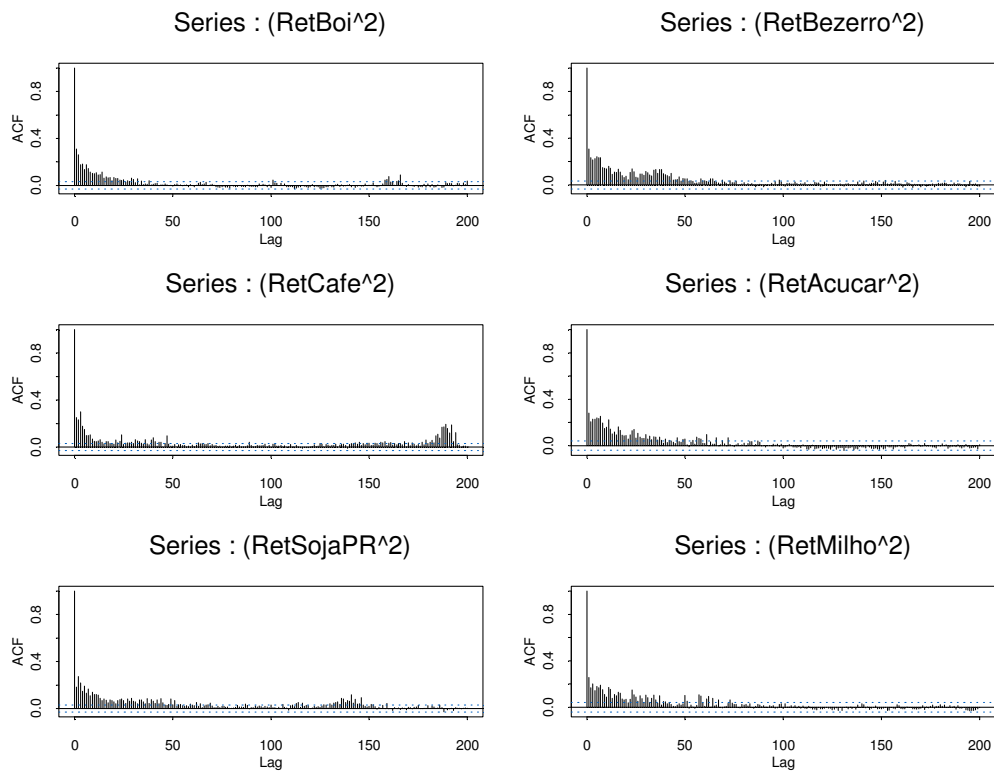
7.10 Funções de Autocorrelação (ACF) das séries de log-Retornos



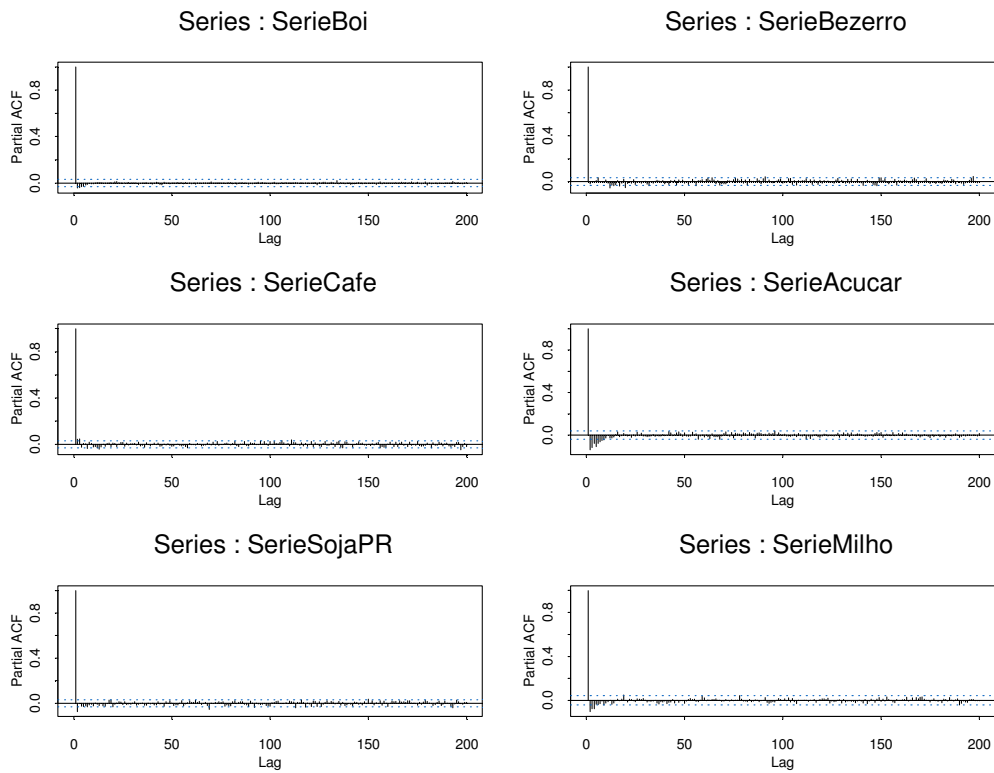
7.11 Funções de Autocorrelação (ACF) dos Retornos Absolutos



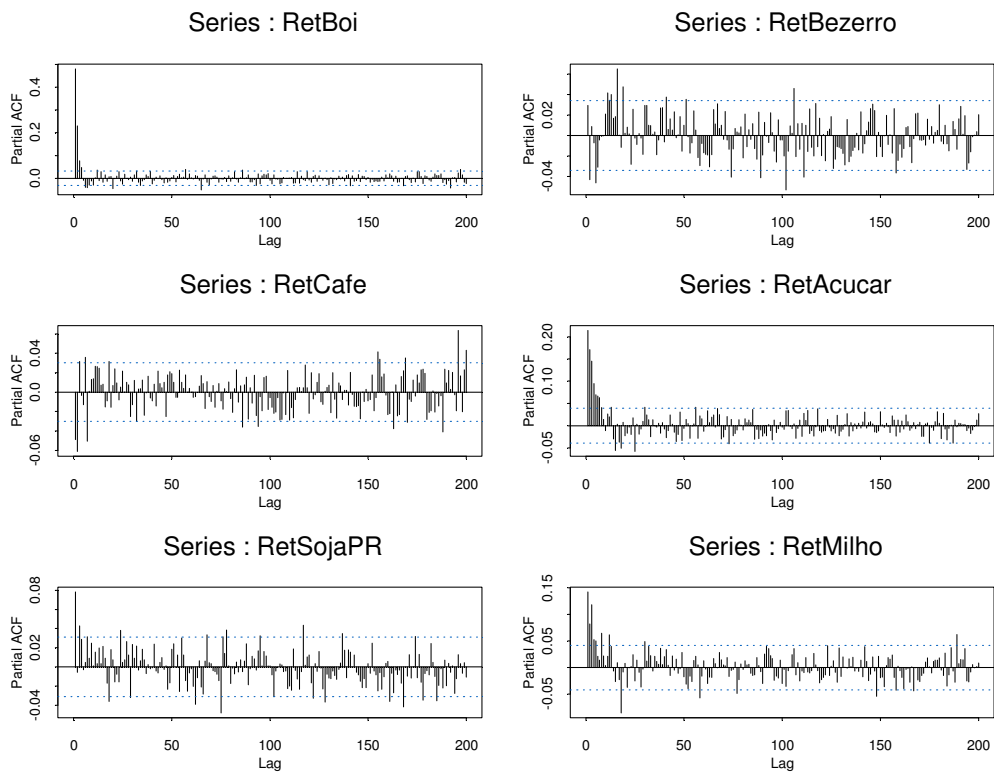
7.12 Funções de Autocorrelação (ACF) dos Retornos ao quadrado



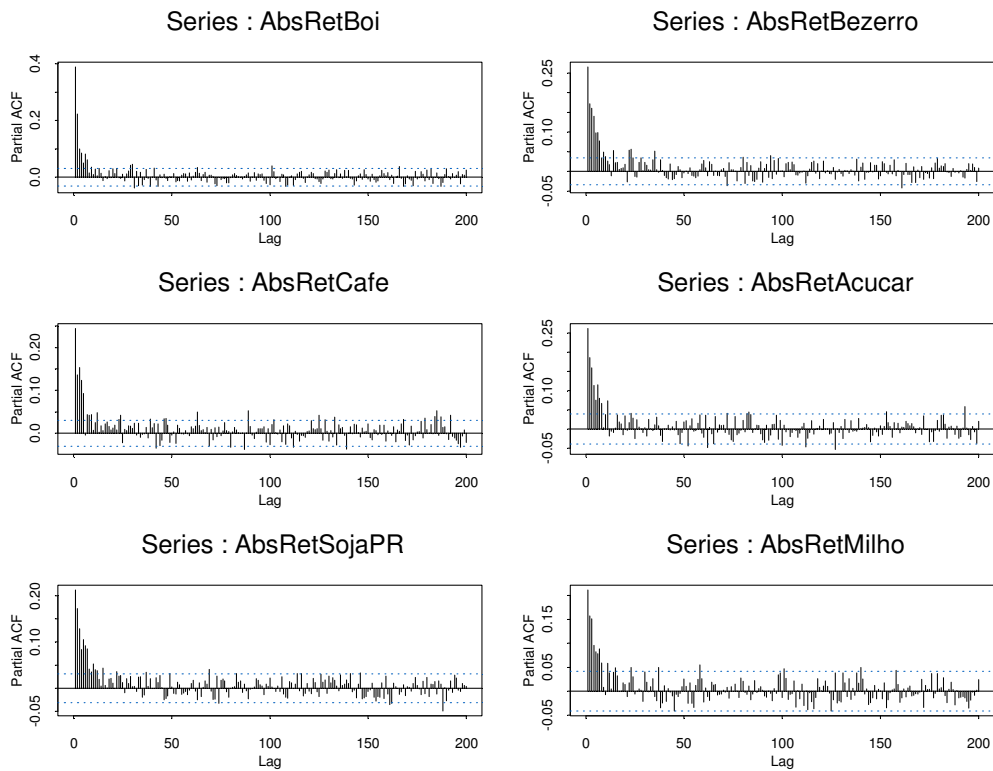
7.13 Funções de Autocorrelação Parcial (PACF) das séries de Preços



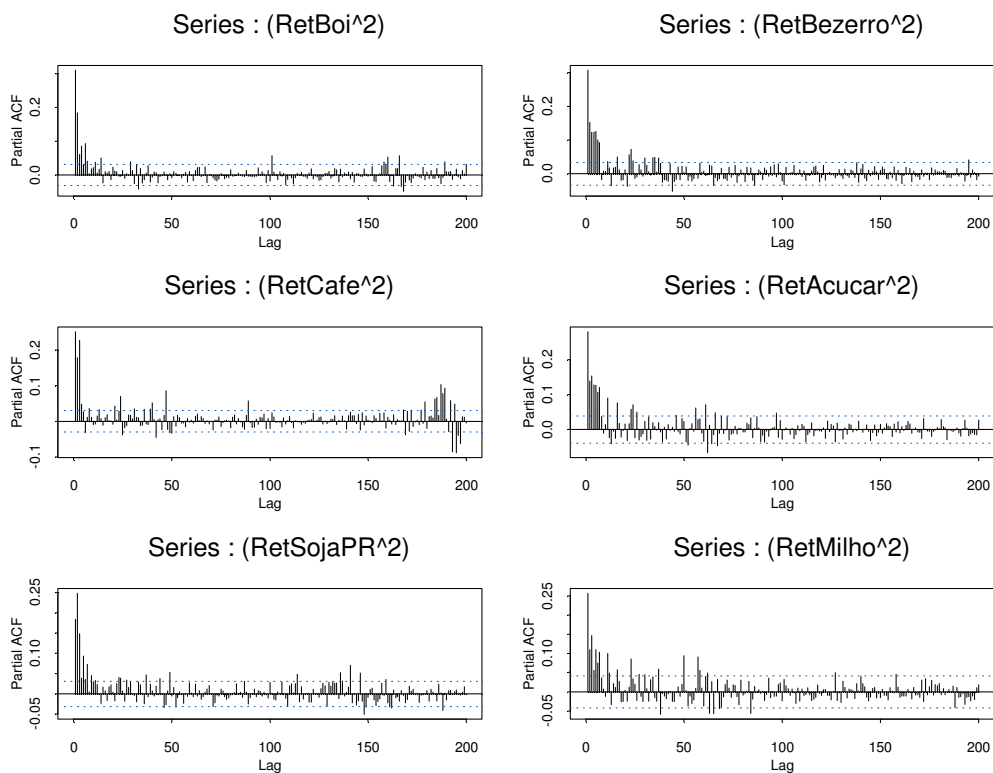
7.14 Funções de Autocorrelação Parcial (PACF) das séries de log-Retornos



7.15 Funções de Autocorrelação Parcial (PACF) dos Retornos Absolutos



7.16 Funções de Autocorrelação Parcial (PACF) dos Retornos ao quadrado



7.17 Resultados dos testes de Razão de Variância

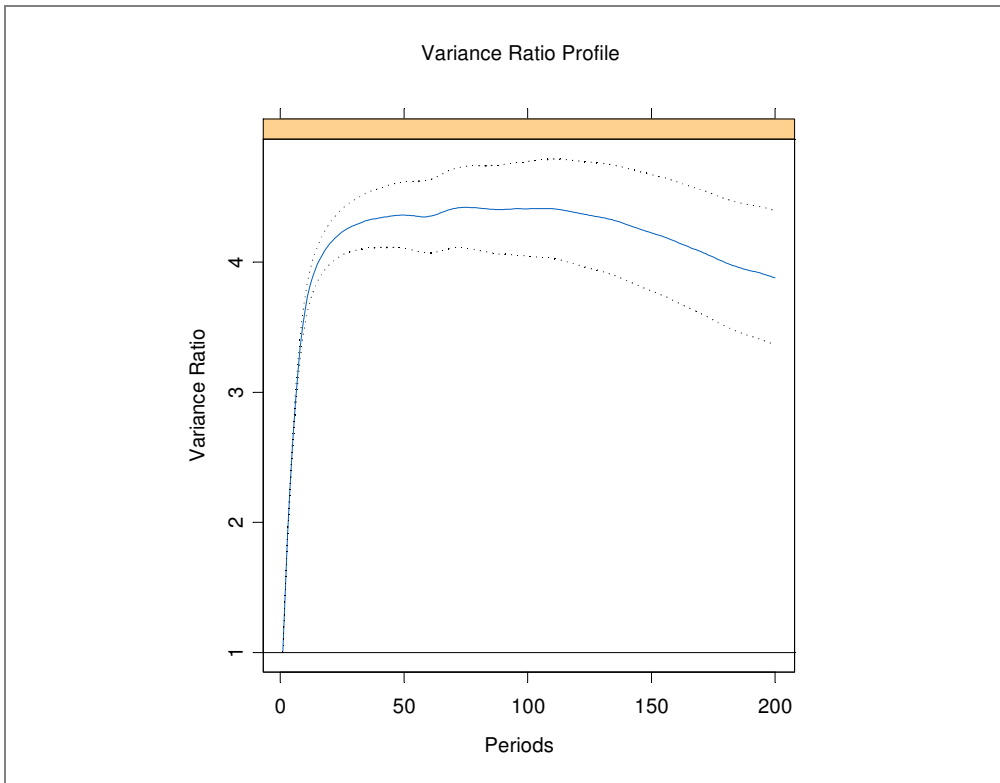


Figura 6.17.a: Resultados do teste de razão de variância para o boi gordo supondo erros homocedásticos.

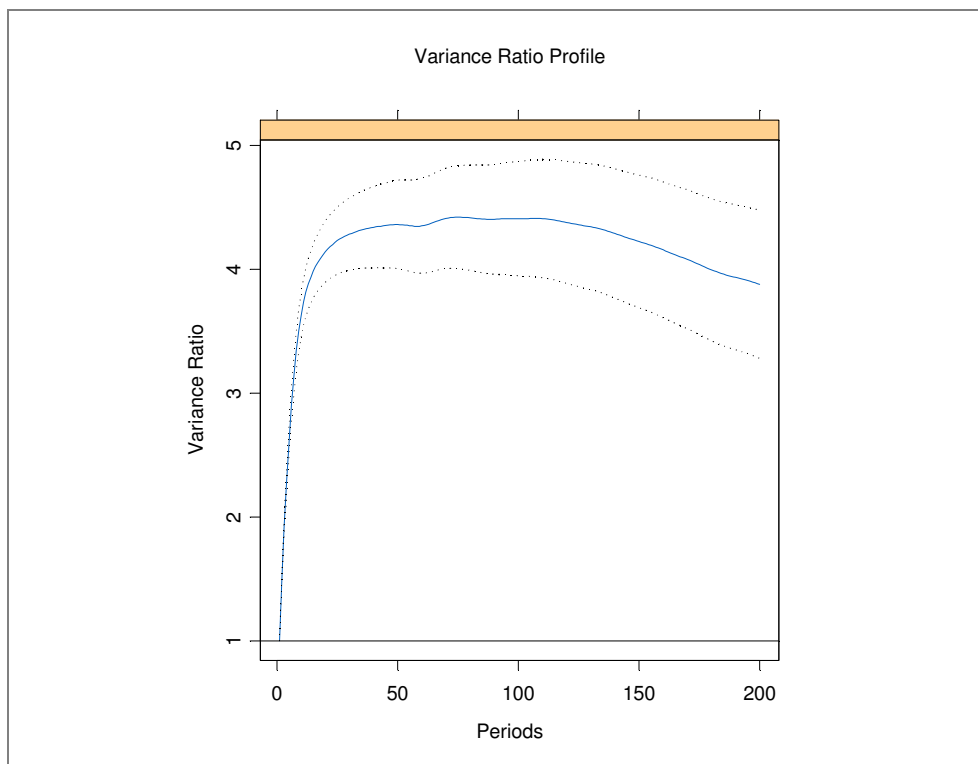


Figura 6.17.b: Resultados do teste de razão de variância para o boi gordo supondo erros heterocedásticos.

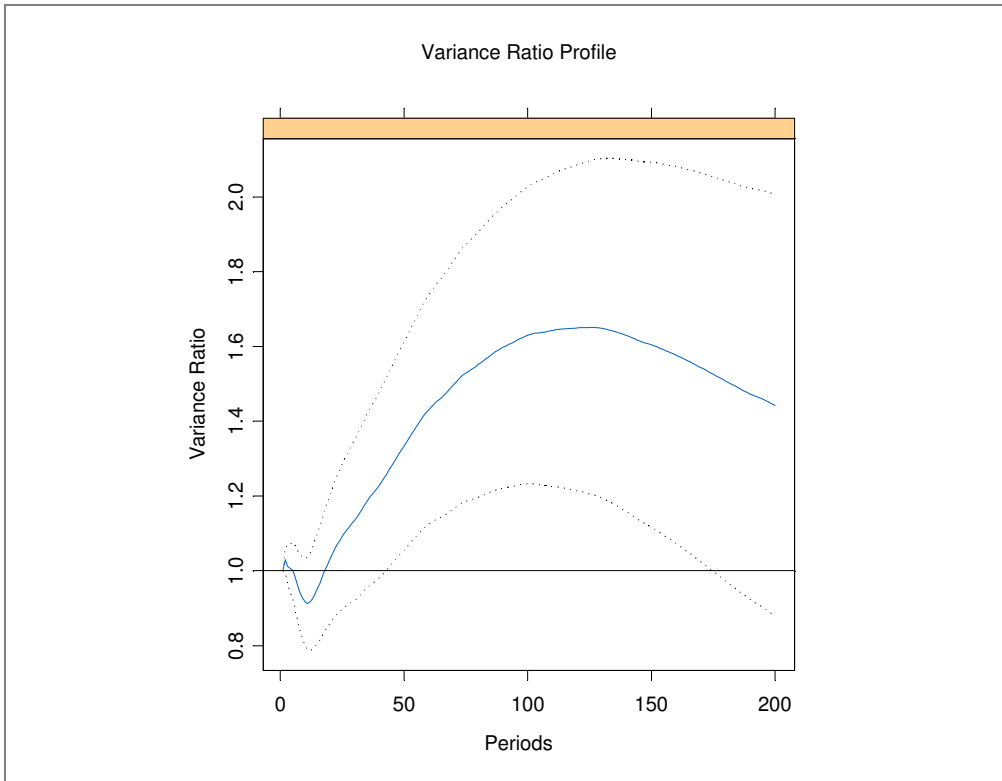


Figura 6.17.c: Resultados do teste de razão de variância do bezerro supondo erros homocedásticos.

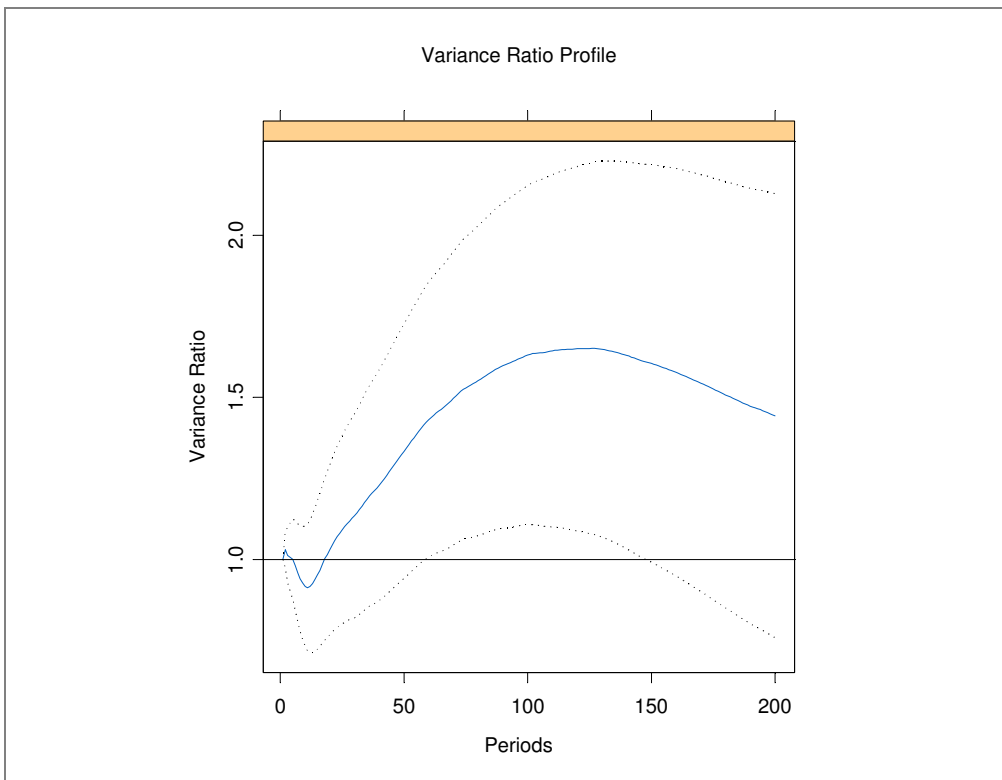


Figura 6.17.d: Resultados do teste de razão de variância do bezerro supondo erros heterocedásticos.

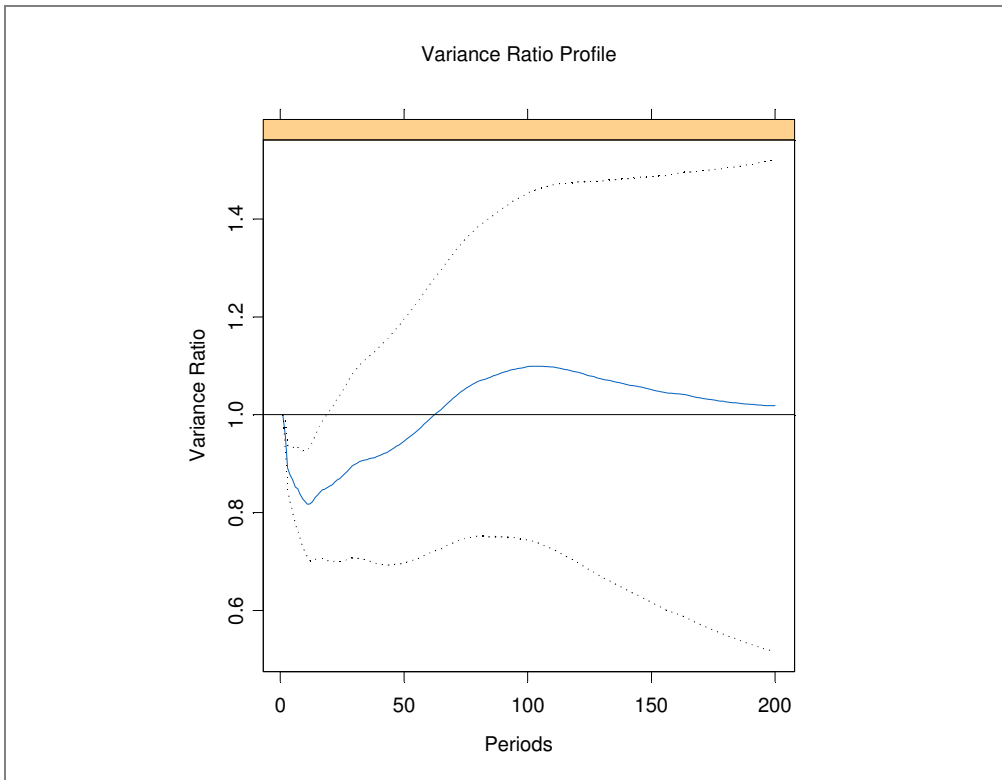


Figura 6.17.e: Resultados do teste de razão de variância do café supondo erros homocedásticos.

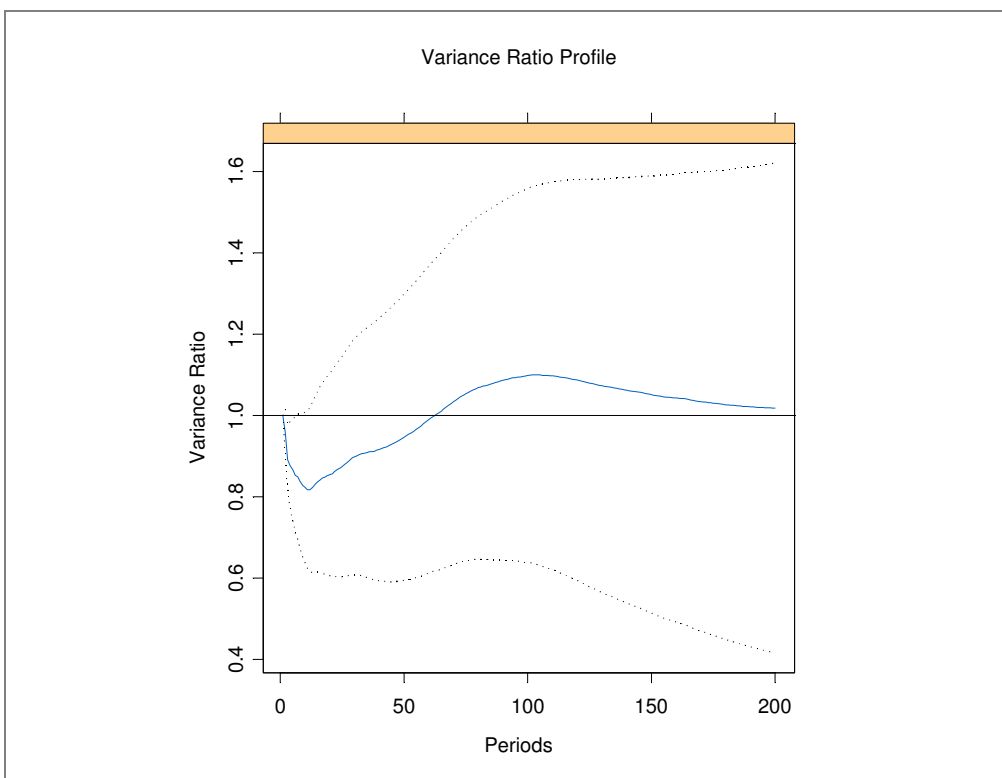


Figura 6.17.f: Resultados do teste de razão de variância do café supondo erros heterocedásticos.

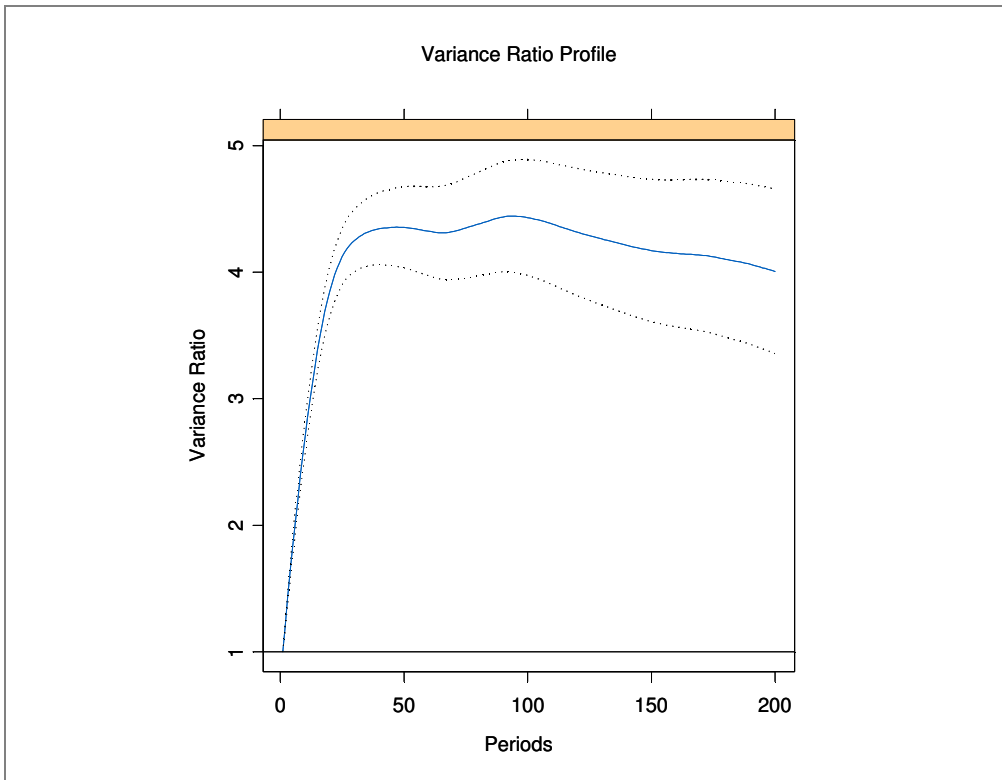


Figura 6.17.g: Resultados do teste de razão de variância do açúcar supondo erros homocedásticos.

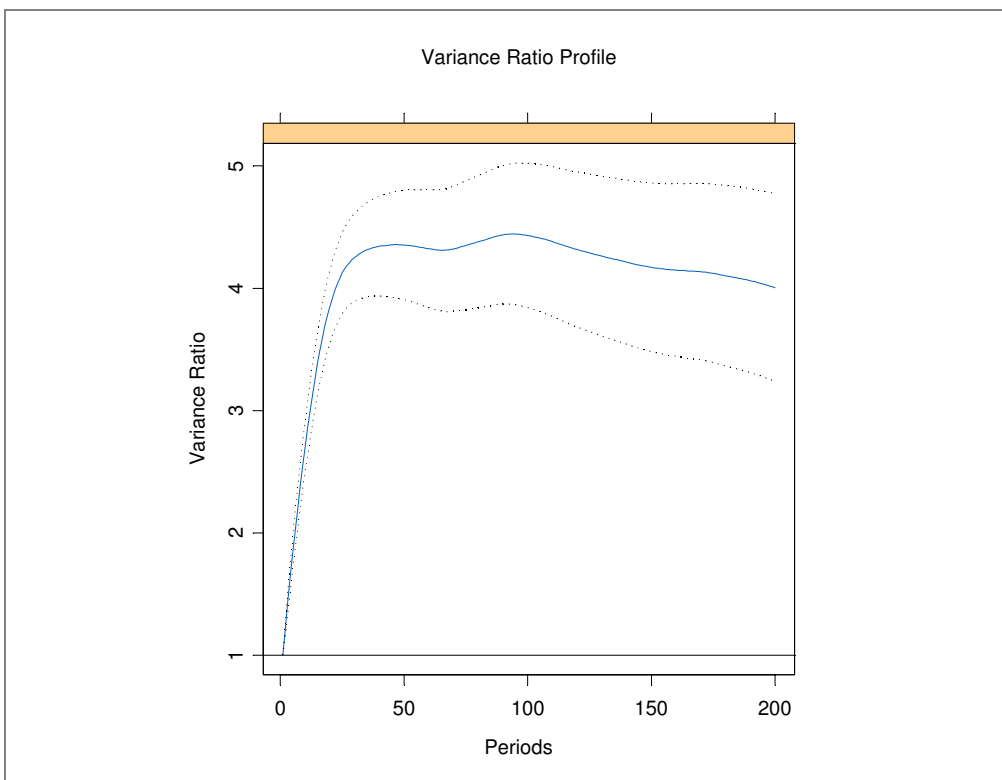


Figura 6.17.h: Resultados do teste de razão de variância do açúcar supondo erros heterocedásticos.

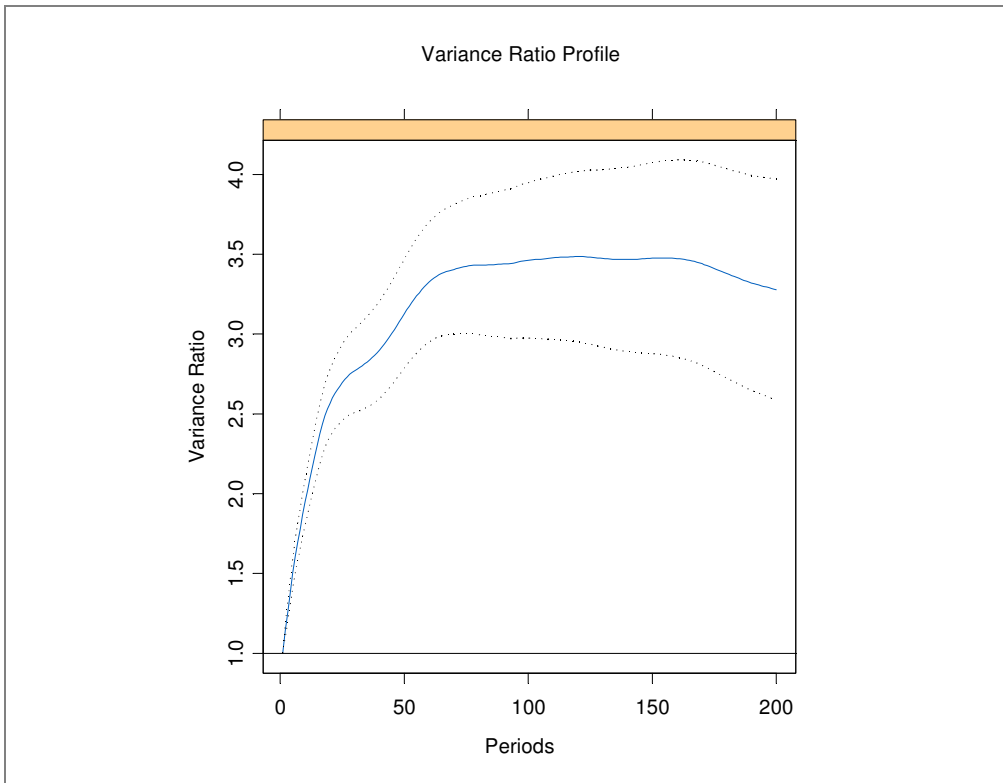


Figura 6.17.i: Resultados do teste de razão de variância do milho supondo erros homocedásticos.

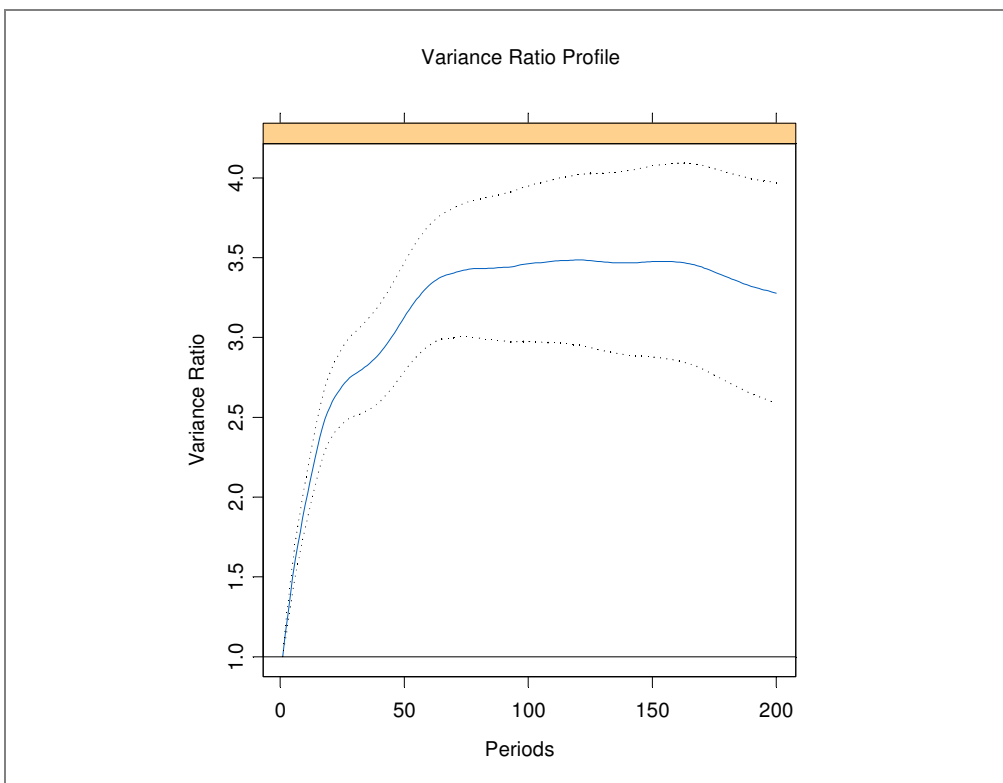


Figura 6.17.j: Resultados do teste de razão de variância do milho supondo erros heterocedásticos.

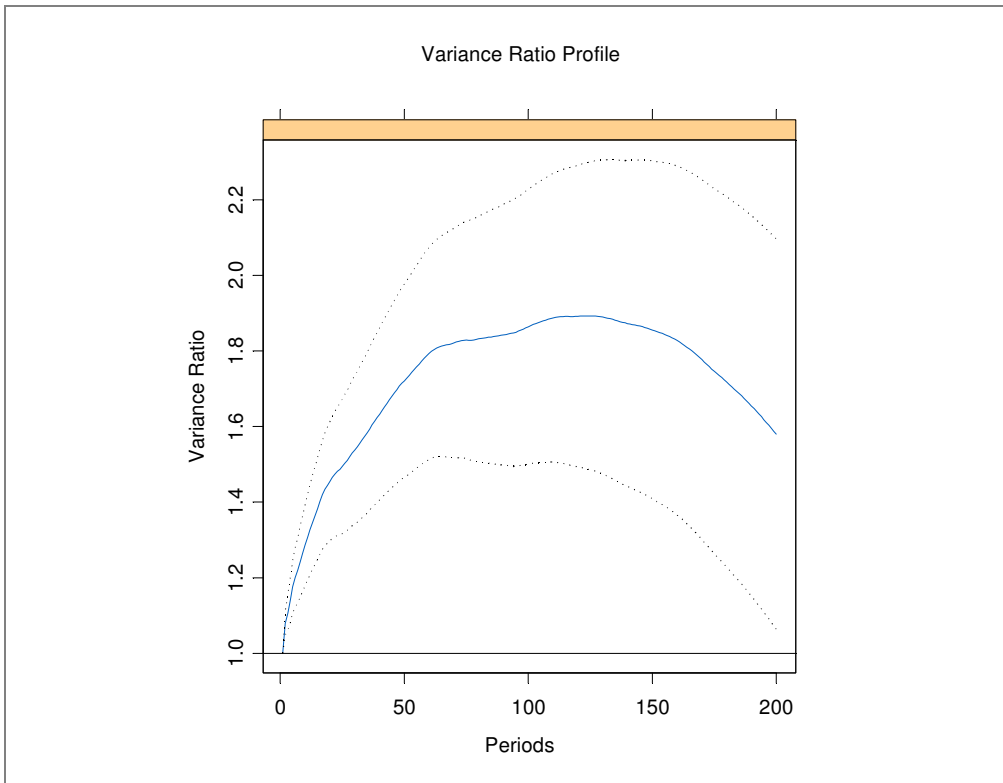


Figura 6.17.k: Resultados do teste de razão de variância da soja supondo erros homocedásticos.

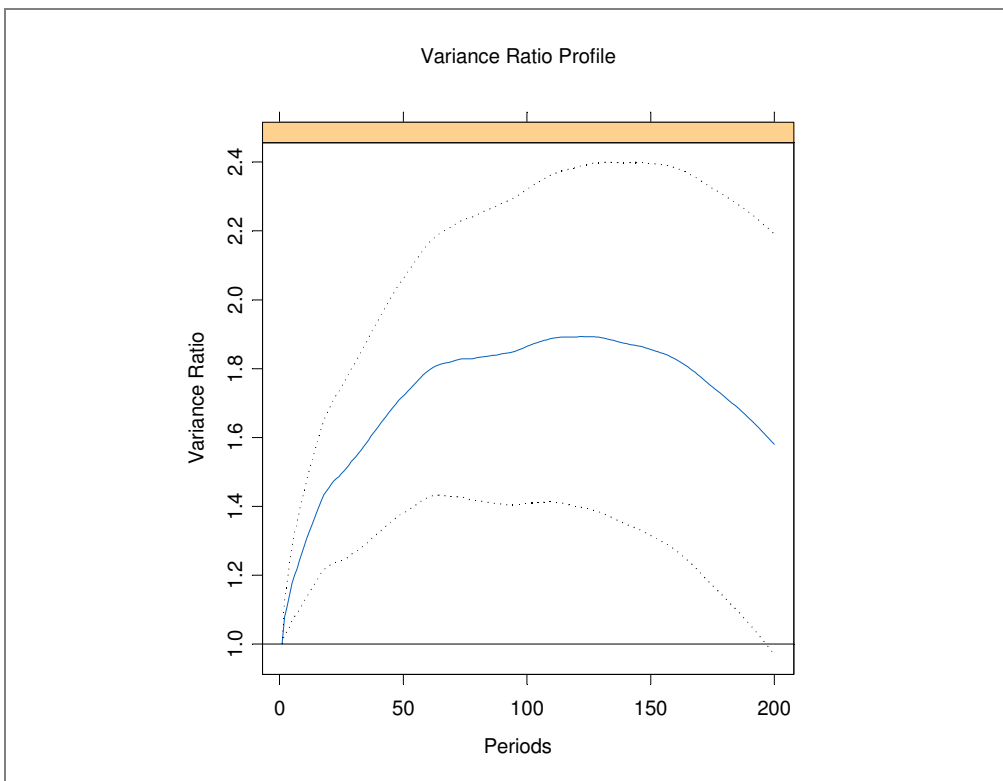


Figura 6.17.l: Resultados do teste de razão de variância do soja (região do Paraná) supondo erros heterocedásticos.

7.18 Output da estimação de um modelo ARFIMA (0,d,0)

7.18.1 Séries de Log-Preços

```
SÉRIE DE BOI GORDO

> LogSerieBoi.d = FARIMA(logSerieBoi, p = 0, q = 0)
> summary(LogSerieBoi.d)

Call:
FARIMA(x = logSerieBoi, p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d    1.3838   0.0124   111.1547  0.0000

Information Criteria:
log-likelihood      BIC
  16163.38         -32318.48

Residual scale estimate: 0.0041

                total residual
Degree of freedom: 3964      3962
```

```
SÉRIE DO BEZERRO

> LogSerieBezerro.d = FARIMA(logSerieBezerro, p = 0, q
= 0)
> summary(LogSerieBezerro.d)

Call:
FARIMA(x = logSerieBezerro, p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d    1.0103   0.0136    74.3373  0.0000

Information Criteria:
log-likelihood      BIC
  9839.819         -19671.529

Residual scale estimate: 0.0126

                total residual
Degree of freedom: 3326      3324
```

SÉRIE DO CAFÉ

```
> LogSerieCafe.d = FARIMA(logSerieCafe, p = 0, q = 0)
> summary(LogSerieCafe.d)
```

```
Call:
FARIMA(x = logSerieCafe, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.9637  0.0121    79.5481  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 10269.35         -20530.35
```

Residual scale estimate: 0.0208

```
                total residual
Degree of freedom:  4186      4184
```

SÉRIE DE AÇÚCAR

```
> LogSerieAcucar.d = FARIMA(logSerieAcucar, p = 0, q = 0)
> summary(LogSerieAcucar.d)
```

```
Call:
FARIMA(x = logSerieAcucar, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  1.1996  0.0157    76.6317  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 7480.099         -14952.371
```

Residual scale estimate: 0.0122

```
                total residual
Degree of freedom:  2507      2505
```

SÉRIE DE SOJA (região do Paraná)

```
> LogSerieSojaPR.d = FARIMA(logSerieSojaPR, p = 0, q = 0)
> summary(LogSerieSojaPR.d)
```

```
Call:
FARIMA(x = logSerieSojaPR, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  1.0652  0.0125    85.5499  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 11375.60         -22742.92
```

Residual scale estimate: 0.0137

```
                total residual
Degree of freedom:  3963      3961
```

```

SÉRIE DO MILHO

> LogSerieMilho.d = FARIMA(logSerieMilho, p = 0, q = 0)
> summary(LogSerieMilho.d)

Call:
FARIMA(x = logSerieMilho, p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d  1.1398  0.0166    68.4870  0.0000

Information Criteria:
log-likelihood      BIC
  6549.574      -13091.445

Residual scale estimate: 0.0126

                                total residual
Degree of freedom:  2218      2216

```

7.18.2 Séries de Log-Retornos

```

SÉRIE DO BOI GORDO

> RetBoi.d = FARIMA(RetBoi, p = 0, q = 0)
> summary(RetBoi.d)

Call:
FARIMA(x = RetBoi, p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.3838  0.0125    30.8249  0.0000

Information Criteria:
log-likelihood      BIC
 16158.82      -32309.35

Residual scale estimate: 0.0041

                                total residual
Degree of freedom:  3963      3961

```

```

SÉRIE DO BEZERRO

> "RetBezerro .d" = FARIMA(RetBezerro, p = 0, q = 0)
> summary(`RetBezerro .d`)

Call:
FARIMA(x = RetBezerro, p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.0103  0.0136    0.7550  0.4503

Information Criteria:
log-likelihood      BIC
  9836.371      -19664.633

Residual scale estimate: 0.0126

                                total residual
Degree of freedom:  3325      3323

```

SÉRIE DO CAFÉ

```
> RetCafe.d = FARIMA(RetCafe, p = 0, q = 0)
> summary(RetCafe.d)
```

```
Call:
FARIMA(x = RetCafe, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d -0.0364  0.0121    -3.0021  0.0027
```

```
Information Criteria:
log-likelihood      BIC
10267.33           -20526.33
```

Residual scale estimate: 0.0208

```
                total residual
Degree of freedom: 4185      4183
```

SÉRIE DE AÇÚCAR

```
> RetAcucar.d = FARIMA(RetAcucar, p = 0, q = 0)
> summary(RetAcucar.d)
```

```
Call:
FARIMA(x = RetAcucar, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.1997  0.0157    12.7564  0.0000
```

```
Information Criteria:
log-likelihood      BIC
7476.694           -14945.561
```

Residual scale estimate: 0.0122

```
                total residual
Degree of freedom: 2506      2504
```

SÉRIE DE SOJA (região do Paraná)

```
> RetSojaPR.d = FARIMA(RetSojaPR, p = 0, q = 0)
> summary(RetSojaPR.d)
```

```
Call:
FARIMA(x = RetSojaPR, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.0652  0.0125     5.2393  0.0000
```

```
Information Criteria:
log-likelihood      BIC
11372.30           -22736.32
```

Residual scale estimate: 0.0137

```
                total residual
Degree of freedom: 3962      3960
```

SÉRIE DO MILHO

```
> RetMilho.d = FARIMA(RetMilho, p = 0, q = 0)
> summary(RetMilho.d)
```

```
Call:
FARIMA(x = RetMilho, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d 0.1399 0.0166      8.4013 0.0000
```

```
Information Criteria:
log-likelihood      BIC
6546.141          -13084.578
```

```
Residual scale estimate: 0.0126
```

```
                total residual
Degree of freedom: 2217      2215
```

7.18.3 Séries de Retornos Absolutos

SÉRIE DO BOI GORDO

```
> AbsRetBoi.d = FARIMA(AbsRetBoi, p = 0, q = 0)
> summary(AbsRetBoi.d)
```

```
Call:
FARIMA(x = AbsRetBoi, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d 0.2915 0.0125     23.4155 0.0000
```

```
Information Criteria:
log-likelihood      BIC
17177.40          -34346.52
```

```
Residual scale estimate: 0.0032
```

```
                total residual
Degree of freedom: 3963      3961
```

```
SÉRIE DO BEZERRO
> AbsRetBezerro.d = FARIMA(AbsRetBezerro, p = 0, q = 0)
> summary(AbsRetBezerro.d)
```

```
Call:
FARIMA(x = AbsRetBezerro, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2146  0.0136    15.7855  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 11279.43          -22550.76
```

```
Residual scale estimate: 0.0081
```

```
                total residual
Degree of freedom: 3325      3323
```

```
SÉRIE DO CAFÉ
```

```
> AbsRetCafe.d = FARIMA(AbsRetCafe, p = 0, q = 0)
> summary(AbsRetCafe.d)
```

```
Call:
FARIMA(x = AbsRetCafe, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2026  0.0121    16.7255  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 11765.70          -23523.06
```

```
Residual scale estimate: 0.0145
```

```
                total residual
Degree of freedom: 4185      4183
```

```
SÉRIE DE AÇÚCAR
```

```
> AbsRetAcucar.d = FARIMA(AbsRetAcucar, p = 0, q = 0)
> summary(AbsRetAcucar.d)
```

```
Call:
FARIMA(x = AbsRetAcucar, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2179  0.0157    13.9173  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 8476.141          -16944.455
```

```
Residual scale estimate: 0.0082
```

```
                total residual
Degree of freedom: 2506      2504
```

SÉRIE DE SOJA (região do Paraná)

```
> AbsRetSojaPR.d = FARIMA(AbsRetSojaPR, p = 0, q = 0)
> summary(AbsRetSojaPR.d)
```

```
Call:
FARIMA(x = AbsRetSojaPR, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.1888  0.0125    15.1636  0.0000
```

```
Information Criteria:
log-likelihood      BIC
 12984.31         -25960.34
```

```
Residual scale estimate: 0.0091
```

```
                total residual
Degree of freedom: 3962      3960
```

SÉRIE DO MILHO

```
> AbsRetMilho.d = FARIMA(AbsRetMilho, p = 0, q = 0)
> summary(AbsRetMilho.d)
```

```
Call:
FARIMA(x = AbsRetMilho, p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.1898  0.0166    11.4006  0.0000
```

```
Information Criteria:
log-likelihood      BIC
  7415.953         -14824.203
```

```
Residual scale estimate: 0.0085
```

```
                total residual
Degree of freedom: 2217      2215
```

7.18.4 Séries de Retornos ao quadrado

```
SÉRIE DO BOI GORDO

> Ret2Boi.d = FARIMA((RetBoi^2), p = 0, q = 0)
> summary(Ret2Boi.d)

Call:
FARIMA(x = (RetBoi^2), p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2481  0.0125     19.9281  0.0000

Information Criteria:
log-likelihood      BIC
  32919.80         -65831.32

Residual scale estimate: 0.0001

                        total residual
Degree of freedom:  3963      3961
```

```
SÉRIE DO BEZERRO

> Ret2Bezerro.d = FARIMA((RetBezerro^2), p = 0, q = 0)
> summary(Ret2Bezerro.d)

Call:
FARIMA(x = (RetBezerro^2), p = 0, q = 0)

Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2374  0.0136     17.4616  0.0000

Information Criteria:
log-likelihood      BIC
  21687.43         -43366.74

Residual scale estimate: 0.0004

                        total residual
Degree of freedom:  3325      3323
```

SÉRIE DO CAFÉ

```
> Ret2Cafe.d = FARIMA((RetCafe^2), p = 0, q = 0)
> summary(Ret2Cafe.d)
```

```
Call:
FARIMA(x = (RetCafe^2), p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2183  0.0121    18.0167  0.0000
```

```
Information Criteria:
log-likelihood      BIC
21189.27          -42370.21
```

```
Residual scale estimate: 0.0015
```

```
                total residual
Degree of freedom: 4185      4183
```

SÉRIE DE AÇÚCAR

```
> Ret2Acucar.d = FARIMA((RetAcucar^2), p = 0, q = 0)
> summary(Ret2Acucar.d)
```

```
Call:
FARIMA(x = (RetAcucar^2), p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2233  0.0157    14.2633  0.0000
```

```
Information Criteria:
log-likelihood      BIC
16258.96          -32510.10
```

```
Residual scale estimate: 0.0004
```

```
                total residual
Degree of freedom: 2506      2504
```

SÉRIE DE SOJA (região do Paraná)

```
> Ret2SojaPR.d = FARIMA((RetSojaPR^2), p = 0, q = 0)
> summary(Ret2SojaPR.d)
```

```
Call:
FARIMA(x = (RetSojaPR^2), p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.1871  0.0125    15.0245  0.0000
```

```
Information Criteria:
log-likelihood      BIC
24907.16           -49806.03
```

Residual scale estimate: 0.0005

```
                total residual
Degree of freedom: 3962      3960
```

SÉRIE DO MILHO

```
> Ret2Milho.d = FARIMA((RetMilho^2), p = 0, q = 0)
> summary(Ret2Milho.d)
```

```
Call:
FARIMA(x = (RetMilho^2), p = 0, q = 0)
```

```
Coefficients:
      Value Std. Error t value Pr(>|t|)
d  0.2070  0.0166    12.4338  0.0000
```

```
Information Criteria:
log-likelihood      BIC
14132.60           -28257.49
```

Residual scale estimate: 0.0004

```
                total residual
Degree of freedom: 2217      2215
```

7.19 Resultados da estimação do componente de *drift* (log-retornos)

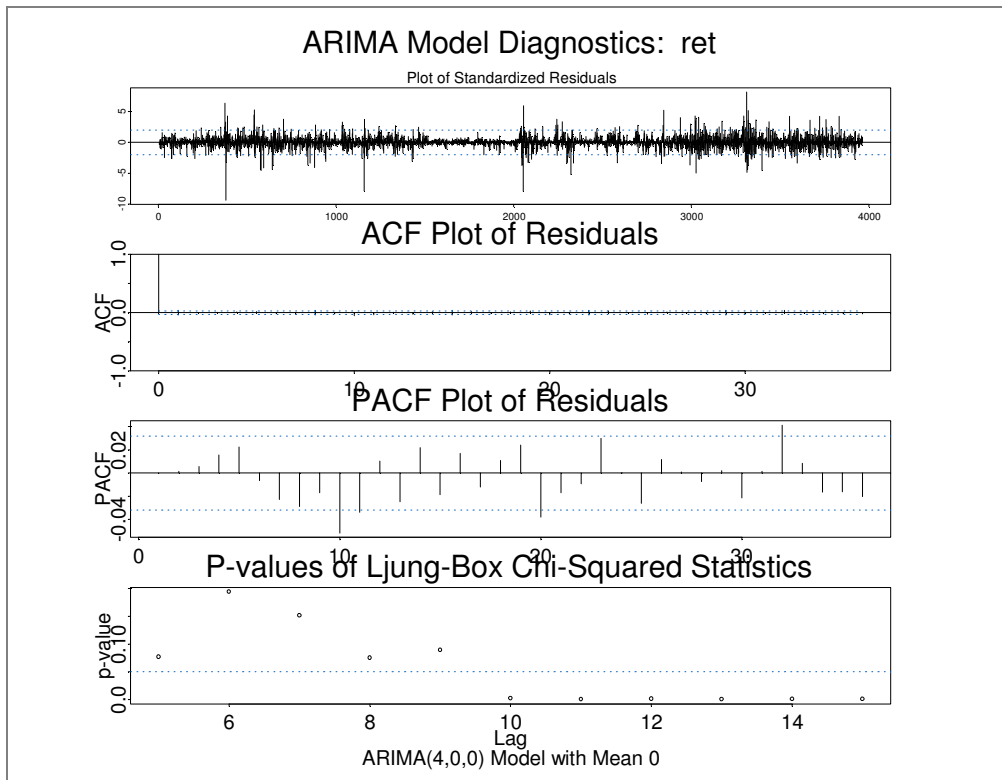


Figura 6.19.a: ajuste de um modelo ARIMA (4,0,0) sobre os log-retornos do boi gordo.

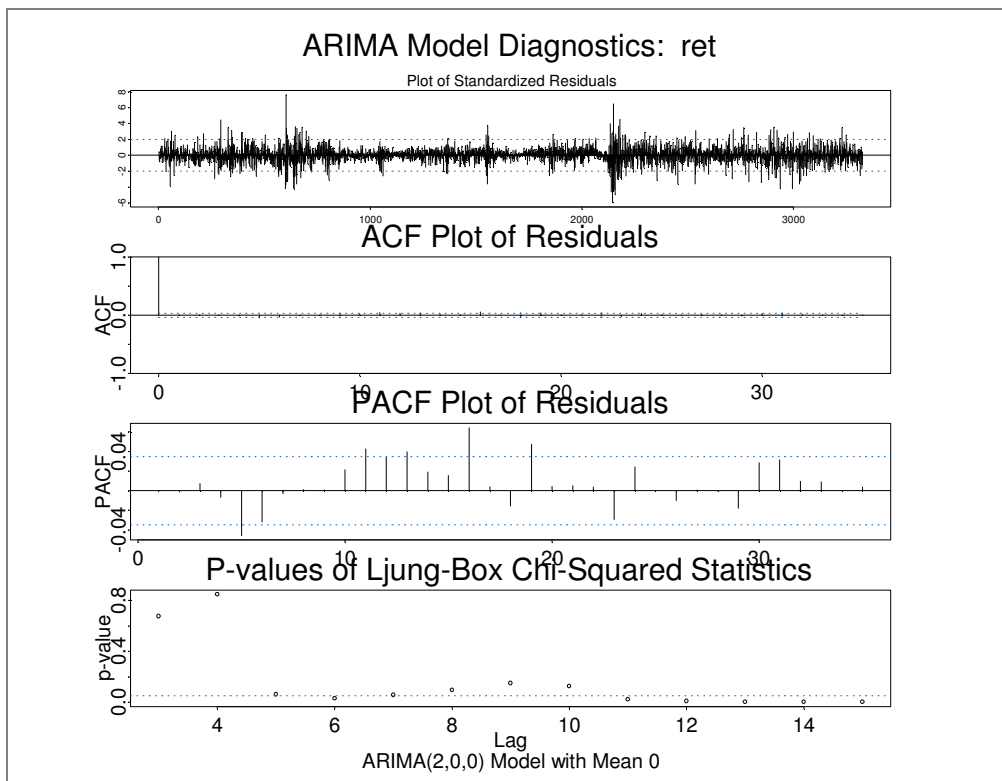


Figura 6.19.b: ajuste de um modelo ARIMA (2,0,0) sobre os log-retornos do bezerro.

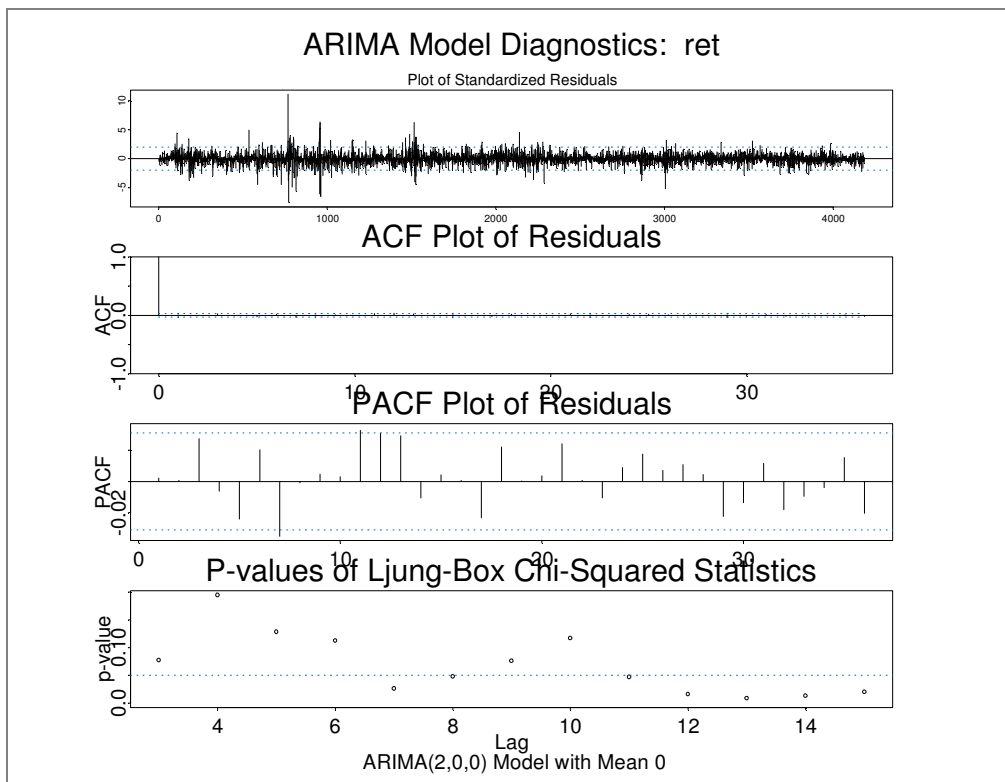


Figura 6.19.c: ajuste de um modelo ARIMA (2,0,0) sobre os log-retornos do café.

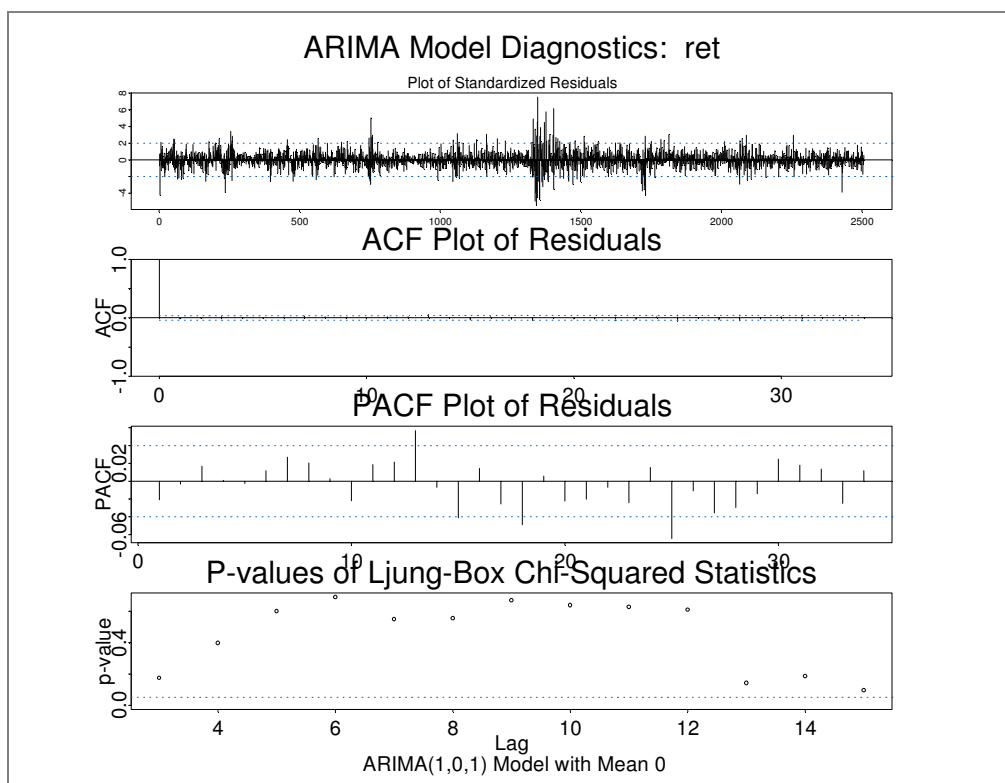


Figura 6.19.d: ajuste de um modelo ARIMA (1,0,1) sobre os log-retornos do açúcar.

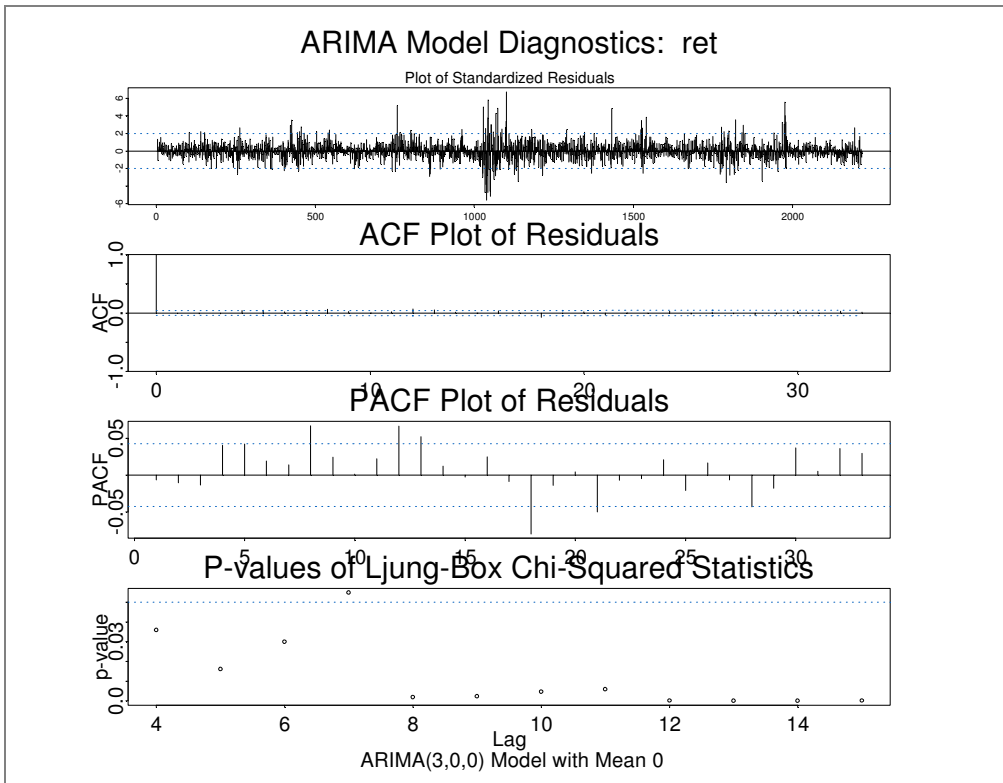


Figura 6.19.e: ajuste de um modelo ARIMA (3,0,0) sobre os log-retornos.

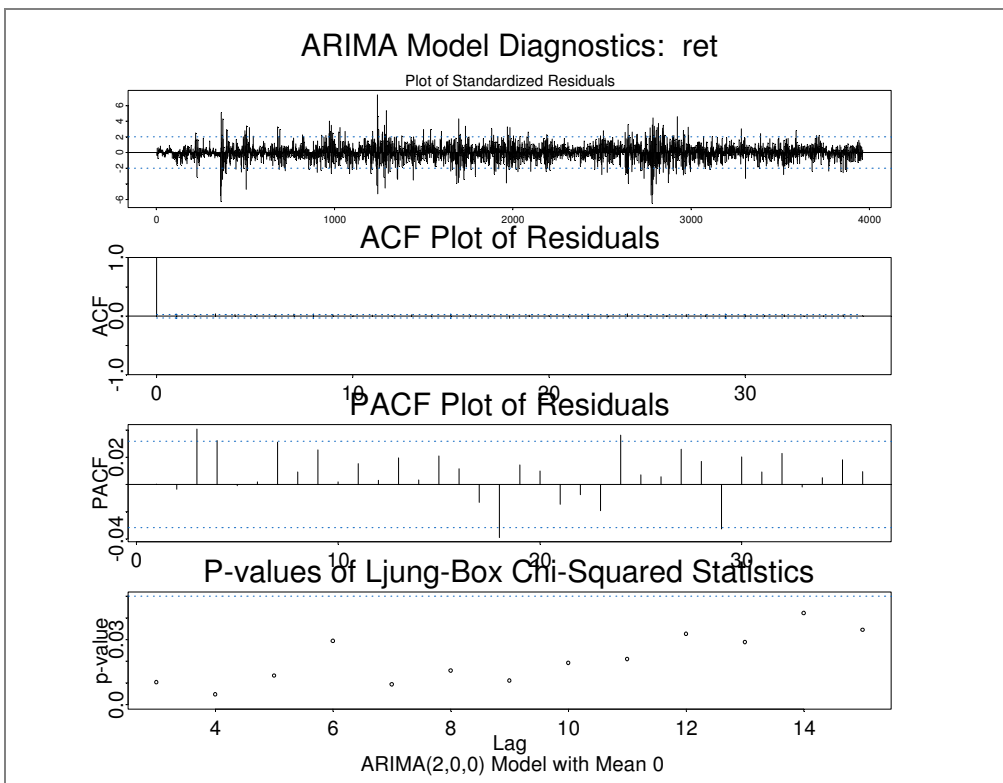


Figura 6.19.f: ajuste de um modelo ARIMA (2,0,0) sobre os log-retornos da soja.

7.20 Output da estimação conjunta do componente de *drift* e da variância

7.20.1 Boi Gordo

```
> boi.figarch = fgarch(RetBoi ~ -1 + arma(4, 0), ~ figarch(1, 1))
Iteration No. 1: log-likelihood=16854.490892

> summary(boi.figarch)
Call:
fgarch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~ figarch(1,
1))

Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ figarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value Std.Error t value Pr(>|t|)
AR(1) 3.811e-001 1.902e-002 20.034 0.000e+000
AR(2) 2.140e-001 1.660e-002 12.894 0.000e+000
AR(3) 6.750e-002 1.702e-002 3.967 3.705e-005
AR(4) 6.385e-002 1.612e-002 3.961 3.793e-005
A 3.339e-007 4.130e-008 8.085 4.441e-016
GARCH(1) 7.000e-001 2.396e-002 29.216 0.000e+000
ARCH(1) 4.000e-001 2.524e-002 15.850 0.000e+000
              Value Std.Error t value Pr(>|t|)
fraction 0.5 0.03244 15.41 0

-----

AIC(8) = -33692.98
BIC(8) = -33642.7

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
3542 0 0.95 9.237e-035

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
20.94 0.05133 12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
24.95 0.01505 12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12
2.836 0.653 1.458 -0.03251 2.052 -0.597 0.1343 -0.209 -0.6918 -0.4387
2.5 0.04711

C
-0.2868

TR^2 P-value F-stat P-value
23.36 0.02483 2.136 0.07642
```

```

> boi.fiegarch = fgarch(RetBoi ~ -1 + arma(4, 0), ~ fiegarch(1, 1),
  leverage = T)
> summary(boi.fiegarch)

Call:
fgarch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~
  fiegarch(1, 1), leverage =
  TRUE)

Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)
Conditional Variance Equation: structure(.Data = ~ fiegarch(1, 1)
, class = "formula"
)
-----

Estimated Coefficients:
-----
              Value Std.Error t value   Pr(>|t|)
AR(1)  0.36780    0.01690  21.768 0.000e+000
AR(2)  0.21964    0.01681  13.065 0.000e+000
AR(3)  0.08726    0.01612   5.415 3.254e-008
AR(4)  0.04457    0.01393   3.200 6.922e-004
A     -0.38845    0.01674 -23.203 0.000e+000
GARCH(1) 0.38072    0.04364   8.725 0.000e+000
ARCH(1)  0.41493    0.01516  27.378 0.000e+000
              Value Std.Error t value   Pr(>|t|)
LEV(1) -0.03175    0.007459  -4.256 0.00001063
fraction 0.53097    0.018226  29.132 0.000000000
-----

AIC(9) = -33750.43
BIC(9) = -33693.86

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      4737         0         0.9454 5.069e-036

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      27.55 0.006435         12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      15.88 0.1966         12

Lagrange multiplier test:
-----
Lag 1  Lag 2 Lag 3  Lag 4 Lag 5  Lag 6  Lag 7 Lag 8  Lag 9  Lag 10 Lag
11  Lag 12      C
0.1134 -0.3121  1.06 -0.328 1.778 0.2485 0.7817 0.237 -0.363 -0.4778
2.914 0.04362 -0.716

TR^2 P-value F-stat P-value
15.17 0.2322 1.385 0.2818

```

```

> boi.garch = garch(RetBoi ~ -1 + arma(4, 0), ~ garch(1, 1), trace = F)
> summary(boi.garch)

Call:
garch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~ garch(1,
1), trace = FALSE)

Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ garch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----

Estimated Coefficients:
-----
              Value  Std.Error t value  Pr(>|t|)
AR(1) 3.811e-001 1.799e-002  21.181 0.0000000
AR(2) 2.140e-001 1.818e-002  11.771 0.0000000
AR(3) 6.750e-002 1.825e-002   3.699 0.0002196
AR(4) 6.385e-002 1.679e-002   3.804 0.0001448
  A 3.339e-007 2.519e-008  13.255 0.0000000
ARCH(1) 1.733e-001 6.403e-003  27.070 0.0000000
GARCH(1) 8.291e-001 5.127e-003 161.709 0.0000000

-----

AIC(7) = -33651.79
BIC(7) = -33607.79

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk      P-value
      4637          0          0.9452 4.499e-036

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      22.34 0.03392          12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      18.31 0.1065          12

Lagrange multiplier test:
-----
Lag 1  Lag 2  Lag 3  Lag 4  Lag 5  Lag 6  Lag 7  Lag 8  Lag 9  Lag 10
Lag 11  Lag 12
0.2941 -0.7664 -0.3186 -0.7435 0.2196 -1.187 -0.5375 -0.5826 -1.115 -1.187
3.436 0.04396

C
0.152

TR^2 P-value F-stat P-value
18.06 0.114 1.649 0.175

```

```

> boi.egarch = garch(RetBoi ~ -1 + arma(4, 0), ~ egarch(1, 1), leverage =
  T, trace = F)
> summary(boi.egarch)
Call:
garch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~ egarch(1,
  1), leverage = TRUE,
  trace = FALSE)

Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)
Conditional Variance Equation: structure(.Data = ~ egarch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian
-----
Estimated Coefficients:
-----
              Value Std.Error  t value  Pr(>|t|)
AR(1)  0.37185  0.016414  22.654 0.000e+000
AR(2)  0.20643  0.016633  12.411 0.000e+000
AR(3)  0.09102  0.017011   5.350 9.271e-008
AR(4)  0.04834  0.014735   3.281 1.044e-003
  A -0.96744  0.045733 -21.154 0.000e+000
ARCH(1) 0.34981  0.010650  32.847 0.000e+000
GARCH(1) 0.93653  0.003637 257.485 0.000e+000
              Value Std.Error t value  Pr(>|t|)
LEV(1) -0.08718  0.01869 -4.664 3.208e-006
-----

AIC(8) = -33714.74
BIC(8) = -33664.46

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      4042      0      0.9485 3.543e-035

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      20.76  0.054      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      15.72  0.2043      12

Lagrange multiplier test:
-----
Lag 1  Lag 2 Lag 3  Lag 4  Lag 5  Lag 6  Lag 7  Lag 8  Lag 9  Lag
10 Lag 11 Lag 12
0.573 -0.6136 0.154 -0.6454 0.9828 -0.6618 -0.04885 -0.1346 -0.8118 -
0.8975 3.314 0.04192

C
-0.2036

TR^2 P-value F-stat P-value
15.33 0.2237 1.399 0.2743

```

```

> boi.comp = compare.mgarch(boi.garch, boi.egarch, boi.figarch,
  boi.fiegarch)
> boi.comp
      boi.garch boi.egarch boi.figarch boi.fiegarch
AIC      -33652   -33715   -33693   -33750
BIC      -33608   -33664   -33643   -33694
Likelihood 16833   16865   16854   16884

```

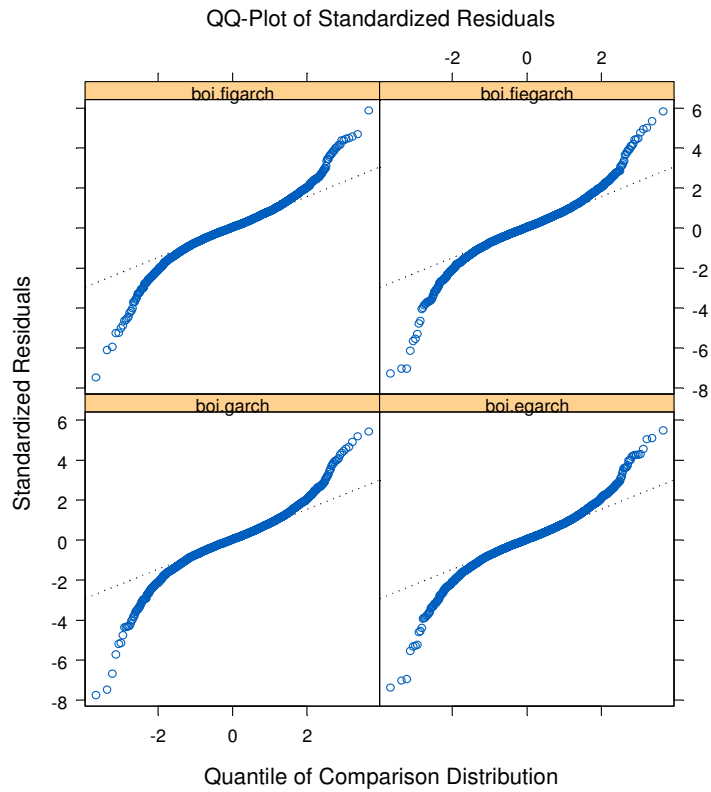


Figura 6.20.1: resultados do ajuste dos modelos da família GARCH para a modelagem da variância em conjunto com o componente de *drift* (dado pelos log-retornos) do boi gordo.

7.20.2 Bezerro

```
> bezerro.figarch = fgarch(RetBezerro ~ -1 + arma(2, 0), ~ figarch(1, 1))
> summary(boi.figarch)

Call:
fgarch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~ figarch(1,
1))

Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ figarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value Std.Error t value Pr(>|t|)
AR(1) 3.811e-001 1.902e-002 20.034 0.000e+000
AR(2) 2.140e-001 1.660e-002 12.894 0.000e+000
AR(3) 6.750e-002 1.702e-002 3.967 3.705e-005
AR(4) 6.385e-002 1.612e-002 3.961 3.793e-005
  A 3.339e-007 4.130e-008 8.085 4.441e-016
GARCH(1) 7.000e-001 2.396e-002 29.216 0.000e+000
ARCH(1) 4.000e-001 2.524e-002 15.850 0.000e+000
              Value Std.Error t value Pr(>|t|)
fraction 0.5 0.03244 15.41 0

-----

AIC(8) = -33692.98
BIC(8) = -33642.7

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
3542 0 0.95 9.237e-035

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
20.94 0.05133 12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
24.95 0.01505 12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12
2.836 0.653 1.458 -0.03251 2.052 -0.597 0.1343 -0.209 -0.6918 -0.4387
2.5 0.04711

C
-0.2868

TR^2 P-value F-stat P-value
23.36 0.02483 2.136 0.07642
```

```
> bezerro.fiegarch = fgarch(RetBezerro ~ -1 + arma(2, 0), ~ fiegarch(1,
  1), leverage = T)
> summary(boi.fiegarch)
```

```
Call:
fgarch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~
  fiegarch(1, 1), leverage =
  TRUE)
```

```
Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)
```

```
Conditional Variance Equation: structure(.Data = ~ fiegarch(1, 1)
, class = "formula"
)
```

 Estimated Coefficients:

	Value	Std.Error	t value	Pr(> t)
AR(1)	0.36780	0.01690	21.768	0.000e+000
AR(2)	0.21964	0.01681	13.065	0.000e+000
AR(3)	0.08726	0.01612	5.415	3.254e-008
AR(4)	0.04457	0.01393	3.200	6.922e-004
A	-0.38845	0.01674	-23.203	0.000e+000
GARCH(1)	0.38072	0.04364	8.725	0.000e+000
ARCH(1)	0.41493	0.01516	27.378	0.000e+000
	Value	Std.Error	t value	Pr(> t)
LEV(1)	-0.03175	0.007459	-4.256	0.00001063
fraction	0.53097	0.018226	29.132	0.00000000

 AIC(9) = -33750.43
 BIC(9) = -33693.86

Normality Test:

Jarque-Bera	P-value	Shapiro-Wilk	P-value
4737	0	0.9454	5.069e-036

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
27.55	0.006435	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
15.88	0.1966	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
0.1134	-0.3121	1.06	-0.328	1.778	0.2485	0.7817	0.237	-0.363	-0.4778		
2.914	0.04362	-0.716									

TR^2	P-value	F-stat	P-value
15.17	0.2322	1.385	0.2818

```

> bezerro.garch = garch(RetBezerro ~ -1 + arma(2, 0), ~ garch(1, 1), trace
= F)
> summary(boi.garch)

Call:
garch(formula.mean = RetBoi ~ -1 + arma(4, 0), formula.var = ~ garch(1,
1), trace = FALSE)

Mean Equation: structure(.Data = RetBoi ~ -1 + arma(4, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ garch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----
Estimated Coefficients:
-----
              Value Std.Error t value Pr(>|t|)
AR(1)  3.811e-001  1.799e-002  21.181 0.0000000
AR(2)  2.140e-001  1.818e-002  11.771 0.0000000
AR(3)  6.750e-002  1.825e-002   3.699 0.0002196
AR(4)  6.385e-002  1.679e-002   3.804 0.0001448
A      3.339e-007  2.519e-008  13.255 0.0000000
ARCH(1) 1.733e-001  6.403e-003  27.070 0.0000000
GARCH(1) 8.291e-001  5.127e-003 161.709 0.0000000

-----

AIC(7) = -33651.79
BIC(7) = -33607.79

Normality Test:
-----
      Jarque-Bera P-value Shapiro-Wilk      P-value
      4637          0          0.9452 4.499e-036

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
  22.34  0.03392          12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
  18.31  0.1065          12

Lagrange multiplier test:
-----
Lag 1  Lag 2  Lag 3  Lag 4  Lag 5  Lag 6  Lag 7  Lag 8  Lag 9  Lag 10
Lag 11 Lag 12
0.2941 -0.7664 -0.3186 -0.7435 0.2196 -1.187 -0.5375 -0.5826 -1.115 -1.187
  3.436 0.04396

C
0.152

TR^2 P-value F-stat P-value
18.06 0.114 1.649 0.175

```

```

> bezerro.egarch = garch(RetBezerro ~ -1 + arma(2, 0), ~ egarch(1, 1),
  leverage = T, trace = F
)
> summary(bezerro.egarch)

Call:
garch(formula.mean = RetBezerro ~ -1 + arma(2, 0), formula.var = ~
  egarch(1, 1), leverage =
  TRUE, trace = FALSE)

Mean Equation: structure(.Data = RetBezerro ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ egarch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian
-----
Estimated Coefficients:
-----
              Value Std.Error   t value   Pr(>|t|)
AR(1)  0.04012  0.017746    2.2608 2.384e-002
AR(2)  0.01574  0.018194    0.8649 3.872e-001
  A -0.45405  0.047997   -9.4599 0.000e+000
ARCH(1)  0.21412  0.016902   12.6689 0.000e+000
GARCH(1)  0.96781  0.004588  210.9599 0.000e+000
LEV(1) -0.25688  0.045335   -5.6662 1.584e-008
-----

AIC(6) = -20632.99
BIC(6) = -20596.33

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      173.4      0      0.9918 6.682e-013

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      18.79 0.09374      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      8.348 0.7573      12

Lagrange multiplier test:
-----
Lag 1  Lag 2  Lag 3  Lag 4  Lag 5  Lag 6  Lag 7  Lag 8  Lag 9  Lag 10
Lag 11  Lag 12
-1.339 0.7723 0.9409 -0.9466 0.6229 -0.5018 -0.9115 0.7413 0.9998 -0.04089
-0.5974 0.07818

C
-0.2591

TR^2 P-value F-stat P-value
7.973 0.7872 0.7266 0.8232

```

```

> bezerro.comp = compare.mgarch(bezerro.garch, bezerro.egarch,
  bezerro.figarch,
  bezerro.fiegarch)
> bezerro.comp
  bezerro.garch bezerro.egarch bezerro.figarch bezerro.fiegarch
    AIC          -20583          -20633          -20573          -20643
    BIC          -20553          -20596          -20536          -20601
Likelihood      10297          10322          10292          10329

```

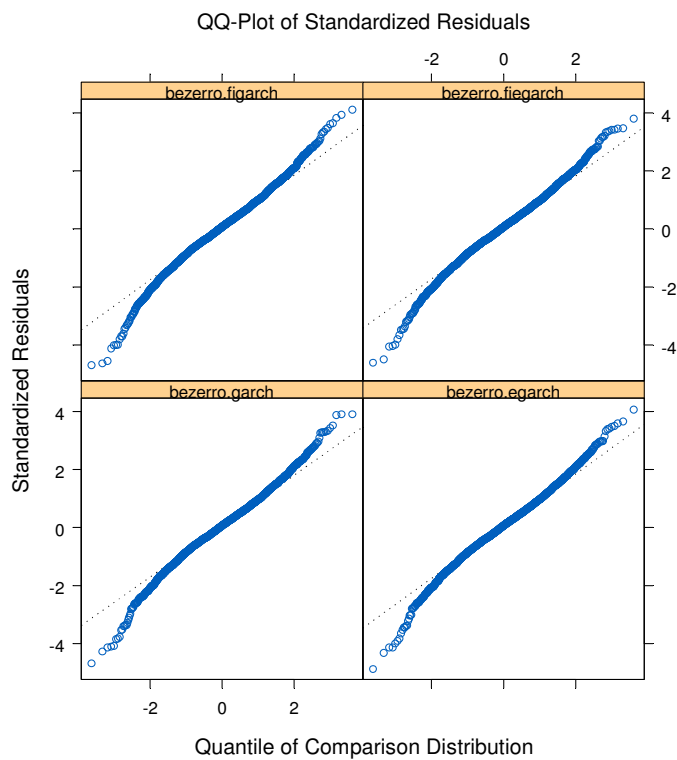


Figura 6.20.2: resultados do ajuste dos modelos da família GARCH para a modelagem da variância em conjunto com o componente de *drift* (dado pelos log-retornos) do bezerro.

7.20.3 Café

```
> cafe.figarch = fgarch(RetCafe ~ -1 + arma(2, 0), ~ figarch(1, 1))
Initializing model parameters.
Iteration No. 1: log-likelihood=10822.428234
Routine aborted. Try rescaling the series or a different initialization.
> summary(cafe.figarch)

Call:
fgarch(formula.mean = RetCafe ~ -1 + arma(2, 0), formula.var = ~
      figarch(1, 1))

Mean Equation: structure(.Data = RetCafe ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ figarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value Std.Error t value Pr(>|t|)
AR(1) -0.04262300 1.921e-002 -2.218 1.329e-002
AR(2) -0.03675984 1.663e-002 -2.210 1.358e-002
      A 0.00001171 1.976e-006  5.924 1.699e-009
GARCH(1) 0.70000000 3.299e-002 21.218 0.000e+000
ARCH(1) 0.40000000 3.322e-002 12.039 0.000e+000
fraction 0.50000000 4.344e-002 11.509 0.000e+000

-----

AIC(6) = -21632.86
BIC(6) = -21594.82

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      1289      0      0.9719 6.506e-028

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      23.89 0.02102      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      14.27 0.2835      12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12
-1.105 -0.992 1.039 1.581 -1.842 -0.5389 0.3469 0.4755 -1.576 -0.5327
 0.1912 0.04546

      C
-0.1416

TR^2 P-value F-stat P-value
13.44 0.3382 1.225 0.376
```

```

> cafe.fiegarch = fgarch(RetCafe ~ -1 + arma(2, 0), ~ fiegarch(1, 1),
  leverage = T)
> summary(cafe.fiegarch)

Call:
fgarch(formula.mean = RetCafe ~ -1 + arma(2, 0), formula.var = ~
  fiegarch(1, 1), leverage =
  TRUE)

Mean Equation: structure(.Data = RetCafe ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ fiegarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value Std.Error t value   Pr(>|t|)
AR(1) -0.04418   0.016677  -2.649 4.050e-003
AR(2) -0.04204   0.016548  -2.540 5.554e-003
A    -0.22825   0.018604 -12.268 0.000e+000
GARCH(1) 0.42230  0.063222   6.680 1.355e-011
ARCH(1) 0.22664   0.011994  18.896 0.000e+000
LEV(1) 0.09679   0.007392  13.094 0.000e+000
fraction 0.53427  0.033922  15.750 0.000e+000

-----

AIC(7) = -21694.18
BIC(7) = -21649.8

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      1252         0       0.9741 6.168e-027

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      22.48  0.0325         12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      21.66 0.04157         12

Lagrange multiplier test:
-----
Lag 1  Lag 2 Lag 3 Lag 4  Lag 5  Lag 6 Lag 7 Lag 8 Lag 9  Lag 10 Lag
11  Lag 12  C
-1.292 -0.5829 1.168 1.852 -1.017 -0.03708 1.26 1.366 -1.41 -0.2099
1.315 0.03274 2.15

TR^2 P-value F-stat P-value
20.6 0.05657 1.882 0.1166

```

```

> cafe.garch = garch(RetCafe ~ -1 + arma(2, 0), ~ garch(1, 1), trace = F)
> summary(cafe.garch)

Call:
garch(formula.mean = RetCafe ~ -1 + arma(2, 0), formula.var = ~ garch(1,
1), trace = FALSE)

Mean Equation: structure(.Data = RetCafe ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ garch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----

Estimated Coefficients:
-----
              Value  Std.Error  t value Pr(>|t|)
AR(1) -0.04262300 1.671e-002   -2.550  0.01080
AR(2) -0.03675984 1.713e-002   -2.145  0.03197
A      0.00001171 1.413e-006    8.287  0.00000
ARCH(1) 0.10322844 4.279e-003   24.124  0.00000
GARCH(1) 0.87115865 6.986e-003  124.704  0.00000

-----

AIC(5) = -21584.37
BIC(5) = -21552.67

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      2040          0      0.9685 2.278e-029

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      24.44 0.01771          12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      23.85 0.02132          12

Lagrange multiplier test:
-----
Lag 1  Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12 C
-1.734 -0.9431 1.025 1.614 -1.858 -0.3051 0.3151 0.08674 -1.378 -0.513
0.181 0.04099 3.056

TR^2 P-value F-stat P-value
23.22 0.02594 2.122 0.07813

```

```
> cafe.egarch = garch(RetCafe ~ -1 + arma(2, 0), ~ egarch(1, 1), leverage
= T, trace = F)
> summary(cafe.egarch)
```

```
Call:
garch(formula.mean = RetCafe ~ -1 + arma(2, 0), formula.var = ~ egarch(1,
1), leverage =
TRUE, trace = FALSE)
```

```
Mean Equation: structure(.Data = RetCafe ~ -1 + arma(2, 0)
, class = "formula"
)
```

```
Conditional Variance Equation: structure(.Data = ~ egarch(1, 1)
, class = "formula"
)
```

```
Conditional Distribution: gaussian
```

```
-----
Estimated Coefficients:
```

```
-----
              Value Std.Error   t value Pr(>|t|)
AR(1) -0.04588   0.015952   -2.876 0.004046
AR(2) -0.03950   0.016217   -2.436 0.014893
  A -0.56531   0.041371  -13.664 0.000000
ARCH(1)  0.21409   0.009734   21.994 0.000000
GARCH(1)  0.94910   0.004713  201.397 0.000000
LEV(1)  0.32636   0.035474    9.200 0.000000
-----
```

```
AIC(6) = -21648.08
BIC(6) = -21610.05
```

```
Normality Test:
```

```
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      1444         0         0.9723 9.344e-028
-----
```

```
Ljung-Box test for standardized residuals:
```

```
-----
Statistic P-value Chi^2-d.f.
      23.89 0.02108         12
-----
```

```
Ljung-Box test for squared standardized residuals:
```

```
-----
Statistic P-value Chi^2-d.f.
      25.3 0.01345         12
-----
```

```
Lagrange multiplier test:
```

```
-----
Lag 1  Lag 2 Lag 3 Lag 4  Lag 5    Lag 6 Lag 7  Lag 8  Lag 9  Lag 10 Lag
11 Lag 12    C
-1.399 -0.5311 1.523 1.807 -1.733 -0.06523 0.532 0.6741 -1.326 -0.3099
0.9994 0.0366 3.051
-----
```

```
TR^2 P-value F-stat P-value
24.43 0.01776 2.234 0.0653
```

```

> cafe.comp = compare.mgarch(cafe.garch, cafe.egarch, cafe.figarch,
  cafe.fiegarch)
> cafe.comp
      cafe.garch  cafe.egarch  cafe.figarch  cafe.fiegarch
AIC      -21584      -21648      -21633      -21694
BIC      -21553      -21610      -21595      -21650
Likelihood 10797      10830      10822      10854

```

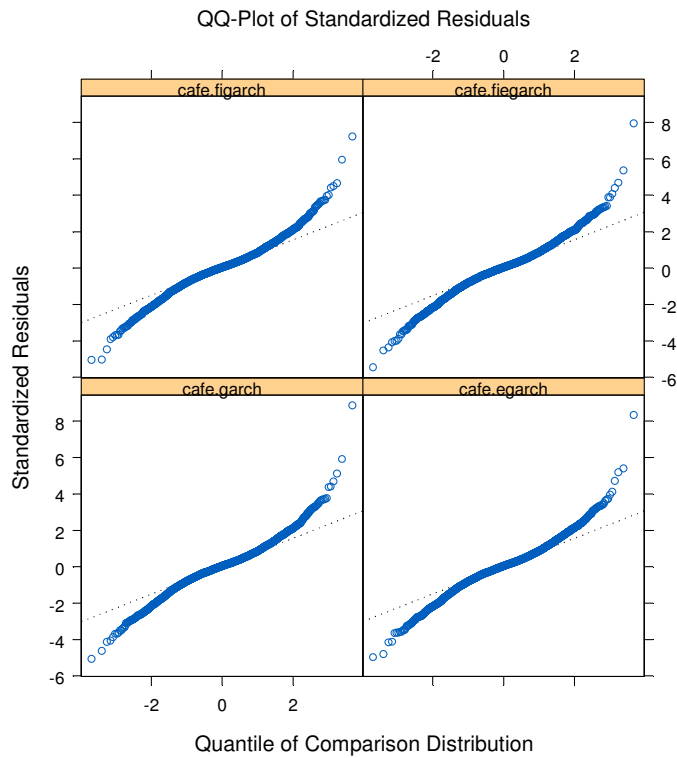


Figura 6.20.3: resultados do ajuste dos modelos da família GARCH para a modelagem da variância em conjunto com o componente de *drift* (dado pelos log-retornos) do café.

7.20.4 Açúcar

```
> acucar.figarch = fgarch(RetAcucar ~ -1 + arma(1, 1), ~ figarch(1, 1))
Initializing model parameters.
Iteration No. 1: log-likelihood=7772.101712
Routine aborted. Try rescaling the series or a different initialization.
> summary(acucar.figarch)

Call:
fgarch(formula.mean = RetAcucar ~ -1 + arma(1, 1), formula.var = ~
      figarch(1, 1))

Mean Equation: structure(.Data = RetAcucar ~ -1 + arma(1, 1)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ figarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value Std.Error t value  Pr(>|t|)
AR(1)  9.211e-001 1.658e-002  55.552 0.000e+000
MA(1) -7.732e-001 2.708e-002 -28.553 0.000e+000
  A  4.917e-006 1.402e-006   3.506 2.317e-004
GARCH(1) 7.000e-001 6.661e-002  10.508 0.000e+000
ARCH(1) 4.000e-001 5.976e-002   6.694 1.337e-011
fraction 5.000e-001 9.016e-002   5.546 1.617e-008

-----

AIC(6) = -15532.2
BIC(6) = -15497.24

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      190.1      0      0.991 2.034e-011

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      10.18  0.6002      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      16  0.1914      12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12 C
0.661 2.016 -1.183 -1.489 0.6123 0.7236 -0.6674 -0.2376 1.642 1.721 -
1.052 -0.2165 -1.23

TR^2 P-value F-stat P-value
17.41 0.135 1.593 0.1934
```

```
> acucar.fiegarch = fgarch(RetAcucar ~ -1 + arma(1, 1), ~ fiegarch(1, 1),
  leverage = T)
> summary(acucar.fiegarch)
```

```
Call:
fgarch(formula.mean = RetAcucar ~ -1 + arma(1, 1), formula.var = ~
  fiegarch(1, 1), leverage
  = TRUE)
```

```
Mean Equation: structure(.Data = RetAcucar ~ -1 + arma(1, 1)
, class = "formula"
)
```

```
Conditional Variance Equation: structure(.Data = ~ fiegarch(1, 1)
, class = "formula"
)
```

```
-----
Estimated Coefficients:
```

```
-----
              Value Std.Error t value   Pr(>|t|)
AR(1)    0.92763    0.01385  66.970 0.000e+000
MA(1)   -0.77775    0.02373 -32.770 0.000e+000
A      -0.22479    0.03160  -7.113 7.407e-013
GARCH(1) 0.29501    0.13347   2.210 1.359e-002
ARCH(1)  0.25072    0.03322   7.548 3.086e-014
LEV(1)  -0.07066    0.01452  -4.866 6.041e-007
fraction 0.64315    0.05148  12.494 0.000e+000
-----
```

```
AIC(7) = -15552.49
```

```
BIC(7) = -15511.7
```

```
Normality Test:
```

```
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      163.5         0         0.9917 8.699e-011
-----
```

```
Ljung-Box test for standardized residuals:
```

```
-----
Statistic P-value Chi^2-d.f.
      10.09  0.6077         12
-----
```

```
Ljung-Box test for squared standardized residuals:
```

```
-----
Statistic P-value Chi^2-d.f.
      13.97  0.3027         12
-----
```

```
Lagrange multiplier test:
```

```
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12 C
0.2656 1.294 -1.374 -1.018 1.307 0.9578 -0.2554 0.1182 2.024 2.082 -
0.6594 -0.2158 0.4918
-----
```

```
TR^2 P-value F-stat P-value
16.1 0.1867 1.473 0.2401
```

```

> acucar.garch = garch(RetAcucar ~ -1 + arma(1, 1), ~ garch(1, 1), trace =
  F)
> summary(acucar.garch)

Call:
garch(formula.mean = RetAcucar ~ -1 + arma(1, 1), formula.var = ~ garch(1,
  1), trace = FALSE
)

Mean Equation: structure(.Data = RetAcucar ~ -1 + arma(1, 1)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ garch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----

Estimated Coefficients:
-----
              Value Std.Error t value Pr(>|t|)
AR(1)  9.211e-001 1.550e-002  59.412 0.000e+000
MA(1) -7.732e-001 2.500e-002 -30.933 0.000e+000
  A  4.917e-006 8.553e-007   5.748 1.011e-008
ARCH(1) 9.554e-002 1.035e-002   9.230 0.000e+000
GARCH(1) 8.675e-001 1.389e-002  62.454 0.000e+000

-----

AIC(5) = -15545.5
BIC(5) = -15516.36

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      171.6      0      0.9915 5.632e-011

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      10.78  0.5476      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      11.44  0.4918      12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12 C
0.2069 1.343 -1.389 -1.626 0.5516 0.4162 -0.7899 -0.3309 1.059 1.462 -
1.22 -0.1956 0.941

TR^2 P-value F-stat P-value
13.04 0.3658 1.192 0.3991

```

```

> acucar.egarch = garch(RetAcucar ~ -1 + arma(1, 1), ~ egarch(1, 1),
  leverage = T, trace = F)
> summary(acucar.egarch)

Call:
garch(formula.mean = RetAcucar ~ -1 + arma(1, 1), formula.var = ~
  egarch(1, 1), leverage =
  TRUE, trace = FALSE)

Mean Equation: structure(.Data = RetAcucar ~ -1 + arma(1, 1)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ egarch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian
-----
Estimated Coefficients:
-----
              Value Std.Error   t value   Pr(>|t|)
AR(1)    0.9269   0.014070   65.873 0.000e+000
MA(1)   -0.7798   0.023911  -32.614 0.000e+000
A      -0.6624   0.088235   -7.508 8.327e-014
ARCH(1)  0.2256   0.022508   10.021 0.000e+000
GARCH(1) 0.9459   0.008468  111.706 0.000e+000
LEV(1)  -0.2259   0.049083   -4.603 4.369e-006
-----

AIC(6) = -15547.14
BIC(6) = -15512.18

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      179.9         0         0.9914 4.322e-011

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      10.36  0.5841         12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      11.02  0.5275         12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12
0.9105 1.282 -1.452 -1.672 0.9959 0.5014 -0.1717 -0.2656 1.048 1.688 -
0.6756 -0.2007

C
0.7137

TR^2 P-value F-stat P-value
13.06 0.3648 1.193 0.3981

```

```

> acucar.comp = compare.mgarch(acucar.garch, acucar.egarch, acucar.figarch,
  acucar.fiegarch)
> acucar.comp
  acucar.garch acucar.egarch acucar.figarch acucar.fiegarch
    AIC        -15545         -15547         -15532         -15552
    BIC        -15516         -15512         -15497         -15512
Likelihood      7778           7780           7772           7783

```

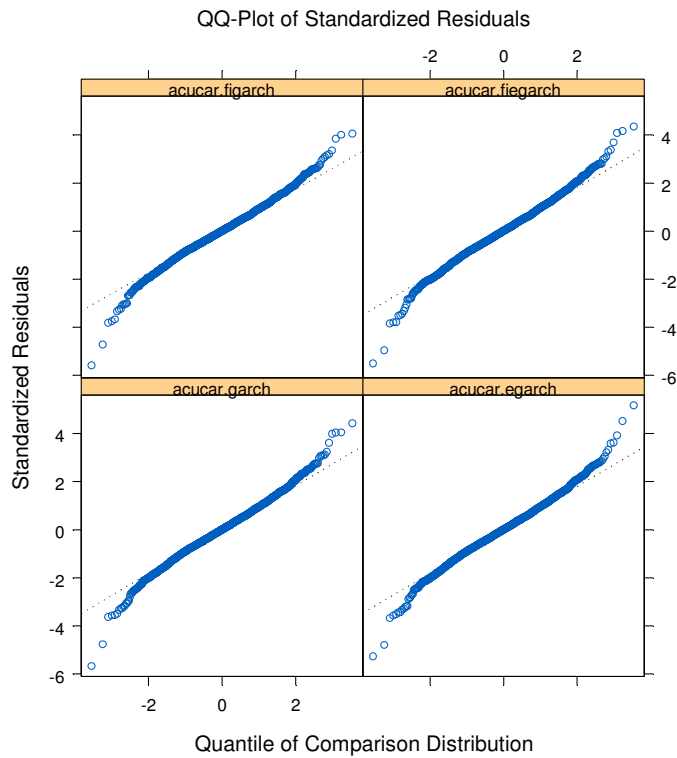


Figura 6.20.4: resultados do ajuste dos modelos da família GARCH para a modelagem da variância em conjunto com o componente de *drift* (dado pelos log-retornos) do açúcar.

7.20.5 Milho

```
> milho.figarch = fgarch(RetMilho ~ -1 + arma(3, 0), ~ figarch(1, 1))
Routine aborted. Try rescaling the series or a different initialization.

> summary(milho.figarch)

Call:
fgarch(formula.mean = RetMilho ~ -1 + arma(3, 0), formula.var = ~
  figarch(1, 1))

Mean Equation: structure(.Data = RetMilho ~ -1 + arma(3, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ figarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value  Std.Error t value  Pr(>|t|)
AR(1)  1.215e-001  2.597e-002   4.678 1.535e-006
AR(2)  9.591e-002  2.336e-002   4.106 2.090e-005
AR(3)  1.048e-001  2.340e-002   4.479 3.932e-006
  A    5.956e-006  1.435e-006   4.151 1.719e-005
GARCH(1) 7.000e-001  5.959e-002  11.746 0.000e+000
ARCH(1)  4.000e-001  5.733e-002   6.977 1.982e-012
fraction 5.000e-001  8.149e-002   6.135 5.016e-010

-----

AIC(7) = -13528.09
BIC(7) = -13488.16

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk    P-value
      177.9         0         0.9897 1.783e-011

Ljung-Box test for standardized residuals:
-----
Statistic    P-value Chi^2-d.f.
      46.89 4.862e-006         12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      18.53 0.1005         12

Lagrange multiplier test:
-----
Lag 1  Lag 2  Lag 3  Lag 4  Lag 5  Lag 6  Lag 7  Lag 8  Lag 9  Lag 10  Lag 11
Lag 12      C
2.479 0.9078 0.594 -1.022 0.7164 1.17 1.981 -1.136 -1.23 0.5614 1.03
0.05645 -1.387

TR^2 P-value F-stat P-value
19.18 0.08421 1.759 0.1443
```

```

> milho.fiegarch = fgarch(RetMilho ~ -1 + arma(3, 0), ~ fiegarch(1, 1),
  leverage = T)

> summary(milho.fiegarch)

Call:
fgarch(formula.mean = RetMilho ~ -1 + arma(3, 0), formula.var = ~
  fiegarch(1, 1), leverage
  = TRUE)

Mean Equation: structure(.Data = RetMilho ~ -1 + arma(3, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ fiegarch(1, 1)
, class = "formula"
)

-----
Estimated Coefficients:
-----
              Value Std.Error t value   Pr(>|t|)
AR(1)  0.11958    0.02255   5.303 6.267e-008
AR(2)  0.10056    0.02180   4.614 2.092e-006
AR(3)  0.10935    0.02134   5.124 1.623e-007
  A -0.23451    0.03012  -7.787 5.218e-015
GARCH(1) 0.27483    0.13553   2.028 2.135e-002
ARCH(1)  0.27604    0.03485   7.920 1.887e-015
LEV(1) -0.05798    0.01497  -3.872 5.551e-005
              Value Std.Error t value Pr(>|t|)
fraction 0.6697    0.04953  13.52      0

-----

AIC(8) = -13548.26
BIC(8) = -13502.62

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
  161.3      0      0.9899 2.255e-011

Ljung-Box test for standardized residuals:
-----
Statistic   P-value Chi^2-d.f.
  47.12 4.451e-006      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
  12.57 0.4008      12

Lagrange multiplier test:
-----
Lag 1  Lag 2  Lag 3   Lag 4 Lag 5 Lag 6 Lag 7  Lag 8  Lag 9 Lag 10 Lag 11
Lag 12      C
1.224 0.5308 0.2142 -0.9874 1.136 1.411 1.832 -1.245 -1.026 0.5584 1.082
0.04292 0.2171

TR^2 P-value F-stat P-value
  13 0.3693 1.189 0.4016

```

```

> milho.garch = garch(RetMilho ~ -1 + arma(3, 0), ~ garch(1, 1), trace =
  F)
> summary(milho.garch)

Call:
garch(formula.mean = RetMilho ~ -1 + arma(3, 0), formula.var = ~ garch(1,
  1), trace = FALSE)

Mean Equation: structure(.Data = RetMilho ~ -1 + arma(3, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ garch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----

Estimated Coefficients:
-----
              Value Std.Error t value Pr(>|t|)
AR(1)  1.215e-001  2.284e-002   5.321 1.136e-007
AR(2)  9.591e-002  2.220e-002   4.319 1.633e-005
AR(3)  1.048e-001  2.241e-002   4.678 3.079e-006
  A    5.956e-006  1.084e-006   5.496 4.334e-008
ARCH(1) 1.241e-001  1.081e-002  11.477 0.000e+000
GARCH(1) 8.396e-001  1.409e-002  59.595 0.000e+000

-----

AIC(6) = -13553.84
BIC(6) = -13519.61

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      149.7      0      0.9905 7.138e-011

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      46.59 5.496e-006      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      10.42 0.5793      12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12
1.051 0.3564 -0.4326 -1.608 0.4487 0.8007 1.363 -1.615 -1.336 -0.1151
0.9324 0.06057

C
-0.1507

TR^2 P-value F-stat P-value
11.14 0.5169 1.018 0.5399

```

```

> milho.egarch = garch(RetMilho ~ -1 + arma(3, 0), ~ egarch(1, 1),
  leverage = T, trace = F)
> summary(milho.egarch)

Call:
garch(formula.mean = RetMilho ~ -1 + arma(3, 0), formula.var = ~ egarch(1,
  1), leverage =
  TRUE, trace = FALSE)

Mean Equation: structure(.Data = RetMilho ~ -1 + arma(3, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ egarch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----

Estimated Coefficients:
-----
              Value Std.Error  t value  Pr(>|t|)
AR(1)  0.12546  0.022530   5.569 2.878e-008
AR(2)  0.09466  0.020805   4.550 5.657e-006
AR(3)  0.10645  0.021733   4.898 1.038e-006
  A -0.70611  0.087339  -8.085 8.882e-016
ARCH(1) 0.26490  0.018680  14.181 0.000e+000
GARCH(1) 0.94361  0.008971 105.179 0.000e+000
  LEV(1) -0.16760  0.041974  -3.993 6.737e-005

-----

AIC(7) = -13545.95
BIC(7) = -13506.02

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk    P-value
      198.4         0         0.9886 2.851e-012

Ljung-Box test for standardized residuals:
-----
Statistic  P-value Chi^2-d.f.
      44.37 0.0000132         12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      13.48 0.3348         12

Lagrange multiplier test:
-----
Lag 1 Lag 2  Lag 3  Lag 4  Lag 5  Lag 6 Lag 7  Lag 8  Lag 9 Lag 10 Lag 11
Lag 12  C
1.585 1.002 -0.2153 -1.375 0.6131 0.9058 1.903 -1.498 -1.272 0.2592 0.9203
0.04647 0.8403

TR^2 P-value F-stat P-value
14.52 0.2686 1.329 0.3118

```

```

> milho.comp = compare.mgarch(milho.garch, milho.egarch, milho.figarch,
  milho.fiegarch)
> milho.comp
      milho.garch milho.egarch milho.figarch milho.fiegarch
AIC      -13554      -13546      -13528      -13548
BIC      -13520      -13506      -13488      -13503
Likelihood 6783      6780      6771      6782

```

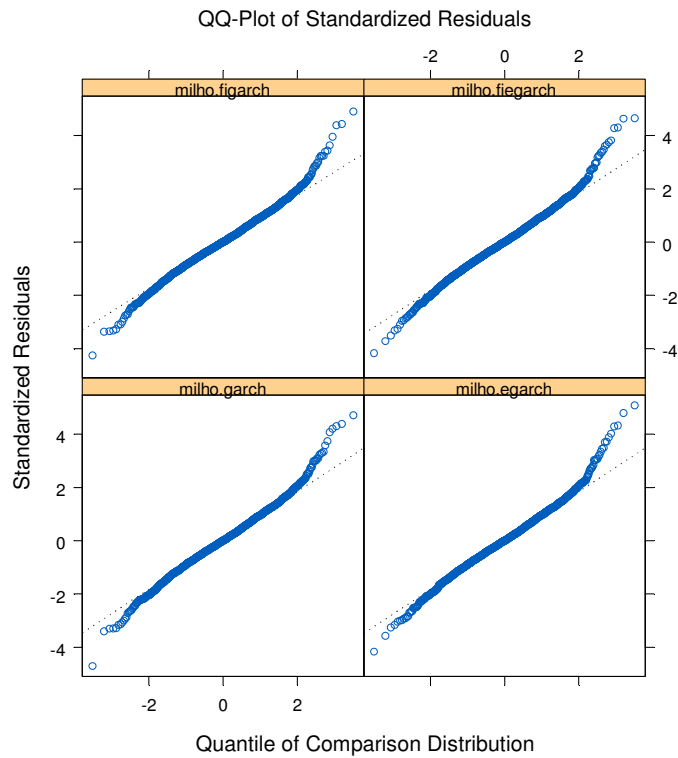


Figura 6.20.5: resultados do ajuste dos modelos da família GARCH para a modelagem da variância em conjunto com o componente de *drift* (dado pelos log-retornos) do milho.

7.20.6 Soja (região do Paraná)

```
> soja.figarch = fgarch(RetSojaPR ~ -1 + arma(2, 0), ~ figarch(1, 1))
> summary(soja.figarch)

Call:
fgarch(formula.mean = RetSojaPR ~ -1 + arma(2, 0), formula.var = ~
      figarch(1, 1))

Mean Equation: structure(.Data = RetSojaPR ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ figarch(1, 1)
, class = "formula"
)

-----

Estimated Coefficients:
-----
              Value  Std.Error  t value  Pr(>|t|)
AR(1)  1.145e-001  1.870e-002   6.124  5.009e-010
AR(2)  3.869e-002  1.591e-002   2.432  7.521e-003
A      3.843e-006  4.431e-007   8.673  0.000e+000
GARCH(1) 7.000e-001  3.624e-002  19.313  0.000e+000
ARCH(1)  4.000e-001  2.882e-002  13.881  0.000e+000
fraction 5.000e-001  5.519e-002   9.060  0.000e+000
-----

AIC(6) = -23571.92
BIC(6) = -23534.21

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk    P-value
      3658          0          0.9733  1.358e-026

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      26.39 0.009463          12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      24.16 0.01933          12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10 Lag
11 Lag 12
  4.11 1.471 0.9007 -0.244 0.3522 0.1793 -0.3662 -0.7775 -0.7787 0.3036
  0.4856 0.04755

C
-1.513

TR^2 P-value F-stat P-value
23.97 0.02051 2.193 0.06977
```

```
> soja.fiegarch = fgarch(RetSojaPR ~ -1 + arma(2, 0), ~ fiegarch(1, 1),
  leverage = T)
> summary(soja.fiegarch)
```

```
Call:
fgarch(formula.mean = RetSojaPR ~ -1 + arma(2, 0), formula.var = ~
  fiegarch(1, 1), leverage
  = TRUE)
```

```
Mean Equation: structure(.Data = RetSojaPR ~ -1 + arma(2, 0)
, class = "formula"
)
```

```
Conditional Variance Equation: structure(.Data = ~ fiegarch(1, 1)
, class = "formula"
)
```

Estimated Coefficients:

	Value	Std.Error	t value	Pr(> t)
AR(1)	0.11526	0.01642	7.021	1.295e-012
AR(2)	0.02963	0.01621	1.828	3.384e-002
A	-0.22788	0.02113	-10.783	0.000e+000
GARCH(1)	0.72373	0.04736	15.281	0.000e+000
ARCH(1)	0.20792	0.01584	13.130	0.000e+000
LEV(1)	-0.04909	0.00596	-8.236	1.110e-016
fraction	0.44058	0.03853	11.435	0.000e+000

AIC(7) = -23607.5
BIC(7) = -23563.51

Normality Test:

Jarque-Bera	P-value	Shapiro-Wilk	P-value
2863	0	0.9759	2.22e-025

Ljung-Box test for standardized residuals:

Statistic	P-value	Chi^2-d.f.
26	0.01072	12

Ljung-Box test for squared standardized residuals:

Statistic	P-value	Chi^2-d.f.
7.758	0.8038	12

Lagrange multiplier test:

Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11	Lag 12
1.208	0.4242	0.2848	-0.1411	0.4947	0.4102	0.05277	-0.6764	-0.7052	1.37	1.746	0.04761

C
-0.092

TR^2	P-value	F-stat	P-value
7.858	0.7961	0.7158	0.8332

```

> soja.garch = garch(RetSojaPR ~ -1 + arma(2, 0), ~ garch(1, 1), trace =
  F)
> summary(soja.garch)

Call:
garch(formula.mean = RetSojaPR ~ -1 + arma(2, 0), formula.var = ~ garch(1,
  1), trace = FALSE
  )

Mean Equation: structure(.Data = RetSojaPR ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ garch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian

-----

Estimated Coefficients:
-----
              Value Std.Error t value   Pr(>|t|)
AR(1)  1.145e-001  1.722e-002   6.651 3.305e-011
AR(2)  3.869e-002  1.677e-002   2.307 2.108e-002
  A    3.843e-006  2.403e-007  15.991 0.000e+000
ARCH(1) 1.018e-001  7.106e-003  14.322 0.000e+000
GARCH(1) 8.804e-001  6.996e-003 125.843 0.000e+000

-----

AIC(5) = -23600.09
BIC(5) = -23568.67

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk   P-value
      2945         0         0.9752 1.065e-025

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      24.51 0.01735         12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      10.74 0.5513         12

Lagrange multiplier test:
-----
Lag 1  Lag 2  Lag 3  Lag 4  Lag 5  Lag 6  Lag 7  Lag 8  Lag 9  Lag 10
Lag 11  Lag 12
2.511 0.9268 0.7229 -0.3702 0.1811 0.1902 -0.6514 -1.034 -0.9482 -0.05427
0.3811 0.05134

C
-0.3517

TR^2 P-value F-stat P-value
10.59 0.5643 0.9654 0.5889

```

```

> soja.egarch = garch(RetSojaPR ~ -1 + arma(2, 0), ~ egarch(1, 1),
  leverage = T, trace = F)
> summary(soja.egarch)

Call:
garch(formula.mean = RetSojaPR ~ -1 + arma(2, 0), formula.var = ~
  egarch(1, 1), leverage =
  TRUE, trace = FALSE)

Mean Equation: structure(.Data = RetSojaPR ~ -1 + arma(2, 0)
, class = "formula"
)

Conditional Variance Equation: structure(.Data = ~ egarch(1, 1)
, class = "formula"
)

Conditional Distribution: gaussian
-----
Estimated Coefficients:
-----
              Value Std.Error  t value  Pr(>|t|)
AR(1)  0.12159  0.016486   7.375 1.990e-013
AR(2)  0.03331  0.015640   2.130 3.323e-002
A    -0.56076  0.038899  -14.416 0.000e+000
ARCH(1) 0.24696  0.013807  17.887 0.000e+000
GARCH(1) 0.95735  0.003629  263.790 0.000e+000
LEV(1) -0.17402  0.026652   -6.529 7.451e-011
-----

AIC(6) = -23612.72
BIC(6) = -23575.01

Normality Test:
-----
Jarque-Bera P-value Shapiro-Wilk P-value
      2068      0      0.9775 1.369e-024

Ljung-Box test for standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      23.19 0.02614      12

Ljung-Box test for squared standardized residuals:
-----
Statistic P-value Chi^2-d.f.
      8.135 0.7745      12

Lagrange multiplier test:
-----
Lag 1 Lag 2 Lag 3 Lag 4 Lag 5 Lag 6 Lag 7 Lag 8 Lag 9 Lag 10
Lag 11 Lag 12
1.886 0.9098 0.4343 -0.1186 0.3172 0.7236 -0.4663 -1.163 -0.8952 0.4635
0.7712 0.05355

C
-0.547

TR^2 P-value F-stat P-value
8.363 0.7562 0.7619 0.7898

```

```

> soja.comp = compare.mgarch(soja.garch, soja.egarch, soja.figarch,
  soja.fiegarch)
> soja.comp
      soja.garch soja.egarch soja.figarch soja.fiegarch
AIC      -23600      -23613      -23572      -23608
BIC      -23569      -23575      -23534      -23564
Likelihood 11805      11812      11792      11811

```

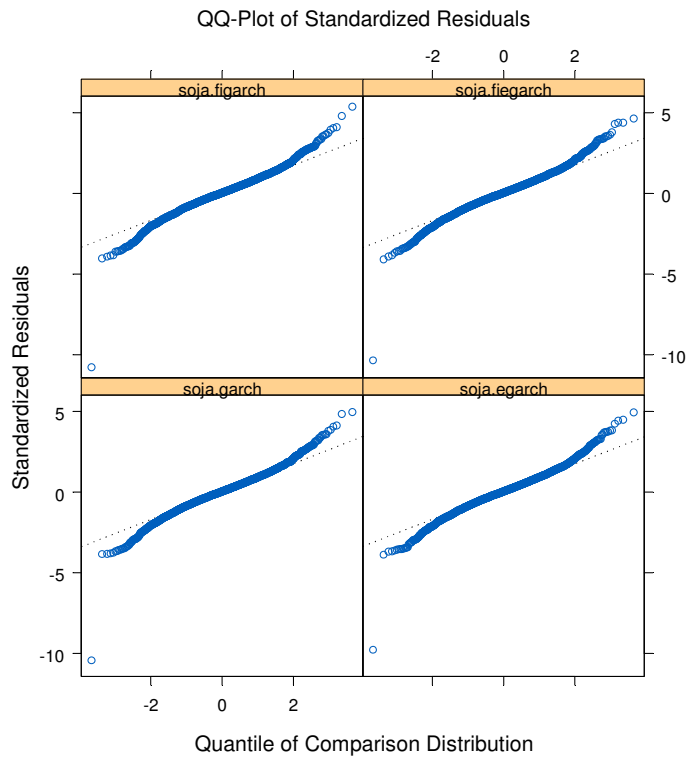


Figura 6.20.6: resultados do ajuste dos modelos da família GARCH para a modelagem da variância em conjunto com o componente de *drift* (dado pelos log-retornos) da soja (região do Paraná).