

WAGNER MOURA LAMOUNIER

**COMPORTAMENTO DOS PREÇOS
NO MERCADO "SPOT" DE CAFÉ DO BRASIL:
ANÁLISE NOS DOMÍNIOS DO TEMPO E DA FREQUÊNCIA**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia Aplicada, para obtenção do título de “Doctor Scientiae”.

VIÇOSA
MINAS GERAIS - BRASIL
2001

“Life is a random walk.”

À minha mãe Odília Marta.

AGRADECIMENTO

A Deus, pela vida.

A São Jorge, pela inspiração.

À Universidade Federal de Viçosa (UFV) e ao Departamento de Economia Rural (DER), pelas oportunidades e pelo acolhimento.

Ao CNPq e à CAPES, pela nobre função que desempenham e pelo financiamento desta pesquisa ao longo do meu Programa de Pós-Graduação.

Ao professor Carlos Antônio Moreira Leite, pela orientação e pela liberdade de ação propiciada ao longo do Programa.

Aos professores João Eustáquio de Lima (DER/UFV) e João Victor Issler (EPGE/FGV-RJ), por terem me apresentado o fascinante “mundo” da *Análise de Séries Temporais*.

Ao professor conselheiro Danilo Rolim Dias de Aguiar e aos demais professores do Departamento de Economia Rural, pelos conhecimentos compartilhados.

Aos amigos Jersone Tasso, Aureliano Bressan, Guilherme Malafaia e aos demais colegas do Programa de Pós-Graduação.

Ao irmão Osvaldo S. Ramos (*in memoriam*) e aos demais amigos de Viçosa.

À Graça e aos demais colaboradores do Departamento de Economia Rural, pela disponibilidade e pelo auxílio.

A todos os colegas e colaboradores da UNA.

Ao Dr. Honório Tomelin, pelas oportunidades e pela confiança.

Aos meus alunos, pelo apoio e interesse.

BIOGRAFIA

WAGNER MOURA LAMOUNIER, filho de Odília Marta Moura Lamounier e Walter Lamounier dos Passos, nasceu em Belo Horizonte, Minas Gerais, em 30 de abril de 1970.

Em dezembro de 1994, bacharelou-se em Ciências Econômicas pela Universidade Federal de Uberlândia (UFU), em Uberlândia, MG.

Em janeiro de 1996, ingressou no Curso de Mestrado em Economia Rural da Universidade Federal de Viçosa (UFV), em Viçosa, MG.

Tendo obtido o conceito “A” em todas as matérias cursadas no Programa de Mestrado em Economia Rural, o que foi considerado um desempenho excelente, foi convidado a ingressar no Programa de Doutorado em Economia Rural da Universidade Federal de Viçosa no 1º semestre de 1997.

De 13 de setembro a 4 de novembro de 1999, realizou e foi aprovado no Exame de Qualificação do Programa de Doutorado em Economia Rural.

Em 27 de julho de 2001, submeteu-se e foi aprovado na defesa de tese de doutorado.

ÍNDICE

	Página
RESUMO	x
ABSTRACT	xiii
1. INTRODUÇÃO	1
1.1. O problema e sua importância	5
1.2. Hipótese	12
1.3. Objetivos	12
1.3.1. Objetivo geral	12
1.3.2. Objetivos específicos	12
2. METODOLOGIA	14
2.1. Referencial teórico	14
2.1.1. Movimentos de longo prazo nos preços das <i>commodities</i> agropecuárias	20
2.1.1.1. Tendências	20

	Página
2.1.1.2. Ciclos	23
2.1.2. Movimentos de curto prazo nos preços das <i>commodities</i> agropecuárias	27
2.1.2.1. Sazonalidade	27
2.1.2.2. Volatilidade	29
2.2. Modelo analítico e procedimentos	33
2.2.1. Análise do componente de tendência	33
2.2.2. Análise do componente cíclico	45
2.2.3. Análise do componente sazonal	63
2.2.4. Análise do componente de volatilidade	71
2.3. Evidência empírica e literatura correlacionada	79
2.4. Dados e <i>softwares</i> computacionais utilizados	81
3. RESULTADOS E DISCUSSÃO	82
3.1. Análise do componente de tendência	82
3.1.1. Análise gráfica	83
3.1.2. Análise da função de autocorrelação amostral	86
3.1.3. Testes de raízes unitárias	89
3.1.3.1. Testes de Dickey-Fuller e de Dickey-Fuller expandido	89
3.1.3.2. Testes de Phillips-Perron	95
3.1.4. Modelagem da tendência determinística nos preços do café ...	98
3.1.5. Modelagem da tendência estocástica nos preços do café	102
3.2. Análise do componente de ciclos	105

	Página
3.2.1. Análise do período de janeiro de 1946 a dezembro de 2000 ...	107
3.2.2. Análise espectral dos preços do café nas décadas de 50, 60, 70, 80 e 90	117
3.3. Análise do componente sazonal	124
3.3.1. Análise gráfica	125
3.3.2. Testes para sazonalidade determinística nos preços	134
3.3.3. Análise do componente sazonal pela função de autocorrelação	140
3.4. Análise do componente de volatilidade	143
3.4.1. Especificação de um modelo para a série de preços do café ...	145
3.4.2. Especificação de um modelo para a variância condicional dos preços do café	154
4. RESUMO E CONCLUSÕES	166
REFERÊNCIAS BIBLIOGRÁFICAS	174
APÊNDICES	180
APÊNDICE A	181
APÊNDICE B	189
APÊNDICE C	190
APÊNDICE D	191
APÊNDICE E	194
APÊNDICE F	199
APÊNDICE G	200
APÊNDICE H	203

RESUMO

LAMOUNIER, Wagner Moura, D.S., Universidade Federal de Viçosa, julho de 2001. **Comportamento dos preços no mercado "spot" de café do Brasil: análise nos domínios do tempo e da frequência.** Orientador: Carlos Antônio Moreira Leite. Conselheiros: Danilo Rolim Dias de Aguiar e João Eustáquio de Lima.

O mercado *spot* de café caracteriza-se por acentuadas flutuações de preços. Essas flutuações implicam que a rentabilidade nessa atividade deverá oscilar, o que pode propiciar aos investidores mais informados um diferencial de lucratividade, pois as suas decisões de *hedge* e de *especulação* estarão baseadas em previsões mais precisas acerca do comportamento do mercado. Contudo, pode-se observar que essas possibilidades de ganho ainda não são plenamente aproveitadas por grande parte dos agentes econômicos envolvidos nesse mercado. Isto porque existe uma imperfeita compreensão acerca da natureza das flutuações dos preços do café, que, se não podem ser evitadas, ao menos deveriam ser utilizadas pelos agentes econômicos ligados a esses mercados, a fim de otimizar suas decisões de *hedge* e de *especulação*. Em termos gerais, pretendeu-se com este trabalho de pesquisa detectar a existência dos componentes estocásticos (e, ou, determinísticos) de *tendência*, *ciclos*,

sazonalidade e volatilidade condicional nos preços do mercado *spot* da principal *commodity* agrícola do Brasil: o café, analisando como a dinâmica desses preços é influenciada por esses componentes, fornecendo assim subsídios teóricos e empíricos aos agentes econômicos envolvidos em operações com essa *commodity* no mercado *spot* e no mercado futuro (*hedgers* e *especuladores*), para otimização dos seus processos de tomada de decisão. Partiu-se do princípio de que as séries históricas, de maneira geral, contêm informações extremamente importantes acerca do comportamento passado da variável em questão (preços, no presente caso). Assim, foi utilizado no presente trabalho o instrumental estatístico conhecido por *Teoria da Análise de Séries Temporais*, para verificar a existência (e analisar a natureza) dos componentes distintos de *tendência*, *ciclos*, *sazonalidade* e *volatilidade* nos preços *spot* de café. Especificamente, a metodologia empregada referiu-se a *análises no domínio do tempo* (para a análise de tendência, sazonalidade e volatilidade) e *análise no domínio da frequência*, também conhecida como *análise espectral*, para o estudo da presença de ciclos de preços. Os resultados encontrados mostraram que a tendência existente nos preços do café se configurou como uma composição de tendências dos tipos determinística e estocástica, ou seja, a série de preços do café brasileiro não se configurou como estacionária no período analisado. Com relação aos ciclos existentes nos preços do café, tem-se que a análise espectral para os dados conjuntos de toda a amostra (jan. de 1946 a dez. de 2000) confirmou a incidência de um ciclo de média duração existente no intervalo de 22 a 44 meses. Subdividindo a análise por décadas, a fim de verificar mudanças na periodicidade dos ciclos, constatou-se a existência de um ciclo bienal (24 meses) como o mais recorrente e que possivelmente esteve ligado diretamente ao ciclo biológico do cafeeiro. Todavia, duas exceções ocorreram. Verificou-se que um ciclo de longo prazo nos preços do café, de aproximadamente 60 meses (cinco anos) de duração, também foi marcante na década de 60 e que na década de 90 pôde-se observar uma mudança no comportamento dos preços em termos do componente de ciclos. Nesta última década, o ciclo bienal não foi significativo, apenas um ciclo com frequências correspondentes ao período de 12 meses, o que

configuraria um ciclo sazonal. Em relação à análise da sazonalidade nos preços, observou-se que para as décadas tomadas individualmente os resultados não diferiram significativamente dos verificados para todo o período. Observou-se que a sazonalidade, enquanto um componente de influência no comportamento dos preços do café no mercado internacional, não é da forma determinística e regular. Pelo contrário, ela é de natureza estocástica, variando em função do período do tempo. Com relação à volatilidade nos preços, os resultados indicaram que a variância condicional do modelo estabelecido para os preços do café possui raiz unitária e que esta não apresentará comportamento de reversão à sua média histórica com o passar do tempo, após um choque, o que significa que choques na volatilidade irão perdurar por muito tempo nos preços em questão; em última instância, indica um elevado grau de risco de preço (e de renda, de forma indireta) associado ao mercado *spot* do café.

ABSTRACT

LAMOUNIER, Wagner Moura, D.S., Universidade Federal de Viçosa, July 2001. **Price behavior in the spot market of Brazilian coffee: analysis in the time and frequency domain.** Adviser: Carlos Antônio Moreira Leite. Committee Members: Danilo Rolim Dias de Aguiar and João Eustáquio de Lima.

The coffee spot market is marked by deep price fluctuations. These fluctuations imply that the incomes in this activity must oscillate, and this can provide a profit differential to the most informed investors. This will happen since their decisions of hedge and speculation will be based in more accurate forecasting about the market behavior. Nevertheless, these opportunities of profit still aren't entirely maximized by the a big part of the economic agents involved in this market. This happens in function of a imperfect knowledge about the nature of the coffee price fluctuations, that if can't be stopped, should be utilized by these agents to optimize their decisions about hedge and speculation. In general, the main goal of this research work was to detect the existence of the stochastic (and/or deterministic) components of *Trend*, *Cycles*, *Seasonality* and *Conditional Volatility* in the spot market prices of the most important Brazilian agricultural commodity: the coffee. Analyzing how the dynamics of these prices

is influenced by these components and providing theoretical and empiric evidences and information to the economic agents involved with operations with this commodity in the spot and future market. The work begun with the principle that time series, in general, contain very important information about the past of the variable in question (coffee prices in this case). So, in the research, the methodology of *time series analysis* was fully employed to try detect and analyze the distinct components of *Trend*, *Cycles*, *Seasonality* and *Conditional Volatility* in the coffee spot market. Specifically, the methodology was the analysis in the *time domain* (to trend, seasonality and volatility) and the analysis in the *frequency domain*, also knew as *spectral analysis* to the study of cycles in the prices. The results found that the trend in the coffee prices is compounded by a mix of deterministic and stochastic trend. The series of Brazilian coffee prices was not stationary in the analyzed period. With relation to the cycles in the prices, the spectral analysis to whole sample data (from January 1946 to December 2000) has confirmed the incidence of a cycle of medium duration between the interval of 22 and 44 months. Subdividing the sample in decades, to find if the periodicity of the cycles changed, the analysis has shown the existence of a biannual cycle (24 months). This cycle was the most statistically significant and reflects the influence of the biological cycle of coffee production in the coffee prices. However, two exceptions has occurred. A long-run cycle of about 5 years of duration was significant too in the decade of 1960. Also, in the decade of 1990, the biannual cycle was not significant, only a cycle with frequencies relative to the period of 12 months was statistically significant what can be characterized a seasonal cycle. With relation to seasonality the research found too that this component of the price series has a very unstable characteristic through the decades analyzed. The seasonality, in the international prices of the Brazilian coffee are not deterministic and regular; it has a stochastic nature, and varies in function of the time. With relation to the volatility in the prices, the results pointed that the conditional variance of the model to the coffee prices series has a unit root and will not return to its historic mean with the time flow, after a shock. This means that shocks in the volatility will persist in these prices, indicating a

very high degree of risk (of price and income) to the economic agents operating in the Brazilian coffee spot market.

1. INTRODUÇÃO

Inerente à toda atividade produtiva humana está inserida a incerteza; conforme aponta DUARTE (1996), qualquer medida numérica dessa incerteza pode ser chamada de risco. Assim, tem-se que na produção de *commodities* agropecuárias a incerteza pode se manifestar sob várias formas, como, por exemplo, na possibilidade da ocorrência de fenômenos naturais, como geadas e secas, que serão prejudiciais à produtividade da atividade; na ocorrência de supersafras; nas intervenções e regulamentações governamentais não antecipadas pelos agentes econômicos; na possibilidade da criação de barreiras à entrada dos produtos nacionais nos mercados importadores; etc. Esses fenômenos, além de outros não mencionados, possivelmente implicarão impactos diretos sobre o nível dos preços das *commodities*.

De acordo com DEBERTIN (1986), alguns mercados de *commodities* agropecuárias operam em sistemas próximos ao modelo teórico da concorrência pura. Isso significa que choques exógenos nos preços dos seus produtos e insumos produtivos terão efeitos diretos na rentabilidade dessas atividades, visto que os agentes econômicos aí envolvidos são tomadores de preços. Nesse sentido, tem-se que o chamado risco de mercado, para esses agentes, irá depender diretamente do comportamento do preço da *commodity* diante das condições

adversas do mercado e de como esses agentes (produtores) irão atuar e formar as suas expectativas perante essas condições.

De maneira similar, conforme observa ATRASAS (1993), existem riscos para aqueles agentes econômicos que utilizam os produtos agropecuários como matérias-primas em seus processos produtivos, como ocorre com as agroindústrias. As possíveis oscilações nos preços, resultantes de mudanças nas condições de oferta e demanda dessas *commodities*, que são insumos para as suas respectivas atividades, poderão levar ao estabelecimento de um horizonte de instabilidade para o planejamento de suas ações, relativo aos seus produtos, clientes e fornecedores, o que em última instância pode significar redução na sua lucratividade.

Os mercados futuros têm-se caracterizado como um "forte instrumento de mercado" na busca da minimização do risco pelos agentes, relativo à oscilação dos preços das principais *commodities* agropecuárias. Isto se dá por meio de uma característica inerente às operações nesses mercados, segundo a qual tem-se numa operação de compra e venda a futuro a possibilidade da transferência do risco de preço dos *hedgers* para os *especuladores*.

Os *hedgers* são os agentes que de fato lidam com o produto no mercado físico (*spot*), como os produtores e os processadores, que visam garantir um preço para o produto, enquanto os especuladores se configuram como aqueles agentes que desejam abrir posições, apostando na alta ou queda dos preços. Deve-se ressaltar que os especuladores possuem papel fundamental para o pleno funcionamento do mercado futuro, visto serem eles que de fato propiciam a liquidez dos contratos transacionados nesses mercados.

Contudo, conforme pode-se observar no Quadro 1, as operações com contratos futuros de *commodities* agropecuárias no Brasil ainda estão um tanto quanto restritas, quando comparadas às operações com outros ativos, chamados financeiros, como ouro, taxa de juros, taxa de câmbio e títulos da dívida na Bolsa de Mercadorias e Futuros do Brasil - BM&F. As principais razões para esse aparente desinteresse pelos contratos agropecuários na BM&F podem ser apontadas como sendo de três naturezas distintas:

- A primeira está ligada ao fato de que os produtores, principalmente, e os processadores de produtos agropecuários no Brasil têm encarado as Bolsas de Futuros e os seus produtos como "mecanismos sofisticados" e de difícil compreensão, principalmente no que tange aos contratos de opções.
- A segunda se refere à relativamente pouca demanda de contratos agropecuários por parte dos especuladores. Nota-se certa falta de compreensão acerca dos mecanismos de formação dos preços dessas *commodities* no mercado *spot*, em razão das características de aleatoriedade mencionadas anteriormente, advindas da natureza particular dos processos produtivos dessas *commodities*.
- Por fim, tem-se que a grande intervenção estatal no *agribusiness* brasileiro, via controle de preços e de estoques, e como o principal financiador da atividade rural, também tem inibido o crescimento das operações com futuros agropecuários no Brasil nas últimas décadas.

Quadro 1 - Volume geral acumulado de janeiro a dezembro de 2000 das operações na Bolsa de Mercadorias e Futuros do Brasil - BM&F

Tipo de contrato	Contratos negociados	Volume financeiro		Participação (%) no total de contratos negociados	Participação (%) no volume financeiro em US\$
		R\$/mil	US\$/mil		
Ouro	247.503	619.153	340.817	0,34	0,01
Índice de ações	6.927.785	343.358.611	188.610.596	9,51	5,37
Taxas de juros	43.562.132	4.234.058.053	2.303.059.656	59,79	65,61
Taxas de câmbio	21.462.283	1.857.773.280	1.012.123.036	29,46	28,83
Títulos da dívida externa	2.600	268.493	146.502	0,00	0,00
<u>Agropecuários</u>	<u>655.194</u>	<u>10.592.648</u>	<u>5.792.802</u>	<u>0,89</u>	<u>0,17</u>
Açúcar cristal futuro	49.291	261.170	141.384	0,07	0,00
Boi gordo futuro	149.348	2.030.260	1.116.172	0,21	0,03
Café arábica futuro	386.929	7.188.680	3.933.905	0,53	0,11
Soja futuro	2.243	19.302	10.770	0,00	0,00
Algodão futuro	306	7.769	4.334	0,00	0,00
Milho futuro	8.018	48.208	26.543	0,01	0,00
Álcool an. fut.	48.973	1.019.580	549.984	0,07	0,02
Total pregão	72.857.497	6.446.670.238	3.510.073.409	100,00	100,00

Fonte: Elaboração do autor a partir de dados da BM&F.

Assim, tem-se configurado um quadro em que o desconhecimento de aspectos ligados aos mercados em que atuam se configura como determinante para a ainda pouca utilização de contratos futuros agropecuários no Brasil.

No que tange a esse último fator, observa-se que a capacidade de intervenção estatal na economia vem se esgotando nos anos recentes. O déficit público cresceu significativamente no período do Plano Real, e, no que se refere às políticas agrícolas de controle de preços e dos estoques e de financiamento da produção agropecuária, observa-se gradativo redirecionamento para instrumentos de mercado como fontes de financiamento da produção, assim como para o controle de riscos nas atividades rurais. No entanto, pouco ainda tem sido feito no sentido de se buscar resolver os dois primeiros fatores de entrave ao desenvolvimento dos mercados de *commodities* agropecuárias no Brasil, citados anteriormente.

1.1. O problema e sua importância

De acordo com BERNSTEIN (1997), a capacidade de definir o que poderá acontecer no futuro e de optar entre várias alternativas é central às sociedades contemporâneas; dessa forma, os agentes deverão buscar a máxima compreensão dos riscos (no presente caso, os riscos de preços) a que estão sujeitos, a fim de tomar as decisões mais corretas possíveis. Isso porque, em um ambiente econômico competitivo tipicamente capitalista, a busca da eficiência produtiva, capaz de garantir a sobrevivência da firma, se torna um ponto-chave. Nesse sentido, o controle do risco, via conhecimento das especificidades dos mercados em que atuam, tem se configurado, cada vez mais, como um dos principais aspectos na elevação da eficiência das diversas atividades econômicas.

Todavia, o que se pode observar é que essas possibilidades de ganho, advindas de uma gestão mais eficiente dos riscos, ainda não são plenamente aproveitadas por grande parte dos agentes econômicos envolvidos nesses mercados, em razão da imperfeita compreensão acerca da natureza das flutuações dos preços, que, se não podem ser evitadas, ao menos deveriam ser mais bem

compreendidas pelos agentes econômicos ligados a esses mercados, a fim de otimizar suas decisões de *hedge* e de especulação.

Em alguns casos, o que se tem é justamente o contrário, com a recorrência de erros sistemáticos nas decisões dos agentes, que de maneira geral culminam com fenômenos como os de supersafras indesejáveis, boi gordo com preço menor que o do boi magro, quedas abruptas nos preços diante da inelasticidade da demanda e, em alguns casos, desperdício e queima de produtos. Esses fatores, em última instância, implicam uma baixa rentabilidade das atividades, ou mesmo na persistência de prejuízos e na saída da atividade dos produtores. Isso, em grande parte, está ligado a um ciclo vicioso de baixa renda, inadimplência recorrente e impossibilidade de poupança, que marca grande parte da atividade agropecuária brasileira, de acordo com o discutido por BUENO (1997).

Conforme aponta AGUIAR (1992:1):

"...a disponibilidade de uma série histórica de preços é um instrumento de grande valia para o planejador da atividade agrícola, mas este nem sempre aproveita esse recurso como deveria...".

As séries históricas, de maneira geral, contêm informações importantes acerca do comportamento passado da variável em questão (preços do café, no presente caso). Todavia, complementando o proposto por AGUIAR (1992), a mera disponibilidade dessas séries para o planejador da atividade agropecuária e para os demais agentes econômicos envolvidos com mercados de *commodities* não será de grande utilidade se estes não forem capazes de "extrair" desses dados informações relevantes, capazes de alavancar a rentabilidade de suas atividades e, ou, aumentar a eficácia das suas decisões e políticas.

Aparentemente, as séries poderão apresentar um comportamento tipicamente aleatório, porém métodos matemáticos e estatísticos, como os propostos pela *Teoria da Análise de Séries Temporais*, poderão revelar a existência de comportamentos (componentes) distintos e relativamente difíceis de serem observados à primeira vista, quais sejam: *tendência, ciclos, sazonalidade* e

volatilidade (também chamada de *variações irregulares*)¹. Esses comportamentos, se devidamente compreendidos e incorporados ao conjunto de informações dos agentes econômicos envolvidos com essas *commodities*, poderão propiciar a estes um diferencial positivo de lucratividade e, ou, eficiência nos seus respectivos processos de tomada de decisões.

Particularmente no que tange às séries históricas de preços, tem-se que a análise do comportamento destas apresenta papel fundamental dentro da economia, visto que praticamente todas as fases das relações econômicas estão diretamente relacionadas aos preços. A análise dos preços agropecuários se configura como ponto de fundamental interesse, dada a ampla gama de inter-relações com outros setores de atividade e agentes econômicos envolvidos com esse setor, principalmente em uma economia com vocação agropecuária como a do Brasil, em que as atividades ligadas ao agronegócio são responsáveis por aproximadamente 32% do PIB, conforme aponta RODRIGUES (2001).

As aplicações das informações acerca do comportamento dos preços de *commodities* agropecuárias são as mais diversas possíveis e, conforme discutido por THOMSEN e FOOTE (1952), podem ser classificadas com relação aos agentes econômicos que poderão se beneficiar da geração dessas informações, a saber:

a) Usos pelos *especuladores* que operam nos mercados futuros:

- na elaboração de planos de investimentos e de sistemas de negociação;
- na determinação da composição de seus *portfólios* de ativos;
- como indicadores da melhor hora para se entrar e sair de um mercado;
- na modelagem e previsão do comportamento futuro dos mercados em que atuam;
- no cálculo do risco envolvido em suas operações.

¹ Esses componentes da variação em uma *série temporal* serão discutidos posteriormente.

b) Usos pelos *produtores individuais*:

- no planejamento de longo prazo dos seus negócios;
- na determinação e no controle do risco de suas atividades;
- na melhor alocação das suas propriedades e recursos produtivos, ou seja, na seleção do que, quando e como produzir;
- na definição da necessidade e forma de realização de seguros de preços para as *commodities*.

c) Usos pelas *cooperativas* de produtores privados:

- no planejamento do tipo, na localização, no tamanho e em outros aspectos relativos aos seus empreendimentos agropecuários;
- na determinação do tempo e local para compra de matérias-primas e no gerenciamento mais eficiente dos seus estoques;
- na seleção de mercados, períodos e esforços promocionais para a venda dos seus produtos;
- na definição de estratégias de minimização do risco de variações na renda de seus cooperados;

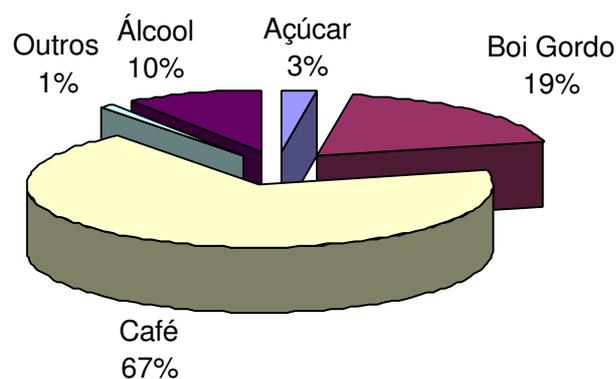
d) Usos por *órgãos governamentais*:

- no planejamento de programas e políticas agrícolas de escopo comunitário, estadual, regional e nacional;
- na processo de disseminação de informações para os produtores rurais e público geral acerca dos resultados dos programas e das políticas agrícolas;
- na avaliação e no ajuste de tais programas;
- na gestão de políticas macroeconômicas, do balanço de pagamentos e do volume de reservas cambiais, para países com alta participação da agropecuária no PIB e na pauta de exportações;
- na definição de mercados e produtos a serem enfatizados pelos programas em função de maiores ou menores riscos associados a estes mercados.

e) Usos pelos *consumidores*:

- os consumidores individuais, principalmente os grandes consumidores, como agroindústrias, cadeias de restaurantes, supermercados e hotéis, podem se beneficiar do conhecimento de informações sobre os preços. Estas serão importantes nas suas decisões acerca da época oportuna para as suas aquisições, do volume dos estoques dos produtos e da utilização de *hedge* de preços nos mercados futuros dessas *commodities*;
- os atacadistas e varejistas podem decidir sobre o volume de transações a serem efetuadas, fundamentados nos conhecimentos do comportamento dos preços.

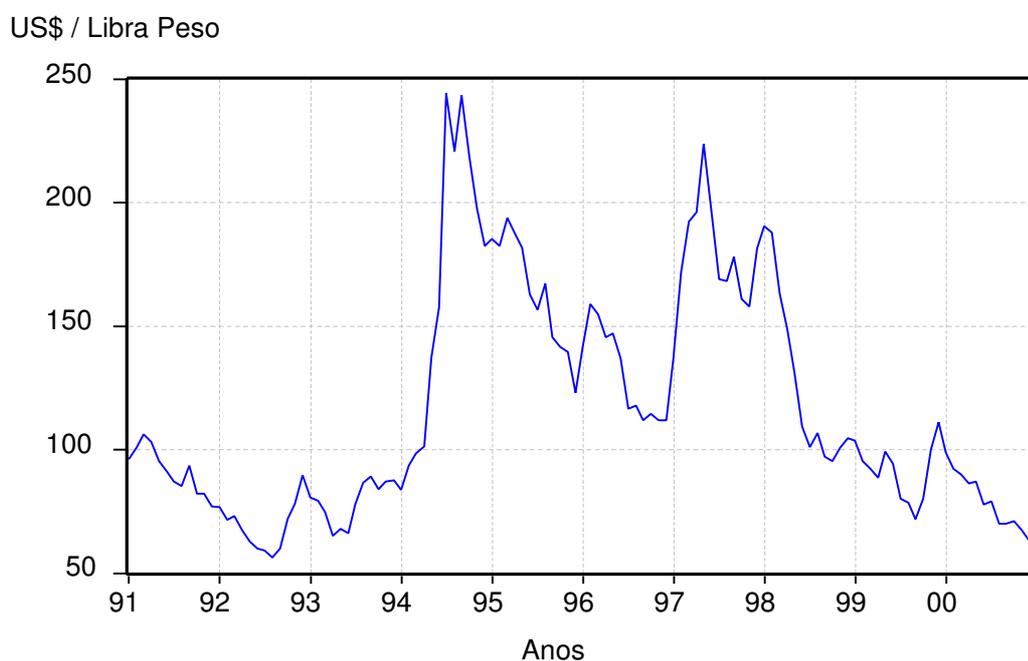
Apesar da pouca liquidez da maioria dos contratos futuros de *commodities* agropecuárias na BM&F (conforme pôde-se observar no Quadro 1), o café se destaca como a principal *commodity* transacionada no mercado de futuros do Brasil. Isso se dá tanto em termos do número de contratos negociados, como em termos do volume financeiro envolvido nessas transações (Figura 1).



Fonte: Dados básicos da BM&F.

Figura 1 - Distribuição dos contratos agropecuários negociados em 2000 na BM&F, em termos do volume financeiro (US\$/mil).

Esse fato, em grande parte, é reflexo das acentuadas flutuações de preços, ao longo do tempo, que caracterizam o mercados *spot* de café (Figura 2). Conforme discutido anteriormente, essas flutuações implicam que a rentabilidade nessa atividade deverá oscilar tanto para os cafeicultores diretamente envolvidos com a produção quanto para os demais agentes que lidam, direta ou indiretamente, com essa *commodity*.



Fonte: Elaboração do autor a partir de dados do Agriannual 2000 - FNP.

Figura 2 - Evolução do preço do café brasileiro no mercado *spot* de Nova Iorque - 1991/2000 (em centavos de US\$/libra-peso²).

² Cada libra-peso tem 453,6 gramas. Para transformar dólares por saca de 60 kg em centavos de dólar por libra-peso, deve-se dividir 60.000 gramas por 453,6 gramas e depois dividir por 100, o que resultará no fator 1,3228. Inversamente, para transformar o preço de Nova Iorque (US\$cent./libra-peso) em cotações da BM&F (US\$/saca), deve-se multiplicar o primeiro por 1,3228.

No que tange aos produtores de café, o conhecimento do comportamento dos preços poderá ser extremamente útil nas suas tomadas de decisões com relação ao planejamento da produção, à manutenção e formação de estoques, dentre outras, podendo assim aproveitar de maneira mais eficiente as fases de baixa e de alta nos preços para a maximização de lucros.

Conforme aponta HIERONYMUS (1976), os *hedgers* e, principalmente, os especuladores deverão acumular o máximo possível de informações acerca das *commodities* que transacionam, no presente caso, o café, principalmente aquelas informações referentes ao comportamento dos preços. Isto ocorre porque uma característica intrínseca de um mercado de futuros é a de que ele pode ser caracterizado como um jogo de soma zero, menos os custos de transações, ou seja, as comissões. Nesse sentido, o ganho de um agente se configura na perda de uma outra parte; assim, cada agente atuando nesses mercados deverá buscar sempre superar os demais em termos do melhor conhecimento do comportamento dos preços e dos mercados das *commodities* que transaciona, a fim de maximizar seus ganhos.

Diante desse contexto, torna-se importante para os agentes econômicos envolvidos com o mercado de café, particularmente os *hedgers* e especuladores, a busca de informações e "mecanismos" capazes de minimizar suas incertezas quanto aos preços futuros dos seus produtos. Análises e informações acerca do comportamento dos preços no mercado *spot* dessa *commodity* se apresentam como pontos fundamentais para a otimização de seus processos decisórios.

Ainda nesse sentido, SCHWAGER (1996) ressalta a importância desse tipo de análise, dentro de um processo de tomada de decisão sob situação de incerteza, que incide sobre os agentes econômicos em diversos ramos de atividade e, particularmente, aqueles envolvidos com os mercados financeiros e de *commodities*. Para ele, tanto os *hedgers*, envolvidos com as transações físicas dos produtos em questão, que querem se proteger de futuras oscilações de preços, quanto os especuladores³, que querem assumir riscos nesses mercados, em

³ E que tem importância fundamental, pois propiciam a liquidez necessária ao pleno funcionamento dos mercados.

função de possíveis ganhos futuros, necessitam de informações sobre as variáveis e os ativos que são fundamentais para a rentabilidade de suas atividades. Ou seja, conforme ressalta ainda o referido autor, modelos estatísticos para a análise dos componentes determinísticos e estocásticos dos preços das *commodities* são de grande importância, no sentido em que podem ser entendidos como *inputs for trading decisions*.

Tem-se, portanto, que a oscilação na renda dos produtores de café, proveniente de flutuações nos preços, se configura como um problema cujas características e causas devem ser amplamente pesquisadas, em vista da importância da *commodity* no agronegócio nacional e das perdas que essas flutuações podem provocar tanto em termos de lucratividade para o setor quanto em termos de empregos e divisas para o país.

1.2. Hipótese

Na tentativa de antecipação, à guisa de hipótese de trabalho, predefiniu-se o seguinte quadro: postula-se que o comportamento dos preços no mercado *spot* (à vista) de café, que é a principal *commodity* agrícola brasileira transacionada em bolsas de futuros, apresenta padrões sistemáticos e repetitivos na forma de componentes estocásticos e, ou, determinísticos, quais sejam: tendências, ciclos, sazonalidade e volatilidade condicional.

1.3. Objetivos

1.3.1. Objetivo geral

Em termos gerais, pretendeu-se verificar a existência dos componentes estocásticos (e, ou, determinísticos, quando for o caso) de tendência, ciclos, sazonalidade e volatilidade condicional na série temporal de preços do mercado *spot* do café, principal *commodity* agrícola do Brasil.

Pretendeu-se ainda analisar como a dinâmica desses preços é influenciada por esses componentes, fornecendo subsídios teóricos e empíricos para os agentes econômicos atuantes nos mercados à vista de café e para os *hedgers* e especuladores que operam nas bolsas de futuros. Essas informações poderão se configurar como relevantes para a otimização dos seus processos de tomada de decisão.

1.3.2. Objetivos específicos

Especificamente, pretendeu-se:

- testar a validade da hipótese que postula que os preços dos produtos agrícolas, em geral, possuem tendência declinante ao longo do tempo⁴, confrontando esta com a *hipótese de passeio aleatório*⁵ (*teoria dos mercados eficientes*) para os preços do café;
- verificar se os componentes de tendência e de sazonalidade nas séries, caso existam, são do tipo *determinístico* ou do tipo *estocástico*, discutindo as suas diferentes implicações para os agentes econômicos envolvidos;
- determinar quais ciclos (em diferentes frequências) têm afetado de forma mais significativa o comportamento dos preços, identificando as suas respectivas amplitudes e verificando se a periodicidade desses ciclos tem sofrido variações ao longo do tempo;
- verificar a incidência de variações sazonais nos preços e a possibilidade de o efeito sazonal ter variado com o passar dos anos, estabelecendo por fim índices sazonais para os preços e obtendo uma série desazonalizada;
- verificar se a volatilidade dos preços é do tipo *aleatória* ou *condicional* e, caso seja condicional, testar a possibilidade de *assimetria*, discutindo ainda os efeitos dessa forma de volatilidade para a análise e previsão do comportamento futuro da série;

⁴ Também conhecida como *hipótese de Schultz*. Todavia, ela será apresentada de forma mais detalhada posteriormente.

⁵ Indica, basicamente, que os preços não deverão apresentar um comportamento previsível. Todavia, esta hipótese também será melhor definida no corpo do trabalho.

- verificar se os possíveis *clusters de volatilidade* tendem a ter períodos de duração similares, discutindo as suas implicações.

2. METODOLOGIA

2.1. Referencial teórico

O marco teórico do presente trabalho se refere à análise de mercados agropecuários no que tange a uma de suas principais "dimensões", qual seja a análise dos preços praticados nesses mercados. Particularmente, tem-se como objeto de análise os preços à vista (*spot*) praticados no mercado de café, que é a principal *commodity* agropecuária transacionada na BM&F, em termos de volume de contratos e de recursos financeiros.

Existe na literatura sobre finanças, particularmente a referente ao estudo dos preços de ações, *commodities* e títulos, uma gama relativamente ampla de abordagens e métodos para análise e previsão desses preços. No entanto, essas abordagens analíticas podem ser agrupadas em dois grupos distintos, classificados genericamente como *Análise Fundamental* e *Análise Técnica*.

De acordo com SCHWAGER (1984), a análise fundamental se refere aos estudos que têm como "insumos" básicos diversos dados e informações econômicos acerca da produção, do consumo, das exportações, das taxas de juros e das políticas macroeconômicas relacionados com o ativo analisado. Ou seja, esse tipo de análise busca determinar relações de causa e efeito entre diversas

variáveis econômicas, a fim de explicar o comportamento dos mercados e preços de interesse.

Por outro lado, para a análise técnica, o principal, e geralmente único⁶, insumo para a elaboração de suas inferências e previsões é o próprio comportamento passado dos preços do ativo em questão. O pressuposto básico desse tipo de análise é o de que os preços exibem comportamentos repetitivos padronizados, e o reconhecimento desses padrões pode ser útil para a identificação de oportunidades de ganho para os agentes econômicos envolvidos com esses ativos.

A mais difundida forma de implementação da análise técnica se refere à *análise gráfica*⁷ das séries para a verificação de padrões e previsão do comportamento desta. Contudo, os modelos univariados de análise de séries temporais também podem ser caracterizados como modelos de análise técnica, pois apresentam como característica principal o fato de que possibilitam inferências e previsões sobre o comportamento das séries tendo como base, principalmente, o comportamento passado das séries.

Os modelos de análise de séries temporais possibilitam uma descrição da natureza aleatória do processo que gerou as observações amostrais estudadas. Essa descrição é dada não em termos de uma relação de causa e efeito (como ocorre na análise fundamental por meio de modelos econométricos "tradicionais" de regressão), mas sim em termos de como essa aleatoriedade está incorporada no processo.

Conforme apontam BOX e JENKINS (1976), o estudo e a elaboração de modelos de análise de séries temporais geralmente estão ligados a três objetivos distintos, porém não necessariamente mutuamente exclusivos, quais sejam:

- *estudo dos padrões comportamentais* das séries, dados os seus componentes não-observáveis;

⁶ A análise técnica poderá incluir e utilizar também dados acerca do volume de negócios realizados e de contratos em aberto.

⁷ Uma análise profunda desse tipo de metodologia pode ser encontrada em SCHWAGER (1995).

- a *previsão* do comportamento futuro das séries, com o uso de *modelos univariados e multivariados*;
- a implementação de métodos para o *controle* do comportamento futuro das séries e demais variáveis correlacionadas.

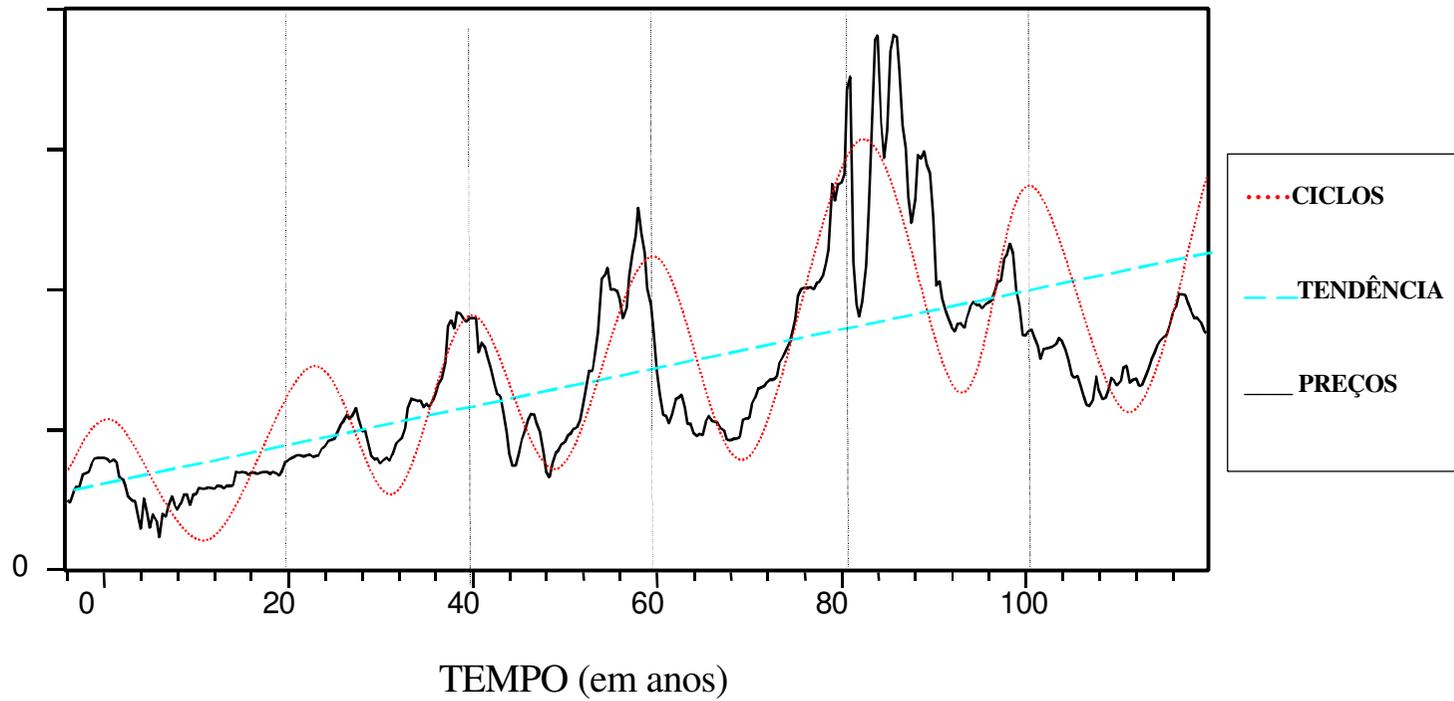
Como o enfoque do presente trabalho de pesquisa se refere ao estudo do comportamento das séries, não serão discutidos diretamente os modelos e procedimentos ligados à previsão⁸ e ao controle das séries.

Uma das principais formas de análise dos preços se dá pela busca da decomposição das séries de preços das *commodities* em termos de seus componentes principais. SCHWAGER (1995) argumenta que análises de tendências, sazonalidades, ciclos e volatilidade em séries econômicas (financeiras) são de extrema importância dentro de um processo de tomada de decisão sob situação de incerteza, que incide sobre os agentes econômicos em diversos ramos de atividade, particularmente aqueles envolvidos com os mercados financeiros e de *commodities*. Esquemáticamente, esses componentes da variação em uma *série temporal* podem ser observados com a ajuda das Figuras 3 e 4.

A *tendência*, em uma série temporal, está ligada, conforme aponta CHATFIELD (1996), a uma mudança no nível médio da série no longo prazo. Ou seja, a tendência irá refletir o declínio, a elevação ou a estabilidade (quando de fato não houver tendência) do valor médio da série temporal no longo prazo; para o caso de variáveis econômicas, pode-se convencionar como sendo referentes a observações relativas a períodos de tempo compreendidos entre 25 e 50 anos, aproximadamente.

⁸ Que é o caso de alguns modelos como os de *espaço de estado* e de *redes neurais*, que são do tipo "caixa preta", mais utilizados para previsão do que para análise das componentes em si.

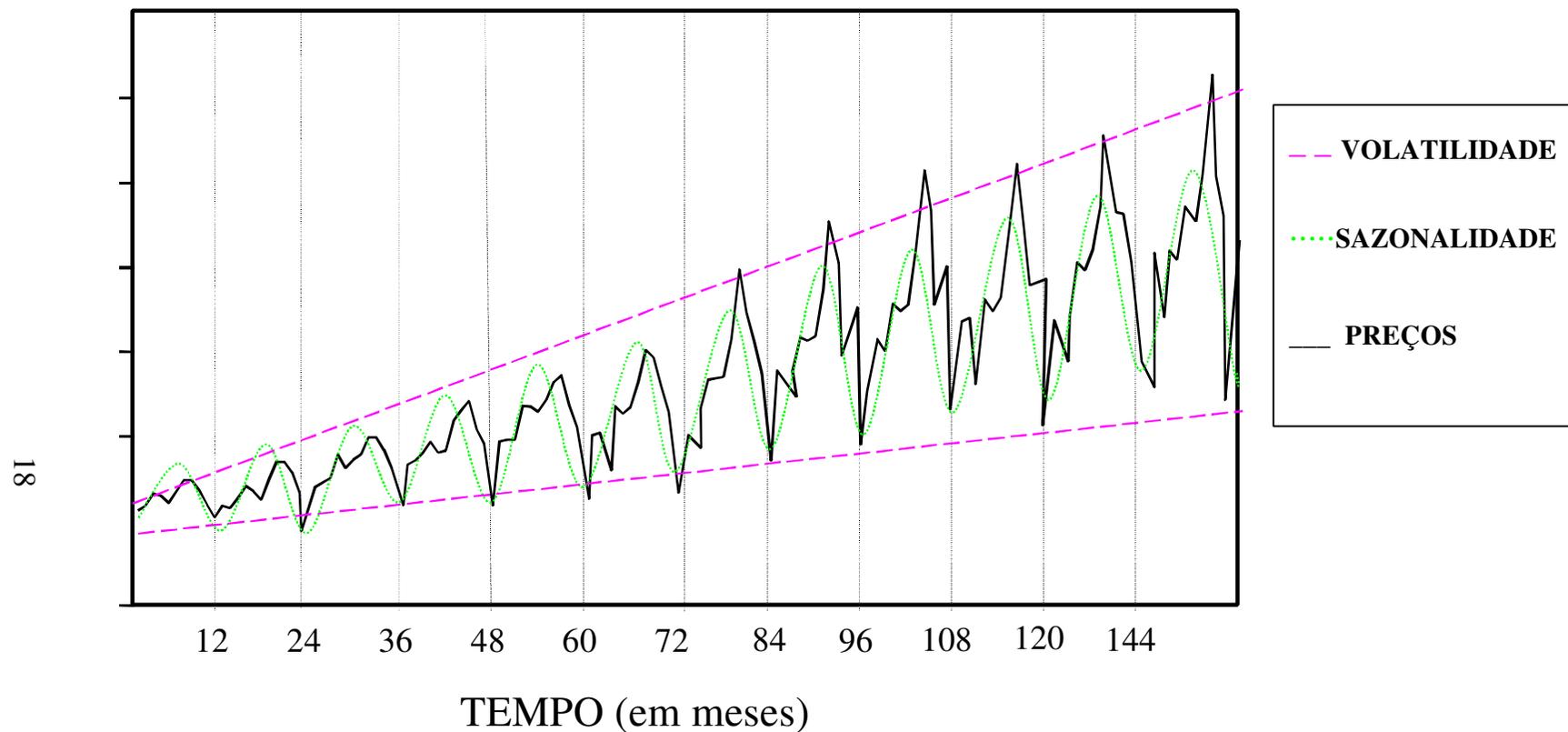
PREÇOS



Fonte: Elaboração do autor.

Figura 3 - Decomposição esquemática dos componentes TENDÊNCIA e CICLOS em uma série temporal de preços fictícia.

PREÇOS



Fonte: Elaboração do autor.

Figura 4 - Decomposição esquemática dos componentes VOLATILIDADE e SAZONALIDADE em uma série temporal de preços fictícia.

Os *ciclos*, em uma série de tempo, se referem a movimentos recorrentes de elevação e queda em torno do nível médio da tendência e, portanto, também são relativos a dados que refletem o longo prazo da variável em questão. De acordo com SOUZA (1989), deve-se ressaltar que os movimentos para cima e para baixo em torno da tendência podem ser estritamente periódicos ou aproximadamente periódicos; assim, eles podem variar em comprimento e magnitude.

O terceiro componente que pode estar presente em uma série temporal é aquele relativo à *sazonalidade*. Assim como o componente cíclico, a sazonalidade também está relacionada a movimentos para cima e para baixo em torno de um valor médio, porém ela difere do primeiro basicamente em dois aspectos: a sazonalidade possui um comprimento constante de 12 meses, repetindo-se nesta base periódica regular; e as variações sazonais podem ser observadas tendo-se por base períodos menores de tempo (médio e curto prazo); dessa forma, os dados a serem analisados não deverão estar distribuídos em observações anuais, mas sim mensais ou trimestrais, para que seja possível a modelagem da sazonalidade.

O componente de *volatilidade* (ou de flutuações irregulares), até há pouco tempo, era descrito como movimentos erráticos em uma série temporal que não seguem um padrão regular ou identificável. Ele seria determinado por uma série de resíduos aleatórios (com média zero e variância uniforme) que resultariam de uma série de tempo, após a "retirada" dos componentes de tendência, ciclos e sazonalidade, conforme a definição dada por BOWERMAN e O'CONNELL (1979).

No entanto, estudos posteriores, como os de ENGLE (1982) e BOLLERSLEV (1986), indicaram que essa série de resíduos poderia não apresentar comportamento aleatório puro, ou seja, eles verificaram que em grande parte das séries temporais, particularmente as relativas a variáveis financeiras, a volatilidade dos dados não é estável, o que implica que a série não apresenta a propriedade desejável de *homocedasticidade*. Assim, a volatilidade poderia ser modelada e descrita em termos de dois componentes distintos: a

volatilidade incondicional, que seria de fato constante; e a *volatilidade condicional*, que poderia oscilar ao longo do tempo e que pode ser identificada e analisada a partir dos modelos de análise de heterocedasticidade condicional.

Esses movimentos e padrões comportamentais dentro de uma série de tempo geralmente estão associados a fenômenos reais (determinísticos ou não). Particularmente, quando da análise de séries de preços agropecuários, tem-se que esses movimentos podem ser explicados em termos de alguns aspectos básicos. Conforme afirmam SHEPHERD (1972), DAHL e HAMMOND (1977) e TOMEK e ROBINSON (1991), os movimentos dos preços de *commodities* agropecuárias são provocados por diferentes forças e fenômenos que se relacionam diretamente com o período de tempo envolvido. Esses preços são de modo geral altamente instáveis, oscilando de ano em ano, de mês em mês, de dia em dia e, no limite, de segundo a segundo. Assim, a análise dos movimentos desses preços ao longo do tempo pode, e deve, ser implementada tendo-se como referência a possibilidade de uma segmentação do tempo em períodos diferenciados.

2.1.1. Movimentos de longo prazo nos preços das *commodities* agropecuárias

2.1.1.1. Tendências

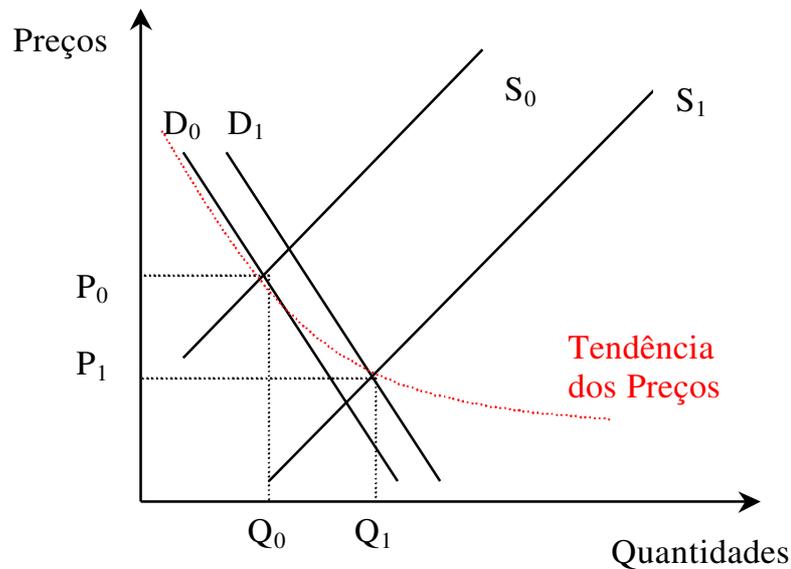
O conceito do que vem a ser curto e longo prazo em economia é um tanto quanto problemático e discutível. SHEPHERD (1972) sugere que o longo prazo para a análise de preços agropecuários se refere a períodos de tempo de aproximadamente 50 anos. Portanto, a análise de longo prazo dos preços pode ter como base observações de dados anuais, semestrais, trimestrais ou mesmo mensais que possibilitem a análise dos componentes de tendência e ciclos nos preços.

Particularmente no que se refere aos produtos de origem agropecuária, tem-se que as tendências nos preços são geralmente associadas às relações intersetoriais na economia e a aspectos específicos da natureza do processo

produtivo e comercial dos produtos desse setor, como: mudanças nos gostos e preferências dos consumidores; crescimento da população e da renda dos indivíduos; modificação nas tecnologias empregadas; e redução nos custos de produção. Assim, conforme apontam TOMEK e ROBINSON (1991), movimentos de tendência de elevação ou redução sistemática nos preços de produtos agropecuários ao longo do tempo podem ser provocados por deslocamentos nas curvas de oferta e de demanda desses bens.

Para SCHULTZ (1945), os preços das *commodities* agropecuárias tenderão a declinar ao longo do tempo. Ele estabeleceu uma relação básica entre a oferta e a demanda de produtos agropecuários, que ficou conhecida na literatura pertinente com *hipótese* (ou fenômeno) *de Schultz*. Essa hipótese estabelece que a oferta agregada de *commodities* agropecuárias é altamente inelástica tanto no curto quanto no longo prazo, pois existem poucas alternativas fora da produção agropecuária para a maioria das comunidades rurais e os investimentos de capital são muito específicos para o setor. Por outro lado, a demanda agregada por esses bens também tende a ser inelástica, uma vez que as pessoas devem se alimentar para sobreviver, mesmo que os preços dos alimentos sejam altos; já que a capacidade limitada do estômago humano impõe uma restrição natural à resposta a possíveis quedas nos preços.

Tanto a oferta como a demanda de produtos agropecuários tendem a crescer ao longo do tempo. A oferta crescerá graças às inovações tecnológicas que elevam a produtividade do setor, e a demanda deverá crescer em função dos aumentos na renda dos indivíduos e do próprio crescimento da população. Entretanto, a taxa de crescimento da oferta tenderá a suplantar a taxa de crescimento da demanda e, dessa forma, os preços dos produtos apresentarão tendência declinante ao longo do tempo. Conforme apontam BROOKS e CARTER (1994), a hipótese de Schultz pode ser esquematizada com a ajuda da Figura 5.



Fonte: BROOKS e CARTER (1994).

Figura 5 - Diagrama do efeito Schultz.

Como se pode observar, à medida que a oferta se desloca de S_0 para S_1 , a demanda se desloca em uma magnitude relativamente menor, de D_0 para D_1 ; assim, os preços de equilíbrio tenderão a cair, de P_0 para P_1 . Com a continuidade desse processo ao longo do tempo, pode-se estabelecer uma *linha de tendência*, ligando os pontos de equilíbrio entre oferta e demanda, que indica o comportamento decrescente dos preços das *commodities* agropecuárias ao longo do tempo.

Por outro lado, existe uma hipótese que pode contradizer aquela proposta por Schultz. Assim como *títulos* e *ações*, muitas *commodities* - dentre as quais as agropecuárias, que são largamente transacionadas nos mercados *spot* (à vista) e que possuem contratos correlacionados nos chamados mercados futuros - podem ter preços e retornos que não seguem uma tendência "tão facilmente" observável como a proposta anteriormente.

Essa hipótese está relacionada à *teoria da eficiência dos mercados*, que implica, de acordo com HELMUTH (1981) e KOPPENHAVER (1983), que os

preços em mercados de *commodities* eficientes deverão seguir um comportamento do tipo *passeio aleatório*. Se esse não fosse o caso, os agentes econômicos envolvidos nesses mercados poderiam incorporar essas informações às suas decisões, esperando obter lucros sobrenormais.

De fato, conforme destacam PINDYCK e RUBINFELD (1998), a maioria dos modelos de finanças ligados a contratos futuros, opções e outros instrumentos financeiros "atrelados" a uma *commodity* são baseados no pressuposto, que contradiz a hipótese de Schultz, de que os preços nos mercados *spot* seguem um comportamento do tipo passeio aleatório (que será discutido posteriormente) e, portanto, não são previsíveis.

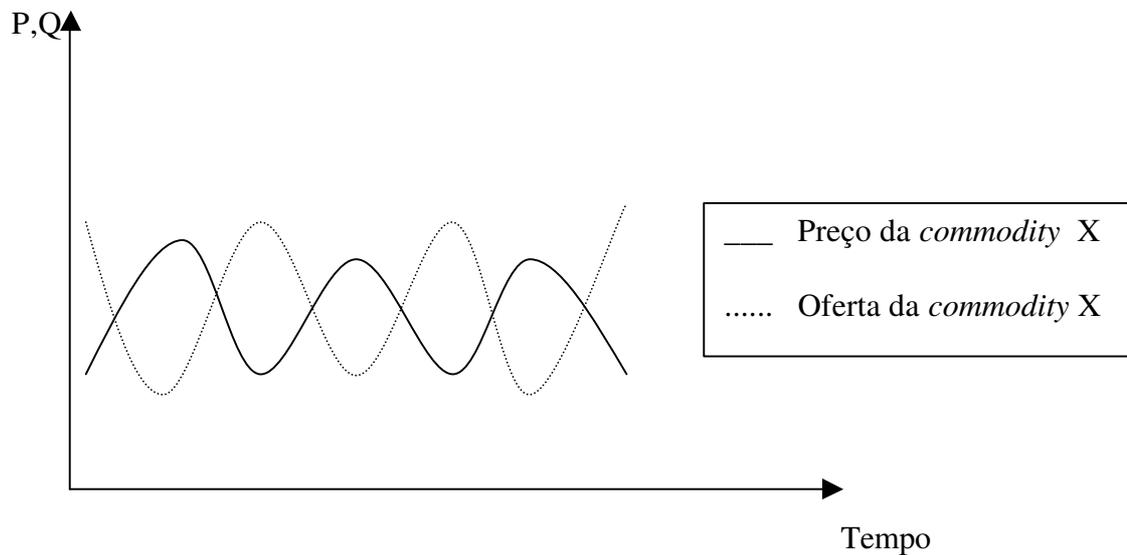
2.1.1.2. Ciclos

Os ciclos de preços, como também os de produção, se configuram como outro tipo de variação de longo prazo que pode estar presente nas séries temporais de preços das *commodities* agropecuárias. Apesar de qualquer movimento recorrente de elevação e queda nos preços poder ser apontado como um ciclo, o que se tem estabelecido como prática comum é a utilização desse termo na referência a mudanças nos preços que ocorrem ao longo dos anos, com certa regularidade e previsibilidade.

Os ciclos nos preços geralmente são inversamente correlacionados com os ciclos da produção das *commodities*, e vice-versa (Figura 6). Altos preços tendem a encorajar a ampliação da produção pelos produtores existentes e por novos produtores que entram nesses mercados, que praticamente não apresentam barreiras expressivas à entrada. Essa ampliação da oferta irá chegar ao mercado somente após determinado espaço de tempo, dada a inelasticidade da oferta dessas *commodities*. Esse intervalo de tempo poderá variar de acordo com alguns aspectos, como:

- as características biológicas das diferentes *commodities*;
- o *estado da arte* das tecnologias empregadas na produção;

- o tipo de insumos produtivos que serão utilizados no processo produtivo para a ampliação da oferta.



Fonte: Elaboração do autor.

Figura 6 - Representação esquemática dos ciclos de preço e produção em uma *commodity* fictícia.

De acordo com o proposto por DAHL e HAMMOND (1977), uma explicação genérica do comportamento cíclico que caracteriza os preços e a produção de diversas *commodities* agropecuárias pode ser dada pelo *teorema da teia de aranha (cobweb theorem)*, segundo o qual a lei da oferta e demanda explica como os preços atingem níveis de equilíbrio nos pontos de interseção das curvas de oferta (S) e demanda (D) para um dado bem econômico. Contudo, para algumas *commodities* tem-se que os preços não se ajustam a esses pontos de equilíbrio, mas flutuam em torno deles.

Para o caso de *commodities* cujas quantidades demandadas e ofertadas podem instantaneamente se ajustar a um choque nos preços, os ciclos não irão ocorrer. Entretanto, esse não é o caso da maioria desse tipo de produtos,

particularmente para o café, que é a *commodity* de interesse do presente estudo, dada a inelasticidade da curva de oferta em relação às condições da demanda.

O teorema da "teia de aranha" apresenta esse nome graças ao padrão dos ajustamentos entre preços e quantidades em um diagrama de oferta e demanda para a *commodity* que apresenta o fenômeno cíclico. Essa teoria envolve uma modificação dos pressupostos básicos dos modelos de oferta e demanda cujos preços e quantidades são determinados simultaneamente. Por exemplo, supondo um modelo simples dado por:

$$S_t = \alpha + \beta P_t \text{ (oferta)} \quad (1)$$

$$D_t = \delta - \gamma P_t \text{ (demanda)} \quad (2)$$

em que t indica o período de tempo, tem-se que o sistema será solucionado por uma única combinação de oferta (S) e demanda (D). Contudo, no modelo advindo do teorema da teia de aranha, as relações comportamentais das variáveis são diferentes, no sentido de que os preços e as quantidades da *commodity* não são determinados simultaneamente. A quantidade ofertada em determinado período (t) não será função dos preços praticados no mesmo período, mas sim uma função dos preços do período anterior ($t-1$), ou de vários períodos anteriores ($t-1, t-2, t-3, \dots, t-n$). No caso mais simples, de apenas uma defasagem, tem-se que a nova equação de oferta da *commodity* será dada por:

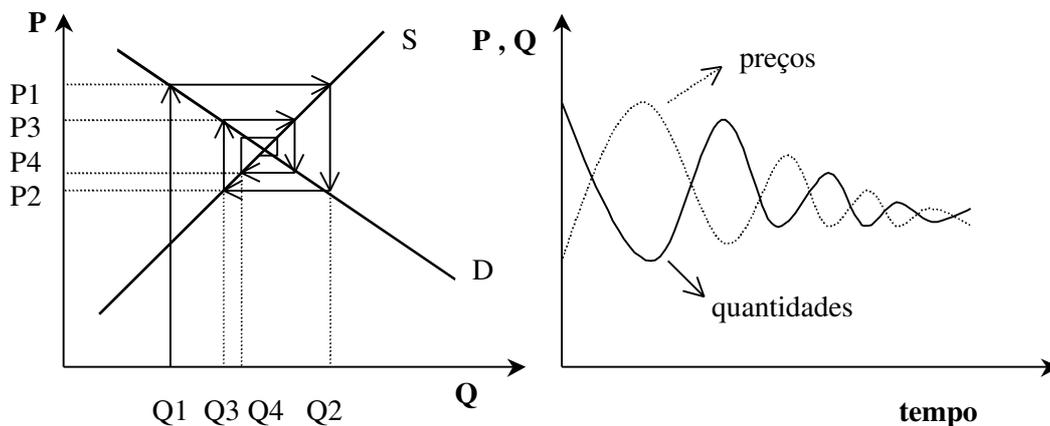
$$S_t = \lambda + \psi P_{t-1} \quad (3)$$

Visto que no período t a quantidade ofertada já está determinada, os preços irão se ajustar de forma residual, de acordo com:

$$P_t = \theta - \eta S_t \quad (4)$$

Esse preço praticado em t será a principal base para as decisões acerca da oferta que deverá chegar ao mercado num próximo momento, $t+1$ por exemplo, e que resultará num outro patamar de preços. Tem-se assim um sistema recursivo de determinação de preços e quantidades, que só será estável se o preço de um

período propiciar uma oferta no próximo período que não resulte em variações nos preços praticados nesse próximo período. No entanto, se esse não for o caso, tem-se como resultado um comportamento cíclico dos preços, que poderá ser *convergente*, *divergente* (explosivo) ou *estável*. Na Figura 7, pode-se observar o caso em que o comportamento cíclico de preços e quantidades é convergente.



Fonte: DAHL e HAMMOND (1977).

Figura 7 - Ajustamentos cíclicos convergentes em um modelo de teia de aranha.

De acordo com o proposto por DAHL e HAMMOND (1977) e SHEPERD (1972), uma quantidade ofertada Q_1 dará origem a um preço P_1 no período 1; os produtores irão se basear nesse preço para ofertar Q_2 num segundo período. Contudo, a essa quantidade corresponde o preço P_2 . Assim, a quantidade ofertada no período 3 irá cair para Q_3 , o que proporciona elevação do preço para P_3 , e assim por diante. Isso implica um movimento cíclico dos preços e das quantidades ao longo do tempo, conforme se pode observar na segunda parte do diagrama. Entretanto, esse comportamento, para o presente caso, irá convergir para um determinado nível médio de longo prazo.

Os casos de movimentos cíclicos divergentes e estáveis podem ser derivados por analogia do proposto anteriormente. A estabilidade, convergência ou divergência dos ciclos dependem diretamente do valor absoluto das elasticidades da oferta e da demanda. Se a elasticidade de oferta for superior à elasticidade da demanda o modelo será divergente; por outro lado, se a elasticidade da oferta for menor que a elasticidade da demanda, o modelo será convergente. Finalmente, tem-se o caso em que as elasticidades são iguais, o que propicia a continuidade dos ciclos ao longo do tempo de forma, aproximadamente, constante e estável.

2.1.2. Movimentos de curto prazo nos preços das *commodities* agropecuárias

2.1.2.1. Sazonalidade

Conforme destacam ABRAHAM e LEDOLTER (1983), diversas séries temporais, entre elas as relativas a preços de *commodities*, exibem um padrão comportamental cíclico que possui tendência a repetir-se em um período fixo de tempo. Esse padrão é chamado de *sazonalidade*, e o tamanho do período em que o ciclo se completa, para uma série Y_t de dados observados mensalmente, será de 12 meses, ou seja, são padrões em séries temporais que se completam em um período de um ano.

Séries que contêm componentes sazonais são comuns nas ciências naturais e particularmente nas ciências gerenciais. Em diversas séries econômicas sobre produção (industrial e agropecuária), vendas, emprego, estoques, demanda e preços, podem ser encontrados comportamentos tipicamente sazonais. Geralmente essa sazonalidade é provocada por variações nas condições climáticas e nos gostos, feriados e dias específicos no ano (Natal, por exemplo), bem como nos demais aspectos específicos da variável econômica em questão.

No que tange à produção agrícola, a sazonalidade geralmente está relacionada a fatores climáticos e aos processos de crescimento biológico das plantas. Para a produção animal, a sazonalidade pode surgir, de acordo com

TOMEK e ROBINSON (1991), por diversas razões, como: variação das condições climáticas e de disponibilidade de pastagens e pelas características biológicas dos processos de produção.

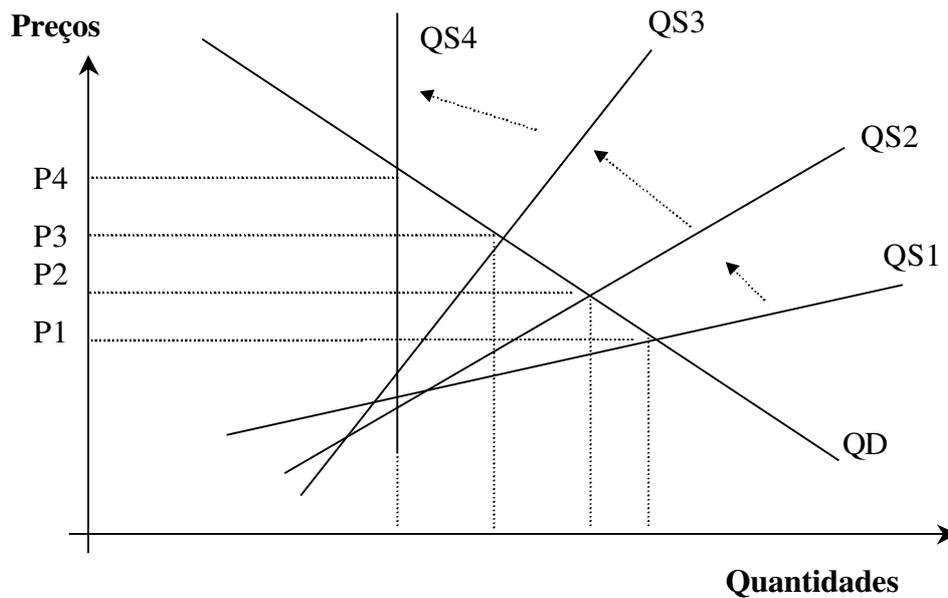
As flutuações sazonais no preço das *commodities* agropecuárias tendem a se relacionar inversamente com suas respectivas ofertas nos mercados. Ou seja, os preços tendem geralmente a atingir níveis mais baixos durante os períodos de safra e passarão a subir gradualmente até um máximo, alguns meses antes que a próxima safra se inicie. VAUGHN et al. (1981:93) apresentam uma definição mais clara e específica ligada ao estudo da sazonalidade dos preços agropecuários, que será a adotada no presente trabalho. Segundo esses autores:

Acredita-se que a maioria das commodities agropecuárias estão sujeitas a notáveis efeitos sazonais atribuíveis a variações nas condições de oferta e demanda desses bens. Efeitos Sazonais serão definidos aqui como repetitivas flutuações cíclicas nos preços, em períodos de 12 meses, que se originam de similares mudanças na oferta e demanda de ano para ano. A forma e a magnitude das variações sazonais dos preços dependerão em grande parte da natureza do ciclo de produção, da demanda pelas commodities e da possibilidade e condições de estocagem dos produtos.

A Figura 8 apresenta um modelo simples para explicar o comportamento sazonal dos preços de *commodities* agropecuárias. A variação sazonal dos preços, nesse caso, resulta da sazonalidade nas condições de oferta do produto em questão. De acordo com o discutido anteriormente, tem-se que os preços das *commodities* apresentarão comportamento sazonal diretamente ligado: aos períodos de safra e entressafra; à possibilidade de estocagem (para o caso de produtos não-perecíveis); e aos custos desta estocagem.

O comportamento dos preços tenderá, a princípio, a evoluir de forma cíclica ao longo dos anos, e a intensidade desse comportamento dependerá dessas variáveis, bem como de outras também importantes, como:

- volume realizado das safras;
- previsões feitas pelos agentes, envolvidos nos mercados dessas *commodities*, para as safras posteriores ;
- evolução dos estoques.



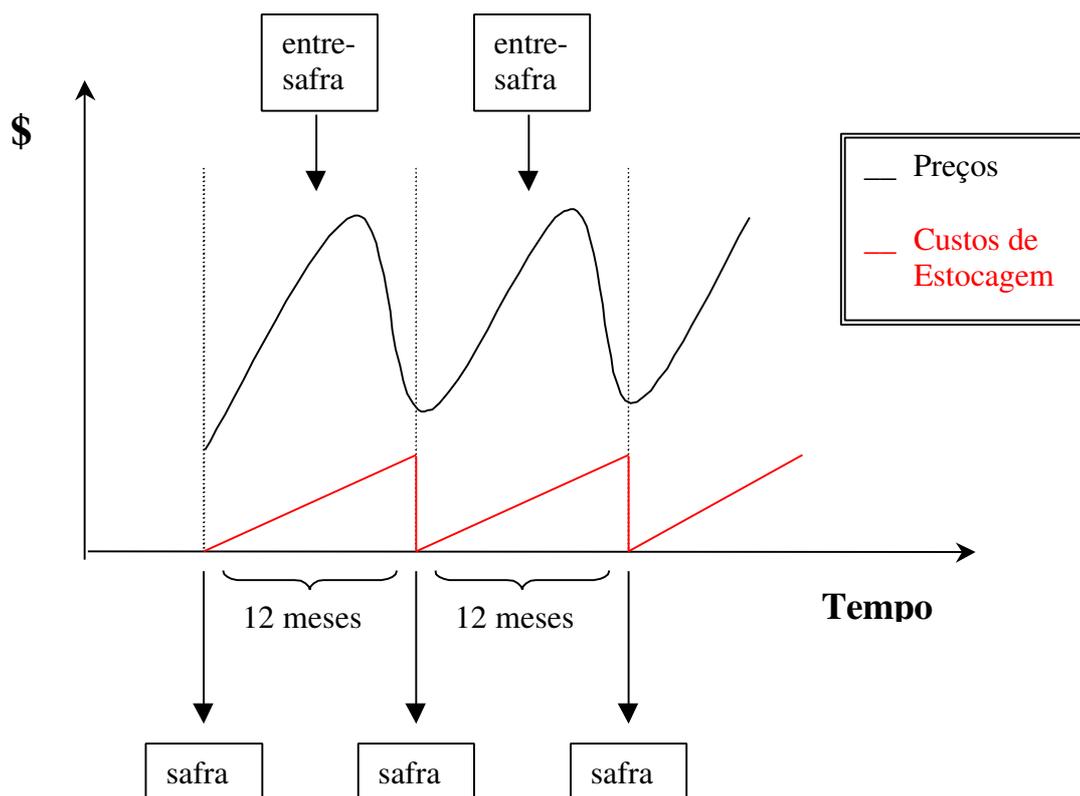
Fonte: TOMEK e ROBINSON (1991).

Figura 8 - Um modelo de oferta e preços sazonais.

Essas relações podem ser mais bem compreendidas com a ajuda do esquema apresentado na Figura 9.

2.1.2.2. Volatilidade

Além das oscilações devidas a fatores sazonais que podem ocorrer nos preços de uma *commodity* ao longo de um ano, pode-se observar que estes também poderão apresentar expressivas oscilações de mês a mês, semanais e, ou, mesmo em intervalos menores de tempo, provocadas por outros fatores determinantes. Esse movimento de curto prazo geralmente é chamado de *Volatilidade*, ou de *Variações Irregulares*, nos preços.



Fonte: Adaptado de TOMEK e ROBINSON (1991).

Figura 9 - Comportamento teórico dos preços e custos de estocagem para *commodities* agropecuárias.

THOMSEN e FOOTE (1952:109-110) argumentam que esse tipo de variação de curto prazo nos preços das *commodities* pode ser causado, entre outros fatores, principalmente pela "experimentação" envolvida no processo de descobrimento das condições de oferta e demanda da referida *commodity*. Segundo suas próprias palavras:

"... a descoberta da real oferta e demanda por uma commodity é uma tarefa árdua. O algodão, por exemplo, é produzido e consumido no mundo todo.(...) Erros de percepção podem ser cometidos, novas informações estão constantemente sendo obtidas e as flutuações de curto prazo nos preços são o resultado. O mesmo ocorre no processo de descobrimento e ajustamento dos preços e quantidades da maioria das outras commodities, ainda que em maior ou menor extensão. Em um leilão de frutas, por exemplo, os primeiros compradores poderão achar que as suas ofertas foram muito altas, enquanto que aqueles que não compraram poderão achar que esperaram tempo demais.

Como resultado os preços dos produtos nos leilões poderão flutuar consideravelmente, mesmo ao longo de um único dia..."

Conforme indicado indiretamente no trecho anterior, esse tipo de movimento oscilatório pode ser explicado com a ajuda da figura do *leiloeiro walrasiano*⁹. Esse leiloeiro agiria nos diversos mercados como um fixador dos preços.

Tem-se como hipótese que os agentes são tomadores de preços e interagem com o leiloeiro da seguinte forma, conforme apontam EATON e EATON (1999): supõe-se inicialmente que o leiloeiro anuncie um dado preço para a *commodity* em questão. Cada comprador escreveria então, em um pedaço de papel, X se estivesse disposto a comprar a *commodity* àquele preço e 0 se não estivesse. Por outro lado, os vendedores da *commodity* escreveriam Y , ao mesmo tempo, se estivessem dispostos a vender por aquele dado preço e 0 se não estivessem. O leiloeiro coletaria então o total de papéis e compararia o número total de X e Y . Se a quantidade demandada pelo preço anunciado exceder a quantidade ofertada, ou seja, se houver excesso de demanda, o leiloeiro irá anunciar um novo preço e repetirá o leilão, porém se a quantidade demandada e ofertada àquele preço forem iguais, o mercado estará em equilíbrio e o leiloeiro irá recolher o dinheiro de cada consumidor disposto a pagar aquele preço e o repassará para os vendedores das *commodities*, que deverão entregá-las.

A repetição desse processo se dará constantemente ao longo do tempo nos mercados à vista (*spot*) das *commodities*, sendo que os preços de equilíbrio em cada leilão não necessariamente serão iguais entre si, pois dependerão da interação dos agentes em cada momento. Isso, em última instância, implicaria o movimento de volatilidade dos preços.

KREPS (1990) formaliza um pouco mais o mecanismo descrito acima. Segundo ele, o leiloeiro apresentaria para a população de uma economia um *vetor de preços* P . Cada indivíduo, baseando-se nas suas preferências,

⁹ Esse termo foi estabelecido na teoria econômica em homenagem ao economista do século XIX Leon Walras.

determinaria o quanto desejaria comprar ou vender àquele patamar de preços. Assim, tem-se que a *troca líquida* entre os agentes seria dada por:

$$Z^i(P) = x^i(P) - e^i \quad (5)$$

em que x representa a demanda do agente i pela *commodity* ao preço p e e representa a *dotação inicial* de cada indivíduo. Cada agente iria reportar ao leiloeiro qual o montante de troca líquida desejada por ele e o leiloeiro iria agregar esses valores. Se o resultado final for:

$$\sum_{i=1}^I Z^i(P) \leq 0 \quad (6)$$

as transações líquidas com a *commodity* seriam efetuadas. Contudo, se o mercado não se ajustar ao nível de preços P , caso em que:

$$\sum_{i=1}^I Z^i(P) > 0 \quad (7)$$

o leiloeiro poderia tentar um novo vetor de preços: P' , por exemplo, e assim por diante. Novamente, tem-se que a repetição sistemática desse processo resultaria no comportamento volátil dos preços da *commodity*.

O modelo anterior, apesar de um tanto quanto simplificador do real processo que ocorre nas interações entre ofertantes e compradores de uma *commodity* no curto prazo, ilustra de maneira satisfatória a essência do mecanismo gerador da volatilidade nos preços dessa *commodity* em períodos menores de tempo. No entanto, outros fatores podem ser levantados como relevantes para a explicação das variações nos preços no curto prazo, como: choques em variáveis econômicas correlacionadas com este produto; fluxos de novas informações acerca do comportamento de preços e quantidades praticados em outros mercados; e variações nas expectativas dos agentes.

2.2. Modelo analítico e procedimentos

2.2.1. Análise do componente de tendência

O primeiro passo na análise da tendência nos preços das *commodities*, que serão aqui denotados genericamente pela variável y_t , é dado verificando-se a *estacionariedade* das séries em questão. Um conjunto composto por vetores aleatórios M-dimensionais $\dots y_{t-1}, y_t, y_{t+1} \dots$ é chamado de *processo estocástico vetorial*. De acordo com JUDGE et al. (1988), esse processo será dito *estacionário* se atender às seguintes condições:

- Todos os vetores aleatórios possuem o mesmo vetor médio μ , ou seja:

$$E[y_t] = \mu \quad \forall t \quad (8)$$

- As variâncias de todas as variáveis aleatórias envolvidas forem finitas e constantes:

$$Var(y_{mt}) < \infty \quad \text{para } m = 1, 2, \dots, M \text{ e } \forall t; \quad (9)$$

- As matrizes de covariância dos vetores y_t e y_{t+k} , que são distantes k períodos, não dependerem de t , mas apenas de k , isto é:

$$\text{cov}(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)] = \Gamma_k \quad \forall t \quad (10)$$

Essas condições indicam, em termos práticos, que as séries temporais (y_t) terão média e variância finitas e serão consideradas estacionárias em covariância. Se esse não for o caso, ou seja, se as séries em questão não forem estacionárias, os métodos a serem empregados para a análise destas irão diferir significativamente e, conseqüentemente, deverão ser considerados, a fim de evitar o problema econométrico de *regressões espúrias*¹⁰. Assim, testes para verificação da estacionariedade ou não das séries (chamados de *testes de raízes unitárias*¹¹) deverão, e serão, implementados com o objetivo de verificar a não-estacionariedade da série devido à presença de tendência nesta.

¹⁰ Esse problema será analisado posteriormente.

¹¹ A forma e os procedimentos desses testes serão apresentados posteriormente.

De acordo com o discutido anteriormente, e que também vai de encontro ao proposto por CHATFIELD (1996) e VANDAELE (1983), a melhor definição, apesar de sua simplicidade, para a *tendência* em uma série de tempo é aquela que estabelece como tendência "... *qualquer mudança sistemática de longo prazo no nível médio de uma série de tempo*". Entretanto, essa mudança pode ser provocada por dois tipos de fenômenos, que apresentam características particulares. Esses fenômenos originam dois conceitos distintos de tendências, que podem estar presentes em um modelo de análise de uma série de tempo, Y_t , por exemplo. De acordo com a sua natureza e o seu padrão comportamental, a tendência poderá ser caracterizada como sendo do tipo *determinística* ou do tipo *estocástica*.

A característica essencial de uma *tendência determinística* se refere ao fato de que a variação no nível médio de uma dada variável se dará, de forma previsível, como uma função do tempo. Ou seja, conforme apontam PINDYCK e RUBINFELD (1998) e BOWERMAN e O'CONNELL (1979), uma série de tempo Y_t poderá evoluir em montantes absolutos em cada período de tempo, de acordo com a seguinte equação:

$$Y_t = T_t + \varepsilon_t \quad (11)$$

em que T_t é o componente de tendência determinística do modelo e ε_t é um termo de erro com média zero e variância igual a σ^2 .

O valor de Y_t poderá ser previsto de acordo com a equação (11), e o componente de tendência (T_t) poderá assumir várias formas, como:

$$T_t = \varphi_0 \quad (12)$$

sendo:

$$Y_t = \varphi_0 + \varepsilon_t \quad (13)$$

Nesse caso, que é o mais simples, não existe a tendência, ou seja, a série de tempo Y_t estará flutuando aleatoriamente em torno de um nível médio (φ_0) ao longo do tempo. Uma segunda possibilidade se refere ao caso em que a tendência determinística assume uma *forma linear*, dada por:

$$T_t = \varphi_0 + \varphi_1 t \quad (14)$$

de maneira que:

$$Y_t = \varphi_0 + \varphi_1 t + \varepsilon_t \quad (15)$$

Nesse caso, a série Y_t evoluirá ao longo do tempo em torno de um nível médio que se modifica de maneira linear. Conforme aponta ENDERS (1995), a tendência determinística poderá ser do tipo polinomial, que, de maneira geral, pode ser especificada como:

$$T_t = \varphi_0 + \varphi_1 t + \varphi_2 t^2 + \dots + \varphi_k t^k \quad (16)$$

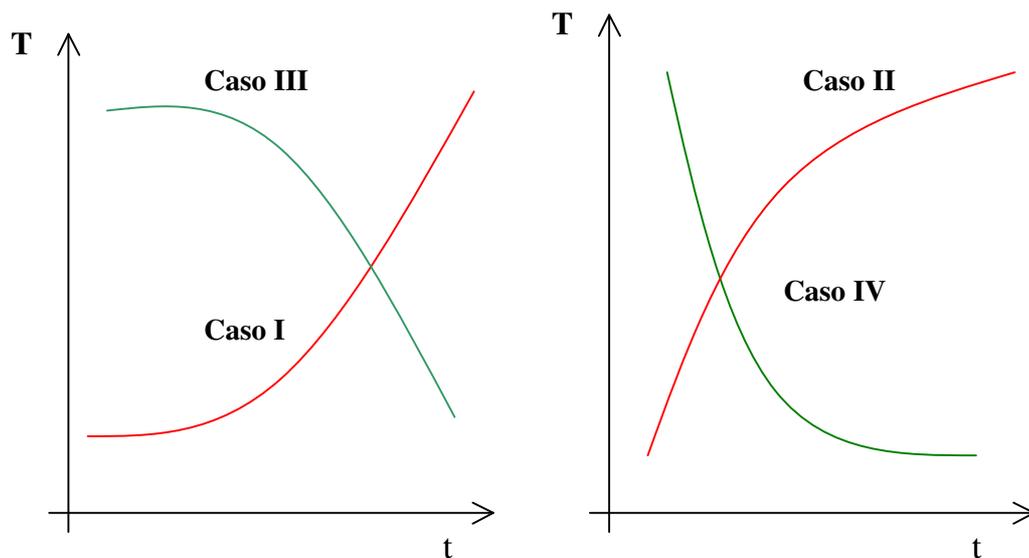
em que k é o grau do polinômio em questão, e assim:

$$Y_t = \varphi_0 + \varphi_1 t + \varphi_2 t^2 + \dots + \varphi_k t^k + \varepsilon_t \quad (17)$$

Nesse tipo de modelo, a média da série evoluirá de forma curvilínea ao longo do tempo, e uma gama enorme de possibilidades poderá ser verificada, de acordo com os sinais dos coeficientes (φ_i) do polinômio. Por exemplo, conforme demonstram BOWERMAN e O'CONNELL (1979), para o caso de uma tendência polinomial de segundo grau (quadrática), do tipo:

$$T_t = \varphi_0 + \varphi_1 t + \varphi_2 t^2 \quad (18)$$

tem-se que, com a ajuda da Figura 10, pode-se "esquematizar" algumas possibilidades, como visto a seguir.



Fonte: BOWERMAN e O'CONNELL (1979).

Figura 10 - Variação do nível médio de uma série provocada por uma tendência quadrática.

No caso I, tem-se que, de acordo com a tendência determinística, o nível médio da série estará crescendo ao longo do tempo a uma taxa crescente, ao passo que no caso II esse crescimento se dará a taxas decrescentes. No caso III, a tendência implica que o nível médio estará diminuindo a uma taxa crescente com o passar do tempo, enquanto no caso IV o nível médio também estará caindo, porém a uma taxa decrescente.

A tendência determinística de uma série também pode assumir formas mais complexas, como a do *modelo de crescimento exponencial*:

$$Y_t = A \cdot e^{\delta t} \varepsilon_t \quad (19)$$

que pode ser estimado após se tomar o logaritmo de ambos os lados de (19), que resulta:

$$\ln Y_t = \varphi_0 + \varphi_1 t + \mu_t \quad (20)$$

em que $\varphi_0 = \ln A$, $\varphi_1 = \delta$, que é a *taxa de crescimento* da série, e $u_t = \ln \varepsilon_t$.

Conforme discutido anteriormente, e também apontado por ENDERS (1995:182), a justificativa teórica (econômica) para a inclusão de um componente de tendência determinística em um modelo se reporta ao fato de que as variáveis econômicas, no caso os preços de *commodities*, estão sujeitas a impactos de mudanças no ambiente econômico (gostos, tecnologias...) que irão levar a um comportamento relativamente previsível ao longo do tempo. De acordo com as suas próprias palavras:

"... existe uma crença difundida que no longo prazo as variáveis macroeconômicas irão crescer a uma taxa constante, sendo que qualquer desvio dessa tendência será eventualmente eliminado pela "mão invisível". A crença de que a tendência não está se modificando ao longo do tempo implica na prática comum de modelar (eliminar) a tendência dos dados macroeconômicos usando uma equação de regressão linear, ou polinomial, determinística."

Contudo, o problema com esse tipo de análise e procedimento é que nem sempre a tendência nos dados em questão é do tipo determinística. Ao contrário, ela pode estar mudando de forma aleatória ao longo do tempo, caracterizando-se como um componente de tendência do tipo estocástica.

Conforme a definição dada por STOCK e WATSON (1988), o conceito de *tendência estocástica* é uma extensão direta do conceito de tendência determinística (que se refere a variações percentuais fixas ao longo de determinado período de tempo). A diferença principal entre essas duas formas de tendência se refere ao fato de que a tendência estocástica implica uma variação percentual média na série em dado período de tempo, porém, ao contrário da determinística, em cada período a mudança provocada pela tendência em relação ao seu nível médio será um montante aleatório e imprevisível, em vez de constante, dado por determinada taxa. Ou seja, no caso da existência de uma tendência do tipo estocástica, tem-se que os coeficientes (φ_k) de uma equação para a descrição da tendência, como os apresentados na equação (17), por exemplo, poderão não ser constantes, mas também variar aleatoriamente ao longo do tempo. Assim, a tendência oscilará de forma aleatória à medida que o tempo evolua e os choques exógenos entrem no sistema.

Existem diversos modelos de séries temporais que incorporam tendências estocásticas e que particularmente são importantes na análise de variáveis do mercado financeiro, como o *retorno de ações* e os *preços de commodities*. Dentro dessa classe de modelos, destaca-se o *modelo de passeio aleatório (random walk)*, dado por:

$$Y_t = Y_{t-1} + \varepsilon_t \quad (21)$$

ou:

$$\Delta Y_t = \varepsilon_t \quad (22)$$

em que ε_t é um erro do tipo *ruído branco*, com média 0 e variância σ^2 , e Δ é um *operador de diferenças finitas de primeira ordem*.

O modelo de passeio aleatório é claramente um caso especial de um processo auto-regressivo de ordem 1 [AR(1)], dado por:

$$Y_t = \gamma_0 + \gamma_1 Y_{t-1} + \varepsilon_t \quad (23)$$

em que $\gamma_0 = 0$ e $\gamma_1 = 1$.

A variância da série Y_t em questão será dependente do tempo, pois os choques exógenos (ε_t) terão efeitos cumulativos com o passar do tempo sobre a série. Assim, o processo passeio aleatório será não-estacionário, pois $t \rightarrow \infty \Rightarrow \text{Var}(Y_t) \rightarrow \infty$.

Conforme aponta ENDERS (1995), tem-se como implicação do modelo de passeio aleatório o fato de que, supondo que se tenha à mão uma amostra de valores da série Y_t e se deseje prever os valores futuros para esta, a melhor previsão de Y_{t+s} que se pode fazer será dada por:

$$E_t(Y_{t+s}) = Y_t \quad (24)$$

ou seja, o valor constante de Y_t será o melhor estimador não-viesado para todos os valores futuros de Y_{t+s} ($s > 0$). Assim, pode-se notar que um choque ε_t terá um

efeito permanente em Y_t e o multiplicador de impactos de ε_t em Y_t (dado por $\frac{\partial Y_t}{\partial \varepsilon_t}$) será o mesmo multiplicador de ε_t em Y_{t+s} . Essa permanência dos efeitos implica que a seqüência Y_t possui tendência estocástica, dada pela expressão $\sum \varepsilon_i$, que impõe mudanças aleatórias no nível médio da série com o passar do tempo.

Um outro modelo que incorpora tendência estocástica, e que é bastante comum na literatura acerca de tendências estocásticas¹², é uma variação do modelo de passeio aleatório que incorpora uma parte determinística ao modelo de análise das mudanças na série Y_t . O *modelo de passeio aleatório com intercepto (random walk plus drift)*, que aumenta o modelo de passeio aleatório com a inclusão de um termo constante, é dado por:

$$Y_t = Y_{t-1} + \gamma_0 + \varepsilon_t \quad (25)$$

Assim, dada uma condição inicial Y_0 , a solução geral para a série Y_t será dada por:

$$Y_t = Y_0 + \gamma_0 t + \sum_{i=1}^t \varepsilon_i \quad (26)$$

De acordo com ENDERS (1995), o comportamento da série Y_t , portanto, será governado por dois componentes não-estacionários, quais sejam: uma tendência linear determinística ($\gamma_0 t$) e uma tendência estocástica ($\sum \varepsilon_i$). A variação determinística em cada nova observação de Y_t será dada por γ_0 e após t períodos de tempo será igual a $\gamma_0 \cdot t$, ao passo que a variação provocada pela tendência estocástica será dada pelo somatório dos diversos choques ε_i que terão efeitos permanentes na média de Y_t .

¹² Conforme se pode observar nos trabalhos de ENDERS (1995), HAMILTON (1994) e CAMPBELL et al. (1997).

A presença de tendência do tipo estocástica em uma série temporal implica a possibilidade da ocorrência do problema econométrico de *regressões espúrias*, caso essa série seja utilizada em um modelo de regressão. Esse problema foi primeiramente detectado e analisado por GRANGER e NEWBOLD (1974) e implica que as conclusões obtidas de um modelo que inclua variáveis com esse tipo de tendência poderão ser extremamente errôneas¹³, pois os estimadores de Mínimos Quadrados Ordinários de tal modelo não apresentarão as propriedades desejáveis para um estimador, ou seja, não serão estimadores do tipo *BLUE*.

Se a tendência presente na série for do tipo estocástica e a análise (e possível "remoção") dessa tendência considerá-la, de forma equivocada, como sendo do tipo determinística, alguns problemas econométricos poderão surgir. Por exemplo, a subtração de uma tendência determinística de uma série que apresenta um comportamento de passeio aleatório com intercepto não terá como resultado uma série estacionária, isto é, a subtração de $Y_0 + \gamma_0 t$ de cada observação não eliminará a tendência estocástica da série. Por outro lado, a aplicação de métodos impróprios para a eliminação de uma tendência estocástica de uma série, cuja tendência seja de fato determinística, também poderá provocar problemas de natureza similar.

A discussão anterior aponta, de acordo com o proposto por STOCK e WATSON (1988), que o primeiro passo na *análise de tendência* nos preços das *commodities* será o de realização de testes estatísticos para verificar se de fato existe tendência; e o segundo, identificar o tipo dessa tendência, se determinística, estocástica ou uma composição das duas, como no modelo de passeio aleatório com intercepto.

Na busca da determinação da presença de um componente de tendência em uma série temporal, primeiramente as observações dos preços de café serão *plotadas* em um gráfico. Os dados empíricos serão dispostos em um *diagrama de dispersão* e buscar-se-á assim identificar "visualmente" se a série apresenta

¹³ Para maiores informações sobre o problema de *regressões espúrias*, ver o artigo original de GRANGER e NEWBOLD (1974) e também GRIFFITHS et al. (1993).

algum tipo de tendência. Apesar do pouco rigor científico em tal procedimento, este é indicado por diversos autores, como CHATFIELD (1996) e VANDAELE (1983), como um procedimento básico e essencial na análise de séries de tempo. Esse procedimento é indicado não só para a identificação do componente de tendência, mas também para a análise de ciclos, sazonalidade e volatilidade.

De acordo com ENDERS (1995), outro procedimento fundamental na detecção do componente de tendência em uma série se refere ao estudo do comportamento da função de autocorrelação da série, dada pela equação a seguir:

$$\rho_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (27)$$

Esta função possibilita o cálculo de valores denominados coeficientes de autocorrelação, os quais mensuram a correlação existente entre os preços em diferentes momentos no tempo. Estes coeficientes também fornecem informações que possibilitam a descoberta do modelo probabilístico que possa ter gerado essa série de tempo.

O gráfico dos coeficientes de autocorrelação amostral, também chamado de correlograma, em que os valores das autocorrelações estimadas ($\hat{\rho}_k$) são plotados em relação às suas diferentes defasagens (k), se constitui num dos principais instrumentos de auxílio para a interpretação desses coeficientes. Se a variável analisada apresentar o componente de tendência os valores das autocorrelações não irão convergir para zero de maneira rápida, apenas após um número relativamente elevado de defasagens (12, por exemplo). Isto ocorrerá, pois as observações de um mesmo lado da média geral dos preços tenderão a ser seguidas por grande número de observações do mesmo lado da média em função do componente de tendência.

Conjuntamente com a análise do gráfico da série e de sua função de autocorrelação, serão implementados testes mais rigorosos, do ponto de vista estatístico, para a verificação da possibilidade de não-estacionariedade da série

(presença de tendência), assim como para identificar se a tendência é do tipo determinística ou estocástica. Um dos testes que serão empregados é o conhecido como *Teste de Dickey-Fuller Expandido* (teste ADF), que possui a forma geral¹⁴:

$$Y_t - Y_{t-1} = \Delta Y_t = \mu + \beta t + (\theta - 1) Y_{t-1} + \sum_{k=1}^{p-1} \phi_k \Delta Y_{t-k} + \varepsilon_t \quad (28)$$

em que ΔY_t é um *operador de diferenças*, μ e β se referem à presença de elementos determinísticos (intercepto e tendência linear¹⁵) no modelo, θ se refere à presença de raiz unitária no processo Y_t e p representa o número de *lags* (defasagens) incluídos na equação, conforme indicam GRIFFITHS et al. (1993), para incorporar a possibilidade de que Y_t siga um processo auto-regressivo de ordem superior. Se esses termos não forem necessários, eles deverão ser excluídos da equação, o que não causa nenhum prejuízo ao teste em questão.

O teste ADF consiste então em testar a hipótese conjunta $\mu = \beta = (\theta - 1) = 0$. Se $(\theta - 1) = 0$, a série possuirá raiz unitária e será considerada como um processo do tipo *passeio aleatório*; assim, diz-se que Y_t possui apenas o componente de tendência do tipo estocástica. Se $\mu \neq 0$, $\beta = 0$ e $(\theta - 1) = 0$, tem-se que a série apresenta raiz unitária e possui o componente determinístico de intercepto (*drift*), mas não possui tendência linear determinística, será então um processo do tipo *passeio aleatório com intercepto*; assim, Y_t possui o componente de tendência do tipo estocástica. Se $\mu \neq 0$, $\beta \neq 0$ e $(\theta - 1) = 0$, tem-se que a série apresenta raiz unitária, possui o componente determinístico de intercepto (*drift*) e possui também a tendência linear determinística, será então um processo do tipo *passeio aleatório com tendência linear e intercepto*. Os resultados anteriores indicam que a série Y_t é um *processo estacionário por diferença* (DSP)¹⁶.

¹⁴ De acordo com GREENE (1997).

¹⁵ A forma funcional do componente de tendência determinística poderá diferir, incorporando termos polinomiais, conforme apresentado anteriormente.

¹⁶ Contudo, nos casos em que a tendência linear for significativa, apenas a tomada da diferença da série poderá não torná-la estacionária.

Para os casos em que $(\theta - 1) \neq 0$, tem-se que a série não possui raiz unitária e, portanto, não é do tipo estacionária por diferença (DSP). Assim, a tendência nesta, caso exista, será exclusivamente do tipo determinística. Nesse caso, o processo em questão será dito um *processo estacionário em tendência* (TSP) e, para o caso de tendência linear, possuirá a forma genérica:

$$Y_t = \mu + \beta t + \varepsilon_t \quad (29)$$

Essa especificação poderá ser modificada para incorporar termos polinomiais de tendência. De acordo com GRIFFITHS et al. (1993), estimando (29) pelo *Método dos Mínimos Quadrados Ordinários* (MQO), tem-se que:

$$\hat{\varepsilon}_t = Y_t - \hat{\mu} - \hat{\beta}_t \quad (30)$$

formará uma série livre do componente determinístico, que poderá ser utilizada em um modelo de regressão, bem como para outras análises que necessitem de séries estacionárias.

Para verificação da significância dos coeficientes estimados em (28), podem-se verificar as tabelas estatísticas elaboradas por FULLER (1976) e DICKEY (1976), dado que as estatísticas dadas pela tabela t de Student não podem ser utilizadas para esse fim. Mais recentemente, MacKINNON (1991)¹⁷ implementou um conjunto de replicações maior que o dado pelas tabelas de Dickey-Fuller. Esse trabalho tornou possível o cálculo de valores críticos de Dickey-Fuller para amostras de qualquer tamanho e qualquer escolha de variáveis independentes, simplificando assim os testes de significância das estimativas dadas em (28). Dessa forma, serão utilizados neste estudo os valores tabelados por MacKINNON (1991).

O Teste de Dickey-Fuller Expandido (ADF) nem sempre é o mais indicado para a verificação da presença de raízes unitárias, pois se baseia no pressuposto de que os erros sejam não correlacionados e apresentam variância

¹⁷ Conforme QMS (1995).

uniforme, ou seja, pressupõe que $\varepsilon_t \sim N(0, \sigma^2)$. Se esse não for o caso, tem-se como alternativa o *Teste de Raízes Unitárias de Phillips-Perron*, que também pode ser utilizado para a verificação da estacionariedade das séries e a eventual necessidade de diferenciação destas. Este teste se baseia na seguinte equação (QMS, 1995):

$$\Delta Y_t = \mu + \rho Y_{t-1} + \varepsilon_t \quad (31)$$

que pode ser estimada por MQO. A hipótese nula a ser testada é:

$$H_0: \rho = 1 \quad (32)$$

ou seja, a hipótese nula é de que a série Y_t segue um processo do tipo passeio aleatório.

Ao contrário do teste ADF, o teste de *Phillips-Perron* não inclui termos de diferenças defasadas, mas pode também incorporar termos de tendência determinística e um intercepto. No entanto, deve-se especificar um termo chamado de *truncation lag* (defasagem de truncamento), que é relacionado ao número de períodos de correlações seriais que serão incluídas no modelo¹⁸.

Portanto, se for verificado que os erros são correlacionados e, ou, apresentam heterocedasticidade, o teste a ser empregado neste trabalho será o de *Phillips-Perron*.

¹⁸ Para maiores detalhes acerca dos procedimentos desse teste, ver HAMILTON (1994), ENDERS (1995) e QMS (1995).

2.2.2. Análise do componente cíclico

A análise de séries temporais pode ser implementada segundo diversas abordagens, que nem sempre são concorrentes entre si. Esse é o caso da *análise no domínio da frequência* (ou *Análise Espectral*), que representa uma forma de análise de séries temporais que fornece informações complementares àquelas propiciadas pela *análise no domínio do tempo*, informações estas que geralmente estão ligadas ao estudo da ocorrência de ciclos nas séries.

A essência da análise espectral reside na representação de uma série temporal $\{Y_t\}_{t=-\infty}^{\infty}$ como uma soma ponderada de funções periódicas do tipo seno (ωt) e cosseno (ωt), em que ω representa uma frequência particular; e na determinação de quão importantes são os ciclos de diferentes frequências, caso eles existam, para explicar a variância de $\{Y_t\}_{t=-\infty}^{\infty}$ e, conseqüentemente, o comportamento dessa série. Daí o porquê de esse tipo de análise ser denominada análise no domínio da frequência, ao contrário dos modelos no domínio do tempo, que tem como dados os valores observados da série dispostos ao longo do tempo. Contudo, a base teórica da metodologia da análise espectral é um tanto quanto complexa, envolvendo alguns passos que serão apresentados a seguir.

Para o melhor entendimento de como é desenvolvida a análise espectral, deve-se ter como ponto de partida um modelo conhecido como *Modelo Senoidal Simples*, que representa uma série de tempo, $\{Y_t\}_{t=-\infty}^{\infty}$ por exemplo, que apresente um comportamento periódico, ou seja, as observações $\dots Y_{t-3}, Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, Y_{t+3}, \dots$ oscilam em torno de um valor médio (μ), com uma certa periodicidade. De acordo com CHATFIELD (1996), uma maneira natural de representar o comportamento dessa série seria dada pelo modelo a seguir:

$$Y_t = \mu + R \cos(\omega t + \theta) + Z_t \quad (33)$$

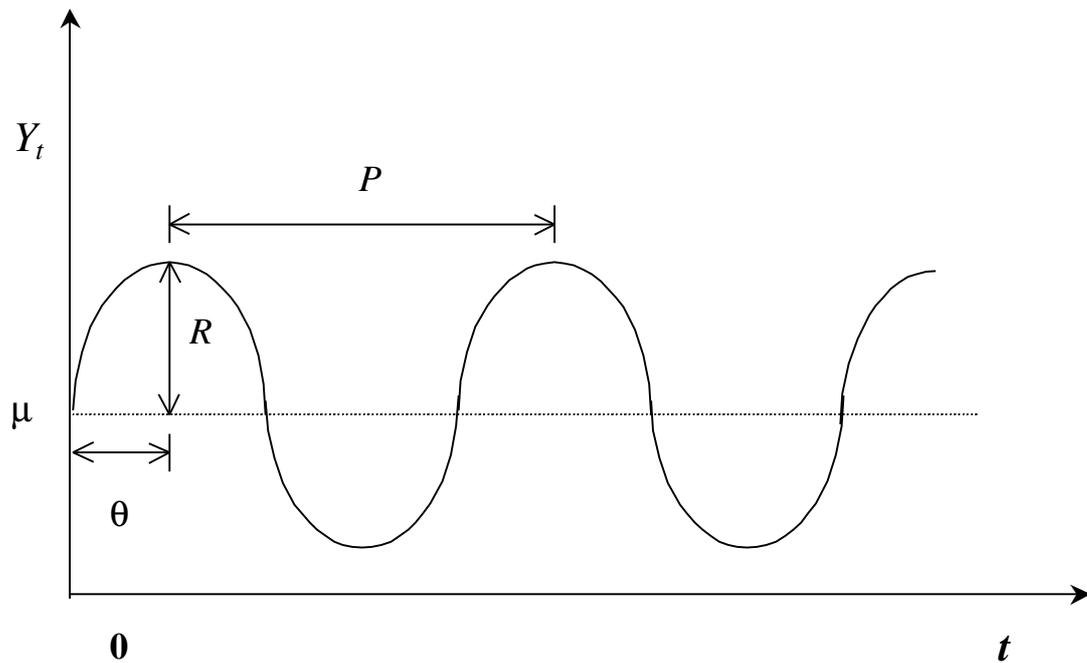
em que μ é um *Valor Médio* qualquer; R é chamada de *Amplitude* e indica que a variável Y_t irá atingir um *ponto de máximo* com $\mu + R$ e um *ponto de mínimo*

com $\mu - R$; ω é chamado de *freqüência angular da variação senoidal* e representa quão rapidamente a variável Y_t oscila ciclicamente no tempo; θ é chamada de *Fase* e indica onde, no ciclo, Y_t estará no ponto $t = 0$, ou seja, é a distância entre a origem da série temporal, no caso $t = 0$, e o pico (ou “fundo do vale”) mais próximo do ciclo; Z_t é um erro do tipo *ruído branco*, com as suas propriedades usuais de média 0 e variância constante; finalmente, pode-se definir o *Período* (P) do ciclo senoidal (também chamado de *tamanho da onda*) como o tempo necessário para a série temporal começar a repetir um ciclo completo, que será igual a $\frac{2\pi}{\omega}$.

Deve-se observar que ω é chamado de freqüência angular, pois sua unidade de medida é dada em termos de número de radianos por unidade de tempo. Contudo, alguns autores, como JENKINS e WATTS (1968), referem-se à *freqüência* (f) como sendo dada por $f = \frac{\omega}{2\pi}$, que representaria o número de ciclos por unidade de tempo, isto é, o número de repetições em um espaço fixo de tempo (uma década, por exemplo). Esse conceito torna a interpretação dos dados e resultados mais direta e também será utilizado neste trabalho. Assim, tem-se que o período é o recíproco da freqüência, ou seja, $P = \frac{1}{f}$.

Esses conceitos podem ser mais bem compreendidos com a ajuda da Figura 11, que representa graficamente o comportamento de uma série de tempo, Y_t , que tem seu comportamento dado pelo modelo senoidal simples. Pode-se observar que os valores da série variam em torno da média μ , e passam a se repetir após um período P de tempo.

Conforme se pode observar, todos os elementos do modelo podem ser interpretados com a ajuda dessa figura. No entanto, o Modelo Senoidal Simples dado pela equação (33) é extremamente simples, como o próprio nome indica, para “explicar” o comportamento de uma série de tempo real (empírica), o que denota a necessidade de um modelo mais abrangente.



Fonte: HAMILTON (1994).

Figura 11 - Gráfico de $Y_t = \mu + R \cos(\omega t + \theta)$.

Conforme apontam PEREIRA et al. (1986) e CHATFIELD (1996), "na prática", a variação em uma série temporal é causada por variações e ciclos em diversas frequências. Por exemplo, dados sobre *vendas* de algum bem econômico podem conter variações cíclicas *semanais*, *mensais* e *anuais*, que representariam variações em *altas*, *médias* e *baixas* frequências, respectivamente.

Deve-se, portanto, generalizar o modelo da equação (33), a fim de se obter um modelo mais abrangente e capaz de melhor determinar o comportamento das séries temporais empíricas. Dessa forma, uma generalização natural seria dada por:

$$Y_t = \mu + \sum_{j=1}^k R_j \cos(\omega_j t + \theta_j) + Z_t \quad (34)$$

em que:

$$t = 1, 2, 3, \dots, T;$$

$$k = 2, 3, 4, \dots, [T/2];$$

e R_j é a amplitude na frequência angular ω_j , que também pode ser apresentada¹⁹ como:

$$Y_t = \mu + \sum_{j=1}^k (\alpha_j \cos \omega_j t + \beta_j \sen \omega_j t) + Z_t \quad (35)$$

em que $\alpha_j = R_j \cdot \cos \theta_j$ e $\beta_j = -R_j \cdot \sen \theta_j$.

Deve-se observar que se R , $\{R_j\}$, θ e $\{\theta_j\}$ nos modelos (33) e (34) forem considerados termos constantes, o processo $Y(t)$ não será estacionário, apresentando raiz unitária, pois $E[Y(t)]$ irá variar com o tempo; assim, tem-se como pressupostos desses modelos, a fim de se aplicar a teoria de processos estacionários, que os termos R e $\{R_j\}$ são considerados variáveis aleatórias não correlacionadas com média zero e que os termos θ e $\{\theta_j\}$ são variáveis aleatórias com distribuição uniforme no intervalo $[0, 2\pi]$, que são fixas para uma única realização do processo. Analogamente, para a equação (35), devem-se considerar os termos α_j e β_j ($j = 1, 2, 3, \dots, k$), como variáveis aleatórias não correlacionadas e que possuem média zero, o que implica que $E[Y(t)] = \mu$ para todo t .

Apesar de representar um avanço em relação ao Modelo Senoidal Simples, dado pela equação (33), o modelo da equação (35) ainda apresenta uma limitação importante; ele considera apenas um número finito de frequências, dado que k é um número finito. Contudo, conforme aponta HAMILTON (1994), existe um teorema conhecido como *Teorema da Representação Espectral*²⁰, segundo o qual qualquer processo estacionário discreto $Y(t)$, com frequências

¹⁹ Ver prova 1 no Apêndice A.

²⁰ Esse teorema não será demonstrado aqui, pois envolve *Cálculo Integral Estocástico*, que é um tópico avançado de análise matemática e foge do escopo do presente trabalho. Para maiores detalhes desse teorema, ver CRAMÉR e LEADBETTER (1967).

pertencentes ao intervalo $[0, \pi]$, pode ser representado por uma generalização da equação (35), com $k \rightarrow \infty$, da seguinte maneira:

$$Y_t = \mu + \int_0^\pi [\alpha(\omega) \cdot \cos(\omega t)] d\omega + \int_0^\pi [\beta(\omega) \cdot \sin(\omega t)] d\omega \quad (36)$$

que é a forma mais comum para expressar a *Representação Espectral do Processo* Y_t .

Da equação (36) surge, portanto, o conceito de que a análise espectral descreve o valor de uma variável Y_t como uma soma ponderada de funções periódicas da forma $\cos(\omega t)$ e $\sin(\omega t)$, em que ω representa uma frequência particular no intervalo $[0, \pi]$ ²¹.

A frequência angular $\omega = \pi$, medida em termos de radianos por unidade de tempo, é conhecida como a *Frequência de Nyquist* e denota o valor mais alto com que se pode obter informações sobre ciclos periódicos em um conjunto de dados (esse conceito será retomado posteriormente). Portanto, tem-se que o ponto principal da introdução da equação (36) se refere ao fato de que ela indica que cada frequência no intervalo $[0, \pi]$ pode contribuir para a variação de uma série de tempo Y_t qualquer.

Assim como na análise de uma série temporal no domínio do tempo, o conhecimento da *Função de Autocovariância*²² é de valor imprescindível para descrever a estrutura de tal processo; na análise no domínio da frequência existe uma função chamada *Função de Densidade Espectral*, ou simplesmente *Espectro Populacional* do processo $\{Y_t\}_{t=-\infty}^\infty$, cuja integral definida em $[0, \pi]$, chamada de *Função de Distribuição Espectral*, é a representação espectral da função de autocovariância. Esse conceito também se configura num desenvolvimento teórico de grande importância para a compreensão da estrutura variacional do

²¹ O intervalo de integração é $[0, \pi]$, e não $[0, \infty]$, devido à seguinte característica das funções trigonométricas do tipo coseno: $\cos[(\omega + n\pi)t] = \{\cos \omega t \ \forall n \text{ e } t \text{ inteiros, com } n \text{ par}\}$ e $\cos[(\omega + n\pi)t] = \{\cos(\pi - \omega)t \ \forall n \text{ e } t \text{ inteiros, com } n \text{ impar}\}$. Assim, variações em frequências maiores do que π não poderão ser distinguidas de variações nas frequências correspondentes no intervalo $[0, \pi]$, pois estas irão se sobrepor.

²² Dada por $\gamma(k) = E\{[Y(t) - \mu][Y(t+k) - \mu]\} = \text{Cov}[Y_t, Y_{t+k}]$.

processo. O *Espectro Populacional* do processo $\{Y_t\}_{t=-\infty}^{\infty}$ pode ser escrito como²³:

$$f_Y(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(\omega j) \right] \quad (37)$$

Conforme aponta HAMILTON (1994), como o espectro representa uma função contínua de ω , definida no conjunto dos números reais, e γ_j representa as autocovariâncias de um processo estacionário, tem-se que $f_Y(\omega)$ terá um valor positivo para qualquer ω . Além disso, como o $\cos(\omega j) = \cos(-\omega j) = \cos[(\omega + 2\pi n)j] \forall n$ e j inteiros, o espectro será simétrico em torno da frequência $\omega = 0$ e será uma função periódica de ω ; o que implica que $f_Y(\omega + 2\pi n) = f_Y(\omega)$. Dessa forma, o conhecimento de $f_Y(\omega)$ para todos os valores de ω entre 0 e π implica a possibilidade de inferência sobre o valor de $f_Y(\omega)$ para qualquer ω .

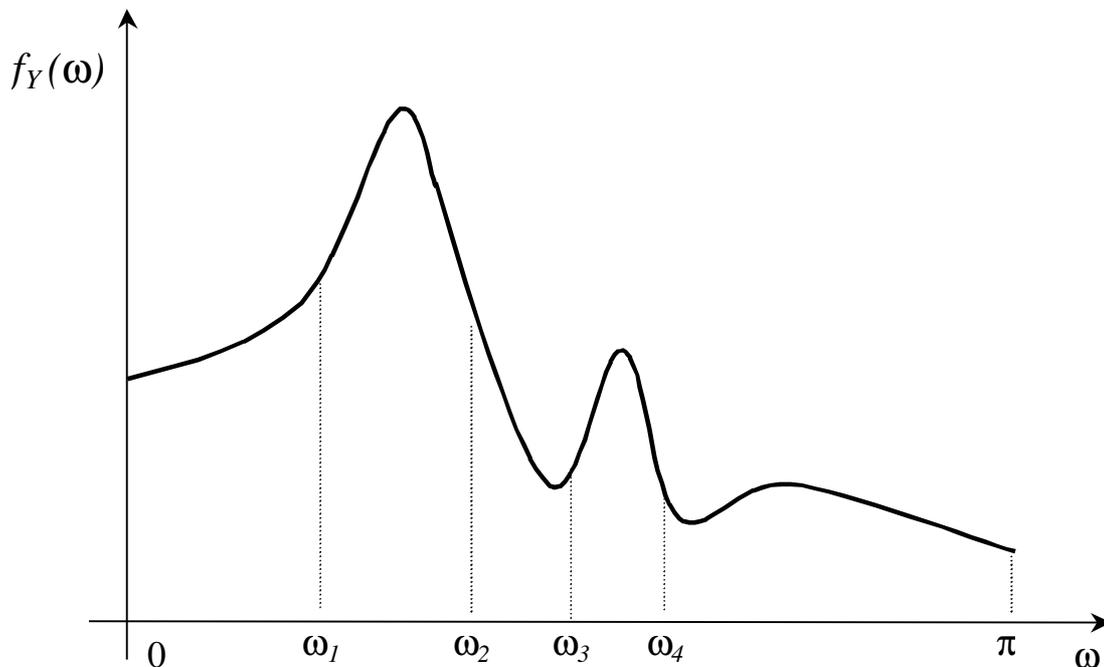
Assim como ocorre na análise no domínio do tempo, em que a importância prática da função de autocorrelação não se dá com a interpretação direta dos seus coeficientes, mas sim com a análise do *correlograma* da série, na análise espectral os coeficientes do espectro não representam o principal ponto de interesse para análise. É o *gráfico do espectro*, que também é chamado de espectro, que fornecerá as informações relevantes acerca da *variância* da série Y_t e sobre quais as frequências que contribuem para um comportamento cíclico dessa série. Portanto, será o principal objeto de análise neste trabalho para a verificação da ocorrência de ciclos nos preços das *commodities*.

Um exemplo de um espectro populacional pode ser observado na Figura 12. Conforme aponta KENNEDY (1996), o espectro na forma gráfica será dado pela representação dos valores da função $f_Y(\omega)$ no eixo vertical, contra os valores das frequências ω representadas no eixo horizontal. A área total sob a curva na Figura 12 é igual à variância do processo²⁴, e "picos" nessa curva indicam uma

²³ Prova 2 do Apêndice A.

²⁴ Prova 3 no Apêndice A.

importante contribuição para a variância das frequências pertencentes ao intervalo em que os picos se situam.



Fonte: KENNEDY (1996).

Figura 12 - Exemplo de um espectro.

Nesse caso, tem-se que a função $f_Y(\omega)$ apresenta valores elevados entre os intervalos de frequência dados por $[\omega_1, \omega_2]$ e $[\omega_3, \omega_4]$, o que indica que, do número infinito de cossenos adicionados para formar o processo $\{Y_t\}_{t=-\infty}^{\infty}$, aqueles com frequência pertencentes ao intervalo $[\omega_1, \omega_2]$ e ao intervalo $[\omega_3, \omega_4]$ possuem amplitude relativamente maior que os demais e, conseqüentemente, têm maior participação na explicação da variância da série de tempo Y_t . Se o gráfico da função $f_Y(\omega)$ fosse plano, ou seja, não apresentasse picos, poder-se-ia concluir

que a série de tempo analisada não possui variações cíclicas, podendo ser caracterizada como um ruído branco.

Conforme ressaltam FRANÇA (1990) e CHATFIELD (1996), a função de autocovariância e a função de densidade espectral representam formas equivalentes para descrever um mesmo processo estocástico estacionário. Ambas contêm a mesma informação, embora expressas de formas distintas, e devem ser utilizadas de acordo com o objetivo da análise. Como um dos objetivos da análise deste estudo será o de verificar a ocorrência de ciclos nos preços do café, tem-se claramente que serão implementadas análises relativas à função de densidade espectral, dado que essa é uma de suas principais importâncias.

Conforme foi discutido anteriormente, a análise espectral visa descrever o valor de uma variável Y_t como uma soma ponderada de funções periódicas da forma $\cos(\omega t)$ e $\sin(\omega t)$, de modo a determinar quais as frequências ω que apresentam maior contribuição relativa para "explicar" o comportamento de Y_t ; sendo que ω representa frequências particulares no intervalo $[0, \pi]$. Variações em frequências maiores do que π não poderão ser distinguidas de variações nas frequências correspondentes no intervalo $[0, \pi]$, pois estas irão se sobrepor. Dessa forma, uma série temporal pode ser representada pela seguinte modificação da equação (35):

$$Y_t = \mu + \left(\sum_{j=1}^{\infty} \alpha_j \cos \omega_{jt} + \sum_{j=1}^{\infty} \beta_j \sin \omega_{jt} \right) + Z_t \quad (38)$$

PEREIRA et al. (1986) apontam que cada componente é uma função periódica que pode ser identificada com a sua frequência angular ω_j (com $j = 1, 2, 3, \dots, \infty$). Contudo, as séries temporais a serem analisadas empiricamente geralmente são *amostras finitas* da população $\{Y_t\}_{t=-\infty}^{\infty}$; portanto, não se pode estimar, a partir delas, o número infinito de parâmetros da equação (38).

Dadas N observações da série Y_t (em que N é um número par), denotados pelo vetor $Y' = [y_1, y_2, y_3, \dots, y_N]$, pode-se estimar no máximo N parâmetros para a

equação anterior. Isso implica que se deve trabalhar com uma *representação amostral* análoga à equação (38), que é dada por:

$$\hat{Y}_t = \sum_{j=0}^n \alpha_j \cos \omega_j t + \sum_{j=0}^n \beta_j \text{sen} \omega_j t + Z_t \quad (39)$$

em que $t = 1, 2, 3, \dots, N$ e $\mu = \alpha_0$, e o último termo será $\text{sen} \omega_n t = \text{sen} \pi_t = 0$.

Os estimadores dos coeficientes da equação (39) podem ser obtidos pelo *Método dos Mínimos Quadrados Ordinários (MQO)*²⁵. PEREIRA et al. (1986) demonstram que esses estimadores serão dados pelas seguintes equações:

$$\hat{\alpha}_j = \begin{cases} \frac{\sum_{t=1}^N y_t}{N} \quad \forall j = 0 \\ \frac{2 \sum_{t=1}^N y_t \cdot \cos \omega_j t}{N} \quad \forall j = 1, 2, 3, \dots, n-1 \\ \frac{\sum_{t=1}^N y_t \cdot \cos \omega_j t}{N} \quad \forall j = n \end{cases} \quad (40)$$

e,

$$\hat{\beta}_j = \begin{cases} \frac{2 \sum_{t=1}^N y_t \cdot \text{sen} \omega_j t}{N} ; \forall j = 1, 2, 3, \dots, n-1 \\ \frac{\sum_{t=1}^N y_t \cdot \text{sen} \omega_j t}{N} = 0 ; \forall j = 0, n \end{cases} \quad (41)$$

em que $n = \frac{N}{2}$ e $\omega_j = \frac{2\pi j}{N}$.

Os dois termos anteriores, n e ω_j , indicam que, com unidades de tempo constante, *anos* por exemplo, a maior curva senóide periódica (que é a mais lenta) que pode ser observada no processo é aquela com um período de N meses,

²⁵ Prova 4 do Apêndice A.

ou de frequência angular igual a $\frac{2\pi}{N}$. Isso se dá porque essa onda só pode se repetir uma vez em N observações se ela possui período N .

Por outro lado, a menor curva senóide periódica, que é a mais rápida, que pode se observar é aquela cujo período é de dois anos, dado que são necessários pelo menos dois anos para a curva completar o ciclo. Assim, a curva mais rápida que pode ser observada tem frequência angular igual a $\frac{2\pi}{2} = \pi$ radianos por ano.

Essa frequência é chamada de *Frequência Nyquist* e denota o valor mais alto com que se podem obter informações sobre ciclos periódicos em um conjunto de dados. Isso se dá pelo fato de que, para se identificar uma função senoidal, deve-se possuir pelo menos dois pontos: um pico alto e um pico baixo.

No presente caso, em que N é um número par, ou seja, $N = 2n$, a frequência angular da j -ésima onda será $\omega_j = \frac{2\pi j}{N}$, com $j = 0, 1, 2, \dots, n$, em que se destacam os casos: $\omega_0 = 0$, quando $j = 0$, que representa a onda de frequência zero, que é um valor médio; $\omega_1 = \frac{2\pi}{N}$, quando $j = 1$, que representa a onda mais lenta; e $\omega_n = \frac{2\pi n}{N} = \pi$, quando $j = n$, que representa a onda mais rápida que se pode observar na série.

O conceito de espectro populacional (ou função de densidade espectral), definido anteriormente pela equação:

$$f_Y(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(\omega j) \right] \quad (42)$$

em que $\gamma_j = E[(Y_t - \mu)(Y_{t+j} - \mu)]$ e $\mu = E(Y_t)$, conforme o próprio nome ressalta, se refere ao caso em que os dados populacionais do processo $\{Y_t\}_{t=-\infty}^{\infty}$ e dos seus segundos momentos, dados por $\{\gamma_j\}_{j=0}^{\infty}$, estão disponíveis. Contudo, assim como nos modelos de análise estatística e econométrica "tradicionais", geralmente o

que se tem em mão, para a implementação da análise espectral, são amostras finitas dos dados populacionais das séries de tempo a serem analisadas.

Assim, deve-se trabalhar com um conceito análogo ao do espectro populacional para a análise espectral de dados amostrais, que é conhecido na literatura como o *Periodograma Amostral*, ou simplesmente *Periodograma*. Para se chegar ao periodograma amostral, deve-se partir de uma amostra de N observações do processo $\{Y_t\}_{t=-\infty}^{\infty}$, denotados pelo vetor $\mathbf{y} = [y_1, y_2, y_3, \dots, y_N]$. Com essa amostra podem ser obtidas $N-1$ autocovariâncias amostrais, a partir da seguinte expressão:

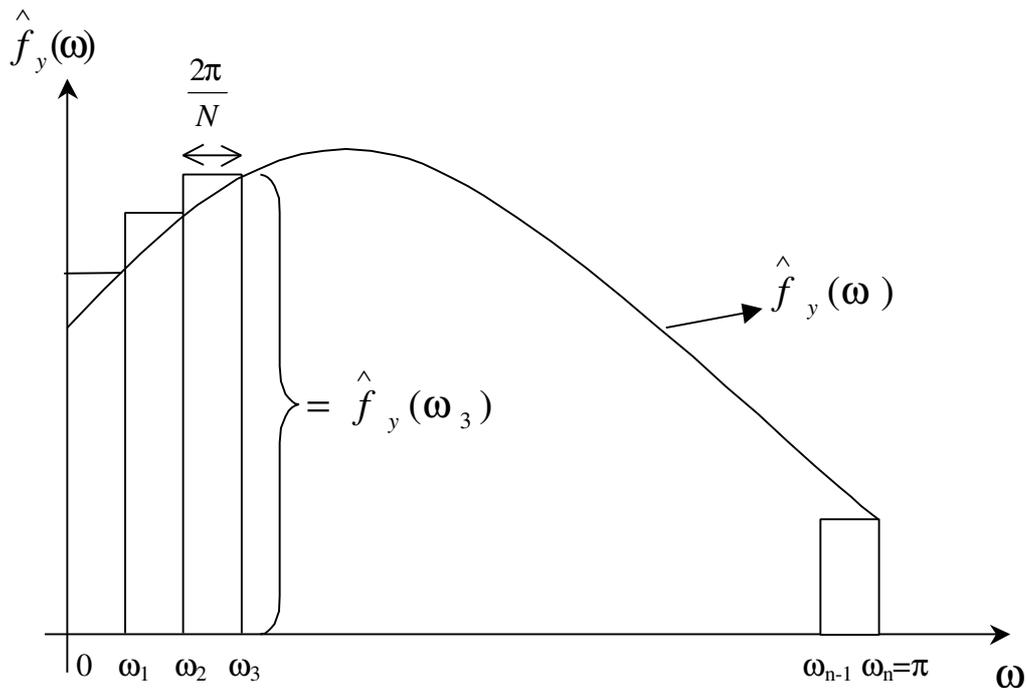
$$\hat{\gamma}_j = c_j = \begin{cases} c_{-j}; \forall j = -1, -2, -3, \dots, -N+1 \\ \frac{1}{N} \sum_{t=j+1}^N (y_t - \bar{y})(y_{t-j} - \bar{y}); \forall j = 0, 1, 2, 3, \dots, N-1 \end{cases} \quad (43)$$

em que $\bar{y} = \frac{\sum_{t=1}^N y_t}{N}$ é a *Média Amostral* de Y_t .

Substituindo os valores estimados de γ_j , dados pelos termos c_j da equação (43), na equação (42), obtém-se o periodograma amostral, denotado por $\hat{f}_y(\omega)$, para um dado valor de ω , que será:

$$\hat{f}_y(\omega) = \frac{1}{2\pi} \left(c_0 + 2 \sum_{j=1}^{N-1} c_j \cdot \cos(\omega j) \right) \quad (44)$$

Assim como foi anteriormente observado que a área abaixo do espectro representava a variância do processo Y_t , tem-se que $\int_{-\pi}^{\pi} \hat{f}_y(\omega) d\omega = c_0 = \hat{\sigma}_Y^2$, ou seja, a área abaixo do periodograma, no intervalo $[-\pi, \pi]$, representará a variância amostral de Y_t . Essas idéias podem ser mais bem compreendidas com a ajuda da Figura 13 a seguir, obtida de HAMILTON (1994), que mostra a área abaixo do periodograma amostral é a porção da variância que pode ser atribuída a ciclos de diferentes frequências.



Fonte: HAMILTON (1994).

Figura 13 - O periodograma e as porções da variância atribuíveis a ciclos de diferentes frequências.

Como se pode observar na Figura 13, o que de fato será calculado com os dados empíricos não é a parcela da variância de Y_t referente a ciclos em uma frequência exata ω_j , mas sim a parcela referente a ciclos de frequência próximos a ω_j , dados por intervalos de frequências $[\omega_{j-1}, \omega_j]$.

De acordo com uma proposição de HAMILTON (1994), pode-se estimar a parcela da variância de Y_t referente a ciclos com frequência entre $[\omega_{j-1}, \omega_j]$ por meio do cálculo da área do retângulo (na Figura 13) com largura igual a $\frac{2\pi}{N}$ e altura igual a $\hat{f}_Y(\omega)$, que será, portanto, igual a $\left(\frac{2\pi}{N}\right) \cdot \hat{f}_y(\omega_j)$. Além disso, pode-

se demonstrar²⁶ que uma fórmula conveniente para o cálculo do periodograma amostral na frequência amostral ω_j será:

$$\hat{f}_y(\omega_j) = \left[\frac{N}{8\pi} \right] (\hat{\alpha}_j^2 + \hat{\beta}_j^2) \quad (45)$$

Ainda que a equação (44) represente um modo direto para a estimação do periodograma amostral, ela deve ser modificada, pois não atende uma das propriedades desejáveis dos estimadores estatísticos.

Conforme demonstram CHATFIELD (1996) e PEREIRA et al. (1986), apesar de o periodograma amostral ser um estimador assintoticamente não-viesado do espectro populacional, a sua variância não diminui à medida que o tamanho da amostra (N) aumenta, ou seja, ele não é um *estimador consistente*. Isso é uma consequência direta do fato de que é estimado um número muito elevado de parâmetros a partir dos dados amostrais. Formalmente, tem-se que:

$$\mathbf{E} \left[\hat{f}_y(\omega) \right] \xrightarrow{N \rightarrow \infty} f_y(\omega) \quad (46)$$

mas,

$$\mathbf{Var} \left[\hat{f}_y(\omega) \right] \xrightarrow{N \rightarrow \infty} \neq 0 \quad (47)$$

Para superar esse problema, diversos métodos foram desenvolvidos, e a maioria deles consiste em processos de "suavização" do periodograma. Essa suavização é obtida tomando-se, para cada frequência, uma média ponderada dos valores do periodograma, com os pesos sendo dados por uma *janela espectral*. O termo janela está ligado ao fato de que a ponderação do sistema por alguns pesos definidos determina qual a parte do periodograma será "observada" pelo estimador.

²⁶ Ver HAMILTON (1994:160-163).

Conforme aponta KENNEDY (1996), a ponderação da janela espectral, que é chamada de *largura da banda* e que determinará o "tamanho" dessa janela, deverá ser escolhida com cuidado, pois uma banda muito larga pode encobrir informações (picos) importantes do espectro, enquanto uma banda muito estreita pode implicar uma variância muito grande, produzindo um espectro estimado muito volátil.

A forma geral dos estimadores consistentes do espectro $f_Y(\omega)$ da equação (37), apontada por JUDGE et al. (1985), é dada pela seguinte expressão:

$$\tilde{f}_y(\omega) = \frac{1}{2\pi} \left(\lambda_0 c_0 + 2 \sum_{j=1}^{M_T} \lambda_j c_j \cos \omega_j \right) \quad (48)$$

em que a seqüência de pesos $\lambda_j = \{\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{M_T}\}$ é chamada de *janela de defasagem* (ou de suavização) e o termo $M_T (< N)$ é chamado de *ponto de truncamento*. Comparando com a equação (42), pode-se observar que os valores estimados das autocovariâncias c_j para $M_T < j < N$ não são utilizados, enquanto os valores de c_j para $j \leq M_T$ são ponderados por um fator λ_j .

A escolha do ponto de truncamento M_T é um tanto quanto problemática, pois a sua determinação rigorosa ainda não foi estabelecida na literatura especializada disponível, e a maioria dos métodos indicados para essa escolha é um tanto quanto subjetiva. Se o valor de M_T for pequeno, a variância de $\tilde{f}_y(\omega)$ será pequena, mas o viés do estimador será maior; em contrapartida, se o valor de M_T for grande, o viés será pequeno, mas a variância de $\tilde{f}_y(\omega)$ será grande, tornando o espectro estimado muito "picotado". Uma regra sugerida por CHATFIELD (1996), e que será a adotada neste trabalho, é a de que se deve escolher um ponto de truncamento (M_T) dado por:

$$M_T = 2 \cdot \sqrt{N} \quad (49)$$

As principais janelas de defasagem utilizadas na suavização do periodograma amostral, e que estão presentes na maioria dos *softwares* que operam com análise espectral²⁷, são as *Janelas de Parzen* e de *Tukey*.

A seqüência de pesos λ_j da *Janela de Parzen* será dada por:

$$\lambda_j = \begin{cases} 1 - 6 \cdot \left(\frac{j}{M_T}\right)^2 + 6 \cdot \left(\frac{j}{M_T}\right)^3; \forall 0 \leq j \leq \frac{M_T}{2} \\ 2 \left(\frac{1-j}{M_T}\right)^3; \forall \frac{M_T}{2} \leq j \leq M_T \end{cases} \quad (50)$$

e a da *Janela de Tukey* será:

$$\lambda_j = \frac{1}{2} \left(1 + \cos \frac{\pi j}{M_T} \right); \forall j = 0, 1, 2, \dots, M_T \quad (51)$$

As janelas de Parzen e de Tukey originarão espectros estimados muito semelhantes, porém a janela de Parzen possui a vantagem de não gerar estimativas negativas (que possuem difícil interpretação), o que pode ocorrer caso seja utilizada a janela de Tukey. Assim, ela será a utilizada neste trabalho.

Outro ponto que será levado em conta na análise espectral a ser implementada se refere ao conceito de *pré-branqueamento* do processo das séries analisadas (Y_t). O pré-branqueamento implica que, para se analisar especificamente os ciclos em uma série temporal, deve-se primeiramente remover o componente de *tendência* da série (que poderá ter sido analisada previamente) e possivelmente subtrair a média desta. Isso se dá porque, para se implementar a análise espectral em uma série de tempo, esta deve ser estacionária; caso contrário, o periodograma e a função de densidade espectral serão "inflados" pelo primeiro coeficiente do coseno (frequência 0,0), que será muito maior que os demais, impossibilitando a análise. Esse problema, que surge caso sejam utilizadas séries de tempo com tendência para estimação do espectro

²⁷ Como o *SPSS* e o *STATISTICA*.

com a janela de Parzen, é denominado *Leakage* (vazamento) e torna o estimador do espectro viesado.

Em alguns casos, também é útil suavizar a série antes da análise, para minimizar as oscilações aleatórias que podem encobrir significativos ciclos periódicos no periodograma. A utilização de dados na forma logarítmica é comum para resolver esse problema.

O tamanho da amostra (N) para a aplicação da análise espectral, a fim de que se tenha razoável estimativa do espectro, deve estar entre $100 < N < 200$, segundo CHATFIELD (1996); no entanto, GRANGER e HUGHES (1968) indicam que $N \approx 100$ observações pode ser um número significativo, principalmente se os dados forem pré-branqueados.

A periodicidade dos dados (semanais, mensais, anuais...), ou *intervalo de frequência*, a serem analisados depende dos objetivos da pesquisa. Como o que de fato se deseja é analisar a possibilidade de ciclos ao longo dos anos nas séries (por exemplo, ciclos de seis em seis anos), as observações deverão ser anuais. Caso se julgar necessária a análise da existência de ciclos sazonais com a análise espectral, o que também é possível, serão trabalhados dados mensais. Um exemplo²⁸ dado por CHATFIELD (1996:108), e também citado por FRANÇA (1990), pode melhor ilustrar essa questão:

"... suponha-se que as leituras de temperatura são tomadas ao meio dia de cada dia em uma certa cidade. É claro que estas observações não nos dirão nada acerca da variação da temperatura dentro do dia. Em particular, não dirão se as noites são mais quentes ou frias que os dias. Somente com uma observação por dia a frequência de Nyquist é $\omega_n = \pi$ radianos por dia, ou $f_n = 1$ ciclo por dois dias. Esta é a mais baixa das frequências que explicariam a variação dentro do dia. Por exemplo, a variação com uma longitude de onda de um dia tem frequência angular $\omega = 2\pi$ radianos por dia, ou $f = 1$ ciclo por dia. Para obter informações acerca da variação da temperatura dentro do dia, dever-se-iam tomar duas ou mais observações por dia."

De acordo com o sugerido por RAUSSER e CARGILL (1970), para verificar se em um espectro estimado existem picos estatisticamente significativos, podem-se empregar dois critérios: o primeiro se refere a uma simples *análise visual* da função de densidade espectral estimada buscando

²⁸ Tradução do autor.

identificar, intuitivamente, se alguns picos se destacam nesse espectro em relação aos demais; o segundo critério, que é cientificamente mais rigoroso, é baseado no *teste da hipótese* (nula) de que o espectro estimado não é estatisticamente diferente do espectro de um *processo puramente aleatório* (ruído branco).

Conforme mencionado anteriormente, o espectro teórico de um processo puramente aleatório será uma linha horizontal, indicando que a contribuição de todas as frequências para a variância do processo será igual e que, portanto, não existem ciclos na série. Assim, se a hipótese nula for rejeitada, pode-se concluir que o espectro estimado apresenta picos relevantes para explicar a variância da série.

Um processo estocástico será dito um *processo puramente aleatório* (*ruído branco*) se ele for composto por uma seqüência de variáveis aleatórias $\{\varepsilon_t\}$ que são mutuamente independentes e identicamente distribuídas, com uma média zero e variância igual a σ^2 , ou seja:

$$Y_t = \varepsilon_t \quad (52)$$

em que ε_t é uma seqüência de variáveis aleatórias e:

$$\begin{aligned} E[\varepsilon_t] &= 0; \\ \text{Var}[\varepsilon_t] &= \sigma^2 \\ E[\varepsilon_t \varepsilon_k] &= 0 \quad \forall t \neq k \end{aligned}$$

A função de autocovariâncias para o ruído branco será:

$$\gamma_j = \begin{cases} \sigma^2, & \forall j = 0 \\ 0, & \forall j \neq 0 \end{cases} \quad (53)$$

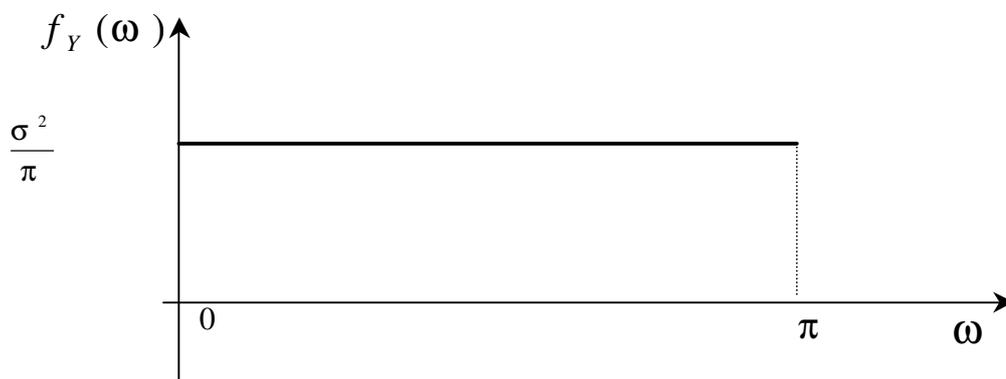
Dessa forma, substituindo esses valores na equação (42), tem-se:

$$f_Y(\omega) = \frac{1}{2\pi} \left[\sigma^2 + 2 \sum_{i=1}^{\infty} 0 \cdot \cos(\omega i) \right] = \frac{\sigma^2}{2\pi} \quad \forall -\pi \leq \omega \leq \pi \quad (54)$$

ou:

$$f_Y(\omega) = \frac{2(Y_0)}{2\pi} = \frac{\sigma^2}{\pi} \quad \forall 0 \leq \omega \leq \pi \quad (55)$$

dado que o espectro é simétrico em torno de $\omega = 0$. Assim, tem-se que o espectro será uma linha horizontal, como se pode observar na Figura 14.



Fonte: RAUSSER e CARGILL (1970).

Figura 14 - Espectro de um ruído branco.

Difícilmente o espectro estimado para uma variável econômica irá apresentar comportamento suave como o da Figura 14. De maneira geral, alguns picos de diferentes amplitudes serão verificados, e a significância deles deverá ser testada.

Se as observações na série analisada (dos preços do café no presente caso) forem independentes entre si, não apresentando repetições cíclicas, os valores calculados em seu periodograma deverão se comportar como um ruído branco e seguir uma *distribuição exponencial*.

Pode-se implementar um teste que compare os valores estimados do periodograma do café com uma distribuição exponencial para indicar se estes

preços apresentam ciclos ou se comportam de maneira similar a uma série do tipo ruído branco, caso em que não se configuram ciclos estatisticamente significativos.

Esse teste permite se encontrar uma *faixa de confiança superior*, que fornece uma linha paralela dada pela distribuição exponencial, indicando que nas frequências com picos que ficarem acima dessa linha ocorrem ciclos estatisticamente significativos, ao passo que os picos de frequências que se encontrarem abaixo desta linha não podem ser considerados significativos.

2.2.3. Análise do componente sazonal

Diversos métodos podem ser empregados para a análise do componente sazonal em uma série de tempo, desde a *análise visual* do gráfico da série, que se apresenta como a mais simples, porém útil, opção, passando pelos *Modelos de Regressão Linear com Variáveis Independentes Binárias (Variáveis Dummy)*, até a utilização de métodos mais sofisticados como a propiciada pelos *Modelos de Análise Espectral* (que trata a sazonalidade como um ciclo de período fixo igual a 12 meses) e pelos *Modelos de Box e Jenkins Sazonais – SARIMA*.

No entanto, a utilização desses respectivos modelos está ligada basicamente a dois aspectos fundamentais: o primeiro se refere aos objetivos do pesquisador quanto ao componente sazonal, ou seja, pode-se ter como objetivo a análise do componente sazonal, que é o caso do presente trabalho, e, ou, elaborar um modelo de previsão que incorpore a sazonalidade para ganhar eficácia preditiva. Entretanto, pode-se ter também como finalidade apenas a desazonalização dos dados, a fim de observar com maior precisão o comportamento de outros componentes. Por outro lado, o próprio comportamento do componente sazonal é decisivo para determinar o método a ser empregado para a sua “modelagem”.

Conforme aponta VANDAELE (1983), assim como o componente de tendência em uma série de tempo pode ser classificado, de acordo com a sua natureza, como do tipo *determinístico* ou do tipo *estocástico*, o componente

sazonal também poderá ser conceituado como *sazonalidade determinística* e *sazonalidade estocástica*²⁹. Assim, a utilização do método errado para a análise e eliminação desse componente poderá implicar inferências e conclusões imprecisas.

A visualização gráfica do componente sazonal em uma série de preços, por exemplo, que é a variável de interesse no presente estudo, pode e será implementada por meio da elaboração de gráficos que fixem no eixo das abcissas o período de um ano, para o caso de sazonalidade anual, e representem no eixo vertical os preços praticados em diferentes anos. Esse procedimento pode ser mais bem compreendido com a ajuda da Figura 15.

Como se pode observar, a sobreposição das repartições da série pode indicar que um determinado comportamento sazonal exista. No caso da Figura 15, tem-se que a série tende a apresentar níveis mais elevados nos meses próximos a julho, que poderia ser considerado o período de entressafra, e níveis mais baixos nos meses próximos à passagem do ano calendário, ou seja, nos meses de janeiro e dezembro.

Caso o comportamento das sobreposições dos diferentes anos não apresentasse padrão regular, poder-se-ia *a priori* inferir que a série em questão não apresenta sazonalidade. Contudo, isso nem sempre é verdade, conforme aponta SCHWAGER (1995). O padrão do comportamento sazonal pode estar mudando ao longo dos anos, e essa mudança pode estar sendo provocada por uma *relação multiplicativa* entre o componente sazonal e o componente de tendência na série. Assim, tem-se como um procedimento usual, conforme indicado por VANDAELE (1983), a remoção da tendência³⁰ da série antes de se iniciar a análise do componente de sazonalidade, o que de fato será feito no trabalho de pesquisa.

²⁹ Esses conceitos serão discutidos posteriormente.

³⁰ Com a utilização dos métodos para eliminação da tendência de uma série apresentados anteriormente.

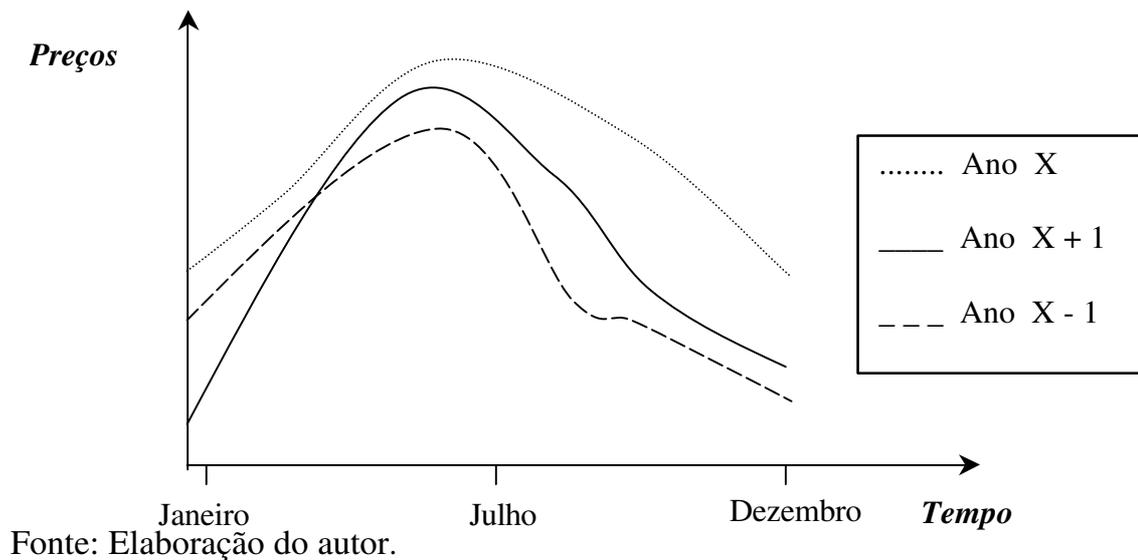


Figura 15 - Exemplo de um gráfico que pode ajudar a detectar a componente sazonal em um série de preços.

Se a série em questão flutuar acentuadamente, a verificação da presença de sazonalidade pela visualização direta do gráfico da série pode não ser possível. PINDYCK e RUBINFELD (1998) argumentam que a análise da *função de autocorrelação* para uma série de tempo pode se configurar num instrumento alternativo eficiente para a verificação da ocorrência de sazonalidade nessa série. Essa função, na *forma amostral*, será dada por:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (56)$$

Para uma série y_t mensal que apresente o componente sazonal, os diversos pontos amostrais deverão apresentar algum grau de correlação com os dados correspondentes com *lags* e *leads*³¹ de 12 meses. Ou seja, espera-se a existência de correlação entre y_{t-12} e y_{t+12} com os valores de y_t . Essas correlações

³¹ *Lags* se referem a valores defasados e *leads* se referem a valores futuros.

deverão se manifestar na função de autocorrelação amostral, dada por (56), pela presença de picos significativos para os valores estimados de $\hat{\rho}_k$, em que $k = 12, 24, 36, 48...$ Esse procedimento também será implementado neste trabalho.

A mensuração dos efeitos, ou impactos, que os diferentes meses do ano exercem sobre a série de tempo em questão pode ser efetuada com a estimação de *índices de sazonalidade*, que de forma agregada representam o componente sazonal. ABRAHAM e LEDOLTER (1983) descrevem o componente sazonal, e os referidos índices, com a ajuda da seguinte equação:

$$S_t = \sum_{i=1}^s \delta_i \text{IND}_{ti} \quad (57)$$

em que $\text{IND}_{ti} = 1$ se t se refere ao período sazonal i e $\text{IND}_{ti} = 0$ para os demais períodos; δ_i representa os valores dos índices que deverão ser adicionados ou subtraídos da série sazonal, a fim de se eliminar esse efeito; e s representa o número de períodos de sazonalidade que poderão ocorrer. Por exemplo, s será igual a 12, se o fenômeno sazonal a ser analisado ocorrer mensalmente ao longo do ano, e s será igual a 4, se o fenômeno sazonal estiver ligado às estações do ano (trimestral).

De maneira geral, a implementação do cálculo de tais índices consistirá da especificação de uma equação de regressão, que será estimada por MQO, tendo como variável dependente a série y_t , livre do componente de tendência. Como variáveis independentes seriam especificadas 12 *variáveis binárias*, para o caso de análise mensal. Assim, a equação a ser estimada será:

$$y_t = \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + \dots + \beta_{12} D_{12t} + \varepsilon_t \quad (58)$$

ou, de maneira mais compacta:

$$y_t = \beta_0 + \sum_{i=1}^{12} \beta_i D_{it} + \varepsilon_t \quad (59)$$

em que D_{it} ($i = 1, 2, 3, \dots, 12$ e $t = 1, 2, 3, \dots, n$) são variáveis *dummy* assumindo o papel das variáveis *IND* na equação (57); ε_t é um termo de erro aleatório; e β_i representa o efeito sazonal no i -ésimo período.

Como pode ser observado, a equação (59) possui intercepto. Esse modelo não poderá ser estimado com essa especificação, pois ocorrerá aí o problema econométrico de *multicolinearidade perfeita*³². Assim, um dos valores de β_i deverá ser igualado a zero (β_{12} por exemplo) e, portanto, a interpretação dos índices sazonais será modificada. Cada um dos índices β_i (para todo $i \neq 12$) representará o efeito sazonal do i -ésimo mês em relação ao valor observado no 12.º mês, ou seja, será a diferença que deverá ser acrescida ou subtraída do valor observado no *mês-base*, a fim de eliminar o efeito sazonal. Neste trabalho será estimada uma equação similar à dada em (59), com a exclusão do intercepto.

A significância estatística dos coeficientes estimados na equação (59) será testada com a utilização do *teste-F* e pela análise dos *valores de probabilidade* estimados (*p-values*) para cada coeficiente. De forma geral, no presente trabalho será adotado o nível de significância $\alpha = 0,10$, isto é, 10%.

Se o componente de sazonalidade presente na série analisada for do tipo *determinístico*, os resíduos do modelo (59) representarão os valores das séries livres do componente de sazonalidade, isto é, serão dados desazonalizados que serão utilizados na análise de outros componentes, como o de volatilidade, por exemplo.

A *sazonalidade determinística* se refere ao efeito sazonal que apresenta comportamento relativamente estável e previsível ao longo dos anos e geralmente está ligada a séries que sofrem influência da ocorrência de fenômenos climáticos e naturais, como também séries ligadas a datas específicas que se repetem de ano em ano, com intensidade regular. Alguns exemplos que podem ser citados, de acordo com ABRAHAM e LEDOLTER (1983), são: demanda de bens consumidos em períodos de temperatura elevada (ventiladores, sorvete, ...);

³² O somatório das colunas das *dummies* apresentará colinearidade perfeita com a coluna de números um que multiplicam o coeficiente de intercepto. A matriz $(X'X)^{-1}$ será singular.

demanda de brinquedos perto do Natal; oferta de produtos agrícolas; ganho de peso de bovinos, dada a maior disponibilidade de pastagens em alguns períodos do ano; e preços de *commodities* agropecuárias, visto que esses preços, de maneira geral, estão inversamente correlacionados com a oferta das *commodities*.

Contudo, conforme o discutido por VANDAELE (1983), o padrão da sazonalidade na série y_t poderá ser do tipo *estocástico*. Isso ocorrerá nos casos em que o componente sazonal possuir comportamento instável ao longo do tempo, modificando-se de forma imprevisível. A *sazonalidade estocástica* em séries de tempo geralmente ocorrerá em séries financeiras e macroeconômicas, que são influenciadas por diversos fatores que não necessariamente irão se repetir de forma previsível de período sazonal para período sazonal. Assim, a estimação de um modelo como o dado pela equação (59) não irá ter nos seus resíduos uma série y_t livre do fenômeno sazonal.

Se a sazonalidade for do tipo estocástica, tem-se como principal método para eliminar o componente sazonal, e que será utilizado neste trabalho, a implementação de um procedimento conhecido como *diferenciação sazonal*. De acordo com CHATFIELD (1996), os efeitos sazonais em uma série mensal serão eliminados pelo emprego do seguinte operador:

$$\nabla_{12}y_t = y_t - y_{t-12} \quad (60)$$

em que o operador ∇_{12} é chamado de *operador de diferença sazonal*, e a sua aplicação possibilitará a obtenção de uma série estacionária.

Alguns modelos de séries temporais, como o SARIMA, desenvolvido por BOX e JENKINS (1976), incorporam a sazonalidade na forma estocástica em suas estruturas, para propiciar análises e previsões mais precisas dos valores futuros da série. Essa classe de modelos será apresentada brevemente a seguir³³.

Os modelos sazonais de BOX e JENKINS (1976), também conhecidos como modelos SARIMA $(p,d,q) \times (P,D,Q)_s$, visam explorar a correlação serial

³³ Para detalhes acerca da metodologia desses modelos, ver o trabalho original de BOX e JENKINS (1976) e ABRAHAM e LEDOLTER (1983).

existente entre diferentes valores de uma série temporal Y_t . A equação fundamental deste modelo é dada pela seguinte expressão:

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_p B^{Ps})(1 - B)^d (1 - B^s)^D y_t = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) \varepsilon_t \quad (61)$$

Na primeira parte desta expressão, o primeiro termo entre parênteses³⁴ corresponde aos coeficientes de um modelo auto-regressivo -AR(p); o segundo termo corresponde aos coeficientes de um modelo auto-regressivo sazonal - SAR(P); o termo d corresponde à ordem de integração da série, ou seja, o número de diferenças tomadas; e D corresponde ao número de eventuais diferenças sazonais tomadas.

Na segunda parte desta expressão, após o sinal de igualdade, o primeiro termo entre parênteses corresponde aos coeficientes de um modelo do tipo média móvel - MA (q) e o segundo corresponde aos coeficientes de um modelo de médias móveis sazonais - SMA (Q).

Utilizando-se o operador de diferenças ∇ , este modelo pode ser reescrito, de maneira mais compacta, como:

$$\varphi(B)\Phi(B^s)\nabla^d \nabla_s^D y_t = \theta(B)\Theta(B^s)\varepsilon_t \quad (62)$$

Para se identificar um modelo SARIMA (p,d,q) x (P,D,Q)_s, deve-se inicialmente determinar os valores de d e D por meio da análise do gráfico das funções de autocorrelação e de autocorrelação parcial. Se for observado comportamento não-estacionário com persistência de elevados valores para os coeficientes de autocorrelação e de autocorrelação parciais, pode-se inferir a necessidade de diferenciação da série.

Da mesma forma, se valores elevados e persistentes nos *lags* com defasagens próximas aos múltiplos do período sazonal (próximas de 12, 24 e 36,

³⁴ Em que B é um operador de defasagens. Por exemplo, $BY_t = Y_{t-1}$ e $B^k Y_t = Y_{t-k}$.

por exemplo), pode-se inferir também a necessidade de diferenciação sazonal da série.

Conforme aponta Fava, em VASCONCELLOS e ALVES (2000), para determinar os valores de p , P , q e Q também se deve observar o comportamento das funções de autocorrelação e de autocorrelação parcial da série estacionarizada após a possível aplicação de diferenças consecutivas e, ou, sazonais. O comportamento dos coeficientes de autocorrelação de *lags* menores fornecerá subsídios para a determinação dos valores de p e q , e os coeficientes ligados aos períodos sazonais indicarão os possíveis valores de P e Q . O Quadro 2, desenvolvido a partir do proposto por GUJARATI (1995), pode sintetizar as possibilidades existentes.

Um outro método muito utilizado para desazonalizar uma série de tempo é o desenvolvido pelo departamento de censos do governo norte-americano (*Census bureau*), conhecido como *programa de ajustamento sazonal X-11*. No entanto, esse programa, de acordo com VAUGHN et al. (1981), apresenta algumas desvantagens fundamentais em relação às propostas do presente estudo. Isto porque no “mecanismo” interno de eliminação da sazonalidade o programa já possui incorporados procedimentos para tratamento das componentes de tendência e ciclos. Ou seja, o programa elimina a sazonalidade implementando também um tratamento para os demais componentes, como uma espécie de “caixa preta”, e esses demais componentes são justamente os pontos de interesse do presente estudo. Além disso, este programa não opera com séries temporais longas, como as que serão analisadas no presente trabalho.

Quadro 2 - Padrões teóricos típicos das funções de autocorrelação

Tipo de Modelo	Padrão Típico da FAC	Padrão Típico da FAC Parcial
AR(p)	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> iniciais.	Picos significativos nos “p” <i>lags</i> do modelo
MA(q)	Picos significativos nos “p” <i>lags</i> do modelo	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> iniciais.
ARMA(p,q)	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> iniciais.	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> iniciais.
SAR(p)	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> sazonais.	Picos significativos nos “P” <i>lags</i> sazonais do modelo
SMA(q)	Picos significativos nos “P” <i>lags</i> sazonais do modelo	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> sazonais.
SARMA(p,q)	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> sazonais.	Queda exponencial ou na forma de uma senóide amortecida nos <i>lags</i> sazonais.

Fonte: Elaboração do autor a partir de proposição inicial de GUJARATI (1995).

2.2.4. Análise do componente de volatilidade

Uma das principais pressuposições do modelo de regressão linear clássico é a de que os termos de erro possuem variância uniforme e não são correlacionados uns com os outros. De forma geral, essas características são descritas em termos da matriz de variâncias e covariâncias do vetor de erros. Esta matriz apresenta em sua diagonal principal as variâncias dos erros individuais das observações amostrais e fora da diagonal principal as covariâncias entre eles. Se todos os termos desta diagonal principal são iguais, tem-se que os erros

apresentam variância uniforme, ou seja, são ditos *homocedásticos*; caso contrário, são chamados de erros *heterocedásticos*. Por outro lado, se os termos fora da diagonal principal são iguais a zero, os erros são ditos não-correlacionados, isto é, o erro ocorrido em um período não é correlacionado com o termo de erro de outro período.

Se esses pressupostos de homocedasticidade e não-correlação serial dos resíduos forem violados, a matriz de Variâncias e Covariâncias não apresentará mais a forma $\sigma^2 I$, em que σ^2 é a variância uniforme, anteriormente referida, e I é uma matriz identidade de ordem T . Com essa violação, os estimadores de Mínimos Quadrados Ordinários (MQO) perdem sua característica de mínima variância (deixam de ser estimadores *BLUE*) e, com isso, o estimador a ser empregado é o de Mínimos Quadrados Generalizados (MQG).

De maneira geral, dadas as características intrínsecas dos dados, a presença de heterocedasticidade está associada a dados extraídos de cortes seccionais (*cross-sections*), e a presença de autocorrelação está ligada a dados de séries de tempo, o que, conforme GUJARATI (1995), já se tornou uma espécie de "senso comum". Entretanto, estudos relativamente recentes, como os de ENGLE (1982, 1983) e CRAGG (1982), mostraram que em análises de dados macroeconômicos e de séries de tempo relativas a dados financeiros, como preços de ações nas bolsas de valores e índices de preços, foram encontradas diversas evidências de variâncias menos estáveis do que as normalmente assumidas nos modelos tradicionais. Eles apontam que a instabilidade das variâncias desses dados se caracteriza como regra geral, e não como exceção, e que os modelos econométricos empregados para análise e previsão dessas séries deveriam incorporar essas características irregulares das variâncias.

É nesse contexto que se originaram os *Modelos de Heterocedasticidade Condicional Auto-regressiva* (ARCH) e duas de suas principais variações, quais sejam: o *Modelo de Heterocedasticidade Condicional Auto-regressiva Generalizado* (GARCH); e o *Modelo de Heterocedasticidade Condicional Auto-regressiva com Limiar* (*Threshold-ARCH* ou *TARCH*), que apresenta a

característica de detectar assimetrias nas variâncias condicionais. Esses modelos serão empregados na tese para a análise da volatilidade dos preços do café.

O surgimento desses modelos se deu com ENGLE (1982), que, ao analisar modelos de inflação, detectou que erros de previsão tendiam a ocorrer na forma de *clusters*, sugerindo uma forma de heterocedasticidade na qual a variância dos erros de previsão dependesse do tamanho do erro anterior. Com o intuito de captar essa relação, ele desenvolveu o modelo ARCH, que, de acordo com GUJARATI (1995), possui como idéia principal o fato de que a variância de 'e' no período de tempo $t(=\sigma_t^2)$ depende do tamanho do quadrado do termo de erro no período $t-1$, ou seja, depende de e_{t-1}^2 . Dado um modelo de regressão linear de $k + 1$ variáveis, como:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \dots + \beta_k X_{kt} + e_t \quad (63)$$

o termo de erro e_t , condicionado à informação disponível no período $(t-1)$, seria distribuído como:

$$e_t \sim N[0, (\alpha_0 + \alpha_1 e_{t-1}^2)] \quad (64)$$

Portanto, tem-se que a variância de e no período t irá depender de um termo constante mais o quadrado do erro no período $t-1$. Esse seria o chamado processo ARCH(1), que pode ser generalizado para 'm' defasagens de e^2 , gerando o processo ARCH (m) conforme se segue:

$$\text{Var}(e_t) = \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \dots + \alpha_m e_{t-m}^2 \quad (65)$$

No modelo ARCH (m) a variância de e no período t dependerá de um termo constante mais o quadrado dos erros nos períodos $t-1$ até $t-m$. Caso não exista a correlação entre as variâncias dos erros, os coeficientes (*alfas*) estimados em (65) não serão estatisticamente diferentes de zero, e o modelo apresentará homocedasticidade, retornando ao caso clássico.

Uma forma proposta por ENGLE (1982), para testar a hipótese nula (de não-correlação das variâncias), seria pelo *Teste do Multiplicador de Lagrange*, que será o utilizado neste estudo. Este teste consiste basicamente em tomar-se o R^2 da regressão (65) e multiplicá-lo pelo número de observações T da amostra; o coeficiente encontrado possuirá distribuição Qui-quadrado com um grau de liberdade. Comparando o valor encontrado com o tabelado, pode-se definir pela significância ou não dos coeficientes estimados.

Conforme aponta GREENE (1997), para o modelo ARCH (1), e similarmente para os de ordem (m), caso $|\alpha_1| < 1$, os estimadores de MQO para (1) continuam sendo *Blue*, porém os estimadores não-lineares de MQG são mais eficientes. Portanto, a estimação deverá ser realizada pelo Método da Máxima Verossimilhança (MMV), onde se assume que os erros são distribuídos normalmente.

Em diversas aplicações, o modelo ARCH (m) apresentou longa estrutura de defasagens, porém BOLLERSLEV (1986) desenvolveu um refinamento do modelo ARCH, proposto por ENGEL, onde seria possível, com número menor de parâmetros (de acordo com o *princípio da parcimônia* em econometria), se estimar um modelo capaz de descrever o comportamento irregular da variância de uma série. Seu modelo ficou conhecido como Modelo de Heterocedasticidade Condicional Auto-regressiva Generalizado de ordem (r,m) - [GARCH (r,m)].

No caso do modelo GARCH de ordem (1,1), tem-se que a variância dos erros de um modelo, econométrico ou de séries temporais, no período t dependerá basicamente de três termos, a saber:

- a) Um termo médio ou constante ω .
- b) Inovações (choques) acerca da volatilidade, que é determinada pelo quadrado dos resíduos (e_{t-1}^2) do período $t-1$, que é o termo ARCH.
- c) revisão da volatilidade feita no último período ($\hat{\sigma}_{t-1}^2$), que é o termo GARCH.

Conforme HALL et al. (1995) mencionam, essa especificação faz sentido, pois em análises de finanças, por exemplo, um agente econômico pode prever a variância de um ativo hoje como sendo formada por uma média

ponderada de uma *variância média*, ou constante de longo prazo, a *previsão da variância* feita ontem e "*o que foi aprendido*" ontem.

As relações mencionadas anteriormente são formalizadas pelas seguintes equações que compõem o modelo GARCH (1,1):

$$Y_t = X_t\pi + e_t \quad (66)$$

$$\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (67)$$

A equação (66) representa um modelo genérico, que tanto pode ser um modelo econométrico "tradicional" como um modelo de séries temporais do tipo SARIMA (p,d,q) x (P,D,Q)_s, descrito anteriormente. O ponto importante a se ressaltar refere-se ao fato de que este modelo deve ser bem ajustado, no sentido de garantir que os resíduos (e_t) possuam média zero e não sejam correlacionados. Contudo, a *variância condicional* de e_t será dada pela equação (67) e indica que ela pode estar mudando ao longo do tempo, o que é o motivo principal do surgimento e desenvolvimento dos modelos ARCH, e distinguir os componentes condicionais da variância dos resíduos dos componentes puramente aleatórios.

A equação (67) pode ser generalizada, a fim de se obter o modelo GARCH de ordem (r,m), que apresenta a seguinte forma:

$$\sigma_t^2 = \omega + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \dots + \alpha_m e_{t-m}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_r \sigma_{t-r}^2 \quad (68)$$

ou, de forma mais compacta:

$$\sigma_t^2 = \omega + \sum_{i=1}^m \alpha_i e_{t-i}^2 + \sum_{j=1}^r \beta_j \sigma_{t-j}^2 \quad (69)$$

em que r representa a ordem dos termos GARCH e m representa a ordem dos termos ARCH do modelo.

Contudo, conforme afirmam GREENE (1997) e HALL et al. (1995), na prática, a maioria dos processos de volatilidade condicional tem sido modelada como processos GARCH (1,1), sendo poucos os casos em que modelos de ordem superior apresentam resultados significativamente melhores.

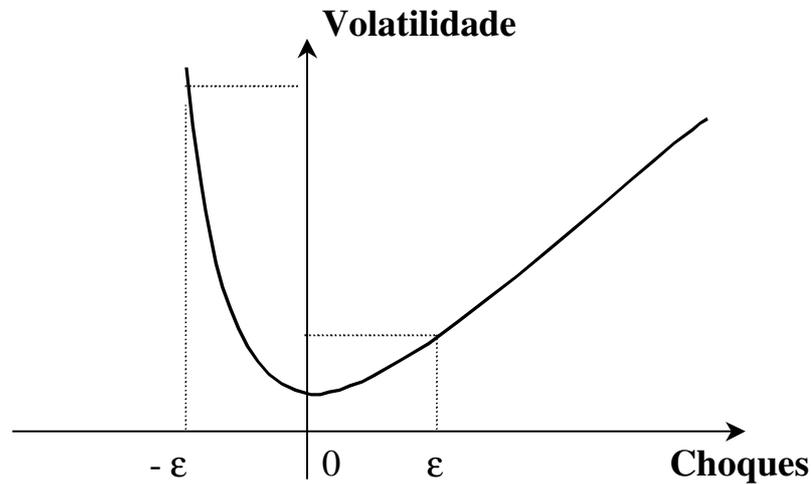
A soma dos coeficientes de um modelo GARCH é chamada de *coeficiente de persistência da volatilidade*. Por exemplo, no modelo GARCH(1,1) dado pela equação (67), a persistência será dada pela soma $(\alpha+\beta)$. Valores baixos (próximos de 0) para o coeficiente de persistência indicam que um choque inicial sobre a volatilidade irá provocar efeitos rápidos sobre o comportamento das séries e que, após curto período de tempo, a variância da série deverá convergir à sua média histórica.

Por outro lado, quanto maior (mais próximo de 1) for o valor do coeficiente de persistência, mais vagarosamente o choque sobre a volatilidade irá se enfraquecer e, portanto, maior será a demora do processo de reversão à média para a variância. Finalmente, conforme apontam CUDDINGTON e LIANG (1999), se o valor do coeficiente de persistência for maior ou igual a 1, os choques na volatilidade irão perdurar por um período extremamente longo na série em questão. Nesse caso, diz-se que a variância condicional de e_t possui raiz unitária e a variância permanecerá elevada, não apresentando reversão à sua média histórica. Isso, de acordo com NELSON (1990), não implica que a série de resíduos e_t não seja estacionária, mas sim que a variância incondicional de e_t não será.

Na análise do comportamento da variância condicional de algumas séries de ativos financeiros (como o preço de ações), alguns autores, como GLOSTEN et al. (1993) e NELSON (1991), observaram comportamento assimétrico da volatilidade em relação a choques positivos e negativos no comportamento das séries. De maneira geral, a evidência empírica indicava que choques negativos nos preços de ações levam a um maior crescimento na volatilidade do que choques positivos, ou seja, movimentos "baixistas" nos mercados tendem a gerar volatilidades de maior magnitude do que movimentos "altistas" da mesma magnitude.

De acordo com ENGLE e NG (1993), esse fenômeno de assimetria na volatilidade pode ser representado graficamente com a ajuda de uma *Curva Assimétrica de Impactos de Choques*, dada pela Figura 16. Como se pode observar, choques (inovações) positivos tendem a ter impactos menos

expressivos sobre a volatilidade das séries do que choques negativos da mesma magnitude.



Fonte: QMS (1995).

Figura 16 - Curva assimétrica de impactos de choques.

De acordo com a especificação de GLOSTEN et al. (1993), a *variância condicional assimétrica* pode ser representada formalmente pela seguinte variação do modelo GARCH (1,1), conhecida como modelo TARARCH (1,1):

$$\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \gamma e_{t-1}^2 d_{t-1} + \beta \sigma_{t-1}^2 \quad (70)$$

em que $d_t = 1 \forall e_t < 0$ e $d_t = 0 \forall e_t > 0$. Dessa forma, choques positivos ($e_t > 0$) e choques negativos ($e_t < 0$) terão diferentes efeitos na variância condicional: choques positivos terão impacto dado por α e choques negativos terão impacto dado por $\alpha + \gamma$. Se o coeficiente γ for significativamente diferente de zero, diz-se que a variância condicional é assimétrica. Assim, $\gamma < 0$ indica que um choque

positivo implica menor volatilidade do que um choque negativo de mesma magnitude.

A especificação do modelo TARARCH de ordem superior, TARARCH (r,m), é dada pela seguinte expressão:

$$\sigma_t^2 = \omega + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \dots + \alpha_m e_{t-m}^2 + \gamma e_{t-1}^2 d_{t-1} + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_r \sigma_{t-r}^2 \quad (71)$$

ou, de maneira mais compacta:

$$\sigma_t^2 = \omega + \sum_{i=1}^m \alpha_i e_{t-i}^2 + \gamma e_{t-1}^2 d_{t-1} + \sum_{j=1}^r \beta_j \sigma_{t-j}^2 \quad (72)$$

Uma maneira alternativa para se detectar a existência de efeitos assimétricos na volatilidade pode e será implementada neste trabalho com a utilização de um gráfico das correlações entre o quadrado dos resíduos e os resíduos defasados do modelo GARCH estimado. Essas correlações cruzadas deverão ser iguais a zero para o caso de um modelo GARCH simétrico e negativas para o caso de um modelo TARARCH, que incorpora essa assimetria.

Conforme afirma GREENE (1997), um modelo de regressão com termos de erro que apresentam as características apontadas pelos modelos do tipo ARCH pode ser estimado pelo método tradicional dos *Mínimos Quadrados Ordinários* (MQO), e os estimadores $(\hat{\beta})$ resultantes desse processo serão os *estimadores lineares mais eficientes* de β . Contudo, existe um estimador não-linear que é mais eficiente e, conseqüentemente, mais desejável do ponto de vista estatístico, apesar de envolver um maior esforço computacional. Esse estimador é dado pelo *Método da Máxima Verossimilhança*³⁵ e geralmente é utilizado na estimação dessa classe de modelos, visto que esse método está disponível nos pacotes econométricos que operam com modelos do tipo ARCH.

³⁵ Prova 5 no Apêndice A.

2.3. Evidência empírica e literatura correlacionada

A presente pesquisa busca contribuir para o melhor entendimento do processo de formação de preços no mercado *spot* de café. Em alguns aspectos, este trabalho deverá se mostrar como inédito, porém em outros ele deverá ratificar alguns resultados encontrados por outros autores ou mesmo contestar esses resultados.

De maneira geral, alguns pontos podem ser apontados como os principais aspectos diferenciadores do presente trabalho de pesquisa em relação aos trabalhos preexistentes. O primeiro se refere à sistemática de análise. No trabalho de pesquisa serão analisados todos os componentes estocásticos determinantes das séries de preços do café no mercado internacional. Vários trabalhos, como os de MORICOCCHI et al. (1997), HEIFNER e KINISHITA (1994), CORTEZ e MARTIN (1996), PETTI (1996), DALY (1956), FRANÇA (1990), CAVALCANTI (1978) e HASEGAWA (1995), entre outros, enfatizam a análise do comportamento de um³⁶ dos componentes dos preços, como, por exemplo, o componente sazonal, ou o cíclico.

Em segundo lugar, tem-se que a metodologia a ser utilizada na análise dos problemas levantados, principalmente no tocante à análise de ciclos e da volatilidade dos preços no mercado *spot* da referida *commodity*, se apresenta como um fator diferenciador do presente trabalho. Com relação à análise espectral, apesar do "potencial" desse instrumental, observa-se que apenas alguns trabalhos, como o de FRANÇA et al. (1989) e FRANÇA (1990), o aplicaram para a análise de ciclos de preços de *commodities*, no caso o preço do algodão.

Mesmo para o estudo de ciclos em outras variáveis econômicas, o uso dessa metodologia ainda não está plenamente difundido. Alguns casos que podem ser citados, na literatura brasileira, se referem ao estudo de ciclos em

³⁶ COUTO (1996) seria uma exceção a esta relação, visto que nesse trabalho o autor analisou o comportamento dos componentes de tendência, ciclo e sazonalidade nos preços da bovinocultura de corte no Estado de São Paulo.

exportações (LEMOS, 1983); índices de ações (CONTADOR, 1975); e consumo de energia elétrica (PEREIRA et al., 1986).

No que tange ao estudo da volatilidade de preços de ativos agropecuários, observa-se que a metodologia dos modelos da "família" ARCH praticamente ainda não foi utilizada no Brasil. A grande maioria das aplicações se refere à análise do comportamento da volatilidade de ativos financeiros e de outras variáveis de natureza macroeconômica, como: taxas de juros (ROCQUE e GARCIA, 1997); inflação (ENGLE, 1982; BOLLERSLEV, 1986); taxa de retorno de ações (NELSON, 1991); e índices de bolsas de valores (BERA et al., 1993).

Outro aspecto diferenciador do trabalho de pesquisa, em relação à literatura preexistente, se refere à *commodity* que terá seus preços e mercados analisados. Conforme discutido anteriormente, o café se configura como a principal *commodity* transacionada nas bolsas de futuros do Brasil, respondendo por mais da metade do volume total de contratos agropecuários negociados, em termos de número de negócios e de recursos movimentados. Todavia, nenhum estudo utilizando a metodologia geral proposta foi ainda implementado para a análise do mercado físico desse importante produto brasileiro.

Muitos dos estudos do comportamento de mercados e preços existentes se referem a produtos que possuem relevante expressão na composição do agronegócio nacional, como o leite, frango e arroz, mas que não são transacionados em bolsas de futuros brasileiras, e, ou, *commodities*, que, apesar da sua importância, têm em seus contratos futuros uma demanda, ainda, pouco expressiva. Este é o caso da soja, do açúcar, do milho e do algodão, estudados por FRANÇA (1990).

Assim, o presente trabalho de pesquisa apresenta um diferencial em relação aos trabalhos preexistentes, pois tem como um dos seus principais objetivos a geração de subsídios teóricos e empíricos aos agentes econômicos atuantes nos mercados futuros de café; esses subsídios podem ser extremamente úteis enquanto informações necessárias à otimização dos seus processos de tomada de decisão de *hedge* e especulação.

2.4. Dados e *softwares* computacionais utilizados

Para a implementação do presente projeto de pesquisa foram utilizados dados mensais secundários dos preços reais do café brasileiro no mercado *Spot* de Nova Iorque (New York Board Of Trade – NYBOT) de janeiro de 1946 a dezembro de 2000, cotados em centavos de US\$ por libra-peso de 2000.

Cada libra-peso tem 453,6 gramas. Para transformar dólares por saca de 60 kg em centavos de dólar por libra-peso, deve-se dividir 60.000 gramas por 453,6 gramas e depois dividir por 100, o que resultará no fator 1,3228. Inversamente, para transformar o preço de Nova Iorque (US\$ cent./libra-peso) em cotações da BM&F (US\$/saca), deve-se multiplicar o primeiro por 1,3228.

Os dados foram coletados em anuários estatísticos da produção agrícola brasileira publicados pela FNP Consultoria. As edições consultadas foram: AGRIANUAL 96, AGRIANUAL 97, AGRIANUAL 98, AGRIANUAL 99, AGRIANUAL 2000 e AGRIANUAL 2001. Deve-se ressaltar que os dados referentes aos meses de setembro a dezembro de 2000, que não estão publicados no anuário de 2000, foram obtidos diretamente no escritório da empresa FNP Consultoria.

Os dados, que se encontravam em valores nominais de US\$, foram atualizados para valores de dezembro de 2000 por meio do índice de inflação americana para o consumidor (Consumer Price Index for All Urban Consumers) do U.S. Department of Labor - Bureau of Labor Statistics.

Os principais *softwares* que foram empregados no trabalho de pesquisa para o processamento dos dados e dos modelos de análise nos domínios do tempo e da frequência foram: *STATISTICA 5.0*, da *Stat Soft Inc.*; o *SPSS 8.0*, da *SPSS Inc.*; e *EViews 3.1*, da *Quantitative Micro Software*.

3. RESULTADOS E DISCUSSÃO

3.1. Análise do componente de tendência

A verificação da ocorrência ou não do componente de tendência (não-estacionariedade) nos preços reais do café se configura como um ponto de profunda importância para análises empíricas que envolvam o mercado cafeeiro, em particular as análises implementadas por meio de modelos econométricos. Isso ocorre porque, caso esses aspectos não sejam levados em consideração, as análises poderão ser equivocadas em razão do problema econométrico de regressões espúrias.

Regressões de uma variável, que apresente tendência em função de outra(s) que também apresente(m) comportamento não-estacionário, poderão levar a elevados valores para o coeficiente de determinação dos modelos (R^2), mesmo que essas variáveis não tenham nenhuma relação econômica entre si. Esse problema ocorre porque, se ambas as séries exibirem componentes de tendência (tanto de elevação como de queda), o valor alto para o R^2 se dará em função da presença da tendência e não de uma verdadeira relação entre as variáveis.

Com base no exposto, foram realizados diversos testes para determinação da existência e influência do componente de tendência nos preços reais do café

brasileiro no mercado *spot* de Nova Iorque. Esses testes se iniciaram com análises mais simples e diretas, como as propiciadas pela verificação do gráfico dos preços para o período selecionado como amostra, passando pela análise da função de autocorrelação e do respectivo correlograma da série de preços, bem como o de sua primeira diferença. Posteriormente, foram realizados os testes de raízes unitárias de Dickey-Fuller (DF), de Dickey-Fuller Expandido (ADF) e o teste de Phillips-Perron, para uma análise mais consistente da estacionariedade dos preços do café. Modelos determinísticos também foram elaborados e testados.

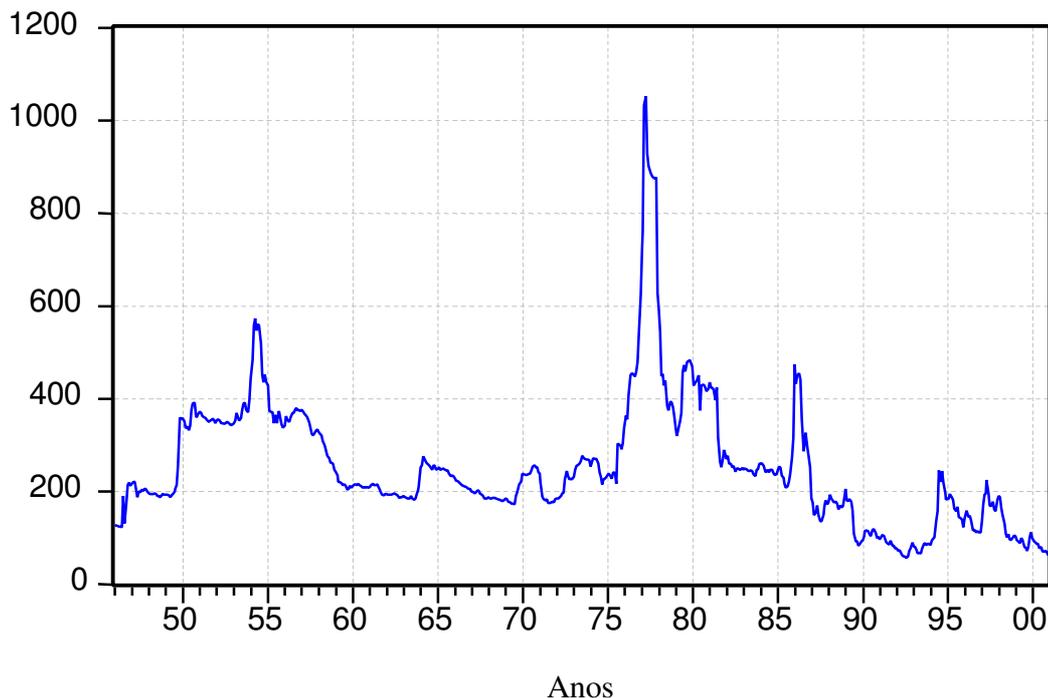
3.1.1. Análise gráfica

Na Figura 17, que representa o comportamento da cotação do café brasileiro no mercado internacional, pode-se observar que, quanto ao componente de tendência³⁷, o gráfico indica que os preços da referida *commodity* apresentaram, de maneira geral, comportamento de razoável declínio no período analisado, compreendido entre os anos de 1946 e 2000.

Deve-se ressaltar que os preços apresentaram alguns momentos de alta abrupta, em razão de fatores climáticos adversos ocorridos no período da amostra, a saber: uma elevação significativa em relação ao seu nível anterior nos anos de 1976 e principalmente 1977, em razão de geada forte ocorrida no ano de 1975; uma seca forte ocorrida em 1985 que implicou quebra de safra e consecutiva elevação de preços em 1986; e, além disso, a junção de geadas e secas ocorridas em 1994 provocou novo pico nos preços em 1995.

³⁷ Aspectos ligados a ciclos, sazonalidade e volatilidade dos preços serão discutidos posteriormente.

Centavos de
US\$ / Libra-Peso



Fonte: Elaboração própria a partir de dados do USDA – Departamento de Agricultura dos Estados Unidos.

Figura 17 - Preços reais do café brasileiro no mercado *Spot* de Nova Iorque – 1946/2000 (em centavos de US\$/libra-peso de 2000).

Estes anos, a princípio atípicos, não foram excluídos da análise dos dados, visto que um dos principais interesses do presente trabalho de pesquisa se constitui na busca de comportamentos periódicos nos preços, que podem ser provocados tanto por variáveis econômicas quanto naturais, como é o caso de geadas e secas, que podem ocorrer com certa periodicidade. Assim, o tratamento dos referidos anos como *outliers* e a sua conseqüente exclusão da amostra poderiam, na verdade, se configurar como procedimento contrário aos objetivos da pesquisa.

Como se pode ainda visualizar na Figura 17, os anos de 1946 a 1955 representaram um período de crescimento para os preços do café brasileiro,

porém foram seguidos por um período de igual duração, em que os preços apresentaram processo de reversão de comportamento. Os 10 anos seguintes, de 1966 a 1975, foram caracterizados por relativa estabilidade com alguma oscilação em torno do valor médio para o período, que foi de US\$ 219,34 por libra-peso.

Após a forte queda ocorrida em 1975, que alavancou os preços do café em 1977 para patamares nunca antes vistos no setor, pôde-se verificar que os mesmos voltaram a apresentar uma tendência de queda ainda mais forte que a apresentada na primeira metade da amostra estudada. Essa tendência, que se iniciou em 1978, perdurou pelos 15 anos seguintes, com exclusão do ano de 1986, que sofreu os reflexos da seca de 1985, conforme mencionado anteriormente, culminando com o período de maior depreciação nos preços para a *commodity* brasileira de toda a sua história recente. Nos anos de 1992 e 1993, o café apresentou preço real médio de US\$ 73,85 por libra-peso, e em agosto de 1992 o produto chegou a ser cotado a US\$ 56,34 por libra-peso.

Apesar da relativa recuperação dos preços em 1995 e 1997, a análise dos cinco últimos anos da amostra estudada indica que os produtores de café deverão buscar se tornar cada vez mais competitivos em suas atividades, visando operar com sistemas mais eficientes em termos de custos e produtividade, uma vez que os preços do café estão em um processo de queda que vem se intensificando, principalmente nos últimos 25 anos.

Tem-se, portanto, que a análise visual da série temporal dos preços do café indica que ela se comporta como uma série não-estacionária, pois pode-se observar a presença do componente de tendência (de baixa no presente caso) e uma relativa inconstância na média da série e, principalmente, na sua variância, o que são pré-condições necessárias, porém não suficientes, para a estacionariedade da série. Ou seja, os referidos momentos da série de preços do café apresentam-se “visualmente” como variantes em função do tempo.

Todavia, a simples análise do gráfico dos preços do café não pode ser considerada como definitiva para determinação da estacionariedade e da presença de tendências determinísticas e, ou, estocásticas para a série em questão. Isso

porque as informações obtidas da visualização de um gráfico de uma série de tempo são meramente indicativas de possíveis padrões sistemáticos e repetitivos de comportamento para esta. Assim, análises e testes mais rigorosos serão apresentados a seguir, para analisar aspectos ligados ao componente de tendência, que aparentemente se verifica como presente nos preços reais do café brasileiro no mercado *spot* de Nova Iorque.

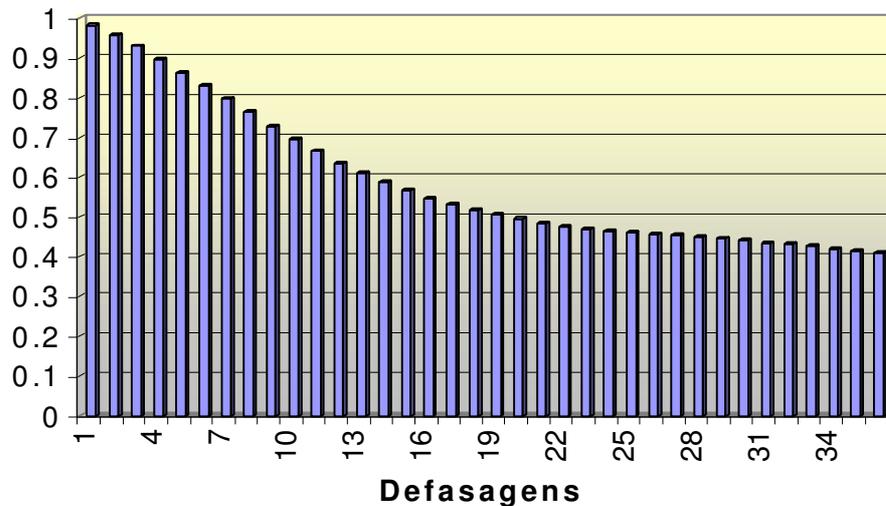
3.1.2. Análise da função de autocorrelação amostral

Conforme discutido anteriormente, um importante guia para a descoberta de propriedades e características da série de preços do café é dado pelo cálculo de valores denominados coeficientes de autocorrelação amostral, por meio da equação (56). Esses coeficientes mensuram a correlação existente entre os preços em diferentes momentos no tempo. Eles também fornecem informações que possibilitam a descoberta do modelo probabilístico que possa ter gerado essa série de tempo.

O gráfico dos coeficientes de autocorrelação amostral, também chamado de correlograma, em que os valores das autocorrelações ($\hat{\rho}_k$) são plotados em relação às suas diferentes defasagens (k), se constitui num dos principais instrumentos de auxílio para a interpretação destes coeficientes. Se os preços do café apresentarem o componente de tendência, os valores das autocorrelações não irão convergir para zero de maneira rápida, apenas após um número relativamente elevado de defasagens (12, por exemplo). Isso ocorrerá, pois as observações de um mesmo lado da média geral dos preços tenderão a ser seguidas por grande número de observações do mesmo lado da média, em função do componente de tendência.

Dessa forma, estimou-se a função de autocorrelação amostral para os preços do café brasileiro a partir da equação (56). Na Figura 18, é representado o correlograma para a série de preços.

Autocorrelações



Fonte: Resultados da pesquisa.

Figura 18 - Correlograma amostral para os preços reais do café.

Um primeiro aspecto que deve ser ressaltado se refere ao fato de que o valor de primeiro coeficiente de autocorrelação é extremamente alto (0,982), e este é seguido por uma série longa de valores das autocorrelações que apresentam comportamento de queda bastante lento. O coeficiente de autocorrelação demora 19 defasagens até cair abaixo de 0,5, o que é um indicador de que os preços do café se comportam como um processo não-estacionário e que possivelmente pode apresentar raízes unitárias.

Esse tipo de padrão comportamental pode ser contrastado com um em que a série analisada seria puramente aleatória. Nesse caso, as autocorrelações para qualquer defasagem maior que zero seriam iguais a zero, indicando a inexistência de qualquer tipo de dependência temporal entre os valores realizados para a série.

Os resultados encontrados apontam ainda que os coeficientes de autocorrelação dos preços do café são estatisticamente significativos até a defasagem de 64 meses (valores assinalados com asteriscos). Como esses valores

encontram-se fora do intervalo de confiança do teste, deve-se rejeitar a hipótese nula de que o verdadeiro valor do coeficiente de autocorrelação seja de fato igual a zero, em favor da hipótese alternativa de que ele seja estatisticamente significativo.

Tendo por base a análise do correlograma, pode-se concluir que a série temporal dos preços do café brasileiro no mercado *spot* de Nova Iorque, para o período analisado, não é estacionária, ou seja, apresenta comportamento de tendência sistemática ao longo do tempo. Todavia, conforme discutido na metodologia do presente trabalho de pesquisa, a não-estacionariedade dos preços do café pode estar ligada a dois tipos específicos de tendência que podem estar presentes de maneira isolada ou conjunta nessa série. Assim, deve-se verificar se essa tendência é do tipo determinística, do tipo estocástica ou uma forma conjugada das duas.

A diferença principal entre essas duas formas de tendência que podem estar presentes nos preços do café se refere ao fato de que a tendência estocástica implica uma variação percentual média em dado período de tempo; entretanto, ao contrário da determinística, em cada período a mudança provocada pela tendência em relação ao seu nível médio será um montante aleatório e imprevisível, em vez de constante, dado por determinada taxa. Assim, a tendência oscilará de forma aleatória à medida que o tempo evolua e choques exógenos afetem os preços do café. No caso conjugado, tem-se uma soma dos dois tipos de comportamento.

A determinação do tipo de tendência que está presente nos preços do café se configura como um ponto-chave em sua análise, pois, caso a tendência presente seja do tipo estocástica e a análise e remoção dessa tendência considerá-la, de forma equivocada, como do tipo determinística, alguns problemas econométricos poderão surgir. Por exemplo, se na verdade os preços do café apresentarem comportamento do tipo passeio aleatório com intercepto, a subtração de uma tendência determinística dessa série não a tornará uma série estacionária. Por outro lado, a aplicação de métodos para a eliminação de

tendências estocásticas nesta série, se a tendência for de fato determinística, também poderá provocar problemas de natureza similar.

3.1.3. Testes de raízes unitárias

Tendo por base a discussão anterior, serão apresentados a seguir os resultados dos chamados testes de raízes unitárias. Estes testes se apresentam como uma outra alternativa para a análise da estacionariedade dos preços do café. Eles possuem vantagem significativa em relação à análise do correlograma, pois, além de testar a estacionariedade dos preços, são capazes de indicar de maneira relativamente precisa o tipo de tendência presente nos preços do café, se determinística, estocástica e, ou, uma composição das duas.

As possíveis implicações do fato de que os preços do café possam apresentar raízes unitárias são potencialmente profundas. Se os preços reais da referida *commodity* forem verdadeiramente I (1), isto é, integrados de ordem um, tem-se que choques nesses preços terão efeitos permanentes. Dessa forma, caso confirmada, esta observação implicaria sérias reconsiderações para a análise de políticas agrícolas e macroeconômicas que afetem o setor cafeeiro, visto que estas teriam impactos na cafeicultura de maneira profunda e permanente. Por exemplo, o argumento de que uma mudança na tributação do setor poderia ter efeito apenas transitório sobre a produção poderia vir abaixo.

3.1.3.1. Testes de Dickey-Fuller e de Dickey-Fuller expandido

De acordo com o apresentado anteriormente, foram estimadas três equações similares³⁸ à equação (28), que caracterizam o teste de Dickey-Fuller: uma equação em que a diferença dos preços do café ($\Delta\text{PRCAFÉ}_t$) é considerada função dos preços de um período anterior (PRCAFÉ_{t-1}) e de um erro do tipo ruído branco (ε_t); uma outra, em que a diferença dos preços é função desses

³⁸ A principal diferença se refere à não-inclusão de diferenças defasadas da série no lado direito das equações.

mesmos componentes mais um termo de intercepto (μ); e, finalmente, uma terceira equação, em que a diferença dos preços seria função também de uma tendência determinística linear (t).

Os resultados encontrados foram os seguintes:

$$\Delta PRCAFÉ_t = -0,003762 PRCAFÉ_{t-1} \quad d = 1,5820 \quad (73)$$

(0,2588)

$$\Delta PRCAFÉ_t = 3,7499 - 0,0152 PRCAFÉ_{t-1} \quad d = 1,5723 \quad (74)$$

(0,0610) (0,0289)

$$\Delta PRCAFÉ_t = 7,6220 - 0,0088 t - 0,0192 PRCAFÉ_{t-1} \quad d = 1,5728 \quad (75)$$

(0,0133) (0,0975) (0,0093)

Os valores entre parênteses indicam os valores exatos de probabilidade (*P-Values*) para os coeficientes estimados, e os valores à sua direita são os valores da estatística d de Durbin-Watson. As equações estimadas, juntamente com suas respectivas estatísticas do teste de Dickey-Fuller, no Quadro 3, indicam, a princípio, que a série de preços do café não se apresenta como estacionária para os valores críticos, em níveis de significância estatística de 1%, 5% e 10%, isto é, indicam que ela apresenta raízes unitárias.

Quadro 3 - Resultados dos testes de Dickey-Fuller

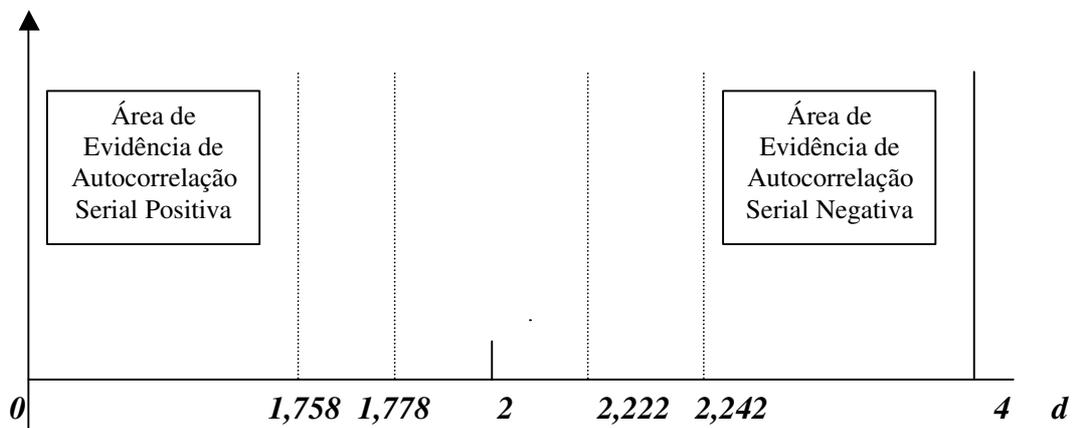
	Equação (73)	Equação (74)	Equação (75)
Estatística do teste de Dickey-Fuller	-1,1301	-2,1897	-2,6079
Valor crítico a 1%	-2,5688	-3,4428	-3,9766
Valor crítico a 5%	-1,9399	-2,8663	-3,4188
Valor crítico a 10%	-1,6159	-2,5693	-3,1316

Fonte: Resultados da pesquisa.

Como os valores de teste calculados, em termos absolutos, são menores que os valores críticos computados por MacKINNON (1991), tem-se que estes não indicam que se deve rejeitar a hipótese nula de presença de raízes unitárias, mesmo em níveis de significância menos rigorosos como o de 10%.

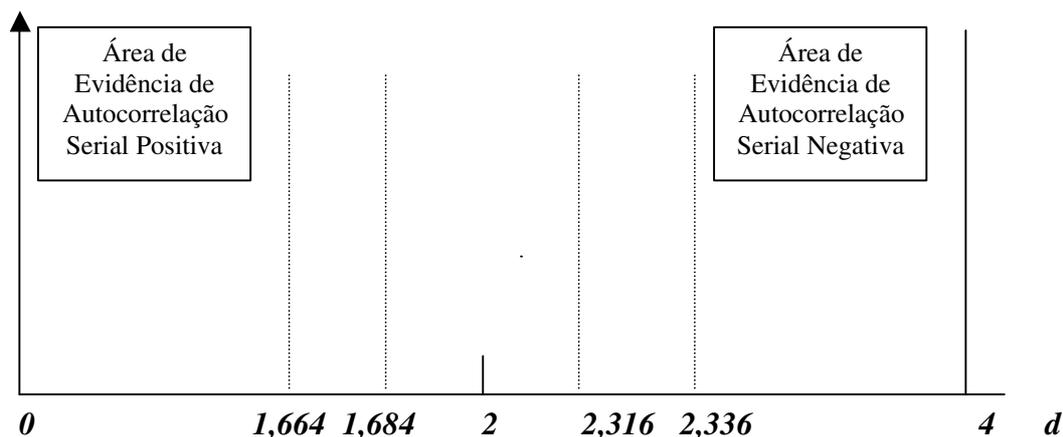
Esses resultados preliminares apontam ainda que o modelo que melhor representa o processo gerador da série de preços do café seria o modelo de passeio aleatório com intercepto e tendência linear dado pela equação (75). Isso porque, no nível de significância estatística de 10%, todos os coeficientes da equação (75) podem ser considerados significativamente diferentes de zero.

Todavia, esses resultados não podem ser considerados definitivos, pois o teste de Dickey-Fuller tem por base o pressuposto de que os erros das equações (73), (74) e (75) sejam não-correlacionados. Assim, foram comparados os valores calculados para as estatísticas d de Durbin-Watson com os valores tabelados tanto em nível de significância de 5% quanto de 1%, como se pode observar nas Figuras 19 e 20.



Fonte: Resultados da pesquisa.

Figura 19 - Resultados do teste de Durbin-Watson em nível de significância de 5%.



Fonte: Resultados da pesquisa.

Figura 20 - Resultados do teste de Durbin-Watson em nível de significância de 1%.

Os valores tabelados, quando comparados aos valores d calculados, indicaram que os resíduos das três equações apresentaram indícios de correlação serial positiva em ambos os testes. Isso implica que se deve utilizar outras versões de testes de raízes unitárias, como o de Dickey-Fuller Expandido (e, ou, o de Phillips-Perron), a fim de obter resultados mais confiáveis e que não sejam afetados por resíduos correlacionados.

Assim como no teste de Dickey-Fuller, na implementação do teste de Dickey-Fuller Expandido deve-se primeiramente se ater ao problema de incluir ou não apenas um termo constante, um termo constante e uma tendência linear, ou nenhum deles, na regressão do teste. De acordo com HAMILTON (1994), uma possibilidade seria realizar o teste com todos os termos, visto que os outros casos são casos particulares desta especificação, que é mais abrangente. Todavia, a inclusão de variáveis irrelevantes poderia reduzir o poder do teste, levando a conclusões erradas, como a da não-existência de raízes unitárias quando de fato elas ocorrerem e vice-versa. Deve-se ter como princípio básico escolher uma especificação do teste condizente com o tipo dos dados analisados. Se os dados

“aparentam” possuir uma tendência e a sua média é diferente de zero, deve-se incluir o termo de tendência e de intercepto no teste.

Assim, optou-se pelo teste com a inclusão do termo de intercepto (μ) e com o componente de tendência determinística (t), pois os seus respectivos coeficientes foram significativos em nível de confiança estatística de 10%, adotado no presente trabalho. Além disso, conforme discutido anteriormente, existem evidências de que os preços do café apresentam comportamento de tendência; portanto, foi utilizada esta versão do teste de Dickey-Fuller, uma vez que ela possibilita a verificação da ocorrência de tendência tanto do tipo determinística como do tipo estocástica. A inclusão do termo de intercepto no teste se deu pelo fato óbvio de que a média histórica dos preços do café é um valor positivo e diferente de zero.

Outro problema ligado à realização do teste se refere ao fato de como determinar o número ideal de defasagens (p) para a equação (28) do teste de Dickey-Fuller Expandido. Foram estimados modelos com número de defasagens crescentes, até que se encontrasse um modelo em que os coeficientes das diferenças mais altas não se mostrassem significativos. Dessa forma, realizaram-se testes para diferenças com uma, duas, três e quatro defasagens. Para o modelo com quatro defasagens, o coeficiente da quarta diferença dos preços não foi significativo, conforme se pode observar nos resultados do Quadro 4. Assim, optou-se pela especificação dada pelo modelo com três defasagens.

Pode-se observar também que o valor da estatística d de Durbin-Watson, que foi de 2,0006, indica que o problema de autocorrelação serial foi resolvido para este modelo. Outro aspecto que pode ser apontado como decisivo para a escolha do modelo com três defasagens se refere ao fato de que este foi o modelo estimado que minimizou simultaneamente os critérios de informação de Akaike e de Schwarz³⁹.

³⁹ Todavia, deve-se ressaltar que não foi encontrada nas diversas bibliografias consultadas, que discutem a questão do número de defasagens a se incluir no teste de Dickey-Fuller Expandido, nenhuma referência à minimização destes critérios como um indicador da escolha do valor de p . O uso desses indicadores se deu por iniciativa própria do autor.

Quadro 4 - Resultados do teste de Dickey-Fuller expandido

Def.	μ	β	$(\theta-1)$	φ_1	φ_2	φ_3	φ_4	d	AIC	Sch.
1	8,5741 (0,0045)	-0,0087 (0,0918)	-0,0228 (0,0016)	0,2172 (0,0000)	-	-	-	2,0387	9,1895	9,2172
2	9,1787 (0,0026)	-0,0090 (0,0828)	-0,0248 (0,0007)	0,1997 (0,0000)	0,0889 (0,0227)	-	-	2,0218	9,1864	9,2206
3	10,0641 (0,0010)	-0,0094 (0,0698)	-0,0277 (0,0002)	0,1916 (0,0000)	0,0673 (0,0874)	0,1222 (0,0017)	-	2,0006	9,1758	9,2168
4	10,1719 (0,0010)	-0,0095 (0,0672)	-0,0279 (0,0002)	0,1914 (0,0000)	0,0672 (0,0885)	0,1219 (0,0021)	0,0023 (0,9530)	1,9999	9,1803	9,2282

Fonte: Resultados da pesquisa.

Os valores críticos dos testes realizados, para determinação da ocorrência ou não de raízes unitárias nos preços do café, estão sistematizados no Quadro 5. Pode-se observar que os valores apresentados, de maneira geral, indicam que a série de preços reais do café brasileiro no mercado *spot* de Nova Iorque apresenta raiz unitária, isto é, ela é uma série não-estacionária em função da incidência de tendência do tipo estocástica e do tipo determinística, como se pode observar pela significância estatística dos termos $(\theta - 1)$ e β , dada pelos valores de probabilidade (*P-Values*) entre parênteses no Quadro 4.

Quadro 5 - Estatísticas dos testes de Dickey-Fuller expandido

Estatísticas do teste de Dickey-Fuller expandido	1 defasagem	2 defasagens	3 defasagens	4 defasagens
Modelo com intercepto e tendência determinística	-3,1641	-3,4198	-3,8097	-3,7871

Fonte: Resultados da pesquisa.

Obs.: Os valores críticos de MacKinnon para a rejeição da hipótese nula de presença de raízes unitárias em níveis de significância de 1% e 5% foram, respectivamente, de -3,9767 e -3,4188.

Para esse teste foi adotado um nível de significância estatística de 1%, devido ao fato de que os resultados encontrados em níveis menores foram pouco conclusivos no sentido de se apontar a presença ou não de raízes unitárias para os preços do café. Isso porque as conclusões podem variar em função do número de defasagens adotado. Assim, optou-se por um nível de significância mais rigoroso, a fim de se minimizar a possibilidade de incidência em erro do tipo I.

Os resultados encontrados indicam que os preços do café apresentam tendência estocástica (raiz unitária) e tendência determinística, implicando que a série de preços reais do café é estacionária por diferença (DSP) e em tendência (TSP). Assim, a tendência existente nos preços do café se configura como do tipo conjugado de tendência determinística e do tipo estocástica, isto é, ela é uma composição das duas formas.

No sentido de confirmar os resultados encontrados anteriormente, optou-se também pela realização de um outro tipo de teste de Raízes Unitárias, denominado Teste de Phillips-Perron. Os resultados desses testes serão apresentados a seguir.

3.1.3.2. Testes de Phillips-Perron

O teste de raízes unitárias de Dickey-Fuller Expandido se apresenta como uma importante alternativa ao teste simples de Dickey-Fuller, no sentido de que não tem como pressuposto que os erros sejam não-correlacionados. Pelo contrário, ele possibilita a correção do problema pela inclusão de termos de diferenças defasadas do preço do café no lado direito da equação.

Todavia, esse teste tem como pressuposto que os erros apresentam variância constante, isto é, não apresentem o problema de heterocedasticidade. A ocorrência desse problema pode lançar dúvidas sobre a confiabilidade dos resultados dos testes de raízes unitárias. Assim, foi testada a sua ocorrência.

Utilizando o teste de White (Apêndice B) para verificação da ocorrência do problema de heterocedasticidade nos resíduos da equação (75) do teste de

Dickey-Fuller, pôde-se constatar que, além do problema de correlação serial, os estes não apresentam variância constante.

O valor do coeficiente de determinação observado ($R^2*Observado$) foi de 87,1122, enquanto o valor da estatística χ^2 (Qui-quadrado) tabelada com cinco graus de liberdade⁴⁰ em nível de significância de 10% foi igual a 9,2363. Isso indica que existe o problema de heterocedasticidade nos resíduos da referida equação, ou seja, a variância dos resíduos da equação (75) estimada para a diferença dos preços do café não é constante para o período analisado.

Mesmo em testes com níveis de significância estatística mais rigorosos, como os de 5% e de 1%, verificou-se a ocorrência do problema de variância não-uniforme dos erros. Isso porque o valor das estatísticas χ^2 de qui-quadrado tabeladas com cinco graus de liberdade em nível de significância de 5% e de 1% foram, respectivamente, iguais a 11,0705 e 15,0863, que ainda assim são menores que o valor do $R^2*Observado$ de 87,1122, o que indica a incidência do problema.

Dessa forma, foram implementados também no presente trabalho testes de raízes unitárias de Phillips-Perron, a fim de confirmar, ou não, os resultados obtidos no teste de Dickey-Fuller Expandido. Conforme discutido anteriormente, este teste se constitui de uma generalização do teste de Dickey-Fuller, que, em vez de acrescentar termos de diferenças defasadas nas equações de teste, faz correções nas estatísticas-t dos coeficientes estimados. Tem-se, portanto, que o teste de Phillips-Perron “permite” que os resíduos sejam dependentes e com variância não-uniforme, pois as suas estatísticas de teste são modificações das estatísticas de Dickey-Fuller que levam em consideração uma natureza menos restritiva para os erros⁴¹.

Com relação ao número de defasagens de truncamento (q) adotado, optou-se pelo valor sugerido pela função proposta por NEWEY-WEST (1994). O valor calculado (Apêndice C) foi igual a 6, todavia também foram apresentados

⁴⁰ Dado que a equação estimada no teste possui cinco variáveis explicativas, sem a inclusão do intercepto.

⁴¹ Para mais detalhes, ver HAMILTON (1994).

resultados para valores de q maiores, dado que, conforme aponta HAMILTON (1994), com amostras grandes, como é o caso da série de preços do café, podem-se utilizar valores maiores para o número de defasagens de truncamento.

Os principais resultados dos testes de Phillips-Perron realizados estão representados no Quadro 6. Observa-se que os resultados encontrados confirmam as conclusões obtidas a partir dos testes de Dickey-Fuller Expandido, indicando que a série de preços reais do café apresenta raiz unitária em nível de significância estatística de 1%; mesmo em nível de significância menos rigoroso, de 5%, não se pode rejeitar a hipótese de existência de raízes unitárias nos preços do café.

Os resultados encontrados mostram que a tendência existente nos preços do café se configura como uma composição de tendências do tipo determinística e do tipo estocástica, ou seja, a série de preços do café brasileiro não é uma série estacionária e o seu comportamento de longo prazo é marcado pela junção de tendências do tipo determinística e do tipo estocástica. Assim, na sua modelagem e eliminação da série de preços, esse resultado deve ser levado em consideração.

Quadro 6 - Resultados dos testes de raízes unitárias de Phillips-Perron

Defasagens de truncamento	$q = 5$	$q = 6$	$q = 7$	$q = 8$
Estatística de P-P calculada (teste com intercepto e tendência determinística)	-3,2870	-3,3219	-3,3662	-3,4114
Estatística de P-P calculada (teste com intercepto e sem tendência determinística)	-2,9025	-2,9386	-2,9845	-3,0311

Fonte: Resultados da pesquisa.

Obs.: Os valores críticos de MacKinnon para a rejeição da hipótese nula de presença de raízes unitárias em níveis de significância de 1 e 5% foram, respectivamente, de -3,9767 e -3,4188 para o teste com intercepto e tendência determinística. Para o teste sem a tendência determinística, foram iguais a -3,4428 e -2,8663, respectivamente.

Conclui-se que, de acordo com os testes de raízes unitárias realizados para uma amostra de 660 meses (de janeiro de 1946 a dezembro de 2000) dos preços reais do café brasileiro no mercado *spot* de Nova Iorque, os mesmos apresentam raiz unitária. Assim, para tornar a série de preços do café estacionária, deve-se retirar a sua tendência determinística e tomar diferenças desta até que se obtenha uma série estacionária.

3.1.4. Modelagem da tendência determinística nos preços do café

Visto que uma parcela da tendência identificada como presente nos preços reais do café brasileiro no mercado *spot* da bolsa de Nova Iorque é do tipo determinística, buscou-se encontrar o melhor modelo para representá-la e, posteriormente, retirá-la da série para a análise dos demais componentes (ciclos, sazonalidade e volatilidade) que possivelmente possam estar nela presentes. Dessa forma, foram estimados (Apêndice D) os diversos modelos de tendência determinística apresentados anteriormente na metodologia do presente trabalho.

Primeiramente, estimou-se o modelo de tendência linear determinística dado pela equação (15). Neste modelo, a variável t , chamada de variável de tendência, será multiplicada por um coeficiente angular que irá indicar a direção da tendência nos preços do café. Se o valor do coeficiente for positivo, tem-se o caso de uma tendência de crescimento nos preços ao longo do período analisado e futuro; todavia, se o valor do coeficiente for negativo, tem-se o caso de uma tendência de queda nos preços.

Os resultados encontrados para a estimativa da equação (15) são apresentados a seguir:

$$\begin{aligned}
 PRCAFÉ_t &= 328,0305 - 0,2314 t && (76) \\
 &\quad (0,0000) \quad (0,0000) \\
 \bar{R}^2 &= 0,1024 \quad F = 76,1397
 \end{aligned}$$

Como se pode observar, tanto o termo constante de intercepto quanto o coeficiente angular da tendência foram significativos em nível de significância de 10%, dados os valores de probabilidade (*P-Values*) extremamente baixos,

calculados para esses coeficientes (entre parênteses). Além disso, o valor da estatística F também indica que os coeficientes estimados são, em conjunto, significativos.

O valor estimado para o coeficiente angular da tendência indica que no período amostral estudado, isto é, de janeiro de 1946 a dezembro de 2000, os preços reais do café brasileiro no mercado *spot* de Nova Iorque decresceram, em média, à taxa absoluta de 0,2314 US¢/libra-peso ao mês. Segundo este modelo, durante o período analisado teria prevalecido uma tendência determinística de queda nos preços da *commodity* brasileira no mercado mundial.

Todavia, conforme se poderia esperar, o valor do coeficiente de determinação ajustado encontrado (\bar{R}^2) para este modelo foi baixo, indicando que apenas 10% das variações nos preços do café são explicadas pela tendência linear, o que sugere que um modelo melhor especificado possa ser estimado.

A seguir, foi estimado o modelo de crescimento (ou de decréscimo) exponencial, representado pelas equações (19) e (20). Neste modelo, o coeficiente angular (φ_1) da equação (20) mede a variação relativa nos preços do café (Y) resultante de uma variação absoluta no tempo⁴²; isto é:

$$\varphi_1 = \frac{\frac{\Delta Y}{Y}}{\Delta t} \quad (77)$$

Multiplicando-se a mudança relativa em Y por 100, a expressão (77) fornecerá a taxa de mudança percentual, ou taxa de crescimento (decréscimo), em Y para uma mudança no tempo. Os resultados encontrados na estimativa da equação (20) para os preços do café foram os seguintes:

⁴² Conforme GUJARATI (1995), utilizando-se de cálculo diferencial, pode-se demonstrar que

$$\varphi_1 = \frac{d(\ln Y)}{dX} = \left(\frac{1}{Y} \right) \left(\frac{dY}{dX} \right) = \left(\frac{dY}{Y} \right) / dX, \text{ que é igual à razão entre a mudança relativa no regressando de (20) sobre a mudança absoluta no regressor, dada pela expressão (77).}$$

$$\begin{aligned} \ln PRCAF\acute{E}_t &= 5,8280 - 0,0013 t & (78) \\ & (0,0000) \quad (0,0000) \\ \bar{R}^2 &= 0,2299 \quad F = 197,7413 \end{aligned}$$

Observando os resultados encontrados, tem-se que tanto o intercepto quanto a taxa de decréscimo relativo foram significativos em nível de significância de 10%, dados os valores de probabilidade (*P-Values*) próximos de zero. Apesar de ser mais elevado do que o encontrado nos resultados da equação (76), os coeficientes de determinação ajustados (\bar{R}^2) dessas equações não podem ser diretamente comparados, pois os regressandos não são os mesmos. Na equação (76) estimou-se um modelo de tendência determinística para os preços reais do café, enquanto na equação (78) o regressando é o logaritmo natural desses preços.

Multiplicando a taxa de decréscimo relativo (-0,0013) por 100, obteve-se a taxa de decréscimo percentual para os preços do café. Assim, no período amostral estudado, os preços reais do café brasileiro no mercado *spot* de Nova Iorque decresceram, em média, à taxa de 0,13% ao mês. Portanto, o modelo (78) estimado também indica que nos últimos 55 anos prevaleceu uma tendência determinística de queda nos preços da *commodity* brasileira no mercado *spot* de Nova Iorque, confirmando a hipótese de Schultz acerca do comportamento de tendência sistemática de queda dos preços das *commodities* agrícolas para os preços do café brasileiro no mercado internacional.

Finalmente, foram estimados modelos de tendência determinística polinomial, dados por variações da equação (17), apresentada anteriormente. Os resultados obtidos para o modelo de tendência quadrática e para o modelo de tendência cúbica estão representados, respectivamente, pelas equações (79) e (80) a seguir.

$$\begin{aligned} PRCAF\acute{E}_t &= 230,6373 + 0,6567 t - 0,0013 t^2 & (79) \\ & (0,0000) \quad (0,0000) \quad (0,0000) \\ \bar{R}^2 &= 0,2035 \quad F = 85,1738 \end{aligned}$$

$$\begin{aligned}
PRCAFÉ_t = & 277,8819 - 0,2068 t + 0,0019 t^2 - 0,0000003 t^3 & (80) \\
& (0,0000) & (0,4015) & (0,0266) & (0,0001) \\
\bar{R}^2 = & 0,2197 & & & F = 62,8403
\end{aligned}$$

Os resultados do modelo de tendência determinística quadrática, dado pela equação (79), indicam que tanto o coeficiente ligado ao termo de tendência de grau um quanto o de grau dois são significativos em nível de significância estatística de 10%. Além disso, o termo constante de intercepto também se apresentou como significativo. Os valores estimados para os parâmetros do modelo (negativo para o termo de maior grau e positivo para o de menor) indicam que a tendência determinística dos preços do café possui a forma de uma parábola com a concavidade voltada para baixo.

Tem-se, portanto, que o nível médio dos preços da *commodity* brasileira apresentou, no início do período analisado, um comportamento de crescimento a taxas decrescentes. Este processo foi então revertido, e o nível médio dos preços passou a apresentar na segunda metade do período analisado um comportamento de queda a taxas crescentes, o que indica que a situação do setor cafeeiro pode-se tornar problemática nos próximos anos, caso essa tendência não seja revertida.

Como se pode observar, o valor de \bar{R}^2 da equação (80) foi um pouco maior que o apresentado pela equação (79), mesmo com o surgimento de coeficientes não-significativos, como o ligado ao termo de tendência de grau um, e com a ocorrência de um valor estimado próximo de zero (-0,0000003), para o coeficiente ligado ao termo de tendência de maior grau. Isso pode ter ocorrido, pois, de maneira geral, elevando o grau (k) do polinômio dos modelos de tendências polinomiais, podem-se obter resultados mais satisfatórios em termos de coeficientes de determinação ajustados (\bar{R}^2) mais elevados, ou seja, os modelos de maior grau de polinômio são relativamente mais eficientes na modelagem da tendência da série para o período selecionado como amostra; no entanto, eles apresentam sérios problemas quanto à sua capacidade de extrapolação para valores futuros desta.

Dessa forma, optou-se pelo modelo de tendência quadrática dado pela equação (79) como o mais representativo para o comportamento da tendência

determinística presente nos preços reais do café brasileiro no mercado *spot* de Nova Iorque, em razão de, neste modelo, todos os coeficientes serem significativos e nenhum deles ser tão próximo de zero como ocorre com o coeficiente ligado ao termo de maior potência na equação (80).

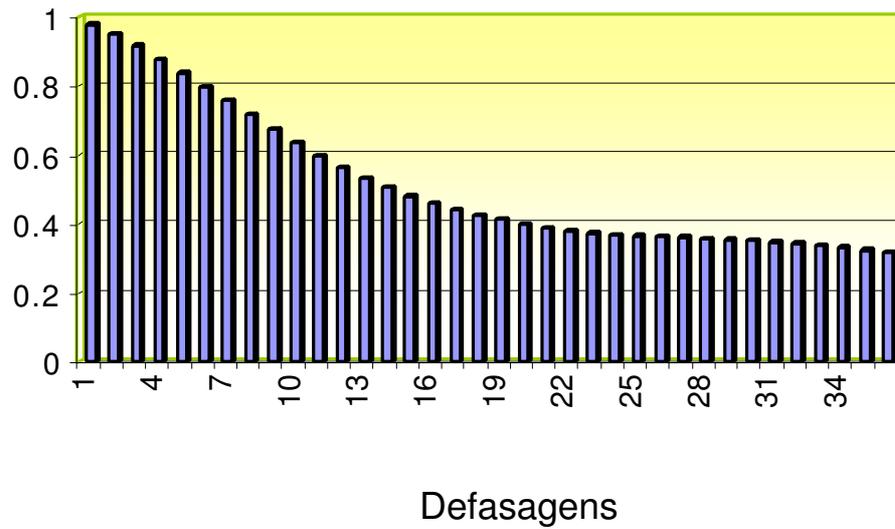
3.1.5. Modelagem da tendência estocástica nos preços do café

A análise dos demais tipos de componentes que podem estar presentes nos preços do café (sazonalidade, ciclos e volatilidade) deve ser implementada após a eliminação do componente de tendência dessa série. Isso deve ser feito a fim de que não ocorra sobreposição dos efeitos dos diferentes componentes, ou mesmo a impossibilidade de identificação destes, além de outros problemas possíveis discutidos anteriormente.

Tomou-se como ponto de partida para o prosseguimento da análise os resíduos da equação (79) que estão livres do comportamento de tendência determinística que havia sido diagnosticada como presente nos preços da referida *commodity*. Todavia, a tendência estocástica ainda se faz presente nessa nova série, como se pode observar pelo gráfico do correlograma dos resíduos da equação (79) na Figura 21.

Conforme se pode observar, o valor de primeiro coeficiente de autocorrelação é elevado (0,979), sendo seguido por uma série longa de valores das autocorrelações que apresentam comportamento de queda bastante lento. O coeficiente de autocorrelação demora 15 defasagens até cair abaixo de 0,5, o que é um indicador de que os preços do café sem a tendência determinística ainda se comportam como um processo não-estacionário, o que era esperado, já que a série apresenta raiz unitária.

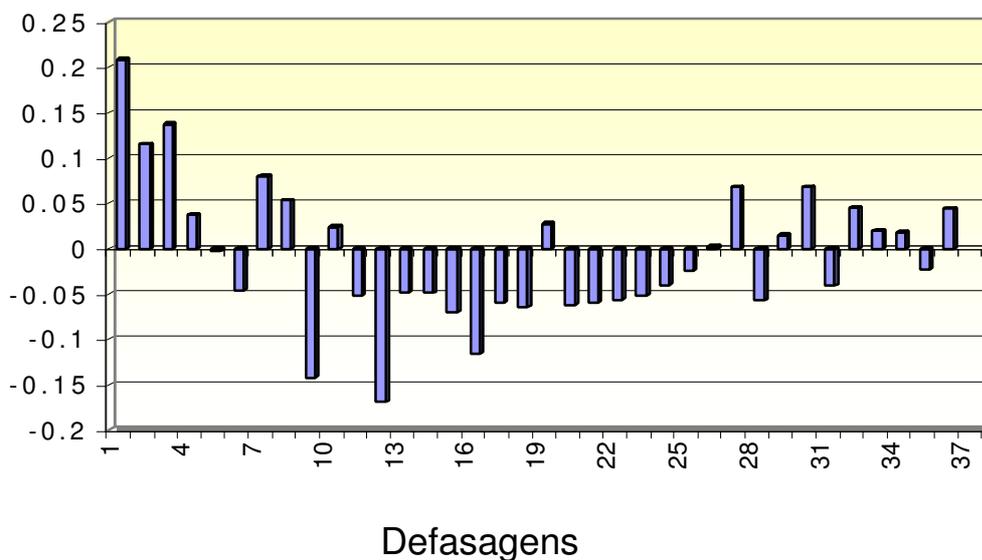
Autocorrelações



Fonte: Resultados da pesquisa.

Figura 21 - Correlograma amostral dos preços do café sem tendência determinística.

Autocorrelações



Fonte: Resultados da pesquisa.

Figura 22 - Correlograma amostral da série DPRCAFÉ.

Assim, foi tomada a primeira diferença da nova série de preços do café, já livre da tendência determinística, a fim de se eliminar também a tendência estocástica desta e torná-la uma série estacionária⁴³. Isso porque, após a modelagem da tendência, a série resultante deve ser estacionária. Não foi necessário diferenciar a série mais de uma vez, visto que apenas uma diferença tornou a mesma estacionária.

Foram realizados também novos testes de raízes unitárias, os quais confirmaram que esta nova série (DPRCAFÉ) é de fato estacionária e não possui raiz unitária. Os resultados dos novos testes podem ser visualizados no Quadro 7.

Quadro 7 - Resultados dos testes de raízes unitárias para DPRCAFÉ

	Teste de Dickey-Fuller expandido	Teste de Phillips-Perron
Resultado encontrado	-10.1816	-21.1731
Valor crítico a 1%	-2.5688	-2.5688
Valor crítico a 5%	-1.9399	-1.9399
Valor crítico a 10%	-1.6159	-1.6159

Fonte: Resultados da pesquisa.

Como se pode verificar, tanto o teste de Dickey-Fuller Expandido quanto o teste de Phillips-Perron indicaram que a série DPRCAFÉ não apresenta raiz unitária, sendo, portanto, uma representação estacionária dos preços do café no mercado *spot* da bolsa de Nova Iorque. Esta série será utilizada no desenvolvimento do trabalho nas seções subsequentes.

⁴³ Esta nova série será chamada de DPRCAFÉ a partir desse ponto.

3.2. Análise do componente de ciclos

Após o estudo e a remoção da tendência estocástica e da tendência determinística da série de preços do café brasileiro no mercado *spot* da bolsa de Nova Iorque, passou-se à análise do componente de ciclos que esses preços possam apresentar. O processo implementado anteriormente, de eliminação da tendência na série, é tratado na metodologia da análise de séries temporais no domínio da frequência (Análise Espectral) como processo de *pré-branqueamento* da série, e faz-se necessário sempre que a série a ser analisada não seja estacionária.

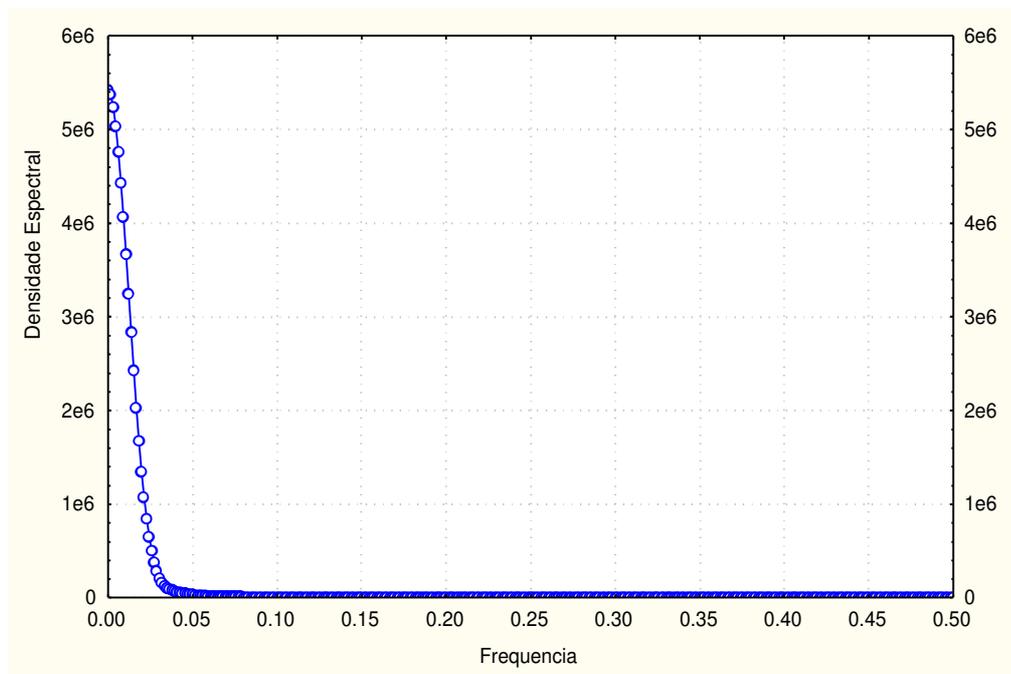
O pré-branqueamento da série é necessário porque a análise espectral se baseia na suposição geral de que as séries temporais são periódicas, e a estimação de um espectro para um intervalo finito de tempo resulta na caracterização de uma série que corresponde àquele intervalo, mas que se repete periodicamente além do intervalo de tempo tomado como amostra. Isso pode se tornar problemático à medida que os dados dessa amostra possuam tendência, porque no final do intervalo ocorrerá um “salto” de volta ao ponto inicial, indicando que ocorre aí um ciclo de duração igual ao tamanho do intervalo de tempo (n).

Assim, caso o componente de tendência não seja retirado da série antes da implementação da análise espectral, o espectro estimado tenderá a apresentar um pico extremamente elevado na frequência (ω) zero, e essa frequência irá “contaminar” o espectro desejado, inviabilizando a visualização e análise dos verdadeiros ciclos que possam estar presentes na série estudada.

O espectro estimado para os preços reais do café (PRCAFE) no mercado *spot* de Nova Iorque (Figura 23) confirma essa observação. Como pode ser observado, o pico no espectro no intervalo de frequências próximas de zero se destaca, impossibilitando a visualização de ciclos em outras frequências⁴⁴. Esse

⁴⁴ Pode-se concluir da discussão anterior que a estimação do espectro de uma série pode se configurar como um teste alternativo aos utilizados tradicionalmente para se verificar se esta é estacionária ou não (tal como os testes de raízes unitárias e a análise do correlograma da série). O procedimento do teste seria tal que, sempre que estimado o espectro da série e detectado que existe um pico extremamente significativo na frequência zero, concluir-se que ela é não-estacionária.

resultado reforça a idéia de que, se a remoção da tendência na série de preços do café não tivesse sido implementada previamente, o estudo dos ciclos nesses preços poderia se tornar inviável.



Fonte: Resultados da pesquisa.

Figura 23 - Espectro estimado dos preços do café antes do pré-branqueamento.

Este problema denominado *Leakage* se constitui numa violação de um dos pressupostos básicos para a implementação da análise de séries no domínio da frequência, que é a de que a série deve ser estacionária. Portanto, esse problema deve ser evitado tornando-se a série que será analisada estacionária, o que de fato foi feito para os preços do café na seção anterior. Nesse sentido, para a efetiva implementação da análise espectral dos preços do café na busca e análise de ciclos de preços, tomou-se a série DPRCAFÉ, resultante da eliminação da componente de tendência obtida no item 3.1 do presente trabalho.

A discussão que segue será dividida em duas partes: a primeira compreende uma análise da ocorrência de ciclos tendo por base todos os dados

disponíveis para o período entre janeiro de 1946 e dezembro de 2000. Na segunda parte, os dados serão divididos em décadas (de 50, de 60, de 70, de 80 e de 90⁴⁵), que serão analisadas individualmente na busca da determinação de mudanças na periodicidade e estrutura dos ciclos que possam estar presentes nos preços do café. Isso se torna importante, pois pode evidenciar possíveis mudanças na periodicidade e amplitude dos ciclos provocadas por: impactos de novas tecnologias; mudanças nas políticas agrícolas; novas relações concorrenciais e econômicas entre os agentes atuantes no setor cafeeiro; e por possíveis mudanças provocadas pelo início das negociações de contratos futuros de café na Bolsa de Mercadorias e Futuros brasileira (BM&F).

3.2.1. Análise do período de janeiro de 1946 a dezembro de 2000

Um aspecto inicial ligado à análise espectral dos preços do café se refere à determinação do ponto de truncamento (M) a ser adotado para a estimação do espectro da série. Conforme discutido anteriormente, um valor muito baixo para M pode levar à estimativa de um espectro demasiadamente “picotado”, que dificultará a distinção entre os ciclos verdadeiros e de maior importância e os ciclos de menor importância ou meramente frutos do acaso, pois a variância do espectro estimado com um valor pequeno para o ponto de truncamento (M) será alta, e este apresentará semelhança muito grande com o periodograma amostral, o que o tornará de difícil interpretação.

Por outro lado, a utilização de um valor demasiadamente alto para M poderá propiciar a estimativa de um espectro muito “suave”, implicando o encobrimento de informações (picos) importantes que os preços do café possam apresentar. Em última instância, isso pode implicar a conclusão de que não há ocorrência de ciclos nos preços quando eles de fato existem.

⁴⁵ Os dados referentes aos anos de 1946, 1947, 1948 e 1949 serão descartados nesta parte da análise, o que possivelmente não comprometerá a qualidade das análises, dado que o número de observações para cada década ainda é elevado ($n > 100$).

Optou-se, portanto, pela adoção do procedimento proposto por Parzen, que é discutido em FRANÇA (1990), de se trabalhar com espectros estimados não apenas com um ponto de truncamento, mas com três, a fim de se evitar a possibilidade de se incorrer em um dos problemas mencionados anteriormente. Parzen sugere que o primeiro ponto ($M1$) deve ser um valor entre 5 e 10% do número de observações (n) utilizado⁴⁶; o segundo ponto deve ser igual a duas vezes o valor do primeiro ($M2=M1 \times 2$); e, finalmente, o terceiro valor a ser utilizado para o ponto de truncamento deve ser igual ao dobro do valor do segundo ponto ($M3 = M2 \times 2$). Como no presente caso tem-se uma amostra de 660 observações⁴⁷, os valores definidos para os pontos de truncamento utilizados na estimativa dos espectros dos preços do café foram: $M1 = 33$, $M2 = 65$ e $M3 = 131$.

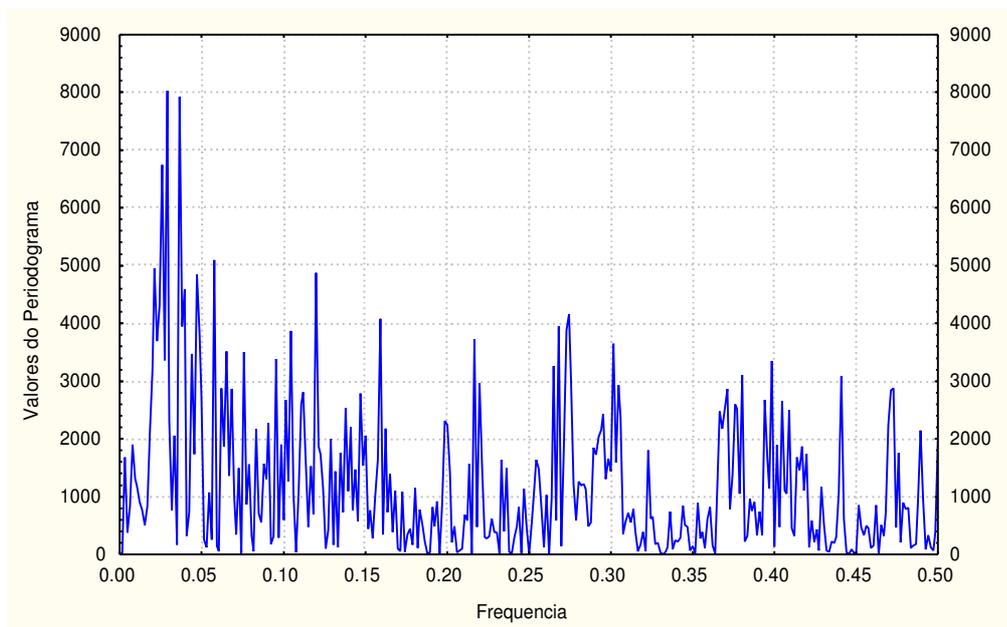
Outro aspecto que deve ser mencionado se refere à escolha da janela de defasagem a ser utilizada para a suavização do periodograma amostral e estimativa dos espectros. Conforme discutido na metodologia deste trabalho, existem diversas opções de janelas de defasagem que podem ser utilizadas para esse fim, como a janela de Danniell, a de Bartlett, a de Tukey e a de Parzen.

Optou-se pela utilização da janela de Parzen devido a três pontos distintos, a saber: as janelas de defasagem apresentam, de maneira geral, resultados semelhantes em termos dos espectros estimados; a janela de Parzen é uma das mais tradicionais, sendo utilizada na maioria dos trabalhos encontrados sobre análise espectral (além de estar disponível em praticamente todos os *softwares* estatísticos que possuem a rotina de análise espectral); e, principalmente, a janela de Parzen possui a vantagem de não gerar estimativas negativas para as densidades espectrais, que não têm interpretação razoável e que podem surgir em estimativas feitas por outras janelas, como a de Tukey.

⁴⁶ De acordo com o *Software Statistica 5.0*, utilizado no presente trabalho, M deve ser um valor ímpar e n deve ser um valor par.

⁴⁷ A observação perdida após a diferenciação da série foi substituída por um valor 0. Isso foi feito no sentido de evitar problemas na estimativa dos espectros, como o mencionado na nota de rodapé anterior, e de seguir uma proposição feita por Parzen, em FRANÇA (1990). Segundo ele, o número de pontos no intervalo de frequência $[0, \pi]$ de um espectro estimado, originário de uma série temporal mensal de dados econômicos, deve ser um número múltiplo de 12.

Observando a estimativa do periodograma para os preços do café apresentada na Figura 24, pode-se concluir que a suavização deste de fato se faz necessária, pois a sua análise direta se torna inviável, dada a grande volatilidade que ele apresenta, conforme discutido anteriormente. Aparentemente esse periodograma sugere a presença de uma série de ciclos diferentes existentes nos preços do café; todavia, os seus resultados devem ser filtrados, a fim de identificar quais são as frequências referentes a ciclos relevantes e quais as relativas a ciclos pouco significativos.



Fonte: Resultados da pesquisa.

Figura 24 - Periodograma estimado para os preços do café.

Tomando os valores definidos anteriormente para os pontos de truncamento da janela de Parzen, como $M1 = 33$, $M2 = 65$ e $M3 = 131$, estimou-se o espectro dos preços do café, com o objetivo de identificar se existem ciclos periódicos nestes. O espectro foi estimado inicialmente em função das frequências (ω), visando identificar entre quais intervalos de frequências $[\omega_0, \omega_1]$ se situam os ciclos mais significativos.

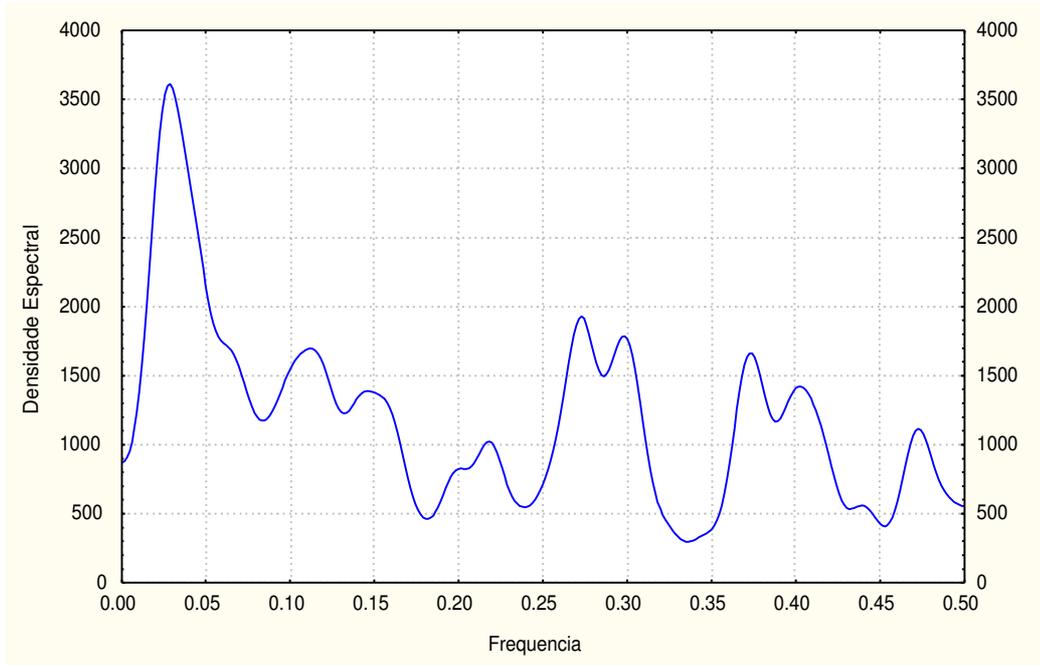
No entanto, estimou-se também o espectro em função dos períodos (P), pois ele possibilita uma observação direta da duração dos diferentes ciclos que possam estar presentes nos preços do café. Assim, têm-se nas Figuras 25 e 26 os resultados encontrados para esses dois espectros.

Apesar de um tanto quanto picotado, o espectro estimado em função das frequências indica que as do intervalo dado aproximadamente por $[0,022; 0,045]$ são as que mais contribuem para a variância dos preços do café. Isso porque existe um pico nesse espectro correspondente a esse intervalo de frequências que se destaca e que, aparentemente, é significativo. Podem-se notar também dois picos menores no intervalo de frequências dado por $[0,25; 0,30]$, porém estes, aparentemente, não são de fato relevantes.

Em termos de duração, os ciclos correspondentes a essas frequências teriam um período mínimo de aproximadamente 22 meses (correspondente à frequência de 0,045) e máximo de 44 meses (definido pela frequência de 0,022) para o pico mais alto; e de aproximadamente quatro meses para os picos entre as frequências do intervalo $[0,25; 0,30]$. Isso indica, portanto, a existência de um ciclo de duração mais longa e um ciclo de curto prazo para esses preços.

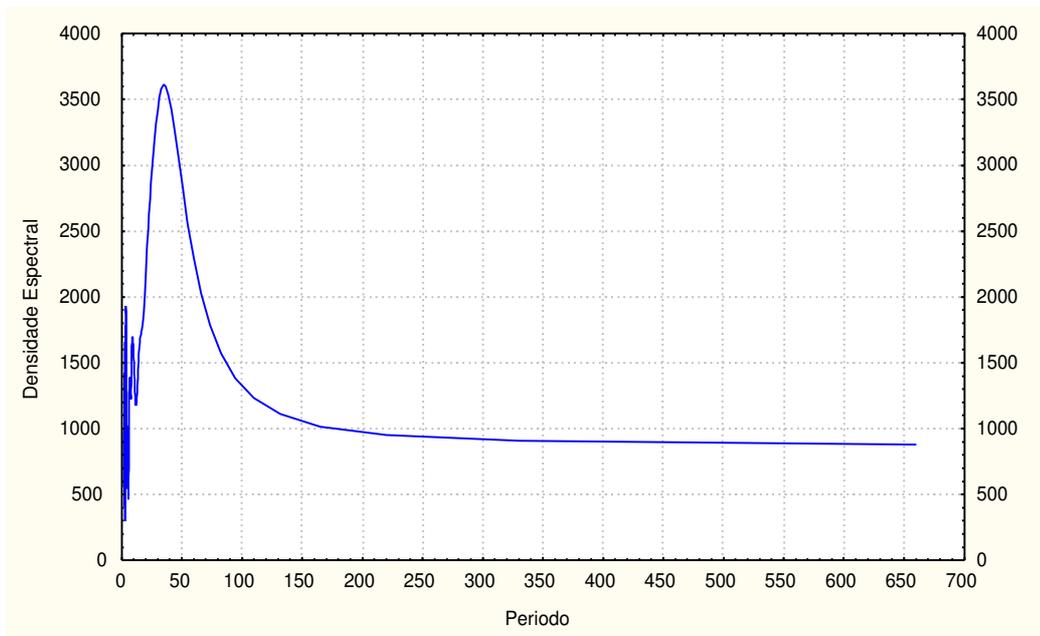
O espectro estimado tendo como eixo horizontal os valores dos períodos dado na Figura 26 representa uma outra maneira de se visualizar a existência de ciclos na série de preços do café. Neste, o período de duração de cada ciclo é diretamente observável, sem a necessidade de conversão de frequência em períodos.

Este espectro confirmou a incidência de um pico maior existente no intervalo de 22 meses a 44 meses, o que também pode ser comprovado observando-se os valores para as densidades espectrais encontrados no Quadro 8.



Fonte: Resultados da pesquisa.

Figura 25 - Espectro estimado para os preços do café de janeiro de 1946 a dezembro de 2000, com $M = 33$ (em função de ω).



Fonte: Resultados da pesquisa.

Figura 26 - Espectro estimado para os preços do café de janeiro de 1946 a dezembro de 2000, com $M = 33$ (em função de P).

Quadro 8 - Maiores valores estimados para as densidades espectrais dos preços do café com $M = 33$

Frequência (ω)	Período (P)	Coeficientes		Valores do periodograma	Densidades espectrais
		(Coseno)	(Seno)		
.028788	34.73684	-4.12124	-2.70638	8022.011	3613.118
.027273	36.66667	3.05165	-.93544	3361.915	3600.334
.030303	33.00000	-.98035	2.46485	2322.068	3579.865
.025758	38.82353	-3.13242	3.25916	6743.291	3537.723
.031818	31.42857	1.21399	-.91259	761.171	3511.782
.024242	41.25000	3.35968	-1.32788	4306.731	3424.561
.033333	30.00000	-1.04550	2.26920	2059.975	3421.767
.034848	28.69565	.46999	-.54684	171.576	3319.319
.022727	44.00000	1.25145	3.10459	3697.532	3263.928
.036364	27.50000	-4.47815	1.99434	7930.306	3210.528
.037879	26.40000	3.07311	1.58256	3943.002	3096.626
.021212	47.14286	1.22541	-3.67690	4957.009	3059.862
.039394	25.38462	-.67297	-3.66616	4584.885	2979.712
.040909	24.44444	.86082	.48340	321.644	2861.436
.019697	50.76923	1.78291	2.55726	3207.051	2820.495

Fonte: Resultados da pesquisa.

De acordo com as estimativas dos 15 maiores valores das densidades espectrais, os preços do café apresentaram no conjunto do período analisado um comportamento cíclico de duração dada pelo intervalo de 22 a 44 meses aproximadamente. Esses resultados indicam que, em média, poder-se-ia se estabelecer um período de 36 meses de duração para este ciclo. De acordo com o espectro estimado, com uma defasagem de truncamento $M = 33$ para o conjunto de observações compreendidas entre janeiro de 1946 e dezembro de 2000, os preços do café brasileiro no mercado *Spot* de Nova Iorque apresentaram um padrão de comportamento cíclico que se completou de três em três anos aproximadamente.

Conforme discutido anteriormente na metodologia referente à análise espectral, se as observações na série de preços do café fossem de fato independentes entre si, não apresentando repetições cíclicas, os valores

calculados em seu periodograma deveriam se comportar como um ruído branco e apresentar um espectro dado por uma linha relativamente constante.

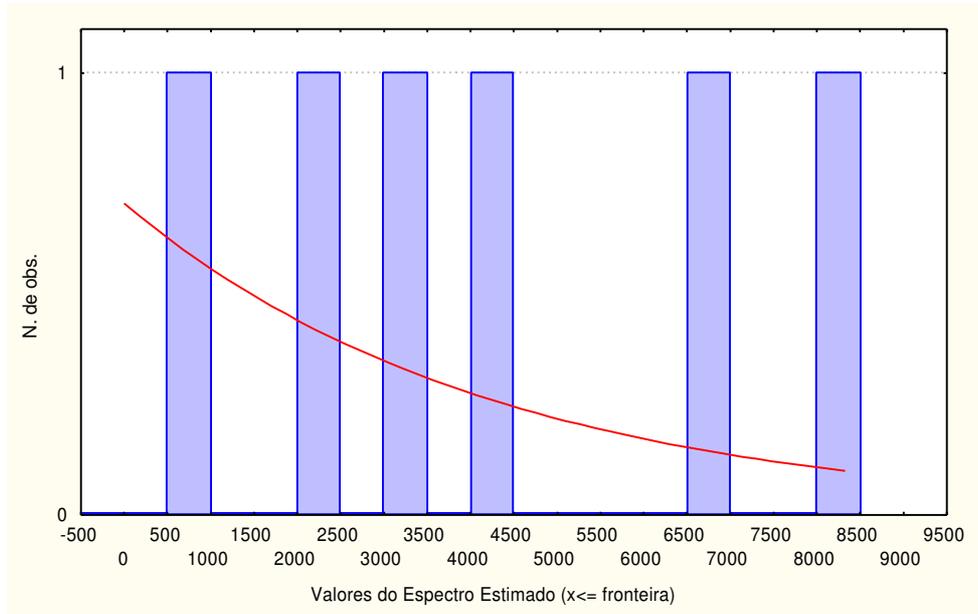
Foram, portanto, implementados testes não-paramétricos, descritos anteriormente, para comparar os valores estimados do periodograma do café com uma distribuição exponencial, a fim de identificar se os picos encontrados nos espectros estimados representam ciclos reais nos preços, ou se estes se comportam de maneira similar a uma série do tipo ruído branco, caso no qual não se configuram como ciclos estatisticamente significativos. Nas Figuras 27 e 28 estão indicados os resultados dos testes implementados.

Nessas figuras, a linha vermelha indica o comportamento esperado dos valores do espectro estimado caso os ciclos compreendidos entre as frequências e os períodos testados não fossem significativos. Isto é, de acordo com este teste, os valores estimados, caso fossem relativos a uma série do tipo ruído branco, deveriam seguir a forma de uma distribuição exponencial que é representada pela linha, o que, para os valores encontrados, não ocorre.

Isso implica que, de acordo com este teste, os ciclos encontrados de duração aproximada de três anos (que caracteriza uma média frequência), e o ciclo de frequência mais alta, correspondente a um período de aproximadamente quatro meses, são significativos, e a evidência mais forte se dá para o primeiro.

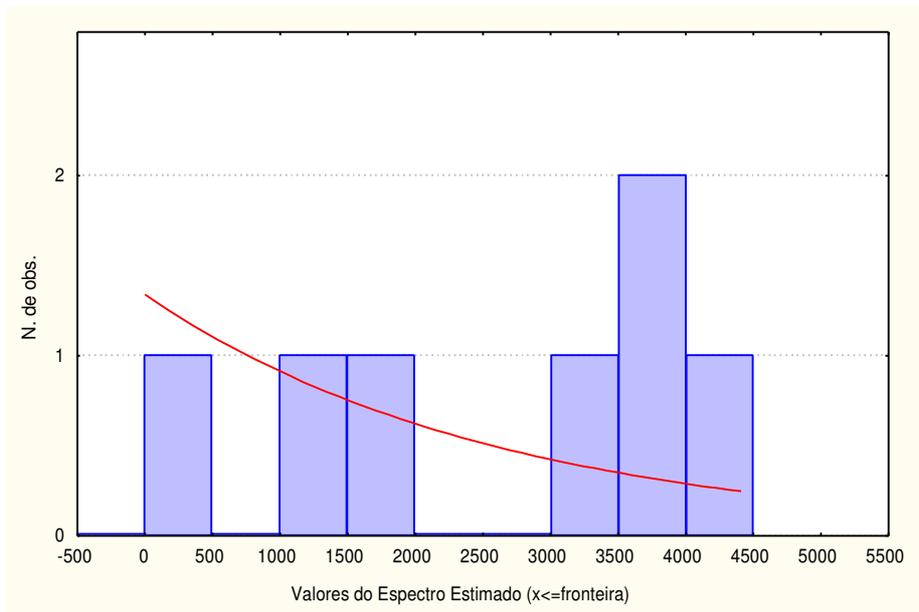
Os resultados encontrados para os espectros estimados com defasagem de truncamento (M) igual a 65 (Figuras 29 e 30) e para defasagem de truncamento igual a 131 (Figuras 31 e 32), de maneira geral, ratificam os resultados anteriores.

No entanto, deve-se ressaltar que os espectros estimados com $M = 131$ apresentaram comportamento demasiadamente suave, o que dificulta a identificação das frequências e períodos mais importantes para a caracterização dos ciclos nos preços do café.



Fonte: Resultados da pesquisa.

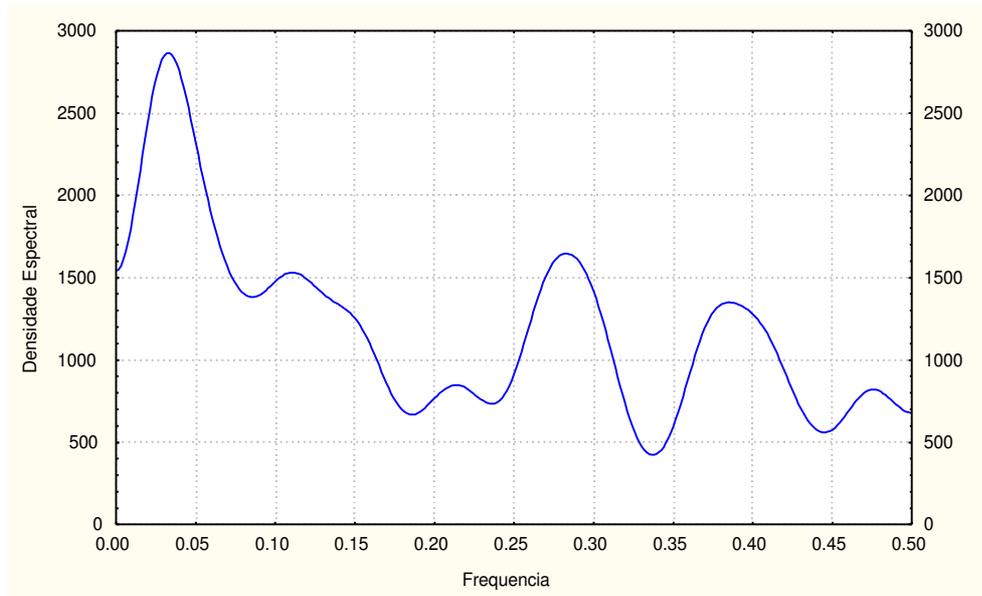
Figura 27 - Histograma do periodograma para teste dos ciclos de longo prazo (entre 22 e 44 meses) nos preços do café.



Fonte: Resultados da pesquisa.

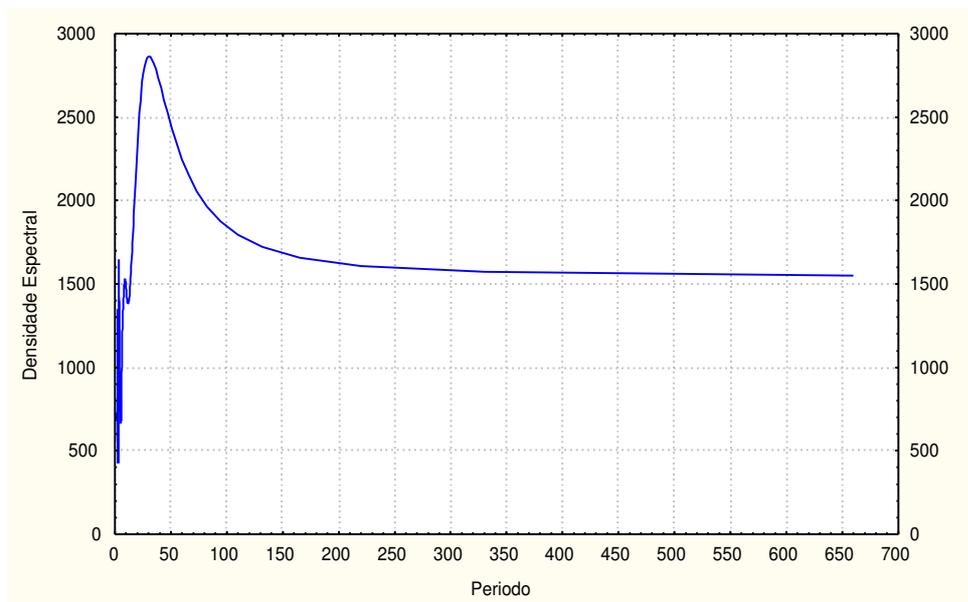
Figura 28 - Histograma do periodograma para teste dos ciclos de curto-prazo⁴⁸ nos preços do café.

⁴⁸ Compreendidos no intervalo de frequências [0,268; 0,277] e de período aproximado de quatro meses.



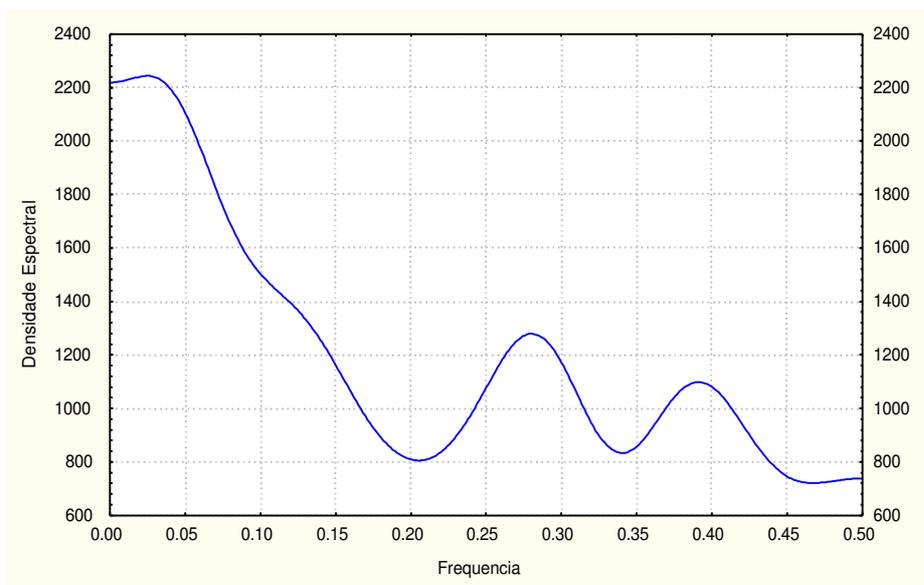
Fonte: Resultados da pesquisa.

Figura 29 - Espectro estimado para os preços do café de janeiro de 1946 a dezembro de 2000, com $M = 65$ (em função de ω).



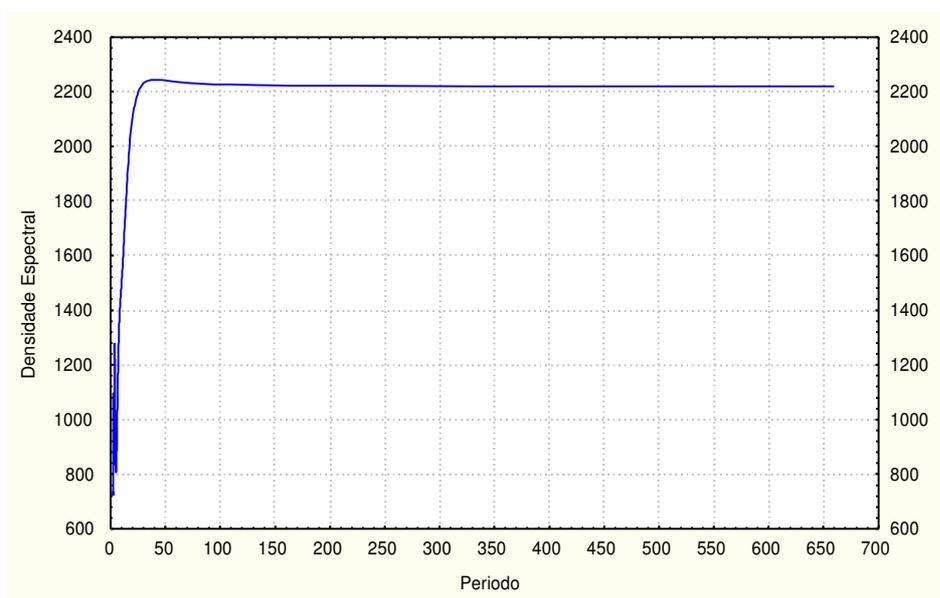
Fonte: Resultados da pesquisa.

Figura 30 - Espectro estimado para os preços do café de janeiro de 1946 a dezembro de 2000, com $M = 65$ (em função de P).



Fonte: Resultados da pesquisa.

Figura 31 - Espectro estimado para os preços do café de janeiro de 1946 a dezembro de 2000, com $M = 131$ (em função de ω).



Fonte: Resultados da pesquisa.

Figura 32 - Espectro estimado para os preços do café de janeiro de 1946 a dezembro de 2000, com $M = 131$ (em função de P).

Por outro lado, os espectros estimados para $M = 65$ tornaram mais nítido o pico referente ao ciclo de alta frequência, compreendido no intervalo de frequência $[0,268; 0,277]$ e de período aproximadamente igual a quatro meses, sem contudo evidenciá-lo como significativo, como ocorreu com o ciclo de longo prazo. Pelo contrário, esses espectros estimados com $M = 65$ e $M = 131$ tornaram esse pico menor, em relação ao estimado com $M = 33$.

3.2.2. Análise espectral dos preços do café nas décadas de 50, 60, 70, 80 e 90

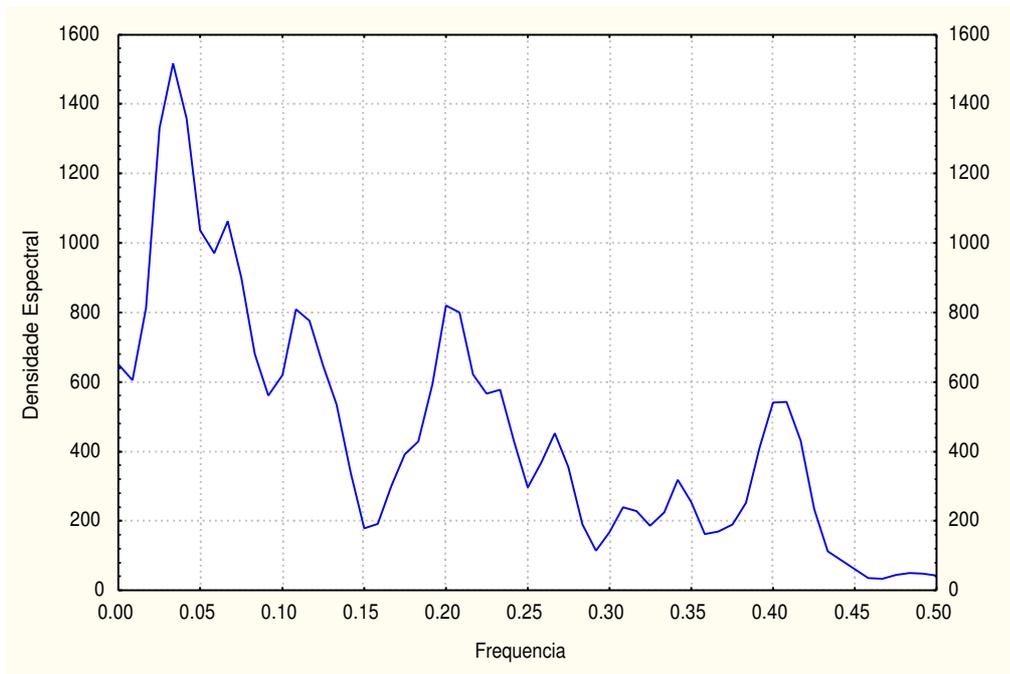
Conforme ressaltado anteriormente, os dados também foram divididos em décadas (de 50, de 60, de 70, de 80 e de 90), visando analisar se os ciclos nos preços do café apresentaram comportamento diferenciado entre as décadas distintas, o que implicaria mudanças estruturais no comportamento cíclico da série analisada. Isso poderia ter ocorrido em função de fatores particulares referentes a cada década, como: impactos de novas tecnologias sobre o processo produtivo do café; mudanças nas políticas agrícolas de tempos em tempos; mudanças nas relações concorrenciais no mercado internacional entre os agentes atuantes no setor cafeeiro; e, particularmente na década de 90, possíveis mudanças provocadas pelo início das negociações de contratos futuros de café na Bolsa de Mercadorias e Futuros brasileira (BM&F).

Na análise individual das décadas de 50 até a década de 90, as amostras de dados apresentaram 120 observações cada. Assim, os valores definidos para os novos pontos de truncamento utilizados na estimativa dos espectros dos preços do café foram, de acordo com a regra proposta por Parzen: $M1 = 7$, $M2 = 13$ e $M3 = 27$.

Na Figura 33 está representado o espectro estimado para os preços na década de 50 em função das frequências para $M1 = 7$ ⁴⁹. O espectro estimado aponta que as frequências próximas a 0,04 são as que mais contribuem para a variância dos preços do café, visto que o maior pico que se pode observar neste

⁴⁹ Os espectros para as defasagens de truncamento $M2 = 13$ e $M3 = 27$ se tornaram muito suavizados.

espectro se encontra no intervalo próximo a esta frequência. Podem-se notar também picos menores em frequências próximas a 0,07, 0,11 e 0,21. Apesar de menores, os testes de significância estatística realizados (Apêndice E) indicam que estes picos também são significativos.



Fonte: Resultados da pesquisa.

Figura 33 - Espectro estimado para os preços do café na década de 50, com $M I = 7$ (em função de ω).

Em termos de duração, os ciclos correspondentes a essas frequências seriam, respectivamente, de 24 meses para o pico mais alto com frequências próximas a 0,04; de 12 meses para o pico com frequências próximas a 0,07; de nove meses para o pico com frequências próximas a 0,11; e de aproximadamente cinco meses para o pico entre as frequências do intervalo [0,19; 0,22].

O resultado ligado ao pico de período igual a 24 meses indica que para a década de 50 os preços do café apresentaram ciclo bienal que possivelmente

esteve ligado diretamente ao ciclo biológico do cafeeiro. Este ciclo biológico consiste do fato de que, de maneira geral, os cafeeiros apresentam em um ano uma grande florada e no outro a seguir uma florada pequena, em razão do esgotamento das plantas, que necessitam da ocorrência de um intervalo de um ano para poder apresentar bom rendimento. Portanto, fatores biológicos implicam que, caso não ocorram geadas nem grandes oscilações nos índices pluviométricos, uma grande safra de café tenderá a ser seguida por uma safra menor no ano subsequente, o que se caracteriza como um ciclo de dois anos (24 meses) de duração na produção de café.

Conforme ressaltado na metodologia do presente trabalho, os ciclos na produção de *commodities* agrícolas geralmente têm como contrapartida ciclos nos preços desses produtos, o que é representado de maneira formal pelo teorema da teia de aranha. Dessa forma, o ciclo bienal existente nos preços do café na década de 50 pode ser configurado como uma sustentação empírica para esse teorema e para a hipótese do ciclo bienal na produção de café.

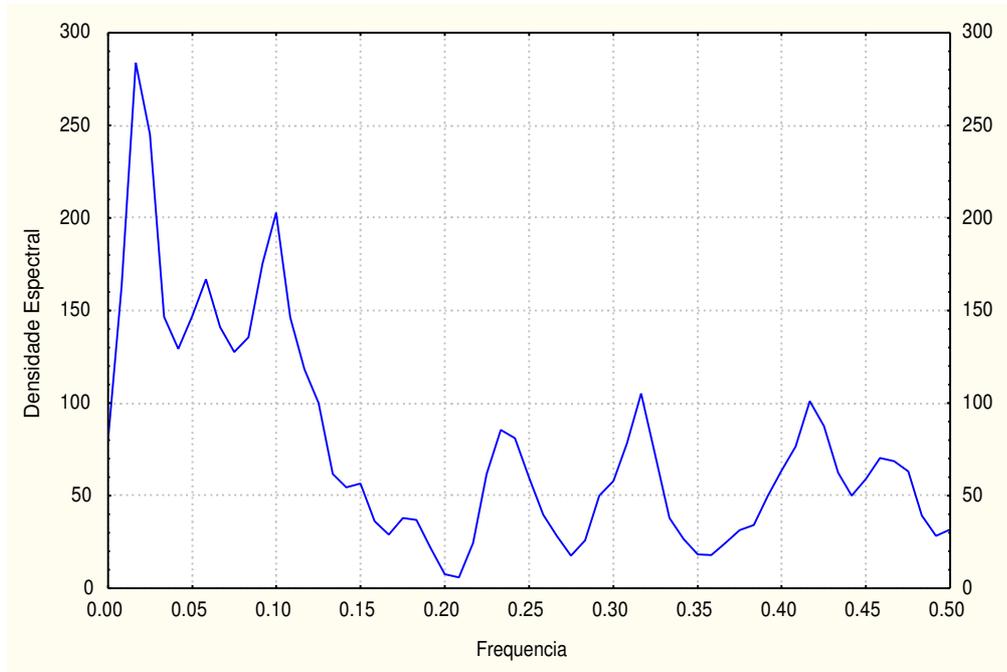
O pico próximo à frequência $\omega = 0,07$, apesar de pouco nítido, possivelmente em função do valor de MI , indica que na década de 50 os preços também apresentaram ciclo de duração de um ano (12 meses), que pode ser relacionado a um possível padrão sazonal existente no comportamento desses preços⁵⁰.

No entanto, os ciclos de frequência mais altas, relativos aos períodos de nove e de cinco meses, não possuem, a princípio, qualquer conotação com processos naturais e biológicos de produção dos cafeeiros. Possivelmente, eles devem estar ligados a outros fatores conjunturais de natureza econômica que possam ter ocorrido nos anos 50, nos mercados interno e externo, ou mesmo a fatores estruturais que determinaram esse comportamento para os preços.

Para a década de 60, os resultados encontrados (Figura 34) divergiram, em certa medida, dos resultados verificados para a década de 50. Nesse período, o pico mais marcante e significativo ocorreu nas frequências próximas a 0,02, o

⁵⁰ Todavia, dada a importância que este componente de uma série temporal possui, ele será discutido separadamente a seguir, conforme definido anteriormente.

que, em termos de período, corresponderia a um ciclo nos preços do café de aproximadamente 60 meses (cinco anos) de duração.



Fonte: Resultados da pesquisa.

Figura 34 - Espectro estimado para os preços do café na década de 60, com $M I = 7$ (em função de ω).

Esse ciclo de prazo mais longo no comportamento dos preços do café pode se caracterizar como uma comprovação empírica do chamado *ciclo plurianual* dos preços do café, destacado por PEROBELLI e SCHOUCHANA (2000:13), que seria determinante do comportamento desses preços. Esses autores ressaltam que:

“Considerando-se inicialmente uma fase de redução da produção, que, associada a um consumo constante ou crescente, implicará a necessidade de utilização de estoques a fim de suprir tal demanda. A utilização destes estoques acarretará elevação dos preços, que estimulará a expansão de novos cafezais. Essa expansão levará ao aumento dos estoques e, depois, a uma queda dos

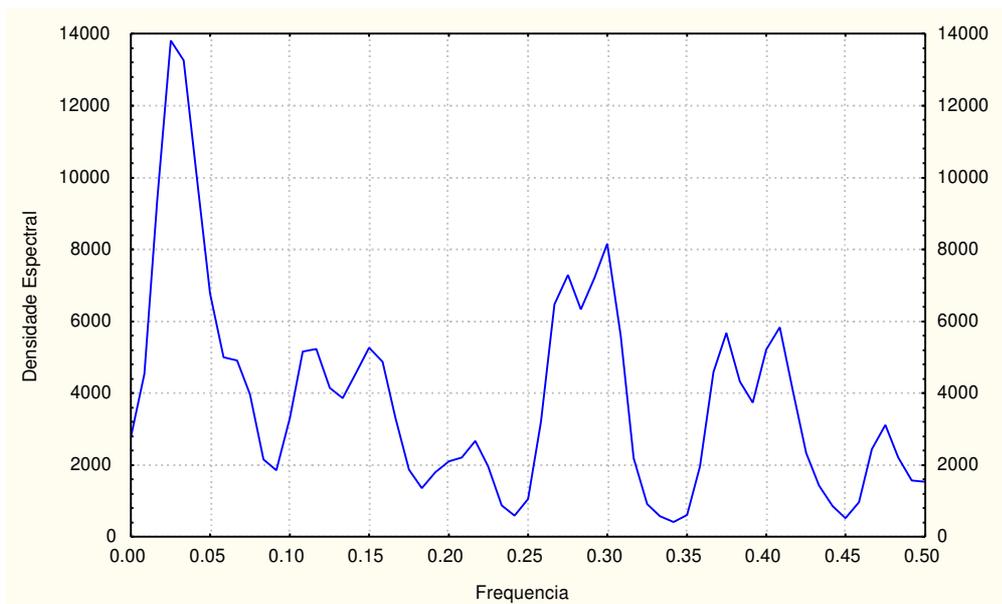
preços, desestimulando a produção cafeeira. E então, está estabelecido um novo ciclo.”

Pode-se observar ainda que o ciclo sazonal (de 12 meses) e o ciclo de nove meses, correspondentes aos picos nas frequências próximas a 0,07 e 0,11, respectivamente, também foram significativos para os dados da década de 60. Por outro lado, o pico próximo à frequência de 0,2, que era marcante no espectro dos preços nos anos 50 e que corresponderia a um ciclo de duração de cinco meses, não se mostrou evidente para esse período.

Com relação ao ciclo bienal, apesar de não estar evidente nos resultados estimados para a defasagem de truncamento $M1=7$, tanto o periodograma estimado como o espectro com $M2=13$ (Apêndice E) indicam que ele é importante para explicar a variância dos preços do café nos anos 60, isto é, de acordo com esses resultados, este ciclo também foi determinante no período em questão.

O espectro estimado para a década de 70 (Figura 35) indica novamente que as frequências próximas a 0,04 são as que mais contribuem para a variância dos preços do café. O maior pico verificado nesse espectro se encontra no intervalo próximo a essa frequência, indicando que um ciclo bienal (de período igual a 24 meses) foi também marcante para a década de 70, ciclo este que está ligado diretamente ao ciclo biológico do cafeeiro. Pode-se notar também um pico menor nas frequências próximas a 0,3, que corresponderia a um período de aproximadamente quatro meses.

Observa-se ainda que o ciclo sazonal (de 12 meses) correspondente a um pico nas frequências próximas a 0,07 não foi verificado para esta década, ao contrário do que ocorreu nas décadas de 50 e de 60. Por outro lado, o pico próximo à frequência de 0,3 que corresponderia a um ciclo de período aproximado igual a quatro meses, que não havia sido constatado para as décadas anteriores, se mostrou marcante no espectro dos preços nos anos 70.

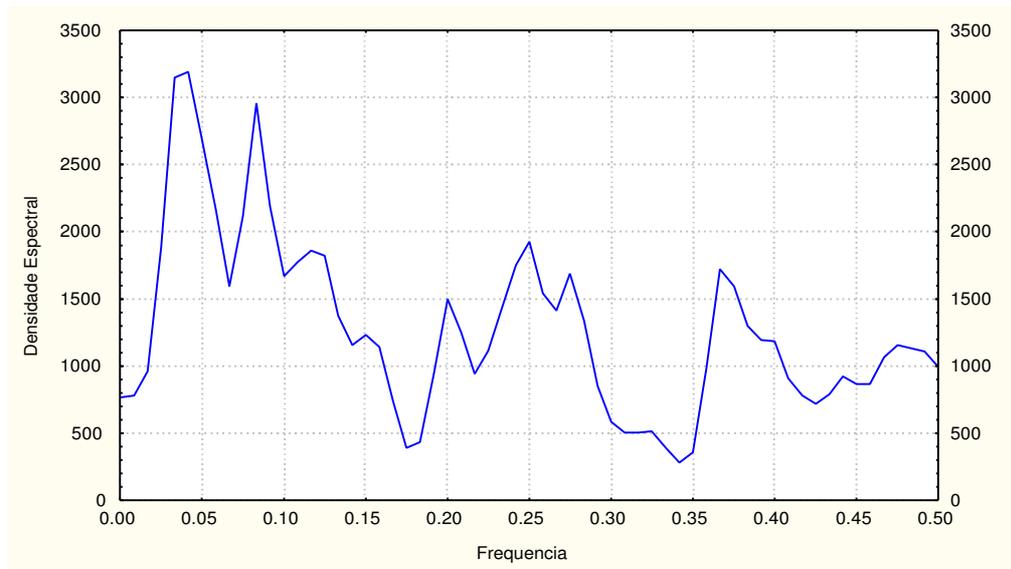


Fonte: Resultados da pesquisa.

Figura 35 - Espectro estimado para os preços do café na década de 70, com $M I = 7$ (em função de ω).

Os resultados encontrados para o espectro estimado dos preços do café nos anos 80 (Figura 36) apresentaram duas características marcantes. Primeiramente, pode-se observar que o ciclo de 24 meses, representado pelo pico nas frequências próximas a 0,04, é novamente o mais marcante em todo o conjunto de densidades espectrais estimadas, indicando que o ciclo bienal se configurou como um importante determinante do comportamento dos preços do café até o final da década de 80, pois foi um ciclo que prevaleceu ao longo das quatro primeiras décadas analisadas.

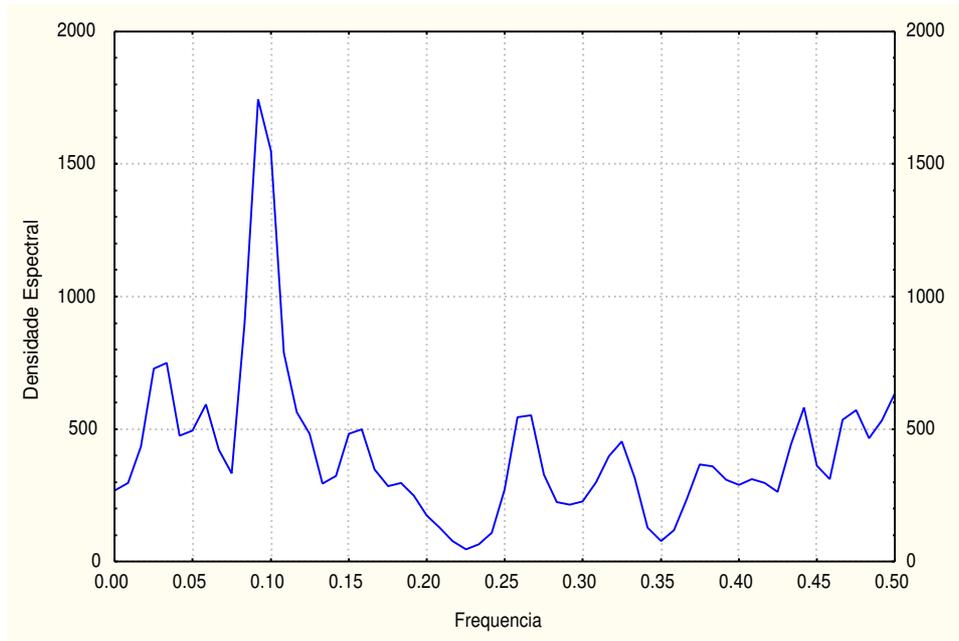
Em segundo lugar, torna-se evidente, com este novo espectro estimado, que o ciclo sazonal de frequências próximas a 0,07 novamente foi muito significativo nessa década, ao contrário do ocorrido na década de 70. Isto é, a repetição dos meses dos anos na década de 80 foi seguida de um comportamento relativamente repetitivo por parte dos preços do café.



Fonte: Resultados da pesquisa.

Figura 36 - Espectro estimado para os preços do café na década de 80, com $M I = 7$ (em função de ω).

Finalmente, para a década de 90 pode-se observar mudança no comportamento dos preços, em termos do componente de ciclos, de acordo com o indicado pelo espectro estimado para os preços nestes anos, dado na Figura 37. Apesar de um tanto quanto *flat*, o espectro estimado em função das frequências indica que aquelas pertencentes ao intervalo dado por $[0,07; 0,10]$ são as que mais contribuíram para a variância dos preços do café, uma vez que existe um pico neste espectro que se destaca em relação aos demais. Essas frequências correspondem ao período de 12 meses, o que configura o ciclo sazonal como o mais marcante ciclo nos preços para a década de 90.



Fonte: Resultados da pesquisa.

Figura 37 - Espectro estimado para os preços do café na década de 90, com $M I = 7$ (em função de ω).

Pode-se notar também que o pico próximo à frequência de 0,04, que para as outras décadas era marcante e configurava um ciclo bienal nos preços do café, nesta última década analisada, não se destacou no espectro estimado, assim como os demais picos verificados anteriormente, como o de período igual a quatro meses, por exemplo.

3.3. Análise do componente sazonal

De acordo com o proposto na metodologia deste estudo, vários métodos podem ser empregados na análise do componente sazonal dos preços do café brasileiro no mercado *spot* de Nova Iorque, desde a *análise visual* do gráfico dos preços, que se configura como uma abordagem mais rápida e simples, passando pelos Modelos de Regressão com Variáveis *Dummy*, até a utilização de métodos mais sofisticados, como o propiciado pelos *Modelos de Análise Espectral*

(segundo os quais a sazonalidade se caracterizaria como um ciclo de período fixo igual a 12 meses) e *pelos Modelos de Box e Jenkins Sazonais - SARIMA*.

Contudo, é o próprio comportamento do componente sazonal que será decisivo para determinar o método a ser empregado para a sua modelagem e eliminação da série dos preços do café. Isso porque, assim como o componente de tendência em uma série de tempo pode ser classificado, de acordo com a sua natureza, como do tipo *determinístico* ou do tipo *estocástico*, o componente sazonal também poderá ser conceituado como *sazonalidade determinística* e *sazonalidade estocástica*. Deve-se, portanto, ao analisar a sazonalidade, identificar em qual categoria ela se encontra para evitar inferências e conclusões imprecisas sobre o comportamento desses preços.

No presente trabalho foram implementados os procedimentos mencionados anteriormente para verificar se os preços do café apresentam comportamento sistemático sazonal, buscando também identificar em qual caso esta sazonalidade (caso confirmada) se enquadra, isto é, se ela é de natureza determinística ou de natureza estocástica.

3.3.1. Análise gráfica

Na tentativa de determinação da incidência de fenômenos sazonais nos preços do café, teve-se como ponto de partida a busca da visualização gráfica do componente sazonal nesses preços. Isso foi feito por meio da montagem de gráficos tridimensionais, em que foram fixados em um eixo os meses de um ano, em outro eixo foram representados os anos referentes às décadas analisadas e, finalmente, no eixo ortogonal ao plano formado pelos dois primeiros eixos, foram representados os valores das repartições da série de preços analisada. Ou seja, nesses gráficos foram representados espacialmente os preços praticados nas diferentes décadas tomadas como amostras neste trabalho, a saber: as décadas de 50, 60, 70, 80 e 90.

As superfícies formadas pelas partições dos preços podem indicar que um determinado comportamento sazonal exista. Caso o comportamento sazonal

se verifique, os preços tenderão a apresentar, sistematicamente, níveis mais elevados nos mesmos meses dos diferentes anos, o que pode ser considerado como um período de entressafra, e níveis mais baixos nos mesmos meses dos anos, caracterizando o período de safra.

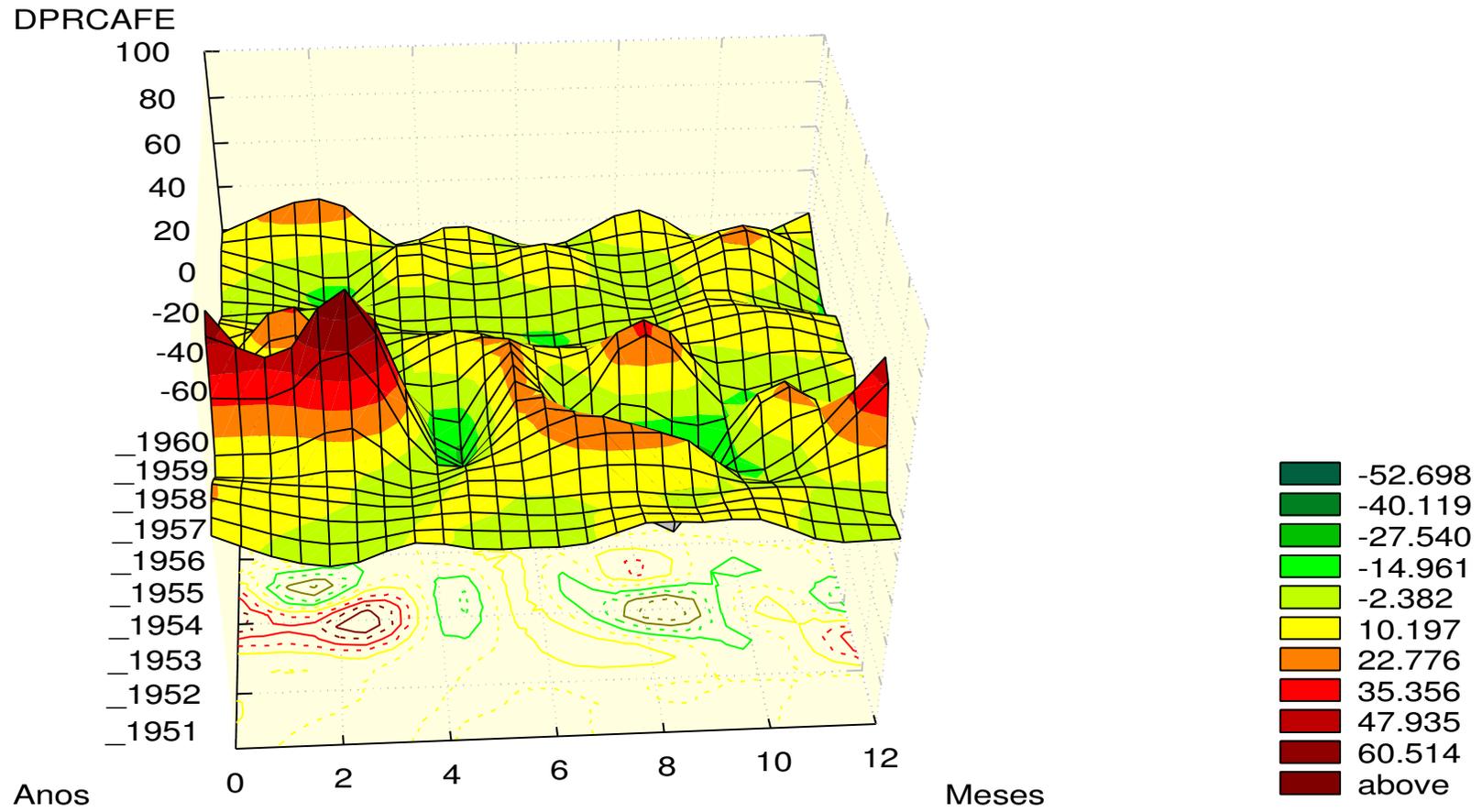
Todavia, o padrão do comportamento sazonal poderá estar mudando ao longo dos anos em função de uma possível relação *multiplicativa* entre o componente sazonal e o componente de tendência na série. Assim, teve-se como procedimento inicial, conforme discutido anteriormente, a remoção da tendência determinística e estocástica da série, antes de se iniciar a análise do componente de sazonalidade.

Nas Figuras 38, 39, 40, 41 e 42 estão representados os gráficos tridimensionais dos preços do café (representados pela série DPRCAFE⁵¹) nas décadas de 50, 60, 70, 80 e 90, respectivamente, em função dos meses do ano, sendo que o ponto 0 corresponderia ao início do ano.

Caso o padrão sazonal esteja incidindo sobre os preços do café ao longo das décadas, os gráficos deveriam apresentar formas semelhantes para os mesmos meses ao longo dos anos, isto é, picos para os meses de alta nos preços e vales nos meses de queda nos preços.

Uma superfície irregular que não obedecesse a esse padrão será um indicativo de os preços estão oscilando ao longo dos anos, mas não em função de uma sazonalidade determinística. Por outro lado, uma superfície relativamente plana ao longo de todo o gráfico seria um indicativo de estabilidade nos preços e ausência do fenômeno sazonal.

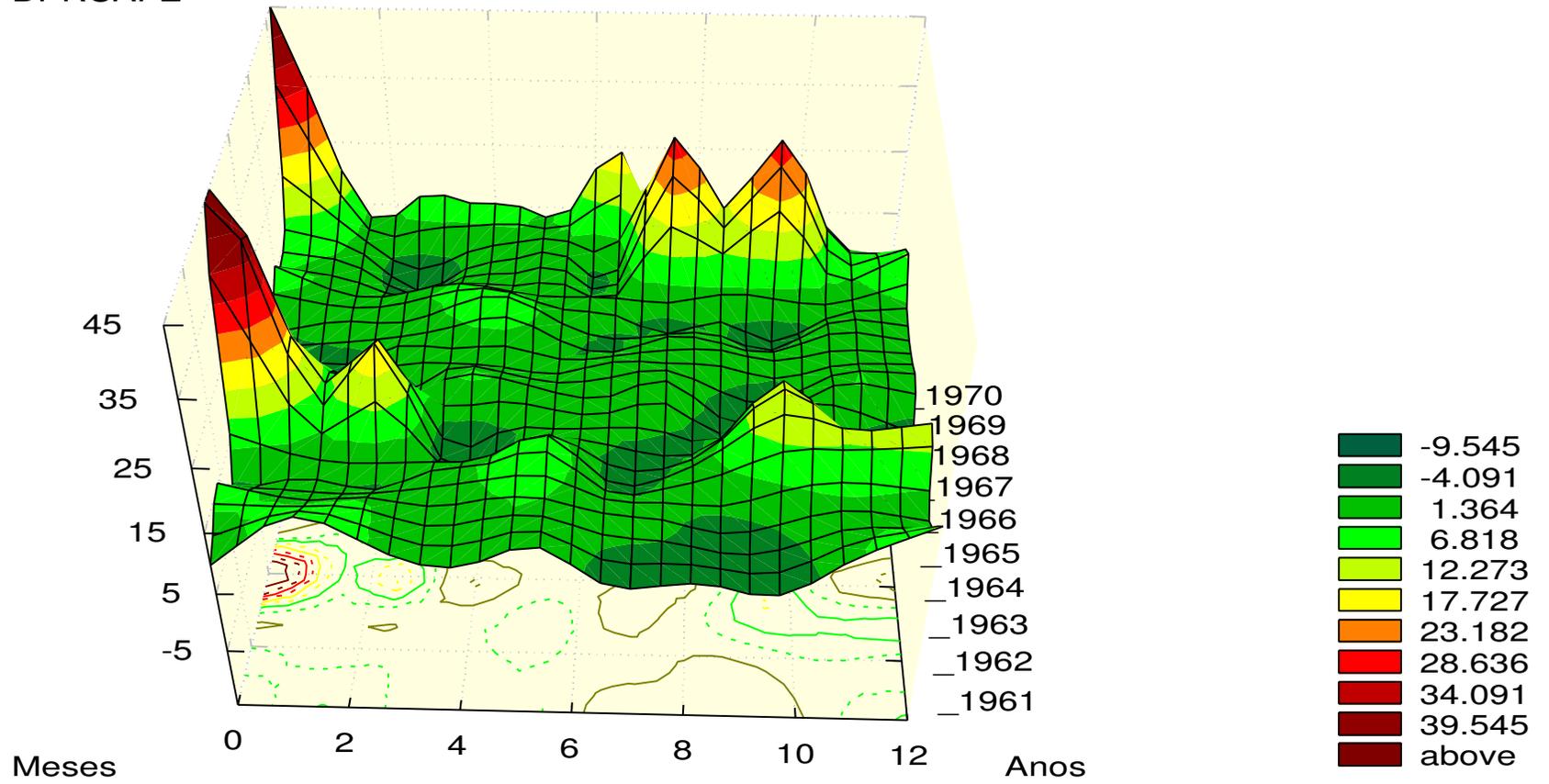
⁵¹ Que é a série dos preços do café tornada estacionária.



Fonte: Resultados da pesquisa.

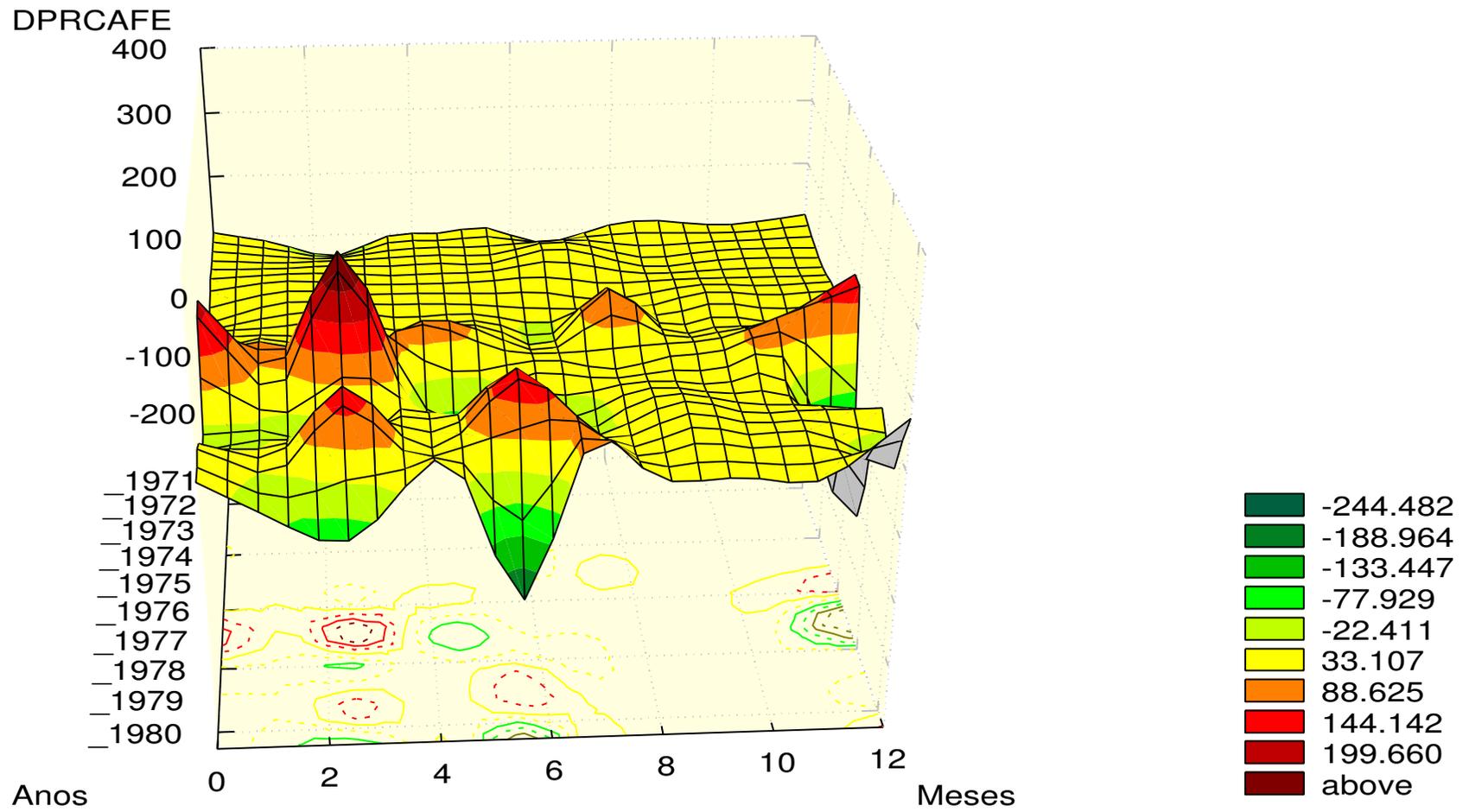
Figura 38 - Gráfico de superfície dos preços do café na década de 50.

DPRCAFE



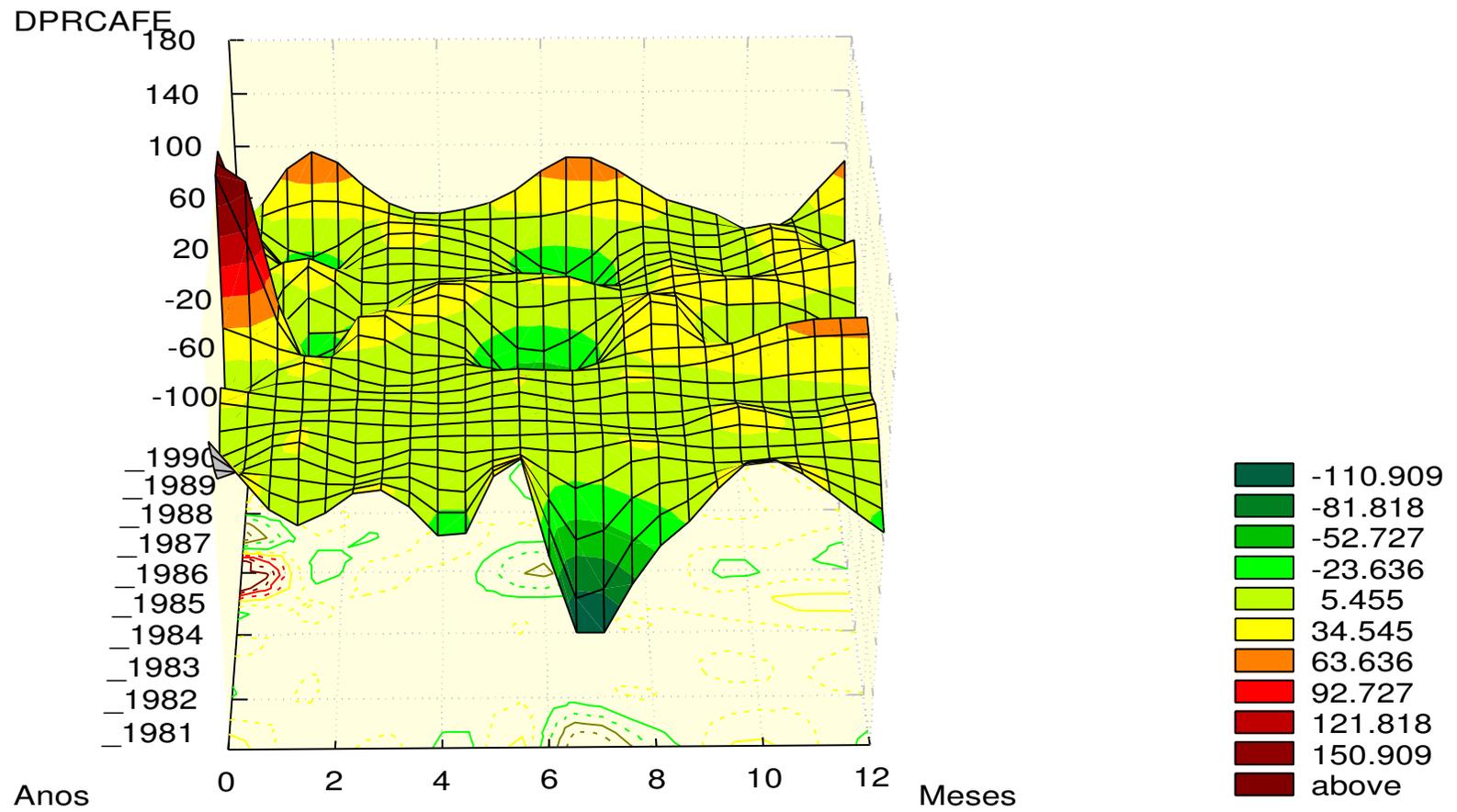
Fonte: Resultados da pesquisa.

Figura 39 - Gráfico de superfície dos preços do café na década de 60.



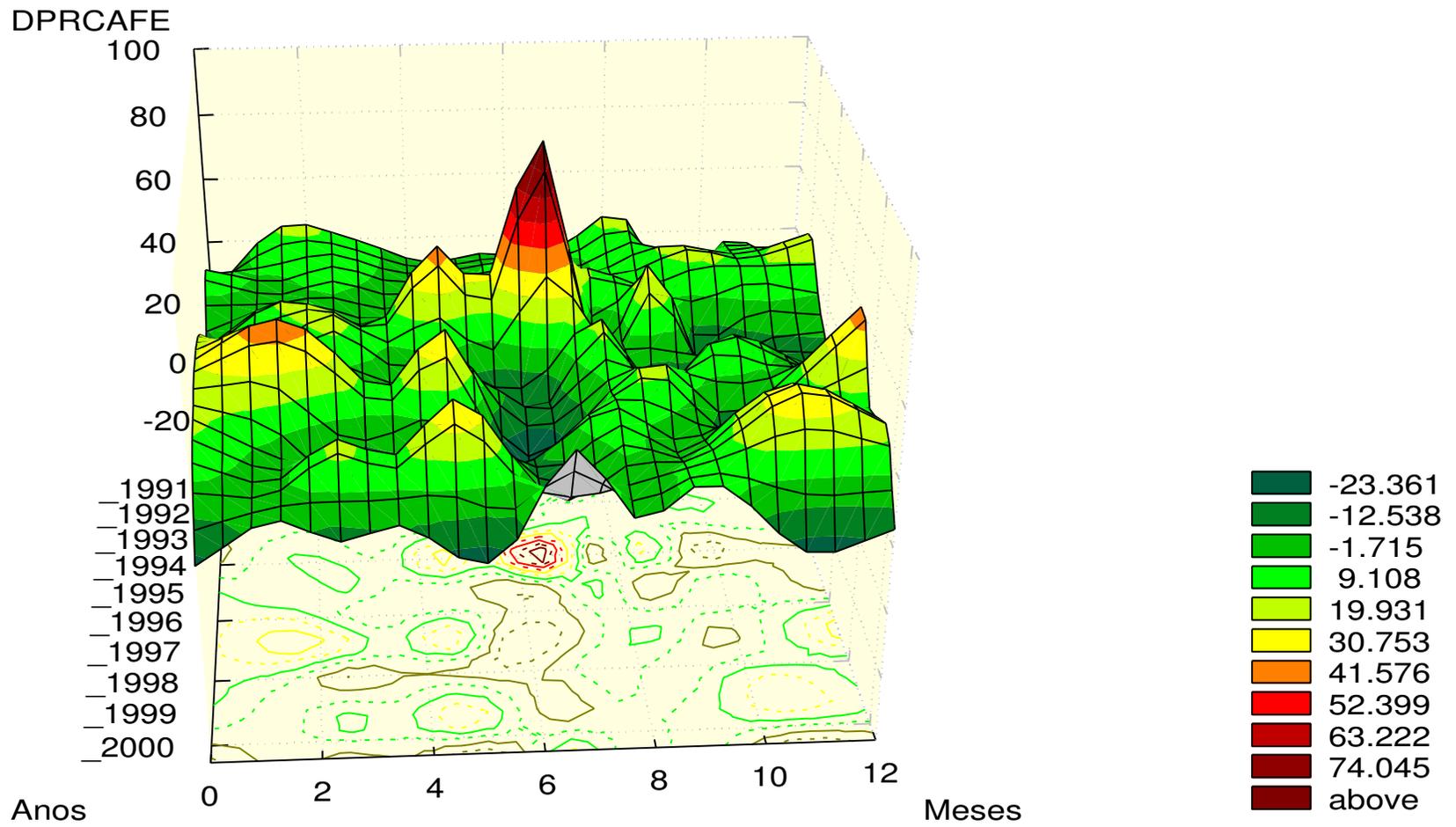
Fonte: Resultados da pesquisa.

Figura 40 - Gráfico de superfície dos preços do café na década de 70.



Fonte: Resultados da pesquisa.

Figura 41 - Gráfico de superfície dos preços do café na década de 80.



Fonte: Resultados da pesquisa.

Figura 42 - Gráfico de superfície dos preços do café na década de 90.

Os resultados encontrados para a década de 50 (Figura 38) apontam, de maneira geral, que os meses de janeiro a abril corresponderam, nesse período, a meses de alta nos preços da *commodity* brasileira no mercado *spot* de Nova Iorque, com exceção para os anos de 1958 e de 1959, que apresentaram nesses meses quedas consecutivas nos preços de aproximadamente 15 centavos de US\$/libra-peso, e, principalmente, para o ano de 1952, que apresentou nesses meses quedas nos preços do café de aproximadamente 55 centavos de US\$/libra-peso.

Por outro lado, os meses de julho a outubro foram aqueles em que os preços do café apresentaram maiores quedas, como foi o caso desses meses no ano de 1954, período em que o produto chegou a cair aproximadamente 130 centavos de US\$/libra-peso. No entanto, deve-se ressaltar que nos anos de 1955 e 1960 esses meses representaram meses de altas de até 14 centavos de US\$/libra-peso nos preços.

A respeito do comportamento dos preços na década de 50, pode-se ressaltar ainda que, apesar dos resultados discutidos anteriormente, não se pode afirmar que o componente sazonal foi marcante nesse período, porque não se tornou plenamente nítido esse padrão comportamental na visualização do gráfico desses preços.

Para a década de 60, o gráfico (Figura 39) confirma que os meses de janeiro a abril representam meses de alta nos preços da *commodity*. As exceções nessa década são para os anos de 1964 e de 1969, que apresentaram nesses meses quedas nos preços de aproximadamente 7 centavos de US\$/libra-peso. Por outro lado, os meses de julho a outubro novamente se mostraram como meses de queda nos preços, com exceção dos anos de 1963, 1969 e 1970, que apresentaram nesses meses comportamento contrário. Pode-se observar ainda que para a década de 60 o padrão de comportamento sazonal nos preços se mostrou um pouco mais evidente, pela observação do gráfico dos preços, quando comparado ao gráfico da década anterior.

O gráfico da década de 70 se apresentou como um dos mais paradoxais no que tange ao comportamento sazonal dos preços. Como se pode verificar

(Figura 40), nessa década os preços apresentaram dois padrões distintos de comportamento. Primeiramente, eles se mantiveram razoavelmente estáveis de 1971 a 1975, e nenhum padrão sazonal pôde ser verificado. Todavia, nos cinco anos posteriores eles apresentaram comportamento bastante diferenciado do que se verificou nas décadas anteriores.

Os meses de janeiro a abril, que de maneira geral eram meses correspondentes a altas nos preços ao longo dos anos, nesse período representaram sucessivamente meses de alta em um ano e meses de baixa no outro, isto é, se em um ano eles corresponderam a períodos de elevação nos preços, no outro ano essa situação se revertia, conforme indica o gráfico. Esse fenômeno pode ter sido provocado pela geada forte ocorrida em 1975, que teve forte impacto no comportamento dos preços da *commodity* brasileira no mercado *spot* de Nova Iorque. Os preços do café nesse período atingiram as maiores cotações de sua história, chegando ao patamar de 1.052,31 centavos de US\$ por libra-peso em abril de 1977⁵².

Os meses de julho a outubro, que nas outras décadas eram meses de relativa queda nos preços, na década de 70 não apresentaram essa característica, isto é, se apresentaram como meses de estabilidade nos preços, indicando a ausência (ou pouca importância) do fenômeno sazonal na determinação do comportamento dos preços nessa década.

A visualização do gráfico dos preços nos anos 80 (Figura 41) indica que, ao contrário da década de 70, o fenômeno sazonal se apresentou de maneira mais nítida. Apesar das exceções verificadas para os anos de 1987 e de 1989, os quatro primeiros meses do ano voltaram a se caracterizar como meses de preços em elevação.

O mesmo ocorreu com os meses de baixa, de julho a outubro, que, exceto para 1990, voltaram a ser assim identificados. Nesses meses, os preços

⁵² Deve-se ressaltar que essas observações poderiam ser consideradas como *outliers* e possivelmente retiradas da amostra dos preços. Todavia, optou-se pela sua manutenção, dado que esses preços estão intrinsecamente ligados a quebras de safras na produção de café, e a detecção de comportamentos cíclicos nessas quebras por meio da verificação de ciclos nos preços do café se configurou como uma hipótese testada.

apresentaram quedas significativas, como as observadas nos anos de 1981 e 1986. Nesses anos, os preços chegaram a apresentar quedas de aproximadamente 170,00 centavos de US\$ por libra-peso nestes meses de baixa, o que pode ser visualizado também no gráfico da Figura 41. Deve-se ressaltar ainda que, nesta década, fica mais evidente que o período de alta nos preços se inicia no último mês de cada ano, o que indica uma ligação mais direta entre altas de preços no mercado internacional e o período do verão⁵³ no Brasil.

Para a década de 90, a visualização do comportamento sazonal por meio de gráficos se tornou um pouco menos direta, conforme se pode observar no último gráfico (Figura 42) da série apresentada. Entretanto, podem-se identificar aí algumas características marcantes da década. Nela, os meses de alta nos preços avançaram até o mês de maio e em alguns anos evoluíram até os meses de junho e julho, que eram, nas décadas anteriores, tipicamente, meses de preços em queda. O principal caso observado se deu para o ano de 1994, ano em que os preços apresentaram crescimento elevado até o mês de julho.

3.3.2. Testes para sazonalidade determinística nos preços

A estimativa de modelos de sazonalidade determinística, como o de variáveis *dummy*, dado pela equação (58), apresentada anteriormente, pode ser utilizada como um instrumento para se comprovar a existência ou não do fenômeno sazonal, na forma determinística, nos preços do café. Se os coeficientes da equação (58) não forem, em conjunto, estatisticamente significativos, tem-se que, de fato, a sazonalidade nos preços do café, se existir, não é do tipo determinística, e vice-versa.

Assim, procedeu-se a testes de sazonalidade determinística para todo o período tomado como amostra (Apêndice F). Também foram estimadas equações para as décadas de 50, 60, 70, 80 e 90, tomadas separadamente, a fim de se

⁵³ Que se inicia em 21 de dezembro e termina em 21 de março.

verificar se dentro dessas partições do período amostral a sazonalidade determinística pôde ser observada.

A mensuração dos efeitos, ou impactos, que os diferentes meses do ano exercem sobre os preços do café na bolsa de Nova Iorque pode ser implementada com a estimação de *índices de sazonalidade*⁵⁴ dados pelos coeficientes da equação (58). Eles representarão, conjuntamente, o componente de sazonalidade determinística que possa estar presente nestes preços.

O número de índices que deverão ser estimados depende da natureza do fenômeno sazonal que os preços podem apresentar. Eles serão iguais a 12, se a sazonalidade estiver atuando mensalmente ao longo do ano, e iguais a 4, se o fenômeno sazonal estiver diretamente ligado às estações do ano.

O cálculo de tais índices no presente trabalho se deu pela estimação de equações de formas semelhantes à da equação (58), com 11 e 3 variáveis *dummy*, respectivamente. Estas foram estimadas pelo Método de Mínimos Quadrados Ordinários, tendo como variável dependente a série estacionária *DPRCAFE*.

O valor do primeiro coeficiente β dessas equações foi igualado a zero para evitar o problema de multicolinearidade perfeita com o termo de intercepto da equação. Assim, cada um dos índices β_i (para todo $i \neq 1$) representará o efeito sazonal do i -ésimo mês em relação ao valor observado no primeiro mês (janeiro) para o modelo de sazonalidade mensal e o efeito da respectiva estação do ano em relação ao período de verão, ou seja, será a diferença que deverá ser acrescida ou subtraída do valor observado no período *base*, a fim de eliminar o efeito sazonal.

A significância estatística dos coeficientes estimados foi testada com a utilização do valor de probabilidade da estatística- F e pela análise dos *valores de probabilidade* estimados (*p-values*) para cada um dos índices sazonais. Foi adotado o nível de significância estatística (α) de 10% nos testes.

Os principais resultados encontrados para as equações para todo o período amostral (de 1946 a 2000) e para as décadas de 50, 60, 70, 80 e 90, tomadas individualmente, podem ser observados no Quadro 9. Como pode ser

⁵⁴ Os valores desses índices deverão ser adicionados ou subtraídos dos preços, para se eliminar a sazonalidade determinística destes.

verificado, os resultados estimados apresentam-se amplamente diferenciados de acordo com o período de tempo utilizado como amostra na regressão. Tomando todo o período, de janeiro de 1946 a dezembro de 2000, observou-se que apenas alguns índices sazonais⁵⁵ se apresentaram como estatisticamente significativos (valores em negrito).

Quadro 9 - Resultados dos testes de sazonalidade determinística mensal

Período/ Variável	1946 a 2000		Década de 50		Década de 60		Década de 70		Década de 80		Década de 90	
	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.
Janeiro	5,99	0,07	5,05	0,30	6,24	0,00	6,18	0,67	12,33	0,12	3,12	0,50
Fevereiro	-6,51	0,16	-5,48	0,43	-6,47	0,02	-9,17	0,65	-15,97	0,15	1,89	0,77
Março	-2,45	0,60	-0,63	0,93	-6,20	0,03	12,32	0,55	-15,08	0,17	-1,71	0,80
Abril	-6,41	0,17	-7,62	0,27	-7,41	0,01	2,47	0,90	-13,75	0,22	-7,06	0,28
Maiο	-8,56	0,07	-11,84	0,09	-6,44	0,02	-11,91	0,56	-16,95	0,13	2,20	0,74
Junho	-8,53	0,07	-4,06	0,56	-6,49	0,02	-6,06	0,77	-21,78	0,05	-10,27	0,12
Julho	-9,72	0,04	-8,46	0,22	-8,14	0,00	-6,74	0,74	-39,86	0,00	-0,36	0,95
Agosto	-6,79	0,15	-7,71	0,27	-5,81	0,04	-0,34	0,99	-16,33	0,14	-3,44	0,60
Setembro	-5,15	0,27	-12,30	0,08	-7,36	0,01	-4,33	0,83	-7,11	0,52	-1,89	0,77
Outubro	-5,91	0,20	-9,93	0,15	-5,22	0,06	-4,56	0,82	-11,95	0,28	-6,15	0,35
Novembro	-1,90	0,69	-5,22	0,45	-5,25	0,06	-1,66	0,94	-6,89	0,53	-1,96	0,77
Dezembro	-8,28	0,08	-7,28	0,29	-6,91	0,01	-22,39	0,28	-6,66	0,55	-1,17	0,86
Estatística-F	0,82	0,62	0,64	0,79	1,11	0,36	0,34	0,97	1,64	0,10	0,64	0,79

Fonte: Resultados da pesquisa.

Obs.: Valores em negrito indicam significância estatística em nível de 10%.

O índice do mês-base (janeiro) e os índices dos meses de maio, junho, julho e dezembro, de maneira geral, confirmam os resultados discutidos na análise de sazonalidade pelos gráficos dos preços, isto é, os valores estimados indicam que o mês de janeiro se caracteriza como um mês de preços em elevação, com uma valorização média de 5,99 centavos de US\$ por libra-peso, enquanto os meses de maio, junho e julho correspondem a meses de desvalorização em relação às cotações observadas para janeiro. As desvalorizações médias na cotação verificadas, relacionadas a esses meses,

⁵⁵ Coeficientes das variáveis *dummies* das equações estimadas.

foram, respectivamente, de 8,56; 8,53; e 9,72 centavos de US\$ por libra-peso, em relação aos valores do mês-base de janeiro. Ao contrário do que era aparente nos gráficos, o mês de dezembro também se caracterizou como um mês de desvalorização da *commodity* para o período analisado como um todo.

No entanto, deve-se ressaltar que o valor de probabilidade calculado para a estatística-F estimada para esta equação (igual a 0,62) indica que esta não é significativa e, portanto, não se pode rejeitar a hipótese de que todos os índices sazonais são de fato iguais a zero. Isso implica portanto que, de acordo com esse teste, a sazonalidade determinística não pode ser confirmada como estando presente nos preços do café brasileiro no mercado *spot* de Nova Iorque para o período analisado, de 1946 até o ano 2000.

A análise individual por décadas da sazonalidade determinística nos preços, de maneira geral, leva à mesma conclusão. Todavia, deve-se ressaltar que os resultados verificados para a década de 60 foram, em certa medida, diferentes do verificado para as demais décadas. Na equação estimada para esses anos todos, os coeficientes estimados (índices sazonais) foram estatisticamente significativos em nível de significância de 10%, porém o teste-F implementado não confirmou esses resultados.

Foi testada também a possibilidade de que a sazonalidade determinística se manifestasse de maneira mais perceptível, tomando-se os dados dispostos em termos das estações do ano. Foram estimados modelos de sazonalidade determinística em que as variáveis *dummy* e os respectivos índices sazonais se dessem de maneira a representar possíveis efeitos das estações do ano no Brasil (que estão ligadas ao comportamento da produção) sobre as cotações do café no mercado internacional.

Assim, tomou-se como período-base o verão, que foi representado pelos valores observados para os meses de janeiro, fevereiro e março⁵⁶. Os meses de abril a junho representaram o período de outono; de julho a setembro, o inverno;

⁵⁶ Na verdade o período de verão vai de 22 de dezembro a 21 de março; o outono, de 22 de março a 21 de junho; o inverno, de 22 de junho a 21 de setembro; e a primavera, de 22 de setembro a 21 de dezembro, nos países do hemisfério sul.

e de outubro a dezembro, o período da primavera. Os principais resultados encontrados para as equações estimadas podem ser verificados no Quadro 10.

Quadro 10 - Resultados dos testes de sazonalidade determinística trimestral

Período/ Variável	1946 a 2000		Década de 50		Década de 60		Década de 70		Década de 80		Década de 90	
	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.	Ín.saz.	Prob.
Verão (jan.\mar.)	3,00	0,12	3,01	0,28	2,01	0,08	7,23	0,38	1,98	0,67	3,18	0,23
Outono (abr.\jun.)	-4,85	0,07	-5,80	0,14	-2,56	0,11	-6,21	0,59	-7,14	0,27	-5,10	0,18
Inverno (jul.\set.)	-4,24	0,12	-7,45	0,06	-2,88	0,08	-4,85	0,68	-10,75	0,10	-1,96	0,60
Primavera (out.\dez.)	-2,38	0,38	-5,44	0,16	-1,57	0,33	-10,58	0,36	1,85	0,78	-3,15	0,40
Estatística-F	1,31	0,27	1,38	0,25	1,30	0,28	0,29	0,84	1,66	0,18	0,65	0,58

Fonte: Resultados da pesquisa.

Pode-se verificar que, no nível de significância estatística de 10%, os índices de sazonalidade em função das estações do ano também não apresentaram significância em nenhum dos casos. Para a equação com dados de 1946 a 2000, apenas para níveis de significância maiores que 27% se poderia rejeitar a hipótese de que todos os índices de sazonalidade determinística são significativos.

Analisando a significância individual dos índices de sazonalidade estimados e tomando-se um valor maior e menos rigoroso para α , igual a 20% por exemplo, poder-se-ia concluir, de acordo com os resultados da regressão para o período de 1946 a 2000, que de fato os primeiros meses do ano (verão) representam um período de entressafra e conseqüente valorização para o café. O mesmo apresentou neste período um comportamento médio de alta de 3,00 centavos de US\$ por libra-peso nos preços, sendo seguido por períodos de queda nas cotações de 4,85 centavos de US\$ por libra-peso em média no outono e de

4,24 centavos de US\$ por libra-peso em média no inverno, em relação aos níveis observados no verão.

Todavia, mesmo em nível maior, e menos rigoroso do ponto de vista científico, para o nível de significância estatística (α) de 20%, o índice sazonal referente ao período da primavera não se mostrou significativo neste teste. O mesmo pode ser dito para a estatística-F calculada. Portanto, os resultados anteriores devem ser tomados com cautela.

Para as décadas tomadas individualmente, os resultados não diferem significativamente dos verificados para todo o período. Isso se constitui em um forte indício de que a sazonalidade, caso seja um componente de influência no comportamento dos preços do café no mercado internacional, não apresenta forma determinística e regular. Pelo contrário, caso ela de fato se verifique, deverá ser de natureza estocástica, variando em função do período do tempo, conforme indicaram alguns dos resultados da análise espectral apresentados anteriormente.

Deve-se ressaltar que os resultados encontrados com relação ao componente sazonal diferem em certa medida do que era esperado *a priori* e do que é um tanto quanto consensual na literatura existente. Esta última indica que, de maneira geral, a oferta e os preços de *commodities* agrícolas tendem a apresentar, no que se refere ao componente sazonal, um comportamento relativamente estável e previsível ao longo dos anos em função da influência da ocorrência de fenômenos climáticos e naturais repetitivos dados pelas diferentes estações do ano. Assim, tem-se configurado no caso dos preços do café, para o período analisado, um exemplo contrário a esse consenso.

Os resultados encontrados apontam ainda que não há a necessidade de se trabalhar com os resíduos dos modelos de sazonalidade determinística estimados para se analisar a volatilidade dos preços do café. Isto porque esses preços, de acordo com os testes implementados, estão livres de sazonalidade do tipo determinística, restando apenas verificar a incidência de sazonalidade estocástica nestes, o que será feito na seção seguinte.

3.3.3. Análise do componente sazonal pela função de autocorrelação

Como se pôde verificar nas seções anteriores, a detecção do componente sazonal nos preços do café por meio da visualização de gráficos da série se configurou como um procedimento útil, porém um tanto quanto impreciso. Principalmente nos anos em que os preços flutuaram acentuadamente, a verificação da presença de sazonalidade pela análise direta do gráfico de superfície dos preços se tornou pouco conclusiva. Por outro lado, os resultados dos testes para a sazonalidade determinística indicaram que a sazonalidade, caso seja um componente que de fato influencia o comportamento dos preços do café, não é da forma determinística e regular.

O padrão da sazonalidade nos preços do café poderá, portanto, ser do tipo estocástico. Isso ocorrerá se o componente sazonal possuir comportamento instável ao longo do tempo, se modificando de forma imprevisível. De maneira geral, a sazonalidade estocástica é verificada de maneira mais freqüente em séries temporais ligadas a mercados financeiros⁵⁷, que são influenciadas por diversos fatores que não necessariamente irão se repetir de forma previsível de período sazonal para período sazonal. Dessa forma, modelos determinísticos como os discutidos anteriormente não apresentarão em seus resíduos uma série para os preços do café livre do fenômeno sazonal.

De acordo com o discutido na metodologia do presente trabalho, a análise da função de autocorrelação dos preços do café se configura como um instrumento alternativo eficiente para a verificação da ocorrência de sazonalidade nesta variável⁵⁸. Caso os preços do café apresentem o componente sazonal, os diversos pontos amostrais deverão apresentar algum grau de correlação com os dados correspondentes com *lags* e *leads* de 12 meses. Ou seja, espera-se a existência de correlação entre as cotações do café de 12 meses atrás e 12 meses à frente com os valores de um dado mês, e assim por diante. Essas correlações

⁵⁷ Que é o caso dos preços do café no mercado *spot* da bolsa de Nova York.

⁵⁸ Deve-se ressaltar a necessidade de se ter a série na forma estacionária, como de fato é a série DPRCAFÉ, que será analisada.

deverão se manifestar, na função de autocorrelação amostral, pela presença de picos significativos para os valores estimados de $\hat{\rho}_k$ em que $k = 12, 24, 36$, e assim por diante.

Assim, foi estimada a função de autocorrelação amostral para os preços do café de janeiro de 1946 a dezembro de 2000, e os principais resultados encontrados podem ser verificados no Quadro 11.

Foi estimada uma função de autocorrelação para 48 defasagens, que corresponde a quatro anos de defasagens dos preços. Os valores observados para $\hat{\rho}_k$, em que $k = 12, 24, 36$ e 48, que seriam as defasagens que indicariam a ocorrência de sazonalidade nos preços, estão assinalados em negrito, e os valores estatisticamente significativos estão demarcados por asteriscos.

Conforme apontam os resultados, apenas os coeficientes de autocorrelação (CAC) e de autocorrelação parcial (CACparc.) para uma defasagem de 12 meses se mostraram significativos, dentre os ligados ao componente sazonal. Para $k = 24$ e 48, nenhum dos dois se mostrou significativo, e para $k = 36$ apenas o coeficiente de autocorrelação parcial esteve fora do intervalo de confiança.

Esses resultados apresentam duas características marcantes: primeiramente, indicam que não há necessidade de se implementar o processo de diferenciação sazonal, discutido anteriormente, pois, conforme aponta ENDERS (1995), a diferenciação sazonal da série só seria indicada se os coeficientes de autocorrelação nos *lags* 12, 24, 36 e 48 apresentassem valores próximos de 1 e os picos sazonais decaíssem lentamente, implicando *raízes unitárias sazonais*, o que não foi verificado para os preços do café. Em segundo lugar, pode-se inferir que esses preços tendem a ser correlacionados de maneira sazonal, mas que esta relação só é forte de ano para ano e não para períodos mais distantes no tempo.

Quadro 11 - Resultados para a função de autocorrelação amostral

Autocorrelações	Correlações parciais	Def.	Coefficiente de autocorrelação	Coefficiente de autocorrelação parcial	Estatística Q	Valor P
.l**	.l**	1	0,207	0,207	28,347	0,000
.l*	.l*	2	0,114	0,074	36,941	0,000
.l*	.l*	3	0,136	0,104	49,295	0,000
.l.	.l.	4	0,036	-0,018	50,142	0,000
.l.	.l.	5	-0,004	-0,029	50,155	0,000
.l.	*l.	6	-0,047	-0,059	51,641	0,000
.l*	.l*	7	0,080	0,107	55,876	0,000
.l.	.l.	8	0,052	0,034	57,685	0,000
*l.	*l.	9	-0,143	-0,171	71,357	0,000
.l.	.l.	10	0,023	0,058	71,728	0,000
.l.	*l.	11	-0,053	-0,062	73,635	0,000
*l.	*l.	12	-0,170	-0,133	93,072	0,000
.l.	.l.	13	-0,049	0,027	94,680	0,000
.l.	.l.	14	-0,050	-0,019	96,389	0,000
*l.	*l.	15	-0,071	-0,060	99,826	0,000
*l.	*l.	16	-0,117	-0,062	109,07	0,000
*l.	.l.	17	-0,060	-0,013	111,55	0,000
*l.	*l.	18	-0,065	-0,075	114,43	0,000
.l.	.l*	19	0,026	0,137	114,90	0,000
*l.	*l.	20	-0,063	-0,088	117,61	0,000
*l.	*l.	21	-0,061	-0,097	120,18	0,000
.l.	.l.	22	-0,057	-0,017	122,38	0,000
.l.	.l.	23	-0,053	-0,023	124,32	0,000
.l.	.l.	24	-0,040	-0,040	125,44	0,000
.l.	.l.	25	-0,025	0,010	125,87	0,000
.l.	.l.	26	0,001	-0,013	125,87	0,000
.l*	.l.	27	0,067	0,025	128,99	0,000
*l.	*l.	28	-0,057	-0,063	131,27	0,000
.l.	.l.	29	0,014	-0,007	131,41	0,000
.l*	.l.	30	0,067	0,046	134,55	0,000
.l.	.l.	31	-0,042	-0,042	135,76	0,000
.l.	.l.	32	0,045	0,019	137,15	0,000
.l.	.l.	33	0,019	-0,033	137,40	0,000
.l.	.l.	34	0,017	-0,022	137,60	0,000
.l.	.l.	35	-0,024	-0,033	137,99	0,000
.l.	.l*	36	0,044	0,074	139,34	0,000
.l.	.l.	37	0,042	-0,038	140,55	0,000
.l.	.l.	38	0,003	-0,016	140,56	0,000
*l.	*l.	39	-0,059	-0,057	142,99	0,000
.l*	.l.	40	0,085	0,062	148,11	0,000
.l.	.l.	41	0,053	0,060	150,09	0,000
.l.	.l.	42	-0,001	-0,008	150,09	0,000
.l*	.l.	43	0,077	0,048	154,31	0,000
.l.	.l.	44	0,036	-0,029	155,25	0,000
.l.	.l.	45	0,032	0,037	155,99	0,000
.l.	.l.	46	0,018	-0,008	156,23	0,000
.l.	.l.	47	0,014	0,018	156,38	0,000
.l.	.l.	48	0,013	-0,011	156,50	0,000

Fonte: Resultados da pesquisa.

Assim, pode-se inferir que, de fato, a sazonalidade nos preços do café brasileiro no mercado *spot* de Nova Iorque foi comprovada. No entanto, deve-se ressaltar que essa relação está se modificando ao longo do tempo, o que caracterizaria uma sazonalidade do tipo estocástica e não do tipo determinística para os preços dessa *commodity* no mercado internacional.

Deve-se ressaltar que, de acordo com os testes implementados, esse componente não se configura como tão evidente ou altamente significativo para esses preços no período analisado, não implicando, portanto, a necessidade de diferenciação sazonal dos dados.

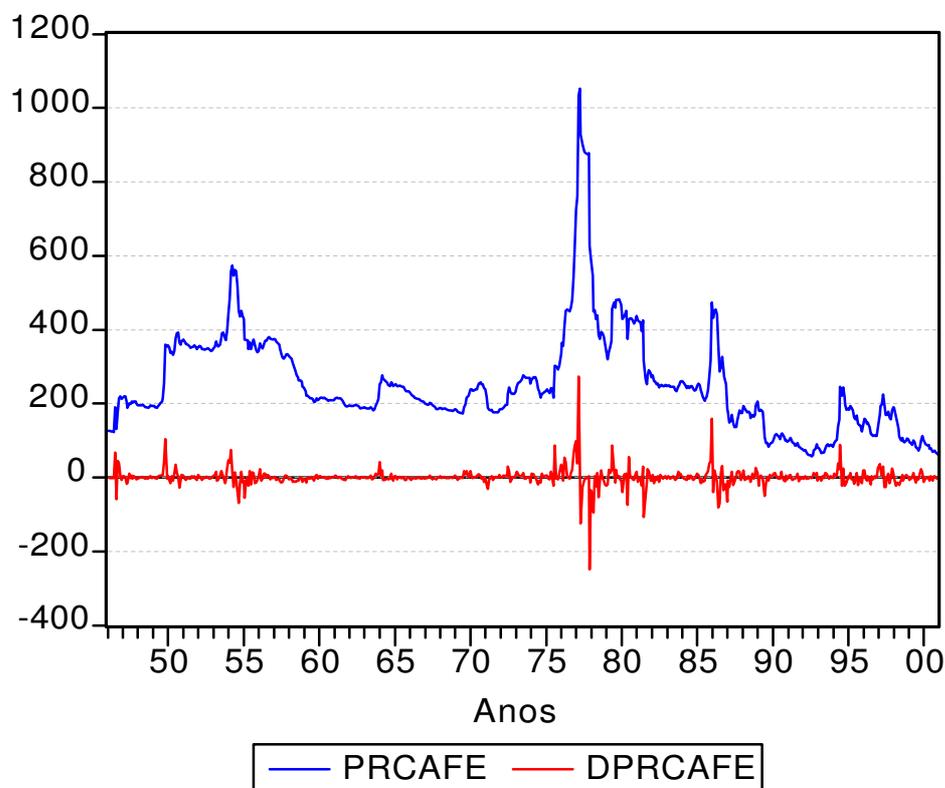
3.4. Análise do componente de volatilidade

De acordo com o detalhado anteriormente na metodologia específica, para se implementar a análise do componente de volatilidade nos preços do café brasileiro no mercado *spot* de Nova Iorque, no período de janeiro de 1946 a dezembro de 2000, devem-se seguir alguns passos distintos. Primeiramente, pode-se analisar o gráfico⁵⁹ dos preços, a fim de identificar visualmente se ele apresenta comportamento de alta volatilidade ou se apresenta relativa estabilidade em seus valores observados. Assim, o gráfico dos preços reais do café e o da série DPRCAFE, na Figura 43, demonstram um comportamento extremamente volátil para os preços do café no período analisado.

O comportamento dessas variáveis indica, *a priori*, que os preços do café são marcados por períodos de grande volatilidade, como a verificada entre os anos de 1975 e 1980, e por momentos de relativa estabilidade, como a ocorrida entre os anos de 1960 e 1975.

⁵⁹ Assim como ocorre na análise dos demais componentes de uma série temporal, em que o gráfico da série se configura como um importante, apesar de pouco preciso, instrumento de apoio à análise.

US\$ cent./Libra-Peso



Fonte: Elaboração do autor a partir de dados da NYBOT.

Figura 43 - Comportamento dos preços do café na NYBOT.

Esse comportamento de volatilidade inconstante (heterocedasticidade) presente nos preços do café pode ser melhor analisado por meio de modelos mais apropriados. De acordo com o apresentado anteriormente, os modelos de heterocedasticidade condicional auto-regressiva generalizados da “família GARCH” têm se constituído em um dos principais instrumentos de apoio à tomada de decisão e de análise e controle do risco⁶⁰ associado a séries temporais.

Na análise e modelagem da volatilidade dos preços do café por meio de um modelo do tipo GARCH, deve-se primeiramente especificar um modelo para

⁶⁰ Que está ligado intrinsecamente à volatilidade dos preços das *commodities* e dos *retornos* dos ativos financeiros em geral.

a série desses preços por meio de uma equação genérica, como a dada pela expressão (66), vista anteriormente. Este modelo tanto pode ser um modelo de regressão múltipla usual, como pode ser um modelo de séries temporais do tipo ARIMA(p,d,q). Entretanto, esse modelo deve ser bem especificado, no sentido de garantir que os resíduos (e_t) possuam média zero e não sejam correlacionados.

Posteriormente, deve-se testar se os resíduos deste modelo apresentam comportamento de heterocedasticidade condicional por meio de testes específicos, como o *teste-Q*⁶¹ e o *teste do multiplicador de Lagrange* (LM), cujos procedimentos foram explicados previamente.

Caso seja verificado que de fato há presença de termos ARCH nos resíduos do modelo estimado, deve-se passar à modelagem do componente de volatilidade por meio da especificação e do teste de modelos de variância condicional. Tendo isso em vista, na seção a seguir buscar-se-á especificar um modelo apropriado para descrever o comportamento dos preços do café, para que se possa posteriormente modelar a volatilidade dessa série.

3.4.1. Especificação de um modelo para a série de preços do café

A análise univariada de séries temporais, de acordo com a metodologia proposta por BOX e JENKINS (1976) e conforme a própria expressão indica, busca determinar e modelar o comportamento de uma série temporal como função de seus próprios valores defasados no tempo e de valores correntes e defasados de termos de erros incorporados nesses modelos. Nesse sentido, esse tipo de análise implica que o comportamento da variável analisada pode ser explicado de maneira independente do comportamento de outras variáveis econômicas.

Tudo isso se configura em um princípio oposto ao dos modelos tradicionais da econometria, que, de maneira geral, buscam relacionar diferentes variáveis econômicas. Nestes últimos, algumas dessas variáveis (ditas

⁶¹ Para o quadrado dos resíduos.

explicativas ou independentes) serão relacionadas por meio de equações de regressão com a variável cujo comportamento médio se pretende explicar (chamada variável explicada ou dependente).

Visto que no presente trabalho de pesquisa teve-se como princípio metodológico básico analisar o comportamento dos preços do café no mercado internacional por meio de *modelos univariados de séries temporais no domínio do tempo e da frequência*; o modelo que será especificado para determinar o comportamento médio da variável de trabalho (DPRCAFÉ) se enquadrará nessa categoria.

No entanto, deve-se ressaltar que um modelo econométrico tradicional de regressão linear também poderia ser implementado, desde que fosse levado em consideração o problema de *regressões espúrias* que poderia ocorrer ao se utilizar a série de preços do café nesse modelo, sem antes torná-la estacionária.

Na primeira etapa da modelagem da série DPRCAFÉ, buscou-se identificar quais os valores apropriados para as defasagens “ p ” dos termos auto-regressivos (AR) e a ordem “ q ” dos termos de média móvel (MA) do modelo ARIMA (p,d,q) a ser estimado. Também nessa etapa buscou-se identificar a ordem “ P ” de possíveis termos sazonais auto-regressivos (SAR) e a ordem “ Q ” de possíveis termos de média móvel sazonal (SMA), caso em que o modelo passaria a ser denominado SARIMA (p,d,q) \times (P,D,Q)₁₂, para dados mensais, como é o caso dos preços do café analisados.

Com relação aos termos “ d ” e “ D ”, eles se referem às ordens de integração da série, indicando o número de *diferenças consecutivas* e de *diferenças sazonais*, respectivamente, que foram aplicadas à série dos preços do café para torná-la estacionária. Conforme já apresentado nos resultados anteriores, verificou-se que a ordem para d será igual a 1, pois uma única diferença da série foi o suficiente para torná-la estacionária.

Com relação ao valor para a ordem de integração sazonal D , pôde-se verificar que o seu valor é igual a zero, pois, conforme aponta ENDERS (1995), a diferenciação sazonal de uma série deve ser aplicada caso esta apresente valores elevados (acima de 0,5) nos coeficientes de autocorrelação e de

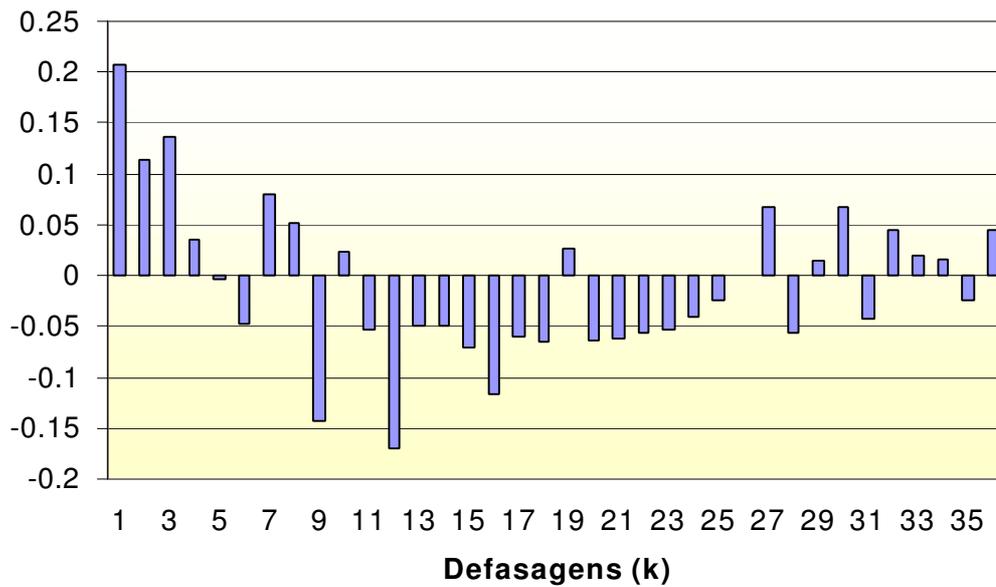
autocorrelação parcial nos diversos *lags* ligados ao fenômeno sazonal ($k = 12, 24, 36, 48...$) e estes não tenderem a se reduzir rapidamente.

Para o caso dos preços do café, o que se verificou no correlograma estimado (Quadro 11) foi a presença de picos sazonais significativos apenas para $k = 12$, isto é, conforme apontaram os resultados, apenas os coeficientes de autocorrelação (CAC) e de autocorrelação parcial (CACparc.) para uma defasagem de 12 meses se mostraram significativos, dentre os ligados ao componente sazonal.

Para $k = 24$ e 48 , nenhum dos dois se mostrou significativo, e para $k = 36$ apenas o coeficiente de autocorrelação parcial foi significativo. Assim, concluiu-se pela não-necessidade de diferenciação sazonal dos preços do café.

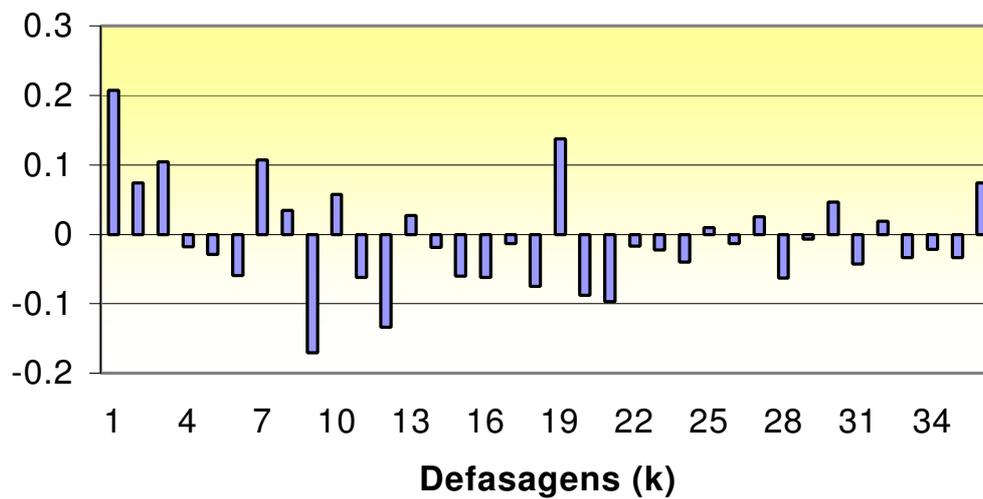
O comportamento dos correlogramas estimados para a série é de fundamental importância na identificação dos valores apropriados para p e q do modelo. Os resultados encontrados para o correlograma e correlograma parcial (Figuras 44 e 45) dos preços do café no mercado *spot* de Nova Iorque indicam que os coeficientes de autocorrelação estimados apresentam comportamento similar, tendo a forma aproximada de senóides amortecidas.

De acordo com o proposto anteriormente no Quadro 2, como ambos os correlogramas tiveram a forma de senóides amortecidas, tem-se um forte indício de que o tipo de especificação que possivelmente será mais apropriada para a modelagem dos preços será dado por uma equação composta por componentes auto-regressivos (AR) e por componentes de médias móveis (MA), e não por um modelo com apenas um desses tipos de termos, o que poderia ser o caso se outros padrões de comportamento para os correlogramas tivessem sido verificados.



Fonte: Resultados da pesquisa.

Figura 44 - Função de autocorrelação para DPRCAFE, com $k = 36$.



Fonte: Resultados da pesquisa.

Figura 45 - Função de autocorrelação parcial para DPRCAFE, com $k = 36$.

No entanto, os valores exatos para p e q não podem ser determinados com tanta certeza. A respeito desse problema dos modelos do tipo ARIMA, MAKRIDAKIS e HIBON (1997:150) mencionam:

“A ordem de um modelo ARIMA é encontrada pelo exame das autocorrelações e autocorrelações parciais de uma série estacionária. BOX e JENKINS (1976) providenciaram uma estrutura teórica e regras práticas para a determinação de valores apropriados para p e q , assim como para as suas contrapartidas sazonais P e Q . A possível dificuldade é a de que geralmente mais de um modelo pode ser considerado, requerendo que o usuário escolha um deles sem nenhum conhecimento das implicações de sua escolha na precisão das previsões pós-amostrais dado que, de acordo com a metodologia de Box-Jenkins, qualquer modelo que resulte em resíduos aleatórios é um modelo apropriado. Box e Jenkins recomendam o princípio da parcimônia, significando que um modelo mais simples (com menos parâmetros) deveria ser selecionado no caso de mais de um modelo possível...”

Tendo isso em vista, foram estimados diferentes configurações de modelos ARIMA e SARIMA, buscando-se selecionar alguns que apresentassem as características necessárias para a estimação do modelo GARCH para o estudo da volatilidade dos preços do café, que é o objetivo final dessa parte da análise.

Os modelos mais parcimoniosos, com menores defasagens para os termos (AR) e (MA), não se apresentaram satisfatórios, pois não foram capazes de gerar resíduos aleatórios. Essa é uma pré-condição para se assumir que um dado modelo explica de forma satisfatória o comportamento dos dados analisados e que pode ser utilizado para previsões e tomadas de decisões.

Todavia, alguns modelos com maiores defasagens foram capazes de gerar resíduos não-correlacionados e com média zero⁶². No Quadro 12 e no Apêndice G) podem-se observar os principais resultados dos modelos que apresentaram os melhores ajustamentos.

⁶² O pressuposto de homocedasticidade será testado e discutido amplamente a seguir, dada a sua importância e ligação direta com a análise de volatilidade propiciada pelos modelos do tipo GARCH.

Quadro 12 - Resultados dos principais modelos estimados para os preços do café no mercado *spot* de Nova Iorque de janeiro de 1946 a dezembro de 2000

Modelo	Variável	Coefficiente	Valores "p"	AIC	Sch.	N.º de iter.	Obs.	Análise dos resíduos
I SARIMA (9,1,10)x(0,0,1) ₁₂	AR(1)	0,244456	0,00	9,0832	9,1383	6	651	A partir de k = 28 os valores-P do teste Q passam a indicar a presença de autocorrelação
	AR(2)	0,069219	0,07					
	AR(7)	0,110631	0,00					
	AR(9)	-0,211293	0,00					
	MA(6)	-0,081601	0,04					
	MA(8)	0,073327	0,07					
	MA(10)	0,155942	0,00					
SMA(12)	-0,177190	0,00						
II SARIMA (7,1,9)x(0,0,1) ₁₂	AR(2)	0,124683	0,00	9,0963	9,1512	6	653	Apenas para k=32 e k=36 os valores-P do teste Q foram significativos
	AR(3)	0,067763	0,08					
	AR(7)	0,066266	0,10					
	MA(1)	0,249349	0,00					
	MA(6)	-0,097578	0,01					
	MA(8)	0,090340	0,02					
	MA(9)	-0,194762	0,00					
SMA(12)	-0,188594	0,00						
III ARIMA (0,1,12)	MA(1)	0,205526	0,00	9,1085	9,1697	7	660	Para k=28, k=32, k=33 e k=36 os valores-P do teste Q foram significativos
	MA(2)	0,116066	0,00					
	MA(3)	0,113545	0,00					
	MA(6)	-0,074961	0,04					
	MA(7)	0,066633	0,08					
	MA(8)	0,107063	0,01					
	MA(9)	-0,153362	0,00					
MA(10)	0,067461	0,08						
MA(12)	-0,205752	0,00						

Fonte: Resultados da pesquisa.

Uma possível razão para o fato de que os modelos que apresentaram resíduos aleatórios sejam modelos com grande número de defasagens pode estar ligada ao fato de que o período de tempo que está sendo analisado é relativamente extenso (mais de 50 anos) e a periodicidade dos dados é mensal. Como foram detectados padrões de comportamento cíclico e sazonal nos preços do café, era de se esperar que ocorressem correlações significativas entre observações relativamente distantes no tempo.

De acordo com os critérios de informação de Akaike (AIC) e de Schwarz (Sch.) estimados, o modelo I seria o mais apropriado para a modelagem dos preços do café no período analisado, pois foi a especificação que apresentou os menores valores para ambos. A segunda melhor opção seria dada pelo modelo II, pois este apresentou valores um pouco maiores que o modelo I para esses dois critérios. Dentre os três modelos, o III foi o que apresentou pior ajustamento, de acordo com essas estatísticas.

A análise do correlograma dos resíduos mostrou que os originados pelo modelo I não estão totalmente livres do problema de autocorrelação, pois a partir da vigésima oitava defasagem os valores de probabilidade calculados para a estatística Q de Ljung-Box, até 36 defasagens, indicaram a presença de autocorrelações estatisticamente significativas.

De acordo com esse teste, o modelo que apresentou melhores resultados foi o II. Para este modelo apenas as autocorrelações com defasagens $k = 32$ e $k = 36$ apresentaram valores-P, para o teste Q, que foram significativos em nível de confiança de 10%. Entretanto, em nível de 5%, por exemplo, não se poderia concluir pela presença de autocorrelações significativas para o mesmo, como se pode observar no Quadro 13. O modelo III também apresentou resultados para este teste melhores que os verificados para o modelo I, porém estes não foram mais satisfatórios que os resultados do modelo II.

Como a aleatoriedade dos erros é uma pré-condição fundamental para se concluir pela relevância de um modelo do tipo ARIMA para a análise e previsão de uma série, optou-se pelo modelo II para ser utilizado na análise e modelagem da volatilidade do preços do café. Assim, tem-se que o modelo II fará o papel da equação genérica dada pela expressão (66) discutida anteriormente, no contexto mais amplo da modelagem da volatilidade condicional por meio de um modelo do tipo GARCH.

Quadro 13 - Correlogramas do modelo II para 36 defasagens

Autocorrelação	Autocorrelação parcial	d.	CAC	CAC parc.	Est.-Q	Prob.
. .	. .	1	-0,019	-0,019	0,2332	
. .	. .	2	-0,009	-0,009	0,2845	
. .	. .	3	-0,002	-0,003	0,2878	
. .	. .	4	0,004	0,003	0,2965	
. .	. .	5	-0,016	-0,016	0,4628	
. .	. .	6	-0,005	-0,005	0,4774	
. .	. .	7	0,003	0,002	0,4819	
. .	. .	8	-0,002	-0,002	0,4854	
. .	. .	9	0,019	0,019	0,7323	0,392
. .	. .	10	0,028	0,029	1,2643	0,531
. .	. .	11	-0,035	-0,034	2,0794	0,556
. .	. .	12	0,007	0,007	2,1149	0,715
. .	. .	13	0,011	0,010	2,1899	0,822
. .	. .	14	-0,018	-0,018	2,4158	0,878
. .	. .	15	-0,012	-0,011	2,5065	0,927
* .	* .	16	-0,106	-0,108	10,109	0,257
. .	. .	17	-0,037	-0,042	11,024	0,274
. .	. .	18	-0,031	-0,035	11,677	0,307
. *	. *	19	0,077	0,074	15,713	0,152
. .	. .	20	-0,046	-0,043	17,131	0,145
. .	. .	21	-0,040	-0,043	18,203	0,150
. .	* .	22	-0,050	-0,058	19,875	0,134
. .	. .	23	-0,028	-0,033	20,399	0,157
. .	. .	24	-0,013	-0,011	20,516	0,198
. .	. .	25	-0,013	-0,012	20,639	0,243
. .	. .	26	-0,019	-0,016	20,876	0,286
. *	. *	27	0,072	0,067	24,441	0,180
* .	* .	28	-0,073	-0,077	28,100	0,107
. .	. .	29	-0,001	-0,006	28,100	0,137
. .	. .	30	0,048	0,053	29,698	0,126
. .	. .	31	-0,036	-0,040	30,608	0,133
. *	. .	32	0,067	0,059	33,740	0,089
. .	. .	33	-0,025	-0,035	34,186	0,104
. .	. .	34	0,012	0,003	34,288	0,128
. .	. .	35	-0,054	-0,041	36,332	0,108
. .	. .	36	0,062	0,052	38,973	0,081

Fonte: Resultados da pesquisa.

Obs.: As probabilidades da Estatística-Q foram ajustadas para OITO termos ARIMA.

O princípio da parcimônia também foi levado em questão na escolha do modelo II, pois este apresenta menos parâmetros que o modelo III, além de menores valores para AIC e Sch., conforme verificado. Pode-se ressaltar que o coeficiente do termo AR(7) estimado para o modelo II está no limite do nível de significância adotado no presente trabalho, que é de 10%, o que não é desejável; entretanto, os resíduos do modelo estimado sem este termo indicaram que este é importante para tornar a série resultante aleatória e que, portanto, ele não deve ser eliminado do modelo.

Outro teste importante que foi implementado para os resíduos do modelo II estimado diz respeito à verificação do valor médio dos resíduos. Se o modelo SARIMA (7,1,9) x (0,0,1)₁₂ estiver bem especificado, os seus resíduos (ϵ_t) devem possuir média que estatisticamente não seja diferente de zero, além de não serem correlacionados conforme se verificou. Para este teste foi utilizada a distribuição-*t* de *Student*, e os resultados encontrados podem ser visualizados no Quadro 14.

Quadro 14 - Teste de hipóteses para a média dos resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂

Teste da Hipótese Nula: Média dos Resíduos = 0.000000		
Amostra: 1946:01 2000:12		
Observações incluídas: 660		
Média amostral = 0,100619		
Desvio-padrão amostral = 22,59926		
<u>Método</u>	<u>Valor</u>	<u>Probabilidade</u>
Estatística - <i>t</i>	0,114382	0,9090

Fonte: Resultados da pesquisa.

Como os resultados indicam, existe forte evidência estatística de que a média dos resíduos do modelo SARIMA estimado para os preços do café seja de fato igual a zero. Em nível de significância de 10%, não se pode rejeitar a hipótese nula de que a média desses resíduos seja igual a zero. Mesmo em testes menos rigorosos, com $\alpha = 25\%$, por exemplo, essa hipótese não poderia ser rejeitada, dado o elevado valor de probabilidade calculado (0,9090).

A ausência de correlação serial nos resíduos deste modelo também foi testada, a fim de confirmar os resultados do teste-Q, pela implementação do teste do multiplicador de Lagrange de Breusch-Godfrey para diferentes defasagens dos resíduos. A hipótese nula neste teste é a de que não existe correlação serial nos resíduos do modelo SARIMA para os preços do café, com diferentes defasagens. Valores de probabilidade significativos (próximos de zero) associados às *estatísticas-F* e aos R^2 observados ($n.R^2$) são indicativos de que a hipótese nula é falsa. Os resultados encontrados para esses testes se encontram no Quadro 15.

Os resultados encontrados para os testes do multiplicador de Lagrange indicam de maneira unânime que, de fato, os resíduos do modelo SARIMA $(7,1,9) \times (0,0,1)_{12}$ não apresentam o problema de correlação serial, o que confirma o resultado do teste-Q implementado anteriormente.

3.4.2. Especificação de um modelo para a variância condicional dos preços do café

Após a especificação do modelo SARIMA $(7,1,9) \times (0,0,1)_{12}$ para descrever o comportamento dos preços, foram realizados testes para a média e para detectar a presença de autocorrelação nos resíduos deste, e os resultados encontrados indicaram que eles se comportam como ruído branco.

Quadro 15 - Resultados do teste do multiplicador de Lagrange (LM) para os resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂

Número de defasagens	Estatísticas	Significância	
<u>3 Defasagens</u>			
Estatística-F	1,484091	Valor-P	0,217673
R ² Observado	4,497005	Valor-P	0,212558
<u>6 Defasagens</u>			
Estatística-F	1,248349	Valor-P	0,279574
R ² Observado	7,565155	Valor-P	0,271722
<u>12 Defasagens</u>			
Estatística-F	0,895160	Valor-P	0,551698
R ² Observado	10,89605	Valor-P	0,537848
<u>24 Defasagens</u>			
Estatística-F	1,273948	Valor-P	0,172769
R ² Observado	30,64128	Valor-P	0,164397
<u>36 Defasagens</u>			
Estatística-F	1,307735	Valor-P	0,111052
R ² Observado	46,85722	Valor-P	0,106284

Fonte: Resultados da pesquisa.

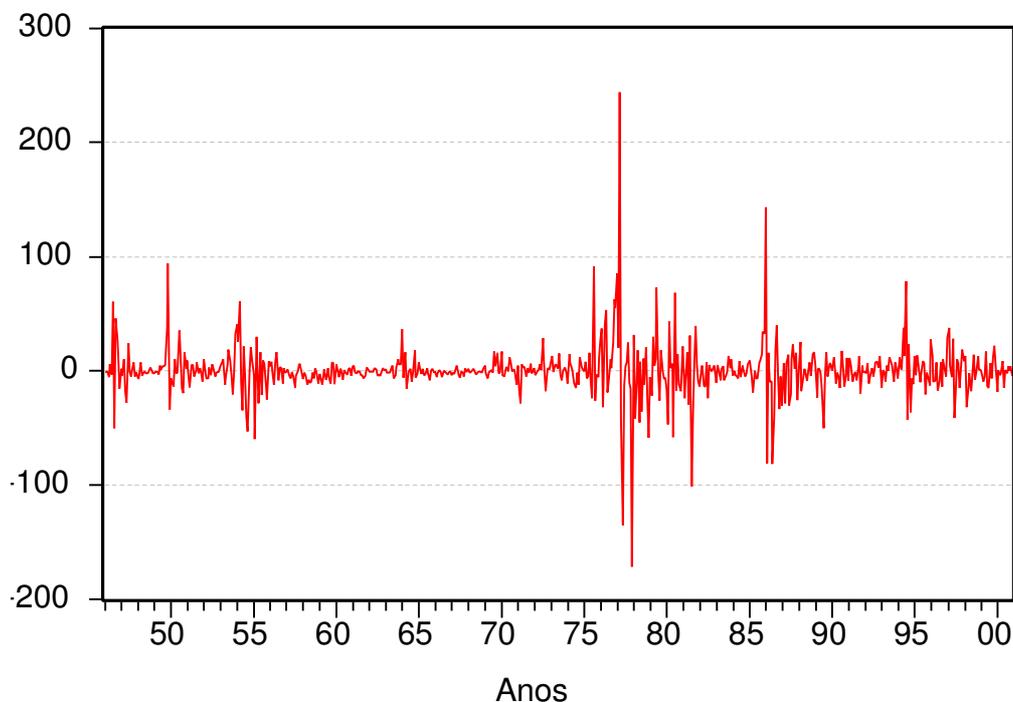
Todavia, importantes testes acerca do comportamento da variância dos resíduos também devem ser implementados. Primeiramente deve-se verificar se esses resíduos apresentam variância constante σ^2 (caso em que se diz que estes apresentam homocedasticidade); se não, a variância estaria intrinsecamente ligada ao período de tempo (σ_t^2), e neste caso diz-se que os resíduos da equação apresentam o problema econométrico heterocedasticidade.

Caso a heterocedasticidade seja detectada, pode-se modelar a variância desses resíduos por meio de uma equação do tipo da equação (67), discutida

anteriormente, que se refere a um modelo de Heterocedasticidade Condicional Auto-regressiva Generalizado (GARCH).

Assim, foram implementados alguns testes para a variância dos resíduos do modelo SARIMA estimado, visando verificar se ela apresenta comportamento de heterocedasticidade condicional auto-regressiva, ou seja, se apresenta um padrão de comportamento do tipo GARCH.

A análise visual do comportamento dos resíduos e de sua dispersão, por meio da Figura 46, se constitui em um instrumento de apoio no processo de identificação da estrutura da variância. Conforme se pode verificar, a dispersão desses resíduos não é constante durante o período analisado; pelo contrário, ela varia de maneira significativa de acordo com o período de tempo em questão. De 1946 a 1956, os resíduos do modelo apresentaram relativa volatilidade quando comparados com seu comportamento nos 20 anos seguintes, que foram os anos em que a variância apresentou menor intervalo.



Fonte: Resultados da pesquisa.

Figura 46 - Resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂.

De 1975 a 1982, o intervalo de oscilação dos resíduos ampliou-se significativamente, caracterizando-se como o período de maior volatilidade para eles. A partir daí a variância dos resíduos passou a apresentar repetidas mudanças de padrão, em que momentos de relativa estabilidade e pequena variância foram alternados com momentos de alta volatilidade, como os verificados em 1986 e no ano de 1994.

De maneira geral, pode-se detectar, pela análise do gráfico dos resíduos do modelo, a incidência de heterocedasticidade. A instabilidade da variância desses dados se caracterizou como regra geral e não como exceção para o período analisado, o que indica que os modelos econométricos empregados para análise e previsão dos preços do café brasileiro no mercado internacional deveriam incorporar essas características irregulares da variância.

No entanto, os resultados inferidos pela análise gráfica dos resíduos devem ser corroborados por meio de testes mais formais, para detecção de heterocedasticidade condicional nos resíduos estimados. Nesse sentido, tem-se como um procedimento usual, conforme aponta ENDERS (1995), estimar as funções de autocorrelação e autocorrelação parcial do quadrado dos resíduos do modelo e verificar a presença de valores significativos por meio de estatísticas como a do teste Q. Os resultados encontrados para este teste podem ser verificados no Quadro 16.

Pode-se verificar que os valores das autocorrelações com duas e nove defasagens foram altamente significativos, indicando a existência do fenômeno de heterocedasticidade condicional nos resíduos do modelo. Os valores de probabilidade das estatísticas Q foram todos indicativos de que se deve rejeitar a hipótese de que os $\hat{\varepsilon}_t^2$ não são correlacionados, o que é equivalente a se rejeitar a hipótese de que os erros não possuem estrutura do tipo GARCH. A hipótese de ausência de correlação no quadrado dos resíduos estimados pode ser então rejeitada em nível de significância de 10% e mesmo em nível de 1%.

Quadro 16 - Correlogramas do quadrado dos resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂ para k = 36

Autocorrelação	Autocorrelação parcial	d.	CAC	CAC parc.	Est.-Q	Prob.
.I*		1	0,070	0,070	3,2423	
.I**		2	0,276	0,272	53,281	
.I.		3	0,034	0,001	54,062	
.I.		4	0,060	-0,019	56,397	
.I.		5	0,043	0,033	57,599	
.I.		6	0,018	0,003	57,802	
.I*		7	0,095	0,081	63,821	
.I.		8	0,016	0,002	63,991	
.I**		9	0,313	0,289	129,03	0,000
.I.		10	0,034	-0,003	129,79	0,000
.I*		11	0,066	-0,109	132,73	0,000
.I.		12	0,060	0,053	135,11	0,000
.I.		13	0,039	0,041	136,14	0,000
.I.		14	0,028	-0,023	136,68	0,000
.I.		15	0,000	-0,023	136,68	0,000
.I.		16	0,010	-0,022	136,75	0,000
.I.		17	0,002	0,007	136,75	0,000
.I.		18	0,026	-0,073	137,20	0,000
.I*		19	0,101	0,112	144,03	0,000
.I.		20	0,012	0,033	144,14	0,000
.I.		21	0,051	-0,044	145,90	0,000
.I.		22	-0,006	-0,038	145,92	0,000
.I.		23	-0,005	-0,008	145,94	0,000
.I.		24	-0,007	0,009	145,98	0,000
.I.		25	0,014	0,027	146,11	0,000
.I.		26	-0,004	-0,011	146,12	0,000
.I.		27	0,048	0,070	147,72	0,000
.I.		28	0,025	-0,050	148,14	0,000
.I.		29	-0,001	-0,042	148,14	0,000
.I.		30	0,008	0,024	148,18	0,000
.I.		31	0,015	0,036	148,33	0,000
.I.		32	-0,014	-0,026	148,46	0,000
.I.		33	-0,010	-0,029	148,53	0,000
.I.		34	-0,009	-0,006	148,58	0,000
.I.		35	0,006	0,031	148,60	0,000
.I.		36	0,000	-0,037	148,60	0,000

Fonte: Resultados da pesquisa.

O teste do Multiplicador de Lagrange, proposto por ENGLE (1982) para testar a hipótese nula (de não-correlação das variâncias), também foi utilizado para confirmação dos resultados anteriores. O Quadro 17 mostra um resumo dos valores encontrados para este teste.

Quadro 17 - Resultados do teste do Multiplicador de Lagrange (LM) para os resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂

Defasagens de $\hat{\varepsilon}_t^2$	Estatística-F	Valor-P	R ² Observado	Valor-P
1	3,23	0,07	3,23	0,07
3	18,42	0,00	51,24	0,00
5	11,21	0,00	52,04	0,00
7	8,61	0,00	55,76	0,00
9	13,68	0,00	104,73	0,00
11	11,94	0,00	110,80	0,00
13	10,31	0,00	112,82	0,00

Fonte: Resultados da pesquisa.

Conforme discutido anteriormente, este teste consiste basicamente em tomar-se o R² da equação (65) e multiplicá-lo pelo número de observações T da amostra; o coeficiente encontrado possuirá distribuição Qui-quadrado com 1 grau de liberdade. Comparando o valor encontrado com o tabelado, pode-se definir pela significância ou não dos coeficientes estimados. Caso um ou mais de um dos coeficientes dessa equação sejam significativos, tem-se que a variância (volatilidade) dos resíduos do modelo estimado apresenta padrão comportamental do tipo GARCH, isto é, a heterocedasticidade dos resíduos será do tipo condicional.

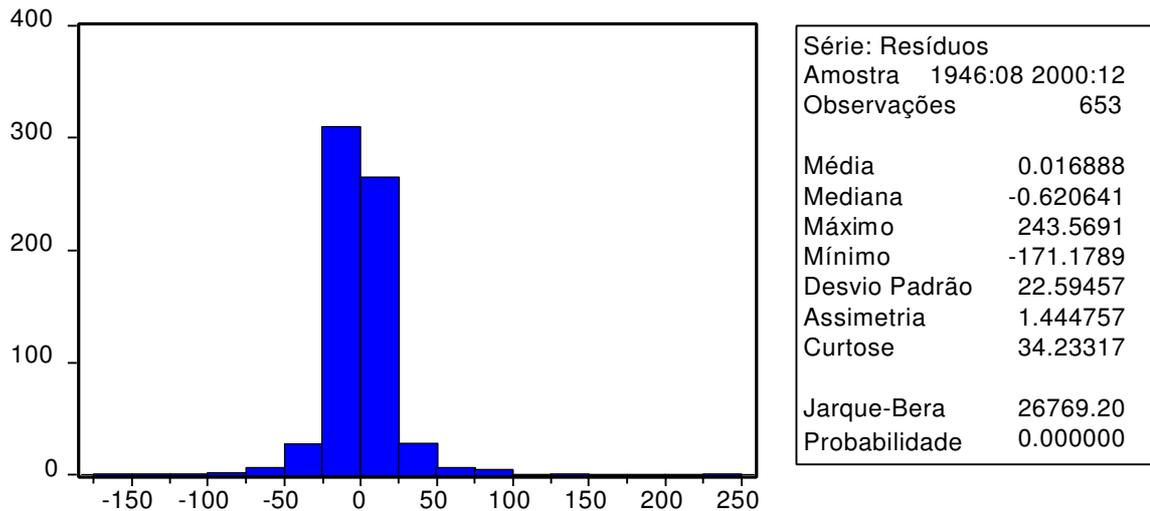
Os resultados encontrados, para diferentes defasagens do quadrado dos resíduos, foram unânimes no sentido de apontarem para a conclusão de que, no caso dos preços do café brasileiro no mercado *spot* de Nova Iorque, um modelo que incorporasse termos de volatilidade condicional (modelos do tipo GARCH) seria mais apropriado para utilizar em análises e previsões desta série do que um modelo em que a variância da série não fosse modelada. Isso porque, como se pôde verificar, todas as *estatísticas-F* e todos os R^2 *obs.* calculados foram significativos em nível de 10% de significância, implicando que existe a presença de termos do tipo GARCH nos resíduos do modelo SARIMA (7,1,9)x(0,0,1)₁₂, estimado para os preços do café. Esses resultados confirmam aqueles encontrados anteriormente na análise dos valores de probabilidade dos testes-Q da função de autocorrelação do quadrado dos resíduos deste modelo.

Constatada a necessidade de se modelar a volatilidade dos preços por meio de um modelo do tipo GARCH, deve-se passar ao processo de identificação e estimação do melhor modelo dentre as várias possibilidades existentes dadas pelos modelos da chamada “família GARCH”⁶³.

Deve-se ressaltar que na estimativa dos modelos GARCH para os preços do café foi adotada a opção de Covariâncias Consistentes de Heterocedasticidade (*Heteroskedasticity Consistent Covariances*) para o cálculo dos erros-padrões e das covariâncias pelo método de quase-máxima verossimilhança. Isso foi feito, pois os resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂ estimados para os preços do café não se apresentaram como sendo normalmente distribuídos, apesar de possuírem média zero, conforme apontam os resultados dos testes de normalidade dados na Figura 47.

O valor de probabilidade da estatística de Jarque-Bera indica que se deve rejeitar a hipótese de normalidade dos resíduos. Os coeficientes de assimetria e de curtose, que para uma distribuição normal deveriam ser próximos a 0 e 3, respectivamente, foram iguais a 1,44 e 34,23, indicando que a distribuição desses resíduos é bem menos “achatada” do que a distribuição normal.

⁶³ ARCH(*p*), GARCH(*p,q*), TARARCH, E-GARCH, ARCH-M, dentre outros, conforme descritos na metodologia do presente trabalho.



Fonte: Resultados da pesquisa.

Figura 47 - Resultados do teste de normalidade dos resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂.

Dada a estimativa do correlograma do quadrado dos resíduos do modelo SARIMA (7,1,9) x (0,0,1)₁₂, apresentado no Quadro 16, pode-se verificar que os valores das autocorrelações com duas e nove defasagens foram os mais significativos, o que indica que uma especificação possível para o modelo de heterocedasticidade condicional poderia ser dada pelo modelo ARCH(9). Assim, tem-se que os resultados deste modelo encontrados para a volatilidade dos preços do café no mercado *spot* de Nova Iorque estão apresentados na equação (81):

$$\sigma_t^2 = 39,33 + 0,57e^2_{t-1} + 0,09e^2_{t-2} + 0,05e^2_{t-3} + 0,01e^2_{t-4} + 0,06e^2_{t-5} - 0,01e^2_{t-6} + 0,01e^2_{t-7} + 0,02e^2_{t-8} + 0,008e^2_{t-9} \quad (81)$$

(0,00) (0,04) (0,70) (0,63) (0,64) (0,12) (0,07) (0,66) (0,43) (0,35)

AIC = 8,0274 Sch. = 8,1509

Os valores entre parênteses, abaixo dos coeficientes do modelo ARCH(9) estimado, representam os valores exatos de probabilidade (*P-values*) calculados. Eles indicam que a maioria dos coeficientes estimados não é significativa, o que sugere que uma especificação mais parcimoniosa deveria ser levada em questão.

Buscou-se, portanto, encontrar um modelo com coeficientes significativos e com menores valores para os critérios de informação de Akaike e de Schwarz, o que foi feito a partir da estimação de modelos ARCH (p) com defasagens decrescentes para os termos e^2_t , que representam as inovações na volatilidade verificadas nos p períodos anteriores. As principais especificações estimadas foram as representadas pelos modelos ARCH(6) e ARCH(3), cujos resultados foram, respectivamente:

$$\sigma^2_t = 7,81 + 1,22 e^2_{t-1} + 0,38 e^2_{t-2} + 0,12 e^2_{t-3} - 0,03 e^2_{t-4} + 0,04 e^2_{t-5} + 0,15 e^2_{t-6} \quad (82)$$

(0,00) (0,00) (0,03) (0,02) (0,00) (0,16) (0,00)

AIC = 7,8832 Sch. = 7,9861

$$\sigma^2_t = 13,66 + 1,16 e^2_{t-1} + 0,44 e^2_{t-2} + 0,36 e^2_{t-3} \quad (83)$$

(0,00) (0,00) (0,01) (0,00)

AIC = 7,9261 Sch. = 8,0084

Tomando-se por base os valores encontrados para os critérios de informação de Akaike e de Schwarz, o modelo ARCH (6) seria o indicado para se representar a volatilidade condicional do preço do café no mercado internacional, porém o coeficiente do termo e^2_{t-5} não apresentou significância estatística em nível de 10%. Por outro lado, no modelo ARCH (3), todos os coeficientes foram estatisticamente significativos, indicando que este modelo seria, portanto, mais apropriado que os anteriores para a modelagem da volatilidade.

Uma outra especificação que foi testada é a representada pelos modelos do tipo GARCH (p, q), que representam um refinamento do modelo ARCH (p). Nestes, seria possível estimar um modelo capaz de descrever o comportamento irregular da variância dos preços do café com um número ainda menor de parâmetros (o que atenderia ao *princípio da parcimônia*). De acordo com o discutido anteriormente, neste modelo a variância dos erros (σ^2_t) do modelo estimado para os preços do café no período t dependerá basicamente de seguintes termos: um termo médio ou constante ω ; inovações (choques) na volatilidade

observadas para períodos anteriores, dadas pelo quadrado dos resíduos (e_{t-i}^2) dos períodos $t-i$, que representaria a mesma estrutura do modelo ARCH (p); e, finalmente, as previsões da volatilidade feitas nos últimos períodos ($\hat{\sigma}_{t-i}^2$), que representariam o termo GARCH.

O modelo GARCH (1,1) por exemplo, implicaria, portanto, que a volatilidade dos preços do café hoje (no mês t) poderia ser definida como formada por uma média ponderada de uma *variância média*, ou constante de longo prazo, pela *previsão da variância* feita nos meses anteriores e "o que foi aprendido" no mês passado.

Na maioria dos estudos empíricos, a especificação GARCH (1,1) se apresenta como suficiente para modelar o comportamento da volatilidade das séries econômicas, e, assim, esta se torna uma especificação inicial natural para o presente trabalho de pesquisa⁶⁴. No entanto, outras especificações também foram testadas, a fim de se verificar o "melhor modelo" para a volatilidade dos preços do café brasileiro no mercado internacional, dentre as quais algumas do modelo TARARCH (p,q). As equações a seguir representam de maneira resumida os modelos GARCH (1,1), GARCH (2,1), GARCH (2,2) e TARARCH (2,1), respectivamente, que foram os que apresentaram os melhores resultados⁶⁵.

$$\begin{aligned} \sigma_t^2 &= 5,88 + 0,86 e_{t-1}^2 + 0,50 \sigma_{t-1}^2 & (84) \\ &(0,03) \quad (0,00) \quad (0,00) \\ &\text{AIC} = 7,9255 \quad \text{Sch.} = 8,0010 \end{aligned}$$

$$\begin{aligned} \sigma_t^2 &= 3,56 + 1,14 e_{t-1}^2 - 0,73 e_{t-2}^2 + 0,74 \sigma_{t-1}^2 & (85) \\ &(0,10) \quad (0,00) \quad (0,00) \quad (0,00) \\ &\text{AIC} = 7,9071 \quad \text{Sch.} = 7,9895 \end{aligned}$$

⁶⁴ De acordo com GREENE (1997) e HALL et al. (1995), a maioria dos processos de volatilidade condicional tem sido modelada como processos GARCH (1,1), e poucos são os casos em que modelos de ordem superior apresentam resultados significativamente melhores.

⁶⁵ Para os resultados completos, ver Apêndice.

$$\sigma^2_t = 9,55 + 0,83 e^2_{t-1} + 0,42 e^2_{t-2} - 0,18\sigma^2_{t-1} + 0,40\sigma^2_{t-2} \quad (86)$$

(0,09) (0,00) (0,09) (0,12) (0,00)

AIC = 7,9011 Sch. = 7,9903

$$\sigma^2_t = 3,35 + 1,05 e^2_{t-1} - 0,42 e^2_{t-2} - 0,47 e^2_{t-1} \cdot d_{t-1} + 0,73\sigma^2_{t-1} \quad (87)$$

(0,02) (0,00) (0,00) (0,00) (0,00)

AIC = 7,8318 Sch. = 7,9211

Os valores calculados para os critérios de informação de Akaike e de Schwarz para esses modelos foram relativamente menores que os verificados para os modelos ARCH (3), o que poderia apontá-los como mais apropriados para se representar a volatilidade do preço do café, principalmente o modelo TARARCH (2,1). Além disso, todos os coeficientes deste modelo apresentaram significância estatística em nível de 10%, inclusive o coeficiente de assimetria na volatilidade (γ). Os resultados deste modelo e dos demais modelos da família GARCH estimados para os preços do café (Apêndice H) apresentaram um comportamento que se constitui em um resultado extremamente importante para os agentes econômicos envolvidos com os mercados físicos (*spot*) e futuros de café, sejam eles produtores, processadores, *hedgers* e/ou especuladores.

Como se pode verificar, os *coeficientes de persistência da volatilidade* dados pelas somas ($\Sigma\alpha_i + \Sigma\beta_i$) dos coeficientes dos modelos GARCH e pelas somas ($\Sigma\alpha_i + \Sigma\beta_i + \gamma/2$), para o caso do modelo TARARCH, foram todos valores maiores ou próximos de um. Esses coeficientes mensuram a duração dos choques de volatilidade; conforme CUDDINGTON e LIANG (1999), valores maiores ou iguais a um para estes significam que os choques na volatilidade irão perdurar por muito tempo na série em questão.

Portanto, quanto maior (mais próximo de 1) for o valor do coeficiente de persistência, mais vagarosamente o choque sobre a volatilidade irá se enfraquecer, e, assim, maior será a demora do processo de reversão à média para a variância. No caso dos preços do café em que os valores dos coeficientes de

persistência foram maiores ou aproximadamente iguais a 1, tem-se que os choques na volatilidade irão perdurar por um período extremamente longo.

Esses resultados mostram que a variância condicional de e_t para os preços do café possui raiz unitária e esta não apresentará comportamento de reversão à sua média histórica com o passar do tempo, após um choque. Todavia, deve-se ressaltar que isso, de acordo com NELSON (1990), não significa que a série de resíduos e_t não seja estacionária, mas sim que a variância incondicional de e_t não será.

Esses resultados demonstram, portanto, que as informações correntes sobre a volatilidade permanecerão importantes para as previsões da variância condicional dos preços do café por um horizonte indefinido de tempo, o que, em última instância, indica elevado grau de risco de preço (e de renda, de forma indireta) associado ao mercado *spot* do café. Talvez esta seja, inclusive, uma das principais razões para o sucesso dos contratos futuros e de opções de café na BM&F (com participação de 68% no volume financeiro total movimentado pelos contratos agropecuários), o que ainda não ocorre com outras *commodities*, à exceção dos contratos de boi gordo (com 19% do volume financeiro transacionado) e do recente contrato de álcool anidro, que responde por 9,5% do total dos contratos agropecuários negociados.

De maneira geral, quanto maior o risco de preço de uma *commodity*, maior a necessidade que os agentes envolvidos com esse produto terão de se precaver contra possíveis perdas; sendo o produto em questão um produto com grande participação no PIB agropecuário de um país (como é o caso do café no Brasil), maior a probabilidade de se ter um volume expressivo de transações no mercado futuro para essa *commodity*.

4. RESUMO E CONCLUSÕES

O mercado *spot* do café caracteriza-se como sendo marcado por acentuadas flutuações de preços. Essas flutuações implicam que a rentabilidade nessa atividade deverá oscilar de maneira significativa, o que pode propiciar para os investidores mais informados um diferencial de lucratividade, pois as suas decisões de *hedge* e de *especulação* estarão baseadas em previsões mais precisas acerca do comportamento do mercado.

No entanto, pode-se observar que essas possibilidades de ganho ainda não são plenamente aproveitadas por grande parte dos agentes econômicos envolvidos nesses mercados, dada a imperfeita compreensão acerca da natureza das flutuações dos preços no mercado de café, que, se não podem ser evitadas, ao menos deveriam ser utilizadas pelos agentes econômicos ligados a esse mercado para otimizar suas decisões e maximizar seus lucros individuais.

Em termos gerais, pretendeu-se neste trabalho de pesquisa detectar e analisar a existência e a magnitude dos componentes estocásticos (e, ou, determinísticos) de *tendência*, *ciclos*, *sazonalidade* e *volatilidade condicional* na série temporal dos preços do mercado *spot* do café brasileiro na bolsa de Nova Iorque (NYBOT) para a segunda metade do século passado, mais especificamente o período compreendido entre janeiro de 1946 e dezembro de 2000.

Optou-se pelo estudo do mercado de café, uma vez que esta é a principal *commodity* agropecuária do Brasil transacionada em bolsas de mercadorias e de futuros (aproximadamente 67% dos contratos agropecuários transacionados na BM&F são de café). Optou-se também pela análise dos preços no mercado internacional, pelo fato de o Brasil ser o maior produtor e exportador mundial de café (30% da produção mundial em média); dessa forma, trata-se de uma *commodity* cuja formação do preço deve contemplar principalmente o mercado internacional.

Do ponto de vista teórico, uma grande motivação para a implementação deste trabalho se deu em função do fato de que as séries históricas, de maneira geral, contêm informações de grande valia para o planejador da atividade agrícola e demais agentes envolvidos com o setor, no que se refere ao comportamento passado da variável em questão: preços do café, no presente caso.

Assim, utilizou-se neste trabalho o instrumental estatístico conhecido por *Teoria da Análise de Séries Temporais* para testar a existência e analisar as implicações da presença dos componentes distintos de *tendência*, *ciclos*, *sazonalidade* e *volatilidade* nos preços *spot* do café brasileiro. Especificamente, a metodologia empregada referiu-se a *análises no domínio do tempo* (para a análise de tendência, sazonalidade e volatilidade) e à *análise no domínio da frequência*, também chamada de *análise espectral*, para o estudo da presença de ciclos nos preços.

Com relação ao componente de tendência nos preços do café no mercado *spot* de Nova Iorque, pode-se verificar na análise que o nível médio dos preços da *commodity* brasileira apresentou, no início do período analisado, comportamento de crescimento a taxas decrescentes. Esse processo foi então revertido, e o nível médio dos preços passou a apresentar na segunda metade do período um comportamento de queda a taxas crescentes, indicando que a situação do setor cafeeiro nacional pode-se tornar problemática nos próximos anos, caso essa tendência não seja revertida. Isso porque, apesar da relativa recuperação dos preços em 1995 e 1997, a análise dos cinco últimos anos da amostra estudada indica agravamento da tendência de queda nos preços; portanto, os produtores de

café deverão buscar se tornar cada vez mais eficientes em suas atividades, visando operar com sistemas otimizados em termos de custos e produtividade, visto que os preços do café estão em um processo de queda relativa nos últimos 25 anos, o que confirma, em parte, a hipótese de Schultz de tendência de queda para os preços do café.

Constatou-se também que, além de apresentar uma tendência determinística, os preços do café apresentam raízes unitárias e são integrados de ordem um (1). Tem-se, portanto, que choques nesses preços terão efeitos permanentes, significando que políticas agrícolas e macroeconômicas que afetem diretamente a cadeia produtiva do café terão impactos duradouros.

Os resultados encontrados mostram que a tendência existente nos preços do café se configurou como uma composição de tendências dos tipos determinística e estocástica, ou seja, a série de preços do café brasileiro não é uma série estacionária e o seu comportamento de longo prazo é marcado pela junção de tendências dos tipos determinística e estocástica.

Ainda em relação à tendência declinante dos preços no mercado externo, algumas medidas para reversão dessa situação deveriam ser implementadas pelos produtores individuais e de maneira coordenada nos âmbitos nacional e internacional (quando de interesse conjunto). Algumas medidas que podem ser apontadas são:

- Busca de diferenciação do produto com ênfase em cafés especiais e do tipo *gourmet*, para atender a demanda de consumidores mais exigentes e de alto poder aquisitivo, como os europeus e americanos.
- Esforços de marketing para promoção dos cafés do Brasil no mercado internacional, sinalizando para a qualidade do produto brasileiro.
- Maior integração dos principais países produtores para definição de metas conjuntas que beneficiem a todos na negociação do produto, dado o cenário de oligopsônio que caracteriza a demanda no mercado internacional, com poucas empresas comprando a maior parte da produção mundial.

Com relação à hipótese de ciclos existentes nos preços do café, a análise espectral para os dados conjuntos de todo o período confirmou a incidência de um ciclo de média duração existente no intervalo de 22 a 44 meses.

Subdividindo a análise por décadas, a fim de verificar mudanças na periodicidade dos ciclos, constatou-se para a década de 50 a existência de um ciclo bienal que possivelmente esteve ligado diretamente ao ciclo biológico do cafeeiro. Este ciclo biológico consiste do fato de que, de maneira geral, os cafeeiros apresentam em um ano uma grande florada e no outro, a seguir, uma florada pequena, em razão do esgotamento das plantas, que necessitam da ocorrência de um intervalo de um ano para poder apresentar bom rendimento.

Para a década de 60, os resultados encontrados divergiram, em certa medida, dos resultados verificados para a década de 50. Neste período, o pico mais marcante e significativo ocorreu nas frequências próximas a 0,02, o que, em termos de período, correspondeu a um ciclo de longo prazo nos preços do café de aproximadamente 60 meses (cinco anos) de duração.

Os espectros estimados para a década de 70 indicaram novamente que as frequências próximas a 0,04 são as que mais contribuem para a variância dos preços do café. O maior pico verificado neste espectro se encontrou no intervalo próximo a esta frequência, indicando que um ciclo bienal (de período igual a 24 meses) também foi marcante para a década de 70. Pode-se notar também um pico menor nas frequências próximas a 0,3, que correspondeu a um período de aproximadamente quatro meses.

Os resultados encontrados para o espectro estimado dos preços do café nos anos 80 apresentaram duas características marcantes. Primeiramente, pode-se observar que o ciclo de 24 meses foi o mais marcante em todo o conjunto de densidades espectrais estimadas, indicando que o ciclo bienal se configurou como um importante determinante do comportamento dos preços do café até o final da década de 80, pois foi um ciclo que prevaleceu ao longo das quatro primeiras décadas analisadas. Em segundo lugar, tornou-se evidente que o ciclo sazonal, de frequências próximas a 0,07, novamente foi significativo nesta década, ao contrário do ocorrido na década de 70.

Finalmente, para a década de 90 pode-se observar uma mudança no comportamento dos preços, em termos do componente de ciclos. O espectro estimado em função das frequências indica que aquelas pertencentes ao intervalo dado por $[0,07; 0,10]$ são as que mais contribuíram para a variância dos preços do café, visto que existe um pico neste espectro que se destaca em relação aos demais. Essas frequências correspondem ao período de 12 meses, o que configura o ciclo sazonal como o mais marcante ciclo nos preços para a década de 90.

Os resultados encontrados na análise de ciclos nos preços do café apontam para a conclusão de que os ciclos nos preços são fenômenos do passado, mas que não mais afetam significativamente o setor, porque na última década os ciclos (em sentido estrito) de fato inexistiram; apenas o “ciclo” sazonal foi detectado. Assim, a persistir esse comportamento, para os anos vindouros os agentes envolvidos com esta *commodity* não têm razões para esperar flutuações nos preços e na rentabilidade de suas atividades em função do componente cíclico.

Em relação à análise da sazonalidade nos preços, observou-se que, para as décadas tomadas individualmente, os resultados não diferiram significativamente dos verificados para todo o período. Constatou-se que a sazonalidade, enquanto componente de influência no comportamento dos preços do café no mercado internacional, não é da forma determinística e regular. Pelo contrário, ela é de natureza estocástica, variando em função do período do tempo, conforme também indicaram alguns dos resultados da análise espectral apresentados anteriormente. Esse resultado se justifica em parte pela formação de estoques, tanto no mercado interno quanto no externo, que são liberados nos períodos de entressafra e que impedem oscilações mais acentuadas nos preços e a caracterização de um padrão sazonal regular para estes.

No que se refere à natureza estocástica da sazonalidade, os resultados apresentaram duas características marcantes: primeiramente, indicaram que não há necessidade de se implementar o processo de diferenciação sazonal na série, pois não foram verificadas *raízes unitárias sazonais* para os preços do café. Em

segundo lugar, pode-se verificar que esses preços tendem a ser correlacionados de maneira sazonal, mas que essa relação só é forte de um ano para o outro e não para períodos mais distantes no tempo.

Portanto, pode-se inferir que, de fato, a sazonalidade nos preços do café brasileiro no mercado *spot* de Nova Iorque foi comprovada. Todavia, deve-se ressaltar que essa relação está se modificando ao longo do tempo, o que caracterizaria uma sazonalidade do tipo estocástica e não do tipo determinística para os preços desta *commodity* no mercado internacional.

Dados os resultados acerca da sazonalidade dos preços, tem-se que os produtores devem ter certa atenção no planejamento da venda de sua safra, no que tange aos meses mais favoráveis para essa venda. Verificou-se que, em média, os meses de junho a setembro se configuram como meses de safra no Brasil e o nível dos preços da *commodity* nesse período tendem a cair significativamente.

Com relação à volatilidade nos preços, os resultados encontrados foram indicativos no sentido de apontarem para a conclusão de que, no caso dos preços do café brasileiro no mercado *spot* de Nova Iorque, um modelo que incorporasse termos de *volatilidade condicional* (modelos do tipo GARCH) seria mais apropriado para se utilizar em análises e previsões desta série do que um modelo em que a variância da série não fosse modelada.

Dentre os modelos testados, os valores calculados para os critérios de informação de Akaike e de Schwarz apontaram o modelo TARARCH (2,1) como o mais apropriado para se representar a volatilidade do preço do café. Todos os coeficientes deste modelo apresentaram significância estatística em nível de 10%, inclusive o coeficiente de assimetria na volatilidade (γ).

Os resultados deste modelo e dos demais modelos da família GARCH estimados para os preços do café apresentaram um comportamento que se constitui em um resultado significativo para os agentes econômicos envolvidos com os mercados físicos (*spot*) e futuros de café, sejam eles produtores, processadores, *hedgers* e, ou, especuladores. Como se pode verificar, os *coeficientes de persistência da volatilidade* foram todos valores maiores ou

próximos de um, implicando que os choques na volatilidade irão perdurar por muito tempo na série em questão. Isto é, os resultados encontrados mostram que a variância condicional de e_t para os preços do café possui raiz unitária e esta não apresentará comportamento de reversão à sua média histórica com o passar do tempo, após um choque.

Verificou-se, portanto, que as informações correntes sobre a volatilidade permanecerão importantes para as previsões da variância condicional dos preços do café por um horizonte indefinido de tempo, o que, em última instância, indica um elevado risco de preço (e de renda, de forma indireta) associado ao mercado *spot* do café brasileiro.

De maneira geral, todos os componentes estruturais (*tendência, ciclos, sazonalidade e volatilidade*) da série temporal dos preços do café brasileiro no mercado *spot* da bolsa de Nova York apresentaram comportamentos instáveis e com natureza eminentemente estocástica, apesar de em alguns casos apresentar uma composição com componentes determinísticos. Isso se reflete, para os produtores e demais agentes econômicos envolvidos com o mercado de café, na forma de um elevado grau de *risco de preço e de renda* nessas atividades. Esses riscos podem proporcionar grandes lucros a esses agentes em determinados momentos, mas também enormes prejuízos e mesmo a saída do mercado em situações adversas.

No entanto, esse alto patamar de risco pode ser minimizado com maior utilização de contratos em mercados futuros de café, em que operações de *hedge* possam ser efetuadas tanto por produtores quanto por processadores da *commodity*. Esses contratos se configuram como o principal “mecanismo de mercado” na atualidade para a negociação do risco com aqueles agentes que de fato desejam corrê-lo. Isso porque, na atual conjuntura econômica interna (de restrição orçamentária do governo) e externa (de combate a políticas protecionistas no comércio internacional), a intervenção governamental no mercado de café por meio de políticas agrícolas, como formação e controle de estoques, definição de preços mínimos, entre outras, se torna impraticável.

Contudo, no Brasil, o volume de negociações efetuadas com futuros agropecuários em geral e mais especificamente com o café ainda está muito aquém do seu potencial. Portanto, uma maior utilização desses instrumentos poderia se configurar como uma opção para a garantia de minimização dos efeitos prejudiciais das oscilações de preços na forma de *tendências, ciclos, sazonalidade e volatilidade*, detectadas no presente trabalho de pesquisa como sendo profundamente marcantes no comportamento dos preços do café do Brasil no mercado internacional.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAHAM, B., LEDOLTER, J. **Statistical methods for forecasting**. New York: John Wiley & Sons, 1983.
- AGRIANUAL 96. **Anuário estatístico da produção agrícola**. São Paulo: FNP, 1996.
- AGUIAR, D.R.D. Séries históricas de preços e a tomada de decisão na agricultura. **Preços Agrícolas**, Piracicaba, n. 67, p. 1-1a, maio 1992.
- ANUALPEC 96. **Anuário estatístico da produção animal**. São Paulo: FNP, 1996. 312 p.
- ATRASAS, A.L. **Formação de preços no mercado futuro brasileiro**. Viçosa: UFV, 1993. Dissertação (Mestrado em Economia Rural) - Universidade Federal de Viçosa, 1993.
- BERA, A., BUBNYS, E.L., PARK, H.Y. Arch effects and efficient estimation of hedge ratios for stock index futures. **Advances in Futures and Options Research**, Boston, v. 6, p. 313-328, 1993.
- BERNSTEIN, P.L. **Desafio aos deuses: a fascinante história do risco**. Rio de Janeiro: Campus, 1997.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, n. 31, p. 307-327, 1986.
- BOWERMAN, B.L., O'CONNEL, R.T. **Time series and forecasting**. California Duxbury Press, 1979.

- BOX, G.E.P., JENKINS, G.M. **Time series analysis: forecasting and control.** 1976.
- BROOKS, J.C., CARTER, C.A. **The political economy of US agriculture.** Canberra, Australia, 1994. (ABARE Research Report, 8).
- BUENO, A.J.T. A transferência do risco de preço no agribusiness brasileiro. In: VIEIRA, W.C., CARVALHO, F.M.A. **Mercosul: agronegócios e desenvolvimento econômico.** Viçosa: UFV, 1997. 296 p.
- CAMPBELL, J.Y., LO, A.W., MacKINLAY, A.C. **The econometrics of financial markets.** New Jersey: Princeton University, 1997.
- CAVALCANTI, J.E.A. **Análise harmônica aplicada às quantidades e aos preços de produtos agrícolas selecionados no Estado de São Paulo.** Viçosa: UFV, 1978. Dissertação (Mestrado em Economia Rural) - Universidade Federal de Viçosa, 1978.
- CHATFIELD, C. **The analysis of time series - an introduction.** 5.ed. London: Chapman & Hall, 1996.
- CONTADOR, C.R. Uma análise espectral dos movimentos da Bolsa de Valores do Rio de Janeiro. **Revista Brasileira de Mercado de Capitais**, Rio de Janeiro, v. 1, n. 1, p. 67-92, 1975
- CORTEZ, J.V., MARTIN, N.B. A sazonalidade da produção da seringueira e a política brasileira de contingenciamento da importação de borracha natural. **Informações Econômicas**, São Paulo, v. 26, n. 7, jul. 1996.
- COUTO, M.T. **Modelos estruturais de séries temporais: o caso dos preços da bovinocultura de corte no Estado de São Paulo.** Piracicaba: ESALQ, 1996. Dissertação (Mestrado em Economia Aplicada) - Escola Superior de Agricultura "Luiz de Queiroz", 1996.
- CRAGG, J. Estimation and testing in time series regressions models with heteroscedastic disturbances. **Journal of Econometrics**, v. 20, p. 135-157, 1982.
- CRAMER, H., LEADBETTER, M.R. **Stationary and related stochastic processes.** New York: Wiley, 1967.
- CUDDINGTON, J.T., LIANG, H. **Appendices for commodity price volatility across exchange rate regimes.** 1999. (Not for publication).
- DAHL, D.C., HAMMOND, J.W. **Market and price analysis - the agricultural industries.** New York: McGraw-Hill, 1977.

- DALY, R.F. The long-run demand for farm products. **Agricultural Economics Research**, v. 8, n. 3, July 1956.
- DEBERTIN, D.L. **Agricultural production economics**. New York: Macmillan, 1986.
- DICKEY, D. **Estimation and hypothesis testing in nonstationary time series**. Iowa State University, 1976. Thesis (Ph.D.) - Iowa State University, 1976.
- DUARTE JR., A.M. **Risco: definições, tipos, medição e recomendações para o seu gerenciamento**. São Paulo: BM&F, 1996. (Resenha BM&F, 114).
- EATON, B.C., EATON, D.F. **Microeconomia**. São Paulo: Saraiva, 1999.
- ENDERS, W. **Applied econometric time series**. New York: John Wiley & Sons, 1995.
- ENGLE, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. **Econometrica**, v. 50, p. 987-1008, 1982.
- ENGLE, R.F., NG, V.K. Measuring and testing the impact of news on volatility. **Journal of Finance**, v. 48, p. 1022-1082, 1993.
- FAVA, V.L. **Metodologia de Box-Jenkins para modelos univariados, em manual de econometria - nível intermediário**. In: VASCONCELOS, M.A.S., ALVES, D. São Paulo: Atlas, 2000.
- FRANÇA, F.M.C. **Técnica de análise espectral univariada: aplicações aos preços do algodão**. Fortaleza: BNB/ETENE, 1990.
- FRANÇA, F.M.C., NETO, J.B.F.G., LEMOS, J.J.S. Análise estacional de preços de algodão para diferentes regiões do Ceará: abordagem pela análise espectral. **Revista de Economia e Sociologia Rural**, Brasília, v. 27, n. 1, jan./mar. 1989.
- FRICK, O. Mercados derivativos no Mercosul. In: VIEIRA, W.C., CARVALHO, F.M.A. **Mercosul: agronegócios e desenvolvimento econômico**. Viçosa: UFV, 1997.
- FULLER, W.A. **Introduction to statistical time series**. New York: Wiley, 1976.
- GLOSTEN, L.R., JAGANNATHAN, R., RUNKLE, D. On the relation between the expected value and the volatility of the normal excess return on stocks. **Journal of Finance**, v. 48, p. 1779-1801, 1993.

- GRANGER, C., HUGHES, A.O. **Spectral analysis of short series - a simulation study**. New York: J.R. Statistical Society, 1968.
- GRANGER, C., NEWBOLD, P. Spurious regressions in econometrics. **Journal of Econometrics**, v. 2, p. 111-120, 1974.
- GREENE, W.H. **Econometric analysis**. 3.ed. New Jersey: Prentice-Hall, 1997.
- GRIFFITHS, W.E., HILL, C., JUDGE, G.G. **Learning and practicing econometrics**. New York: John Wiley & Sons, 1993.
- GUJARATI, D.N. **Basic econometrics**. 3.ed. New York: McGraw-Hill, 1995.
- HALL, R.E. et al. **QMS: eviews 2.0 - user's guide**. Irvine, California, 1995.
- HAMILTON, J.D. **Time series analysis**. Princeton, New Jersey: Princeton University, 1994.
- HASEGAWA, M.M. **O mercado de reposição da pecuária bovina de corte no Estado de São Paulo**. Piracicaba: ESALQ, 1995. Dissertação (Mestrado em Economia Aplicada) - Escola Superior de Agricultura "Luiz de Queiroz", 1995.
- HEIFNER, R., KINOSHITA, R. Differences among commodities in real price variability and drift. **The Journal of Agricultural Economics Research**, v. 45, n. 3, 1994.
- HELMUT, J.W. A report on the systematic downward bias in live cattle futures prices. **Journal of Futures Markets**, v. 1, n. 3, p. 347-358, 1981.
- HIERONYMUS, T.A. **Economics of futures trading for commercial and personal profit**. New York: Commodity Research Bureau, 1976.
- JENKINS, G.M., WATTS, D.C. **Spectral analysis and its applications**. Holden Day, 1968.
- JUDGE, G.G. et al. **Introduction to the theory and practice of econometrics**. 2.ed. New York: Wiley, 1988.
- KENNEDY, P. **A guide to econometrics**. 3.ed. Cambridge: The Mit Press, 1996. 410 p.
- KOPPENHAVER, G.D. The forward pricing efficiency of live cattle futures market. **Journal of Futures Markets**, v. 3, n. 3, p. 307-319, 1983.

- KREPS, D.M. **A course in microeconomic theory**. Harvester Wheatsheaf, 1990.
- MacKINNON, J.G. Critical values for cointegration tests. In: ENGLE, R.F., GRANGER, C.W.J. **Long-run economic relationships: readings in cointegration**. Oxford University, 1991.
- MAKRIDAKIS, S., HIBON, M. ARMA - models and the Box-Jenkins methodology. **Journal of Forecasting**, v. 16, p. 147-163, 1997.
- MARQUES, P.V., CAFFAGNI, L.C., SOUSA, E.L. Mercados futuros e financiamentos agropecuários. In: VIEIRA, W.C., CARVALHO, F.M.A. **Mercosul: agronegócios e desenvolvimento econômico**. Viçosa: UFV, 1997.
- MORICOCCHI, L., MARTIN, N.B., VEGRO, C.L.R. Produção de café nos países concorrentes do Brasil e tendências do consumo mundial. **Informações Econômicas**, São Paulo, v. 27, n. 5, 1997.
- NELSON, D.B. Stationarity and persistence in the GARCH(1,1) model. **Econometric Theory**, v. 6, p. 318-334, 1990.
- NELSON, D.B. Conditional heteroskedasticity in asset returns: a new approach. **Econometrica**, v. 59, p. 347-370, 1991.
- NEWBY, W., WEST, K. Automatic lag selection in covariance matrix estimation. **Review of Economic Studies**, v. 61, p. 631-653, 1994.
- PEREIRA, B.B., PAIS, M.B.Z., SALES, P.R.H. **Análise espectral de séries temporais - uma introdução para engenharia, economia e estatística**. Rio de Janeiro: ELETROBRÁS, 1986.
- PEROBELLI, F.S., SCHOUCHANA, F. **Formação do preço do café na BM&F**. São Paulo: BM&F, 2000.
- PETTI, R.H.V. Abate bovino: mudança no perfil sazonal. **Informações Econômicas**, São Paulo, v. 26, n. 7, jul. 1996.
- PINDYCK, R.S., RUBINFELD, D.L. **Econometric models and economic forecasts**. 4.ed. New York: Irwin/McGraw-Hill, 1998.
- RAUSSER, G.C., CARGILL, T.F. The existence of broiler cycles: an application of spectral analysis. **American Journal of Agricultural Economics**, v. 52, n. 1, p. 109-121, 1970.

- ROCQUE, E.C., GARCIA, M.G.P. **Um estudo sobre a volatilidade do mercado futuro de taxas de juros no Brasil**. São Paulo: BM&F, 1997. (Resenha da BM&F, 117).
- RODRIGUES, R., O papel do setor privado e os novos desafios do abastecimento nacional. **Revista de Política Agrícola**, v. 10, 2001.
- SCHULTZ, T.W. **Agriculture in an unstable economy**. New York: McGraw-Hill, 1945.
- SCHWAGER, J.D. **Fundamental analysis**. New York: John Wiley & Sons, 1995.
- SHEPHERD, G.S. **Agricultural price analysis**. 5.ed. Iowa State University, 1972.
- SIMON, C., BLUME, L. **Mathematics for economists**. New York: W.W. Norton, 1994.
- SOUZA, R.C. Modelos estruturais para previsão de séries temporais: abordagens clássica e bayesiana. In: COLÓQUIO BRASILEIRO DE MATEMÁTICA, 17, 1989, Rio de Janeiro. **Anais...** Rio de Janeiro: IMPA, 1989.
- STOCK, J.H., WATSON, M.W. Variable trends in economic time series. **Journal of Economic Perspectives**, v. 2, n. 3, p. 147-174, 1988.
- THOMSEN, F.L., FOOTE, R.J. **Agricultural prices**. New York: McGraw-Hill, 1952.
- TOMEK, W.G., ROBINSON, K.L. **Agricultural product prices**. 3.ed. London: Cornell University, 1991.
- VANDAELE, W. **Applied time series and Box-Jenkins models**. New York: Academic Press, 1983.
- VAUGHN, R., KELLY, M., HOCHHEIMER, F. Identifying seasonality in futures prices using X-11. **Journal of Future Markets**, v. 1, n. 1, p. 93-101, 1981.

APÊNDICES

APÊNDICE A

Prova 1

Como existe uma identidade de funções trigonométricas que demonstra que se pode fazer $\cos(a + b) = \cos a \cdot \cos b - \sin a \cdot \sin b = \cos a \cdot b - \sin a \cdot b$, tem-se que:

$$Y_t = \mu + \sum_{j=1}^k R_j \cos(\omega_j t + \theta_j) + Z_t$$

$$Y_t = \mu + \sum_{j=1}^k R_j [\cos \omega_j t \cdot \cos \theta_j - \sin \omega_j t \cdot \sin \theta_j] + Z_t$$

$$Y_t = \mu + \sum_{j=1}^k R_j [\cos(\omega_j t \theta_j) - \sin(\omega_j t \theta_j)] + Z_t$$

Prova 2

Para se chegar ao espectro populacional de um processo estacionário $\{Y_t\}_{t=-\infty}^{\infty}$, com média $E(Y_t) = \mu$ e com uma seqüência de autocovariâncias $\{Y_j\}_{j=-\infty}^{\infty}$, em que a j -ésima autocovariância é igual a $\gamma_j = E(Y_t - \mu)(Y_{t-j} - \mu)$ e que

$\sum_{j=0}^{\infty} |\gamma_j| < \infty$, deve-se partir da *Função Geratriz de Autocovariâncias* de Y_t , que é

dada por:

$$\Gamma_Y(z) \equiv \sum_{j=-\infty}^{\infty} \gamma_j z^j \quad (\text{A1})$$

em que z é um *número complexo*. A função geratriz de autocovariâncias é um modo de resumir as autocovariâncias da série temporal dada, sendo que um caso de importância fundamental se dá quando o valor de z se encontra dentro do chamado *círculo unitário complexo*, ou seja:

$$z = \cos(\omega) - i \cdot \text{sen}(\omega) = e^{-i\omega} \quad (\text{A2})$$

em que i é a *unidade imaginária* denotada por $i = \sqrt{-1}$ e ω é um escalar que representa o ângulo em radianos que z forma com o eixo dos números reais. A equação (A2) advém de uma identidade matemática conhecida como *Equação de Euler*, cuja definição e demonstração podem ser encontradas em SIMON e BLUME (1994).

Substituindo (A2) na equação (A1) e dividindo o resultado por 2π , tem-se a seguinte equação:

$$f_Y(\omega) = \frac{1}{2\pi} \Gamma_Y(e^{-i\omega}) = \frac{1}{2\pi} \left(\sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} \right) \quad (\text{A3})$$

que é o *Espectro Populacional* (ou *Função de Densidade Espectral*) do processo $\{Y_t\}_{t=-\infty}^{\infty}$, mencionado anteriormente. Deve ser observado que o espectro é uma função de ω , o que implica que, para um dado valor de ω e uma seqüência de autocovariâncias $\{\gamma_j\}_{j=-\infty}^{\infty}$, pode-se calcular o valor de $f_Y(\omega)$.

Utilizando novamente da equação de Euler, tem-se que o espectro pode ser rescrito da seguinte maneira:

$$f_Y(\omega) = \frac{1}{2\pi} \left\{ \sum_{j=-\infty}^{\infty} \gamma_j [\cos(\omega j) - i \cdot \text{sen}(\omega j)] \right\} \quad (\text{A4})$$

Porém, para processos estacionários, tem-se que $\gamma_j = \gamma_{-j}$, o que implica que a equação (d2) poderá ser representada por:

$$f_Y(\omega) = \frac{1}{2\pi} \gamma_0 [\cos(0) + i \cdot \text{sen}(0)] + \frac{1}{2\pi} \left\{ \sum_{j=1}^{\infty} \gamma_j [\cos(\omega j) + \cos(-\omega j) - i \cdot \text{sen}(\omega j) - i \cdot \text{sen}(-\omega j)] \right\}$$

$$f_Y(\omega) = \frac{1}{2\pi} \gamma_0 \cdot (1) + \frac{1}{2\pi} \left\{ \sum_{j=1}^{\infty} \gamma_j [2\cos(\omega j) - i \cdot \text{sen}(\omega j) + i \cdot \text{sen}(\omega j)] \right\}$$

$$f_Y(\omega) = \frac{1}{2\pi} \gamma_0 + \frac{1}{2\pi} \left\{ \sum_{j=1}^{\infty} \gamma_j [2\cos(\omega j)] \right\}$$

No desenvolvimento acima foram utilizados os seguintes resultados trigonométricos: $\cos(0) = 1$;

$$\text{sen}(0) = 0; \text{sen}(-\alpha) = -\text{sen}(\alpha); \cos(-\alpha) = \cos(\alpha).$$

Prova 3

A integral da *Função de Densidade Espectral (Espectro Populacional)* do processo $\{Y_t\}_{t=-\infty}^{\infty}$, definida em $[0, \pi]$, chamada de *Função de Distribuição Espectral*, e denotada por $F_Y(\omega)$, é a representação espectral da função de autocovariância. Isso pode ser observado por meio da seguinte relação:

$$f_Y(\omega) = \frac{dF_Y(\omega)}{d\omega} \Rightarrow F_Y(\omega) = \int_0^{\pi} f_Y(\omega) d\omega \quad (\text{A5})$$

dado que para processos discretos, mensurados em intervalos unitários de tempo, a frequência mais alta possível é π , e, portanto, ω está definido no intervalo $[0, \pi]$.

Para se chegar a essa função, deve-se aplicar um conceito matemático conhecido como *Transformada de Fourier*⁶⁶ à função $f_Y(\omega)$ dada na forma

⁶⁶A *Transformada de Fourier* é uma "ferramenta matemática" que indica que uma função $h(t)$ de uma variável real t terá uma *função transformada* definida como: $H(\omega) = \int_{-\infty}^{\infty} h(t) \cdot e^{-i\omega t} dt$, desde que a integral exista para cada ω real. Além disso, tem-se, pela inversão da fórmula, que existe a função

apresentada na equação (A5), o que implica que se pode calcular a j -ésima autocovariância γ_j do processo $\{Y_t\}_{t=-\infty}^{\infty}$, por meio da equação:

$$\gamma_j = \int_{-\pi}^{\pi} f_Y(\omega) \cdot e^{i\omega j} d\omega \quad (A6)$$

Para o caso em que $j = 0$, tem-se que a variância (σ_Y^2) do processo $\{Y_t\}_{t=-\infty}^{\infty}$ será:

$$\sigma_Y^2 = \gamma_0 = \int_{-\pi}^{\pi} f_Y(\omega) d\omega \quad (A7)$$

A equação (A7) implica que a área abaixo do espectro populacional entre as frequências $[-\pi, \pi]$ será a variância da série temporal analisada. A equação (A6) é equivalente a:

$$\gamma_j = \int_{-\pi}^{\pi} f_Y(\omega) \cos(\omega j) d\omega \quad (A8)$$

ou mesmo, dado que $f_Y(\omega)$ é simétrica, a:

$$\gamma_j = 2 \cdot \int_0^{\pi} f_Y(\omega) \cos(\omega j) d\omega \quad (A9)$$

Uma importante relação que se pode inferir das equações anteriores se refere ao fato de que, como a função $f_Y(\omega)$ é não-negativa, se for calculado $\int_{-\omega_1}^{\omega_1} f_Y(\omega) d\omega$, para qualquer valor ω_1 no intervalo $[0, \pi]$, o resultado será um

inversa, $h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) \cdot e^{i\omega t} d\omega$. Conforme aponta CHATFIELD (1996), na análise de séries temporais, utiliza-se a *Transformada de Fourier na Forma Discreta*, quando $h(t)$ é definida apenas para valores inteiros de t . Assim, dada a função: $h(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\omega) \cdot e^{i\omega t} d\omega$, a sua transformada discreta de Fourier será: $H(\omega) = \sum_{t=-\infty}^{\infty} h(t) \cdot e^{-i\omega t}$; $-\pi \leq \omega \leq \pi$.

número positivo, que pode ser interpretado como a parcela da variância de Y_t que é associada com as frequências ω que são menores que ω_l em termos absolutos.

Prova 4

Os estimadores de MQO serão obtidos reescrevendo-se o sistema na forma:

$$Y = \Omega \cdot \Theta + \varepsilon \tag{A10}$$

em que $Y = [y_1, y_2, y_3, \dots, y_N]'$; $\Theta = [\alpha_0, \alpha_1, \beta_1, \alpha_2, \beta_2, \dots, \alpha_{n-1}, \beta_{n-1}, \alpha_n]'$ e,

$$\Omega = \begin{bmatrix} 1 & \cos\omega_1 & \text{sen}\omega_1 & \cos\omega_2 & \text{sen}\omega_2 & \dots & \cos\omega_{n-1} & \text{sen}\omega_{n-1} & 1 \\ 1 & \cos\omega_1 2 & \text{sen}\omega_1 2 & \cos\omega_2 2 & \text{sen}\omega_2 2 & \dots & \cos\omega_{n-1} 2 & \text{sen}\omega_{n-1} 2 & -1 \\ 1 & \cos\omega_1 3 & \text{sen}\omega_1 3 & \cos\omega_2 3 & \text{sen}\omega_2 3 & \dots & \cos\omega_{n-1} 3 & \text{sen}\omega_{n-1} 3 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ 1 & \cos\omega_1(N-1) & \text{sen}\omega_1(N-1) & \cos\omega_2(N-1) & \text{sen}\omega_2(N-1) & \dots & \cos\omega_{n-1}(N-1) & \text{sen}\omega_{n-1}(N-1) & -1 \\ 1 & 1 & 0 & 1 & 0 & \dots & (-1)^{n-1} & 0 & 1 \end{bmatrix}$$

sendo que $n = \frac{N}{2}$.

Assim, tem-se que os estimadores de M.Q.O. de Θ , que minimizam as equações normais,

$$\sum_{t=1}^N (Y_t - \alpha_0 - \alpha_1 \cdot \cos \omega_1 t - \beta_1 \text{sen} \omega_1 t - \alpha_2 \cdot \cos \omega_2 t - \beta_2 \text{sen} \omega_2 t - \dots - \alpha_{n-1} \cos \omega_{n-1} t - \beta_{n-1} \text{sen} \omega_{n-1} t - \alpha_n)^2$$

serão dados pela fórmula usual:

$$\hat{\Theta} = (\Omega' \Omega)^{-1} \Omega' Y \tag{A11}$$

e representados por:

$$\hat{\Theta} = [\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\alpha}_3, \hat{\beta}_3, \dots, \hat{\alpha}_{n-1}, \hat{\beta}_{n-1}, \hat{\alpha}_n].$$

Prova 5

Os estimadores de MMV podem ser obtidos partindo-se de uma equação de regressão com resíduos apresentando um comportamento dado no modelo ARCH, do tipo:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \dots + \beta_k X_{kt} + e_t \tag{A12}$$

que pode ser reescrita como:

$$Y_t = X_t' \beta + e_t \tag{A13}$$

em que X_t representa um vetor de variáveis explicativas predeterminadas, que podem incluir termos defasados de Y . Tem-se como pressuposto que o termo de erro e_t possua as seguintes características:

$$e_t = \sqrt{\sigma_t^2} \cdot v_t \tag{A14}$$

em que $\{v_t\}$ é uma seqüência distribuída estatisticamente de maneira independente e identicamente (i.i.d.) com média zero e variância unitária, ou seja, $\{v_t\}$ apresenta as seguintes características:

$$E(v_t) = 0; \quad E(v_t^2) = 1 \tag{A15}$$

e, além disso:

$$\sigma_t^2 = \omega + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \dots + \alpha_m e_{t-m}^2 \tag{A16}$$

A representação alternativa do termo de erro dada pela equação (A14) para um processo ARCH(m), segundo apontam HAMILTON (1994) e ENDERS (1995), apresenta algumas vantagens estatísticas, pois impõe alguns pressupostos acerca da representação da dependência serial de e_t , sendo por isso utilizada na maioria das especificações de modelos do tipo ARCH⁶⁷. Assim, partindo-se de um vetor Ψ_t de observações obtidas até uma data t , dado por $\Psi_t = (y_t, y_{t-1}, \dots, y_1, y_0, \dots, y_{-m+1}, X_t', X_{t-1}', \dots, X_1', X_0', \dots, X_{-m+1}')$, se $\{v_t\}$ atende às características apontadas em (A16) e é independente de X_t e de Ψ_{t-1} , tem-se que a *distribuição condicional* de Y_t será *normal*, com média $X_t' \beta$ e variância σ_t^2 , e será dada por:

$$f(Y_t | X_t, \Psi_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp \left[\frac{-(Y_t - X_t' \beta)^2}{2\sigma_t^2} \right] \quad (A17)$$

em que

$$\sigma_t^2 = \omega + \alpha_1 (y_{t-1} - X_{t-1}' \beta)^2 + \alpha_2 (y_{t-2} - X_{t-2}' \beta)^2 + \dots + \alpha_m (y_{t-m} - X_{t-m}' \beta)^2 \quad (A18)$$

que pode ser reescrita de forma resumida como:

$$\sigma_t^2 = [\varphi(\beta)]' \Omega \quad (A19)$$

em que

$$\Omega \equiv (\omega, \alpha_1, \alpha_2, \alpha_3, \dots, \alpha_m)'$$

e

$$[\varphi(\beta)]' \equiv [1, (y_{t-1} - X_{t-1}' \beta)^2, (y_{t-2} - X_{t-2}' \beta)^2, \dots, (y_{t-m} - X_{t-m}' \beta)^2]$$

⁶⁷ Para maiores detalhes, ver HAMILTON (1994:659) e ENDERS (1995:142).

Agrupando-se em um vetor os parâmetros desconhecidos a ser estimados: $\Theta = (\beta', \Omega')$, tem-se que a Função de Verossimilhança Amostral, na forma logarítmica, a ser maximizada será então:

$$L(\Theta) = \sum_{t=1}^T \log f(Y_t | X_t, \Psi_{t-1}; \Theta) \quad (\text{A20})$$

que é igual a:

$$L(\Theta) = -\left(\frac{T}{2}\right) \log(2\pi) - \left(\frac{1}{2}\right) \sum_{t=1}^T \log(\sigma_t^2) - \left(\frac{1}{2}\right) \sum_{t=1}^T \left[\frac{(Y_t - X_t' \beta)^2}{\sigma_t^2} \right] \quad (\text{A21})$$

Conforme demonstra HAMILTON (1994), a derivada da Função de Verossimilhança, na forma logarítmica, com respeito ao vetor de parâmetros Θ , será:

$$S_t(\Theta)_{(a \times 1)} = \frac{\partial \log f(Y_t | X_t, \Psi_{t-1}; \Theta)}{\partial \Theta} \quad (\text{A22})$$

que será igual a:

$$S_t(\Theta)_{(a \times 1)} = \left\{ \frac{e_t^2 - \sigma_t^2}{2(\sigma_t^2)^2} \right\} \begin{bmatrix} \sum_{j=1}^m -2\alpha_j e_{t-j} X_{t-j} \\ \varphi_t(\beta) \end{bmatrix} + \begin{bmatrix} \frac{X_t e_t}{\sigma_t^2} \\ 0 \end{bmatrix} \quad (\text{A23})$$

o que irá fornecer as estimativas de máxima verossimilhança para os parâmetros desconhecidos: $\Theta = (\beta', \Omega')$.

APÊNDICE B

TESTE DE HETEROCEDASTICIDADE DE WHITE

White Heteroskedasticity Test:

F-statistic	19.89351	Probability	0.000000
Obs*R-squared	87.11223	Probability	0.000000

Test Equation:

Dependent Variable: RESID²

Method: Least Squares

Date: 05/04/01 Time: 06:32

Sample: 1946:02 2000:12

Included observations: 659

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1453.439	1157.835	1.255307	0.2098
@TREND	-7.586519	5.003984	-1.516096	0.1300
(@TREND) ²	0.007068	0.005851	1.208064	0.2275
(@TREND)*PRCAFE(-1)	0.020785	0.008877	2.341282	0.0195
PRCAFE(-1)	-8.256113	4.450618	-1.855049	0.0640
PRCAFE(-1) ²	0.013903	0.004070	3.415874	0.0007
R-squared	0.132189	Mean dependent var	594.0238	
Adjusted R-squared	0.125544	S.D. dependent var	4108.176	
S.E. of regression	3841.652	Akaike info criterion	19.35426	
Sum squared resid	9.64E+09	Schwarz criterion	19.39514	
Log likelihood	-6371.227	F-statistic	19.89351	
Durbin-Watson stat	2.236464	Prob(F-statistic)	0.000000	

APÊNDICE C

CÁLCULO DO NÚMERO DE DEFASAGENS DE TRUNCAMENTO DO TESTE DE PHILLIPS-PERRON

A fórmula proposta por Newey-West em HALL (1995) para o cálculo de q é dada por:

$$q = \text{floor} \left(4 \left(\frac{T}{100} \right)^{\frac{2}{9}} \right)$$

em que T = número de observações; e *floor* = indica um arredondamento do número encontrado.

Como o número de observações dos preços do café foi de $T = 660$, tem-se que:

$$q = \text{floor} \left(4 \left(\frac{660}{100} \right)^{\frac{2}{9}} \right) = 6,0839 \cong 6$$

APÊNDICE D

RESULTADOS DOS MODELOS DE TENDÊNCIA DETERMINÍSTICA

Equação 15

Dependent Variable: PRCAFE

Method: Least Squares

Date: 05/07/01 Time: 05:20

Sample: 1946:01 2000:12

Included observations: 660

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	328.0305	10.09195	32.50419	0.0000
@TREND	-0.231361	0.026515	-8.725805	0.0000
R-squared	0.103713	Mean dependent var		251.7969
Adjusted R-squared	0.102351	S.D. dependent var		136.9799
S.E. of regression	129.7807	Akaike info criterion		12.57260
Sum squared resid	11082721	Schwarz criterion		12.58621
Log likelihood	-4146.957	F-statistic		76.13967
Durbin-Watson stat	0.035730	Prob(F-statistic)		0.000000

Equação 20

Dependent Variable: LOG(PRCAFE)

Method: Least Squares

Date: 05/07/01 Time: 05:26

Sample: 1946:01 2000:12

Included observations: 660

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.828014	0.035287	165.1607	0.0000
@TREND	-0.001304	9.27E-05	-14.06205	0.0000
R-squared	0.231076	Mean dependent var		5.398449
Adjusted R-squared	0.229907	S.D. dependent var		0.517104
S.E. of regression	0.453784	Akaike info criterion		1.260634
Sum squared resid	135.4952	Schwarz criterion		1.274247
Log likelihood	-414.0093	F-statistic		197.7413
Durbin-Watson stat	0.028857	Prob(F-statistic)		0.000000

Modelos Polinomiais

Equação 18

Dependent Variable: PRCAFE

Method: Least Squares

Date: 05/07/01 Time: 05:28

Sample: 1946:01 2000:12

Included observations: 660

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	230.6373	14.23279	16.20464	0.0000
@TREND	0.656723	0.099764	6.582741	0.0000
@TREND*@TREND	-0.001348	0.000147	-9.194630	0.0000
R-squared	0.205896	Mean dependent var		251.7969
Adjusted R-squared	0.203479	S.D. dependent var		136.9799
S.E. of regression	122.2519	Akaike info criterion		12.45458
Sum squared resid	9819208.	Schwarz criterion		12.47500
Log likelihood	-4107.011	F-statistic		85.17388
Durbin-Watson stat	0.040288	Prob(F-statistic)		0.000000

Modelo de Tendência Cúbica

Dependent Variable: PRCAFE

Method: Least Squares

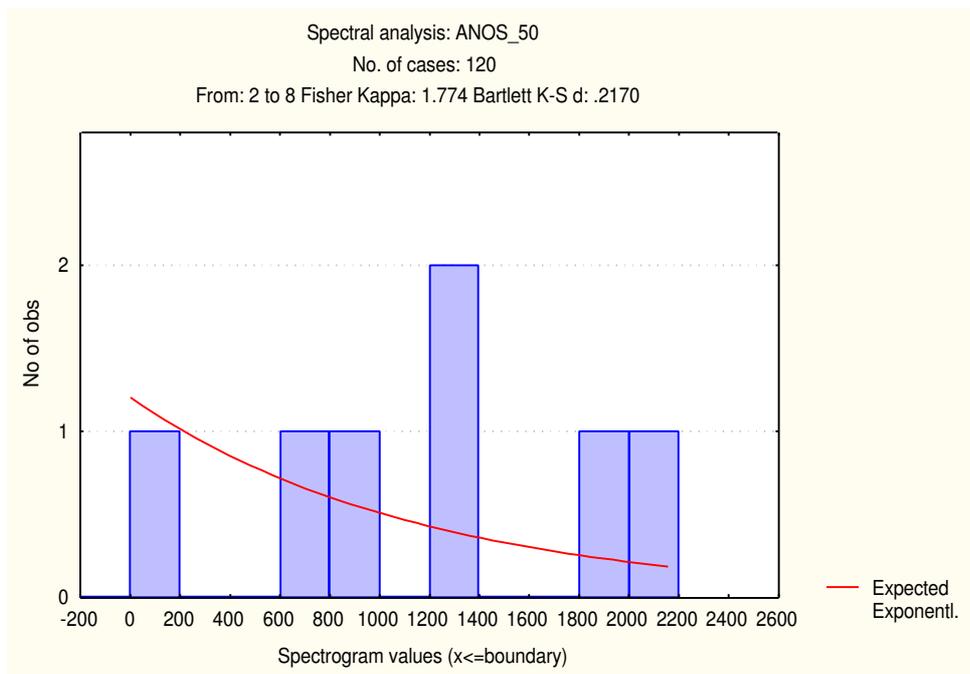
Date: 05/07/01 Time: 05:31

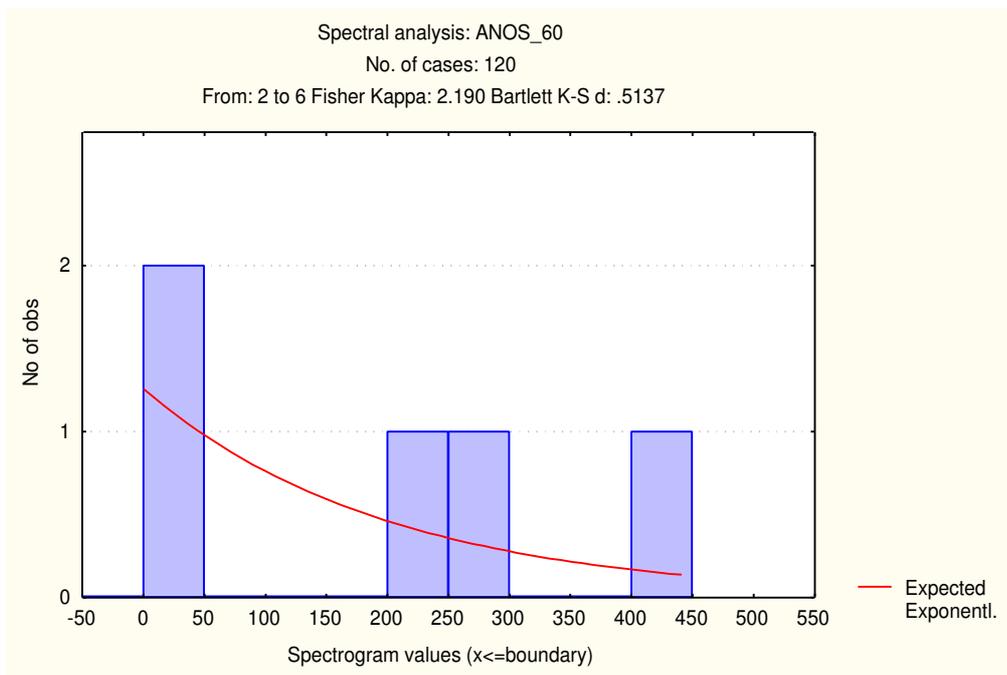
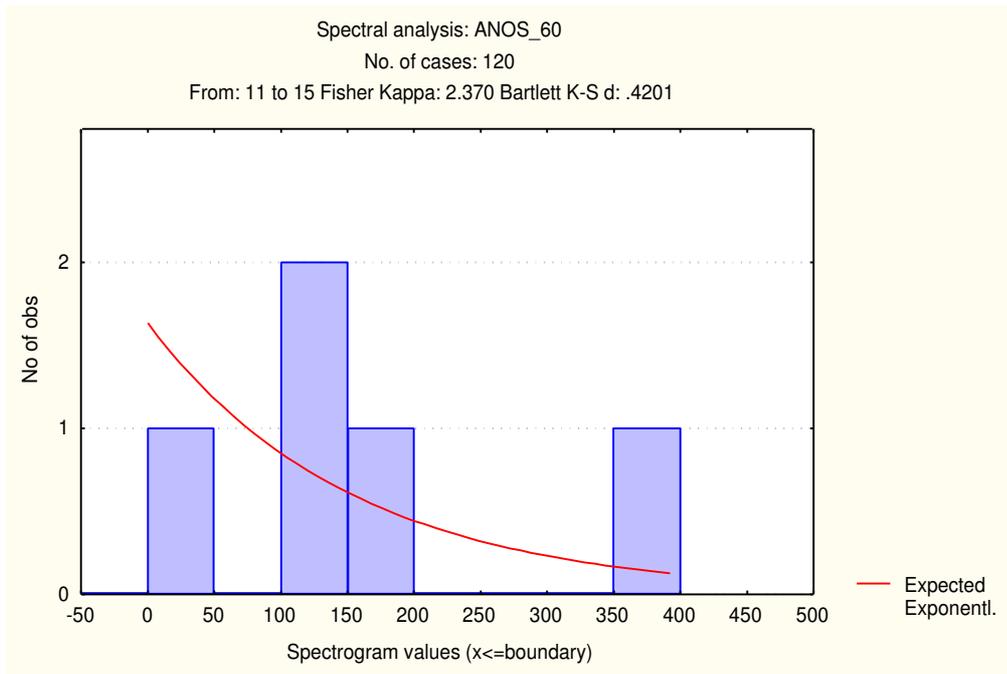
Sample: 1946:01 2000:12

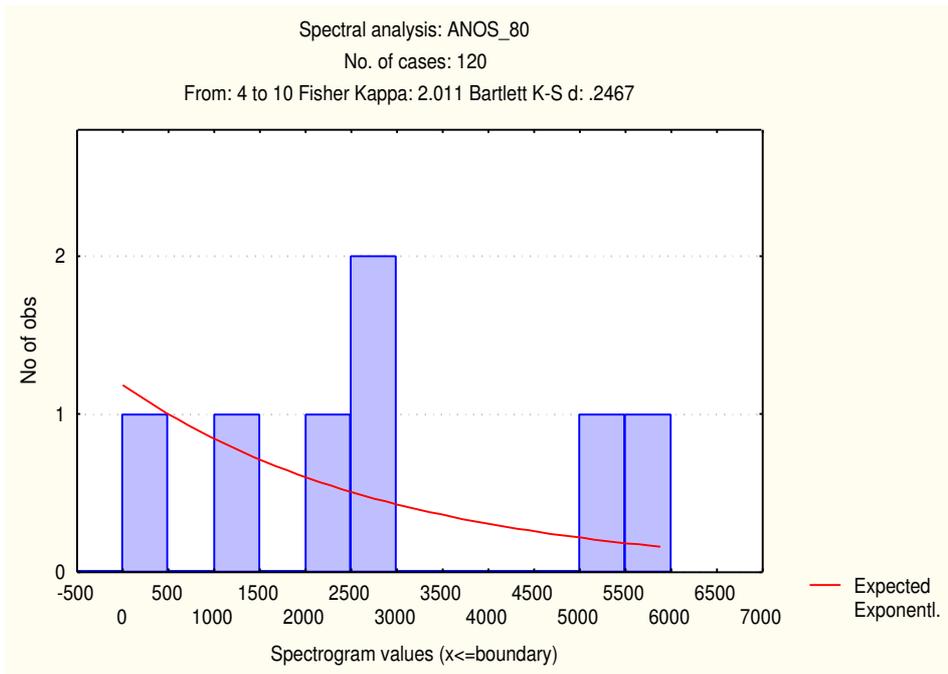
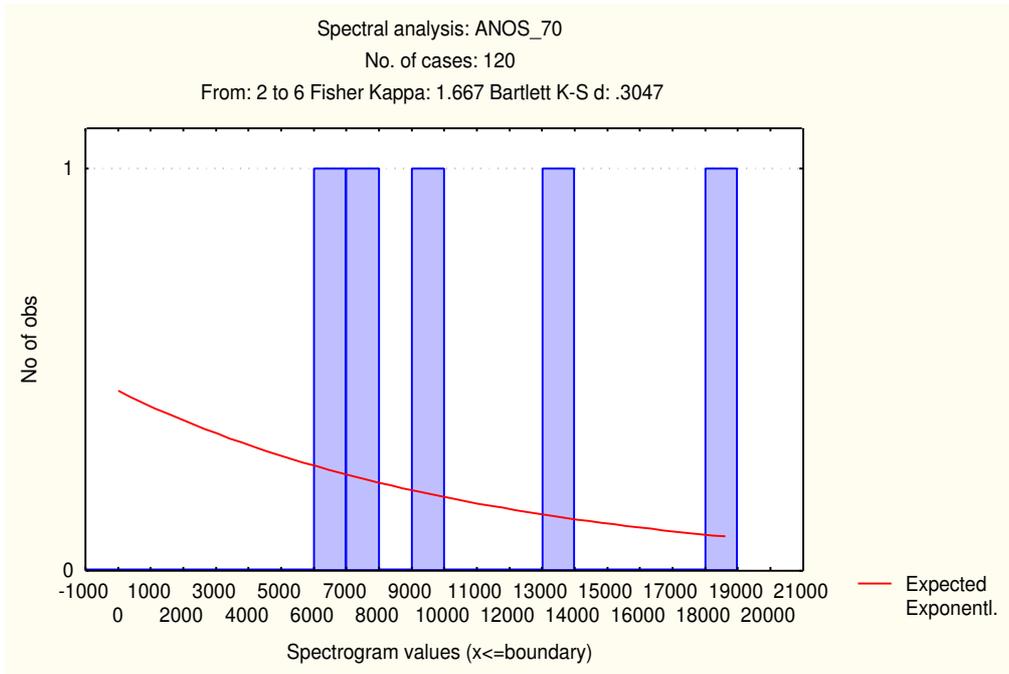
Included observations: 660

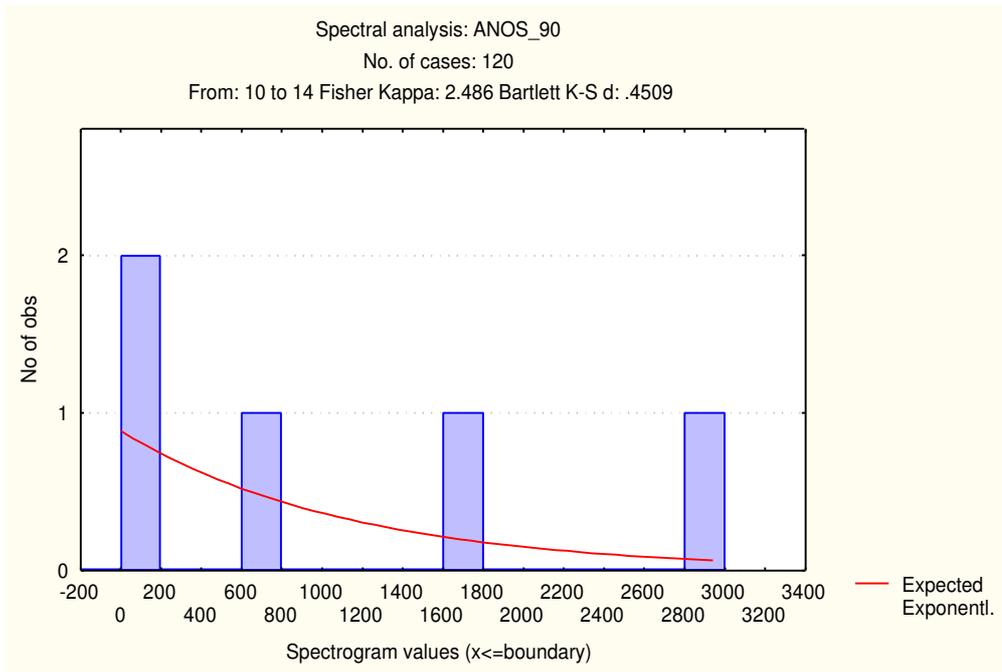
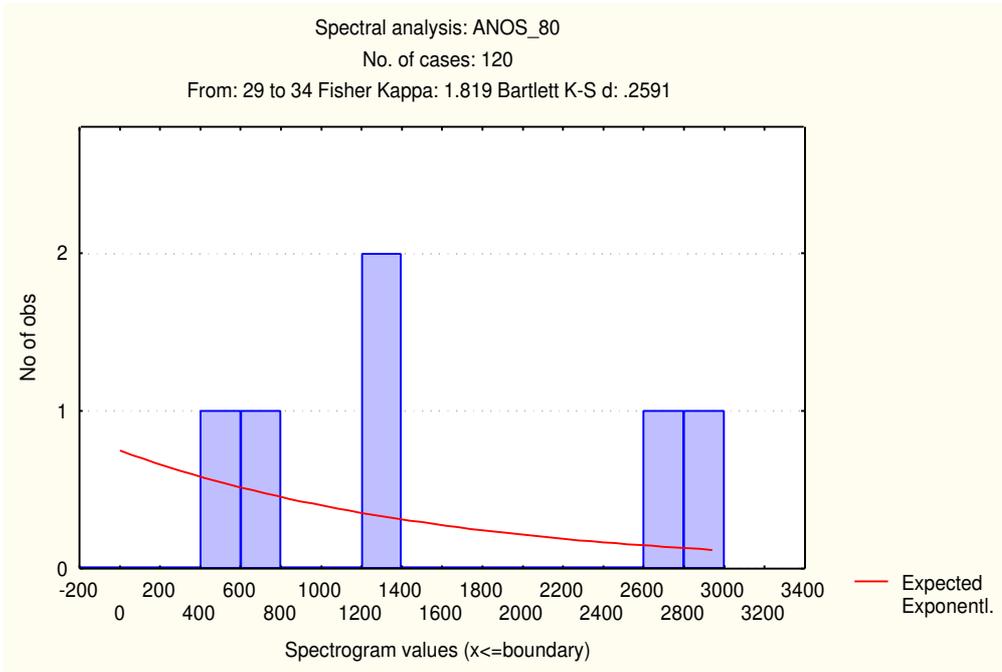
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	277.8819	18.73353	14.83340	0.0000
@TREND	-0.206849	0.246374	-0.839576	0.4015
@TREND*@TREND	0.001931	0.000869	2.221663	0.0266
@TREND*@TREND *@TREND	-3.32E-06	8.67E-07	-3.825861	0.0001
R-squared	0.223228	Mean dependent var	251.7969	
Adjusted R-squared	0.219676	S.D. dependent var	136.9799	
S.E. of regression	121.0025	Akaike info criterion	12.43554	
Sum squared resid	9604896.	Schwarz criterion	12.46277	
Log likelihood	-4099.729	F-statistic	62.84028	
Durbin-Watson stat	0.041212	Prob(F-statistic)	0.000000	

APÊNDICE E

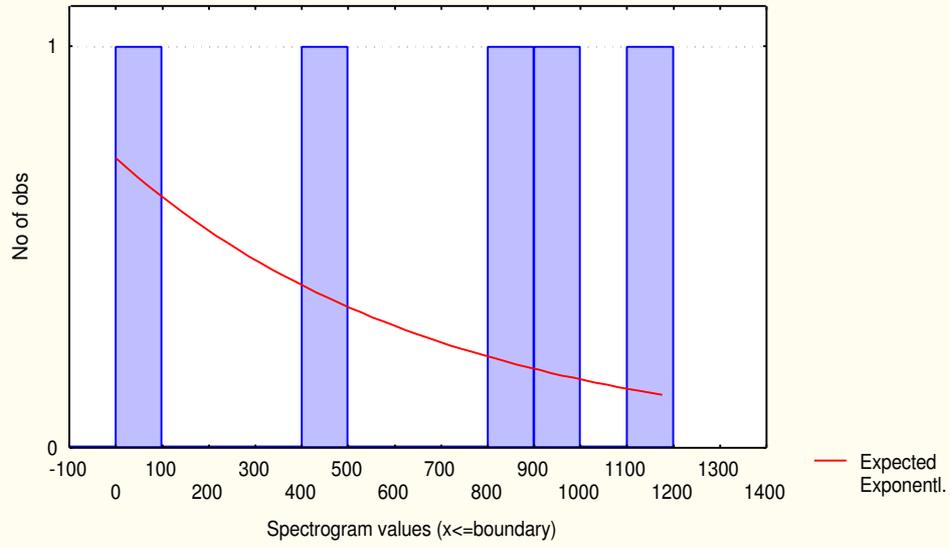








Spectral analysis: ANOS_90
No. of cases: 120
From: 3 to 7 Fisher Kappa: 1.593 Bartlett K-S d: .3854



APÊNDICE F

Dependent Variable: DPRCAFE

Method: Least Squares

Date: 05/23/01 Time: 04:55

Sample: 1946:01 2000:12

Included observations: 660

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.985851	3.308588	1.809186	0.0709
DM2	-6.507889	4.679050	-1.390857	0.1647
DM3	-2.445260	4.679050	-0.522598	0.6014
DM4	-6.408717	4.679050	-1.369662	0.1713
DM5	-8.557639	4.679050	-1.828927	0.0679
DM6	-8.531343	4.679050	-1.823307	0.0687
DM7	-9.719696	4.679050	-2.077280	0.0382
DM8	-6.784734	4.679050	-1.450024	0.1475
DM9	-5.153768	4.679050	-1.101456	0.2711
DM10	-5.913983	4.679050	-1.263928	0.2067
DM11	-1.894665	4.679050	-0.404925	0.6857
DM12	-8.283075	4.679050	-1.770247	0.0772
R-squared	0.013762	Mean dependent var		0.135787
Adjusted R-squared	-0.002980	S.D. dependent var		24.50067
S.E. of regression	24.53715	Akaike info criterion		9.256268
Sum squared resid	390142.3	Schwarz criterion		9.337945
Log likelihood	-3042.568	F-statistic		0.822014
Durbin-Watson stat	1.577794	Prob(F-statistic)		0.618031

APÊNDICE G

Modelo ARIMA (0,1,12)

Dependent Variable: DPRCAFE

Method: Least Squares

Date: 06/09/01 Time: 06:30

Sample: 1946:01 2000:12

Included observations: 660

Convergence achieved after 7 iterations

Backcast: 1945:01 1945:12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	0.205526	0.037911	5.421283	0.0000
MA(2)	0.116066	0.038345	3.026894	0.0026
MA(3)	0.113545	0.037739	3.008731	0.0027
MA(6)	-0.074961	0.037155	-2.017555	0.0440
MA(7)	0.066633	0.038210	1.743866	0.0817
MA(8)	0.107063	0.038696	2.766788	0.0058
MA(9)	-0.153362	0.038633	-3.969752	0.0001
MA(10)	0.067461	0.038072	1.771906	0.0769
MA(12)	-0.205752	0.037260	-5.522119	0.0000
R-squared	0.141499	Mean dependent var		0.135787
Adjusted R-squared	0.130949	S.D. dependent var		24.50067
S.E. of regression	22.84023	Akaike info criterion		9.108467
Sum squared resid	339611.2	Schwarz criterion		9.169725
Log likelihood	-2996.794	F-statistic		13.41232
Durbin-Watson stat	1.988277	Prob(F-statistic)		0.000000
Inverted MA Roots	.84	.73+.45i	.73 -.45i	.40 -.77i
	.40+.77i	.05+.88i	.05 -.88i	-.48+.81i
	-.48 -.81i	-.80 -.36i	-.80+.36i	-.83

Modelo SARIMA (9,1,10)x(0,0,1)

Dependent Variable: DPRCAFE

Method: Least Squares

Date: 06/07/01 Time: 08:40

Sample(adjusted): 1946:10 2000:12

Included observations: 651 after adjusting endpoints

Convergence achieved after 6 iterations

Backcast: 1944:12 1946:09

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.244456	0.039147	6.244517	0.0000
AR(7)	0.110631	0.037881	2.920463	0.0036
AR(9)	-0.211293	0.038117	-5.543239	0.0000
AR(2)	0.069219	0.038367	1.804136	0.0717
MA(6)	-0.081601	0.038836	-2.101148	0.0360
MA(8)	0.073327	0.040176	1.825162	0.0684
MA(10)	0.155942	0.040844	3.817981	0.0001
SMA(12)	-0.177190	0.039329	-4.505380	0.0000
R-squared	0.150606	Mean dependent var		0.068118
Adjusted R-squared	0.141359	S.D. dependent var		24.35589
S.E. of regression	22.56887	Akaike info criterion		9.083233
Sum squared resid	327514.5	Schwarz criterion		9.138268
Log likelihood	-2948.592	F-statistic		16.28715
Durbin-Watson stat	2.022774	Prob(F-statistic)		0.000000
Inverted AR Roots	.81 -.25i -.14 -.85i -.79	.81+.25i -.14+.85i	.49 -.72i -.65+.51i	.49+.72i -.65 -.51i
Inverted MA Roots	.87 .75 -.43i .43 -.75i -.00+.87i -.46 -.70i -.80 -.26i	.80 -.26i .46+.70i .00+.78i -.43+.75i -.75+.43i -.87	.80+.26i .46 -.70i .00 -.78i -.43 -.75i -.75 -.43i	.75+.43i .43+.75i .00 -.87i -.46+.70i -.80+.26i

Modelo SARIMA (7,1,9)x(0,0,1)

Dependent Variable: DPRCAFE

Method: Least Squares

Date: 06/08/01 Time: 23:06

Sample(adjusted): 1946:08 2000:12

Included observations: 653 after adjusting endpoints

Convergence achieved after 6 iterations

Backcast: 1944:11 1946:07

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(2)	0.124683	0.039977	3.118900	0.0019
AR(3)	0.067763	0.038730	1.749631	0.0807
AR(7)	0.066266	0.040287	1.644848	0.1005
MA(1)	0.249349	0.038349	6.502078	0.0000
MA(6)	-0.097578	0.038125	-2.559435	0.0107
MA(8)	0.090340	0.038104	2.370883	0.0180
MA(9)	-0.194762	0.038517	-5.056521	0.0000
SMA(12)	-0.188594	0.038893	-4.849007	0.0000
R-squared	0.148876	Mean dependent var		0.044548
Adjusted R-squared	0.139639	S.D. dependent var		24.49107
S.E. of regression	22.71685	Akaike info criterion		9.096267
Sum squared resid	332855.8	Schwarz criterion		9.151171
Log likelihood	-2961.931	F-statistic		16.11735
Durbin-Watson stat	2.030073	Prob(F-statistic)		0.000000
Inverted AR Roots	.73	.43+.49i	.43 -.49i	-.17+.64i
	-.17 -.64i	-.62 -.30i	-.62+.30i	
Inverted MA Roots	.87	.80	.75+.44i	.75 -.44i
	.58+.50i	.58 -.50i	.44 -.75i	.44+.75i
	.18 -.80i	.18+.80i	.00+.87i	-.00 -.87i
	-.44 -.75i	-.44+.75i	-.45 -.77i	-.45+.77i
	-.75+.44i	-.75 -.44i	-.84 -.28i	-.84+.28i
	-.87			

APÊNDICE H

GARCH (1,1)

Dependent Variable: DPRCAFE

Method: ML - ARCH

Date: 06/14/01 Time: 09:42

Sample(adjusted): 1946:08 2000:12

Included observations: 653 after adjusting endpoints

Convergence achieved after 60 iterations

Bollerslev-Wooldrige robust standard errors & covariance

Backcast: 1944:11 1946:07

	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-0.057230	0.059383	-0.963729	0.3352
AR(3)	0.065497	0.046080	1.421374	0.1552
AR(7)	0.062449	0.045055	1.386050	0.1657
MA(1)	0.167069	0.061333	2.723977	0.0065
MA(6)	-0.016459	0.052667	-0.312500	0.7547
MA(8)	0.149555	0.094827	1.577128	0.1148
MA(9)	-0.001501	0.027820	-0.053959	0.9570
SMA(12)	-0.007428	0.027965	-0.265630	0.7905

Variance Equation				
C	5.877190	2.720886	2.160028	0.0308
ARCH(1)	0.860568	0.194844	4.416705	0.0000
GARCH(1)	0.503240	0.055459	9.074131	0.0000
R-squared	0.061778	Mean dependent var		0.044548
Adjusted R-squared	0.047164	S.D. dependent var		24.49107
S.E. of regression	23.90656	Akaike info criterion		7.925554
Sum squared resid	366918.0	Schwarz criterion		8.001048
Log likelihood	-2576.694	F-statistic		4.227274
Durbin-Watson stat	1.852913	Prob(F-statistic)		0.000010
Inverted AR Roots	.68	.41 -.52i	.41+.52i	-.17 -.68i
	-.17+.68i	-.58+.31i	-.58 -.31i	
Inverted MA Roots	.71+.30i	.71 -.30i	.66	.58+.33i
	.58 -.33i	.33 -.58i	.33+.58i	.29 -.73i
	.29+.73i	.01	.00+.66i	-.00 -.66i
	-.33 -.73i	-.33+.73i	-.33 -.58i	-.33+.58i
	-.58+.33i	-.58 -.33i	-.66	-.75+.29i
	-.75 -.29i			

GARCH (2,1)

Dependent Variable: DPRCAFE

Method: ML - ARCH

Date: 06/14/01 Time: 10:00

Sample(adjusted): 1946:08 2000:12

Included observations: 653 after adjusting endpoints

Failure to improve Likelihood after 13 iterations

Bollerslev-Wooldrige robust standard errors & covariance

Backcast: 1944:11 1946:07

	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-0.092553	0.059104	-1.565940	0.1174
AR(3)	0.047626	0.050508	0.942931	0.3457
AR(7)	0.016591	0.037952	0.437147	0.6620
MA(1)	0.069866	0.076777	0.909994	0.3628
MA(6)	-0.003922	0.049312	-0.079537	0.9366
MA(8)	0.100416	0.074474	1.348332	0.1776
MA(9)	-0.025961	0.039577	-0.655943	0.5119
SMA(12)	-0.016452	0.033179	-0.495837	0.6200
Variance Equation				
C	3.555913	2.100334	1.693023	0.0905
ARCH(1)	1.140506	0.248289	4.593470	0.0000
ARCH(2)	-0.726293	0.206394	-3.518962	0.0004
GARCH(1)	0.736863	0.072444	10.17152	0.0000
R-squared	0.030631	Mean dependent var		0.044548
Adjusted R-squared	0.013996	S.D. dependent var		24.49107
S.E. of regression	24.31908	Akaike info criterion		7.907143
Sum squared resid	379098.6	Schwarz criterion		7.989500
Log likelihood	-2569.682	F-statistic		1.841385
Durbin-Watson stat	1.642649	Prob(F-statistic)		0.044240
Inverted AR Roots	.56	.33 -.43i	.33+.43i	-.14 -.58i
	-.14+.58i	-.47+.26i	-.47 -.26i	
Inverted MA Roots	.71	.64+.29i	.64 -.29i	.62+.36i
	.62 -.36i	.36+.62i	.36 -.62i	.26
	.25+.70i	.25 -.70i	.00 -.71i	-.00+.71i
	-.33 -.70i	-.33+.70i	-.36+.62i	-.36 -.62i
	-.62+.36i	-.62 -.36i	-.71	-.73+.29i
	-.73 -.29i			

GARCH (2,2)

Dependent Variable: DPRCAFE

Method: ML - ARCH

Date: 06/14/01 Time: 09:59

Sample(adjusted): 1946:08 2000:12

Included observations: 653 after adjusting endpoints

Failure to improve Likelihood after 27 iterations

Bollerslev-Wooldrige robust standard errors & covariance

Backcast: 1944:11 1946:07

	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-0.028968	0.057030	-0.507935	0.6115
AR(3)	0.085763	0.050278	1.705791	0.0880
AR(7)	0.069889	0.055262	1.264668	0.2060
MA(1)	0.216537	0.068634	3.154936	0.0016
MA(6)	0.034328	0.030653	1.119892	0.2628
MA(8)	0.022946	0.017082	1.343269	0.1792
MA(9)	-0.038000	0.033424	-1.136929	0.2556
SMA(12)	-0.005636	0.028364	-0.198694	0.8425
Variance Equation				
C	9.554013	5.579475	1.712350	0.0868
ARCH(1)	0.826245	0.241070	3.427402	0.0006
ARCH(2)	0.422124	0.244385	1.727293	0.0841
GARCH(1)	-0.183682	0.118289	-1.552824	0.1205
GARCH(2)	0.407911	0.125767	3.243394	0.0012
R-squared	0.067365	Mean dependent var		0.044548
Adjusted R-squared	0.049878	S.D. dependent var		24.49107
S.E. of regression	23.87248	Akaike info criterion		7.901137
Sum squared resid	364733.1	Schwarz criterion		7.990357
Log likelihood	-2566.721	F-statistic		3.852287
Durbin-Watson stat	1.999525	Prob(F-statistic)		0.000010
Inverted AR Roots	.70	.42+.51i	.42 -.51i	-.18+.68i
	-.18 -.68i	-.59 -.32i	-.59+.32i	
Inverted MA Roots	.65	.61	.56+.32i	.56 -.32i
	.53 -.41i	.53+.41i	.32+.56i	.32 -.56i
	.11 -.69i	.11+.69i	.00+.65i	-.00 -.65i
	-.32+.56i	-.32 -.56i	-.35 -.61i	-.35+.61i
	-.56 -.32i	-.56+.32i	-.65	-.71 -.27i
	-.71+.27i			

TARCH (2,1)

Dependent Variable: DPRCAFE

Method: ML - ARCH

Date: 06/14/01 Time: 10:00

Sample(adjusted): 1946:08 2000:12

Included observations: 653 after adjusting endpoints

Failure to improve Likelihood after 16 iterations

Bollerslev-Wooldrige robust standard errors & covariance

Backcast: 1944:11 1946:07

	Coefficient	Std. Error	z-Statistic	Prob.
AR(2)	-0.033281	0.046294	-0.718903	0.4722
AR(3)	0.075629	0.042599	1.775356	0.0758
AR(7)	0.007537	0.033646	0.224005	0.8228
MA(1)	0.137307	0.053690	2.557415	0.0105
MA(6)	0.003209	0.029439	0.109018	0.9132
MA(8)	0.014477	0.028979	0.499561	0.6174
MA(9)	-0.011301	0.027065	-0.417543	0.6763
SMA(12)	-0.010466	0.032418	-0.322833	0.7468
Variance Equation				
C	3.347405	1.480445	2.261081	0.0238
ARCH(1)	1.052932	0.251360	4.188938	0.0000
ARCH(2)	-0.418085	0.181381	-2.305007	0.0212
(RESID<0)*ARCH (1)	-0.470588	0.174633	-2.694719	0.0070
GARCH(1)	0.728771	0.055170	13.20956	0.0000
R-squared	0.054372	Mean dependent var		0.044548
Adjusted R-squared	0.036642	S.D. dependent var		24.49107
S.E. of regression	24.03819	Akaike info criterion		7.833786
Sum squared resid	369814.1	Schwarz criterion		7.923006
Log likelihood	-2544.731	F-statistic		3.066591
Durbin-Watson stat	1.835934	Prob(F-statistic)		0.000317
Inverted AR Roots	.53	.30 -.36i	.30+.36i	-.15 -.51i
	-.15+.51i	-.41+.25i	-.41 -.25i	
Inverted MA Roots	.68	.59 -.34i	.59+.34i	.52
	.46+.33i	.46 -.33i	.34+.59i	.34 -.59i
	.14 -.59i	.14+.59i	.00 -.68i	-.00+.68i
	-.30+.57i	-.30 -.57i	-.34 -.59i	-.34+.59i
	-.59+.34i	-.59 -.34i	-.62+.24i	-.62 -.24i
	-.68			