

**Testes HEGY de raízes unitárias sazonais:  
efeitos de observações atípicas,  
erros de medida e quebras estruturais**

**Karen Elisa do Vale Nogueira Penna**

Dissertação apresentada  
ao  
Instituto de Matemática e Estatística  
da  
Universidade de São Paulo  
para  
Obtenção do Título  
de  
Mestre em Ciências

Programa: **Estatística**

Orientadora: **Profa. Dra. Clélia Maria de Castro Tolo**

**São Paulo, Junho de 2009**

## **Testes HEGY de raízes unitárias sazonais: efeitos de observações atípicas, erros de medida e quebras estruturais**

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Karen Elisa do Vale Nogueira Penna e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dra. Clélia Maria de Castro Tolo (Orientadora) – IME-USP
- Prof. Dra. Airlane Pereira Alencar – IME-USP
- Prof. Dra. Sílvia Regina Costa Lopes – UFRGS

## Agradecimentos

Agradeço a **Deus** que me permitiu concluir mais essa etapa da minha vida.

Agradeço imensamente aos meus pais, **Maria Nadiége do Vale Nogueira** e **Fernando Nogueira** por todo amor, incentivo e dedicação. Mãe, Pai, a nossa família é o que há de melhor na minha vida e eu sempre vou estar com vocês.

Aos meus irmãos **Fernando Júnior, Sílvia** e **Ricardo** por tornarem a minha vida mais feliz. Vocês são os melhores irmãos que alguém pode ter. Como irmã caçula, só tenho a agradecer por todo o amor, união e, sobretudo, pela amizade que temos. Por mais que a vida nos traga distância e saudade, ainda que em pensamento, sempre estarei junto a vocês.

Ao meu marido **Gustavo** pela paciência, apoio, carinho e confiança de sempre, sem os quais certamente esse passo não seria possível. Gustavo, em você encontrei amor, equilíbrio, paz e felicidade. Obrigada por ter escolhido caminhar ao meu lado.

Agradeço a professora **Clélia** pelo apoio, compreensão e confiança. Professora, obrigada pelos ensinamentos e, sobretudo, pela amizade construída.

Pais, irmãos, professores e amigos, vocês são parte dessa conquista. Muito obrigada!

## Resumo

O presente trabalho trata-se de um estudo de testes de sazonalidade em séries temporais lineares. O texto é resultado de uma vasta pesquisa abrangendo diversos artigos e livros relacionados ao tema.

Inicialmente, um teste de raízes unitárias correspondente a frequências sazonais para dados trimestrais é apresentado. Esse teste é denominado procedimento HEGY e permite que raízes unitárias sejam testadas em algumas frequências sazonais isoladas, sem assumir que raízes unitárias estejam presentes em todas as frequências sazonais. Em seguida, a extensão desse teste para dados mensais é analisada e um teste mais poderoso que combina o procedimento trimestral e mensal é apresentado. Observações atípicas e erros de medida são estudados, a fim de avaliar os efeitos que causam no comportamento dos testes de raízes unitárias e de apresentar um teste que contemple a correção de atipicidades. Por fim, é analisada a robustez assintótica do procedimento HEGY na presença de quebras estruturais (mudança de magnitude finita na média sazonal) e um teste mais poderoso que incorpora a quebra estrutural é apresentado.

A fim de ilustrar e complementar as análises teóricas apresentadas, algumas aplicações em séries temporais reais são desenvolvidas. Os dados analisados foram obtidos do Instituto Brasileiro de Geografia e Estatística (IBGE) e o software utilizado para análise foi o R.

## Abstract

This current work is a testing study for seasonality in time linear series. The text is the result of an extensive search covering several articles and books related to this subject.

Initially, a test of unit roots corresponding to seasonal frequencies in quarterly data is presented. This test is called procedure HEGY and it allows the unit roots to be tested in some isolated seasonal frequencies, without assuming that the unit roots are present at all seasonal frequencies. Then, the extension of this test to monthly data is analyzed and a test that combines the most powerful procedure is presented quarterly and monthly. Atypical observations and errors of measurement are studied to evaluate the effects that they can cause on the behavior of the unit root tests and to present a test that includes the correction of this difference. Finally, we examined the asymptotic robustness procedure HEGY in the presence of structural breaks (finite magnitude of change in seasonal average) and a more powerful test that incorporates a structural break is introduced.

In order to illustrate and complement the presented theoretical analysis, some applications in real time series are developed. The data were obtained from the Instituto Brasileiro de Geografia e Estatística (IBGE) and the R software was used for the analysis.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Preliminares</b>	<b>3</b>
2.1	Processos Estocásticos	3
2.2	Estacionariedade	4
2.3	Modelos ARMA	5
2.3.1	Processos Autorregressivos	5
2.3.2	Processos de Médias-Móveis	6
2.3.3	Processos Autorregressivos e de Médias-Móveis	6
2.3.4	Função Densidade Espectral	7
2.4	Processos Sazonais	7
2.5	Teste de Dickey- Fuller	10
<b>3</b>	<b>Testes HEGY para Raízes Unitárias Sazonais</b>	<b>13</b>
3.1	O Procedimento HEGY Trimestral	13
3.2	Distribuição Assintótica das Estatísticas do Teste HEGY Trimestral	17
3.3	Valores Críticos para as Estatísticas do Teste HEGY Trimestral	18
3.4	O Procedimento HEGY Mensal	19
3.5	Distribuição Assintótica das Estatísticas do Teste HEGY Mensal	25
3.6	Valores Críticos para Amostras Finitas das Estatísticas do Teste HEGY Mensal	30
<b>4</b>	<b>Teste HEGY Mensal / Trimestral</b>	<b>33</b>
4.1	O Procedimento de Pons	33
4.2	Efeitos da Agregação Temporal	34
4.3	Estatísticas e correspondentes Valores Críticos	39

<b>5</b>	<b>Efeitos de Erros de Medida e Valores Atípicos nos Testes HEGY</b> .....	<b>48</b>
5.1	<i>Outliers</i> Aditivo e de Mudança Temporária .....	48
5.2	O Teste HEGY para Séries com Ruído .....	49
5.3	Distribuição Assintótica das Estatísticas do Teste HEGY com Ruído .....	50
5.4	Simulações de Monte Carlo .....	53
5.5	O Teste HEGY com Correção para <i>Outliers</i> .....	55
<b>6</b>	<b>Efeitos de Quebras Estruturais nos Testes HEGY</b> .....	<b>63</b>
6.1	Testes para Raízes Unitárias Sazonais sob Quebras Estruturais .....	63
6.2	Distribuição Assintótica das Estatísticas do Teste LM – HEGY – AO .....	67
6.3	Impactos da não consideração de Quebras Estruturais .....	70
6.4	Impactos da consideração errônea de Quebras Estruturais .....	73
6.5	Simulações de Monte Carlo .....	73
<b>7</b>	<b>Aplicações e Considerações Finais</b> .....	<b>80</b>
7.1	Análise Descritiva .....	80
7.2	Aplicações dos Testes HEGY .....	82
7.2.1	O Teste HEGY Mensal .....	82
7.2.2	O Teste HEGY Mensal / Trimestral .....	85
7.2.3	O Teste HEGY com Correção para <i>Outliers</i> .....	90
7.2.4	O Teste HEGY com Quebras Estruturais .....	93
7.3	Considerações Finais .....	94
	<b>Apêndice</b>	
	Códigos na linguagem R para aplicação dos testes HEGY .....	97
	<b>Referências Bibliográficas</b> .....	<b>103</b>

# Capítulo 1

## Introdução

A utilização de séries temporais com raízes unitárias tem tido participação expressiva no âmbito das análises de séries econômicas financeiras. Os primeiros testes de raízes unitárias foram propostos por Dickey e Fuller (1979, 1981) e, posteriormente, foram generalizados por Engle e Granger (1987). Todos esses trabalhos assumem que a raiz de interesse não somente tem módulo igual a um, mas vale exatamente um. Essa raiz corresponde a um pico no espectro na frequência zero, indicando a existência de uma tendência estocástica na série analisada.

Não obstante, sabemos que, na prática, muitas séries temporais também apresentam um comportamento sazonal, indicando a presença de raízes unitárias correspondentes a picos no espectro em frequências sazonais.

Tendências e sazonalidades são componentes de suma importância para se avaliar o comportamento de uma série temporal. Nesse sentido, faz-se necessário saber se o processo autorregressivo que gera uma série temporal tem raízes unitárias em valor absoluto na frequência zero ou em frequências sazonais. O presente trabalho explora o tema de sazonalidade e tem como objetivo apresentar procedimentos que determinem se há raízes unitárias sazonais em uma série temporal, além de determinar qual a classe de processo responsável pela sazonalidade.

O Capítulo 2 introduz alguns conceitos preliminares necessários ao desenvolvimento do presente trabalho, tais como modelos ARIMA, de grande utilização na modelagem de séries estacionárias e não estacionárias integradas e os testes de Dickey-Fuller, necessários à compreensão dos testes de raízes unitárias sazonais apresentados.

O Capítulo 3 apresenta o procedimento HEGY, cuja denominação é derivada das iniciais dos autores que o propuseram (Hilleberg, Engle, Granger e Yoo) e cuja aplicação está vinculada a séries temporais lineares que tem raízes com módulo um, mas que correspondem às frequências sazonais. O Capítulo 3 apresenta também a extensão do procedimento HEGY, desenvolvido originalmente para dados trimestrais, para dados mensais e as distribuições assintóticas das estatísticas de ambos os testes (trimestral e mensal) também são apresentadas. Ainda nesse capítulo, temos os valores críticos obtidos por simulações de Monte Carlo para os casos de dados trimestrais e mensais.

O Capítulo 4 contém um procedimento resultante da combinação dos testes de raízes unitárias sazonais para dados trimestrais e mensais. Essa nova técnica utiliza as raízes unitárias trimestrais rejeitadas para reduzir o conjunto de frequências mensais nas quais as

raízes unitárias deverão ser testadas e se apresenta com maior poder do que o teste de raízes unitárias sazonais para dados mensais. No Capítulo 4 apresentamos também a distribuição assintótica das estatísticas desse teste.

O Capítulo 5 estuda os efeitos da presença de observações atípicas e erros de medida no comportamento dos testes de raízes unitárias sazonais. Para tal, primeiramente apresenta-se um procedimento capaz de identificar a presença de atipicidades em uma série temporal e, na seqüência, apresentam-se sugestões de como minimizar os impactos causados nos testes de raízes unitárias sazonais. Com esse objetivo, um teste que contempla a correção das atipicidades é apresentado.

O Capítulo 6 avalia a robustez assintótica do procedimento HEGY na presença de quebras estruturais (mudança de magnitude finita na média sazonal). Nesse capítulo, alguns testes que consideram a presença de quebras estruturais são apresentados e, por intermédio de experimentos de Monte Carlo, esses testes são comparados a fim de se verificar qual deles melhor se aplica a séries com quebras estruturais.

Finalmente, o Capítulo 7 contém algumas aplicações dos testes apresentados. Foram utilizados dados reais obtidos do Instituto Brasileiro de Geografia e Estatística (IBGE) para ilustrar as análises teóricas apresentadas. Nesse Capítulo, há ainda as considerações finais acerca do tema estudado. Os programas desenvolvidos em linguagem R se encontram no Apêndice.

## Capítulo 2

### Preliminares

Neste Capítulo alguns conceitos preliminares necessários ao desenvolvimento dos testes HEGY serão apresentados, tais como modelos ARIMA, de grande utilização na modelagem de séries estacionárias e não estacionárias integradas e os testes de Dickey-Fuller.

#### 2.1 Processos Estocásticos

Um processo estocástico é uma família  $\{X(t), t \in T\}$ , tal que  $T$  é um conjunto arbitrário e para cada  $t \in T$ ,  $X(t)$  é uma variável aleatória. Assim, temos que um processo estocástico é uma família de variáveis aleatórias que supomos definidas em um mesmo espaço de probabilidades  $(\Omega, A, P)$ . Normalmente, supõe-se que as variáveis aleatórias envolvidas são reais, porém, elas podem ser complexas. O conjunto  $T$  é, em geral, um conjunto de números inteiros ou reais.

Como exemplo de um processo estocástico, vamos apresentar o movimento browniano padrão (ou processo de Wiener).

##### ➤ Movimento Browniano Padrão (MBP)

Trata-se do processo contínuo  $W = \{W(t), t \geq 0\}$ , cujo nome deve-se ao biólogo Robert Brown e tal que:

- (i)  $W(0) = 0$ ;
- (ii) para quaisquer instantes  $0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq 1$ , as variáveis aleatórias  $W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_k) - W(t_{k-1})$  são independentes;
- (iii) para quaisquer  $s, t$  e  $\tau$  não negativos, as variáveis aleatórias  $W(t) - W(s)$  e  $W(t + \tau) - W(s + \tau)$  tem a mesma distribuição;
- (iv) para todo  $t > 0$ ,  $W(t) \sim \mathcal{N}(0, t)$ ;
- (v) as trajetórias de  $W(t)$  são contínuas com probabilidade um.

De (ii) vem que o MBP tem incrementos independentes e de (iii) segue que tem incrementos estacionários. Como as distribuições finito-dimensionais de  $W$  são normais multivariadas, o processo  $W$  é gaussiano. Como  $W(t) - W(s)$  tem a mesma distribuição de  $W(t-s) - W(0) = W(t-s)$ , temos que  $W(t) - W(s)$  e  $W(t-s)$  tem distribuição  $\mathcal{N}(0, t-s)$ , em que  $s < t$ .

## 2.2 Estacionariedade

Uma das suposições mais frequentes feita na análise de séries temporais é a de que o processo estocástico gerador dos dados seja um processo estacionário, ou seja, oscila ao redor de uma média constante, com uma variância também constante, refletindo uma forma de equilíbrio.

Há dois tipos de estacionariedade – forte (estrita) e fraca (de segunda ordem). No presente trabalho, estamos interessados principalmente na classe de processos estacionários de segunda ordem, aos quais chamaremos simplesmente de processos estacionários.

As propriedades que caracterizam um processo estocástico  $(X(t))$  estacionário de segunda ordem são:

- (i)  $E\{X(t)\} = \mu(t) = \mu$ , constante, para todo  $t \in \mathcal{T}$ ;
- (ii)  $E\{X^2(t)\} < \infty$ , para todo  $t \in \mathcal{T}$ ;
- (iii)  $\gamma(t_1, t_2) = \text{Cov}\{X(t_1), X(t_2)\}$  é uma função apenas de  $|t_1 - t_2|$ .

Como exemplo de um processo estacionário, vamos apresentar o ruído branco.

### ➤ Ruído Branco (RB)

Dizemos que  $\{\varepsilon_t, T \in Z\}$  é um ruído branco discreto estacionário se as variáveis aleatórias  $\varepsilon_t$  são não correlacionadas, isto é,  $\text{Cov}\{\varepsilon_t, \varepsilon_s\} = 0$ ,  $t \neq s$  e se  $E\{\varepsilon_t\} = \mu$  e  $\text{Var}\{\varepsilon_t\} = \sigma^2$ , para todo  $t$ . Para simplificar, denotaremos  $\varepsilon_t \sim \text{RB}(\mu, \sigma^2)$ .

Para mais detalhes acerca do tema estacionariedade, sugerimos Morettin e Toloi (2006).

## 2.3 Modelos ARMA

### 2.3.1 Processos Autorregressivos

Dizemos que  $\{X_t, t \in Z\}$  é um processo autorregressivo de ordem  $p$  denotado por  $X_t \sim \text{AR}(p)$ , se  $X_t$  satisfizer à equação de diferenças

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \varepsilon_t, \quad (2.1)$$

em que  $\mu, \phi_1, \dots, \phi_p$  são parâmetros reais e  $\varepsilon_t \sim \text{RB}(0, \sigma^2)$ . Temos que  $E(X_t) = \mu$  e, se escrevermos o processo na forma

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t,$$

então  $\mu = E(X_t) = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}$ .

Seja  $B$  o operador retroativo definido por  $B^s X_t = X_{t-s}$ ,  $s \geq 1$ . Dessa forma, (2.1) pode ser escrita como

$$\phi(B)\tilde{X}_t = \varepsilon_t,$$

em que  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  é o operador autorregressivo de ordem  $p$  e

$\tilde{X}_t = X_t - \mu$ . Sem perda de generalidade, vamos supor que  $\mu = 0$ .

Um caso particular importante é o processo  $\text{AR}(1)$ , de onde vem que

$$\tilde{X}_t = \phi \tilde{X}_{t-1} + \varepsilon_t$$

e  $\phi(B) = 1 - \phi B$ .

Pode-se demonstrar (ver Box, Jenkins e Reinsel, 1994) que a condição para que  $X_t$  seja estacionário é que todas as raízes de  $\phi(B) = 0$  estejam fora do círculo unitário. Em particular, para  $p = 1$ ,  $\phi(B) = 1 - \phi B = 0$  implica em  $B = \phi^{-1}$  e, portanto,  $X_t$  será estacionário se  $|\phi| < 1$ .

Por outro lado, para que  $X_t$  seja invertível é necessário que o operador  $\phi(B)$  convirja para  $|B| \leq 1$ . Para o caso particular AR(1), como  $\phi(B) = 1 - \phi B$ , o processo é sempre invertível.

### 2.3.2 Processos de Médias-Móveis

Dizemos que  $\{X_t, t \in Z\}$  é um processo de médias-móveis de ordem  $q$ , denotado por MA( $q$ ), se  $X_t$  satisfizer a equação de diferenças

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q},$$

em que  $\mu, \theta_1, \dots, \theta_q$  são constantes reais e  $\varepsilon_t \sim \text{RB}(0, \sigma_\varepsilon^2)$ .

Considerando que  $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ , não há restrições sobre os parâmetros  $\theta_j$  para que o processo seja estacionário. Por outro lado, para que  $X_t$  seja invertível é necessário que as raízes da equação característica  $\theta(B) = 0$  estejam fora do círculo unitário. Vale citar que, nestas condições, um modelo MA( $q$ ) é equivalente a um modelo AR de ordem infinita.

### 2.3.3 Processos Autorregressivos e de Médias-Móveis

Dizemos que  $\{X_t, t \in Z\}$  é um processo autorregressivo e de médias-móveis de ordem ( $p, q$ ) denotado por  $X_t \sim \text{ARMA}(p, q)$ , se  $X_t$  satisfizer à equação de diferenças

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \quad (2.2)$$

em que  $\mu$  é a média do processo e  $\varepsilon_t \sim \text{RB}(0, \sigma^2)$ . Utilizando os operadores autorregressivo e de médias móveis descritos anteriormente, podemos escrever (3.4) como

$$\phi(B)\tilde{X}_t = \theta(B)\varepsilon_t,$$

em que  $\tilde{X}_t = X_t - \mu$ .

O processo ARMA( $p, q$ ) será estacionário se as raízes de  $\phi(B) = 0$  estiverem fora do círculo unitário e será invertível se as raízes de  $\theta(B) = 0$  estiverem fora do círculo unitário.

### 2.3.4 Função Densidade Espectral ou Espectro

A função densidade espectral de um processo estacionário,  $Z_t = \psi(B)a_t$ ,  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ , é dada por:

$$f(\lambda) = \frac{\sigma_a^2}{2\pi} |\psi(e^{-i\lambda})|^2, \quad -\pi \leq \lambda \leq \pi,$$

$$\text{em que } \psi(e^{-i\lambda}) = 1 + \psi_1 e^{-i\lambda} + \psi_2 e^{-2i\lambda} + \psi_3 e^{-3i\lambda} + \dots,$$

isto é, o polinômio linear geral aplicado em  $e^{-i\lambda}$ .

Definamos, agora, o espectro de um processo AR(p). Como  $\psi(B) = \phi^{-1}(B)$ , temos que esse espectro será dado por:

$$f(\lambda) = \frac{\sigma_a^2}{2\pi |\phi(e^{-i\lambda})|^2}, \quad -\pi \leq \lambda \leq \pi,$$

$$f(\lambda) = \frac{\sigma_a^2}{2\pi |1 - \phi_1 e^{-i\lambda} - \phi_2 e^{-2i\lambda} - \dots - \phi_p e^{-pi\lambda}|^2}.$$

## 2.4 Processos Sazonais

Muitas séries temporais exibem um comportamento sazonal, indicando a possibilidade de existir raízes unitárias correspondentes a picos no espectro em frequências sazonais (do tipo  $w_s = \frac{2\pi j}{s}$ ,  $j = 1, \dots, s-1$ , em que  $s =$  período da série).

As três principais classes de processos que podem ser utilizadas para modelar uma série temporal sazonal são:

- 1) Processo Sazonal Puramente Determinístico;
- 2) Processo Sazonal Estacionário;
- 3) Processo Sazonal Integrado.

O primeiro processo é gerado por variáveis *dummies* sazonais. Considerando, por exemplo, uma série trimestral teremos:

$$x_t = u_t$$

$$u_t = w_0 + w_1 S_1 + w_2 S_2 + w_3 S_3$$

com  $w_i$ ,  $i = 0, 1, 2, 3$  constantes e  $S_i = 1$ , se a observação corresponder ao  $i$ -ésimo trimestre e  $S_i = 0$ , caso contrário.

O segundo tipo, Processo Sazonal Estacionário, pode ser gerado por um modelo autorregressivo:

$$\varphi(B)x_t = \varepsilon_t, \quad \varepsilon_t \text{ iid}(0, \sigma_\varepsilon^2) \quad (2.3)$$

com todas as raízes de  $\varphi(B) = 0$  fora do círculo unitário e com, no mínimo, um par de raízes complexas com periodicidade sazonal. Dessa forma, o espectro do processo é dado por:

$$f(w) = \frac{\sigma^2}{|\varphi(e^{iw})|^2}$$

que, por hipótese, tem picos em frequências sazonais  $w_s$ . Vale citar que  $\sigma_\varepsilon^2$  é a variância da série  $\varepsilon_t$ . No caso de dados trimestrais, segue um exemplo:

$$x_t = \rho x_{t-4} + \varepsilon_t, \quad |\rho| < 1$$

Nesse caso, há picos nas frequências sazonais  $\frac{\pi}{2}$  (um ciclo por ano) e  $\pi$  (dois ciclos por ano).

Finalmente, dizemos que um Processo é Sazonal Integrado de ordem  $d$  na frequência  $\theta$ , se o seu espectro é da forma:

$$f(w) = c(w - \theta)^{-2d}, \quad w \text{ próximo de } \theta.$$

A notação utilizada para esse tipo de processo é  $x_t \sim I_\theta(d)$ . O caso mais comum é  $d = 1$ .

➤ **Processo Trimestral**

Um exemplo de um processo trimestral integrado na frequência correspondente a dois ciclos por ano é:

$$x_t = -x_{t-1} + \varepsilon_t$$

e a um ciclo por ano:

$$x_t = -x_{t-2} + \varepsilon_t$$

O operador diferença sazonal trimestral pode ser escrito como:

$$(1 - B^4) = (1 - B)(1 + B)(1 - iB)(1 + iB) = (1 - B)(1 + B)(1 + B^2). \quad (2.4)$$

Portanto, há quatro raízes unitárias com módulo um. As raízes são: 1, -1, i e -i (uma na frequência zero, uma em dois ciclos por ano e um par complexo em um ciclo por ano).

➤ **Processo Mensal**

Suponha, agora, que a série de interesse  $x_t$  tenha observações mensais e seja gerada pelo processo autorregressivo dado por (2.3), com  $\varphi(B) = (1 - B^{12})$ , denominado operador diferença sazonal, que foi introduzido por Box e Jenkins (1970). Esse operador pode ser fatorado da seguinte forma:

$$(1 - B^{12}) = (1 - B)(1 + B)(1 + B^2)(1 + B + B^2)(1 - B + B^2)(1 + \sqrt{3}B + B^2)(1 - \sqrt{3}B + B^2). \quad (2.5)$$

Portanto, as raízes unitárias sazonais são dadas por:

$$-1, \quad \pm i, \quad -\frac{1}{2}(1 \pm \sqrt{3}i), \quad \frac{1}{2}(1 \pm \sqrt{3}i), \quad -\frac{1}{2}(\sqrt{3} \pm i), \quad \frac{1}{2}(\sqrt{3} \pm i)$$

e correspondem a 6, 3, 9, 8, 4, 2, 10, 7, 5, 1 e 11 ciclos por ano, respectivamente. As frequências dessas raízes são  $\pi$ ,  $\pm \frac{\pi}{2}$ ,  $\pm \frac{2\pi}{3}$ ,  $\pm \frac{\pi}{3}$ ,  $\pm \frac{5\pi}{6}$  e  $\pm \frac{\pi}{6}$ , respectivamente. Assim, o operador diferença sazonal tem uma raiz unitária na frequência zero e onze raízes unitárias sazonais.

Na prática, pode acontecer de uma série temporal sazonal ser sazonalmente integrada, ter uma sazonalidade determinística e uma sazonalidade estacionária ou alguma combinação entre esses processos. Os modelos de séries temporais que exibem formas complexas de sazonalidades podem ser escritos como:

$$d(B)a(B)(x_t - \mu_t) = \varepsilon_t,$$

onde todas as raízes de  $a(z) = 0$  estão fora do círculo unitário e, em contrapartida, todas as raízes de  $d(z) = 0$  estão sobre o círculo unitário. A sazonalidade estacionária e outras componentes de estacionariedade de  $x_t$  são contempladas por  $a(B)$ , enquanto a sazonalidade determinística está em  $\mu_t$  quando não há raízes unitárias sazonais em  $d(B)$ .

## 2.5 Testes de Dickey-Fuller

Trata-se de testes para raízes unitárias em modelos da família ARIMA. Para efeito de ilustração, consideremos o modelo AR(1) estacionário com média zero dado por:

$$X_t = \phi X_{t-1} + \varepsilon_t, \varepsilon_t \sim RB(0, \sigma^2) \quad (2.6)$$

em que  $|\phi| < 1$ .

Se  $\hat{\phi}_{MV}$  indica o Estimador de Máxima Verossimilhança (EMV) de  $\phi$ , então considerando T observações desse processo, temos:

$$\hat{\phi}_{MV} \sim \mathcal{N}\left(\phi, \frac{(1 - \phi^2)}{T}\right)$$

Supondo que queiramos testar a hipótese de que  $\phi$  é igual a um determinado  $\phi_0$ . Dessa forma, teremos o seguinte teste:

$$\begin{aligned} H_0: \phi &= \phi_0 \\ H_a: \phi &\neq \phi_0 \end{aligned}$$

A estatística do teste será dada por:

$$\frac{\hat{\phi}_{MV} - \phi_0}{\widehat{ep}(\hat{\phi}_{MV})} \quad (2.7)$$

em que  $\widehat{ep}(\widehat{\phi}_{MV})$  é o erro padrão estimado de  $\widehat{\phi}_{MV}$ .

Sob a hipótese nula, temos que a distribuição da estatística em (2.7) é dada por:

$$\sqrt{T}(\widehat{\phi}_{MV} - \phi) \sim \mathcal{N}(0, 1 - \phi^2).$$

No caso de raízes unitárias, essa aproximação normal não se aplica, logo não podemos utilizar a distribuição t para realizarmos o seguinte teste:

$$\begin{aligned} H_0: \phi &= 1 \\ H_a: \phi &< 1 \end{aligned} \tag{2.8}$$

Sabemos que os EMV são assintoticamente equivalentes aos EMQ, então assumindo que  $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ , temos:

$$\begin{aligned} \widehat{\phi}_{MQ} &= \frac{\sum_{t=2}^T X_{t-1} X_t}{\sum_{t=2}^T X_{t-1}^2} \\ \text{e } \widehat{\phi}_{MQ} - \phi &= \frac{\sum_{t=2}^T X_{t-1} \epsilon_t}{\sum_{t=2}^T X_{t-1}^2} \end{aligned}$$

Logo, para testar (2.8) temos que estudar o comportamento de

$$\widehat{\phi}_{MQ} - 1 = \frac{\sum_{t=2}^T X_{t-1} \epsilon_t}{\sum_{t=2}^T X_{t-1}^2}$$

Considerando o modelo expresso em (2.6), temos:

$$\Delta X_t = \phi^* X_{t-1} + \epsilon_t, \tag{2.9}$$

em que  $\phi^* = \phi - 1$ .

Para obtermos o Estimador de Mínimos Quadrados (EMQ) de  $\phi^*$ , basta fazermos a regressão de mínimos quadrados de  $\Delta X_t$  sobre  $X_{t-1}$ . Dessa forma, o teste em (2.8) é equivalente a testar:

$$\begin{aligned} H_0: \phi^* &= 0 \\ H_a: \phi^* &< 0 \end{aligned} \tag{2.10}$$

**Teorema 2.1.** Considere o modelo em (2.6) com  $X_0 = 0$  e suponha que  $\varepsilon_t \sim i.i.d. (0, \sigma^2)$ . Assim, temos:

$$T(\hat{\phi}_{MQ} - 1) \rightarrow \frac{\frac{1}{2}([W(1)]^2 - 1)}{\int_0^1 [W(r)]^2 dr}, \quad (2.11)$$

em que  $W(r)$  é o Movimento Browniano Padrão, ou seja, para cada  $z$ , temos que  $W(z) \sim N(0, z)$ .

Para testarmos (2.8) ou (2.10) podemos utilizar a seguinte estatística:

$$\hat{t} = \frac{\hat{\phi}_{MQ}^*}{\widehat{ep}(\hat{\phi}_{MQ}^*)} \quad (2.12)$$

em que  $\widehat{ep}(\hat{\phi}_{MQ}^*) = \frac{s}{\sqrt{(\sum_{t=1}^T X_{t-1}^2)}}$  e  $S^2 = \frac{1}{T-2} \sum_{t=2}^T (\Delta X_t - \hat{\phi}_{MQ}^* X_{t-1})^2$  é o estimador de  $\sigma^2$  na regressão expressa em (2.9).

Vale citar que a estatística dada em (2.12) pode ser escrita na forma:

$$\hat{t} = \frac{T^{-1} \sum X_{t-1} \varepsilon_t}{\sqrt{S(T-2) \sum X_{t-1}^2}}$$

**Teorema 2.2.** Sob as mesmas suposições do teorema 2.1., temos:

$$\hat{t} \rightarrow \frac{\frac{1}{2}([W(1)]^2 - 1)}{\int_0^1 [W(r)]^2 dr} \quad (2.13)$$

Os testes de Dickey-Fuller são os testes que utilizam (2.11) ou (2.13). Vale comentar que existem extensões do teste de Dickey-Fuller, que consideram que a série de interesse pode ser representada por um processo AR(p) ou por um processo ARMA(p, q) e, também, extensões que permitem a incorporação de uma tendência determinística e erros correlacionados. Para mais detalhes, recomendamos ver Dickey e Fuller (1979, 1981).

## Capítulo 3

# Testes HEGY para Raízes Unitárias Sazonais

Neste Capítulo apresentaremos o procedimento HEGY desenvolvido originalmente para dados trimestrais, a extensão do procedimento HEGY para dados mensais e as distribuições assintóticas das estatísticas de ambos os testes (trimestral e mensal).

### 3.1 O Procedimento HEGY Trimestral

Hilleberg, Engle, Granger and Yoo (1990) desenvolveram um teste para determinar se há raízes unitárias em determinadas frequências sazonais em séries univariadas, ao qual chamaram de Procedimento HEGY. Trata-se de testes em séries temporais lineares que possuem raízes com módulo um, mas que correspondem a frequências sazonais.

Na literatura podemos encontrar alguns testes desenvolvidos com o propósito de avaliar a presença de raízes unitárias em série temporais univariadas. No entanto, a maior parte deles possui a desvantagem de não permitir testar se existem raízes unitárias em apenas algumas frequências sazonais.

O procedimento HEGY, por sua vez, trata-se de uma estratégia de teste que avalia a presença de raízes unitárias isoladamente em todas as frequências sazonais assim como na frequência zero, isto é, permite testar a existência de uma particular raiz unitária sem que o teste seja comprometido se outras raízes unitárias estiverem presentes na frequência zero ou em frequências sazonais.

Considerando dados provenientes de um processo trimestral integrado, o polinômio  $(1 - B^4)$  descrito em (2.4) pode ser expresso por:

$$(1 - B^4) = (1 - B)(1 + B)(1 - iB)(1 + iB) = (1 - B)(1 + B)(1 + B)^2$$

tal que as raízes unitárias são 1, -1, i e -i correspondentes a frequência zero,  $\frac{1}{2}$  ciclo por trimestre (dois ciclos por ano) e  $\frac{1}{4}$  ciclo por trimestre (um ciclo por ano). Vale citar que ambas as raízes i e -i correspondem a um ciclo anual.

O procedimento HEGY tem por objetivo testar se as raízes do polinômio autorregressivo pertencem ao círculo unitário, contra a alternativa de que estão fora do círculo unitário.

Para a realização dos testes é conveniente reescrever o polinômio autorregressivo de acordo com o resultado a seguir, devido a Lagrange e usado em teoria de aproximação.

**Resultado 3.1.** (Heggy, 1990) *Qualquer polinômio  $\varphi(B)$  que assume valores finitos nos pontos  $\theta_1, \theta_2, \dots, \theta_p$  distintos, diferentes de zero e possivelmente complexos, pode ser expresso em termos de polinômios elementares e um resto como a seguir:*

$$\varphi(B) = \sum_{k=1}^p \frac{\lambda_k \Delta(B)}{\delta_k(B)} + \Delta(B) \varphi^{**}(B), \quad (3.1)$$

em que  $\lambda_k$  é um conjunto de constantes determinadas por:

$$\lambda_k = \frac{\varphi(\theta_k)}{\prod_{j \neq k} \delta_j(\theta_k)},$$

$\varphi^{**}(B)$  é um polinômio possivelmente infinito ou racional,

$$\delta_k(B) = 1 - \frac{1}{\theta_k} B$$

$$\text{e } \Delta(B) = \prod_{k=1}^p \delta_k(B).$$

A expressão (3.1) é uma linearização do polinômio  $\varphi(B)$  em torno de  $\theta_1, \theta_2, \dots, \theta_p$ .

Adicionando e subtraindo  $\Delta(B) \sum \lambda_k$  da expressão dada em (3.1), obtemos:

$$\varphi(B) = \sum_{k=1}^p \frac{\lambda_k \Delta(B)(1 - \delta_k(B))}{\delta_k(B)} + \Delta(B) \varphi^*(B), \quad (3.2)$$

$$\text{em que } \varphi^*(B) = \varphi^{**}(B) + \sum \lambda_k$$

e, portanto, o polinômio  $\varphi(B)$  terá uma raiz em  $\theta_k$  se e somente se  $\lambda_k = 0$ , indicando que testes para raízes unitárias podem ser equivalentes a testar se os parâmetros  $\lambda_k$  são iguais a zero.

Para aplicar o resultado 3.1 para testar raízes unitárias em dados trimestrais, expandimos o polinômio  $\varphi(B)$  em torno das raízes +1, -1, i e -i. Assim, de (3.2) obtemos:

$$\begin{aligned}\varphi(B) = & \lambda_1 B(1+B)(1+B^2) + \lambda_2(-B)(1-B)(1+B^2) + \lambda_3(-iB)(1-B)(1+B)(1-iB) \\ & + \lambda_4(iB)(1-B)(1+B)(1+iB) + \varphi^*(B)(1-B^4)\end{aligned}\quad (3.3)$$

Claramente,  $\lambda_3$  e  $\lambda_4$  são conjugados complexos, uma vez que  $\varphi(B)$  é real.

Substituindo  $\pi_1 = -\lambda_1$ ,  $\pi_2 = -\lambda_2$ ,  $2\lambda_3 = -\pi_3 + i\pi_4$  e  $2\lambda_4 = -\pi_3 - i\pi_4$ , temos:

$$\begin{aligned}\varphi(B) = & -\pi_1 B(1+B+B^2+B^3) - \pi_2(-B)(1-B+B^2-B^3) - (\pi_4 + \pi_3 B)(-B)(1-B^2) \\ & + \varphi^*(B)(1-B^4)\end{aligned}\quad (3.4)$$

Assumindo que os dados são gerados por um modelo autorregressivo, temos:

$$\varphi(B)x_t = \varepsilon_t \quad (3.5)$$

e substituindo  $\varphi(B)$  por (3.4), temos o modelo que será utilizado para realizar o teste HEGY:

$$\varphi^*(B)y_{4t} = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \varepsilon_t, \quad (3.6)$$

em que,

$$\begin{aligned}y_{1,t} &= (1+B+B^2+B^3)x_t = S(B)x_t, \\ y_{2,t} &= -(1-B+B^2-B^3)x_t, \\ y_{3,t} &= -(1-B^2)x_t, \\ y_{4,t} &= (1-B^4)x_t = \Delta_4 x_t \text{ e}\end{aligned}\quad (3.7)$$

$x_t$  é a série original (possivelmente sazonal).

O modelo em (3.6) possui variáveis regressoras assintoticamente mutuamente ortogonais, cujos coeficientes podem ser estimados por mínimos quadrados, possivelmente com lags adicionais de  $y_4$  para branquear os erros. Assim, a estratégia dos testes se torna aparente e podemos reescrever as hipóteses dos testes conforme as raízes que serão avaliadas. A fim de obter os testes correspondentes, analisemos cada uma das raízes. Vamos iniciar pela raiz 1.

**(i) Teste para a raiz 1:**

Substituindo B por 1 em (3.4) obtemos  $\varphi(1) = -4\pi_1$ . Portanto, para que 1 seja raiz do polinômio  $\varphi(B)$  é necessário que  $\pi_1$  seja igual a zero e, como  $\pi_1 = -\lambda_1$ , então equivale a testar se  $\lambda_1 = 0$ . A hipótese alternativa natural é a estacionariedade, o que significa  $\varphi(1) > 0$  ou, equivalentemente,  $\pi_1 < 0$ . Assim, as hipóteses a serem testadas são:

$$\begin{aligned} H_0: \pi_1 &= 0 \\ H_a: \pi_1 &< 0 \end{aligned} \tag{3.8}$$

**(ii) Teste para a raiz -1:**

De forma análoga, podemos obter as hipóteses do teste para a raiz -1. Nesse caso, as hipóteses são:

$$\begin{aligned} H_0: \pi_2 &= 0 \\ H_a: \pi_2 &< 0 \end{aligned} \tag{3.12}$$

**(iii) Testes para as raízes +i e -i:**

Para o par de raízes complexas, substituindo B por  $\pm i$  em (3.4) temos que  $|\varphi(i)| = 0$  equivale a  $\pi_3 = \pi_4 = 0$  e, conseqüentemente,  $\lambda_3$  assumirá valor zero se e somente se  $\pi_3$  e  $\pi_4$  forem iguais a zero. Portanto, as hipóteses do teste para as raízes  $\pm i$  são:

$$\begin{aligned} H_0: \pi_3 = \pi_4 &= 0 \\ H_a: \pi_3 \text{ e/ou } \pi_4 &\neq 0 \end{aligned} \tag{3.13}$$

O teste em (3.13) possui uma hipótese nula bidimensional, de tal forma que devemos calcular uma estatística “F” para testar a hipótese de que  $\pi_3 = \pi_4 = 0$ . Uma estratégia é fazer um teste bilateral de  $\pi_4 = 0$  e, se o resultado for positivo, testar unilateralmente se  $\pi_3 = 0$  (contra a alternativa de  $\pi_3 < 0$ ).

Note que não haverá raízes unitárias sazonais se  $\pi_2$  e qualquer um entre  $\pi_3$  e  $\pi_4$  for diferente de zero. É importante ressaltar que para concluir que uma série não tem raízes unitárias e é, portanto, estacionária, devemos estabelecer que todos os  $\pi$ 's são diferentes de zero (salvo possivelmente um entre  $\pi_3$  e  $\pi_4$ ).

## 3.2 Distribuição Assintótica das Estatísticas do Teste HEGY Trimestral

A distribuição assintótica das estatísticas “t” da regressão expressa em (3.6) foram estudadas por Chan e Wei (1988). Eles observaram que a distribuição assintótica para os testes de  $\pi_1$  e  $\pi_2$  pode ser extraída de Dickey e Fuller (1979) e de Fuller (1976), enquanto para  $\pi_3$  pode ser extraída de Dickey, Hasza e Fuller (1984), assumindo que  $\pi_4$  é igual a zero. Os testes são assintoticamente similares ou invariantes com respeito aos parâmetros de perturbação. Os resultados obtidos para amostras finitas se aproximam dos obtidos pela teoria assintótica e os testes tem poder razoável contra as alternativas especificadas.

Para mostrar como as distribuições limites das estatísticas “t” da regressão estão relacionadas com os testes de raízes unitárias padrão, consideremos  $\varphi^*(B) = 1$  em (3.6). O teste para  $\pi_1 = 0$  terá a distribuição usual de Dickey-Fuller se  $\pi_2 = \pi_3 = \pi_4 = 0$ , uma vez que o modelo poderá ser escrito na forma:

$$y_{1t} = (1 + \pi_1)y_{1t-1} + \varepsilon_t.$$

De forma similar, temos:

$$y_{2t} = -(1 + \pi_2)y_{2t-1} + \varepsilon_t,$$

se os outros  $\pi$ 's forem iguais a zero. Se  $y_{2t}$  estiver relacionado com  $y_{2t-1}$ , então a distribuição de Dickey e Fuller será apropriada. O terceiro teste pode ser escrito como:

$$y_{3t} = -(1 + \pi_3)y_{3t-2} + \varepsilon_t,$$

assumindo  $\pi_4 = 0$ .

Para mostrar que as mesmas distribuições são obtidas quando não é conhecido a priori que algum dos  $\pi$ 's são iguais a zero, dois casos devem ser considerados. Primeiramente, se os outros  $\pi$ 's que estão sendo testados não são zeros, então o processo não tem raízes unitárias nessas frequências e os correspondentes  $y$ 's são estacionários. No entanto, se alguns dos outros  $\pi$ 's são iguais a zero, então há outras raízes unitárias na regressão.

Entretanto, é exatamente sob essa condição que os correspondentes  $y$ 's são assintoticamente não correlacionados. A distribuição da estatística do teste não será afetada pela inclusão de uma variável com um coeficiente zero que é ortogonal às variáveis incluídas. Por exemplo, quando testamos  $\pi_1 = 0$ , supomos  $\pi_2 = 0$ , mas  $y_2$  ainda está incluído na regressão. Então,  $y_1$  e  $y_2$  serão assintoticamente não correlacionados desde que eles tenham raízes unitárias em diferentes frequências e ambos serão assintoticamente não correlacionados com lags de  $y_4$  que é estacionário. O teste para  $\pi_1 = 0$  terá a mesma distribuição limite independentemente de  $y_2$  estar incluído na regressão.

Em modelos mais complexos onde a alternativa inclui componentes determinísticas é necessário permitir que  $\mu_t \neq 0$ . Nesse caso, para os testes termos o seguinte modelo:

$$\varphi^*(B)y_{4t} = \mu_t + \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{3t-2} + \pi_4 y_{3t-1} + \varepsilon_t \quad (3.14)$$

que também pode ser estimado por mínimos quadrados e as estatísticas relativas aos  $\pi$ 's podem ser usadas para inferência.

Quando componentes determinísticas estão presentes na regressão, temos que as distribuições se modificam. O intercepto e a tendência da média determinística influenciam somente a distribuição de  $\pi_1$ . Isso ocorre porque esses componentes tem todos os seus espectros concentrados na frequência zero. Se o intercepto for incluído na regressão, as três *dummies* sazonais restantes não influenciarão na distribuição limite de  $\pi_1$ . Entretanto, as *dummies* sazonais afetam as distribuições de  $\pi_2, \pi_3$  e  $\pi_4$ .

### 3.3 Valores Críticos de Amostras Finitas para as Estatísticas do Teste HEGY Trimestral

A Tabela 3.1, extraída de Hylleberg, Engle, Granger and Yoo (1990), apresenta os valores críticos obtidos por simulações de Monte Carlo, para os testes t unilaterais sobre  $\pi_1, \pi_2$  e  $\pi_3$ , nos casos mais importantes (considerando diferentes combinações das seguintes componentes na regressão: intercepto, *dummies* sazonais e tendência). Vale citar que os valores críticos da Tabela 3.1 são muito próximos dos valores críticos obtidos por simulações de Monte Carlo em Dickey-Fuller (1979) e Dickey-Hasza-Fuller (1984), para as situações em que as estatísticas foram tabuladas.

Tabela 3.1 – Valores críticos para as estatísticas do teste HEGY trimestral ( $\pi_1, \pi_2$  e  $\pi_3$ )

Critical values from the small-sample distributions of test statistics for seasonal unit roots on 24000 Monte Carlo replications; data generating process $\Delta_4 x_t = \varepsilon_t \sim \text{nid}(0,1)$ .													
Auxiliary regressions		Fractiles											
		T	$t^*: \pi_1$				$t^*: \pi_2$				$t^*: \pi_3$		
	0.01		0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10
No intercept	48	-2.72	-2.29	-1.95	-1.59	-2.67	-2.27	-1.95	-1.60	-2.66	-2.23	-1.93	-1.52
No seas. dum.	100	-2.60	-2.26	-1.97	-1.61	-2.61	-2.22	-1.92	-1.57	-2.55	-2.18	-1.90	-1.53
No trend	136	-2.62	-2.25	-1.93	-1.59	-2.60	-2.23	-1.94	-1.61	-2.58	-2.21	-1.92	-1.56
	200	-2.62	-2.23	-1.94	-1.62	-2.60	-2.24	-1.95	-1.61	-2.58	-2.24	-1.92	-1.55
Intercept	48	-3.66	-3.25	-2.96	-2.62	-2.68	-2.27	-1.95	-1.60	-2.64	-2.23	-1.90	-1.52
No seas. dum.	100	-3.47	-3.14	-2.88	-2.58	-2.61	-2.24	-1.95	-1.60	-2.61	-2.23	-1.90	-1.54
No trend	136	-3.51	-3.17	-2.89	-2.58	-2.60	-2.21	-1.91	-1.58	-2.53	-2.18	-1.88	-1.53
	200	-3.48	-3.13	-2.87	-2.57	-2.58	-2.22	-1.92	-1.59	-2.57	-2.21	-1.90	-1.53
Intercept	48	-3.77	-3.39	-3.08	-2.72	-3.75	-3.37	-3.04	-2.69	-4.31	-3.92	-3.61	-3.24
Seas. dum.	100	-3.55	-3.22	-2.95	-2.63	-3.60	-3.22	-2.94	-2.63	-4.06	-3.72	-3.44	-3.14
No trend	136	-3.56	-3.23	-2.94	-2.62	-3.49	-3.15	-2.90	-2.59	-4.06	-3.72	-3.44	-3.11
	200	-3.51	-3.18	-2.91	-2.59	-3.50	-3.16	-2.89	-2.60	-4.00	-3.67	-3.38	-3.07
Intercept	48	-4.23	-3.85	-3.56	-3.21	-2.65	-2.24	-1.91	-1.57	-2.68	-2.27	-1.92	-1.52
No seas. dum.	100	-4.07	-3.73	-3.47	-3.16	-2.58	-2.24	-1.94	-1.60	-2.56	-2.19	-1.89	-1.54
Trend	136	-4.09	-3.75	-3.46	-3.16	-2.65	-2.25	-1.96	-1.63	-2.56	-2.20	-1.90	-1.52
	200	-4.05	-3.70	-3.44	-3.15	-2.59	-2.25	-1.95	-1.62	-2.58	-2.21	-1.92	-1.56
Intercept	48	-4.46	-4.04	-3.71	-3.37	-3.80	-3.41	-3.08	-2.73	-4.46	-4.02	-3.66	-3.28
Seas. dum.	100	-4.09	-3.80	-3.53	-3.22	-3.60	-3.22	-2.94	-2.63	-4.12	-3.76	-3.48	-3.14
Trend	136	-4.15	-3.80	-3.52	-3.21	-3.57	-3.18	-2.93	-2.61	-4.05	-3.72	-3.44	-3.12
	200	-4.05	-3.74	-3.49	-3.18	-3.52	-3.18	-2.91	-2.60	-4.04	-3.69	-3.41	-3.10

Tabela extraída de Hilleberg, Engle, Granger and Yoo (1990)

Na Tabela 3.2, também extraída de Hylleberg, Engle, Granger and Yoo (1990), temos os valores críticos para os testes t bilaterais sobre  $\pi_4 = 0$  e os valores críticos para o teste F sobre  $\pi_3 \cap \pi_4 = 0$ . Note que a distribuição da estatística “t” é muito similar a normal padrão, exceto quando a regressão auxiliar contém *dummies* sazonais, nos casos em que a cauda da distribuição aumenta. A distribuição da estatística “F” é bastante similar a uma distribuição F com graus de liberdade iguais a dois e T menos o número de variáveis regressoras em (3.6). Entretanto, quando *dummies* sazonais estão presentes, a cauda da distribuição também se torna mais pesada.

Tabela 3.2 – Valores críticos para as estatísticas do teste HEGY trimestral ( $\pi_3$  e  $\pi_4$ )

Critical values from the small-sample distributions of test statistics for seasonal unit roots on 24000 Monte Carlo replications: data-generating process $\Delta_4 x_t = \epsilon_t \sim \text{nid}(0, 1)$ .													
Auxiliary regressions	T	Fractiles											
		‘t’: $\pi_4$							‘F’: $\pi_3 \cap \pi_4$				
		0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99	0.90	0.95	0.975	0.99
No Intercept	48	-2.51	-2.11	-1.76	-1.35	1.33	1.72	2.05	2.49	2.45	3.26	4.04	5.02
No seas. dum.	100	-2.43	-2.01	-1.68	-1.32	1.31	1.67	2.00	2.40	2.39	3.12	3.89	4.89
No trend	136	-2.44	-1.99	-1.68	-1.31	1.30	1.66	1.99	2.38	2.41	3.14	3.86	4.81
	200	-2.43	-1.98	-1.65	-1.30	1.29	1.67	1.97	2.36	2.42	3.16	3.92	4.81
Intercept	48	-2.44	-2.06	-1.72	-1.33	1.30	1.68	2.04	2.41	2.32	3.04	3.78	4.78
No seas. dum.	100	-2.38	-1.99	-1.68	-1.30	1.28	1.65	1.97	2.32	2.35	3.08	3.81	4.77
No trend	136	-2.36	-1.98	-1.68	-1.31	1.27	1.65	1.97	2.31	2.36	3.00	3.70	4.73
	200	-2.36	-1.98	-1.66	-1.29	1.28	1.65	1.96	2.30	2.37	3.12	3.86	4.76
Intercept	48	-2.86	-2.37	-1.98	-1.53	1.54	1.96	2.35	2.81	5.50	6.60	7.68	9.22
Seas. dum.	100	-2.78	-2.32	-1.96	-1.53	1.52	1.93	2.29	2.73	5.56	6.57	7.72	8.74
No trend	136	-2.72	-2.31	-1.96	-1.52	1.51	1.92	2.28	2.71	5.56	6.63	7.66	8.92
	200	-2.74	-2.33	-1.96	-1.54	1.53	1.95	2.32	2.78	5.56	6.61	7.53	8.93
Intercept	48	-2.41	-2.05	-1.70	-1.33	1.26	1.64	1.96	2.37	2.23	2.95	3.70	4.64
No seas. dum.	100	-2.38	-1.97	-1.65	-1.28	1.28	1.65	1.98	2.32	2.31	2.98	3.71	4.70
Trend	136	-2.36	-1.97	-1.64	-1.29	1.26	1.62	1.92	2.31	2.33	3.04	3.69	4.57
	200	-2.35	-1.97	-1.66	-1.29	1.26	1.64	1.96	2.30	2.34	3.07	3.76	4.66
Intercept	48	-2.75	-2.26	-1.91	-1.48	1.51	1.97	2.34	2.78	5.37	6.55	7.70	9.27
Seas. dum.	100	-2.76	-2.32	-1.94	-1.51	1.51	1.92	2.28	2.69	5.52	6.60	7.52	8.79
Trend	136	-2.71	-2.78	-1.94	-1.51	1.53	1.96	2.31	2.78	5.55	6.62	7.59	8.77
	200	-2.65	-2.27	-1.92	-1.48	1.55	1.97	2.31	2.71	5.56	6.57	7.56	8.96

Tabela extraída de Hylleberg, Engle, Granger and Yoo (1990)

### 3.4 O Procedimento HEGY Mensal

Beaulieu e Miron (1993) estenderam o procedimento HEGY, desenvolvido para dados trimestrais, a dados mensais. De forma análoga ao apresentado para o procedimento HEGY trimestral, o objetivo é saber se o polinômio autorregressivo  $\varphi(B)$  tem raízes iguais a um em valor absoluto na frequência zero ou em frequências sazonais. Este procedimento também permite testar as hipóteses sobre uma particular raiz unitária sazonal, sem impor que todas as raízes unitárias sazonais estejam presentes.

O procedimento HEGY mensal também se utiliza da técnica de linearização do polinômio autorregressivo  $\varphi(B)$  em torno da raiz unitária na frequência zero e das (S-1) raízes unitárias sazonais, que podem estar presentes em dados mensais, a saber:

$$-1, \quad \pm i, \quad -\frac{1}{2}(1 \pm \sqrt{3}i), \quad \frac{1}{2}(1 \pm \sqrt{3}i), \quad -\frac{1}{2}(\sqrt{3} \pm i), \quad \frac{1}{2}(\sqrt{3} \pm i),$$

que correspondem a 6, 3, 9, 8, 4, 2, 10, 7, 5, 1 e 11 ciclos por ano, respectivamente.

Dessa forma, podemos escrever:

$$\varphi(B) = \sum_{k=1}^S \frac{\lambda_k \Delta(B)(1 - \delta_k(B))}{\delta_k(B)} + \Delta(B)\varphi^*(B), \quad (3.15)$$

em que

$$\delta_k(B) = 1 - \frac{1}{\theta_k} B, \quad \lambda_k = \frac{\varphi(\theta_k)}{\prod_{j \neq k} \delta_j(\theta_k)}, \quad \Delta(B) = \prod_{k=1}^S \delta_k(B),$$

$\varphi^*(B)$  é um resto com raízes fora do círculo unitário e  $\theta_k$  são as S-1 raízes unitárias sazonais mais a raiz unitária na frequência zero.

Expandindo o polinômio  $\varphi(B)$  em torno das raízes

$$1, \quad -1, \quad \pm i, \quad -\frac{1}{2}(1 \pm \sqrt{3}i), \quad \frac{1}{2}(1 \pm \sqrt{3}i), \quad -\frac{1}{2}(\sqrt{3} \pm i) \quad \text{e} \quad \frac{1}{2}(\sqrt{3} \pm i), \quad \text{obtemos:}$$

$$\begin{aligned} \varphi(B) = & \lambda_1 B(1+B)(1+B^2)(1+B^4+B^8) + \\ & \lambda_2 (-B)(1-B)(1+B^2)(1+B^4+B^8) + \\ & \lambda_3 (-iB)(1-iB)(1-B^2)(1+B^4+B^8) + \\ & \lambda_4 (iB)(1+iB)(1-B^2)(1+B^4+B^8) + \\ & \lambda_5 \left(-\frac{1}{2}B\right)(1-\sqrt{3}i+2B)(1-B+B^2)(1-B^2+B^6-B^8) + \\ & \lambda_6 \left(-\frac{1}{2}B\right)(1+\sqrt{3}i+2B)(1-B+B^2)(1-B^2+B^6-B^8) + \\ & \lambda_7 \left(\frac{1}{2}B\right)(1-\sqrt{3}i-2B)(1+B+B^2)(1-B^2+B^6-B^8) + \\ & \lambda_8 \left(\frac{1}{2}B\right)(1+\sqrt{3}i-2B)(1+B+B^2)(1-B^2+B^6-B^8)(1+B^2-B^6-B^8) + \end{aligned}$$

$$\begin{aligned}
& \lambda_9 \left(-\frac{1}{2}B\right) (\sqrt{3} - i + 2B)(1 - \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) + \\
& \lambda_{10} \left(-\frac{1}{2}B\right) (\sqrt{3} + i + 2B)(1 - \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) + \\
& \lambda_{11} \left(\frac{1}{2}B\right) (\sqrt{3} - i - 2B)(1 + \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) + \\
& \lambda_{12} \left(\frac{1}{2}B\right) (\sqrt{3} + i - 2B)(1 + \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) + \\
& \varphi^*(B)(1 - B^{12}).
\end{aligned} \tag{3.16}$$

Claramente, os pares  $(\lambda_3, \lambda_4)$ ,  $(\lambda_5, \lambda_6)$ ,  $(\lambda_7, \lambda_8)$  e  $(\lambda_9, \lambda_{10})$  são complexos conjugados, uma vez que  $\varphi(B)$  é real.

Sejam  $\pi_k$  definidos por:

$$\begin{aligned}
\lambda_1 &= -\pi_1, \\
\lambda_2 &= -\pi_2, \\
\lambda_3 &= \frac{1}{2}(-\pi_3 + i\pi_4), \\
\lambda_4 &= \frac{1}{2}(-\pi_3 - i\pi_4), \\
\lambda_5 &= \frac{1}{2}(-\pi_5 + i\pi_6), \\
\lambda_6 &= \frac{1}{2}(-\pi_5 - i\pi_6), \\
\lambda_7 &= \frac{1}{2}(-\pi_7 + i\pi_8), \\
\lambda_8 &= \frac{1}{2}(-\pi_7 - i\pi_8), \\
\lambda_9 &= \frac{1}{2}(-\pi_9 + i\pi_{10}), \\
\lambda_{10} &= \frac{1}{2}(-\pi_9 - i\pi_{10}), \\
\lambda_{11} &= \frac{1}{2}(-\pi_{11} + i\pi_{12}), \\
\lambda_{12} &= \frac{1}{2}(-\pi_{11} - i\pi_{12}).
\end{aligned}$$

Dessa forma, substituindo  $\pi_k$  por  $\lambda_k$  em (3.16), temos:

$$\begin{aligned}
\varphi(B) = & -\pi_1 B(1+B)(1+B^2)(1+B^4+B^8) - \\
& \pi_2(-B)(1-B)(1+B^2)(1+B^4+B^8) - \\
& (\pi_4 + \pi_3 B)(-B)(1-B^2)(1+B^4+B^8) - \\
& \frac{1}{2}(\sqrt{3}\pi_6 - (1+2B)\pi_5 B(1-B+B^2)(1-B^2+B^6-B^8) - \\
& \frac{1}{2}(\sqrt{3}\pi_8 - (1-2B)\pi_7)(-B)(1+B+B^2)(1-B^2+B^6-B^8) - \\
& \frac{1}{2}(\pi_{10} - (\sqrt{3}+2B)\pi_9)B(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8) - \\
& \frac{1}{2}(\pi_{12} - (\sqrt{3}-2B)\pi_{11})(-B)(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8) + \\
& \varphi^*(B)(1-B^{12}).
\end{aligned} \tag{3.17}$$

Assumindo que os dados são gerados por um modelo autorregressivo geral, temos que a substituição de  $\varphi(B)$  exposto em (3.17) em  $\varphi(B)x_t = \varepsilon_t$  fornece o seguinte modelo que será utilizado para realizar o teste HEGY:

$$\varphi(B)^* y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + \varepsilon_t, \tag{3.18}$$

em que

$$\begin{aligned}
y_{1t} &= (1+B+B^2+B^3+B^4+B^5+B^6+B^7+B^8+B^9+B^{10}+B^{11})x_t, \\
y_{2t} &= -(1-B+B^2-B^3+B^4-B^5+B^6-B^7+B^8-B^9+B^{10}-B^{11})x_t, \\
y_{3t} &= -(B-B^3+B^5-B^7+B^9-B^{11})x_t, \\
y_{4t} &= -(1-B^2+B^4-B^6+B^8-B^{10})x_t, \\
y_{5t} &= -\frac{1}{2}(1+B-2B^2+B^3+B^4-2B^5+B^6+B^7-2B^8+B^9+B^{10}-2B^{11})x_t, \\
y_{6t} &= \frac{\sqrt{3}}{2}(1-B+B^3-B^4+B^6-B^7+B^9-B^{10})x_t,
\end{aligned}$$

$$y_{7t} = \frac{1}{2}(1 - B - 2B^2 - B^3 + B^4 + 2B^5 + B^6 - B^7 - 2B^8 - B^9 + B^{10} + 2B^{11})x_t ,$$

$$y_{8t} = -\frac{\sqrt{3}}{2}(1 + B - B^3 - B^4 + B^6 + B^7 - B^9 - B^{10})x_t ,$$

$$y_{9t} = -\frac{1}{2}(\sqrt{3} - B + B^3 - \sqrt{3}B^4 + 2B^5 - \sqrt{3}B^6 + B^7 - B^9 + \sqrt{3}B^{10} - 2B^{11})x_t ,$$

$$y_{10t} = \frac{1}{2}(1 - \sqrt{3}B + 2B^2 - \sqrt{3}B^3 + B^4 - B^6 + \sqrt{3}B^7 - 2B^8 + \sqrt{3}B^9 - B^{10})x_t ,$$

$$y_{11t} = \frac{1}{2}(\sqrt{3} + B - B^3 - \sqrt{3}B^4 - 2B^5 - \sqrt{3}B^6 - B^7 + B^9 + \sqrt{3}B^{10} + 2B^{11})x_t ,$$

$$y_{12t} = -\frac{1}{2}(1 + \sqrt{3}B + 2B^2 + \sqrt{3}B^3 + B^4 - B^6 - \sqrt{3}B^7 - 2B^8 - \sqrt{3}B^9 - B^{10})x_t ,$$

$$y_{13t} = (1 - B^{12})x_t \text{ e}$$

$x_t$  é a série original (possivelmente sazonal). (3.19)

Novamente, pode-se demonstrar que o conjunto de variáveis regressoras em (3.18) são assintoticamente mutuamente ortogonais.

Para testar as hipóteses de que as raízes unitárias existem, estimamos (3.18) por mínimos quadrados e, então, comparamos as estatísticas dos testes com os valores críticos obtidos para amostras finitas, baseados em simulações de Monte Carlo. Assim, temos:

### (i) Testes para as raízes 1 e -1:

Para verificar a existência das raízes 1 e -1 correspondentes às frequências 0 e  $\pi$ , respectivamente, faremos o seguinte teste:

$$\begin{aligned} H_0: \pi_k &= 0 \\ H_a: \pi_k &< 0 \end{aligned} , k=1, 2. \quad (3.20)$$

### (ii) Testes para as demais raízes:

Para as demais raízes, testaremos  $\pi_k = 0$  (para k par) através de um teste bilateral.

$$\begin{aligned} H_0: \pi_k &= 0 \\ H_a: \pi_k &\neq 0 \end{aligned} , k = \text{par}. \quad (3.21)$$

O coeficiente par será zero se a série contiver uma raiz unitária na frequência que está sendo testada e, por outro lado, não será zero para as frequências sazonais com exceção da frequência  $\frac{\pi}{2}$ . Para  $\frac{\pi}{2}$ , o coeficiente não será zero se não existir raiz nesta frequência. Sob a alternativa, o coeficiente par pode ser positivo ou negativo. Se o teste não rejeitar  $\pi_k = 0$ , então testaremos  $\pi_{k-1} = 0$  contra a alternativa  $\pi_{k-1} < 0$ . O teste é unilateral porque, sob a hipótese alternativa, a série contém uma raiz fora do círculo unitário. Sob a condição de estacionariedade, o verdadeiro coeficiente será menor do que zero. Outra estratégia é testar  $\pi_{k-1} = \pi_k = 0$  calculando uma estatística “F”. Para concluir que não existem raízes unitárias em frequências sazonais,  $\pi_k$  não deve ser igual a zero para  $k = 2$  e para pelo menos um membro de cada um dos conjuntos  $\{3,4\}$ ,  $\{5,6\}$ ,  $\{7,8\}$ ,  $\{9,10\}$ ,  $\{11,12\}$ . Dessa forma, faremos o seguinte teste:

$$\begin{aligned} H_0: \pi_{k-1} = \pi_k = 0 \\ H_a: \pi_{k-1} \text{ e/ou } \pi_k \neq 0 \end{aligned} \quad (3.22)$$

Quando uma constante, *dummies* sazonais e tendência estão presentes, então a expressão em (3.18) poderá ser re-escrita como:

$$\varphi(B)^* y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + m_0 t + m_1 + \sum_{k=2}^{12} m_k S_{kt} + \varepsilon_t \quad (3.23)$$

A equação (3.23) também será estimada por mínimos quadrados, porém, a distribuição assintótica e a distribuição para amostras finitas se modificarão.

Hasza e Fuller (1982) consideraram, como hipótese nula, o seguinte modelo:

$$(1 - B)(1 - B^d)y_t = \varepsilon_t,$$

em que  $d = 2, 4, 6, 12$  dependendo do número de observações por ano. Eles sugeriram estimar as seguintes equações:

$$\begin{aligned} y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-d} + \alpha_3 y_{t-d-1} + \varepsilon_t, \\ y_t &= \phi_1 y_{t-1} + \phi_2 (y_{t-d} - \phi_1 y_{t-d-1}) + \varepsilon_t \end{aligned}$$

e, então, testar a restrição  $(\alpha_1, \alpha_2, \alpha_3) = (1, 1, -1)$  ou  $(\phi_1, \phi_2) = (1, 1)$ , calculando a estatística “F” padrão e utilizando uma distribuição apropriada.

Os resultados de Hasza e Fuller assumem que, sob a hipótese nula, existem duas raízes unitárias na frequência zero e não é claro identificar como esse teste se modifica quando apenas algumas frequências sazonais possuem raízes unitárias. Por essa razão, a interpretação dos resultados dos testes de Hasza e Fuller não é trivial. A rejeição não prova que não existem raízes em quaisquer frequências, assim como também não ajuda a identificar as frequências que são integradas. Uma alternativa é colocar na hipótese nula que todas as frequências possuem raízes unitárias e duas tem frequência zero. Nesse caso, no teste estará imposta a condição de que todas as raízes sazonais possuem o mesmo módulo. Se o processo que se está avaliando não possui raízes unitárias, mas as raízes nas frequências sazonais tem módulos diferentes, então a autocorrelação do resíduo em diversos lags será significativa e, portanto, o resíduo não será ruído branco. Dessa forma, esse teste é prejudicado pelo fato de não se mostrar muito poderoso e ser passível de apresentar autocorrelação significativa nos resíduos.

Dickey, Hasza e Fuller (1984), por sua vez, não assumiram a existência de duas raízes unitárias na frequência zero, considerando, sob a hipótese nula, o modelo:

$$(1 - B^d)y_t = \varepsilon_t.$$

Eles sugeriram estimar a equação:

$$y_t = \rho y_{t-d} + b_t + \varepsilon_t,$$

e testar se  $\rho = 1$ , onde  $b_t$  ou é uma constante ou são *dummies* sazonais. Eles publicaram a distribuição de várias estatísticas quando a equação de regressão inclui uma constante e *dummies* sazonais. Assim como no teste proposto por Hasza e Fuller, a interpretação da rejeição desse teste também não é trivial e também apresenta baixo poder e autocorrelação residual.

O teste HEGY apresentado evita esses problemas incluindo uma variável regressora correspondente a cada uma das potenciais raízes unitárias.

### 3.5 Distribuição Assintótica das Estatísticas do Teste Mensal

Para o caso de dados mensais, as distribuições assintóticas das estatísticas dos testes HEGY são derivadas de Stock (1988) e Phillips (1987). Considerando que a constante, as *dummies* sazonais e a tendência não estão incluídas no modelo – modelo (3.18) – o resultado 3.2 apresenta a distribuição assintótica das estatísticas “t”.

**Resultado 3.2.** (Beaulieu e Miron, 1993) Seja  $W_k$  um Movimento Browniano Padrão independente de todos os outros  $W_j$ , para  $k \neq j$ . Para facilitar a notação diremos que  $k$  é ímpar se  $k \neq 1$  e  $k \in \{3, 5, 7, 9, 11\}$  e que  $k$  é par se  $k \neq 2$  e  $k \in \{4, 6, 8, 10, 12\}$ . Então, as estatísticas  $t_k$ , onde  $t_k = \frac{\sum_{t=2}^T y_{kt-1} \varepsilon_t}{\hat{\sigma} \sqrt{\sum_{t=2}^T y_{kt-1}^2}}$  convergem em distribuição ( $\overset{D}{\rightarrow}$ ) para

$$\frac{\int_0^1 W_1(r) dW_1(r)}{\sqrt{\int_0^1 W_1(r)^2 dr}}, \quad \text{se } k = 1,$$

$$\frac{\int_0^1 W_2(r) dW_2(r)}{\sqrt{\int_0^1 W_2(r)^2 dr}}, \quad \text{se } k = 2,$$

$$\frac{\int_0^1 W_k(r) dW_k(r) + \int_0^1 W_{k+1}(r) dW_{k+1}(r)}{\sqrt{\int_0^1 W_k(r)^2 dr + \int_0^1 W_{k+1}(r)^2 dr}}, \quad \text{se } k \text{ é ímpar},$$

$$\frac{\int_0^1 W_k(r) dW_{k-1}(r) - \int_0^1 W_{k-1}(r) dW_k(r)}{\sqrt{\int_0^1 W_{k-1}(r)^2 dr + \int_0^1 W_k(r)^2 dr}}, \quad \text{se } k \text{ é par}.$$

Assim, temos que  $t_1$  e  $t_2$  tem a mesma distribuição assintótica, que todas as estatísticas sazonais pares tem a mesma distribuição e que todas as estatísticas sazonais ímpares também tem a mesma distribuição.

A construção de  $\pi_1$  e  $\pi_2$  é exatamente a mesma de Franses (1990). As distribuições das estatísticas “t” sazonais ímpares são iguais à distribuição de  $t_3$  no teste HEGY trimestral, enquanto as distribuições das correspondentes estatísticas pares são as mesmas de  $t_4$  no teste HEGY trimestral. A construção de  $\pi_3$  e  $\pi_4$  em Franses (1990) também é a mesma para o teste mensal. Como Chan e Wei (1988) provaram, a distribuição de  $t_3$  é a mesma para  $\hat{t}_d$  com  $d = 2$  em Dickey-Hasza-Fuller (1984).

Considerando que termos determinísticos são incluídos na regressão, a partir do resultado 3.3, obtemos as distribuições das estatísticas “t”.

**Resultado 3.3.** (Beaulieu e Miron, 1993) Seja  $y_{kt}^\mu$  o resíduo da regressão de  $y_{kt}$  sobre uma constante. Sejam  $y_{kt}^\tau$ ,  $y_{kt}^\varepsilon$  e  $y_{kt}^{\varepsilon\tau}$ , analogamente, onde  $\tau$  representa a constante mais a tendência,  $\varepsilon$  representa a constante mais as onze dummies e  $\varepsilon\tau$  representa a constante, as onze dummies e a tendência. Os numeradores e denominadores das respectivas regressões estão parcialmente definidos no Resultado 3.2. Seja  $N_k^x$  a parte do numerador diferente do correspondente numerador no Resultado 3.2,  $x \in \{\mu, \tau, \varepsilon, \varepsilon\tau\}$  e seja  $D_k^x$  a parte da raiz do denominador diferente do correspondente denominador no Resultado 3.2.

$$(\mu) N_k^\mu = -W_1(1) \int_0^1 W_1(r) dr, \quad \text{se } k = 1,$$

$$N_k^\mu = 0, \quad \text{se } k \geq 2.$$

$$(\tau) N_k^\tau = -4 W_1(1) \int_0^1 W_1(r) dr + 6 \int_0^1 W_1(r) dr \int_0^1 r dW_1(r) - 12 \int_0^1 W_1(r) dr \int_0^1 r dW_1(r) + \\ + 6 W_1(1) \int_0^1 W_1(r) dr, \quad \text{se } k = 1,$$

$$N_k^\tau = 0, \quad \text{se } k \geq 2.$$

$$(\varepsilon) N_k^\varepsilon = -W_1(1) \int_0^1 W_1(r) dr, \quad \text{se } k = 1,$$

$$N_k^\varepsilon = -W_2(1) \int_0^1 W_2(r) dr, \quad \text{se } k = 2,$$

$$N_k^\varepsilon = -\frac{1}{2} W_k(1) \int_0^1 W_k(r) dr - \frac{1}{2} W_{k+1}(1) \int_0^1 W_{k+1}(r) dr, \quad \text{se } k \text{ é ímpar},$$

$$N_k^\varepsilon = -\frac{1}{2} W_k(1) \int_0^1 W_{k-1}(r) dr + \frac{1}{2} W_{k-1}(1) \int_0^1 W_k(r) dr, \quad \text{se } k \text{ é par}.$$

$$(\varepsilon\tau) N_k^{\varepsilon\tau} = -4 W_1(1) \int_0^1 W_1(r) dr + 6 \int_0^1 W_1(r) dr \int_0^1 r dW_1(r) - 12 \int_0^1 r W_1(r) dr \int_0^1 r dW_1(r) + \\ + 6 W_1(1) \int_0^1 r W_1(r) dr, \quad \text{se } k = 1,$$

$$N_k^{\varepsilon\tau} = -W_2(1) \int_0^1 W_2(r) dr, \quad \text{se } k = 2,$$

$$N_k^{\varepsilon\tau} = -\frac{1}{2} W_k(1) \int_0^1 W_k(r) dr - \frac{1}{2} W_{k+1}(1) \int_0^1 W_{k+1}(r) dr, \quad \text{se } k \text{ é ímpar},$$

$$N_k^{\varepsilon\tau} = -\frac{1}{2} W_k(1) \int_0^1 W_{k-1}(r) dr + \frac{1}{2} W_{k-1}(1) \int_0^1 W_k(r) dr, \quad \text{se } k \text{ é par}.$$

$$(\mu) D_k^\mu = -\left(\int_0^1 W_1(r) dr\right)^2, \quad \text{se } k = 1,$$

$$D_k^\mu = 0, \quad \text{se } k \geq 2.$$

$$(\tau) D_k^\tau = -4\left(\int_0^1 W_1(r) dr\right)^2 + 12 \int_0^1 W_1(r) dr \int_0^1 r W_1(r) dr - 12 \left(\int_0^1 r W_1(r) dr\right)^2, \quad \text{se } k = 1,$$

$$D_k^\tau = 0, \quad \text{se } k \geq 2.$$

$$(\epsilon) D_k^\epsilon = -\left(\int_0^1 W_1(r) dr\right)^2, \quad \text{se } k = 1,$$

$$D_k^\epsilon = -\left(\int_0^1 W_2(r) dr\right)^2, \quad \text{se } k = 2,$$

$$D_k^\epsilon = -\frac{1}{4}\left(\int_0^1 W_k(r) dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_{k+1}(r) dr\right)^2, \quad \text{se } k \text{ é ímpar},$$

$$D_k^\epsilon = -\frac{1}{4}\left(\int_0^1 W_{k-1}(r) dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_k(r) dr\right)^2, \quad \text{se } k \text{ é par}.$$

$$(\epsilon\tau) D_k^{\epsilon\tau} = -4\left(\int_0^1 W_1(r) dr\right)^2 + 12 \int_0^1 W_1(r) dr \int_0^1 r W_1(r) dr - 12\left(\int_0^1 r W_1(r) dr\right)^2, \quad \text{se } k = 1,$$

$$D_k^{\epsilon\tau} = -\left(\int_0^1 W_2(r) dr\right)^2, \quad \text{se } k = 2,$$

$$D_k^{\epsilon\tau} = -\frac{1}{4}\left(\int_0^1 W_k(r) dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_{k+1}(r) dr\right)^2, \quad \text{se } k \text{ é ímpar},$$

$$D_k^{\epsilon\tau} = -\frac{1}{4}\left(\int_0^1 W_{k-1}(r) dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_k(r) dr\right)^2, \quad \text{se } k \text{ é par}.$$

**Resultado 3.4.** (Beaulieu e Miron, 1993) As estatísticas  $t_k = \frac{\sum_{t=2}^T y_{kt}^x \epsilon_t^x}{\hat{\sigma} \sqrt{\sum_{t=2}^T y_{kt-1}^x}}$  convergem em distribuição para

$$\frac{\int_0^1 W_1(r) dW_1(r) + N_1^x}{\sqrt{\int_0^1 W_1(r)^2 dr + D_1^x}}, \quad \text{se } k = 1,$$

$$\frac{\int_0^1 W_2(r) dW_2(r) + N_2^x}{\sqrt{\int_0^1 W_2(r)^2 dr + D_2^x}}, \quad \text{se } k = 2,$$

$$\frac{\frac{1}{2} \int_0^1 W_k(r) dW_k(r) + \frac{1}{2} \int_0^1 W_{k+1}(r) dW_{k+1}(r) + N_k^x}{\sqrt{\frac{1}{4} \left( \int_0^1 W_k(r)^2 dr + \int_0^1 W_{k+1}(r)^2 dr + D_k^x \right)}}, \quad \text{se } k \text{ é ímpar},$$

$$\frac{\frac{1}{2} \int_0^1 W_k(r) dW_{k-1}(r) - \frac{1}{2} \int_0^1 W_{k-1}(r) dW_k(r) + N_k^x}{\sqrt{\frac{1}{4} \left( \int_0^1 W_{k-1}(r)^2 dr + \int_0^1 W_k(r)^2 dr + D_k^x \right)}}, \quad \text{se } k \text{ é par},$$

em que  $x = \mu, \tau, \varepsilon, \varepsilon\tau$ .

A partir do Resultado 3.4 temos que todas as estatísticas, para  $k$  par ou ímpar, tem as mesmas distribuições quando diferentes termos determinísticos são incluídos na regressão. As distribuições de  $t_2, \dots, t_{12}$  independem da constante e dos termos de tendência. Isso se justifica pelo fato de que os termos de  $y_{kt}$  ( $k \geq 2$ ) podem ser escritos como funções de cossenos e senos, que se repetem a cada 12 períodos. Esses termos são assintoticamente ortogonais aos termos que não são periódicos, tais como uma constante ou uma tendência. Nota-se também que a distribuição de  $t_2$  quando *dummies* são incluídas na regressão é a mesma de  $t_1$  quando somente uma constante é incluída no modelo.

**Resultado 3.5.** (Beaulieu e Miron, 1993) Seja  $F_k = \left( \frac{1}{2\sigma^2} \right) b' (X'X) b$ , com  $b = \begin{pmatrix} \hat{\beta}_k \\ \hat{\beta}_{k+1} \end{pmatrix}$  e  $X =$

$$\begin{bmatrix} y_{k,1} & y_{k+1,1} \\ \vdots & \vdots \\ y_{k,T} & y_{k+1,T} \end{bmatrix}. \text{ A partir dos Resultados 3.2, 3.3 e 3.4 temos que}$$

$$F_k^x \stackrel{D}{\rightarrow} \frac{1}{2} (t_k^{x2} + t_{k+1}^{x2}),$$

em que  $x = \mu, \tau, \varepsilon, \varepsilon\tau$  e  $k \in \{3, 5, 7, 9, 11\}$ .

Note que devido a todas as estatísticas “t” ímpares terem a mesma distribuição e todas as estatísticas pares também terem a mesma distribuição, então para qualquer conjunto de termos determinísticos incluídos, todas as estatísticas “F” também terão a mesma distribuição.

### 3.6 Valores Críticos para Amostras Finitas das Estatísticas do Teste HEGY Mensal

A Tabela 3.3 contem os valores críticos das distribuições de amostras finitas das estatísticas “t” e “F” necessárias ao procedimento HEGY para dados mensais. Os valores críticos foram obtidos simulando 24.000 regressões da forma (3.16), com várias combinações de constantes, *dummies* sazonais e tendências incluídas. As séries fundamentais foram geradas considerando  $y_t = y_{t-12} + e_t$ , com  $e_t$  proveniente de uma normal padrão.

Tabela 3.3 – Valores críticos para as estatísticas do teste HEGY Mensal

Critical values from the distributions of test statistics for seasonal unit roots; data-generating process $\Delta_{12}x_t = \varepsilon_t$ , NID (0, 1).													
Auxiliary regressions	T	‘t’: $\pi_1$				Fractiles ‘t’: $\pi_2$				‘t’: $\pi_{\text{odd}}$			
		0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10
No intercept	240	-2.51	-2.18	-1.89	-1.58	-2.53	-2.16	-1.87	-1.57	-2.50	-2.16	-1.88	-1.55
No seas.dum.	480	-2.52	-2.21	-1.91	-1.59	-2.52	-2.20	-1.91	-1.59	-2.52	-2.18	-1.90	-1.57
No trend	$\infty$	-2.57	-2.24	-1.95	-1.62	-2.57	-2.24	-1.95	-1.62	-2.56	-2.23	-1.95	-1.59
Intercept	240	-3.35	-3.06	-2.80	-2.51	-2.48	-2.15	-1.89	-1.57	-2.51	-2.16	-1.87	-1.54
No seas.dum.	480	-3.40	-3.11	-2.85	-2.55	-2.54	-2.20	-1.91	-1.59	-2.56	-2.20	-1.90	-1.57
No trend	$\infty$	-3.41	-3.12	-2.86	-2.57	-2.57	-2.24	-1.95	-1.62	-2.56	-2.23	-1.95	-1.59
Intercept	240	-3.32	-3.02	-2.76	-2.47	-3.28	-3.01	-2.76	-2.48	-3.83	-3.51	-3.25	-2.95
Seas.dum.	480	-3.37	-3.06	-2.81	-2.53	-3.37	-3.07	-2.81	-2.52	-3.86	-3.55	-3.29	-2.99
No trend	$\infty$	-3.41	-3.12	-2.86	-2.57	-3.41	-3.12	-2.86	-2.57	-3.91	-3.61	-3.35	-3.05
Intercept	240	-3.87	-3.58	-3.32	-3.06	-2.52	-2.18	-1.88	-1.55	-2.49	-2.16	-1.88	-1.54
No seas.dum.	480	-3.92	-3.63	-3.37	-3.09	-2.55	-2.20	-1.93	-1.60	-2.53	-2.20	-1.91	-1.57
Trend	$\infty$	-3.97	-3.67	-3.40	-3.12	-2.57	-2.24	-1.95	-1.62	-2.56	-2.23	-1.95	-1.59
Intercept	240	-3.83	-3.54	-3.28	-2.99	-3.31	-3.02	-2.75	-2.47	-3.79	-3.50	-3.24	-2.95
Seas.dum.	480	-3.85	-3.57	-3.32	-3.04	-3.40	-3.08	-2.84	-2.54	-3.85	-3.55	-3.29	-3.00
Trend	$\infty$	-3.97	-3.67	-3.40	-3.12	-3.41	-3.12	-2.86	-2.57	-3.91	-3.61	-3.35	-3.05

Tabela extraída de Beaulieu e Miron (1993)

Tabela 3.3 – Continuação dos Valores críticos para as estatísticas do teste HEGY Mensal

Auxiliary regressions	T	Fractiles											
		$t'$ : $\pi_{\text{even}}$						$F'$ : $\pi_{\text{odd}}, \pi_{\text{even}}$					
		0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99	0.90	0.95	0.975	0.99
No intercept	240	-2.31	-1.95	-1.63	-1.27	1.25	1.61	1.93	2.29	2.34	3.03	3.71	4.60
No seas.dum.	480	-2.33	-1.96	-1.65	-1.28	1.27	1.63	1.94	2.32	2.38	3.08	3.78	4.70
No trend	$\infty$	-2.30	-1.94	-1.63	-1.28	1.27	1.63	1.94	2.32	2.40	3.10	3.79	4.68
Intercept	240	-2.30	-1.93	-1.62	-1.27	1.24	1.60	1.91	2.28	2.32	3.01	3.68	4.60
No seas.dum.	480	-2.32	-1.95	-1.63	-1.27	1.27	1.62	1.93	2.30	2.36	3.06	3.76	4.66
No trend	$\infty$	-2.30	-1.94	-1.63	-1.28	1.27	1.63	1.94	2.32	2.40	3.10	3.79	4.68
Intercept	240	-2.61	-2.21	-1.85	-1.45	1.46	1.86	2.20	2.60	5.27	6.26	7.19	8.35
Seas.dum.	480	-2.65	-2.25	-1.90	-1.49	1.49	1.91	2.25	2.63	5.42	6.42	7.38	8.60
No trend	$\infty$	-2.72	-2.31	-1.95	-1.54	1.53	1.95	2.30	2.72	5.64	6.67	7.63	8.79
Intercept	240	-2.28	-1.93	-1.61	-1.25	1.24	1.59	1.90	2.26	2.30	2.97	3.64	4.53
No seas.dum.	480	-2.30	-1.94	-1.63	-1.27	1.25	1.61	1.92	2.28	2.36	3.05	3.72	4.62
Trend	$\infty$	-2.30	-1.94	-1.63	-1.28	1.27	1.63	1.94	2.32	2.40	3.10	3.79	4.68
Intercept	240	-2.57	-2.18	-1.85	-1.45	1.45	1.86	2.19	2.60	5.25	6.23	7.14	8.33
Seas.dum.	480	-2.66	-2.27	-1.91	-1.49	1.49	1.90	2.25	2.64	5.44	6.43	7.35	8.52
Trend	$\infty$	-2.72	-2.31	-1.95	-1.54	1.53	1.95	2.30	2.72	5.64	6.67	7.63	8.79

Tabela extraída de Beaulieu e Miron (1993)

Os valores críticos para  $t_{\text{par}}$ ,  $t_{\text{impar}}$  e F foram calculados combinando as observações sobre todas as cinco estatísticas similares. Por exemplo,  $t_{\text{impar}}$  foi calculado empilhando  $t_3$ ,  $t_5$ ,  $t_7$ ,  $t_9$  e  $t_{11}$  e calculando a estatística de ordem do vetor de dimensão (120.000 X 1). Assim,  $t_1$  e  $t_2$  são baseados em 24.000 observações, enquanto  $t_{\text{par}}$ ,  $t_{\text{impar}}$  e F são baseados em 120.000 observações. Esse procedimento é eficiente porque as distribuições assintóticas dessas estatísticas são as mesmas para cada um dos cinco pares de coeficientes. Além disso, investigações das distribuições amostrais indicaram que essas distribuições são similares para um dado número de simulações e que convergem quando o número de simulações aumenta, reforçando a eficiência desse procedimento.

Para calcular os valores críticos assintóticos das estatísticas do teste HEGY mensal, Beaulieu e Miron (1993) aproximaram as funções do movimento browniano usando somas parciais de variáveis aleatórias normais. Por exemplo,  $W(1)$  foi aproximado por:

$$W(1) = \frac{1}{\sqrt{5000}} \sum_{j=1}^{5000} \varepsilon_j$$

em que os  $\varepsilon_j$ 's são variáveis aleatórias independentes normais padrão. Como as distribuições das estatísticas convergem quando o número de simulações é grande, então é conveniente repetir esse procedimento muitas vezes (nesse caso, o procedimento foi repetido cem mil vezes). Para mais detalhes sobre os valores críticos das estatísticas do teste HEGY mensal, ver Beaulieu e Miron (1993).

A Tabela 3.3 indica que as distribuições das amostras finitas tem todas as características das distribuições assintóticas que descrevemos. Primeiramente, as distribuições de cada um dos onze coeficientes sazonais não são afetadas pela presença de uma constante ou de uma constante com uma tendência – esses termos somente afetam a distribuição de  $\pi_1$ . Um segundo ponto a ser observado é que a distribuição de  $\pi_1$  não será afetada pela presença de *dummies* sazonais, contanto que uma constante esteja incluída na regressão. Por fim, podemos observar também que a distribuição de  $t_2$  quando *dummies* estão incluídas é similar a

de  $t_1$ , quando uma constante está incluída. Portanto, os resultados na Tabela 3.3 indicam que não há grandes diferenças entre amostras finitas e distribuições assintóticas.

Os erros padrões dos valores críticos estimados são baseados na distribuição assintótica das estatísticas de ordem normalizadas (Bickel and Doksum (1977, p. 400)). No cálculo dos erros padrões, em Beaulieu e Miron (1993), foi utilizada uma estimativa da densidade através do kernel Epanechnikov com amplitude igual a  $(40\sqrt{\pi})^{0.2}N^{-0.2}$  ( $N = 24.000, 120.000$  ou  $100.000$ ). Os erros padrões dos valores críticos estimados utilizando amostras finitas foram menores ou iguais a 0,02 para  $t_1$  e  $t_2$ , menores ou iguais a 0,01 para  $t_{impar}$  e  $t_{par}$  e menores ou iguais a 0,03 para F. Os erros padrões dos valores críticos assintóticos estimados foram menores ou iguais a 0,01 para as estatísticas “t” e menores ou iguais a 0,03 para as estatísticas “F”.

Vale citar que um problema derivado dos testes HEGY é o tratamento da autocorrelação do resíduo de  $x_t$ . Se  $\varphi(B)$  tiver ordem superior a S (onde S é o número de observações por ano), então  $\varphi^*(B) \neq 1$  e, portanto, lags adicionais de  $y_{13t}$  devem ser incluídos do lado direito da expressão em (3.16). Esses lags extras não afetam a distribuição assintótica das estatísticas do teste (contanto que o resto do componente AR de  $x_t$  seja estimado corretamente), mas afetam as distribuições das amostras finitas. Em particular, se o modelo tiver uma componente MA, então o número de lags corretos em  $\varphi(B)$  é infinito. Isso implica em um viés de amostra finita que desaparece quando o número de lags cresce para infinito.

## Capítulo 4

### O Teste HEGY Mensal / Trimestral

Neste Capítulo apresentaremos um procedimento resultante da combinação dos testes de raízes unitárias sazonais para dados trimestrais e mensais. Essa nova técnica utiliza as raízes unitárias trimestrais rejeitadas para reduzir o conjunto de frequências mensais nas quais as raízes unitárias deverão ser testadas. Neste Capítulo apresentaremos, também, a distribuição assintótica das estatísticas do teste HEGY Mensal / Trimestral.

#### 4.1 O Procedimento de Pons

Trata-se de um método para testar raízes unitárias sazonais que combina os testes HEGY mensal e trimestral. Esse método foi proposto por Pons (2005) e se baseia nos efeitos da agregação temporal sobre os modelos sazonais. Ele foi construído a partir dos estudos de Rodrigues e Franses (2003), estudos estes que basicamente utilizam as raízes unitárias trimestrais rejeitadas da série agregada ( $X_T^a$ ) para reduzir o conjunto de frequências mensais nas quais serão testadas as raízes unitárias. Veremos que o procedimento de Pons se apresenta mais poderoso do que o teste HEGY mensal.

A agregação temporal tem implicações importantes para análise de séries temporais sazonais, uma vez que dados transformados confundem ciclos sazonais não observados com ciclos sazonais observados. Esse efeito conhecido como confundimento (Koopmans, 1974) implica em não ser possível concluir se uma raiz unitária está presente na frequência mensal com o mesmo período trimestral – quando uma particular raiz unitária trimestral é detectada – ou em outra frequência mensal com um período não observado no intervalo trimestral. Isso não ocorre somente para componentes sazonais, mas também para componentes de tendência, uma vez que uma raiz unitária na frequência zero trimestral pode aparecer quando amostramos uma raiz unitária em dados mensais na frequência  $\frac{2\pi}{3}$  (Granger e Siklos, 1995).

Em alguns casos, a teoria econômica pode nos direcionar a unidade de tempo correta que permite identificar os ciclos sazonais da série de interesse, porém, em muitos casos não há a priori uma explicação teórica sobre as ocorrências de particulares ciclos sazonais. Se os dados estão disponíveis no intervalo onde todos os ciclos sazonais são gerados, então a sazonalidade é passível de ser identificada. Entretanto, essa não é uma situação típica dos testes de raízes unitárias sazonais, normalmente aplicados para dados mensais ou séries temporais macroeconômicas trimestrais.

Primeiramente, vamos considerar as implicações que a amostragem sistemática trimestral causa em testes de raízes unitárias sazonais. Para esse propósito, discutiremos as relações entre modelos sazonais mensais e trimestrais usados como hipóteses nula e alternativa dos testes HEGY.

A Tabela 4.1 apresenta a relação entre ciclos sazonais mensais e trimestrais. As frequências trimestrais  $\frac{\pi K}{2}$  estão relacionadas com as frequências mensais  $\frac{\pi k}{6}$  associadas com o mesmo período, isto é, o mesmo número de ciclos por ano ( $K = k$ ). A partir de  $k = 3$ , temos ciclos mensais não observados em intervalos trimestrais com seu período real ( $k = 3, 4, 5, 6$ ), mas sim com um período maior ( $K = 1, 0, 1, 2$ ).

Tabela 4.1 - Relações entre ciclos mensais e trimestrais (extraída de Pons, 2005)

Intervalo Mensal		Intervalo Trimestral	
Frequência ( $\frac{\pi k}{6}$ )	Ciclos por ano (k)	Frequência ( $\frac{\pi K}{2}$ )	Ciclos por ano (K)
0	0	0	0
$\frac{\pi}{6}$	1	$\frac{\pi}{2}$	1
$\frac{\pi}{3}$	2	$\pi$	2
$\frac{\pi}{2}$	3	$\frac{\pi}{2}$	1
$\frac{2\pi}{3}$	4	0	0
$\frac{5\pi}{6}$	5	$\frac{\pi}{2}$	1
$\pi$	6	$\pi$	2

## 4.2 Efeitos da Agregação Temporal

### 4.2.1 Efeito da Agregação Temporal Trimestral sobre o Processo Sazonal Determinístico Mensal

Consideremos amostras trimestrais de um modelo sazonal determinístico mensal (Miron, 1996), cuja representação utilizando variáveis *dummies* sazonais é dada por:

$$x_t = \sum_{j=1}^{12} \psi_j d_{j,t} + u_t \quad (4.1)$$

em que  $d_{j,t}$  é a variável *dummy* sazonal para o mês  $j$ ,  $\psi_j$  é a média de  $x_t$  no  $j$ -ésimo mês e  $u_t$  é  $RB(0, \sigma^2)$ .

Os testes para raízes unitárias sazonais são baseados em uma representação diferente de sazonalidade – a representação espectral (veja Hannan et al., 1970), que para o modelo expresso em (4.1) é dada por:

$$x_t = \gamma_0 + \sum_{k=1}^5 \left[ \gamma_k \cos\left(\frac{\pi k}{6} t\right) + \lambda_k \sin\left(\frac{\pi k}{6} t\right) \right] + \gamma_6 \cos(\pi t) + u_t$$

em que  $\gamma_0$  é a média geral de  $x_t$ ,  $\gamma_k$  e  $\lambda_k$  medem a contribuição do ciclo sazonal, associado com a frequência  $\pm \frac{\pi k}{6}$ , na média sazonal incondicional.

➤ **Processos Trimestrais gerados por Amostragem Sistemática de Processos Mensais**

Há três diferentes processos trimestrais que podem ser gerados por amostragem sistemática de processos mensais. Esses processos trimestrais serão denotados por:

$$X_T^s = x_{3(T-1)+s} \quad (s = 1, 2, 3).$$

No caso de sazonalidade determinística, os três processos  $X_T^s$  dependem de diferentes parâmetros mensais. A representação, utilizando variável *dummy*, para  $X_T^s$  ( $s = 1, 2, 3$ ) é dada por:

$$X_T^s = \sum_{J=1}^4 \Psi_J^s D_{J,T} + U_T^s$$

$D_{J,T}$  é a variável *dummy* sazonal referente ao trimestre J,

$\Psi_J^s = \psi_{3(J-1)+s}$  ( $s = 1, 2, 3$ ), tal que cada  $X_T^s$  absorve médias sazonais diferentes de um processo mensal e  $U_T^s$  é RB.

A representação espectral de  $X_T^s$  ( $s = 1, 2, 3$ ) é dada por:

$$X_T^s = \Gamma_0^s + \Gamma_1^s \cos\left(\frac{\pi}{2} T\right) + \Lambda_1^s \sin\left(\frac{\pi}{2} T\right) + \Gamma_2^s \cos(\pi T) + U_T^s,$$

em que as médias trimestrais gerais são dadas por:

$$\Gamma_0^1 = \gamma_0 + \frac{1}{2}(\sqrt{3}\lambda_4 - \gamma_4),$$

$$\Gamma_0^2 = \gamma_0 - \frac{1}{2}(\gamma_4 + \sqrt{3}\lambda_4) \text{ e}$$

$$\Gamma_0^3 = \gamma_0 + \gamma_4,$$

os coeficientes associados com as frequências trimestrais  $\pm\frac{\pi}{2}$  são dados por:

$$\Gamma_1^1 = \frac{1}{2}(\gamma_1 - \sqrt{3}\lambda_1 - 2\gamma_3 + \gamma_5 + \sqrt{3}\lambda_5),$$

$$\Gamma_1^2 = \frac{1}{2}(\sqrt{3}\gamma_1 - \lambda_1 - 2\lambda_3 - \sqrt{3}\gamma_5 - \lambda_5),$$

$$\Gamma_1^3 = \gamma_1 + \gamma_3 + \gamma_5,$$

$$\Lambda_1^1 = \frac{1}{2}(\sqrt{3}\gamma_1 + \lambda_1 + 2\lambda_3 - \sqrt{3}\gamma_5 + \lambda_5),$$

$$\Lambda_1^2 = \frac{1}{2}(\gamma_1 + \sqrt{3}\lambda_1 - 2\gamma_3 + \gamma_5 - \sqrt{3}\lambda_5),$$

$$\Lambda_1^3 = \lambda_1 + \lambda_3 + \lambda_5$$

e os coeficientes da componente trimestral associados com a frequência  $\pi$  são dados por:

$$\Gamma_2^1 = \frac{1}{2}(-\gamma_2 - \sqrt{3}\lambda_2 + 2\gamma_6),$$

$$\Gamma_2^2 = \frac{1}{2}(\gamma_2 - \sqrt{3}\lambda_2 - 2\gamma_6),$$

$$\Gamma_2^3 = \gamma_2 + 2\gamma_6, U_T^s = u_{3(T-1)+s}.$$

A dependência dos coeficientes espectrais trimestrais  $\Gamma_j^s$ ,  $\Lambda_1^s$  ( $j = 0, 1, 2$ ;  $s = 1, 2, 3$ ) em relação aos coeficientes espectrais mensais  $\gamma_l$  ( $l = 0, \dots, 6$ ),  $\lambda_l$  ( $l = 1, \dots, 5$ ) é determinada pela relação entre ciclos sazonais mensais e trimestrais expressa na Tabela 4.1.

### ➤ Processos Trimestrais gerados por Amostragem Agregada de Processos Mensais

Consideremos o processo gerado por agregação de dados dado por:

$$X_T^a = \sum_{s=1}^3 X_T^s \quad (4.2)$$

tal que, dada a equação (4.1), pode ser representado por:

$$X_T^a = \sum_{j=1}^4 \Psi_j^a D_{j,T} + U_T^a, \quad \text{em que } \Psi_j^a = \sum_{s=1}^3 \psi_j^s$$

ou por:

$$X_T^a = \Gamma_0^a + \Gamma_1^a \cos\left(\frac{\pi}{2}T\right) + \Lambda_1^a \sin\left(\frac{\pi}{2}T\right) + \Gamma_2^a \cos(\pi T) + U_T^a$$

em que

$$\Gamma_j^a = \sum_{s=1}^3 \Gamma_j^s, \quad j = 0, 1, 2 \quad \text{e} \quad \Lambda_1^a = \sum_{s=1}^3 \Lambda_1^s.$$

É importante enfatizar que os processos trimestrais  $X_T^1$ ,  $X_T^2$  e  $X_T^3$  associados com os processos determinísticos sazonais mensais podem ter padrões diferentes de sazonalidade e isso não ocorre com processos integrados sazonalmente. Por exemplo, quando  $\psi_1 = \psi_7$  e  $\psi_4 = \psi_{10}$ ,  $X_T^1$  não possui um ciclo sazonal anual. O mesmo ocorre quando  $\gamma_2 = -2\gamma_6$  e, nesse caso,  $X_T^3$  não possui nenhum ciclo sazonal semi-anual.

O padrão de sazonalidade do processo mensal é separado em três processos amostrados trimestrais e, cada um deles,  $X_T^1$ ,  $X_T^2$ ,  $X_T^3$  ou  $X_T^a$ , conterá somente parte da informação referente à componente sazonal mensal.

#### 4.2.2 Efeito da Agregação Temporal Trimestral sobre o Processo Sazonal Integrado Mensal

**Teorema 4.1.** *Seja  $x_t$  um processo sazonal integrado mensal:*

$$(1 - B)^{d_0} (1 - \sqrt{3}B + B^2)^{d_1} (1 - B + B^2)^{d_2} (1 + B^2)^{d_3} (1 + B + B^2)^{d_4} \\ (1 + \sqrt{3}B + B^2)^{d_5} (1 + B)^{d_6} x_t = u_t$$

em que  $d_k = 0, 1$  e  $u_t$  é um processo autorregressivo e de médias móveis (ARMA) estacionário e invertível. Os processos trimestrais amostrados sistematicamente  $X_T^s$  ( $s = 1, 2, 3$ ) são processos integrados sazonalmente tais que:

$$(1 - B^3)^{D_0}(1 + B^6)^{D_1}(1 + B^3)^{D_2}X_T^s = U_T^s,$$

em que

$$B^3 X_T^s = X_{T-1}^s,$$

$$D_0 = \max\{d_0, d_4\},$$

$$D_1 = \max\{d_1, d_3, d_5\},$$

$$D_2 = \max\{d_2, d_6\} \text{ e}$$

$U_T^s = u_{3(T-1)+s}$  é um processo ARMA estacionário e invertível.

**Corolário 4.1.** O processo amostral trimestralmente agregado,  $X_T^a$ , é o processo sazonal integrado

$$(1 - B^3)^{D_0}(1 + B^6)^{D_1}(1 + B^3)^{D_2}X_T^a = U_T^a,$$

em que

$$D_0 = d_0, D_1 = \max\{d_1, d_3, d_5\},$$

$$D_2 = \max\{d_2, d_6\} \text{ e}$$

$U_T^a = \sum_{s=1}^3 U_T^s$  é um processo ARMA estacionário e invertível.

O Teorema 4.1 mostra que a amostragem sistemática distribui as doze raízes unitárias mensais nos três processos trimestrais amostrados  $X_T^s$ ,  $s = 1, 2$  e  $3$ . A ordem de integração na frequência zero ( $D_0$ ) dos processos amostrados está relacionada à ordem de integração na frequência zero do processo mensal ( $d_0$ ) e à ordem de integração na frequência  $\frac{2\pi}{3}$  ( $d_4$ ). Além disso, a raiz unitária real positiva mensal e o par de raízes unitárias complexas conjugadas mensais se transformam em três diferentes raízes unitárias na frequência zero trimestral depois da amostragem sistemática. Entretanto, como vemos no Corolário 4.1, a ordem de integração na frequência zero de  $X_T^a$  está relacionada somente com a ordem de integração da frequência zero mensal ( $d_0$ ), uma vez que o filtro soma remove as raízes unitárias das frequências  $\pm \frac{2\pi}{3}$  de  $x_t$ .

No caso das ordens  $D_1$  e  $D_2$  de integração sazonal, a relação é a mesma para o processo obtido por amostragem sistemática e para o processo obtido por agregação. A raiz unitária na frequência trimestral  $\frac{\pi}{2}$  ( $D_1$ ) está relacionada às raízes unitárias mensais na frequência  $\frac{\pi}{6}$  ( $d_1$ ), o ciclo mensal com o mesmo período que o ciclo trimestral e as raízes

unitárias nas frequências  $\frac{\pi}{2}$  ( $d_3$ ) e  $\frac{5\pi}{6}$  ( $d_5$ ). Finalmente, a raiz unitária na frequência  $\pi$  trimestral está relacionada com as raízes unitárias na frequência  $\pi$  e  $\frac{\pi}{3}$  mensal.

As relações entre as ordens de integração possibilitam a interpretação dos testes trimestrais em termos dos testes mensais e a contrapartida também é verdadeira. Assim, temos que o teste trimestral pode ser utilizado para testar a presença de algumas raízes unitárias mensais associadas com as frequências trimestrais. Portanto, considerando os testes HEGY, a rejeição de uma particular raiz unitária trimestral implica na rejeição de todas as raízes unitárias mensais associadas à essa raiz trimestral, enquanto a não rejeição da raiz unitária trimestral implica na não rejeição de pelo menos uma das raízes unitárias mensais associadas. Por exemplo, a rejeição de uma raiz unitária na frequência  $\pi$  trimestral implica na rejeição da integração nas frequências mensais  $\frac{\pi}{3}$  e  $\pi$ .

Por fim, de forma análoga, temos que o teste mensal pode ser utilizado como um teste trimestral. Por exemplo, se não rejeitamos a presença de uma raiz unitária mensal na frequência  $\pi$ , então não rejeitaremos uma raiz unitária na frequência  $\pi$  trimestral nas séries trimestrais  $X_T^s$  ( $s = 1, 2, 3$ ) e na série  $X_T^a$ .

Para as demonstrações do Teorema e Corolário 4.1, sugerimos ler o autor Pons (2005).

### 4.3 Estatísticas e correspondentes Valores Críticos

#### 4.3.1 Teste HEGY Trimestral – Obtenção das Estatísticas do Teste para as Séries Trimestrais $X_T^1$ , $X_T^2$ e $X_T^3$

Primeiramente, vamos definir as estatísticas dos testes.

- $T_0^s$  ( $s = 1, 2, 3$ ): estatísticas “t” para a hipótese de existir uma raiz unitária na frequência zero de  $X_T^s$  ( $s = 1, 2, 3$ );
- $F_1^s$  ( $s = 1, 2, 3$ ): estatísticas “F” para a hipótese de existir o par de raízes unitárias complexas conjugadas nas frequências  $\pm \frac{\pi}{2}$  de  $X_T^s$  ( $s = 1, 2, 3$ );
- $T_2^s$  ( $s = 1, 2, 3$ ): estatísticas t para a hipótese de existir uma raiz unitária na frequência  $\pi$  de  $X_T^s$  ( $s = 1, 2, 3$ ).

As estatísticas trimestrais são definidas por:

$$\begin{aligned}
T_0^I &= \max_s(|T_0^s|), & T_0^{II} &= \text{med}_s(|T_0^s|), & T_0^{III} &= \min_s(|T_0^s|), \\
F_1^I &= \max_s(|F_1^s|), & F_1^{II} &= \text{med}_s(|F_1^s|), & F_1^{III} &= \min_s(|F_1^s|), \\
T_2^I &= \max_s(|T_2^s|), & T_2^{II} &= \text{med}_s(|T_2^s|), & T_2^{III} &= \min_s(|T_2^s|),
\end{aligned}$$

em que  $\max_s(|\cdot|)$  denota a estatística do teste com o maior valor absoluto,  $\text{med}_s(|\cdot|)$  é a estatística do teste com o valor mediano absoluto e  $\min_s(|\cdot|)$  é a estatística do teste com o menor valor absoluto.

Do Teorema 4.1 vem que os processos trimestrais  $X_T^s$  não podem ter propriedades diferentes de raízes unitárias, de onde segue que:

- (i)  $T_0^I$ ,  $F_1^I$  e  $T_2^I$  podem ser interpretados como testes para uma raiz unitária nas frequências trimestrais  $\pm \frac{\pi k}{2}$  ( $k = 0, 1, 2$ ) contra a estacionariedade de uma das séries trimestrais  $X_T^s$ ;
- (ii)  $T_0^{II}$ ,  $F_1^{II}$  e  $T_2^{II}$  são testes para uma raiz unitária nas frequências trimestrais  $\pm \frac{\pi k}{2}$  ( $k = 0, 1, 2$ ) contra a estacionariedade de duas das séries trimestrais  $X_T^s$ ;
- (iii)  $T_0^{III}$ ,  $F_1^{III}$  e  $T_2^{III}$  são testes para uma raiz unitária trimestral nas frequências  $\pm \frac{\pi k}{2}$  ( $k = 0, 1, 2$ ) contra a estacionariedade nas três séries trimestrais  $X_T^s$ .

#### 4.3.2 Teste HEGY Mensal – Obtenção das Estatísticas do Teste para as Séries Mensais $x_t$

Chamemos  $t_0$  a estatística “t” para a integração na frequência zero,  $f_k$  a estatística “F” para a integração nas frequências  $\pm \frac{\pi k}{6}$  e  $t_6$  a estatística “t” para a integração na frequência  $\pi$ .

As estatísticas do teste HEGY mensal foram descritas no Capítulo 3, na Seção 3.5.

#### 4.3.3 Teste HEGY Mensal / Trimestral – Combinação das Estatísticas dos Testes HEGY Mensal e Trimestral

Considerando o Teorema 4.1 e a independência assintótica dos testes HEGY trimestrais e mensais, sob a hipótese nula, podemos obter um teste para a integração na frequência mensal  $\frac{\pi k}{6}$  ( $k = 0, \dots, 6$ ), ao nível  $\alpha_{QM}$ , combinando o teste por integração na frequência trimestral  $\frac{\pi K}{2}$  com a frequência mensal  $\frac{\pi k}{6}$ .

A independência assintótica dos testes mensais e trimestrais é devida a forma particular da regressão auxiliar do teste HEGY, em que o teste de integração em uma particular frequência não é realizado utilizando a série original e sim os dados filtrados (veja Chan e Wei, 1988). Assim, sob a hipótese nula, as variáveis regressoras da regressão trimestral estão filtradas das raízes unitárias trimestrais ausentes. Por exemplo, quando testamos a raiz unitária na frequência zero mensal, a variável regressora correspondente trimestral está filtrada, isto é, não contém as raízes unitárias trimestrais nas frequências  $\frac{\pi}{2}$  e  $\pi$  associadas com as raízes unitárias mensais  $\frac{\pi}{6}$ ,  $\frac{\pi}{2}$ ,  $\frac{5\pi}{6}$ ,  $\frac{\pi}{3}$ ,  $\pi$  e  $\frac{2\pi}{3}$ .

Devido à independência assintótica dos testes HEGY mensal e trimestral, sob a presença de todas as raízes unitárias sazonais mensais, para garantir um nível global de significância igual a  $\alpha_{QM}$ , os níveis dos testes trimestral ( $\alpha_Q$ ) e mensal ( $\alpha_M$ ) devem satisfazer:

$$\alpha_{QM} = 1 - (1 - \alpha_M)(1 - \alpha_Q),$$

tal que para níveis pequenos  $\alpha_{QM} \approx \alpha_Q + \alpha_M$ . Os níveis individuais  $\alpha_Q$  e  $\alpha_M$  podem ser interpretados como pesos. Em casos extremos, quando  $\alpha_Q = 0$ , a informação trimestral não é usada para testar a raiz unitária mensal (HEGY mensal), enquanto quando  $\alpha_Q = \alpha_{QM}$ , somente a informação trimestral é utilizada para testar a raiz unitária mensal (HEGY trimestral).

Conforme exposto anteriormente, devido ao efeito de confundimento, não é possível utilizar o teste HEGY trimestral para testar raízes unitárias sazonais mensais específicas. Portanto, não rejeitamos a existência da raiz unitária  $\frac{\pi k}{6}$  ao nível  $\alpha_{QM}$ , quando para algum  $\alpha_Q$ , tal que  $0 \leq \alpha_Q < \alpha_{QM}$ , as seguintes relações forem satisfeitas:

$$T_0^i > VC_0^i(\alpha_Q) \quad \text{e} \quad t_0 > VC_0(\alpha_{QM} - \alpha_Q),$$

$$F_1^i < VC_1^i(\alpha_Q) \quad \text{e} \quad f_1 < VC_1(\alpha_{QM} - \alpha_Q),$$

$$T_2^i > VC_2^i(\alpha_Q) \quad \text{e} \quad f_2 < VC_2(\alpha_{QM} - \alpha_Q),$$

$$F_1^i < VC_1^i(\alpha_Q) \quad \text{e} \quad f_3 < VC_3(\alpha_{QM} - \alpha_Q),$$

$$T_0^i > VC_0^i(\alpha_Q) \quad \text{e} \quad f_4 < VC_4(\alpha_{QM} - \alpha_Q),$$

$$F_1^i < VC_1^i(\alpha_Q) \quad \text{e} \quad f_5 < VC_5(\alpha_{QM} - \alpha_Q),$$

$$T_2^i > VC_2^i(\alpha_Q) \quad \text{e} \quad t_6 > VC_6(\alpha_{QM} - \alpha_Q),$$

em que,  $i = I, II$  e  $III$  e  $VC$  são os correspondentes valores críticos.

Vamos denotar esse teste por  $QM - HEGY^i$ ,  $i = I, II$  e  $III$ .

#### 4.3.4 Valores Críticos do Teste HEGY Mensal / Trimestral

Os valores críticos para o teste HEGY Mensal / Trimestral foram obtidos por simulações de Monte Carlo, considerando a presença de tendência linear e *dummies* sazonais na regressão auxiliar. Foram gerados quantis de amostras finitas para as distribuições de  $T_0^i$  ( $i = I, II, III$ ) e  $F_1^i$ , satisfazendo  $\alpha_Q + \alpha_M = 5\%$ .

As Tabelas 4.2 a 4.5 contem os valores críticos mensais e trimestrais, em que  $T =$  número de anos completos de observações das séries analisadas.

Tabela 4.2 – Valores críticos para as estatísticas do teste  $QM - HEGY^i$  ( $T = 10$ )

CRITICAL VALUES FOR THE QM-HEGY TEST ( $T = 10$ )									
		$\alpha_Q$							
		0.00	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Quarterly 0 frequency–Monthly 0 and $2\pi/3$ frequencies									
$T_0^I$	$-\infty$	-4.47	-4.30	-4.19	-4.11	-4.03	-3.96	-3.90	
$T_0^{II}$	$-\infty$	-3.25	-3.15	-3.08	-3.02	-2.98	-2.94	-2.90	
$T_0^{III}$	$-\infty$	-2.58	-2.51	-2.45	-2.40	-2.36	-2.33	-2.29	
$t_0$	-3.19	-3.24	-3.29	-3.35	-3.42	-3.50	-3.61	-3.75	
$f_4$	5.84	6.02	6.18	6.39	6.60	6.89	7.23	7.77	
Quarterly $\pi/2$ frequency–Monthly $\pi/6$ , $\pi/2$ and $5\pi/6$ frequencies									
$F_1^I$	$\infty$	11.40	10.63	10.08	9.68	9.34	9.08	8.83	
$F_1^{II}$	$\infty$	6.25	5.89	5.63	5.45	5.28	5.15	5.03	
$F_1^{III}$	$\infty$	3.97	3.74	3.57	3.42	3.31	3.21	3.12	
$f_1$	5.77	6.02	6.20	6.42	6.66	6.94	7.35	7.89	
$f_3$	5.77	5.99	6.16	6.36	6.60	6.91	7.25	7.80	
$f_5$	5.77	5.96	6.13	6.33	6.57	6.83	7.20	7.74	
Quarterly $\pi$ frequency–Monthly $\pi/3$ and $\pi$ frequencies									
$T_2^I$	$-\infty$	-3.84	-3.69	-3.58	-3.49	-3.42	-3.37	-3.32	
$T_2^{II}$	$-\infty$	-2.68	-2.58	-2.52	-2.46	-2.42	-2.38	-2.35	
$T_2^{III}$	$-\infty$	-2.04	-1.96	-1.90	-1.85	-1.81	-1.78	-1.75	
$f_2$	5.82	6.02	6.19	6.42	6.66	6.95	7.33	7.84	
$t_6$	-2.65	-2.74	-2.79	-2.84	-2.91	-2.98	-3.08	-3.21	

Tabela extraída de Pons (2005)

Tabela 4.3 – Valores críticos para as estatísticas do teste  $QM - HEGY^i$  ( $T = 20$ )

CRITICAL VALUES FOR THE QM-HEGY TEST ( $T = 20$ )								
	$\alpha_Q$							
	0.00	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Quarterly 0 frequency–Monthly 0 and $2\pi/3$ frequencies								
$T_0^I$	$-\infty$	-4.30	-4.17	-4.08	-4.01	-3.95	-3.89	-3.85
$T_0^{II}$	$-\infty$	-3.29	-3.21	-3.15	-3.09	-3.05	-3.01	-2.98
$T_0^{III}$	$-\infty$	-2.68	-2.60	-2.54	-2.50	-2.46	-2.43	-2.40
$t_0$	-3.29	-3.39	-3.44	-3.49	-3.55	-3.63	-3.72	-3.85
$f_4$	6.20	6.47	6.64	6.84	7.07	7.35	7.72	8.24
Quarterly $\pi/2$ frequency–Monthly $\pi/6$ , $\pi/2$ and $5\pi/6$ frequencies								
$F_1^I$	$\infty$	10.59	9.98	9.55	9.17	8.91	8.69	8.49
$F_1^{II}$	$\infty$	6.29	5.98	5.76	5.58	5.42	5.30	5.18
$F_1^{III}$	$\infty$	4.17	3.94	3.78	3.64	3.53	3.44	3.35
$f_1$	6.18	6.50	6.67	6.87	7.11	7.42	7.75	8.28
$f_3$	6.24	6.46	6.66	6.86	7.10	7.41	7.79	8.25
$f_5$	6.26	6.47	6.64	6.84	7.07	7.36	7.70	8.20
Quarterly $\pi$ frequency–Monthly $\pi/3$ and $\pi$ frequencies								
$T_2^I$	$-\infty$	-3.77	-3.63	-3.55	-3.47	-3.42	-3.36	-3.31
$T_2^{II}$	$-\infty$	-2.75	-2.65	-2.59	-2.53	-2.48	-2.45	-2.41
$T_2^{III}$	$-\infty$	-2.11	-2.03	-1.97	-1.93	-1.89	-1.85	-1.82
$f_2$	6.46	6.48	6.65	6.85	7.09	7.40	7.77	8.28
$t_6$	-2.83	-2.84	-2.88	-2.95	-3.01	-3.09	-3.19	-3.31

Tabela extraída de Pons (2005)

Tabela 4.4 – Valores críticos para as estatísticas do teste  $QM - HEGY^i$  ( $T = 30$ )

CRITICAL VALUES FOR THE QM-HEGY TEST ( $T = 30$ )								
	$\alpha_Q$							
	0.00	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Quarterly 0 frequency–Monthly 0 and $2\pi/3$ frequencies								
$T_0^I$	$-\infty$	-2.70	-2.62	-2.56	-2.51	-2.48	-2.45	-2.42
$T_0^{II}$	$-\infty$	-3.31	-3.23	-3.17	-3.11	-3.07	-3.04	-3.00
$T_0^{III}$	$-\infty$	-4.31	-4.19	-4.10	-4.03	-3.96	-3.90	-3.85
$t_0$	-3.32	-3.40	-3.45	-3.50	-3.57	-3.64	-3.72	-3.85
$f_4$	6.37	6.64	6.83	7.03	7.27	7.57	7.89	8.44
Quarterly $\pi/2$ frequency–Monthly $\pi/6$ , $\pi/2$ and $5\pi/6$ frequencies								
$F_1^I$	$\infty$	4.23	4.01	3.84	3.69	3.59	3.49	3.41
$F_1^{II}$	$\infty$	6.36	6.04	5.79	5.61	5.47	5.34	5.23
$F_1^{III}$	$\infty$	10.52	9.98	9.56	9.23	8.97	8.74	8.55
$f_1$	6.30	6.63	6.82	7.02	7.26	7.55	7.92	8.47
$f_3$	6.35	6.65	6.84	7.05	7.29	7.59	7.96	8.45
$f_5$	6.34	6.65	6.83	7.00	7.27	7.55	7.92	8.44
Quarterly $\pi$ frequency–Monthly $\pi/3$ and $\pi$ frequencies								
$T_2^I$	$-\infty$	-2.12	-2.05	-1.98	-1.94	-1.90	-1.87	-1.84
$T_2^{II}$	$-\infty$	-2.74	-2.66	-2.60	-2.54	-2.50	-2.46	-2.43
$T_2^{III}$	$-\infty$	-3.79	-3.65	-3.56	-3.48	-3.42	-3.36	-3.31
$f_2$	6.31	6.60	6.81	7.03	7.27	7.56	7.95	8.47
$t_6$	-2.79	-2.88	-2.93	-2.99	-3.06	-3.13	-3.23	-3.36

Tabela extraída de Pons (2005)

Tabela 4.5 – Valores críticos para as estatísticas do teste  $QM - HEGY^i$  ( $T = 40$ )

CRITICAL VALUES FOR THE QM-HEGY TEST ( $T = 40$ )								
	$\alpha_Q$							
	0.00	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Quarterly 0 frequency–Monthly 0 and $2\pi/3$ frequencies								
$T_0^I$	$-\infty$	-4.29	-4.18	-4.09	-4.02	-3.95	-3.91	-3.86
$T_0^{II}$	$-\infty$	-3.32	-3.23	-3.17	-3.12	-3.08	-3.05	-3.01
$T_0^{III}$	$-\infty$	-2.71	-2.64	-2.59	-2.54	-2.50	-2.46	-2.43
$t_0$	-3.35	-3.44	-3.49	-3.54	-3.60	-3.68	-3.76	-3.89
$f_4$	6.40	6.74	6.92	7.15	7.37	7.66	8.01	8.50
Quarterly $\pi/2$ frequency–Monthly $\pi/6$ , $\pi/2$ and $5\pi/6$ frequencies								
$F_1^I$	$\infty$	10.40	9.81	9.43	9.10	8.85	8.64	8.45
$F_1^{II}$	$\infty$	6.35	6.02	5.81	5.63	5.48	5.37	5.26
$F_1^{III}$	$\infty$	4.29	4.04	3.87	3.73	3.62	3.53	3.44
$f_1$	6.30	6.66	6.85	7.06	7.30	7.60	7.97	8.48
$f_3$	6.35	6.69	6.90	7.11	7.35	7.67	8.03	8.59
$f_5$	6.48	6.69	6.87	7.08	7.31	7.64	8.04	8.53
Quarterly $\pi$ frequency–Monthly $\pi/3$ and $\pi$ frequencies								
$T_2^I$	$-\infty$	-3.77	-3.66	-3.56	-3.48	-3.42	-3.37	-3.32
$T_2^{II}$	$-\infty$	-2.77	-2.69	-2.62	-2.57	-2.53	-2.49	-2.46
$T_2^{III}$	$-\infty$	-2.14	-2.07	-2.01	-1.96	-1.92	-1.88	-1.85
$f_2$	6.46	6.77	6.94	7.14	7.40	7.68	8.05	8.56
$t_6$	-2.81	-2.91	-2.96	-3.01	-3.08	-3.16	-3.26	-3.39

Tabela extraída de Pons (2005)

Uma das vantagens de se combinar os testes HEGY trimestral e mensal é reduzir o impacto dos filtros das raízes unitárias ausentes, no poder das frequências presentes. Por exemplo, o filtro soma  $S_0(B) = 1 + B + \dots + B^{s-1}$ , que se refere a todas as raízes unitárias sazonais, com exceção da raiz unitária na frequência zero, aumenta a persistência na frequência zero, reduzindo o poder do teste HEGY em comparação com o teste de Dickey-Fuller, que testa a raiz unitária na frequência zero com dados não filtrados.

O impacto dos filtros de raízes unitárias é similar para as outras frequências. Os efeitos desses filtros sobre a persistência na frequência  $\theta_k$  podem ser mensurados por intermédio da função filtro  $|S_k(e^{iw})|^2$  avaliada na frequência  $w = \theta_k$ , onde  $S_k(B)$  denota o filtro que contém todas as raízes unitárias sazonais, com exceção daquela na frequência  $\theta_k$ . No caso de filtros mensais,  $|S_k(e^{iw})|^2$  assume os seguintes valores:

$$|S_k(e^{iw})|_{w=\theta_k}^2 = 144 \text{ nas frequências zero, } \frac{\pi}{6}, \frac{5\pi}{6} \text{ e } \pi \text{ (valor máximo);}$$

$$|S_k(e^{iw})|_{w=\theta_k}^2 = 48 \text{ para as frequências } \frac{\pi}{3} \text{ e } \frac{2\pi}{3} \text{ e}$$

$$|S_k(e^{iw})|_{w=\frac{\pi}{2}}^2 = 36 \text{ para a frequência } \frac{\pi}{2}.$$

De maneira análoga, o impacto dos filtros das raízes unitárias trimestrais é maior nas frequências zero e  $\pi$  ( $|S_k(e^{iw})|_{w=\theta_k}^2 = 16$  para  $k = 0, 2$ ) do que na frequência  $\frac{\pi}{2}$  ( $|S_1(e^{iw})|_{w=\frac{\pi}{2}}^2 = 4$ ). Em qualquer caso, o aumento da persistência dos filtros é muito menor com o teste HEGY trimestral do que com o teste HEGY mensal.

Além disso, uma característica importante dos testes de raízes unitárias nos três processos trimestrais  $X_T^s$  é que valores muito diferentes das estatísticas  $T_k^I (F_1^I)$ ,  $T_k^{II} (F_1^{II})$  e  $T_k^{III} (F_1^{III})$  revelam a presença de características como quebras estruturais ou valores atípicos que podem afetar o desempenho dos testes HEGY. Características tais como alteração da média sazonal, *outliers* aditivos ou erros de medida geralmente não afetam todos os meses ou todos os ciclos mensais com a mesma intensidade e, portanto, poderão afetar as médias sazonais trimestrais e os ciclos dos três processos amostrados  $X_T^s$  com intensidades diferentes. Por exemplo, uma ou múltiplas quebras no parâmetro  $\psi_7$  ou múltiplos *outliers* na observação referente ao mês de Julho, somente afetará o processo trimestral  $X_T^1$ , mas não os processos  $X_T^2$  e  $X_T^3$ .

A presença de mudanças na média sazonal, sob a alternativa, reduzirá artificialmente o valor absoluto das estatísticas HEGY (veja Smith e Otero, 1997; Franses e Vogelsang, 1998; da Silva Lopes, 2001; Hassler e Rodrigues, 2004) de tal forma que as estatísticas trimestrais  $T_k^{III}$  e  $F_1^{III}$  serão mais distorcidas do que  $T_k^{II}$  e  $F_1^{II}$  e muito mais distorcidas do que  $T_k^I$  e  $F_1^I$ . Em outra situação, a presença de *outliers* ou erros de medida torna o teste HEGY predisposto a rejeitar a hipótese nula. Portanto, nesse caso  $T_k^{III}$  e  $F_1^{III}$  costumam ser mais robustos do que  $T_k^I$  e  $F_1^I$ . Por fim, a presença de mudanças na média sazonal pode super-rejeitar ou sub-rejeitar a hipótese nula, dependendo do processo gerador dos dados, de tal forma que a estatística trimestral  $T_k^i$  e  $F_1^i$  é a priori superior a qualquer outra nessa situação.

Assim, avaliando o gráfico das séries amostradas trimestrais e comparando as estatísticas trimestrais, podemos ter evidências acerca da presença de quebras ou *outliers* que não são detectados somente com a informação mensal ou com a média dos dados trimestrais amostrados.

Para mais detalhes, sugerimos ver Pons (2005).

### ➤ Experimento de Monte Carlo

Pons (2005) conduziu um experimento de Monte Carlo que permitiu comparar o tamanho e o poder do teste HEGY mensal com os tamanhos e os poderes das diferentes versões dos testes HEGY resultantes da combinação mensal e trimestral – QM-HEGY. Também foi avaliado o desempenho do teste HEGY utilizando o processo trimestral gerado por amostras agregadas – teste *QM – HEGY<sup>a</sup>*.

Para tal, foram utilizadas duas variações para cada teste  $QM - HEGY^i$ , sendo uma com um peso menor para a informação trimestral ( $\alpha_Q = 1\%$ ) e outra com um peso maior ( $\alpha_Q = 4\%$ ). Foi considerado o processo de geração de dados dado por:

$$(1 - \rho L^{12})x_t = u_t,$$

em que  $\rho = \{0.7, 0.8, 0.9, 1\}$ ,  $u_t \sim RB(0, 1)$ ,  $t = 1, \dots, 12N$ , com  $N = 20, 40$ .

Foram feitas 10.000 réplicas desse processo para calcular a porcentagem de rejeição da hipótese nula ao nível  $\alpha_{QM} = 5\%$ .

Tabela 4.6 – Porcentagens de rejeição do teste na frequência zero

PERCENTAGES OF REJECTION (ZERO FREQUENCY)								
<i>T</i> :	20	20	20	20	40	40	40	40
<i>ρ</i> :	0.70	0.80	0.90	1.00	0.70	0.80	0.90	1.00
Monthly HEGY	0.12	0.08	0.06	0.05	0.33	0.16	0.08	0.05
QM-HEGY <sup>a</sup> (1)	0.12	0.08	0.06	0.05	0.33	0.16	0.08	0.05
QM-HEGY <sup>a</sup> (4)	0.12	0.08	0.06	0.05	0.33	0.15	0.08	0.05
QM-HEGY <sup>III</sup> (1)	0.14	0.09	0.06	0.05	0.51	0.22	0.08	0.05
QM-HEGY <sup>III</sup> (4)	0.21	0.12	0.07	0.05	0.69	0.32	0.11	0.05
QM-HEGY <sup>II</sup> (1)	0.14	0.09	0.06	0.05	0.43	0.19	0.08	0.05
QM-HEGY <sup>II</sup> (4)	0.19	0.11	0.07	0.05	0.60	0.26	0.10	0.05
QM-HEGY <sup>I</sup> (1)	0.12	0.08	0.06	0.05	0.34	0.16	0.08	0.05
QM-HEGY <sup>I</sup> (4)	0.13	0.09	0.07	0.05	0.37	0.17	0.08	0.05

Tabela extraída de Pons (2005)

Tabela 4.7 – Porcentagens de rejeição do teste na frequência  $\frac{\pi}{2}$

PERCENTAGES OF REJECTION (3 FREQUENCY)								
<i>T</i> :	20	20	20	20	40	40	40	40
<i>ρ</i> :	0.70	0.80	0.90	1.00	0.70	0.80	0.90	1.00
Monthly HEGY	0.30	0.15	0.08	0.05	0.85	0.45	0.15	0.05
QM-HEGY <sup>a</sup> (1)	0.31	0.16	0.09	0.05	0.85	0.46	0.14	0.05
QM-HEGY <sup>a</sup> (4)	0.31	0.16	0.09	0.05	0.83	0.44	0.14	0.05
QM-HEGY <sup>III</sup> (1)	0.46	0.22	0.10	0.05	0.98	0.68	0.19	0.05
QM-HEGY <sup>III</sup> (4)	0.63	0.32	0.14	0.05	0.99	0.83	0.30	0.04
QM-HEGY <sup>II</sup> (1)	0.39	0.19	0.10	0.05	0.96	0.59	0.16	0.05
QM-HEGY <sup>II</sup> (4)	0.53	0.26	0.12	0.05	0.99	0.75	0.23	0.05
QM-HEGY <sup>I</sup> (1)	0.31	0.16	0.09	0.05	0.87	0.47	0.14	0.05
QM-HEGY <sup>I</sup> (4)	0.34	0.17	0.09	0.05	0.92	0.50	0.15	0.05

Tabela extraída de Pons (2005)

As Tabelas 4.6 e 4.7 contem os resultados obtidos para as frequências mensais 0 e  $\frac{\pi}{2}$ . A análise dessas tabelas mostra que o poder dos testes de raízes unitárias mensais aumenta quando utilizamos a informação trimestral. A versão mais poderosa do teste QM-HEGY observada é o  $QM - HEGY^{III}$ , isto é, quando a rejeição da raiz unitária trimestral requer a rejeição de  $X_T^1$ ,  $X_T^2$  e  $X_T^3$ . Também é interessante notar que o teste  $QM - HEGY^i$  (para qualquer  $i$ ) rejeita mais frequentemente o nível trimestral utilizando  $\alpha_Q = 4\%$ . Quando a melhor versão dos testes QM-HEGY é utilizada, o incremento no poder desse teste comparado ao poder do teste HEGY mensal é considerável. Por exemplo, quando  $N = 40$  e  $\rho = 0.7$ , o teste HEGY mensal tem poder empírico igual a 33% e 85% nas frequências 0 e  $\frac{\pi}{2}$ , respectivamente, enquanto o teste  $QM - HEGY^{III}$  tem poder igual a 69% e 99%, respectivamente. A diferença observada entre os poderes nas frequências 0 e  $\frac{\pi}{2}$  está relacionada provavelmente com a distorção introduzida pelos filtros de raízes unitárias ausentes em ambas as séries mensais e trimestrais.

Quanto ao processo  $X_T^a$ , por utilizar somente parte da informação trimestral, as Tabelas 4.6 e 4.7 apontam que não há diferenças significativas entre o  $QM - HEGY^a$  e o HEGY mensal. Nota-se também que não há ganho considerável no poder do teste  $QM - HEGY^a$ , quando diferentes níveis de significância trimestrais são utilizados.

## Capítulo 5

# Efeitos de Erros de Medida e Valores Atípicos no Teste HEGY

Os valores de uma série temporal podem, muitas vezes, serem afetados por eventos inesperados tais como mudanças governamentais ou crises econômicas, ondas inesperadas de frio ou de calor, erros de medida ou de digitação, etc. Como consequência desses tipos de eventos, a série temporal pode apresentar observações espúrias, que são denominadas valores atípicos ou *outliers*.

A presença desses *outliers* faz com que os testes para raízes unitárias sazonais sofram impactos cuja intensidade varia conforme a frequência, o tamanho, a persistência dos *outliers*, bem como a razão sinal ruído na série que está sendo analisada. O teste HEGY, por exemplo, tem suas estatísticas impactadas quando a série observada contém erros de medida e *outliers* aditivos e de mudança temporária de nível.

### 5.1 *Outliers* Aditivo e de Mudança Temporária

Considere o processo univariado sazonal dado por:

$$\Delta_s y_t = \epsilon_t, \quad t = 1, 2, \dots, T, \quad (5.1)$$

em que  $\epsilon_t, \sim \text{i.i.d. } (0, \sigma_\epsilon^2)$  e  $\Delta_s = 1 - B^s$  é o filtro de diferença sazonal.

Na prática temos  $s = 1, 2, 4$  e  $12$  correspondendo, respectivamente, a observações amostradas anualmente, semestralmente, trimestralmente e mensalmente.

Para avaliar as implicações dos erros de medida e *outliers* no teste HEGY, vamos assumir que a série observada possa ser escrita da seguinte forma:

$$z_t = y_t + v_t, \quad (5.2)$$

em que  $v_t$  é o termo de erro que representa a contaminação de  $y_t$ . Em particular, consideremos  $v_t$  gerado por:

$$v_t = \frac{\theta}{(1 - \alpha B)} \delta_t + \eta_t, \quad (5.3)$$

em que  $\eta_t \sim \text{i.i.d. } (0, \sigma_\eta^2)$  é o erro de medida. O primeiro termo em (5.3) é a componente geral do *outlier* na qual assumimos  $|\alpha| < 1$ . Um primeiro tipo de *outlier*, denominado aditivo, ocorre quando  $\alpha = 0$ . Neste caso,  $\theta\delta_t$  é o termo de ruído gerado pelo *outlier* aditivo (AO) observado irregularmente. O parâmetro  $\theta$  denota a magnitude do *outlier*, enquanto  $\delta_t$  é uma variável indicadora Bernoulli que pode assumir valor 1 ou -1 com probabilidade  $\frac{p}{2}$ . Dessa forma, temos que os *outliers* aditivos são caracterizados por alguns eventos que não se repetem, ocorrem irregularmente e não são afetados pela dinâmica do processo  $y_t$ .

Um outro tipo de *outlier* ocorre quando  $\alpha \neq 0$ . Nessa situação, os *outliers* também aparecem irregularmente, porém, os efeitos que causam tendem a persistir ao longo da série. Esses *outliers* causam mudanças de níveis temporárias e serão denotados por *outliers* TC's (IO's).

Quando  $\theta = 0$ , então temos a ocorrência de um erro de medida.

Da forma como foi estipulado o modelo, os impactos causados pelos *outliers* no teste HEGY podem ser controlados quando a frequência, a magnitude relativa e a persistência dos *outliers* se modificam. De forma análoga, o ruído resultante dos erros de medida pode ser controlado utilizando o inverso da razão ruído  $\left(\frac{\sigma_\eta}{\sigma_\epsilon}\right)^2$ .

## 5.2 O Teste HEGY para Séries com Ruído

Como visto anteriormente, o teste HEGY se baseia em uma regressão auxiliar, na qual as variáveis regressoras são transformadas tais que se tornam assintoticamente ortogonais, o que possibilita testarmos separadamente as raízes em específicas frequências do processo autorregressivo. Neste contexto, a regressão auxiliar utilizando a série contaminada  $z_t$  pode ser escrita na forma:

$$\Delta_s z_t = \sum_{j=1}^s \pi_s^j z_{j,t-1}^s + u_t^s = z_t^{s'} \boldsymbol{\pi}^s + u_t, \quad (5.4)$$

em que

$z_{jt}^s$ ,  $j = 1, 2, \dots, s$  são as séries  $z_t$  filtradas,

$z_t^s = (z_{1t}^s, z_{2t}^s, \dots, z_{st}^s)'$  e  $\pi^s = (\pi_1^s, \pi_2^s, \dots, \pi_s^s)'$  são os parâmetros do modelo.

Se uma raiz unitária real estiver presente em uma particular frequência, então o coeficiente  $\pi_j^s$  associado será zero. Vale citar que raízes unitárias imaginárias implicam em pares de coeficientes  $\pi_j^s$  iguais a zero, em particular  $\pi_j^s = \pi_{j+1}^s = 0$  para  $j$  ímpar e maior do que um. Neste caso, para testar a presença de raízes unitárias sazonais devemos realizar os testes F conjuntos na forma  $F_{\pi_j^s, \pi_{j+1}^s}$ , para  $j$  ímpar e maior do que um.

### 5.3 Distribuição Assintótica das Estatísticas do Teste HEGY com Ruído

Quando os dados da série que está sendo avaliada apresentam ruídos, isto é, podem ser representados pelas equações (5.1), (5.2) e (5.3), as distribuições das estatísticas “t” e “F” são aquelas apresentadas na Seção 3.2, para o caso de dados trimestrais e 3.5, para dados mensais, acrescidas de um fator de escala dado por:

$$\sqrt{\left[1 + p \left(\frac{\theta}{\sigma_\epsilon}\right)^2 \frac{(2 - \alpha^s)}{(1 - \alpha^2)} + 2 \left(\frac{\sigma_n}{\sigma_\epsilon}\right)^2\right]}$$

e de um fator de localização dado por:

$$\sqrt{s} \frac{(\mathcal{K}_j - \sigma_n^2)}{\sigma_\epsilon^2} \quad \text{ou} \quad \sqrt{s} \frac{\mathcal{K}_j}{\sigma_\epsilon^2},$$

em que

$$\mathcal{K}_j = \begin{cases} -\theta^2 p \left( \frac{1 - \alpha^s}{1 - \alpha^2} \right), & \text{para } j = 1, 2, 3, 5, 7, \dots, s-1, \\ -2\theta^2 p \alpha \left( \frac{1 - \alpha^s}{1 - \alpha^4} \right), & \text{para } j = 4, \\ -\sqrt{3}\theta^2 p \alpha (\alpha^8 - \alpha^7 + \alpha^6 + \alpha^2 - \alpha + 1), & \text{para } j = 6, \\ -\sqrt{3}\theta^2 p \alpha (\alpha^8 + \alpha^7 + \alpha^6 + \alpha^2 + \alpha + 1), & \text{para } j = 8, \\ -\sqrt{3}\theta^2 p \alpha (\alpha^8 - \sqrt{3}\alpha^7 + 3\alpha^6 - 2\sqrt{3}\alpha^5 + 4\alpha^4 - 2\sqrt{3}\alpha^3 + 3\alpha^2 - \sqrt{3}\alpha + 1), & \text{para } j = 10, \\ -\sqrt{3}\theta^2 p \alpha (\alpha^8 + \sqrt{3}\alpha^7 + 3\alpha^6 + 2\sqrt{3}\alpha^5 + 4\alpha^4 + 2\sqrt{3}\alpha^3 + 3\alpha^2 + \sqrt{3}\alpha + 1), & \text{para } j = 12. \end{cases}$$

A fim de estudar as implicações causadas por esses fatores de escala e de localização nas estatísticas dos testes HEGY, Haldrup et al (2005) avaliaram os seguintes casos: séries contendo erros de medida, *outliers* aditivos e *outliers* de mudança temporária (ou de inovação).

### ➤ Erros de Medida

Na presença de erros de medida, assumimos que  $\theta = 0$  e  $\sigma_n^2 \neq 0$ , o que implica em:

(i) A distribuição de  $t_{\pi_j^s}$  é deslocada para a esquerda para  $j = 1, 2, 3, 5, \dots, s-1$ , enquanto para valores pares de  $j$  maiores do que 2, não há deslocamento.

(ii) O efeito do fator de escala tende a concentrar as distribuições das estatísticas  $t$ . Quanto maior a taxa de ruído  $\left(\frac{\sigma_n}{\sigma_e}\right)^2$ , mais concentrada se torna a distribuição. Portanto, para  $j$  par e maior do que 2, os resultados indicam que os testes  $t$  rejeitarão as hipóteses nulas com menor frequência do que o esperado pelo nível de significância.

(iii) Para  $j = 1, 2, 3, 5, \dots, s-1$ , os efeitos de localização e mudança tem sinais opostos e, portanto, é difícil prever o efeito para um particular valor da razão de ruídos. Contudo, na medida em que essa razão aumenta, o efeito da localização é claramente dominante e, portanto, aumenta o tamanho do teste comparado com o nível de significância. Dessa forma, rejeitaremos frequentemente a raiz unitária sazonal associada à essas particulares estatísticas.

Como vimos no Capítulo 3, sugerimos testar os pares de raízes unitárias complexas através dos testes  $F$ ,  $F_{\pi_j^s, \pi_{j+1}^s}$ , para  $j = 3, 5, 7, \dots, 11$ . As implicações para esses testes seguem

aquelas de  $t_{\pi_j^s}$  (com  $j$  ímpar). Entretanto, pelo fato de ser esperado que  $t_{\pi_j^s}$  tenha tamanho superior ao tamanho nominal quando  $j$  é par, a distorção no tamanho total de  $F_{\pi_j^s, \pi_{j+1}^s}$  deve estar compreendida entre os correspondentes tamanhos dos testes  $t$ .

É interessante explicar a natureza dos erros de medida no contexto dos testes HEGY. O erro populacional da regressão HEGY terá a seguinte forma:

$$\Delta_s z_t = \epsilon_t + \eta_t - \eta_{t-s},$$

que é equivalente a um processo MA(s) definido por  $\Delta_s z_t = (1 + \lambda_s B^s) \epsilon_t$ , em que  $\epsilon_t \sim i.i.d. (0, \sigma_\epsilon^2)$ . Por exemplo, quando a razão ruído estiver entre 0,5 e 1, então o parâmetro MA associado estará entre -0,38 e -0,46 e, no limite, quando a razão ruído tende a infinito, o parâmetro se aproxima de -1.

### ➤ **Outliers Aditivos (AO)**

Nesse caso, assumimos que  $\theta \neq 0$  e  $\sigma_n^2 = \alpha = 0$ . Os resultados da distribuição (efeitos de localização e de escala) das estatísticas são qualitativamente similares aqueles considerados nos erros de medida. Portanto, o tamanho do teste será afetado da seguinte forma:

- (i) Para  $j = 1, 2$  e todos os valores ímpares de  $j$ , o tamanho do teste tenderá a ser maior do que o nível de significância;
- (ii) Para  $j$  par e maior do que 2, o tamanho do teste será inferior ao nível de significância.

Quanto à interpretação, é evidente que erros do tipo média-móvel possam aparecer. Em particular, temos:

$$\Delta_s z_t = \epsilon_t + \theta \Delta_s \delta_t,$$

em que o parâmetro de média-móvel é como o caso de erros de medida com  $\theta^2 p$  substituindo  $\sigma_n^2$ .

### ➤ **Outliers de Mudança Temporária (TC ou IO)**

Nesse caso, assumimos que  $\theta \neq 0$  com  $0 < |\alpha| < 1$  e  $\sigma_n^2 = 0$ . *Outliers* de mudança temporária (TC), assim como *outliers* aditivos, ocorrem irregularmente, mas tendem a persistir por algum tempo. Qualitativamente, todas as estatísticas são deslocadas para a esquerda,

enquanto quantitativamente, o deslocamento dependerá de qual estatística for considerada. De forma contrária aos erros de medida e aos *outliers* aditivos, para os *outliers* TC, o tamanho de todos os testes excederá o nível nominal, se os outliers não estiverem sendo tratados corretamente. Portanto, frequentemente tenderemos a rejeitar a presença de raízes unitárias.

Para mais detalhes, sugerimos ver Haldrup et al (2005).

## 5.4 Simulações de Monte Carlo

Haldrup et al (2005), a fim de avaliar as implicações resultantes dos efeitos causados pelos *outliers* e erros de medida nas estatísticas do teste HEGY, conduziram um experimento de Monte Carlo para dados trimestrais ( $s = 4$ ).

O desenho experimental utilizado é baseado nas equações (5.1), (5.2) e (5.3), com valores iniciais zero e  $\varepsilon_t \sim N(0, 1)$ . Os modelos de AO, TC e erro de medida foram analisados separadamente. Como vimos, AO's e erros de medida possuem tamanhos assintóticos idênticos para um determinado valor de  $\left(\frac{\sigma_\eta}{\sigma_\varepsilon}\right)^2 = p\left(\frac{\theta}{\sigma_\varepsilon}\right)^2$ . Porém, como as séries contendo erros de medida são muito diferentes das séries contendo AO's, foram conduzidas simulações para ambos os casos. Os valores dos parâmetros do modelo AO foram escolhidos tais que  $p\theta^2 = \{0,4, 0,8\}$ . Para comparar com o modelo de erro de medida foi considerado o mecanismo gerador de ruído dado por  $\eta_t \sim N(0, \sigma_\eta^2 = p\theta^2)$ . Os valores para os parâmetros do modelo de *outliers* TC foram os mesmos utilizados para o modelo AO. Em particular, foram utilizados os valores  $\alpha = \{0,25, 0,75\}$  para o parâmetro autorregressivo.

Para cada repetição de Monte Carlo, todas as estatísticas “t” do teste HEGY ( $t_{\pi_1}, t_{\pi_2}, t_{\pi_3}, t_{\pi_4}$ ), assim como a estatística “F” ( $F_{\pi_3, \pi_4}$ ) foram calculadas e, a partir delas, as frequências com que os testes foram rejeitados foram obtidas. A Tabela 5.1 apresenta os painéis A e B contendo, respectivamente, as frequências de rejeição para os modelos de erro de medida e AO.

Tabela 5.1 – Porcentagens de rejeição dos testes HEGY considerando modelos de erros de medida e AO's

Sizes at nominal 5% level of the HEGY tests for quarterly data with measurement errors and additive outliers. Data is generated according to the data generating process  $z_t = y_t + \theta/(1 - \alpha L)\delta_t + \eta_t$  with  $\Delta_4 y_t = \varepsilon_t, \alpha = 0, \eta_t \sim N(0, \sigma_\eta^2)$ , and  $\varepsilon_t \sim N(0, 1)$ .

$\sigma_\eta^2$	$T$	$t_{\pi_1^4}$	$t_{\pi_2^4}$	$t_{\pi_3^4}$	$t_{\pi_4^4}$	$F_{\pi_3^4, \pi_4^4}$
<i>Panel A: measurement errors</i>						
0.4	48	0.097	0.101	0.133	0.015	0.072
	100	0.121	0.133	0.17	0.009	0.103
	200	0.145	0.147	0.157	0.009	0.081
	400	0.155	0.175	0.175	0.010	0.097
0.8	48	0.156	0.145	0.24	0.009	0.135
	100	0.216	0.208	0.299	0.007	0.183
	200	0.224	0.248	0.293	0.006	0.158
	400	0.252	0.260	0.300	0.003	0.181
<i>Panel B: additive outliers</i>						
0.4	48	0.105	0.095	0.145	0.017	0.087
	100	0.101	0.12	0.168	0.010	0.100
	200	0.143	0.134	0.170	0.009	0.100
	400	0.149	0.162	0.173	0.012	0.093
0.8	48	0.159	0.153	0.242	0.014	0.131
	100	0.189	0.215	0.287	0.01	0.178
	200	0.229	0.222	0.296	0.006	0.186
	400	0.239	0.256	0.287	0.005	0.178

Tabela extraída de Haldrup et al (2005)

Embora os erros de medida e os AO's sejam muito diferentes graficamente, as suas implicações com relação às distribuições de amostra finita são bastante similares (assim como também são similares assintoticamente). Se nenhuma providência for tomada para ajustar os *outliers* ou erros de medida, então as distorções nos tamanhos dos testes podem ser enormes. As distorções de amostras finitas parecem similares para as estatísticas  $t_{\pi_1^4}$  e  $t_{\pi_2^4}$ , o que foi sugerido nos resultados assintóticos apresentados. Também de acordo com os resultados assintóticos, a estatística  $t_{\pi_4^4}$  terá um tamanho menor do que o tamanho nominal e, como o teste  $F_{\pi_3^4, \pi_4^4}$  é construído através de uma combinação de uma estatística “over-sized” com outra “under-sized”, então a distorção do tamanho desse teste será uma combinação convexa dos tamanhos dos testes t.

Os resultados para o modelo TC se encontram na Tabela 5.2, onde se pode avaliar as frequências de rejeição obtidas dos experimentos.

Tabela 5.2 – Porcentagens de rejeição dos testes HEGY considerando modelos de TC

Sizes at nominal 5% level of the HEGY tests for quarterly data with temporary change outliers. Data is generated according to the data generating process  $z_t = y_t + \theta/(1 - \alpha L)\delta_t$  with  $\Delta_4 y_t = \varepsilon_t$ , and  $\varepsilon_t \sim N(0, 1)$ .

$p\theta^2$	$\alpha$	$T$	$t_{\pi_1^A}$	$t_{\pi_2^A}$	$t_{\pi_3^A}$	$t_{\pi_4^A}$	$F_{\pi_3^A, \pi_4^A}$
0.4	0.25	48	0.094	0.112	0.149	0.027	0.090
		100	0.102	0.139	0.170	0.019	0.113
		200	0.146	0.148	0.178	0.024	0.109
		400	0.155	0.172	0.180	0.017	0.111
	0.75	48	0.103	0.146	0.182	0.057	0.165
		100	0.114	0.188	0.208	0.079	0.198
		200	0.190	0.193	0.241	0.106	0.213
		400	0.199	0.218	0.239	0.088	0.221
0.8	0.25	48	0.147	0.180	0.243	0.024	0.162
		100	0.193	0.228	0.301	0.035	0.205
		200	0.235	0.232	0.313	0.029	0.217
		400	0.243	0.270	0.308	0.024	0.217
	0.75	48	0.162	0.260	0.293	0.106	0.277
		100	0.228	0.297	0.382	0.194	0.368
		200	0.295	0.314	0.397	0.253	0.402
		400	0.312	0.348	0.417	0.267	0.424

Tabela extraída de Haldrup et al(2005)

Para valores moderados do parâmetro de persistência do *outlier* ( $\alpha = 0,25$ ), observamos que as distorções nos tamanhos são muito similares ao caso em que  $\alpha = 0$ . Entretanto, as distorções do teste  $t_{\pi_4^A}$  parecem ser moderadas e as distorções para  $F_{\pi_3, \pi_4}$  tendem a se agravar comparadas com o caso do modelo AO. Geralmente, os tamanhos atuais se tornam maiores quando o parâmetro de persistência aumenta ( $\alpha = 0,75$ ).

## 5.5 O Teste HEGY com Correção para *Outliers*

### ➤ Identificação de *outliers* em dados sazonais

Primeiramente, faz-se necessário apresentar um procedimento capaz de identificar *outliers* em uma série temporal sazonal. Vogelsang (1999) sugeriu um procedimento iterativo e Perron e Rodriguez (2003) provaram que esse procedimento era inconsistente e, como alternativa, sugeriram um esquema também iterativo que traz consistência ao teste para *outliers* aditivos com tamanho indicado. Em Haldrup et al. (2003) esse teste foi estendido para

o caso de dados sazonais e foram apresentadas simulações do seu desempenho. A idéia do teste e a sua subsequente ligação com o procedimento HEGY são muito similares, porém, devido à amostragem sazonal dos dados, o teste precisa ser ligeiramente modificado.

O teste modificado de Perron-Rodriguez é baseado na seguinte regressão auxiliar:

$$\Delta_s z_t = \theta [D(T_{AO})_t - D(T_{AO})_{t-s}] + v_t,$$

em que  $D(T_{AO})_{t-j} = 1$ , quando  $t = T_{AO} + j$  e 0, caso contrário.

Assim, sob a hipótese nula de  $\theta = 0$ , temos:

$$\hat{\theta}(T_{AO}) = \begin{cases} -v_{T_{AO}+s}, & T_{AO} \leq s, \\ \frac{1}{2}(v_{T_{AO}} - v_{T_{AO}+s}), & s < T_{AO} \leq T - s, \\ v_{T_{AO}}, & T_{AO} > T - s. \end{cases}$$

Haldrup et al. (2005) demonstraram que os resíduos da regressão de  $\Delta_s z_t$  sobre componentes determinísticas tais como constante, tendência e *dummies* sazonais podem ser usados para substituir  $\Delta_s z_t$  sem afetar a distribuição do teste de detecção de *outlier*.

A estatística desse teste é dada por:

$$\tau^{PR} = \sup_{T_{AO}} |\tau^{PR}(T_{AO})|$$

em que

$$\tau^{PR}(T_{AO}) = \begin{cases} \hat{\theta}(T_{AO}) \hat{R}_v(0)^{-\frac{1}{2}}, & T_{AO} \leq s, \\ \sqrt{2} \hat{\theta}(T_{AO}) (\hat{R}_v(0) - \hat{R}_v(s))^{-\frac{1}{2}}, & s < T_{AO} \leq T - s, \\ \hat{\theta}(T_{AO}) \hat{R}_v(0)^{-\frac{1}{2}}, & T_{AO} > T - s \end{cases}$$

$$e \quad \hat{R}_v(j) = T^{-1} \sum_{t=s+j+1}^T \hat{v}_t \hat{v}_{t-j}$$

Esse teste é utilizado de forma iterativa. Primeiramente,  $\tau^{PR}$  é calculado considerando todas as observações e, na sequência, essa estatística é comparada ao correspondente valor crítico que se encontra na Tabela 5.3, extraída de Perron e Rodriguez (2003). Quando o primeiro *outlier* é identificado, Perron e Rodriguez sugerem excluir essa observação e recalculer o teste para a amostra reduzida. Entretanto, para dados sazonais, essa ação destruirá a estrutura de auto-correlação dos dados e sua sazonalidade, então a sugestão é substituir o *outlier*  $z_{T_{AO}}$  por  $z_{T_{AO}-s}$ , de forma que na próxima iteração temos  $\Delta_s z_{T_{AO}} = 0$ , onde  $T_{AO}$  é o *outlier* detectado na iteração anterior. O procedimento iterativo deve continuar até que não haja mais rejeições e, assim, todos os *outliers* terão sido identificados.

Tabela 5.3 – Valores críticos para identificação de *outliers*

FINITE SAMPLE CRITICAL VALUES OF THE TEST $\tau_d$				
Level of significance	Model 1 $z_t = \{1\}$		Model 2 $z_t = \{1, t\}$	
	$T = 100$	$T = 200$	$T = 100$	$T = 200$
1.0%	4.14	4.20	4.13	4.19
2.5%	3.87	3.95	3.85	3.94
5.0%	3.65	3.75	3.63	3.74
10.0%	3.44	3.56	3.42	3.55

Tabela extraída de Perron e Rodriguez (2003)

➤ **O teste HEGY com correção para *outlier***

Uma vez que os *outliers* foram identificados, podemos escrever a regressão do teste HEGY da seguinte forma:

$$\Delta_s z_t = \mu_s + \sum_{j=1}^{s-1} \mu_j D_{jt} + \beta t + \sum_{j=1}^s \pi_j^s z_{j,t-1}^s + \sum_{j=1}^k \alpha_j \Delta_s z_{t-j} + \sum_{i=0}^{s+k} \sum_{j=1}^q \gamma_{ij} D(T_0^j)_{t-i} + u_t \quad (5.5)$$

em que  $D_{jt}$  são as variáveis *dummies* sazonais e  $D(T_0^j)_t$ ,  $j = 1, \dots, q$  são as variáveis *dummies* capturando os  $q$  *outliers* que foram encontrados e excluindo os lags insignificantes de  $\Delta_s z_t$ , a fim de aumentar o poder. Esta estratégia implica em, para cada *outlier* detectado, temos que descartar  $(s+1+k)$  observações. Por exemplo, para dados trimestrais com oito lags significantes, cada *outlier* implica na exclusão de treze observações ( $s = 4$  e  $k = 8$ ). Se houver muitos *outliers* na amostra, muitos dados serão descartados, o que causará uma relevante perda de poder, mesmo que o tamanho se aproxime do correto.

Uma alternativa é considerar a correspondente observação  $z_{T_0^j}$  como “perdida” e substituí-la pela previsão  $\hat{z}_{T_0^j}$ , após os *outliers* terem sido detectados. Dessa forma, teremos a nova série:

$$x_t = \begin{cases} z_t, & \text{se } t \notin \{T_0^1, T_0^2, \dots, T_0^q\} \\ \hat{z}_{T_0^j}, & \text{se } t \in \{T_0^1, T_0^2, \dots, T_0^q\} \end{cases} \quad t = 1, \dots, T$$

Assim, a regressão do teste HEGY para raízes unitárias sazonais poderá ser escrita como:

$$\Delta_s x_t = \mu_s + \sum_{j=1}^{s-1} \mu_j D_{jt} + \beta t + \sum_{j=1}^s \pi_j^s x_{j,t-1}^s + \sum_{j=1}^k \alpha_j \Delta_s x_{t-j} + u_t \quad (5.6)$$

que difere de (5.5) por não conter o penúltimo termo de variáveis *dummies*.

### ➤ Análise do tamanho e do poder do teste HEGY com correção para *outliers*

A fim de avaliar os poderes e os tamanhos dos testes HEGY quando *outliers* estão presentes, Haldrup et al (2005) conduziram um experimento de Monte Carlo, utilizando quatro métodos para construir  $x_t$  e a regressão auxiliar do teste HEGY. A saber:

**Método 1 (M1)** – Baseado em (5.6) com  $x_t = z_t$ , para todo t. Este método é equivalente a executar o teste HEGY sem considerar que *outliers* estão presentes. Os lags de  $\Delta_s x_{t-j}$  incluídos na regressão auxiliar foram os que se apresentaram significativos quando o modelo (5.6) foi estimado.

**Método 2 (M2)** – Baseado em (5.5), após terem sido identificados os *outliers* através do teste modificado de Perron-Rodriguez. Os lags incluídos na regressão auxiliar foram os que se apresentaram significativos quando o modelo (5.5) foi estimado.

**Método 3 (M3)** – Baseado em (5.6) com as previsões das observações “perdidas” dadas pelo modelo, sob a hipótese nula, dado por:  $\hat{z}_{T_0^j} = z_{T_0^j-s} + c_t$ , em que  $c_t$  é um termo que depende de termos determinísticos. Por exemplo, se o processo tem uma tendência, então sobre a hipótese nula,  $z_t - z_{t-s} = \mu + u_t$  e c é estimado como a média de  $(z_t - z_{t-s})$ , uma vez que os *outliers*

foram excluídos (isto é, para cada outlier  $T_0^j$  não utilizaremos nem  $z_{T_0^j} - z_{T_0^j-s}$  nem  $z_{T_0^j+s} - z_{T_0^j}$ ). Os lags foram selecionados como no Método 1.

**Método 4 (M4)** – Baseado em (5.6) com as previsões das observações “perdidas” dadas por um modelo  $AR(1)_s$  ( $\hat{z}_{T_0^j} = \hat{\phi} z_{T_0^j-s} + c_t$ ).  $\hat{\phi}$  e  $c_t$  foram calculados sem considerar os *outliers*.

O poder e o tamanho dos testes resultantes das simulações de Monte Carlo foram avaliados considerando os métodos de 1 a 4 no caso de *outliers* aditivos. É esperado que os resultados obtidos sejam similares aos resultados que seriam obtidos para *outliers* do tipo mudança temporária. O seguinte processo gerador dos dados trimestrais ( $s = 4$ ) foi utilizado:

$$(1 - \phi_1 L)(1 + \phi_2 L)(1 + \phi_3 L^2)y_t = \varepsilon_t$$

$$\varepsilon_t \sim N(0,1), z_t = y_t + v_t \text{ e } v_t = \theta \delta_t$$

O tamanho da amostra utilizado foi  $T = 120$  e, na implementação dos métodos, foi utilizado  $k = 12$  como valor prévio do lag no teste HEGY. As regressões auxiliares incluíram um intercepto e variáveis *dummies* sazonais. O procedimento de Monte Carlo foi repetido 3.000 vezes.

A Tabela 5.4 apresenta os tamanhos dos testes considerando níveis nominais de 5%,  $\phi_1 = \phi_2 = \phi_3 = 1$ ,  $\theta = 5$  e ocorrências de *outliers* com probabilidade  $\frac{p}{2} = 0,025$  cada, que corresponde a  $p\theta^2 = 1,25$ . Dessa forma, considerando  $T = 120$  observações, eram esperados 6 *outliers* aditivos. Vale citar que os métodos M3 e M4, isto é, os métodos nos quais os valores *missings* foram removidos e substituídos por previsões de passeio aleatório sazonal ou por previsões estimadas pelo modelo  $(1 - \phi B^4)z_t$ , produzem resultados quase idênticos e, portanto, somente os resultados para M3 são apresentados na Tabela 5.4.

Tabela 5.4 – Tamanhos dos testes HEGY com correções para *outliers*

Sizes at nominal 5% levels of the HEGY testing schemes M1–M3 in case of additive outliers. $T = 120$ observations, $p = 0.05$ , and $\theta = 5$ For the “no outlier”-case $\theta = 0$ .			
	$t_1$	$t_2$	$F$
M1	0.214	0.213	0.229
M2	0.101	0.112	0.114
M3	0.111	0.112	0.111
No outliers	0.073	0.091	0.071

Tabela extraída de Haldrup et al (2005)

A análise da Tabela 5.4 permite observar que as distorções nos tamanhos dos testes são maiores quando executamos os testes HEGY sem considerar os *outliers* presentes na série – trata-se do método 1. A razão pela qual os tamanhos obtidos para o caso “sem *outliers*” serem bastante diferentes de 5% é devido ao pré-teste de seleção dos lags na regressão auxiliar do teste HEGY. Comparados ao M1, os métodos M2 e M3 obtêm sucesso ao ajustar o tamanho ao tamanho nominal, embora os testes ainda sejam distorcidos.

Nas Figuras 5.1, 5.2 e 5.3 temos as curvas de poder para os testes HEGY  $t_1$ ,  $t_2$  e F, onde os parâmetros  $\phi$  variam, permanecendo as raízes sazonais restantes iguais a um.

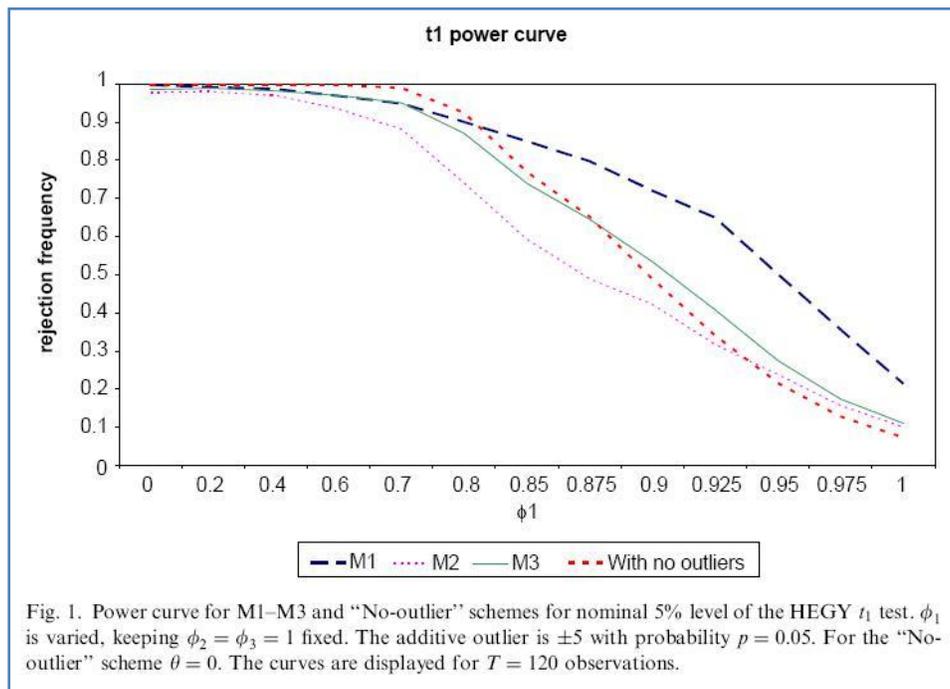


Figura 5.1 – Curvas de poder para a estatística  $t_1$  dos testes HEGY baseados nos diferentes métodos

Figura extraída de Haldrup, Montanés e Sanso (2005)

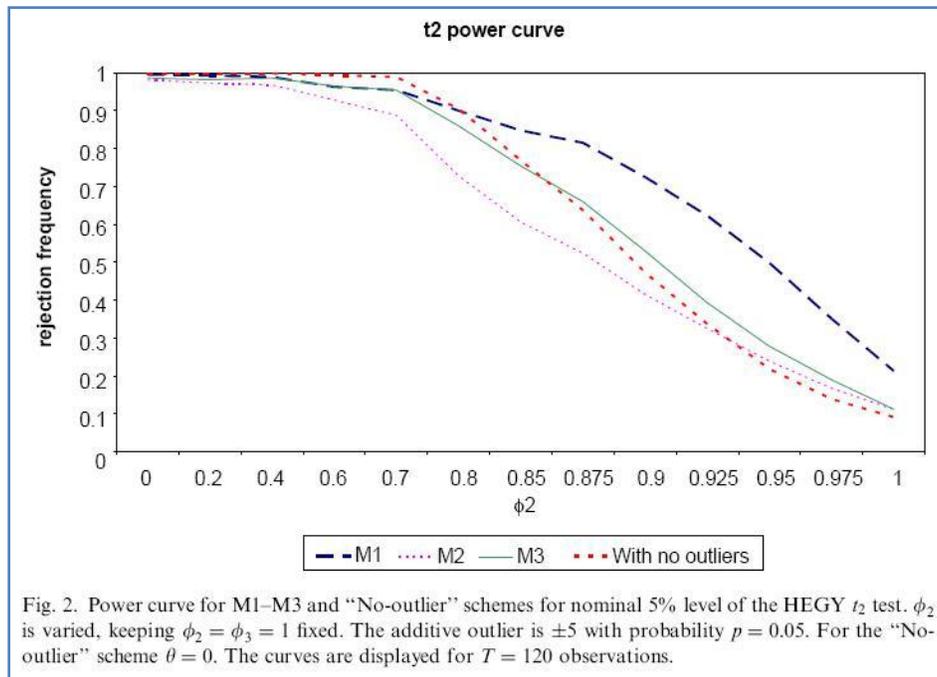


Figura 5.2 – Curvas de poder para a estatística  $t_2$  dos testes HEGY baseados nos diferentes métodos

Figura extraída de Haldrup, Montanés e Sanso (2005)

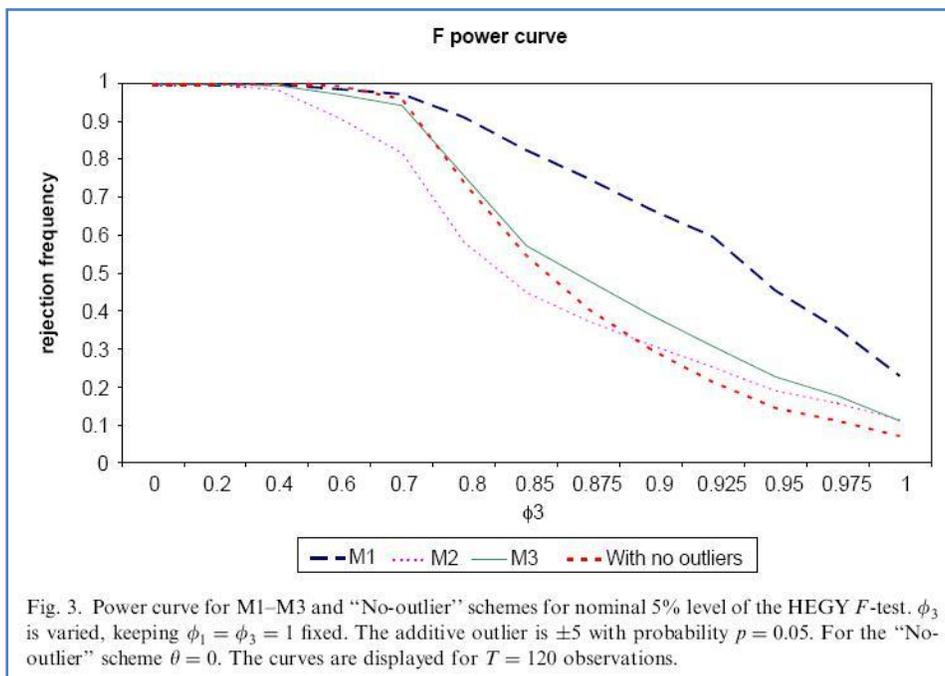


Figura 5.3 – Curvas de poder para a estatística  $F$  dos testes HEGY baseados nos diferentes métodos

Figura extraída de Haldrup, Montanés e Sanso (2005)

Quanto ao poder, o método M1 se apresenta mais favorável, devido à distorção do tamanho desse procedimento. O modelo “benchmark”, para efeito de comparação, é o modelo sem *outlier*. Comparado a ele, o método M2 é o menos favorável. Vale citar que todos os testes apresentam uma perda de poder considerável por incluírem variáveis *dummies* para substituir os *outliers* que foram identificados. Substituir os *outliers* por previsões razoáveis (como em M3 e em M4) melhora muito o poder e torna os testes comparáveis ao caso sem *outlier*.

## Capítulo 6

# Efeitos de Quebras Estruturais nos Testes HEGY

Neste Capítulo avaliaremos a robustez assintótica do procedimento HEGY na presença de quebras estruturais (mudança de magnitude finita na média sazonal) e apresentaremos alguns testes que consideram a presença dessas quebras.

### 6.1 Testes para Raízes Unitárias Sazonais sob Quebras Estruturais

É bastante comum séries temporais econômicas apresentarem alterações no comportamento da sazonalidade. Essas alterações, as quais chamaremos de quebras estruturais, se não forem consideradas podem implicar em distorções dos testes de raízes unitárias sazonais, como foi demonstrado por Hassler e Rodrigues (2004).

Consideremos, sob a hipótese nula, a existência de quebras nas médias sazonais de tamanho  $\tau_s$  no tempo  $T_B = \lambda T$ , tal que a série trimestral  $y_t$  pode ser representada pelo seguinte modelo:

$$y_t = \sum_{s=1}^4 \delta_s D_{s,t} + \frac{\gamma}{4} t + \sum_{s=1}^4 \tau_s D(T_B)_{s,t} + x_t \quad (6.1)$$

com  $\Delta_4 x_t = \varepsilon_t$ ,  $t = 1, 2, \dots, T$ ,  $D_{s,t}$  são variáveis *dummies* sazonais com  $D_{s,t} = 1$ , quando  $t$  se refere ao  $s$ -ésimo trimestre do ano e  $D_{s,t} = 0$ , caso contrário,  $\varepsilon_t$  é iid  $(0, \sigma^2)$  e  $\Delta_4 = 1 - B^4$ . Além disso,  $D(T_B)_{s,t} = D_{s,t}$ , se  $t > T_B$  e  $D(T_B)_{s,t} = 0$ , caso contrário.

A equação em (6.1) pode ser escrita na forma:

$$\Delta_4 y_t = \gamma + \sum_{s=1}^4 \tau_s \Delta_4 D(T_B)_{s,t} + \varepsilon_t$$

em que  $\Delta_4 D(T_B)_{s,t} = 0$  para todo  $t$  com exceção de um.

Primeiramente, vamos apresentar alguns testes desenvolvidos com o propósito de se considerar as quebras estruturais existentes em uma série temporal.

### 6.1.1 O Teste HEGY AO

Smith e Otero (1997) propuseram um procedimento de teste que se baseia em uma correção do teste HEGY, considerando quebras estruturais sazonais em unidades de tempo conhecidas a priori. Eles sugeriram remover todas as componentes determinísticas em uma primeira regressão, utilizando o seguinte modelo:

$$y_t = \sum_{s=1}^4 \hat{\delta}_s D_{s,t} + \sum_{s=1}^4 \hat{\tau}_s D(T_B)_{s,t} + \hat{\gamma}t + \tilde{y}_t \quad (6.2)$$

em que  $D_{s,t}$  e  $D(T_B)_{s,t}$  são variáveis *dummies* sazonais como definidas em (6.1).

Na sequência, aplica-se as diferenças usuais do teste HEGY aos resíduos obtidos de (6.2), de tal forma que:

$$\tilde{x}_{1,t} = (1 + B + B^2 + B^3)\tilde{y}_t,$$

$$\tilde{x}_{2,t} = -(1 - B + B^2 - B^3)\tilde{y}_t,$$

$$\tilde{x}_{3,t} = -(1 - B^2)\tilde{y}_t.$$

A equação do teste será dada por:

$$\Delta_4 \tilde{y}_t = \hat{\pi}_1 \tilde{x}_{1,t-1} + \hat{\pi}_2 \tilde{x}_{2,t-1} + \hat{\pi}_3 \tilde{x}_{3,t-2} + \hat{\pi}_4 \tilde{x}_{3,t-1} + \hat{\epsilon}_t.$$

Smith e Otero (1997, Tabela 3, p. 18) apresentaram os valores críticos para o teste baseado em amostras de 1.000 observações, considerando diferentes instantes de quebras estruturais.

### 6.1.2 O Teste HEGY IO

Franses e Hobijn (1997) propuseram um procedimento de teste que se baseia em um modelo IO, onde quebras estruturais ocorrem em unidades de tempo conhecidas. A regressão auxiliar do procedimento HEGY, nesse caso, é dada por:

$$\Delta_4 y_t = \sum_{s=1}^4 \hat{d}_s D_{s,t} + \sum_{s=1}^4 \hat{t}_s D(T_B)_{s,t} + \hat{c}t + \hat{\pi}_1 x_{1,t-1} + \hat{\pi}_2 x_{2,t-1} + \hat{\pi}_3 x_{3,t-2} + \hat{\pi}_4 x_{3,t-1} + \hat{\epsilon}_t. \quad (6.3)$$

Os valores críticos foram obtidos utilizando amostras de tamanho  $T = 80$  e podem ser encontrados em Franses e Hoijn (1997, Tabelas 14-18, pp. 41-42).

É importante comentar que, na prática, não sabemos a priori qual modelo AO ou IO é mais adequado e, por isso, é necessário calcular, utilizando Monte Carlo, o efeito de uma má especificação.

### 6.1.3 O Teste LM - HEGY

Breitung e Franses (1998) propuseram um procedimento de teste que se baseia em uma variante de Multiplicador de Lagrange (LM) do teste HEGY. Vamos apresentar o procedimento do teste para dados trimestrais ( $s = 4$ ).

O parâmetro de tendência linear  $\gamma$  é estimado como a média de diferenças anuais,  $\Delta_4 y_t$ , utilizando o modelo do teste HEGY sem considerar quebras estruturais dado por:

$$y_t = \sum_{s=1}^4 \delta_s D_{s,t} + \frac{\gamma}{4} t + x_t$$

com  $\Delta_4 x_t = \epsilon_t$ ,  $t = 1, 2, \dots, T$ .

Seja  $\tilde{x}_t$  a série de diferenças corrigidas pela média, tal que:

$$\tilde{x}_t = \Delta_4 y - \overline{\Delta_4 y}$$

e sejam as variáveis regressoras calculadas recursivamente, tais que:

$$\tilde{x}_{1,t} = \tilde{x}_{1,t-1} + \tilde{x}_t,$$

$$\tilde{x}_{2,t} = -\tilde{x}_{2,t-1} - \tilde{x}_t, \quad (6.4)$$

$$\tilde{x}_{3,t} = -\tilde{x}_{3,t-2} - \tilde{x}_t,$$

em que  $\tilde{x}_{k,t} = 0$ , para  $t \leq 4$  e  $k = 1, 2, 3$ . Dessa forma, a regressão HEGY é dada por:

$$\tilde{x}_t = \hat{\pi}_1 \tilde{x}_{1,t-1} + \hat{\pi}_2 \tilde{x}_{2,t-1} + \hat{\pi}_3 \tilde{x}_{3,t-2} + \hat{\pi}_4 \tilde{x}_{3,t-1} + \hat{\epsilon}_t, \quad t = 5, 6, \dots, T \quad (6.5)$$

em que as estatísticas  $t$  sobre  $\hat{\pi}_k$ ,  $k = 1, 2, 3, 4$  são definidas como  $t_{\pi_k}$ .

Para testar o par de raízes unitárias complexas, Breitung e Franses (1998) sugerem calcular a seguinte estatística de teste:

$$\Phi_{\pi_3, \pi_4} = t_{\pi_3}^2 + t_{\pi_4}^2$$

que equivale, assintoticamente, a duas vezes a estatística “F” para o teste padrão HEGY exposto em (3.13).

Schmidt e Lee (1991) e Schmidt e Phillips (1992) sugerem que o poder dos testes de raízes unitárias LM pode ser aumentado acrescentando variáveis regressoras na regressão auxiliar. Portanto, o seguinte modelo é sugerido:

$$\tilde{x}_t = \sum_{s=1}^4 \hat{d}_s D_{s,t} + \hat{\pi}_1 \tilde{x}_{1,t-1} + \hat{\pi}_2 \tilde{x}_{2,t-1} + \hat{\pi}_3 \tilde{x}_{3,t-2} + \hat{\pi}_4 \tilde{x}_{3,t-1} + \hat{\epsilon}_t. \quad (6.6)$$

Breitung e Franses (1998, Tabela 1, p. 210) apresentaram os valores críticos para alguns tamanhos de amostras.

#### 6.1.4 O Teste LM – HEGY – AO

Hassler e Rodrigues (2004), a partir do teste LM-HEGY, desenvolveram um procedimento de teste que considera quebras estruturais presentes em instantes conhecidos a priori.

O procedimento de teste LM – HEGY – AO, consiste em, primeiramente, estimar por mínimos quadrados o parâmetro de tendência linear  $\gamma$  e as mudanças na média  $\tau_s$ ,  $s = 1, 2, 3, 4$ , considerando a equação expressa em (6.1). Assim, temos:

$$\Delta_4 y_t = \tilde{\gamma} + \sum_{s=1}^4 \tilde{\tau}_s \Delta_4 D(T_B)_{s,t} + \tilde{x}_t \quad t = 5, 6, \dots, T \quad (6.7)$$

Na sequência, os resíduos são utilizados para calcular as variáveis regressoras do procedimento HEGY dadas em (6.4), com valores iniciais  $\tilde{x}_{k,t} = 0$ , para  $t \leq 4$  e  $k = 1, 2, 3$ . Finalmente, a regressão em (6.6) é estimada por mínimos quadrados, de acordo com a nova definição de  $\tilde{x}_t$ .

É importante citar que a eliminação de quatro *outliers* através de  $\sum_{s=1}^4 \tilde{\tau}_s \Delta_4 D(T_B)_{s,t}$  em (6.7) não afeta a distribuição limite das estatísticas do teste, isto é, as estatísticas  $t_{\pi_1}$ ,  $t_{\pi_2}$  e  $F_{\pi_3, \pi_4}$  tem as mesmas distribuições estudadas no Capítulo 3 e, portanto, com os valores críticos das Tabelas 3.1 e 3.2. Para mais detalhes, sugerimos ver Hassler e Rodrigues (2004).

## 6.2 Distribuição Assintótica das Estatísticas do Teste LM – HEGY – AO

### 6.2.1 Resultados Preliminares

Consideremos o modelo dado por:

$$y_t = \sum_{s=1}^4 \delta_s D_{s,t} + \frac{\gamma}{4} t + \sum_{s=1}^4 \tau_s D(T_B)_{s,t} + x_t, \quad t = 1, 2, \dots, T \quad (6.8)$$

$$e \quad x_t = \alpha x_{t-4} + \varepsilon_t \quad (6.9)$$

com  $\varepsilon_t \sim iid(0, \sigma^2)$ , em que  $D_{s,t}$  e  $D(T_B)_{s,t}$  são definidos como em (6.1). Para simplificar, assumi-se que as quebras estruturais ocorrem no início do ano, isto é,  $N_B + 1 = \frac{T_B}{4}$ .

O modelo em (6.8) pode ser escrito considerando a representação anual, usando  $t = 4(n - 1) + s$ , onde  $n$  denota o ano e  $s$  o trimestre. Dessa forma, temos:

$$y_{s,n} = \delta_s + \gamma(n - 1) + \gamma \frac{s}{4} + \tau_s D(N_B)_{s,n} + x_{s,n}, \quad n = 1, 2, \dots, N = \frac{T}{4} \quad e$$

$$x_{s,n} = \alpha x_{s,n-1} + \varepsilon_{s,n}$$

em que  $D(N_B)_{s,n}$  é uma função indicadora que assume valor um quando  $n > N_B$ ,  $s = 1, \dots, 4$  e  $n = \left\lfloor \frac{(t-1)}{4} \right\rfloor + 1$ .

Sob a hipótese nula ( $\alpha = 1$ ) em (6.9) e assumindo que  $y_1$  ocorre no primeiro trimestre do primeiro ano, obtemos para  $n = 1$ :

$$y_{s,1} = \delta_s + \gamma \frac{S}{4} + \tau_s D(N_B)_{s,1} + x_{s,0} + \varepsilon_{s,1} \quad (6.10)$$

e, para  $n = 2, 3, \dots, N$ , obtemos:

$$y_{s,n} = \gamma + \tau_s \Delta D(N_B)_{s,n} + y_{s,n-1} + \varepsilon_{s,n} \quad (6.11)$$

em que o operador diferença agora se aplica ao índice anual:

$$\Delta x_{s,n} = x_{s,n} - x_{s,n-1}$$

Para mais detalhes, ver Hassler e Rodrigues (2004).

## 6.2.2 Resultados Assintóticos

Vamos apresentar os limites assintóticos do teste LM – HEGY - AO no contexto de quebras de magnitude crescente,  $\tau_s = k_s \sqrt{N}$ . Note que os resultados para quebras de magnitude finita podem ser obtidos eliminando os termos que envolvem  $k_s$  das expressões assintóticas. Para as demonstrações dos resultados que seguirão, sugerimos ler os autores Hassler e Rodrigues (2004).

**Resultado 6.1.** (Hassler e Rodrigues, 2004) *Assumindo os modelos expressos em (6.10) e (6.11), quando  $T = 4N \rightarrow \infty$ , podemos estabelecer que:*

$$(i) \quad \frac{1}{\sqrt{N}} \tilde{y}_{s,[rN]} \xrightarrow{D} \sigma \left\{ [W_s(r) - rW(1)] + \left[ \frac{k_s}{\sigma} d(\lambda, r) - \frac{r}{4\sigma} \sum_{s=1}^4 k_s \right] \right\}$$

$$\begin{aligned}
\text{(ii)} \quad & \frac{1}{N^2} \sum_{n=2}^N \tilde{y}_{s,n} \xrightarrow{D} \sigma \int_0^1 \left\{ V_s(r) + \left[ \frac{k_s}{\sigma} d(\lambda, r) - \frac{r}{4\sigma} \sum_{s=1}^4 k_s \right] \right\} dr, \\
\text{(iii)} \quad & \frac{1}{N^2} \sum_{n=2}^N \tilde{y}_{s,n}^2 \xrightarrow{D} \sigma^2 \int_0^1 \left\{ V_s(r) + \left[ \frac{k_s}{\sigma} d(\lambda, r) - \frac{r}{4\sigma} \sum_{s=1}^4 k_s \right] \right\}^2 dr, \\
\text{(iv)} \quad & \frac{1}{N} \sum_{n=3}^N \Delta \tilde{y}_{s,n} \tilde{y}_{s,n-1} \xrightarrow{D} \sigma^2 \int_0^1 \left\{ V_s(r) + \left[ \frac{k_s}{\sigma} d(\lambda, r) - \frac{r}{4\sigma} \sum_{s=1}^4 k_s \right] \right\} dW_s(r),
\end{aligned}$$

$$\tilde{y}_{s,n} = y_{s,n} - y_{s,1} - \tilde{\gamma}(n-1) - \tilde{\tau}_s [D(N_B)_{s,n} - D(N_B)_{s,1}],$$

$$\Delta \tilde{y}_{s,n} = \Delta y_{s,n} - \tilde{\gamma} - \tilde{\tau}_s \Delta D(N_B)_{s,n},$$

$$V_s(r) \equiv W_s(r) - rW(1), \quad W(1) = \frac{1}{4} \sum_{s=1}^4 W_s(1),$$

$W_s(r)$  para  $s = 1, 2, 3, 4$  são Movimentos Brownianos Padrões independentes e  $d(\lambda, r) = 1$ , quando  $r > \lambda$ .

### 6.2.3 O Teste LM – HEGY – AO

A regressão do teste para dados trimestrais pode ser escrita como:

$$\Delta_4 \tilde{y}_t = \pi_1 \tilde{x}_{1,t-1} + \pi_2 \tilde{x}_{2,t-1} + \pi_3 \tilde{x}_{3,t-2} + \pi_4 \tilde{x}_{3,t-1} + \epsilon_t. \quad (6.12)$$

em que

$$\tilde{x}_{1,t} = (1 + B + B^2 + B^3) \tilde{y}_t,$$

$$\tilde{x}_{2,t} = -(1 - B + B^2 - B^3) \tilde{y}_t \text{ e}$$

$$\tilde{x}_{3,t} = -(1 - B^2) \tilde{y}_t,$$

que podem ser reescritas como:

$$\begin{aligned}\tilde{x}_{1,t} &= \tilde{y}_{1,n} + \tilde{y}_{2,n} + \tilde{y}_{3,n} + \tilde{y}_{4,n} + Op(1), \\ \tilde{x}_{2,t} &= \tilde{y}_{1,n} - \tilde{y}_{2,n} + \tilde{y}_{3,n} - \tilde{y}_{4,n} + Op(1), \text{ t mod } 2 = 1, \\ &= -\tilde{y}_{1,n} + \tilde{y}_{2,n} - \tilde{y}_{3,n} + \tilde{y}_{4,n} + Op(1), \text{ t mod } 2 = 0, \\ \tilde{x}_{3,t} &= -\tilde{y}_{1,n} + \tilde{y}_{3,n} + Op(1), \text{ t mod } 4 = 1, \\ &= -\tilde{y}_{2,n} + \tilde{y}_{4,n} + Op(1), \text{ t mod } 4 = 2, \\ &= \tilde{y}_{1,n} - \tilde{y}_{3,n} + Op(1), \text{ t mod } 4 = 3, \\ &= \tilde{y}_{2,n} - \tilde{y}_{4,n} + Op(1), \text{ t mod } 4 = 0,\end{aligned}$$

em que  $n = \left\lfloor \frac{(t-1)}{4} \right\rfloor + 1$ .

Hassler e Rodrigues (2004) mostraram que as distribuições das estatísticas “t” e “F” do parâmetro  $\pi_1$  é obtida de Schmidt e Lee (1991, p. 263), enquanto as distribuições das estatísticas “t” e “F” dos parâmetros  $\pi_2$ ,  $\pi_3$  e  $\pi_4$  são equivalentes às distribuições obtidas da regressão do teste HEGY sem componentes determinísticas, apresentadas na Seção 3.2.

### 6.3 Impactos da Não Consideração de Quebras Estruturais

Nessa seção, veremos as implicações causadas pela negligência de quebras nos procedimentos HEGY e LM-HEGY. Além disso, serão apresentados os efeitos assintóticos causados por não assumir a existência de pontos de quebras no teste LM – HEGY - AO. É importante comentar que serão apresentados somente os limites de quebras de magnitude crescente  $\tau_s = k_s \sqrt{N}$ , porém, os resultados para quebras de magnitude finita podem ser obtidos eliminando os termos que envolvem  $k_s$  das expressões assintóticas.

#### 6.3.1 Impactos no Teste HEGY

Vamos considerar as expressões em (6.10) e (6.11) com  $\delta_s = 0, s = 1, \dots, 4$  e  $\gamma = 0$ , para facilitar a notação. Assumamos que há uma quebra cuja magnitude cresce conforme o tamanho da amostra, isto é,  $\tau_s = k_s \sqrt{N}$ . Assim,  $y_{s,n}$  pode ser escrito como:

$$y_{s,n} = k_s \sqrt{N} \sum_{i=2}^n \Delta D(N_B)_{s,i} + y_{s,0} + \sum_{i=2}^n \varepsilon_{s,i} \quad n = 2, 3, \dots, N.$$

Consequentemente, as implicações assintóticas dessa quebra se encontram no Teorema a seguir.

**Teorema 6.1.** (Hassler e Rodrigues, 2004) *Considerando as expressões em (6.10) e (6.11) com  $\tau_s = k_s \sqrt{N}$ , então para  $N \rightarrow \infty$ , temos:*

$$(i) \quad \frac{1}{N^2} \sum_{n=2}^N y_{s,n} \xrightarrow{\mathcal{D}} \sigma \left[ k_s \lambda + \int_0^1 W_s(r) dr \right],$$

$$(ii) \quad \frac{1}{N^2} \sum_{n=2}^N y_{s,n}^2 \xrightarrow{\mathcal{D}} \sigma^2 \left[ k_s^2 \lambda + \int_0^1 W_s^2(r) dr + 2k_s \int_\lambda^1 W_s(r) dr \right],$$

$$(iii) \quad \frac{1}{N} \sum_{n=2}^N \Delta y_{s,n} y_{s,n-1} \xrightarrow{\mathcal{D}} \sigma^2 \left[ k_s^2 + k_s W_s(r) + k_s [W_s(1) - W_s(\lambda)] + \int_0^1 W_s(r) dW_s(r) \right],$$

em que  $W_s(r)$ ,  $s = 1, \dots, 4$ , são Movimentos Brownianos Padrões independentes e  $\lambda$  representa a posição onde a quebra ocorre na série que está sendo avaliada.

Do Teorema 6.1 vem que as distribuições assintóticas das estatísticas do teste HEGY são impactadas quando estão presentes na série quebras estruturais de magnitude crescente. No entanto, se excluirmos os termos nos quais estão presentes as quantidades  $k_s$ , temos os resultados para as quebras de magnitudes finitas e, portanto, podemos observar que estas não impactam as distribuições assintóticas das estatísticas do teste HEGY.

### 6.3.2 Impactos no Teste LM - HEGY

Consideremos o modelo expresso em (6.8) e (6.9) e  $\tau_s = k_s \sqrt{N}$ . Assim, quando  $\alpha = 1$ , temos:

$$\tilde{\gamma} = \gamma + \frac{\sqrt{N}}{4(N-1)} \sum_{s=1}^4 k_s + \bar{\varepsilon}_s \quad e$$

$$\tilde{y}_{s,n} = \sum_{k=2}^n (\varepsilon_{s,k} - \bar{\varepsilon}_s) + \sum_{k=2}^n k_s \sqrt{N} \Delta D(N_B)_{s,k} - \frac{(n-1)\sqrt{N}}{4(N-1)} \sum_{s=1}^4 k_s$$

As implicações assintóticas dessa quebra nos testes LM - HEGY se encontram no Teorema a seguir.

**Teorema 6.2.** (Hassler e Rodrigues, 2004) *Considerando as expressões em (6.10) e (6.11) com  $\tau_s = k_s \sqrt{N}$ , então quando  $N \rightarrow \infty$ , temos:*

$$(i) \quad \frac{1}{\sqrt{N}} \tilde{y}_{s[rN]} \xrightarrow{D} \sigma \left[ V_s(r) + k_s d(\lambda, r) - \frac{1}{4} \sum_{s=1}^4 r k_s \right],$$

$$(ii) \quad \frac{1}{N^2} \sum_{n=2}^N \tilde{y}_{s,n} \xrightarrow{D} \sigma \left[ \int_0^1 V_s(r) dr + k_s \lambda - \frac{1}{8} \sum_{s=1}^4 k_s \right],$$

$$(iii) \quad \frac{1}{N^2} \sum_{n=2}^N (\tilde{y}_{s,n})^2 \rightarrow \sigma^2 \left[ \int_0^1 V_s^2(r) dr + 2k_s \int_\lambda^1 V_s(r) dr - \frac{1}{2} \sum_{s=1}^4 k_s \int_0^1 r V_s(r) dr + k_s^2 \lambda - \frac{1}{2} (1 - \lambda^2) k_s \sum_{s=1}^4 k_s + \frac{1}{12} \sum_{s=1}^4 k_s^2 \right],$$

$$(iv) \quad \frac{1}{N} \sum_{n=2}^N \Delta \tilde{y}_{s,n} \tilde{y}_{s,n-1} \\ \rightarrow \sigma^2 \left[ \int_0^1 V_s(r) dV_s(r) + k_s [V_s(1) - V_s(\lambda)] - \frac{1}{4} \sum_{s=1}^4 k_s \int_0^1 r dV_s(r) + k_s V_s(\lambda) + k_s^2 - \lambda \frac{1}{2} k_s \sum_{s=1}^4 k_s \right. \\ \left. - \frac{1}{4} \sum_{s=1}^4 k_s \int_0^1 V_s(r) dr + \frac{1}{2} \left( \frac{1}{4} \sum_{s=1}^4 k_s \right)^2 \right],$$

em que

$$V_s(r) = W_s(r) - rW(1) \quad e$$

$$W(1) = \frac{1}{4} \sum_{s=1}^4 W_s(1).$$

Assim, temos que de forma análoga ao observado para o teste HEGY, o Teorema 6.2 mostra que as distribuições limites do teste LM – HEGY também são impactadas pelas quebras de magnitude crescentes, assim como também não são impactadas pelas quebras finitas.

## 6.4 Impactos da Consideração Errônea de Quebras Estruturais no Teste LM – HEGY - AO

Estudemos, agora, os efeitos assintóticos causados pela especificação errônea de uma quebra estrutural no teste LM – HEGY - AO, isto é, as consequências de se considerar que existe uma quebra em um determinado instante, em que não há quebras relevantes.

**Teorema 6.3.** (Hassler e Rodrigues, 2004) *Considerando o modelo expresso em (6.8) e (6.9) com  $\alpha = 1$  e quebra verdadeira ocorrendo em  $T_B = \lambda T$  e supondo que estimamos (6.12) considerando  $T_A = lT$ , com  $\lambda \neq l$ . Então, quando  $T \rightarrow \infty$ , os resultados apresentados nos Teoremas 6.1 e 6.2 ainda são verdadeiros.*

De acordo com os resultados do Teorema 6.3, as estatísticas do teste LM – HEGY - AO não são afetadas pela especificação errônea de quebras estruturais finitas, mesmo que o ponto exato da quebra não seja conhecido a priori.

Para as demonstrações dos Teoremas e Resultados apresentados nesse Capítulo, sugerimos ler os autores Hassler e Rodrigues (2004).

## 6.5 Simulações de Monte Carlo

Hassler e Rodrigues (2004) conduziram um experimento de Monte Carlo para avaliar o desempenho do teste LM – HEGY - AO quanto às propriedades de amostras finitas. Para tal, um modelo AR(1) com AO no tempo  $T_B = \lambda T$  foi simulado como a seguir:

$$y_t = x_t + \sum_{s=1}^4 \tau_s D(T_B)_{s,t},$$

$$x_t = \alpha x_{t-4} + \varepsilon_t \quad \text{e} \quad (6.13)$$

$$\varepsilon_t \sim iid N(0,1), \quad t = 1, 2, \dots, T.$$

Se  $\alpha = 1$ , então  $\pi_1 = 0$  e existe uma raiz na frequência zero (presença de tendência estocástica). De forma similar, as mudanças na média podem ser modeladas pelo parâmetro  $\tau$ . Avaliaremos seis casos, a saber:

**(i)**  $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau$ ;

**(ii)**  $\tau_1 = \tau, \tau_2 = \tau_3 = \tau_4 = 0$ ;

**(iii)**  $\tau_1 = -\tau_2 = \tau, \tau_3 = \tau_4 = 0$ ;

**(iv)**  $\tau_1 = \tau_2 = \tau, \tau_3 = \tau_4 = 0$ ;

**(v)**  $\tau_1 = \tau_2 = \tau_3 = \tau, \tau_4 = 0$ ;

**(vi)**  $\tau_1 = -\tau_2 = \tau_3 = -\tau_4 = \tau$ .

Em (i), todos os trimestres são afetados pela mesma quebra, enquanto em (ii), somente um trimestre é afetado. Nos casos (iii) e (iv), dois trimestres são afetados com sinais opostos e dois são afetados com sinais iguais, respectivamente. No caso (v), com exceção de um trimestre, todos os demais são afetados identicamente e, no caso (vi), todos os trimestres são afetados com sinais opostos, de forma que não há efeito sobre a média geral.

A Tabela 6.1 contem os percentuais de rejeições obtidos considerando um nível de significância de 5% para testar a hipótese nula de  $\alpha = 1$  para  $\lambda = 0,5$  e  $\tau = 3$ . Os resultados para  $T = 100$  e  $T = 200$  foram obtidos por intermédio de 5.000 réplicas do processo e os testes utilizados foram o HEGY, LM – HEGY, HEGY – AO, HEGY – IO e LM – HEGY – AO, a fim de compará-los quanto ao desempenho em amostras finitas.

O teste HEGY utilizado se baseia em uma regressão auxiliar contendo variáveis *dummies* de sazonalidade e tendência, dada por:

$$\varphi(B)^* y_{13t} = \sum_{k=1}^4 \pi_k y_{k,t-1} + m_0 t + m_1 + \sum_{k=2}^4 m_k S_{kt} + \varepsilon_t.$$

Os valores críticos para as estatísticas do teste HEGY foram obtidos das Tabelas 3.1 e 3.2.

Para o teste LM – HEGY, as frequências de rejeição são baseadas em (6.5) e os valores críticos de amostras finitas foram obtidos de Schmidt e Phillips (1992) para  $\pi_1$  e das Tabelas 3.1 e 3.2 para  $\pi_2$ ,  $\pi_3$  e  $\pi_4$ .

Os testes HEGY – AO e HEGY – IO são baseados em (6.2) e (6.3), respectivamente, com valores críticos obtidos de Smith e Otero (1997) e Franses e Hobijn (1997) para  $T = 1000$  e  $T = 80$ , respectivamente.

O teste LM – HEGY – AO foi construído baseado em (6.6) e (6.9) e os mesmos valores críticos do teste LM – HEGY foram utilizados.

Os resultados da Tabela 6.1 para  $\tau = 3$  (que é uma quebra moderada de três vezes o desvio padrão) indicam que o modelo IO claramente não é apropriado quando a série apresenta AO's. Na maior parte dos casos, esse teste resultou em frequências de rejeição superiores a 10%, enquanto o nível de significância nominal do teste era de 5%. Embora os testes HEGY e LM – HEGY sejam assintoticamente robustos aos AO's, podem apresentar problemas de tamanho em amostras finitas. Em particular, para os casos (i), (v) e (vi), pode-se notar que esses testes são muito conservadores na frequência zero e  $\frac{\pi}{2}$ , enquanto super-rejeitam nas outras frequências. Para os casos (iii) e (iv), os testes são conservadores nas frequências  $\pi$  e zero, respectivamente. Conforme o tamanho da amostra cresce de  $T = 100$  a  $T = 200$ , o tamanho dos testes também se aproxima do nominal, embora distorções possam ser observadas.

Tabela 6.1 – Tamanhos dos testes considerando  $\alpha = 1$

5% EXPERIMENTAL LEVEL, $\tau = 3, \lambda = 0.5^*$						
	T = 100			T = 200		
	$t_{n_1}$	$t_{n_2}$	$F_{n_3, n_4}$	$t_{n_1}$	$t_{n_2}$	$F_{n_3, n_4}$
(i): $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau$						
HEGY	0.12	7.82	9.72	0.30	6.32	7.54
LM-HEGY	0.06	8.86	10.88	0.60	6.72	7.58
HEGY-AO	6.12	6.66	7.12	4.98	5.78	5.76
HEGY-IO	70.02	10.80	13.36	54.32	8.66	8.68
LM-HEGY-AO	4.40	4.64	5.90	4.62	5.06	5.40
(ii): $\tau_1 = \tau, \tau_2 = \tau_3 = \tau_4 = 0$						
HEGY	4.12	3.82	4.66	4.70	4.04	4.98
LM-HEGY	4.18	4.36	4.88	4.26	4.56	4.68
HEGY-AO	6.00	6.50	7.36	5.60	5.86	6.72
HEGY-IO	12.08	11.68	14.94	9.94	9.34	8.98
LM-HEGY-AO	5.10	5.40	5.86	4.44	5.38	5.40
(iii): $\tau_1 = -\tau_2 = \tau, \tau_3 = \tau_4 = 0$						
HEGY	6.78	2.42	4.12	5.76	2.84	4.62
LM-HEGY	7.44	2.50	4.80	6.14	3.00	4.78
HEGY-AO	5.70	6.74	7.08	5.58	5.62	6.16
HEGY-IO	9.28	28.22	22.56	7.52	18.74	12.94
LM-HEGY-AO	4.02	5.10	5.96	4.72	5.14	5.56
(iv): $\tau_1 = \tau_2 = \tau, \tau_3 = \tau_4 = 0$						
HEGY	2.28	5.80	3.92	3.06	5.32	4.42
LM-HEGY	2.68	6.52	4.74	3.46	5.42	5.00
HEGY-AO	6.08	6.68	7.12	5.40	5.90	5.78
HEGY-IO	26.06	8.00	24.78	20.24	7.68	13.12
LM-HEGY-AO	5.30	5.30	6.02	4.68	4.74	5.50
(v): $\tau_1 = \tau_2 = \tau_3 = \tau, \tau_4 = 0$						
HEGY	0.78	4.86	4.96	1.58	5.58	5.82
LM-HEGY	0.84	5.84	5.92	2.04	6.04	6.16
HEGY-AO	6.54	6.42	7.12	5.44	5.58	6.02
HEGY-IO	49.66	12.64	18.82	34.10	10.36	11.24
LM-HEGY-AO	4.66	5.22	5.44	4.82	5.20	5.34
(vi): $\tau_1 = -\tau_2 = \tau_3 = -\tau_4 = \tau$						
HEGY	9.88	0.12	9.74	7.78	0.70	7.00
LM-HEGY	11.66	0.12	10.76	8.64	0.78	7.08
HEGY-AO	5.86	6.44	7.26	5.96	5.76	5.62
HEGY-IO	12.14	69.52	13.12	9.44	47.44	8.76
LM-HEGY-AO	4.88	5.30	5.76	4.68	4.64	5.50
Critical values						
HEGY	-3.45	-2.89	6.60	-3.43	-2.88	6.57
LM-HEGY	-3.06	-2.89	6.60	-3.04	-2.88	6.61
HEGY-AO	-4.02	-3.50	10.19	-4.02	-3.50	10.19
HEGY-IO	-3.64	-3.25	9.01	-3.64	-3.25	9.01
LM-HEGY-AO	-3.06	-2.89	6.60	-3.04	-2.88	6.61

Tabela extraída de Hassler e Rodrigues (2004)

A Tabela 6.2 apresenta os poderes dos testes considerando  $\alpha = 0,8$  em (6.13), com exceção do teste HEGY – IO, que vimos não ser adequado aos dados.

Tabela 6.2 – Tamanhos dos testes considerando  $\alpha = 0,8$ 

5% EXPERIMENTAL POWER, $\tau = 3, \lambda = 0.5^*$						
	T = 100			T = 200		
	$I_{n_1}$	$I_{n_2}$	$F_{n_3, n_4}$	$I_{n_1}$	$I_{n_2}$	$F_{n_3, n_4}$
(i): $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau$						
HEGY	0.28	21.52	38.66	1.58	45.10	76.86
LM-HEGY	0.34	23.12	40.80	2.33	46.18	77.06
HEGY-AO	6.36	12.64	17.68	11.54	22.64	39.22
LM-HEGY-AO	9.88	12.28	18.98	24.72	30.70	54.32
(ii): $\tau_1 = \tau, \tau_2 = \tau_3 = \tau_4 = 0$						
HEGY	7.22	9.02	12.84	18.36	23.96	42.56
LM-HEGY	9.40	9.84	14.60	23.00	24.84	42.98
HEGY-AO	6.44	11.56	17.18	11.80	22.04	38.66
LM-HEGY-AO	10.50	11.84	18.86	24.62	30.30	54.02
(iii): $\tau_1 = -\tau_2 = \tau, \tau_3 = \tau_4 = 0$						
HEGY	10.98	2.28	8.54	23.24	6.32	27.78
LM-HEGY	13.44	2.84	9.52	28.10	6.84	28.42
HEGY-AO	6.92	11.74	17.70	11.58	22.28	39.26
LM-HEGY-AO	9.18	11.96	19.32	23.94	29.66	51.74
(iv): $\tau_1 = \tau_2 = \tau, \tau_3 = \tau_4 = 0$						
HEGY	3.38	15.94	8.62	10.88	38.80	28.65
LM-HEGY	4.32	16.84	9.78	14.90	39.40	28.60
HEGY-AO	6.50	11.80	17.56	11.14	22.40	38.80
LM-HEGY-AO	9.72	12.32	18.76	24.98	29.74	52.22
(v): $\tau_1 = \tau_2 = \tau_3 = \tau, \tau_4 = 0$						
HEGY	1.14	11.28	17.30	4.92	27.20	50.60
LM-HEGY	1.60	12.14	18.74	6.40	27.44	50.42
HEGY-AO	6.62	11.24	18.04	11.32	22.24	39.30
LM-HEGY-AO	9.76	11.40	19.60	24.48	29.56	53.42
(vi): $\tau_1 = -\tau_2 = \tau_3 = -\tau_4 = \tau$						
HEGY	14.80	0.00	35.92	26.26	0.10	74.68
LM-HEGY	17.96	0.00	38.72	33.36	0.08	75.06
HEGY-AO	6.62	11.38	17.78	10.68	22.70	38.22
LM-HEGY-AO	9.12	10.66	19.36	23.70	30.96	52.06

Tabela extraída de Hassler e Rodrigues (2004)

Essas simulações foram realizadas na tentativa de se encontrar melhores propriedades de poder resultantes dos testes HEGY e LM – HEGY, por serem conservadores sob a hipótese nula. Entretanto, a Tabela 6.2 mostra que em todas as vezes que esses testes se mostraram conservadores, sob a hipótese nula, eles se apresentaram com menor poder do que o teste LM – HEGY – AO. Além disso, o teste LM – HEGY – AO rejeita mais frequentemente todas as frequências nos casos (ii) e (v). Finalmente, embora o teste HEGY-AO super-rejeite sob a hipótese nula, há poucas situações em que ele se mostra mais poderoso do que LM – HEGY – AO. Em suma, as Tabelas 6.1 e 6.2 sugerem que o teste LM – HEGY – AO tem um desempenho melhor do que os demais, em termos de tamanho e poder, quando testamos a existência de raízes unitárias sob a presença de mudanças na média sazonal.

Para comparar o comportamento dos testes quando quebras são especificadas em séries que não apresentam quebras, Hassler e Rodrigues (2004) conduziram simulações de Monte Carlo considerando  $\tau = 0$ , em (6.13). A Tabela 6.3 apresenta os resultados obtidos.

Tabela 6.3 – Tamanhos dos testes considerando  $\tau = 0$

5% EXPERIMENTAL SIZE AND POWER, $\tau = 0$ , $\lambda = 0.5^*$						
	T = 100			T = 200		
	$t_{n_1}$	$t_{n_2}$	$F_{n_3, n_4}$	$t_{n_1}$	$t_{n_2}$	$F_{n_3, n_4}$
$\alpha = 1$						
HEGY	4.32	4.54	5.06	4.46	3.96	5.16
LM-HEGY	4.42	4.70	5.54	4.66	4.04	5.46
HEGY-AO	5.82	6.38	7.18	5.32	5.14	6.06
LM-HEGY-AO	4.68	4.84	5.88	4.66	4.36	5.38
$\alpha = 0.9$						
HEGY	5.20	6.84	10.06	7.44	11.30	19.54
LM-HEGY	5.44	7.56	11.14	8.96	12.02	20.04
HEGY-AO	4.52	7.88	10.76	4.78	9.34	13.30
LM-HEGY-AO	5.52	7.26	10.64	9.22	11.34	19.22
$\alpha = 0.8$						
HEGY	7.98	12.12	21.14	20.24	34.02	63.38
LM-HEGY	9.82	13.28	22.02	25.68	34.12	63.34
HEGY-AO	6.54	11.18	18.54	10.88	21.50	39.72
LM-HEGY-AO	10.10	11.62	19.72	23.72	29.48	53.58
$\alpha = 0.5$						
HEGY	36.50	60.74	89.60	95.42	99.70	100
LM-HEGY	45.80	61.02	89.24	94.28	99.58	100
HEGY-AO	27.60	46.66	73.76	82.46	95.52	99.96
LM-HEGY-AO	36.96	41.86	66.88	90.08	90.10	98.86

Tabela extraída de Hassler e Rodrigues (2004)

Podemos observar que o teste LM – HEGY é mais poderoso do que o teste HEGY original na frequência zero e o teste LM – HEGY – AO, por sua vez, é mais poderoso do que o HEGY – AO na frequência zero e nas demais frequências, quando  $T = 200$ . Embora o teste LM – HEGY – AO considere a correção da quebra em um processo em que não há quebras estruturais ( $\tau = 0$ ) e, portanto, é esperado que ele se apresente com poder reduzido, o que se observa na Tabela 6.3 é que há situações em que ele se apresenta mais poderoso do que o HEGY e o LM – HEGY.

Para avaliar o comportamento dos testes quando quebras estruturais são erroneamente especificadas, isto é, quando consideramos a fração  $l = 0,5$  de quebra em séries cuja quebra real se apresenta na fração de tempo igual a  $\lambda = 0,4$ , simulações de Monte Carlo também foram realizadas e os resultados se encontram na Tabela 6.4.

Tabela 6.4 – Tamanhos dos testes considerando  $l = 0,5$  quando  $\lambda = 0,4$

5% EXPERIMENTAL SIZE AND POWER, $\tau = 0, \lambda = 0.5^*$						
	T = 100			T = 200		
	$t_{n_1}$	$t_{n_2}$	$F_{n_2, n_4}$	$t_{n_1}$	$t_{n_2}$	$F_{n_2, n_4}$
$\alpha = 1$						
HEGY	4.32	4.54	5.06	4.46	3.96	5.16
LM-HEGY	4.42	4.70	5.54	4.66	4.04	5.46
HEGY-AO	5.82	6.38	7.18	5.32	5.14	6.06
LM-HEGY-AO	4.68	4.84	5.88	4.66	4.36	5.38
$\alpha = 0.9$						
HEGY	5.20	6.84	10.06	7.44	11.30	19.54
LM-HEGY	5.44	7.56	11.14	8.96	12.02	20.04
HEGY-AO	4.52	7.88	10.76	4.78	9.34	13.30
LM-HEGY-AO	5.52	7.26	10.64	9.22	11.34	19.22
$\alpha = 0.8$						
HEGY	7.98	12.12	21.14	20.24	34.02	63.38
LM-HEGY	9.82	13.28	22.02	25.68	34.12	63.34
HEGY-AO	6.54	11.18	18.54	10.88	21.50	39.72
LM-HEGY-AO	10.10	11.62	19.72	23.72	29.48	53.58
$\alpha = 0.5$						
HEGY	36.50	60.74	89.60	95.42	99.70	100
LM-HEGY	45.80	61.02	89.24	94.28	99.58	100
HEGY-AO	27.60	46.66	73.76	82.46	95.52	99.96
LM-HEGY-AO	36.96	41.86	66.88	90.08	90.10	98.86

Tabela extraída de Hassler e Rodrigues (2004)

O teste HEGY-AO se mostra muito sensível quando a quebra estrutural é considerada erroneamente. Em todos os casos, com exceção de (ii), são observadas distorções nos tamanhos desse teste. Entretanto, o teste LM – HEGY – AO se apresenta bastante robusto, comprovando o resultado assintótico apresentado no Teorema 6.3. Portanto, o teste LM – HEGY – AO pode ser utilizado mesmo quando o momento em que ocorre a quebra estrutural não seja conhecido a priori.

Para mais detalhes, ver Hassler e Rodrigues (2004).

## Capítulo 7

### Aplicações e Considerações Finais

A fim de ilustrar os testes apresentados, vamos analisar a série temporal de Índices da Produção Industrial Extrativa obtida do Instituto Brasileiro de Geografia e Estatística e observada no período de Janeiro/1991 a Dez/2007, totalizando 204 unidades amostrais, 17 anos.

#### 7.1 Análise Descritiva

A Figura 7.1 apresenta o gráfico da série Índices da Produção Industrial Extrativa, onde, claramente, podemos notar a presença de uma tendência crescente (raiz unitária na frequência zero).

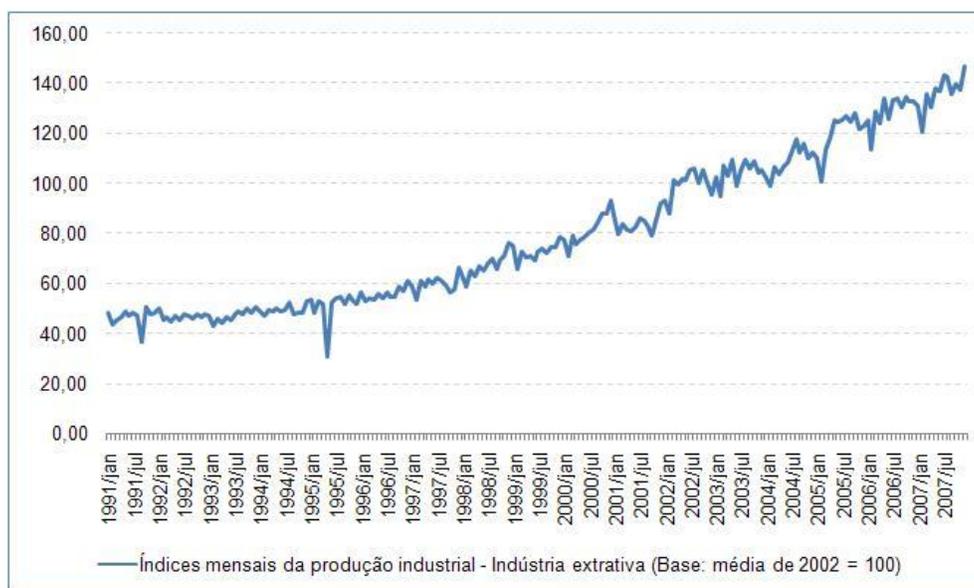


Figura 7.1 – Série Índices Mensais da Produção Industrial Extrativa

Notamos, também, que a variabilidade da série parece não ser constante ao longo do tempo, sugerindo ser necessária uma transformação logarítmica nos dados para estabilizar a variância da série.

A Figura 7.2 apresenta o gráfico da série transformada.

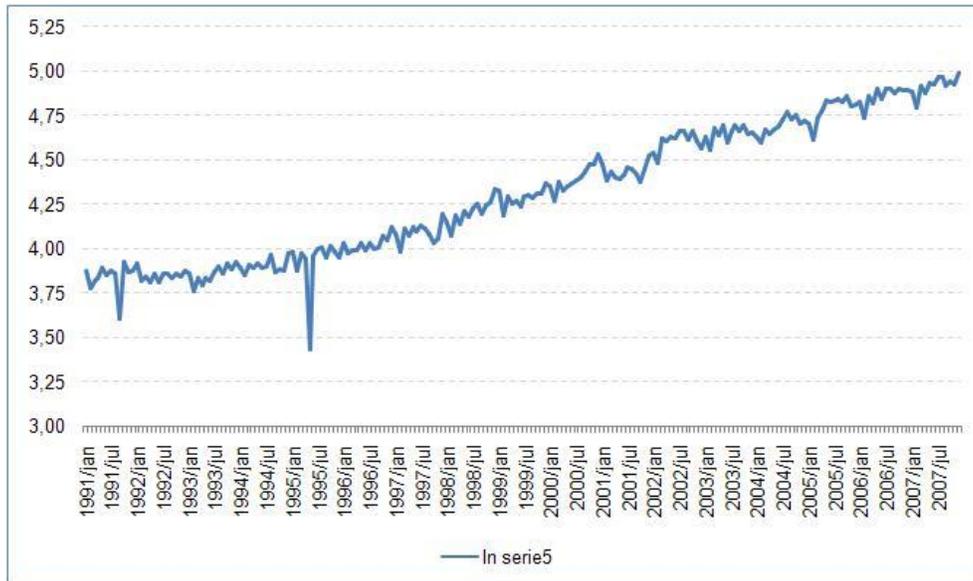


Figura 7.2 – Logaritmo da Série Índices Mensais da Produção Industrial Extrativa

Na Figura 7.2 podemos notar que a transformação parece ter estabilizado a variância da série. Observamos, também, a presença de duas observações que destoam das demais – setembro de 1991 e maio de 1995 – sugerindo a presença de observações atípicas nesses períodos.

A Figura 7.3 apresenta a Função de Autocorrelação (FAC) e a Função de Autocorrelação Parcial (FACP) da série transformada.

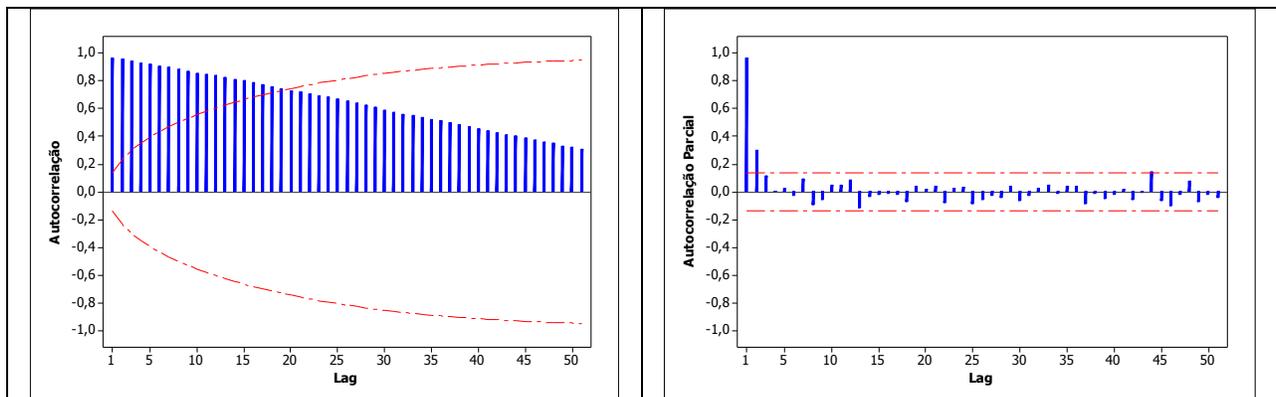


Figura 7.3 – FAC e FACP do Logaritmo da Série Índices Mensais da Produção Industrial Extrativa

A FAC decai vagarosamente para zero, indicando a presença de tendência na série.

Para avaliar se existem componentes sazonais, vamos verificar o comportamento da FAC e da FACP da série transformada e diferenciada (tomaremos uma diferença simples para tentar eliminar a tendência presente na série), apresentadas na Figura 7.4.

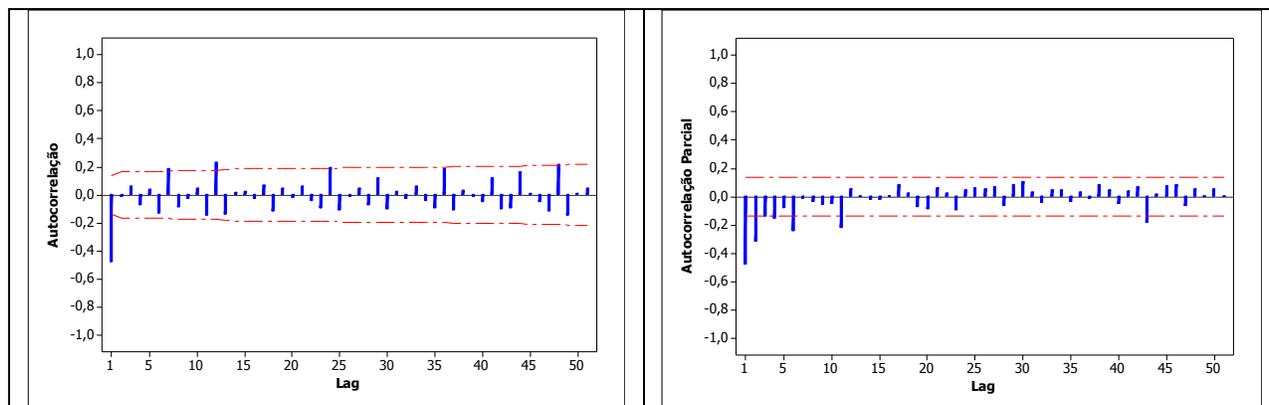


Figura 7.4 – FAC e FACP da Série Diferenciada LN dos índices Mensais da Produção Industrial Extrativa

A FAC da série diferenciada possui as autocorrelações dos lags 1, 12, 24, 36 e 48 significativas, indicando uma possível componente sazonal de ordem 12.

Na prática, é esperado que a produção industrial extrativa exiba um comportamento sazonal, pois a temperatura e o clima das estações do ano exercem influência sobre essa série temporal.

## 7.2 Aplicações dos Testes HEGY

### 7.2.1 O Teste HEGY Mensal

De posse dos resultados descritivos, vamos realizar os testes HEGY. Inicialmente, realizaremos o teste HEGY mensal sem considerar que observações atípicas ou quebras estruturais estejam presentes na série.

Construímos as variáveis regressoras expressas em (3.19) e, a partir delas, obtivemos a matriz de correlação da Figura 9. As correlações são inferiores a 0,09, o que era esperado, uma vez que as variáveis regressoras são assintoticamente ortogonais, permitindo testar a existência de uma particular raiz unitária sem que o teste seja comprometido se outras raízes unitárias estiverem presentes na frequência zero ou em frequências sazonais.

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12
Y1	1,00											
Y2	-0,03	1,00										
Y3	-0,05	-0,01	1,00									
Y4	-0,05	-0,01	0,00	1,00								
Y5	-0,04	0,08	-0,07	-0,04	1,00							
Y6	0,03	-0,03	-0,08	0,08	-0,04	1,00						
Y7	-0,04	0,05	0,09	0,04	0,07	-0,04	1,00					
Y8	-0,09	0,07	0,05	0,08	0,08	-0,04	0,08	1,00				
Y9	-0,05	0,04	0,00	0,00	0,06	0,03	0,04	0,06	1,00			
Y10	0,01	0,04	-0,02	0,02	-0,06	0,02	-0,03	-0,02	-0,02	1,00		
Y11	-0,06	0,02	0,05	0,00	0,01	-0,05	-0,04	0,09	0,03	-0,02	1,00	
Y12	-0,09	0,07	0,08	0,09	0,08	-0,04	0,06	0,09	0,08	-0,02	0,07	1,00

Figura 7.5 – Matriz de Correlação de Pearson

Estimamos, por mínimos quadrados, a regressão HEGY Mensal dada por:

$$y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + m_0 t + m_1 + \sum_{k=2}^{12} m_k S_{kt} + Y_{13(t-1)} + \varepsilon_t$$

em que

$y_{k,t}$ ,  $k = 1, 2, \dots, 13$  são as variáveis expressas em (3.19),

$m_0 t$  representa a tendência determinística,

$m_1$  representa o intercepto do modelo,

$S_{kt}$ ,  $k = 2, \dots, 12$  são as *dummies* sazonais que assumem valor 1, quando  $t$  corresponde ao mês  $k$  e 0, caso contrário e

$\varepsilon_t \sim \text{RB}(0, \sigma_\varepsilon^2)$ .

Optamos por incluir as *dummies* sazonais no modelo porque a perda de poder que resulta da inclusão dessas variáveis, quando elas não são significativas é mínima comparada a perda proveniente da omissão delas, quando são significativas. A inclusão da variável regressora  $Y_{13(t-1)}$  fez-se necessária para tornar ruído branco os resíduos do modelo.

A Figura 7.6 apresenta a FAC dos resíduos do Modelo HEGY Mensal, indicando que os resíduos são ruídos branco.

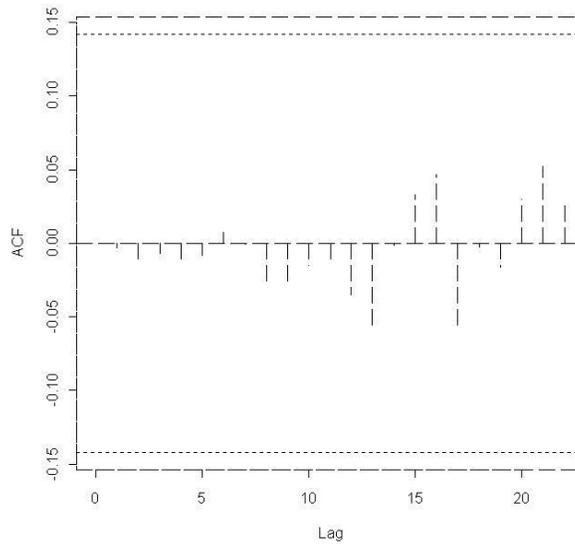


Figura 7.6 – FAC dos resíduos do Modelo HEGY Mensal

Nas Tabelas 7.1 e 7.2, temos os resultados dos testes para raízes unitárias sazonais em dados mensais.

Tabela 7.1 – Estatísticas t dos testes para raízes unitárias sazonais (HEGY Mensal)

0	$\pi$	$\pi/2$		$2\pi/3$		$\pi/3$		$5\pi/6$		$\pi/6$	
$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	$\pi_7$	$\pi_8$	$\pi_9$	$\pi_{10}$	$\pi_{11}$	$\pi_{12}$
0,141 *	3,33 *	-6,00	-0,47	-1,36	-5,12	-3,56	-1,92	3,19	-3,97	-0,08	-4,69

Tabela 7.2 – Estatísticas F dos testes para raízes unitárias sazonais (HEGY Mensal)

$\pi/2$	$2\pi/3$	$\pi/3$	$5\pi/6$	$\pi/6$
F3,4	F5,6	F7,8	F9,10	F11,12
18,11 *	14,03 *	8,18 **	12,97 *	11,01 *

(\*) Raiz unitária significativa ao nível de significância de 1%.

(\*\*) Raiz unitária significativa ao nível de significância de 2,5%.

A Tabela 7.3 sumariza os resultados do teste HEGY Mensal, utilizando os correspondentes valores críticos obtidos das Tabelas 3.2 e 3.3.

Tabela 7.3 – Resultados do teste HEGY Mensal

Raiz Presente	Frequência	Ciclos por ano
1	0	0 *
-1	$\pi$	6
$\pm i$	$\frac{\pi}{2}$	3 e 9
$-\frac{1}{2}(1 \pm \sqrt{3}i)$	$\frac{2\pi}{3}$	4 e 8
$\frac{1}{2}(1 \pm \sqrt{3}i)$	$\frac{\pi}{3}$	2 e 10
$-\frac{1}{2}(\sqrt{3} \pm i)$	$\frac{5\pi}{6}$	5 e 7
$\frac{1}{2}(\sqrt{3}i \pm i)$	$\frac{\pi}{6}$	1 e 11

\* Tendência estocástica

Por intermédio do teste HEGY Mensal concluímos que a série Índices da Produção Industrial Extrativa apresenta raiz unitária na frequência um (tendência estocástica) e em todas as frequências sazonais.

### 7.2.2 O Teste HEGY Mensal / Trimestral

Como vimos no Capítulo 4, o teste HEGY resultante da combinação dos testes HEGY Mensal e Trimestral torna-se mais poderoso do que o teste mensal. Para ilustrar, realizaremos o teste Mensal / Trimestral e comparemos os seus resultados com os obtidos do teste HEGY Mensal.

Inicialmente, vamos construir as séries trimestrais  $X_t^s$  ( $s = 1, 2, 3$ ) por amostragem sistemática. Por exemplo, a série  $X_t^1$  iniciará pela observação de janeiro de 1991, conterà informações de janeiro, abril, julho e outubro e finalizará pela observação de outubro de 2007.

A Figura 7.7 apresenta as três séries amostradas em um mesmo gráfico e as Figuras 7.8 a 7.10 apresentam as séries separadamente.



Figura 7.7 – Séries Amostradas Sistemáticamente

Na Figura 7.7 podemos notar que a tendência crescente se faz presente nas três séries amostradas e que cada uma delas apresenta parte da componente sazonal mensal.

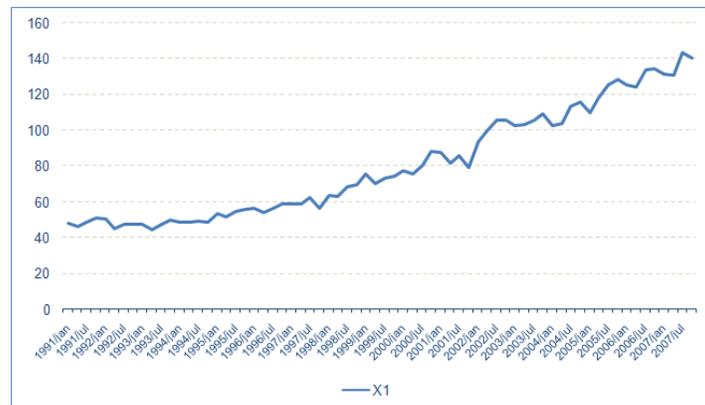


Figura 7.8 – Série  $X_t^1$

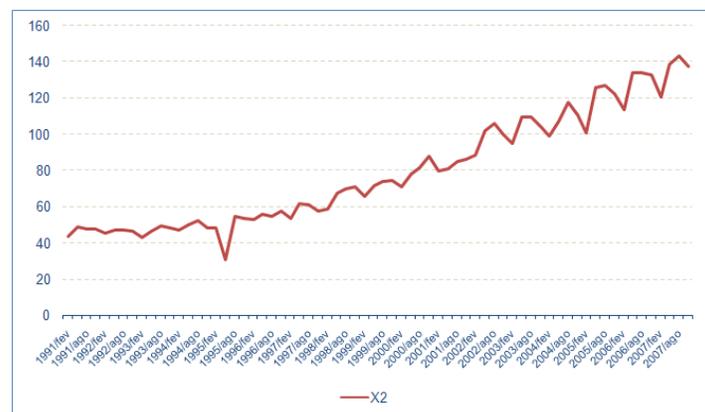


Figura 7.9 – Série  $X_t^2$

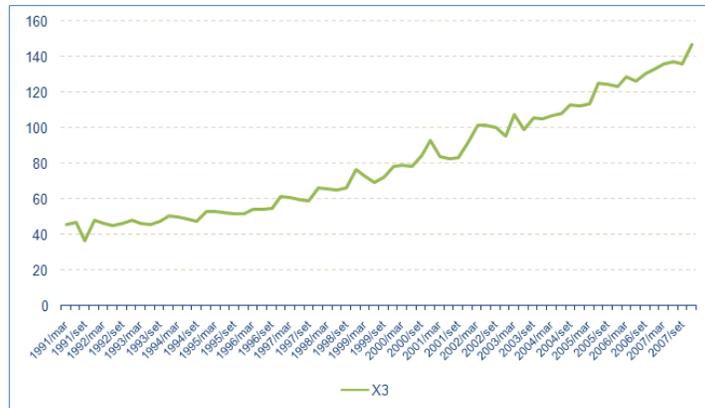


Figura 7.10 – Série  $X_T^3$

➤ **Construção das Estatísticas do Teste HEGY Mensal / Trimestral**

Como vimos no Capítulo 4, as estatísticas dos testes trimestrais são definidas por:

- $T_0^S$  = a estatística “t” correspondente ao teste da raiz unitária na frequência zero da série  $X_T^S$ ;
- $F_1^S$  = a estatística “F” correspondente ao teste do par de raízes unitárias complexas conjugadas nas frequências  $\pm \frac{\pi}{2}$  da série  $X_T^S$ ;
- $T_2^S$  = a estatística “t” correspondente ao teste da raiz unitária na frequência  $\pi$  da série  $X_T^S$ .

A Tabela 7.4 apresenta os resultados do teste HEGY trimestral para cada uma das séries amostradas.

Tabela 7.4 – Estatísticas do teste HEGY Trimestral

Estatística	Série Amostrada (s)		
	1	2	3
$T_0^S$	-3,52	-3,56	-2,83
$F_1^S$	8,94	7,23	13,34
$T_2^S$	-3,53	-3,26	-4,07

A partir dos resultados da Tabela 7.4, calculamos as estatísticas dadas por:

$$\begin{aligned}
T_0^I &= \max_s(|T_0^s|), & T_0^{II} &= \text{med}_s(|T_0^s|), & T_0^{III} &= \min_s(|T_0^s|), \\
F_1^I &= \max_s(|F_1^s|), & F_1^{II} &= \text{med}_s(|F_1^s|), & F_1^{III} &= \min_s(|F_1^s|), \\
T_2^I &= \max_s(|T_2^s|), & T_2^{II} &= \text{med}_s(|T_2^s|), & T_2^{III} &= \min_s(|T_2^s|),
\end{aligned}$$

Fornecendo os resultados:

$$\begin{aligned}
T_0^I &= -3,56, & T_0^{II} &= -3,52, & T_0^{III} &= -2,83, \\
F_1^I &= 13,34, & F_1^{II} &= 8,94, & F_1^{III} &= 7,23, \\
T_2^I &= -4,07, & T_2^{II} &= -3,53, & T_2^{III} &= -3,26.
\end{aligned}$$

### ➤ O Teste HEGY Mensal / Trimestral

Conforme exposto na Seção 4.3.3 do Capítulo 4, o teste HEGY Mensal / Trimestral não rejeita a existência da raiz unitária  $\frac{\pi k}{6}$  ao nível  $\alpha_{QM}$ , quando para algum  $\alpha_Q$ , tal que  $0 \leq \alpha_Q < \alpha_{QM}$ , as seguintes relações forem satisfeitas:

$$\begin{aligned}
T_0^i &> VC_0^i(\alpha_Q) \quad \text{e} \quad t_0 > VC_0(\alpha_{QM} - \alpha_Q), \\
F_1^i &< VC_1^i(\alpha_Q) \quad \text{e} \quad f_1 < VC_1(\alpha_{QM} - \alpha_Q), \\
T_2^i &> VC_2^i(\alpha_Q) \quad \text{e} \quad f_2 < VC_2(\alpha_{QM} - \alpha_Q), \\
F_1^i &< VC_1^i(\alpha_Q) \quad \text{e} \quad f_3 < VC_3(\alpha_{QM} - \alpha_Q), \\
T_0^i &> VC_0^i(\alpha_Q) \quad \text{e} \quad f_4 < VC_4(\alpha_{QM} - \alpha_Q), \\
F_1^i &< VC_1^i(\alpha_Q) \quad \text{e} \quad f_5 < VC_5(\alpha_{QM} - \alpha_Q), \\
T_2^i &> VC_2^i(\alpha_Q) \quad \text{e} \quad t_6 > VC_6(\alpha_{QM} - \alpha_Q),
\end{aligned}$$

em que,  $i = I, II$  e  $III$  e  $VC$  são os correspondentes valores críticos.

Para todos os testes, vamos utilizar  $\alpha_Q = 2,5\%$  e  $\alpha_M = 2,5\%$ , de tal forma que os níveis globais dos testes são dados por  $\alpha_{QM} = 5\%$ .

Note que  $f_1 = F_{3,4}$ ,  $f_2 = F_{5,6}$ ,  $f_3 = F_{7,8}$ ,  $f_4 = F_{9,10}$  e  $f_5 = F_{11,12}$ .

➤ Os Testes QM – HEGY<sup>i</sup> (i = I, II e III)

Tabela 7.5 – Resultados dos testes QM – HEGY<sup>i</sup> (i = I, II e III)

Condição	QM – HEGY <sup>I</sup>	QM – HEGY <sup>II</sup>	QM – HEGY <sup>III</sup>
$T_0^i > VC_0^i(\alpha_Q)$ e $t_0 > VC_0(\alpha_{QM} - \alpha_Q)$	V	V	V
$F_1^i < VC_1^i(\alpha_Q)$ e $f_1 < VC_1(\alpha_{QM} - \alpha_Q)$	F	F	F
$T_2^i > VC_2^i(\alpha_Q)$ e $f_2 < VC_2(\alpha_{QM} - \alpha_Q)$	F	F	V
$F_1^i < VC_1^i(\alpha_Q)$ e $f_3 < VC_3(\alpha_{QM} - \alpha_Q)$	F	F	F
$T_0^i > VC_0^i(\alpha_Q)$ e $f_4 < VC_4(\alpha_{QM} - \alpha_Q)$	V	V	V
$F_1^i < VC_1^i(\alpha_Q)$ e $f_5 < VC_5(\alpha_{QM} - \alpha_Q)$	F	F	F
$T_2^i > VC_2^i(\alpha_Q)$ e $t_6 > VC_6(\alpha_{QM} - \alpha_Q)$	F	F	V

Os resultados na Tabela 7.5 revelam a presença das raízes unitárias nas frequências mensais zero e  $\frac{2\pi}{3}$ . Portanto, o teste QM – HEGY<sup>I</sup> indica que a série Índices Mensais da Produção Industrial Extrativa apresenta tendência estocástica além de 4 e 8 ciclos por ano.

Os resultados obtidos utilizando QM – HEGY<sup>II</sup> são os mesmos obtidos utilizando QM – HEGY<sup>I</sup>. Portanto, novamente a série Índices Mensais da Produção Industrial Extrativa apresenta tendência estocástica, 4 e 8 ciclos por ano.

O teste QM – HEGY<sup>III</sup>, por sua vez, além de identificar a presença das raízes unitárias nas frequências 0 e  $\frac{2\pi}{3}$  identificadas pelos testes QM – HEGY<sup>I</sup> e QM – HEGY<sup>II</sup>, aponta também a presença das raízes unitárias nas frequências  $\frac{\pi}{3}$  e  $\pi$ . Portanto, o teste QM – HEGY<sup>III</sup> indica que a série Índices Mensais da Produção Industrial Extrativa apresenta tendência estocástica, 2, 4, 6, 8 e 10 ciclos por ano.

Vale comentar que as estatísticas  $T_k^I (F_1^I)$ ,  $T_k^{II} (F_1^{II})$  e  $T_k^{III} (F_1^{III})$  se apresentaram muito diferentes, indicando a presença de valores atípicos ou quebras estruturais na série que está sendo avaliada. Como vimos anteriormente, há duas observações que destoam das demais e, na construção das séries amostradas  $X_T^s$ , elas se dividiram entre a série  $X_T^2$  e  $X_T^3$ . Isso certamente impactou em intensidades diferentes nas médias sazonais trimestrais e nos ciclos dos três processos amostrados  $X_T^s$ , afetando o desempenho dos testes HEGY.

### 7.2.3 O Teste HEGY com Correção para Outliers

#### ➤ Procedimento Perron e Rodriguez

De posse dos resultados anteriores, é conveniente realizarmos o teste HEGY com correção para *outliers*. Antes, porém, vamos realizar o procedimento de Perron-Rodriguez (2003) exposto no Capítulo 5 para detectar *outliers*.

Como vimos na Figura 7.2, a série Índices Mensais de Produção Extrativa parece conter observações atípicas nos períodos de setembro de 1991 e maio de 1995. Entretanto, quando realizamos uma diferença de ordem 12, a observação referente a maio/1994 também parece ser atípica, como podemos ver na Figura 7.11.

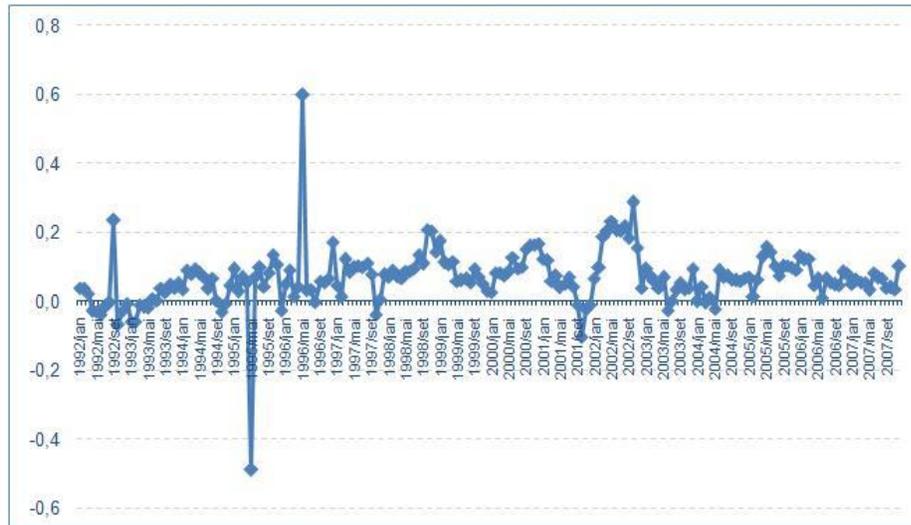


Figura 7.11 – LN da Série Diferenciada Índices Mensais da Produção Industrial Extrativa

Para a realização do procedimento de Perron e Rodriguez, utilizamos o software R e obtivemos os seguintes resultados:

**(i) Primeira iteração:** A primeira iteração resultou em  $\tau^{PR} = 9,50$  que corresponde ao período de maio de 1995. Note que trata-se justamente de uma das observações que se destoavam das demais na Figura 7.2. Considerando um nível de significância de 5%, comparada com o valor crítico obtido de Tabela 5.3, a estatística  $\tau^{PR}$  é significativa e, portanto, maio de 1995 é considerado um *outlier*.

Para que não percamos a estrutura de auto-correlação dos dados, conforme a sugestão de Haldrup, Montanés e Sanso (2004), o próximo passo consiste em substituir a observação referente a maio de 1995 pela observação do ano anterior, ou seja, maio de 1994.

**(i) Segunda iteração:** A segunda iteração resultou em  $\tau^{PR} = 4,65$  que corresponde ao período de outubro de 2001. Considerando um nível de significância de 5%, comparada com o valor crítico obtido de Tabela 5.3, a estatística  $\tau^{PR}$  é significativa e, portanto, outubro de 2001 também é considerado um *outlier*.

De forma análoga ao passo anterior, substituímos a observação de outubro de 2001 pela observação de outubro de 2000.

**(ii) Terceira iteração:** A terceira iteração resultou em  $\tau^{PR} = 3,10$  que corresponde ao período de outubro de 1997. Considerando um nível de significância de 5%, comparada com o valor crítico obtido de Tabela 5.3, essa estatística não é significativa e, portanto, outubro de 1997 não é considerado um *outlier*. Nesse instante, o procedimento de Perron e Rodriguez é finalizado.

Assim, temos que o procedimento de Perron e Rodriguez resultou em atípicas as observações de maio de 1995 e outubro de 2001.

### ➤ O Teste HEGY com Correção para *Outliers*

Conforme exposto na expressão (5.5), a regressão HEGY com correção para *outliers* é dada por:

$$\Delta_s z_t = \mu_s + \sum_{j=1}^{s-1} \mu_j D_{jt} + \beta t + \sum_{j=1}^s \pi_j^s z_{j,t-1}^s + \sum_{j=1}^k \alpha_j \Delta_s z_{t-j} + \sum_{i=0}^{s+k} \sum_{j=1}^2 \gamma_{ij} D(T_0^j)_{t-i} + u_t$$

em que  $D_{jt}$  são as variáveis *dummies* sazonais e  $D(T_0^j)_t$ ,  $j = 1, 2$  são as variáveis *dummies* capturando os dois *outliers* que foram encontrados.

A Figura 7.12 apresenta a FAC dos resíduos do Modelo HEGY Mensal com correção para *outliers*, indicando que os resíduos são RB  $(0, \sigma^2)$ .

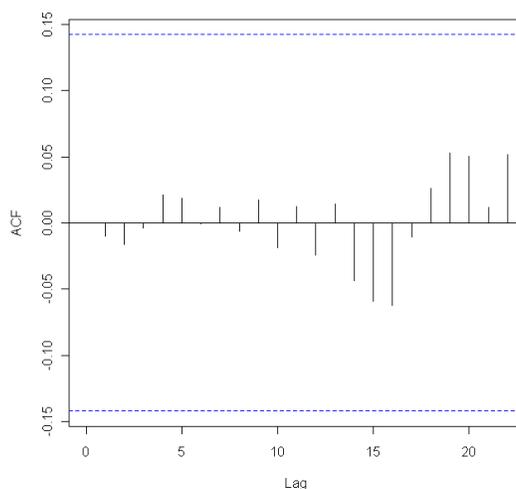


Figura 7.12 – FAC dos resíduos do Modelo HEGY Mensal com Correção para Outliers

Nas Tabelas 7.6 e 7.7, temos as estatísticas dos testes com correções para *outliers* para raízes unitárias sazonais em dados mensais.

Tabela 7.6 – Estatísticas “t” dos testes para raízes unitárias sazonais

0		$\pi$		$\pi/2$		$2\pi/3$		$\pi/3$		$5\pi/6$		$\pi/6$	
$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	$\pi_7$	$\pi_8$	$\pi_9$	$\pi_{10}$	$\pi_{11}$	$\pi_{12}$		
-1,26 *	3,13 *	-4,12	0,24	-1,61	-2,34	-2,87	-1,17	2,75	-3,89	0,01	-3,59		

Tabela 7.7 – Estatísticas F dos testes para raízes unitárias sazonais

$\pi/2$	$2\pi/3$	$\pi/3$	$5\pi/6$	$\pi/6$
F3,4	F5,6	F7,8	F9,10	F11,12
9,32 *	5,12	4,94	12,86 *	6,91 **

(\*) Raiz unitária significativa ao nível de significância de 1%.

(\*\*) Raiz unitária significativa ao nível de significância de 5%.

A Tabela 7.8 sumariza os resultados do teste HEGY Mensal com correções para *outliers*.

Tabela 7.8 – Resultados do teste HEGY Mensal com correções para *outliers*

Raiz Presente	Frequência	Ciclos por ano
1	0	0 *
-1	$\pi$	6
$\pm i$	$\frac{\pi}{2}$	3 e 9
$-\frac{1}{2}(\sqrt{3} \pm i)$	$\frac{5\pi}{6}$	5 e 7
$\frac{1}{2}(\sqrt{3}i \pm i)$	$\frac{\pi}{6}$	1 e 11

\* Tendência estocástica

Portanto, por intermédio do teste HEGY Mensal com correção para *outliers* concluímos que a série Índices da Produção Industrial Extrativa apresenta raiz unitária na frequência um (tendência estocástica) e 1, 3, 5, 6, 7, 9 e 11 ciclos por ano.

É importante comentar que quando incluímos a componente de correção para *outliers* no teste HEGY, as frequências sazonais  $\frac{2\pi}{3}$  e  $\frac{\pi}{3}$  deixaram de ser significativas, mostrando a relevância de se considerar métodos apropriados para detectar e remover os efeitos causados por *outliers*.

#### 7.2.4 O Teste HEGY com Quebras Estruturais

Nessa seção, vamos ilustrar a aplicação do teste LM-HEGY-AO utilizando a série trimestral amostrada  $X_T^2$ .

Inicialmente, o modelo expresso em (6.7) é ajustado:

$$\Delta_4 y_t = \tilde{\gamma} + \sum_{s=1}^4 \tilde{\tau}_s \Delta_4 D(T_B)_{s,t} + \tilde{x}_t \quad t = 5, 6, \dots, T$$

em que  $y_t = X_T^2$ ,  $\Delta_4 = 1 - B^4$ , B é o operador lag,  $\tilde{\gamma}$  é a estimativa do parâmetro  $\gamma$  de tendência linear,  $D(T_B)_{s,t} = D_{s,t}$ , se  $t > T_B$ ,  $D_{s,t}$  são *dummies* sazonais tais que  $D_{s,t} = 1$ , quando t se refere ao s-ésimo trimestre e  $D_{s,t} = 0$ , caso contrário e  $\tilde{x}_t$  é o resíduo do modelo.

O próximo passo consiste em se obter as variáveis regressoras da Regressão LM-HEGY-AO através do resíduo  $\tilde{x}_t$ . Dessa forma, teremos:

$$\tilde{x}_{1,t} = (1 + B + B^2 + B^3)\tilde{y}_t,$$

$$\tilde{x}_{2,t} = -(1 - B + B^2 - B^3)\tilde{y}_t,$$

$$\tilde{x}_{3,t} = -(1 - B^2)\tilde{y}_t.$$

Finalmente, a regressão LM-HEGY-AO é ajustada conforme o modelo expresso em (6.3).

$$\Delta_4 y_t = \sum_{s=1}^4 \hat{d}_s D_{s,t} + \sum_{s=1}^4 \hat{\tau}_s D(T_B)_{s,t} + \hat{c}t + \hat{\pi}_1 x_{1,t-1} + \hat{\pi}_2 x_{2,t-1} + \hat{\pi}_3 x_{3,t-2} + \hat{\pi}_4 x_{3,t-1} + \hat{\epsilon}_t.$$

Para efeito de ilustração, consideramos que uma quebra estrutural ocorreu na observação maio de 1995 ( $T_B$ ). A Tabela 7.9 sumariza os resultados.

Tabela 7.9 – Resultados do teste LM-HEGY-AO Trimestral

0	$\pi$	$\pi/2$		$\pi/2$
$\pi 1$	$\pi 2$	$\pi 3$	$\pi 4$	F3,4
1,23 *	2,64 *	-2,23	-4,13	11,12 *

(\*) Raiz unitária significativa ao nível de significância de 1%.

Portanto, por intermédio do teste LM-HEGY-AO concluímos que a série trimestral amostrada  $X_T^2$  possui tendência estocástica e ciclo semestral e anual.

### 7.3 Considerações Finais

O tema de sazonalidade é amplamente estudado por muitos pesquisadores e muitas parametrizações e abordagens diferentes tem sido utilizadas para explorar a importância e os efeitos da sazonalidade em contextos univariados e multivariados.

Quando uma ou mais séries contem uma raiz unitária, problemas sérios de inferência estatística podem aparecer. A variância de uma série com uma raiz unitária, por exemplo, é infinita, a distribuição da primeira autocorrelação estimada se modifica quando a verdadeira autocorrelação é igual a um e duas séries integradas independentes podem apresentar correlações espúrias. Estes problemas não são solucionados quando aplicamos diferenciações para tornar ambas as séries estacionárias e são mais freqüentes quando raízes unitárias em freqüências sazonais estão presentes. A utilização de variáveis *dummies* sazonais não é apropriada, se a sazonalidade presente na série não é gerada por um processo integrado. Modelos de cointegração sazonal, como os procedimentos HEGY expostos no presente trabalho, necessitam que as séries que estão sendo avaliadas sejam sazonalmente integradas.

Vale citar que a aplicação de filtros da forma  $(1-B^d)$  em séries com componentes sazonais, como proposto por Box e Jenkins (1970), necessita que a série seja integrada na freqüência zero e em todas as freqüências sazonais. Como as raízes unitárias, frequentemente, não estão presentes em todas as freqüências sazonais, então é muito importante avaliarmos, à priori, em quais freqüências as raízes unitárias estão realmente presentes.

Para essa finalidade, o presente trabalho expôs o procedimento HEGY e suas diversas derivações. Esse procedimento permite testar a presença de raízes unitárias isoladamente em todas as freqüências sazonais e na freqüência zero, isto é, permite testar a existência de uma particular raiz unitária sem que o teste seja comprometido se outras raízes unitárias estiverem presentes na freqüência zero ou em freqüências sazonais. Essa particularidade é devida ao Resultado 3.1, a partir do qual um polinômio autorregressivo pode ser reformulado por representações de polinômios racionais, conforme exposto no Capítulo 3.

Algumas aplicações foram feitas em dados obtidos do Instituto Brasileiro de Geografia e Estatísticas (IBGE). Em particular, apresentamos as análises realizadas na série de Índices Mensais de Produção Extrativa. Primeiramente, realizamos o teste HEGY sem considerar que dados atípicos ou quebras estruturais estavam presentes nessa série. Esse teste resultou na confirmação da presença das raízes correspondentes as freqüências  $0, \pi, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{\pi}{3}$  e  $\frac{\pi}{6}$ , concluindo, portanto, que a série Índices Mensais da Produção Industrial Extrativa apresenta tendência estocástica e 1, 2, 3, 4, 6, 8, 9, 10 e 11 ciclos por ano.

Em seguida, realizamos os procedimentos HEGY resultantes da combinação dos testes HEGY Mensal e Trimestral. Para tal, primeiramente, construímos as séries trimestrais  $X_T^s$  ( $s=1, 2, 3$ ) por amostragem sistemática. Notamos que a tendência existente na série completa mensal se faz presente nas três séries amostradas, porém, cada uma delas apresenta parte da componente sazonal mensal.

Os testes QM - HEGY<sup>I</sup> resultaram na presença das raízes unitárias nas freqüências mensais zero e  $\frac{2\pi}{3}$ , o que também se verificou através dos testes QM - HEGY<sup>II</sup>. Portanto, ambos indicaram que a série Índices Mensais da Produção Industrial Extrativa apresenta tendência estocástica e 4 e 8 ciclos por ano. O teste QM - HEGY<sup>III</sup>, por sua vez, além de identificar a

presença das raízes unitárias nas frequências 0 e  $\frac{2\pi}{3}$ , apontou também a presença das raízes unitárias nas frequências  $\frac{\pi}{3}$  e  $\pi$ . Portanto, o teste QM - HEGY<sup>III</sup> indica que a série Índices Mensais da Produção Industrial Extrativa apresenta tendência estocástica e 2, 4, 6, 8 e 10 ciclos por ano. É importante citar que as estatísticas  $T_k^I(F_1^I)$ ,  $T_k^{II}(F_1^{II})$  e  $T_k^{III}(F_1^{III})$  resultantes dos testes QM - HEGY<sup>I</sup>, QM - HEGY<sup>II</sup> e QM - HEGY<sup>III</sup> se apresentam muito diferentes, indicando a presença de valores atípicos ou quebras estruturais na série que está sendo avaliada.

O teste HEGY com correção para *outliers* resultou na presença de tendência estocástica e 1, 3, 5, 6, 7, 9 e 11 ciclos por ano, diferindo do teste HEGY sem correção para *outliers*, que apresentou periodicidades significativas em todas as frequências sazonais. A presença de outliers se não for considerada pode levar a resultados imprecisos sendo, portanto, muito importante considerar métodos apropriados para detectar e remover os efeitos causados por *outliers*.

Por fim, realizamos o teste HEGY com quebras estruturais (LM-HEGY-AO) utilizando a série trimestral  $X_T^2$  para ilustrar. Os resultados obtidos mostraram que essa série contém tendência estocástica e ciclos semestral e anual.

## Apêndice

### Códigos na linguagem R para aplicação dos testes HEGY

Todos os testes HEGY apresentados no Capítulo 7 foram executados utilizando o software R. A seguir, todos os códigos desenvolvidos para a realização dos testes serão apresentados.

Inicialmente, fizemos a leitura das séries temporais utilizadas. Para importar dados em formato txt e transformá-los em série temporal, fazemos:

```
y1 = scan (file = "diretorio/y1.txt")  
y1 = ts (y1, frequency = 12, start = c(10991, 1))
```

#### ➤ Teste HEGY Mensal

O teste HEGY mensal consistiu no ajuste da regressão HEGY dada por:

$$y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + m_0 t + m_1 + \sum_{k=2}^{12} m_k S_{kt} + Y_{13(t-1)} + \varepsilon_t$$

em que

$y_{k,t}$ ,  $k = 1, 2, \dots, 13$  são as variáveis expressas em (3.19),

$m_0 t$  representa a tendência determinística,

$m_1$  representa o intercepto do modelo,

$S_{kt}$ ,  $k = 2, \dots, 12$  são as *dummies* sazonais que assumem valor 1, quando  $t$  corresponde ao mês  $k$  e 0, caso contrário e

$\varepsilon_t \sim \text{RB}(0, \sigma^2)$ .

```
model <- lm (y13 ~ y1 + y2 + y3 + y4 + y5 + y6 + y7 + y8 + y9 + y10 + y11 + y12 + m0t + lag_1_y13 + s2 + s3 +  
s4 + s5 + s6 + s7 + s8 + s9 + s10 + s11 + s12)
```

Vale citar que lag\_1\_y13 é a série  $Y_{13(t-1)}$  incluída na Regressão HEGY para tornar os resíduos Ruído Branco.

Para avaliar se o modelo está bem ajustado aos dados, verificamos se a função de autocorrelação dos resíduos se comporta como um Ruído Branco.

```
res <- residuals (model)
func_fac <- acf (res)
func_fac$acf[1] <- NA # Atribuindo "NA" ao primeiro valor de acf para excluir o lag zero
plot(func_fac)
```

### ➤ Teste HEGY Mensal / Trimestral

A construção do teste HEGY Mensal / Trimestral se deu de forma similar a construção do teste mensal, guardadas as devidas diferenças resultantes das estatísticas de ambos os testes.

Os testes HEGY trimestrais consistiram nos ajustes de regressões HEGY dadas por:

$$y_{4t} = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + m_0 t + m_1 + \sum_{k=2}^{12} m_k S_{kt} + \varepsilon_t$$

em que

$y_{k,t}$ ,  $k = 1, 2, 3, 4$  são as variáveis expressas em (3.7),

$m_0 t$  representa a tendência determinística,

$m_1$  representa o intercepto do modelo,

$S_{kt}$ ,  $k = 2, 3, 4$  são as *dummies* sazonais que assumem valor 1, quando  $t$  corresponde ao mês  $k$  e 0, caso contrário e

$\varepsilon_t \sim \text{RB}(0, \sigma^2)$ .

```
model <- lm (y4 ~ y1 + y2 + y3 + y3_1 + m0*t + s2 + s3 + s4)
```

➤ **Procedimento para Identificação de *Outliers* (Perron e Rodriguez)**

Para encontrar os *outliers* das séries avaliadas, utilizamos o procedimento proposto por Perron e Rodriguez e desenvolvido no aplicativo R através do código a seguir:

```
orig_In = scan (file = "diretorio/orig_In.txt")
orig_In = ts (orig_In, frequency = 12, start = c(1991, 1))
dif_orig_In <- diff (orig_In, lag = 12)
teta <- rep (0, 192)
estatisticas_t <- rep(0,192)
dummy <- rep (0,192)
k <- rep (0,167)
vetor <- rep (0,167)
i <- 1
while (i < 181)
{
  for (m in 1:192)
  {
    dummy[m] <- 0
  }
  dummy[i] <- -1
  dummy[i+12] <- 1
  for (j in 1:192)
  {
    if ( dummy[j] != 1 | dummy[j] != -1)
      dummy[j] <- 0
    else
      dummy[j] = dummy[j]
  }
  model <- lm (dif_orig_In ~ dummy)
  res <- residuals(model)
  if (i <= 12)
```

```

teta[i] <- -res[i+12]

else if (i > 12 & i <= 180)

teta[i] <- 0.5*(res[i]-res[i+12])

else
teta[i] <- res[i]
r_zero <- sum(res*res) / 192

    for (c in 25:192)
    {
        for (k in 1:167)
        {
            vetor[k] <- res[c]*res[c-12]
        }
    }

r_doze <- sum (vetor) / 192

if (i <= 12)
estatisticas_t[i] <- teta[i] / sqrt(r_zero)

else if (i > 12 & i <= 180)
estatisticas_t[i] <- sqrt(2)*teta[i]/sqrt(r_zero-r_doze)

else
estatisticas_t[i] <- teta[i]/sqrt[r_zero]

i <- i+1

}

tal <- max(abs(estatisticas_t))

```

### ➤ Teste HEGY com Correção para *Outliers*

A construção do teste HEGY Mensal / Trimestral se deu de forma similar a construção do teste mensal, porém, na Regressão HEGY foram acrescentadas as dummies regressoras referentes aos *outliers* identificados através do procedimento de Perron e Rodriguez.

O teste HEGY mensal com correção para *outliers* consistiu no ajuste da regressão HEGY dada por:

$$\Delta_s z_t = \mu_s + \sum_{j=1}^{s-1} \mu_j D_{jt} + \beta t + \sum_{j=1}^s \pi_j^s z_{j,t-1}^s + \sum_{j=1}^k \alpha_j \Delta_s z_{t-j} + \sum_{i=0}^{s+k} \sum_{j=1}^q \gamma_{ij} D(T_0^j)_{t-i} + u_t$$

em que

$z_{jt}^s$ ,  $j = 1, 2, \dots, s$  são as séries  $z_t$  filtradas,

$z_t^s = (z_{1t}^s, z_{2t}^s, \dots, z_{st}^s)'$  e  $\pi^s = (\pi_1^s, \pi_2^s, \dots, \pi_s^s)'$  são os parâmetros do modelo,

$D_{jt}$  são as variáveis *dummies* sazonais e  $D(T_0^j)_t$ ,  $j = 1, \dots, q$  são as variáveis *dummies* capturando os  $q$  *outliers* que foram encontrados e excluindo os lags insignificantes de  $\Delta_s z_t$ , a fim de aumentar o poder,

$\beta t$  representa a tendência determinística,

$\mu_s$  representa o intercepto do modelo e

$u_t \sim \text{RB}(0, \sigma^2)$ .

```

model <- lm (y13 ~ y1 + y2 + y3 + y4 + y5 + y6 + y7 + y8 + y9 + y10 + y11 + y12 +  $\beta t$  + s2 + s3 + s4 + s5 + s6 +
s7 + s8 + s9 + s10 + s11 + s12 + D0,1 + D1,1 + D2,1 + D3,1 + D4,1 + D5,1 + D6,1 + D7,1 + D8,1 + D9,1 +
D10,1+D11,1 + D12,1 + D13,1 + D0,2 + D1,2 + D2,2 + D3,2 + D4,2 + D5,2 + D6,2 + D7,2 + D8,2 + D9,2 +
D10,2+D11,2 + D12,2 + D13,2)

```

### ➤ Teste HEGY com Quebras Estruturais

Inicialmente, ajustamos o modelo dado por:

$$\Delta_4 y_t = \tilde{\gamma} + \sum_{s=1}^4 \tilde{\tau}_s \Delta_4 D(T_B)_{s,t} + \tilde{x}_t \quad t = 5, 6, \dots, T$$

em que  $y_t = X_T^2$ ,  $\Delta_4 = 1 - B^4$ ,  $B$  é o operador lag,  $\tilde{\gamma}$  é a estimativa do parâmetro  $\gamma$  de tendência linear,  $D(T_B)_{s,t} = D_{s,t}$ , se  $t > T_B$ ,  $D_{s,t}$  são *dummies* sazonais tais que  $D_{s,t} = 1$ , quando  $t$  se refere ao  $s$ -ésimo trimestre e  $D_{s,t} = 0$ , caso contrário e  $\tilde{x}_t$  é o resíduo do modelo.

```

model <- lm (y ~ d1)

```

Na sequência, obtivemos as variáveis regressoras da Regressão LM-HEGY-AO através do resíduo  $\tilde{x}_t$ :

$$\tilde{x}_{1,t} = (1 + B + B^2 + B^3)\tilde{y}_t,$$

$$\tilde{x}_{2,t} = -(1 - B + B^2 - B^3)\tilde{y}_t,$$

$$\tilde{x}_{3,t} = -(1 - B^2)\tilde{y}_t.$$

`res <- residuals(model)`

Finalmente, ajustamos a regressão LM-HEGY-AO conforme o modelo a seguir:

$$\Delta_4 y_t = \sum_{s=1}^4 \hat{a}_s D_{s,t} + \sum_{s=1}^4 \hat{t}_s D(T_B)_{s,t} + \hat{c}t + \hat{\pi}_1 x_{1,t-1} + \hat{\pi}_2 x_{2,t-1} + \hat{\pi}_3 x_{3,t-2} + \hat{\pi}_4 x_{3,t-1} + \hat{\epsilon}_t$$

em que  $y_t = X_T^2$ ,  $\Delta_4 = 1 - B^4$ ,  $B$  é o operador lag,  $\hat{c}t$  é a estimativa da tendência linear,  $D(T_B)_{s,t} = D_{s,t}$ , se  $t > T_B$ ,  $D_{s,t}$  são *dummies* sazonais tais que  $D_{s,t} = 1$ , quando  $t$  se refere ao  $s$ -ésimo trimestre e  $D_{s,t} = 0$ , caso contrário e  $x_1$ ,  $x_2$  e  $x_3$  são as variáveis da Regressão LM-HEGY-AO.

`model <- lm (y ~ y1 + y2 + y3 + y4 + d1+  $\beta$ t + s2 + s3 + s4)`

## Referências Bibliográficas

Beaulieu, J. J. and Miron, J. A. (1993). Seasonal unit roots in aggregate U.S. data. *Journal of Econometrics*, **55**, 305-328.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. Englewood Cliffs: Prentice Hall.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. Third Edition. Englewood Cliffs: Prentice Hall.

Breitung, J. and Franses, P. H. (1998). On Phillips-Perron-type tests for seasonal unit roots. *Econometric Theory*, **14**, 200-221.

Chan, N. H. and C. Z. Wei. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes, *Annals of Statistics*, **16**, 367-401.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427-431.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1052-1072.

Dickey, D. A., H. P. Hasza and W. A. Fuller. (1984). Testing for unit roots in seasonal times series, *Journal of the American Statistical Association*, **79**, 355-367.

Eagle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, **55**, 251-276.

Franses, P. H. (1990). Testing for seasonal unit roots in montly data, Report n° 9032/A (Econometric Institute, Erasmus University, Rotterdam).

Franses, P. H. and Hobijn, B. (1997). Critical values for unit root tests in seasonal time series. *Journal of Applied Statistics*, **24**, 25-47.

Franses, P. H. and Vogelsang, T. J. (1998). On seasonal cycles, unit roots and mean shifts. *The Review of Economics and Statistics*, **80**, 231-240.

Fuller, W. A. (1976). *Introduction of statistical time series* (Wiley, New York, NY).

Granger, C. W. J. and Siklos, P. L. (1995). Systematic sampling, temporal aggregation, seasonal adjustment and cointegration. Theory and evidence. *Journal of Econometrics*, **66**, 357-69.

Haldrup, N., Montanes, A. and Sanso, A. (2003). Testing for additive outliers in seasonally integrated data. Working paper.

Haldrup, N., Montanés, A. and Sanso, A. (2005). Measurement errors and outliers in seasonal unit root testing. *Journal of Econometrics*, **127**, 103-128.

Hannan, E. J., Terrell, R. D. and Tuckwell, N. (1970). The seasonal adjustment of economic time series. *International Economic Review*, **11**, 24-52.

Hassler, U. and Rodrigues, P. M. M. (2004). Seasonal unit root tests under structural breaks. *Journal of Time Series Analysis*, **25**, 33-53.

Hasza, D. P. and W. Fuller. (1982). Testing for nonstationarity parameter specifications in seasonal time series models. *Annals of Statistics*, **10**, 1209 -1216.

Hylleberg, S., Engle, R. F., Granger, C. W. J. and Yoo, B. S. (1990). Seasonal integration and cointegration. *Journal of Econometrics*, **44**, 215-238.

Koopmans, L. H. (1974). *The Spectral Analysis of Time Series*. New York: Academic Press.

Miron, J. A. (1996). *The Economics of Seasonal Cycles*. Cambridge, MA: MIT Press.

Morettin, P. A. e Toloj, C. M. C. (2006). *Análise de Séries Temporais*. Segunda Edição. São Paulo: Editora E. Blucher - Associação Brasileira de Estatística.

Perron, P. and Rodriguez, G. (2003). Searching for additive outliers in nonstationary time series. *Journal of Time Series Analysis*, **24**, 193-220.

Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica*, **55**, 277-301.

Pons, G. (2005). Testing monthly seasonal unit roots with monthly and quarterly information. *Journal of Time Series Analysis*, **27**, 191-209.

Rodrigues, P. M. M. and Franses, P. H. (2003). A sequential approach to testing seasonal unit roots in high frequency data. Econometric Institute Report 2003 – 14. Erasmus University Rotterdam.

Schmidt, P. and Lee, J. (1991). A modification of the Schmidt and Phillips unit root test. *Economics Letters*, **36**, 285-289.

Schmidt, P. and Phillips, P. C. B. (1992). LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, **54**, 257-287.

da Silva Lopes, A. C. B. (2001). The robustness of tests for seasonal differencing to structural breaks. *Economic Letters*, **71**, 173-179.

Smith, J. and Otero, J. (1997). Structural breaks and seasonal integration. *Economics Letters*, **56**, 13-19.

Stock, J. H. (1988). A class of tests for integration and cointegration, Mimeo. (Kennedy School of Government, Harvard University, Cambridge, MA).

Vogelsang, T. J. (1999). Two simple procedures for testing for a unit root when there are additive outliers. *Journal of Time Series Analysis*, **20**, 137-252.