

# TESTES DE CO-INTEGRAÇÃO UTILIZANDO O SAS<sup>®</sup>: teoria e aplicação<sup>1</sup>

Mario Antonio Margarido<sup>2</sup>

Lilian Cristina Anefalos<sup>3</sup>

## 1 - INTRODUÇÃO

A questão do relacionamento espúrio entre variáveis foi levantada inicialmente em GRANGER e NEWBOLD (1974). Basicamente, esses autores constataram que, apesar de os resultados de um processo estocástico independente apresentarem resultados significativos para o teste  $t$  e de o coeficiente de determinação ( $R^2$ ) ser elevado, ainda assim, esses resultados não tinham significado em termos econômicos. De acordo com GUJARATI (1995, p.509) ao se fazer uma regressão entre duas séries de tempo “esse problema surge porque se ambas séries temporais envolvidas exibirem forte tendência (movimentos sustentados tanto com inclinação positiva, quanto negativa) o alto valor observado de  $R^2$  será devido à presença da tendência, e não ao verdadeiro relacionamento entre as duas séries. Por essa razão torna-se muito importante descobrir se o relacionamento entre variáveis econômicas é verdadeiro ou espúrio”.

Segundo as hipóteses da estatística “tradicional” tanto a média como a variância são constantes ao longo do tempo. No entanto, quando ambas deixam de ser constantes ao longo do tempo, isso implica a presença de raiz unitária<sup>4</sup> na série e, conseqüentemente, os pressupostos da estatística “tradicional” perdem validade e os

resultados da regressão podem ser espúrios. Para verificar se uma série individualmente é estacionária ou não (isto é, possui raiz unitária) é necessária a execução dos testes de raiz unitária<sup>5</sup>. Quando uma variável individualmente tem raiz unitária, pode acontecer que a sua combinação linear com outra variável seja estacionária e, nesse caso, as duas séries são ditas co-integradas, ou seja, possuem um relacionamento de equilíbrio no longo prazo.

## 2 - OBJETIVOS

O principal objetivo deste trabalho é realçar a importância dos testes de co-integração na realização de trabalhos econométricos.

Especificamente, serão delineados os testes de co-integração desenvolvidos por ENGLE e GRANGER (1987 e 1991) e o teste de PHILLIPS e OULIARIS (1990), o qual é derivado em relação ao primeiro.

Também, foi elaborada uma rotina no SAS para fazer o teste de co-integração de Engle-Granger, além de mostrar como o SAS faz o teste de co-integração de Phillips-Ouliaris passo a passo. Visando facilitar o aprendizado por parte do leitor, será utilizado um exemplo contendo dados sobre o comércio internacional de farelo de soja.

## 3 - MATERIAL E MÉTODO

### 3.1 - Material

Foram utilizados dados com periodicidade mensal relativos aos preços Cost Insurance and Freight (CIF) do farelo de soja no porto de

<sup>1</sup>Versão preliminar deste artigo foi apresentada no IX Congresso Brasileiro de Usuários SAS, realizado na cidade do Rio de Janeiro no período de 27 a 30 de maio de 2000.

<sup>2</sup>Economista, Doutor, Pesquisador Científico do Instituto de Economia Agrícola. E-mail: mamargarido@iea.sp.gov.br.

<sup>3</sup>Engenheiro Agrônomo, Mestre, Pesquisador Científico do Instituto de Economia Agrícola.

<sup>4</sup>Conforme HORTON (1998) uma série tem raiz unitária quando após a aplicação de uma diferença ela torna-se estacionária. Num sentido restrito, estacionariedade implica que a série tenha média e variância constantes ao longo do tempo.

<sup>5</sup>A questão relativa à parte teórica e à realização de testes de raiz unitária utilizando o SAS encontra-se detalhada em MARGARIDO; ANEFALOS (1999).

Rotterdam, além dos preços Free on Board (FOB) do farelo de soja no Brasil e Argentina para o período de outubro de 1990 até setembro de 1999. Essas três séries foram extraídas da OILSEEDS (1990/1999).

### 3.2 - Co-integração: aspectos teóricos

O conceito de co-integração procura identificar se duas ou mais variáveis integradas de mesma ordem possuem uma relação de equilíbrio no longo prazo. Especificamente no caso de duas variáveis, se  $x_t$  e  $y_t$  são integradas de ordem 1 (o que é representado por  $I(1)$ ), então, é verdadeiro que a sua combinação linear dada por:

$$z_t = y_t - \alpha x_t \quad (1)$$

também será  $I(1)$ . No entanto, naqueles casos em que ambas as variáveis forem  $I(1)$  e  $z_t$  for  $I(0)$  na expressão (1), isto significa que  $y_t$  e  $\alpha x_t$  devem ter componentes de longo prazo que praticamente se cancelam para produzir  $z_t$ . Nessas circunstâncias diz-se que  $x_t$  e  $y_t$  são co-integradas e  $\alpha$  é o parâmetro de co-integração, ou então, é denominado vetor de co-integração, sendo representado como  $(1, -\alpha)$ . O relacionamento entre co-integração e o conceito de equilíbrio de longo prazo do princípio de que esse equilíbrio é representado como:

$$y_t = \alpha x_t \quad \text{ou} \quad y_t - \alpha x_t = 0 \quad (2)$$

onde  $z_t$  na expressão (1) é o *modelo de erro de equilíbrio*, ou seja, mede quanto o sistema representado por  $(x_t, y_t)$  se encontra distante do ponto de equilíbrio. Se duas variáveis são co-integradas, isso significa que elas convergem para uma condição de equilíbrio no longo prazo, no entanto, no curto prazo, pode ser que essas duas variáveis não consigam atingir uma situação de equilíbrio em função de algum fator que impeça essa convergência<sup>6</sup>. Entretanto, a presença do

<sup>6</sup>Entre os fatores econômicos que possam impedir que o sistema atinja o equilíbrio no curto prazo, tem-se em nível microeconômico a celebração de contratos entre empresas, enquanto que, em termos macroeconômicos, destaca-se a implementação de instrumentos de política econômica, com ênfase na própria política comercial, a qual pode gerar entraves ao fluxo de comércio internacional.

modelo de correção de erro permite que essas duas variáveis possam convergir para uma situação de equilíbrio no longo prazo. Outro aspecto relevante em relação à interface entre co-integração e modelo de correção de erro, segundo ENGLE e GRANGER (1991, p.10), é que "co-integração implica que o sistema segue uma representação de correção de erro e reciprocamente um sistema de correção de erro tem variáveis co-integradas".

ENGLE e GRANGER (1991, p.84) definiram a co-integração de um processo da seguinte forma: "Os componentes de um vetor  $x_t$  é dito ser co-integrado de ordem  $d, b$ , denotado por  $x_t \sim CI(d, b)$ , se: a) todos os componentes de  $x_t$  são  $I(d)$ ; b) existe um vetor  $\alpha (\neq 0)$  tal que  $z_t = \alpha' x_t \sim I(d, b)$ ,  $b > 0$ . O vetor  $\alpha$  é chamado de *vetor de co-integração*". Em outras palavras, se ao ajustar um modelo de regressão entre duas variáveis e ambas forem integradas de ordem 1, isso quer dizer que essas variáveis apresentam uma combinação linear ( $d = b = 1$ ) e como resultado os resíduos dessa regressão são  $I(0)$ , isto é, são estacionários e, conseqüentemente, essas duas variáveis são co-integradas.

Portanto, partindo-se das definições dadas acima, testar a presença de co-integração pelo método desenvolvido por ENGLE e GRANGER (1991) significa verificar se as variáveis  $y_t$  e  $x_t$  derivam de um processo gerador estacionário ( $I(0)$ ) ou não estacionário ( $I(1)$ ). Portanto, o procedimento inicial reside em verificar se tanto  $y_t$  quanto  $x_t$  contêm ou não raiz unitária. Se ambas forem caracterizadas por um processo  $I(1)$ , então para testar se essas variáveis são co-integradas ou não, deve-se estimar a seguinte equação de co-integração:

$$y_t = \alpha + \beta x_t + u_t \quad \text{ou} \quad y_t - \alpha - \beta x_t = u_t \quad (3)$$

e testar posteriormente se o resíduo estimado ( $u_t$ ) é estacionário ( $I(0)$ ) ou não, através da aplicação de teste de raiz unitária. Os valores críticos para os testes de raiz unitária sobre os resíduos da equação de co-integração podem ser encontrados em ENGLE e GRANGER (1991), PHILLIPS e OULIARIS (1990) e MACKINNON (1991).

Assim como no caso dos testes de raiz unitária, os testes de co-integração admitem as hipóteses de que o modelo tenha somente constante; ou tenha constante e tendência. Outra ana-

logia em relação aos testes de raiz unitária é que os testes de co-integração podem ser ampliados para verificar se há presença de correlação serial nos resíduos. Nesse caso, incorporam-se defasagens contendo a variável dependente diferenciada para se obter um ruído branco (*white noise*). Enquanto que esse procedimento, para a aplicação de testes de raiz unitária, denomina-se Dickey-Fuller Aumentado (ADF), no caso do teste de co-integração passa a ser denominado de Engle-Granger Aumentado (AEG). Outra semelhança com testes de raiz unitária diz respeito aos testes não paramétricos propostos por PHILLIPS (1987). Segundo MACKINNON (1991, p.270-271), um "procedimento diferente para obter os testes de raiz unitária que é assintoticamente válido na presença de correlação serial e/ou heterocedasticidade de forma desconhecida foi sugerido por PHILLIPS (1987) e estendido para o caso de co-integração por PHILLIPS e OULIARIS (1990). A distribuição assintótica a qual Phillips e Ouliaris chamam de estatística  $Z_t$  são idênticas àquelas correspondentes aos testes DF, ADF, EG e AEG. Phillips e Ouliaris tabularam valores críticos para duas formas de suas estatísticas (correspondendo às versões sem tendência e com tendência das estatísticas DF e EG) para vários valores de  $N$ ". No entanto, segundo MACKINNON (1991), os valores críticos elaborados por Phillips e Ouliaris sofrem de considerável erro experimental, ou seja, dependendo do tamanho da amostra adotado, seus valores são viesados, isto é, não convergem para as estimativas dos verdadeiros valores críticos assintóticos.

Portanto, dado que o teste de co-integração envolve a realização de teste de raiz unitária do tipo Dickey-Fuller Aumentado (ADF), sobre os resíduos da equação de co-integração, esse teste deve ser efetuado sobre a seguinte equação de resíduos:

$$\nabla \hat{u}_t = \alpha + \rho \hat{u}_{t-1} + \sum_{i=1}^p \gamma_i \nabla \hat{u}_{t-i} + e_t \quad (4)$$

Nesse caso, o termo  $\nabla$  representa a primeira diferença da variável  $\hat{u}_t$ , enquanto que  $\nabla \hat{u}_{t-i}$  é a primeira diferença defasada da variável  $\hat{u}_t$ , sendo que o número total de defasagens corresponde a  $i$  períodos e  $\alpha$  é uma constante. Quan-

do o termo  $\gamma_i$  é igual a zero para todos os  $i$ s, ao invés do teste Dickey-Fuller Aumentado (ADF), tem-se o teste Dickey-Fuller (DF).

O procedimento de verificação de presença ou não de raiz unitária, tendo como base os resíduos da equação de co-integração, testa a hipótese nula de que a série de resíduo tem raiz unitária ( $\rho = 0$ ) contra a hipótese alternativa de que os resíduos são estacionários ( $\rho \neq 0$ ), ou seja, testa a hipótese nula de que não há co-integração *versus* a hipótese alternativa de que existe co-integração entre as variáveis  $y_t$  e  $x_t$ .

#### 4 - TESTES DE CO-INTEGRAÇÃO UTILIZANDO O SOFTWARE SAS®

##### 4.1 - Construindo um Procedimento para fazer o Teste de Co-integração Engle-Granger no SAS

O primeiro passo para fazer o teste de co-integração consiste em verificar a ordem de integração das variáveis via testes de raiz unitária. Neste caso foi utilizado o teste de raiz unitária Dickey-Fuller Aumentado (ADF) conforme apresentado em DICKEY e FULLER (1979) e DICKEY e FULLER (1981). Na versão 6.12 do SAS a execução desse teste é descrito no quadro 1.

Após a realização do teste de raiz unitária para cada variável individualmente, e constatada que as variáveis são integradas de mesma ordem, o próximo passo consiste na estimação de uma regressão com as variáveis em nível (variáveis originais)<sup>7</sup>. A seguir, os resíduos dessa regressão são "capturados" e submetidos ao teste de raiz unitária ADF.

Caso esses resíduos sejam estacionários, ou seja, se o valor calculado do teste de raiz unitária for superior em módulo ao respectivo valor tabelado, isso implica que as variáveis são co-integradas. Dado que as variáveis são co-integradas, então, necessariamente existe um modelo de correção de erro, o qual é representado pelos resíduos defasados de um período da regressão estimada com as variáveis em nível. Na última etapa, esses resíduos defasados em

<sup>7</sup>Neste ponto é necessário destacar que as variáveis utilizadas neste estudo estão no formato logarítmico, logo, os seus respectivos coeficientes fornecem diretamente as elasticidades.

QUADRO 1 - Procedimento para Realização do Teste Dickey-Fuller Aumentado (ADF) na Versão 6.12 do SAS

| Procedimento para realização do teste             | Descrição dos comandos/opções   |
|---|---|
| PROC ARIMA DATA=TEST;                             | Esse procedimento realiza o teste ADF, utilizando os dados do arquivo SAS <i>data set</i> TEST;   |
| IDENTIFY VAR=X<br>MINIC<br>STATIONARITY=(ADF=12); | O comando IDENTIFY identifica as seguintes instruções: VAR=X, que especifica a variável utilizada na análise, que nesse caso é denominada X; MINIC identifica a ordem de um processo ARMA estacionário e inversível; STATIONARITY=(ADF=12), que determina a utilização do teste ADF para doze defasagens ( <i>lags</i> ); |
| RUN;  | Esse comando executa o procedimento ARIMA.  |

Fonte: Adaptado de SAS INSTITUTE (1996, p.5).

um período são incorporados como variável de entrada numa regressão, porém, neste caso todas as demais variáveis são utilizadas diferenciadas.

O modelo de correção de erro torna-se importante pelo fato de permitir a ligação entre aspectos relacionados com a dinâmica de curto com os de longo prazo. Em outras palavras, segundo BANERJEE et al. (1993, p.139) os “mecanismos de correção de erro pretendem fornecer um caminho para combinar as vantagens de se modelar tanto em nível quanto nas diferenças. Em um modelo de correção de erro tanto a dinâmica do processo de ajustamento de curto prazo (variações), quanto de longo prazo (níveis) são modelados simultaneamente”. Quanto maior o valor do coeficiente de correção de erro, mais rápido o sistema converge para o equilíbrio no longo prazo. Matematicamente, o modelo de correção tem o seguinte aspecto:

$$\nabla y_t = \omega_0 + \omega_1 \nabla x_t - \gamma(y_{t-1} - \beta x_{t-1} - \alpha) + \varepsilon_t \quad (5)$$

onde:  $\nabla$  é o operador diferença ( $\nabla y_t = Y_t - Y_{t-1}$ ).

Sendo assim, a rotina construída passo a passo descrevendo o teste Engle-Granger no SAS assume o formato apresentado nos quadros 2 a 5.

O SAS versão 6.12 também permite que se faça o teste de co-integração do tipo Phillips-Ouliaris, o qual é uma derivação do teste Engle-Granger. Nesse caso, a principal vantagem para o usuário reside no fato de que não há necessidade de se construir um procedimento para

fazer o teste de co-integração, como acontece no caso do teste Engle-Granger, dado que o SAS faz esse teste diretamente utilizando o procedimento AUTOREG conforme apresentado em SAS INSTITUTE (1995). No entanto, como no caso anterior, quando o teste indicar que as variáveis são co-integradas, então existe um modelo de correção de erro, o qual tem que ser construído, pois o SAS somente faz o teste de co-integração.

Assim como aconteceu com o teste Engle-Granger, antes de fazer o teste Phillips-Ouliaris é necessário realizar um teste de raiz unitária para averiguar a ordem de integração das variáveis. Só que neste caso, ao invés de se utilizar o teste ADF, deve-se utilizar o teste de raiz unitária desenvolvido por PHILLIPS e PERRON (1988), o qual é denominado de teste PP. Os comandos para fazer o teste PP na versão 6.12 do SAS encontram-se no quadro 6.

Se o teste de raiz unitária PP indicar que as variáveis são integradas de mesma ordem, ou mais precisamente, que elas são diferenças estacionárias (integradas de ordem um), então, o próximo passo consiste na estimação da regressão com as variáveis em nível via o procedimento AUTOREG, o qual fornece diretamente as estimativas dos parâmetros da equação de regressão em nível, além do próprio teste de raiz unitária sobre os resíduos. Apesar de o SAS proporcionar a execução direta de teste de co-integração do tipo Phillips-Ouliaris, a sua desvantagem é de que a sua saída dos resultados não produz os valores de probabilidade em relação a esse teste. Dessa forma, para tomar uma decisão em relação à estacionariedade ou não dos resíduos, o usuário é obrigado a recorrer ao

## QUADRO 2 - Procedimento para Estimação da Regressão com as Variáveis em Nível

| Procedimento para estimar uma regressão com as variáveis em nível | Descrição dos comandos/opções  |
|---|--|
| PROC REG DATA=TEST;   | Início do procedimento REG; DATA=TEST o comando DATA introduz o nome do arquivo SAS <i>data set</i> que contém os dados relativos às variáveis que serão utilizadas nesse modelo;  |
| MODEL YLOG = TIME XLOG;<br>output out = a residual= res;          | O comando MODEL especifica que as variáveis YLOG, TIME (tendência) e XLOG serão utilizadas nesse modelo; o comando OUTPUT gerencia a saída dos resultados, os quais são armazenados no arquivo <i>a</i> , também, é criada uma variável contendo os resíduos da regressão, a qual é denominada de <i>res</i> ; |
| RUN;  | Esse comando executa o procedimento REG.   |

Fonte: Adaptado de BROCKLEBANK e DICKEY (1986, p.92).

## QUADRO 3 - Procedimento para Armazenar os Resíduos da Regressão Estimada com as Variáveis em Nível

| Procedimento para "capturar" os resíduos da regressão estimada com as variáveis em nível | Descrição dos comandos/opções   |
|--|---|
| DATA a1;   | O comando DATA introduz o nome do arquivo SAS <i>data set</i> que contém os dados relativos aos resíduos da regressão estimada com as variáveis em nível que serão armazenadas no arquivo <i>a1</i> ; |
| SET a;   | Utilizar os dados armazenados no arquivo <i>a</i> ;   |
| res1=lag1(res);  | Cria a variável <i>res1</i> , a qual é representada pelos resíduos da regressão estimada com as variáveis em nível e defasada ( <i>lag1</i> ) de um período;  |
| DYLOG=DIF(YLOG);   | O comando DIF permite a criação da variável de saída diferenciada (YLOG) de ordem igual a um;   |
| DXLOG=DIF(XLOG);   | O comando DIF permite a criação da variável de entrada diferenciada (XLOG) de ordem igual a um;   |
| RUN;   | Executa os comandos estabelecidos acima.  |

Fonte: Adaptado de BROCKLEBANK e DICKEY (1986, p.92).

## QUADRO 4 - Procedimento para a Realização do Teste de Raiz Unitária Dickey-Fuller Aumentado na Versão 6.12 do SAS sobre os Resíduos da Regressão Estimada com as Variáveis em Nível

| Procedimento para realização do teste               | Descrição dos comandos/opções   |
|---|---|
| PROC ARIMA DATA=a;                                  | Esse procedimento realiza o teste ADF, utilizando os dados do arquivo SAS <i>data set a</i> ;   |
| IDENTIFY VAR=RES<br>MINIC<br>STATIONARITY=(ADF=12); | O comando IDENTIFY identifica as seguintes instruções: VAR=RES, que especifica a variável utilizada na análise, que nesse caso é denominada RES; MINIC identifica a ordem de um processo ARMA estacionário e inversível; STATIONARITY=(ADF=12), que determina a utilização do teste ADF para doze defasagens ( <i>lags</i> ); |
| RUN;  | Esse comando executa o procedimento ARIMA.  |

Fonte: Adaptado de SAS INSTITUTE (1996, p. 5).

### QUADRO 5 - Procedimento para Estimação da Regressão com as Variáveis Diferenciadas e do Modelo de Correção de Erro

| Procedimento para estimar uma regressão com as variáveis em nível | Descrição dos comandos/opções   |
|---|---|
| PROC REG DATA=a1;   | Início do procedimento REG; DATA=a1 o comando DATA introduz o nome do arquivo SAS <i>data set</i> que contém os dados relativos às variáveis que serão utilizadas nesse modelo;   |
| MODEL DYLOG = DXLOG RES1 / NOINT<br>P CLI;                        | O comando MODEL especifica que as variáveis DYLOG, DXLOG e RES1 serão utilizadas nesse modelo, onde DYLOG e DXLOG são os logaritmos das variáveis diferenciadas e RES1 representa os resíduos defasados de um período, enquanto que NOINT indica que não será introduzido um termo de intercepto e P e CLI são comandos para fazer previsões; |
| RUN;  | Esse comando executa o procedimento REG.  |

Fonte: Adaptado de BROCKLEBANK e DICKEY (1986, p.92).

### QUADRO 6 - Procedimento para Realização do Teste Phillips-Perron na Versão 6.12 do SAS

| Procedimento para realização do teste            | Descrição dos comandos/opções   |
|--|---|
| PROC ARIMA DATA=TEST;                            | Esse procedimento realiza o teste de Phillips-Perron (assim como o teste ADF), utilizando os dados do arquivo SAS <i>data set</i> TEST;   |
| IDENTIFY VAR=X<br>MINIC<br>STATIONARITY=(PP=12); | O comando IDENTIFY identifica os seguintes parâmetros: VAR=X, que especifica a variável utilizada na análise, que nesse caso é denominada X; a opção MINIC identifica a ordem de um processo ARMA estacionário e inversível; STATIONARITY=(PP=12), determina a utilização do teste <i>Phillips-Perron</i> para um total de 12 defasagens ( <i>lags</i> ), ou seja, incorpora as defasagens de ordens zero até 12 conjuntamente; |
| RUN;   | Esse comando executa o procedimento ARIMA.  |

Fonte: Adaptado de SAS INSTITUTE (1996, p.5).

trabalho de MACKINNON (1991) para obter os valores críticos tabulados. A descrição dos comandos para a realização do teste Phillips-Ouliaris encontra-se no quadro 7.

Uma vez que as variáveis são co-integradas, então, necessariamente existe um modelo de correção de erro, sendo assim, o usuário pode repetir os procedimentos descritos nos quadros 3 e 5 para finalizar a análise relativa ao comportamento de curto e longo prazo sobre as variáveis em estudo.

#### 4.2 - Exemplo Numérico: resultados e análise

Apesar de o Brasil ser um dos mais

importantes produtores e exportadores de soja em nível mundial, ainda assim, os preços domésticos da soja no Brasil são formados no mercado internacional. Segundo MARGARIDO et al. (1999), de um lado tem-se a Chicago Board of Trade (CBOT), Bolsa de futuros norte-americana, a qual exerce papel preponderante no que se refere à formação do preço no mercado internacional de soja, pois representa tanto as forças de oferta quanto de demanda desse produto no mercado mundial. Por outro, o complexo portuário de Rotterdam destaca-se por ser um dos principais pontos de exportação e importação de soja e seus derivados em nível mundial, além de ser relevante pólo processador e distribuidor de soja e seus derivados dentro da União Européia (UE).

QUADRO 7 - Procedimento para Realização do Teste de Co-integração Phillips-Ouliaris na Versão 6.12 do SAS

| Procedimento para estimar uma regressão com as variáveis em nível                 | Descrição dos comandos/opções   |
|---|---|
| PROC AUTOREG DATA=TEST;   | Início do procedimento AUTOREG; DATA=TEST o comando DATA introduz o nome do arquivo SAS <i>data set</i> que contém os dados relativos às variáveis que serão utilizadas nesse modelo;   |
| MODEL YLOG = XLOG / STATIONARITY=(PHILLIPS=(1));<br>output out = a residual= res; | O comando MODEL especifica que as variáveis YLOG e XLOG serão utilizadas nesse modelo; o símbolo / indica que a partir dele serão introduzidas as opções desse comando que nesse caso refere-se ao teste de co-integração do tipo Phillips-Ouliaris que tem como base o teste de raiz unitária Phillips-Perron, o comando OUTPUT gerencia a saída dos resultados, os quais são armazenados no arquivo <i>a</i> , também, é criada uma variável contendo os resíduos da regressão, a qual é denominada de <i>res</i> ; |
| RUN;  | Esse comando executa o procedimento AUTOREG.  |

Fonte: Adaptado de SAS INSTITUTE (1995, p.26).

Cerca de 70,0% da soja exportada pelo Brasil e 75,0% da soja argentina têm como destino a UE. Em função dessas características do próprio mercado internacional de soja, os preços domésticos, tanto no Brasil quanto na Argentina, são influenciados pelas variações nos preços internacionais da soja, fato comprovado por diversos estudos entre os quais pode-se citar MARGARIDO et al. (1999), MARGARIDO e SOUSA (1998), LAZZARINI (1997), PINO e ROCHA (1994), NEVES (1993), AGUIAR e BARROS (1991) e PINO; NOGUEIRA JÚNIOR; TOLOI (1983).

Para identificar a ordem de integração das variáveis utilizou-se o teste Dickey-Fuller Aumentado (ADF), o qual está incorporado ao procedimento ARIMA através do comando *stationarity* no SAS. Para facilitar a compreensão por parte do leitor, as saídas do SAS para o teste ADF foram reproduzidas nas tabelas de 1 a 6.

O comando MINIC dentro do procedimento ARIMA permite determinar o número de defasagens a serem utilizadas no teste de raiz unitária, dado que esse comando produz valores para o Critério de Informação de SCHWARZ<sup>8</sup>. O

<sup>8</sup>Basicamente, os critérios de informação, tanto de Schwarz (Schwarz's Bayesian - SBC ou Bayesian Information Criterion - BIC) quanto de Akaike (Akaike's Information Criterion - AIC), são utilizados para a determinação do número de defasagens a serem incorporadas ao teste de raiz unitária e objetivam eliminar a autocorrelação dos resíduos. Maiores detalhes sobre esses dois critérios podem ser obtidos em MARGARIDO; ANEFALOS (1999), PINDYCK; RUBINFELD (1988), DAVIDSON (2000) e MADALA; KIM (1998), entre outros.

número de defasagens a serem incorporadas nos testes de raiz unitária corresponde ao menor valor obtido pelo Critério de Informação de SCHWARZ.

Para a variável relativa ao logaritmo do preço Free on Board (FOB) do farelo de soja no Brasil (*BRLOG*) em nível, o critério de informação obteve valor mínimo para um modelo auto-regressivo de ordem um. Sendo assim, foi utilizada somente uma defasagem no teste de raiz unitária.

Os valores calculados para os testes individuais e conjuntos que correspondem às estatísticas  $\tau$ ,  $\tau_{\mu}$ ,  $\tau_{\tau}$  e  $\phi_3$ ,  $\phi_1$ , respectivamente, mostraram-se não significativos, dado que seus respectivos valores de probabilidades encontram-se acima de 10,0%. Portanto, não se pode rejeitar a hipótese de raiz unitária para *BRLOG* em nível (Tabela 1). Em função desse resultado, o teste de raiz unitária deve ser repetido, só que, desta vez, com a variável diferenciada.

Neste caso, o Critério de Informação de SCHWARZ apresentou menor valor para um modelo sem a necessidade de se incluir defasagens (ordem zero). Todos os valores calculados das estatísticas utilizadas no teste Dickey-Fuller mostraram-se significativos, pois em todos os casos analisados os valores calculados apresentaram valores de probabilidade abaixo de 1,0% (Tabela 2). Portanto, pode-se concluir que a hipótese nula de raiz unitária foi rejeitada em detrimento da hipótese alternativa de que a variável é estacionária nas diferenças, ou seja, *BRLOG* é estacionária na primeira diferença e

TABELA 1 - Resultado no SAS do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação à Variável BRLOG em Nível

| Type                     | Lags | RHO     | Prob<RHO | T                    | Prob<T | F                   | Prob<F |
|--------------------------|------|---------|----------|----------------------|--------|---------------------|--------|
| Zero mean <sup>1</sup>   | 1    | -4.7016 | 0.1339   | -1.4955 <sup>4</sup> | 0.1258 | -                   | -      |
| Single mean <sup>2</sup> | 1    | -4.6985 | 0.4575   | -1.4875 <sup>5</sup> | 0.5366 | 1.109 <sup>7</sup>  | 0.7882 |
| Trend <sup>3</sup>       | 1    | -4.4658 | 0.8531   | -1.3929 <sup>6</sup> | 0.8573 | 1.1697 <sup>8</sup> | 0.9424 |

<sup>1</sup>Modelo sem constante e sem tendência.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>6</sup>Corresponde à estatística  $\tau_{\tau}$ .

<sup>7</sup>Corresponde à estatística  $\phi_1$ .

<sup>8</sup>Corresponde à estatística  $\phi_3$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 2 - Resultado no SAS do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação à Variável BRLOG Diferenciada

| Type                     | Lags | RHO      | Prob<RHO | T                    | Prob<T | F                    | Prob<F |
|--------------------------|------|----------|----------|----------------------|--------|----------------------|--------|
| Zero mean <sup>1</sup>   | 0    | -98.8081 | 0.0001   | -9.4979 <sup>4</sup> | 0.0001 | -                    | -      |
| Single mean <sup>2</sup> | 0    | -98.7964 | 0.001    | -9.4508 <sup>5</sup> | 0.0001 | 44.6812 <sup>7</sup> | 0.001  |
| Trend <sup>3</sup>       | 0    | -99.3257 | 0.0004   | -9.4434 <sup>6</sup> | 0.0001 | 44.5969 <sup>8</sup> | 0.001  |

<sup>1</sup>Modelo sem constante e sem tendência.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>6</sup>Corresponde à estatística  $\tau_{\tau}$ .

<sup>7</sup>Corresponde à estatística  $\phi_1$ .

<sup>8</sup>Corresponde à estatística  $\phi_3$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 3 - Resultado no SAS do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação à Variável ARGLOG em Nível

| Type                     | Lags | RHO     | Prob<RHO | T                    | Prob<T | F                   | Prob<F |
|--------------------------|------|---------|----------|----------------------|--------|---------------------|--------|
| Zero mean <sup>1</sup>   | 2    | -0.0407 | 0.6718   | -0.2756 <sup>4</sup> | 0.5846 | -                   | -      |
| Single mean <sup>2</sup> | 2    | -6.2272 | 0.3208   | -1.6359 <sup>5</sup> | 0.4609 | 1.3602 <sup>7</sup> | 0.7242 |
| Trend <sup>3</sup>       | 2    | -6.1983 | 0.7195   | -1.6378 <sup>6</sup> | 0.7714 | 1.4738 <sup>8</sup> | 0.8832 |

<sup>1</sup>Modelo sem constante e sem tendência.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>6</sup>Corresponde à estatística  $\tau_{\tau}$ .

<sup>7</sup>Corresponde à estatística  $\phi_1$ .

<sup>8</sup>Corresponde à estatística  $\phi_3$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).



TABELA 4 - Resultado no SAS do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação à Variável ARGLOG Diferenciada

| Type                     | Lags | RHO      | Prob<RHO | T                    | Prob<T | F                   | Prob<F |
|--------------------------|------|----------|----------|----------------------|--------|---------------------|--------|
| Zero mean <sup>1</sup>   | 4    | -90.4065 | 0.0001   | -4.204 <sup>4</sup>  | 0.0001 | -                   | -      |
| Single mean <sup>2</sup> | 4    | -89.5805 | 0.001    | -4.1667 <sup>5</sup> | 0.0012 | 8.7775 <sup>7</sup> | 0.001  |
| Trend <sup>3</sup>       | 4    | -101.612 | 0.0001   | -4.2324 <sup>6</sup> | 0.0058 | 8.9723 <sup>8</sup> | 0.001  |

<sup>1</sup>Modelo sem constante e sem tendência.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>6</sup>Corresponde à estatística  $\tau_{\tau}$ .

<sup>7</sup>Corresponde à estatística  $\phi_1$ .

<sup>8</sup>Corresponde à estatística  $\phi_3$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 5 - Resultado no SAS do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação à Variável ROTLOG em Nível

| Type                     | Lags | RHO     | Prob<RHO | T                    | Prob<T | F                   | Prob<F |
|--------------------------|------|---------|----------|----------------------|--------|---------------------|--------|
| Zero mean <sup>1</sup>   | 1    | -6.6448 | 0.0732   | -1.7477 <sup>4</sup> | 0.0765 | -                   | -      |
| Single mean <sup>2</sup> | 1    | -6.6215 | 0.2918   | -1.7359 <sup>5</sup> | 0.4104 | 1.5277 <sup>7</sup> | 0.6815 |
| Trend <sup>3</sup>       | 1    | -6.7828 | 0.6707   | -1.7766 <sup>6</sup> | 0.7098 | 1.6729 <sup>8</sup> | 0.8434 |

<sup>1</sup>Modelo sem constante e sem tendência.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>6</sup>Corresponde à estatística  $\tau_{\tau}$ .

<sup>7</sup>Corresponde à estatística  $\phi_1$ .

<sup>8</sup>Corresponde à estatística  $\phi_3$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 6 - Resultado no SAS do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação à Variável ROTLOG Diferenciada

| Type                     | Lags | RHO      | Prob<RHO | T                    | Prob<T | F                   | Prob<F |
|--------------------------|------|----------|----------|----------------------|--------|---------------------|--------|
| Zero mean <sup>1</sup>   | 3    | -72.0451 | 0.0001   | -4.3385 <sup>4</sup> | 0.0001 | -                   | -      |
| Single mean <sup>2</sup> | 3    | -71.7177 | 0.001    | -4.3044 <sup>5</sup> | 0.0008 | 9.3338 <sup>7</sup> | 0.001  |
| Trend <sup>3</sup>       | 3    | -76.7615 | 0.0003   | -4.3234 <sup>6</sup> | 0.0043 | 9.3999 <sup>8</sup> | 0.001  |

<sup>1</sup>Modelo sem constante e sem tendência.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>6</sup>Corresponde à estatística  $\tau_{\tau}$ .

<sup>7</sup>Corresponde à estatística  $\phi_1$ .

<sup>8</sup>Corresponde à estatística  $\phi_3$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

conseqüentemente é integrada de ordem um.

O mesmo tipo de análise realizada anteriormente se aplica à variável logaritmo do preço *FOB* do farelo de soja na Argentina (*ARGLOG*) e para o logaritmo do preço *Cost Insurance and Freight (CIF)* do farelo de soja no porto de Rotterdam (*ROTLOG*), cujos resultados encontram-se nas tabelas 3 e 5 para as variáveis em nível e nas tabelas 4 e 6 para as variáveis diferenciadas, respectivamente.

Determinada a ordem de integração, o próximo passo consistiu na estimação de regressões com as variáveis em nível, visando, dessa forma, obter os respectivos resíduos para que posteriormente se pudesse fazer o teste de co-integração. Nos dois modelos de regressão estimados, a variável de entrada é *ROTLOG*, enquanto que, na primeira regressão, a variável de saída é *BRLOG* (Tabela 7) e no segundo é *ARGLOG* (Tabela 8).

Os resultados relativos à equação de regressão, tendo como variável de entrada *ROTLOG* e de saída *BRLOG*, mostram que o teste *t* para o intercepto, tendência e *ROTLOG* é significativo para cada um deles, dado que o valor da probabilidade está abaixo do nível de significância de 1,0% (Tabela 7).

Assim como aconteceu no caso anterior, a regressão contendo as variáveis *ARGLOG* (saída) e *RTOLOG* (entrada) em nível também apresentou significância em relação ao teste *t* para todas as estimativas dos parâmetros em nível de 1,0% (Tabela 8).

Uma vez estimadas as regressões com as variáveis em nível, o próximo passo é aplicar o teste de raiz unitária ADF sobre os resíduos de cada regressão anteriormente estimada.

Neste caso, o teste incide sobre os modelos contendo somente constante e contendo constante e tendência. Os resultados mostraram que os resíduos da primeira regressão são estacionários, ou seja, pode-se rejeitar a hipótese nula de raiz unitária, pois os valores calculados para as estatísticas  $\tau_{\mu}$  e  $\tau_{\tau}$  são significativos em relação aos respectivos valores de probabilidade em nível de 1,0% (Tabela 9). Portanto, dado que os resíduos são estacionários, implica que as duas variáveis estimadas na regressão em nível são co-integradas, ou seja, possuem um relacionamento de longo prazo. Sendo assim, ao invés de se utilizar a regressão em nível, é necessário estimá-la com as variáveis diferenciadas

(captar aspectos de curto prazo), mais os resíduos da regressão com as variáveis em nível defasados de um período (termo de correção de erro), e, conseqüentemente, estimar o modelo de correção de erro (MCE) para medir o desequilíbrio entre o período atual e anterior, captando dessa forma aspectos de longo prazo.

Os mesmos resultados foram obtidos para os resíduos da segunda regressão (Tabela 10), e sua análise segue o mesmo caminho apresentado em relação à primeira regressão.

Os resultados estimados para o modelo de correção de erro com as variáveis diferenciadas mais os resíduos defasados de um período mostraram que a estimativa do coeficiente de *DROTLOG* é igual a 0,9841, ou seja, variações nos preços internacionais dos preços de farelo de soja no porto de Rotterdam são repassados menos que proporcionalmente aos preços *FOB* do farelo de soja no Brasil, pois somente 98,41% das variações do preço em Rotterdam são transferidas para os preços no Brasil, isto é, nesse caso a elasticidade de transmissão de preços<sup>9</sup> é inelástica, dado que ela é menor que a unidade (Tabela 11).

Os resíduos defasados de um período (*RES1*) mostram que no caso do Brasil o desequilíbrio existente no período atual relativamente ao período anterior é corrigido na velocidade de 34,29%, ou seja, esse desequilíbrio é corrigido lentamente, pois quanto maior o valor da estimativa do parâmetro do resíduo defasado de um período, mais rápido o equilíbrio é atingido.

No caso do segundo modelo de correção de erro contendo as variáveis diferenciadas e os resíduos defasados de um período, o coeficiente estimado de *DROTLOG* assumiu valor igual a 1,1831. Isso quer dizer que variações nos preços do farelo de soja em Rotterdam são transmitidos mais que proporcionalmente para os preços do farelo no mercado argentino, uma vez que variações dos preços do farelo em Rotterdam são transferidos com magnitude igual a 118,31, implicando que o relacionamento entre os preços internacionais do farelo e os seus respectivos pre-

<sup>9</sup>Conforme BARROS e BURNQUIST (1987), a elasticidade de transmissão de preços refere-se à variação relativa no preço a um nível de mercado em relação à variação relativa no preço a outro nível, mantidos em equilíbrio esses dois níveis de mercado após o choque inicial em um deles.

TABELA 7 - Resultados do SAS<sup>1</sup> para a Regressão Contendo as Variáveis BRLOG e ROTLOG em Nível

| Variable | DF <sup>2</sup> | Parameter estimates |                |                          |           |
|----------|-----------------|---------------------|----------------|--------------------------|-----------|
|          |                 | Parameter estimate  | Standard error | T for H0:<br>Parameter=0 | Prob >  T |
| INTERCEP | 1               | -0.55606            | 0.125274       | -4.439                   | 0.0001    |
| TIME     | 1               | 0.0018              | 0.000145       | 12.422                   | 0.0001    |
| ROTLOG   | 1               | 1.077281            | 0.02331        | 46.215                   | 0.0001    |

<sup>1</sup>É necessário observar que somente parte da saída do SAS relativa ao procedimento REG foi reproduzida.

<sup>2</sup>Grau de Liberdade.

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 8 - Resultados do SAS<sup>1</sup> para a Regressão Contendo as Variáveis ARGLOG e ROTLOG em Nível

| Variable | DF <sup>2</sup> | Parameter estimates |                |                          |           |
|----------|-----------------|---------------------|----------------|--------------------------|-----------|
|          |                 | Parameter estimate  | Standard error | T for H0:<br>Parameter=0 | Prob >  T |
| INTERCEP | 1               | -0.888040           | 0.08915328     | -9.961                   | 0.0001    |
| TIME     | 1               | 0.000374            | 0.00010315     | 3.625                    | 0.0004    |
| ROTLOG   | 1               | 1.138070            | 0.01658930     | 68.603                   | 0.0001    |

<sup>1</sup>É necessário observar que somente parte da saída do SAS relativa ao procedimento REG foi reproduzida.

<sup>2</sup>Grau de Liberdade.

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 9 - Resultado no SAS<sup>1</sup> do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação aos Resíduos da Primeira Equação de Regressão (Tabela 7)

| Type                     | Lags | RHO      | Prob<RHO | T                    | Prob<T |
|--------------------------|------|----------|----------|----------------------|--------|
| Single mean <sup>2</sup> | 1    | -42.3266 | 0.0010   | -4.5212 <sup>4</sup> | 0.0004 |
| Trend <sup>3</sup>       | 1    | -42.3213 | 0.0004   | -4.4981 <sup>5</sup> | 0.0024 |

<sup>1</sup>É necessário observar que somente parte da saída do SAS relativa ao procedimento ARIMA foi reproduzida.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\tau}$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 10 - Resultado no SAS<sup>1</sup> do Teste de Raiz Unitária Utilizando Dickey-Fuller Aumentado (Augmented Dickey-Fuller Unit Roots Tests) em Relação aos Resíduos da Segunda Equação de Regressão (Tabela 8)

| Type                     | Lags | RHO      | Prob<RHO | T                    | Prob<T |
|--------------------------|------|----------|----------|----------------------|--------|
| Single mean <sup>2</sup> | 1    | -42.6237 | 0.0010   | -4.7703 <sup>4</sup> | 0.0002 |
| Trend <sup>3</sup>       | 1    | -42.6466 | 0.0004   | -4.7549 <sup>5</sup> | 0.0010 |

<sup>1</sup>É necessário observar que somente parte da saída do SAS relativa ao procedimento ARIMA foi reproduzida.

<sup>2</sup>Modelo com constante e sem tendência.

<sup>3</sup>Modelo com constante e com tendência.

<sup>4</sup>Corresponde à estatística  $\tau_{\mu}$ .

<sup>5</sup>Corresponde à estatística  $\tau_{\tau}$ .

Fonte: Dados básicos da OILSEEDS (1990-1999).

TABELA 11 - Resultados do SAS<sup>1</sup> para a Regressão (Modelo de Correção de Erro) Contendo as Variáveis BRLOG e ROTLOG Diferenciadas e com Termo de Correção de Erro

| Variable | DF <sup>2</sup> | Parameter estimates |                |             |           |
|----------|-----------------|---------------------|----------------|-------------|-----------|
|          |                 | Parameter estimate  | Standard error | T for H0:   |           |
|          |                 |                     |                | Parameter=0 | Prob >  T |
| DROTLOG  | 1               | 0.984192            | 0.07216012     | 13.639      | 0.0001    |
| RES1     | 1               | -0.342977           | 0.07773755     | -4.412      | 0.0001    |

<sup>1</sup>É necessário observar que somente parte da saída do SAS relativa ao procedimento REG foi reproduzida.

<sup>2</sup>Grau de Liberdade.

Fonte: Dados básicos da OILSEEDS (1990-1999).

ços no mercado doméstico argentino tem um caráter elástico, ou seja, a elasticidade de transmissão de preços é maior que a unidade, indicando que os preços do farelo de soja são mais sensíveis às variações de seu preço no mercado internacional comparativamente ao caso brasileiro (Tabela 12).

Neste caso, a estimativa dos resíduos defasados de um período (RES1) assumiu valor igual a -0,5016, indicando que, em relação ao período anterior, os preços no período atual são corrigidos em cerca de 50,16%, valor superior ao obtido no caso brasileiro (Tabela 12). Novamente, assim como aconteceu com as estimativas dos parâmetros de curto prazo, ao se levar em consideração os aspectos de longo prazo, os preços do farelo de soja na Argentina são mais sensíveis a variações de preços desse produto no mercado internacional relativamente aos preços domésticos praticados no Brasil.

Os resultados obtidos aqui são consistentes em função das características de cada mercado doméstico e de seu relacionamento com o mercado internacional. De forma resumida, no caso do Brasil, existe um grande mercado consumidor doméstico para a soja e seus derivados, pois o País é um importante produtor mundial de frango e o farelo de soja é um dos principais componentes para alimentação das aves, além do que, o esmagamento da soja necessariamente produz o óleo, o qual praticamente não tem demanda no mercado internacional junto aos países desenvolvidos, mas domesticamente tem grande mercado interno em função do seu baixo preço relativamente aos demais tipos de óleos. Logo, os preços domésticos no Brasil são menos sensíveis às variações do preço do farelo no mercado internacional. Já no caso da Argentina,

não existe um relevante mercado consumidor doméstico para a soja e seus derivados, sendo assim, toda a produção interna objetiva ser comercializada no mercado internacional. Portanto, os resultados obtidos pelos testes de co-integração captaram as características de cada mercado doméstico, ou seja, os preços domésticos do farelo de soja na Argentina são mais sensíveis às variações nos preços desse produto no mercado internacional comparativamente aos preços domésticos no Brasil, dado que a Argentina é mais dependente das condições de demanda que prevalecem no mercado mundial de farelo de soja do que o Brasil.

## 5 - CONSIDERAÇÕES FINAIS

Para aqueles que entendem de programação, os procedimentos utilizados neste texto podem ser considerados triviais. No entanto, para aqueles que não entendem de programação, como é o caso da maioria dos usuários, isso pode ser um abismo praticamente intransponível, a menos que eles tenham algum tipo de suporte por parte de programadores.

Dado que os testes de co-integração são importante instrumental para aqueles que trabalham com economia/econometria, este trabalho procurou colocar, numa linguagem direta e simplificada, a parte teórica sobre co-integração envolvendo os métodos desenvolvidos por ENGLE e GRANGER (1987 e 1991) e PHILLIPS e OULLIARIS (1990), além de apresentar ao usuário do SAS como construir uma rotina (procedimento) para realizar o teste de co-integração do tipo Engle-Granger, como executar o teste Phillips-Ouliaris, o qual está devidamente incorporado na

TABELA 12 - Resultados do SAS<sup>1</sup> para a Regressão (Modelo de Correção de Erro) Contendo as Variáveis ARGLOG e ROTLOG Diferenciadas e com Termo de Correção de Erro

| Variable | DF <sup>2</sup> | Parameter estimates |                |             |           |
|----------|-----------------|---------------------|----------------|-------------|-----------|
|          |                 | Parameter estimate  | Standard error | T for H0:   |           |
|          |                 |                     |                | Parameter=0 | Prob >  T |
| DROTLOG  | 1               | 1.183122            | 0.05580832     | 21.200      | 0.0001    |
| RES1     | 1               | -0.501684           | 0.08435173     | -5.948      | 0.0001    |

<sup>1</sup>É necessário observar que somente parte da saída do SAS relativa ao procedimento REG foi reproduzida.

<sup>2</sup>Grau de Liberdade.

Fonte: Dados básicos da OILSEEDS (1990-1999).

versão 6.12 do SAS, como construir uma rotina para o modelo de correção de erro para os dois tipos de testes de co-integração apresentados, assim como, incorporar esse modelo de correção de erro nas equações de co-integração e, finalmente, foi adicionado um exemplo numérico para mostrar ao usuário todo o caminho que ele deve percorrer para executar os testes de co-integra-

ção.

Portanto, a expectativa é de que este artigo seja de grande utilidade não somente para estudantes e professores, como também para aquelas pessoas que trabalhem em empresas e cujas funções estejam diretamente relacionadas com análise econômica tanto em nível estrutural quanto conjuntural.

## LITERATURA CITADA

AGUIAR, D. R. D.; BARROS, G. S. A. C. Causalidade e assimetria na transmissão de preços de soja e derivados no Brasil nos anos 80. **Estudos Econômicos**, São Paulo, v.21, n.1, p.89-103, 1991.

BANERJEE, A. et al. **Co-integration, error-correction, and the econometric analysis of non-stationary data**. New York: Oxford University Press, 1993. 329 p. (Advanced Texts in Econometrics).

BARROS, G. S. A. C.; BURNQUIST, H. L. Causalidade e transmissão de preços agrícolas entre níveis de atacado e varejo. In: ENCONTRO LATINO AMERICANO DA ECONOMETRIC SOCIETY, 7., 1987, São Paulo. **Anais...** São Paulo, 1987. p.175-190.

BROCKLEBANK, J. C.; DICKEY, D. A. **SAS Systems for Forecasting Time Series**. North Carolina: SAS INSTITUTE Inc./SAS Campus Drive, 1986. 240p.

DAVIDSON, J. **Econometric theory**. Great Britain: Blackwell, 2000. 499p.

DICKEY, D. A.; FULLER, W. A. Distribution of the estimator for autoregressive time series with a unit root. **Journal of the American Statistical Association**, v.74, n.366, p.427-31, Jun. 1979.

\_\_\_\_\_. Likelihood ratio statistics for autoregressive time series with a unit root. **Econometrica**, Chicago, v.49, n.4, p.1057-72, Jul. 1981.

ENGLE, R. F.; GRANGER, C. W. J. Cointegration and error correction: representation, estimation and testing. **Econometrica**, Chicago, v. 55, n. 2, p. 251-76, Mar. 1987.

\_\_\_\_\_. **Long-run economic relationship: readings in cointegration**. New York: Oxford University Press, 1991. 301p. (Advanced Texts in Econometrics).

- Margarido; f
- GRANGER, C.; NEWBOLD, P. Spurious regressions in econometrics. **Journal of Econometrics**, Nottingham, v.2, p.111-20, Jul. 1974.
- GUJARATI, D. N. **Basic econometrics**. 3rd ed. New York: McGraw-Hill, 1995. 838p.
- HORTON, G. A. Forecasting and economic impact analysis: glossary of statistical and forecasting terminology. Disponível em: [http://www.comstockbank.com/forecast/econ\\_pg2.htm](http://www.comstockbank.com/forecast/econ_pg2.htm). [Acesso em: 23 jul. 1998].
- LAZZARINI, S. G. **Inovação e organização de bolsas de futuros: teoria e evidências no agribusiness brasileiro**. São Paulo, 1997. 216p. Dissertação (Mestrado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo.
- MACKINNON, James G. Critical values for cointegration tests. In: ENGLE, R. F.; GRANGER, W. J. **Long-run economic relationships: readings in cointegration**. New York: Oxford University Press, 1991. p.267-76.
- MADDALA, G. S.; KIM, In-Moo. **Unit roots, cointegration, and structural change**. United Kingdom: Cambridge University Press, 1998. 505p.
- MARGARIDO, M. A.; ANEFALOS, L. C. Testes de raiz unitária e o software SAS. **Agricultura em São Paulo**, São Paulo, v.46, t.2, p.19-45, 1999.
- \_\_\_\_\_; SOUSA E. L. Leão de. Formação de preços da soja no Brasil. **Agricultura em São Paulo**, São Paulo, v.45, t.2, p.52-61, 1998.
- \_\_\_\_\_. et al. Transmissão de preços no mercado internacional de grãos de soja: uma aplicação da metodologia de séries temporais. In: CONGRESSO BRASILEIRO DE ECONOMIA E SOCIOLOGIA RURAL, 37., 1-5 ago. 1999, Foz do Iguaçu. **Anais...** Brasília: SOBER, 1999.
- NEVES, L. C. **Margens de comercialização e elasticidade de transmissão de preços na indústria de esmagamento de soja**. Piracicaba, 1993. 90p. Dissertação (Mestrado em Agronomia) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo.
- OILSEEDS: World Market and Trade. Washington: USDA, 1990-1999.
- PHILLIPS, P. C. B. Time series with a unit root. **Econometrica**, Chicago, v.55, n.2, p.277-301, Mar. 1987.
- \_\_\_\_\_; OULIARIS, S. Asymptotic properties of residual based tests for cointegration. **Econometrica**, Chicago, v.58, n.1, p.165-193, Jan. 1990.
- \_\_\_\_\_; PERRON, P. Testing for a unit root in time series regression. **Biometrika**, Great Britain, v.75, n.2, p.335-46. 1988.
- PINDYCK, R. S.; RUBINFELD, D. L. **Econometric models and economic forecasts**. 4rd ed. Singapore: McGraw-Hill, 1998. 634p.
- PINO, F. A.; ROCHA, M. B. Transmissão de preços de soja no Brasil. **Revista de Economia e Sociologia Rural**, Brasília, v.32, n.4, p.345-61, out./dez. 1994.
- \_\_\_\_\_; NOGUEIRA JÚNIOR, S.; TOLOI, C. M. C. Relações dinâmicas entre preços da soja brasileira. **Pesquisa Agropecuária Brasileira**, Brasília, v.18, n.11, p.1163-1173, nov. 1983.
- SAS INSTITUTE. **SAS/ETS software: changes and enhancements, release 6.12**. Cary, NC, 1996. 112p.
- \_\_\_\_\_. **SAS/ETS software: changes and enhancements, release 6.11**. Cary, NC, 1995. 114p.

### TESTES DE CO-INTEGRAÇÃO UTILIZANDO O SAS®: teoria e aplicação

**RESUMO:** Um procedimento de grande importância para aqueles que trabalham com séries econômicas, para obter relações estruturais entre variáveis ou modelos de previsão, diz respeito aos testes de co-integração. Ao se realizar uma regressão entre duas ou mais variáveis econômicas, esta regressão pode ser espúria, ou seja, sem significado, apesar de os coeficientes serem significativos em termos de seus respectivos testes  $t$  e do alto valor do coeficiente de determinação ( $R^2$ ). Sendo assim, antes de se estimar uma regressão é necessário executar testes de raiz unitária para se determinar a ordem de integração das variáveis. Quando duas variáveis são integradas de mesma ordem, conforme os resultados obtidos nos testes de raiz unitária, e os seus resíduos são estacionários, isso implica que essas variáveis são co-integradas, ou seja, apesar de individualmente serem integradas (não estacionárias) a combinação linear entre elas as tornam estacionárias e, portanto, existe um relacionamento de longo prazo entre elas e, conseqüentemente, os testes empregados pela estatística tradicional deixam de ser válidos, pois produzem resultados viesados. Este trabalho objetiva mostrar ao usuário como fazer os testes de co-integração e posteriormente ter condições de analisar os resultados gerados pelo Statistical Analysis Software (SAS). Para isso foram apresentados e discutidos aspectos relacionados aos testes de co-integração, conforme as metodologias desenvolvidas por ENGLE e GRANGER (1987) e PHILLIPS e OULIARIS (1990).

**Palavras-chave:** co-integração, teste Engle-Granger, estacionariedade, raiz unitária.

### CO-INTEGRATION TESTS USING SAS®: theory and application

**ABSTRACT:** In order to obtain structural relationships or forecasting models, an important procedure for researchers working with economic time series concerns the co-integration tests. A regression between two or more economic variables may be spurious, i.e., meaningless, even though the coefficients are meaningful in terms of their respective  $t$  tests and the regression coefficient ( $R^2$ ). Thus, before estimating a regression it is necessary to apply unit root tests to determine the integration order of the variables. When two or more variables are integrated in the same order, according to the results obtained from the unit root test, and their residuals are stationary, then they are said to be co-integrated. This means that, although individually integrated (non-stationary), the linear combination among them makes them stationary. Therefore there is a long-term relationship among them and consequently, the tests used by the traditional statistics are no longer valid, as they produce biased results. This paper shows users how to do the co-integration tests and analyze the results using the Statistical Analysis Software (SAS). To this end, several aspects of the ENGLE-GRANGER (1987) and PHILLIPS-OULIARIS (1990) co-integration tests have been shown and discussed.

**Key-words:** co-integration, engle-granger test, stationarity, unit root.

---

Recebido em 27/11/2000. Liberado para publicação em 01/12/2000.