

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS

CENTRO DE ECONOMIA E ADMINISTRAÇÃO

FACULDADE DE CIÊNCIAS ECONÔMICAS

NÍCHOLAS RODRIGUES NEVES LE PETIT RAMOS

TRABALHO FINAL DA DISCIPLINA: ECONOMETRIA II

CAMPINAS

2019

NICHOLAS RODRIGUES NEVES LE PETIT RAMOS

TRABALHO FINAL DA DISCIPLINA: ECONOMETRIA II

Trabalho apresentado ao Prof. Dr. Paulo Ricardo
da Silva Oliveira como exigência parcial para a
aprovação na disciplina Econometria II.

PUC-CAMPINAS

2019

1. INTRODUÇÃO – EXERCÍCIO UM

Foram selecionadas duas séries temporais econômicas para que fosse realizada a atividade. A primeira série temporal é a evolução do câmbio R\$/US\$ - venda - comercial, indexado mensalmente, obtido através de publicação do BACEN retirada do IPEADATA. A segunda série temporal é a evolução do preço da saca de 60kg de milho, conforme publicado e retirado do CEPEA.

O banco de dados referente ao preço do milho estava com o formato de datas em “dias”, começando em 02/08/2004; e tinha 3818 observações. Através dos pacotes “anytime” e “lubridate”, as datas foram indexadas por mês, e foi feita a média das observações para cada mês utilizando os pacotes “zoo” e “tidyverse”. O banco de dados final do milho contém 184 observações. O banco de dados do câmbio foi selecionado para incluir apenas o mesmo número e período temporal das observações encontradas no banco de dados do preço do milho.

O modelo elaborado coloca o preço do milho como função do câmbio, uma vez que a exportação de milho no Brasil é uma atividade de magnitude significativa para a economia e, portanto, espera-se que alterações no câmbio devam exercer influência sobre o preço do commodity, além de sua produção e exportação nacional.

2. ESTIMAÇÃO DO MODELO – EXERCÍCIO DOIS

A Figura 1 apresenta a saída (*output*) do R para o modelo estimado inicialmente:

FIGURA 1 - MODELO RLS DO PREÇO DO MILHO EM FUNÇÃO DO CÂMBIO

```
Time series regression with "ts" data:
Start = 2004(8), End = 2019(11)

Call:
dynlm(formula = pmilho ~ cambio, data = db1.ts)

Residuals:
    Min       1Q   Median       3Q      Max
-12.6818  -4.5476  -0.6069   5.1091  16.8971

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  11.2707    1.6410    6.868 9.95e-11
cambio        6.5864    0.6221   10.587 < 2e-16

(Intercept) ***
cambio      ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.379 on 182 degrees of freedom
Multiple R-squared:  0.3811,    Adjusted R-squared:  0.3777
F-statistic: 112.1 on 1 and 182 DF,  p-value: < 2.2e-16
```

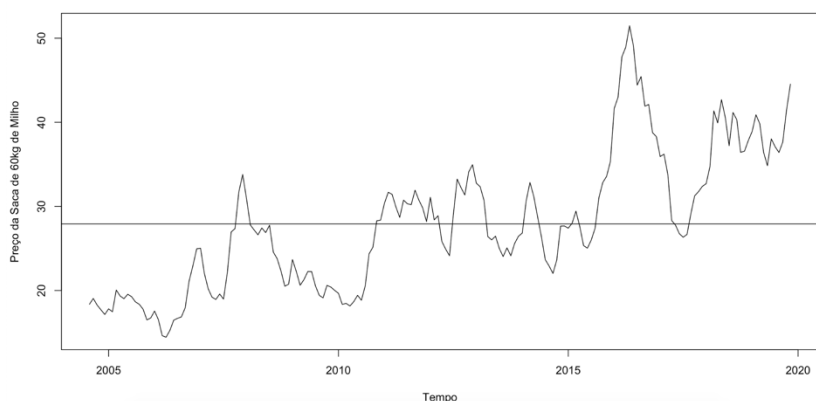
A priori, a variável câmbio é significativa estatisticamente para explicar o preço do milho, uma vez que apresentou um valor de estatística “t” de 10,587, com um p-valor associado menor do que 2^{-16} . O erro padrão do coeficiente da variável “câmbio” representa 9,45% apenas do próprio coeficiente β_1 .

Quanto ao R^2 e o R_{adj}^2 , temos que o modelo não obteve um desempenho notável, uma vez que seus valores são de respectivamente apenas 38,11% e 37,77%. Ou seja, este indicador nos mune do conhecimento de quantas observações presentes no banco de dados utilizado podem ser efetivamente explicadas pela reta de regressão estimada. Apesar deste fato, ainda temos que a reta da regressão é capaz de explicar mais do que 1/3 do banco de dados, agregando coerência para o modelo, mesmo que eventuais previsões que possam ser elaboradas a partir do mesmo não tendem a serem confiáveis ou precisas.

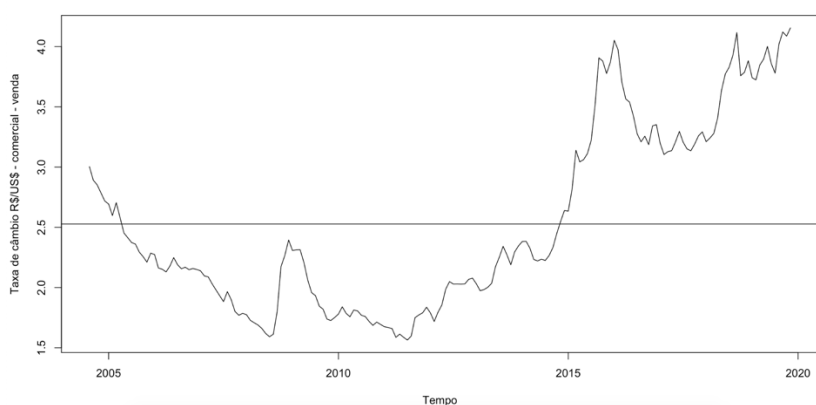
O teste-f resultante da regressão nos indica que as variáveis são significantes como um todo, com um p-valor de $2,2^{-16}$. A análise sumarizada indica que o câmbio tem de fato influência significativa sobre o preço do milho produzido no Brasil, entretanto previsões geradas a partir deste modelo não devem ser precisas, apesar de indicarem uma tendência correta.

3. DEFINIÇÃO DE ESTACIONARIEDADE E INVESTIGAÇÃO GRÁFICA DAS SÉRIES TEMPORAIS – EXERCÍCIO TRÊS

Uma série temporal é não-estacionária se: sua distribuição, em particular sua média, variância ou covariância (no quesito temporal) mudarem com o tempo. Séries não-estacionárias não podem ser usadas em modelos de regressão pois elas quebram as premissas do MQO e podem criar uma regressão espúria – um relacionamento falso entre as variáveis por conta de, por exemplo, uma tendência em comum presente em duas variáveis que não estão relacionadas. Duas ou mais séries não-estacionárias ainda podem estar presentes em um modelo de regressão, mas apenas se elas forem cointegradas, ou seja, que há um relacionamento estacionário de algum tipo. (COLONESCU, 2016).

GRÁFICO 1 - EVOLUÇÃO DO PREÇO DA SACA DE MILHO DE 60KG

A série temporal referente ao preço da saca de milho de 60kg apresenta tanto uma tendência, quanto sintomas de não-estacionariedade e intercepto. Tal conclusão pode ser esperada uma vez que podemos observar que o comportamento evolutivo do preço tende a aumentar de patamar conforme o tempo passa, sendo coerente com a influência de conceitos esperados como a inflação. Ou seja, temos que sua média, variância e covariância apresentam sinais de oscilação ao longo do tempo, sendo esta oscilação a definição de não-estacionariedade e uma violação de uma das premissas do MQO.

GRÁFICO 2 - EVOLUÇÃO DO CÂMBIO R\$/US\$ - COMERCIAL - VENDA

Com o Gráfico 2, temos também indicações para esperar que a série temporal referente a taxa de câmbio R\$/US\$ seja não-estacionária e possua intercepto. Não pode ser observado comportamentos cíclicos atemporais e, apesar da taxa de câmbio oscilar diariamente, quando

observada por uma ótica de média temporal mensal, apresenta resultados indicativos de uma evolução não-estacionária, sendo que esta aparenta ter médias, variâncias ou covariâncias diferentes conforme sua evolução temporal.

4. DEFINIÇÃO DO TESTE DE RAÍZ UNITÁRIA E APLICAÇÃO DO TESTE DICKY-FULLER

O teste de raiz unitária Dickey-Fuller para estacionariedade é baseado em um processo AR(1). Neste processo, se o resultado do beta associado a variável autoregressiva for igual a 1, significa que o valor da variável em um período t é igual ao seu próprio valor no período $t-1$ agregado a um choque aleatório. Desta forma, teríamos um passeio aleatório e uma série que não atende aos requisitos do método de estimadores MQO, o que pode, por sua vez, levar a uma correlação espúria. De acordo com o teste, uma série é não-estacionária quando $\rho = 1$ (no caso do teste de raiz unitária, o coeficiente β de AR(1) é igual ao ρ). As hipóteses do teste são: $H_0: \rho = 1$; $H_A: \rho < 1$. Rejeitar a hipótese nula implica que a série temporal é estacionária. (COLONESCU, 2016).

Foram elaborados dois correlogramas, um para cada série temporal analisada. O Gráfico 3 apresenta o correlograma referente a série do preço do milho e o Gráfico 4 apresenta o correlograma referente a série do câmbio:

GRÁFICO 3 - CORRELOGRAMA MILHO

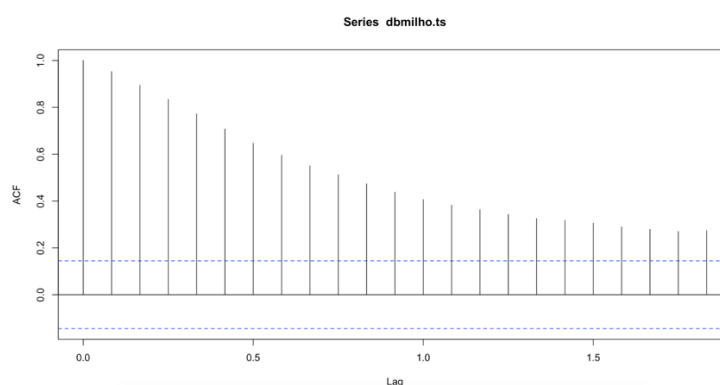
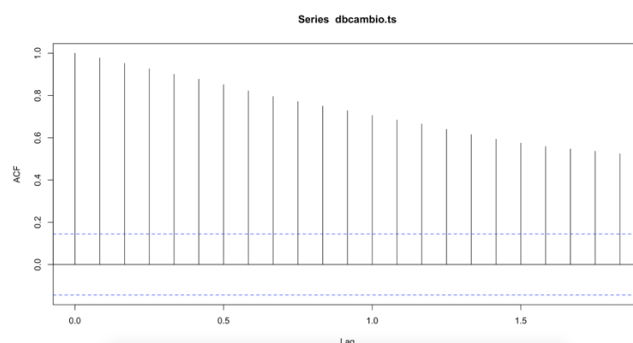


GRÁFICO 4 - CORRELOGRAMA CÂMBIO



Com base nos dois correlogramas elaborados inicialmente não foi possível apontar uma ordem máxima ideal para o teste de Dickey-Fuller aplicado em cada uma das séries. Contudo, foi plotado mais dois correlogramas, um para cada variável, com um “*lag.max*” maior que alcançasse o período em que acabasse a correlação. Com base em testes, foi determinado o “*lag.max*” para a série do preço do milho e do câmbio de, respectivamente, 60 e 50 *lags*. Os Gráficos 6 e 7 a seguir apresentam os correlogramas obtidos com esta modificação:

GRÁFICO 5 - CORRELOGRAMA MILHO (LAG.MAX=60)

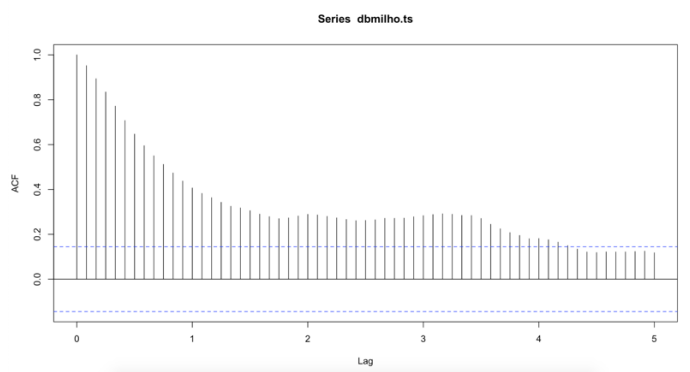
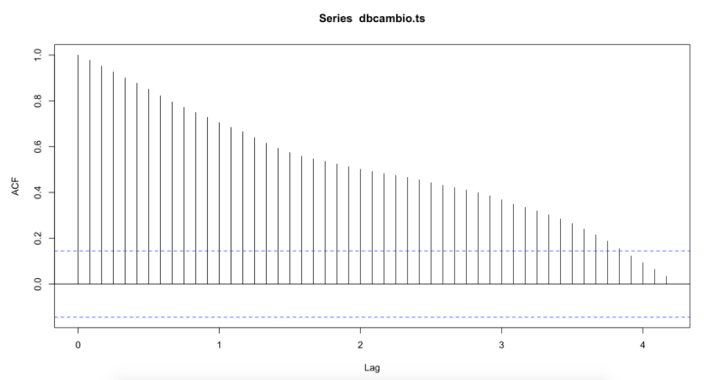


GRÁFICO 6 - CORRELOGRAMA CÂMBIO (LAG.MAX=50)



Com base nos Gráficos 5 e 6, podemos esperar que as melhores ordens para a elaboração do Teste de Dickey-Fuller para as séries do milho e do câmbio sejam de, respectivamente, 50 e 46 *lags* – ou seja, o parâmetro “k” do teste ADF incluso no pacote “t-series” sejam os citados na última frase para cada uma das variáveis. Seguindo para a elaboração do teste de Dickey-Fuller para cada uma das séries, temos apresentadas as Figuras 2 e 3 a seguir:

FIGURA 2 - TESTE DE DICKEY-FULLER PARA A SÉRIE DO PREÇO DO MILHO

```
> adf.test(dbmilho.ts, k=50)

Augmented Dickey-Fuller Test

data: dbmilho.ts
Dickey-Fuller = -1.9549, Lag order = 50, p-value = 0.595
alternative hypothesis: stationary
```

Temos que o Teste de Dickey-Fuller nos indica que a chance de rejeitarmos a hipótese nula de não estacionariedade é muito elevada, sendo de 59,5%; portanto não há evidências estatísticas para rejeitarmos H_0 e assumirmos a estacionariedade da série. O diagnóstico é de não-estacionariedade.

FIGURA 3 - TESTE DE DICKEY-FULLER PARA A SÉRIE DO CÂMBIO

```
> adf.test dbcambio.ts, k=46)

Augmented Dickey-Fuller Test

data: dbcambio.ts
Dickey-Fuller = -1.1027, Lag order = 46, p-value = 0.9195
alternative hypothesis: stationary
```

Continuando a linha de raciocínio do diagnóstico da série do preço do milho, para a série temporal do câmbio o resultado do Teste de Dickey-Fuller indica uma chance de 91,95% de Erro do Tipo I ao rejeitar H_0 . Com isso, temos também na série do câmbio um diagnóstico de não-estacionariedade.

5. DISCUSSÃO DA CORRELAÇÃO ESPÚRIA NO CONTEXTO DAS SÉRIES NÃO-ESTACIONÁRIAS – EXERCÍCIO CINCO

A não-estacionariedade pode levar a uma regressão espúria. A última se trata de um relacionamento entre variáveis que na verdade não estão relacionadas. (COLONESCU, 2016). Uma regressão espúria, conforme descoberto por Yule, retrata uma relação entre as variáveis

que é por definição sem sentido, ou seja, neste fenômeno uma série cuja significância estatística dos betas é ausente, mas que na elaboração do modelo apresenta uma significância falsa, que pode levar a diagnósticos incorretos e inferências estatísticas comprometidas.

6. IDÉIAS PARA A SOLUÇÃO DO PROBLEMA – EXERCÍCIO SEIS

Uma possibilidade para a correção do problema e a obtenção de uma série estacionária que, neste quesito, atenda às premissas do MQO, é a utilização da função “*ndiffs*” do pacote “*forecast*”, que retorna a ordem de integração de uma determinada série. O conceito de Integração representa quantas vezes precisamos diferenciar uma série até que ela se torne estacionária. A série é integrada de ordem 0, ou seja, $I(0)$, quando esta já é estacionária inicialmente. Uma série $I(1)$, ou seja, integrada de ordem 1, é não estacionária em certos níveis, mas as diferenças de primeira ordem são estacionárias. No mesmo raciocínio, uma série $I(2)$ é também não-estacionária em níveis, mas se torna estacionária nas diferenças de segunda ordem.

Portanto foi empregada a função “*ndiffs*” para que fossem obtidas as ordens de integração das séries do preço do milho e do câmbio. A seguir está a Figura 4, que apresenta, respectivamente, o resultado obtido para a série do preço do milho (dbmilho.ts) e do câmbio (dbcambio.ts):

FIGURA 4 - ORDEM DE INTEGRAÇÃO DAS SÉRIES DO PREÇO DO MILHO E DO CÂMBIO

```
> ndiffs(dbmilho.ts)
[1] 1
> ndiffs dbcambio.ts)
[1] 2
```

Com isso, foi aplicada a função “*diff*” em cada uma das séries, empregando diferenças das ordens iguais às obtidas na análise da Figura 4, ou seja, diferenças de primeira ordem para o banco de dados do preço do milho, e de segunda ordem para o banco de dados do câmbio. Segue abaixo os Gráficos 7 e 8, que apresentam o antes e o depois dos seus respectivos bancos de dados, separados por uma linha horizontal. A área superior é referente ao banco de dados original, e a área inferior é referente ao banco de dados obtido a partir da aplicação da função “*diff*”.

GRÁFICO 7 - COMPARAÇÃO ANTES E DEPOIS DA CORREÇÃO - BANCO DE DADOS DO PREÇO DO MILHO

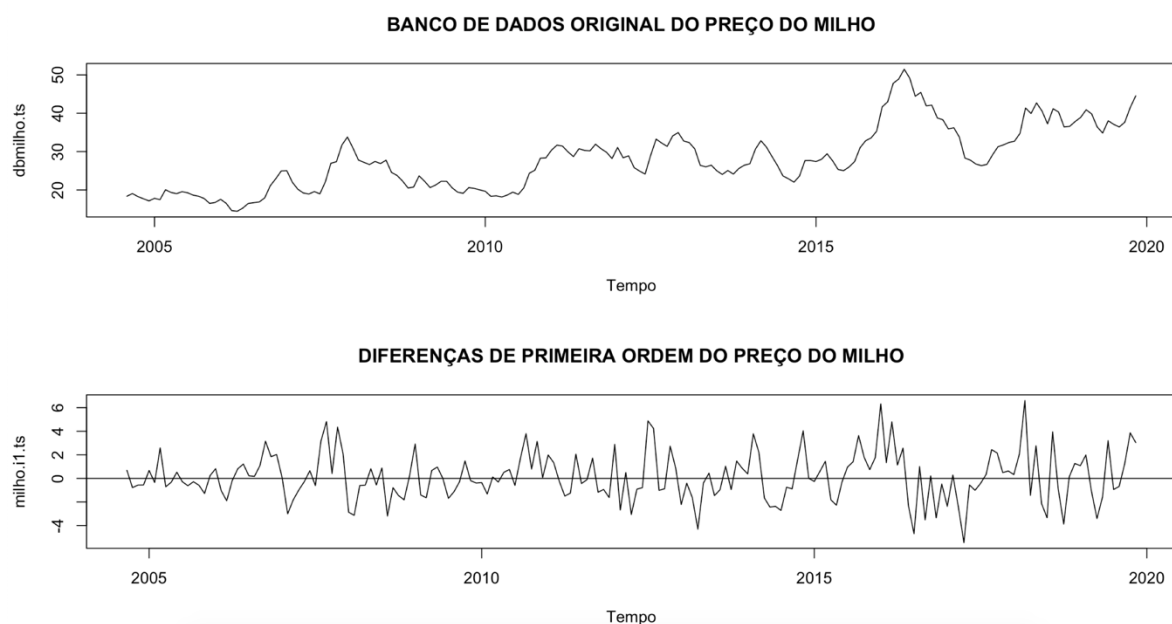
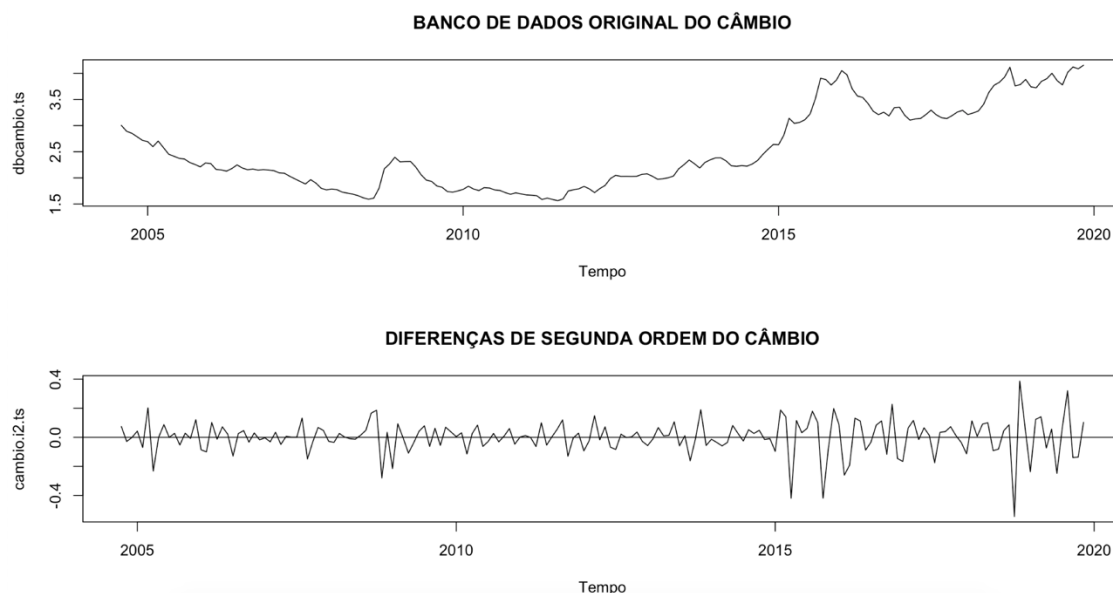


GRÁFICO 8 - COMPARAÇÃO ANTES E DEPOIS DA CORREÇÃO - BANCO DE DADOS CÂMBIO



Graficamente, as correções parecem ter tornado as séries estacionárias, uma vez que podemos observar sua proximidade a uma média 0, conforme a linha horizontal no nível 0 proporciona. Entretanto, abaixo temos as o Gráfico 9 e as Figuras 5 e 6, referentes aos

correlogramas e Testes de Dickey-Fuller respectivos às séries do preço do milho e do câmbio, para que seja verificada a estacionariedade estatisticamente:

GRÁFICO 9 - CORRELOGRAMA DAS SÉRIES INTEGRADAS

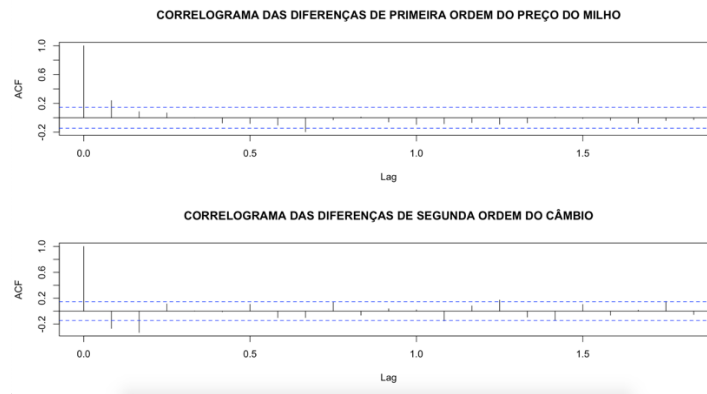


FIGURA 5 - TESTE DE DICKEY-FULLER PARA A SÉRIE I(1) DO PREÇO DO MILHO

```
> adf.test(milho.i1.ts, k=1)

Augmented Dickey-Fuller Test

data: milho.i1.ts
Dickey-Fuller = -7.8035, Lag order = 1, p-value = 0.01
alternative hypothesis: stationary
```

FIGURA 6 - TESTE DE DICKEY-FULLER PARA A SÉRIE I(2) DO CÂMBIO

```
> adf.test(cambio.i2.ts, k=2)

Augmented Dickey-Fuller Test

data: cambio.i2.ts
Dickey-Fuller = -12.352, Lag order = 2, p-value = 0.01
alternative hypothesis: stationary
```

Com isso, rejeitamos a hipótese nula da não-estacionariedade e assumimos a estacionariedade em ambas as séries obtidas. A seguir a Figura 7 apresenta o resultado do modelo calculado com estes novos bancos de dados estacionários:

FIGURA 7 - MODELO RLS COM AS DIFERENÇAS DE RESPECTIVAS ORDENS DO PREÇO DO MILHO - I(1) - E DO CÂMBIO - I(2).

```
> summary(mod2)

Time series regression with "ts" data:
Start = 2004(10), End = 2019(11)

Call:
dynlm(formula = milho.i1.ts ~ cambio.i2.ts)

Residuals:
    Min       1Q   Median       3Q      Max
-5.5767 -1.1475 -0.3063  1.1314  6.4602

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.1393     0.1530   0.911   0.364
cambio.i2.ts    0.7540     1.3237   0.570   0.570

Residual standard error: 2.064 on 180 degrees of freedom
Multiple R-squared:  0.0018,    Adjusted R-squared:  -0.003746
F-statistic: 0.3245 on 1 and 180 DF,  p-value: 0.5696
```

Com o modelo acima, chegamos à conclusão de que se tratava de uma regressão espúria, onde a relação indicada estatisticamente pelo primeiro modelo não existia, as inferências estatísticas estavam distorcidas, uma vez que quebravam as premissas do MQO. Com as correções aplicadas observamos que a significância estatística, os R-quadrados, e a F-estatística diminuem e quase “desaparecem”. Um fato que vale notar é de que apesar da função “i” ter retornado uma ordem de integração igual a 2 para o banco de dados do câmbio, por experimentação foi observado que uma diferença de primeira ordem já rejeitava a H0 de não estacionariedade com probabilidade associada ao Erro do Tipo I de 0,01%, e retornava um modelo que, apesar de ainda estatisticamente não-significante, tinha melhores p-valores e estatísticas associadas, este desenvolvimento pode ser observado no Anexo A.

7. CONCEITO DE COINTEGRAÇÃO E DIAGNÓSTICO DE COINTEGRAÇÃO PARA AS SÉRIES DO PREÇO DO MILHO E CÂMBIO

Duas séries são co-integradas quando suas tendências são relativamente similares. Este conceito vago pode ser estimado de forma precisa utilizando um teste de cointegração. No teste de cointegração, verifica-se se os resíduos da regressão de uma das séries sobre a outra são

estacionários. Se os resíduos forem estacionários, temos que as séries são cointegradas. (COLONESCU, 2016). Após transformar os bancos de dados em uma matriz, onde a primeira coluna é a da variável dependente, foi aplicado o Teste de Phillips-Ouliaris para Cointegração. A Figura 8 apresenta os resultados do teste:

FIGURA 8 - TESTE DE PHILLIPS-OULIARIS PARA COINTEGRAÇÃO

```
> po.test(dbx) # H0: Nao haá cointegracao entre as series.
```

Phillips-Ouliaris Cointegration Test

data: dbx

Phillips-Ouliaris demeaned = -14.218, Truncation lag parameter = 1, p-value = 0.15

De acordo com o teste, temos que não há evidências estatísticas para rejeitarmos a H0 de não-cointegração. Ou seja, assumimos que as duas séries não estão cointegradas.

ANEXO A – TESTE DE DICKEY-FULLER E MODELO ESTIMADOS UTILIZANDO BANCO DE DADOS I(1) DO CÂMBIO – CONTRÁRIO A INDICAÇÃO DA FUNÇÃO “NDIFFS”

```
> adf.test(cambio.i1.ts, k=1)
```

Augmented Dickey-Fuller Test

```
data: cambio.i1.ts
```

```
Dickey-Fuller = -9.2732, Lag order = 1, p-value = 0.01  
alternative hypothesis: stationary
```

```
> mod3 <- dynlm(milho.i1.ts ~ cambio.i1.ts)  
> summary(mod3)
```

```
Time series regression with "ts" data:  
Start = 2004(9), End = 2019(11)
```

```
Call:  
dynlm(formula = milho.i1.ts ~ cambio.i1.ts)
```

```
Residuals:
```

Min	1Q	Median	3Q	Max
-5.5959	-1.2963	-0.2683	1.0828	6.3995

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.1306	0.1519	0.860	0.391
cambio.i1.ts	1.9765	1.5491	1.276	0.204

```
Residual standard error: 2.051 on 181 degrees of freedom  
Multiple R-squared: 0.008913, Adjusted R-squared: 0.003438  
F-statistic: 1.628 on 1 and 181 DF, p-value: 0.2036
```

REFERÊNCIAS

8. Bibliografia

COLONESCU, C. **Principle of Econometrics with R**. [S.l.]: [s.n.], 2016.

GUJARATI, Damodar N.; PORTER, Dawn C. **Econometria Básica**. 5aed. Amgh Editora, 2011.