Pset4 - Econometrics II

Estevão Cardoso

March 2025

1 Question 1 (Testing for Cointegration when the Cointegrating Vector is Known)

Item 1. In this question we answered most in the code file `pset4_code.R`, which is attached in the zip file. In the first item we were asked to just the period of the serie from January 1995 to December 2019, and we did it as required.

```
#adjusting column names
colnames(brazil_data) <- c("date", "exchange_rate", "ipca_index", "drop")
brazil_data <- brazil_data[, 1:3]
colnames(usa_data) <- c("date", "us_cpi_index")

#setting date columns to the same format
brazil_data <- brazil_data %>%
  mutate(
    date = as.yearmon(as.character(date), "%Y.%m"),
    date = as.Date(date)
  )

usa_data <- usa_data %>%
  mutate(
    date = as.Date(date)
  )

#defining analysis period
start_date <- as.Date("1995-01-01")
end_date <- as.Date("2019-12-31")

brazil_data_flt <- brazil_data %>%
  filter(date >= start_date & date <= end_date)

usa_data_flt <- usa_data %>%
```

```
  filter(date >= start_date & date <= end_date)
```

Item 2. In the second item we were asked for each variable $X_{k,t}$, where $k \in \{1, 2, 3\}$, define $Y_{k,t} := 100 \cdot [\log(X_{k,t}) - \log(X_{k,\text{January 1995}})]$, and we did it.

```
#calculating the log-variation since Jan/1995

#brazil
exchange_rate_base <- log(brazil_data_flt$exchange_rate[1])
ipca_index_base <- log(brazil_data_flt$ipca_index[1])

brazil_data_flt <- brazil_data_flt %>%
  mutate(
    log_change_exchange = 100 * (log(exchange_rate) - exchange_rate_base),
    log_change_ipca = 100 * (log(ipca_index) - ipca_index_base)
  )



#EUA
us_cpi_index_base <- log(usa_data_flt$us_cpi_index[1])

usa_data_flt <- usa_data_flt %>%
  mutate(
    log_change_us_cpi = 100 * (log(us_cpi_index) - us_cpi_index_base)
  )
```

Item 3. In this item, we need to justify the value of the cointegration vector based on the Purchasing Power Parity (PPP). We are given three variables: $Y_{1,t}$, the log of the accumulated variation of the exchange rate (BRL/USD); $Y_{2,t}$, the log of the accumulated variation of the Brazilian inflation index (IPCA); and $Y_{3,t}$, the accumulated variation of the U.S. inflation index (CPI). We want to understand whether there exists a vector $a = (a_1, a_2, a_3)$ such that,

$$Z_t = a_1 Y_{1,t} + a_2 Y_{2,t} + a_3 Y_{3,t}$$

is a stationary series.
According to PPP, in the long run, the nominal exchange rate S_t between two countries—disregarding frictions such as transportation costs, pandemics, and conflicts—should be explained by the relative price levels in these countries, such that:

$$S_t = \frac{P_t^{\text{BR}}}{P_t^{\text{US}}} \Rightarrow S_t \times \frac{P_t^{\text{US}}}{P_t^{\text{BR}}} = 1$$

where $S_t = Y_{1,t}$ is the nominal exchange rate (BRL/USD), $Y_{2,t} = P_t^{\text{BR}}$ is the price level in Brazil (IPCA), and $Y_{3,t} = P_t^{\text{US}}$ is the price level in the United States (CPI). With a few manipulations, we obtain

$$\text{Real EXR} = \log(S_t) + \log(P_t^{\text{BR}}) - \log(P_t^{\text{US}}) = 0$$

and rearranging it, we get

$$\log(P_t^{\text{US}}) = \log(P_t^{\text{BR}}) - \log(S_t)$$

From the manipulation above, it follows that the following linear combination of the logs of the three variables should be stationary under PPP

$$Z_t = \log(P_t^{\text{BR}}) - \log(S_t) - \log(P_t^{\text{US}})$$

And, as stated in **Item 2**, when we apply this result to the transformed data, where each variable $Y_{k,t}$ is defined as the cumulative log-change since January 1995, we get

$$Y_{k,t} := 100 \cdot [\log(X_{k,t}) - \log(X_{k,\text{January 1995}})]$$

Since in the real world we face many frictions, we argue that the following form provides a more flexible representation of PPP that accounts for estimation errors

$$Z_t = Y_{1,t} - Y_{2,t} - Y_{3,t}$$

Thus, we find the desired economically motivated cointegration vector

$$a = (1, -1, -1)$$

$$Z_t = a' Y_t$$

This vector reflects the PPP condition and can be tested empirically. If the linear combination Z_t is stationary (i.e., $Z_t \sim I(0)$), then the original series are said to be cointegrated with this known vector. Moreover, note that **Real EXR** is adimensional, so measurement units are a non-issue.

Item 4. We answered this item in the R, so bellow we show the code we made,

```
#we need to combine both dataframes to get the cointegrated variable
zt_data <- brazil_data_flt %>%
  dplyr::select(date, log_change_exchange, log_change_ipca) %>%
  left_join(
    usa_data_flt %>% dplyr::select(date, log_change_us_cpi),
```

```
    by = "date"
  )

zt_data <- zt_data %>%
  mutate(
    z_t =  log_change_ipca - log_change_exchange - log_change_us_cpi
  )
zt_ts <- ts(zt_data$z_t, start = c(1995, 1), frequency = 12)
```

Item 5.

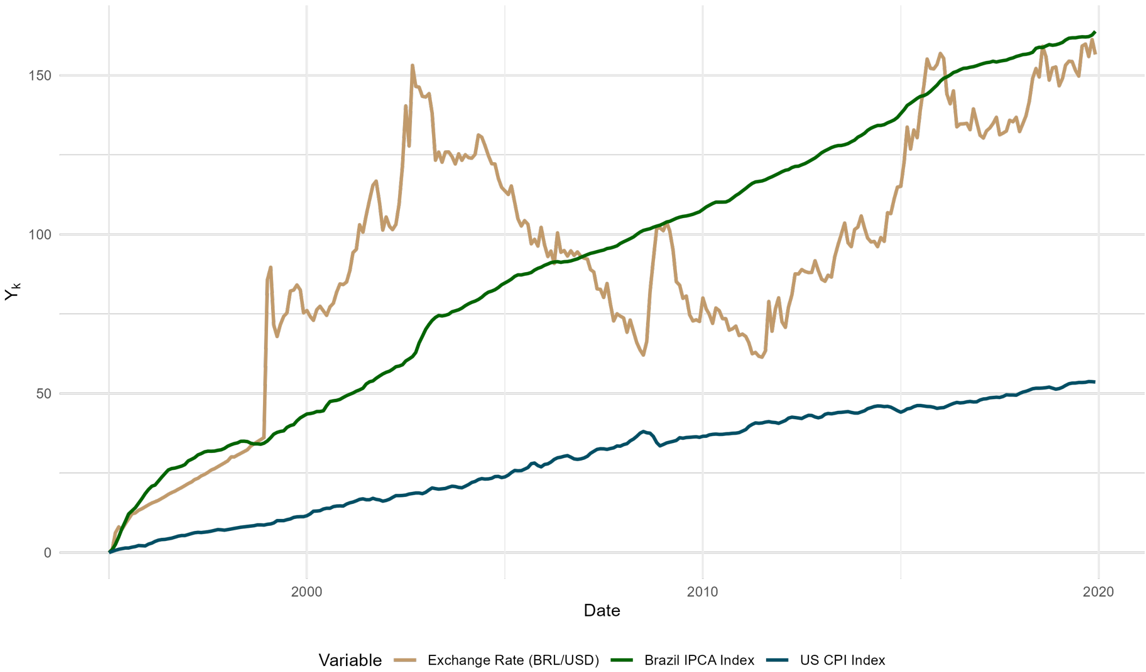


Figure 1: Transformed series: Exchange rate, Brazil IPCA, and US CPI (1995–2019).

Item 6. Our decisions regarding the specification of the ADF Test were based on the visual elements observed in Figure 1 and on the economic characteristics of this type of financial series. First, we see that the exchange rate series does not seem to have a clear trend, but it exhibits very wide movements away from its mean. Therefore, it is reasonable to include a drift to correctly specify this pattern. The series for the Brazilian IPCA and the CPI have a clear positive trend, which is consistent with processes that persist over time, typical of inflationary movements. Thus, we included both a drift and a linear trend in the test, which adjusts the null hypothesis and critical values to account for the presence of deterministic components. We follow the convention of using an intercept and drift when dealing with financial time series transformed by log differences.

To define the number of lags, we implemented the Bayesian Information Criterion (BIC), which helps us understand the model fit and the presence of autocorrelation in the residuals. Moreover, as in the standard ADF Test, we test the null hypothesis $H_0 : \rho = 1$ for the presence of a unit root. In Table 1, we present the ADF test results for each series. The statistics are referred to as τ, which corresponds to the t-test for the unit root hypothesis, and ϕ, referring to the F-tests on deterministic components (intercept and/or trend). In green, we highlight the statistics for which we reject the null hypothesis of the presence of a unit root; in red, the statistics for which we fail to reject it. Hence, we conclude that we fail to reject the null hypothesis of the presence of a unit root in the Exchange Rate and CPI series, which strongly indicates they are integrated of order one, non-stationary, and likely have a unit root under the specifications we implemented. Conversely, we observe some evidence of stationarity in the levels of the IPCA series, which is somewhat surprising, since it is not expected for a price index to behave that way, especially in Brazil, where things are not steady, particularly in the long run.

Table 1: ADF Tests Results

(1) Serie	(2) Test Statistic	(3) Statistic	(4) 1%	(5) 5%	(6) 10%
Exchange Rate	τ_2	-1.74	-3.44	-2.87	-2.57
Exchange Rate	ϕ_1	2.89	6.47	4.61	3.79
Brazil IPCA	τ_3	-3.83	-3.98	-3.42	-3.13
Brazil IPCA	ϕ_2	15.90	6.15	4.71	4.05
Brazil IPCA	ϕ_3	8.80	8.34	6.30	5.36
US CPI	τ_3	-2.42	-3.98	-3.42	-3.13
US CPI	ϕ_2	9.64	6.15	4.71	4.05
US CPI	ϕ_3	3.32	8.34	6.30	5.36

Note: All the series were tested with a AR(2) and with one lag. In column (2), the colors indicate the results of the unit root test: green denotes the series for which we reject the null hypothesis of a unit root (i.e., the series is stationary), while red denotes the series for which we fail to reject the null (i.e., the series is non-stationary). Critical values and test statistics for ADF are refered as: τ refers to the t-test for the unit root hypothesis, and ϕ refers to test on deterministic components (intercept and/or trend). Subscripts denote the test specification: τ_2 and ϕ_1 correspond to models with intercept only; τ_3, ϕ_2, and ϕ_3 correspond to models with intercept and trend.

Item 7. Plot the graph for Z_t

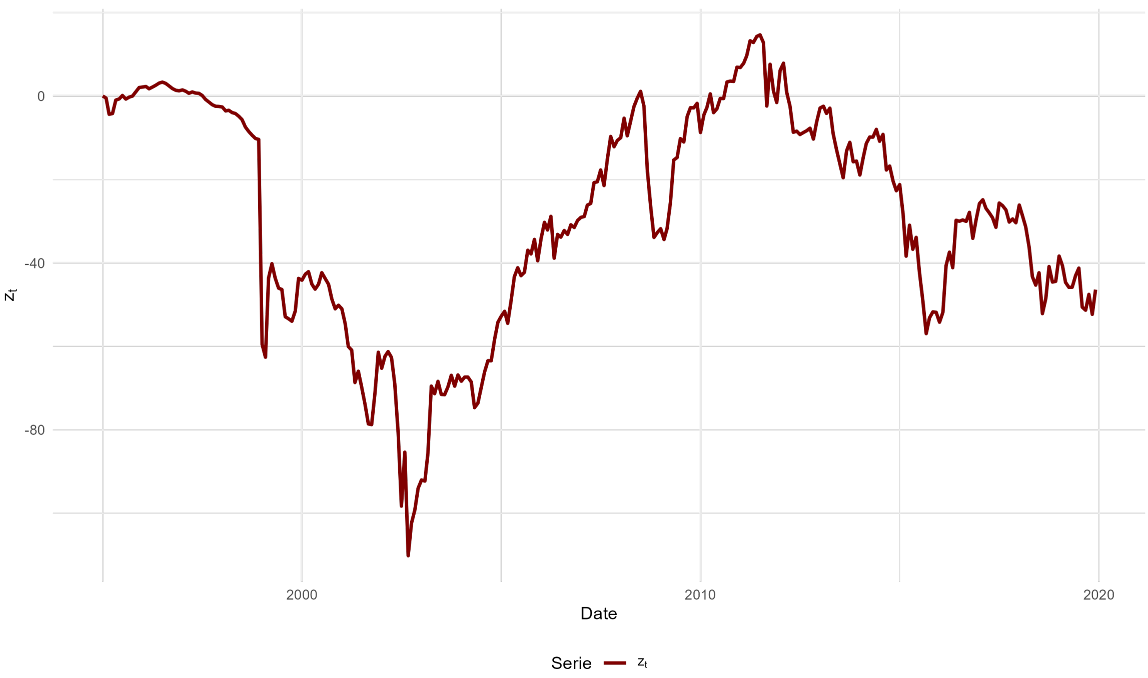


Figure 2: Residual series z_t

Item 8. We want to test whether the cointegrated vector $\{Z_t\}$ generated from the three series has a unit root or is $I(1)$. By a brief visual inspection, we can see that the series does not exhibit any permanent trend throughout the entire period of analysis. However, it does show a clear pattern of deviation from its mean, suggesting that we should include a drift in the ADF test.

Our specification of the ADF test closely follows what we did in **Item 6**. We test the null hypothesis $H_0 : \rho = 1$, that is, the series has a unit root $\Rightarrow Z_t \sim I(1)$. We selected the number of lags for the test based on the BIC criterion, for the same reason of achieving better control over model fitting and autocorrelation in the residuals. The number of lags (1) is automatically selected by the `urca` package in R.

We present the results of the test in Table 2, and we fail to reject the null hypothesis of a unit root at the 5% significance level. From class, we learned that a cointegration vector requires that a linear combination of $I(1)$ series must be $I(0)$. However, we find that Z_t indeed has a unit root, so we argue that there is no cointegration among the transformed series we used in this specific setup. This leads us to be less confident about the full empirical adequacy of the PPP hypothesis in this case.

Table 2: ADF Test Results for Cointegrated Series

(1)	(2)	(3)	(4)	(5)	(6)
Series	Test	Statistic	1%	5%	10%
Cointegrated Z_t	τ_2	-1.85	-3.44	-2.87	-2.57
Cointegrated Z_t	ϕ_1	1.84	6.47	4.61	3.79

Note: The Z_t serie was tested with a AR(2) and with one lag. In column (2), the colors indicate the results of the unit root test: green denotes the series for which we reject the null hypothesis of a unit root (i.e., the series is stationary), while red denotes the series for which we fail to reject the null (i.e., the series is non-stationary). Critical values and test statistics for ADF are refered as: τ refers to the t-test for the unit root hypothesis, and ϕ refers to test on deterministic components (intercept and/or trend). Subscripts denote the test specification: τ_2 and ϕ_1 correspond to models with intercept.

Item 9. Based on all the tests and analysis we realized in the previous items, we do not find (enough) empirical support for the validity of the PPP in this context. We discussed that the PPP theory implies that a linear combination of the log of the exchange rate and the domestic and foreign price levels should result in a stationary series. That is, if PPP holds in the long run, then the constructed residual series Z_t should be stationary. However, our results from the ADF test presented in Table 2 indicate that we fail to reject the null hypothesis of a unit root in Z_t at the 5% significance level. Both the τ_2 and ϕ_1 test statistics fall short of their respective critical values, suggesting for us that Z_t is non-stationary and thus integrated of order one, $I(1)$.
Therefore, we argue that by this body of results, under the specific transformation of the variables and estimation strategy we adopted, the long-run relationship suggested by PPP does not hold. In other words, the deviation of the real exchange rate from its theoretical parity condition does not revert to a constant mean over time, contradicting the main implication of the PPP hypothesis that we established before. Moreover, there might be a several reasons for that, we could say something like, frictions such as transaction costs, capital controls , or trade barriers. Structural breaks in the data. Measurement issues or even the short sample window we choose, which may not capture enough of the long-run patterns of the series.

2 Question 2 (Testing for Cointegration when the Cointegrating Vector is Unknown)

From here, we used all the series in level to run the models and plot the graphs. But we tested all the combinations using the transformed series and the results don't change.

Item 1. We want to define the order of our variables to run the OLS regression and analyise if the residuals are strationary. We know that from previous items the PPP hyptothesises that the exchange rate can be defined as

$$S_t = \frac{P_t^{\text{BR}}}{P_t^{\text{US}}} \Rightarrow S_t = P_t^{\text{BR}} - P_t^{\text{US}}$$

And using this appropriate transformation, we can run a OLS regression using exchange rate S_t as dependent variable, and the domestig price index P_t^{BR} and the foreing price index P_t^{US} as explanatory variables. Thus, we run the following regression,

$$S_t = \alpha + \beta_1 \cdot P_t^{\text{BR}} + \beta_2 P_t^{\text{US}} + U_t$$

Item 2. Recall that here our assumptions are that the coeficients are cointegrated $I(1)$ and the residuals are $I(0)$. We see in Table 3 that the estimated coeficients $(\alpha, 0.0013, -0.0725)$ are somewhat different from the ones predicted by the PPP $[1, -1, 1]'$ for the order $(S_t, P_t^{\text{BR}}, P_t^{\text{US}})$.

Table 3: OLS regression for the unknonwn cointegrated vector

Dependent variable: *log*(Exchange Rate)		
	Coefficient	p-value
Intercept	4.8815**	0.0000
	(0.6248)	
log(IPCA)	0.0013***	0.0000
	(0.0001)	
log(CPI)	-0.0725***	0.0000
	(0.0111)	

Note: Standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

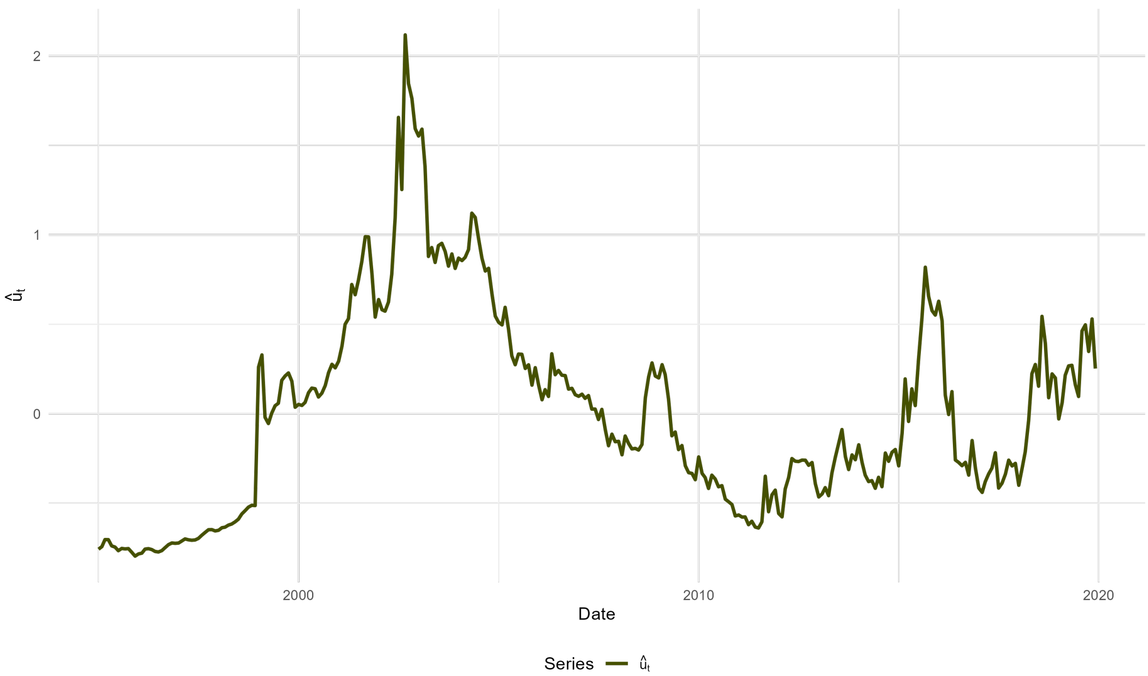


Figure 3: Residuals $\hat{U}_t$ from the OLS regressin of S_t

Item 3. With the Phillips and Ouliaris (1990) test, we assess whether the vector $\{\hat{U}_t\}$ is stationary. The null hypothesis is $H_0 : \{\hat{U}_t\} \sim I(1)$, indicating no cointegration. In R, we use the `ca.po()` function from the `urca` package. As before, rejecting H_0 implies that $\hat{U}_t$ is stationary and the variables are cointegrated. In Table 4 we see the results for the test we run. We fail to reject H_0 and we have no evidences that the three variables are cointegrated.

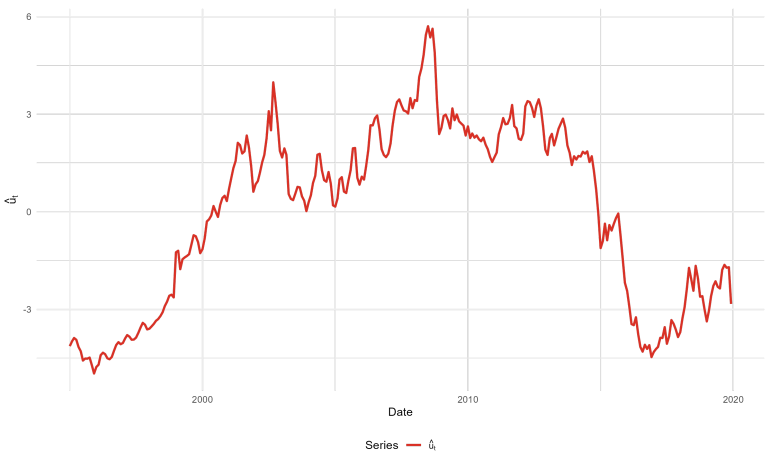
Table 4: Phillips and Ouliaris (1990) Test Results for $\{\hat{U}_t\}$

Series	Test Statistic	1% Crit. Value	5% Crit. Value	10% Crit. Value
$\{\hat{U}_t^{S_t}\}$	16.77	53.87	40.53	33.70

Item 4. Based on the results of the Phillips and Ouliaris (1990) test, all three variable orderings fail to reject the null hypothesis of no cointegration at the 5% significance level. Therefore, even though it may be reasonable to consider the ordering chosen and its empirical relevance, the evidence found for the period under analysis (including previous items) does not provide sufficient support to conclude that the variables are cointegrated. Furthermore, both the visual analysis in Figure 3 and the statistical tests suggest that the residuals do not appear to be stationary.

Item 5.

(a) Residuals from the regression: $P_t^{\text{BR}} = \alpha + \beta_1 S_t + \beta_2 P_t^{\text{US}} + U_t$



(b) Residuals from the regression: $P_t^{\text{US}} = \alpha + \beta_1 S_t + \beta_2 P_t^{\text{BR}} + U_t$

Figure 4: Time series of residuals $\hat{U}_t$ for the two alternative variable orderings.

Table 5: OLS regressions for alternative cointegrating relationships

Dependent variable: log(IPCA)		
	Coefficient	p-value
Intercept	-4584.1640***	0.0000
	(90.4357)	
Exchange Rate	212.8791***	0.0000
	(20.3014)	
CPI	80.8984***	0.0000
	(1.3009)	

Dependent variable: log(CPI)		
	Coefficient	p-value
Intercept	57.1278***	0.0000
	(0.4609)	
Exchange Rate	-1.7246***	0.0000
	(0.2648)	
IPCA	0.0115 ***	0.0000
	(0.0002)	

Note: Standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 6: Phillips and Ouliaris (1990) Test Results for $\{\hat{U}_t^{P_t^{\mathrm{BR}}}\}$ and $\{\hat{U}_t^{P_t^{\mathrm{US}}}\}$

Series	Test Statistic	1% Crit. Value	5% Crit. Value	10% Crit. Value
$\{\hat{U}_t^{P_t^{\mathrm{BR}}}\}$	1.75	53.87	40.53	33.70
$\{\hat{U}_t^{P_t^{\mathrm{US}}}\}$	5.29	53.87	40.53	33.70

We ran all possible alternative combinations to check the robustness of our previous findings. In Figure 4 we present the two plots of the residuals for each alternative reordering. Similarly, in Table 5 there are the actual estimation of each one of the reodering models. These results remain consistent with our initial conclusions. As shown in Table 6, the Phillips and Ouliaris test systematically failed to detect any stationarity in the residuals as we failed to reject the hypothesis H_0 of no cointegration. This provides strong evidence that the variables do not exhibit signs of cointegration, which once again does not support the Purchasing Power Parity (PPP) theory. Instead, it suggests that changes in the exchange rate are fully permanent within the system rather than stationary.

3 Question 3 (Testing for Cointegration: Johansen's Approach)

Item 1. We present the results of our estimation in Table 7. We set the number of lags to $K = 12$ to cover a full year, allowing us to fully capture the dynamic structure (including shocks and seasonal patterns) over a complete annual cycle for each variable. In the `ca.jo` function, we set the parameter `ecdet = "none"` to do not include (no intercept or trend) a deterministic component based on the PPP hypothesis, which assumes that the cointegration relation among the variables does not include a deterministic trend. This specification is consistent with the idea that any deviations from the long-run equilibrium are mean-reverting after shocks. Our results show a eigenvalue test for $r = 0$ ($H_0 : r = 0 \sim I(1)$) of 16.26 which is not above any of the critical values, therefore we fail to reject the null hypothesis of no cointegration between the series. So we lack statistical evidences to say that the three series arecointegrated.

Table 7: Johansen's Cointegration Test (Eigenvalue) – No Deterministic Terms, $K = 12$

Null Hypothesis	Statistic	10%	5%	1%
$r = 0$	16.26	18.90	21.07	25.75
$r \leq 1$	5.16	12.91	14.90	19.19
$r \leq 2$	2.57	6.50	8.18	11.65

Note: This table reports test statistics from Johansen's eigenvalue test for cointegration under the specification with no deterministic terms (i.e., `ecdet = "none"`) and using 12 lags (i.e., `K = 12` in the `ca.jo` function). Critical values at the 10%, 5%, and 1% significance levels are included.

Item 2. Overall, when we look at the body of evidences we have of no cointegration between the variables, it is dificult to say that the PPP hypothesis holds in this framework in the sample period that we analyse. It is even harder to say it holds in the long run. Moreover, this result is consistent with the findings from the Phillips and Ouliaris test in Question 2 and reinforces the idea that the real exchange rate may follow a non-stationary process. One could expect that at least reordering the variables in another set up in order to find a potential cointegration relation could make things work, but this is not the case here for any of our variables. And it goes strongly against the PPP, meaning that given that the cointegration fails, we maight see a patterns of extrem deviations from the mean or even unpredictable behavior in the exchange rate. This results can still be just a reflect of the relation of the two price index of Brazil and EUA, which might be sistematicaly affected by structural conditions within each country, making it harder to correctly especify the relation between the three variables. Therefore, summing up the evidences we found, we lack sufficient evidences to argue in favor of the validity of the PPP in our framework.