

# Assignment 4

## Question 1: Cointegration

The block of code situated below just shows how I imported the data into R and transformed both the r1 and r3 into time series data. However, given the fact that the data given is a stata file I am unsure if it was necessary for us to transform the data into a time series data to begin with. I tested for a unit root in the r1 and r3 series using the augmented Dickey Fuller test on both the untransformed, the original imported data, and the transformed data, that I transformed in the block of code underneath, to see if there was a difference in the results. The results were the same, leading me to conclude that the original data was already a time series data.

```
irates <- read_dta("C:/Users/gaeta/Documents/Github/TimeSeries/Data/irates.dta")
attach(irates)
ts(r1)
ts(r3)
```

A)

Testing for a unit root in r1 using several augmented Dickey Fuller tests.

```
tseries::adf.test(r1,k=0) # Dickey Fuller Test: no unit root
```

```
##
## Augmented Dickey-Fuller Test
##
## data: r1
## Dickey-Fuller = -3.5056, Lag order = 0, p-value = 0.04177
## alternative hypothesis: stationary
```

```
tseries::adf.test(r1,k=1) # Augmented DF Test with 1 lag: no unit root
```

```
##
## Augmented Dickey-Fuller Test
##
## data: r1
## Dickey-Fuller = -3.6428, Lag order = 1, p-value = 0.02858
## alternative hypothesis: stationary
```

```
tseries::adf.test(r1,k=2) # Augmented DF Test with 2 lags: no unit root
```

```
##
## Augmented Dickey-Fuller Test
##
## data: r1
## Dickey-Fuller = -3.6091, Lag order = 2, p-value = 0.03182
## alternative hypothesis: stationary
```

```
tseries::adf.test(r1,k=3) # Augmented DF Test with 3 lags: unit root
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r1  
## Dickey-Fuller = -3.1973, Lag order = 3, p-value = 0.08839  
## alternative hypothesis: stationary
```

```
tseries::adf.test(r1,k=4) # Augmented DF Test with 4 lags: unit root
```

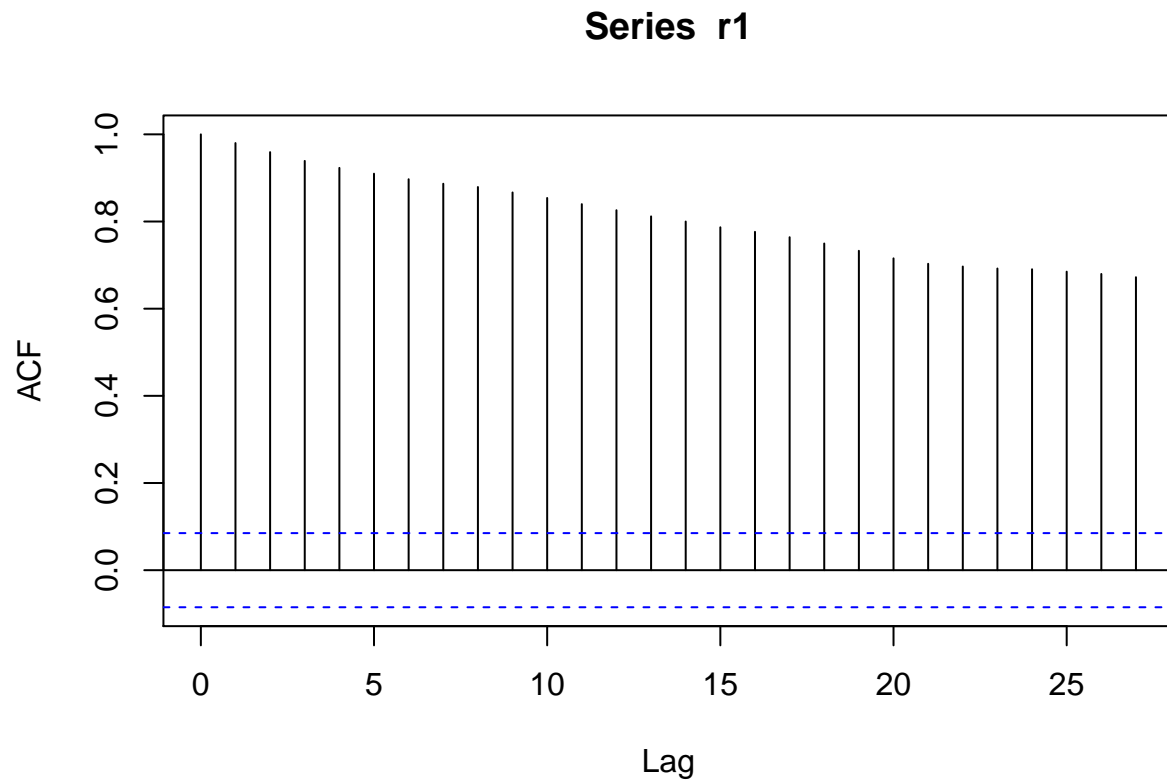
```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r1  
## Dickey-Fuller = -2.9642, Lag order = 4, p-value = 0.1702  
## alternative hypothesis: stationary
```

```
tseries::adf.test(r1) # Augmented DF Test with upper bounds lags: unit root
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r1  
## Dickey-Fuller = -2.8044, Lag order = 8, p-value = 0.2378  
## alternative hypothesis: stationary
```

When we run the augmented Dickey Fuller tests with 0,1 and 2 lags, we reject the null hypothesis that the time series data contains a unit root. However, when we we increase the number of lags to 3 or more, we fail to reject the null that the time series data has a unit root. Looking at the ACF of the time series, the graph below, I would say that the more lags that we have included are better, thus the r1 time series does indeed have a unit root.

```
acf(r1)
```



B)

Testing for a unit root in r3 using several augmented Dickey Fuller tests.

```
tseries::adf.test(r3,k=0) # Dickey Fuller Test: unit root
```

```
##
## Augmented Dickey-Fuller Test
##
## data: r3
## Dickey-Fuller = -3.0286, Lag order = 0, p-value = 0.1429
## alternative hypothesis: stationary
```

```
tseries::adf.test(r3,k=1) # Augmented DF Test with 1 lag: no unit root
```

```
##
## Augmented Dickey-Fuller Test
##
## data: r3
## Dickey-Fuller = -3.484, Lag order = 1, p-value = 0.04384
## alternative hypothesis: stationary
```

```
tseries::adf.test(r3,k=2) # Augmented DF Test with 2 lags: unit root
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r3  
## Dickey-Fuller = -3.2046, Lag order = 2, p-value = 0.08713  
## alternative hypothesis: stationary
```

```
tseries::adf.test(r3,k=3) # Augmented DF Test with 3 lags: unit root
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r3  
## Dickey-Fuller = -3.1842, Lag order = 3, p-value = 0.09065  
## alternative hypothesis: stationary
```

```
tseries::adf.test(r3,k=4) # Augmented DF Test with 4 lags: unit root
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r3  
## Dickey-Fuller = -2.9286, Lag order = 4, p-value = 0.1853  
## alternative hypothesis: stationary
```

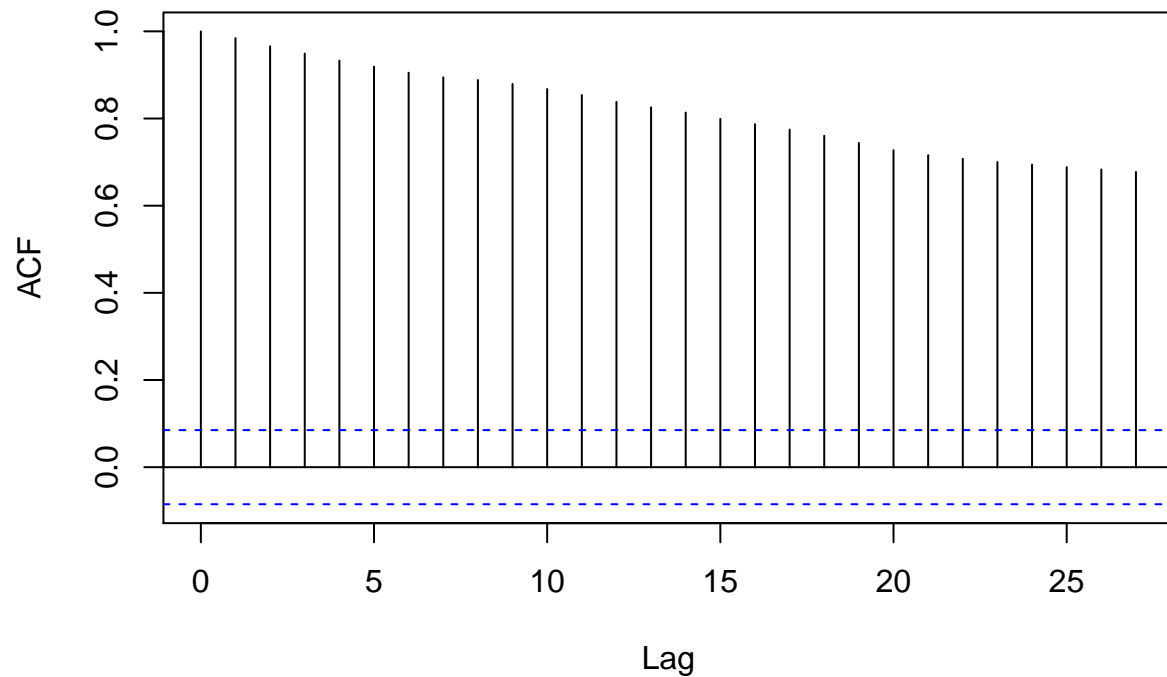
```
tseries::adf.test(r3) # Augmented DF Test with upper bounds lags: unit root
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: r3  
## Dickey-Fuller = -2.5198, Lag order = 8, p-value = 0.3583  
## alternative hypothesis: stationary
```

What is interesting here is that when I run the simple Dickey Fuller test and the augmented DF test with 2 or more lags, I find that I consistently fail to reject the null hypothesis that there is a unit root present in the time series. Looking at the ACF of the time series, I do not see why that would be the case. Either way, we can conclude that there is a unit root in the r3 time series data.

```
acf(r3)
```

## Series r3



Performing a regression by OLS explaining r1 from r3 and testing for co-integration.

```
Reg1 <- lm(r1~r3)
Reg1Resid <- Reg1$residuals
tseries::adf.test(Reg1Resid,k=0) # no unit root
```

```
## Warning in tseries::adf.test(Reg1Resid, k = 0): p-value smaller than printed p-
## value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: Reg1Resid
## Dickey-Fuller = -16.846, Lag order = 0, p-value = 0.01
## alternative hypothesis: stationary
```

```
tseries::adf.test(Reg1Resid,k=1) # no unit root
```

```
## Warning in tseries::adf.test(Reg1Resid, k = 1): p-value smaller than printed p-
## value
```

```
##
## Augmented Dickey-Fuller Test
##
```

```
## data: Reg1Resid
## Dickey-Fuller = -11.598, Lag order = 1, p-value = 0.01
## alternative hypothesis: stationary
```

```
tseries::adf.test(Reg1Resid,k=2) # no unit root
```

```
## Warning in tseries::adf.test(Reg1Resid, k = 2): p-value smaller than printed p-
## value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: Reg1Resid
## Dickey-Fuller = -11.825, Lag order = 2, p-value = 0.01
## alternative hypothesis: stationary
```

```
tseries::adf.test(Reg1Resid,k=3) # no unit root
```

```
## Warning in tseries::adf.test(Reg1Resid, k = 3): p-value smaller than printed p-
## value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: Reg1Resid
## Dickey-Fuller = -9.6504, Lag order = 3, p-value = 0.01
## alternative hypothesis: stationary
```

```
tseries::adf.test(Reg1Resid) # no unit root. Long Run Relationship between r3 and r1.
```

```
## Warning in tseries::adf.test(Reg1Resid): p-value smaller than printed p-value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: Reg1Resid
## Dickey-Fuller = -6.7803, Lag order = 8, p-value = 0.01
## alternative hypothesis: stationary
```

C)

Performing a regression by OLS explaining  $r_3$  from  $r_1$  and testing for co-integration.

```
Reg2 <- lm(r3~r1)
Reg2Resid <- Reg2$residuals
tseries::adf.test(Reg2Resid,k=0) # no unit root
```

```
## Warning in tseries::adf.test(Reg2Resid, k = 0): p-value smaller than printed p-
## value
```

```

##
## Augmented Dickey-Fuller Test
##
## data: Reg2Resid
## Dickey-Fuller = -16.898, Lag order = 0, p-value = 0.01
## alternative hypothesis: stationary

tseries::adf.test(Reg2Resid,k=1) # no unit root

## Warning in tseries::adf.test(Reg2Resid, k = 1): p-value smaller than printed p-
## value

##
## Augmented Dickey-Fuller Test
##
## data: Reg2Resid
## Dickey-Fuller = -11.62, Lag order = 1, p-value = 0.01
## alternative hypothesis: stationary

tseries::adf.test(Reg2Resid,k=2) # no unit root

## Warning in tseries::adf.test(Reg2Resid, k = 2): p-value smaller than printed p-
## value

##
## Augmented Dickey-Fuller Test
##
## data: Reg2Resid
## Dickey-Fuller = -11.827, Lag order = 2, p-value = 0.01
## alternative hypothesis: stationary

tseries::adf.test(Reg2Resid,k=3) # no unit root

## Warning in tseries::adf.test(Reg2Resid, k = 3): p-value smaller than printed p-
## value

##
## Augmented Dickey-Fuller Test
##
## data: Reg2Resid
## Dickey-Fuller = -9.6514, Lag order = 3, p-value = 0.01
## alternative hypothesis: stationary

tseries::adf.test(Reg2Resid) # no unit root. Long Run Relationship between r3 and r1.

## Warning in tseries::adf.test(Reg2Resid): p-value smaller than printed p-value

##
## Augmented Dickey-Fuller Test
##
## data: Reg2Resid
## Dickey-Fuller = -6.775, Lag order = 8, p-value = 0.01
## alternative hypothesis: stationary

```

D)

Cointegration when there is a stable relationship between two unit root time series. In this case we do want to regress the unit root time series on the other one because this regression would not be spurious. The reason why we would want to do this is because the estimate for  $\beta$  with OLS is not only consistent, but it is “super consistent” meaning that it converges to the true  $\beta$  much faster than usual asymptotic. In short, cointegrated series have stochastic trend in common.

In this case, we can see that the r1 and r3 time series data contain unit roots from the adf test. We think that there exists a linear relationship between the two since they are both interest rates. To test this theory, we run a regression of r1 on r3 and vice versa, we take the residuals, which represent the equilibrium error, to see if they reject the null hypothesis that the residuals contain a unit root. As we can see from the results, the residuals we consistently reject the null hypothesis that there is a unit root in the residuals. IF the residuals are stationary, i.e. do not contain a unit root, then if r1 or r3 deviate from the equilibrium then it would eventually get back to it.

However, it is important to note that when we do the augmented DF tests on the residuals we look at the DF value, not necessarily the p-value, and compare it to the Davidson and Mackinnon Cointegrated Regression Dickey Fuller critical value table. The reason is that the OLS chooses small variance in the residuals leading us to reject the null hypothesis more often than not. In either case, no matter the number of lags, we get a DF statistic less than -3.90, meaning that we reject the null hypothesis.

What do you mean interpret results intuitively.

In this context, the interest rate spread moves over time but will eventually settle at its mean.

## Question 2:

### I) Cholesky Decomposition

A. The vector of structural shocks is:

$$\begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

The assumptions that we usually make for those shocks is that they are exogenous to each variable, meaning that they are uncorrelated with each. We also assume that they have a mean of 0 and a variance covariance matrix D written as:

$$\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

B. The reduced form model is written as:

$$y_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 y_{t-1} + B^{-1}\epsilon_t \quad (1)$$

$$= C + \Phi y_{t-1} + e_t \quad (2)$$

Assuming that our structural VAR model is of the form:

$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + \epsilon_t \quad (3)$$

The assumptions that we make is that  $\beta_{12} \neq \beta_{21} \neq |1|$  so that the inverse of the  $B$  matrix, representing the coefficients, exists.

The reduced form shocks in relation to the structural one is given by:

$$e_t = B^{-1}\epsilon_t \quad (4)$$

$$= \frac{1}{1 - \beta_{12}\beta_{21}} \begin{bmatrix} 1 & \beta_{12} \\ \beta_{12} & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \quad (5)$$

$$= \frac{1}{1 - \beta_{12}\beta_{21}} \begin{bmatrix} \epsilon_{1t} & \beta_{12}\epsilon_{2t} \\ \beta_{12}\epsilon_{1t} & \epsilon_{2t} \end{bmatrix} \quad (6)$$



The assumptions that we are making is that  $A_0$  and  $A_0^{-1}$  are lower-triangular. This means that we can identify the reduced form shock recursively once we know  $a_{11}$  etc... The assumptions that we have made are the same as above. What we are assuming is that the shock on the real price does not affect contemporaneously the changes in oil production and neither does the shock to real economic activity. In the next equation, the shock on the real price of oil does not affect contemporaneously real economic activity.

The variance-covariance matrix using those assumptions is. What do you mean by all the entries?

$$\begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix}$$

It is important to note that the  $\Omega$  matrix is the variance-covariance matrices for the forecast error. It is not diagonal as both forecast errors are affected by all the shocks. Since  $\omega_{12} = \omega_{21}$ ,  $\omega_{13} = \omega_{31}$  and  $\omega_{23} = \omega_{32}$  it cannot be the case that all of these are equal to zero. If they were then we would a diagonal matrix.

C. The Cholesky decomposition allows us to find  $A^{-1}$  and the covariance-variance matrix of the structural shocks that I will denote from now on as  $D$  as the decomposition is unique. The Cholesky equation is:

$$\Omega = T\Delta T' \tag{7}$$

where  $\Delta$  is a diagonal matrix containing positive elements and  $T$  is a lower triangular matrix with 1s on the diagonal. Since we know that  $A_0^{-1}$  is lower triangular we can say that  $T = A_0^{-1}$  and that  $D$  is a diagonal matrix with positive elements we can say that  $\Delta = D$ . So from the reduced form VAR and the Cholesky decomposition we can get the structural VAR assuming that  $A^{-1}$  is indeed lower triangular.

It is important to order the variables carefully in order to identify the specific parameters of interest through Cholesky because you will not get the same results if you were to change the order.