University of Duisburg-Essen
Faculty of Business Administration and Economics
Chair of Econometrics

Winter Term 2019/2020

Dr. Yannick Hoga Thilo Reinschlüssel

Multivariate Time Series Analysis

Solution Exercise Sheet 4

1 Exercise 1: Information Criteria

Prove Corollary 4.5 from Slide 4-7.

Solution:

From Theorem 4.4:

$$C(l) = \log\left(\hat{\Sigma}_a(l)\right) + \frac{l}{T} \cdot c_T$$

i) $\lim_{T\to\infty} c_T \longrightarrow \infty$

ii) $\lim_{T\to\infty} \frac{c_T}{T} \longrightarrow 0$

If i) and ii) hold, $C(l)$ chooses the optimal/correct model.

- AIC: $c_T = 2K^2$

$$\lim_{T\to\infty} c_T = 2\ K^2 \not\Longrightarrow\ \infty$$

$\Rightarrow$ not consistent

- BIC: $c_T = \log(T) \cdot K^2$

$$\lim_{T\to\infty} c_T = \log(T)K^2 \implies \infty$$
$$\lim_{T\to\infty} \frac{c_T}{T} = \frac{\log(T)}{T}K^2 \implies 0$$

$\Rightarrow$ consistent

- HQ: $c_T = 2\ \log(\log(T))\ K^2$

$$\lim_{T\to\infty} c_T = 2\ \log(\log(T))\ K^2 \implies \infty$$
$$\lim_{T\to\infty} \frac{c_T}{T} = \frac{2\ \log(\log(T))\ K^2}{T} \implies 0$$

$\Rightarrow$ consistent

2 Exercise 2: VAR(p): Data application

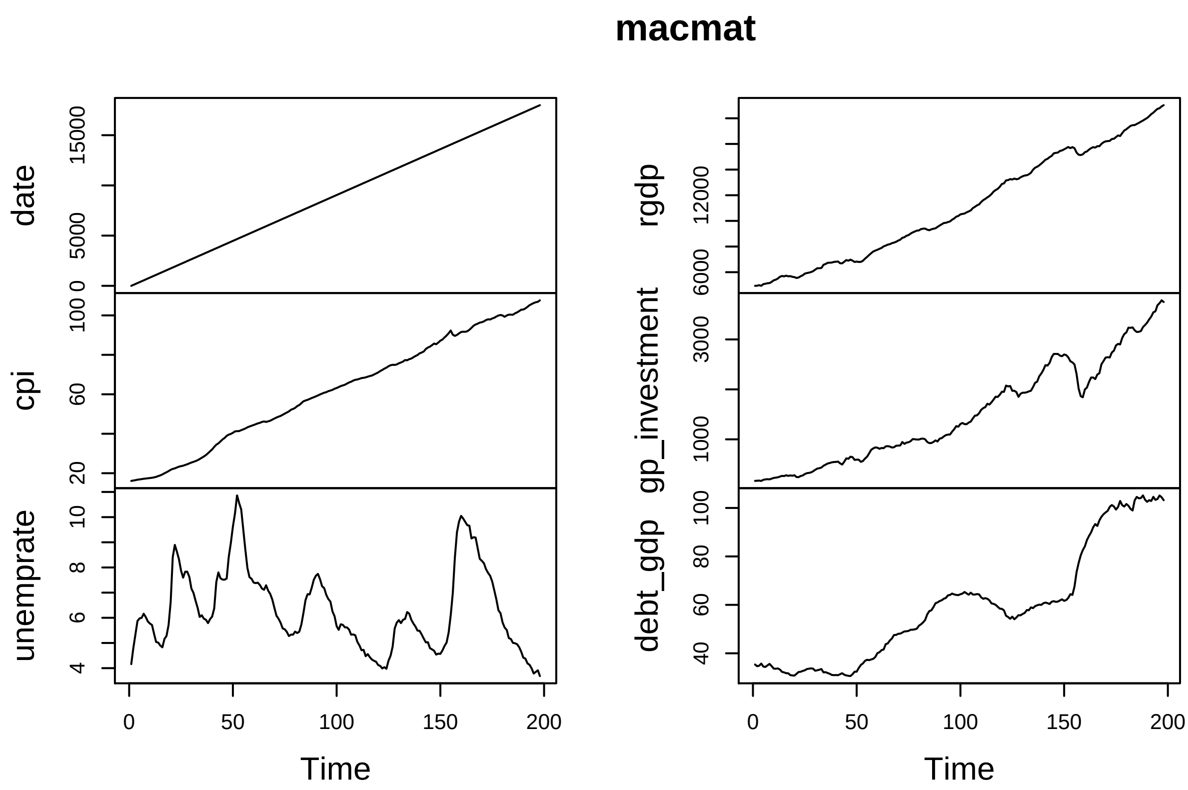
This exercise is concerned with finding an appropriate VAR(p) model for US macroeconomic data. You can find the dataset `us_macrodata.Rda` attached to this exercise sheet in the Moodle folder for this tutorial. Please use the `load` command to import the dataset from your directory into R. There are 5 variables – CPI, Real GDP, the unemployment rate, general private investment and the debt-to-GDP ratio. All series have been sampled quarterly and were seasonally adjusted before downloaded from FRED.

```
# loading data
load(file = here::here("exercise_MTSA/00_data/us_macrodata.Rda"))
# loading the MTS package
library(MTS)
```

a.) Plot all time series and judge which time series seem non-stationary. Proceed to compute growth rates of the non-stationary variables.

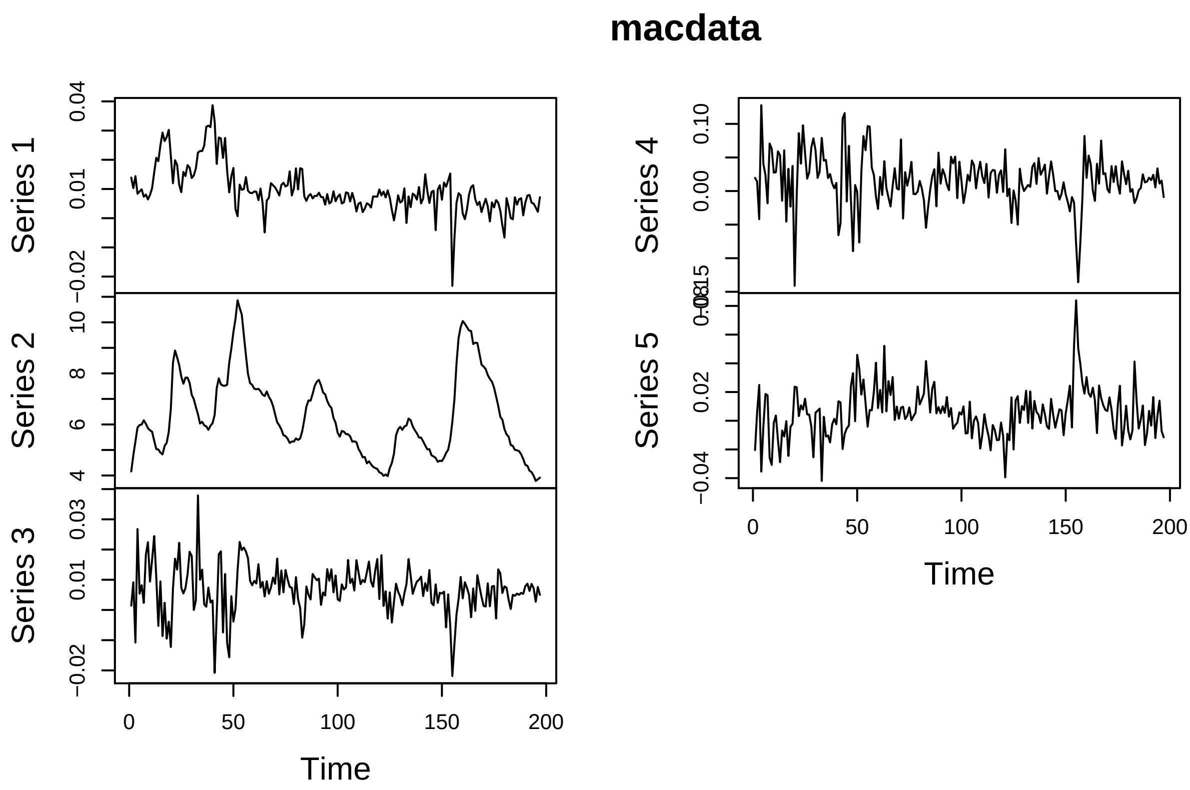
Solution:

```
macmat <- data.matrix(us.macro_series)
plot.ts(macmat)
```



Every series except unemployment looks non-stationary. Regarding the debt-to-gdp ratio, this is surprising, but we better difference it as well.

```
macdata <- cbind(diff(log(us.macro_series$cpi)),
                 us.macro_series$unemprate[-(nrow(macmat))],
                 diff(log(us.macro_series$rgdp)),
                 diff(log(us.macro_series$gp_investment)),
                 diff(log(us.macro_series$debt_gdp)))

plot.ts(macdata)
```



Note that the last observation of "unemp" was dropped for conformable length. Its last and not first due to the date information: measurements are always from the first day of a quarter.

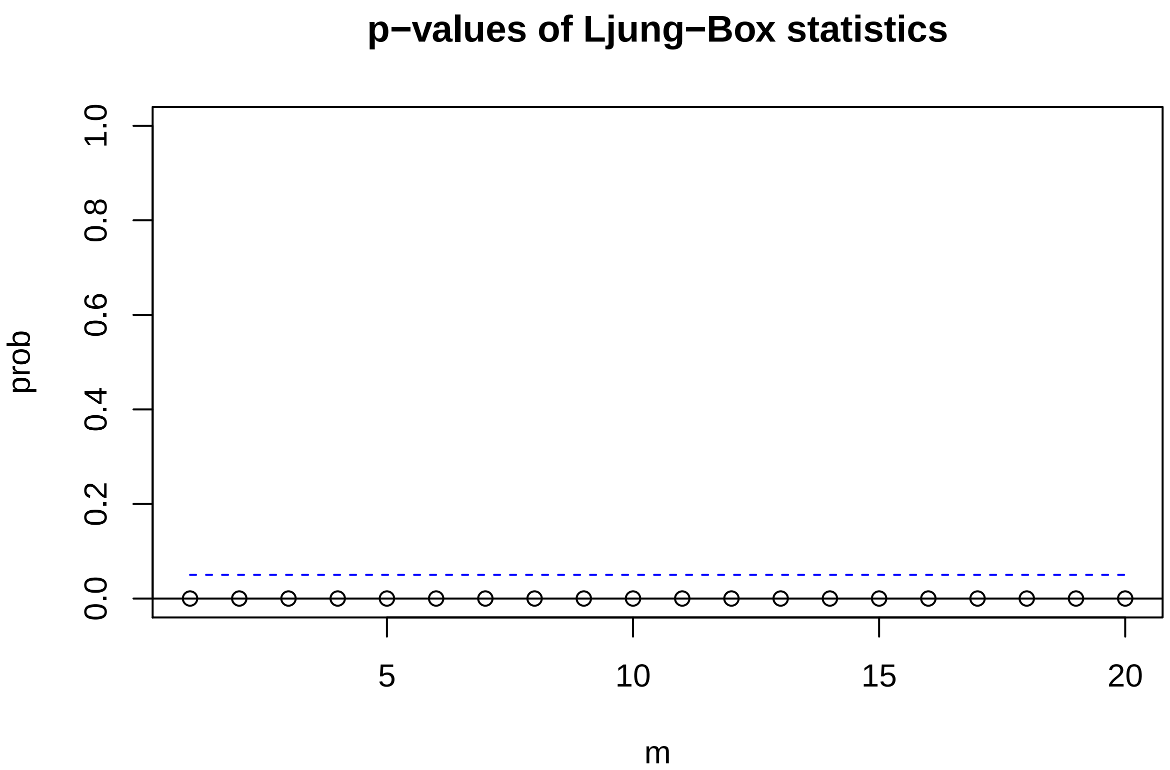
b.) Perform a Ljung-Box test on the dataset. Does it look worthwhile to estimate a VAR(p)

Solution:

```
mq(x = macdata, lag = 20)
```

```
## Ljung-Box Statistics:
##          m       Q(m)     df    p-value
##  [1,]     1       369      25         0
##  [2,]     2       658      50         0
##  [3,]     3       932      75         0
##  [4,]     4      1211     100         0
##  [5,]     5      1430     125         0
##  [6,]     6      1624     150         0
##  [7,]     7      1796     175         0
##  [8,]     8      1953     200         0
##  [9,]     9      2083     225         0
## [10,]    10      2205     250         0
```

```
## [11,]     11        2313      275          0
## [12,]     12        2418      300          0
## [13,]     13        2513      325          0
## [14,]     14        2619      350          0
## [15,]     15        2702      375          0
## [16,]     16        2793      400          0
## [17,]     17        2881      425          0
## [18,]     18        2965      450          0
## [19,]     19        3031      475          0
## [20,]     20        3112      500          0
```

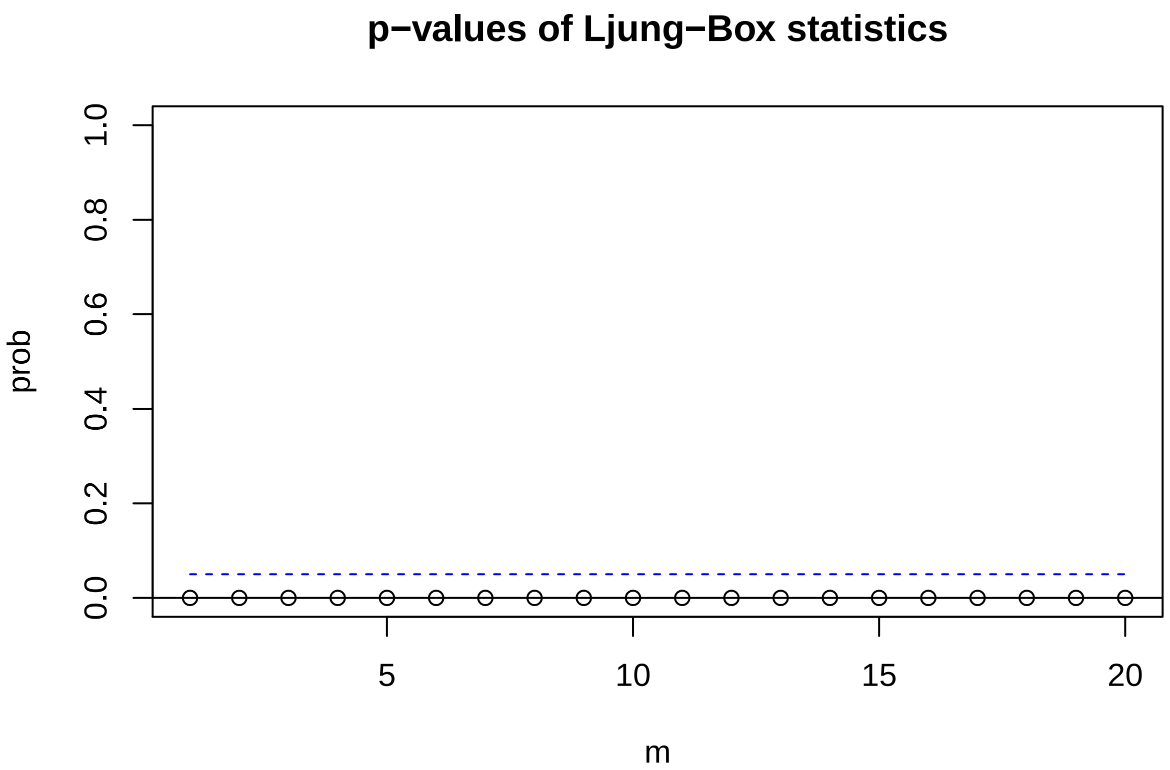


There is some correlation in the dataset.

```r
mq(x = macdata[,-3], lag = 20)
```

```
## Ljung-Box Statistics:
##          m       Q(m)      df     p-value
##  [1,]     1       337      16          0
##  [2,]     2       607      32          0
##  [3,]     3       866      48          0
##  [4,]     4      1121      64          0
```

```
## [5,]   5   1325   80   0
## [6,]   6   1502   96   0
## [7,]   7   1650  112   0
## [8,]   8   1781  128   0
## [9,]   9   1894  144   0
## [10,]  10  2003  160   0
## [11,]  11  2101  176   0
## [12,]  12  2197  192   0
## [13,]  13  2278  208   0
## [14,]  14  2372  224   0
## [15,]  15  2444  240   0
## [16,]  16  2518  256   0
## [17,]  17  2589  272   0
## [18,]  18  2651  288   0
## [19,]  19  2702  304   0
## [20,]  20  2761  320   0
```



Even without unemployment, there is some correlation in the dataset.

c.) Determine the length of the time series. How many coefficients can be estimated and what does it mean for K and p?

Solution:

We have $T \cdot K$ data points and we estimate K^2 parameters for each lag. For the intercept we estimate K parameters. Which leads to the following condition for the maximal number of lag(s) p:

$$\frac{K \cdot (T-1)}{K^2} \geq p$$

```
data_dim <- dim(macdata)

Tmax <- data_dim[1] # observations
K <- data_dim[2] # variables
(max.p <- (Tmax * K - K) / K^2 )
```

```
## [1] 39.2
```

39 lags can be estimated in addition to the intercept.

d.) Consult the AIC, BIC and HQ to determine the optimal lag order for a VAR(p) model for the whole dataset. Plot the values of the three criteria for the lag orders p from 1 to 5 in one plot.

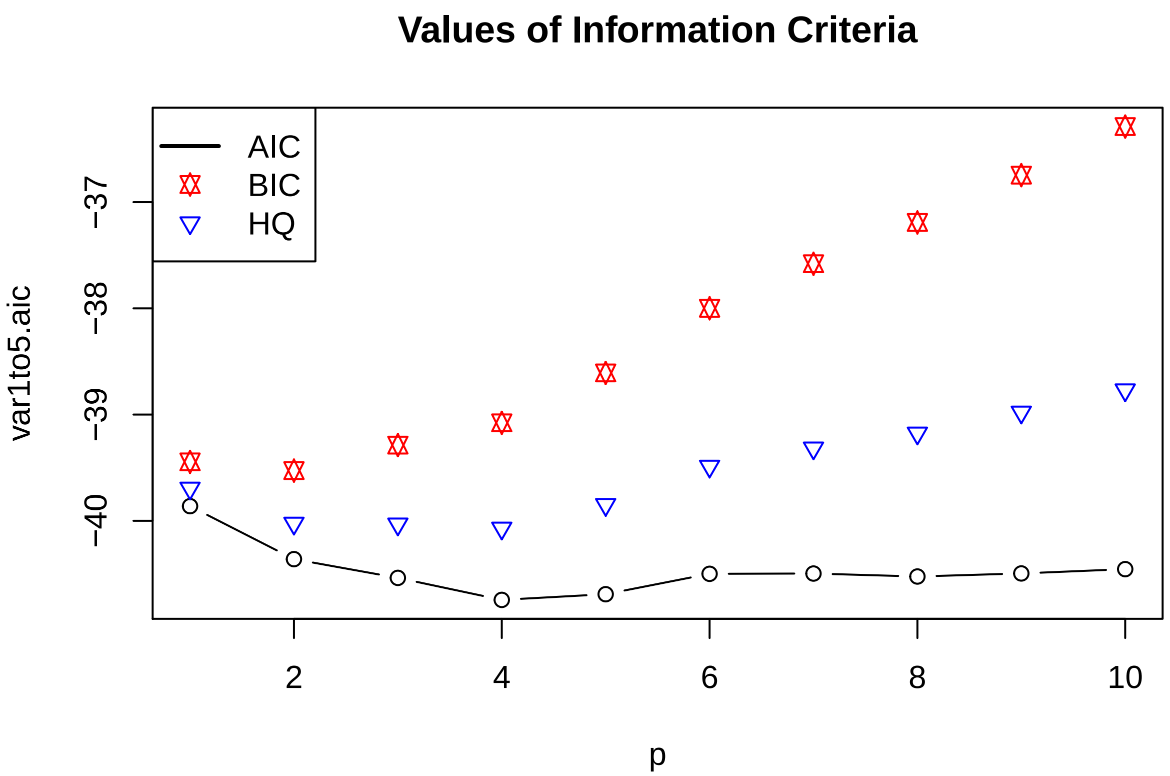
Solution:

```
M <- 10 # maximal p
VARorder(x = macdata, maxp = M)
```

```
## selected order: aic =  4
## selected order: bic =  1
## selected order: hq =  2
## Summary table:
##          p      AIC      BIC       HQ     M(p) p-value
##  [1,]  0 -34.8741 -34.8741 -34.8741   0.0000  0.0000
##  [2,]  1 -40.1273 -39.7107 -39.9587 994.0213  0.0000
##  [3,]  2 -40.4982 -39.6649 -40.1608 109.6255  0.0000
##  [4,]  3 -40.6510 -39.4011 -40.1450  69.3340  0.0000
##  [5,]  4 -40.7551 -39.0885 -40.0805  59.2354  0.0001
##  [6,]  5 -40.6962 -38.6129 -39.8529  31.2764  0.1800
##  [7,]  6 -40.5110 -38.0111 -39.4990  10.6681  0.9944
##  [8,]  7 -40.5487 -37.6321 -39.3680  43.8714  0.0112
```

```
## [9,]  8 -40.5490 -37.2158 -39.1997  36.9772  0.0580
## [10,]  9 -40.4934 -36.7436 -38.9754  27.8481  0.3149
## [11,] 10 -40.4553 -36.2888 -38.7687  29.2269  0.2545
```

```r
# so it's p = 1, 2 or 4.....
var1to5.fit <- lapply(X = 1:M, function(i)
  VAR(x = macdata, include.mean = TRUE, p = i, output = FALSE))
var1to5.aic <- sapply(1:M, function(i) var1to5.fit[[i]]$aic)
var1to5.bic <- sapply(1:M, function(i) var1to5.fit[[i]]$bic)
var1to5.hq <- sapply(1:M, function(i) var1to5.fit[[i]]$hq)
plot(x = 1:M, y = var1to5.aic, type = "b",
     main = "Values of Information Criteria", xlab = "p",
     ylim = c( min(c(var1to5.aic, var1to5.bic, var1to5.hq)),
               max(c(var1to5.aic, var1to5.bic, var1to5.hq)) ) )

points(x = 1:M, y = var1to5.bic, pch = 11, col = "red")
points(x = 1:M, y = var1to5.hq, col = "blue", pch = 25)
legend("topleft", legend = c("AIC", "BIC", "HQ"),
       col = c("black", "red", "blue"),
       pch = c(NA, 11, 25), lwd = c(2, NA, NA))
```



- AIC and HQ are flat around the minima ⇒ no distinct optimum visible.

- Conceivable reasons: persistence, omitted variables, wrong functional form
- The VAR may just work as an approximation
- BIC is the most conservative IC
- Mimima at:

$$p = \begin{cases} 1 & BIC \\ 2 & HQ \\ 4 & AIC \end{cases}$$

e.) Fit VAR(p) models incorporating all variables using the optimal lag order(s) p suggested by each of the information criteria. Apply the Ljung-Box test to inspect the residuals' properties. For which models does the test reject the null hypothesis on one of the first ten lags?

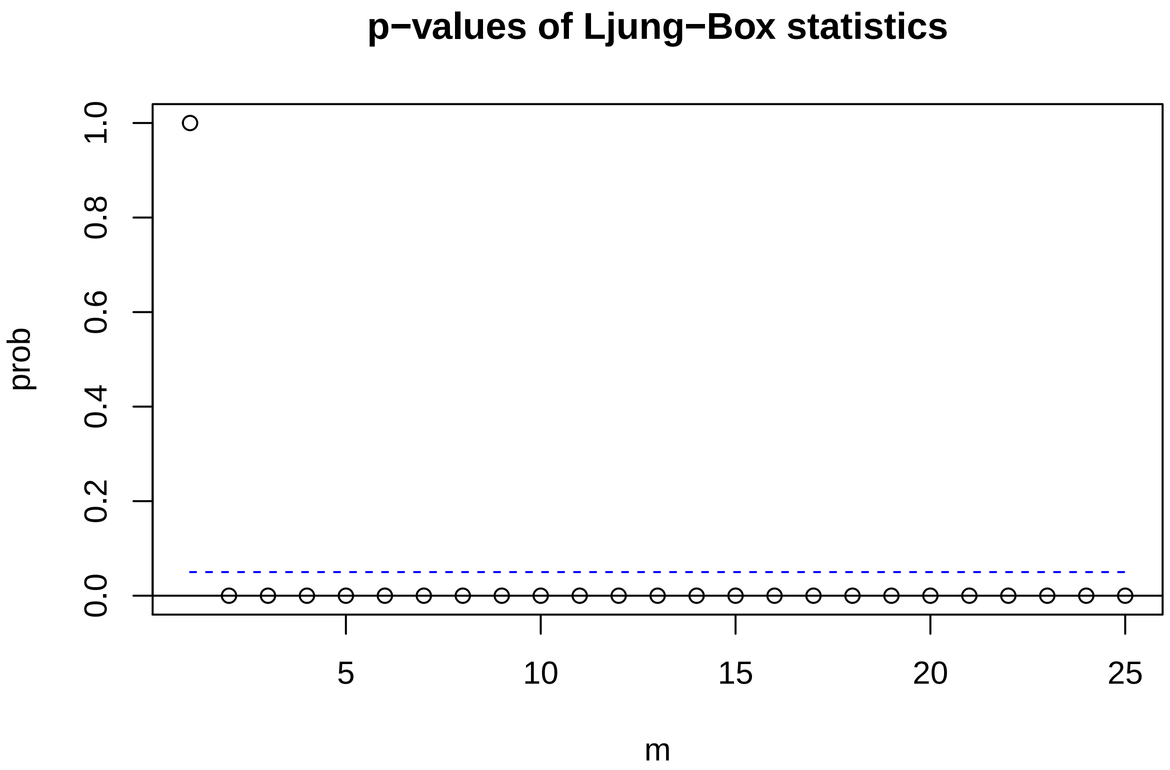
Solution:

First we need to estimated a VAR(p).

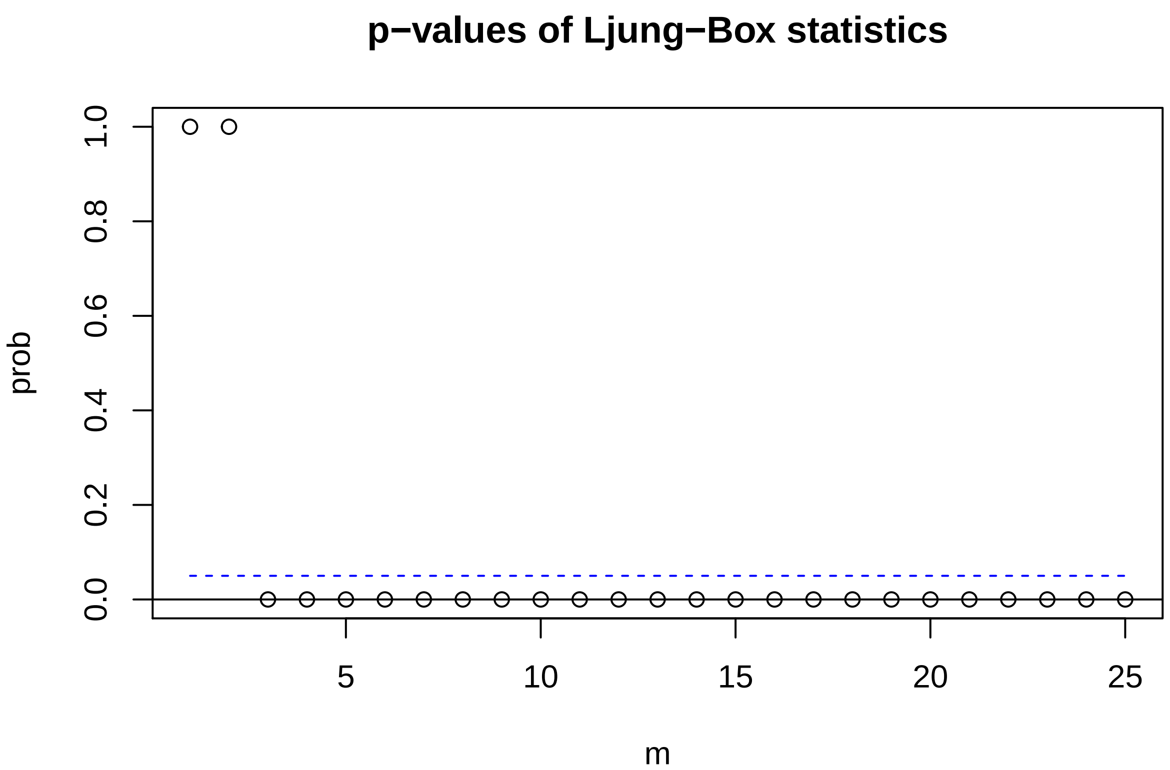
```
var1.fit <- VAR(x = macdata, p = 1, include.mean = TRUE, output = FALSE)
var2.fit <- VAR(x = macdata, p = 2, include.mean = TRUE, output = FALSE)
var4.fit <- VAR(x = macdata, p = 4, include.mean = TRUE, output = FALSE)
```

Now the Ljung-Box test can be performed. We adjust using $K = 5$ with $5^2 \times p$ degree of freedom. Adjustment for the intercept is not necessary, since z_t needs to be demeaned anyway $\left(\Gamma_0 = (z_t - \mu)(z_t - \mu)'\right)$. The Ljun-Box test has $m \times K^2$ degree of freedom (K^2 per lagged cross-correlation matrices), so after adjustment, we have $(m-p)K^2$ degrees of freedom.

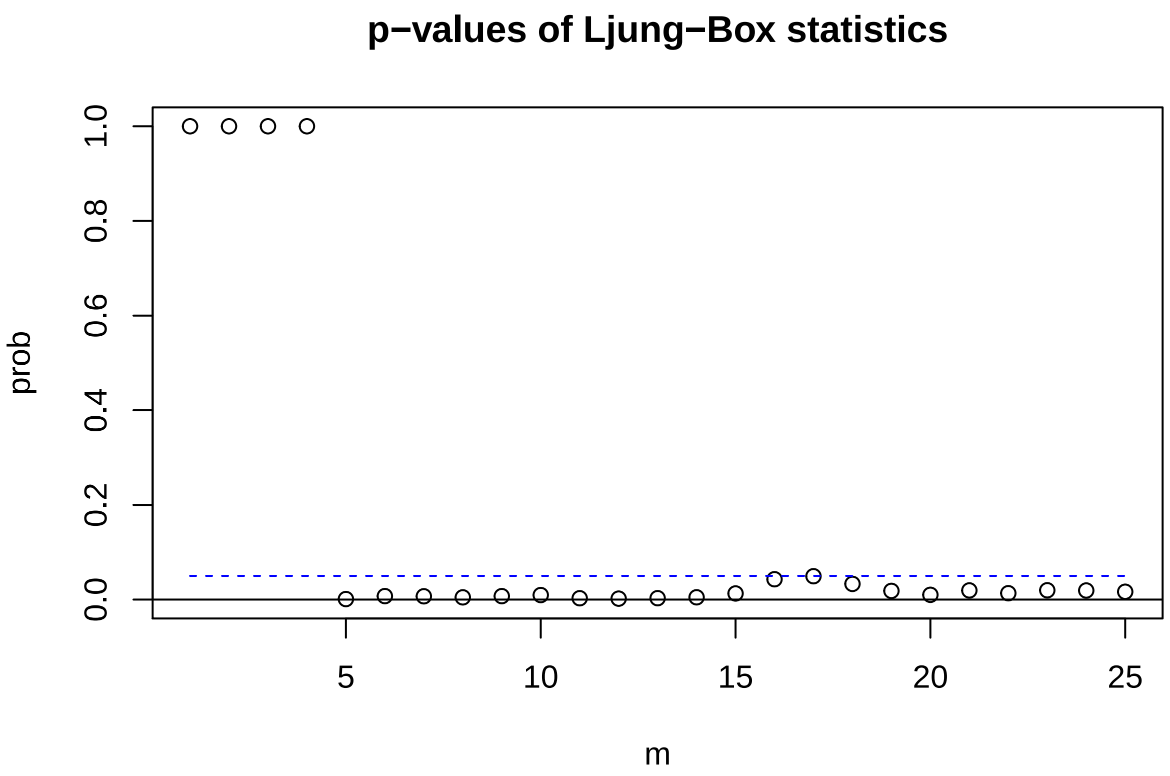
```
mq(x = var1.fit$residuals, lag = 25, adj = 25 * 1)
```




```r
mq(x = var2.fit$residuals, lag = 25, adj = 25 * 2)
```



```r
mq(x = var4.fit$residuals, lag = 25, adj = 25 * 4)
```




```
## Ljung-Box Statistics:
##          m       Q(m)     df    p-value
##  [1,]   1.0      59.7    0.0         1
##  [2,]   2.0     113.3   25.0         0
##  [3,]   3.0     173.6   50.0         0
##  [4,]   4.0     232.8   75.0         0
##  [5,]   5.0     266.1  100.0         0
##  [6,]   6.0     298.1  125.0         0
##  [7,]   7.0     332.9  150.0         0
##  [8,]   8.0     367.1  175.0         0
##  [9,]   9.0     397.3  200.0         0
## [10,]  10.0     450.3  225.0         0
## [11,]  11.0     470.5  250.0         0
## [12,]  12.0     508.8  275.0         0
## [13,]  13.0     535.2  300.0         0
## [14,]  14.0     573.0  325.0         0
## [15,]  15.0     588.0  350.0         0
## [16,]  16.0     612.2  375.0         0
## [17,]  17.0     646.3  400.0         0
```

```
## [18,]  18.0     672.2   425.0        0
## [19,]  19.0     697.7   450.0        0
## [20,]  20.0     740.1   475.0        0
## [21,]  21.0     762.6   500.0        0
## [22,]  22.0     796.2   525.0        0
## [23,]  23.0     818.7   550.0        0
## [24,]  24.0     847.9   575.0        0
## [25,]  25.0     865.4   600.0        0
## Ljung-Box Statistics:
##          m       Q(m)     df    p-value
##  [1,]    1.0      12.7   -25.0        1
##  [2,]    2.0      38.9     0.0        1
##  [3,]    3.0      74.3    25.0        0
##  [4,]    4.0     118.2    50.0        0
##  [5,]    5.0     151.2    75.0        0
##  [6,]    6.0     173.3   100.0        0
##  [7,]    7.0     208.2   125.0        0
##  [8,]    8.0     251.8   150.0        0
##  [9,]    9.0     282.7   175.0        0
## [10,]  10.0     320.9   200.0        0
## [11,]  11.0     346.2   225.0        0
## [12,]  12.0     385.0   250.0        0
## [13,]  13.0     407.7   275.0        0
## [14,]  14.0     444.4   300.0        0
## [15,]  15.0     460.8   325.0        0
## [16,]  16.0     481.2   350.0        0
## [17,]  17.0     511.5   375.0        0
## [18,]  18.0     536.5   400.0        0
## [19,]  19.0     564.8   425.0        0
## [20,]  20.0     600.8   450.0        0
## [21,]  21.0     626.8   475.0        0
## [22,]  22.0     659.7   500.0        0
## [23,]  23.0     686.5   525.0        0
## [24,]  24.0     721.5   550.0        0
## [25,]  25.0     742.9   575.0        0
## Ljung-Box Statistics:
##           m       Q(m)     df    p-value
##  [1,]    1.00      3.38  -75.00     1.00
##  [2,]    2.00     10.06  -50.00     1.00
##  [3,]    3.00     17.79  -25.00     1.00
```

```
## [4,]   4.00   31.52   0.00    1.00
## [5,]   5.00   52.36  25.00    0.00
## [6,]   6.00   77.69  50.00    0.01
## [7,]   7.00  108.50  75.00    0.01
## [8,]   8.00  140.58 100.00    0.00
## [9,]   9.00  166.97 125.00    0.01
## [10,] 10.00  193.58 150.00    0.01
## [11,] 11.00  231.30 175.00    0.00
## [12,] 12.00  262.58 200.00    0.00
## [13,] 13.00  287.90 225.00    0.00
## [14,] 14.00  311.71 250.00    0.00
## [15,] 15.00  329.90 275.00    0.01
## [16,] 16.00  343.34 300.00    0.04
## [17,] 17.00  368.22 325.00    0.05
## [18,] 18.00  400.16 350.00    0.03
## [19,] 19.00  434.46 375.00    0.02
## [20,] 20.00  468.69 400.00    0.01
## [21,] 21.00  487.46 425.00    0.02
## [22,] 22.00  519.00 450.00    0.01
## [23,] 23.00  540.71 475.00    0.02
## [24,] 24.00  567.70 500.00    0.02
## [25,] 25.00  596.44 525.00    0.02
```

Since the tests rejects everywhere else, the VARs do not explain the dynamics entirely.

f.) Now take a VAR(1) and a VAR(4) model with all variables included and an intercept specified.

   (i) How many coefficients are estimated in each case?
   (ii) Look at both estimates of Σ_a - are there major difference?
   (iii) Compare the standard errors associated with the ϕ_1 matrices of the VAR(1) and VAR(4) from above. Do you see the same pattern regarding Σ_a?

Solution:

(i) How many coefficients are estimated in each case?

   The formula for computing the number of parameters is $K^2 \times p + K$. For a VAR(1) we estimate 30 parameters. Whereas for a VAR(4) we already need to estimate 105 parameters.

(ii) Look at both estimates of Σ_a - are there major difference?

Ratio of residual coveriances:

```
var1.fit$Sigma / var4.fit$Sigma
```

```
##           [,1]      [,2]     [,3]     [,4]     [,5]
## [1,] 1.2908794 0.7594541 1.179527 1.161651 1.115374
## [2,] 0.7594541 1.8963595 1.535077 1.242121 1.946967
## [3,] 1.1795272 1.5350769 1.216217 1.282319 1.276541
## [4,] 1.1616514 1.2421212 1.282319 1.272154 1.256466
## [5,] 1.1153745 1.9469672 1.276541 1.256466 1.291657
```

Reisdual (co)variances are higher for the VAR(1). VAR(4) predicts better (in-sample).

(iii) Compare the standard errors associated with the ϕ_1 matrices of the VAR(1) and VAR(4) from above. Do you see the same pattern regarding Σ_a?

Ratio of standard errors:

```
var1.fit$secoef / var4.fit$secoef[1:6,]
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] 0.7270512 0.8812163 0.7057122 0.7217587 0.7272702
## [2,] 0.6486366 0.7861744 0.6295990 0.6439149 0.6488319
## [3,] 0.1268594 0.1537589 0.1231361 0.1259359 0.1268976
## [4,] 0.9832151 1.1916975 0.9543577 0.9760579 0.9835113
## [5,] 0.9100956 1.1030737 0.8833843 0.9034707 0.9103698
## [6,] 0.9956293 1.2067440 0.9664075 0.9883817 0.9959292
```

No, it is exactly the other way around. Estimating more coefficients with the same information leads to less information per coefficient.

g.) Repeat the task from above with CPI and the debt-to-gdp ratio as the only variables (hence $K = 2$). How many coefficients are estimated in this case?

With only two explanatory variables we can estimated maximally 98 lags.

```r
var1red.fit <- VAR(x = macdata[,c(1,5)], p = 1, output = FALSE, include.mean = TRUE)
var4red.fit <- VAR(x = macdata[,c(1,5)], p = 4, output = FALSE, include.mean = TRUE)
```

For a VAR(1) 6 coefficients are estimated, for a VAR(4) 18 coefficients.

Any major differences between the residual covariance matrices?

Ratio of residual coveriances:

```r
var1red.fit$Sigma / var4red.fit$Sigma
```

```
##            [,1]      [,2]
## [1,] 1.1763704 0.9715465
## [2,] 0.9715465 1.2602521
```

Ratio of standard errors:

```r
var1red.fit$secoef / var4red.fit$secoef[1:3,]
```

```
##            [,1]      [,2]
## [1,] 0.9299869 0.9625727
## [2,] 0.6319156 0.6540572
## [3,] 0.9201816 0.9524238
```

h.) At last, go back to VAR(1) and VAR(4) models from task f). Use the standard error matrices to compute t-statistics for each coefficient with the null hypothesis $H_0 : phi_(p, jk) = 0$. How often is the null hypothesis rejected at the 5% level in each of the two models?

To compute the t-statistic the estimated parameters get divided by the standard error of the parameter.

```r
var1.t_ratios <- var1.fit$coef / var1.fit$secoef
var2.t_ratios <- var2.fit$coef / var2.fit$secoef
var4.t_ratios <- var4.fit$coef / var4.fit$secoef
```

To test how often the H_0 is rejected we just count how often the t-value is absolute greater than 1.96 (`sum(abs(var1.t_ratios ) > 1.96)`). For the VAR(1) the H_0 is 16 times which are 0.5333 of all parameters. For a VAR(2) 21 times the null hypothesis is rejected (0.3818) and for a VAR(4) 22 times (0.2095).

Same pattern as in f), but not that pronounced this time.

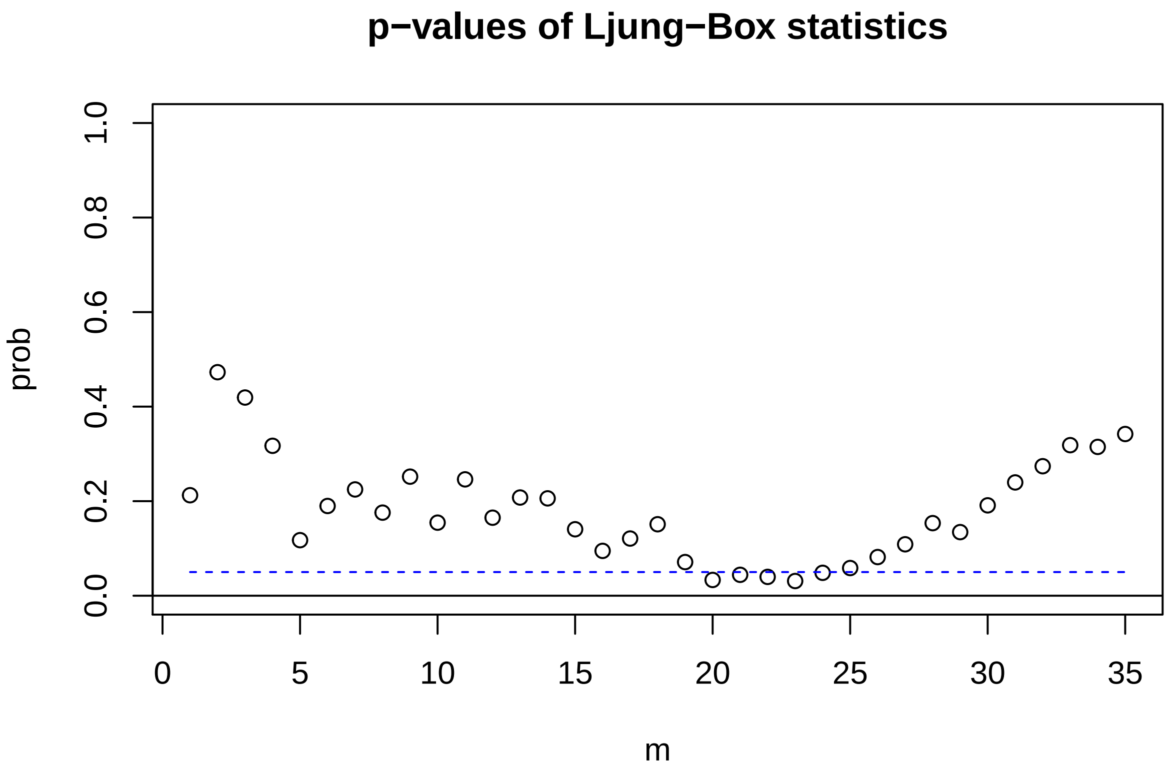
3 Exercise 3: This exercise is concerned with predicting growth rates of ex

To download the data from `Quandl` you need your own API key and execture the following code.

```
library(Quandl)
# Set API key
Quandl.api_key("") # Please enter your key in here.

## Download and prepare data

# Download daily data on Japan/US FX rates
FX.Ja   <- Quandl("FRED/DEXJPUS", start_date = "1998-12-30",
                  end_date = "2018-12-31", type = "xts")
# Download daily data on Euro/US FX rates
FX.Eu   <- Quandl("FRED/DEXUSEU", start_date = "1998-12-30",
                  end_date = "2018-12-31", type = "xts")

# Compute Growth rates
lr.Eu <- diff(log(FX.Eu[, 1]))[-1] # leave out NA in first component via [-1]
lr.Ja <- diff(log(FX.Ja[, 1]))[-1] # leave out NA in first component via [-1]

# only use data when both log-returns are available
V         <- merge.xts(lr.Eu, lr.Ja, all=FALSE)

date      <- index(V)    # save dates for later use
fx_series <- coredata(V)  # raw log-returns
N         <- length(date)
```

a.) Apply the Ljung-Box test on the multivariate time series and comment.

```
mq(x = fx_series, lag = 35)
```




```
## Ljung-Box Statistics:
##             m       Q(m)      df     p-value
##   [1,]    1.00       5.83     4.00      0.21
##   [2,]    2.00       7.61     8.00      0.47
##   [3,]    3.00      12.33    12.00      0.42
##   [4,]    4.00      18.12    16.00      0.32
##   [5,]    5.00      27.67    20.00      0.12
##   [6,]    6.00      29.85    24.00      0.19
##   [7,]    7.00      33.30    28.00      0.22
##   [8,]    8.00      39.29    32.00      0.18
##   [9,]    9.00      41.25    36.00      0.25
##  [10,]   10.00      49.05    40.00      0.15
##  [11,]   11.00      50.04    44.00      0.25
##  [12,]   12.00      57.44    48.00      0.17
##  [13,]   13.00      60.02    52.00      0.21
##  [14,]   14.00      64.41    56.00      0.21
##  [15,]   15.00      71.85    60.00      0.14
##  [16,]   16.00      79.25    64.00      0.09
##  [17,]   17.00      81.83    68.00      0.12
##  [18,]   18.00      84.36    72.00      0.15
##  [19,]   19.00      94.80    76.00      0.07
```

```
## [20,]  20.00    104.72   80.00      0.03
## [21,]  21.00    107.29   84.00      0.04
## [22,]  22.00    112.58   88.00      0.04
## [23,]  23.00    118.91   92.00      0.03
## [24,]  24.00    120.12   96.00      0.05
## [25,]  25.00    123.08  100.00      0.06
## [26,]  26.00    124.67  104.00      0.08
## [27,]  27.00    126.42  108.00      0.11
## [28,]  28.00    127.27  112.00      0.15
## [29,]  29.00    132.94  116.00      0.13
## [30,]  30.00    133.34  120.00      0.19
## [31,]  31.00    134.78  124.00      0.24
## [32,]  32.00    137.16  128.00      0.27
## [33,]  33.00    139.13  132.00      0.32
## [34,]  34.00    143.42  136.00      0.31
## [35,]  35.00    146.23  140.00      0.34
```

Except those correlations around lag 20 to 24, there seems to be no commanding dynamic pattern here.

b.) Do the usual information criteria support the finding of the Ljung-Box test?

```
VARorder(x = fx_series, maxp = 35)
```

```
## selected order: aic =  0
## selected order: bic =  0
## selected order: hq =  0
## Summary table:
##          p      AIC      BIC       HQ    M(p) p-value
##   [1,]   0 -20.3509 -20.3509 -20.3509  0.0000  0.0000
##   [2,]   1 -20.3504 -20.3452 -20.3485  5.0619  0.2810
##   [3,]   2 -20.3493 -20.3389 -20.3456  2.5648  0.6331
##   [4,]   3 -20.3484 -20.3328 -20.3429  3.4809  0.4808
##   [5,]   4 -20.3480 -20.3272 -20.3407  5.8327  0.2120
##   [6,]   5 -20.3485 -20.3226 -20.3394 10.6929  0.0302
##   [7,]   6 -20.3474 -20.3162 -20.3364  2.0955  0.7182
##   [8,]   7 -20.3463 -20.3100 -20.3336  2.8209  0.5882
##   [9,]   8 -20.3460 -20.3045 -20.3315  6.5108  0.1641
## [10,]   9 -20.3448 -20.2980 -20.3284  1.6483  0.8001
## [11,]  10 -20.3447 -20.2928 -20.3265  7.5959  0.1076
```

```
## [12,] 11 -20.3433 -20.2862 -20.3233  0.8605  0.9302
## [13,] 12 -20.3435 -20.2811 -20.3216  8.7104  0.0688
## [14,] 13 -20.3423 -20.2748 -20.3187  2.3603  0.6698
## [15,] 14 -20.3417 -20.2690 -20.3162  4.6531  0.3248
## [16,] 15 -20.3419 -20.2640 -20.3146  9.0812  0.0591
## [17,] 16 -20.3421 -20.2590 -20.3130  8.6706  0.0699
## [18,] 17 -20.3409 -20.2526 -20.3100  2.1661  0.7052
## [19,] 18 -20.3398 -20.2463 -20.3070  2.3256  0.6761
## [20,] 19 -20.3403 -20.2415 -20.3057 10.0891  0.0390
## [21,] 20 -20.3407 -20.2368 -20.3043 10.1670  0.0377
## [22,] 21 -20.3396 -20.2305 -20.3014  2.4550  0.6527
## [23,] 22 -20.3393 -20.2250 -20.2992  6.0724  0.1938
## [24,] 23 -20.3391 -20.2197 -20.2973  7.2942  0.1211
## [25,] 24 -20.3379 -20.2132 -20.2942  1.8071  0.7712
## [26,] 25 -20.3369 -20.2070 -20.2913  2.6354  0.6206
## [27,] 26 -20.3356 -20.2005 -20.2883  1.5773  0.8129
## [28,] 27 -20.3344 -20.1941 -20.2853  2.1157  0.7145
## [29,] 28 -20.3330 -20.1875 -20.2820  0.7476  0.9453
## [30,] 29 -20.3326 -20.1820 -20.2798  6.1498  0.1882
## [31,] 30 -20.3311 -20.1753 -20.2765  0.4671  0.9766
## [32,] 31 -20.3299 -20.1689 -20.2735  1.7841  0.7754
## [33,] 32 -20.3287 -20.1625 -20.2705  2.0704  0.7228
## [34,] 33 -20.3276 -20.1562 -20.2676  2.4413  0.6552
## [35,] 34 -20.3268 -20.1502 -20.2649  3.8463  0.4272
## [36,] 35 -20.3259 -20.1441 -20.2622  3.4624  0.4836
```

The information criterias (AIC, BIC, and HQ) support the findings of the Ljung-Box test.

c.) Regardless of a) and b), fit a VAR(1) to the time series. Compare Σ_a with Γ_0.

```
fx_var1.fit <- VAR(x = fx_series, p = 1, include.mean = TRUE, output = FALSE)
( Gamma_0 <- cov(fx_series))
```

```
##                lr.Eu         lr.Ja
## lr.Eu  3.807917e-05 -1.129738e-05
## lr.Ja -1.129738e-05  4.206437e-05
```

```
(fx_var1.fit$Sigma)
```

```
##                 [,1]          [,2]
## [1,]  3.804083e-05 -1.128310e-05
## [2,] -1.128310e-05  4.202843e-05
```

```r
fx_var1.fit$Sigma / Gamma_0
```

```
##           lr.Eu     lr.Ja
## lr.Eu 0.9989932 0.9987354
## lr.Ja 0.9987354 0.9991455
```

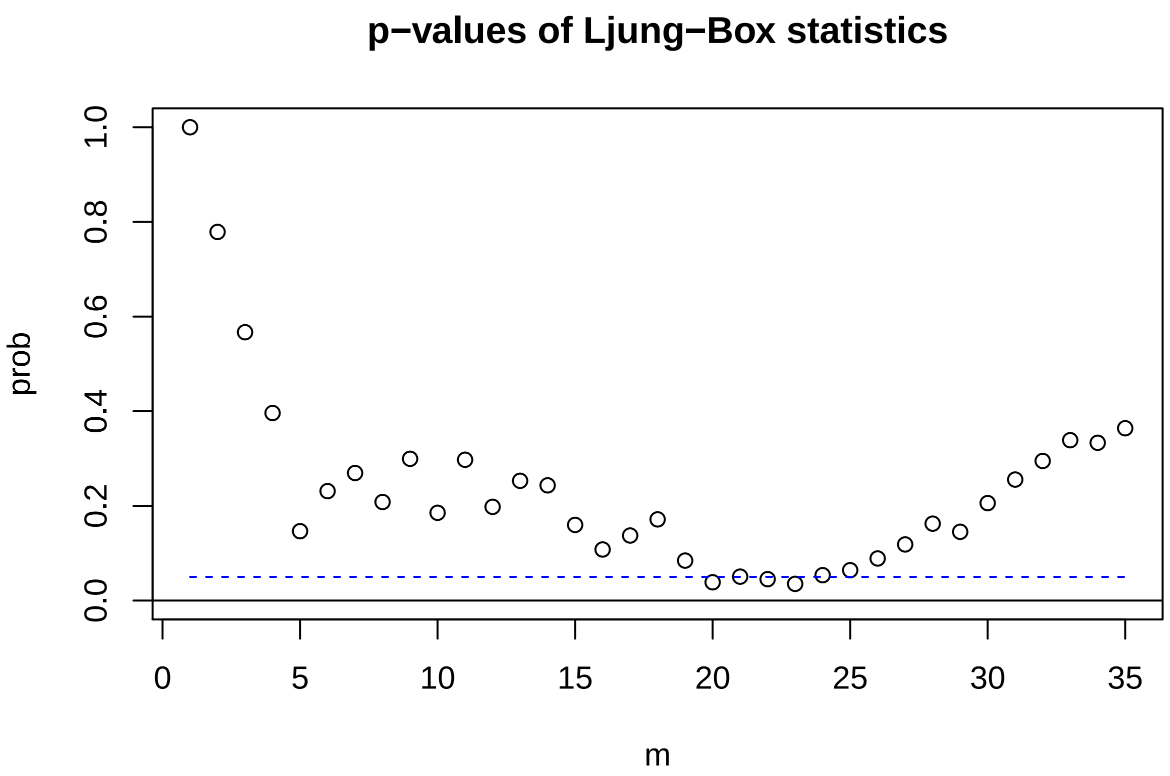
No important difference. here is almost no variation taken away by the VAR(1).

d.) Apply the Ljung-Box test on the residuals. Do the results surprise you?

```r
mq(x = fx_var1.fit$residuals, lag = 35, adj = 2^2 * 1)
```

```
## Ljung-Box Statistics:
##              m       Q(m)       df    p-value
##  [1,] 1.00e+00  3.78e-03 0.00e+00      1.00
##  [2,] 2.00e+00  1.77e+00 4.00e+00      0.78
##  [3,] 3.00e+00  6.72e+00 8.00e+00      0.57
##  [4,] 4.00e+00  1.26e+01 1.20e+01      0.40
##  [5,] 5.00e+00  2.19e+01 1.60e+01      0.15
##  [6,] 6.00e+00  2.43e+01 2.00e+01      0.23
##  [7,] 7.00e+00  2.78e+01 2.40e+01      0.27
##  [8,] 8.00e+00  3.38e+01 2.80e+01      0.21
##  [9,] 9.00e+00  3.57e+01 3.20e+01      0.30
## [10,] 1.00e+01  4.34e+01 3.60e+01      0.19
## [11,] 1.10e+01  4.42e+01 4.00e+01      0.30
## [12,] 1.20e+01  5.17e+01 4.40e+01      0.20
## [13,] 1.30e+01  5.41e+01 4.80e+01      0.25
## [14,] 1.40e+01  5.87e+01 5.20e+01      0.24
## [15,] 1.50e+01  6.65e+01 5.60e+01      0.16
## [16,] 1.60e+01  7.38e+01 6.00e+01      0.11
## [17,] 1.70e+01  7.64e+01 6.40e+01      0.14
## [18,] 1.80e+01  7.89e+01 6.80e+01      0.17
## [19,] 1.90e+01  8.90e+01 7.20e+01      0.08
## [20,] 2.00e+01  9.91e+01 7.60e+01      0.04
## [21,] 2.10e+01  1.02e+02 8.00e+01      0.05
```

```
## [22,] 2.20e+01  1.07e+02 8.40e+01     0.05
## [23,] 2.30e+01  1.13e+02 8.80e+01     0.04
## [24,] 2.40e+01  1.15e+02 9.20e+01     0.05
## [25,] 2.50e+01  1.18e+02 9.60e+01     0.06
## [26,] 2.60e+01  1.20e+02 1.00e+02     0.09
## [27,] 2.70e+01  1.21e+02 1.04e+02     0.12
## [28,] 2.80e+01  1.22e+02 1.08e+02     0.16
## [29,] 2.90e+01  1.28e+02 1.12e+02     0.15
## [30,] 3.00e+01  1.28e+02 1.16e+02     0.21
## [31,] 3.10e+01  1.30e+02 1.20e+02     0.26
## [32,] 3.20e+01  1.32e+02 1.24e+02     0.29
## [33,] 3.30e+01  1.34e+02 1.28e+02     0.34
## [34,] 3.40e+01  1.38e+02 1.32e+02     0.33
## [35,] 3.50e+01  1.41e+02 1.36e+02     0.36
```



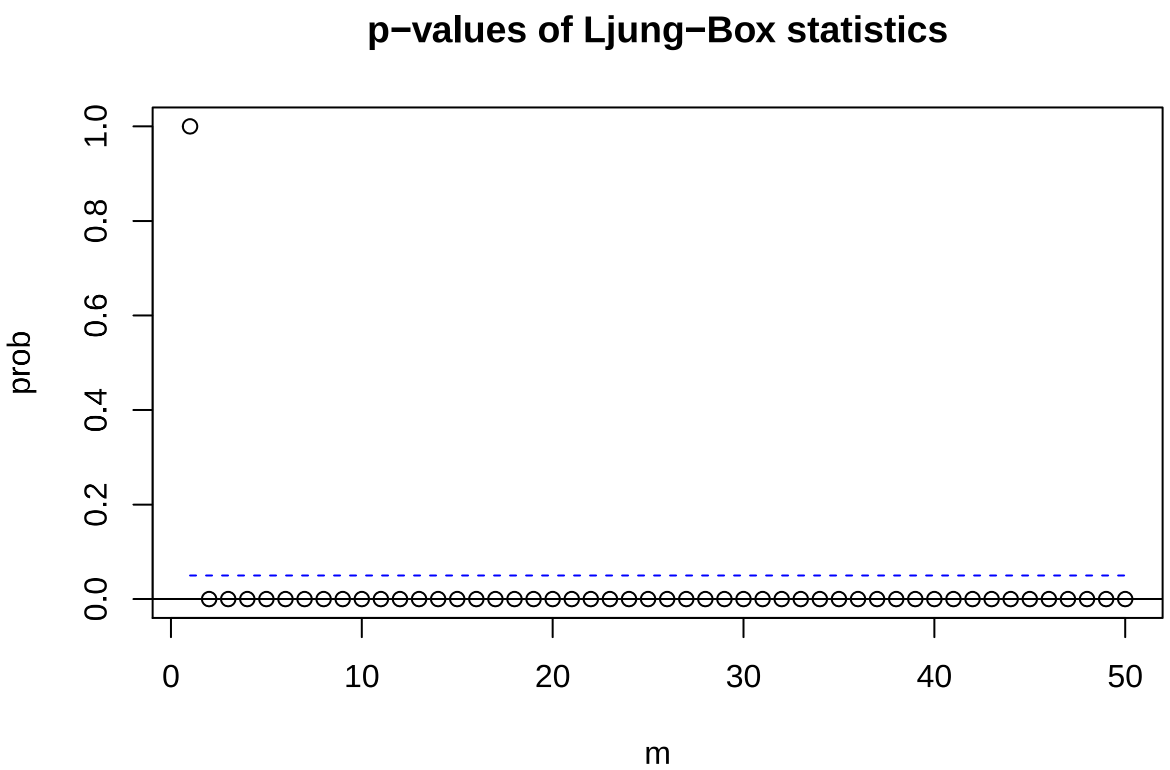
No. c) has shown that nothing has changed at all.

e.) Repeat the Ljung-Box test but this time with the squared residuals. Also have a look at the information criteria.

```r
mq(x = fx_var1.fit$residuals^2, lag = 50, adj = 2^2 * 1)
```

```
## Ljung-Box Statistics:
##            m       Q(m)      df     p-value
##   [1,]     1.0      67.4     0.0          1
##   [2,]     2.0     236.3     4.0          0
##   [3,]     3.0     272.9     8.0          0
##   [4,]     4.0     355.4    12.0          0
##   [5,]     5.0     401.8    16.0          0
##   [6,]     6.0     486.6    20.0          0
##   [7,]     7.0     546.6    24.0          0
##   [8,]     8.0     602.2    28.0          0
##   [9,]     9.0     701.2    32.0          0
##  [10,]    10.0     744.7    36.0          0
##  [11,]    11.0     872.9    40.0          0
##  [12,]    12.0     933.1    44.0          0
##  [13,]    13.0    1085.4    48.0          0
##  [14,]    14.0    1130.8    52.0          0
##  [15,]    15.0    1215.3    56.0          0
##  [16,]    16.0    1271.8    60.0          0
##  [17,]    17.0    1329.8    64.0          0
##  [18,]    18.0    1384.4    68.0          0
##  [19,]    19.0    1414.6    72.0          0
##  [20,]    20.0    1515.5    76.0          0
##  [21,]    21.0    1559.3    80.0          0
##  [22,]    22.0    1665.3    84.0          0
##  [23,]    23.0    1709.9    88.0          0
##  [24,]    24.0    1786.0    92.0          0
##  [25,]    25.0    1823.2    96.0          0
##  [26,]    26.0    1864.1   100.0          0
##  [27,]    27.0    1954.8   104.0          0
##  [28,]    28.0    1991.2   108.0          0
##  [29,]    29.0    2168.8   112.0          0
##  [30,]    30.0    2197.2   116.0          0
##  [31,]    31.0    2236.4   120.0          0
##  [32,]    32.0    2255.5   124.0          0
##  [33,]    33.0    2341.5   128.0          0
##  [34,]    34.0    2379.0   132.0          0
##  [35,]    35.0    2424.5   136.0          0
##  [36,]    36.0    2509.7   140.0          0
```

```
## [37,]   37.0   2581.5   144.0        0
## [38,]   38.0   2640.0   148.0        0
## [39,]   39.0   2655.7   152.0        0
## [40,]   40.0   2739.5   156.0        0
## [41,]   41.0   2784.3   160.0        0
## [42,]   42.0   2825.4   164.0        0
## [43,]   43.0   2835.0   168.0        0
## [44,]   44.0   2890.5   172.0        0
## [45,]   45.0   2921.3   176.0        0
## [46,]   46.0   2949.7   180.0        0
## [47,]   47.0   3034.6   184.0        0
## [48,]   48.0   3053.4   188.0        0
## [49,]   49.0   3180.2   192.0        0
## [50,]   50.0   3219.8   196.0        0
```

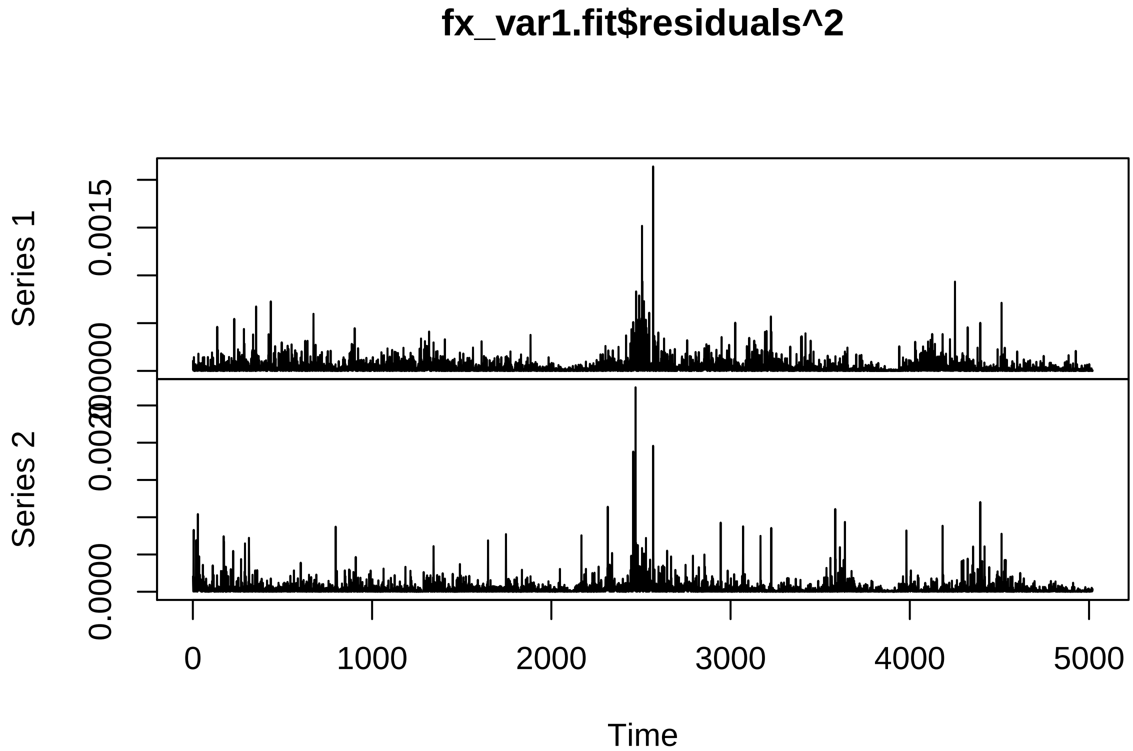



```
VARorder(x = fx_var1.fit$residuals^2, maxp = 20)
```

```
## selected order: aic =  20
## selected order: bic =  13
## selected order: hq =  13
## Summary table:
```

```
##        p       AIC      BIC       HQ     M(p) p-value
##  [1,]  0 -37.4176 -37.4176 -37.4176   0.0000  0.0000
##  [2,]  1 -37.4290 -37.4238 -37.4272  64.6885  0.0000
##  [3,]  2 -37.4592 -37.4488 -37.4556 158.9360  0.0000
##  [4,]  3 -37.4619 -37.4463 -37.4564  21.3438  0.0003
##  [5,]  4 -37.4713 -37.4505 -37.4640  54.6839  0.0000
##  [6,]  5 -37.4750 -37.4490 -37.4659  26.6857  0.0000
##  [7,]  6 -37.4840 -37.4528 -37.4730  52.5478  0.0000
##  [8,]  7 -37.4891 -37.4527 -37.4764  33.6938  0.0000
##  [9,]  8 -37.4923 -37.4507 -37.4777  23.6494  0.0001
## [10,]  9 -37.5008 -37.4540 -37.4844  50.4637  0.0000
## [11,] 10 -37.5012 -37.4492 -37.4830   9.8909  0.0423
## [12,] 11 -37.5119 -37.4548 -37.4919  61.2238  0.0000
## [13,] 12 -37.5143 -37.4519 -37.4925  19.8154  0.0005
## [14,] 13 -37.5278 -37.4602 -37.5041  75.0385  0.0000
## [15,] 14 -37.5287 -37.4559 -37.5032  12.2053  0.0159
## [16,] 15 -37.5307 -37.4527 -37.5034  17.9167  0.0013
## [17,] 16 -37.5310 -37.4479 -37.5019   9.7336  0.0452
## [18,] 17 -37.5322 -37.4438 -37.5012  13.4969  0.0091
## [19,] 18 -37.5317 -37.4381 -37.4989   5.3946  0.2491
## [20,] 19 -37.5325 -37.4338 -37.4979  12.0269  0.0172
## [21,] 20 -37.5391 -37.4352 -37.5027  40.5684  0.0000
```

```r
plot.ts(fx_var1.fit$residuals^2)
```



Plenty of lagged cross-/auto correlations. Information criteria suggest high lag orders. Fitting a VAR appears sensible.

f.) Can you rule out weak stationarity for the growth rates of the exchange rates only based on your findings up to this point?

No. Weak stationarity is about time-invariance regarding the *unconditional* expecation and variance. Heteroscedasticity can be a violation of weak stationarity, simply because the variance is not constant over time. But if we are able to sufficiently model the variation of the variance using a VAR, we are in fact facing *conditional* heteroscedasticity: $E(a_t^2) = \sigma_t * \epsilon_t$ with ϵ_t as white noise. And as we learned before, a stable VAR with white noise-innovations yields a stationary time series. This is analogous to a stationary VAR, where the conditional expectation of the observations may differ from the unconditional expectation (depending on past observations).