Macroeconometrics - 2168 Joao B. Duarte
Nova SBE, Spring 2020

Problem Set 2 - SOLUTIONS

VAR and Nonstationary Models

Problem 1 Suppose we have the following model:

$$
\begin{aligned}
y_t &= 0.5 + 0.1z_t + 0.3y_{t-1} + 0.3z_{t-1} + \varepsilon_{yt} \\
z_t &= 1 - 0.2y_t + 0.1y_{t-1} + 0.4z_{t-1} + \varepsilon_{zt}
\end{aligned}
$$

a) Show the model in a matrix representation whereby you clearly show the elements of each matrix (Γ_0, Γ_1 and B). Show the reduced form model and find its coefficients (A_0 and A_1). Is the model stationary (you can use R or do it by hand following the lecture notes)?

$$
\begin{bmatrix} 1 & -0.1 \\ 0.2 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.3 & 0.3 \\ 0.1 & 0.4 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}
$$

The reduced form model is given by:

$$
x_t = A_0 + A_1 x_{t-1} + e_t
$$

where $A_0 = B^{-1}\Gamma_0$, $A_1 = B^{-1}\Gamma_1$ and $e_t = B^{-1}\varepsilon_t$.

$$
A_0 = \begin{bmatrix} 0.9803922 & 0.09803922 \\ -0.1960784 & 0.98039216 \end{bmatrix} \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5882353 \\ 0.8823529 \end{bmatrix}
$$

$$
A_1 = \begin{bmatrix} 0.9803922 & 0.09803922 \\ -0.1960784 & 0.98039216 \end{bmatrix} \begin{bmatrix} 0.3 & 0.3 \\ 0.1 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.30392157 & 0.3333333 \\ 0.03921569 & 0.3333333 \end{bmatrix}
$$

The eigenvalues of A_1 are: 0.4339017 and 0.2033532. Since all of the eigenvalues are inside the unit circle, the model is stationary.

b) Do we need to identify the model? Compute the responses of y_t and y_{t+1} and to a unit shock in z_t. Interpret them.

Since we are given the structural model, the latter is already identified. The IRF is given by $\phi_j = A_1^j B^{-1}$, whereby the shocks are ordered by columns and variables by rows.

$$
\phi_0 = \begin{bmatrix} 0.9803922 & 0.09803922 \\ -0.1960784 & 0.98039216 \end{bmatrix}
$$

The response of y_t to a unit shock in z_t is then $\phi_0(1,1) = 0.09803922$.

$$\phi_1 = \begin{bmatrix} 0.30392157 & 0.3333333 \\ 0.03921569 & 0.3333333 \end{bmatrix} \begin{bmatrix} 0.9803922 & 0.09803922 \\ -0.1960784 & 0.98039216 \end{bmatrix} = \begin{bmatrix} 0.23260285 & 0.3565936 \\ -0.02691273 & 0.3306421 \end{bmatrix}$$

The response of y_{t+1} to a unit shock in z_t is then $\phi_1(1,1) = 0.3565936$. A unit shock in z_t increases y on impact approximately by 0.1, and 0.36 next period. So the effect rises first. Then, since the model is stationary we know that the effects will eventually fade away, which means the effect of z_t is U-shaped.

c) Suppose you assume a recursive identification where y_t is ordered first. What are the reduced form parameters now? Recalculate the responses of y_t and y_{t+1} to a shock in z_t. How different are they from b)? Explain.

With the recursive identification we have that:

$$B = \begin{bmatrix} 1 & 0 \\ 0.2 & 1 \end{bmatrix}$$

so

$$B^{-1} = \begin{bmatrix} 1 & 0 \\ -0.2 & 1 \end{bmatrix}$$

This changes the reduces form parameters:

$$A_0 = \begin{bmatrix} 1 & 0 \\ -0.2 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.9 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 1 & 0 \\ -0.2 & 1 \end{bmatrix} \begin{bmatrix} 0.3 & 0.3 \\ 0.1 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.3 \\ 0.04 & 0.34 \end{bmatrix}$$

The impulse responses assuming the recursive identification are then:

$$\phi_0' = \begin{bmatrix} 1 & 0 \\ -0.2 & 1 \end{bmatrix}$$

and

$$\phi_1' = \begin{bmatrix} 0.24 & 0.3 \\ -0.028 & 0.34 \end{bmatrix}$$

A unit shock in z_t increases y on impact is now 0 by construction instead of 0.1, and it is 0.3 next period instead of 0.36. Hence, the recursive assumption would strongly underestimate the contemporaneous response of y to a shock in z and would underestimate slightly the response in one period.

d) Let $x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}$. Use R to save matrix B^{-1} and A_1. We know the IRF for step j is given by $\frac{\partial x_{t+j}}{\partial \varepsilon_t} = A_1^j B^{-1}$. Compute the IRFs for 30 steps using this formula in R and plot them (note that in R matrix multiplication of say A with B is done by A%*%B). Do they converge to zero? Explain.

Problem 2

a) Can we apply OLS to the structural VAR? Why or why not?

We cannot apply OLS estimation to the structural model because the estimates would be biased due to endogeneity.

b) Can we interpret the reduced form errors of the VAR? Explain.

We cannot interpret the reduced form errors because they are a linear combination of the structural shocks $e_t = B^{-1}\varepsilon_t$*. Hence, they are a linear combination of surprises in the structural variables, which makes their interpretation impossible without identification.*

c) Can we use VAR models to infer causal relationships? Explain.

VAR models can be used to make statements on causal relationships. These can be done with Granger Causality tests, IRFs and FEVD. However, in terms of IRFs and FEVD, the quality of the causal inference depends directly on the quality of the identification. If the identification weak, we are likely have erroneous conclusions on causal relationships.

Problem 3 Given an initial condition y_0, consider the two following processes:

$$\begin{aligned} I) &\qquad y_t = 0.7y_{t-1} + \varepsilon_t \\ II) &\qquad y_t = 2 + y_{t-1} + 0.5t + \varepsilon_t \end{aligned}$$

a) Find the solution of y_t iterating forward for both I) and II).

I)

$$y_1 = 0.7y_0 + \varepsilon_1$$
$$y_2 = 0.7y_1 + \varepsilon_2$$
$$y_2 = 0.7(0.7y_0 + \varepsilon_1) + \varepsilon_2$$
$$y_2 = 0.7^2 y_0 + 0.7\varepsilon_1 + \varepsilon_2$$
$$y_t = 0.7^t y_0 + \sum_{0}^{t-1} 0.7^i \varepsilon_{t-1}$$

II)

$$y_1 = 2 + y_0 + 0.5 + \varepsilon_1$$
$$y_2 = 2 + y_1 + 0.5 * 2 + \varepsilon_2$$
$$y_2 = 2 + 2 + y_0 + 0.5 + \varepsilon_1 + 0.5 * 2 + \varepsilon_2$$
$$y_t = 2t + y_0 + 0.5\sum_{i=1}^{t} i + \sum_{i=1}^{t} \varepsilon_i$$

b) What transformations of the series should you apply to I) and II) to make them stationary?
I) no transformation. The series is already stationary since the root is inside the unit circle.
II) first take first differences and get:

$$\Delta y_t = 2 + 0.5t + \varepsilon_t$$

Second, detrend the first difference series by running a OLS of Δy_t on a linear trend, and subtract the predicted value from Δy_t.

c) Suppose you take the first differences in II). Is the resulting series stationary? Explain.
No, as explained in b) the series will still exhibit a deterministic trend.

Problem 4 For this problem, we will make use of the data set included in the *ps_2_data.xls*. The data set consists of two series: U.S monthly unemployment rate (unrate) and Average Hourly Earnings (wage) from first month of 1964 to February of 2020. The data are from the Federal Reserve Bank of St. Louis.

a) Test each series for stationarity using the ADF test.

b) Estimate VAR model of the two series in levels selecting the number of lags using the AIC criteria (remember that you should only use the columns that have data and not the column with the dates). Check the residuals of the VAR, are they stationary?

c) Compute and plot the IRF. You will notice that the dimension of the wage raw data does not help with interpretation. Lets transform that series into log so that the IRF are in percentage. Take log of wage and multiply it by 100. Re-estimate the VAR and plot the IRF. Interpret your results.

d) Change the order of the variables. Estimate a VAR model by ordering wage first and plot the IRF. Does your conclusions in b) change?

e) Perform a forecast error variance decomposition. Explain your findings.

f) Does unemployment rate Granger cause wages?

PS2_2020_solutions

May 6, 2020

1 Problem Set 2 Solutions

1.1 Question 1

1.1.1 a)

[9]:
```
# install.packages('expm')
library(expm)

# Define B and Gamma_1 matrices
B = matrix(c(1,-0.1, 0.2, 1), nrow = 2, byrow = TRUE)
gamma_1 = matrix(c(0.3,0.3, 0.1, 0.4), nrow = 2, byrow = TRUE)
gamma_0 = matrix(c(0.5, 1), nrow = 2)

# Invert B and find A_1
B_inv = solve(B)
A_0 = B_inv %*% gamma_0
A_1 = B_inv %*% gamma_1

# Results
print(A_0)
print(A_1)

# Stationarity, both eigenvalues are in modulus below one
eigen(A_1)
```

```
          [,1]
[1,] 0.5882353
[2,] 0.8823529
           [,1]      [,2]
[1,] 0.30392157 0.3333333
[2,] 0.03921569 0.3333333
```

```
eigen() decomposition
$values
[1] 0.4339017 0.2033532

$vectors
```

```
          [,1]       [,2]
[1,] -0.9316733 -0.9573758
[2,] -0.3632971  0.2888452
```

1.1.2 b)

[11]:
```
# Calculating the IRF of y in the first period:
res = A_1 %*% B_inv

print(B_inv[1,2])
print(res[1,2])
```

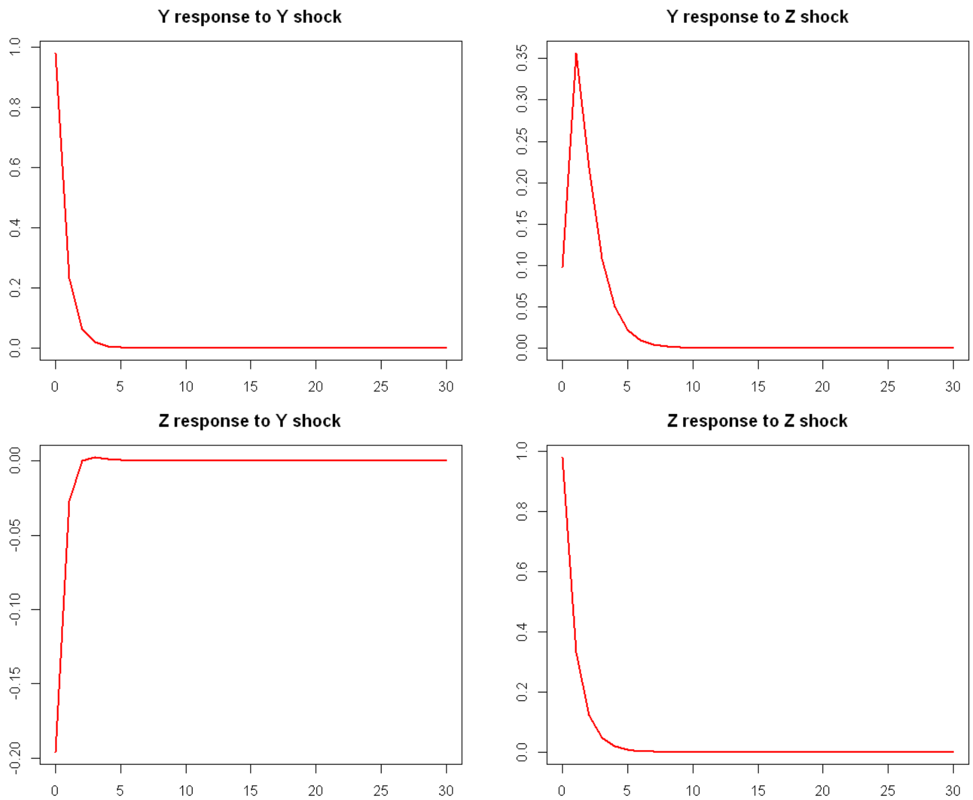
```
[1] 0.09803922
[1] 0.3565936
```

1.1.3 d)

[3]:
```
# Set parameters
steps = 30
n_var = 2
n_shocks = 2

# Build empty array to save IRFS
IRFS = array(dim=c(n_var,n_shocks,steps + 1))

for (j in 0:steps)
{
  IRFS[ , , j + 1] = A_1%^%j %*% B_inv
}

var_names = c("Y", "Z")

options(repr.plot.width=10, repr.plot.height=8)
par(mfrow=c(2,2), oma = c(1,1,0,0),
    mar = c(0,0,1,1) + 2)

for (i in 1:2)
{   for (j in 1:2)
    {
      plot(0:steps, IRFS[i,j,],
      xlab = "Steps",
      ylab = "",
      main = paste(var_names[i], "response to", var_names[j], "shock"),
```

```
      type = "l",
      lwd = 2,
      col = "red")
    }
}
```

```
package 'expm' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\User\AppData\Local\Temp\Rtmp6x5XQu\downloaded_packages

Warning message:
"package 'expm' was built under R version 3.6.3"Loading required package: Matrix

Attaching package: 'expm'

The following object is masked from 'package:Matrix':

    expm


           [,1]      [,2]
[1,] 0.30392157 0.3333333
[2,] 0.03921569 0.3333333
```



1.2 Question 4

[35]:
```r
library(readxl)

data = read_xls("../../Data/ps_2_data.xls")
head(data)
```

A tibble: 6 × 3

observation_date	unrate	wage
<dttm>	<dbl>	<dbl>
1964-01-01	5.6	2.50
1964-02-01	5.4	2.50
1964-03-01	5.4	2.51
1964-04-01	5.3	2.52
1964-05-01	5.1	2.52
1964-06-01	5.2	2.53

1.2.1 a)

[16]:
```
library(tseries)
unrate = ts(data[,2], start = 1964, frequency = 12)
wage = ts(data[,3], start = 1964, frequency = 12)
adf.test(unrate)
adf.test(wage)
```

```
Augmented Dickey-Fuller Test

data:  unrate
Dickey-Fuller = -3.1415, Lag order = 8, p-value = 0.09801
alternative hypothesis: stationary



Augmented Dickey-Fuller Test

data:  wage
Dickey-Fuller = -0.94687, Lag order = 8, p-value = 0.9473
alternative hypothesis: stationary
```

We reject that the unemployment rate is not stationary at 10% confidence level. But we do not reject that wage is not stationarity at 10% confidence level.

1.2.2 b)

[17]:
```
library(vars)

var = VAR(data[,2:3], lag.max=15, ic="AIC")
resid = residuals(var)
adf.test(resid[,1])
adf.test(resid[,2])
var
```

```
Warning message in adf.test(resid[, 1]):
"p-value smaller than printed p-value"
```

```
Augmented Dickey-Fuller Test

data:  resid[, 1]
Dickey-Fuller = -8.1468, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary
```

```
Warning message in adf.test(resid[, 2]):
"p-value smaller than printed p-value"
```

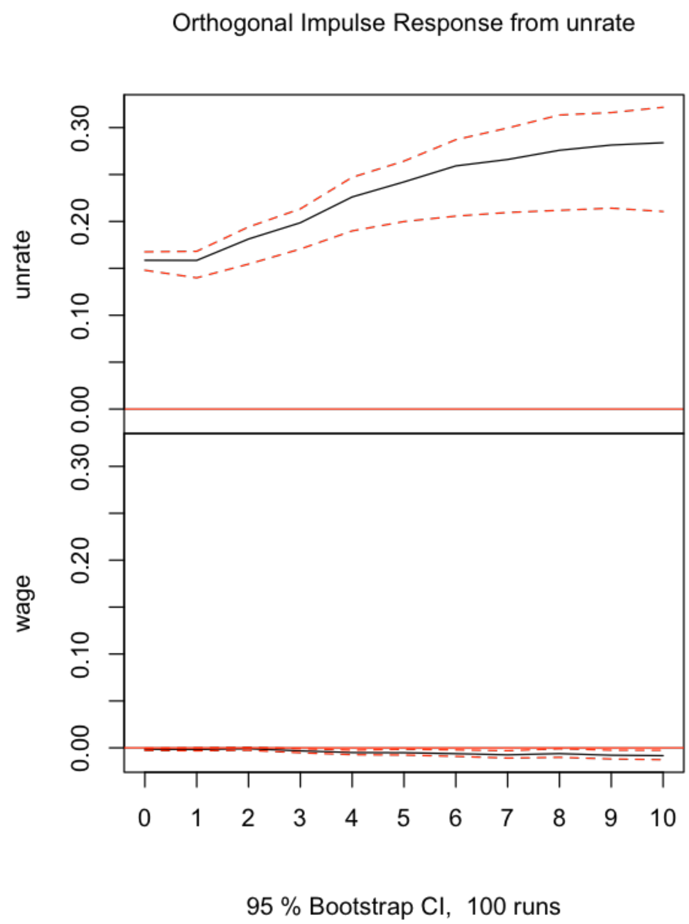
```
Augmented Dickey-Fuller Test

data:  resid[, 2]
Dickey-Fuller = -8.9356, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary
```

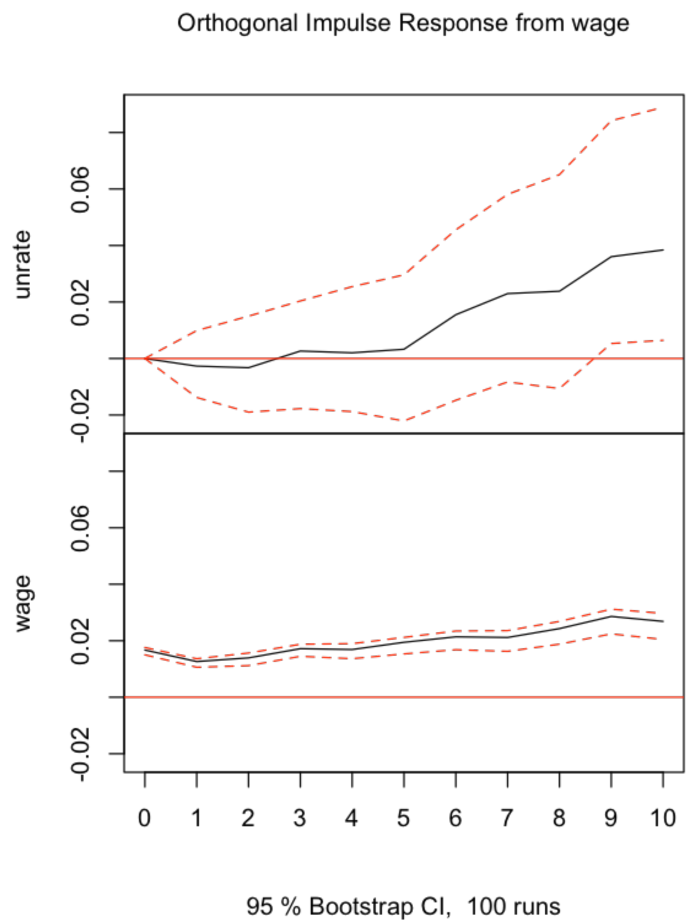
```
VAR Estimation Results:
=======================

Estimated coefficients for equation unrate:
===========================================
Call:
unrate = unrate.l1 + wage.l1 + unrate.l2 + wage.l2 + unrate.l3 + wage.l3 + unrate.l4 + wage.l4

   unrate.l1      wage.l1    unrate.l2      wage.l2    unrate.l3      wage.l3 
 0.997715160 -0.161609462  0.145098129  0.087422518 -0.028651446  0.445054900 
   unrate.l4      wage.l4    unrate.l5      wage.l5    unrate.l6      wage.l6 
 0.038682675 -0.256108281 -0.087279762 -0.050692008 -0.006612815  0.606033550 
   unrate.l7      wage.l7    unrate.l8      wage.l8    unrate.l9      wage.l9 
-0.064995910 -0.096719950 -0.003116134 -0.624989969  0.007338920  0.386963642 
  unrate.l10     wage.l10        const 
-0.008657744 -0.338593054  0.039556636 


Estimated coefficients for equation wage:
=========================================
Call:
wage = unrate.l1 + wage.l1 + unrate.l2 + wage.l2 + unrate.l3 + wage.l3 + unrate.l4 + wage.l4 +

    unrate.l1       wage.l1     unrate.l2       wage.l2     unrate.l3 
-0.0038732148  0.7562178394  0.0088196455  0.2608343374 -0.0151680427 
      wage.l3     unrate.l4       wage.l4     unrate.l5       wage.l5 
 0.2016551242 -0.0038714621 -0.1372962918  0.0124597715  0.0625317685 
    unrate.l6       wage.l6     unrate.l7       wage.l7     unrate.l8 
 0.0019741778 -0.0009406821 -0.0026475831 -0.1122385961  0.0152924751 
      wage.l8     unrate.l9       wage.l9    unrate.l10      wage.l10 
 0.0868318172 -0.0104657897  0.1572949253 -0.0036178871 -0.2745880342 
        const 
 0.0106040129 
```

We reject that both residuals are nonstarionary at the 5% confidence level.

[18]:
```
irfs_level = irf(var)
```

[19]:
```
plot(irfs_level)
```

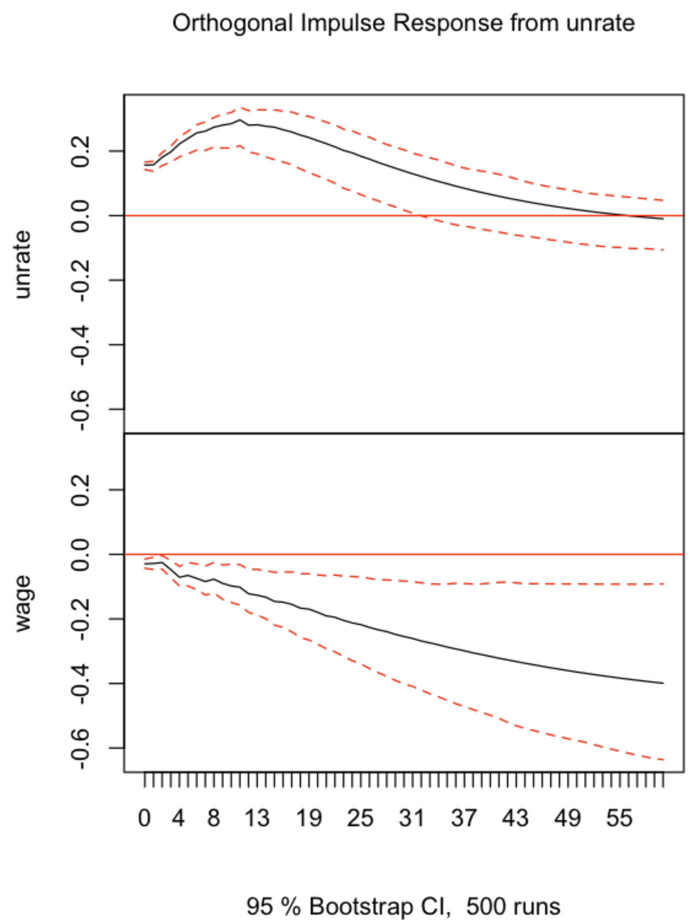


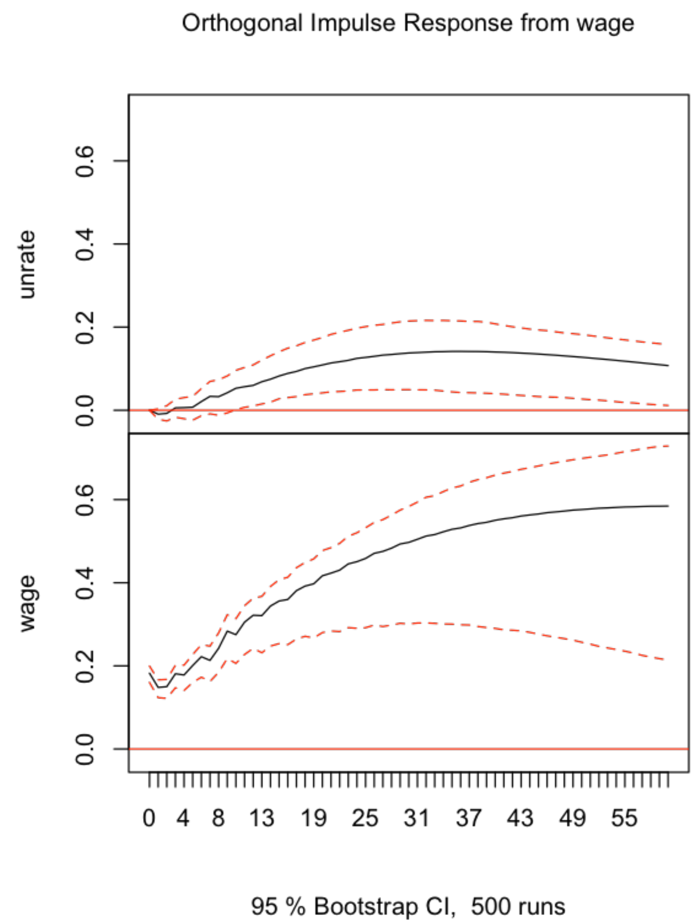


[20]:
```
data[,3] = log(data[,3])*100
```

[21]:
```
var2 = VAR(data[,2:3], lag.max = 15, ic = 'AIC')
irfs_log_level = irf(var2, runs = 500, n.ahead = 60)
```

[22]:
```
plot(irfs_log_level)
```


Orthogonal Impulse Response from unrate
unrate
wage
0.2
0.0
-0.2
-0.4
-0.6
0 4 8 13 19 25 31 37 43 49 55
95 % Bootstrap CI, 500 runs



[23]:
```
irfs_log_level$irf$unrate[1,]
```

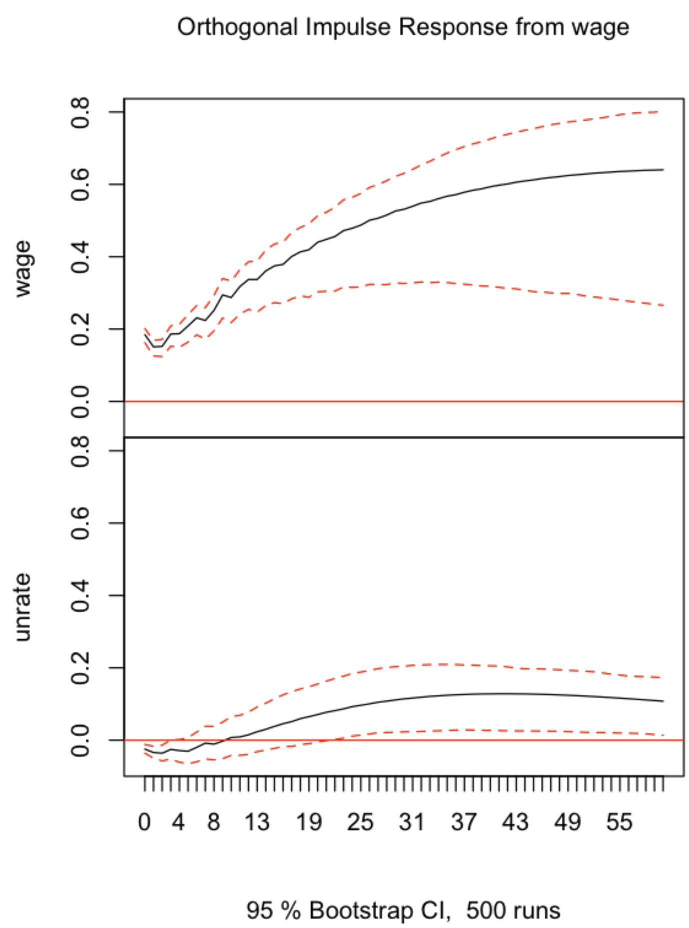
unrate 0.15654244671866 **wage** -0.0292272374389841

An average positive shock of 0.156% in unemployment rate generates a drop of -0.03% in wages on impact, and of -0.4% in five years. This corraborates economic theory that shows how excess supply in the labor market today is associated with lower wages in equilibrium next periods. The IRFs also show that an unexpexted increase in wages of 0.2% is associated with higher unemployment two years into the future. If labor markets are in equilibrium (i.e. no excess supply nor demand), then higher wages will generate excess supply in the future which is consisten with the IRF evidence.

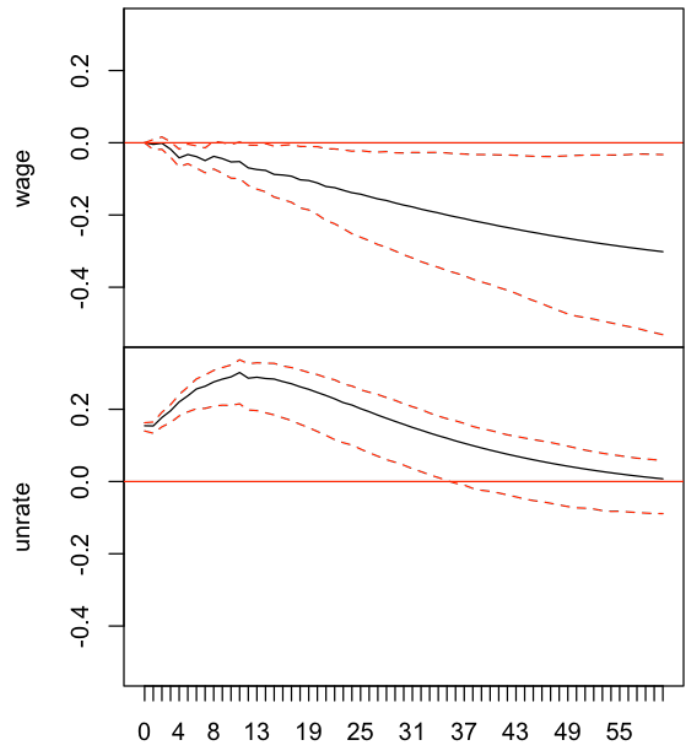
1.2.3 d)

[24]:
```
data2 = data[,c(3,2)]
```

[25]:
```
var3 = VAR(data2, lag.max = 15, ic = 'AIC')
irfs_log_level = irf(var3, runs = 500, n.ahead = 60)
plot(irfs_log_level)
```



Orthogonal Impulse Response from unrate



95 % Bootstrap CI, 500 runs

The mains conclusions remain unaffected by the reordoring the variables which indicates that they do not depend on the identification assumption.

1.2.4 e)

[26]:
```
fevd(var2, n.ahead = 60)
```

```
$unrate
         unrate        wage
 [1,] 1.0000000 0.000000000
 [2,] 0.9982164 0.001783562
 [3,] 0.9982282 0.001771819
```

```
 [4,] 0.9985318 0.001468155
 [5,] 0.9987241 0.001275861
 [6,] 0.9988099 0.001190065
 [7,] 0.9975941 0.002405872
 [8,] 0.9949811 0.005018915
 [9,] 0.9934217 0.006578340
[10,] 0.9910307 0.008969348
[11,] 0.9876385 0.012361474
[12,] 0.9846467 0.015353315
[13,] 0.9816652 0.018334754
[14,] 0.9779855 0.022014457
[15,] 0.9739937 0.026006312
[16,] 0.9694473 0.030552709
[17,] 0.9645078 0.035492184
[18,] 0.9593994 0.040600605
[19,] 0.9537255 0.046274473
[20,] 0.9479045 0.052095542
[21,] 0.9418272 0.058172807
[22,] 0.9354284 0.064571598
[23,] 0.9289210 0.071079045
[24,] 0.9222196 0.077780388
[25,] 0.9151553 0.084844712
[26,] 0.9080159 0.091984130
[27,] 0.9007487 0.099251328
[28,] 0.8932860 0.106713971
[29,] 0.8858036 0.114196381
[30,] 0.8782382 0.121761827
[31,] 0.8706122 0.129387760
[32,] 0.8630090 0.136991019
[33,] 0.8554086 0.144591368
[34,] 0.8478114 0.152188608
[35,] 0.8402935 0.159706494
[36,] 0.8328214 0.167178649
[37,] 0.8254385 0.174561491
[38,] 0.8181788 0.181821171
[39,] 0.8110190 0.188980989
[40,] 0.8039937 0.196006279
[41,] 0.7971258 0.202874179
[42,] 0.7904007 0.209599312
[43,] 0.7838424 0.216157609
[44,] 0.7774645 0.222535501
[45,] 0.7712535 0.228746509
[46,] 0.7652326 0.234767429
[47,] 0.7594031 0.240596940
[48,] 0.7537592 0.246240811
[49,] 0.7483127 0.251687304
[50,] 0.7430624 0.256937600
[51,] 0.7380033 0.261996713
```

```
[52,] 0.7331434 0.266856604
[53,] 0.7284770 0.271523048
[54,] 0.7240004 0.275999643
[55,] 0.7197169 0.280283055
[56,] 0.7156197 0.284380294
[57,] 0.7117056 0.288294426
[58,] 0.7079740 0.292026023
[59,] 0.7044178 0.295582183
[60,] 0.7010335 0.298966546

$wage
          unrate      wage
 [1,] 0.02503058 0.9749694
 [2,] 0.02890487 0.9710951
 [3,] 0.02851231 0.9714877
 [4,] 0.03910311 0.9608969
 [5,] 0.06277157 0.9372284
 [6,] 0.07010377 0.9298962
 [7,] 0.07683710 0.9231629
 [8,] 0.08686489 0.9131351
 [9,] 0.08766762 0.9123324
[10,] 0.08835942 0.9116406
[11,] 0.09211408 0.9078859
[12,] 0.09334561 0.9066544
[13,] 0.09833182 0.9016682
[14,] 0.10321135 0.8967886
[15,] 0.10669165 0.8933083
[16,] 0.11137745 0.8886225
[17,] 0.11522805 0.8847719
[18,] 0.11823928 0.8817607
[19,] 0.12200621 0.8779938
[20,] 0.12522347 0.8747765
[21,] 0.12834095 0.8716590
[22,] 0.13214683 0.8678532
[23,] 0.13553770 0.8644623
[24,] 0.13893268 0.8610673
[25,] 0.14247888 0.8575211
[26,] 0.14576173 0.8542383
[27,] 0.14901321 0.8509868
[28,] 0.15236790 0.8476321
[29,] 0.15553907 0.8444609
[30,] 0.15873002 0.8412700
[31,] 0.16197033 0.8380297
[32,] 0.16508934 0.8349107
[33,] 0.16821968 0.8317803
[34,] 0.17134078 0.8286592
[35,] 0.17436490 0.8256351
[36,] 0.17738879 0.8226112
```

```
[37,] 0.18040117 0.8195988
[38,] 0.18333139 0.8166686
[39,] 0.18626521 0.8137348
[40,] 0.18916822 0.8108318
[41,] 0.19200958 0.8079904
[42,] 0.19484576 0.8051542
[43,] 0.19764392 0.8023561
[44,] 0.20039103 0.7996090
[45,] 0.20312659 0.7968734
[46,] 0.20582202 0.7941780
[47,] 0.20847573 0.7915243
[48,] 0.21111268 0.7888873
[49,] 0.21370743 0.7862926
[50,] 0.21626732 0.7837327
[51,] 0.21880450 0.7811955
[52,] 0.22130177 0.7786982
[53,] 0.22376712 0.7762329
[54,] 0.22620656 0.7737934
[55,] 0.22860805 0.7713919
[56,] 0.23097977 0.7690202
[57,] 0.23332327 0.7666767
[58,] 0.23563100 0.7643690
[59,] 0.23790987 0.7620901
[60,] 0.24015927 0.7598407
```

The unemployment rate is mainly explained by itself in the short run, but in the long run (5 years) 30% of its variation is explained by variation in wages.

On the other hand, 9% of the variation in Wages is explained by unempoyment rate in the short run (1 year). And in the long run, unrate explains 24% of the variation in wages.

1.2.5 f)

[59]:
```
wage_d = diff(log(ts(data$wage)))
data3 = data.frame(unrate = data$unrate[-1], wage_d)
```

[61]:
```
adf.test(data3[,2])

var = VAR(data3, lag.max=15, ic="AIC")
causality(var, cause = c('unrate'))
```

```
Augmented Dickey-Fuller Test

data:  data3[, 2]
Dickey-Fuller = -3.6622, Lag order = 8, p-value = 0.02671
alternative hypothesis: stationary
```

```
$Granger

Granger causality H0: unrate do not Granger-cause wage_d

data:  VAR object var
F-Test = 2.5689, df1 = 9, df2 = 1290, p-value = 0.006247


$Instant

H0: No instantaneous causality between: unrate and wage_d

data:  VAR object var
Chi-squared = 15.149, df = 1, p-value = 9.935e-05
```

Granger-causality indicates that the unemployment rate Granger cause wages, since null hypothesis of no causality is rejected.

[ ]: