Econometrics III HW - part 2

A. Schmidt and P. Assunção

March, 2020

Assignment 2

See the source code if interested in all functions (chunks were ommited unless relevant for the assignment). Click here to access the code.

Introduction

By the beginning of 2014, in the face of fast rising unemployment, the Dutch Ministry of Social Affairs is concerned that the provisions for future government expenditure with social pensions may be severely under-estimated. This is particularly true if the economy is hit again by a large negative shock. Suppose that you have been asked to analyze alternative unemployment scenarios.

Importing and checking data

```
urlRemote  <- "https://raw.githubusercontent.com/aishameriane"
pathGithub <- "/Mphil/master/EconIII/data_assign_p2.csv"
token      <- "?token=AAVGJTRVJIKVBISR37WEY3K6RIY4E"


url        <- paste0(urlRemote, pathGithub, token)
dfData21 <- read.csv2(url, sep = ",", dec = ".", header = TRUE)
```

As in Part I, we make an initial check on the dataset. The head and tail of the imported dataset indeed show the quarterly data ranging from from the second quarter of 1987 until the first quarter of 2014.

| obs	GDP_QGR	UN_RATE	obs	GDP_QGR	UN_RATE
1987Q2	2.7404816	7.4	2012Q4	0.4536453	6.2
1987Q3	2.0275355	7.5	2013Q1	0.4088997	6.8
1987Q4	0.7983716	7.5	2013Q2	-0.1027331	7.1
1988Q1	0.9903613	7.5	2013Q3	0.3907661	7.5
1988Q2	0.3396408	7.4	2013Q4	0.2676572	7.6
1988Q3	0.5052618	7.3	2014Q1	-0.3744039	7.8

As for the descriptive statistics for GDP and unemployment, we see there is no missing information and all values are numeric (there are no problems of formatting). Furthermore, the statistics show values reasonable with the variables pourpose. That is, we do not observe unreasonable high/low levels of GDP growth and unemployment.

	Quart. GDP	Un. Rate
Observations	108.0000	108.0000
Minimum	-2.7051	3.1000
1st quartile	0.4944	4.5750
Mean	1.0670	5.6185
Median	1.0058	5.7000
3rd quartile	1.6310	6.7250
Maximum	4.0908	8.4000
Desv. Pad.	0.9739	1.4074

Question 1

Plot the sample of Dutch quarterly unemployment rates and Dutch GDP quarterly growth rates that you have at your disposal. Estimate an AR model for the GDP growth rate and an ADL model for the unemployment rate using the GDP growth rate as an exogenous explanatory variable. Please adopt a general-to-specific methodology for both models by eliminating insignificant lags. For the AR model start with four lags of GDP. For the ADL model start with four lags of each variable. Report the final estimated AR and ADL models, working at a 5% significance level. Comment on the estimated coefficients. What do these coefficients tell you about the dynamic properties of the unemployment rate and the GDP growth rate?

Dutch quarterly GDP growth and unemployment rate
1987Q2 to 2014Q1
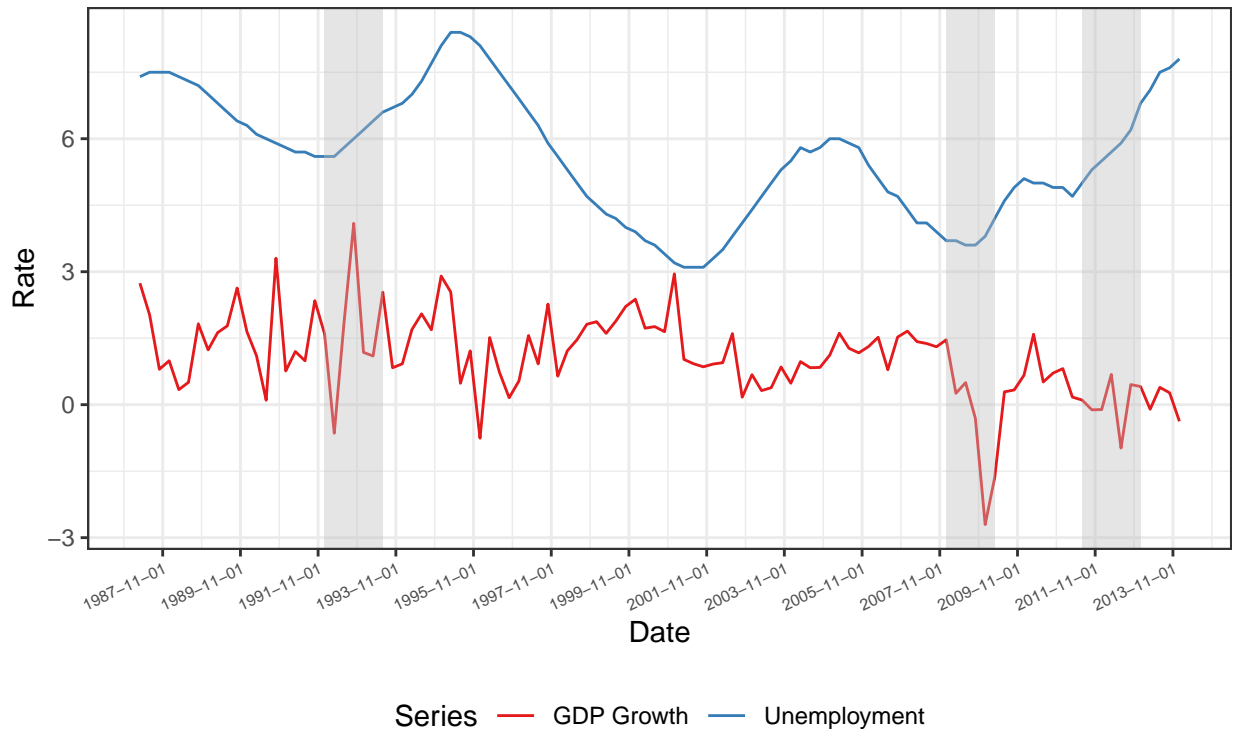
Comments on the plots: We start by analysing the GDP series. As discussed in Part I, there seems to be some change in the time series behavior after the turn of the millenium, with a large drop during the 2008 financial crisis. Now, with data up to 2014, we observe a recovery of the GDP growth to levels pre crisis around 2 years after the crisis. However, after the 2010s, GDP growth seems to be oscilating around zero (or very low) growth, with a volatily smaller than the one observed in the 90s.

2

As for the unemployment rate, we observe some sort of cycical behaviour that seems to last for (roughly) 5 years. One exception to this behaviour is around the financial crisis period. Although unemployment was indeed smaller in 2010 than in 2005, one would expect that the continuous drop from 2005 until 2010. That is, that the favorable cenario for umemployment rate, where the economoy is experiencing diminishing unemployment, would last for at least a couple more years. Instead, we observe that such drop occurs up to 2008, where the unemployment trend is reversed and the economy faces more unemployment as the crisis goes on. After the crises, we observe a mostly increasing unemployment trend until the end of the data sample, with a short drop after 2010.

The gray areas in the graph represents the recessions in the Euro Area for the period (link for the source: Euro Area Business Cycle Network: https://eabcn.org/dc/chronology-euro-area-business-cycles). Notice that indeed the Dutch GDP growth rate was in decline during the last two periods flagges as recessions, but it was not case for the first one (from 1992Q1 to 1993Q3), but this is not an error: Netherlands only adopted the Euro in 1999.

PALOMA: nao sei mais o que adicionar aqui, ja que nem é solicitado explicitamente uma análise desses gráficos. *AISHA* Acho que está bom. Eu fiz um gráfico onde as duas séries ficam juntas e adicionei as recessões com uma explicação no final.

AR(p) model for GDP - AR(4)

Comments about the estimation procedure: As in Part I, we opted by a model including the intercept and we used the function `Arima()` from the `forecast` package.

To avoid too many code in the document, only the final model code is appearing below and the remaining is omited — the original code is available in the link provided at the begining of this document.

Variable	Estimate	Std. Error	t-statistic	P-Value
Lag 1	0.33	0.10	3.42	0.00
Lag 2	0.06	0.10	0.56	0.57
Lag 3	0.21	0.10	2.08	0.04
Lag 4	0.05	0.10	0.47	0.64
Intercept	1.06	0.22	4.83	0.00

AR(p) model for GDP - AR(3)

With all lags

Variable	Estimate	Std. Error	t-statistic	P-Value
Lag 1	0.34	0.09	3.61	0.00
Lag 2	0.06	0.10	0.59	0.55
Lag 3	0.22	0.10	2.34	0.02
Intercept	1.07	0.21	5.05	0.00

Removing the second lag

```r
mAR3              <- Arima(tsData21[,1], order = c(3,0,0), fixed = c(NA,0,NA, NA))

AR3coef           <- tidy(coeftest(mAR3), stringsAsFactors = FALSE)
AR3coef           <- cbind(AR3coef[, 1], round(AR3coef[, 2:5], digits = 2))
colnames(AR3coef) <- c('Variable', 'Estimate', 'Std. Error', 't-statistic', 'P-Value')
AR3coef[,1]       <- c("Lag 1", "Lag 3", "Intercept")
```

3

```
AR3coef %>%
kable("latex") %>%
kable_styling(bootstrap_options = c("striped", "hover"))
```

Variable	Estimate	Std. Error	t-statistic	P-Value
Lag 1	0.36	0.09	4.05	0.00
Lag 3	0.24	0.09	2.69	0.01
Intercept	1.07	0.20	5.31	0.00

Comments on the results: We started by estimating and AR(4) and we use a significance level of 5% in our analysis. Since the coefficient for the fourth lag is statistically insignificant, we droped the fourth lag and re-estimated the model as an AR(3). From the results of this second model, the coefficient for the fourth lag is significant. However, the second lag p-value shows evidence of non-sisgnificance. Therefore, we reestimated the AR(3) with a zero coefficient for the second lag and obtained the final model:

$$GDP\_QGR_t = 1.07 + 0.36 GDP\_QGR_{t-1} + 0.24 GDP\_QGR_{t-3} + \epsilon_t,$$

$$\text{with} \quad \{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2).$$

According to this model, we have a persistence of the past GDP growth levels on the current growth. Also, since the coefficients are positive, we observe a "positive persistence". That is, past GDP growth tends to influence growth today in the same direction: growth would be related to some more growth in the future, and decline would be related with more future decline. Such persistence would last for about three quarters.

This lag persistence is consistent with the definition from the Euro Area Business Cycle Dating Committee (EABCDC) for a recession. Accordingly to the EABCDC, a recession is characterized as *"a significant decline in the level of economic activity, spread across the economy of the euro area, usually visible in **two or more** consecutive quarters of negative growth in GDP, employment and other measures of aggregate economic activity for the euro area as a whole."* (link for the source: https://eabcn.org/dc/methodology).

ADL(p,q) model for unemployment (GDP growth as exogenous regressor) - ADL(4,4)

Comments about the estimation procedure:

To estimate the model, we used the function `dynlm()` from the package `dynlm`. This is a wrapper for the class of `lm` functions in R, that offers some advantages by allowing to easily add lags from the series (by using the operator L). The same result would have been achieved by using the `lag()` function, but without the convenience of being able to preserve the objects as time series objects. If the model had seasonal or other components, it would be more easy to incorporate as well. The estimation procedure is the same as the one used in `lm()`, which can be consulted here.

Now, we proceed with the estimation of an ADL model for quarterly unemployment rate, using quarterly GDP growth as regressor. At first, we attempted to follow the general-to-specific method for selecting the ADL model with best fit. However, by sequentially removing insignificant lags (respecting a chronological ordering), we ended up with a model with three lags for unemployment an no lags and no current value of GDP. Since we are interesting in investigating some dynamics between unemployment and economic activity, we opted for alternative ways of selecting the best fit. The procedure we opted for is to eliminate the most insignificant lags (now we are not following a chronological ordering).

Sequentially, those were the models estimated and the lags removed (based on the p-values):

```
* We started with an ADL(4,4) model and realized that the 4th lag of GDP growth is
the one witht the highest p-value. Thus, this lag was removed.
```

```
* Next, we estimated an ADL(4,3) and decided to eliminate the 2nd lag of GDP
growth, and the the current GDP growth.
* After estimating this model, we decided to eliminate the fourth lag of unemployment.
By doing this, we were left with an ADL(3,3) with zero coefficient for the coefficient
of current and second lag of GDP growth.
* Finally, we proceed by eliminating the third lag of GDP growth and, since we are
considering a significance level of 5%, we also eliminate the second lag of unemployment rate.
```

As before, to avoid too many code in the document, only the final model code is appearing below and the remaining is omited — the original code is available in the link provided at the begining of this document.

Variable	Estimate	Std. Error	t-statistic	P-Value
Intercept	0.17	0.06	2.90	0.00
UN_RATE(-1)	1.60	0.10	15.57	0.00
UN_RATE(-2)	-0.30	0.19	-1.60	0.11
UN_RATE(-3)	-0.44	0.19	-2.24	0.03
UN_RATE(-4)	0.11	0.11	1.05	0.29
GDP_QGR	-0.01	0.01	-0.94	0.35
GDP_QGR(-1)	-0.02	0.02	-1.33	0.19
GDP_QGR(-2)	-0.01	0.02	-0.56	0.58
GDP_QGR(-3)	0.02	0.02	1.34	0.18
GDP_QGR(-4)	-0.01	0.01	-0.47	0.64

ADL(4,3)

Variable	Estimate	Std. Error	t-statistic	P-Value
Intercept	0.16	0.06	2.89	0.00
UN_RATE(-1)	1.60	0.10	15.69	0.00
UN_RATE(-2)	-0.29	0.19	-1.55	0.12
UN_RATE(-3)	-0.44	0.19	-2.26	0.03
UN_RATE(-4)	0.11	0.10	1.02	0.31
GDP_QGR	-0.01	0.01	-0.95	0.35
GDP_QGR(-1)	-0.02	0.01	-1.45	0.15
GDP_QGR(-2)	-0.01	0.02	-0.59	0.56
GDP_QGR(-3)	0.02	0.01	1.26	0.21

ADL(4,3) with zero coefficient for the second lag of GDP growth

Variable	Estimate	Std. Error	t-statistic P-Value	NA
Intercept	0.16	0.06	2.85	0.01
UN_RATE(-1)	1.61	0.10	16.01	0.00
UN_RATE(-2)	-0.30	0.19	-1.61	0.11
UN_RATE(-3)	-0.44	0.19	-2.28	0.02
UN_RATE(-4)	0.11	0.10	1.06	0.29
GDP_QGR	-0.01	0.01	-1.00	0.32
GDP_QGR(-1)	-0.02	0.01	-1.68	0.10
GDP_QGR(-3)	0.02	0.01	1.15	0.25

ADL(4,3) with zero coefficient for the current and second lag of GDP growth

Variable	Estimate	Std. Error	t-statistic	P-Value
Intercept	0.1621085	0.0562086	2.8840491	0.0048366
UN_RATE(-1)	1.5991858	0.1000723	15.9803049	0.0000000
UN_RATE(-2)	-0.3015861	0.1872108	-1.6109438	0.1104414
UN_RATE(-3)	-0.4120201	0.1906222	-2.1614486	0.0331218
UN_RATE(-4)	0.0885453	0.1017021	0.8706338	0.3861037
GDP_QGR(-1)	-0.0287103	0.0133990	-2.1427213	0.0346380
GDP_QGR(-3)	0.0127485	0.0137022	0.9304018	0.3544728

ADL(3,3) with zero coefficient for the current and second lag of GDP growth

Variable	Estimate	Std. Error	t-statistic	P-Value
Intercept	0.18	0.05	3.33	0.00
UN_RATE(-1)	1.58	0.10	16.35	0.00
UN_RATE(-2)	-0.34	0.18	-1.86	0.07
UN_RATE(-3)	-0.27	0.10	-2.80	0.01
GDP_QGR(-1)	-0.03	0.01	-2.12	0.04
GDP_QGR(-3)	0.01	0.01	0.90	0.37

ADL(3,1) with coefficient zero for current GDP

Variable	Estimate	Std. Error	t-statistic	P-Value
Intercept	0.18	0.05	3.54	0.00
UN_RATE(-1)	1.57	0.10	16.34	0.00
UN_RATE(-2)	-0.34	0.18	-1.88	0.06
UN_RATE(-3)	-0.26	0.10	-2.73	0.01
GDP_QGR(-1)	-0.03	0.01	-1.98	0.05

ADL(3,1) with coefficient zero for current GDP and second lag of unemployment

```
mADL31_0 <- dynlm(ts(tsData21[,2]) ~ L(ts(tsData21[,2])) +L(ts(tsData21[,2]),3)
              + L(ts(tsData21[,1])))

ADL31_0coef            <- tidy(coeftest(mADL31_0), stringsAsFactors = FALSE)
ADL31_0coef            <- cbind(ADL31_0coef[, 1], round(ADL31_0coef[, 2:5], digits = 2))
colnames(ADL31_0coef) <- c('Variable', 'Estimate', 'Std. Error', 't-statistic',
                          'P-Value')
ADL31_0coef[,1]        <- c("Intercept", "UN_RATE(-1)", "UN_RATE(-3)", "GDP_QGR(-1)")

ADL31_0coef %>%
kable("latex") %>%
kable_styling(bootstrap_options = c("striped", "hover"))
```

Variable	Estimate	Std. Error	t-statistic	P-Value
Intercept	0.21	0.05	4.04	0.00
UN_RATE(-1)	1.40	0.03	48.71	0.00
UN_RATE(-3)	-0.43	0.03	-14.88	0.00
GDP_QGR(-1)	-0.03	0.01	-2.07	0.04

Comments on the results: The resulting model from this procedure for model selection is an ADL(3,1) with no secong lag for unemployment and no current GDP growth:

$$UN\_RATE_t = 0.21 + 1.40 UN\_RATE_{t-1} - 0.43 UN\_RATE_{t-3} - 0.03 GDP\_QGR_{t-1} + \varepsilon_t,$$

$$\text{with} \quad \{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2).$$

From the coefficients, we note a very small, but significant, negative effect of previous period GDP growth on current unemployment. This effect makes sense since it is expected that, as the economic activity declines, employers would find it more difficult to keep employing the same amount of workers. There is also a second effect, which is more observed in recessions: during uncertain times, people who otherwise would be outside the labor force (for example studying or as domestic partner), starts to look for jobs, increasing the unemployment rate.

As for the coefficients of unemployment lags, they show a positive effect of the unemployment rate from the period immediately before, but a negative effect of unemployment three periods before. This means that the trend for unemployment will most likely persist until next period, but then it will start changing a few quarters from now. This behavior is somewhat consistent with the cycles observed in the graph analysis.

Question 2

Use the estimated ADL model to calculate and interpret the short-run multiplier, the 2-step-ahead multiplier, and the long-run multiplier. Finally, report and interpret the long-run relation between the unemployment rate and the GDP growth rate.

Let the unemployment rate be u_t, the gdp quartely growth be gdp_t. The final model will be given by:

$$u_t = \alpha + \phi_1 u_{t-1} + \phi_2 u_{t-3} + \beta gdp_{t-1} + \varepsilon_t$$

$$\{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$$

The estimated model will be:

$$u_t = 0.21 + 1.40 u_{t-1} - 0.43 u_{t-3} - 0.03 gdp_{t-1} + \varepsilon_t$$

$$\{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$$

The short-run multiplier tells us how the current value of the exogenous variable (quarterly GDP growth in our case) explains the endogenous variable (quarterly unemployment). Since we do not have gdp_t as an exogenous regressor in the final model, the short-run multiplier will be zero (percentual increases over GDP lead to no changes in current unemployment. The derivation is as follows:

$$u_t^* = \alpha + \phi_1 u_{t-1} + \phi_2 u_{t-3} + \beta gdp_{t-1} + \varepsilon_t = u_t$$

The 2-step-ahead multiplier will give us the impact on unemployment two-step ahead of a one unit temporary increase in current GDP growth. This multiplier will be equal to $\phi_1 \beta = 1.4 \cdot (-0.03) = -0.042$. That is, a

percentage increase in GDP is expected to result in a fall of 0.042% in unemployment. The derivation is as follows:

$$\begin{aligned}
u_{t+1}^* &= \alpha + \phi_1 u_t + \phi_2 u_{t-2} + \beta(gdp_t + 1) + \varepsilon_{t+1} \\
&= \alpha + \phi_1 u_t + \phi_2 u_{t-2} + \beta gdp_t + \beta + \varepsilon_{t+1} \\
&= u_{t+1} + \beta \\
u_{t+2}^* &= \alpha + \phi_1 u_{t+1} + \phi_2 u_{t-1} + \beta gdp_{t+1} + \varepsilon_{t+2} \\
&= \alpha + \phi_1(u_{t+1} + \beta) + \phi_2 u_{t-1} + \beta gdp_{t+1} + \varepsilon_{t+2} \\
&= u_{t+2} + \phi_1 \beta
\end{aligned}$$

The long-run multiplier will give us the effect over unemployment of a permanent change in GDP growth, and it is obtained from the long run relationship between these two variables. Such relatioship is obtained from the fact that, in the long-run, variables should be equal (or very close to) their historical averages. Thus, the long-run relatioship between unemployment and GDP growth will be:

$$\overline{u} = \alpha + \phi_1 \overline{u} + \phi_2 \overline{u} + \beta \overline{gdp}$$

$$\overline{u} = \frac{\alpha}{1 - \phi_1 - \phi_2} + \frac{\beta}{1 - \phi_1 - \phi_2} \overline{gdp}.$$

Thus, the long-run multiplier will be $\frac{\beta}{1-\phi_1-\phi_2} = \frac{-0.03}{1-1.40+0.43} = -1$: 1% increase in GDP is expected to result in 1% fall on unemployment.

Question 3

Please provide a detailed comment on the following statement: "An increase in the GDP growth rate causes a reduction in the unemployment rate."

From the multipliers in the previous question, we can breakdown the effects of a change in GDP growth over unemployment rate by analyzing the multipliers for the short and long run. Since the short-run multiplier β_0 (the coefficient of GPD at time t in the equation of unemployment) is zero, there is no immediate effect of the GDP growth rate over unemployment.

As time goes by, the effect of GPD over unemployment tends to get stronger and more negative. In the previous question, we found that there is no immediate effect of the GDP growth rate over unemployment (zero short-run multiplier). As part of the solution for the 2-step multiplier, however, we found the 1-step multiplier, which equals $\beta_1 = -0.03$. Thus, we have that changes in GDP start to have an effect over unemployment only after 1 period. The 2-step ahead multiplier, which is given by $\phi \beta_1 = -0.042$, tells us that the effect of a positive GDP shock today will lead to even more decrease in unemployment two steps ahead, which indicates that the effect tends to become stronger (i.e. more negative) as time goes by. That is, there is some sort of persistence of temporary GDP shocks over unemployment rate that would lead to smaller and smaller unemployment. This result extends to the effects of a permanent shock on GDP growth, which according to the long-run multiplier, is of a proportional decrease in unemployment.

Note that these numbers tell us something about how the dynamics of GDP growth would affect unemployment. This is not a causal analysis. Basically, what we found is that, for example, increasing GDP growth tends to be associated with decreasing unemployment, but we cannot tell yet whether increasing GDP causes decreasing unemployment. To be able to draw conclusions about the causality, one needs to make a deeper investigation using not only statistical methods but also more economic theory to sustain these conclusions. In the latter, one could discuss the relationship between Okun's Law and the model results.

Question 4

Suppose that the innovations are iid Gaussian. What is the probability of the unemployment rate rising above 7.8% in the second quarter of 2014? What is the probability that it drops below 7.8%? Do you trust the iid Gaussian assumption?

Residuals from the ADL model
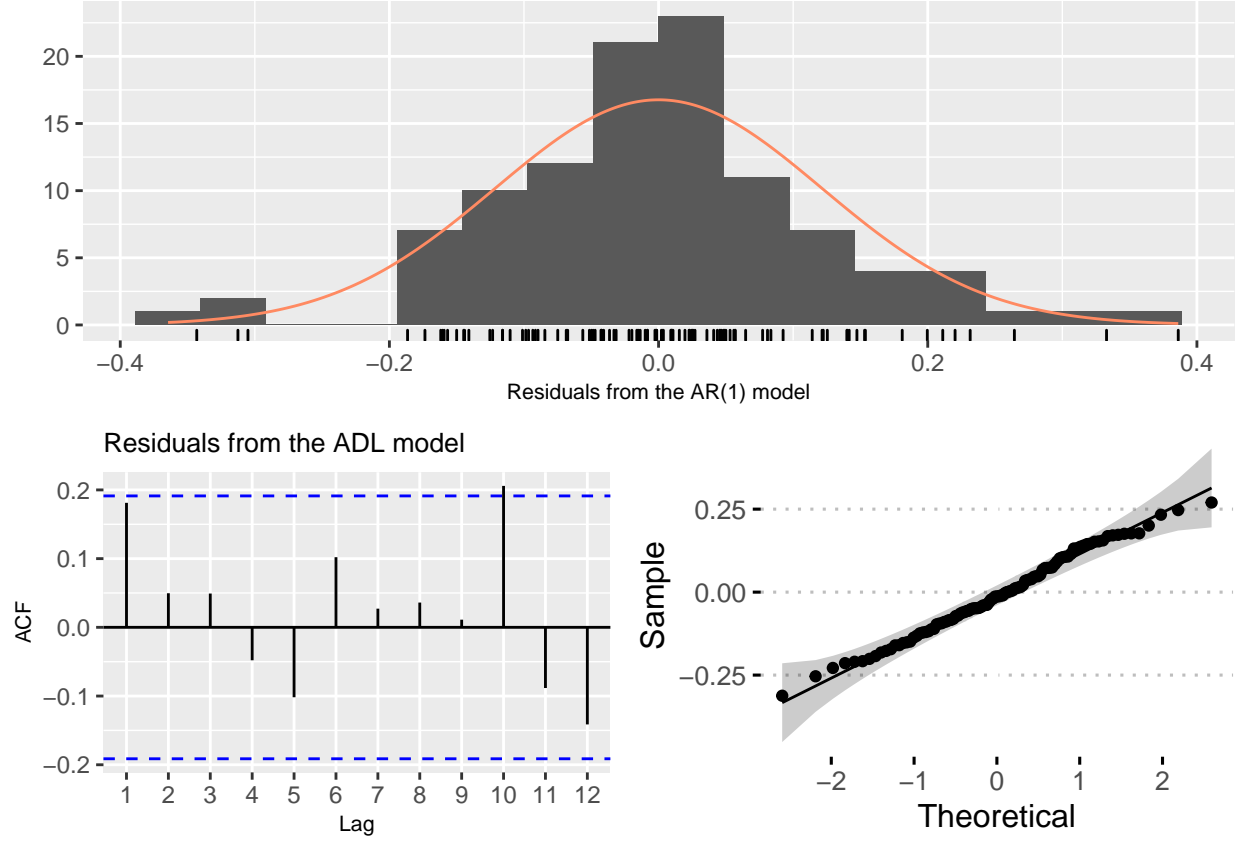
To compute these probabilities, we first find the one step ahead estimator for UN_RATE. Again, let the unemployment rate be u_t, the gdp quartely growth be gdp_t and the one step ahead forecast be $\hat{u}_{T+1}$. Since we are assuming $\{\varepsilon_t\} \sim iid(0, \sigma_\varepsilon^2)$, we have that:

$$
\begin{aligned}
\hat{u}_{T+1} &= E[u_{T+1}|D_T = (u_1, \ldots, u_T, gdp_1, \ldots, gdp_T)] \\
&= E[\alpha + \phi_1 u_T + \phi_2 u_{T-2} + \beta gdp_T + \varepsilon_{T+1}|D_T] \\
&= \alpha + \phi_1 u_T + \phi_2 u_{T-2} + \beta gdp_T + E[\varepsilon_{T+1}|D_T] \\
&= \alpha + \phi_1 u_T + \phi_2 u_{T-2} + \beta gdp_T \\
&= u_{T+1} - \varepsilon_{T+1}
\end{aligned}
$$

The last step follows from the innovations being iid Gaussian distributed.

The forecast error term will be given by

$$
e_{T+1} = u_{t+1} - \hat{u}_{T+1} = \varepsilon_{T+1}.
$$

Therefore, we have that $e_T \sim NID(0, \sigma_\varepsilon^2)$.

From this result, it follows that the probability that unemployment in the second quarter of 2014 will be

$$
\begin{aligned}
P(u_{T+1} > 7.8) &= P(u_{T+1} - \hat{u}_{T+1} > 7.8 - \hat{u}_{T+1}) \\
&= P(e_{T+1} > 7.8 - \hat{u}_{T+1}) \\
&= P\left(\frac{e_{T+1}}{\sigma_\varepsilon} > \frac{7.8 - \hat{u}_{T+1}}{\sigma_\varepsilon}\right)
\end{aligned}
$$

From the historical values and model estimates for the parameters and the residual standard error (0.1233),

we get:

$$\hat{u}_{T+1} = 0.21 + 1.40 \cdot 7.8 - 0.43 \cdot 7.5 - 0.03 \cdot (-0.37) = 7.9161$$
$$\hat{\sigma}_{\varepsilon} = 0.1233$$

Thus, the probability that the unemployment exeeceds 7.8% in the second quarter of 2014 is

$$P(u_{T+1} > 7.8) = P\left(\frac{e_{T+1}}{\hat{\sigma}_{\varepsilon}} > \frac{7.8 - 7.9161}{0.1233}\right) = P\left(\frac{e_{T+1}}{\hat{\sigma}_{\varepsilon}} > -0.94\right) = 0.8264 = 82.64\%$$

and the probability that unemployment will be smaller than 7.8% is $1 - 0.8264 = 17.36\%$.

As for the last question, we look at the histogram and the p-value of Kolmogorov-Smirnov test. The null hypothesis of such test is that the data comes from a normal distribution, with mean 0 and standard deviation 0 (as remark, the mean of the residuals is virtually zero, as we would expect). Since the p-value for the KS test is higher than 5 (KS statistic = 0.103; p-value = 0.2154), we have evidence to not reject the null hypothesis. That is, we have evidence that the assumption of iid Gaussian distrubution for the innovations does seem reasonable. From the histogram of the residuals, we see that their distribution indeed seems to be bell shaped, but with somewhat heavy tails, speccialy on the left of the mean, but this was not significant in terms of the test and also doesn't show in the qqplot.

Question 5

Make use of your estimated AR and ADL models to produce a 2-year (8 quarter) forecast for the Unemployment rate that spans until the first quarter of 2016. Report the obtained values.

First, we estimate values for quartely GDP growth. We used the same procedure as the one used for the forecasts in part 1, the `forecast()` function from the package `forecast`.

```
h = 8
AR3forecast  <- forecast::forecast(mAR3, h, level = 95)
vdNewDate    <- c("2014Q2", "2014Q3", "2014Q4", "2015Q1", "2015Q2", "2015Q3", "2015Q4",
                  "2016Q1")
AR3forecast  <- cbind(vdNewDate, data.frame(AR3forecast))

colnames(AR3forecast) <- c("Date", "Forecast", "L95", "H95")

autoplot(forecast::forecast(mAR3, h, level = 95),
         main = "12 step ahead forecast using the AR(3) model",
         xlab = "Quarter", ylab = "Dutch GDP q. growth")
```

12 step ahead forecast using the AR(3) model
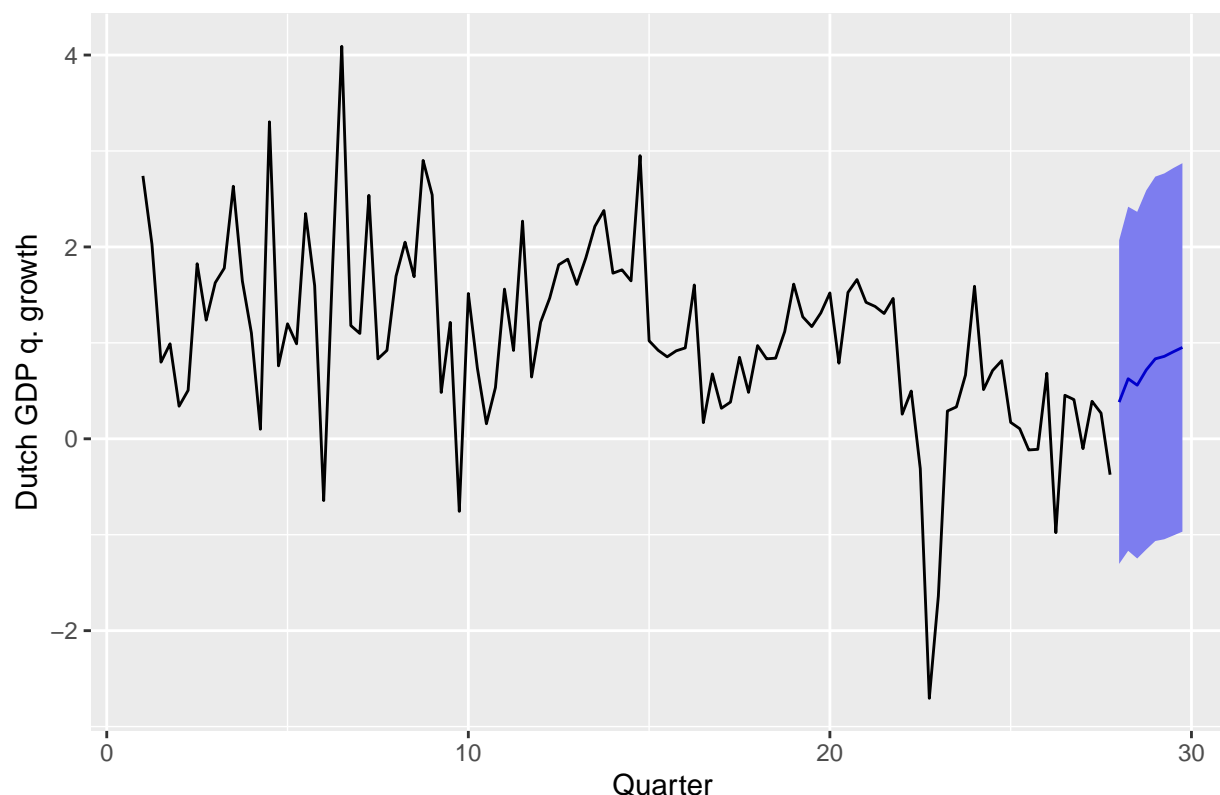

```
AR3forecast %>%
  kable("latex") %>%
  kable_styling(bootstrap_options = c("striped", "hover"))
```

	Date	Forecast	L95	H95
28 Q1	2014Q2	0.3820861	-1.3043609	2.068533
28 Q2	2014Q3	0.6258468	-1.1674324	2.419126
28 Q3	2014Q4	0.5588832	-1.2478931	2.365659
28 Q4	2015Q1	0.7174058	-1.1538686	2.588680
29 Q1	2015Q2	0.8335985	-1.0654103	2.732607
29 Q2	2015Q3	0.8594311	-1.0471856	2.766048
29 Q3	2015Q4	0.9070621	-1.0079544	2.822079
29 Q4	2016Q1	0.9523491	-0.9679906	2.872689

From the forecasted results, we see that the model expects, on average, a gradual increase in GDP growth for the next 2 years. However, it is important to notice that the 95% confidence region goes below the last observed point, covering negative values for the forecasted results.

Now, we use these forecasted values for GDP growth in the unemployment forecast. We used the approach to compute the forecasts "step-by-step", i.e., we used the estimated model from question (1) to predict what will be the Unemployment for 2014 Q2 using the forecasted value for the GPD. Next, we use this forecasted Unemployment to compute a forecast for 2014 Q3 and so on until 2016 Q1. We couldn't find any reliable source on how to compute the confidence intervals – more specifically, given that we have the uncertainty from the forecasts for GDP, we were unsure how to compute the confidence bands for this model. The next chunk of code contains the steps used to assemble the data and make the forecast.

```r
# Setting up the data frame with GDP growth forecasts
vdNewDate <- c("2014Q2", "2014Q3", "2014Q4", "2015Q1",
               "2015Q2", "2015Q3", "2015Q4", "2016Q1")

dfNewData <- data.frame(vdNewDate, AR3forecast[,2])
 names(dfNewData) <- names(dfData21[,1:2])
 dfData22 <- rbind(dfData21[,1:2], dfNewData)
 dfData22 <- merge(dfData22, dfData21[,3], by = "row.names", all = TRUE)

 # This will result in two columns, where the last 8 values
 # for Unemployment are NA.
 tsData22 <- xts(dfData22[,3:4], order.by = as.yearqtr(dfData22[,2]), frequency = 4)
 names(tsData22) <- c("GDP_QGR", "UN_RATE")

# We are doing the forecast step by step
# Recovers the index for the row where the first forecast of GDP is located
indexts        <- which(index(tsData22) == "2014 Q2")
ADL22_forecast    <- tsData22

for (i in indexts:nrow(tsData22)){
  ADL22_forecast[i,2] <- coef(mADL31_0) %*% c(1, ADL22_forecast[(i-1), 2],
                                              ADL22_forecast[(i-3), 2],
                                              ADL22_forecast[(i-1), 1])

}
```

The last two observed periods (2013 Q3 and 2014 Q1) as well the 8 forecast results are displayed in the table below.

	GDP_QGR	UN_RATE
2013 Q4	0.2676572	7.600000
2014 Q1	-0.3744039	7.800000
2014 Q2	0.3820861	7.899095
2014 Q3	0.6258468	7.974145
2014 Q4	0.5588832	7.986278
2015 Q1	0.7174058	7.962313
2015 Q2	0.8335985	7.892082
2015 Q3	0.8594311	7.785393
2015 Q4	0.9070621	7.645696
2016 Q1	0.9523491	7.479175

Question 6

Use impulse response functions (IRFs) to analyze two different scenarios for the Dutch un-employment rate:

(a) In the 'good scenario' the GDP quarterly growth rate is hit by a positive shock of 2%.

(b) In the 'bad scenario' the GDP quarterly growth rate suffers a negative shock of 2%.

Please use the last observed value of the unemployment rate and gdp growth rate as the origin of your IRFs. In particular, set the origin to -0.37% for GDP and 7.8% for the unemployment rate.

We start by computing the IRF for the AR(3) model. In general terms, when we have the following model

$$X_t = \alpha + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \varepsilon_t, \text{ with } \varepsilon_t \sim \text{i.i.d.}(0, \sigma^2),$$

to evaluate the IRF, we are interested in determine the above system:

$$\tilde{x}_t = x, \quad t < s$$
$$\tilde{x}_t = x + \epsilon, \quad t = s$$
$$\tilde{x}_t = x + \frac{\partial X_t}{\partial \varepsilon_s}\epsilon, \quad \forall t > s$$

Therefore, are interested in knowing $\frac{\partial X_t}{\partial \varepsilon_s}$.

We know that for an AR(3) model, we have the following:

$$\frac{\partial X_{s+1}}{\partial \varepsilon_s} = \phi_1 \frac{\partial X_s}{\partial \varepsilon_s}$$
$$\frac{\partial X_{s+2}}{\partial \varepsilon_s} = \phi_1 \frac{\partial X_{s+1}}{\partial \varepsilon_s} + \phi_2 \frac{\partial X_s}{\partial \varepsilon_s} = (\phi_1^2 + \phi_2)\frac{\partial X_s}{\partial \varepsilon_s}$$
$$\frac{\partial X_{s+3}}{\partial \varepsilon_s} = \phi_1 \frac{\partial X_{s+2}}{\partial \varepsilon_s} + \phi_2 \frac{\partial X_{s+1}}{\partial \varepsilon_s} + \phi_3 \frac{\partial X_s}{\partial \varepsilon_s} = (\phi_1^3 + 2\phi_1\phi_2 + \phi_3)\frac{\partial X_s}{\partial \varepsilon_s}$$
$$\frac{\partial X_{s+4}}{\partial \varepsilon_s} = \phi_1 \frac{\partial X_{s+3}}{\partial \varepsilon_s} + \phi_2 \frac{\partial X_{s+2}}{\partial \varepsilon_s} + \phi_3 \frac{\partial X_{s+1}}{\partial \varepsilon_s} = (\phi_1^4 + \phi_2^2 + 3\phi_1^2\phi_2 + 2\phi_1\phi_3)\frac{\partial X_s}{\partial \varepsilon_s}$$
$$\frac{\partial X_{s+5}}{\partial \varepsilon_s} = \phi_1 \frac{\partial X_{s+4}}{\partial \varepsilon_s} + \phi_2 \frac{\partial X_{s+3}}{\partial \varepsilon_s} + \phi_3 \frac{\partial X_{s+2}}{\partial \varepsilon_s},$$
$$\text{etc}$$

To implement this, we only need to compute separately the first three components. From there, we can use a recursion to obtain a vector with $\frac{\partial X_{s+i}}{\partial \varepsilon_s}$, $\forall\ i > 3$.

Now that we have a recursion for the AR part, we need to write the recursion for the ADL part. Using a general structure for an arbitrary ADL(3,1)-AR(3) system given by

$$Y_t = \gamma + \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + \psi_3 Y_{t-3} + \beta X_{t-1} + \varepsilon_t,$$
$$X_t = \alpha + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + v_t$$

we have the same equations above for the IRF of the x (which represents the AR(3) part) plus the following:

$$\tilde{y}_{s-1} = y,$$
$$\tilde{y}_s = y + 0,$$
$$\tilde{y}_{s+1} = y + \beta \frac{\partial X_s}{\partial v_s}\epsilon$$
$$\tilde{y}_{s+2} = y + \psi_1 \frac{\partial Y_{s+1}}{\partial v_s}\epsilon + \beta \frac{\partial X_{s+1}}{\partial v_s}\epsilon$$
$$\tilde{y}_{s+3} = y + \psi_1 \frac{\partial Y_{s+2}}{\partial v_s}\epsilon + \psi_2 \frac{\partial Y_{s+1}}{\partial v_s}\epsilon + \beta \frac{\partial X_{s+2}}{\partial v_s}\epsilon$$
$$\tilde{y}_{s+j} = y + \psi_1 \frac{\partial Y_{s+j-1}}{\partial v_s}\epsilon + \psi_2 \frac{\partial Y_{s+j-2}}{\partial v_s}\epsilon + \psi_3 \frac{\partial Y_{s+j-3}}{\partial v_s}\epsilon + \beta \frac{\partial X_{s+j-1}}{\partial v_s}\epsilon,\ j > 3$$

Given this, we can write a function that will do the recursions:

```
# The function below computes the IRF for an ADL(3,1)-AR(3) system
fIRFADL31 <- function(param, shock = 0.02*100, h = 20,
                      lastgdp = -0.0037*100, lastun = 0.078*100){

  # To avoid errors
```

```r
  if (length(param) != 7){
    param <- rep(0,7)
  }

  # To look pretty (because the first value is at the moment of the shock)
  h <- h -1

  # Extract parameters
  phi1 <- param[1]
  phi2 <- param[2]
  phi3 <- param[3]

  psi1 <- param[4]
  psi2 <- param[5]
  psi3 <- param[6]

  beta <- param[7]

  # Create empty vectors
  derivativesx <- c(shock, rep(NA, h))
  derivativesy <- c(shock, rep(NA, h))

  # Make the recursion
  for (i in 1:h){
    if (i == 1){
      derivativesx[i+1] <- phi1 * shock
      derivativesy[i+1] <- 0
    } else if (i == 2) {
      derivativesx[i+1] <- phi1 * derivativesx[i] + phi2 * shock
      #derivativesy[i+1] <- psi1 * derivativesy[i] * shock + beta * derivativesx[i] * shock
      derivativesy[i+1] <- psi1 * derivativesy[i] + beta * derivativesx[i]
    } else if (i == 3){
      derivativesx[i+1] <- phi1 * derivativesx[i] + phi2 * derivativesx[i-1] + phi3 * shock
      #derivativesy[i+1] <- psi1 * derivativesy[i] * shock + psi2 * derivativesy[i-1] * shock
      #+ beta * derivativesx[i] * shock
      derivativesy[i+1] <- psi1 * derivativesy[i] + psi2 * derivativesy[i-1] + beta * derivativesx[i]
    } else if (i > 3) {
      derivativesx[i+1] <- phi1 * derivativesx[i] + phi2 * derivativesx[i-1] + phi3 * derivativesx[i-2]
      #derivativesy[i+1] <- psi1 * derivativesy[i] * shock + psi2 * derivativesy[i-1] * shock +
      #psi3 * derivativesy[i-2] * shock + beta * derivativesx[i] * shock
      derivativesy[i+1] <- psi1 * derivativesy[i] + psi2 * derivativesy[i-1] + psi3 * derivativesy[i-2]
    }
  }

  irfx <- lastgdp + derivativesx
  irfy <- lastun + derivativesy

  dfIRF <- data.frame(1:(h+1), irfx, irfy)
  names(dfIRF) <- c("Lag", "GDP_QGR", "UN_RATE")

  return(dfIRF)
}
```
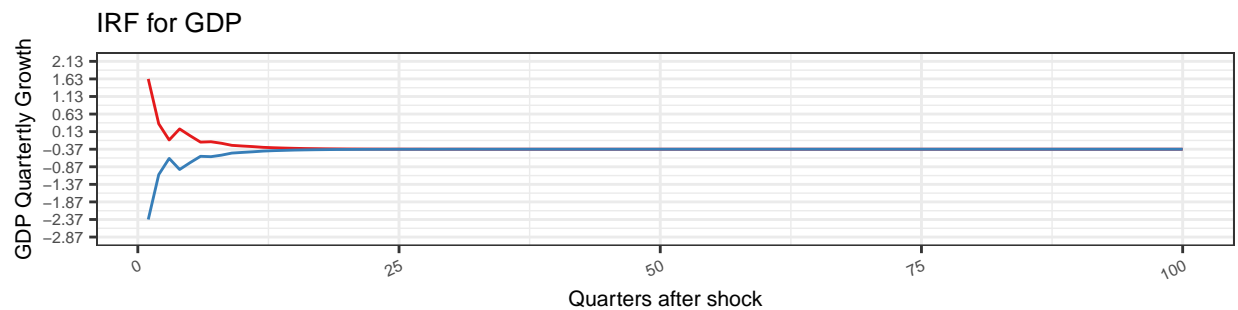
And we can compute the IRF to plot the graphs.

```r
# Get the coefficients
phi1  <- coef(mAR3)[1]
phi2  <- coef(mAR3)[2]
phi3  <- coef(mAR3)[3]
psi1  <- coef(mADL31_0)[2]
psi2  <- 0
psi3  <- coef(mADL31_0)[3]
beta  <- coef(mADL31_0)[4]

# Create a vector with parameters
parameters <- c(phi1, phi2, phi3, psi1, psi2, psi3, beta)
janela = 100

# If you want, may also create variables for shock, last observation, etc.
dfIRF1 <- fIRFADL31(parameters, shock = 0.02*100, h = janela,
                    lastgdp = -0.0037*100, lastun = 0.078*100)
dfIRF2 <- fIRFADL31(parameters, shock = -0.02*100, h = janela,
                    lastgdp = -0.0037*100, lastun = 0.078*100)
```