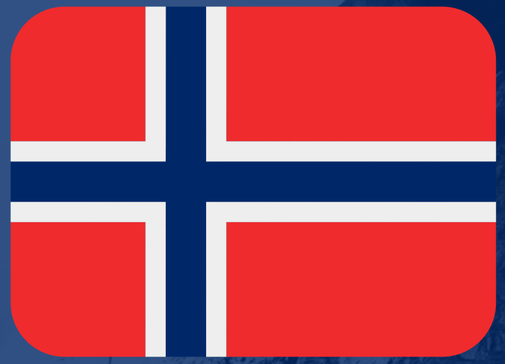


NORWAY

Presentation by Luca Colaci (3176608), Rocco Totaro (3159435), Giovanni Parlangeli(3152699)

Econometrics 1st Assignment

A.A. 2022/23

THE VARIABLES

TARGET INFLATION

Obtained through Norway Central Bank historical news

ACTUAL INFLATION

Obtained through OECD datasets

GDP

Quarterly, obtained through Eurostat datasets

EXCHANGE RATE USD

Quarterly, obtained through Eurostat datasets

UNEMPLOYMENT RATE

Quarterly, obtained through Eurostat datasets

POTENTIAL GDP

OECD datasets give a measure of the yearly output gap as a percentage of potential output: we inverted the equation to obtain the nominal potential output and then interpolated it linearly to obtain the quarterly potential output

LIBRARIES

```
1
2  ###LOADING LIBRARIES###
3  library("dplyr")
4  library("lmtest")
5  library("tsoutliers")
6  library("readxl")
7  library("ggplot2")
8  library("aTSA")
9  library("knitr")
10 library("dynlm")
11 library("car")
12 library("readxl")
13 library("tidyverse")
14 library("broom")
15 library("performance")
16 library("leaps")
17 library("glmnet")
```

DATASET

```
> summary(df)
      time          target            actual              gdp          potential          rate          log_gdp
 Min.   :2001   Min.   :0.02000   Min.   :-0.01427   Min.   :564395   Min.   :571624   Min.   :0.00000   Min.   :5.752
 1st Qu.:2007   1st Qu.:0.02000   1st Qu.: 0.01376   1st Qu.:637590   1st Qu.:627782   1st Qu.:0.00750   1st Qu.:5.805
 Median :2012   Median :0.02500   Median : 0.02035   Median :670576   Median :681837   Median :0.01571   Median :5.826
 Mean   :2012   Mean   :0.02368   Mean   : 0.02231   Mean   :679692   Mean   :688099   Mean   :0.02202   Mean   :5.830
 3rd Qu.:2017   3rd Qu.:0.02500   3rd Qu.: 0.02960   3rd Qu.:737288   3rd Qu.:753146   3rd Qu.:0.02446   3rd Qu.:5.868
 Max.   :2023   Max.   :0.02500   Max.   : 0.06741   Max.   :806724   Max.   :824133   Max.   :0.07000   Max.   :5.907
 log_potential   unemployement_rate  exchange_usd
 Min.   :5.757   Min.   :0.003652   Min.   :0.09985
 1st Qu.:5.798   1st Qu.:0.032022   1st Qu.:0.11731
 Median :5.834   Median :0.036989   Median :0.14488
 Mean   :5.835   Mean   :0.035987   Mean   :0.14185
 3rd Qu.:5.877   3rd Qu.:0.042905   3rd Qu.:0.16460
 Max.   :5.916   Max.   :0.051049   Max.   :0.19661
```

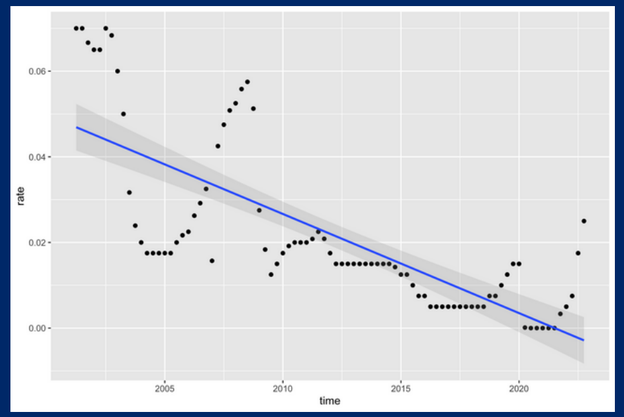
The dataset referred to the period from the second quarter of 2001 to the third quarter of 2022. This choice was dictated by the fact that from that quarter the Norwegian central bank began setting their policy rate based on target inflation.

Potential output was interpolated linearly due to the fact that is an estimate, therefore no particular growth trend can be imposed on it.

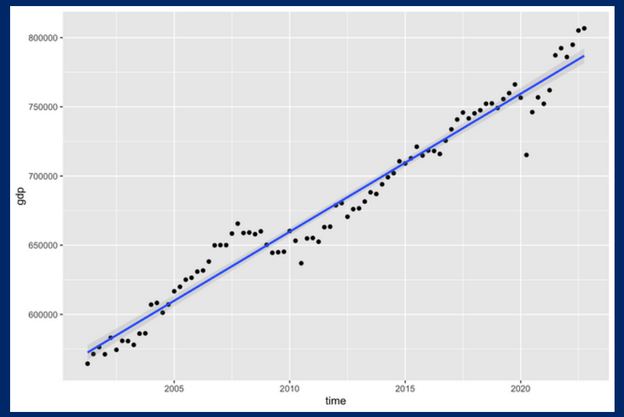
Data sources:

- OECD Datasets
- Eurostat Datasets
- Norway Central Bank

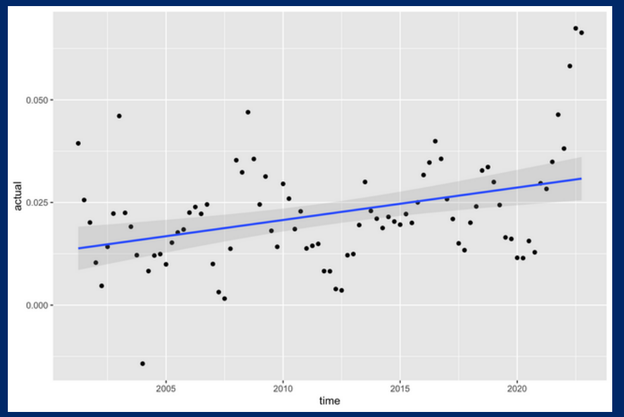
PRELIMINARY VISUALIZATION



POLICY RATE



GDP



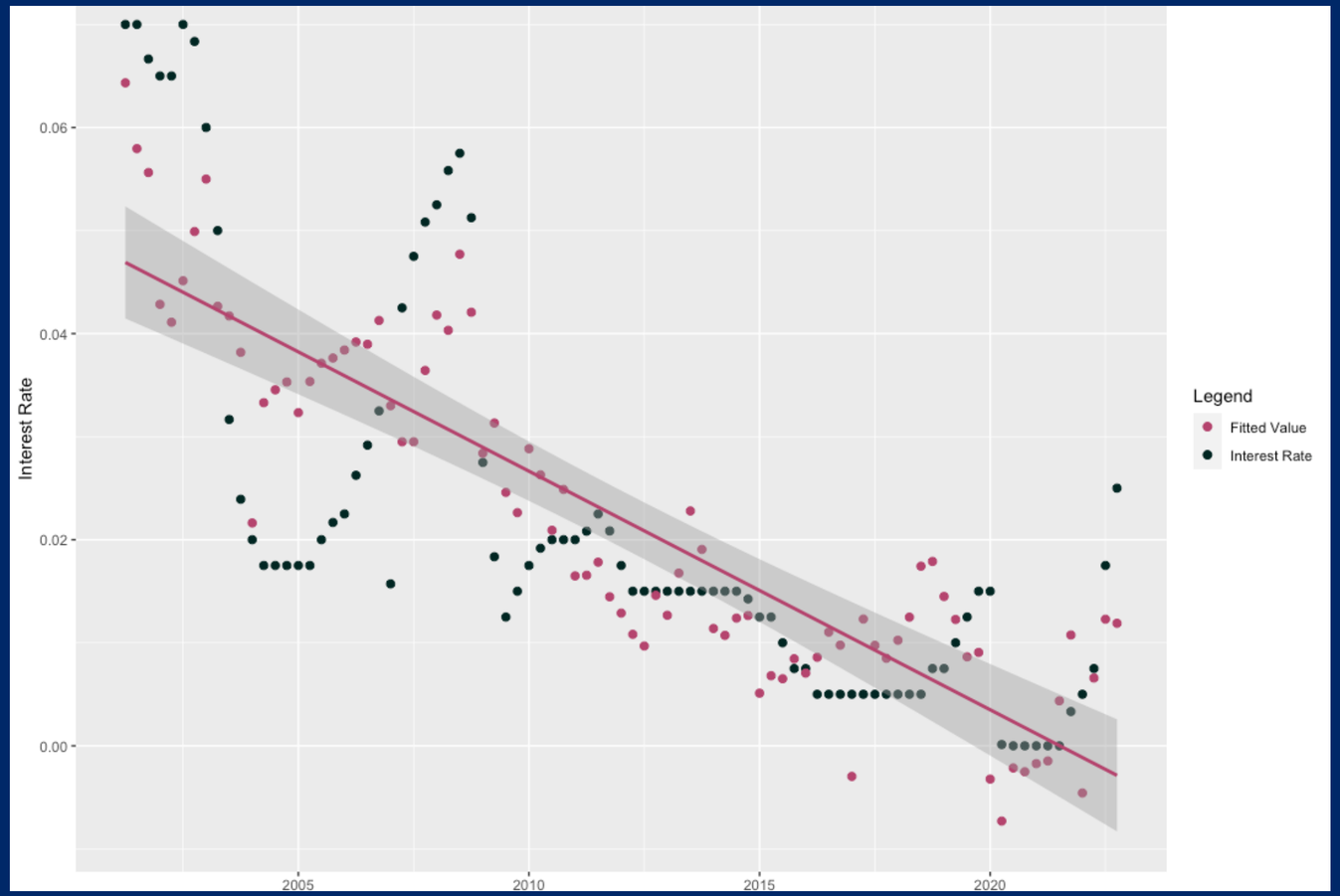
INFLATION

```
###PRELIMINARY VISUALIZATION###
#policy rate
df %>% ggplot(aes(time, rate)) + geom_point() + stat_smooth(method = "lm", formula = y ~ x, alpha = 0.2)
#GDP
df %>% ggplot(aes(time, gdp)) + geom_point() + stat_smooth(method = "lm", formula = y ~ x, alpha = 0.2)
#inflation
df %>% ggplot(aes(time, actual)) + geom_point() + stat_smooth(method = "lm", formula = y ~ x, alpha = 0.2)
```

REGRESSION

It is possible to see that target&actual and gdp&potential have inverse signs recalling the Taylor rule.
Everything is significant other than GDP.
With a fast check of the model, we can notice collinearity between GDP and Potential.

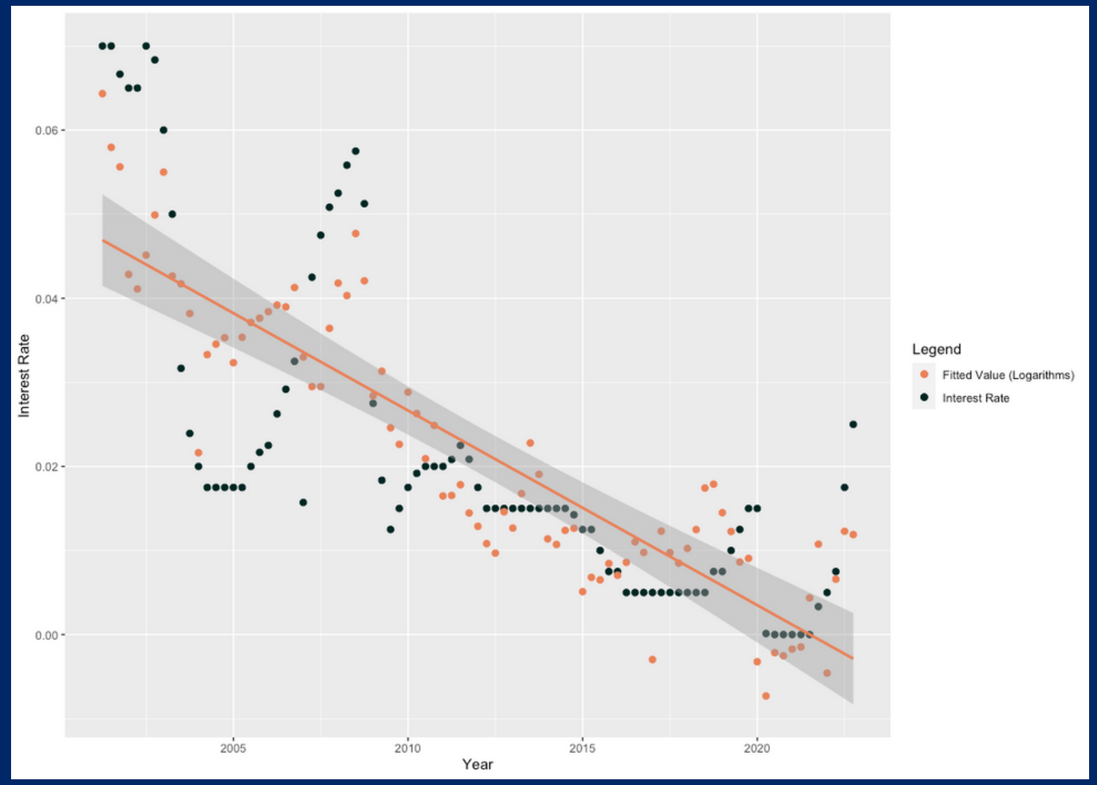
```
Call:
lm(formula = df$rate ~ df$target + df$actual + df$gdp + df$potential)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0178547 -0.0074576  0.0009197  0.0058181  0.0248705

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.164e-01  3.760e-02   8.416 1.04e-12 ***
df$target    -3.411e+00  8.654e-01  -3.941  0.00017 ***
df$actual     5.119e-01  9.386e-02   5.454 5.08e-07 ***
df$gdp        9.762e-08  8.131e-08   1.201  0.23340
df$potential -4.235e-07  7.983e-08  -5.305 9.36e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01058 on 82 degrees of freedom
Multiple R-squared:  0.717,     Adjusted R-squared:  0.7032
F-statistic: 51.94 on 4 and 82 DF,  p-value: < 2.2e-16
```



REGRESSION

As an alternative formulation, we can use logarithms for GDP and potential.

```
Call:
lm(formula = rate ~ target + actual + log_gdp + log_potential,
    data = df)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0182450 -0.0068181  0.0007241  0.0062321  0.0240215

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.02275    0.26360  11.467  < 2e-16 ***
target        -3.04726    0.80895  -3.767  0.00031 ***
actual         0.48042    0.09027   5.322 8.74e-07 ***
log_gdp        0.19227    0.12158   1.582  0.11761
log_potential -0.69580    0.12203  -5.702 1.81e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01028 on 82 degrees of freedom
Multiple R-squared:  0.7333,    Adjusted R-squared:  0.7203
F-statistic: 56.36 on 4 and 82 DF,  p-value: < 2.2e-16
```



We have lost the sign match with the Taylor rule but R^2 is higher, so the model is more precise

HOSCEDASTICITY

BREUSCH-PAGAN TEST

```
data:  logfit
BP = 31.792, df = 4, p-value = 2.109e-06
```

The model failed in rejecting the null hypothesis, it is not correctly specified

GOLDFELT-QUANDT TEST

```
data:  logfit
GQ = 0.1546, df1 = 30, df2 = 30, p-value = 1
alternative hypothesis: variance increases from segment 1 to 2
```

We reject the null hypothesis, we have heteroscedasticity

WHITE TEST

```
data:  logfit
BP = 31.792, df = 4, p-value = 2.109e-06
```

It failed in rejecting the null hypothesis, it is homoscedastic

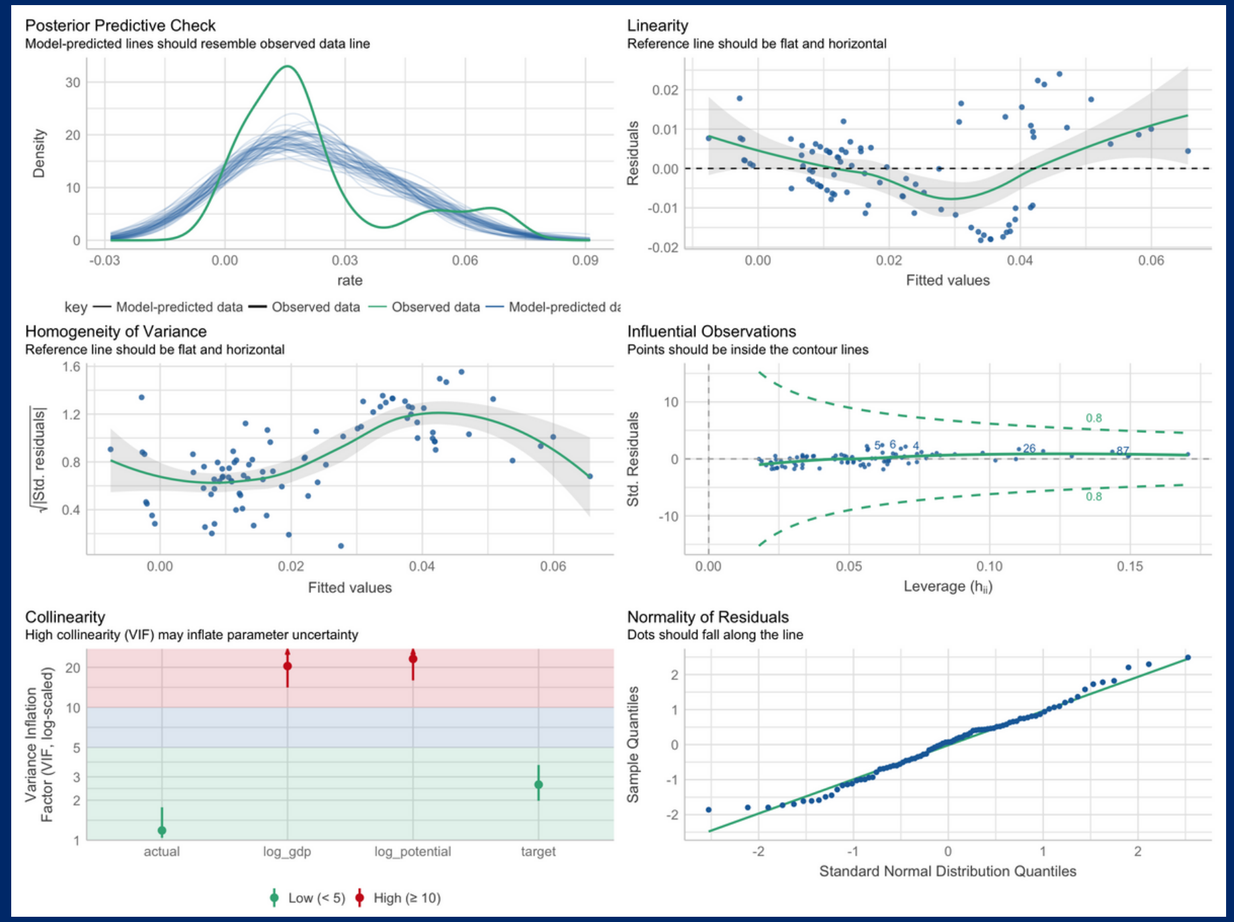
As we can see, heteroscedasticity is shown in only on one of three tests. This is either because of

- Low significance of log_gdp, but inclusion in the model itself
- Collinearity of log_gdp and log_pot (most probable)

OTHER OLS TESTS

Running the command check_model() of the fitted variables in the logarithmic scale we receive this output.

We can instantly notice collinearity between log_gdp and log_potential; which is understandable due to the fact that also the original variables are collinear.



OTHER OLS TESTS

JARQUE-BRERA TEST

```
        Jarque Bera Test

data:  logfit$residuals
X-squared = 0.85228, df = 2, p-value = 0.653


        Skewness

data:  logfit$residuals
statistic = 0.09307, p-value = 0.723


        Kurtosis

data:  logfit$residuals
statistic = 2.5523, p-value = 0.394
```

We reject H0 in all three tests, namely errors are not equally distributed

LM TEST

```
data:  logfit
LM test = 52.949, df = 2, p-value = 3.18e-12
```

Failed to reject H0qNo serial correlation

RESET TEST

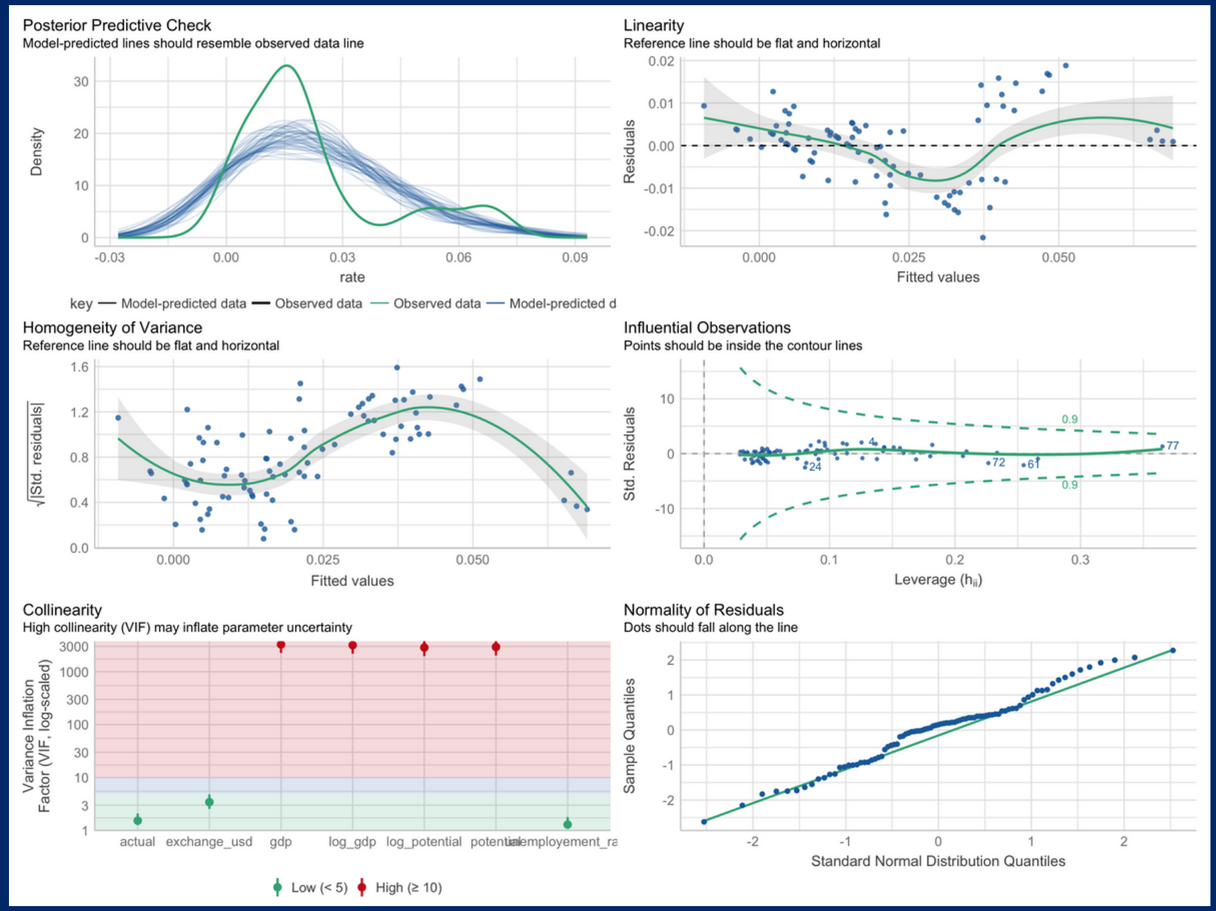
```
data:  logfit
RESET = 10.183, df1 = 2, df2 = 80, p-value = 0.0001148
```

The model failed in rejecting the null hypothesis, it is not correctly specified

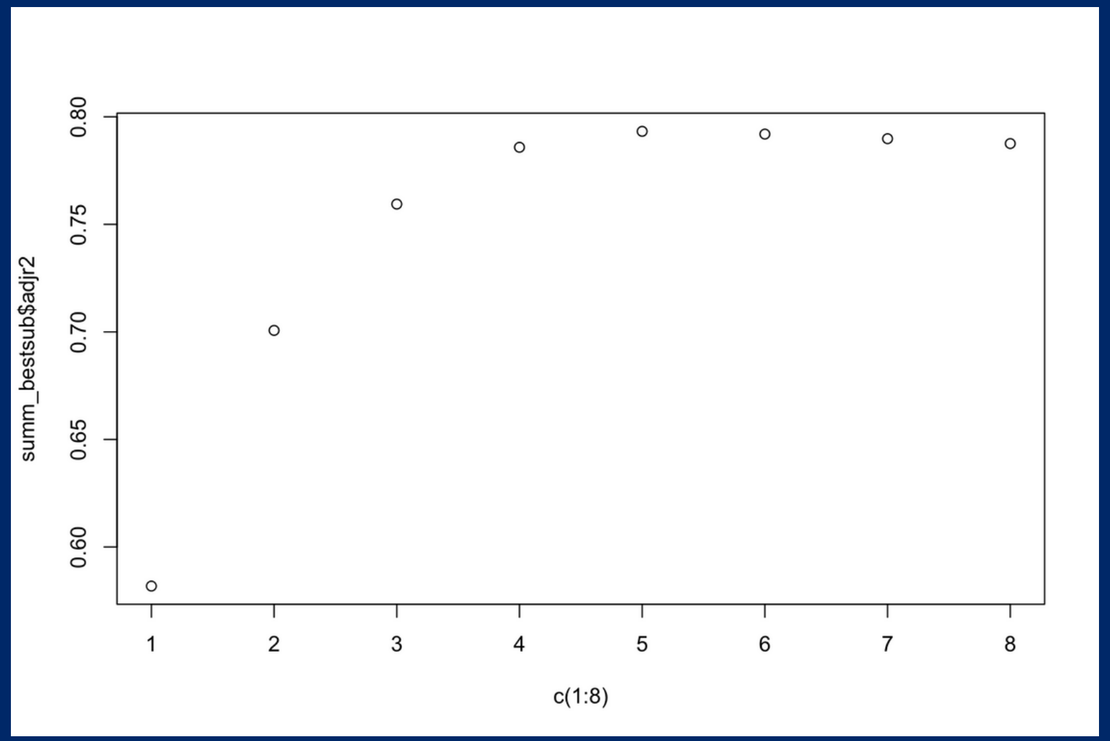
EXTENDED MODEL

We are now performing feature selection on extended model, adding unemployment rate and the rate of exchange NOK/USD. Preliminary tests do not show unexpected problems.

```
Call:
lm(formula = rate ~ actual + log_gdp + log_potential + gdp +
    potential + unemployement_rate + exchange_usd, data = df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.021624 -0.006876  0.001333  0.004298  0.018825

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         2.312e+01  4.327e+00   5.344 8.56e-07 ***
actual              2.387e-01  8.898e-02   2.683 0.008894 **
log_gdp            -5.646e-01  1.312e+00  -0.430 0.668089
log_potential      -3.687e+00  1.175e+00  -3.138 0.002388 **
gdp                 5.110e-07  8.556e-07   0.597 0.552061
potential           1.975e-06  7.470e-07   2.644 0.009887 **
unemployement_rate -4.109e-01  1.125e-01  -3.653 0.000465 ***
exchange_usd        5.418e-02  6.715e-02   0.807 0.422215
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.008906 on 79 degrees of freedom
Multiple R-squared:  0.8069,    Adjusted R-squared:  0.7898
F-statistic: 47.17 on 7 and 79 DF,  p-value: < 2.2e-16
```



EXTENDED MODEL



Performing the best subset selection, we have that for small numbers it is a feasible option
Then, best subset is selected as the one maximizing the adjusted R^2
In this case, maximum value is obtained at 5 features.
The command below allow us to retrieve the selected features and coefficients.

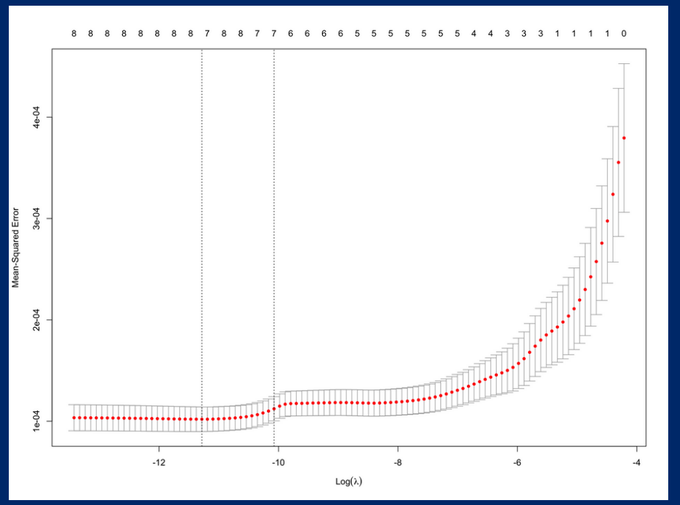
```
> coef(bestsub, 5) #This gets the value and the features
      (Intercept)            actual               gdp         potential     log_potential unemployement_rate
     2.044469e+01      2.544481e-01      1.503713e-07      2.003596e-06     -3.751899e+00     -4.204109e-01
```

LASSO SELECTION

For a double-check, we perform a LASSO selection

We find a lasso_lambda, namely a value that minimizes the MSE.
Them, that lambda is used to compue the coeffieint for the evaluated subsets, trying to drive them completeley to 0 in order to discard them.

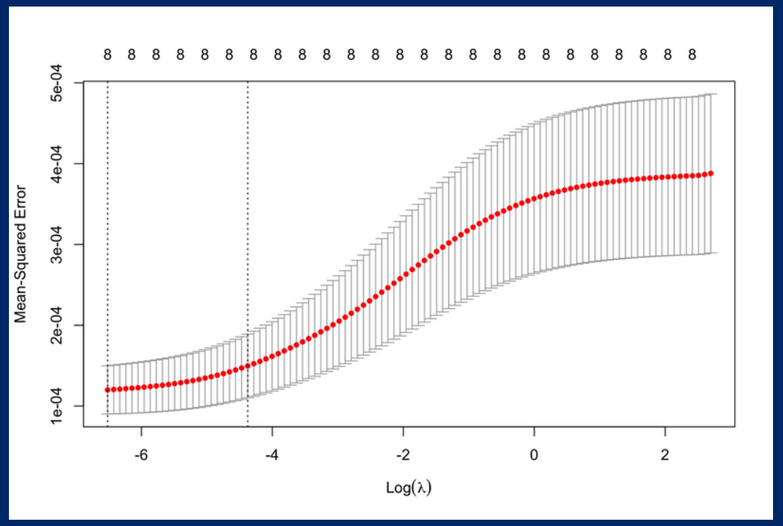
	s0
(Intercept)	2.009955e+01
target	-7.306473e-02
actual	2.648406e-01
gdp	5.972628e-07
potential	1.526477e-06
log_gdp	-7.072841e-01
log_potential	-2.982248e+00
unemployement_rate	-4.025189e-01
exchange_usd	2.125332e-02




```
>
> #find optimal lambda value that minimizes test MSE
> lasso_lambda <- lasso_model$lambda.min
> lasso_lambda
[1] 6.554949e-06
>
> #produce plot of test MSE by lambda value
> plot(lasso_model)
>
> #find coefficients of best model
> best_lasso <- glmnet(x, y, alpha = 1, lambda = lasso_lambda)
> coef(best_lasso)
9 x 1 sparse Matrix of class "dgCMatrix"
                              s0
(Intercept)         2.280840e+01
target              1.972848e-01
actual              2.376344e-01
gdp                 7.923488e-07
potential           1.666387e-06
log_gdp            -9.970352e-01
log_potential      -3.197651e+00
unemployement_rate -4.105175e-01
exchange_usd        4.316473e-02
>
> #calculating R^2: use fitted best model to make predictions
> y_lasso <- predict(lasso_model, s = lasso_lambda, newx = x)
>
> #find SST and SSE
> lasso_sst <- sum((y - mean(y))^2)
> lasso_sse <- sum((y_lasso - y)^2)
>
> #find R-Squared
> lasso_rsq <- 1 - lasso_sse/lasso_sst
> lasso_rsq
[1] 0.8057613
```

RIDGE SELECTION

As double-check, we can use the ridge model, with no discarded coefficient:
the proceding is the same, with the computation of the lambda through k-fold regression. Then the method tries to bring coefficient to 0 so that they can be discarded.

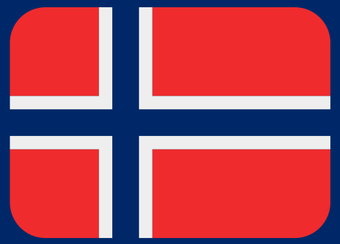



```
(Intercept)          1.693990e+00
target              -1.450366e+00
actual               4.188987e-01
gdp                 -2.301862e-08
potential           -9.217041e-08
log_gdp             -7.106791e-02
log_potential       -1.930986e-01
unemployement_rate  -3.917215e-01
exchange_usd        -8.906520e-02
```

```
> ridge_lambda
[1] 0.001479465
```

```r
#calculating R^2
y_ridge <- predict(ridge_model, s = ridge_lambda, newx = x)

ridge_sst <- sum((y - mean(y))^2)
ridge_sse <- sum((y_ridge - y)^2)

ridge_rsq <- 1 - ridge_sse/ridge_sst
ridge_rsq
```

```
> ridge_rsq
[1] 0.7379339
```



THANK YOU

You can find here the complete R code and the dataset used

Sources

- https://stats.oecd.org/#
- https://ec.europa.eu/eurostat/data
- https://www.norges-bank.no/en/topics/Monetary-policy/Policy-rate/
- https://www.norges-bank.no/en/topics/Monetary-policy/

https://roccototaro.notion.site/Econometrics-Group-Project-ceb7f00500e0485098dd979bdafdf8b6