HOMEWORK

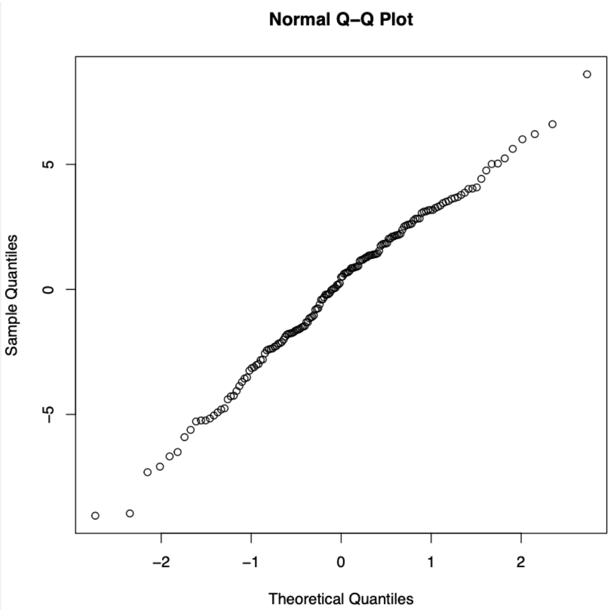
5.5

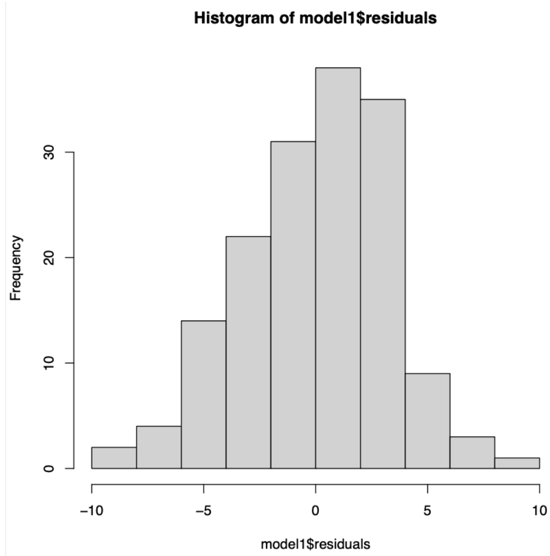
a)

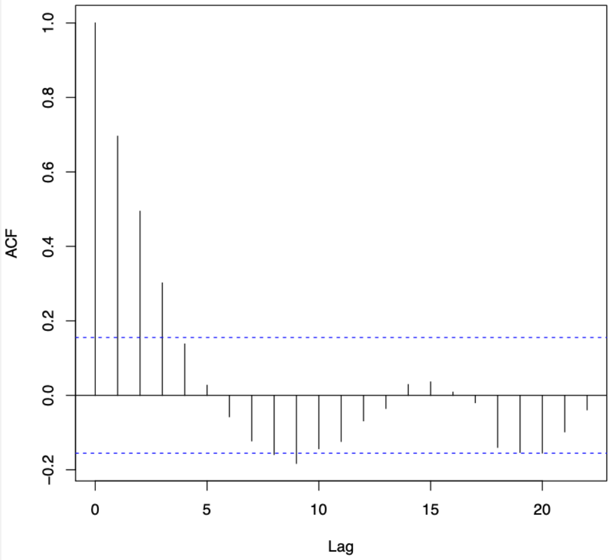
```
library(car)
library(forecast)
library(lmtest)
t38 = read.table("t38.txt",header=T,sep="\ ")
consumption = ts(t38$"CONS")
income = ts(t38$"INC")
inflation = ts(t38$"INFLAT")

model1 = lm(consumption~income + inflation)
summary(model1)

qqnorm(model1$residuals)
hist(model1$residuals)
acf(model1$residuals,main="")
plot(model1$residuals,type="p", ylab="Residuals", xlab="Quarter", pch=16, xaxt="n")
fit <- ts(predict(model1))

plot(consumption,col="red",xlab="Time",ylab="")
lines(fit,col="blue")
plot(cooks.distance(model1),ylab="Cook's distances",xlab="Index",
    pch=16)
vif(model1)

#blue is the fitted model
```

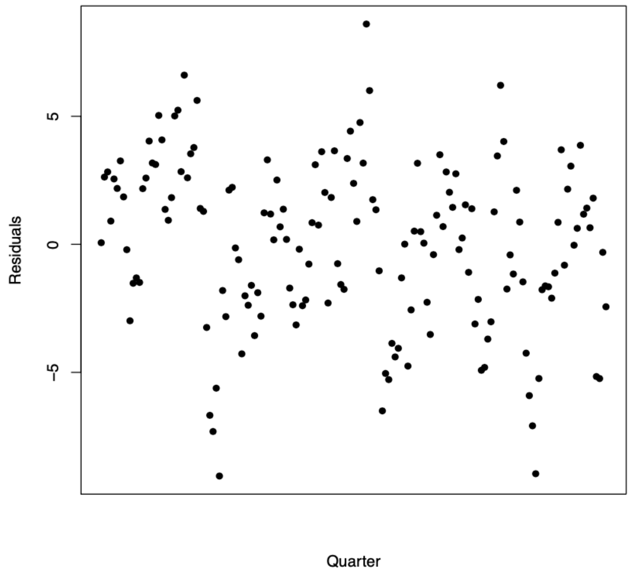
```
Call:
lm(formula = consumption ~ income + inflation)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.0491 -2.1273  0.4948  2.3026  8.6025 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -147.38977   21.71937  -6.786 2.26e-10 ***
income         1.15263    0.02431  47.420  < 2e-16 ***
inflation     -2.47468    0.20268 -12.210  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 3.267 on 156 degrees of freedom
Multiple R-squared:  0.9425,  Adjusted R-squared:  0.9418 
F-statistic:  1279 on 2 and 156 DF,  p-value: < 2.2e-16
```

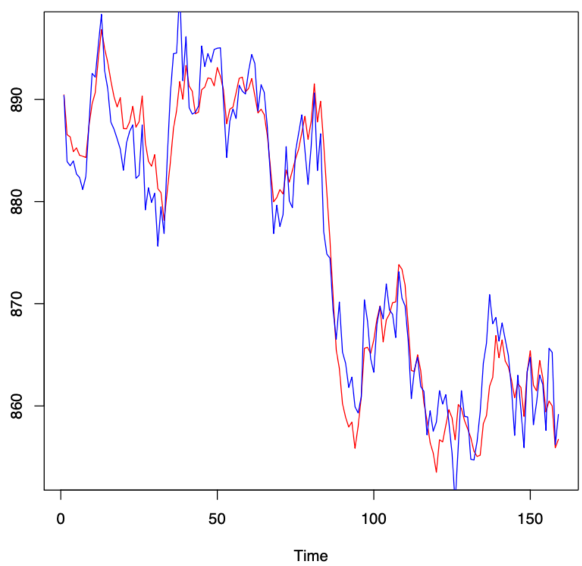
We can see that the explanatory variables are statistically significant with 5% level of significance. Since we are explaining consumption, as expected the sign of inflation is negative and positive for income. The coefficient of determination is 0.9418. If income goes up 1 %, consumption goes up 1.15%. Contextually interpretation for inflation is quite complicated since even though the model says that inflation drives consumption down we know that that isn't usually the case since ~2% yearly inflation is optimal.

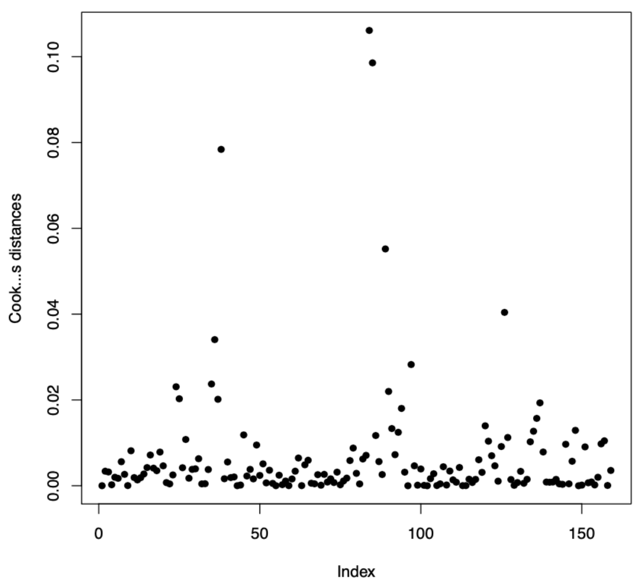

Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles


Histogram of model1$residuals
Frequency
model1$residuals


ACF
Lag


Residuals
Quarter


Time


Cook...s distances
Index

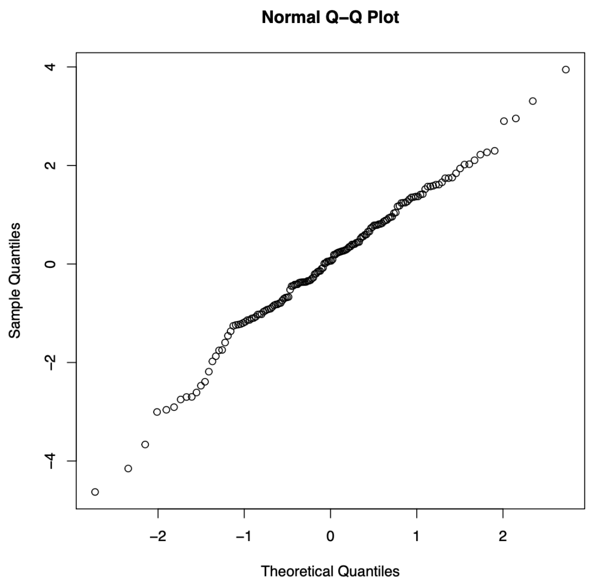
Residuals seem correlated and maybe normally distributed but snot sure.

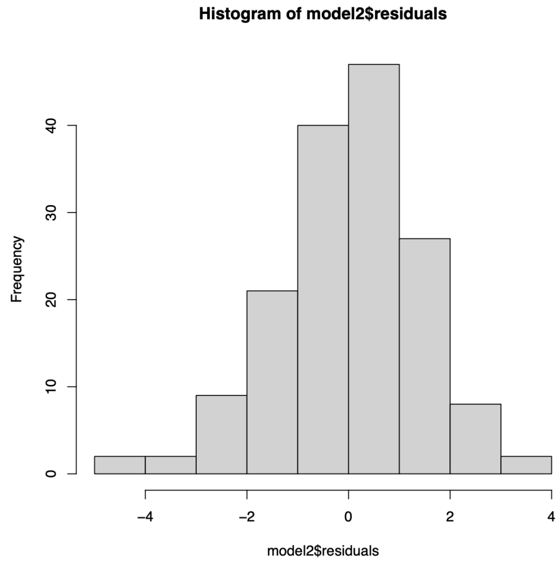
b)

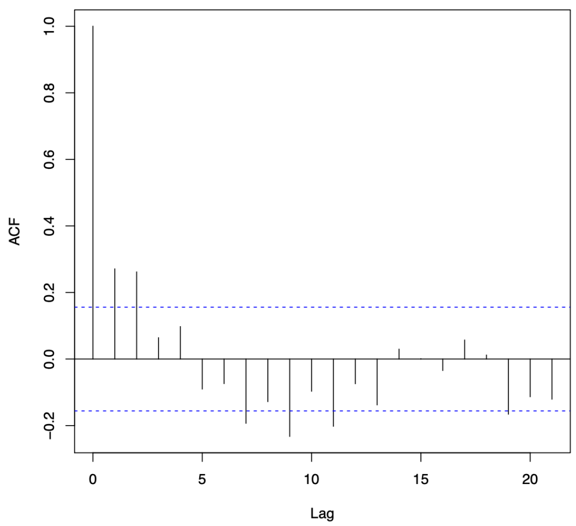
```
dconsumption = diff(consumption)
dincome = diff(income)
dinflation = diff(inflation)

model2 = lm(dconsumption~dincome + dinflation)
summary(model2)

qqnorm(model2$residuals)
hist(model2$residuals)
acf(model2$residuals,main="")
plot(model2$residuals,type="p",ylab="Residuals",xlab="Quarter",pch=16,
    xaxt="n")
fit = ts(predict(model2))

plot(dconsumption,col="red",xlab="Time",ylab="")
lines(fit,col="blue")
plot(cooks.distance(model2),ylab="Cook's distances",xlab="Index",
    pch=16)
vif(model2)
```

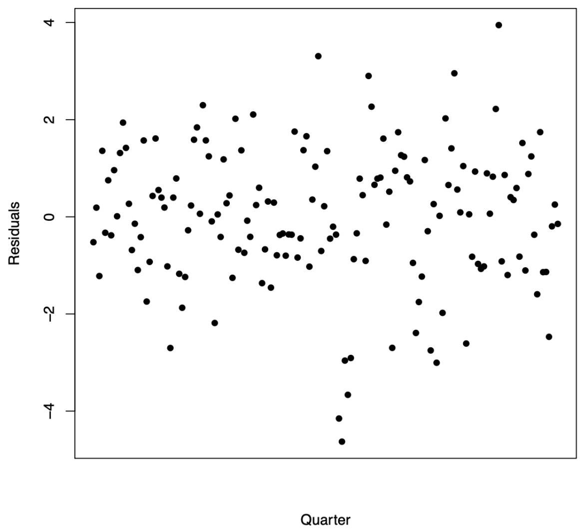
```
Call:
lm(formula = dconsumption ~ dincome + dinflation)

Residuals:
    Min      1Q  Median      3Q     Max
-4.6284 -0.8637  0.0631  0.9223  3.9466

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.11968    0.11453  -1.045  0.29764
dincome      0.51830    0.03527  14.696  < 2e-16 ***
dinflation  -0.71594    0.23934  -2.991  0.00323 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.437 on 155 degrees of freedom
Multiple R-squared:  0.5826,  Adjusted R-squared:  0.5772
F-statistic: 108.2 on 2 and 155 DF,  p-value: < 2.2e-16
```

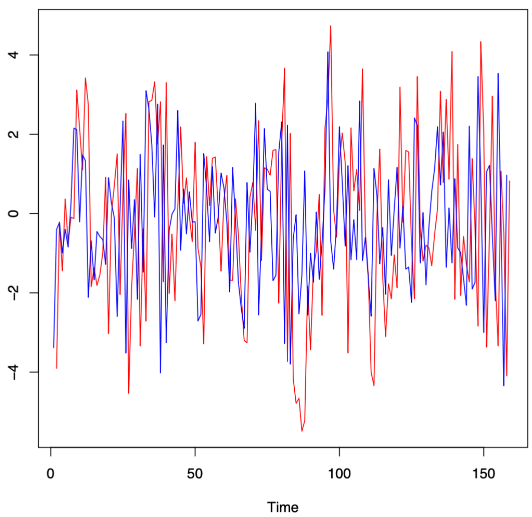
We can see that the explanatory variables are still statistically significant with 5% level of significance. The signs of coefficients for income and inflation are still the same. The coefficient of determination is now however 0.5772. Correlation of the residuals is apparent and they seem quite normally distributed.

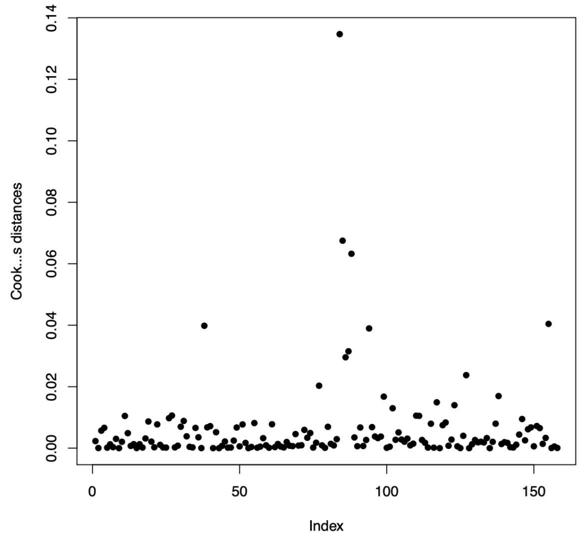

Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles


Histogram of model2$residuals
Frequency
model2$residuals


ACF
Lag


Residuals
Quarter

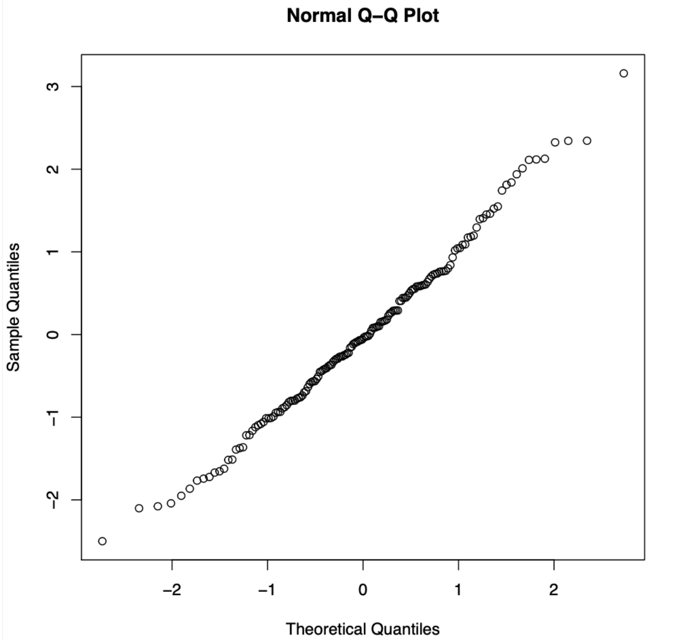

Time

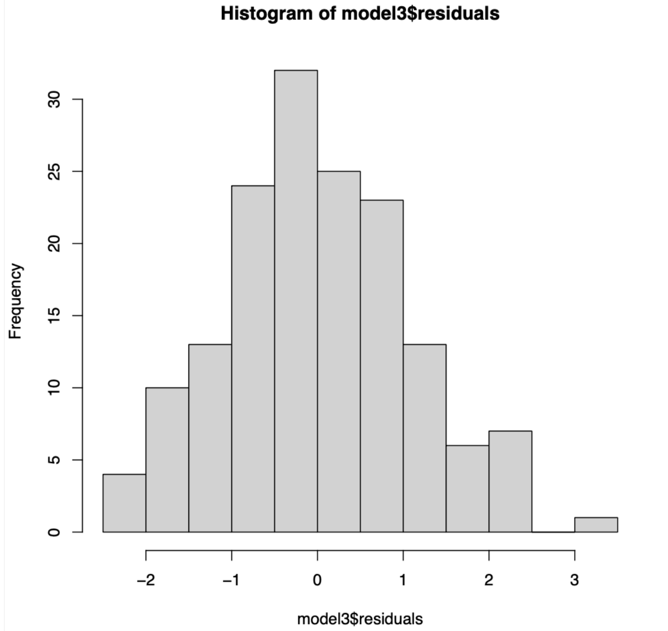

Cook...s distances
Index

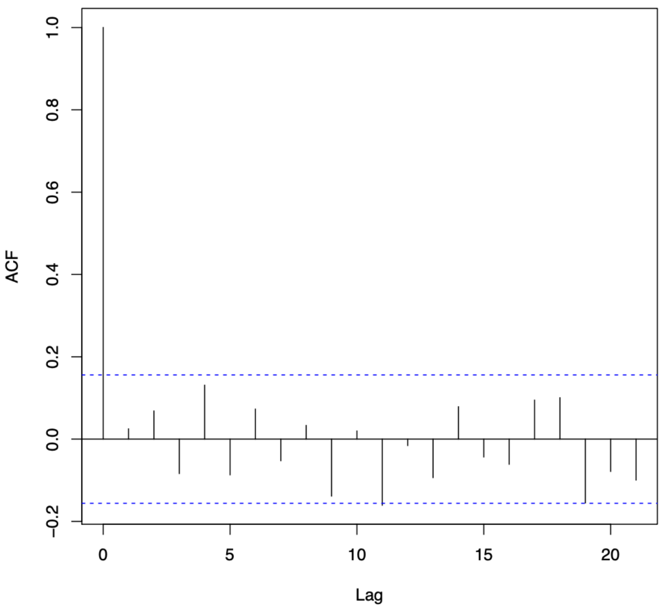
c)

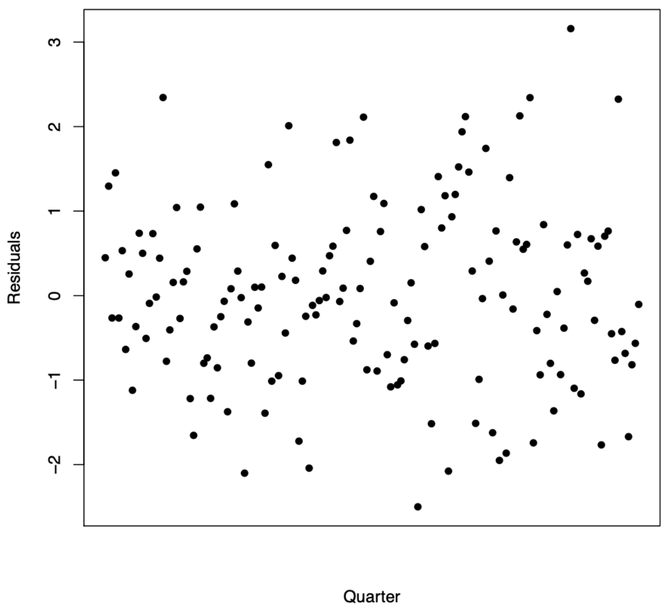
```
n = nrow(t38)
model3 = lm(consumption[-1]~ consumption[-n]+income[-1] + income[-n] + inflation[-1]+ inflation[-
n])
summary(model3)

qqnorm(model3$residuals)
hist(model3$residuals)
acf(model3$residuals,main="")
plot(model3$residuals,type="p",ylab="Residuals",xlab="Quarter",pch=16,
    xaxt="n")
fit = ts(predict(model3))

plot(consumption,col="red",xlab="Time",ylab="")
lines(fit,col="blue")
plot(cooks.distance(model3),ylab="Cook's distances",xlab="Index",
    pch=16)
vif(model3)

Call:
lm(formula = consumption[-1] ~ consumption[-n] + income[-1] +
    income[-n] + inflation[-1] + inflation[-n])

Residuals:
     Min       1Q   Median       3Q      Max
-2.49953 -0.76349 -0.04695  0.62801  3.15931

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     -20.26950    8.52628  -2.377   0.0187 *
consumption[-n]   0.79831    0.02716  29.393  < 2e-16 ***
income[-1]        0.49894    0.02833  17.611  < 2e-16 ***
income[-n]       -0.27611    0.03788  -7.290 1.59e-11 ***
inflation[-1]    -0.79309    0.18395  -4.311 2.90e-05 ***
inflation[-n]    -0.25061    0.20310  -1.234   0.2191
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.078 on 152 degrees of freedom
Multiple R-squared:  0.9939,	Adjusted R-squared:  0.9937
F-statistic:  4915 on 5 and 152 DF,  p-value: < 2.2e-16
```

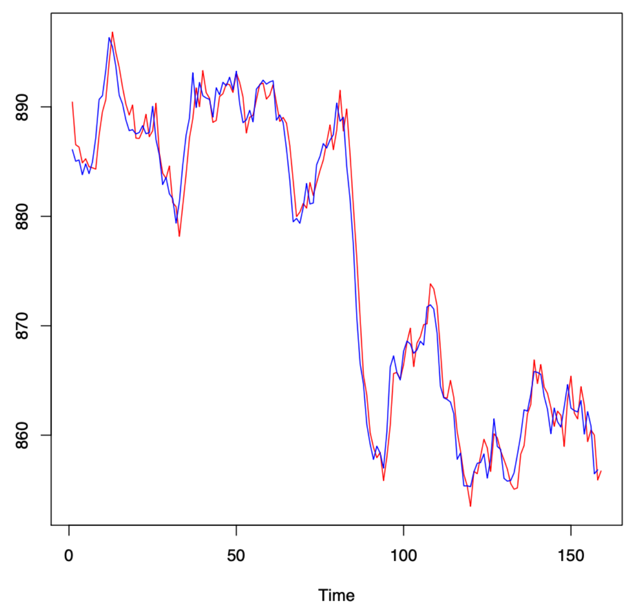
We can see that the explanatory variables are statistically significant with 5% level of significance. The signs of coefficients for income and inflation are now a bit more interesting. The sign for income even changes from t-1 to t which can tell something about seasonality. The coefficient of determination is now 0.9937 Correlation of the residuals is apparent and they seem quite normally distributed.

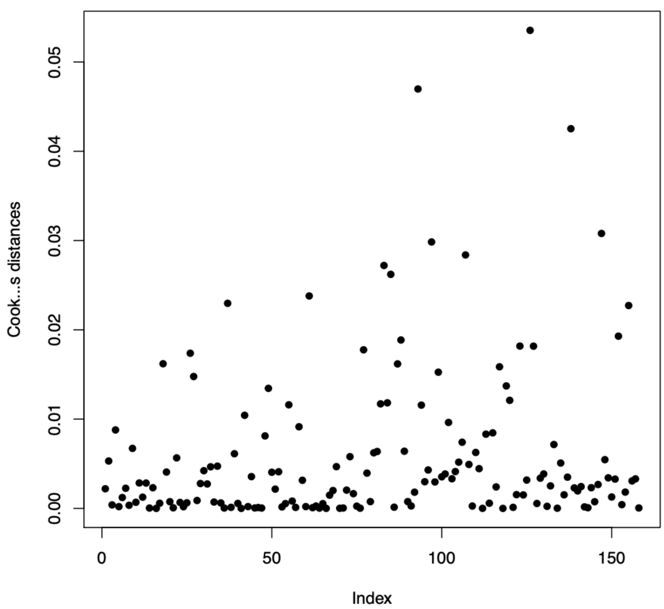

Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles


Histogram of model3$residuals
Frequency
model3$residuals


ACF
Lag


Residuals
Quarter


Time


Cook...s distances
Index

d)

In my opinion the last model is the best in explaining the response variable. I would choose it for the sake of the longer “memory” due to lags. All the plots seem reasonably well fitting and there isn’t anything suspicious going on.