```r
coen <-read.csv(file.choose(),sep = "")
library("dynlm","lmtest")
coen <- ts(coen, start = c(1948,1), end = c(1966,4), frequency = 4)
```

Let's start with the regression of UK government security index as regressor and UK car production index as dependent variable:

```r
regr <- dynlm(CAR~GSX,coen)
summary(regr)

##
## Time series regression with "ts" data:
## Start = 1948(1), End = 1966(4)
##
## Call:
## dynlm(formula = CAR ~ GSX, data = coen)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -142.519  -55.545   -3.289   57.204  180.658
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 948.3934    70.7993   13.40  < 2e-16 ***
## GSX          -7.7624     0.7752  -10.01 2.09e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 78.29 on 74 degrees of freedom
## Multiple R-squared:  0.5753,Adjusted R-squared:  0.5696
## F-statistic: 100.3 on 1 and 74 DF,  p-value: 2.086e-15

AIC(regr)

## [1] 882.4261
```

Let's run the Breusch-Godfrey test:

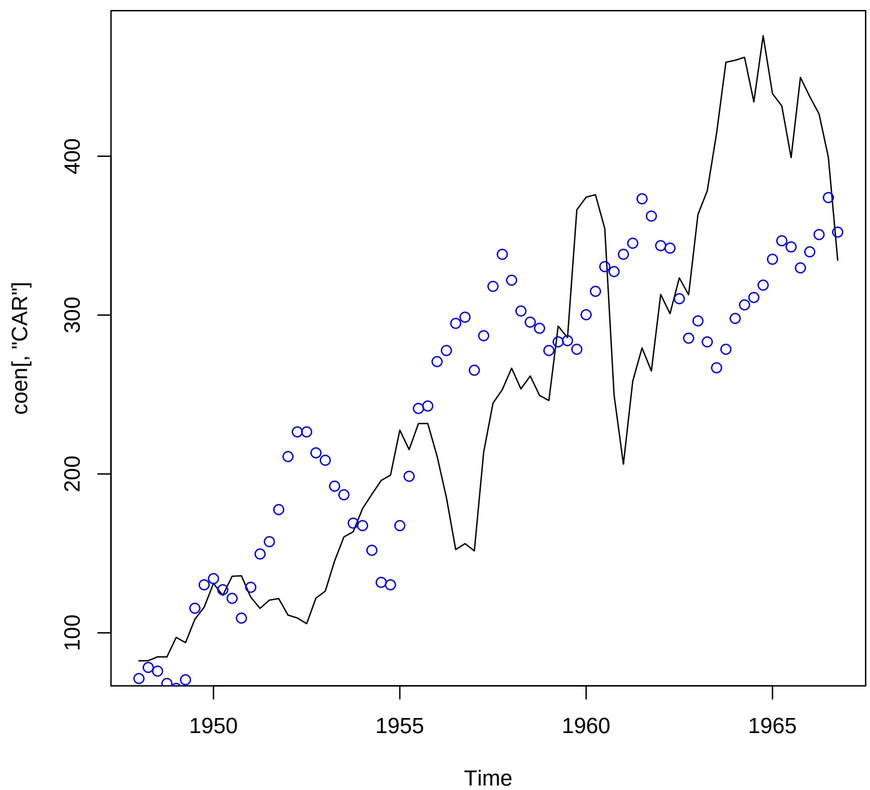
```r
lmtest::bgtest(regr)

##
##  Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  regr
## LM test = 63.275, df = 1, p-value = 1.798e-15
```

The test shows that null hypothesis of no autocorrelation should be rejected.

Let's give visual interpretation:

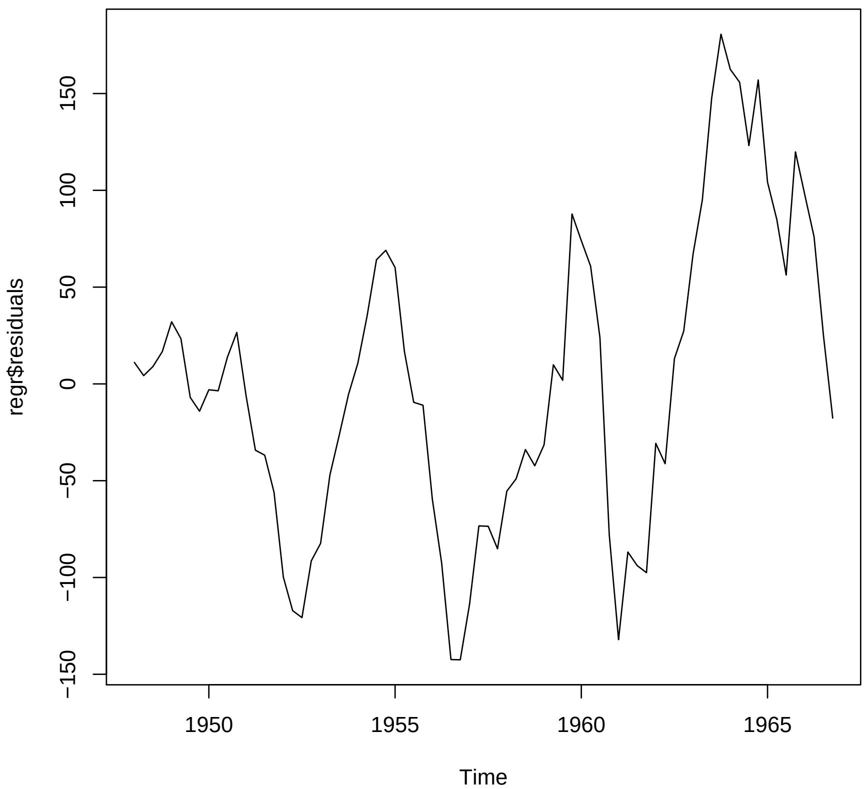
```
plot(coen[,"CAR"])
points(regr$fitted.values,col="blue")
```




It's obvious that the fitted values do not correspond well to the actual values. The revealed autocorrelation is a good candidate for the reason of it.

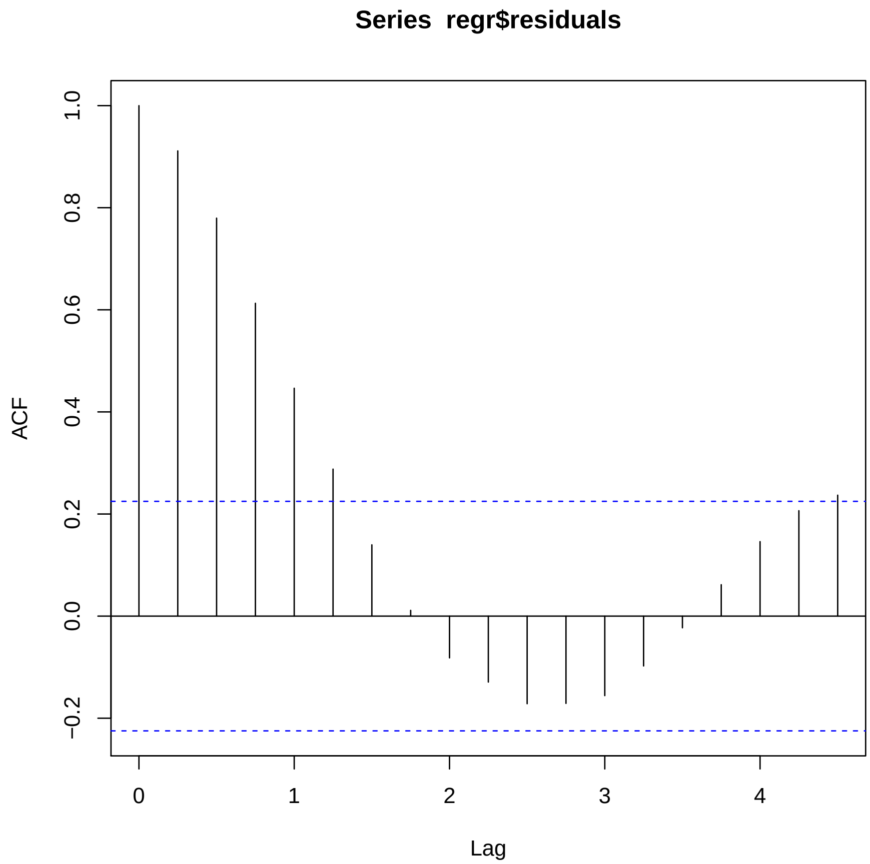
So let's look at the plot of residuals of the model:

```
plot(regr$residuals)
```

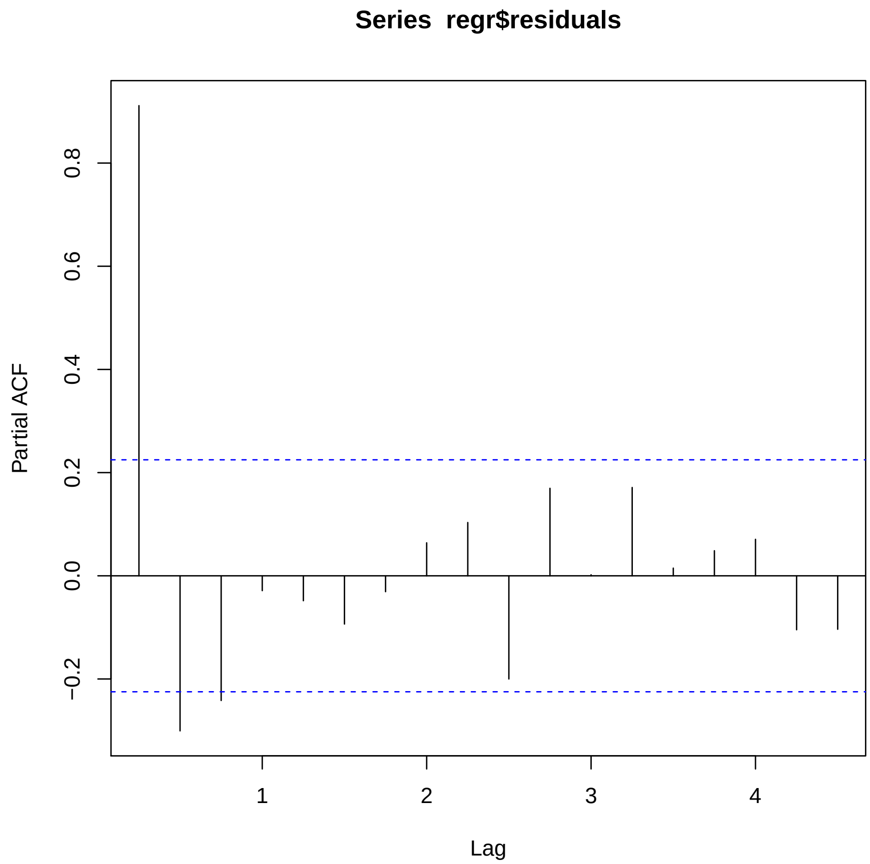



It can be seen that the residuals are not stationary and they have different trends depending on the time period. This graph also shows the autocorrelation among lags of residuals:

```
acf(regr$residuals)
```




```
pacf(regr$residuals)
```




To overcome autocorrelation, let's include lagged dependent variables in our model. Based on acf and pacf results, we can conclude that lags 1 and 2 periods could be used in the model. Let's try.

```
regr1 <- dynlm(CAR~L(CAR,1) + L(GSX,1) + L(GSX,2),coen)
summary(regr1)

##
## Time series regression with "ts" data:
## Start = 1948(3), End = 1966(4)
##
## Call:
## dynlm(formula = CAR ~ L(CAR, 1) + L(GSX, 1) + L(GSX, 2), data = coen)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -99.108 -13.828  -2.089  16.281  76.614
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 28.84608   46.02878   0.627   0.5329
## L(CAR, 1)    0.95361    0.04024  23.698   <2e-16 ***
## L(GSX, 1)    3.67342    1.48339   2.476   0.0157 *
## L(GSX, 2)   -3.80620    1.44140  -2.641   0.0102 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 27.08 on 70 degrees of freedom
## Multiple R-squared:  0.9494,Adjusted R-squared:  0.9472
## F-statistic: 437.5 on 3 and 70 DF,  p-value: < 2.2e-16
```

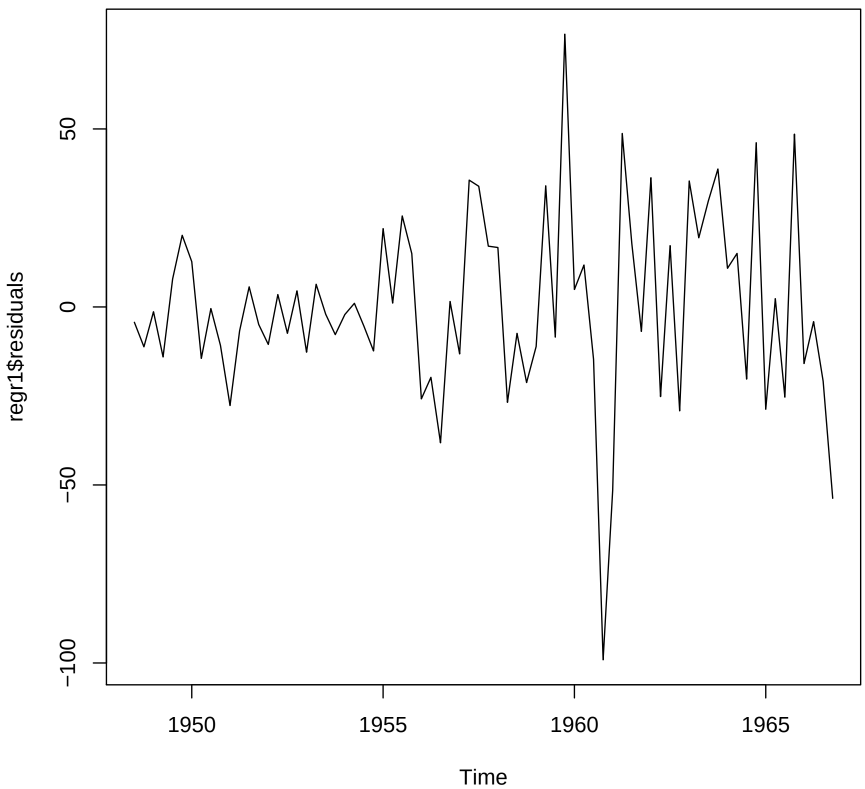
```
lmtest::bgtest(regr1)
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  regr1
## LM test = 0.39642, df = 1, p-value = 0.5289
```

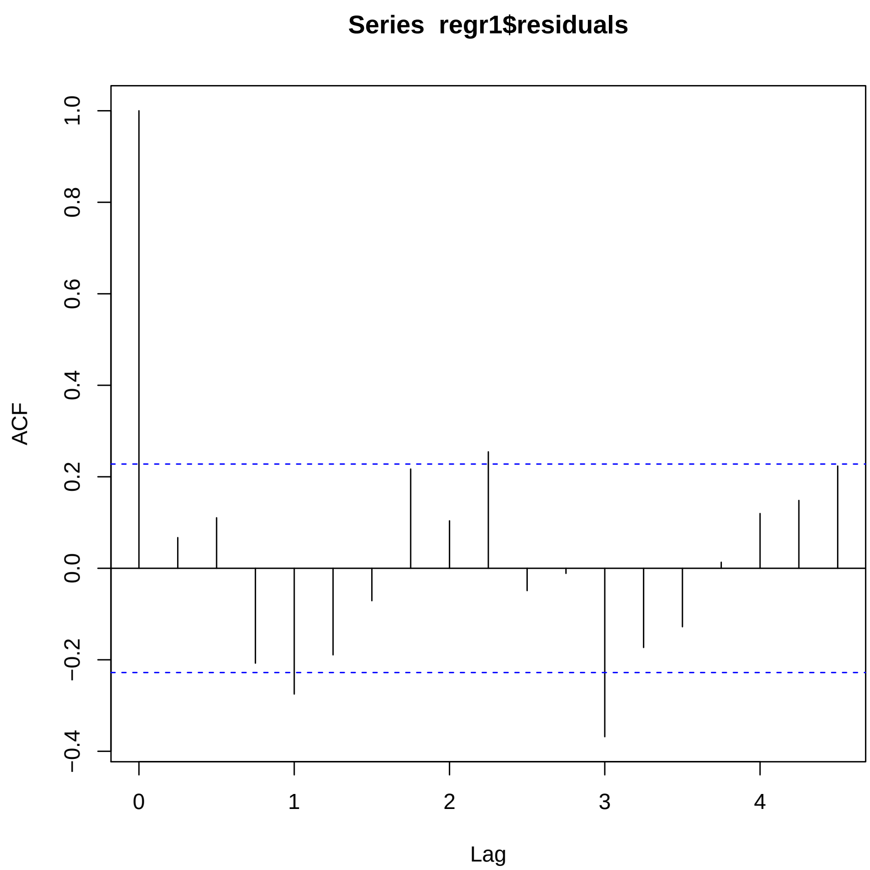
```
AIC(regr1)
```

```
## [1] 704.0896
```

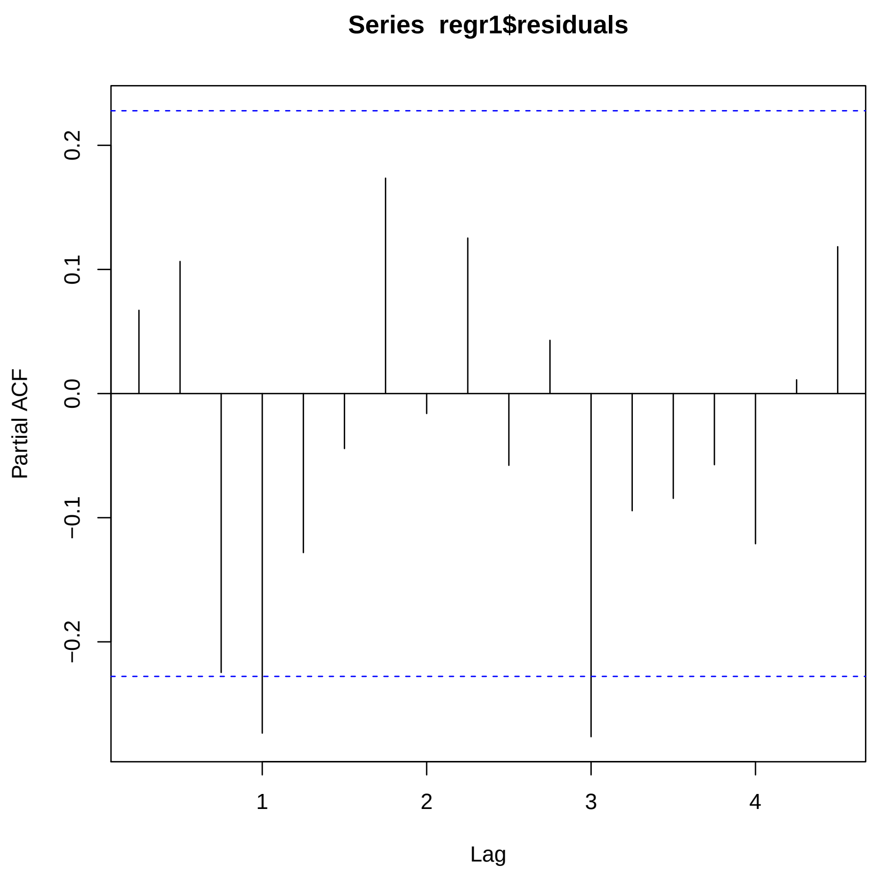
```
plot(regr1$residuals)
```




```r
acf(regr1$residuals)
```



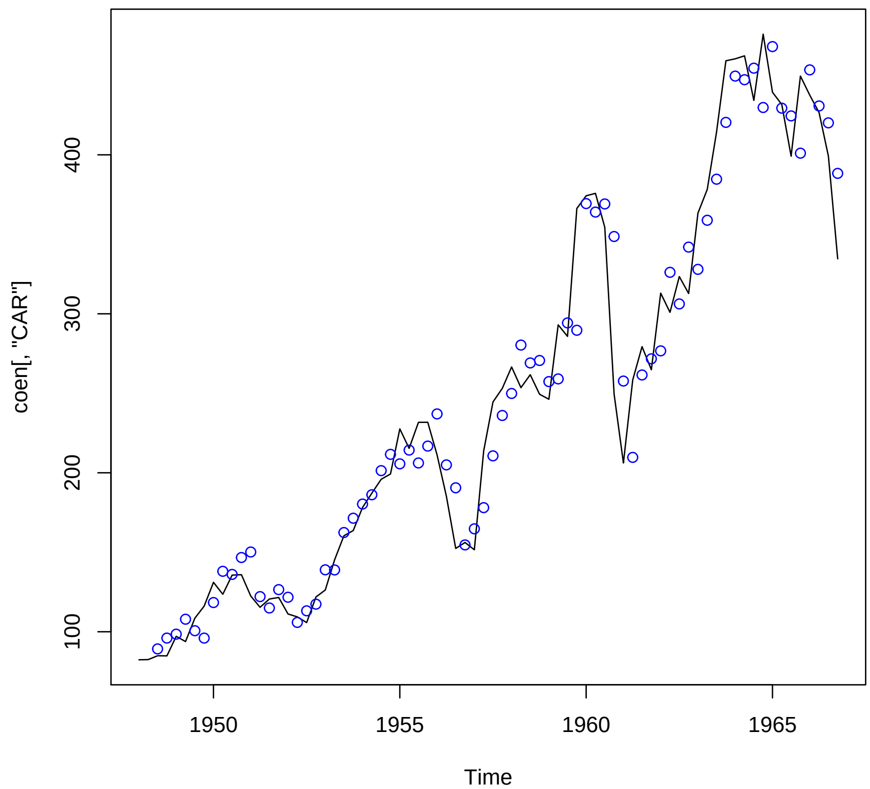

```
pacf(regr1$residuals)
```




The Breusch-Godfrey test now shows that autocorrelation is eliminated. The graphs show the same. It is also noticeable that AIC decreased, so model became better.

Let's examine how fitted values correspond to the actual values after including lagged dependent variables:

```
plot(coen[,"CAR"])
points(regr1$fitted.values,col="blue")
```




Clearly, fitted values now are very close to the actual, so the model was improved considerably.