Final Exam

Aliaksandr Panko

August 22, 2018

Contents

1 Task 1: Stationarity of L

To check stationarity Augmented Dickey-Fuller test is used. HO - nonstationary(unit root exists)

```
adf_test <- tseries::adf.test(data.fe$L,alternative = 'stationary')
print(adf_test)
```

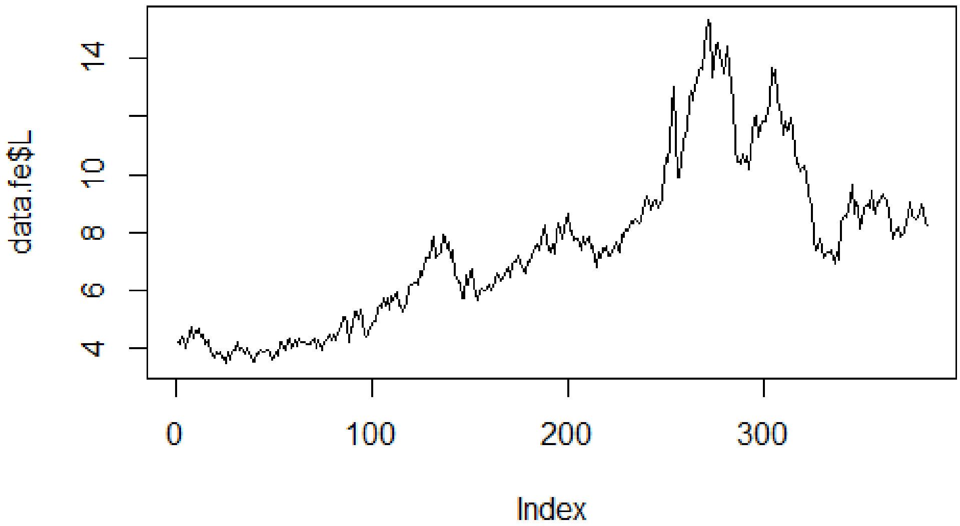
```
        Augmented Dickey-Fuller Test

data:  data.fe$L
Dickey-Fuller = -1.6791, Lag order = 7, p-value = 0.7126
alternative hypothesis: stationary
```

Based on the test we cannot reject H0. This means that we assume that the data are not stationary.

Let's hava a look at the serias:

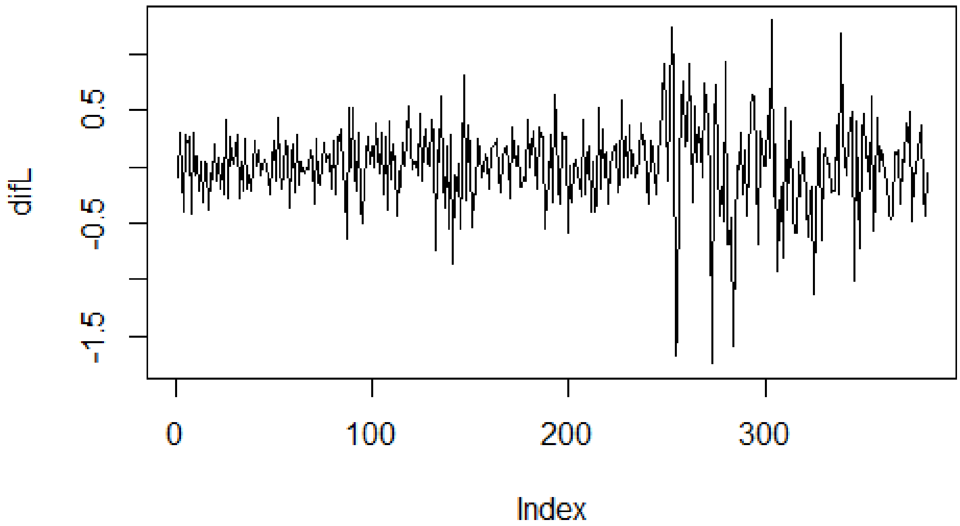
```
plot(data.fe$L, type = "l")
```



We can clearly see two trends, which proves nonstationarity.

To continue we need to take differences: Take differencies:

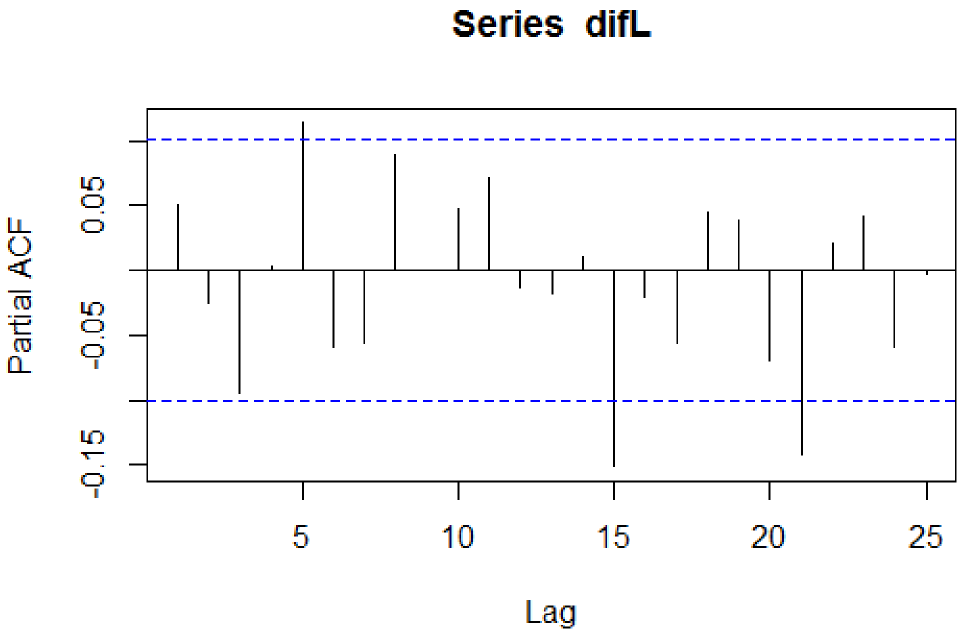
```
difL <- diff(data.fe$L)
plot(difL, type = "l")
```

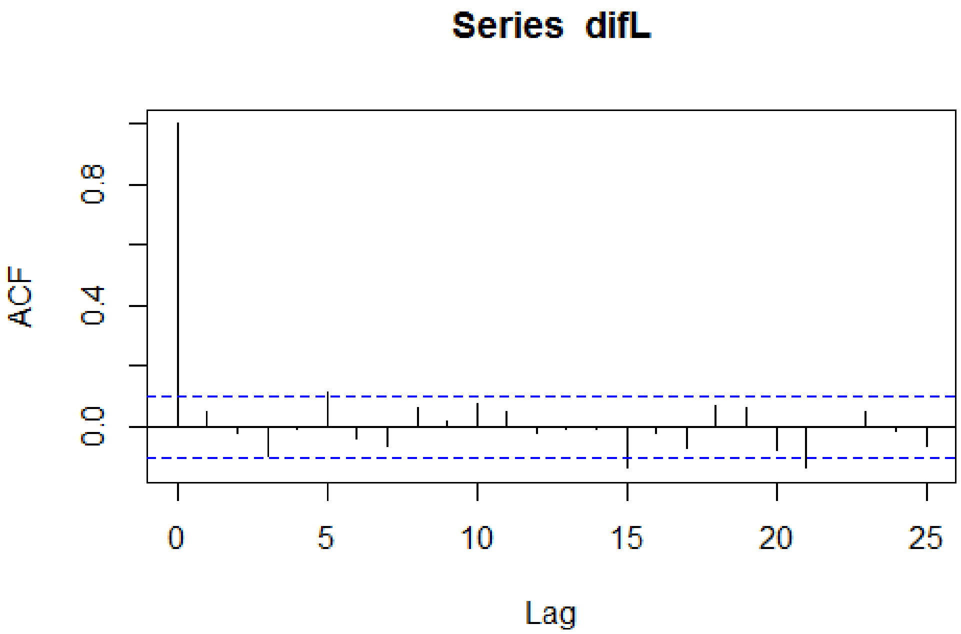


Is looks much better.

Let's have a look at acf and pacf

```
acf(difL)
pacf(difL)
```





Choose the best model automatically

```
forecast::auto.arima(difL, seasonal=FALSE)
```

```
Series: difL
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.1439:  log likelihood=-171.78
AIC=345.56   AICc=345.57   BIC=349.51
```

The function gives us the same result as we could see ARIMA (0,0,0). So, basically the difference of long run rates is noise.

2 Task 2 : Model for a stock price index

Firstly, we need to determine all regressors that could be important for the model since omitted regressors have critical impact.

In my opinion all of these factors are potentially possible, so I will use iterative process to eliminate insignificant ones.

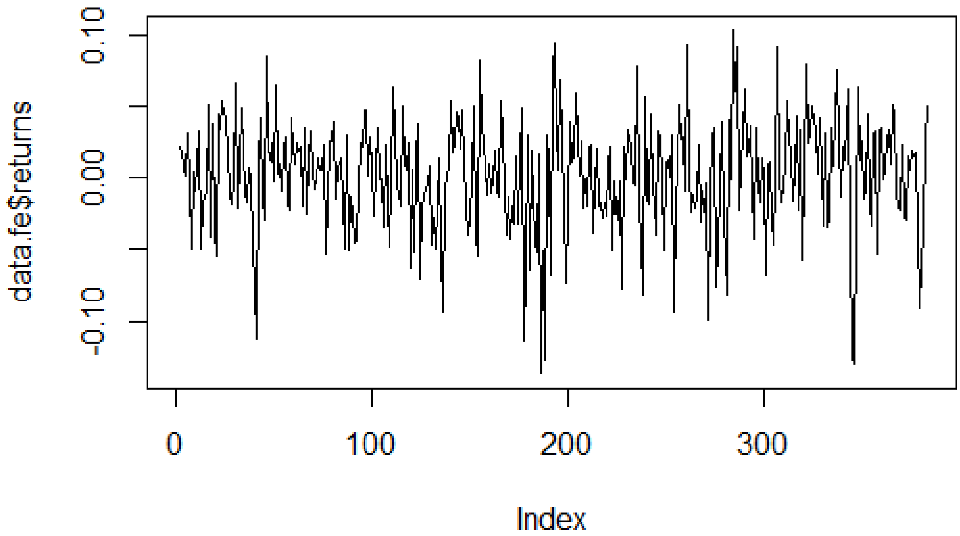
So, initially use all the regressors to avoid ommited ones.

Since all questions are about returns we will derive returns from the price and work with them.

Calculate returns using price X

```
data.fe$returns = NA
for(i in 2:nrow(data.fe))
{
data.fe$returns[i] = (data.fe$X[i] - data.fe$X[i-1]) / data.fe$X[i-1]
}

plot(data.fe$returns, type = "l")
```



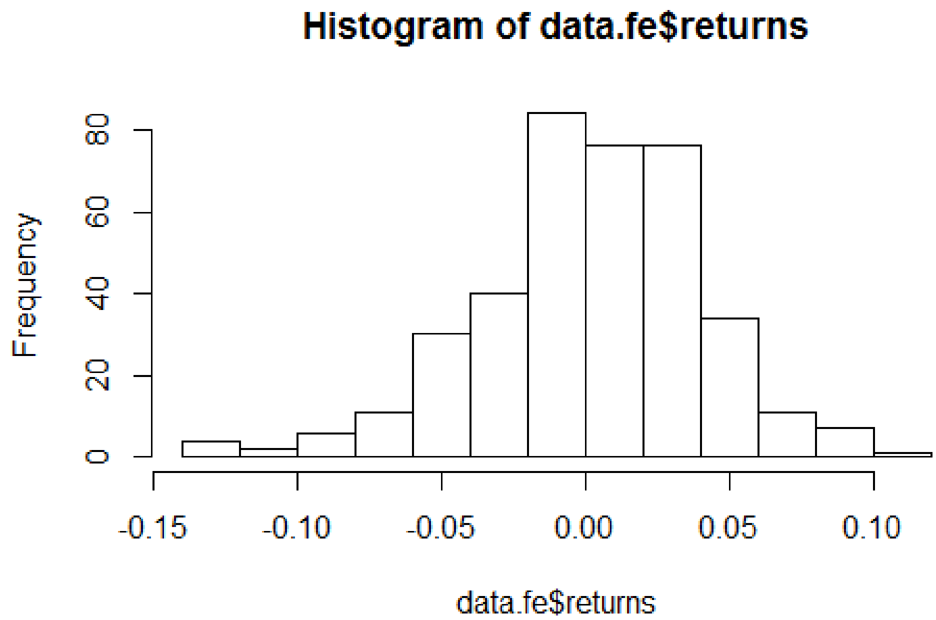
Check Normal Distribution. H0 - distribution is normal

```
shapiro.test(data.fe$returns)
```

```
        Shapiro-Wilk normality test

data:  data.fe$returns
W = 0.98132, p-value = 7.53e-05
```

```
hist(data.fe$returns)
```



We can see that they are skewed to the right.

So, we can see that returns are not normal.

Based on iterative elimination of regressors we have the next model:

```
model = lm(data = data.fe, formula = data.fe$returns ~ data.fe$L +
#data.fe$D +
#data.fe$E +
data.fe$C : data.fe$P +
data.fe$S)
summary(model)
```

```
Call:
lm(formula = data.fe$returns ~ data.fe$L + data.fe$C:data.fe$P +
    data.fe$S, data = data.fe)

Residuals:
      Min        1Q    Median        3Q       Max
-0.143677 -0.023212  0.003197  0.024922  0.109108

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           0.040767   0.020816   1.958  0.05091 .
data.fe$L             0.004133   0.002800   1.476  0.14073
data.fe$S            -0.006893   0.002169  -3.177  0.00161 **
data.fe$C:data.fe$P0 -0.031408   0.014862  -2.113  0.03523 *
data.fe$C:data.fe$P1 -0.025285   0.015044  -1.681  0.09364 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03829 on 377 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.05535,   Adjusted R-squared:  0.04533
F-statistic: 5.523 on 4 and 377 DF,  p-value: 0.0002491
```

The model is fairly bad wirh $R\hat{2}$ only around 4%

Let's execute Breusch-Pagan test to check the homoscedasticity: HO - homoscedasticity

```
lmtest::bptest(model)
```

```
        studentized Breusch-Pagan test

data:  model
BP = 4.6039, df = 4, p-value = 0.3304
```

We cannot reject H0 of homoscedasticity at 5% level of significance

Check autocorrelation. HO - no autocorrelation

```
Box.test(resid(model1),type="Ljung",lag=20,fitdf=1)
```

```
        Box-Ljung test

data:  resid(model)
X-squared = 17.892, df = 19, p-value = 0.5297
```

The test says we don't have autocorelation here at 5% level of significance

So the result is: President is not significant and the claim from 3b is not true too