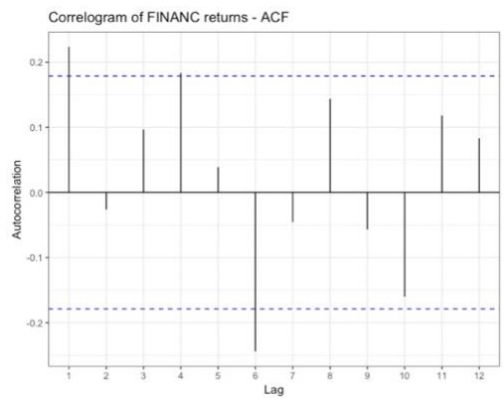
8,154,1.00 financial econometrics

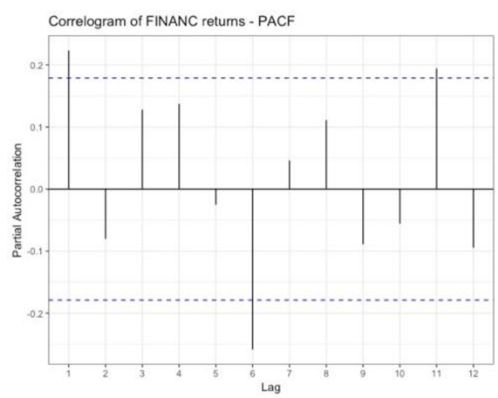
hw3: answer sheet

Nicolò Ceneda, Claire Fromholz, Pietro Gnecco

0. Before starting the exercise, we computed the excess returns of SP500 and FINANC and the logarithm of the price-dividend ratio. We then removed all data points before year 2000. We report the magnitude of the results consistent with the given data.

1. Using data until December 2009, we constructed the correlogram of the excess return of FINANC, for both the autocorrelation and the partial autocorrelation functions, up to 12 lags. The excess return of the S&P Financial Select Sector index displays a statistically significant (at the 95% confidence level) positive autocorrelation coefficient at lags 1 and 4, while it displays a negative coefficient at lag 6. The values are also economically significant, so that they may allow to make some inferences about the level of future returns. The correlogram of the PACF displays significant positive coefficients at lags 1 and 11, while it displays a negative coefficient at lag 6. As can be noticed from the two plots, autocorrelation and partial autocorrelation are identical at displacement 1, as there are no earlier lags to control for when computing the sample partial autocorrelation. Moreover, since we now consider an autoregressive model, we can interpret the last significant lag as the order of the model, which in this this case is 11. By increasing the number of lags included, it is possible to verify that this result does not change.





2. The Akaike's Information Criterion is a measure of goodness of a model that allows us to identify the optimal number of parameters, by accounting for the trade-off between accuracy (as measured by the variance of the fitted residuals) and complexity of the regression (as measured by the number of parameters). Here we report output of the function implemented in R:

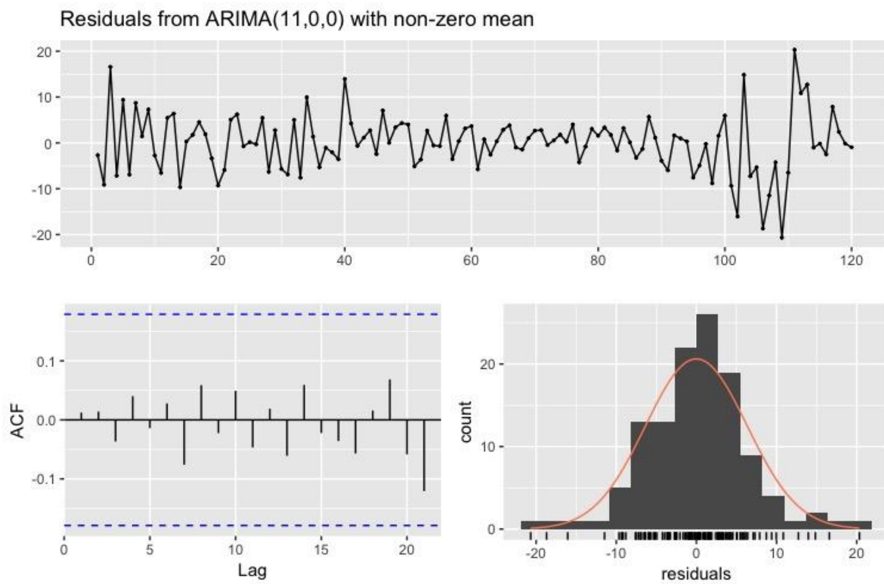
```
Order selected 11  sigma^2 estimated as  39.96
```

As expected, the result obtained is the same as the one inferred from the PACF plot: 11. This means that an autoregressive model with 11 lags describe the data best. However, we should keep in mind that, generally speaking, the AIC often exaggerates the lag length.

3. We now estimate a demeaned autoregressive model of order 11.

	ar1	ar2	ar3	ar4	ar5	ar6	ar7	ar8	ar9	ar10	ar11	mean
Coefficients	0.282	-0.006	0.068	0.023	0.149	-0.329	0.084	0.048	-0.047	-0.216	0.308	-0.689
AIC	811.43											

We then plotted the following charts for the residuals of the AR(11) model:



The model seems to be adequate, as the residuals are clearly uncorrelated, as visible from their autocorrelation function which is null at all lags greater than zero and centered around zero. Therefore, it looks like the residuals can be described by a white noise, although not Gaussian, as visible from the histogram of the residuals. Notice however, that the assumption of normality (and constant variance) is not necessary for the maximum likelihood method to work well, as this estimator behaves asymptotically nice even under these conditions (quasi maximum likelihood).

We then run a Ljung-Box test for the joint significance of the autocorrelation coefficients of the residuals of the autoregressive model:

Q*	df	p-value	Model df	Total lags used
3.527	3	0.317	12	15

The null hypothesis of the test is that the data are independently distributed (i.e. the correlations are 0. The Ljung-Box test statistic can be compared to the critical value of 7.261 for a 5% significance level and 15 degrees of freedom. Since Q* is smaller than this value, we fail to reject the null hypotheses of randomness. Therefore, we confirm what we inferred from the correlogram plot.

4. We use the AR(11) model that we have estimated over the time span from 2000:01 to 2009:12 to predict the value of the excess return the following month (i.e. on 2010:01).

	Lower c. bound	Forecast	Higher c. bound
2010:01	-23.568	-12.607	-1.647

5. We now gradually increase the estimation window length until we include the whole sample. The root mean squared error is: 5.415. It is the sample standard deviation of the prediction errors, i.e. it is informative of the differences between the values predicted by a model and the values actually observed.

6. In order to test for a unit root, we run two different tests. For the Augmented Dickey-Fuller test, we imposed a number of lags derived from the AIC test. In the simplest case, what the test does is running the following regression:

$$y_t = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \varepsilon_{2t} \; or \; \Delta y_t \; = \delta + (\theta_1 + \theta_2 - 1) y_{t-1} - \theta_2 \Delta y_{t-1} + \varepsilon_{2t}$$

and then testing the coefficient on y_{t-1} against the alternative that it is less than zero. Instead, what the KPSS test does is running the following regression:

$$y_t = a + \varepsilon_t$$

then defining:

$$S_t = \sum_{s=1}^{t} \hat{\varepsilon}_s \; for \; t = 1 \ldots T \; and \; \hat{\sigma}^2 = Var(\hat{\varepsilon}_s)$$

and then computing the test statistic as:

$$KPSS = \frac{1}{T} \sum_{t=1}^{T} S_t^2 / \hat{\sigma}^2$$

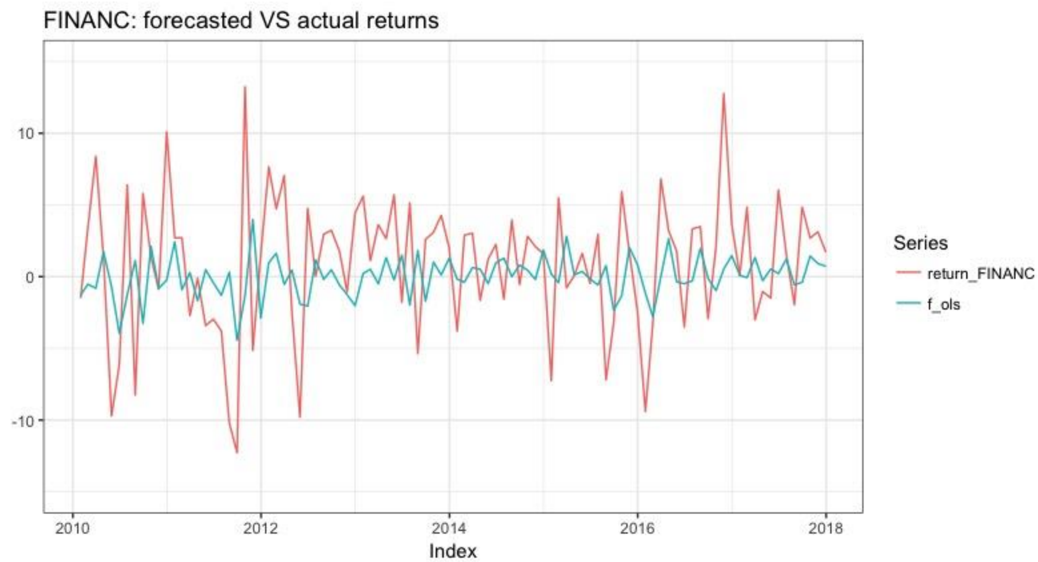
	ADF test	KPSS test
Test statistic	-2.444	2.114
Critical value	-3.46	0.739
Conclusion	Fail to reject the null hypothesis of non-stationarity	Reject the null hypothesis of stationarity

Since the DF test statistic is not less than the critical value, we cannot reject the null hypothesis of non-stationarity. Conversely, for the KPSS test we can reject the null hypothesis of stationarity (and therefore derive the same conclusion for both tests) since the test statistic is greater than the critical value.

7. We estimate the given regression model over the time span from 2000:01 to 2009:12 to predict the value of the excess return the following month (i.e. on 2010:01).

	Lower c. bound	Forecast	Higher c. bound
2010:01	-15.870	-1.304	13.262

8. We now gradually increase the estimation window length until we include the whole sample. The root mean squared error is: 4.947. Moreover, we produce the following plot for the predicted (green line) and realized (red line) values.



9. We implement a two-sided Diebold and Mariano test for the performance of the forecast, by using a quadratic loss function.

Test statistic (p-value)	2.480 (0.014)
Conclusion	Reject the null hypothesis of equal forecast errors

Since the test statistic is greater than the critical value we reject the null hypothesis of equal forecast errors (i.e. of same accuracy) at the 5% confidence level.

10. Finally, we combine the two forecast models with a simple average and we derive a root mean squared error of 5.017. Since this value is higher than the RMSE of the second regression model (4.947) but lower than the RMSE of the first regression model (5.415) we conclude that this new model has medium accuracy in forecasting.