Problem Set #5

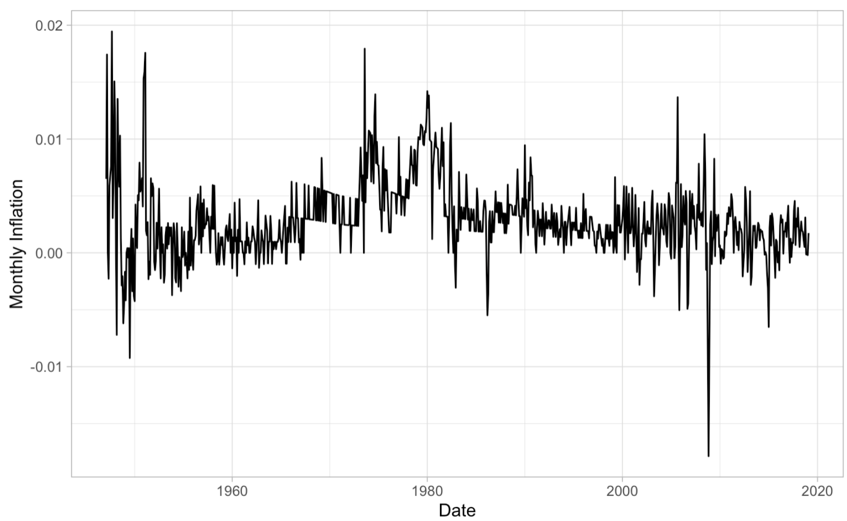
Elina Harutyunyan
15.457 - Advanced Analytics of Finance
elinahar@mit.edu

April 23, 2020

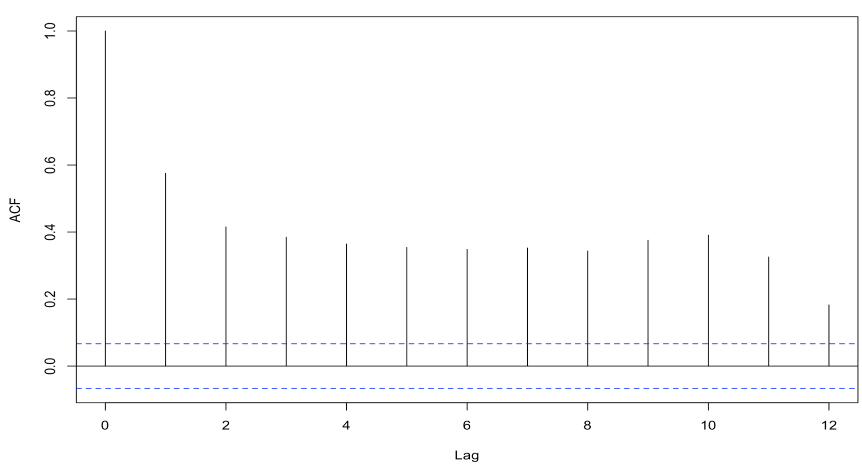
Question 1

Modeling Inflation

a) The graph for monthly inflation is plotted below:



The autocorrelation function for the inflation is plotted below:

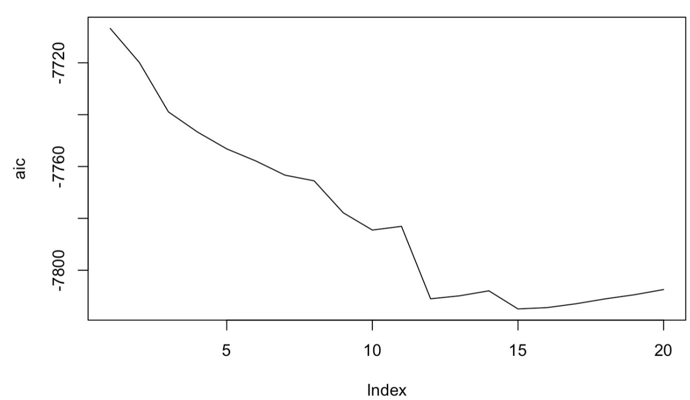


We can see from the ACF graph that it decreases gradually and all 12 lags are significant, hence we would expect the series to behave like an AR(12) model. In addition, the monthly inflation graph shows that the series is relatively distributed around the same mean however has periods of high and low volatility.

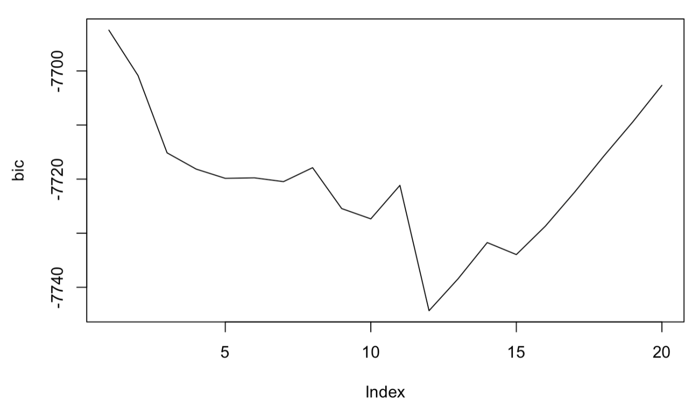
b) To determine the optimal number of lags to use in the AR model, we perform two types of tests. The first one is the auto.arima function from R as well as a manual function that minimizes the negative log-likelihood with AIC or BIC penalty. We do this allowing maximum of 20 lags and we obtain the following results.

	Auto.Arima Function	**Manual Function**
AIC	p = 15	p = 15
BIC	p = 5	p = 12

We also plot the AIC and BIC values for all AR models up to 20 lags.



(a)



(b)

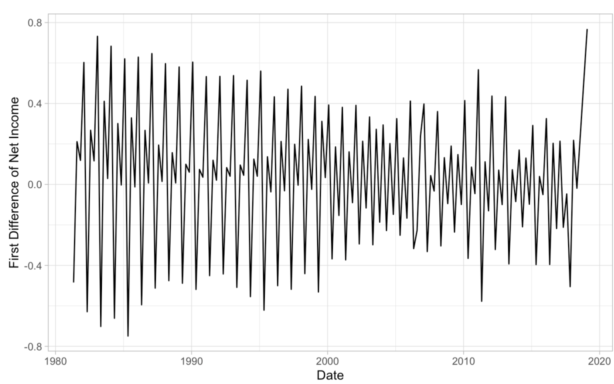
Figure 1: AIC and BIC Values

We can see there is a discrepancy between the auto.arima function and the manual minimization problem that we posed in case of using BIC. However, the graphs confirm AR(15) using AIC and AR(12) using BIC.

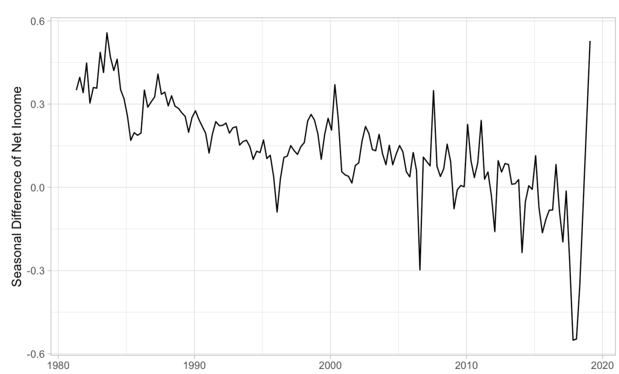
Question 2

Forecasting Corporate Earnings

a) We denote x_t as log quarterly earnings, hence the first difference $x_t - x_{t-1}$ represents the log growth rate of earnings on a quarterly basis. Using this instead of the log of quarterly earnings is done, since in general first difference of non-stationary series would make them stationary. We can see from the graph of the first difference that it is close to being stationary although the volatility seems to change with time. Given that quarterly earnings exhibit seasonal behavior, the seasonal difference $x_t - x_{t-4}$ is the log growth rate of quarterly earnings on an annual basis. However, as we can see from its graph, it does not seem to have stationary properties. We would be more interested to compare growth rates year over year instead of quarter over quarter which is done using $(x_t - x_{t-1}) - (x_t - x_{t-4})$.



(a)



(b)

Figure 2: First and Seasonal Difference of Quarterly NI

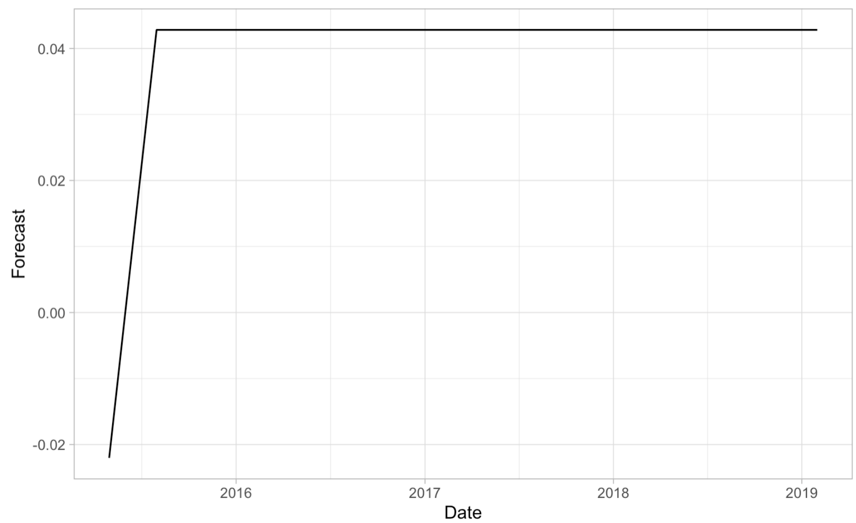
b) We build an ARIMA(0,1,1) model for the log quarterly earnings (x_t) for the period from 1980Q1 to 2014Q4:

$$x_{t+1} - x_t = \alpha_0 + \epsilon_{t+1} - \theta\epsilon_t \tag{1}$$

We obtain the following estimates for the model, which we can see that both are significant at 1% level:

	$\boldsymbol{\alpha_0}$	$\boldsymbol{\theta}$
Estimate	0.043	-0.707
Standard Error	0.007	0.041

c) We generate one-quarter ahead forecast given by the ARIMA(0,1,1) model of the first difference of Wal-Mart's earnings.



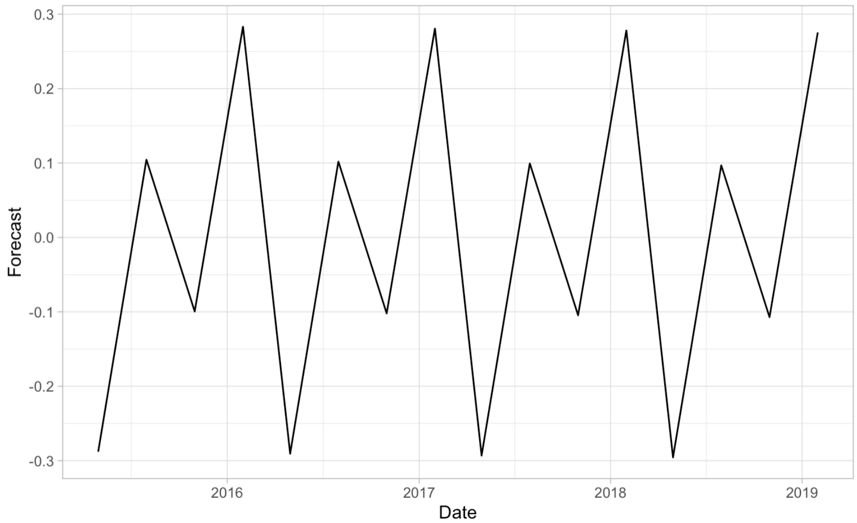
d) We build an "airline" model for the Wal-Mart's earnings with the following equation:

$$(x_t - x_{t-1}) - (x_{t-4} - x_{t-5}) = (\epsilon_t - \theta_1\epsilon_{t-1}) - \theta_4(\epsilon_{t-4} - \theta_1\epsilon_{t-5}) \tag{2}$$

We obtain the following results with both estimates significant at 1% level:

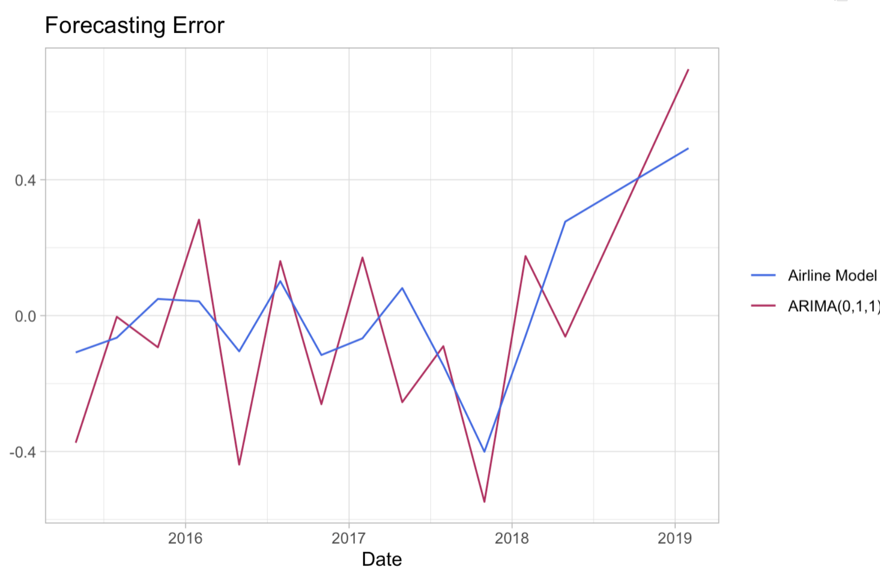
	$\boldsymbol{\theta_1}$	$\boldsymbol{\theta_4}$
Estimate	-1.000	-0.412
Standard Error	0.019	0.116

The one-quarter ahead forecasts of this model are plotted below:



θ_1, θ_4 can be used to back out the coefficients form the AR models. In a sense that $\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$ and $\rho_4 = \frac{-\theta_4}{1+\theta_4^2}$ are the coefficients of regressing x_t on x_{t-1} and x_{t-4} respectively.

e) We obtain MSE of 0.1056 in case of ARIMA(0,1,1) model and 0.040785 for the "airline" model. This and the graph below show us that forecasting error for both models varies during the test period, with ARIMA performing better in some quarters and "airline" in others. In addition, we can see that as the time period approaches the end of testing sample, forecasting error of both models increases meaning that forecasts too ahead in the future are less accurate than in the immediate future.



Question 3

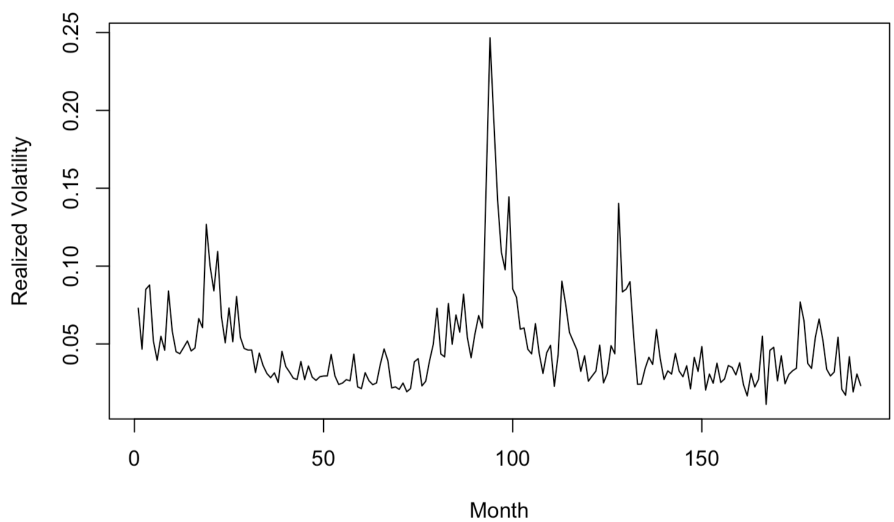
Portfolio Risk Model and CPPI

a) Our portfolio has the following characteristics in annualized terms:

Sharpe Ratio	0.398
CAPM Alpha	1.389%
Information Ratio	0.81

The alpha produced from the CAPM model is significant at 1% level. Overall, both the Sharpe Ratio and Information Ratio are low, which translates in a low annualized alpha in excess of market returns.

b) We then compute the monthly realized volatility of the portfolio by multiplying standard deviation of daily returns with the square root of days in that given month. The graph below shows the realized monthly volatility of our portfolio for the sample period.



As can be observed from the graph, the monthly realized volatility of the portfolio ranges from 5% to 25% the latter corresponding to the global financial crisis period. We can see that overall excluding the months of crisis, the volatility is somewhat stable and ranges between 5% and 10%.

c) Log portfolio return series $r_1, r_2...r_n$ follows AR(1)-GARCH(1,1) model with corresponding equations:

$$r_t = \mu + \phi_1 r_{t-1} + x_t \tag{3}$$

$$x_t = \sigma_t \epsilon_t \quad \epsilon_t\ N(0,1) \tag{4}$$

$$\sigma_t^2 = a_0 + a_1 x_{t-1}^2 + b_1 \sigma_{t-1}^2 \tag{5}$$

Given that the distribution function for x_t given σ_t is:

$$p(x_t|\sigma_t;\theta)\ N(0,\sigma_t^2) \tag{6}$$

The expectation and variance of r_t would equal to:

$$E_t(r_t|\sigma_t) = \mu + \phi_1 r_{t-1} \tag{7}$$
$$Var_t(r_t|\sigma_t) = \sigma_t^2 \tag{8}$$

We obtain the log-likelihood function for r_t:

$$L(\theta) = \sum_{i=1}^{T} lnp(r_t|\sigma_t;\theta) \tag{9}$$
$$= \sum_{i=1}^{T} -ln\sqrt{2\pi} - \frac{(x_t-\mu-\phi_1 r_{t-1})^2}{2\sigma_t^2} - \frac{1}{2}ln(\sigma_t^2) \tag{10}$$

To obtain the values of the model, we initialize $\sigma_0 = \frac{a_0}{1-a_1-b_1}$, from there we would obtain x_0, then r_0 and so forth.

d) We estimate AR(1)-GARCH(1,1) model and obtain the following coefficients that are significant at 1% level:

	Estimate	**St. Error**
$\boldsymbol{\mu}$	0.00069	0.00013
$\boldsymbol{\phi_1}$	-0.06379	0.01687
$\boldsymbol{a_0}$	0	0
$\boldsymbol{a_1}$	0.10063	0.00958
$\boldsymbol{b_1}$	0.88104	0.01052

e) Using the parameter estimates from above, we initialize returns at the last value of training period(2001-12-31), initialize sigma as mentioned above($\sigma_0 = \frac{a_0}{1-a_1-b_1}$) and set the initial value of $x_0 = r_0 - mu - \phi_1 r_1$ as the residual value of realized returns and the predicted returns. For each day onward we calculate the expected r_t by equation (7), x_t as the residual of realized and predicted returns by (3) and σ_t by equation (5). After obtaining the time series of those variables we compute the 1-day VaR at 1% as follows:

$$VaR(\text{1-day}, 1\%) = V_t(r_t - 2.33\sigma_t) \tag{11}$$

To calculate the 10-day VaR at 1% since above we got that r_t is assumed to follow $r_t\ N(\mu + \phi_1 r_{t-1}, \sigma_t^2)$ that means that $r_{t,t+k}\ N(\mu + \phi_1 r_{t-1,t+k}, k\sigma_t)$, hence we obtain:

$$VaR(10, 1\%) = \sqrt{10} VaR(1, 1\%) \tag{12}$$

Below we graph the two VaR series for the period 2002 to 2016. As we can see, the most that the portfolio will lose with 99% probability in case of 1 day time horizon is very small as expected ranging from 0 to 15 and is much higher with 10-day horizon with the biggest losses corresponding to crisis period and reaching almost 40.