FINANCIAL ECONOMETRICS AND EMPIRICAL FINANCE (20192) – HOMEWORK 1

***GROUP 9***

Filippo Cambiaghi – 3036161

Federico Brunelli – 3043110

Federico Buizza – 3050918

Marc Gehring – 3130862

Luca Amedeo Giacardi – 3043914

***GENERAL INDEX***

**TASK 1**3

**TASK 2**4

**TASK 3**5

**TASK 4**6

**TASK 5**7

**TASK 6**9

**TASK 7**10

**TASK 8**13

**TASK 9**16

**TASK 10**18

***Task 1***

***Histogram of Energy Returns:***



The graph above suggests that the energy returns are not distributed according to a Gaussian distribution. In fact, the distribution looks skewed to the right and has fatter tails than expected. In a normal distribution, 99.73% of the observations should fall inside the mean ± 3 SD interval. In our case, this interval is (-0.2530; 0.2664), and given 360 observations, we would expect just one observation to fall outside this interval if the data were normally distributed. It is clear from the graph that many more observations fall outside this interval.

***Table of Summary Statistics:***



Strikingly, the kurtosis value in excess of 3 is a clear indication of fat tails. We call this a “leptokurtic distribution”. Moreover, from the positive skewness, we can infer that the distribution is slightly asymmetrical towards the right. This rather informal analysis is confirmed by the Jarque-Bera test, by which we can reject the null hypothesis of skewness = 0 and kurtosis – 3 = 0. In fact, the test statistic is remarkably high, and considering that its P-value basically equals zero, we can reject the null hypothesis at virtually any confidence level, thus confirming the absence of normality in the data.

***Task 2***

***Autocorrelogram for Energy Returns (29 lags):***

From the graph, we can infer that the time series is stationary as the first lag is statistically different from zero and among the higher-order lags, there are just a few that are barely statistically significant. However, this is consistent with our 95% confidence interval, for which, on 29 lags, we would expect one or two exceptions (i.e., spurious lags). From this result, we can conclude that, since the SACF does not show a geometrical decay towards zero, but an immediate reversion to zero after the first lag, we are facing a stationary time series.

Looking solely at the SACF, the absence of slowly fading autocorrelation suggests that this time series model might be a MA(q). Specifically, since autocorrelation coefficients become non-significant immediately after the first lag, we can conclude that q is equal to 1. Therefore, we find evidence for an MA(1) process.

Moreover, we can see from the low P-values of the Q-test that the series is unlikely to be a white noise process. This is evidence against an efficient market, in which we would expect no predictable pattern whatsoever between returns. Still, we see that the first lag is significantly different from zero, meaning that each return is affected (only) by shocks from its respective previous period, which is potentially explainable by some structural characteristics of the market. This implies that past shocks are quickly forgotten, therefore suggesting that the market is efficient.

***Task 3***

***CER model for Energy Returns:***

***Task 4***

***Estimation by Maximum Likelihood of MA(9) with first 5 coefficients equal to zero:***

Comparing the two models by their R-squared values, the second one has more explanatory power while the first one exhibits an R-squared value close to 0. Therefore, looking just at this variable, we should prefer the second model. Looking at the information criteria, however, we observe that all of them are greater for the second model. Hence, according to the ICs, the model with only the constant (CER model, task 3) is preferred to the MA(9) model (task 4). It is important to note that the Schwarz criterion has the greatest penalty factor for every added parameter. Accordingly, a lower value for this criterion in the second model would be expected. In this case, the positive effect of adding a new variable to the decrease of the sum of squares is more than offset by the negative effect of overfitting (lack of parsimony). Finally, if we consider the F-value of the MA(9) model, we notice a high P-value, which allows us to reject the statistical significance of the whole model.

***Task 5***

We investigate which ARIMA model, restricted to a maximum p of 8, q of 8, and d of 1, delivers the lowest AIC. Running the test, we notice that all reported models do not include differencing. This is only intuitive since we saw early that the series is stationary. We further observe that the ARMA (6,6) model exhibits the lowest AIC (*figure to the left*). Plotting the top 20 AIC models, we see that an ARMA (5,4) model shows a similar AIC value (*second figure below*) and better BIC and HQIC values, but all its MA coefficients are highly *statically insignificant.* The subsequent models fall off steeply. Lower-order models, however, perform better in BIC and HQIC, but these criteria are naturally favorable towards such models (*figure to the left*). Basing our model selection solely on the AIC, we would select the ARMA (6,6) model (*first figure below*). We can see that all ARMA components, but the constant, are statistically significant at the 3% level. Concerning the different AR and MA coefficients, we see a recursive pattern, where MA and AR coefficients almost equal their negative self. This anomaly leads us to believe that there might be better descriptive models.

***ARMA(6,6) Model:***

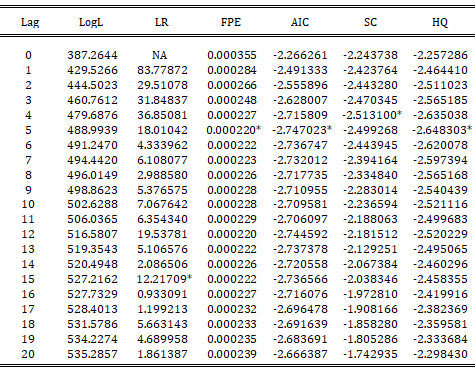


***AIC for the best 20 models:***



***Task 6***

***Model Selection Analysis (for max 20 lags):***



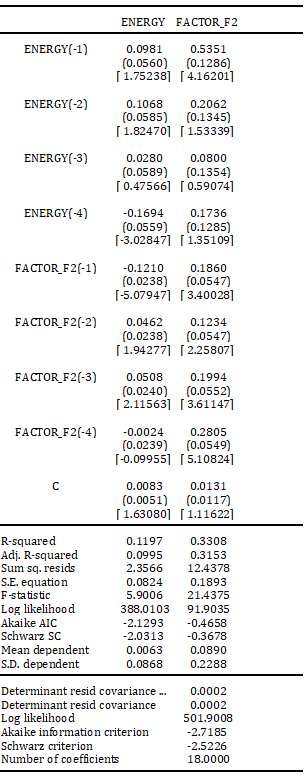
Looking at the different criteria, we notice that they all reach their minimum at different numbers of lags, except for the AIC and the HQIC, which both suggest using a model with 5 lags. On the other hand, according to the SC, we would use a model with 4 lags and according to the likelihood ratio one with 15 lags.

Usually, it happens that the AIC recommends models with more lags, while in this case, we can notice that the criterion suggests using a model with rather few lags. In fact, its recommendation is almost identical to the ones provided by the other criteria.

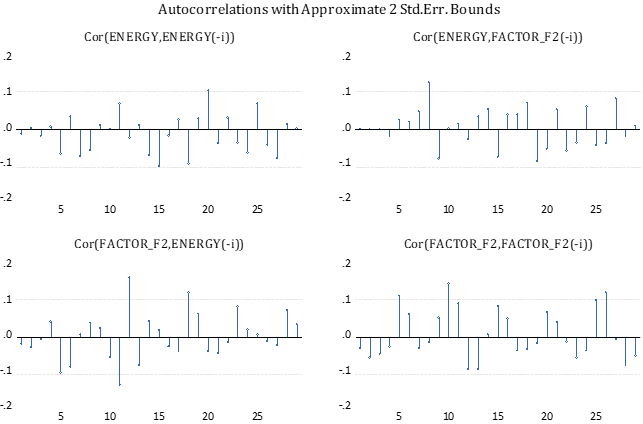
The modified likelihood ratio test is carried out starting from the maximum number of lags (20). It tests the hypothesis of whether the coefficients up until lag *l* are jointly equal to zero using a chi-squared distribution. The test keeps running until it obtains a rejection. In this case, the likelihood ratio leads to a model, which has too many lags for our purposes.

***TASK 7***

***VAR(4) model with Energy and Factor\_F2:***

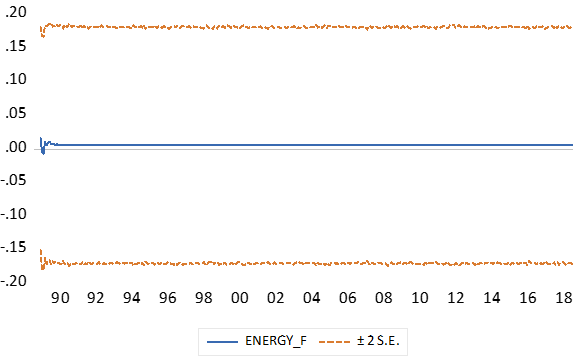


The model suggested by the SC is based on 4 lags. Looking at P-values (values in brackets), we notice that on a total of 18 coefficients only 6 are statistically significant. In order to increase parsimony, we could impose restrictions, but in this case, it would become a restricted VAR. It is also necessary to underline the fact that an assumption of normality must be made in order to arrive at our conclusions. Finally, observing the R-squared values, we can notice that they are both quite low (even if FACTOR\_F2’s one is higher).

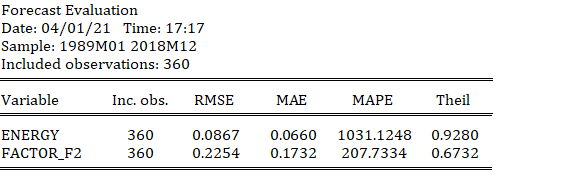


These graphs above represent the SACF of the VAR residuals. If the model is correct all the sample ACFs should be non-significant (residuals should be White Noise). In the bottom-right graph, we see that the number of coefficients exceeding the 95% confidence interval is too high (3 out of 29). Hence, this model is doing a poor job at modeling the Factor 2 behavior. On the contrary, the development of the energy returns is certainly better explained, as only one coefficient is exceeding the confidence interval. This means that we do not reject the null hypothesis that the residuals are white noise processes, at least for the energy returns.

***Plot of forecasts of energy returns:***

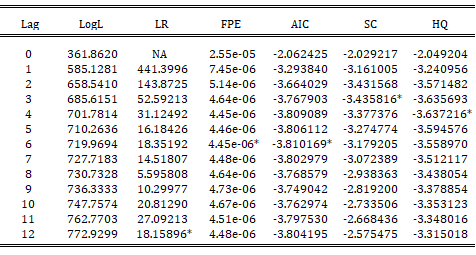


***Table of forecasting accuracy measures:***

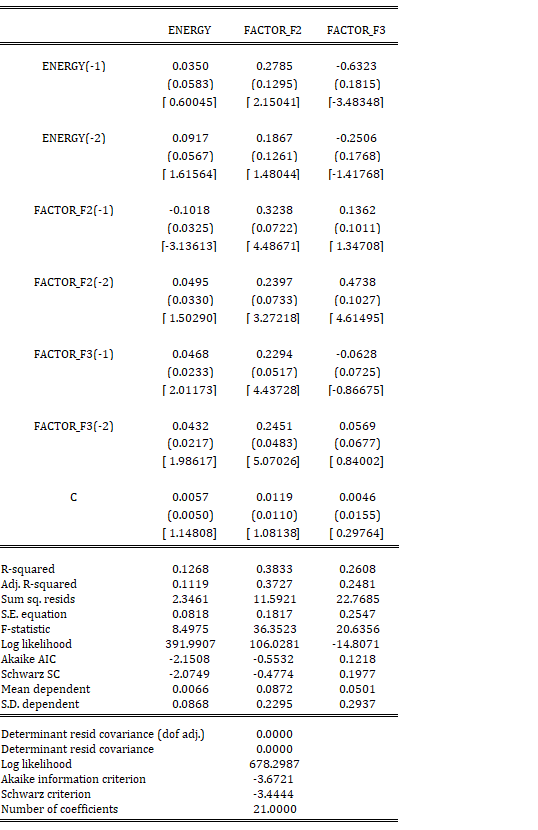


***TASK 8***

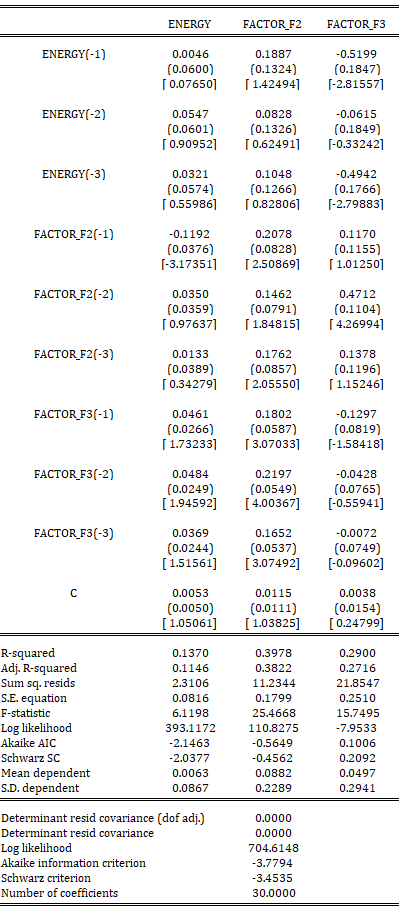
***Optimal lag selection table:***



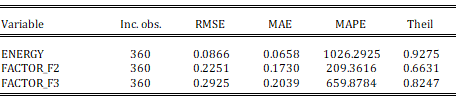
***VAR(2) Estimate:***



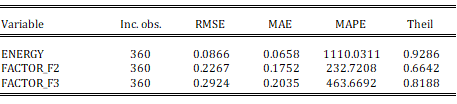
***VAR(3) model, suggested by the Bayesian criterion:***



***VAR(2) forecast accuracy measures:***

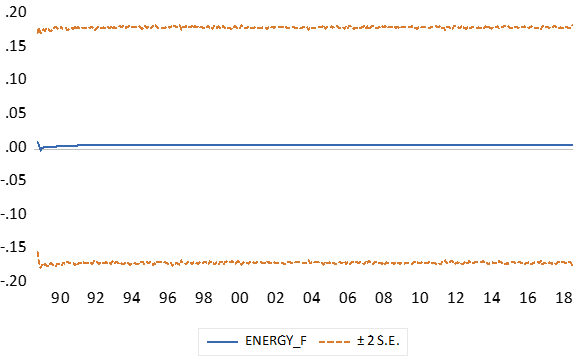


***VAR(3) BIC forecast accuracy measures:***



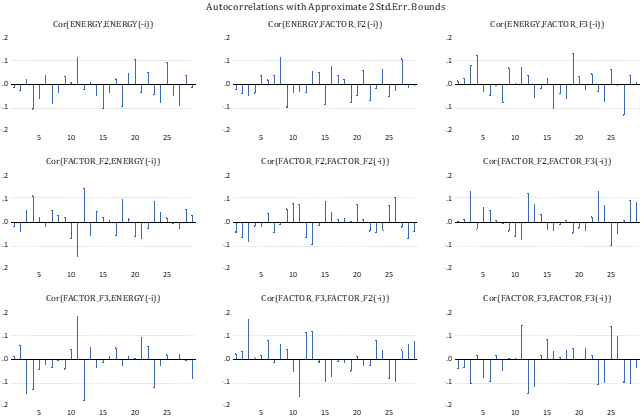
Looking at the two tables above, the VAR(2) model does a better job than the VAR(3) model, which is selected by the BIC, at forecasting the energy returns. It is evident that the VAR(2) and the VAR(3) models display very similar forecasting accuracy, though the VAR(2) slightly outperforms the VAR(2) according to some of the errors. In particular, even though the RMSE and the MAE show the same value for both models, the MAPE and Theil’s accuracy coefficient are minimized in the VAR(2) model. For this reason, this model should be preferred. Moreover, considering that models with more lags tend to increase the goodness of fit but reduce the parsimony of the model, we find another reason to choose the VAR(2) rather than the VAR(3) model.

***VAR(2) energy forecasts plot:***



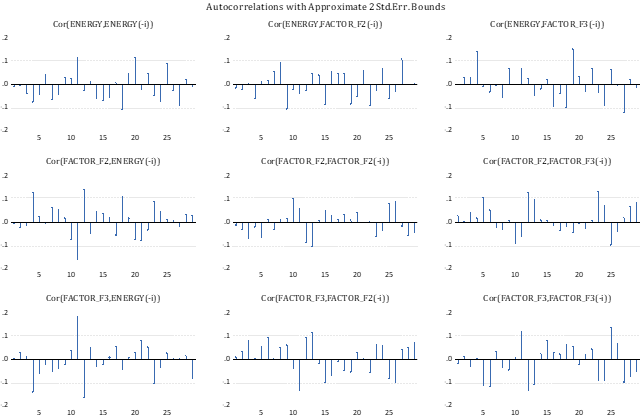
***TASK 9***

***Residuals correlograms of VAR(2) model:***



Looking at the correlograms of the VAR(2) model we can say that Factor 3’s behavior is not well explained by the model since it exhibits a relatively high number of coefficients exceeding the 95% confidence interval (6 out of 29). The same thing applies to the energy return since we can see that 3 out of 29 coefficients are exceeding the confidence interval. Instead, the model gives a quite good explanation for Factor 2 as only one coefficient exceeds the confidence interval. Overall, there should be concerns about the specifications of this model since the residuals seem not to be white noise processes.

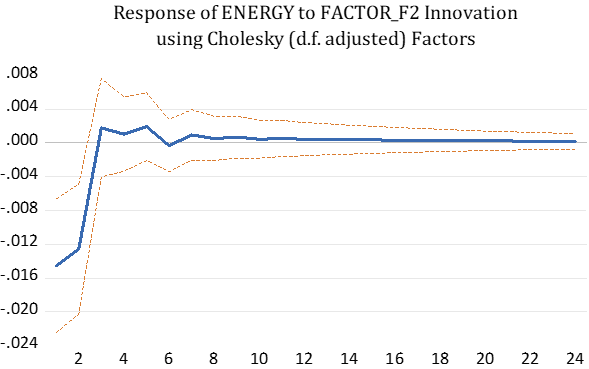
***Residuals correlograms of VAR(3) BIC model:***



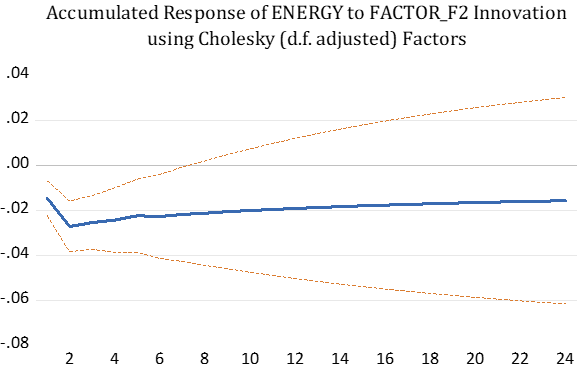
Looking at the correlograms of the VAR(3) model we can say that the behavior of Factor 3 is not well explained by the model since it also has a relatively high number of coefficients exceeding the 95% confidence interval (6 out of 29). The Factor 3–Factor 2 correlogram is not correctly explained, too. The same thing applies to the Energy-Factor 2 correlogram, as we can see that 2 out of 29 coefficients are exceeding the confidence interval. Instead, the model gives a quite good explanation for Factor 2, as only one coefficient is exceeding the confidence interval. Overall, there are concerns about the specifications of this model. In order to better specify this model, we should change the number of lags.

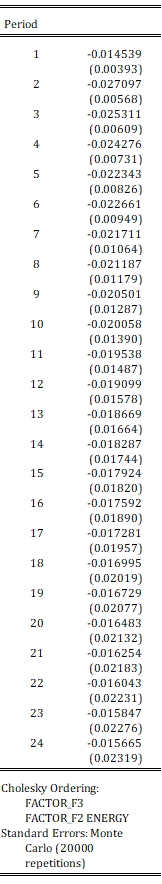
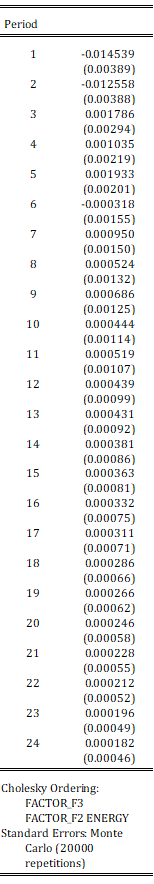
***TASK 10***

***Impulse response of model VAR(2). Marginal plot of the response of energy to a shock to factor 2:***



***Impulse response of model VAR(2). Cumulative plot of the response of energy to a shock to factor 2:***



***Marginal IRF table: Accumulated IRF table:***