WorldQuant University
MScFE 610 – Econometrics
Group 5P Submission 2

Constructed Response Questions

Step 7: Modelling Cointegration

Explain the connection between linear regression, VAR, and VEC.

A Vector Autoregressive (VAR) model is a statistical model used in case of multivariate time series data which are stationary. It is used to model 2 or more variables which are not only linearly dependent on their own lagged values but also on the lagged values of other variables to which they are related.

For instance, a VAR model of order 2 for two variables would include 2 equations, one for each variable wherein the current value of the variable is linearly related to 2 lagged values of itself and 2 lagged values of the other variable. So, we get the following system of equations:

$$Y_{1,t} = \alpha_1 + \beta_{11,1} Y_{1,t-1} + \beta_{12,1} Y_{2,t-1} + \beta_{11,2} Y_{1,t-2} + \beta_{12,2} Y_{2,t-2} + \epsilon_{1,t}$$
$$Y_{2,t} = \alpha_2 + \beta_{21,1} Y_{1,t-1} + \beta_{22,1} Y_{2,t-1} + \beta_{21,2} Y_{1,t-2} + \beta_{22,2} Y_{2,t-2} + \epsilon_{2,t}$$

By extension, a VAR(p) model will have the following vector equation:

$$x_t = A_0 + A_1 x_{t-1} + \cdots + A_p x_{t-p} + \varepsilon_t$$

where A_0 is an *(n x 1)* vector of regression constants and A_i is an *(n x n)* matrix of coefficients linking each of the *n* variables' current values to each of the *n* variables' past values at lag *i*.

At heart, the VAR model is essentially a system of linear regression models because they satisfy the assumptions of a classic linear regression model:

- Study the dependence of a several dependent variables on explanatory variables, which in this case are lagged values of those variables (explanatory variables)
- The parameters linking the dependent variables to the explanatory variables are linear and estimated using OLS

Because the VAR model is only applicable on stationary time series, there is a possibility of loss of information about important long-term relationships between levels. If there is a long-term relationship between levels (cointegration) of nonstationary series, we use a Vector Error Correction Model (VECM). Once equilibrium conditions are specified, the model describes the adjustment in each time period towards the long-run equilibrium. Cointegration of variables causes deviations from long-run equilibrium in the short-run to feed back into the dependent variables until they reach the long-run equilibrium.

Explain limitations of Engle-Granger methodology and Johansen test. Are methods sensitive to certain factors/parameters?

Limitations of Engle-Granger Methodology:

- Since Engle-Granger is a two-step procedure wherein the first step involves generating the residuals and the second step involves regressing first-differenced residuals on lagged residuals, any error occurred in the first step will be carried into the second step

- In conducting the test, one may find a cointegration relationship from residuals of the first regression but not from other regressions. So, the test for unit root in the error term sequences may not be equivalent across regressions

- Because of a finite sample size, there may be a lack of power in unit root and cointegration tests

- There may be more than one cointegrating relationship

- It is not possible to perform hypothesis tests on the actual cointegrating relationship estimated in step 1

Limitations of Johansen Test:

- Researchers Gonzalo & Lee (1997) report that Engle-Granger was more robust than Johansen's likelihood ratio test for several situations

- Much more complex than the Engle-Granger methodology in terms of formulating hypotheses and interpreting results, since all variables are treated symmetrically and there is no distinction between endogenous and exogenous variables

Explain speed of adjustment parameters in VEC model.

Because the VECM specifies cointegration relations built into the model, these relations restrict the long-run behaviour of the endogenous variables. Over the short-term, these variables dynamically adjust and move towards the long-run equilibrium. This movement is modelled through the cointegration term, which is also known as the error correction term. The size of this error correction term indicates the speed of adjustment towards the long-run equilibrium state in the VEC model.

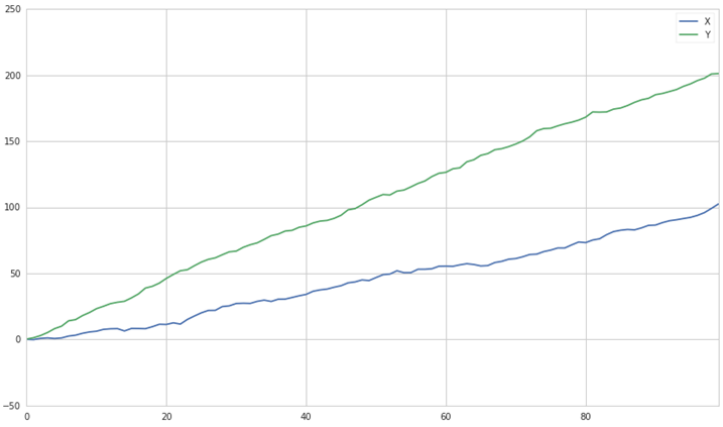
Step 8: Does correlation imply cointegration?

Correlation describes the strength of relationship between two or more variables. It is a measure of how well two or more variables move in tandem together over time. Popular correlation measures are Pearson's coefficient and Spearman's ranks-order coeffcient. Both measures range from -1 (perfect negative correlation), to 0 (no correlation) and to 1 (perfect positive correlation). Positive correlation means that the variables move in tandem in the same direction while negative correlation means that they move in tandem but in opposite directions.

On the other hand, cointegration is not a measure of how well two variables move together, but rather whether the difference between their means remains constant or stationary. So in other words, cointegartion tests can let us know whether the distance between two or more time-series(s) remains

the same over time. This helps in identifying variables(times series(s)) that would not drift too far away from each other in the longer term and would revert to a mean distance between them, if they are cointegtrated. So usually cointegration applies to long-term properties of time series(s).

There are cases where two time-series are highly correlated and also cointegrated and vice versa but that is not always the case. Also, as opposed to correlation, while performing cointegration tests we tend to use asset prices rather than their returns as we are more interested in identifying the the trend between the variables means over time, than in their individual price movements.



In the figure alongside, we can observe that stock prices of X and Y are highly positively correlated. So they are moving in tandem in the same direction. However, spread between the price of two stocks has increased from 0 to 100. Thus, they are not cointegrated. This is a perfect example of “Cointegration without Correlation”.

Alternative Investments Performance Report (Bitcoin)

Figure 1:

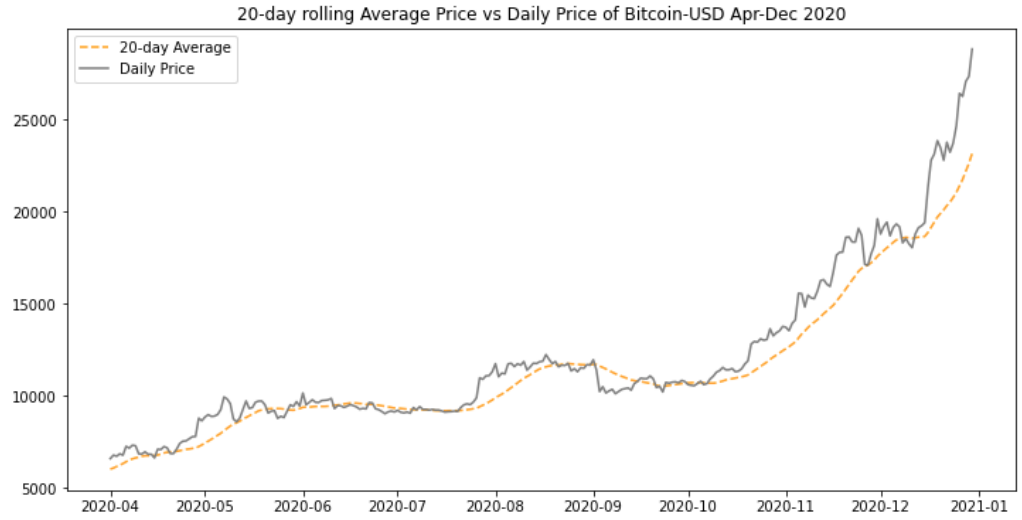
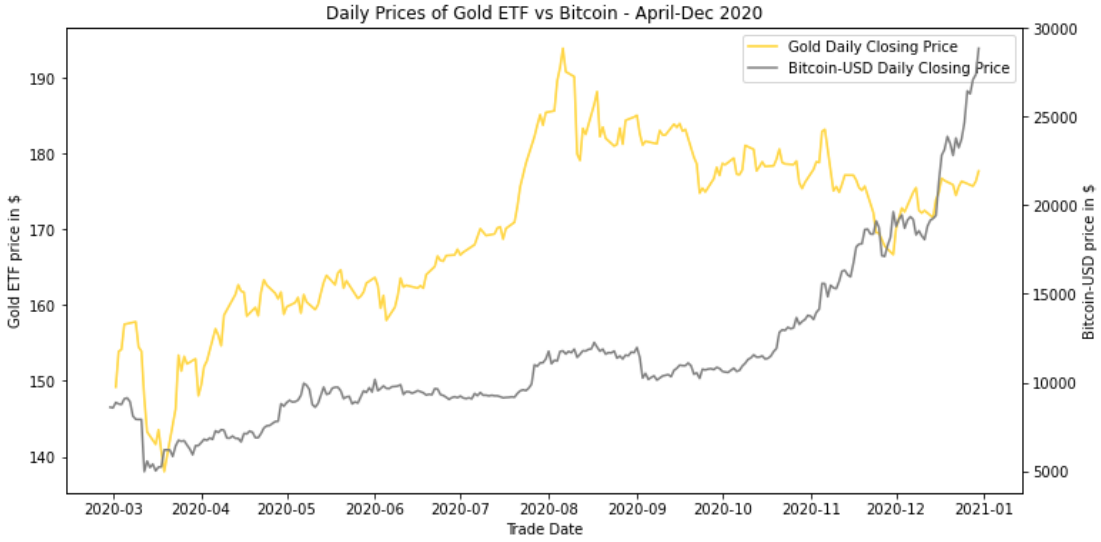


Figure 2:



We notice seasonality in the bitcoin returns with a cyclicality of 5 days. The significant value of beta (0.9699) and extremely low value of alpha in the GARCH(1,1) model indicates high persistence and a stable market. Probably, this is because of the confidence of the investors in their bitcoin investment.

Using the cointegration vector of Bitcoin and Gold ETF in Q2 we notice that there is significant volatility in Bitcoin ETF which balance the low volatility of Gold ETF and produces a stationary time series. Thus, bitcoin could be used for risk management purposes thereby ensuring the right amount of volatility.

The cointegration components could help us to conduct Pair trading for Bitcoin and Gold ETF where portfolio manager could leverage arbitrage opportunities.

Alternative Investments Performance Report (As compared to Equity Investments)

Figure 1:

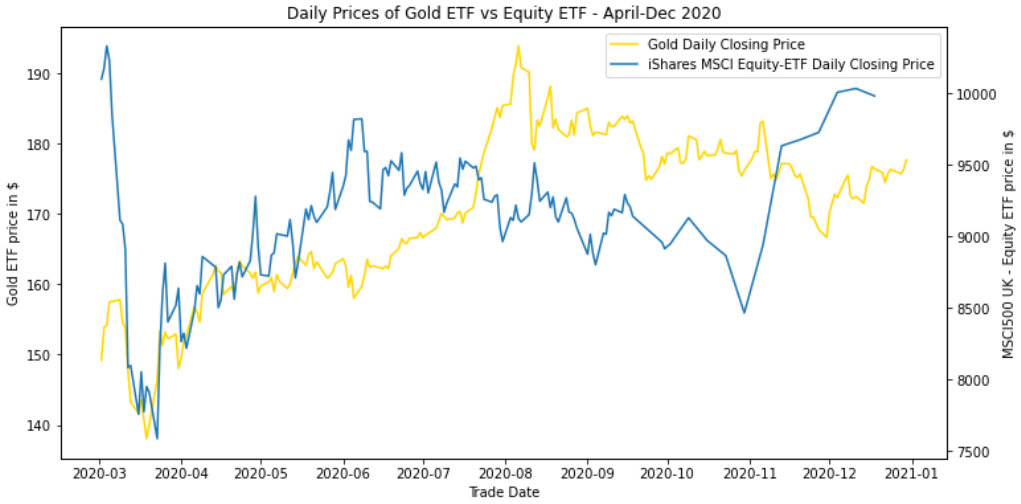
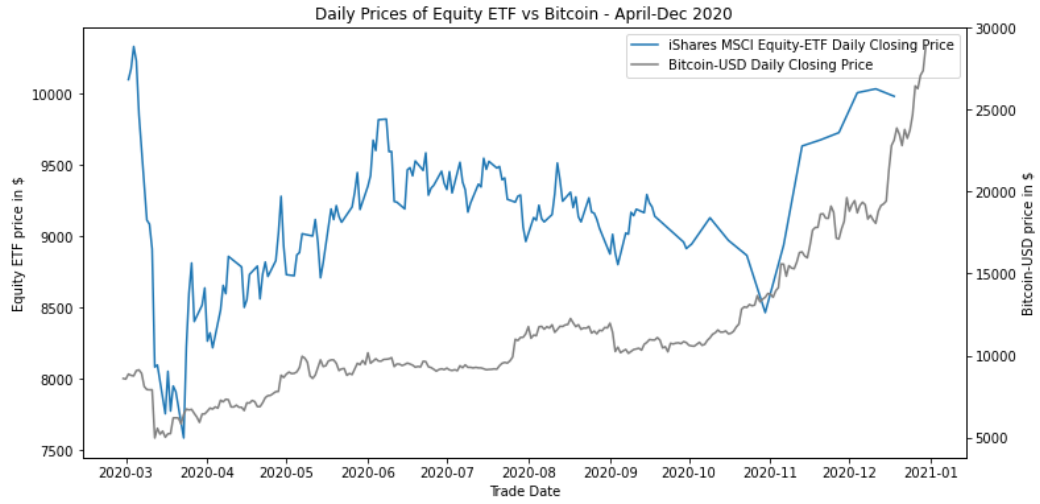


Figure 2:



Correlation Matrix:

	Pearson's Correlation with Nifty ETF		
Asset	**Q2**	**Q3**	**Q4**
GLD	0.585	-0.551	-0.623
BTC	0.813	-0.479	0.81

Returns Matrix:

	Log Returns		
Asset	**Q2**	**Q3**	**Q4**
GLD	11.3%	6.1%	-0.6%
BTC	32.4%	15.6%	99.9%
Nifty	12.6%	-4.5%	11.1%

We can observe from Figure 1 and correlation matrix that GLD ETF can be an attractive alternative investment option to equity investments. They are negatively correlated and also GOLD ETF returns

exceeds equity ETF's return significantly in Q2, when Equity ETF gave a negative return. Gold is also viewed as safe haven by investors in times of uncertainity. So, considering the damage to world-wide economies caused by COVID-19 pandemic, it is advisable to be 30 – 40 percent of portfolio invested in GLD ETFS until mass vaccination programme is completed.

We can observe from Figure 1 and correlation matrix, BTC ETF can also be an attractive alternative investment option to equity investments. Correlation with Equity ETFs are not stable as BTC as an asset class is quite new. So, there is no historical evidence to support these calculated correlations. Bitcoin returns as very much lucrative than Equity investments. It beat the returns of Equity investments in every quarter of year 2020. However, we need to consider risk-adjusted return for Bitcoin as it is very volatile.

Thus, following **"barbell strategy"**, which is an investment concept that suggests that the best way to strike a balance between reward and risk is to invest in the two extremes of high risk and no risk assets while avoiding middle-of-the-road choices, we would advise investors to hold 5-10 percent of their portfolio in Bitcoin and rest in GLD ETF and equity investments.

Report Summarising Distribution of Work in Group 5-H

- All 3 members were active and shared the workload. In terms of the distribution, all members, both independently and together, wrote the code and reports for this project

- Since each part of the coding assignment was to some extent dependent on other parts (notably with regard to creation of data frames, calculating summary measures etc.), there were many overlaps in the code which each of us individually wrote. Further, the subparts of each part were also shared among the team members

- For this reason, it is difficult to provide a precise description of who did each step. Concerning the reports and constructed response essays, all of them contain ideas of all active team members