TIME SERIRES ANALYSIS:

CRYPTOCURRENCY INFLUENCE ON TRADITIONAL STOCK MARKET INDICES: A COMPARATIVE MULTIPLE REGRESSION ANALYSIS OF VOLATILITY AND CORRELATION

WITH THE INTEGRATION OF INSTITUTIONAL INVOLVEMENT

BASIC ECONOMETRICS

PRESENTED TO: SIR MANSOOR MUSHTAQ

Presented by:

MUSSAB BIN UMAIR (221-9809)
RAED ALV (221-7460)I

SEMESTER PROJECT

Table of Contents

Abstract

By examining the cross-correlations between the two most valuable cryptocurrencies, Bitcoin (BTC) and Ethereum (ETH), this research paper focuses on how cryptocurrencies affect traditional stock market indices. However, Litecoin (LTC) has also been a significant factor in the study. By examining the increasing potential for institutional involvement in the cryptocurrency region, we hope to analyse the impact of the cryptocurrency field's growing capitalization on the stock market as a whole.

This study looks at the long-term relationship between cryptocurrencies and the traditional stock market. As exchange rates have been rising and falling over time, and as institutional involvement in this industry has grown, we have looked at various models that explain and describe the relationships between these variables that help us with our research. Since more and more individuals and businesses are investing in cryptocurrencies, institutional involvement has been a significant factor. In the near future, we can also anticipate cryptocurrencies playing a significant role in the economy due to their rapid, exponential growth. Our analysis looks at how cryptocurrencies have the potential to be a significant turning point and how institutional investment in this space will impact the conventional stock market as well.

INTRODUCTION

The data that we have sourced from investing.com, is of the following:

1. NASDAQ 100 (NDX) – Traditional stock market representation
2. BITCOIN (BTC) – The largest cryptocurrency in the market
3. ETHEREUM (ETH) – A leading blockchain platform with smart contracts
4. LITECOIN (LTC) – One of the earliest Bitcoin alternatives
5. Grayscale Bitcoin ETF (GBTC) – One of the biggest crypto ETFs in the world.
6. EUR/USD – The exchange rate of Euros and US Dollar.

With a timeframe spanning from May 2017 to October 2024, the data offers a solid framework for examining the potential effects of cryptocurrencies on the conventional stock market over time. The data's weekly frequency guarantees that long-term trends are prioritized while short-term market noise is eliminated.

Understanding the relationship between cryptocurrencies and the conventional stock market is the aim of this investigation, which also looks at whether cryptocurrencies have a discernible and substantial influence on stock market indices such as the NASDAQ 100. This study investigates the interactions and influences that have occurred over time between the cryptocurrency environment and the conventional stock market.

VARIABLE EXPLANATION:

Functional Form of the Model:

The econometric model can be expressed as:

NASDAQ=$\beta 0+\beta 1(BTC)+\beta 2(ETH)+\beta 3(LTC)+\beta 4(DOGE)+ U$

Where:

- **NASDAQ:** Dependent variable, representing the stock market index.
- **β0:**Constant term (intercept), capturing the baseline level of the NASDAQ index when all other variables are zero.
- **BTC:** Explanatory variable for Bitcoin prices.
- **ETH:** Explanatory variable for Ethereum prices.
- **LTC**: Explanatory variable for Litecoin prices.
- **DOGE**: Explanatory variable for Dogecoin prices.

- **$\beta1,\beta2,\beta3,\beta4$**: Coefficients of the explanatory variables, representing the marginal impact of each cryptocurrency on the NASDAQ index.
- **U:** Error term, accounting for unobserved factors affecting NASDAQ.

Hypothesis

- **Null Hypothesis** (Ho: The explanatory variables (BTC, ETH, LTC, DOGE) have no significant impact on the NASDAQ index.

- **H0**: $\beta BTC=\beta ETH=\beta LTC=\beta DOGE=0$
- **Alternative Hypothesis H1**: At least one of the explanatory variables has a significant impact on the NASDAQ index.
- **H1**:At least one $\beta \neq 0$

DESCRIPTIVE STATISTICS:

In this section of the report, we discuss the descriptive statistics of the variables taken in our study and analyse all of them to see if they are acceptable or meet set standard which is, Normality.

Normality refers to whether the data follows a normal distribution (bell shaped curve). It also involves checking if the data has properties like symmetry around the mean and specific kurtosis and skewness values. Normality is important to further proceed with our analysis and moving on to liner regression and other models.

For better understanding and results being consistent, we take logs of the variables and use all the information in that order so we can achieve optimal skewness and kurtosis, which infers that the variables are normal and are a fit for this model for further analysis.

Name	Mean	Median	Std. Deviation	Skewness	Kurtosis
NDX	9.285207	9.370836	0.370466	-0.113446	1.690794
BTC	9.745074	9.853644	0.925030	-0.160306	1.856700
ETH	6.693582	7.105831	1.144797	-0.224778	1.551545
LTC	4.389632	4.288105	0.511719	0.502912	2.832180
GBTC	2.803711	2.656405	0.733260	0.133763	2.064442
EUR/USD	0.116854	0.114266	0.050394	-0.238765	2.840106

Interpretation of Mean vs. Median:

A near mean and median indicate that the data is reasonably symmetrical. Skewness is indicated by larger discrepancies.
The mean and median for all of the variables (NDX, BTC, ETH, GBTC, EUR/USD and LTC) are rather close, suggesting moderate symmetry.

For Skewness:

- Skewness = 0: A distribution that is perfectly symmetrical.
- Skewness < 0: Left-skewed, meaning the left tail is longer.
- Skewness > 0: Right-skewed, meaning the right tail is longer

Here,

- NDX: Shows a slight inclination toward lower values, with a slightly left-skewed value of -0.113.
- BTC: Shows a slight asymmetry, with a slightly left-skewed value of -0.160.

- ETH: Displays a longer left tail and is moderately left-skewed (-0.224).
- LTC: More extreme high values are present because it is right-skewed (0.503).
- GBTC: Near symmetry, as indicated by being quite close to 0 (0.133).
- EUR/USD: Displays a longer left tail and is moderately left-skewed (-0.238765).

For Kurtosis:

- Kurtosis = 3: Normal distribution (mesokurtic).
- Kurtosis > 3: Leptokurtic (sharper peak, heavier tails).
- Kurtosis < 3: Platykurtic (flatter peak, lighter tails)

Here,

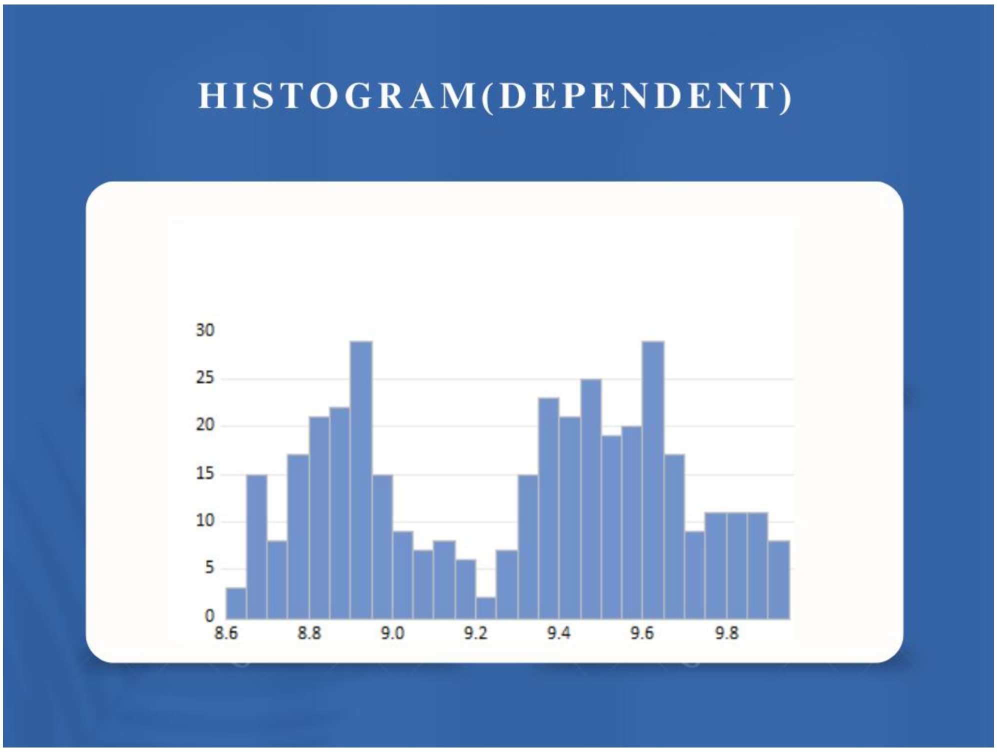
- NDX: 1.69: Platykurtic, which denotes a flatter peak and lighter tails.
- Bitcoin: Platykurtic, 1.86, with comparable traits.
- ETH: 1.55: Platykurtic, meaning tails are lighter.
- LTC: 2.83: This value is close to normal but leptokurtic, indicating slightly heavier tails.
- GBTC: 2.840106: This value is close to normal but leptokurtic, indicating slightly heavier tails.

So far, these are the descriptive statistics for every variable in this model.

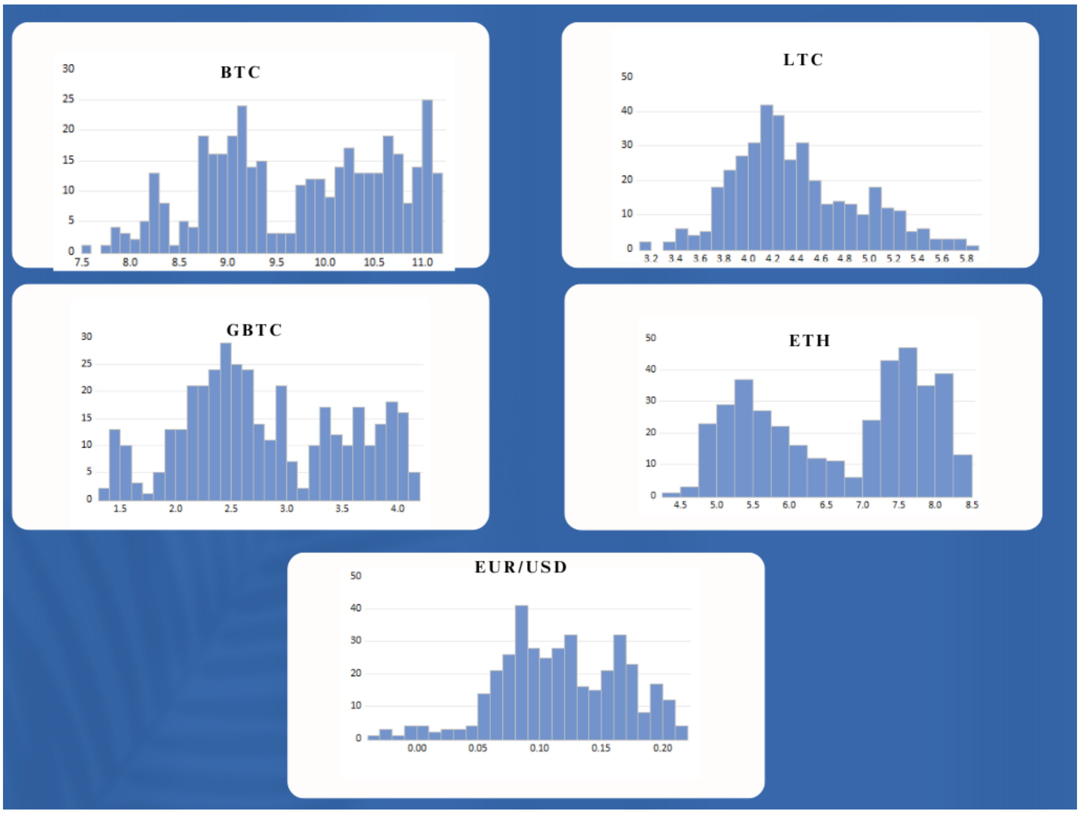
Moving on we have the histogram tables for all the variables in the next figure.

HISTOGRAM TABLES:

Dependent Variable

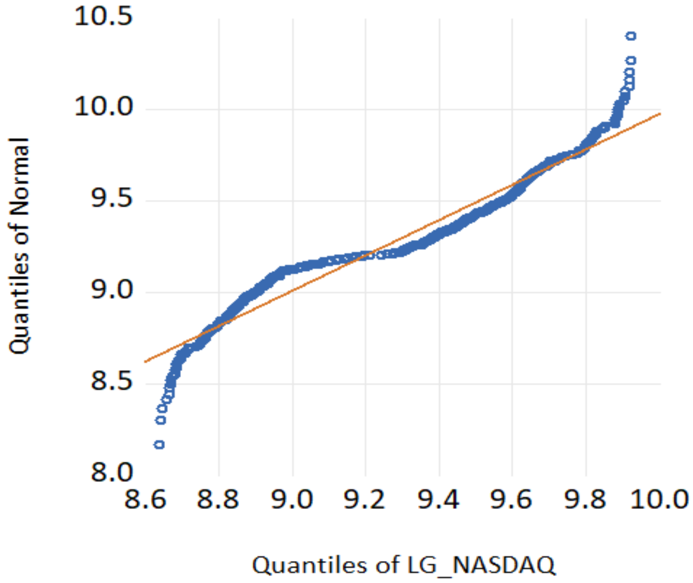


Independent Variables:

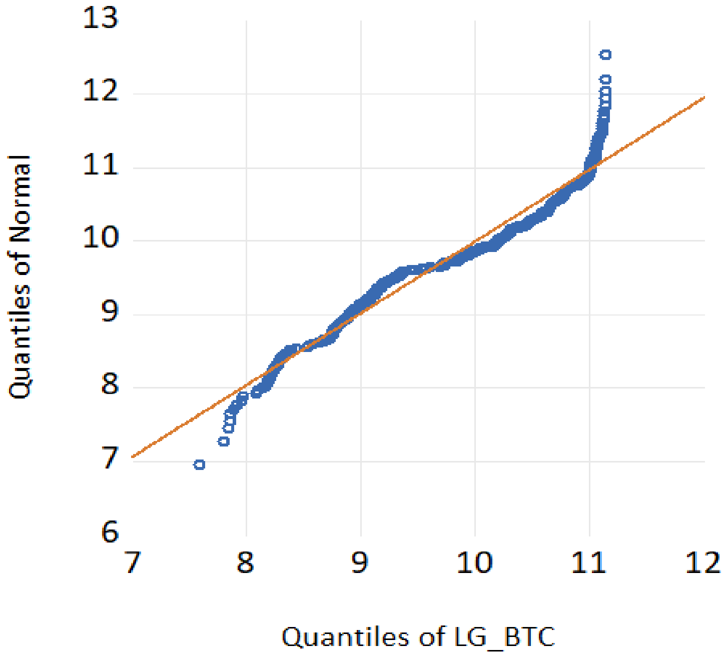


Q-Q PLOTS:

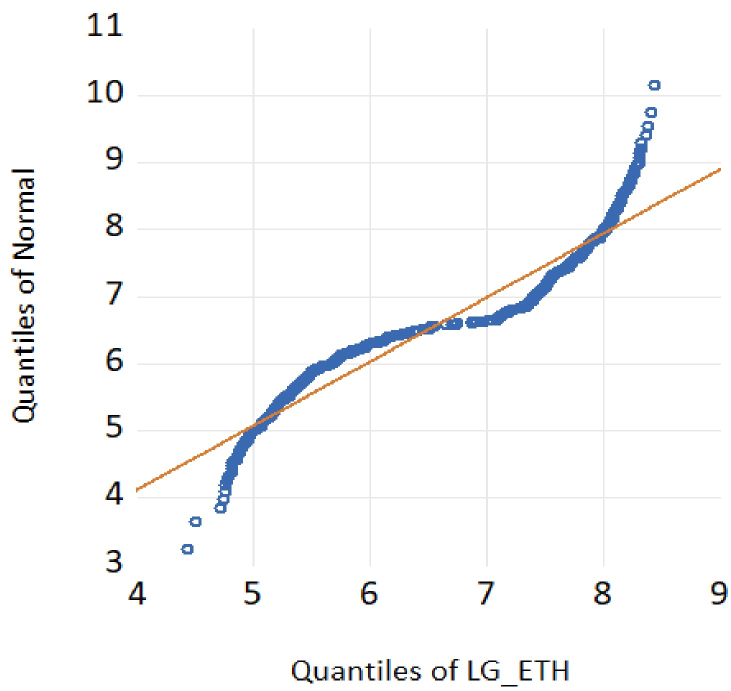
NASDAQ



BITCOIN

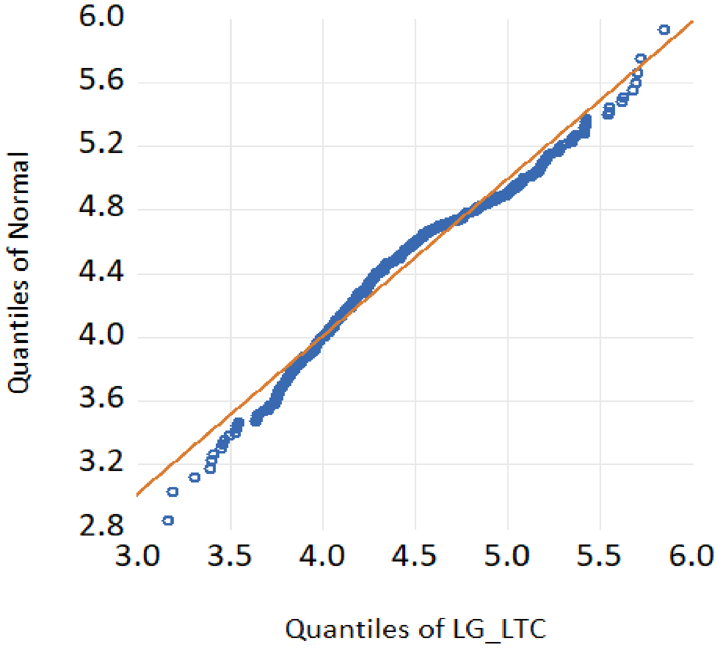


ETHEREUM

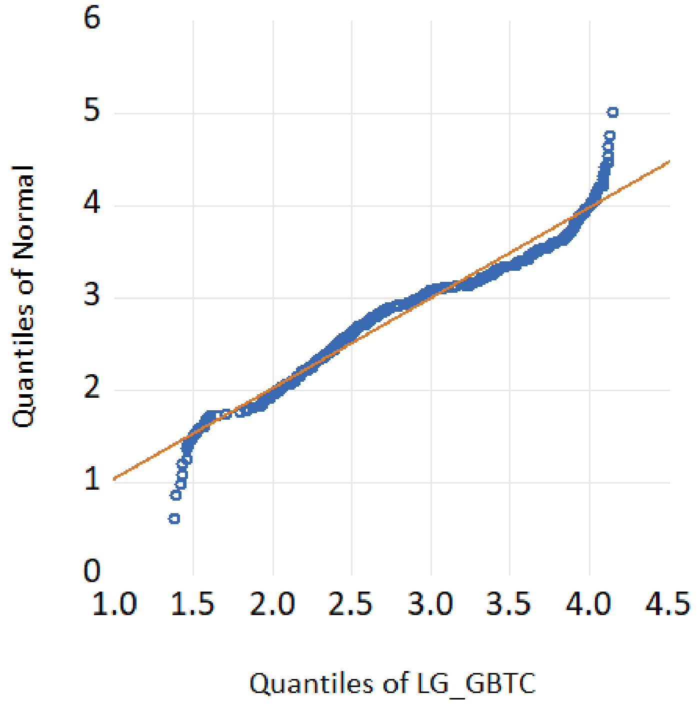


11
10
9
8
7
6
5
4
3
4
5
6
7
8
9
Quantiles of Normal
Quantiles of LG_ETH

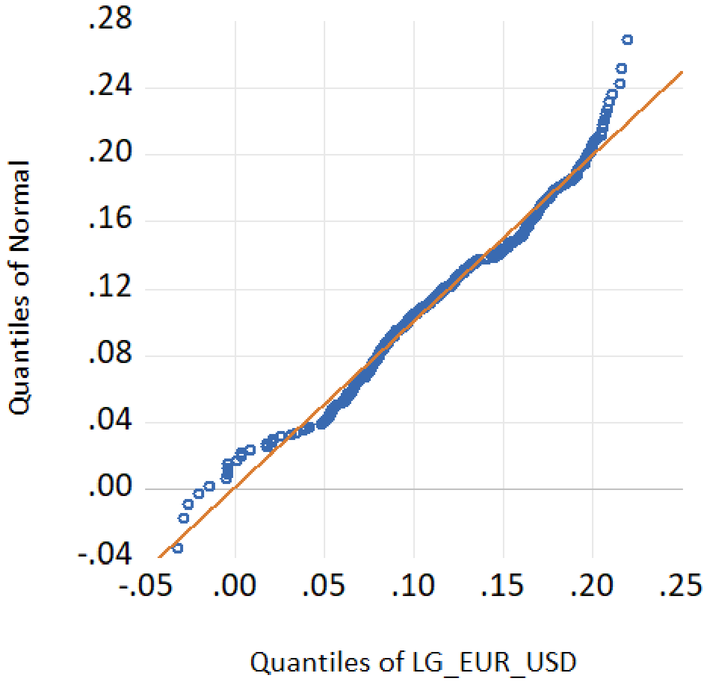
LITECOIN



GBTC



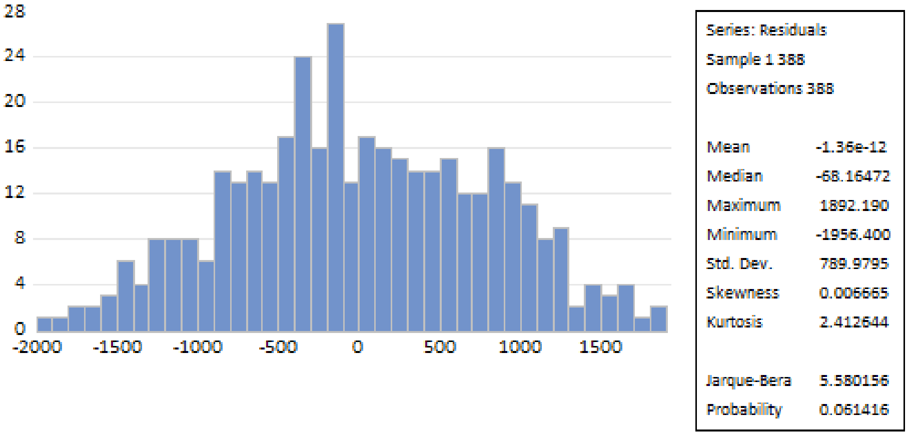
EUR/USD



JARQUE-BERA TEST:

Now we move on to the Jarque-Bera test, for this, we take estimate our model and then take the residuals and check their normality to see if they are normally distributed.

We go into E-views and estimate our model.



After estimating our model, we go into the descriptive stats and look the Jarque-Bera tabulated value and the probability of the residuals.

Here, the following two things need to happen to so that we can assess the following hypothesis.

H_0: The residuals are normally distributed.

H_1: The residuals are not normally distributed.

So, we look into the values given and compare them with what is a set standard, for the Jarque-Bera test we need the Jarque-Bera factor to be less than 1.76 and the probability value needs to be greater than 0.05.

In this case, we come to the conclusion that the residuals are in fact normal and we fail to reject the null hypothesis, which states, the residuals are normally distributed, and they follow a bell-shaped curve.

CORRELATION ANALYSIS:

CORRELATION MATRIX:

	NASDAQ	LG_BTC	ETH	LTC	GBTC	EUR_USD
NASDAQ	1.000000	0.925418	0.859858	0.163167	0.856291	-0.329473
LG_BTC	0.925418	1.000000	0.895691	0.437440	0.724788	-0.219723
ETH	0.859858	0.895691	1.000000	0.448356	0.641422	-0.192364
LTC	0.163167	0.437440	0.448356	1.000000	-0.135629	0.496422
GBTC	0.856291	0.724788	0.641422	-0.135629	1.000000	-0.576778
EUR_...	-0.329473	-0.219723	-0.192364	0.496422	-0.576778	1.000000

CORRELATION MATRIX EXPLAINED:

For correlation, it explains the linear relationship between the variables and this is indicated by the values of the coefficients and it is ranged between -1 to 1 where:

- **+1**: Perfect positive correlation (as one variable increases, the other increases proportionally).
- **0**: No correlation.
- **-1**: Perfect negative correlation (as one variable increases, the other decreases proportionally).

CORRELATION BETWEEN VARIABLES:

Now, after getting out results we will be explaining the correlation matrix and the values obtained from it as follows:

NASDAQ vs. CRYPTOCURRENCIES CORRELATION:

- **NASDAQ & BTC (0.925418)**:
  There is a substantial positive link between the movements of the NASDAQ 100 and the weekly logarithmic returns of Bitcoin. This implies that there is a

substantial correlation between the price trends of Bitcoin and the success of the stock market.

- **NASDAQ & ETH (0.859858)**:
  Although not as strong as Bitcoin's, a substantial positive correlation indicates Ethereum's close link with the NASDAQ.

- **NASDAQ & LTC (0.163167)**:
  A weak positive correlation, indicating Litecoin's movements are less synchronized with NASDAQ compared to other cryptocurrencies.

INTER-CRYPTOCURRENCY CORRELATION:

- **BTC & ETH (0.895691)**:
  Bitcoin and Ethereum share a **very strong positive correlation**, indicating that their price trends move in tandem most of the time. This is expected as they are the two largest cryptocurrencies by market capitalization and are often influenced by similar market dynamics, news, and investor sentiment.

- **BTC & LTC (0.4347440)**:
  Bitcoin and Litecoin have a **moderate positive correlation**, meaning they share some similarities in price movements but are not as closely aligned as BTC with ETH or DOGE. Litecoin, being an earlier altcoin, may have its price influenced by unique factors (e.g., its use as a testbed for Bitcoin innovations).

- **ETH & LTC (0.448356)**:
  Ethereum and Litecoin have a **moderate positive correlation**, indicating less price dependency compared to ETH’s relationship with BTC and DOGE. This suggests that Litecoin's market behavior is less influenced by Ethereum's trends.

STATIONARITY OF VARIABLES:

Stationarity at I (1) via ADF (Augmented Dicky Fuller) Testing:

Time Series	ADF stat	at 1%	at 5%	at 10%	p-value
Nasdaq (NDX)	-20.35719	-3.982011	-3.421508	-3.133535	0.0000
Bitcoin (BTC)	-18.31172	-3.982011	-3.421508	-3.133535	0.0000
Ethereum (ETH)	-15.47034	-3.982074	-3.421539	-3.133553	0.0000
Litecoin (LTC)	-17.59367	-3.982074	-3.421539	-3.133553	0.0000
Grayscale Bitcoin ETF (GBTC)	-17.46934	-3.982011	-3.421508	-3.133535	0.0000
EUR/USD	-22.17754	-3.982011	-3.421508	-3.133535	0.0000

After assessing these, we come to the conclusion that all the variables are stationary at first difference I (1).

MULTIPLE REGRESSION ANALYSIS:

Dependent Variable: NASDAQ
Method: Least Squares
Date: 12/11/24 Time: 10:30
Sample: 1 388
Included observations: 388
Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LG_BTC	3005.882	136.2210	22.06621	0.0000
ETH	0.843489	0.077499	10.88385	0.0000
LTC	-21.30266	1.501086	-14.19149	0.0000
GBTC	67.76639	5.525758	12.26373	0.0000
EUR_USD	11638.16	1029.026	11.30987	0.0000
C	-31505.43	1621.120	-19.43436	0.0000

R-squared	0.961671	Mean dependent var	11525.89
Adjusted R-squared	0.961169	S.D. dependent var	4115.657
S.E. of regression	811.0105	Akaike info criterion	16.24978
Sum squared resid	2.51E+08	Schwarz criterion	16.31103
Log likelihood	-3146.458	Hannan-Quinn criter.	16.27407
F-statistic	1916.871	Durbin-Watson stat	0.301517
Prob(F-statistic)	0.000000	Wald F-statistic	2163.703
Prob(Wald F-statistic)	0.000000		

Interpretation:

Coefficients:

- **Intercept (C):** The constant term, -31505.43, represents the NASDAQ's predicted value when all independent variables are zero.
- **BTC (Bitcoin):** A coefficient of 3005.882 indicates that a 1-unit increase in BTC leads to a 3005.882 -unit increase in NASDAQ, holding other variables constant.
- **ETH (Ethereum):** A coefficient of 0.843489 suggests that a 1-unit increase in ETH results in a 0.843489 -unit increase in NASDAQ, holding other variables constant.
- **LTC (Litecoin):** A coefficient of -21.30266 indicates that a 1-unit increase in LTC leads to a 21.30266-unit decrease in NASDAQ, holding other variables constant.
- **GBTC (Grayscale Bitcoin):** A coefficient of 67.76639 suggests that a 1-unit increase in DOGE results in a 67.76639-unit increase in NASDAQ.
- **EUR/USD**: A coefficient of 11638.16 suggests that a 1-unit increase in DOGE results in a 11638.16-unit increase in NASDAQ.

"A key thing to note here would be to understand that in order to get significant results, log of GBTC has been taken to make sure it does not give us problems later on like multi-collinearity and other problems."

Significance:

Individual Hypothesis

- **Null Hypothesis (H_0):**
  - $\beta 1=0$
  - $\beta 2=0$

- $\beta3=0$
- $\beta4=0$

- **Alternative Hypothesis (H_1):**
  - $\beta1\neq0$
  - $\beta2\neq0$
  - $\beta3\neq0$
  - $\beta4\neq0$

a. T-values:

- **BTC:**
  - tcal|=22.06621>2 (benchmark).
  - We reject H0 and conclude that there is a significant relationship between NASDAQ and BTC.
- **ETH:**
  - tcal|=10.88385>2 (benchmark).
  - We reject H0 and conclude that there is a significant relationship between NASDAQ and ETH.
- **LTC:**
  - |tcal|=-14.19149>2 (benchmark).
  - We reject H0 and conclude that there is a significant relationship between NASDAQ and LTC.
- **GBTC:**
  - tcal|=12.26373>2 (benchmark).
  - We reject H0 and conclude that there is a significant relationship between NASDAQ and GBTC.

- **EUR/USD:**
  - tcal|=11.30987>2 (benchmark).
  - We reject H0 and conclude that there is a significant relationship between NASDAQ and EUR/USD.

b. P-values at 5% level of significance:

- **BTC:** p=0.0000<0.05, so we reject H0 and conclude that there is a significant relationship between NASDAQ and BTC.
- **ETH:**p=0.0000<0.05, so we reject H0 and conclude that there is a significant relationship between NASDAQ and ETH.
- **LTC:** p=0.0000<0.05, so we reject H0 and conclude that there is a significant relationship between NASDAQ and LTC.
- **GBTC:** p=0.0000<0.05, so we reject H0 and conclude that there is a significant relationship between NASDAQ and GBTC.
- **EUR/USD:** p=0.0000<0.05, so we reject H0 and conclude that there is a significant relationship between NASDAQ and EUR/USD.

Model Fit:

- **R^2 = 0.961671 or 96.1671%**
  This indicates that **96.1671**% of the variation in NASDAQ is explained by BTC, ETH, LTC, and DOGE, while the remaining is explained by other factors not included in the model.
- **Adjusted R^2 = 0.961169**
  This adjusted value accounts for the number of predictors in the model and confirms a strong model fit.

DIAGNOSTIC TESTS

MULTICOLLINEARITY

When independent variables in your regression model have a high degree of correlation with one another, this is known as multicollinearity, and it can skew the estimates of your regression coefficients.

VARIANCE INFLATION FACTOR (VIF)

TABLE FOR VIF

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
LG_BTC	18556.16	1122.689	9.943675
ETH	0.006006	13.10171	5.319776
LTC	2.253260	14.29306	3.533292
GBTC	30.53400	16.45677	5.146750
EUR_USD	1058895.	817.7363	1.958466
C	2628030.	1631.990	NA

EXPLANATION

Interpreting VIF Values

- **VIF < 10**: Indicates there is low multicollinearity and it's not a concern.
- **VIF > 10**: Suggests high multicollinearity and may require action to address it.

Here,

- **BTC**: The Centered VIF is 9.790880, which is below 10, indicating that BTC is not experiencing severe multicollinearity.

- **ETH**: The Centered VIF is 8.031357, which is below 10, suggesting that ETH has low multicollinearity and it is not a concern.

- **LTC**: The Centered VIF is 2.364753, which is well below 10, indicating no multicollinearity issue.

- **GBTC**: The Centered VIF is 5.452328, below 10, so there is no concern of multicollinearity.

- **EUR/USD:** The Centered VIF is 2.971264, below 10, so there is no concern of multicollinearity.

HETEROSCEDASTICITY

When the variance of the residuals varies across all levels of the independent variables, this is known as heteroskedasticity. Hypothesis tests may become unreliable as a result of this violation of the homoscedasticity assumption.

Heteroskedasticity causes inefficient coefficient estimates and invalid standard errors, leading to unreliable t- and F-tests.

WHITE TEST

Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	3.503942	Prob. F(20,367)	0.0000
Obs*R-squared	62.20981	Prob. Chi-Square(20)	0.0000
Scaled explained SS	44.43614	Prob. Chi-Square(20)	0.0013

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 12/11/24 Time: 10:05
Sample: 1 388
Included observations: 388

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-63641473	47648204	-1.335653	0.1825
LG_BTC^2	-8781.434	438743.2	-0.020015	0.9840
LG_BTC*ETH	-310.2896	429.9922	-0.721617	0.4710
LG_BTC*LTC	-989.4735	6703.438	-0.147607	0.8827
LG_BTC*GBTC	-56551.21	59411.02	-0.951864	0.3418
LG_BTC*EUR_USD	-1865401.	4899125.	-0.380762	0.7036
LG_BTC	3534179.	6579734.	0.537131	0.5915
ETH^2	-0.306026	0.151167	-2.024421	0.0437
ETH*LTC	9.664104	2.872438	3.364426	0.0008
ETH*GBTC	39.17647	17.31756	2.262240	0.0243
ETH*EUR_USD	-6908.152	3076.324	-2.245587	0.0253
ETH	10481.07	4248.904	2.466771	0.0141
LTC^2	-24.56818	30.33287	-0.809952	0.4185
LTC*GBTC	-15.78533	538.3052	-0.029324	0.9766
LTC*EUR_USD	-9700.116	45577.58	-0.212826	0.8316
LTC	16389.13	78149.03	0.209716	0.8340
GBTC^2	2259.857	1636.614	1.380813	0.1682
GBTC*EUR_USD	26920.36	238948.9	0.112662	0.9104
GBTC	324651.6	485401.7	0.668831	0.5040
EUR_USD^2	-21134190	22355769	-0.945357	0.3451
EUR_USD	71467985	60325498	1.184706	0.2369

R-squared	0.160335	Mean dependent var	647566.8
Adjusted R-squared	0.114576	S.D. dependent var	787167.3
S.E. of regression	740700.5	Akaike info criterion	29.92118
Sum squared resid	2.01E+14	Schwarz criterion	30.13557
Log likelihood	-5783.710	Hannan-Quinn criter.	30.00618
F-statistic	3.503942	Durbin-Watson stat	0.854311

Since the p-value is less than 0.05, that means we accept the Null Hypothesis stating that there is heterokedasticity present in our model.

This infers that there is no constant mean and constant variance across the model.

However, in order to tackle this problem we will have used robust mean method to cater the issue of heteroscedasticity here.

It will be shown in an MLS output below after all the tests.

BREUSCH-PAGAN-GODFREY TEST

Heteroskedasticity Test: Breusch-Pagan-Godfrey
Null hypothesis: Homoskedasticity

F-statistic	2.690594	Prob. F(5,382)	0.0209
Obs*R-squared	13.19943	Prob. Chi-Square(5)	0.0216
Scaled explained SS	9.428282	Prob. Chi-Square(5)	0.0932

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 12/11/24 Time: 10:06
Sample: 1 388
Included observations: 388

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1042650.	1551854.	-0.671874	0.5021
LG_BTC	-42403.37	124877.6	-0.339559	0.7344
ETH	37.08359	76.01843	0.487824	0.6260
LTC	-128.0006	1309.537	-0.097745	0.9222
GBTC	-3500.148	5475.500	-0.639238	0.5231
EUR_USD	1901222.	1012803.	1.877189	0.0613

R-squared	0.034019	Mean dependent var	647566.8
Adjusted R-squared	0.021375	S.D. dependent var	787167.3
S.E. of regression	778708.8	Akaike info criterion	29.98401
Sum squared resid	2.32E+14	Schwarz criterion	30.04526
Log likelihood	-5810.897	Hannan-Quinn criter.	30.00829
F-statistic	2.690594	Durbin-Watson stat	0.795120
Prob(F-statistic)	0.020949		

As we can see that the Chi-Squared Probability is greater than 0.05, this tells us that we should be accepting the Null Hypothesis and claiming that homoskedastcity exists in this model.

This infers that there is constant mean and constant variance across the model.

The functional form for this test will be:

Resid^2 = C+ LG_BTC + ETH + LTC + GBTC + EUR_USD

HARVEY TEST

Heteroskedasticity Test: Harvey
Null hypothesis: Homoskedasticity

F-statistic	2.891751	Prob. F(5,382)	0.0141
Obs*R-squared	14.15027	Prob. Chi-Square(5)	0.0147
Scaled explained SS	7.697453	Prob. Chi-Square(5)	0.1737

Test Equation:
Dependent Variable: LRESID2
Method: Least Squares
Date: 12/11/24 Time: 10:09
Sample: 1 388
Included observations: 388

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.637280	3.230121	2.054808	0.0406
LG_BTC	-0.163310	0.259928	-0.628288	0.5302
ETH	0.000137	0.000158	0.868174	0.3858
LTC	-0.001675	0.002726	-0.614526	0.5392
GBTC	0.003346	0.011397	0.293612	0.7692
EUR_USD	6.577825	2.108108	3.120251	0.0019

R-squared	0.036470	Mean dependent var	12.55387
Adjusted R-squared	0.023858	S.D. dependent var	1.640539
S.E. of regression	1.620850	Akaike info criterion	3.819122
Sum squared resid	1003.574	Schwarz criterion	3.880375
Log likelihood	-734.9097	Hannan-Quinn criter.	3.843408
F-statistic	2.891751	Durbin-Watson stat	1.096867
Prob(F-statistic)	0.014113		

Here we also see that the Chi-Square Probability is greater than 0.05 which leads us to the conclusion that Heteroskedasticity does not exist in the model.

Harvey test is distinct as it tests Heteroskedasticity by estimating the model by taking logarithmic residual^2 as the dependent variable and keeping the other independent variables the same as the regular OLS ones.

The function form for this becomes:

Log (Rsesid^2) =C + LG_BTC+ ETH + LTC+ GBTC + EUR_USD

As for the concern of White test and it not proving significant for homoscedasticity, we apply robust mean methods which makes an exception for the heteroskedasticity scenario.

Since, the white test is more robust and provides a non-parametric idea for heteroskedasticity, we apply robust mean method and use the following as our model, this will be used as our model which has its own VIFs (controlled) and the residuals are normal as well and the model is also consistent (will be proven in Ramsey RESET test)

We run the regular OLS model and make a few changes.

Dependent Variable: NASDAQ
Method: Least Squares
Date: 12/11/24 Time: 10:39
Sample: 1 388
Included observations: 388
Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LG_BTC	3005.882	136.2210	22.06621	0.0000
ETH	0.843489	0.077499	10.88385	0.0000
LTC	-21.30266	1.501086	-14.19149	0.0000
GBTC	67.76639	5.525758	12.26373	0.0000
EUR_USD	11638.16	1029.026	11.30987	0.0000
C	-31505.43	1621.120	-19.43436	0.0000

R-squared	0.961671	Mean dependent var	11525.89
Adjusted R-squared	0.961169	S.D. dependent var	4115.657
S.E. of regression	811.0105	Akaike info criterion	16.24978
Sum squared resid	2.51E+08	Schwarz criterion	16.31103
Log likelihood	-3146.458	Hannan-Quinn criter.	16.27407
F-statistic	1916.871	Durbin-Watson stat	0.301517
Prob(F-statistic)	0.000000	Wald F-statistic	2163.703
Prob(Wald F-statistic)	0.000000		

Here, it clearly explains that it caters for the heteroskedasticity consistent standard errors and covariance.

AUTOCORRELATION

When the residuals are not independent of one another, autocorrelation arises, which goes against the presumption that errors are random.

We can detect Autocorrelation in our model by running the following tests on our model.

DURBIN WATSON TEST

Dependent Variable: NASDAQ
Method: Least Squares
Date: 12/11/24 Time: 10:39
Sample: 1 388
Included observations: 388
Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LG_BTC	3005.882	136.2210	22.06621	0.0000
ETH	0.843489	0.077499	10.88385	0.0000
LTC	-21.30266	1.501086	-14.19149	0.0000
GBTC	67.76639	5.525758	12.26373	0.0000
EUR_USD	11638.16	1029.026	11.30987	0.0000
C	-31505.43	1621.120	-19.43436	0.0000

R-squared	0.961671	Mean dependent var	11525.89
Adjusted R-squared	0.961169	S.D. dependent var	4115.657
S.E. of regression	811.0105	Akaike info criterion	16.24978
Sum squared resid	2.51E+08	Schwarz criterion	16.31103
Log likelihood	-3146.458	Hannan-Quinn criter.	16.27407
F-statistic	1916.871	Durbin-Watson stat	0.301517
Prob(F-statistic)	0.000000	Wald F-statistic	2163.703
Prob(Wald F-statistic)	0.000000		

Here, we look at the Durbin Watson statistic and as we can see that it is close to zero which makes it near positive autocorrelation according to the following threshold:

DW ≈ 2: No autocorrelation.

DW = 2: Positive autocorrelation.

DW = 4: Negative autocorrelation.

Since, Heteroskedasticity exists here we go a step further to make our results more concise and consistent.

DURBIN WATSON -h TEST

To apply Durbin h test, we use to following formula:

h= (1- d/2) √(n-1)/(1-n*var)

Here,

d= regular Durbin Watson stat (0.301517)

n= number of observations (388)

var = is the estimated variance of the coefficient of the lagged dependent variable (0.000731323849)

And to achieve these values we run the serial correlation LM test which gives us this answer.

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 1 lag

F-statistic	996.6966	Prob. F(1,381)	0.0000
Obs*R-squared	280.6992	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 12/11/24 Time: 11:40
Sample: 1 388
Included observations: 388
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LG_BTC	-124.5307	68.59779	-1.815374	0.0703
ETH	0.020637	0.041694	0.494964	0.6209
LTC	1.386062	0.719505	1.926409	0.0548
GBTC	3.992011	3.005486	1.328241	0.1849
EUR_USD	-358.3802	555.5483	-0.645093	0.5193
C	1375.811	852.1693	1.614481	0.1073
RESID(-1)	0.853748	0.027043	31.57050	0.0000

R-squared	0.723451	Mean dependent var	-1.79E-12
Adjusted R-squared	0.719096	S.D. dependent var	805.7543
S.E. of regression	427.0526	Akaike info criterion	14.96957
Sum squared resid	69484457	Schwarz criterion	15.04103
Log likelihood	-2897.096	Hannan-Quinn criter.	14.99790
F-statistic	166.1161	Durbin-Watson stat	1.880047
Prob(F-statistic)	0.000000		

Now we take the values and get the h stat:

h= 14.138974417

Now, as the threshold gives us the reasoning that the lagged variable's probability should be less than 0.005 and it is 0.0000,

AND

The h stat value is greater than 1.96 which gives us proof that serial correlation exists and we reject the Null Hypothesis saying:

H0: No serial correlation at up to 1 lag.

Ramsey RESET Test

Now we apply the Ramsey RESET test in order to check any misspecifications in our model.

After application of the test, we get:

Ramsey RESET Test
Equation: UNTITLED
Omitted Variables: Squares of fitted values
Specification: NASDAQ LG_BTC ETH LTC GBTC EUR_USD C

	Value	df	Probability
t-statistic	0.105633	381	0.9159
F-statistic	0.011158	(1, 381)	0.9159
Likelihood ratio	0.011363	1	0.9151

This table shows a p-value of 0.9151 which means that there is no misspecification in our model and this model is optimal.

CORRELOGRAM OF RESIDUALS

Correlogram of Residuals

Date: 12/11/24 Time: 12:19
Sample: 1 388
Included observations: 388

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.847	0.847	280.77	0.000
		2	0.705	-0.048	475.39	0.000
		3	0.610	0.088	621.59	0.000
		4	0.520	-0.036	728.08	0.000
		5	0.455	0.055	809.95	0.000
		6	0.424	0.082	881.17	0.000
		7	0.406	0.052	946.52	0.000
		8	0.399	0.066	1009.9	0.000
		9	0.383	-0.009	1068.4	0.000
		10	0.338	-0.072	1114.2	0.000
		11	0.299	0.007	1150.0	0.000
		12	0.257	-0.034	1176.5	0.000
		13	0.216	-0.006	1195.4	0.000
		14	0.182	-0.021	1208.7	0.000
		15	0.153	-0.015	1218.3	0.000
		16	0.112	-0.078	1223.4	0.000
		17	0.082	0.001	1226.1	0.000
		18	0.065	0.010	1227.9	0.000
		19	0.059	0.035	1229.3	0.000
		20	0.065	0.043	1231.0	0.000
		21	0.043	-0.089	1231.8	0.000
		22	0.023	0.005	1232.0	0.000
		23	0.015	0.014	1232.1	0.000
		24	0.001	-0.006	1232.1	0.000
		25	0.007	0.082	1232.1	0.000
		26	0.018	0.005	1232.3	0.000
		27	0.021	-0.007	1232.5	0.000

The result of the correlogram shows us that autocorrelation exists in our model, this explains that the residuals are not independent of each other which leads us to believe that the lagged variables are connected and it is by one another.

We look at the p-value here, it is visible that they are less than 0.05 which indicates the presence of Autocorrelation.

Generally, autocorrelation can never be zero in a model as dependency of events across the model is something which is inevitable.

As autocorrelation is present in our residuals:

1. **Include Lagged Terms**:
   - Add lagged dependent or independent variables to the model.

This would be our go to remedy to decrease the autocorrelation problem.

As this can create this problem:

"Autocorrelation invalidates standard errors and makes coefficient estimates inefficient."

After including lagged term AR (1) in our model we found out that the autocorrelation controlled and there is no fluctuated spike in the ACF PACF.

Correlogram of Residuals

Date: 12/11/24 Time: 12:13
Sample: 1 388
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.046	-0.046	0.8119	
		2	0.004	0.001	0.8167	0.366
		3	0.013	0.013	0.8818	0.643
		4	-0.060	-0.059	2.2807	0.516
		5	-0.062	-0.068	3.8227	0.431
		6	0.050	0.045	4.8230	0.438
		7	-0.060	-0.055	6.2709	0.394
		8	-0.024	-0.032	6.4909	0.484
		9	0.049	0.038	7.4306	0.491
		10	0.013	0.020	7.4937	0.586
		11	0.005	0.005	7.5024	0.677
		12	-0.082	-0.097	10.233	0.510
		13	0.005	0.004	10.242	0.595
		14	0.033	0.042	10.678	0.638
		15	0.115	0.118	15.997	0.314
		16	-0.046	-0.046	16.858	0.327
		17	0.010	-0.005	16.898	0.392
		18	0.032	0.046	17.320	0.433
		19	-0.044	-0.034	18.131	0.447
		20	0.145	0.147	26.719	0.111
		21	0.034	0.043	27.185	0.130
		22	-0.029	0.001	27.527	0.154
		23	-0.050	-0.063	28.582	0.157
		24	0.026	0.008	28.853	0.185

JOHANSEN CO-INTEGRATION TEST

As the variables are all stationary at first difference I (1), we use the Johansen Cointegration test and see the long-term relationships in our model. As the Ramsey RESET test also showed that there are no misspecifications in our model, this model is sound to be tested by the Johansen Cointegration Test.

TRACE TEST

No of CE's	Eigenvalue	Trace Stat	0.05 Critical Value	Prob
None*	0.237912	195.155	103.8473	0.0000
At most 1*	0.105792	90.55312	76.97277	0.0032
At most 2	0.054954	47.50341	54.07904	0.1692
At most 3	0.034845	25.74259	35.199275	0.3565
At most 4	0.019229	12.08813	20.26184	0.441
At most 5	0.01191	4.612992	9.164546	0.3285

After the Johansen Cointegration test is applied, we look at Trace Test here which tells us how many cointegrating equations exist.

According to the trace test, we can see that there are at most 2 co-integrating equations.

This output from E-views gives the trace test table. Trace test indicates 2 cointegrating eqn(s) at the 0.05 level.

RANK TEST

In the rank test, we look at the maximum Eigen Value and look at our table and see for cointegrating equations.

No of CE's	Eigenvalue	Trace Stat	0.05 Critical Value	Prob
None*	0.237912	104.6019	40.9568	0.0000
At most 1*	0.105792	43.04971	34.80587	0.0042
At most 2	0.054954	21.76083	28.58808	0.2897
At most 3	0.034845	13.65445	22.29962	0.4941
At most 4	0.019229	7.47514	15.8921	0.6113
At most 5	0.01191	4.612992	9.164546	0.3285

Max-eigenvalue test indicates 2 cointegrating equations at the 0.05 level.

"*" denotes the rejection of the hypothesis at the 0.05 level.

Now, we look at the cointegrating equation:

1 Cointegrating Equation(s):	Log likelihood	-5190.758				
Normalized cointegrating coefficients (standard error in parentheses)						
NASDAQ	LG_BTC	ETH	GBTC	EUR_USD	LTC	C
1.000000	-3094.991	-1.005364	-211.9446	-9197.858	22.24454	29326.57
	(870.956)	(0.52571)	(38.7634)	(7086.09)	(9.50925)	(11028.0)

Here, we can obtain the t-stat values from this table and see if the variables are significant or not which in this case are significant, which tells us that there exists a long-term relationship between our variables in the model one way or another.

Vector Error Correction Model (VECM)

In order to see our model exhibits long-term relationship, we need to see if the ECT (Error Correction Term) is stationary at level I (0) and all the other variables are stationary at first difference. So, we run the stationarity test on our error term and we find out that:

Augmented Dickey-Fuller Unit Root Test on VECM

Null Hypothesis: VECM has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=16)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.574405	0.0000
Test critical values:	1% level	-3.981949	
	5% level	-3.421478	
	10% level	-3.133517	

Here the value is significant and we reject the null hypothesis which states, VECM has a unit root, this means that our error correction term is stationary at level and therefore exhibits a long-term relationship in our model.

Granger Casualty Test

Now, we move onto the next thing, which is Granger Casualty Test which we obtained from E-Views as follows.

Pairwise Granger Causality Tests
Date: 12/11/24   Time: 13:56
Sample: 1 388
Lags: 1

Null Hypothesis:	Obs	F-Statistic	Prob.
LG_BTC does not Granger Cause NASDAQ	387	4.07745	0.0442
NASDAQ does not Granger Cause LG_BTC		3.11340	0.0784
ETH does not Granger Cause NASDAQ	387	5.09070	0.0246
NASDAQ does not Granger Cause ETH		2.64422	0.1047
GBTC does not Granger Cause NASDAQ	387	0.49558	0.4819
NASDAQ does not Granger Cause GBTC		0.34821	0.5555
EUR_USD does not Granger Cause NASDAQ	387	2.92783	0.0879
NASDAQ does not Granger Cause EUR_USD		0.07626	0.7826
LTC does not Granger Cause NASDAQ	387	0.05119	0.8211
NASDAQ does not Granger Cause LTC		0.52973	0.4672
ETH does not Granger Cause LG_BTC	387	6.05274	0.0143
LG_BTC does not Granger Cause ETH		0.04968	0.8237
GBTC does not Granger Cause LG_BTC	387	1.06556	0.3026
LG_BTC does not Granger Cause GBTC		0.25454	0.6142
EUR_USD does not Granger Cause LG_BTC	387	2.10026	0.1481
LG_BTC does not Granger Cause EUR_USD		0.11747	0.7320
LTC does not Granger Cause LG_BTC	387	0.48898	0.4848
LG_BTC does not Granger Cause LTC		0.75395	0.3858
GBTC does not Granger Cause ETH	387	2.44957	0.1184
ETH does not Granger Cause GBTC		0.23872	0.6254
EUR_USD does not Granger Cause ETH	387	3.32589	0.0690
ETH does not Granger Cause EUR_USD		0.31322	0.5760
LTC does not Granger Cause ETH	387	10.2805	0.0015
ETH does not Granger Cause LTC		2.66191	0.1036
EUR_USD does not Granger Cause GBTC	387	0.27377	0.6011
GBTC does not Granger Cause EUR_USD		1.87094	0.1722
LTC does not Granger Cause GBTC	387	0.89704	0.3442
GBTC does not Granger Cause LTC		0.01895	0.8906
LTC does not Granger Cause EUR_USD	387	1.30218	0.2545
EUR_USD does not Granger Cause LTC		2.22288	0.1368

This table explains how the variables have predictive causality between one another, these results show if the causal relationship is bidirectional, no causality and unidirectional causality and it can be measured based on the following criteria.

H0: There is no Granger Cause.

H1: There is Granger Cause.

If the p-value is less than 0.05 then we reject Ho and if it is greater than 0.05, there is a granger cause.

MODEL SELECTION:

ARMA/ARIMA Modelling

So, for the selection of ARMA/ARIMA models, there is a procedure that needs to be followed to get there, we incorporated the following steps for the model selection.

We will be assessing Nasdaq (NDX) for the modelling basis. Since we used the log values in the normality procedure, we will take log of NDX.

Before the procedure, the rule of thumb for ARMA/ARIMA model is as follows:

INSTRUCTIONS:

Model Selection (ARMA vs. ARIMA):

- **ARMA (p, q)**: Use when the time series is stationary (no trends or seasonality).
- ARIMA (p, d, q): Use when the time series is non-stationary (trend or seasonality present). The "d" (differencing) removes non-stationarity.

Checking Stationarity:

- Use the **Augmented Dickey-Fuller (ADF) test** or **visual inspection**.
  - If $p>0.05$, the series is non-stationary → Apply differencing (d).
  - If $p\leq 0.05$, the series is stationary → No differencing needed ($d=0$).

Parameter Selection:

- Autoregressive (AR, p): Determine from the Partial Autocorrelation Function (PACF).
    1. Significant spikes at lag k in the PACF suggest AR terms up to p=k.

- **Moving Average (MA, q)**: Determine from the **Autocorrelation Function (ACF)**
    2. Significant spikes at lag k in the ACF suggest MA terms up to q=k.

- **Differencing (d)**: Number of times the series is differenced to achieve stationarity.
    3. Start with d=1 and test stationarity after each differencing.

APPLICATION:

Nasdaq (NDX):

We took log (NDX) first, and we checked the stationarity:

At level I (0):

Augmented Dickey-Fuller Unit Root Test on LG_NASDAQ

Null Hypothesis: LG_NASDAQ has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=16)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.226742	0.4728
Test critical values:	1% level	-3.981949	
	5% level	-3.421478	
	10% level	-3.133517	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LG_NASDAQ)
Method: Least Squares
Date: 12/01/24 Time: 18:24
Sample (adjusted): 2 388
Included observations: 387 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LG_NASDAQ(-1)	-0.025466	0.011437	-2.226742	0.0265
C	0.248521	0.113115	2.197065	0.0286
@TREND("1")	-7.82E-05	3.78E-05	-2.069194	0.0392

R-squared	0.012752	Mean dependent var	-0.003167
Adjusted R-squared	0.007610	S.D. dependent var	0.029461
S.E. of regression	0.029348	Akaike info criterion	-4.211435
Sum squared resid	0.330751	Schwarz criterion	-4.180749
Log likelihood	817.9126	Hannan-Quinn criter.	-4.199267
F-statistic	2.479964	Durbin-Watson stat	2.086152
Prob(F-statistic)	0.085087		

Since it proves to be insignificant as the p-value is greater than 0.05 and the tabulated value is also less than the critical values.

At First Difference I (1):

Augmented Dickey-Fuller Unit Root Test on D(LG_NASDAQ)

Null Hypothesis: D(LG_NASDAQ) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=16)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-20.73082	0.0000
Test critical values:	1% level	-3.982011	
	5% level	-3.421508	
	10% level	-3.133535	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LG_NASDAQ,2)
Method: Least Squares
Date: 12/01/24 Time: 18:28
Sample (adjusted): 3 388
Included observations: 386 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LG_NASDAQ(-1))	-1.058099	0.051040	-20.73082	0.0000
C	-0.003650	0.003026	-1.206178	0.2285
@TREND("1")	1.26E-06	1.35E-05	0.093648	0.9254

R-squared	0.528781	Mean dependent var	2.13E-05
Adjusted R-squared	0.526321	S.D. dependent var	0.042877
S.E. of regression	0.029510	Akaike info criterion	-4.200451
Sum squared resid	0.333526	Schwarz criterion	-4.169706
Log likelihood	813.6871	Hannan-Quinn criter.	-4.188259
F-statistic	214.8930	Durbin-Watson stat	1.997063
Prob(F-statistic)	0.000000		

As the p-value is significant, being less than 0.05 and the tabulated value is also greater than the critical values, this shows us that the series is stationary at first difference.

By having our series stationary at first difference, we come to the conclusion that we will have an ARIMA model which is going to be ARIMA (p ,1, q) model.

Now,

We take the differenced series and run a correlogram on it.

Date: 12/01/24 Time: 18:35
Sample (adjusted): 2 388
Included observations: 387 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.058	-0.058	1.3109	0.252
		2	0.012	0.008	1.3639	0.506
		3	0.034	0.035	1.8131	0.612
		4	-0.068	-0.064	3.6044	0.462
		5	-0.105	-0.114	7.9387	0.160
		6	0.053	0.041	9.0502	0.171
		7	-0.032	-0.019	9.4538	0.222
		8	-0.057	-0.061	10.747	0.216
		9	0.049	0.026	11.704	0.231
		10	-0.021	-0.019	11.882	0.293
		11	0.045	0.052	12.688	0.314
		12	-0.064	-0.078	14.333	0.280
		13	-0.010	-0.024	14.372	0.348
		14	0.041	0.051	15.047	0.375
		15	0.050	0.058	16.075	0.377
		16	-0.043	-0.040	16.822	0.397
		17	0.021	-0.007	17.004	0.454
		18	0.029	0.041	17.350	0.499
		19	-0.059	-0.032	18.752	0.473
		20	0.071	0.053	20.805	0.409
		21	0.034	0.037	21.276	0.442
		22	-0.041	-0.022	21.965	0.462
		23	-0.074	-0.079	24.204	0.393
		24	0.025	0.001	24.453	0.436
		25	0.074	0.111	26.704	0.371
		26	-0.033	-0.024	27.168	0.401
		27	-0.019	-0.046	27.314	0.447
		28	0.016	0.006	27.427	0.495
		29	-0.012	0.010	27.486	0.545
		30	0.003	0.018	27.491	0.597
		31	0.066	0.034	29.330	0.552
		32	0.082	0.106	32.175	0.458
		33	-0.034	-0.006	32.659	0.484
		34	0.017	-0.007	32.778	0.527
		35	0.116	0.108	38.537	0.313
		36	-0.085	-0.049	41.658	0.238

Looking at significant spikes we see that, AR (5) & MA (5)

AR (35) & MA (35)

These will turn into:

ARIMA (5,1,5)

ARIMA (35,1,35)

Now, we check these models by running them on E-views by estimating them one by one.

ARIMA (5,1,5)

By estimating:

d (lg_nasdaq) ar (5) ma (5) c

we get:

Dependent Variable: D(LG_NASDAQ)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/01/24 Time: 18:43
Sample: 2 388
Included observations: 387
Convergence achieved after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.003186	0.001375	-2.317600	0.0210
AR(5)	0.027365	0.454562	0.060200	0.9520
MA(5)	-0.134535	0.452110	-0.297571	0.7662
SIGMASQ	0.000856	4.65E-05	18.38723	0.0000

R-squared	0.011572	Mean dependent var	-0.003167
Adjusted R-squared	0.003830	S.D. dependent var	0.029461
S.E. of regression	0.029404	Akaike info criterion	-4.204922
Sum squared resid	0.331146	Schwarz criterion	-4.164008
Log likelihood	817.6524	Hannan-Quinn criter.	-4.188699
F-statistic	1.494666	Durbin-Watson stat	2.116280
Prob(F-statistic)	0.215546		

Inverted AR Roots	.49	.15+.46i	.15-.46i	-.39-.29i
	-.39+.29i			
Inverted MA Roots	.67	.21+.64i	.21-.64i	-.54-.39i
	-.54+.39i			

Findings:

From this estimation we see that the AR and MA variables are not significant at all so we cannot be using this model.

ARIMA (35,1,35)

By estimating:

d(lg_nasdaq) ar(35) ma(35) c

we get:

Dependent Variable: D(LG_NASDAQ)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/01/24 Time: 18:46
Sample: 2 388
Included observations: 387
Convergence achieved after 10 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.003174	0.001643	-1.931920	0.0541
AR(35)	-0.456646	0.313620	-1.456049	0.1462
MA(35)	0.585331	0.295993	1.977519	0.0487
SIGMASQ	0.000846	4.84E-05	17.49502	0.0000

R-squared	0.022721	Mean dependent var	-0.003167
Adjusted R-squared	0.015066	S.D. dependent var	0.029461
S.E. of regression	0.029238	Akaike info criterion	-4.213582
Sum squared resid	0.327411	Schwarz criterion	-4.172668
Log likelihood	819.3281	Hannan-Quinn criter.	-4.197359
F-statistic	2.968104	Durbin-Watson stat	2.079116
Prob(F-statistic)	0.031871		

Inverted AR Roots	.97-.09i	.97+.09i	.94-.26i	.94+.26i
	.88+.42i	.88-.42i	.79+.57i	.79-.57i
	.68+.71i	.68-.71i	.54-.82i	.54+.82i
	.38+.90i	.38-.90i	.22-.95i	.22+.95i
	.04-.98i	.04+.98i	-.13+.97i	-.13-.97i
	-.30+.93i	-.30-.93i	-.46-.86i	-.46+.86i
	-.61+.76i	-.61-.76i	-.74-.64i	-.74+.64i
	-.84+.50i	-.84-.50i	-.92-.34i	-.92+.34i
	-.96+.17i	-.96-.17i	-.98	
Inverted MA Roots	.98-.09i	.98+.09i	.95-.26i	.95+.26i
	.89-.43i	.89+.43i	.80+.58i	.80-.58i
	.68-.71i	.68+.71i	.54+.82i	.54-.82i
	.39+.91i	.39-.91i	.22-.96i	.22+.96i
	.04-.98i	.04+.98i	-.13+.98i	-.13-.98i
	-.30+.94i	-.30-.94i	-.47-.87i	-.47+.87i
	-.61+.77i	-.61-.77i	-.74-.65i	-.74+.65i
	-.85+.51i	-.85-.51i	-.92-.35i	-.92+.35i
	-.97+.18i	-.97-.18i	-.98	

Findings:

We see here that only the MA variable is significant and we AR is not.

To move forward, we make sure to use a combination of both to see how it plays out, by keeping the MA (35) and replacing the AR (35) with AR (5).

ARIMA (5,1,35)

By estimating:

d(lg_nasdaq) ar(5) ma(35) c

we get:

Dependent Variable: D(LG_NASDAQ)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/01/24 Time: 18:50
Sample: 2 388
Included observations: 387
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.003190	0.001545	-2.065261	0.0396
AR(5)	-0.109781	0.052801	-2.079141	0.0383
MA(35)	0.139900	0.052114	2.684499	0.0076
SIGMASQ	0.000840	4.74E-05	17.73196	0.0000

R-squared	0.029552	Mean dependent var	-0.003167
Adjusted R-squared	0.021950	S.D. dependent var	0.029461
S.E. of regression	0.029136	Akaike info criterion	-4.221486
Sum squared resid	0.325122	Schwarz criterion	-4.180572
Log likelihood	820.8575	Hannan-Quinn criter.	-4.205263
F-statistic	3.887675	Durbin-Watson stat	2.100193
Prob(F-statistic)	0.009291		

Inverted AR Roots	.52-.38i	.52+.38i	-.20+.61i	-.20-.61i
	-.64			
Inverted MA Roots	.94+.08i	.94-.08i	.91+.25i	.91-.25i
	.85-.41i	.85+.41i	.76+.56i	.76-.56i
	.65+.68i	.65-.68i	.52+.79i	.52-.79i
	.37-.87i	.37+.87i	.21-.92i	.21+.92i
	.04-.94i	.04+.94i	-.13-.94i	-.13+.94i
	-.29-.90i	-.29+.90i	-.45+.83i	-.45-.83i
	-.59-.74i	-.59+.74i	-.71+.62i	-.71-.62i
	-.81+.48i	-.81-.48i	-.89-.33i	-.89+.33i
	-.93+.17i	-.93-.17i	-.95	

Findings:

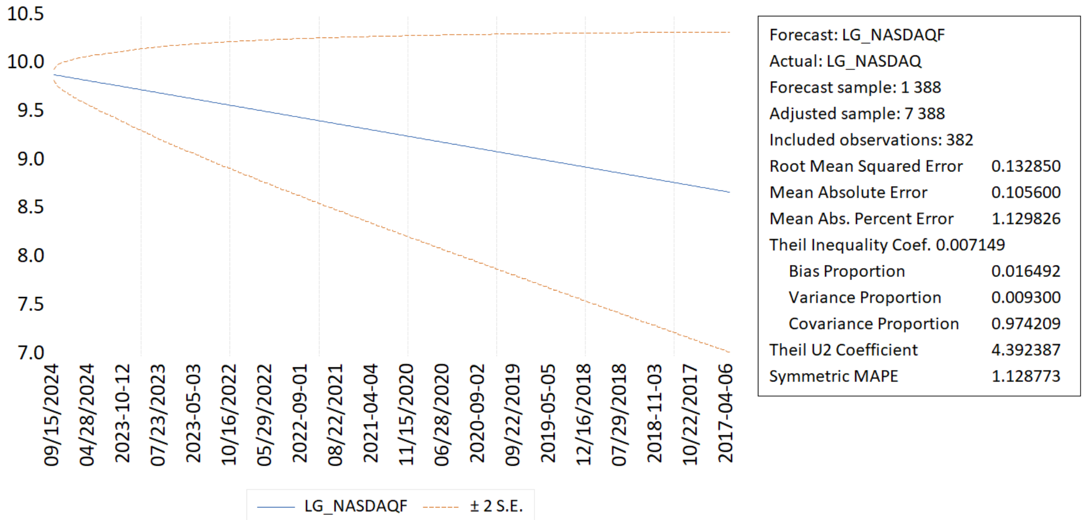
Here, we get both AR and MA to be significant and hence we can use this as our model for forecasting and moving on to the next steps.

FORECASTING:

Moving on to the next step, we will be forecasting another period for the model we selected.

We select the forecasting option from e-views and then:

We get these results:



The forecast gives us the following the insights:

GRAPH EXPLANATION

The ARMA/ARIMA model's predicted values (blue line) and 95% confidence interval (dashed orange lines) for the variable LG_NASDAQ are displayed in the chart. Using historical data up to the most recent observation, a forecast is created for the future (starting in September 2024).

The forecasted values stretch from the most recent observation, displaying the anticipated values for the future, while the blue line represents the actual data**.**

- **CONFIDENCE INTERVAL:**
  The range that the actual future numbers are expected to fall within is represented by the dashed lines surrounding the forecast, which are the 95% confidence interval (± 2 Standard Errors). As the confidence interval around the predicted line gets smaller, the forecast's dependability rises.

TABLE EXPLANATION

- RMSE, or root mean squared error
  0.132850 means that, on average, there is a 0.1328-unit difference between the predicted and actual values. The better the model fits the data, the lower the RMSE.

- Mean Absolute Error
  0.105600 indicates that there is an average absolute error of 0.1056 units between the predicted and actual values.
- Mean Absolute Percentage Error
  1.129826% indicates that, on average, there is a 1.13% discrepancy between the predicted and actual values. Good forecasting performance is indicated by this comparatively low percentage.

- Theil U2 Coefficient

A relative indicator of predicted performance is 4.392387. Although the model's performance might yet be enhanced, a result nearer 1 would signify perfect predictions.

CONCLUSION:

This study examined the effect of cryptocurrency markets on traditional stock market indices using a comprehensive multiple regression analysis of volatility and correlation. By using GBTC ETF data to reflect institutional engagement, the study provided deep insights into the growing relationship between these two financial sectors.

Significant findings revealed a robust correlation between cryptocurrency volatility and shifts in traditional stock indices, pointing to a shifting dynamic fueled by increased institutional participation. The results demonstrate how cryptocurrencies are altering investment strategies and market dynamics.

Future research could expand on these findings by utilizing larger datasets, looking into alternative institutional proxies, and analyzing how specific macroeconomic events affect the relationships between cryptocurrencies and the stock market. Examining geographic variations in these relationships may also yield a deeper understanding of regional market dynamics.

This study concludes by emphasizing the importance of understanding the integration of new financial products into current markets, which paves the way for more informed frameworks for investment and regulatory decisions.

REFERENCES:

Following were the links used to gather data for our study:

https://www.investing.com/indices/nq-100

https://www.investing.com/crypto/bitcoin/btc-usd-historical-data

https://www.investing.com/crypto/ethereum/eth-usd

https://www.investing.com/crypto/litecoin/ltc-usd

https://www.investing.com/crypto/dogecoin/doge-usd