

Bitcoin is dead: Long live Bitcoin?

Stephen Lee

2018

1 Introduction

In 2008, the Bitcoin white paper was published under the pseudonym Satoshi Nakamoto. The paper combined cryptography with a game theoretic incentive structure to provide secure peer-to-peer financial transactions without needing a trusted 3rd party. Since then, other projects have modified the Bitcoin rules to create new protocols to handle these transactions. Colloquially referred to as “cryptocurrencies”, these projects have captured the imagination of many. As of May 3, 2018, the three largest cryptocurrencies by market capitalization are Bitcoin, Ether, and Ripple. This time series analysis looks at those three coins, along with another project, Litecoin, to see if there is any exploitable market inefficiency. Notably, I test for Granger causality and co-integration to see if price movements of one coin can act as an early indicator for another coin. While these results are inconclusive at this point, the appearance of a large “bubble” from December 2017 to February 2018 is likely responsible. Preliminary work with the “post-bubble” data indicate significant unidirectional causality for each coin, with Bitcoin acting as the most exogenous series. Further, I fit TGARCH models to Bitcoin, Ether, and Ripple and show that investors respond with increased volatility after a negative price shock than for a positive price shock.

Further work is needed to explain the dramatic rise, and subsequent fall, in price during the two month span from December 2017 to February 2018. In particular, I plan to explore a sentiment analysis approach utilizing Twitter data during those periods.

2 Data

I gathered transaction-level data on Bitcoin, Ether, Ripple, and Litecoin from the European coin exchange Bitstamp. The raw data includes information about each transaction in the history of the exchange, including the unix-timestamp¹, trade price, and trade quantity. Several points to consider:

¹The unix-timestamp is given by the number of seconds that have passed since January 1, 1970 at 12:00:00 am GMT

- The different coins were offered at different points in time. For example, Bitcoin was the first coin offered on the exchange, while Ether was the most recent addition. Thus, for consistency of the analysis I used the earliest date such that trade information exists for each coin.
- Each transaction is marked to the second. In order to use a smooth time series without missing or repeated observations, I found the smallest period of time such that at least one transaction took place for each series. In this case, Litecoin was the limiting factor in calculating the time periods - there was a maximum gap of 5525 seconds between subsequent transactions. Thus, I made a dataset using this as the period length as well as a second dataset that used one-hour as the period length². Ultimately, for the analysis I used the hourly dataset.

The final series consisted of hourly periods ranging from Thursday, August 17, 2017 2:00:03 pm to Tuesday, April 24, 2018 4:00:03 am GMT. Figure 1 show the raw series, and table 1 shows the descriptive statistics.

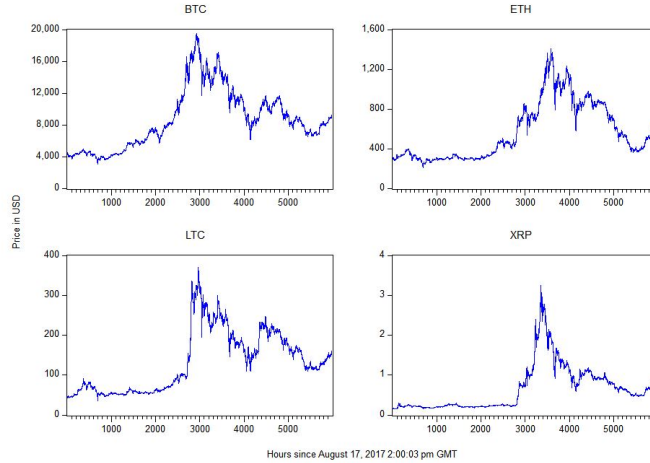


Figure 1: Time series of each coin in dataset.

| | Bitcoin | Ethereum | Ripple | Litecoin |
|--------------|------------|-----------|--------|----------|
| Maximum | \$19537.83 | \$1405.34 | \$3.25 | \$370.28 |
| Minimum | \$3009.53 | \$204.85 | \$0.15 | \$36.17 |
| Mean | \$8648.10 | \$566.08 | \$0.65 | \$133.05 |
| St. Dev | \$3749.17 | \$284.70 | \$0.57 | \$76.66 |
| Observations | 5991 | 5991 | 5991 | 5990 |

Table 1: Descriptive statistics. Prices in USD.

²Using one hour as the period length created a single missing observation in the Litecoin series.

Interestingly, the price for each coin are on different orders of magnitude and include an order of magnitude variation from the minimum to maximum. As the descriptive statistics show in table 1, Bitcoin has a maximum price of \$19,537.83 per BTC and minimum price of \$3009.53 per BTC. On the other end of the spectrum, Ripple has a maximum price of \$3.25 per XRP, and a minimum price of \$0.15 per XRP. Ethereum and Litecoin have similar relative variation. In order to better understand how the series moved in relative terms, I created a normalized series, $\{s^n\}$, by dividing the price at a given point in time by the series mean, μ_s . Mathematically this is to say:

$$s_t^n = \frac{s_t}{\mu_s} \quad (1)$$

Since this helps to normalize the series deviations, we can view this graphically below:

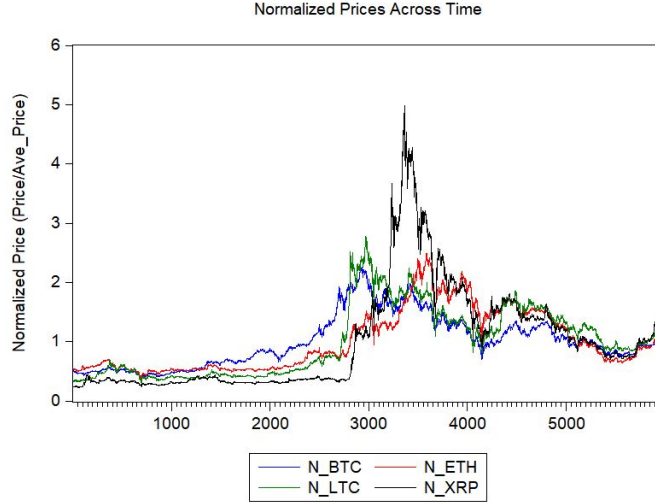


Figure 2: Overlay of each normalized series in dataset.

While the graph is a bit cluttered, what stands out is that Ripple experienced the largest percent increase at its peak, while Litecoin, Ether, and Bitcoin are all very similar. To get a clearer idea of how the series respond to each other, I restrict the graph to only the two largest coins by market cap - Bitcoin and Ether:

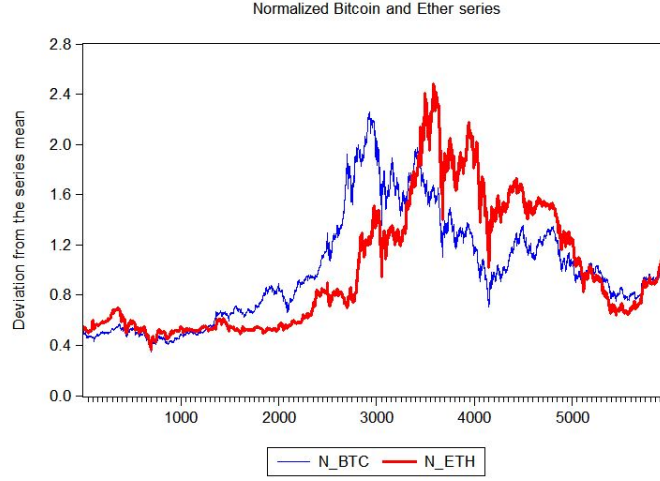


Figure 3: Overlay of Bitcoin and Ether's normalized series.

Here we see that, in normalized terms, Bitcoin's largest surge and drop seems to occur about a month (i.e. approximately 750 hours) before Ether. Interesting, however, the last month or so in the dataset show much tighter price movements, suggesting a potential change in investor perspective.

3 Analysis and Results

Since the series levels are clearly not stationary, I calculated the series of first differences ($\Delta s_t = s_t - s_{t-1}$) and performed the Augmented Dicky-Fuller test for unit roots. In each differenced series, I am able to reject the null hypothesis that a unit root exists and accept that the level series are all integrated of degree one (i.e. $I(1)$). Visually we can see these differenced series below:

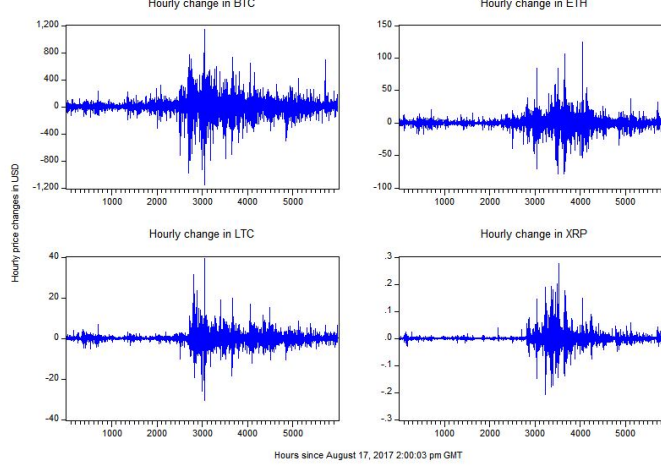


Figure 4: Time series of each differenced coin series in dataset.

3.1 Individual Series Analysis

While the differenced series do seem to have a time constant mean, the variance is definitely not constant across time. With that in mind, I tested various specifications using ARCH and GARCH errors to see how the data moves. For each series, the ACF and PACF suggest significant autocorrelation for two-periods (i.e. two hours). However, after finding the best-fit ARCH and GARCH, the series of squared residuals still indicates significant autocorrelation and partial autocorrelation. Thus, I finally settled on an AR model with TGARCH errors, suggesting that traders are affected by whether the prices are rising or falling. The results for Bitcoin, Ether, and Ripple are presented below.

3.1.1 Bitcoin

The best fitting model for the Bitcoin data is an AR(3) with TGARCH error terms. After fitting I find the following model³:

$$\Delta BTC_t = 0.20\Delta BTC_{t-1} - 0.12\Delta BTC_{t-2} + 0.03\Delta BTC_{t-3} + \epsilon_t \quad (2)$$

(15.34) (-8.25) (2.36)

With an error process given by:

$$E_{t-1}(\epsilon_t^2) = 23.25 + 0.068\epsilon_{t-1}^2 + 0.02D_{t-1}\epsilon_{t-1}^2 + 0.926h_{t-1} \quad (3)$$

(14.23) (23.13) (3.82) (530.00)

³Z-statistics in parentheses.

Next we look at the ACF and PACF of the residual and squared residual series from this estimation. We find that they mirror a white noise process.

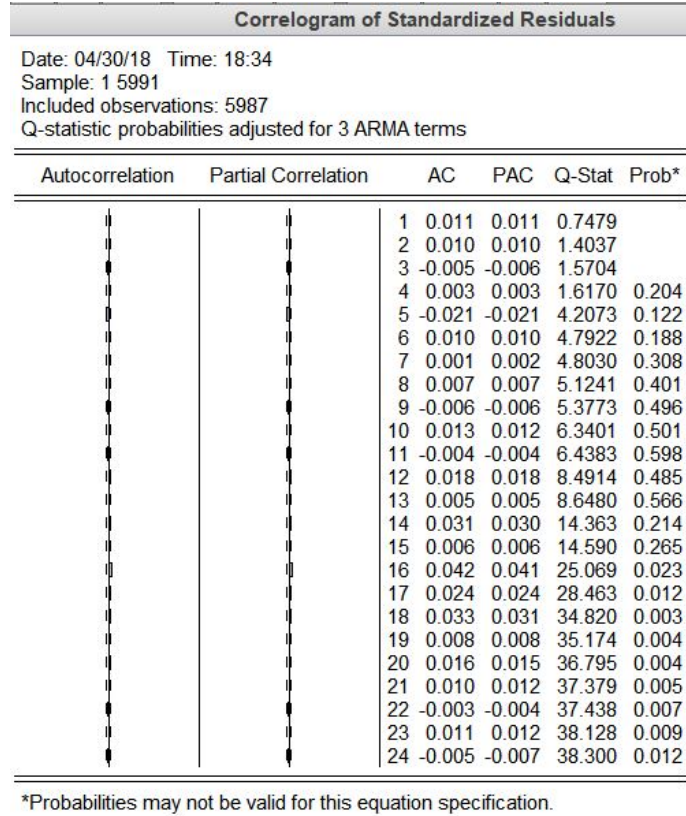


Figure 5: ACF and PACF of the residuals from the above AR(2) model with TGARCH errors on the differenced Bitcoin series.

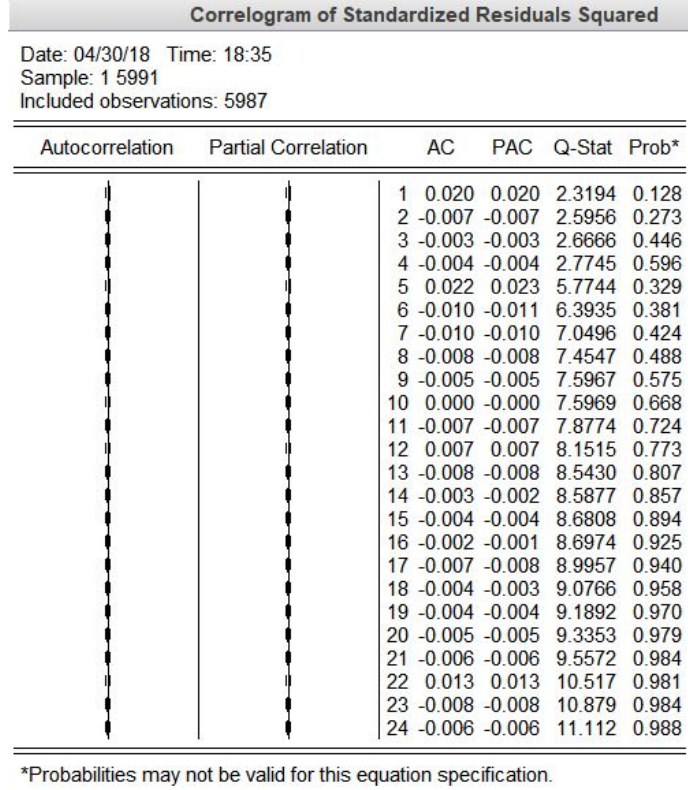


Figure 6: ACF and PACF of the squared residuals from AR(2) model with TGARCH errors on the differenced Bitcoin series.

3.1.2 Ether

The best fitting model for the Ether data is an AR(2) with TGARCH error terms. After fitting I find the following model⁴:

$$\Delta ETH_t = 0.15\Delta ETH_{t-1} - 0.10\Delta ETH_{t-2} + \epsilon_t \quad (4)$$

(12.02) (-6.71)

With an error process given by:

$$E_{t-1}(\epsilon_t^2) = 0.12 + 0.11\epsilon_{t-1}^2 + 0.02D_{t-1}\epsilon_{t-1}^2 + 0.89h_{t-1} \quad (5)$$

(12.72) (21.36) (2.84) (270.10)

We again look at the ACF and PACF of the residual and squared residual series from this estimation and confirm that they mirror a white noise process.

⁴Z-statistics in parentheses.

Date: 05/01/18 Time: 10:03
Sample: 1 5991
Included observations: 5988
Q-statistic probabilities adjusted for 2 ARMA terms

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob* |
|-----------------|---------------------|----|--------|--------|--------|-------|
| | | 1 | 0.038 | 0.038 | 8.7696 | |
| | | 2 | 0.021 | 0.020 | 11.537 | |
| | | 3 | -0.000 | -0.002 | 11.537 | 0.001 |
| | | 4 | 0.022 | 0.021 | 14.369 | 0.001 |
| | | 5 | -0.016 | -0.018 | 15.898 | 0.001 |
| | | 6 | 0.012 | 0.013 | 16.811 | 0.002 |
| | | 7 | 0.014 | 0.014 | 18.055 | 0.003 |
| | | 8 | -0.004 | -0.006 | 18.137 | 0.006 |
| | | 9 | -0.002 | -0.002 | 18.166 | 0.011 |
| | | 10 | 0.023 | 0.022 | 21.253 | 0.007 |
| | | 11 | 0.011 | 0.010 | 22.024 | 0.009 |
| | | 12 | 0.007 | 0.006 | 22.303 | 0.014 |
| | | 13 | 0.009 | 0.008 | 22.775 | 0.019 |
| | | 14 | 0.032 | 0.030 | 28.964 | 0.004 |
| | | 15 | 0.014 | 0.012 | 30.221 | 0.004 |
| | | 16 | 0.034 | 0.032 | 37.226 | 0.001 |
| | | 17 | 0.027 | 0.024 | 41.743 | 0.000 |
| | | 18 | 0.015 | 0.011 | 43.178 | 0.000 |
| | | 19 | 0.015 | 0.014 | 44.616 | 0.000 |
| | | 20 | 0.048 | 0.045 | 58.514 | 0.000 |
| | | 21 | 0.015 | 0.010 | 59.846 | 0.000 |
| | | 22 | -0.006 | -0.010 | 60.098 | 0.000 |
| | | 23 | -0.001 | -0.002 | 60.104 | 0.000 |
| | | 24 | -0.029 | -0.032 | 64.999 | 0.000 |

*Probabilities may not be valid for this equation specification.

Figure 7: ACF and PACF of the residuals from the above AR(2) model with TGARCH errors on the differenced Ether series.

Date: 05/01/18 Time: 10:03
Sample: 1 5991
Included observations: 5988

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob* |
|-----------------|---------------------|-----------|--------|--------|-------|
| | | 1 0.040 | 0.040 | 9.5561 | 0.002 |
| | | 2 -0.009 | -0.010 | 10.009 | 0.007 |
| | | 3 -0.014 | -0.013 | 11.128 | 0.011 |
| | | 4 -0.021 | -0.020 | 13.828 | 0.008 |
| | | 5 0.005 | 0.006 | 13.960 | 0.016 |
| | | 6 -0.012 | -0.013 | 14.756 | 0.022 |
| | | 7 -0.014 | -0.013 | 15.882 | 0.026 |
| | | 8 -0.003 | -0.002 | 15.938 | 0.043 |
| | | 9 -0.010 | -0.010 | 16.536 | 0.057 |
| | | 10 -0.002 | -0.002 | 16.569 | 0.084 |
| | | 11 0.001 | 0.000 | 16.571 | 0.121 |
| | | 12 0.003 | 0.003 | 16.632 | 0.164 |
| | | 13 -0.019 | -0.020 | 18.836 | 0.128 |
| | | 14 0.002 | 0.004 | 18.868 | 0.170 |
| | | 15 0.005 | 0.004 | 19.025 | 0.213 |
| | | 16 0.004 | 0.003 | 19.102 | 0.263 |
| | | 17 -0.017 | -0.018 | 20.788 | 0.236 |
| | | 18 -0.014 | -0.013 | 22.039 | 0.230 |
| | | 19 0.012 | 0.013 | 22.947 | 0.240 |
| | | 20 -0.009 | -0.011 | 23.453 | 0.267 |
| | | 21 0.004 | 0.004 | 23.558 | 0.315 |
| | | 22 -0.003 | -0.004 | 23.630 | 0.367 |
| | | 23 -0.000 | 0.000 | 23.631 | 0.424 |
| | | 24 0.001 | 0.000 | 23.641 | 0.482 |

*Probabilities may not be valid for this equation specification.

Figure 8: ACF and PACF of the squared residuals from AR(2) model with TGARCH errors on the differenced Ether series.

3.1.3 Ripple

The best fitting model for the Ether data is an AR(2) with TGARCH error terms. After fitting I find the following model⁵:

$$\Delta XRP_t = 0.13\Delta XRP_{t-1} - 0.10\Delta XRP_{t-2} + \epsilon_t \quad (6)$$

(9.67) (-6.68)

With an error process given by:

$$E_{t-1}(\epsilon_t^2) = 0.000000312 + 0.14\epsilon_{t-1}^2 + 0.014D_{t-1}\epsilon_{t-1}^2 + 0.87h_{t-1} \quad (7)$$

(35.14) (26.60) (1.97) (261.57)

We again look at the ACF and PACF of the residual and squared residual series from this estimation and confirm that they mirror a white noise process.

⁵Z-statistics in parentheses.

Date: 05/01/18 Time: 10:10

Sample: 1 5991

Included observations: 5988

Q-statistic probabilities adjusted for 2 ARMA terms

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob* |
|-----------------|---------------------|-----------|--------|--------|-------|
| | | 1 0.026 | 0.026 | 4.1606 | |
| | | 2 0.026 | 0.025 | 8.2374 | |
| | | 3 0.009 | 0.008 | 8.7675 | 0.003 |
| | | 4 0.015 | 0.014 | 10.143 | 0.006 |
| | | 5 -0.001 | -0.002 | 10.149 | 0.017 |
| | | 6 0.017 | 0.016 | 11.906 | 0.018 |
| | | 7 0.013 | 0.012 | 12.887 | 0.024 |
| | | 8 -0.011 | -0.013 | 13.671 | 0.034 |
| | | 9 0.020 | 0.020 | 16.104 | 0.024 |
| | | 10 0.019 | 0.018 | 18.292 | 0.019 |
| | | 11 -0.006 | -0.008 | 18.496 | 0.030 |
| | | 12 -0.002 | -0.003 | 18.519 | 0.047 |
| | | 13 0.017 | 0.017 | 20.349 | 0.041 |
| | | 14 0.046 | 0.046 | 33.197 | 0.001 |
| | | 15 0.022 | 0.019 | 36.002 | 0.001 |
| | | 16 0.036 | 0.031 | 43.774 | 0.000 |
| | | 17 0.028 | 0.024 | 48.405 | 0.000 |
| | | 18 0.025 | 0.021 | 52.226 | 0.000 |
| | | 19 0.015 | 0.010 | 53.558 | 0.000 |
| | | 20 0.035 | 0.030 | 60.754 | 0.000 |
| | | 21 0.009 | 0.005 | 61.238 | 0.000 |
| | | 22 0.023 | 0.020 | 64.503 | 0.000 |
| | | 23 0.027 | 0.022 | 68.859 | 0.000 |
| | | 24 -0.027 | -0.032 | 73.179 | 0.000 |

*Probabilities may not be valid for this equation specification.

Figure 9: ACF and PACF of the residuals from the above AR(2) model with TGARCH errors on the differenced Ripple series.

Date: 05/01/18 Time: 10:10
Sample: 1 5991
Included observations: 5988

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob* |
|-----------------|---------------------|-----------|--------|--------|-------|
| | | 1 0.001 | 0.001 | 0.0042 | 0.948 |
| | | 2 -0.002 | -0.002 | 0.0281 | 0.986 |
| | | 3 -0.001 | -0.001 | 0.0376 | 0.998 |
| | | 4 -0.005 | -0.005 | 0.1760 | 0.996 |
| | | 5 -0.004 | -0.004 | 0.2817 | 0.998 |
| | | 6 -0.005 | -0.005 | 0.4317 | 0.999 |
| | | 7 -0.002 | -0.002 | 0.4503 | 1.000 |
| | | 8 -0.003 | -0.003 | 0.5028 | 1.000 |
| | | 9 -0.001 | -0.001 | 0.5115 | 1.000 |
| | | 10 -0.002 | -0.002 | 0.5301 | 1.000 |
| | | 11 -0.002 | -0.002 | 0.5627 | 1.000 |
| | | 12 0.001 | 0.001 | 0.5687 | 1.000 |
| | | 13 0.001 | 0.001 | 0.5708 | 1.000 |
| | | 14 -0.002 | -0.002 | 0.5872 | 1.000 |
| | | 15 -0.001 | -0.002 | 0.6007 | 1.000 |
| | | 16 -0.001 | -0.001 | 0.6073 | 1.000 |
| | | 17 -0.002 | -0.002 | 0.6356 | 1.000 |
| | | 18 -0.002 | -0.002 | 0.6697 | 1.000 |
| | | 19 -0.003 | -0.003 | 0.7376 | 1.000 |
| | | 20 -0.001 | -0.001 | 0.7433 | 1.000 |
| | | 21 0.001 | 0.000 | 0.7452 | 1.000 |
| | | 22 0.004 | 0.004 | 0.8374 | 1.000 |
| | | 23 -0.001 | -0.001 | 0.8489 | 1.000 |
| | | 24 -0.002 | -0.002 | 0.8826 | 1.000 |

*Probabilities may not be valid for this equation specification.

Figure 10: ACF and PACF of the squared residuals from AR(2) model with TGARCH errors on the differenced Ripple series.

3.2 Cointegrated Series Analysis

After finding a model for Bitcoin, Ether, and Ripple that only rely on their own history, next, I test for Granger causality in the level-series⁶ to determine if information of the other coins can improve estimates for a given coin. I find significant bi-directional causality in most of the major coin pairs. The only notable unidirectional exception is Bitcoin causing Litecoin. We see the results below:

⁶The lag length for the level series was determined by optimizing the AIC in a standard VAR.

Pairwise Granger Causality Tests
Date: 04/30/18 Time: 19:08
Sample: 1 5991
Lags: 2

| Null Hypothesis: | Obs | F-Statistic | Prob. |
|------------------------------------|------|-------------|--------|
| D_ETH does not Granger Cause D_BTC | 5988 | 10.3061 | 3.E-05 |
| D_BTC does not Granger Cause D_ETH | | 6.76282 | 0.0012 |
| D_LTC does not Granger Cause D_BTC | 5984 | 1.94498 | 0.1431 |
| D_BTC does not Granger Cause D_LTC | | 10.2860 | 3.E-05 |
| D_XRP does not Granger Cause D_BTC | 5988 | 0.90736 | 0.4036 |
| D_BTC does not Granger Cause D_XRP | | 0.03955 | 0.9612 |
| D_LTC does not Granger Cause D_ETH | 5984 | 6.99999 | 0.0009 |
| D_ETH does not Granger Cause D_LTC | | 11.4547 | 1.E-05 |
| D_XRP does not Granger Cause D_ETH | 5988 | 3.03104 | 0.0483 |
| D_ETH does not Granger Cause D_XRP | | 4.84651 | 0.0079 |
| D_XRP does not Granger Cause D_LTC | 5984 | 1.40093 | 0.2464 |
| D_LTC does not Granger Cause D_XRP | | 0.71664 | 0.4884 |

Figure 11: Granger causality tests using the stationary differenced series.

Note, however, that these results are for the entire series. Restricting the sample to “post-bubble” data results in significant unidirectional Granger causality for each coin, with Bitcoin acting as the first mover.

Next we look for co-integrating equations using the Johansen system test with two lags as described above. The results from both the trace and maximum eigenvalue statistics indicate three co-integrating equations at the 5% level. To further motivate and understand the possibility of co-integration, we can look at the relative prices of each series. Specifically, I calculate the price of a given coin as denominated in another coin. For example, the Bitcoin and Ether pair would tell us how many Bitcoin it costs to buy a single Ether across our sample. The idea is to check if there exists a relatively consistent relationship between the coins. A graph of each pair is presented below:

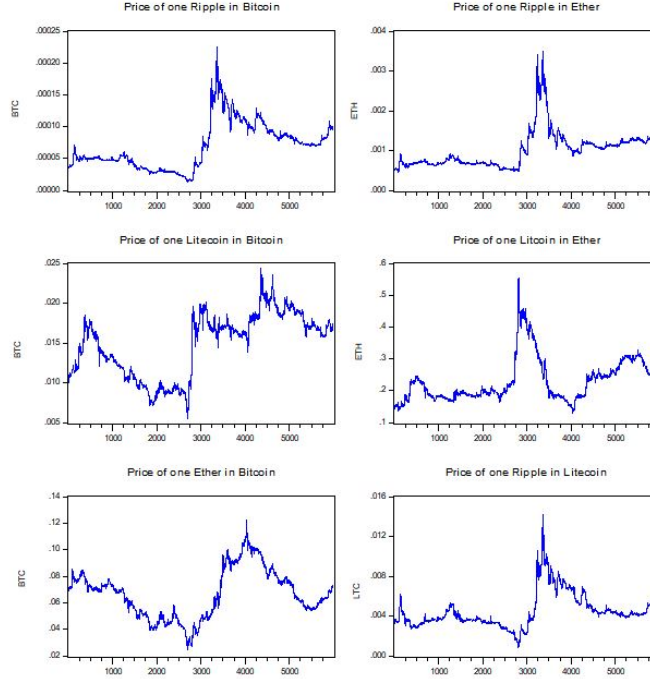


Figure 12: The relative price for each coin pair.

These graphs, while undoubtedly noisy, tell the story of different relative price jumps, but also of a potential correction to a long-run relationship. Specifically, we see that each coin was “cheap” to buy with Bitcoin up until about the 2750th hour in the dataset (this corresponds to approximately December 10, 2017). At this point, Litecoin, Ripple, and Ether all began to increase in price (and in that respective order) relative to Bitcoin. The order of the price increases are exemplified in the graphs on the right-hand-side, which show that Ripple and Litecoin both became relatively “expensive” to buy with Ether, before returning to the previous level. This is to say that first, the price of Litecoin increased away from the long-run equilibrium. This was followed by a more dramatic increase in the price of Ripple, and finally a slower, but substantial, increase in the price of Ether. Finally we see that, through some mechanism, the prices correct back to what appears to be an average ratio between the pairs. Thus, while volatile indeed, there does appear to be the possibility of persistent long-run relationships.

With that said, I run the vector error correction model with three co-integrating equations, and no trend or intercept, as suggested by the Johansen system test. The results are presented in Figure 13 below:

Vector Error Correction Estimates
Date: 05/01/18 Time: 20:57
Sample (adjusted): 5 5991
Included observations: 5982 after adjustments
Standard errors in () & t-statistics in []

| Cointegrating Eq: | CointEq1 | CointEq2 | CointEq3 |
|-------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| ETH(-1) | 1.000000 | 0.000000 | 0.000000 |
| XRP(-1) | 0.000000 | 1.000000 | 0.000000 |
| LTC(-1) | 0.000000 | 0.000000 | 1.000000 |
| BTC(-1) | -0.068185 (0.00479) [-14.2310] | -8.21E-05 (8.7E-06) [-9.42743] | -0.015995 (0.00083) [-19.1596] |

| Error Correction: | D(ETH) | D(XRP) | D(LTC) | D(BTC) |
|-------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| CointEq1 | -0.003404 (0.00079) [-4.32193] | -6.91E-06 (1.8E-06) [-3.91229] | -0.000360 (0.00023) [-1.54972] | -0.008474 (0.01028) [-0.82448] |
| CointEq2 | 1.400112 (0.39368) [3.55650] | -0.000194 (0.00088) [-0.21929] | -0.007361 (0.11607) [-0.06342] | 1.825967 (5.13791) [0.35539] |
| CointEq3 | -0.002783 (0.00396) [-0.70263] | 2.22E-05 (8.9E-06) [2.49730] | -0.002850 (0.00117) [-2.44053] | -0.122030 (0.05170) [-2.36041] |

Figure 13: Vector Error Correction model with three co-integrating equations and no time-trend or intercept.

While this table will take me some time to understand and explain in detail, the parts that stand out is the non-zero speed of adjustment for both Litecoin and Bitcoin in cointegrating equation 3. Besides that, nothing stands out as statistically and economically significant. Based on the graphs above, the next steps are to explore longer lag lengths, or by using a restricted dataset.

4 Conclusion

A preliminary analysis of Bitcoin, Ether, Ripple, and Litecoin suggest the following:

- **An apparent bubble began for each coin starting around December 2017 until early February 2018.** Figure 2 shows the ratio of each coin's price to it's series mean. We see that Ripple (XRP) experienced the most dramatic price increase during the bubble with prices of nearly 5 times the average price. In contrast, Bitcoin, Ether, and Litecoin all saw a maximum price of about 2.5 times their mean price during this period.
- **The maximum price of each coin occurred at different points in time, possibly indicating causality with a lag length of a couple weeks during the bubble period.** Specifically, Bitcoin was the first to

reach it's maximum, followed closely by Litecoin, then Ripple, and finally Ether. This tells us that Bitcoin is likely the most exogenous coin, with the others following it's movement at various lag lengths. If, however, we restrict the dataset to the "post-bubble" period, we find significant uni-directional granger causality indicating that Bitcoin is the first mover with all other coin's following it's movement 2 hours later.

- **We cannot reject the possibility of a long-run relationship between coin pairs.** Based on the exchange rates in Figure 12, there was a dramatic change in the exchange rates from early December 2017 to early February 2018. However, after this apparent bubble, the exchange rates have all returned to levels that are consistent with "pre-bubble" rates. More analysis and data are needed to confidently investigate co-integration.

4.0.1 Further Research

The next steps in this analysis are to compare various subsamples of this data set. Specifically, I will test whether the "pre-bubble" and "post-bubble" series have similar behavior, or if the "bubble mainia" changed investor perspective. Further research is also needed to explain the "bubble" that started in late-November/early-December 2017 and ended in late-January/early-February 2018. For that, I intend to perform a sentiment analysis using Twitter data to see if quantity of positive and negative tweets correlate with large price changes.