Reversion and Location Trends in the Bitcoin Market

Dominic Lewinski [*], Yiming Yang[†], Zhongxue Chen [‡], Guoyi Zhang[§]

Abstract

The cryptocurrency market is different from traditional markets due to its unique property, which allows global trading around the clock. It is of interest to investigate if some traditional stock market phenomena still exist in the cryptocurrency markets. In this research, we studied the application of the 75 percent reversion rule in cryptocurrency markets. Using local linear regression, we identified active markets at certain time and location, and examined government regulations and news' influence on the cryptocurrency markets.

[*]Dominic Lewinski, Graduate Student, Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM, 87131-0001(*d*lewinsk@unm.edu)

[†]Yiming Yang, Graduate Student, Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM, 87131-0001, (*y*ang.yihming@gmail.com)

[‡]Zhongxue Chen: Associate Professor of Biostatistics, Department of Epidemiology and Biostatistics School of Public Health, Indiana University, Bloomington, IN, 47405, (*zc3@indiana.edu*)

[§]Guoyi Zhang, Associate Professor, Department of Mathematics and Statistics, University of New Mexico, Albuquerque,NM, 87131-0001, (*g*zhang123@gmail.com)

Keywords: Bitcoin, cryptocurrency market, 75% reversion rule, local linear regression, geographic location identification.

1 Introduction

Cryptocurrencies are digital currencies built on the blockchain (Economist, 2015) technology and peer-to-peer exchanges, which allows verification of payments in the absence of a centralized custodian. As the first cryptocurrency, Bitcoin (Brito & Castillo, 2013; Tschorsch & Scheuermann, 2016) was introduced by Nakamoto (2008) and was created in 2009. Bitcoins have been traded around the clock 24 hours every day in more than 100 exchanges worldwide starting from 2010. At the end of February 2018, the entire cryptocurrency market is about $40 billion in value.

As an asset class, cryptocurrency has received significant attention in recent years. We now review some of the recent research work. Yermack (2015) pointed out several problems of the cryptocurrency market, such as high volatilities, zero correlation with widely used currencies, and daily hacking and theft risks. All of these make Bitcoin appear to behave more like a speculative investment than a currency. Bianchi (2018) empirically investigated some of the key features of cryptocurrency returns and volatilities such as their relationship with traditional asset classes, as well as the main driving factors behind market activities. Liu and Tsyvinski (2018) showed that the risk return trade-off of cryptocurrencies (Bitcoin, Ripple, and Ethereum) is distinct from those of stocks, currencies, and precious metals. Makarov and Schoar (2018) studied the efficiency and price formation of bitcoin and other cryptocurrency markets by trading and arbitrage. Borri (2018) investigated conditional Tail-Risk in Cryptocurrency Markets. In this research, we will investigate some other properties in the Cryptocurrency Markets.

The 75 percent reversion rule (Pole, 2007) has been used in stock markets for a long time. Assume that the spread of price (defined as the difference between the highest and lowest price within a time interval) is independent and identically distributed (iid) Gaussian and the market is relatively stable. The 75 percent reversion rule says that the following events will happen 75 percent of the time: (1) the next day's spread is less than today's spread given that today's spread is greater than median of the spread distribution; (2) the next day's spread is greater than today's spread given that today's spread is less than the median of the spread distribution. Section 2 gives a detailed review of the 75 percent reversion theorem and proof.

The cryptocurrency market is different from the traditional markets because it allows global trading around the clock. Due to this unique property, it is of great interest to investigate if some traditional stock market phenomena still exist in the cryptocurrency markets and how global trading around the clock affect these phenomena. For example, we believe that the 75% reversion rule exist outside of the stock markets. Does it hold in cryptocurrency market? How does traditional market open/close time affect the cryptocurrency markets as different regions become active at different time? How does government regulations and news influence the cryptocurrency markets? This research is motivated by these interesting questions and is intended to answer them.

This paper is organized as follows: in Section 2, we give a background review of the 75% reversion rule and local polynomial regression; in Section 3, we describe the data set, methodology and data analysis results; and in Section 4, we give conclusions and discussions.

2 Background

In this section, we review some background knowledge related to our research. First, we review the 75% reversion rule in stock price market. Next, we review local polynomial regression estimator that we use in data analysis for pattern recognition. Last, we review an optimal bandwidth selection method: cross validation (CV) that is used in polynomial regression estimation.

2.1 The 75% reversion rule

Assume that data is independent and identically distributed and market is relatively stable. In practice, stock price spreads approximately meet these conditions. The 75 percent reversion rule says that the probability of the next day's spread less/greater than today's spread given that today's spread is greater/less than the median of spread distribution is 75%. In mathematical language, it can be stated as follows:
Let $d_1, d_2, \cdots$ be a sequence of iid continuous random variables with median $\tilde{d}$. The 75 percent reversion rule says that

$$P(d_{t+1} > d_t | d_t < \tilde{d}) = \frac{3}{4}, \tag{1}$$

and

$$P(d_{t+1} < d_t | d_t > \tilde{d}) = \frac{3}{4}. \tag{2}$$

Below is a short proof of the 75 percent reversion rule.

Proof. Let $X = d_t$, $Y = d_{t+1}$, and let $F(x)$ be the cumulative distribution function of the random variable d_i.

$$\begin{aligned}
P(d_{t+1} < d_t | d_t > \tilde{d}) &= P(Y < X | X > \tilde{d}) \\
&= \frac{P(Y < X \text{ and } X > \tilde{d})}{P(X > \tilde{d})} \\
&= 2P(Y < X \text{ and } X > \tilde{d}), \text{ since } P(X > \tilde{d}) = 1/2 \\
&= 2\int_{\tilde{d}}^{\infty} \int_{-\infty}^{x} f_{X,Y}(x,y)dxdy \\
&= 2\int_{\tilde{d}}^{\infty} \int_{-\infty}^{x} f(x)f(y)dydx, \text{ by independence of } f_{X,Y}(x,y) \\
&= 2\int_{\tilde{d}}^{\infty} f(x) \left(\int_{-\infty}^{x} f(y) \right) dx \\
&= 2\int_{\tilde{d}}^{\infty} f(x)F(x)dx \\
&= 2\int_{\tilde{d}}^{\infty} F(x)dF(x) \\
&= (F(x))^2|_{\tilde{d}}^{\infty} = 1 - F(\tilde{d})^2 \\
&= 1 - \left(\frac{1}{2}\right)^2 = \frac{3}{4}
\end{aligned}$$

Similarly, we can prove that $P(d_{t+1} > d_t | d_t < \tilde{d}) = \frac{3}{4}$. □

2.2 Local Polynomial Regression Estimator

Consider a general nonparametric regression model

$$y_i = u(t_i) + \varepsilon_i \ , \qquad i = 1, 2, \ldots, n, \tag{3}$$

where $\{\varepsilon_i\}$ is a sequence of uncorrelated and identically distributed random variables with $E(\varepsilon_i) = 0$ and $E(\varepsilon_i^2) = 1$; $u(\cdot)$ is an unknown smooth regression curve; $\{(t_i, y_i)\}$ is a sequence of observations. Estimators and properties of local polynomial regression have been studied by many researchers, such as Stone (1977, 1980, 1982), Cleveland (1979), Fan (1992, 1993), Fan and Gijbels (1992) and Ruppert and Wand (1994).

By using Taylor's expansion, the regression function $u(\cdot)$ can be approximated by

$$u(z) \approx \sum_{j=0}^{p} \frac{u^{(j)}(t)}{j!}(z-t)^j = \sum_{j=0}^{p} \beta_j(t)(z-t)^j \tag{4}$$

for z in a neighborhood of t, where $\beta_j(t) = u^j(t)/j!$. Equation (4) models $u(z)$ locally by a simple polynomial model. The $\beta_j(t)$'s are chosen to minimize the weighted least square error

$$\sum_{i=1}^{n} \left\{ y_i - \sum_{j=0}^{p} \beta_j(t)(t_i - t)^j \right\}^2 K\left\{ \frac{(t-t_i)}{\lambda} \right\}, \tag{5}$$

where $K(\cdot)$ is the epanechnikov kernel and λ is the bandwidth. The weighted least square estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{Y},$$

where $\mathbf{Y} = (y_1, y_2, \cdots, y_n)^T$, $\mathbf{X}$ is an $n \times (p+1)$ matrix with ith row $[1, (t_i - t), \cdots, (t_i - t)^p]$ and $\mathbf{W} = \mathrm{diag}\left\{ \frac{1}{h} K\left(\frac{t_i - t}{\lambda} \right) \right\}$.

N-W kernel estimator is a special case of the local polynomial kernel estimator in case of fitting degree zero polynomials i.e., constant. Local linear estimator is another special case of polynomial estimator in the case of fitting degree 1 polynomials $\beta_0+\beta_1 t$. Fan (1992) proved that local linear estimator has the same variance as the N-W kernel estimator but with a smaller bias. In our research, local linear estimators are applied to analyze data.

2.3 Optimal Bandwidth Selection: Cross Validation (CV)

The remaining problem is to choose the smoothing parameter λ in Equation (5), which controls the trade-off between smoothness and goodness-of-fit. A bigger λ will give a smooth estimator with small variance and usually large bias. On the other hand, if λ is too small, the fitted curve is choppy with large variance and usually small bias. Two commonly used techniques for smoothing parameter selection are the cross validation (CV) method and the generalized cross validation (GCV) method (Craven & Wahba, 1979) . In this research, we use CV method to select the optimal bandwidth.

The vector of fitted values from (3) can be written as

$$\mu_\lambda = \mathbf{S}_\lambda \mathbf{Y}, \tag{6}$$

where $\mathbf{S}_\lambda = \mathbf{X}(\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}$.

The CV criterion is defined as

$$\begin{aligned} CV(\lambda) &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu_{\lambda(i)}(t_i))^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \mu_\lambda(t_i)}{1 - s_{ii}}\right)^2 \end{aligned}$$

where $\mu_{\lambda(i)}$ is the estimate of μ_λ computed without using the ith observation (t_i, y_i), and s_{ii} is the ith diagonal element of $\mathbf{S}_\lambda$. The idea behind the CV criterion is that the ith observation is treated like an additional observation for prediction and $CV(\lambda)$ measures the quality of prediction. The optimal λ_{opt} is the best smoothing parameter that minimizes the prediction error via $CV(\lambda)$. It can be obtained by a searching algorithm with different λs. More details can be found from (Craven & Wahba, 1979).

3 Data Analysis

In this section, we first describe the data set. Next we examine the application of the 75% reversion rule in cryptocurrency market. Last, we use a local linear estimator to identify the active geographic locations/markets, and to investigate the influence of government regulation and important news on the cryptocurrency market.

3.1 Dataset

The data source is from Global Digital Assets Exchange (GDAX). We use Coordinated Universal Time (UTC) from 0:00 to 23:59 to define the 24 hours clock. We wrote a program in Python to collect data on Bitcoin price from February 18th, 2018 at

UTC 0:00 to March 20th, 2018 at UTC 0:00, which constitutes a total of 43199 observations (indexed from 1 to 43199) recorded every minute. Data is installed in Structured Query Language (SQL) local server through PHPMyAdmin. All data analysis is done in R. At the time of writing this project, one minute data is the finest price data available from GDAX. This dataset has one missing entry of one minute in the data table coming from an error in the database GDAX, which is negligible given the big dataset. The dataset contains information on trading volume and low, high, open and close prices of Bitcoin on GDAX for each minute.

Due to the fact that GDAX returns one minute data, high and low price often have same numbers, and volumes are often close to zero. A three-minute window is created with the "new high" and "new low" as the maximum and minimum values taken from the three-minute time frame, and "new volume" be the total volume over the three-minute window. Table 1 illustrate the transformation from one minute data to three-minute data. For the three-minute (UTC 0:00 to UTC 0:02) data, the new high is $10,600, new low is $10,050, and the total volume over the three minutes is 39,900 bitcoins exchanged.

Table 1: Illustration of transforming one minute data to three minute time frame

High	Low	UTC	Volume		New High	New Low	New UTC	New Volume
10,500	10,250	0:00	10					
10,600	10,100	0:01	16.5	=	10,600	10,050	0:02	39,900
10,400	10,050	0:02	13.4					

In our research, we use local linear regression to smooth out the noise in order to reveal the hidden patterns to help understand the market. Our data is from 24 hours non-stop continuous market, which supports the assumption that the underlying function "$u(\cdot)$" is a smooth curve. The uncorrelated, identically distributed assumption of the error terms are commonly assumed to be valid in financial market, and could be diagnosed by using residual plots.

3.2 Evaluating the 75% reversion rule in cryptocurrency market

In this section, we want to investigate if the 75% reversion rule exists in the cryptocurrency market. The following procedure is used to evaluate result in Equation (1). The spread of data is calculated by "high" minus "low" within the specified time frame. Median of the spread distribution at time t is approximated by the median of the spreads from $t - m$ to $t - 1$, where m is set up to be some certain number such as 20. If the current spread is less than the median, and if next spread is greater than the current spread, we set count as 1, otherwise 0. The sum of the counts is then divided by the total number of events such that the spread is less than the median, which gives a percent of how often result in Equation (1) occurs. A similar procedure is done for evaluating result in Equation (2). These percentages are expected to be close to 75%.

Table 2 reports empirical results of the above procedures under various time frames

Table 2: Evaluating 75% reversion rule by using various time frames with $m = 20$ [a]

Minute	1	2	3	5	10	20	30	60
Percentage 1	56.32	62.08	64.61	66.36	67.88	68.47	67.76	68.22
Percentage 2	73.13	74.92	75.4	75.77	74.31	76.07	77.05	75.40
Percentage total	58.93	66.93	69.35	70.66	71.44	71.06	71.67	72.36

[a] $m = 20$: using the past 20 observations to approximate the median; Percentage 1 = $P(d_{t+1} < d_t | d_t > \tilde{d})$, Percentage 2 = $P(d_{t+1} > d_t | d_t < \tilde{d})$, Percentage total = sum of number of events $(d_{t+1} < d_t | d_t > \tilde{d})$ and $(d_{t+1} > d_t | d_t < \tilde{d})$/total number of time windows.

with median approximated by the past 20 observations ($m = 20$ is determined by evaluating numbers of m ranged from 15 to 25 regarding estimation of the true median). We define "Percentage 1" as the percentage that the next spread was lower than the current spread given the current spread was higher than the median, i.e.,$P(d_{t+1} < d_t | d_t > \tilde{d})$; "Percentage 2" as the percentage that the next spread was higher than the current spread given the current spread was lower than the median, i.e., $P(d_{t+1} > d_t | d_t < \tilde{d})$; and "Percentage total" as the number of events that the next spread was lower given the current spread was higher than the median plus the number of events that next spread was higher given the current spread was lower than the median, divided by the total number of the observations. We can see from Table 2 that one minute and two-minute data are not very meaningful, since some of the data points may have same value for high and low due to short time window. It seems that three minute data works reasonably well. One may observe that as the time window increases from 3 minutes to 60 minutes, "% higher" converges to 68%, "%

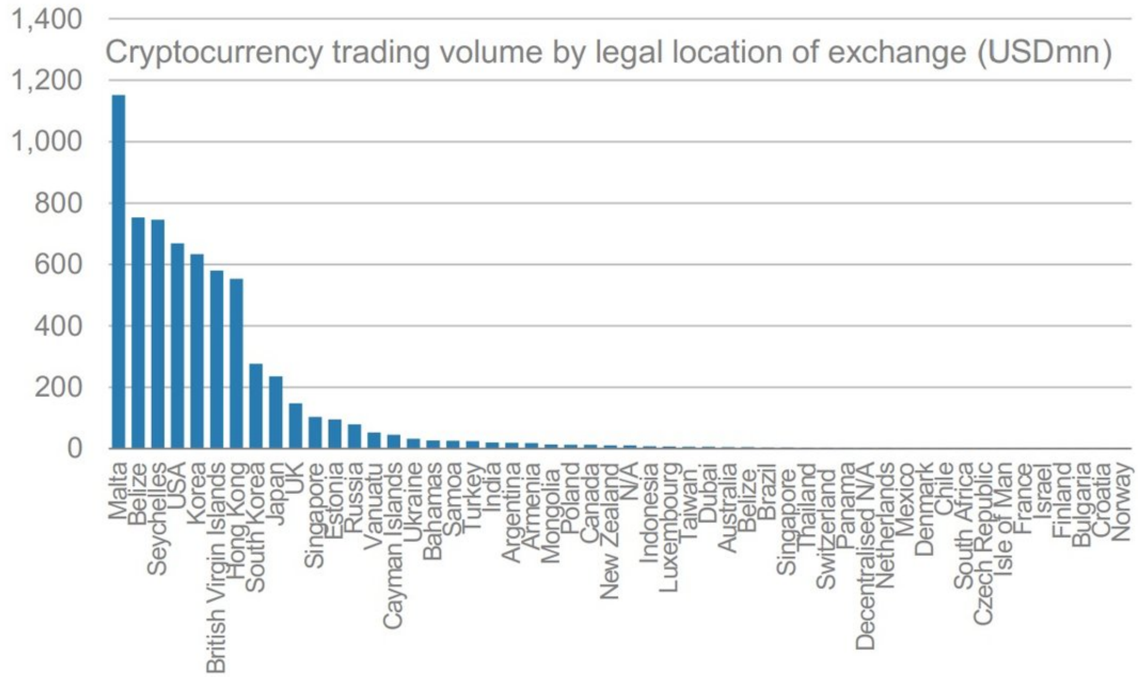
lower" converges to 75%, and "% total" converges to 72%. These numbers indicates that the 75% reversion rule still exist in the cryptocurrency market, although they are a little off from the target number 75%. This may be due to the fact that market was trending down over the month when the data was taken.

3.3 Identifying active geographic locations/markets

Suppose that there are T trading time points with trading volume $\{V_i\}$, $i = 1, 2, \cdots, T$. Recall that we use three-minute data for analysis as illustrated in Table 1. For example, from Feb 26, 2018, UTC 12:00 to Feb 27, 2018, UTC 12:00, there are $T = 480$ data points each corresponding to a 3 minute frame. To identify a pattern, data within each window with size λ (selected by CV criterion) is fitted by local linear regression. The smoothed curve provides the spikes/peaks needed to identify the patterns.

Let's take a look at the active countries with high trading volumes. Exhibit 1 displays the trading volume from one day in April 2018. The top five countries by trading volumes are, Malta, Belize, Seychelles, United States and Korea. Notice that Malta is in European markets, Belize and United States are in North American markets, and Seychelles and Korea are in Asian markets.

Exhibit 1: Most cryptocurrency trading is moving to a base in Malta



Source: CoinMarketCap.com, company websites, Morgan Stanley Research; For this chart Binance is in Malta and OKEx in Belize. Volume from one day in April but the relative exchange rankings are still comparable if several days are averaged.

Now we will fit a local linear regression to the daily data to examine the spikes in trading volumes, and to see if they are related to open or close of the traditional financial markets. We randomly select three days: Feb 26 UTC 12:00 to Feb 27 UTC 12:00, Mar 13 UTC 12:00 to Mar 14 UTC 12:00, and Mar 18 UTC 12:00 to Mar 19 UTC 12:00 in 2018 for daily data investigation. Figures 1, 2 and 3 plot the fitted smooth curves for these three days respectively. The number labeled in X axis is the nth observation from the entire dataset (Feb 18 to Mar 20, 2018). For example, in Figure 1, 4100 means the 4100th observation of the entire data. The numbers labeled in Y axis are trading volumes in thousand.

The red vertical lines are at time UTC 12:00 and the grey line is at UTC 0:00.

Figure 1: Fitted regression curve: Feb 26 UTC 12:00 to Feb 27 UTC 12:00

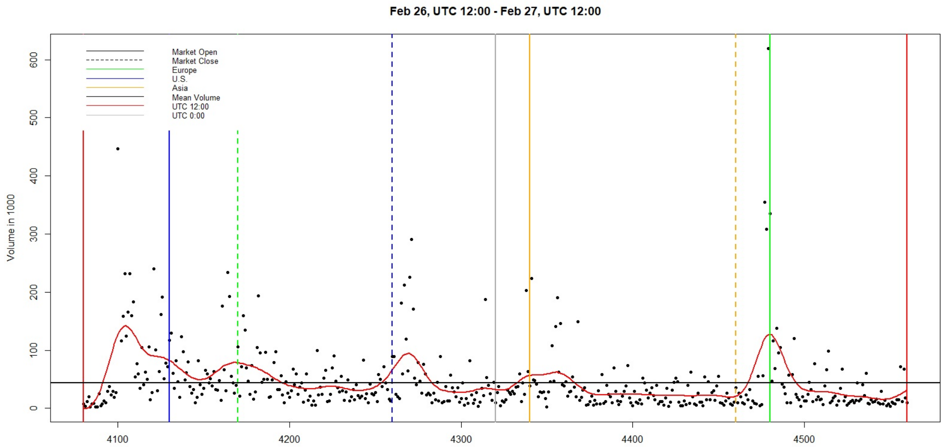


Figure 2: Fitted regression curve: Mar 13 UTC 12:00 to Mar 14 UTC 12:00

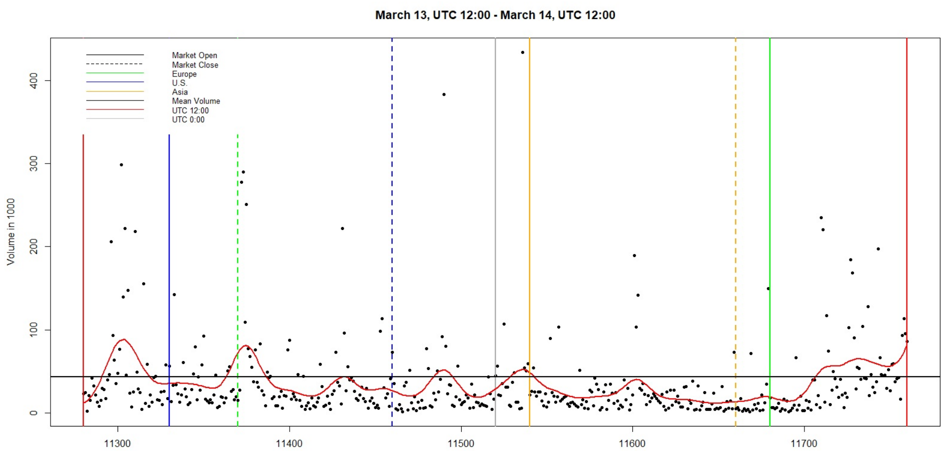
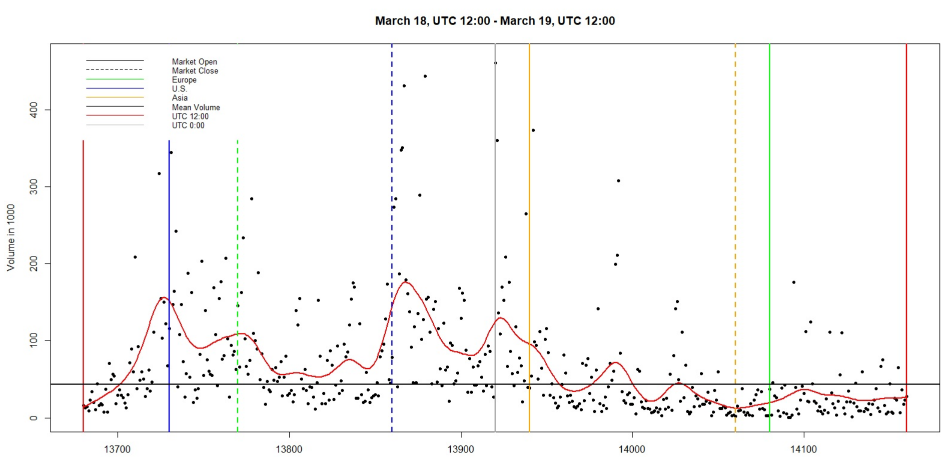


Figure 3: Fitted regression curve: Mar 18 UTC 12:00 to Mar 19 UTC 12:00



The black horizontal line is at 43,500, the mean volume of the entire data set (Feb 18 to Mar 20). The blue, green and yellow lines are the opening (solid line) or closing (dotted line) time of North American, European and Asian traditional financial markets respectively. Asia has markets that open between UTC 0:00 and UTC 3:00. Therefore, Asian market open time is approximated as UTC 1:00 (Japan traditional financial market open time). Japan traditional financial market close time is used to approximate the Asian market close time. Similarly, London and U.S. (eastern time) traditional financial market open/close time are used to approximate the European and North American market open/close time respectively. Starting from UTC 12:00, open of the North American market (the blue solid line) is followed by close of the European market (the green dotted line), next is followed by close of the the North American market (the blue dotted line), next is followed by open of the the Asian market (the yellow solid line), next is followed by close of the the Asian market (the yellow dotted line), and lastly is followed by the open of the European market (the green solid line) until 0:00 to the next day.

From Figure 1, we see that there are three spikes/peaks (high volume, active transactions) appear around $X = 4100, 4300$ and 4500, which are either close to the market opening or to the market closing. The first peak happens around the solid blue line (open of the North American market); the second peak happens around the dotted blue line (close of the North American market, and open of Asian market); and the third peak happens around the green solid line (open of the European market).

These peaks occur within plus or minus one and a half hours of the open/close time, when the institutions may have pre-market and after-hours meetings for buying or selling. Figures 2 and 3 reveals similar findings that the peaks usually appear around market opening or closing.

We've observed a pattern that peaks are around open or close of the major markets by examining daily data. Now we want to study the average volume of the 29 (Feb 18 to Mar 19) daily data to see if there is a similar pattern. Figure 4 plots the fitted regression curve by using average volume data. As can be seen from Figure 4, the largest peak happens after the opening of the U.S. market and the closing of the European market. After the U.S. market opens (blue solid line), average trading volume first decreases a little bit, then increases gradually until reaching its maximum right after the closing of European market. Other local peaks also happen around open or close of the markets as we observed from daily plot. The average trading volume reaches its minimum in the swoop of Figure 4, when the European market is the only active market (the Asian market is closed, and the North American market is not active yet).

3.4 The influence of government regulation and important news

The effect of regulation and media can cause the market to have a drastic shift in volume and price. In this section, we fitted a local linear regression on the data from

Figure 4: Fitted regression curve by using average volume from Feb 18th UTC 0:00 to March 19th UTC 23:59pm

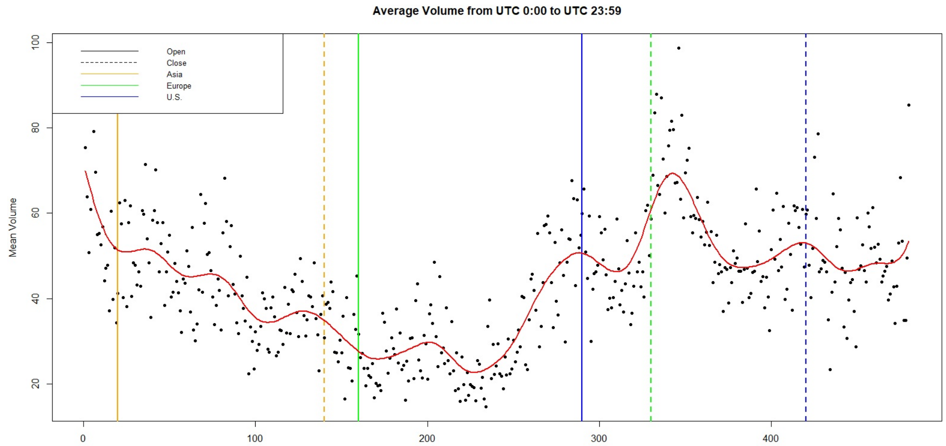
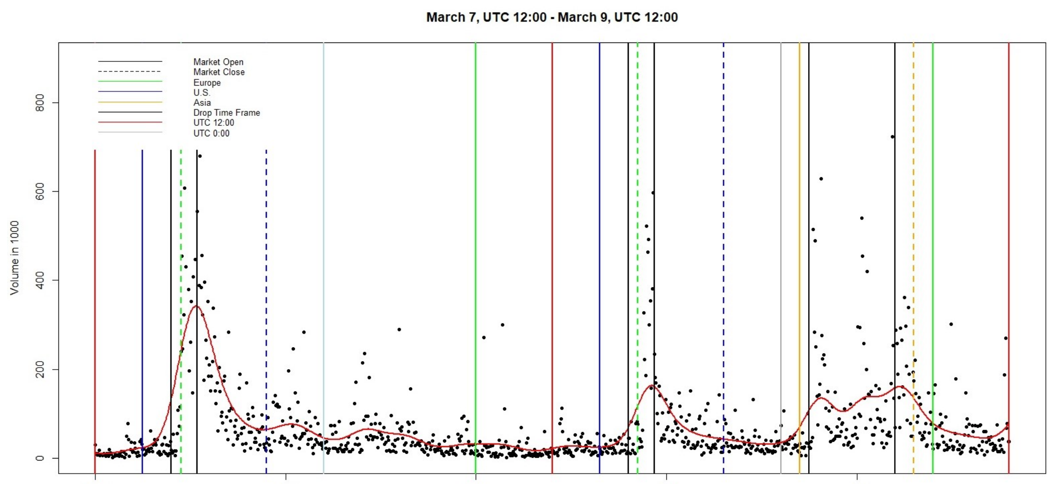


Figure 5: Fitted regression curve: Mar 7 UTC 12:00 to Mar 9, 2018, UTC 12:00



March 7, UTC 12:00 to Mar 9, UTC 12:00 to investigate how regulations and news affect the market.

On March 7, the United States Security and Exchange Committee released a public statement saying that bitcoin was unlawful and should be regulated. The investors and institutions were panic. As can be seen from Figure 5, a huge number of transaction volumes occurred around $X = 8500$ after the open of the North American market (the blue line) and the close of the European market (dotted green line). This

Figure 6: Price change of the Bitcoin during Mar 07 to Mar 10 2018

cause the Bitcoin market to shed nearly one thousand points (from \$10500 to \$9600) in one hour (from UTC 16:00 to UTC 17:20) as can be seen from Figure 6. In this time frame the trading volume of bitcoin jumped to 310,250 in one hour. After receiving the signals from North American market, Asia and European market were watching and were panic. They responded with conservative trading volumes maintained around 25,000 with price fluctuated around \$9600 until the following day. On March 8, the whole market reached their decision after a good deal of consideration. The second high volume of transactions occurred around $X = 9000$ (Figure 5), after the open of the North American market (the blue line) and the close of the European market (dotted green line). Correspondingly, the price dropped nearly eight hundred points (Figure 6) and the trading volume jumped to 250,700 in less than an hour (from UTC 16:40 to UTC 17:20).

On the following day March 9, the Bitcoin market was hit again in Asia because it was reported around UTC 1:30 that the exchange Binance was hacked. After receiving the news, several Asian countries made the decision to drop their stakes. Large trading volumes around 134,700 can be observed from the three spikes between the open (yellow line) and close (dotted yellow line) of the Asian markets (Figure 5). Meanwhile, media release caused the start of the price decline from $9400 at UTC 1:30 to $8500 at UTC 6:00 (Figure 6) when Asian markets begin to close.

4 Conclusions and discussion

In this research, we have studied some traditional stock market phenomena in the field of cryptocurrency markets. The empirical studies show that the 75% reversion rule works well in cryptocurrency markets. By using local linear regression estimator, we observe that the active market trading (represented by peaks from the fitted regression curve) usually occur around the open or close of the markets. We also observe that the stance of a government on cryptocurrency and news can affect a specific region and potentially have a ripple effect into other regions. These findings can help investors and institutions understand the cryptocurrency market, so they can make decisions before the next geographic location comes active.

When the trading volume is low and variability is small during a trading period, it becomes hard to identify a peak. One possible solution is to consider using a moving average volume instead of the fixed average volume to help identify the peaks.

On the other hand, the leading institutions in the amount of Bitcoin being traded worldwide changes over time. Therefore, the volume spikes in each region will also change over time. Exchanges are sometimes leaving countries where regulation and questions regarding the legality of bitcoin arises. For instance, Binance moved its operations from Asia to Malta due to such issues. Researchers should be careful when interpreting the results related to regions. Future studies may also consider developing statistical models for the bitcoin price formation by incorporating region information, analyzing market forces of supply and demand by different regions, and modeling intra-daily volume and prediction based on region effects.

Acknowledgements

The authors thank the referees for their careful reading of the manuscript and would like to express sincere appreciation for their helpful comments.

References

Bianchi, D. (2018). Cryptocurrencies as an asset class? an empirical assessment. *University of Warwick - Finance Group*, *June 6*.

Borri, N. (2018). Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance*, *Forthcoming*.

Brito, J., & Castillo, A. (2013). Bitcoin: A primer for policymakers. *Mercatus Center. George Mason University*, *September 21*.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of American Statistical Association*, *74*, 829-836.

Craven, P., & Wahba, G. (1979). Smoothing noisy data with spline function: Estimating the correct degree of smoothing by the method of cross-validation. *Numerische Mathematik*, *3*, 377–403.

Economist, T. (2015). The great chain of being sure about things. *The Economist*, *October 31*.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of American Statistical Association*, *87*, 998-1004.

Fan, J. (1993). local linear regression smoothers and their minimax efficiency. *Annals of Statistics*, *21*, 196-216.

Fan, J., & Gijbels, I. (1992). variable bandwidth and local linear regression smoothers. *Annals of Statistics*, *20*, 2008-2036.

Liu, Y., & Tsyvinski, A. (2018). Risks and returns of cryptocurrency. *NBER WORKING PAPER SERIES, NATIONAL BUREAU OF ECONOMIC RESEARCH*, 2084-2123.

Lo, A. W., Mamaysky, H., & Wang, J. (2000, March). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *Journal of Finance*, *55*(4), 1705-1765.

Makarov, I., & Schoar, A. (2018, 01). Trading and arbitrage in cryptocurrency markets. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3171204

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.

Pole, A. (2007). *Statistical arbitrage, algorithmic trading insights and techniques.* John Wiley and Sons, Inc.

Ruppert, D., & Wand, M. P. (1994). multivariate weighted least squares regression. *Annals of Statistics*, *22*, 1346-1370.

Stone, C. J. (1977). Consistent nonparmetric regression. *Annals of Statistics*, *5*, 595-645.

Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Annals of Statistics*, *8*, 1348-1360.

Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *Annals of Statistics*, *10*, 1040-1053.

Tschorsch, F., & Scheuermann, B. (2016). Bitcoin and beyond: A technical survey on decentralized digital currencies. *IEEE Communications Surveys and Tutorials*, *18*(3), 2084-2123.

Wahba, G. (1990). *Spline models for observational data.* Society for Industrial and Applied Mathematics.

Yermack, D. (2015). Is bitcoin a real currency? an economic appraisal. *In Handbook of digital currency. Elsevier*, 31-43.

Zhang, G. (2011). *Smoothing splines using compactly supported, positive definite, radial basis functions.* Ann Arbor: ProQuest, UMI Dissertation Publishing.