The Impact of Blockchain Bubble and Related Prediction Based on Investor's Sentiment*

Yuxin Guo

School of Business Administration, Northeastern University, Shenyang, China

20170891@stu.neu.edu.cn

Keywords: bitcioin; volatility; sentiment index; blockchain bubble

Abstract: Bitcoin investing has become increasingly popular recently. A myriad of researches has centred on it, while the conclusions are disparate. In this paper, the unary multiple linear regression model is introduced to investigate the relationship between Bitcoin volatility and investor sentiment, as well as the macro investing market, with the trading data tracing back to the beginning of 2004 from Coinmarketcap. An abnormal instability and significance level in the course of empirical research and robustness test are found surprisingly. Consequently, it is assumed, a priori, that this abnormity could be ascribed to the Blockchain industry upheaval in 2018. This assumption is proved in the supplementary period segment experiment that both of the roles of synthesized sentiment index and investing index appear to be conspicuously reversed before and after the collapse.

1. Introduction

Most assume ex post facto that the original intention of Nakamoto Satoshi strived for is more about an idealist and antagonist political pursuit, targeting the centralized economy manipulation that was widely accused of the financial crisis in 2008[1]. Bitcoin is much more popular and better known as a maverick financial asset [2, 3], attracting a various host of ambitious investors [4]. The capriciousness of the Bitcoin market has brought enviable wealth for some of the wise and lucky, while also spell doom for many of the rest. The enigmatic volatility characteristic is appealing to many economic and financial researchers to explore the delicate heterogeneous methods and encyclopedic, academic theories, albeit disparate and contradictory results come out.

At the outset, the mainstream of inchoate research focused on sentiment and public opinion analysis with orthodox computer technology. For example, Kim, Lee, et al. used the modified keywords sentimental analysis method to predict value fluctuation of the comments posted in the Bitcoin online forum over 2.8 years from December 2013 to September 2016, and proved its conduciveness [5]. Subsequently, in the same direction, Evita and Jacob considered Bitcoin as a kind of pure free-market commodity and predicted the price with Twitter sentiment analysis, getting a satisfying accuracy rate [6]. Consequently, further orthodox financial empirical researches were involved, seeming to corroborate those conclusions at first. One of them could be Sovbeto's study on investigating factors influencing Crypto-currencies. Namely, the attractiveness of tokens is a significant factor in value fluctuation. It also indicated that market beta and trading volume are determinants, echoed in the pervasiveness of the financial investigating method applied in the Bitcoin volatility study [7]. However, the results contradicted what had been proved before and appeared into normalcy for the subsequent studies that cast doubt on the precedent. Aalborg et al. introduced the HAR-type model in the research of Bitcoin market, thoroughly considering returns, volatility, trading volume, transaction volume, inter alia, change in the number of unique Bitcoin addresses, the Volatility (VIX) Index, and Google searches, yet found none of the considered variables could predict Bitcoin returns [8]. Nonetheless, Lyócsa et al. kept the idea that the volatility of Bitcoin reacted strongly to positive investor sentiment regarding Bitcoin regulation extracted using Google searches [9]. Based on the GARCH-MIDAS framework, Walther et al. endorsed a global real economic activity as a crucial factor among the tested exogenous drivers, which was at variance with the precedent research [10]. López-Cabarcos et al. gave credence that Standard & Poor (S&P) 500 returns, and VIX returns did forge a close connection with Bitcoin volatility variance insofar as relatively stable periods [11]. And again, Wang, Liu, and Hsu qualified this conclusion to concentrate the economic activity factor into the stock market of Europe and the US [12].

Disparate conclusions exist in this academic sector, stemming from plausible researching methods of common and rarefied, concise and complicated, alleging their own reliability and rationality. While the apparent divergence indicates that there is something being lost sight of, biasing the aberrant results made by those meticulous models.

In this research, an empirical volatility study is resumed with a unary multiple linear regression model [13, 14], considering the macro investing market and investor sentiment under a relatively macroscopic perspective roughly, being echoed in the mainstream financial inspection of Bitcoin. Nonetheless, the differences in sample selection on time dimension may profoundly sway the final result even with an extremely analogous model, and this study concentrates on this variable.

This sort of variation is caused by single period selection, which presents the considerably unreasonable significance level and abnormal instability of its robustness, and should be considered as an anomaly worthy of study. This counterintuitive phenomenon is further explored with a supplementary period segment experiment. The results present that both the synthesized investing index and sentiment index appear to be conspicuous contrast before and after the collapse. This research may shed light on the irregularity and unpredictability of the Bitcoin market, and play a guide rule to the Bitcoin investing strategy making.

This manuscript is structured as infra: In Section 2, the data used in this research are described. Then in Section 3, the economic model that the whole research bases is specifically elaborated. In Section 4, the empirical result and the abnormity underlying are described. Consequently, the Bitcoin price collapse is discussed based on the model proposed in Section 3. And consequently, we investigate the influence of Blockchain Bubble on Bitcoin investing market based on the proposed model and supplementary period segment experiment in Section 5. Eventually, Section 6 gives a research conclusion.

2. Data

Bitcoin has a rather long transaction history that could date back to 2008, approaching the moment it was created. Nonetheless, in this paper, a relatively limited period is adopted for pursuing the accuracy and precision of finance research at the expense of advisable reduction of sample size. To be specific, the closing price data of Bitcoin trading from January 1st in 2014 to August 31st in 2020 are collected, and the counterparts of S&P 500 index are used for processing into Investment Index, VIX, Google index of 'Bitcoin'. The trading volume of Bitcoin is used to be synthesized into the Sentiment Index. All the data related to Bitcoin trading come from Coinmarketcap, while other essential data are gathered from the most original resources approachable.

In deference to the immaturity of Crypto-currency exchanges, the frequent abnormal needle-shaped curve phenomenon, inter alia consequent conventional cross-platform arbitrage, a day as a single-window phase is simply used in volatility calculation out of the consideration of the veracity and authenticity. A commonly used method [15] is introduced that augment the original figure 100 times in the actual implementation for better observation as infra.

$$r_{t,i} = \left(\ln P_{t,i} - \ln P_{t,i-1}\right) \times 100\% \tag{1}$$

The subsequent modified calculating method is shown below.

$$V_t = (\ln \frac{P_{t,i}}{P_{t-1,i}} \times 100\%)^2 \tag{2}$$

3. Economic Model

The relationship between possible underlying influence factors and the variation trend of Bitcoin price is described with a linear regression model from a probabilistic perspective. Except for public concern, which is the mainstream of empirical study and forecast of status quo, the variation of Bitcoin price is attributed to the macroeconomic environment, i.e., both sentiment index and investment index are employed as the two independent variables of the economic model. To be specific, the volatility of S&P500 is used to denote the investment variable, and synthesize the sentiment variable with the volatility of VIX, Bitcoin trading volume, as well as Google index of Bitcoin. In section 3.1, the sentiment index establishing method is expounded in detail, and a whole description of the regression model is described.

3.1 Sentiment Index Establishing

Though the qualification of Bitcoin as a rightful virtual financial asset is still under discussion, the impact of investor sentiment on this maverick is nonetheless keeping unassailable and sometimes over-estimated in some researches that take the sentiment of public comment as the only orientation of Bitcoin price.

In this research, public concern is taken as a no mean indicator, yet also the sentiment index is valued from a more comprehensive perspective with a comparably classical financial investigating method, treating Bitcoin as an actual financial res in commercium that influenced by macro investing emotion and substantial trading situation. To be specific, the volatility of VIX Close, Bitcoin trading volume, inter alia, Google index are introduced as the three proxy variables of sentiment index and synthesize the Investor Sentiment index with Principal Component Analysis (PCA) [16], reflecting the variation of macro investment market motion, Bitcoin investing enthusiasm as well as the public attention on Bitcoin. In Table1, the component matrix of the PCA process is demonstrated as a result, and Figure 1 exhibits the variation curve of the synthesized sentiment index.

Table 1. Component Matrix

	Component	
	1	2
Volatility of VIX	0.228	0.839
Trading Volume of Bitcoin	0.801	0.204
Google Index	-0.632	0.560

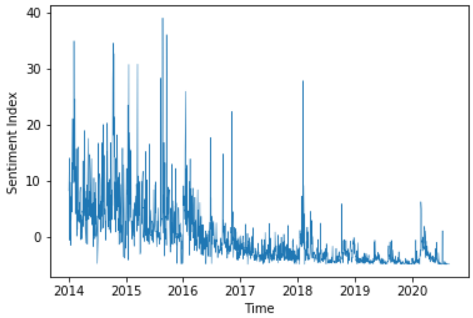


Fig. 1. Sentiment Index.

3.2 Regression Model

In this research, a unary multiple linear regression model is introduced to investigate the relationship between Bitcoin volatility and investing sentiment, as well as the macroeconomic market, with ordinary least squares as regression parameter estimation method. $\mathrm{Sen_i}$ is denoted to express the sentiment index of the ith trading day, and $\mathrm{Inv_i}$, the corresponding Investment index. The specific regression model is exhibited below.

$$\mathrm{RV_t} = \beta_0 + \beta_{\mathrm{Sen}} \cdot \mathrm{Sen_i} + \beta_{\mathrm{Inv}} \cdot \mathrm{Inv_i} \tag{3}$$

The calculation method of $\mathrm{Sen_i}$ has been separately discussed in section 3.1, while the index $\mathrm{Inv_i}$ is equal to the volatility of S&P 500, which has been calculated in section 2.

4. Empirical Result

In this section, the empirical result deriving from the model delineated in Section 3 is elaborated with the data from Section 2. With meticulous experimentation notwithstanding, the abnormal instability and significance level are surprisingly observed and their characteristics are further explored. In section 4.1, a brief statistical description of the data used is given to show the distribution and trend variation of each important data set in general terms. Subsequently, Section 4.2 we presents and expands on the empirical result and the abnormal underlying in a systematic way, and Section 4.3 substantiates this anomaly with a random test.

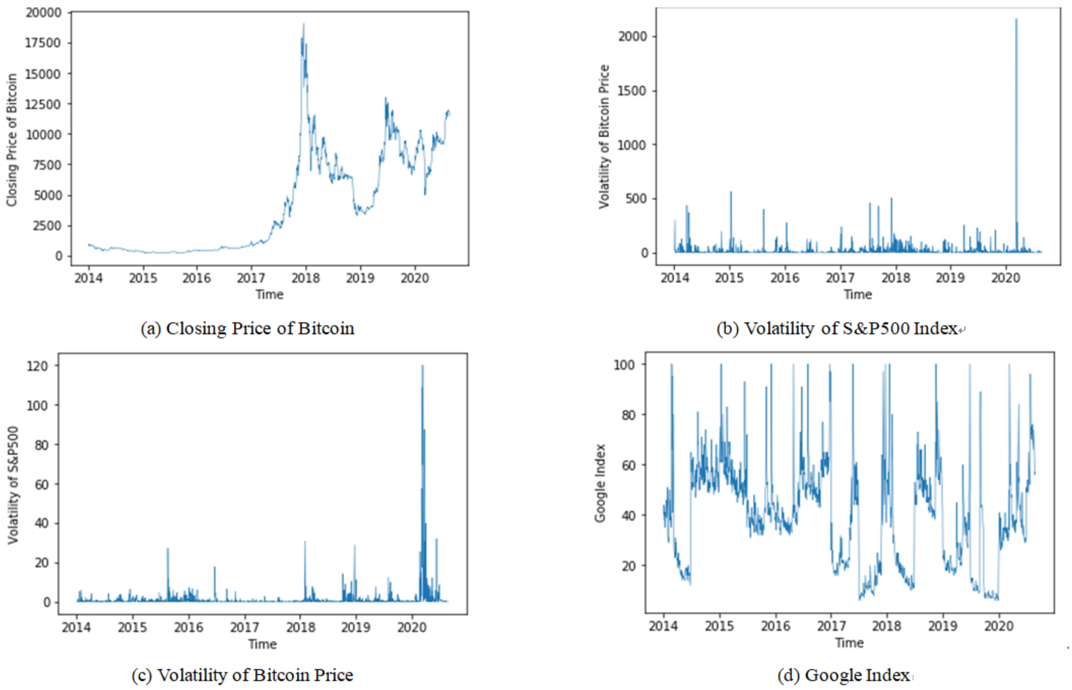
4.1 Statistical Description

In this research, the volatility calculation is used [17] to process most germane data and the corresponding trend variation is further denoted, i.e., except for Google Index and Bitcoin trading volume, other indispensable figures is constructed with the volatility calculation results of related data. Figure 2 shows a sketchy trend of Bitcoin price, Volatility of S&P500 Index, Volatility of S&P500 Index, and Investor Sentiment Index. The figures exhibit the violent fluctuation of related figures in a crystal clear way that anchored change points do exist, albeit the overall trends of four curves do not fully correspond to each other.

Table 2. Descriptive Statistics of Variables

Variable	Volatility of VIX	Volatility of S & P500	Volatility of Bitcoin Price	Google Index
Minimum value	0	0	0	6
Maximum value	5902.004	120.1135	2159.738	100
Mean value	69.48505	1.245	17.18776	38.23454
Standard Deviation	211.282	5.757556	68.81945	19.56789
Kurtosis	366.3187	236.0187	583.3655	-0.20215
Skewness	15.1166	13.86233	20.13362	0.324672

And in Table 2, the data characteristics are elaborated with several commonly used statistical indicators, including minimum value, maximum value, mean value, standard deviation, kurtosis, and skewness. The disparate distributions of data also prove the suitability of introduced data sets to continue the research.



(a) Closing Price of Bitcoin
(b) Volatility of S&P500 Index
(c) Volatility of Bitcoin Price
(d) Google Index

Fig. 2. Statistical Description.

With the statistical description supra, both Figure 2 and Table 2 contribute to exhibiting the prodigious and gripping trading history of Bitcoin, the ever-changing concentration public pay on, contracting with the considerably stable trend of S&P500 volatility. The violent fluctuation is strewing with the line of Bitcoin variation, albeit the Blockchain upheaval is outstandingly speculative among.

4.2 Experimental Result

The experimental regression result is exhibited in Table 3, illustrating that both of the chosen factors, i.e., investor sentiment and macro investing market influence, are in a positive relationship with Bitcoin price volatility. The P-Value of $\beta_{Inv}$is notwithstanding at a rather satisfying significance level, indicating an obvious causal connection. The ones of Sentiment Index are far short of expectations and hard to give credence to any substantial validity, not least considering the subtle limited magnitude of the coefficient.

Table 3. Parameter Estimation Result

	β_0	β_{Sen}	β_{Inv}
Coefficient	11.6429***	0.1494	4.4537***
P-Value	(7.195)	(0.542)	(16.211)
T-Value	0.000	0.588	0.000

Given that the prevalence of introducing sentiment analysis to predict Bitcoin price variation of previous studies, the experimental result is grossly counterintuitive and abnormal, since the investor sentiment should theoretically be even more effective on financial assets without an orthodox credit base like Crypto-currencies. To corroborate the existence of this phenomenon, the model with a robustness test is subsequently assessed in section 4.3.

4.3 Robustness Test

In the robustness test, the random function of program is employed to select a subset of the original sample defined to be longer than 2 years, and the result afresh with the new data set is calculated. The test is repeated two times with different, randomly selected subsets, and the estimation result is presented in Table 4.

Table 4. Parameter Estimation Result of Robustness Test

	Whole Sample	Random1	Random2
β_0	11.6429***	15.9942***	12.4892**
	(7.195)	(9.332)	(2.152)
	0.000	0.000	0.032
β_{Sen}	0.1494	0.0322	0.3297
	(0.542)	(0.077)	(0.227)
	0.588	0.938	0.821
β_{Inv}	4.4537***	0.0981	4.6086***
	(16.211)	(1.143)	(12.644)
	0.000	0.253	0.000
r^2	0.139	0.001	0.186

Note: numbers in each line paralleling to parameters denote Coefficient P-Value and T-Value from up to down, same with Table 5 infra.

The robustness test result further validates the existence of the abnormal phenomenon in Bitcoin volatility regression model. The parameter estimation results in the second random test, i.e., Random 2 is highly consistent with the original experimental result of the whole sample. Both of the two random tests show a disillusionary high P value, indicating the synthesized investor sentiment proved barren.

The significance level is feeble, albeit the overall results of three models in robustness tests differ widely in an unusual way. First of all, even though the rather limited r^2 of three tests show a general feeble predictive ability, the one of test Random 1 is still considerably shockingly low. To be more specific, the experimental result of Random 1 shows a conspicuous reduction in the coefficient magnitude compared with the whole-sample one, and additionally, an aberrant decrease in the Investing index as while as its corresponding significance level demonstrates the instability of the model. And on the contrary, most evaluation index of Random 2 is improved because of the further selection of data set. The r^2 of Random 2 manifest that the selected partition is more predictive. The T values of two introduced variables both increase, and their magnitudes of coefficients show that the sentiment index and investing index are growingly influential.

In view of the fact that time series is the only alteration in this test, it is reasonable to ascribe this counterintuitive phenomenon into the period segment selection. Such a conclusion conjures up the impressive price upheaval caused by the Blockchain bubble event presented in section 4.1 ex post facto. Consequently, it is a priori assumed that this Bitcoin price collapse leads to the unreasonable empirical result.

5. Discussion of Blockchain Bubble

To explain the counterintuitive phenomenon, the influence of Blockchain bubble and consequent Bitcoin price collapse is further explored by subdividing the investigated trading period and recalculating the regression model separately, ceteris paribus. In this period-segmented experiment, the collapse period is taken out, identifying and endorsing the Bitcoin collapse from November 1st in 2018 to the last day of this year, and reconstructing the linear regression model with the rest dimidiate time.

Table 5. Parameter Estimation Result of Period-segmented Experiment

	Whole Sample	Before Collapse	After Collapse
Time	2014/01/01-2020/08/31	2014/01/01-2018/10/31	2019/01/01-2020/08/31
β_0	11.6429***	16.5681***	-25.5391
	(7.195)	(11.271)	(-1.540)
	0.000	0.000	0.124
β_{Sen}	0.1494	0.1193	-7.1555 *
	(0.542)	(0.528)	(-1.821)
	0.588	0.597	0.069
β_{Inv}	4.4537***	0.1626	5.3968 1***
	(16.211)	(0.211)	(10.751)
	0.000	0.833	0.000
r^2	0.139	0.000	0.244

The supplementary period segment experiment result presented in Table 5 demonstrates that the problem of disqualification of significance level and coefficient magnitude variance could be solved by and large. The variance of r^2 indicates that the Bitcoin price is much more regulative and predictive after the price collapse, while the market is nearly straightforward restive before the industry upheaval. The obvious divergence of the β_{Sen} and β_{Inv} also contribute to the qualitative transformation of Bitcoin investing market before and after the collapse. It is a priori assumed that these variations may highly possibly be capable of explaining the counterintuitive phenomenon illustrated in Section 4.2 and 4.3. The variances of two coefficient and corresponding possible interpretation will be elaborated separately in Section 5.1 and 5.2.

5.1 Coefficient of Sentiment Index

The significance level of sentiment index in the whole-sample model indicates the relationship between the synthesized investor index and Bitcoin is subtle, which is consisted of the result of the pre-collapse regression model. However, the sharp reduction of P-value in the subsequent post-collapse period reforges the connection. The change of coefficient itself is of subversiveness. β_{Sen} in both the model of the whole sample and before the collapse is limited positive, while the coefficient is, to a large extent, conversely amplified into nearly 70 times and becomes negative. This reversal manifests that the synthesized investor sentiment index in this research does not become actual effective until the Blockchain event is finished. And in the midst of this limited effective period, the influence of investors is substantially negative, which is at variance with the common truism.

An inferable latent interpretation could be that, from a parochial perspective, the components of the synthesized sentiment index lead to the anomaly. To be specific, the recent economic crisis and COVID-19 have made Bitcoin contrarily became a kind of safe-haven asset, which lead to a negative correlation with the VIX index. Google index may also lose its effectiveness since people are much more familiar with Bitcoin after Blockchain bubble event, and the public concern is no more with regard to the investing willingness, but linked with provocative price catastrophes. Another explain in hindsight is under a rather macroscopic perspective, which is with the maturing of Bitcoin investing market, professional industry insight of those actual investors is becoming more and more separated from general thoughts.

5.2 Coefficient of Investing Index

Even though the significance level of the Investing index in the whole sample regression experiment is considerably ideal in toto, the P value of the pre-collapse period indicates a feeble connection with S&P 500 volatility, and the Investing Index should be considered as actual effective de facto only after the Blockchain industry upheaval, not least considering the conspicuous similarity between the original regression result and the one of the post-collapse.

This phenomenon is fascinating but elusive. The result may be attributed to heterogeneous forces and convoluted casual connections notwithstanding. A revelation could be that the macro investing environment is probably, after the American government yoke Crypto-currencies into securities regulatory system, exerting a much more dominating influence on the Bitcoin trading market and accounting for the spectacularly increasing wieldy.

6. Conclusion

This paper reintroduces the widely used regression model on Bitcoin volatility investigation with a random robustness test, presenting the abnormal instability and unreasonable significance level of the empirical and test results. The Bitcoin price collapse caused by Blockchain industry upheaval in 2018 is concentrated, with further supplementary period segment experiments. The speculative effect of this event on Bitcoin volatility variance characteristics is proved. The additional experiment results systematically present that both the macro investing market factors and compositive sentiment factors are significantly qualified before the reshaping of Blockchain industry, and have a satisfactory significance level after the event. This research may contribute to the interpretation of the capriciousness of Bitcoin market, and offer effective revelation to Bitcoin investors and policy-makers.

References

[1] Y. Chen, "Decentralized finance: blockchain technology and the quest for an open financial system," SSRN Electron. J., July 2019.

[2] H. Elendner, S. Trimborn, B. Ong, and T.M. Lee, "The cross-section of crypto-currencies as financial assets[1]," in Handbook of Blockchain, Digital Finance, and Inclusion, vol. I, D.L.K. Chuen and R. Deng, Eds. New York: Academic Press, 2016, pp. 145-173.

[3] L. Wang and Y. Liu, "Exploring miner evolution in bitcoin network. international conference on passive and active network measurement," International Conference on Passive and Active Network Measurement, New York, NY, USA, pp. 290-302, March 2015.

[4] L.L. Xu and S.O. Finance, "Analysis on the functions and risk of virtual currency——take

bitcoin as an example," J. Econ. Manage. Strat., no. 1, pp. 98-102, July 2016.

[5] Y.B. Kim, J. Lee, N. Park, J. Choo, J.H. Kim, and C.H. Kim, "When bitcoin encounters information in an online forum: using text mining to analyse user opinions and predict value fluctuation," PLoS One, vol. 12, no. 5, p. e0177630, May 2017.

[6] E. Stenqvist and L. Jacob, Predicting Bitcoin price fluctuation with Twitter sentiment analysis. Stockholm, Sweden: KTH Royal Institute of Technology, 2017.

[7] Y. Sovbetov, "Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero," J. Financ. Econ., vol. 2, no. 2, pp. 1-27, February 2018.

[8] A.H. Aarhus, P. Molnár, and D.V.J. Erik, "What can explain the price, volatility and trading volume of bitcoin?," Financ. Res. Lett., vol. 29, pp. 255-265, August 2019.

[9] Š. Lyócsa, P. Molnár, T. Plíhal, and M. Širaňová, "Impact of macroeconomic news, regulation and hacking exchange markets on the volatility of bitcoin," J. Econ. Dyn. Control, vol. 119, p. 103980, October 2020.

[10] T. Klein, I. Mathur, and C.J. Neely, "Exogenous drivers of Bitcoin and Cryptocurrency volatility – A mixed data sampling approach to forecasting," J. Int. Financ. Mark. I., vol. 63, p. 101133, November 2019.

[11] M.Á. López-Cabarcos, A.M. Pérez-Pico, J. Piñeiro-Chousa, and A. Šević́c, "Bitcoin volatility, stock market and investor sentiment. are they connected?," Financ. Res. Lett., p. 101399, December 2019.

[12] J.N. Wang, H.C. Liu, and Y.T. Hsu, "Time-of-day periodicities of trading volume and volatility in bitcoin exchange: Does the stock market matter?," Financ. Res. Lett., vol. 34, p. 101243, May 2020.

[13] P.C.B. Phillips and H.R. Moon, "Linear regression limit theory for nonstationary panel data," Econometrica, vol. 67, no. 5, pp. 1057-1111, February 1999.

[14] P. Katsiampa, "Volatility estimation for bitcoin: A comparison of garch models," Econ. Lett., vol. 158, pp. 3-6, September 2017.

[15] W. Li, Y.X. Cheng, and Q. Fang, "Forecast on silver futures linked with structural breaks and day-of-the-week effect," N. Am. J. Econ. Financ., vol. 53, p. 101192, July 2020.

[16] B. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction," IEEE Trans. Automat. Contr., vol. 26, no. 1, pp. 17-32, March 1981.

[17] T.G. Andersen and T. Bollerslev, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," Int. Econ. Rev., vol. 39, no. 4, pp. 885-905, April 1998.