PRICE AND PREJUDICE: AN EMPIRICAL LOOK AT THE VALUE FORMATION OF BITCOIN

by

David Doster

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Economics

Charlotte

2018

Approved by:

____

Dr. Craig Depken

____

Dr. Steven Clark

____

Dr. Paul Gaggl

ABSTRACT

DAVID DOSTER. Price and Prejudice: An Empirical Look at the Value Formation of Bitcoin. (Under the direction of DR. CRAIG DEPKEN)

Bitcoin price formation has been the topic of many studies due to the recent rise in popularity of cryptocurrencies around the globe. The problem not only lies with attempting to find how the value of this currency is established, but finding a framework that best describes how Bitcoin is created. In this paper, a modified version of Barro's framework is used, along with other prior frameworks, in an attempt to model pricing variation and formation for Bitcoin. To find causality, a simple VAR(p) model is used as a starting point, where the lag-order is selected based on BIC. This model includes various network statistics, Bitcoin popularity measures, commodity prices, and financial markets to identify potential pricing factors which could be argued to cause changes in Bitcoin price. A multivariate GARCH approach (MGARCH) is then used to fortify this model by not only modeling these causal relationships but modeling changes in volatility, eventually using Google trends to explain volatility changes in Bitcoin prices. According to these models, Bitcoin price formation follows an AR(1) process with ARCH/GARCH effects where these ARCH/GARCH effects can be explained by Bitcoin popularity. Due to the returns of Bitcoin relying on past information, a violation of the efficient market hypothesis may be present, meaning that arbitrage may exist in the Bitcoin market.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Bitcoin has recently become a rising star in the public eye because of its large price swings and volatility rarely seen in any sort of financial instrument. Unlike other financial instruments or currencies, Bitcoin itself has no government to stabilize its price or business assets to back its value. The question then becomes, how is such an instrument priced or who is determining its value? Since one cannot rely on the net present value of its future cash flows or possibly predict any future value due to lack of interest rates, what is causing such large swings in its price? In this paper, these questions are addressed from an empirical point of view. What do the data tell us? To begin, we must first understand the mechanics of Bitcoin and provide a little background to the problem.

Contrary to popular belief, the idea of digital currency is not a new idea. Since banks have been keeping computerized ledgers and digital accounts, digital currency has been part of our everyday life for over two decades. The innovation comes in the form of a decentralized currency, a purely unregulated medium of exchange which can be traded freely beyond the control of banks, governments, or other outside agents. Technology and the rise of the internet have provided the ability to accomplish this task with a simple but effective means of confirming transactions: the blockchain.

The blockchain is truly the backbone of any digital currency and is nothing more than a publicly held ledger that posts all past and present transactions. These transactions are confirmed by a network of computers, called miners, that meticulously check the validity of each transaction. This check consists of three parts: (i) whether the specified sender, in most cases anonymous, has the correct amount of Bitcoins in their wallet to make the transaction, (ii) the receiver, also anonymous in most cases, exists and can receive the Bitcoins, and (iii) past confirmations of this same transaction have successfully performed these checks. (Bitcoin, 2008) As thousands of computers check these transactions it is then considered confirmed and the appropriate amount of Bitcoin is then transferred. Because hundreds of these transactions can be combined into a "block" and the network is able to confirm these transactions fairly quickly, the time and cost of sending Bitcoin is ostensibly very low.

To incentivize these miners to confirm transactions, senders of Bitcoin attach a fee associated with their transactions, usually some fraction of a Bitcoin. On top of these transaction fees a miner also adds a reward block to the blockchain. Reward blocks award a certain amount of Bitcoin to the miner who has successful "mined" the block. The ability for the network of miners to mine a block is dependent on the network difficulty. This difficulty is automatically adjusted every eight days to ensure that a block is mined every ten minutes, on average. If the processing power of the network rises so the average time to mine a block is shorter than ten minutes, the difficulty of mining a block is increased. This system creates a very steady flow of new Bitcoins and encourages miners to confirm as many transactions as possible.

1.1 A Bitcoin Mining Example

Say there is a game where there are two players. Each player has a button in front of them and a basket next to them. To win the prize, in this case Bitcoins, each player must press their button to draw a series of random numbers in an attempt to match a given series of random numbers. Each player is able to simultaneously press their button and put things in their basket, which is representative of transactions. The first player to match the given series of random numbers wins the prize, Bitcoins, along with any Bitcoins that were part of the transactions which were put into the basket. The amount of random numbers the players have to match is determined by the software which is used to govern the rules of the game. The software adjusts the difficulty, the length of numbers that a player has to match, to insure that a player finds the sequence of random numbers every ten minutes. How fast a player can hit their button is determined by their hashrate which is representative of a player's processing speed on the Bitcoin network. The higher a player's processing speed the quicker the player is able to press their button and generate higher volumes of random numbers.

CHAPTER 2: LITERATURE REVIEW

Because Bitcoin and digital currencies are relatively new concepts, there is not a large literature that analyzes the economic aspects of Bitcoin. The few that are available answer questions that help develop a model that explains what is driving the price of Bitcoin.

The focus of several papers is the nature of Bitcoin and whether it should be considered as a currency. Yermack (2014) argues against Bitcoin being considered as a currency for several reasons. One argument is that Bitcoin does not act as a sufficient medium of exchange. This conclusion comes from the fact that Bitcoin is rarely used to actually purchase goods and services; on average showing far less than one purchase per day (Yermack, 2014). Since this paper was written in 2014, this statement still seems to hold weight. Recent articles from Bloomberg and other news agencies report there is actually a decline in the number of online retailers who accept Bitcoin as payment (Boomberg, 2017) and hardly any brick-and-mortar establishments accept Bitcoin to purchase goods. Due to the volatility surrounding Bitcoin and its lack of correlation with other commodities, Bitcoin is also not effective for hedging or risk management (Yermack, 2014). If Bitcoin is not a currency, then how should a model be structured to capture price fluctuations?

Many papers start by comparing Bitcoin mining to that of mining a commodity, such as gold. Many authors modify an approach in Barro (1979), relating the price of a given commodity to the demand and supply of that commodity along with some factor of government intervention (Barro, 1979). Since Bitcoin supply and demand factors are not controlled by any government entity, the modified framework simply measures the supply and demand factors of the model. The issue with measuring the supply and demand as factors for pricing Bitcoin comes from the known supply-side of Bitcoin. Since there are a known amount of Bitcoin in circulation, the rate at which they are produced is fixed (on average), and a known total amount of Bitcoin that can ever be produced, pricing fluctuations should not occur due to the supply of Bitcoin, in an efficient market. Another problem with Barro's framework is that it assumes there is some cost for holding Bitcoin, such as inflation or interest rates. With Bitcoin, due to the fact that it is not tied directly to a specific nation or government, inflation and interest rates should not play a roll in price formation. Where one nation may be experiencing high inflation or interest rates, another may be experiencing the opposite. Also, Bitcoin exchanges offer a wide variety of currencies in which a person may convert Bitcoin to and there are no limits to how much are bought and sold in one day.

Following the same framework as Barro (1979), Smith (2016) argues that Bitcoin's value should be measured in the form of exchange rate dynamics if it is to be modeled as a commodity. In this paper, Smith uses a Vector Error Correction Model (VECM) to measure relative price dynamics of Bitcoin across different nominal exchange rates, finding that these dynamics explain long-run changes in Bitcoin price. (Smiith, 2016)

CHAPTER 3: DATA AND METHODOLOGY

Following previous studies, the data used for the model are divided into four parts: (i) network statistics, (ii) BTC popularity, (iii) global commodities, and (iv) global financial indices. Each of these parts will be combined to form four models which will be measured using weekly data. Table 3.1 describes the variables associated with each of these measures. Weekly data is used because Google trend data is measured

Table 3.1: Variable List for BTC Pricing Models

(i) Network	(ii) Popularity	(iii) Commodities	(iv) Financial Indices
Total Bitcoin	Google Trends	Gold	Dow Jones
Network Hashrate	# of Transactions	Oil	Nikkei 225
Network Difficulty	# of Unique Addresses	Natural Gas	Hang Seng
Cost per Transaction			SHCOMP

on a weekly basis. To account for the weekly timespan, data for each of our variables is averaged each week. This helps eliminate some of the volatility associated with the network and price and allows for a much more stable view of the variables over time, where short-lived, large price swings are averaged out.

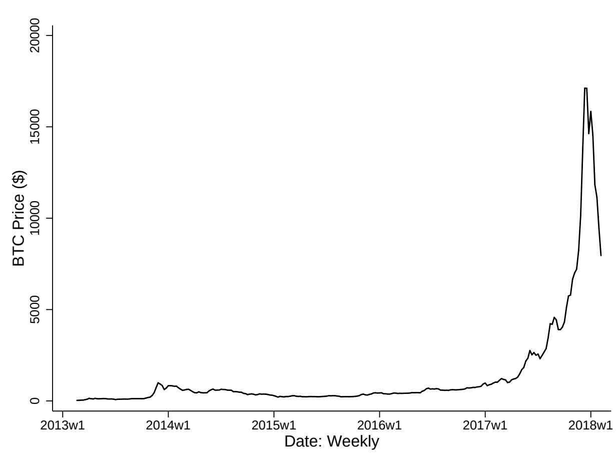

Figure 3.1: BTC Price

3.1 Variables

To begin describing the model, consider the dependent variable: Bitcoin price. Bitcoin price is measured as the exchange rate between Bitcoin and the U.S. dollar (USD). Though Bitcoin is considered a global asset, its value can be described in terms of a more commonly used global currency. Since the U.S. dollar is the most widely used and traded global currency[1], measuring Bitcoin in terms of the USD provides an easily interpretable measure for a commodity that may otherwise have a nebulous interpretation. Once the dependent variable has been defined, a definition of the independent variables is necessary.

Though Bitcoin has a certain anonymity behind its use, the data behind the network is quite transparent. Because all Bitcoin transactions and inner workings are publicly posted, data on the Bitcoin network is readily available and quite reliable. Data from the network can be accessed through a number of websites or directly from the network via the public ledger. For ease of use, all data collected for the Bitcoin network statistics are collected directly from Blockchain[2].

To be consistent with prior studies, total Bitcoin in circulation will be used as an explanatory variable. This variable is a somewhat precise measure of Bitcoin supply and can be used to capture how changes in supply may effect Bitcoin's price. The reason the total Bitcoin measure is somewhat precise is that around 20% of total Bitcoin has been lost, the majority of this due to misplaced wallets. One estimate is that approximately 2.56 million BTC worth approximately 20.0 billion USD has been lost (Fortune, 2017). These BTC's are then classified as "out of circulation" and are almost impossible to recover. This is an important fact but should not have an effect in the model. The majority of these lost BTC were mined in the early years of BTC when the price was less than $1.00 and has no impact on future creation

[1] According to www.investopedia.com and global foreign exchange market data
[2] www.blockchain.info

of BTC. There have been no measured losses of BTC in the last few years with an estimated zero BTC lost in 2017 (Fortune, 2017). This estimate comes from tracing BTCs that are mined to addresses that receive these BTC. If these addresses have sent or received BTC then it is classified as active and the BTCs are then considered to be in circulation. Also, these estimates cannot distinguish between wallets that are lost and wallets that are not. If an early adopter of BTC simply bought BTC when it was cheap and has held onto it without any activity, this BTC would be classified as "out of circulation" when it is not. For these reasons BTC supply will be taken as is but further studies may look into the effect of lost BTC on the price.

The next two variables in the network statistics are the network hashrate and the network difficulty. Network hashrate measures the total combined processing power of the entire Bitcoin network. Network difficulty is used to measure the difficulty of the algorithm solved by miners in an effort to add a block to the blockchain. As the network hashrate increases, the network difficulty is adjusted to ensure that Bitcoin miners are able to mine a Bitcoin block every ten minutes, on average. Using both of these measures will show the effects of how BTC miners react to price fluctuations and thereby examining how the network adjusts to these changes.

The network hashrate is also a measure of how often new BTC are created and circulated. Since the difficulty is adjusted to ensure a consistent circulation of new Bitcoins, one would expect to see barely any effect of the hashrate on total BTC but should see a positive effect on difficulty as the hashrate rises[3]. Knowing that network hashrate is a measure of the entire processing power of the BTC network also identifies some cost measure on the miners. If the hashrate is a function of some cost, theoretically electricity costs and computer equipment, to increase a miners hashrate would mean higher electricity costs and more mining equipment. Since it is in the best interest of miners to have the highest possible hashrate compared to

[3]See the Appendix for a proposed theoretical model for measuring supply-side behavior

the entire network, one would not expect the hashrate to fall unless it is no longer profitable for the miners to mine BTC and exit the network.

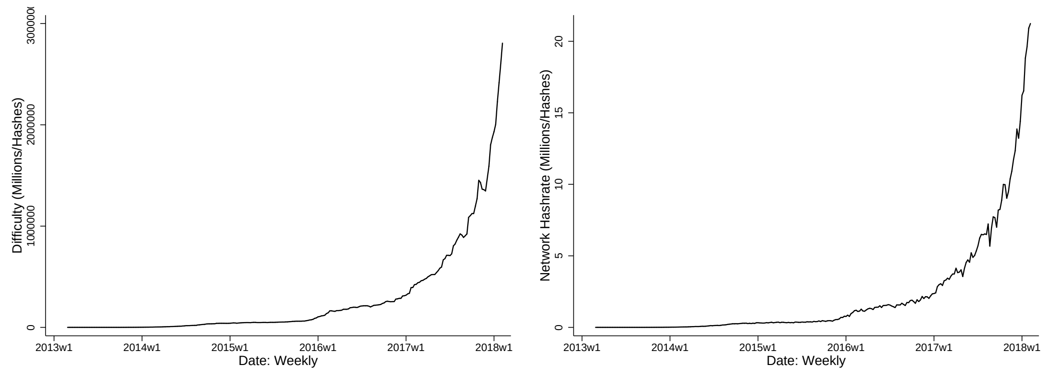

Figure 3.2: Network Hashrate & Difficulty

In terms of the difficulty, the difficulty with rise or fall solely dependent on how quickly a block can be solved. The software controlling the BTC network restricts a block to be solved every ten minutes, on average, and therefore the difficulty will adjust to ensure this. Changes in these two measures, network hashrate and difficulty, should lead to some increase or decrease in the transaction costs. Since transaction costs are determined by the sender, if a BTC miner is attempting to send BTC they may impose no cost if they have faith that they are able to solve the block in which they are adding their transaction, where as a user attempting to send to another user must choose a transaction cost that would incentivize miners to include their transaction in the block they are attempting to mine.

Transaction costs or cost per transaction is measured in USD and is calculated by dividing the total transaction fees by the total number of transactions of Bitcoin for that week. This measures the average price senders are willing to pay for their transactions to go through, over time. Since a miner is only able to include a certain amount of transactions in one block, as the number of transactions increase senders must offer higher fees so their transaction is more attractive to miners. Since these transaction fees are initially paid in fractions of Bitcoin rather than in a set amount of

USD, it would be expected that these fees increase as the price of purchasing a Bitcoin increases, and vice versa. It can also be expected that as the network difficulty rises the cost of mining will also increase leading to an increase in the cost per transaction.

The next set of variables are associated with the popularity of Bitcoin. These variables help test whether the popularity behind Bitcoin is a leading cause in the shifts in price. If these variables are found to be significant in Bitcoin price formulation then this may answer whether demand and supply interactions are important in determining Bitcoin's price. Economic demand and supply side fundamentals would predict that as Bitcoin demand increases with a fixed supply, the price of Bitcoin should rise. Therefore, there is an expectation that a positive relationship between the popularity measures and BTC price exists. To distinguish between fundamentals and speculation, an examination of the causal effects of these variables is analyzed, looking to see if there is a feedback loop between the two. If Bitcoin price and the popularity of Bitcoin cause each other, speculation may be the main explanation behind price shifts.

The three measures of popularity used to explain any speculation in price formation are search trends on Google[4] associated with Bitcoin and Bitcoin price, the number of transactions on the Bitcoin network, excluding exchanges, and the number of new unique addresses on the BTC network. The reasoning behind this is to measure those that are interested in BTC and then entered the market to buy, sell, or mine BTC. The combination of Google trends and number of unique addresses on the BTC network will provide a measure to distinguish between simple searches and those that have such interest to enter the marketplace.

The number of transactions measures the amount of transactions that have been confirmed by the BTC network, excluding popular addresses. It would be expected that as the demand or popularity rises for BTC then the number of transactions will

[4] Data report by trends.google.com

also increase. The number of transactions may also be tied to the number of unique addresses, where an increase in the number of unique addresses in the BTC market will also increase the number of transactions that take place. Another relationship that is expected ties in with the cost of the transactions. As the number of transactions increases an increase in the cost per transaction should also be prevalent. The rationale behind this is senders of BTC want to prioritize their transaction over other transactions and therefore will offer a higher price to ensure their transaction is sent.

The last two sets of variables deal with global commodity prices and global financial market indicators, respectively. The global commodities used are gold prices, oil prices, and natural gas prices, all measured in USD. These are used to test if Bitcoin prices moves with any of these commodity prices and also if the supply-side of Bitcoin creation is effected by oil and natural gas prices due to the cost of mining Bitcoin.

If the cost of mining BTC is a function of the hashrate then it could be expected that fluctuations in oil and natural gas prices, commodities used in the production of electricity, could cause fluctuations in the cost of mining BTC. One may see this cost translate to BTC price. Gold may have an effect on Bitcoin price because of the risk-profile of investors. Commodity prices tend to be more volatile than other assets and investors may adjust their portfolio by selling or buying gold and selling or buying Bitcoin.

The global financial indicators are used to explain any price fluctuations in Bitcoin associated with price fluctuations in global market indices. The financial indicators used are the Dow Jones Industrial Index (DJII), Nikkei 225 Index (N225), Hang-Seng Index (HSI), and the Shanghai Composite Index (SHCOMP). These indicators were used because of where the majority of BTC transactions and miners are located, with China accounting for approximately 60% of all Bitcoins mined[5].

The DJII is used because it is representative of the global financial market, as a

[5]Information taken from Blockchain.

whole. The other three financial indices are related to Asia-based financial markets in Japan and China. This is because BTC is heavily used and mined in these countries and fluctuations in these markets may have more of an effect on BTC than the DJII.

3.2 Data

The data are measured over five years on a weekly basis between 8Feb2013 and 11Feb2018. Table 3.2 describes the summary statistics associated with each of the variables:

Table 3.2: Summary Statistics

		Level				Log			
Variable	Units	Mean	σ	Min	Max	Mean	σ	Min	Max
BTC Price	$	1381.71	2836.30	29.53	17114.13	6.30	1.20	3.39	9.75
Total Bitcoin	mil/BTC	14.30	1.81	10.81	16.85	2.65	0.13	2.38	2.82
Net. Hashrate	mil/hashes	2.13	3.81	.0000327	21.24	-1.52	3.36	-10.33	3.06
Net. Difficulty	mil/hashes	275485.9	482157.6	3.65	2806773	10.24	3.40	1.30	14.85
Cost per Trans.	$	1.45	4.96	0.30	41.12	-1.64	1.63	-3.50	3.72
Google Trend	Indexed	7.18	13.38	1	100	1.34	0.93	0	4.61
# of Trans.	k/trans.	157.40	94.81	33.50	399.62	11.76	0.66	10.42	12.90
# of Unique Add.	k/add.	304.75	194.86	41.48	931.54	12.38	0.76	10.63	13.74
Gold	$	1259.37	95.27	1060	1608.64	7.14	0.07	6.97	7.38
Oil	$	70.33	27.30	27.42	116.09	4.18	0.38	3.31	4.75
Natural Gas	$	3.26	0.86	1.60	7.37	1.15	0.26	0.47	2.00
Dow Jones	$	18093.22	2562.32	13925.71	26337.41	9.79	0.13	9.54	10.18
Nikkei 225	$	17373.08	2650.12	11439.64	23836.53	9.75	0.15	9.34	10.08
Hang Seng	$	23739.87	2589.04	18916.98	32818.27	10.07	0.10	9.85	10.40
SHCOMP	$	2902.72	650.441	1967.093	5064.33	7.94	0.22	7.58	8.53

Looking at the dependent variable, BTC Price, there is large variation between the minimum and maximum values over the five year time span. The average price of BTC during this time is $1381.71 with a standard deviation of $2836.30 and the minimum and maximum prices being $29.53 and $17114.13, respectively.

The next set of variables measures network statistics. Total bitcoin in circulation is measured in millions of Bitcoin with a mean of 14.3 million in circulation and a standard deviation of 1.81 million over the five years. During the beginning of the sample there were 10.81 million BTC in circulation and the sample ends with 16.85 million BTC. Network hashrate is measured in millions of tera hashes per second (TH/s) with an average hashrate of 2.13 million and a standard deviation of 3.81 million. The minimum network hashrate in the sample is 32.7 TH/s and the maximum is 21.24 million TH/s. Since the network hashrate measures the entire processing power of the BTC network, this low number can be explained in two ways:

either there were a small number of miners in the BTC network during the early years of BTC and/or the processing power in early 2013 was much less than that of today.

Network difficulty also has a large amount of variation between the beginning and ending values in the sample. The network difficulty is measured in millions of hashes, with a minimum value of 32.7 hashes and maximum of 2.81 trillion hashes, the average being 275 billion for the timespan of the sample. The cost per transaction is measured in USD with an average of $1.45 per transaction, the minimum being $0.30 and the maximum being $41.12.

The next set of variables, popularity measures, includes trends in Google searches. This variable is measured as an index from 1 to 100, 100 being the most searches during a week and 1 being the least searches during a week. The average popularity is 7.18 with the minimum being 1 and the maximum being 100. Number of transactions is measured in thousands of transactions and has an average of approximately 157,000 transactions with a minimum of 33,500 and a maximum of 399,620. The number of unique addresses is also measured in thousands, with an average of 304,750 unique addresses, minimum of 41,480 and a maximum of 931,540.

Commodity measures are measured in USD and show much more stable variations. The average gold price for the timespan of the sample is $1259.37 with a minimum of $1060 and maximum of $1608.64. Oil price is the global price of Brent crude, reported by the St. Louis Federal Reserve Economic Data (FRED) and is measured in USD dollar price per barrel. It has an average price of $70.33 with a minimum price of $27.42 and a maximum price of $116.09. Natural gas is measured by the Henry Hub Natural Gas Spot Price as reported by FRED, which is measured in dollars per million BTU. The average price of the sample is $3.26 with a minimum price of $1.60 and maximum of $7.37.

Financial market indicators are all measured in USD and are taken from the FRED. The first of the financial market indicators is the Dow Jones Industrial Average (DJII).

The average price of the DJII is \$18,093.22 with a minimum of \$13,925.71 and a maximum reaching \$26,337.41 towards the end of the sample period. The Nikkei 225 (NI225) is the stock market index for Tokyo, Japan. The average price of the NI225 is \$17,373.08 with a minimum of \$11,439.64 and a maximum of \$23,836.53 during the sample. The Hang Seng Index (HSI) is a market index which includes major companies specifically in Hong Kong, China. The HSI had an average price of \$23,739.87 with a minimum of \$1,816.98 and a maximum of \$32,818.27. Lastly, the Shanghai Composite Index (SHCOMP) is an index which is widely used to represent the entire Chinese economy, much like the DJII or S&P 500 in the United States. The average price of the SHCOMP during the sample period is \$2902.72 with a minimum of \$1,967.093 and a maximum of \$5,064.33.

Table 3.2, also includes the summary statistics for the log transformation of the variables. The log transformations will be used to reduce some of the skewness in the data and make interpretations of the results easier. When using a log-log model, a percent change in the independent variable will lead to a percent change in the dependent variable i.e., an elasticity. This has a much easier interpretation then, for example, attempting to convert changes in the billions of hashes to changes in dollars of BTC price.

3.2.1 Testing for Structural Breaks

As seen from the summary statistics (Table 3.2), there are large variations in BTC price and other variables associated with the BTC network. Using time series data, this variation over time may cause breaks in the structure of the data, leading to unreliable tests, test statistics, coefficients, and an overall model. To counter any potential breaks, first detrend the data by regressing each variable on different measures of time. Initially, a standard time trend is utilized and further tests determine whether any further time trends may be causing issues with the data, such as an exponential time trend. After the data have been detrended, two different tests are used to test whether there are structural breaks: a test whether the coefficients are stable over time and a supremum Wald test. Both of these tests are used to determine if there is a structural break in the data at an unknown break date and examines the entire sample period. If the log of a variable has a structural break, convert the variable to log first-difference form and then retest the variable. The results of the Wald test on different forms of each variable are listed in Table 3.3 and the graphs associated with the stability test of the coefficients are included in the Appendix.

The null hypothesis for this test is there is no structural break in the data, the alternative being a structural break at a given time (t). From Table 3.3, every variable has a structural break at different times using the log form of the variables. When converting the variables to log first-difference forms, these structural breaks drop out and we fail to reject the null hypothesis at the 5% level in each case.

3.2.2 Testing for Stationarity

To ensure the model and test statistics are accurate, it is proper to first test for stationarity in the data. Each model relies on the assumption that the data is stationary, or the mean and variance do not vary with time. If these assumptions are violated the regression may be spurious and the resulting model is inaccurate. The

Table 3.3: Wald Test: Unknown Break Date

Variable	Log Break Date	D.Log Break Date
BTC Price	2017w18***	None
Total Bitcoin	2015w9***	None
Net. Hashrate	2014w8***	None
Net. Difficulty	2014w9***	None
Cost per Trans.	2017w20***	None
Google Trend	2017w19***	None
# of Trans.	2015w36***	None
# of Unique Add.	2015w7***	None
Gold	2013w47***	None
Oil	2014w49***	None
Natural Gas	2014w52***	None
Dow Jones	2016w49***	None
Nikkei 225	2014w45***	None
Hang Seng	2017w20***	None
SHCOMP	2014w48***	None

* p-value $<$ 0.10 ** p-value $<$ 0.05 *** p-value $<$ 0.01

most accurate measure to test whether the data is stationary is the DF-GLS test using ERS critical values (Elliot, Rosenberg, Stock, 1996). To ensure the data are stationary at a high-order, ten lags for each variable are included.

The null hypothesis for the DF-GLS test is that the variable contains a unit root, the alternative being that there is no unit root. The *tau* statistic is compared to critical values of 3.840 for the 1% level and 2.890 for the 5% level. If the absolute value of the *tau* statistic is greater than the absolute value of the critical values we can then reject the null hypothesis that the series contains a unit root. From Table 3.4, we fail to reject the null in log form but we can reject the null at the 1% level using first-differences of the log. To control for non-stationarity and unit root problems, first-differences will be used in the price models developed below.

3.3 Methodology

Since the tests show that the data are structured in a way to provide consistent and reliable estimates, a discussion on how to model the data is necessary. Since the goal

Table 3.4: DF-GLS Test for Stationarity

	Tau Statistic	
Variable	Log	D.Log
BTC Price	-1.525	-5.728
Total BTC	-0.867	-3.256
Net. Difficulty	0.014	-4.660
Net. Hashrate	0.099	-6.480
Trans. Cost	-0.858	-7.763
Google Trend	-2.472	-9.095
# of Trans.	-2.275	-9.092
# of Unique Add.	-1.405	-9.130
Gold	-1.227	-9.115
Oil	-1.050	-7.424
Natural Gas	-2.266	-7.348
Dow Jones	-2.012	-6.251
Nikkei 225	-1.740	-8.116
Hang Seng	-1.752	-6.765
SHCOMP	-1.789	-6.957

Ten lags were used for each of the variables.
D.Log represents the first-difference of the variable.

is to measure the causal effects of the explanatory variables on BTC price, the models should be structured in a way to capture these effects. One way to measure causal effects is a Granger-Causality test, first established by Granger (1969). If past levels in the independent variables tells a story about future values of the dependent variable then the independent variables are said to "Granger-cause" that variable. This would mean that the model should treat the dependent variable as a function of a certain number of lagged terms of itself and the independent variables. To look at how each of the variables interact with each other, a model that treats all variables as endogenous will be used. To do this, a system of equations will be constructed that will be solved simultaneously. To construct this system of equations one must define the equation for the dependent variable, BTC price. The equation to define the general form of the model for BTC price as a function of a subset of the variables which represent: (i) network statistics (ii) popularity (iii) commodities and (iv) financial indices. This form of equation is called an Autoregressive Distributed Lag (ADL) model, taking the form shown in Equation 3.1.

$$P_t = \beta_0 + \beta_1 P_{t-p} + \delta_h \mathrm{NS}_{t-p} + \gamma_i \mathrm{POP}_{t-p} + \psi_j \mathrm{COMM}_{t-p} + \phi_k \mathrm{FI}_{t-p} + \mu_t \tag{3.1}$$

Where P_t represents the BTC price at time t, NS represents the variables associated with network statistics, POP represents the variables associated with popularity measures, COMM represents the variables associated with commodities, and FI represents the variables associated with global financial indices.

To convert this ADL model into a system of equations to represent a VAR(p) model, matrix algebra can be used (Lutkepohl, 2005).

$$\mathbf{y}_t = \mathbf{A}\mathbf{Y}_{t-1} + \mathbf{B}_0\mathbf{x}_t + \mathbf{u}_t \tag{3.2}$$

Where y_t is a $(K \times 1)$ vector of endogenous variables, A is the coefficient matrix for the endogenous variables which contains $(K \times K_p)$, $\mathbf{B}_0$ is a $(K \times M)$ which contains the coefficients for the exogenous variables. $\mathbf{x}_t$ is a $(M \times 1)$ vector of exogenous variables. $\mathbf{u}_t$ is a $(K \times 1)$ vector containing white noise innovations. $\mathbf{Y}_t$ is a $(K_p \times 1)$ matrix of the form,

$$\mathbf{Y}_t = \begin{bmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix}$$

Once the initial set of models are created, a test to see if there is serial correlation in the residuals and a test to examine these residuals to look for Autoregressive Conditional Heteroskedastic (ARCH) effects is appropriate. If these effects are found, a multivariate GARCH approach to adjust for these effects should be used. The multivariate GARCH approach (MGARCH) uses a Gaussian likelihood estimation by first using least squares estimation to fit the VAR(p) model then fitting the GARCH model to the residuals at the first stage of estimation (Lutkepohl, 2005). When using an

MGARCH model, the adjustments made to the model are either dynamic or constant over time. These two types of MGARCH models are Constant Conditional Correlation (CCC) or Dynamic Conditional Correlation (DCC)[6]. The MGARCH CCC model states that adjustments to the estimates are based on constant correlations between the variable's residuals over time where the DCC model states that this is a dynamic factor and these correlations changes with time. To determine which model should be used, a Wald test for significance can be done on the λ adjustments of the model. These λ adjustments describe how the correlations between the residuals of the variables changes over time. If these adjustments are no different than zero, a CCC MGARCH model can be chosen. The DCC MGARCH model can be written in general form by the following equations (Engle, 2002).

$$\mathbf{Y}_t = \mathbf{C}\mathbf{X}_t + \epsilon_t \tag{3.3}$$

$$\epsilon_t = \mathbf{H}_t^{1/2}\mathbf{v}_t \tag{3.4}$$

$$\mathbf{H}_t = \mathbf{D}_t^{1/2}\mathbf{R}_t\mathbf{D}_t^{1/2} \tag{3.5}$$

$$\mathbf{R}_t = \text{diag}(\text{Q})_t^{-1/2}\mathbf{Q}_t\text{diag}(\mathbf{Q})_t^{-1/2} \tag{3.6}$$

$$\mathbf{Q}_t = (1-\lambda_1-\lambda_2)\mathbf{R} + \lambda_1\widetilde{\epsilon}_{t-1}\widetilde{\epsilon}'_{t-1} + \lambda_2\mathbf{Q}_{t-1} \tag{3.7}$$

$\mathbf{Y}_t$ is a $(M \times 1)$ vector of the dependent variables, $\mathbf{C}$ is an $(M \times K)$ matrix for the parameter coefficients. $\mathbf{X}_t$ is a $(K \times 1)$ vector that contains the independent variables. $\mathbf{H}_t^{1/2}$ is the Cholesky factor of the time-varying conditional covariance matrix $\mathbf{H}_t$. $\mathbf{v}_t$ is a $(M \times 1)$ vector which contains normal, i.i.d innovations. The conditional variances

[6]For more information on MGARCH CCC & DCC see Aielli(2009), Engle (2002), and Lutkepohl, 2005

are described by the matrix $\mathbf{D}_t$ that takes the form,

$$\mathbf{D}_t = \begin{bmatrix} \sigma_{1,t}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{2,t}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{m,t}^2 \end{bmatrix}$$

where each $\sigma_{i,t}^2$ is modeled using a GARCH form of

$$\sigma_{i,t}^2 = \exp(\gamma_i \mathbf{z}_{i,t}) + \sum_{j=1}^{p_i} \alpha_j \epsilon_{i,t-j}^2 + \sum_{j=1}^{q_i} \beta_j \sigma_{i,t-j}^2 \tag{3.8}$$

where γ_t is a $(1\times p)$ vector of parameters, $\mathbf{z}_i$ is a $(p\times 1)$ vector of independent variables used to model the heteroskedasticity within each of the dependent variables, and the α_j and β_j's are the ARCH and GARCH parameters, respectively. The matrix which represents the conditional quasicorrelations takes the following form,

$$\mathbf{R}_t = \begin{bmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1m,t} \\ \rho_{12,t} & 1 & \cdots & \rho_{2m,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1m,t} & \rho_{2m,t} & \cdots & 1 \end{bmatrix}$$

$\widetilde{\epsilon}_t$ is a $(M \times 1)$ vector of the standardized residuals computed from the conditional variance matrix, $\mathbf{D}_t^{-1/2}\epsilon_t$. The λ_1 and λ_2 adjusted parameters are used to model the dynamics of the conditional quasicorrelations (Aielli, 2009, Engle, 2002).

CHAPTER 4: RESULTS

The first model is a VAR(1) model for which all variables are treated as endogenous, including the commodity prices and financial market indicators. This model is used to find any interesting relationships that may be found, such as BTC popularity measures lead to price changes in financial indicators and/or commodity prices. The results of this model are displayed in Table 4.1 and Table 4.2 below. At first glance, one can see that the commodity variables and financial market variables have no effect on BTC price and very little role in explaining anything about the BTC network. BTC network statistics and popularity measures also have very little effect on commodity and financial market variables. This result is consistent with prior studies and provides evidence that both commodities and financial market variables should be treated as exogenous.

Table 4.1: VAR(1) Model Results (a)
All Variables

		Network Statistics				Popularity		
Variable	L.ΔPrice	L.ΔT. BTC	L.ΔHashrate	L.ΔDiff.	L.ΔT. Cost	L.ΔTrend	L.Δ# Trans.	L.Δ# Add.
Price	0.258**	35.389	-0.543	0.120	0.011	0.058**	0.054	-0.073
	(0.123)	(46.237)	(0.136)	(0.161)	(0.107)	(0.023)	(0.141)	(0.083)
Total BTC	-0.0004***	0.649***	-0.0001	0.0001	0.0004***	0.00001	0.0002	0.0001
	(0.0001)	(0.050)	(0.0001)	(0.0002)	(0.0001)	(0.00003)	(0.0002)	(0.00009)
Hashrate	-0.196	-4.032	-0.407***	0.273**	0.143*	0.004	0.080	0.051
	(0.916)	(34.383)	(0.101)	(0.120)	(0.080)	(0.017)	(0.105)	(0.062)
Difficulty	-0.167***	135.765***	-0.009	-0.068	0.102**	0.006	0.052	0.038
	(0.053)	(19.842)	(0.059)	(0.069)	(0.046)	(0.010)	(0.061)	(0.036)
Trans. Cost	-0.361**	-97.059	-0.921***	1.078***	0.511***	0.019	0.931***	-0.080
	(0.167)	(62.796)	(0.185)	(0.219)	(0.146)	(0.032)	(0.192)	(0.113)
Google Trend	0.400	139.263	0.050	0.657	0.231	-0.152**	0.380	-0.090
	(0.336)	(126.228)	(0.372)	(0.440)	(0.293)	(0.064)	(0.385)	(0.227)
# Trans.	0.121	-58.391	0.070	0.058	0.049	0.039*	-0.286**	-0.023
	(0.120)	(45.056)	(0.133)	(0.157)	(0.105)	(0.023)	(0.137)	(0.081)
# Add.	-0.025	-29.593	-0.072	0.334**	0.216**	0.068***	0.288**	-0.557***
	(0.120)	(44.867)	(0.132)	(0.156)	(0.104)	(0.023)	(0.137)	(0.081)
Gold	-0.051**	9.759	-0.035	0.033	0.045	-0.004	0.030	0.008
	(0.022)	(8.188)	(0.024)	(0.029)	(0.019)	(0.004)	(0.025)	(0.015)
Oil	-0.033	-21.992	0.055	0.025	-0.005	-0.004	0.037	0.013
	(0.047)	(17.687)	(0.052)	(0.062)	(0.041)	(0.009)	(0.054)	(0.032)
N. Gas	-0.186*	84.685**	-0.185*	0.048	0.157*	0.004	0.012	0.041
	(0.100)	(37.711)	(0.111)	(0.131)	(0.088)	(0.019)	(0.115)	(0.068)
DJII	0.002	1.202	0.016	-0.011	-0.001	-0.002	0.009	-0.003
	(0.017)	(6.373)	(0.019)	(0.022)	(0.015)	(0.003)	(0.019)	(0.011)
N225	0.013	0.022	0.014	-0.033	-0.023	0.001	-0.003	0.012
	(0.028)	(10.580)	(0.031)	(0.037)	(0.025)	(0.005)	(0.032)	(0.019)
HSI	0.003	9.789	0.011	-0.054*	-0.011	0.001	-0.008	0.006
	(0.024)	(8.955)	(0.026)	(0.031)	(0.021)	(0.005)	(0.027)	(0.016)
SHCOMP	-0.001	16.097	-0.003	-0.071	-0.002	0.00003	(0.043	-0.015
	(0.035)	(13.037)	(0.038)	(0.045)	(0.030)	(0.007)	(0.040)	(0.023)

Standard errors are represented by parenthesis below the coefficients
* p-value $<$ 0.10 ** p-value $<$ 0.05 *** p-value $<$ 0.01

To ensure that the results for this model are accurate, a test for serial correlation and ARCH/GARCH effects in the residuals is needed. When testing for serial correlation, the Breusch-Godfrey test, which can detect high-order serial correlation within the residuals, and the Engle's Lagrange Multiple Test (ARCH LM) to test for ARCH effects. The results of the tests are display in Table 4.3.

Table 4.2: VAR(1) Model Results (b)
All Variables

	Commodities			Financial Markets			
Variable	L.ΔGold	L.ΔOil	L.ΔN. Gas	L.ΔDJII	L.ΔN225	L.ΔHSI	L.ΔSHCOMP
Price	-0.417	0.222	0.076	0.512	-0.330	0.593	-0.240
	(0.371)	(0.165)	(0.074)	(0.675)	(0.386)	(0.468)	(0.251)
Total BTC	0.00003	0.0002	0.00005	0.0009	-0.0009**	0.0001	0.0002
	(0.0004)	(0.0002)	(0.00008)	(0.0007)	(0.0004)	(0.0005)	(0.0003)
Hashrate	-0.125	-0.076	-0.024	-0.241	-0.271	0.159	0.113
	(0.276)	(0.123)	(0.055)	(0.502)	(0.287)	(0.348)	(0.187)
Diffuclty	-0.165	0.012	-0.036	0.544*	-0.178	-0.061	-0.002
	(0.159)	(0.071)	(0.032)	(0.289)	(0.166)	(0.201)	(0.108)
Trans. Cost	-0.745	0.359	0.088	-0.330	0.196	0.031	0.139
	(0.504)	(0.224)	(0.010)	(0.916)	(0.524)	(0.635)	(0.342)
Google Trend	-0.532	0.032	0.100	0.122	-0.301	1.802	-0.481
	(1.012)	(0.450)	(0.201)	(1.842)	(1.053)	(1.277)	(0.687)
# Trans.	0.040	-0.158	-0.018	-0.323	-0.507	0.630	-0.210
	(0.361)	(0.161)	(0.072)	(0.657)	(0.376)	(0.456)	(0.245)
# Add.	-0.727**	-0.226	0.056	-0.275	-0.671*	1.044**	-0.213
	(0.360)	(0.160)	(0.071)	(0.655)	(0.374)	(0.454)	(0.244)
Gold	0.236***	-0.046	0.018	-0.077	-0.044	0.248***	-0.009
	(0.066)	(0.029)	(0.013)	(0.119)	(0.068)	(0.083)	(0.045)
Oil	-0.147	0.324***	-0.009	0.066	-0.266*	0.213	-0.155
	(0.142)	(0.063)	(0.028)	(0.258)	(0.148)	(0.179)	(0.096)
N. Gas	-0.916***	0.277**	-0.224***	-0.978*	0.064	0.602	-0.136
	(0.302)	(0.135)	(0.060)	(0.550)	(0.315)	(0.381)	(0.205)
DJII	-0.001	0.031	0.002	0.267***	-0.071	-0.020	-0.003
	(0.051)	(0.023)	(0.010)	(0.093)	(0.053)	(0.064)	(0.035)
N225	-0.235***	0.061	-0.004	0.354**	0.035	-0.084	0.018
	(0.085)	(0.038)	(0.017)	(0.154)	(0.089)	(0.107)	(0.058)
HSI	-0.043	0.046	0.002	0.061	-0.080	0.246***	0.059
	(0.072)	(0.032)	(0.014)	(0.131)	(0.075)	(0.091)	(0.049)
SHCOMP	-0.013	-0.008	-0.019	0.014	0.125	-0.018	0.271***
	(0.105)	(0.047)	(0.021)	(0.190)	(0.109)	(0.132)	(0.071)

Standard errors are represented by parenthesis below the coefficients
* p-value < 0.10 ** p-value < 0.05 *** p-value < 0.01

Table 4.3: BG & ARCH LM Test:
VAR(1) - All Factors

Lags	BG Test Prob $> \chi^2$	ARCHLM Prob $> \chi^2$
1	0.0351	0.0000
2	0.0786	0.0000
3	0.1455	0.0000
4	0.1842	0.0000
5	0.2819	0.0000
6	0.3634	0.0000
7	0.4756	0.0000
8	0.5778	0.0000
9	0.6759	0.0000
10	0.6461	0.0000

The null hypothesis for the Breusch-Godfrey Test is that there is no serial correlation at the lag order specified. From the results, we fail to reject the null hypothesis at all lags except for lag one, where there is sufficient evidence of serial correlation. For the ARCH LM Test, the null hypothesis is that there are no ARCH effects at the specified lag order. From the results, we reject the null hypothesis at all lag orders, providing sufficient evidence that there are ARCH effects within the residuals. Since serial correlation and a changing variance due to the ARCH effects can provide biased estimates, inaccurate test statistics, and an overall inaccurate model, this model cannot be accepted as an appropriate model for BTC price.

As the previous models suggest, treating commodity prices and financial markets as exogenous may provide a more accurate model for explaining BTC price formation. Table 4.4 and 4.5 below provides the results of the VAR(1) model, using commodity prices and financial markets as having exogenous effects on BTC price, network statistics, and BTC popularity measures. As with the previous model, tests for serial

Table 4.4: VAR(1) Model Results (a):
Exog. Factors

	Endogenous							
	Network Statistics					Popularity		
Variable	L.ΔPrice	L.ΔT. BTC	L.ΔHashrate	L.ΔDiff.	L.ΔT. Cost	L.ΔTrend	L.Δ# Trans.	L.Δ# Add.
Price	0.288**	27.532	-0.058	0.118	-0.011	0.055**	0.023	-0.081
	(0.123)	(45.612)	(0.137)	(0.162)	(0.108)	(0.023)	(0.141)	(0.083)
Total BTC	-0.0004***	0.662***	-0.00008	-0.000007	0.0003***	0.00002	0.0002	0.0001
	(0.0001)	(0.049)	(0.0001)	(0.0002)	(0.0001)	(0.00002)	(0.0001)	(0.0001)
Hashrate	-0.157*	-15.770	-0.370***	0.218*	0.093	0.010	0.040	0.055
	(0.090)	(33.627)	(0.101)	(0.119)	(0.079)	(0.017)	(0.104)	(0.061)
Difficulty	-0.156***	138.850***	-0.002	-0.093	0.097**	0.006	0.053	0.031
	(0.053)	(19.664)	(0.059)	(0.070)	(0.046)	(0.010)	(0.061)	(0.036)
Trans. Cost	-0.313*	-108.561*	-0.937***	1.047***	0.486***	0.018	0.828***	-0.064
	(0.165)	(61.536)	(0.184)	(0.218)	(0.145)	(0.031)	(0.190)	(0.112)
Google Trend	0.459	110.987	0.074	0.584	0.154	-0.164***	0.348	-0.096
	(0.334)	(124.348)	(0.373)	(0.441)	(0.293)	(0.063)	(0.383)	(0.226)
# Trans.	0.132	-78.121*	0.114	0.051	0.007	0.045**	-0.257*	-0.035
	(0.119)	(44.303)	(0.133)	(0.157)	(0.104)	(0.023)	(0.136)	(0.081)
# Add.	0.011	-55.588	-0.006	0.312**	0.152	0.070***	0.283**	-0.569***
	(0.120)	(44.675)	(0.134)	(0.158)	(0.105)	(0.023)	(0.138)	(0.081)

Standard errors are represented by parenthesis below the coefficients
* p-value < 0.10 ** p-value < 0.05 *** p-value < 0.01

correlation and ARCH/GARCH effects are conducted. Table 4.6 shows the results of that test for the exogenous factor model.

The results of these tests show that serial correlation has been accounted for, failing to reject the null hypothesis of no serial correlation at all lags. What has not been controlled for are the ARCH/GARCH effects persistent in the model. To control these ARCH effects a MGARCH model, specified earlier, will be used to adjust the model, variances, and standard errors accordingly. The first of the MGARCH models will

Table 4.5: VAR(1) Model Results (b):
Exog. Factors

	Exogenous						
	Commodities			Financial Markets			
Variable	ΔGold	ΔOil	ΔN. Gas	ΔDJII	ΔN225	ΔHSI	ΔSHCOMP
Price	0.325	0.0117	0.023	0.742	0.539	-0.630	0.224
	(0.367)	(0.166)	(0.073)	(0.656)	(0.386)	(0.458)	(0.249)
Total BTC	0.0011***	0.0001	-0.00002	0.0010	0.0003	-0.001***	0.0002
	(0.0004)	(0.0002)	(0.0001)	(0.0007)	(0.0004)	(0.0005)	(0.0003)
Hashrate	0.473*	-0.032	0.027	0.642	-0.295	-0.431	0.490
	(0.270)	(0.122)	(0.054)	(0.483)	(0.285)	(0.338)	(0.184)
Difficulty	-0.024	-0.033	0.018	-0.315	0.181	0.069	-0.154
	(0.158)	(0.071)	(0.032)	(0.283)	(0.167)	(0.197)	(0.107)
Trans. Cost	0.604	0.367	-0.053	0.754	0.628	-1.558**	0.624*
	(0.494)	(0.224)	(0.099)	(0.884)	(0.521)	(0.618)	(0.336)
Google Trend	0.067	-0.567	-0.072	1.822	-0.582	-0.730	-0.322
	(0.999)	(0.452)	(0.200)	(1.787)	(1.053)	(1.249)	(0.680)
# Trans.	0.328	-0.223	0.115	0.832	-0.432	0.382	-0.351
	(0.356)	(0.161)	(0.071)	(0.637)	(0.375)	(0.445)	(0.242)
# Add.	0.186	-0.117	0.123*	0.384	0.243	-0.392	-0.118
	(0.359)	(0.162)	(0.072)	(0.642)	(0.378)	(0.449)	(0.244)

Standard errors are represented by parenthesis below the coefficients
* p-value < 0.10 ** p-value < 0.05 *** p-value < 0.01

take the same form as the previous VAR(1) model, treating commodity prices and financial market indicators as exogenous factors that effect the model. The results of this MGARCH model are displayed in Table 4.7 and Table 4.8.

When comparing the two models, the VAR(1) and MGARCH, slightly different results can be seen in both the coefficients and standard errors. Since Robust Standard Error measurements are used in the MGARCH model and ARCH/GARCH effects are also adjusted for, the MGARCH model should be chosen over the VAR(1) model. The order of the ARCH and GARCH effects for the MGARCH model is a MGARCH(1,1), where there is one ARCH term and one GARCH term. This order was selected using Schwartz-Bayesian Criteria (BIC), where the statistic is defined by:

$$\text{BIC} = -2\ln L + k\ln N \tag{4.1}$$

Where N is the sample size and L is the log-likelihood value of the model and the lowest BIC value represents the model that best fits the sample (Schwarz, 1978).

Table 4.6: BG & ARCH LM Test:
VAR(1) - Exog. Factors

Lags	BG Test Prob > χ^2	ARCHLM Prob > χ^2
1	0.9211	0.0000
2	0.6212	0.0000
3	0.8118	0.0000
4	0.8234	0.0000
5	0.9106	0.0000
6	0.8974	0.0000
7	0.9454	0.0000
8	0.9653	0.0000
9	0.9828	0.0000
10	0.9757	0.0000

The variable that is the main focus is the BTC price, or BTC returns in this case, where the only variable from the model that is effecting the returns of BTC deals with trends in Google. This is consistent with the previous VAR(1) models and the sign is positive, as would be expected. What is intriguing is that the returns of BTC do not effect changes in trends or popularity. Logic would suggest that as the returns rise, especially to the levels that have been seen recently, the popularity of BTC would rise along with it. Because no significance can be found in the trend variable, people may become more interested in BTC for other reasons that are not specified in the model. Since no other factors effect trends in Google besides prior trends in Google, there may be better uses for this popularity measure, such as using it to explain the volatility that is going on in Bitcoin's price. This will be the specification for the final model.

The model represented in Table 4.9 and 4.10 uses the current trends in Google and past weeks trends in Google to explain volatility in the network statistics and other network popularity measures. As can be seen, this trend variable is highly significant and positive for almost all of the network statistics, excluding the network hashrate and transaction costs. The lack of significance for the hashrate can easily be explained by what the hashrate is measuring, the processing power of the network. Since miners only care about confirming transactions and mining blocks to confirm those transactions, trends in the BTC network would have no effect on whether these miners increase or decrease their hashrate. Also, to be competitive with other BTC miners, miners would not have incentive to decrease their hashrate as long as transactions are waiting to be confirmed and the price of BTC is high enough to provide a profit. This logic is also bolstered by the fact that the ARCH and GARCH terms for the hashrate measure are both insignificant as well, further showing that volatility within the hashrate measure tend to be consistent.

Focusing now on BTC price, changes in trends on Google during the past week no longer have any significance in determining future returns for BTC price but are very significant when explaining volatility in BTC price. With both the past and current week changes in Google trends being positive and statistically significant at the 99% level, changes in Google trends cause increases in the volatility of BTC price but no longer have a casual effect on BTC price. What can be found when looking for changes in BTC price are that current returns in the Hang-Seng Index now explain variations in current returns of BTC, with significance at a 10% level. A 1% increase in the HSI leads to a -0.5865% decrease in returns for BTC. This may be due to several factors, such as investor preference or investment strategy. Investor's may sell BTC to purchase shares in the HSI when the return is high or market outlook is promising, or when BTC is too volatile for investors. Since the majority of BTC is mined out of the Chinese market, this may be large-scale mining operations selling

BTC to purchase shares in the HSI for a more strategic investment.

Digging further into the results of the model, interesting results on how this Google trend variable effects other popularity measures can be seen. Firstly, the number of addresses increases as the trends for BTC increase, along with the volatility in the number of addresses actively on the BTC network. A 1% increase in past changes in the Google Trend variable leads to approximately 0.046% increase in the number of unique addresses in the following week. This may seem like a small percentage but this change is roughly 13,866 unique users on average. To see if this change effects the number of transactions, one can see if the number of addresses on the network effects the number of transactions that take place.

As can be seen, the number of addresses does not effect the number of transactions that take place on the network which could be due to several reasons. For one, the transaction measure used in the data excludes BTC exchanges and other large BTC transaction hubs. These new addresses may be joining the network to purchase a few BTC and then hold onto these BTC in hopes that the price will continue to rise. This can be confirmed by the fact that the number of transactions and the Google trend variable are significant and positive at a 90% level of significance, but is positive and very significant when explaining volatility in the number of transactions.

When examining the cost of transactions, the BTC price, network difficulty, network hashrate, number of transactions, and the Hang-Seng Index all play a role in determining the transaction cost. Examining the coefficients of these significant variables further, a 1% increase in past network difficulty and network hashrate cause opposite and almost equal effects in the transaction cost. The signs for these variables are both within logic. When network difficulty rises, it is harder to mine a block and confirm transactions so the miners will require a higher transaction cost to include these transactions into their block. On the other hand, when the hashrate, which is a measure of the networks total processing power, rises it becomes easier for a miner to mine a block and therefore confirm a transaction, requiring less of a fee in the form of transaction costs. The number of transactions also plays a role in forming the transaction costs in the BTC network, where a 1% increase in the number of transactions increases the transaction costs by 0.823%. As the BTC network becomes flooded with transactions, senders of BTC will have to compete to have their transaction sent in a timely fashion, therefore offering a higher price to have their transaction confirmed.

The relationship between number of transactions and transaction costs are quite interesting. Logic would suggest there would be some sort of feedback loop associated with these two variables. As the transaction costs increase, the number of transactions on the network would decrease and therefore adjusting the transaction costs, but this is not the case. There are several reasons why this could happen. Since price has a positive effect on the number of transactions, though only significant at the 10% level, the rise in the transaction costs does not matter to an investor who wants to sell their BTC when they believe the price is high. Since BTC exchanges make up the majority of BTCs sent, they will offer the lowest fee possible to miners to have their transactions confirmed.

From the model, it looks as if past returns in Bitcoin may predict future returns in Bitcoin. Bitcoin price can be modeled as an AR(1) process with ARCH/GARCH effects using past and present changes in popularity to explain volatility. To confirm, the BIC of the AR(1) model of this form is compared to the BIC of the MGARCH model specified above.

Table 4.7: MGARCH Model Results (a):
Exog. Factors

		Endogenous						
		Network Statistics				Popularity		
Variable	L.ΔPrice	L.ΔT. BTC	L.ΔHashrate	L.ΔDiff.	L.ΔT. Cost	L.ΔTrend	L.Δ# Trans.	L.Δ# Add.
Price	0.120	39.423	-0.088	0.067	0.015	0.045*	0.074	-0.092
	[0.121]	[40.971]	[0.120]	[0.170]	[0.090]	[0.026]	[0.112]	[0.075]
ARCH	0.104**							
	[0.043]							
GARCH	0.591***							
	[0.120]							
Total BTC	-0.0002***	0.699***	0.00008	-0.000009	0.0002***	0.00002	0.00006	0.00014
	[0.00007]	[0.044]	[0.00009]	[0.0001]	[0.00007]	[0.00002]	[0.00012]	[0.00009]
ARCH	0.025							
	[0.024]							
GARCH	0.982***							
	[0.019]							
Hashrate	-0.092	-1.772	-0.323***	0.208*	0.072	0.007	0.015	0.047
	[0.079]	[30.436]	[0.091]	[0.108]	[0.074]	[0.015]	[0.102]	[0.051]
ARCH	0.062**							
	[0.030]							
GARCH	0.857***							
	[0.050]							
Difficulty	-0.291***	172.972***	-0.176*	0.004	0.216**	-0.003	0.174**	0.032
	[0.096]	[40.912]	[0.101]	[0.115]	[0.087]	[0.012]	[0.086]	[0.025]
ARCH	0.201*							
	[0.107]							
GARCH	0.799***							
	[0.062]							
Trans. Cost	-0.212	-187.541***	-0.764***	1.018***	0.464**	0.028	0.837***	-0.089
	[0.192]	[55.532]	[0.213]	[0.211]	[0.192]	[0.029]	[0.208]	[0.098]
ARCH	0.177							
	[0.116]							
GARCH	0.178							
	[0.109]							
Google Trend	0.581	-7.724	0.150	0.382	0.009	-0.239***	0.221	-0.062
	[0.386]	[133.770]	[0.372]	[0.513]	[0.338]	[0.075]	[0.373]	[0.120]
ARCH	0.125							
	[0.079]							
GARCH	0.702***							
	[0.253]							
# Trans.	0.254*	-59.491*	0.253*	-0.144	-0.202*	0.033*	-0.491***	-0.029
	[0.133]	[32.844]	[0.146]	[0.144]	[0.116]	[0.020]	[0.150]	[0.081]
ARCH	0.128**							
	[0.053]							
GARCH	0.763***							
	[0.108]							
# Add.	0.087	-42.426	0.078	0.053	0.006	0.039	0.118	-0.571***
	[0.114]	[32.732]	[0.126]	[0.158]	[0.095]	[0.024]	[0.130]	[0.106]
ARCH	0.093***							
	[0.043]							
GARCH	0.901***							
	[0.036]							
λ_1	0.053***							
	[0.011]							
λ_2	0.868***							
	[0.033]							

Robust Standard errors are represented by brackets below the coefficients
* p-value $<$ 0.10 ** p-value $<$ 0.05 *** p-value $<$ 0.01

Table 4.8: MGARCH Model Results (b):
Exog. Factors

	Exogenous						
	Commodities			Financial Markets			
Variable	ΔGold	ΔOil	ΔN. Gas	ΔDJII	ΔN225	ΔHSI	ΔSHCOMP
Price	-0.371	0.131	0.019	0.768	0.213	-0.435	0.166
	[0.429]	[0.120]	[0.073]	[0.532]	[0.394]	[0.421]	[0.176]
ARCH	0.104**						
	[0.043]						
GARCH	0.591***						
	[0.120]						
Total BTC	0.00075**	0.00008	-0.00003	0.0015***	-0.00002	-0.0016***	0.00001
	[0.00031]	[0.00013]	[0.00003]	[0.00048]	[0.00037]	[0.00039]	[0.00024]
ARCH	0.025						
	[0.024]						
GARCH	0.982***						
	[0.019]						
Hashrate	0.223	0.047	0.033	0.943**	-0.593**	-0.547*	0.034
	[0.273]	[0.110]	[0.045]	[0.440]	[0.238]	[0.326]	[0.160]
ARCH	0.062**						
	[0.030]						
GARCH	0.857***						
	[0.050]						
Difficulty	-0.167	-0.047	0.048	-0.149	0.141	0.153	-0.131***
	[0.130]	[0.071]	[0.032]	[0.205]	[0.197]	[0.152]	[0.043]
ARCH	0.201*						
	[0.107]						
GARCH	0.799***						
	[0.062]						
Trans. Cost	-0.066	0.327*	0.024	1.089	-0.051	-1.010**	0.184
	[0.633]	[0.178]	[0.085]	[0.813]	[0.393]	[0.477]	[0.317]
ARCH	0.177						
	[0.116]						
GARCH	0.178						
	[0.109]						
Google Trend	-0.782	-0.661	-0.252	1.625	-0.827	-0.178	-0.572
	[1.058]	[0.419]	[0.205]	[1.785]	[1.008]	[1.139]	[0.511]
ARCH	0.125						
	[0.079]						
GARCH	0.702***						
	[0.253]						
# Trans.	0.168	-0.104	-0.024	0.400	-0.256	0.174	-0.066
	[0.342]	[0.149]	[0.094]	[0.542]	[0.365]	[0.342]	[0.270]
ARCH	0.128**						
	[0.053]						
GARCH	0.763***						
	[0.108]						
# Add.	-0.303	-0.017	0.001	0.218	-0.184	-0.071	0.025
	[0.361]	[0.154]	[0.093]	[0.550]	[0.407]	[0.330]	[0.254]
ARCH	0.093***						
	[0.043]						
GARCH	0.901***						
	[0.036]						
λ_1	0.053***						
	[0.011]						
λ_2	0.868***						
	[0.033]						

Robust Standard errors are represented by brackets below the coefficients

* p-value < 0.10 ** p-value < 0.05 *** p-value < 0.01

Table 4.9: MGARCH Model Results (a): Conditional Variance

		Endogenous						
		Network Statistics				Popularity		
Variable	L.ΔPrice	L.ΔT. BTC	L.ΔHashrate	L.ΔDiff.	L.ΔT. Cost	L.ΔTrend	L.Δ# Trans.	L.Δ# Add.
Price	0.2283***	40.4654	-0.0388	0.0134	-0.0000	0.0353	0.0208	-0.0749
	[0.1018]	[30.0338]	[0.1013]	[0.1383]	[0.0699]	[0.0238]	[0.0870]	[0.0677]
ARCH	0.0400							
	[0.0296]							
GARCH	0.8576***							
	[0.0581]							
Google Trend	4.2716***							
	[0.6259]							
L.Google Trend	1.0259***							
	[1.5276]							
Total BTC	-0.00018	0.8018***	-0.000002	-0.00018	0.00016	0.000021	0.000007	0.00016*
	[0.00012]	[0.6190]	[0.00011]	[0.00016]	[0.0001]	[0.00002]	[0.00014]	[0.00009]
ARCH	0.3442							
	[0.2459]							
GARCH	0.4957**							
	[0.2032]							
Google Trend	-0.8894							
	[1.2028]							
L.Google Trend	4.9826***							
	[1.2799]							
Difficulty	-0.1178*	107.7139***	0.0243	-0.0636	0.0403	-0.0155	0.0283	-0.0022
	[0.0627]	[21.8118]	[0.0658]	[0.0798]	[0.0262]	[0.0603]	[0.0533]	[0.0113]
ARCH	0.0309							
	0.0375							
GARCH	0.9089***							
	0.0260							
Google Trend	5.2411***							
	0.8687							
L.Google Trend	2.7528***							
	[0.6085]							
Hashrate	-0.0184	-20.4090	-0.1971	0.0883	0.0893	-0.1115	-0.0572	0.0145
	[0.0974]	[30.1429]	[0.1211]	[0.1322]	[0.0563]	[0.1186]	[0.1003]	[0.0169]
ARCH	0.0711							
	[0.0654]							
GARCH	0.0460							
	[0.3572]							
Google Trend	-0.8152							
	[0.5008]							
L.Google Trend	0.1544							
	[0.4754]							
Trans. Cost	-0.2974	-72.7024	-0.9085***	0.9069***	0.5097***	0.0028	0.8232***	-0.0174
	[0.1534]	[62.0443]	[0.1630]	[0.2430]	[0.1149]	[0.0397]	[0.1486]	[0.1148]
ARCH	0.0212							
	[0.0360]							
GARCH	-0.9457***							
	[0.0909]							
Google Trend	-0.2367*							
	[0.1360]							
L.Google Trend	-0.1444							
	[0.1505]							
# Trans.	0.1586*	-72.4176**	0.1599	0.0625	-0.0603	0.0362*	-0.3682***	-0.0766
	[0.0897]	[29.9472]	[0.1003]	[0.1253]	[0.0733]	[0.0208]	[0.1155]	[0.0785]
ARCH	0.1320***							
	[0.0541]							
GARCH	0.7761***							
	[0.0964]							
Google Trend	2.3347***							
	[0.8185]							
L.Google Trend	0.2603							
	[1.6587]							
# Add.	0.0350	-52.6802	0.0472	0.1378	0.0689	0.0455**	0.1572	-0.6119***
	[0.0889]	[32.1295]	[0.0974]	[0.1352]	[0.0713]	[0.0231]	[0.1140]	[0.0999]
ARCH	0.1010***							
	[0.0458]							
GARCH	0.8718***							
	[0.0532]							
Google Trend	4.3145***							
	[1.3496]							
L.Google Trend	1.2522							
	[2.4484]							
λ_1	0.0345**							
	[0.0152]							
λ_2	0.9333***							
	[0.0271]							

Robust Standard errors are represented by brackets below the coefficients
* p-value < 0.10 ** p-value < 0.05 *** p-value < 0.01

Table 4.10: MGARCH Model Results (b): Conditional Variance

	Exogenous						
	Commodities			Financial Markets			
Variable	ΔGold	ΔOil	ΔN. Gas	ΔDJII	ΔN225	ΔHSI	ΔSHCOMP
Price	0.0422	0.1287	0.0215	0.8820	0.3495	-0.5865*	0.1830
	[0.3282]	[0.0595]	[0.4464]	[0.2804]	[0.3269]	[0.1451]	
ARCH	0.0400						
	[0.0296]						
GARCH	0.8576***						
	[0.0581]						
Google Trend	4.2716***						
	[0.6259]						
L.Google Trend	1.0259***						
	[1.5276]						
Total BTC	0.00032	0.00011	-0.00002	0.00027	0.00013	-0.00062	0.00010
	[0.00034]	[0.0001]	[0.00004]	[0.00065]	[0.00041]	[0.00040]	[0.00020]
ARCH	0.3442						
	[0.2459]						
GARCH	0.4957**						
	[0.2032]						
Google Trend	-0.8894						
	[1.2028]						
L.Google Trend	4.9826***						
	[1.2799]						
Difficulty	-0.0633	-0.0001	0.0244	-0.1880	0.2128	-0.0174	-0.1077**
	[0.1417]	[0.0510]	[0.0311]	[0.1924]	[0.1625]	[0.1629]	[0.0486]
ARCH	0.0309						
	0.0375						
GARCH	0.9089***						
	0.0260						
Google Trend	5.2411***						
	[0.8687]						
L.Google Trend	2.7528***						
	[0.6085]						
Hashrate	0.4637	-0.0672	0.0390	0.6284	-0.3330	-0.5270	0.1049
	[0.3060]	[0.1163]	[0.048]	[0.4424]	[0.2589]	[0.3438]	[0.1674]
ARCH	0.0711						
	[0.0654]						
GARCH	0.0460						
	[0.3572]						
Google Trend	-0.8152						
	[0.5008]						
L.Google Trend	0.1544						
	[0.4754]						
Trans. Cost	0.5895	0.2707	-0.0347	1.3778	0.3831	-1.3258**	0.2131
	[0.4988]	[0.2066]	[0.0862]	[0.8695]	[0.5020]	[0.6353]	[0.3514]
ARCH	0.0212						
	[0.0360]						
GARCH	-0.9457***						
	[0.0909]						
Google Trend	-0.2367*						
	[0.1360]						
L.Google Trend	-0.1444						
	[0.1505]						
# Trans.	0.0740	-0.1191	0.0425	0.4312	-0.4092	0.2575	0.0532
	[0.3757]	[0.1448]	[0.0723]	[0.5906]	[0.3789]	[0.3575]	[0.3106]
ARCH	0.1320***						
	[0.0541]						
GARCH	0.7761***						
	[0.0964]						
Google Trend	2.3347***						
	[0.8185]						
L.Google Trend	0.2603						
	[1.6587]						
# Add.	-0.1874	-0.0385	0.0290	0.3935	-0.2276	-0.1521	0.1138
	[0.3814]	[0.1571]	[0.0784]	[0.5555]	[0.4065]	[0.3263]	[0.2623]
ARCH	0.1010***						
	[0.0458]						
GARCH	0.8718***						
	[0.0532]						
Google Trend	4.3145***						
	[1.3496]						
L.Google Trend	1.2522						
	[2.4484]						
λ_1	0.0345**						
	[0.0152]						
λ_2	0.9333***						
	[0.0271]						

Robust Standard errors are represented by brackets below the coefficients

* p-value < 0.10 ** p-value < 0.05 *** p-value < 0.01

CHAPTER 5: CONCLUSIONS

This paper focuses on modeling and describing the price formation of Bitcoin using prior frameworks with a more robust variable selection. The variables used can be divided into four categories: (i) Bitcoin network measures, (ii) Bitcoin popularity measures, (iii) commodity prices, and (iv) global financial markets.

To gain insight into causal effects of these variables, a Vector Autoregressive Model (VAR) is used as a starting point, where all the variables are treated as endogenous factors. From the results, it can be seen that commodity prices and global financial markets may play a role but should be treated as exogenous factors, which would be consistent with prior logic and expectations. Using these two variable sets as exogenous, there are similar effects across the model coefficients but underlying problems associated Autoregressive Conditional Heteroscedasticity (ARCH) effects persist. To account for these variables and maintain the same structure of the initial VAR model, a Multivariate GARCH (MGARCH) approach is used to capture these effects.

In modeling BTC price and the rest of the variables using MGARCH, both the ARCH and GARCH effects are statistically significant and prevalent throughout the model. When modeling these effects, the coefficients and relationships from the prior models show a significant change. The initial MGARCH model shows that past trends on Google explain changes in BTC price but no other variables within the model have an effect. Since the trends in Google are only significant at the 90% level, an attempt to use this trend variable to explain volatility may be a better use, leading to the final MGARCH model.

The final MGARCH model uses Google trends to explain the heteroscedasticity and volatility in the BTC price and other variables of the model. Past and present values of the Google trend variable are found to be statistically significant for a majority of the variables in the model and also removes the ARCH effects that were prevalent in BTC price, though GARCH effects still persist. This leads to the conclusion that BTC price follows a simple stochastic autoregressive process with volatility being driven by popularity for BTC. To confirm this finding, the models are compared using Schwartz-Bayesian Criteria (SBIC), where it can be shown that an AR(1) model using Google trends and ARCH/GARCH effects to explain heteroscedasticity is the best performing model. The AR(1) process in returns shows there is momentum in the price of BTC. That is, as price rises or falls, it can be expected that the same rise or fall has happened in the prior week, where the momentum then fades.

These results are not consistent with prior works, where supply and demand factors were significant in the formation of BTC price. Since the supply of BTC is strictly controlled by the network and demand for BTC is quite volatile, there may be inefficiency in the market or the demand may be a result of short periods of speculation where large spikes in demand drive significant changes in the entire BTC network.

REFERENCES

[1] G. P. Aielli, "Dynamic Conditional Correlations: On Properties and Estimation," 2009.

[2] R. J. Barro, "Money and the Price Level Under the Gold Standard," March 1979.

[3] P. Ciaian, M. Rajcaniova, and d'Artis Kancs, "The Economics of BitCoin Price Formation," November 2015.

[4] Elliot, Rothenberg, and Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, vol. 64, pp. 813–836, 1996.

[5] R. F. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models.," *Journal of Business and Economic Statistics*, vol. 20, pp. 339–350, 2002.

[6] I. Georgoula, "Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices," May 2015.

[7] C. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, pp. 424–438, August 1969.

[8] H. Lutkepoh, *New Introduction to Multiple Time Series Analysis*. 2005.

[9] J. Roberts and N. Rapp, "Exclusive: Nearly 4 Million Bitcoins Lost Forever, New Study Says." http://fortune.com/2017/11/25/lost-bitcoins/, 2017.

[10] G. Schwarz, "Estimating the Dimension of a Model," *The Annals of Statistics*, vol. 6, pp. 461–464, March 1978.

[11] J. Smith, "An Analysis of Bitcoin Exchange Rates," May 2016.

[12] D. Yermack, "Is Bitcoin a Real Currency? An Economic Appraisal," April 2014.

APPENDIX A: Variable Graphics

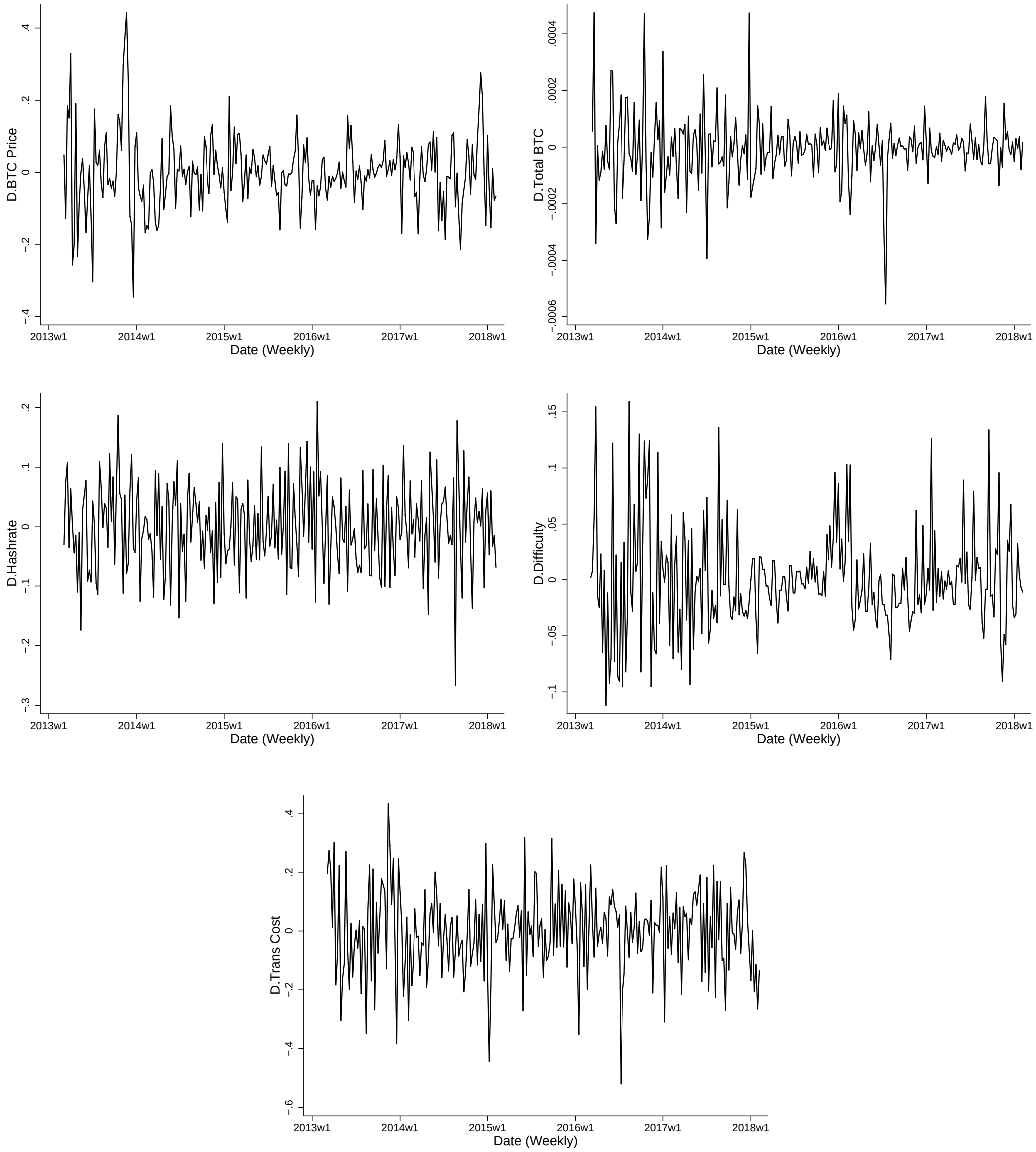

Figure A.1: Network Statistics

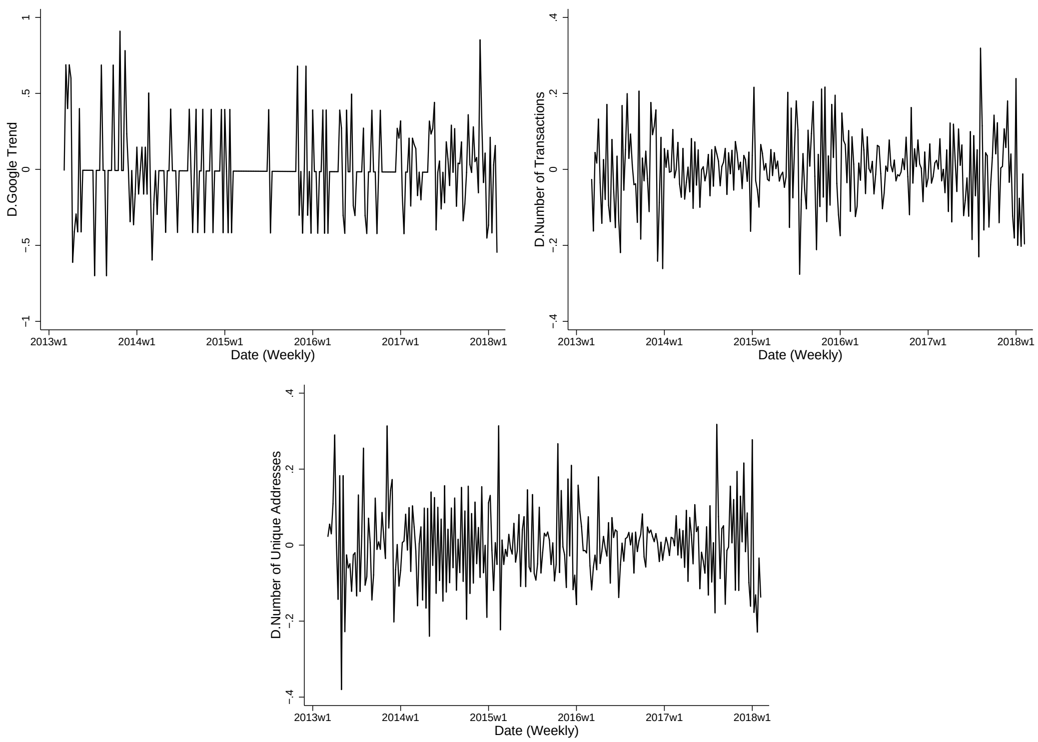

Figure A.2: Popularity Measures

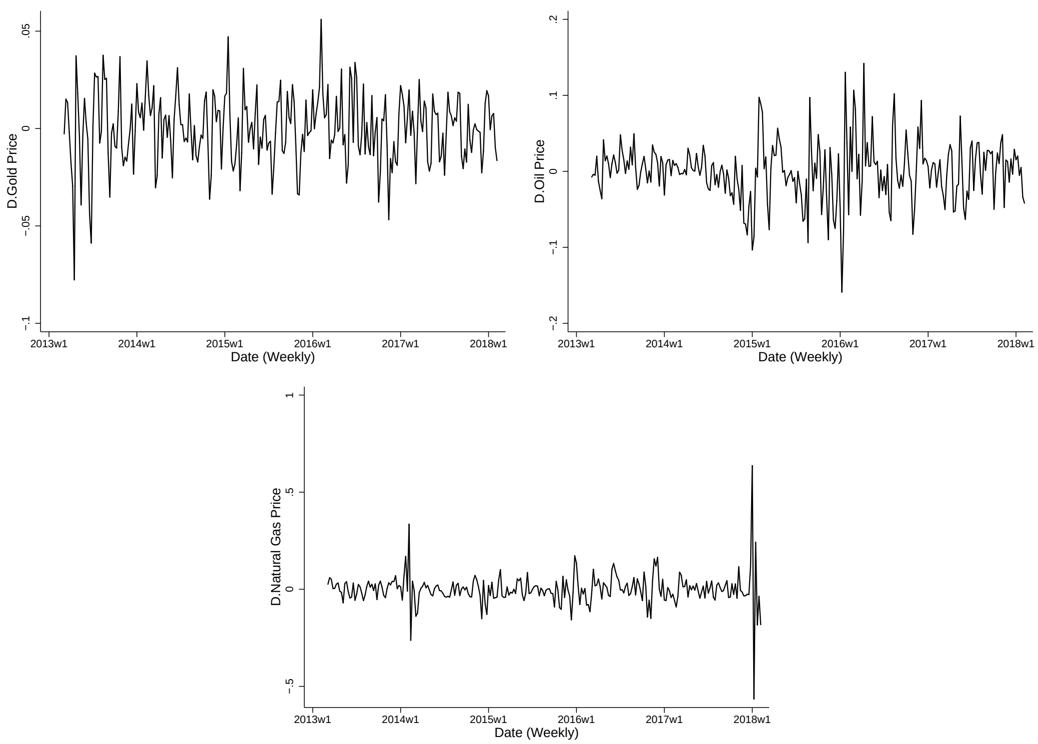

Figure A.3: Commodity Prices

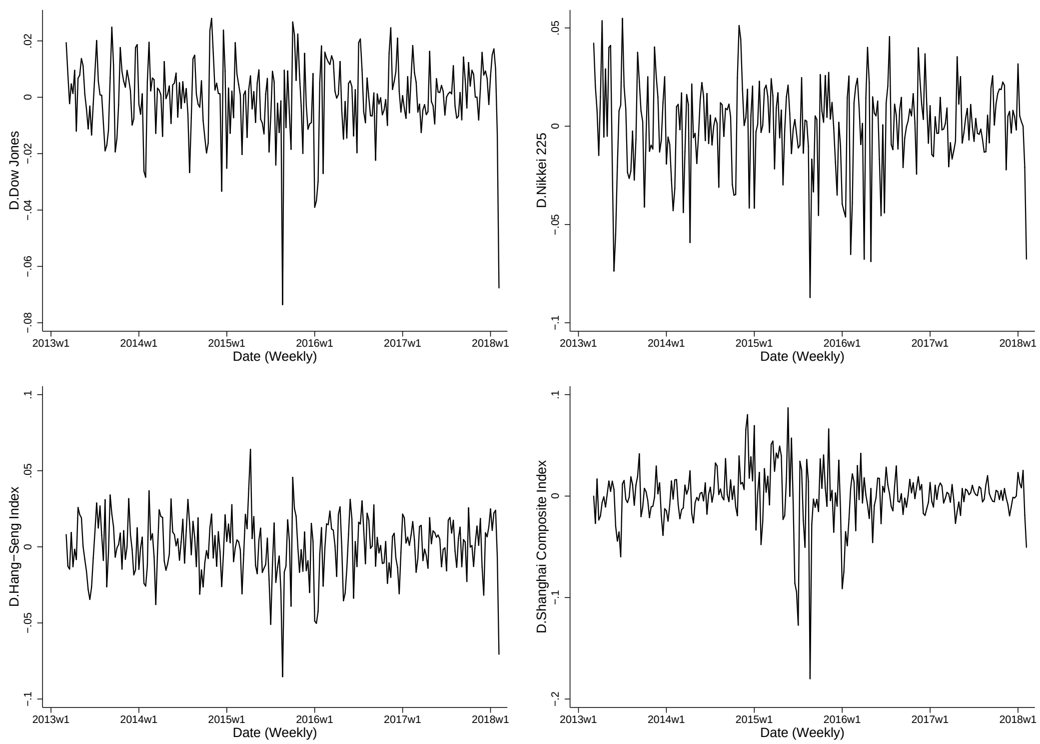

Figure A.4: Financial Markets

APPENDIX B: Residual Stability Graphs

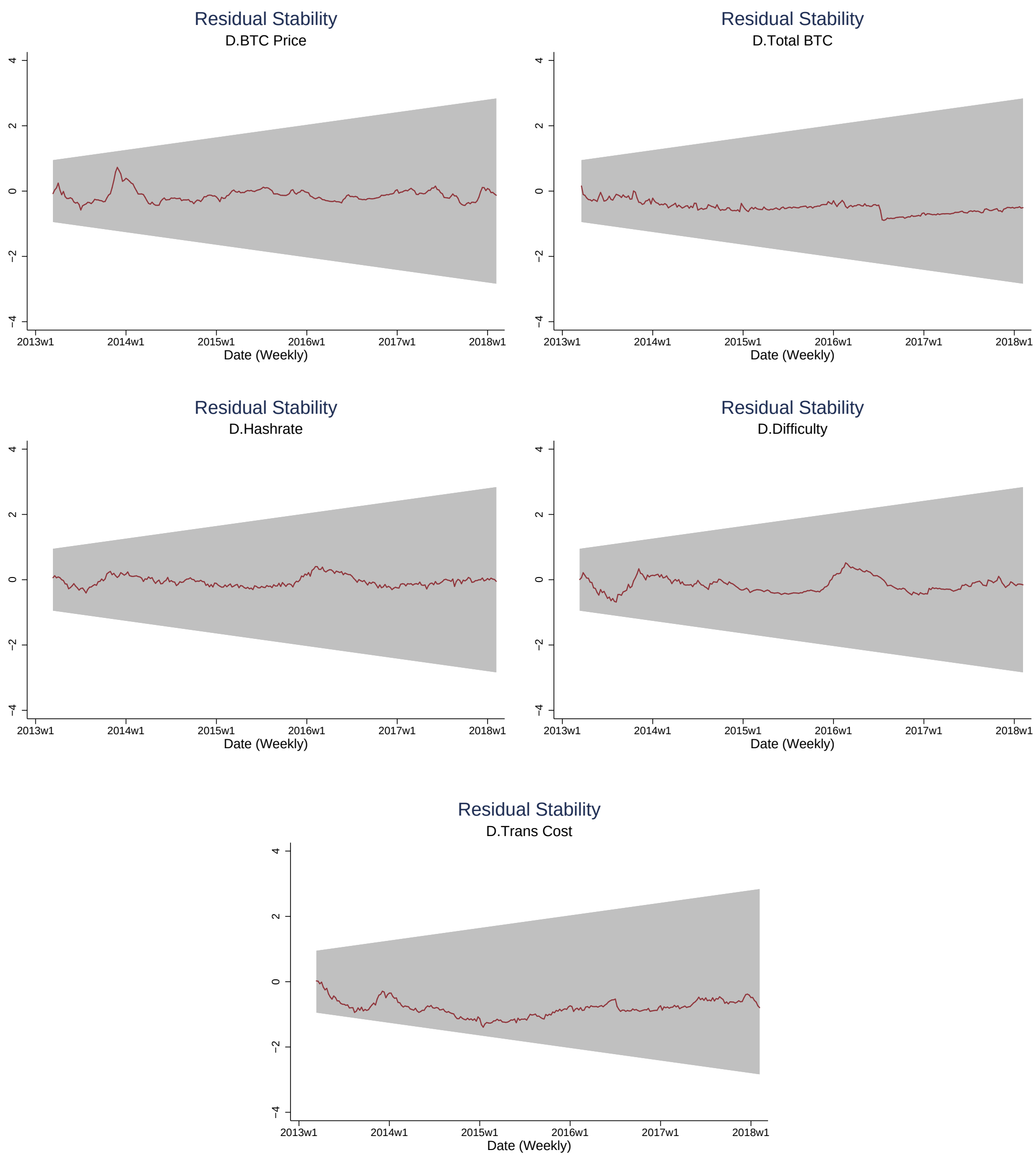

Figure B.1: Residual Stability: Network Statistics

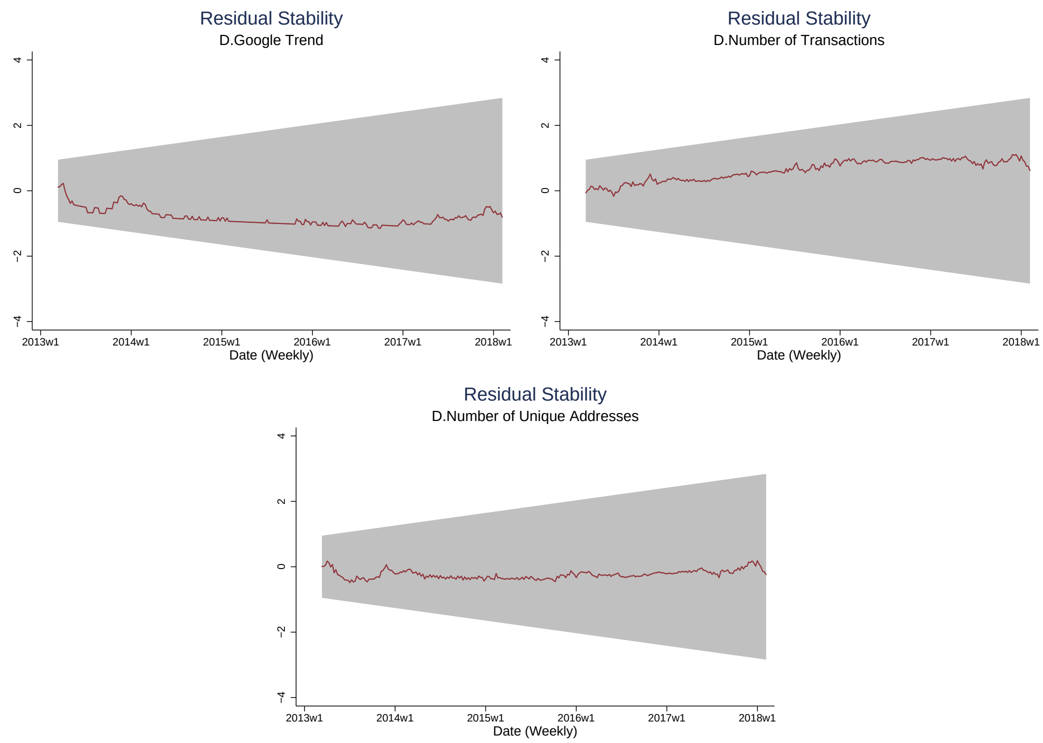

Figure B.2: Residual Stability: Popularity Measures

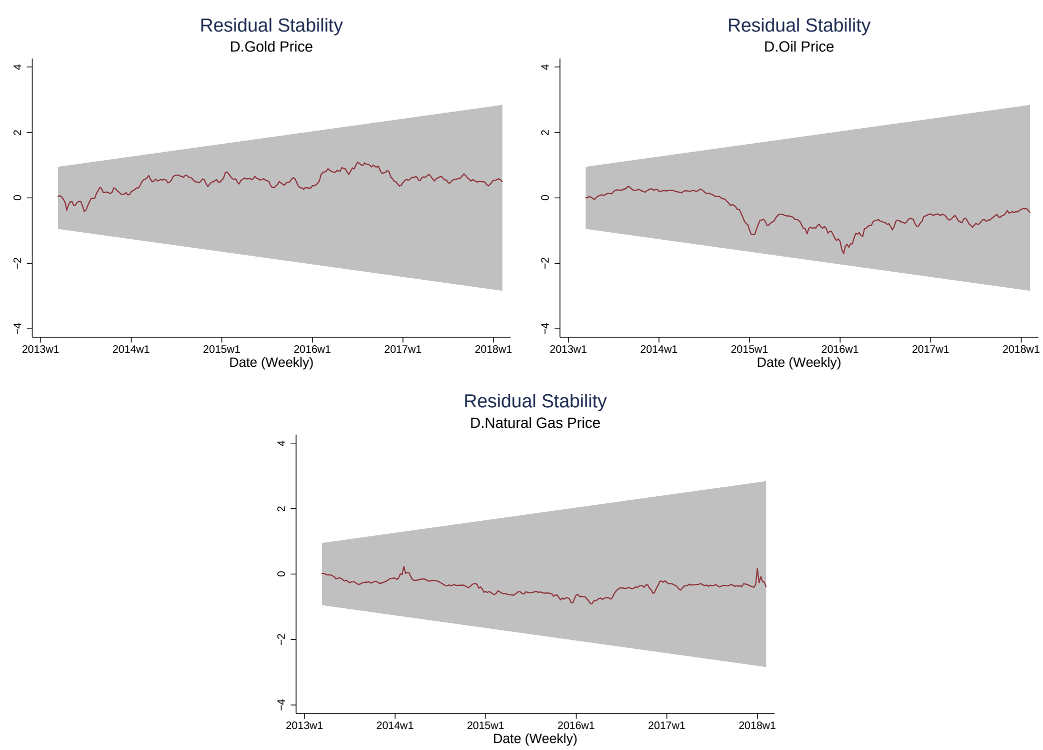

Figure B.3: Residual Stability: Commodity Prices

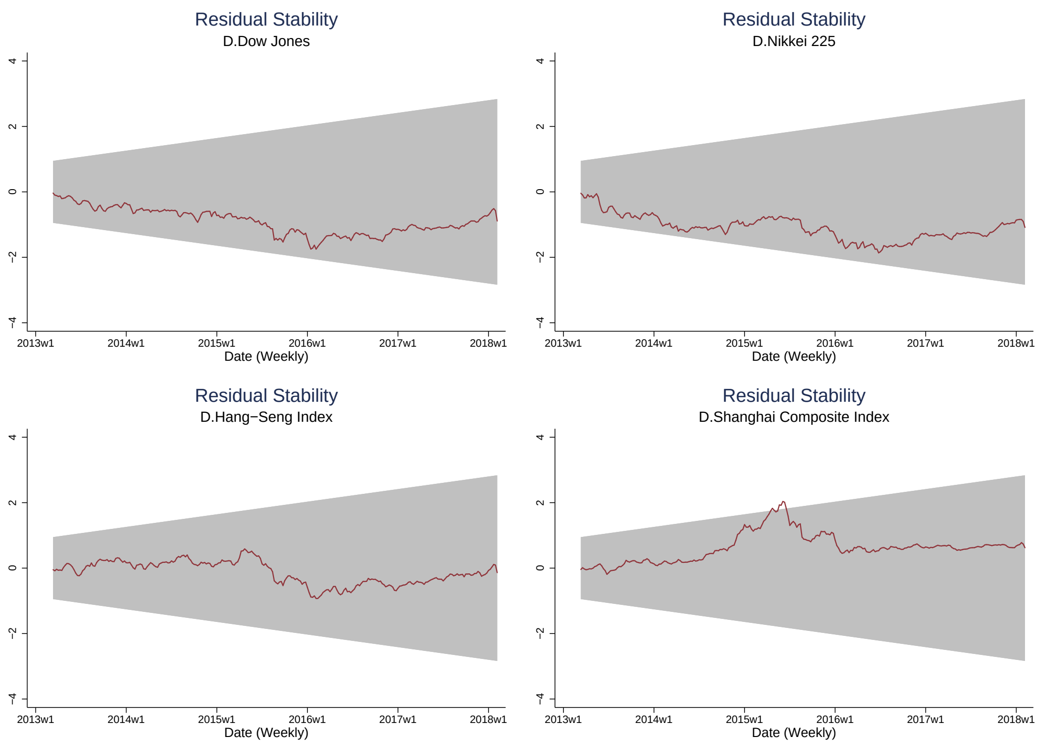

Figure B.4: Residual Stability: Financial Markets

APPENDIX C: Theoretical Supply-Side Model for BTC Returns

We start by defining the probability that a miner will solve a block:

$$\rho_i = \frac{H_i}{H_n - H_i} \tag{C.1}$$

Where H_i is the *hashrate* of miner X_i, and H_n is the *hashrate* of the entire Bitcoin network:

$$H_n = \sum_{i=1}^{\infty} X_i \tag{C.2}$$

Taking into account that a miner may combine their hashrate with a group of other miners, defined as a mining pool, the hashrate (H_i) can be represented as the sum of all hashrates within a given pool, i, where H_i would then represent:

$$H_i = \sum_{m=1}^{N} H_m \tag{C.3}$$

The probability (ρ_i) of a given miner (X_i) solving a block times the price of a Bitcoin at a given time, P_t, can be defined as some revenue, TR_i.

$$TR_i = \rho_i[P_t(w_m B_t + T_i)] \tag{C.4}$$

Where B_t is the number of Bitcoins awarded for solving a block, T_i is the sum of all transaction fees within the block that was solved $\sum_{n=1}^{350} T_n$, and w_m is the weight of an individual miners hashrate compared to the hashrate of the entire pool. If a miner is not part of a pool then $w_m = 1$. This revenue equation can capture the revenue (TR_i) at a given time, t, or at all time periods by representing B_t as $\sum_{t=1}^{\infty} B_t$.

Since we are able to define a revenue function for a given miner (TR_i), we can also define a cost function for this same miner:

$$TC_i = VC_i(B) + FC_i \tag{C.5}$$

One can then assume that the variable cost (VC_i) is zero. This is due to the nature of mining Bitcoin. Intuition would define the variable cost as some function of the cost of energy consumption. For this to be correct, energy consumption would have to show variation for the number of Bitcoins mined. Since energy consumption is only a function of the hashrate and a miner would want to maximize their hashrate by running their mining equipment at the maximum safe level it would allow, energy consumption would not vary.

Taking our TR and TC functions (Equations C.4 and C.5 respectively), we can then develop our Total Profit function (π) as:

$$\pi = \rho_i[P_t(w_m B + T_i)] - FC \tag{C.6}$$

We can then take this equation and solve for when a miner would no longer mine Bitcoin:

$$\begin{aligned} 0 &= \rho_i[P_t w_m(B + T_i)] - FC \\ FC &= \rho_i[P_t w_m(B + T_i)] \\ P_t &= \frac{FC}{\rho_i w_m \left[B + T_i\right]} \end{aligned} \tag{C.7}$$

Since a combination of all miners on the network have a 100% probability of mining a BTC ($p_n = \sum \rho_i = 1$) then the equation would simply be:

$$P_t = \frac{FC}{(B + T_n)} \tag{C.8}$$

For ease, we can then convert our equation to log form:

$$ln(P_t) = ln(FC_t) - ln(B_t + T_{n,t}) \tag{C.9}$$

Now, looking at the return of BTC which is the log change of price from one period to the next, we can then see that FC_t and B_t will be 0, on average. The FC_t will shift with energy prices and depending on how often these energy prices shift from period to period will determine how often our FC_t are non-zero. The amount of Bitcoins mined, B_t will tend to be zero due to the design of the BTC network. The difficulty adjustment will rise or fall to create a constant supply of a BTC mined every ten minutes, on average. Therefore, our equation then simplifies to:

$$ln(P_t) - ln(P_{t-1}) = -ln(B_t + T_{n,t}) - ln(B_{t-1} + T_{n,t-1}) \tag{C.10}$$

Taking the negative of both sides to provide a more intuitive problem yields:

$$-ln(P_t) + ln(P_{t-1}) = ln(B_t + T_{n,t}) + ln(B_{t-l} + T_{n,t-1}) \tag{C.11}$$

Which says that a proposed return on BTC at time t should be equal to the change in the amount of BTC mined and transacted on the network.