THE COOPER UNION ALBERT NERKEN SCHOOL
OF ENGINEERING

Applying a Bayesian Structural Time Series Model to Infer Causal Impact in the Crypto Market

By

Philip Blumin

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Engineering

December 2022

Advisor

Professor Sam Keene

THE COOPER UNION FOR THE ADVANCEMENT OF
SCIENCE AND ART

ALBERT NERKEN SCHOOL OF ENGINEERING

This thesis was prepared under the direction of the Candidate's Thesis Advisor and has received approval. It was submitted to the Dean of the School of Engineering and the full Faculty, and was approved as partial fulfillment of the requirements for the degree of Master of Engineering.

12.7.2022

Barry L. Shoop, Ph.D., P.E. *Date*

Dean, Albert Nerken School of Engineering

12.7.2022

Prof. Sam Keene, Thesis Advisor *Date*

Acknowledgements

I would like to thank my family and friends for motivating and supporting me throughout my time at the Cooper Union. I would also like to thank Dr. Sam Keene for being an amazing advisor and making the process of creating this thesis as smooth as possible. Additionally, I would like to acknowledge Joshua Riess for proofreading and providing feedback on this thesis. Finally, I would like to acknowledge Josh Yoon for helping me with Latex.

Abstract

Cryptocurrency is the 'Wild West' of investing, with no regulatory oversight and little insight into the future of widely-owned cryptocurrency. Current analysis techniques are unable to make sense of price trends in the cryptocurrency market. This thesis proposes that determining the causal impact of events on cryptocurrencies will allow analysts to more easily predict prices and trajectories based on their knowledge of similar situations. Inferences on the causal impacts of events on cryptocurrencies were analyzed using a Bayesian structural time series (BSTS) model. A BSTS model utilizes prior knowledge of trends from the variable it analyzes, and multiple control markets to determine the impact of treatment (an action at a point in time) on the variable's value. Different analyses are conducted in this thesis using a Python implementation of Google's Causal Impact R library. This paper uses a BSTS model to run various causal inference analyses. The benefits and limitations of this approach are explored through the lens of cryptocurrencies. Additionally, this thesis examines the potential impact of social media on the prices of cryptocurrency. After running through several experiments this thesis demonstrates the viability of using a BSTS model on volatile data such as cryptocurrency. The results show the model is able to reveal what effects a treatment has on price data and the conclusion suggests avenues for improvement.

Contents

1 Introduction

Understanding the stock market has always been one of the most complex tasks in statistical and time series analysis. There are many daily events and news that can possibly influence the value of a stock or the entire stock market. With the arrival of Bitcoin in 2009, the first successful decentralized cryptocurrency, analyzing the market became even more strenuous. Bitcoin is extremely volatile [19], which made it difficult to make sense of with the leading analysis methods that were being used on the market at the time. As Bitcoin grew in popularity, many more cryptocurrencies were created with equally challenging behaviors to predict. When trying to predict something so unstable, analysts must determine if it is possible to narrow down what causes a dramatic shift in the price. This leads to the question: Do certain events have a causal impact on the price of Bitcoin and other cryptocurrencies? Uncovering the events that impact the price of cryptocurrencies may allow analysts to more easily predict prices and trajectories based on their knowledge of similar situations.

Over the years, quantitative research and analysis has helped develop many methods and models that can make sense of shifts in the market, as well as forecast prices. One of the most well-known techniques developed to determine the causal impact of an event is difference-in-differences [4, 9]. Although there are many advantages to using difference-in-differences, there are many cases where this method falls short. A more detailed discussion of the benefits, limitations, and an example of difference-in-differences being applied will be discussed in section 2.3.2. In more recent years, endlessly growing data has led researchers to examine machine learning techniques as a more advanced method of understanding and forecasting the market. They have developed many time series forecasting approaches, such as the Auto-regressive Integrated Moving Average (ARIMA) model. In 2015, a very powerful library was developed by Google in R which uses a Bayesian structural time series (BSTS) model to infer the causal impact of an event [1].

This approach was introduced in a paper by Broderson et al. that showed the power of this tool when analyzing the causal impact of an advertising campaign. The model utilizes prior knowledge of trends from the variable it is analyzing and multiple control markets (predictors) to determine the impact of a point in time on the variable's price or value. A detailed discussion of the model as well as the original paper will be in section 2.3.3.

The goal of this thesis is to determine whether a BSTS model can be applied to a highly volatile variable such as Bitcoin and other cryptocurrencies. This paper used a Python version of Google's causal impact library to perform various analyses. This paper explores the library's benefits and limitations through the lens of cryptocurrencies. Additionally, this thesis explores the potential impact of social media on the prices of cryptocurrency.

Recent events have indicated that social media could have an impact on the price of crypto. Memes and tweets appear to correlate to shifts in cryptocurrency prices. For example, Elon Musk's tweets in April 2021 praising Dogecoin seemed to lead to the cryptocurrency skyrocketing in price. In general, Elon Musk has demonstrated his impact on the prices of cryptocurrencies, whether it is through his tweets [5] or other forms of social media. Being able to infer the causal impact of an event, such as a tweet from Elon Musk, can have a crucial effect on the forecasts of crypto prices.

2 Background

2.1 Cryptocurrency Overview

One of cryptocurrency's most well-known features is its extreme price volatility [19]. The high volatility of crypto makes it difficult to forecast and analyze, thus making people cautious about investing. Bitcoin, the first successful decentralized cryptocurrency, is famous for rapidly gaining and losing much of its value. For example in 2017, Bitcoin was shocked the world when its price skyrocketed from around $1,000 to over $19,000 by December.

After the boom of Bitcoin, the popularity of cryptocurrency investing grew and many other coins were created. By July 2022, there were over 20 thousand cryptocurrencies in circulation. With time and effort, anyone can create their own cryptocurrency. This has led to a lot of 'memecoins' (cryptocurrencies created as elaborate jokes) being created, such as Dogecoin.

The oversaturation of cryptocurrencies has created a challenging environment for researchers to understand and analyze. Getting an accurate forecast for any cryptocurrency is nearly impossible. Every analyst seems to have a different opinion or calculation on the future prices of crypto. Researchers broadly agree (all other popular cryptocurrencies) markets are interdependent [6]. A shift in the price of Bitcoin will cause a shift in the price of all altcoins, and vice versa. This could potentially simplifies analyzing and forecasting altcoins.

The key to forecasting crypto could lay in the ability to understand what impacts the price in the long and short-term. Ascertaining the impact of certain types of events allow analysts to create more accurate forecasts based on future events. For example, if analysts find that the price of Bitcoin goes down about 5% every winter, they can factor this expected decrease into their forecast.

2.2 Impacts in the Cryptocurrency Market

At its core, the cryptocurrency market is very similar to the stock market. As established earlier, almost all cryptocurrency prices follow Bitcoin. This is similar to the behavior the stocks of smaller companies have when compared to the Dow Jones Industrial Average. Just like the stock market, there are countless events that have had an impact on cryptocurrencies over the past few years.

2.2.1 Global Events and Regulations

A recurring event that has been found to impact crypto prices is regulations [20]. Crypto's high volatility combined with its correlation to the stock market could lead to a global economic collapse. This has caused public authorities and regulators to often attempt to regulate the crypto market. A recent study analyzed the long- and short-term impact of these regulations and discovered a correlation between the crypto prices and these regulations [7].

The creators of cryptocurrencies aspired to create a decentralized global currency, ensuring that global events directly lead to price fluctuations for large cryptocurrencies. Examples of global events with major price impacts are the COVID-19 pandemic and El Salvador adopting Bitcoin as one of its official currencies [21]. Some of the effects from these events are easy to visualize and understand. For instance, the global recession caused by the COVID-19 pandemic caused traditional stocks and cryptocurrencies to crash. Looking at the trend of Bitcoin, the price nearly halved between mid-February 2020 and mid-March 2020. However, the impact of some events, such as El Salvador adopting Bitcoin, cannot be seen as easily on a graph. As seen in Figure 11, the price of Bitcoin fluctuated so much that it is difficult to draw a conclusion about how this event impacted its price. Many events may impact the prices of cryptocurrencies, so it is not easy to know if an event has a causal impact on the price of crypto. In order to determine if an event has impacted a cryptocurrency's price it is possible to use causal impact techniques.

2.2.2 Elon Musk and Social Media

Analyses of social media sentiments demonstrate Twitter and Facebook's impact on cryptocurrency prices. [8]. Due to this, the events in this thesis are classified as either relating to social media or not.

As a result of the COVID-19 pandemic, many people started to use more social media since they could not physically interact with others. Social media quickly became one of the main places for people to disseminate and create news. During this time, Elon Musk, founder and CEO of Tesla Motors, cultivated his massive presence on social media, specifically Twitter. He tweeted multiple times a day and gained millions of followers (currently over 100 million). Elon Musk's Twitter fame and perceived knowledge of cryptocurrencies allowed him to greatly impact the prices of Bitcoin, Dogecoin, and many other altcoins. A recent study from Blockchain Research Lab (BRL) done by Lennart Ante found that some of Elon Musk's Tweets regarding cryptocurrency had an impact on the price of crypto. The conclusion Lennart found was that Elon Musk's Dogecoin-related tweets had a significant impact, while the Bitcoin-related tweets often varied in impact [5].

One of the most famous examples of Elon Musk's tweets moving crypto was when he tweeted "... going to moon very soon" on April 10th, 2021. This tweet was meant to be a joke regarding the ‘meme’ cryptocurrency, Dogecoin. Prior to April 2021 Dogecoin sat consistently under 10 cents, but suddenly skyrocketed to nearly 70 cents by May 2021. The reason for this large jump can be solely attributed to Elon Musk’s influence on Twitter and other media platforms. This event is just one example of many tweets from Elon Musk about crypto that had a demonstrated impact on their prices. As with global events, it may be difficult to determine how impactful a certain tweet can be. An analyst needs to consider other factors going on at the time of the tweet relating to crypto. There are several techniques that can be used to determine the causal impact social media posts have on the price of crypto.

2.3 Techniques for Inferring Causal Impact

As mentioned in the previous section, determining whether an event has a causal impact on the price of crypto is difficult. There could be many other variables at play through the duration of an event. For example, studying the effect of a random Elon Musk tweet from February 2020, effects of the global shutdown, due to the COVID-19 pandemic, needs to be considered. Additionally, cryptocurrency’s innate volatility can make it difficult to perform simple analysis techniques to infer the causal impact.

2.3.1 Randomized Experiments

The gold standard method for estimating causal effects is a randomized experiment. This involves randomly splitting a group into two groups, and giving one group the treatment (treatment group) and nothing to the other (control group). The effect of the treatment on the treatment group is observed and a conclusion is drawn about the possibility of a causal impact. However, in many situations it is not possible to run a randomized experiment because it is too expensive, too difficult, or unethical.

In the case of analyzing the impacts from a specific event in the crypto market, a random experiment would be impossible to do because in the real world if an event (treatment) occurs it will affect the entirety of the stock market. Ideally, there would be a way to randomly split up stocks, apply a treatment to only one group, and observe the impacts. Unfortunately, there is no feasible way to do this, so in the case of finding causal impacts in the market, a different approach must be used.

2.3.2 Difference in Differences

When a randomized experiment cannot be done, a popular approach to inferring the causal impact of a treatment is different-in-differences (DD). The main question causal analysis aims to answer is determining the effect of the treatment on the observed outcome. In a DD approach this is done by looking at the counterfactual, which is the outcome had the treatment not been applied. The causal effect of the treatment can be estimated by taking the difference between the counterfactual and the observed outcome.

Figure 2 shows a simple example of difference in difference, where the causal effect of the intervention can be found with the formula (Y_{T2} - Y_{C2}) - (Y_{T1} - Y_{C1}). In this formula Y_{T2} and Y_{C2} represents the value at the end of the time frame for the treatment and control variables, respectively. Similarly, Y_{T1} - Y_{C1} is the difference between the treatment and control variable at the start of the time frame (T = 0). This is the average change over time from the post-treatment to the pre-treatment.

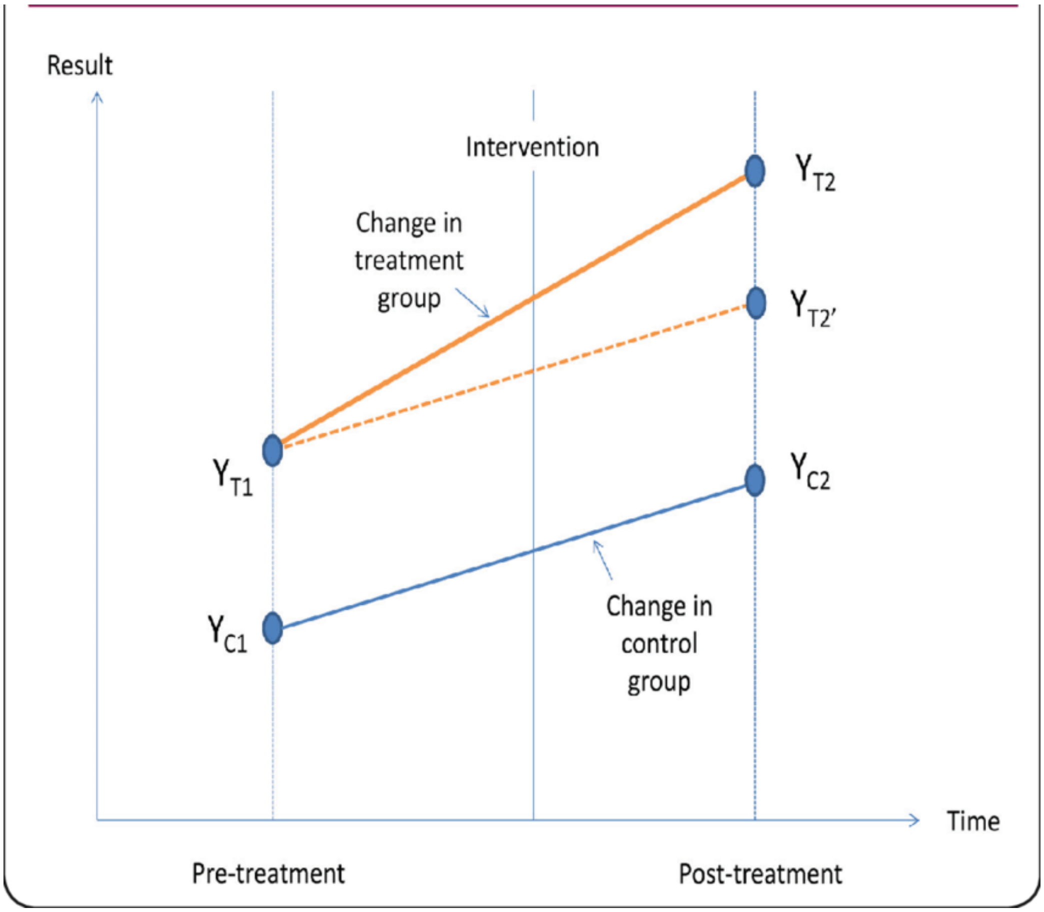

Figure 1: Visualization of the different-in-differences(DD) method [22]

A study done in 2013 by Shiqing Xie demonstrated a great example of using the difference-in-difference approach to examine the impact of index future trading had on stock market volatility in China. In the study a multi-period DD analysis was done and only a short-term impact was found to be significant [4].

Using a traditional DD approach has a few problems. A problem arises with the fact that a causal effect can never be observed because there is no way to observe both the outcome and counterfactual at the same time. One solution to this problem is finding controls or predictors, which follow the same trend as the main variable, but are not affected by the treatment. In the crypto market, Bitcoin can be used as an excellent control for other coins. As discussed in section 2.1, many popular altcoins tend to follow the price trends set by Bitcoin. Similarly, studies have found that Bitcoin follows similar trends to the stock market [10], so stocks could potentially be great controls when analyzing the impact of an event on bitcoin. In general, as many good controls as possible are needed to estimate the counterfactual. Once the counterfactual is estimated a traditional DD approach seen in Figure 2 can be used to simply find the causal effect.

There are a few assumptions of the difference-in-differences approach. The first assumption is that the observations are independent and identically distributed. The second one is known as the parallel trend assumption, which states that the trend of the variable with the treatment and the control are identical [9]. With real world time series data, these assumptions almost never hold. Taking any stock as an example, there will constantly be countless factors or variables that affect its price. As a result, using a stock as a control would violate the parallel trend assumption. Another limitation of DD is that the model does not reveal anything about the change in effect over time. As seen from the example in figure X, there are only two points in time that are being observed, which are the pre- and post-values. Anything that occurs in between would be a black box when using this approach.

2.3.3 Bayesian Structural Time Series

Another approach for inferring the causal impact of an event is using a Bayesian Structural Time Series (BSTS) model. This method involves fitting a BSTS with several controls to predict a counterfactual post-treatment time series. This time series showcases the variable of interest's trajectory in a reality where the treatment never took place. Then, similarly to the DD approach, the estimated causal impact of the treatment can be found by taking the difference between the observed and counterfactual post-treatment time series.

Using a machine learning technique, rather than the traditional DD approach has many advantages. A BSTS approach reduces many of the disadvantages of DD and allows for a lot more flexibility. Unlike DD, BSTS allows the possibility to see the impact of the treatment over a period of time. Furthermore, since this is a fully Bayesian approach, all parameters have empirical priors, which prevents overfitting. Another advantage of BSTS is that it can account for large variations in trends, such as seasonality [1, 11].

Structural time series models are state-space models for time series data [1]. A state-space model is a model of a system as a set of input, output and state variables related by first-order differential equations. The following two equations can be used to define a structural time-series model:

$$y_t = Z_t^T \alpha_t + \varepsilon_t \tag{1}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \tag{2}$$

where $\varepsilon_t \sim$ N(0, σ_t^2) and $\eta_t \sim$ N(0, Q_t) are independent of all other unknowns. Equation 1 is the *observation equation*, which links the observed data y_t to a d-dimensional state vector α_t. Equation 2 is the *state equation*, which controls the change of the state vector α_t over time. In these equations Z_t is a d-dimensional output vector, T_t is a d x d transition matrix, R_t is a d x q control matrix, and η_t is a q-dimensional system vector [1].

Many models used for forecasting, such as ARIMA [23], can be written based on equations 1 and 2. The state vector α_t allows for the flexibility to account for trends and seasonality. This is based on the assumption that the different state component model errors are independent, thus permitting vector α_t to be formed by concatenating the individual state components [1]. A regression component allows for the counterfactual prediction to be derived by creating one synthetic control based on a combination of controls not affected by the treatment.

The first component of the model is a linear trend that can be defined by the following equations:

$$\mu_{t+1} = \mu_t + \delta_t + \eta_{\mu,t} \tag{3}$$

$$\delta_{t+1} = \delta_t + \eta_{\delta,t} \tag{4}$$

where $\eta_{\mu,t} \sim$ N(0, σ^2_μ) and $\eta_{\delta,t} \sim$ N(0, σ^2_δ). Here the variable μ_t is the value of the trend at time t, δ_t is the expected change in μ between times t and t + *1* (the *slope* at time t). This is a commonly used local linear trend model because it can quickly adapt to local variations, which is great for short-term predictions [1].

The next step is to add components to the state space that represent the controls that have similar trends, but are not affected by the treatment. As mentioned before, these controls will be combined to predict the counterfactual. These controls are incorporated using linear regression. Given controls j = 1,2,...,J the regressed component is defined as

$$X_t^T \beta_t = \sum_{j=1}^{J} x_{j,t}\beta_{j,t}, \tag{5}$$

$$\beta_{j,t+1} = \beta_{j,t} + \eta\beta_{j,t} \tag{6}$$

where $\eta\beta_{j,t} \sim$ N(0, $\sigma^2_{\beta_j}$). In these equations $\beta_{j,t}$ is the coefficient of the *jth* control. This equation can be written into state-space form as $Z_t = \beta^T X_t$ and $\alpha_t = 1$, where

X_t represents all of the controls.

Since the approach is fully Bayesian, a spike-and-slab prior can be used to determine which controls should be used and how heavy they should be weighted. Spike-and-slab is a type of Bayesian regression method that is used to help pick what variables (controls) are the most important and should be used in the final prediction. The model shapes two prior distributions. The first one is a 'spike' which reveals the probability of a variable being chosen for the model. The second is the 'slab' which shows the variable coefficient values, or the weights [12].This approach allows the user to avoid overfitting by confirming the controls within the BSTS model. In other words, using this method ensures that the controls being used are not actually significantly affected by the treatment, which other methods such as DD do not support.

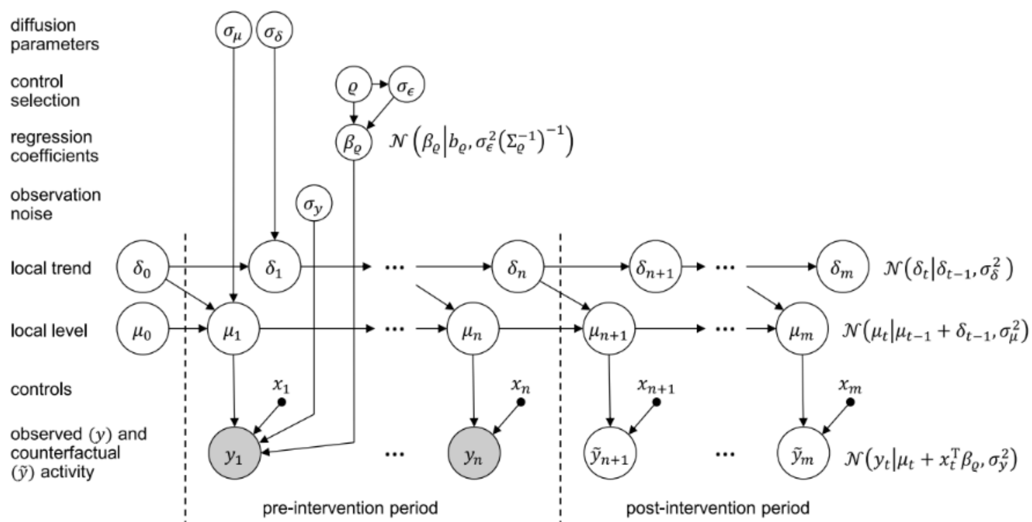

Figure 2: Graphical representation of BSTS model taken from the original Causal Impact paper by Brodersen, et al. [1]

3 Application of BSTS to the Crypto Market

3.1 Problem Statement

The main goal of this thesis has two elements to it: a technical component and an analytical component. Its primary objective is to demonstrate that a BSTS model can be applied to infer causality in the crypto market. This is achieved by running several experiments to observe causality using a Python version of Google's CasualImpact [2] library, developed by Willian Fuks [3]. The analyses use global and social media events as treatments on the cryptocurrency price. If the BSTS approach proves to be feasible, many different events can be run using the model to give analysts a better understanding of the type of events that cause significant movement in the crypto market.

3.2 Data

This thesis used three different data sources to perform its experiments: Cryptocurrency, Stock, and Twitter data. The data was taken from investing.com [14] by selecting all the data since January 1st, 2019. Then, the daily closing price for each crypto was cleaned from this data. At first, the idea was to see if the various other altcoins could be used as controls for each other, however, after performing various analyzes, it was found that all the altcoins were extremely correlated and followed bitcoin [6]. This would make the altcoins poor controls for most experiments.

As a result, this thesis decided to use the stock market and various individual stocks as controls. Using stocks as a control for inferring causal impacts of crypto is viable because major cryptocurrencies, such as Bitcoin, follow similar trends to the market, but the individual stocks do not have huge price swings like crypto. The individual stocks selected to be used in the experiments were Amazon, Apple, Facebook, Google, Tesla, and Twitter. Additionally, the Dow Jones Industrial Average was used as a control. All of these controls are used to predict a counterfactual for each experiment. The data for each of these was obtained by using the *GOOGLEFINANCE* function [24] on Google Sheets, which provides daily data of the closing price for each stock. The following equation shows an example of how this function was used to obtain the time series data for Amazon:

```
GOOGLEFINANCE("AMZN","price", DATE(2019,1,1), DATE(2022,6,30), "DAILY")
```

Since stock prices were decided to be used as the controls for the experiments rather than altcoin prices, this study selected just 3 cryptocurrencies as main variables: Bitcoin, Dogecoin, and Cardano. Only 3 different coins were used, due to the similarity in the trends explained in section 2.1.

3.2.1 Bitcoin

Bitcoin (BTC) was created in 2009 and is the first successful decentralized digital currency. Being the most popular and trend defining cryptocurrencies, Bitcoin was selected as the treatment variable for most of the experiments. Furthermore, since most altcoins follow Bitcoin, it is used as a control for causal impact experiments done with altcoins.

3.2.2 Dogecoin

Dogecoin (DOGE) was created in 2013 by 2 software engineers as a joke to make fun of the idea of using a digital currency such as Bitcoin. This crypto was selected

Figure 3: Price of Bitcoin since January 2019

to be used because of its connections to social media. Throughout 2021, Dogecoin was a popular subject to joke about on Twitter, but the price of Dogecoin seemed to shift in tune with Twitter-driven frenzy. In the analysis, Elon Musk's tweets were used as a treatment to infer a causal impact on Dogecoin.

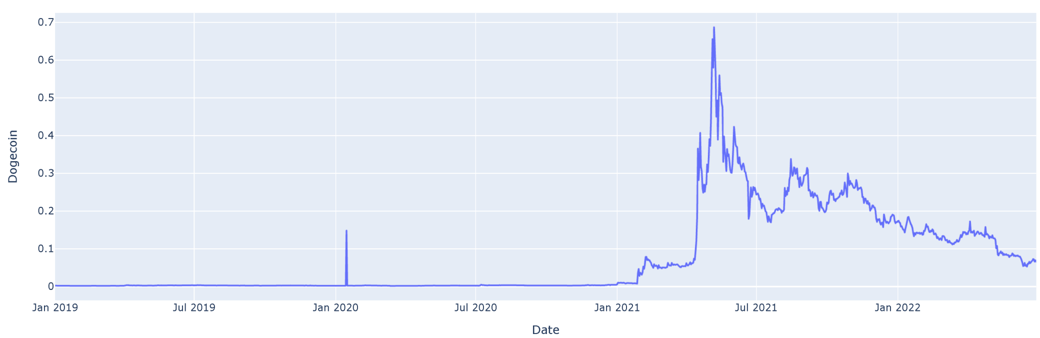

Figure 4: Price of Dogecoin since January 2019

3.2.3 Cardano

Cardano (ADA) is a popular altcoin that was founded in 2015. The goal of Cardano is to find a greener digital currency solution by using a proof-of-stake block chain. Cardano was selected as a secondary altcoin to observe effects of social media on crypto.

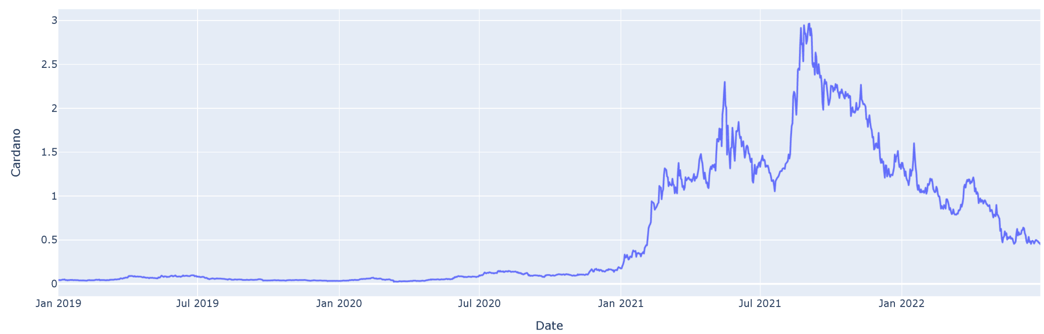

Figure 5: Price of Cardano since January 2019

3.3 Events Investigated

In order to test the BSTS model, the effects of 3 global events and 4 tweets from Elon Musk were investigated. This study chose three world events from different years in order to see short- and long-term effects on the crypto market. The effect of the world events were analyzed specifically on Bitcoin because other altcoins tend to follow Bitcoin prices. On the social media side of the analysis, two of Elon Musk's tweets relating to crypto and two tweets not relating were chosen. The reason for this was to explore if tweets unrelated to crypto would have an effect on the crypto market as well. The tweets that discussed crypto are some of the most well known and controversial tweets that Elon Musk made in 2021, during the crypto boom. The other two tweets were picked from a dataset found on Kaggle which contained all of Musk's tweets from 2010 to 2021 [13]. Two tweets with the most likes, retweets, and replies were since 2019 selected (the crypto dataset only had data after January 1*st* 2019). The effects of each tweet were studied on Bitcoin, Dogecoin, and Cardano.

Below is a list with more detailed descriptions of all the events chosen:

- **Facebook Announces Libra** (June 17*th*, 2019): Facebook announces that they will launch their own cryptocurrency in 2020, which they will integrate into Facebook's apps and services. Unfortunately the Libra (now known as Diem) project had a lot of backlash and ended up failing.

- **The Onset of the COVID-19 Pandemic** (February 14*th*, 2020): Around February 2020, most companies started to close down and transitioned to fully remote work due to the global pandemic. This led to huge swings in the markets that had not been seen in a long time.

- **El Salvador Adopts Bitcoin** (September 7*th*, 2021): El Salvador became the first country to officially adopt Bitcoin as a legal tender. This was huge news in the crypto world and the effects of this are analyzed.

- **"... going to moon very soon" - Elon Musk** (April 10*th*, 2021): One of the most well-known tweets made by Elon Musk concerning crypto. After this tweet crypto was at the forefront of peoples' minds. Many people began to invest in Dogecoin either as a joke or believing Elon Musk's conjectures.

- **"Cryptocurrency is promising, but please invest with caution!" - Elon Musk** (May 7*th*, 2021): After this tweet Elon Musk announced that Tesla would no longer be accepting Bitcoin for purchases.

- **"The coronavirus panic is dumb" - Elon Musk** (May 6*th*, 2020): Elon Musk's most liked and retweeted tweet.

- **"Me in my sick new car (left him the money)" - Elon Musk** (April $3rd$, 2021): Elon Musk's most replied-to tweet. The tweet also has a photograph of a car he had recently purchased.

3.4 BSTS Analysis

The experiments in this thesis used a Bayesian Structural Time Series model to infer the causal impacts of various events (treatments) on Bitcoin, Dogecoin, and Cardano. Before the model could be applied the first step in the procedure was to gather and format the data so it could be properly fed into the model. The format required a dataframe with dates in the first column, the treatment variable in the second column, and controls in the remaining columns. Additionally, based on the pre- and post-period values the dates needed to be filtered out to contain the proper time period. Appendix B.1 showcases the functions that were developed to generate the input dataframe. Once the preprocessing of the data was complete, the dataframe was able to be fed into the causal impact model along with inputs for the treatment date, number of days before the event, and number of days after the event.

3.4.1 Python Implementation

Several different Python packages were considered for the implementation. It was found that the *tfcausalimpact* Python package developed by Willian Fuk [3] most closely resembled the Causal Impact R package created by Google. This Python package has nearly all the same features and outputs of the R package. Unfortunately, one crucial feature the Python package does not have is the ability to get the probability of controls. This is a slight limitation of the experiment and is discussed in more detail in the future work section.

Using this Python library a pipeline was developed to easily apply the BSTS model on various events relating to cryptocurrency. The inputs required are a dataframe with the treatment and control variables, a treatment date, a number of days before, and a number of days after. This allows a computer to be able to perform an experiment on the data based on any event and time frame, as long as the time frame is within the data. Within the pipeline based on inputs, the data is masked to create separate datasets for the pre- and post-periods. Then after the data is masked, it is fed into the *CausalImpact* function from the library. After the function is run, the p-value and control weights are analyzed. For all experiments, the α chosen was 0.05. The control weights showcase how much each control contributed to the prediction of the counterfactual. If the weight of a control is very small, then it is most likely a useless control for the experiment.

When using the default parameters of the *CausalImpact* function, it was found that slightly different results were being outputted for identical experiments. This was a particular concern in some cases when the function would jump between a significant and insignificant p-value. Changing the method used to fit the Bayesian model was the best solution to remedy this error. The default fit method used in the function is Variational Inference (VI), but has the option to use Hamiltonian Monte Carlo (HMC), which is a fit method that takes more samples and iteration. The HMC method leads to more precise results, but at the cost of a longer run time. This study ran a comparison between HMC and VI and discovered that HMC works better for the specific world events in this experiment. In most cases this only increased run time by a few minutes, which was not a significant issue. The run time difference becomes more noticeable when analyzing effects over larger time frames.

The *CausalImpact* function also outputs a summary with various numerical statistics and a graphical visualization of the causal impact. The summary allows for users to be able to see numerically how much of an impact an event had on the treatment variable. Another useful feature of this library provides the ability to generate the summary into text and gives users a full textual analysis report.

3.4.2 Varying Time Frames

For the world events three different pre- and post-event time periods were taken:

- **Short Time Frame**: 45 days before, 30 days after
- **Medium Time Frame**: 100 days before, 70 days after
- **Long Time Frame**: 150 days before, 120 days after

When preforming an analysis on the global events, results did not show major differences between the medium and long time frames. As a result, for the Elon Musk tweets only two different time frames were analyzed for each experiment:

- **Short Time Frame**: 45 days before, 30 days after
- **Long Time Frame**: 100 days before, 70 days after

4 Results

For every experiment done, the causal impact library outputs a **posterior probability**, a **p-value**, and statistical results. The table below summarizes all the statistical values that which are outputted:

	Average	Cumulative
Actual	#	#
Prediction (s.d) 95% Confidence Interval	# (#) [#,#]	# (#) [#,#]
Absolute effect (s.d) 95% Confidence Interval	# (#) [#,#]	# (#) [#,#]
Relative effect (s.d) 95% Confidence Interval	% (%) [%,%]	% (%) [%,%]

For the purposes of this experiment, the actual and predicted values are enough to understand the impact of the treatments. When summarizing the results from each experiment, the following outputs are provided: p-value, posterior probability of effect, actual average after treatment, and predicted average after treatment.

4.1 Facebook Announces Libra (June 17*th*, 2019)

Figure 6: Bitcoin prices around the time Facebook announced Libra. Towards the end of June there is a peak of about $13,000 then the price drops for the reminder of the year. On June 17*th*, 2019 the price of Bitcoin was about $9,500

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (VI)	0.0	100%	\$11001.63	\$7712.81 (342.89)
Short (HMC)	0.0	100%	\$11001.63	\$8186.03 (384.67)
Medium (VI)	0.0002	99.9%	\$10756.04	\$8204.01 (630.12)
Medium (HMC)	0.0	100.0%	\$10756.04	\$6956.81 (217.59)
Long (VI)	0.0149	98.5%	\$10162.68	\$8341.54 (836.63)
Long (HMC)	0.001	99.9%	\$10162.68	\$6378.55 (752.85)

Stocks	Short	Medium	Long
Dow Jones	0.54	0.39	0.15
Google	-1.27	-1.24	-0.70
Apple	-1.24	-1.01	-0.71
Tesla	-0.39	-1.33	-1.24
Twitter	-1.21	-0.62	-0.18
Amazon	1.33	1.21	0.66

Table 1: Results table when applying Facebook announcing Libra as the treatment. The second table has the average control weights over 3 runs with the HMC fit method. Facebook is not included in the control weights table, due to potential effects from the treatment

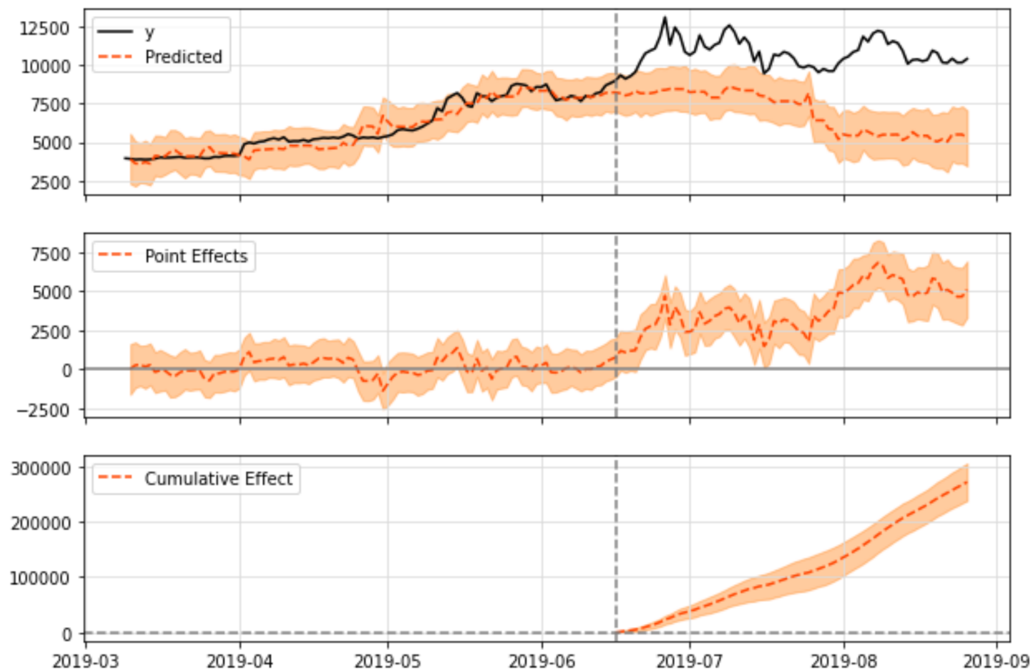

Figure 7: Causal impact output for Bitcoin prices with the date of Facebook announcing Libra as the treatment

As seen in Table 1 and Figure 7 there is a significant impact of the date Libra was announced on the price of Bitcoin. In general, based on the controls the expected average Bitcoin price was significantly greater. The longest time frame had the highest standard deviation for the predicted average, which means the model is not as sure about the long term values compared to shorter time frames. When comparing the model fit methods, the HMC method had a smaller standard deviation, which means the predicted results are more accurate.

This event having a significant impact is expected because this was major in the crypto industry. There was nothing else major that occurred during this time relating to the crypto, so this event likely results in the Bitcoin price shifts seen.

4.2 The Onset of the COVID-19 Pandemic(February 14*th*, 2020)

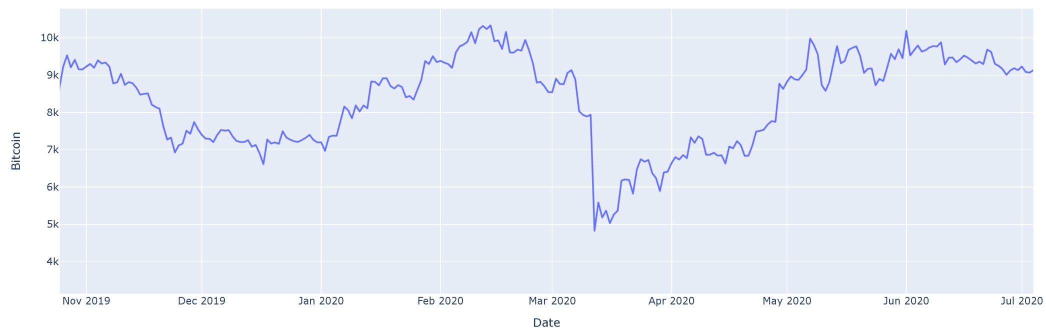

Figure 8: Bitcoin prices around the time frame of everything shutting down due to COVID-19. Prices go steady until a sudden drop around mid-February. On February 14*th*, 2020 the price of Bitcoin was about $10,300.

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (VI)	0.3377	66.23%	$8618.3	$8786.71 (647.12)
Short (HMC)	0.2607	79.9%	$8618.3	$9115.47 (576.75)
Medium (VI)	0.0020	99.8%	$7511.59	$9812.6 (616.05)
Medium (HMC)	0.0259	97.4%	$7511.59	$9158.39 (837.67))
Long (VI)	0.0300	97.0%	$8216.39	$10056.88 (976.81)
Long (HMC)	0.0370	96.3%	$8216.39	$9948.23 (944.82)

Stocks	Short	Medium	Long
Dow Jones	-0.64	-0.25	0.09
Google	1.13	1.16	-0.11
Apple	1.07	0.28	0.10
Facebook	-0.44	-0.26	0.71
Tesla	0.94	0.73	0.43
Twitter	0.76	0.41	-0.16
Amazon	0.37	0.55	0.61

Table 2: Results table when applying everything shutting down due to COVID-19 as the treatment. The second table has the average control weights over 3 runs with the HMC fit method

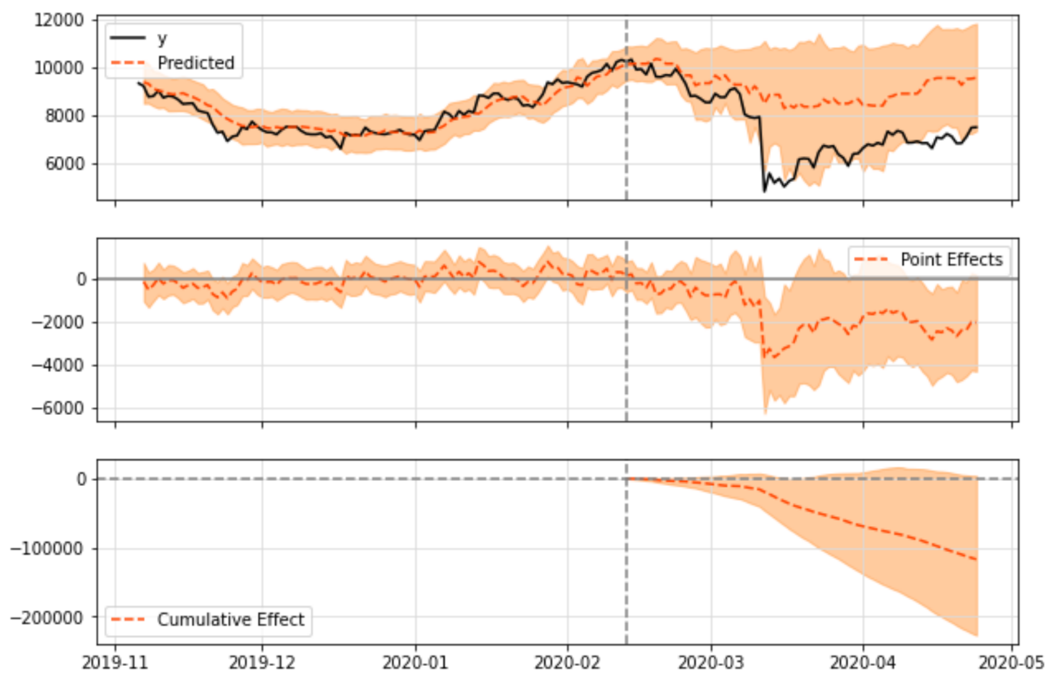

Figure 9: Causal impact output for Bitcoin prices with the date of everything shutting down due to COVID-19 as the treatment

As seen in Table 2 and Figure 9 there is a significant impact only for the medium and long time frames. The varying results make sense due to the fact that the stock market was affected by the pandemic. This would make all the controls used in the experiment very poor. In general, for this time it is difficult to find good controls because pretty much everything was affected by COVID-19 in some way. Thus, using this date for the experiment is a poor choice.

4.3 El Salvador Adopts Bitcoin (September 7*th*, 2021)

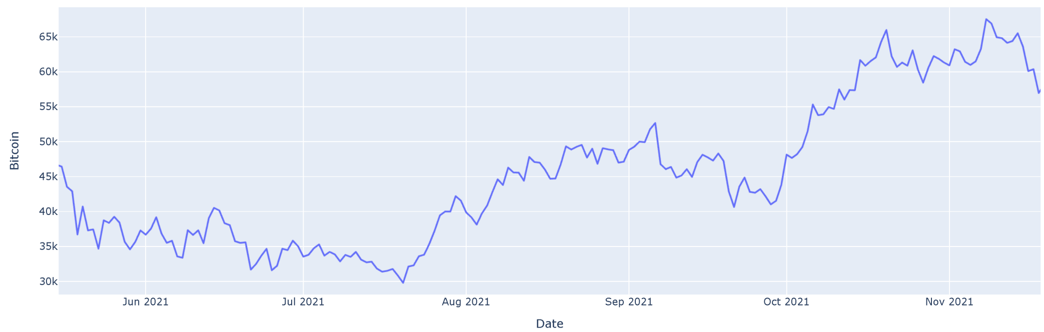

Figure 10: Bitcoin prices around the time El Salvador announced they would adopt Bitcoin as a national currency. On September 7*th*, 2021 the price of Bitcoin was about $46,800.

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (VI)	0.0270	97.3%	$46098.45	$49678.18 (1415.13)
Short (HMC)	0.0679	93.2%	$46098.45	$49404.91 (1451.6)
Medium (VI)	0.0489	95.1%	$54734.66	$49354.02 (3456.48)
Medium (HMC)	0.1409	85.9%	$54734.66	$51776.07 (3297.4)
Long (VI)	0.4575	54.3%	$53300.8	$52661.46 (6445.15)
Long (HMC)	0.2398	76.0%	$53300.8	$58316.92 (7538.01)

Stocks	Short	Medium	Long
Dow Jones	0.04	0.02	0.48
Google	0.60	1.22	0.15
Apple	0.29	0.05	0.12
Facebook	0.36	1.29	0.53
Tesla	0.84	0.86	0.93
Twitter	-1.38	-1.15	-0.68
Amazon	-0.61	-1.08	0.07

Table 3: Results when applying El Salvador adopting Bitcoin as the treatment

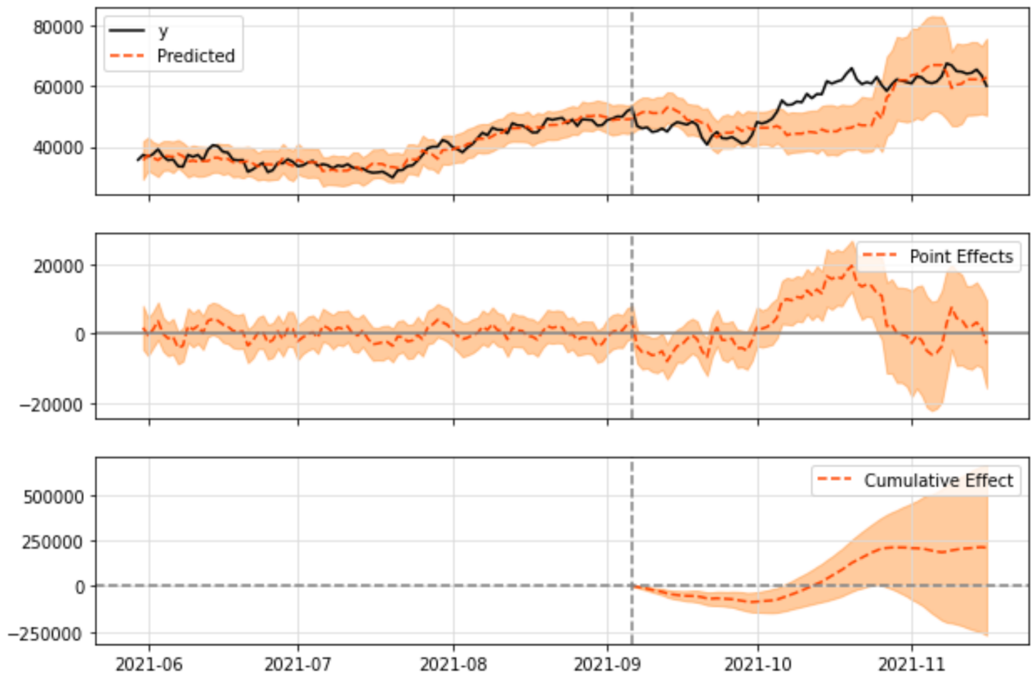

Figure 11: Causal impact output for Bitcoin prices with El Salvador adopted Bitcoin as the treatment. The second table has the average control weights over 3 runs with the HMC fit method

As seen in Table 3 and Figure 11 there is no significant impact of El Salvador adopting Bitcoin on the price of Bitcoin. The p-value for every time frame is less than the α value (0.05), except for short and medium (VI). A possible explanation for this is that the VI fit method is less accurate and p-value outputs are more variable compared to the HMC method. Just as before, as the time frame increases, the predicted average is less accurate. For this event it is clear that the HMC fit method works better than the VI method.

The overall result of this event not having a significant impact is somewhat surprising because cryptocurrency tends to have fluctuations in price, especially for major events such as this. A possible explanation for why the model did not pick up a significant effect on this date is because there could be other events at the time that are counteracting the treatment. For instance, around this time China started to crack down on crypto [25].

4.4 "... going to moon very soon" - Elon Musk (April 10*th*, 2021)

Figure 12: Prices of Bitcoin, Cardano, and Dogecoin around the time frame that Elon Musk sent this tweet. The effect of this tweet is most noticeable for Dogecoin.

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (BTC)	0.0270	97.3%	$56667.21	$59271.14 (1338.47)
Long (BTC)	0.0	100.0%	$46932.41	$59979.34 (2030.1)
Short (DOGE)	0.0	100%	$0.24	$0.06 (0.0)
Long (DOGE)	0.0	100.0%	$0.35	$0.07 (0.01)
Short (ADA)	0.0160	98.4%	$1.28	$1.17 (0.05)
Long (ADA)	0.0	100.0 %	$1.52	$1.22 (0.09)

Table 4: Results when applying this tweet as the treatment

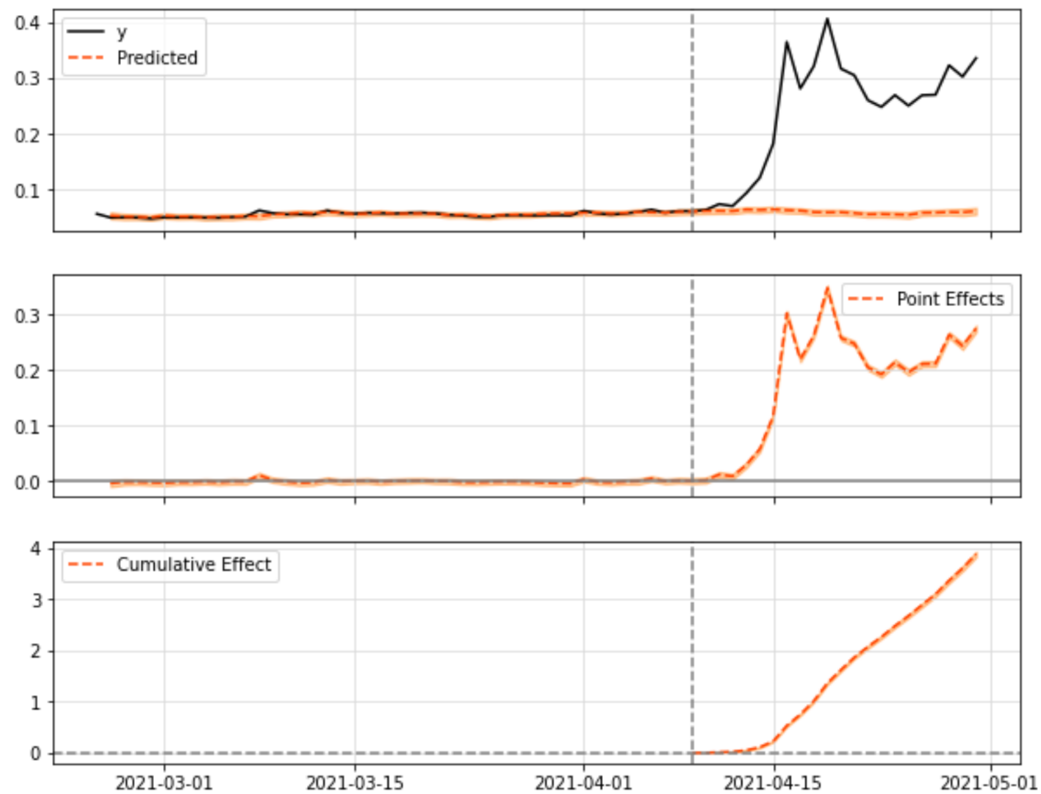

Figure 13: Causal impact output for Doge prices with this treatment

As seen in Table 4 and Figure 13 Elon Musk's tweet had a significant impact on all three coins being analyzed. The effect is the largest for Dogecoin as the actual average price was much larger than the predicted average. This is likely because the tweet is directed towards Dogecoin. This result does not come as a surprise because this tweet is famous for causing people to invest in Dogecoin as a joke. During this time Elon Musk also made his voice heard through more tweets and talk shows. This tweet alone might not be the reason for the price explosions, but this time period, where Musk was actively supporting crypto, definitely had a causal impact on crypto prices.

4.5 “Cryptocurrency is promising, but please invest with caution!” - Elon Musk (May 7*th*, 2021)

Figure 14: Prices of Bitcoin, Cardano, and Dogecoin around the time frame that Elon Musk sent this tweet. From all the graphs there is a clear downward price trajectory around the time the tweet was made

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (BTC)	0.0	100.0%	\$42643.12	\$53237.38 (2156.81)
Long (BTC)	0.0	100.0%	\$38175.72	\$57857.95 (1358.57)
Short (DOGE)	0.0040	99.6%	\$0.42	\$0.58 (0.08)
Long (DOGE)	0.0	100.0%	\$0.33	\$0.43 (0.05)
Short (ADA)	0.0	100.0%	\$1.72	\$1.14 (0.07)
Long (ADA)	0.0220	100.0 %	\$1.53	\$1.3 (0.1)

Table 5: Results when applying this tweet as the treatment

As seen in Table 5 and Figure 15 again Elon Musk's tweet had a significant impact on all three coins being analyzed. For Dogecoin and Bitcoin the predicted counterfactual price is much higher than the actual average price. These results make sense when looking at Figure 18, the trends the coins have around May $7th$. It is important to note that when using the longer time frame for the analysis, there is an overlap with the tweet made in Section 4.4. This means that only the short term result will be abstained from other majors events.

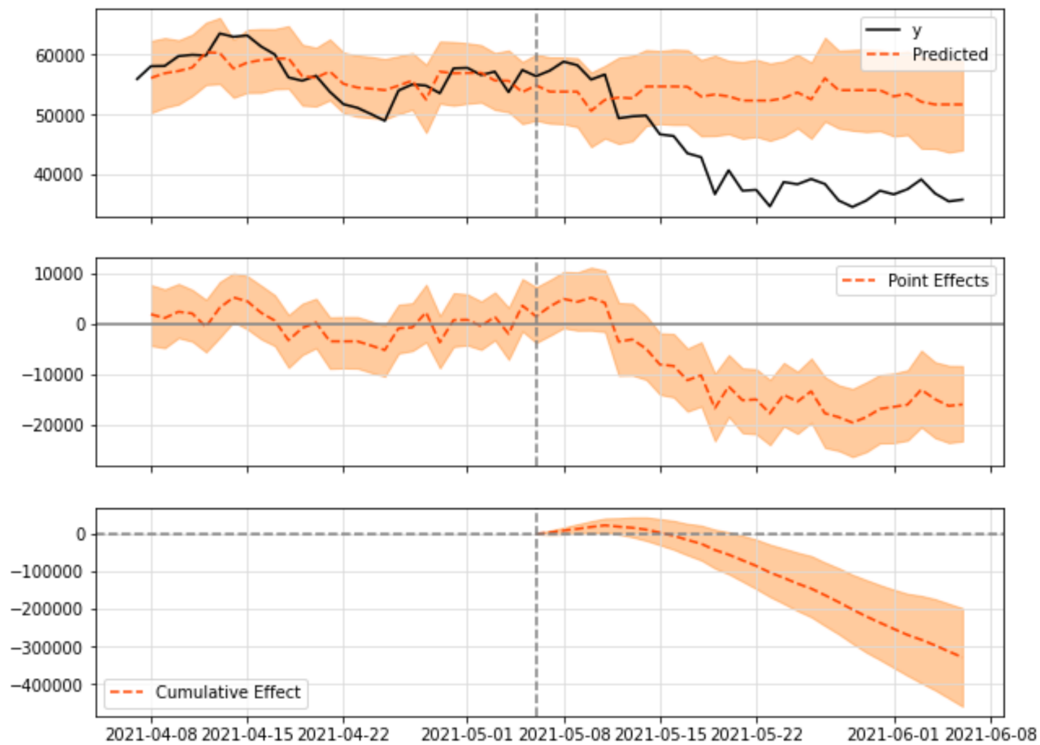

Figure 15: Causal impact output for Bitcoin prices with the date Elon Musk made this tweet as the treatment

4.6 "The coronavirus panic is dumb" - Elon Musk (May *6th*, 2020)

Figure 16: Prices of Bitcoin, Cardano, and Dogecoin around the time frame that Elon Musk sent this tweet

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (BTC)	0.0	100%	$6533.45	$7832.22 (530.47)
Long (BTC)	0.0050	99.5%	$7361.01	$8451.42 (534.71)
Short (DOGE)	0.0050	99.5%	$0.0	$0.00 (0.0)
Long (DOGE)	0.4855	51.5%	$0.0	$0.00 (0.01)
Short (ADA)	0.0220	97.8%	$0.03	$0.02 (0.0)
Long (ADA)	0.1738	82.6%	$0.04	$0.03 (0.0)

Table 6: Results when applying this tweet as the treatment

Table 6 and Figure 17 show mixed results when analyzing the causal impact with the date of this tweet being used as the treatment. For BTC both short and long time frames revealed to have a significant impact, while for the other coins only the short time frame had a significant impact. During this time period the COVID-19 pandemic was just starting to ramp up, so markets were beginning to fall. This explains why in the results actual average was much less than the predicted average.

It is important to remember that this tweet was tested to see results for a social media event not relating to crypto. Due to this, it is likely this tweet had nothing to do with the results. Most likely, around the date this tweet was published there were other significant events that caused this price shift, such as COVID-19 beginning to cause panic. This reveals a limitation of using this approach to infer causality.

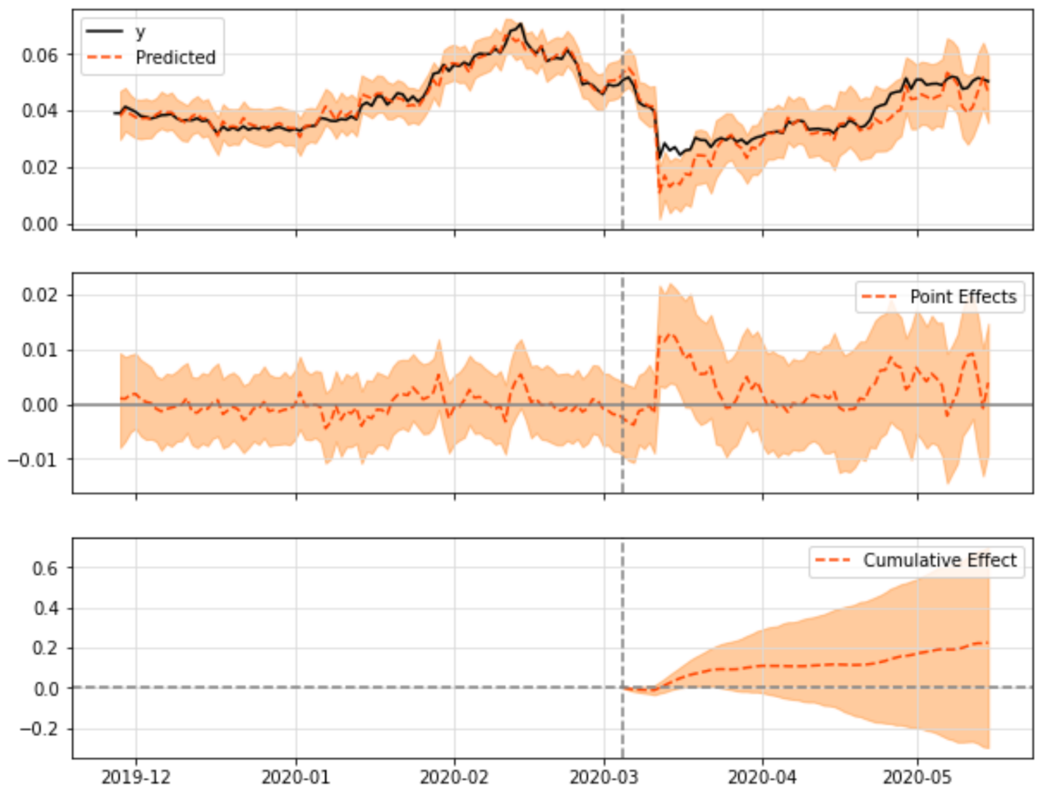

Figure 17: Causal impact output for Cardano prices with the date Elon Musk made this tweet as the treatment

4.7 “Me in my sick new car (left him the money)” - Elon Musk (April 3*rd*, 2021)

Figure 18: Prices of Bitcoin, Cardano, and Dogecoin around the time frame that Elon Musk sent this tweet

Time Frames	P-value	Posterior Prob.	Actual Avg	Predicted Avg (s.d.)
Short (BTC)	0.1519	84.8%	\$56967.3	\$59739.25 (2894.01)
Long (BTC)	0.0	100.0%	\$48862.48	\$62603.59 (1886.12)
Short (DOGE)	0.0	100.0%	\$0.22	\$0.05 (0.0)
Long (DOGE)	0.0	100.0%	\$0.33	\$0.08 (0.01)
Short (ADA)	0.2188	78.1%	\$1.27	\$1.22 (0.07)
Long (ADA)	0.0060	99.4%	\$1.50	\$1.21 (0.1)

Table 7: Results when applying this tweet as the treatment

The results in Table 7 show that the date Elon Musk made this tweet there was a significant impact for the long time frame, but not the shorter time frame. In general, the actual average prices for the coins were a lot higher than the predicted average based on the counterfactual. As seen Figure 19, the significant price shift did not occur until a few days after the treatment date. This explains why the shorter time frame did not have a significant impact for BTC and ADA.

Just like in the previous example, this tweet was simply a test and does not relate to crypto in any way. The results are likely due to events that follows a few days after this tweet, where Elon Musk was actively supporting crypto and joking about Dogecoin. The analysis for this date was done in section 4.4 and a significant impact was found. Due to this tweet occurring too close to the tweet made on April 10*th*, the results are likely a shadow of this larger event. This again showcases one of the limitations of using the BSTS model to determine causality.

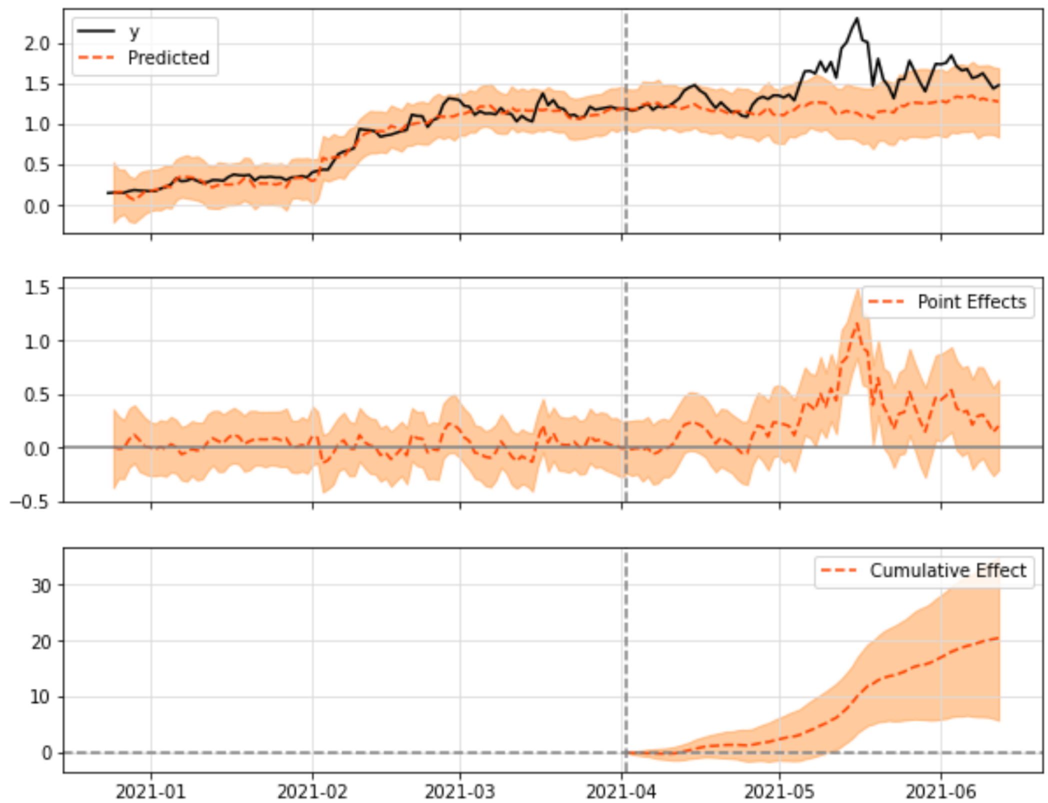

Figure 19: Causal impact output for Cardano prices with the date Elon Musk made this tweet as the treatment

5 Conclusions and Future Work

5.1 Conclusions

The BSTS model is an extremely powerful tool for measuring the causal impact of events on market prices. The results of this thesis demonstrate the viability of this approach on volatile data such as cryptocurrency. The model offers flexibility and pertinent information detailing the impact of a treatment. While more improvements can be made to the Python library, the process to get results from an analysis is simple and quick.

In general, the results for each event studied were easy to understand. When comparing the fit methods for the model, the Hamiltonian Monte Carlo (HMC) proved to be more precise and accurate. The HMC fit method was also not as punishing as expected, as the run time was only a few minutes slower for each event. If a much larger time frame had been used, the slower run time with the HMC fit method would have been more noticeable. Fortunately, large time frames should not be used often when using this approach because the precision goes down as the time frame increases.

These results showcase some limitations of the BSTS model and the Python library. The analyses done in sections 4.6 and 4.7 demonstrate that significant events occurring close to the treatment date could have an impact, which could cause misleading results. Additionally, the BSTS model assumes that the controls were not affected by the treatment. The original R causal impact package developed by Google offered a way to check the probability of a control being good. Unfortunately, the Python version has yet to implement this feature, so there is no efficient way to validate controls. Using the Python package, the best way to validate the controls was to look at the weights and test out dropping controls.

5.2 Future Work

There are a few improvements that could be made to the overall pipeline of inferring the causal impact of an event in the crypto market. The causal impact Python package offers a few more parameters and features that were not explored in this thesis. One parameter that could be useful to look into is the option to add seasonality. Telling the model that there are seasons in the data provided could be helpful and lead to more accurate results. There are a few seasonality effects seen in the crypto market that could be useful to investigate, such as "Crypto Winter".

In addition, finding more valid controls would be an impactful way to get a more accurate outcome. In the future, when the control probability feature is added to the Python package, it will be a lot easier to pick and choose controls for each analysis. In this thesis all the controls used were different stock prices, however, it might be a good idea to look into other controls, such as Google searches.

Once the pipeline for finding causal impacts on the crypto market is optimized, an analysis could be done on many more events, which has the possibility to unveil many of new connections. This will eventually lead to a better understanding of the crypto market and help analysts forecast prices more accurately.

An interesting experiment can be done by finding the causal impact whenever Elon Musk sends a tweet. Using a natural language processing (NLP) technique, tweets could be organized into positive or negative tones. Based on this, researchers could use the analysis techniques demonstrated in this thesis to see if the tone of Elon Musk's tweet has some correlation to the causal impact on the price of Bitcoin, or any altcoin.

A Data Samples

A.1 Raw Bitcoin Data Sample

Date	Price	Open	High	Low	Vol.	Change %
Jun 30, 2022	19,089.90	20,111.30	20,155.10	18,786.70	110.22K	-5.08%
Jun 29, 2022	20,111.30	20,278.00	20,415.80	19,880.70	98.97K	-0.82%
Jun 28, 2022	20,278.00	20,727.90	21,200.20	20,210.50	83.83K	-2.18%
Jun 27, 2022	20,730.20	21,037.70	21,497.50	20,568.00	84.69K	-1.49%
Jun 26, 2022	21,043.50	21,489.90	21,837.40	20,989.70	67.21K	-2.08%
Jun 25, 2022	21,489.90	21,226.90	21,553.40	20,917.50	66.63K	1.24%
Jun 24, 2022	21,226.90	21,100.90	21,488.70	20,743.90	104.56K	0.60%
Jun 23, 2022	21,100.70	19,967.30	21,189.80	19,907.50	110.70K	5.68%
Jun 22, 2022	19,965.80	20,720.20	20,864.30	19,780.20	125.94K	-3.64%
Jun 21, 2022	20,720.40	20,571.60	21,689.30	20,371.70	136.32K	0.72%
Jun 20, 2022	20,572.30	20,576.90	20,996.80	19,658.80	140.60K	-0.02%
Jun 19, 2022	20,577.20	18,983.40	20,763.50	17,983.70	169.39K	8.38%
Jun 18, 2022	18,986.50	20,446.40	20,744.70	17,630.50	267.25K	-7.13%
Jun 17, 2022	20,444.60	20,391.30	21,315.40	20,244.10	136.27K	0.28%
Jun 16, 2022	20,386.60	22,577.90	22,942.10	20,231.10	144.00K	-9.71%
Jun 15, 2022	22,577.90	22,137.50	22,754.40	20,125.80	280.41K	1.90%
Jun 14, 2022	22,157.30	22,449.10	23,200.30	20,860.90	251.01K	-1.29%
Jun 13, 2022	22,448.00	26,606.30	26,857.60	22,006.30	379.26K	-15.63%
Jun 12, 2022	26,606.30	28,404.00	28,534.80	26,606.30	120.02K	-6.33%
Jun 11, 2022	28,403.40	29,083.30	29,426.60	28,161.80	82.92K	-2.34%
Jun 10, 2022	29,083.30	30,097.40	30,325.60	28,884.90	104.26K	-3.37%
Jun 09, 2022	30,097.80	30,202.10	30,691.40	29,953.80	61.72K	-0.34%
Jun 08, 2022	30,201.60	31,127.20	31,312.10	29,874.80	87.96K	-2.98%
Jun 07, 2022	31,128.80	31,370.30	31,556.60	29,235.00	140.32K	-0.76%
Jun 06, 2022	31,367.60	29,911.20	31,753.40	29,888.60	94.01K	4.86%

A.2 Merged Crypto Prices Dataframe Sample

Date	Bitcoin	Ethereum	XRP	Cardano	Solana	Dogecoin
2019-01-01	3809.4	139.61	0.36326	0.042		0.002347
2019-01-02	3873.8	152.95	0.37138	0.0443		0.00235
2019-01-03	3780.1	146.94	0.35434	0.042		0.002293
2019-01-04	3802.7	152.86	0.3556	0.0431		0.002243
2019-01-05	3785.4	153.49	0.35106	0.044		0.002249
2019-01-06	4004.1	154.96	0.36412	0.0483		0.002257
2019-01-07	3985.9	149.96	0.36186	0.0475		0.00223
2019-01-08	3971	148.41	0.36311	0.0478		0.002223
2019-01-09	3978	148.79	0.36664	0.0514		0.002215
2019-01-10	3603.7	125.63	0.32724	0.0436		0.002079
2019-01-11	3616.5	125.15	0.32749	0.0433		0.002067

A.3 Merged Stock Prices Dataframe Sample

Date	Dow_Jones	Google	Apple	Facebook	Tesla	Twitter	Amazon
2019-01-01	23346.24	52.29	39.48	135.68	62.02	28.81	76.96
2019-01-02	23346.24	52.29	39.48	135.68	62.02	28.81	76.96
2019-01-03	22686.22	50.8	35.55	131.74	60.07	27.99	75.01
2019-01-04	23433.16	53.54	37.07	137.95	63.54	29.95	78.77
2019-01-05	23433.16	53.54	37.07	137.95	63.54	29.95	78.77
2019-01-06	23433.16	53.54	37.07	137.95	63.54	29.95	78.77
2019-01-07	23531.35	53.42	36.98	138.05	66.99	31.34	81.48
2019-01-08	23787.45	53.81	37.69	142.53	67.07	31.8	82.83
2019-01-09	23879.12	53.73	38.33	144.23	67.71	32.25	82.97
2019-01-10	24001.92	53.52	38.45	144.2	68.99	33.09	82.81
2019-01-11	23995.95	52.86	38.07	143.8	69.45	32.87	82.03

B Codes Samples

B.1 Functions

```
def generate_df(coin_to_analyze, cypro_df, stock_df):
    # based on provided coin creates a dataframe with predictors
    coin = cypro_df[['Date', coin_to_analyze]]
    df = pd.merge(coin, stock_df, on= "Date")

    return df
```

```
def ci_analysis(data, event_date, num_days_before, num_days_after):
    date_format = '20%y-%m-%d'
    analyze_date = datetime.strptime(event_date, date_format)
    # shift from requested based on inputs
    # subtract n days
    start_date = analyze_date - timedelta(days=num_days_before)
    # subtract 1 day
    before_analyze_date = analyze_date - timedelta(days=1)
    # add n days
    end_date = analyze_date + timedelta(days=num_days_after)

    # convert dates to strings
    start_date = start_date.strftime(date_format)
    before_analyze_date = before_analyze_date.strftime(date_format)
```

```python
    end_date = end_date.strftime(date_format)

    pre_period = [start_date, before_analyze_date]
    post_period = [event_date, end_date]

    # filter dataset for the proper time range
    mask = (data['Date'] >= start_date) & (data['Date'] <= end_date)
    filtered_data = data.loc[mask]
    filtered_data = filtered_data.set_index('Date')
    filtered_data.sort_index(inplace=True)
    # BIG DIFFERENCE IS NOW USING HMC TO FIT MODEL
    ci = CausalImpact(filtered_data, pre_period, post_period,
        model_args= {'fit_method': 'hmc'})

    return ci
```

```python
def ci_get_weights(ci, orignal_data):
    result_stats = dict()
    weights = np.mean(np.array(ci.model_samples[6]), axis= 0)
    item = {'Weights': weights}
    result_stats.update(item)

    weights_df = pd.DataFrame(result_stats,
        index=orignal_data.columns[2:])
    return weights_df
```

B.2 Sample Run

```
num_days_before = 100
num_days_after = 70
event_date = '2019-06-17'
ci_bitcoin_06_17_2019 = ci_analysis(bitcoin_analysis_df, event_date,
    num_days_before, num_days_after)
weights_v1 = ci_get_weights(ci_bitcoin_06_17_2019, bitcoin_analysis_df)
```

C Output Samples

C.1 Raw Summary Output

```
Posterior Inference {Causal Impact}
                          Average            Cumulative
Actual                    48862.48           3469236.5
Prediction (s.d.)         62603.59 (1886.12) 4444854.5 (133914.18)
95% CI                    [58899.5, 66292.94][4181864.69, 4706798.76]

Absolute effect (s.d.)    -13741.1 (1886.12) -975618.0 (133914.18)
95% CI                    [-17430.46, -10037.02][-1237562.26, -712628.19]

Relative effect (s.d.)    -21.95% (3.01%)    -21.95% (3.01%)
95% CI                    [-27.84%, -16.03%] [-27.84%, -16.03%]

Posterior tail-area probability p: 0.0
Posterior prob. of a causal effect: 100.0%

For more details run the command: print(impact.summary('report'))
Analysis report {CausalImpact}
```

C.2 Textual Summary Output

During the post-intervention period, the response variable had an average value of approx. 48862.48. By contrast, in the absence of an intervention, we would have expected an average response of 62603.59. The 95% interval of this counterfactual prediction is [58899.5, 66292.94]. Subtracting this prediction from the observed response yields an estimate of the causal effect the intervention had on the response variable. This effect is -13741.1 with a 95% interval of [-17430.46, -10037.02]. For a discussion of the significance of this effect, see below.

Summing up the individual data points during the post-intervention period (which can only sometimes be meaningfully interpreted), the response variable had an overall value of 3469236.5. By contrast, had the intervention not taken place, we would have expected a sum of 4444854.5. The 95% interval of this prediction is [4181864.69, 4706798.76].

The above results are given in terms of absolute numbers. In relative terms, the response variable showed a decrease of -21.95%. The 95interval of this percentage is [-27.84%, -16.03%].

This means that the negative effect observed during the intervention period is statistically significant. If the experimenter had expected a positive effect, it is recommended to double-check whether anomalies in the control variables may have caused an overly optimistic expectation of what should have happened in the response variable in the absence of the intervention.

The probability of obtaining this effect by chance is very small (Bayesian one-sided tail-area probability p = 0.0). This means the causal effect can be considered statistically significant.

References

[1] K. Brodersen, F. Gallusser, J. Koehler, N. Remy, and S. Scott, “Inferring causal impact using bayesian structural time-series models,” The Annals of Applied Statistics, vol. 9, 2015, http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/41854.pdf.

[2] Brodersen, Kay H., and Alain Hauser. “CausalImpact.” Github, Google, Inc., 2017, http://google.github.io/CausalImpact/.

[3] Fuks, Willian, et al. “Tfcausalimpact.” Github, v0.0.12, 2022, https://github.com/WillianFuks/tfcausalimpact.

[4] Xie, Shiqing, and Taiping Mo. “Index Futures Trading and Stock Market Volatility in China: A Difference-in-Difference Approach.” Journal of Futures Markets, vol. 34, no. 3, 27 Dec. 2013, pp. 282–297., https://doi.org/10.1002/fut.21650.

[5] Ante, Lennart. “How Elon Musk’s Twitter Activity Moves Cryptocurrency Markets.” SSRN Electronic Journal, 12 Jan. 2022, https://doi.org/10.2139/ssrn.4065313.

[6] Ciaian, Pavel, et al. “Virtual Relationships: Short- and Long-Run Evidence from Bitcoin and Altcoin Markets.” Journal of International Financial Markets, Institutions and Money, vol. 52, Jan. 2018, pp. 173–195., https://doi.org/10.1016/j.intfin.2017.11.001.

[7] Alfieri, Elise. "Long and Short-Term Impacts of Regulation in the Cryptocurrency Market." The Quarterly Review of Economics and Finance, vol. 81, Aug. 2021, pp. 157–173., https://doi.org/10.1016/j.qref.2021.05.005.

[8] Narman, Husnu S., and Alymbek Damir Uulu. "Impacts of Positive and Negative Comments of Social Media Users to Cryptocurrency." 2020 International Conference on Computing, Networking and Communications (ICNC), Feb. 2020, https://doi.org/10.1109/icnc47757.2020.9049693.

[9] Schwerdt, Guido, and Ludger Woessmann. "Empirical Methods in the Economics of Education." The Economics of Education, 2020, pp. 3–20., https://doi.org/10.1016/b978-0-12-815391-8.00001-x.

[10] Sharma, Rakesh. "Is There a Cryptocurrency Price Correlation to the Stock Market?" Investopedia, Investopedia, 4 Oct. 2022, https://www.investopedia.com/news/are-bitcoin-price-and-equity-markets-returns-correlated

[11] Aggarwal, Neema. "Inferring the Causal Impact of Super Bowl Marketing Campaigns Using a Bayesian Structural Time Series Model." THE COOPER UNION ALBERT NERKEN SCHOOL OF ENGINEERING, 2017. 1

[12] Agarwal, Anjali, et al. "Bayesian Variable Selection with Spike-and-Slab Priors." The Ohio State University, 2016.

[13] Olteanu, Andrada. "All Elon Musk's Tweets." https://www.kaggle.com/Datasets/Andradaolteanu/All-Elon-Musks-Tweets, 2020.

[14] Investing.com. “Cryptocurrency Data.” https://www.investing.com/Crypto/.

[15] Oosterbaan, Edward. “The Elon Effect: How Musk’s Tweets Move Crypto Markets.” CoinDesk Latest Headlines RSS, CoinDesk, 14 Dec. 2021, https://www.coindesk.com/layer2/culture-week/2021/12/14/the-elon-effect-how-musks-tweets-move-crypto-markets

[16] Shahzad, Syed Jawad, et al. “Price Explosiveness in Cryptocurrencies and Elon Musk’s Tweets.” Finance Research Letters, vol. 47, June 2022, p. 102695., https://doi.org/10.1016/j.frl.2022.102695.

[17] Fuks, Willian. “Pycausalimpact” Pypi, 0.1.1, 11 May 2020, https://pypi.org/project/pycausalimpact/. Accessed 2022.

[18] He, Pinglin, et al. “Covid–19’s Impact on Stock Prices across Different Sectors—an Event Study Based on the Chinese Stock Market.” Research on Pandemics, 2021, pp. 66–80., https://doi.org/10.4324/9781003214687-6.

[19] Kristoufek, Ladislav. “Will Bitcoin Ever Become Less Volatile?” Finance Research Letters, vol. 51, Jan. 2023, p. 103353., https://doi.org/10.1016/j.frl.2022.103353.

[20] Auer, Raphael, and Stijn Claessens. “Cryptocurrency Market Reactions to Regulatory News 1.” The Routledge Handbook of FinTech, 2021, pp. 455–468., https://doi.org/10.4324/9780429292903-31.

[21] Delcid, Merlin. “El Salvador Buys Bitcoin as the Digital Currency Becomes Legal Tender — CNN Business.” CNN, Cable News Network, 7 Sept. 2021, https://www.cnn.com/2021/09/06/business/bitcoin-price-el-salvador-intl-hnk/index.html.

[22] Morris, Stephen Tödtling-Schönhofer, Herta Wiseman, Michael. (2013). Design and Commissioning of Counterfactual Impact Evaluations - A Practical Guidance for ESF Managing Authorities. 10.2767/94454.

[23] Ho, S.L., and M. Xie. “The Use of Arima Models for Reliability Forecasting and Analysis.” Computers amp; Industrial Engineering, vol. 35, no. 1-2, 1998, pp. 213–216., https://doi.org/10.1016/s0360-8352(98)00066-7.

[24] “Googlefinance.” Google Docs Editors Help, Google, https://support.google.com/docs/answer/3093281?hl=en.

[25] John, Alun, et al. “China’s Top Regulators Ban Crypto Trading and Mining, Sending Bitcoin Tumbling.” Reuters, Thomson Reuters, 24 Sept. 2021, https://www.reuters.com/world/china/china-central-bank-vows-crackdown-cryptocurrency-trading-2021-09-24/.