2023 S.T. Yau High School Science Award

Research Report

The Team

Name of team member: Jonathan Jin

School: Phillips Academy Andover

City, Country: Andover, USA

Name of supervising teacher: Xiangyu Shi

Job Title: PhD candidate

School/Institution: Yale university

City, Country: New Haven, USA

Title of Research Report

Market sentiment and cryptocurrency prices: The case of Twitter

Date

08/19/2023

Market sentiment and cryptocurrency prices: The case of Twitter

Jonathan Jin*

August 20, 2023

Abstract

In this paper, we study the effects of market sentiments on the price of cryptocurrency, using Ethereum as a case and point. We find that market sentiments, especially positive ones, have positive and significant effects on the price. We also explore the relevant interactions. We find that the COVID pandemic, oil prices, the stock market, politicians' Twitter, and currency exchange rates are related interactions at play.

JEL Codes: D4, D9, E22, E6, E7, G1, G4
Keywords: Cryptocurrency, Twitter, Sentiment, Rationality

*I thank Dr.Shi for his guidance on this project

Declaration of Academic Integrity

The participating team declares that the paper submitted is comprised of original research and results obtained under the guidance of the instructor. To the team's best knowledge, the paper does not contain research results, published or not, from a person who is not a team member, except for the content listed in the references and the acknowledgment. If there is any misinformation, we are willing to take all the related responsibilities.

Names of team members: Jonathan Jin

Signatures of team members Jonathan Jin

Name of the instructor: Xiangyu Shi

Signature of the instructor Xiangyu Shi

Date: 19/08/2023

Contents

1 Introduction

Cryptocurrency is designed to work as a method of exchange through a decentralized network free from the influence of banks and governments to uphold or maintain it. In this paper, we study how market sentiments reflected by social media, specifically Twitter, impact cryptocurrency prices, using Ethereum as a case and point. We also explore relevant interactions including COVID deaths, oil prices, politician's Twitter, the stock market, and currency exchange rates.

This research question addresses the behavior of a relatively new market. Cryptocurrency is one of the first digital and unregulated currencies. Countries with unstable currencies and economies have started adapting to accepting cryptocurrency. A variety of modern companies accept cryptocurrency as a form of payment. It is important for the market to understand how cryptocurrency prices react to external stimuli. Furthermore, social media provides a good medium through which to study this price, due to the younger user base as well as its common usage among the cryptocurrency community. By conducting this study, we can gain a better understanding of first how cryptocurrency markets react to different stimuli, but also how social media reflects upon that change and sentiment within that market.

The empirical results show that positive word count is positively correlated with price. Negative word count is also positively correlated with price. While the results may not be fully robust, the positive market sentiment is still positively correlated to price. In regards to interactions, results of regressions on COVID deaths indicate that the partial effects of positive word count are more positive when there are more COVID deaths in the United States. Regarding oil prices, the coefficient on the interaction term between positive word count and oil cost is positive and statistically significant. Our results with the stock market using the Dow Jones index as a case and point show that the coefficient for the interaction term between the positive word count and the closing price of the DOW Jones index is statistically significant and positive.

This paper speaks to four lines of literature. First, this paper speaks to the literature on cryptocurrency. Liu and Tsyvinski (2021) establish that cryptocurrency returns are driven and can be predicted by factors that are specific to the cryptocurrency market. Market sentiments, which are manifested by social media such as Twitter, are one of the relevant factors. Liu et al. (2022) find that three factors, cryptocurrency market, size, and return, capture the cross-sectional cryptocurrency returns. Market sentiment is a factor that belongs to the cryptocurrency market. Moreover, some papers study several specific cryptocurrencies, such as Bitcoin. For example, John et al. (2022) provides a survey of the literature on the economics of Bitcoin and blockchain fundamentals with particular emphasis on Bitcoin, proof of work, and proof of state. Griffin and Shams (2020) investigate whether Tether, a digital currency pegged to the US dollar affected Bitcoin and other cryptocurrency prices during the 2017 boom. On the theoretical side, Cong et al. (2021) developed a dynamic asset pricing model of cryptocurrencies that allowed users to conduct peer-to-peer transactions on digital platforms. This paper contributes to this line of the literature by taking into consideration the effects of market sentiments reflected in social media on the price of cryptocurrencies.

Second, this paper is related to the literature on sentiment and the market. Liu (2015) examines how investor sentiment affects the time-series variation in stock market liquidity. This paper uses two measures of sentiment, a direct one and an indirect one. The results show that the stock market is more liquid when investor sentiment is higher. Bandopadhyaya and Jones (2006) develop a measure that can be used in a stock market setting by studying the price movements of a group of firms that represent a stock market index. News events that affect the underlying study are quickly captured by changes in this measure of investor sentiment, and the sentiment measure is capable of explaining a significant proportion of the changes in the stock market index. Compared to this paper, my paper measures sentiment using Twitter data. I extract the sentiment from each Tweet using NLPs, which is different from the existing literature. On the theoretical side, some papers explain why sentiments matter for the financial market and the macroeconomy. For example, Benhabib et al. (2016) study how financial information frictions can generate sentiment-driven fluctuations in asset prices and self-fulfilling business cycles. In addition, Maxted (2023) incorporates diagnostic expectations into a general equilibrium macroeconomic model with a financial intermediary sector. The model features sentiment-driven financial crises characterized by low pre-crises risk premia and neglected risk. In this area of the literature, no paper so far has examined the effects of sentiments on the price of cryptocurrencies. Therefore, my paper fills the gap.

Third, this paper builds on the literature surrounding social media. Allcott et al. (2020) exploit a randomized experiment and find that social media has significant effects on people's welfare and political outcomes. Deactivating Facebook for the four weeks before the 2018 US midterm election induced people to participate more in offline activities and socialize more with family and friends as well as increase their subjective well-being. Levy (2021) estimates the effects of social media exposure by conducting a large field experiment randomly offering participants subscriptions to conservative or liberal news outlets on Facebook. He finds that random variation in exposure to news on social media substantially affects the slant of news sites that individuals visit. The results suggest that social media algorithms may limit exposure to counter-attitudinal news and thus increases polarization. Braghieri et al. (2022) provide quasi-experimental estimates of the impact of social media on mental health by leveraging a unique natural experiment: the staggered introduction of Facebook across U.S. colleges. Using a generalized difference-in-differences empirical strategy, they find that the rollout of Facebook at a college increased symptoms of poor mental health, especially depression. Additional evidence on interactions suggests that the results are due to Facebook fostering unfavorable social comparisons. Some other papers also measure sentiment using Twitter data. For example, Polyzos (2022) proposes the use of social media information as a real-time decision-making tool for significant events, using the war in Ukraine as a case study. He proxies the public's perception of the progression of events using sentiment analysis on 42 million tweets and calculates impulse response functions on 5-min data for 15 economic and financial indicators (stock markets, commodities, interest rates, currencies, and cryptocurrencies). Similar to this paper, I also examine the effects of sentiments on Twitter on the financial market, especially on the prices of cryptocurrencies. However, I am using U.S. data.

Finally, this paper contributes to the literature on behavioral finance. Barberis and Thaler (2003) provides an excellent survey of behavioral finance. They argue that some financial phenomena can be plausibly understood using models in which some agents are not fully rational. Twitter and the sentiments in the tweets embody such irrationality. Hirshleifer (2015) argues that the frontier of behavioral finance is to study social finance, in particular the structure of social interactions and how financial ideas spread and evolve, and how social processes affect financial outcomes. Twitter is a platform for social interaction and therefore the tweets are an outcome of such social processes that may affect financial outcomes. My paper provides evidence of the irrational behavior reflected in tweets on the financial market, especially the price of cryptocurrencies. Therefore, generally speaking, my paper contributes to the field of behavioral finance.

The remainder of this paper proceeds as follows. Section 2 discusses the background of this paper. Section 3 establishes the conceptual framework. Section 4 describes the data. Section 5 presents and analyzes the empirical results. Finally, Section 6 concludes this paper.

2 Background

According to Wik (2023), Cryptocurrencies are a form of digital currency designed to function on an online network completely disconnected from any central government or agency. It is mostly used to due the fact that it is a decentralized currency, and transaction, mining, and ownership data is stored in digital ledgers, an online database with strong protection. Its popularity in the modern day is due to the confidential and fast transactions that decentralized currencies offer. Figure 1 demonstrates the growth that Ethereum, the currency we chose to examine, experienced during the selected time period. Cryptocurrency differs from typical currencies due to its completely digital nature as well as its classification as a currency, security, and commodity. Bellia (2023) states that Cryptocurrency can typically only be mined in a limited amount which is typically predetermined and requires significant processing power as well as electricity in order to do so. Cryptocurrency prices are determined through the market's supply and demand, and influenced by oil prices, government regulatory changes, and the supply of mining equipment available. With a majority of investors being in the 18-25 age range, the cryptocurrency market is heavily influenced by social media, which affects market demand as well as introduces new investors. Figure 2 and Figure 3 demonstrates the time trend of positive and negative tweets, in reaction to some of these factors. In this paper, we will study a specific cryptocurrency called Ethereum, conceptually created in 2013 and created in order to inspire app and web development while using a central verification system for payments. Being a more recent coin, the effects of social media on Ethereum prices are greater compared to older coins, prompting me to choose to study this coin.

Figure 1: Time Trend for Ethereum price

Smith (2023) states that market sentiment is an overall gauge of public attitude towards a product such as a commodity or in this case cryptocurrency. It is commonly used to predict the future state of the market and provides valuable information in estimating the profitability of certain investments during different time periods. Additionally, market sentiment is also an important factor in price changes within assets. Baker and Wurgler (2007) defines investor sentiment as a belief about future cash flows and investment risks that are not justified by the facts at hand. In order to determine the effects of this sentiment, they take an investment sentiment approach that is distinctly "top-down" and macroeconomic: they take the origin of investor sentiment as exogenous and focus on its empirical effects. In this paper, will be doing a similar analysis with a comparison of Twitter's overall sentiment with its effects on the price of our given cryptocurrency, Ethereum.

The measures which this paper uses will center around the positive and negative tweets on any given day. By measuring the amount of positive and negative tweets per day as well as examining the ratio between these two different texts, we can gain insight into the sentiment of the market as well as how the market reacts to that.

Figure 2: Time Trend of Positive and Negative Words in Tweets

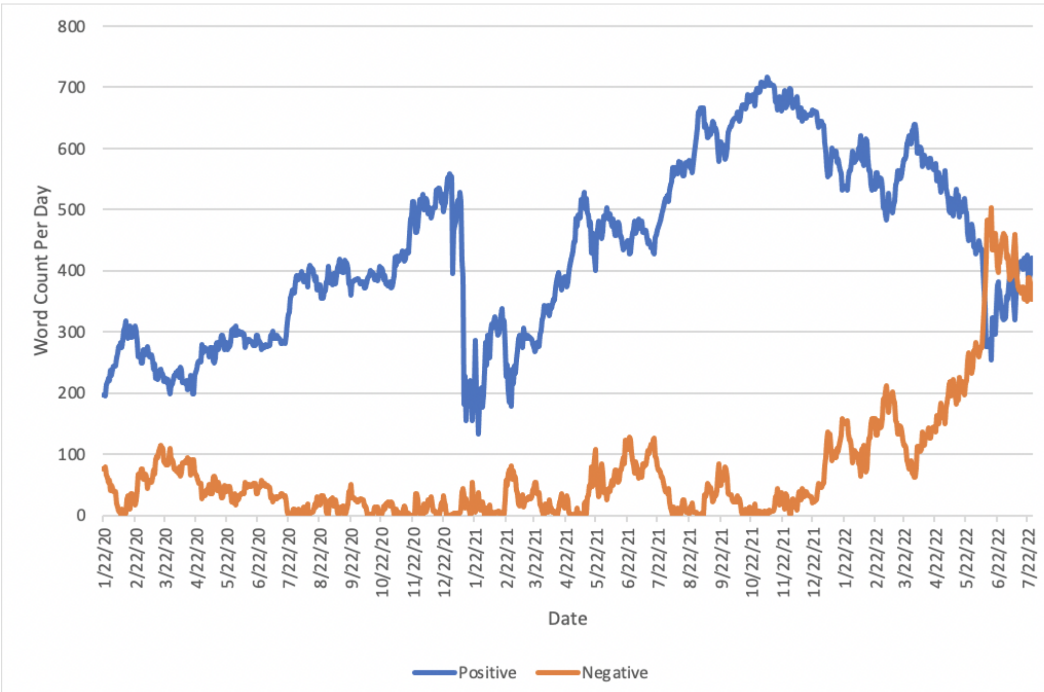
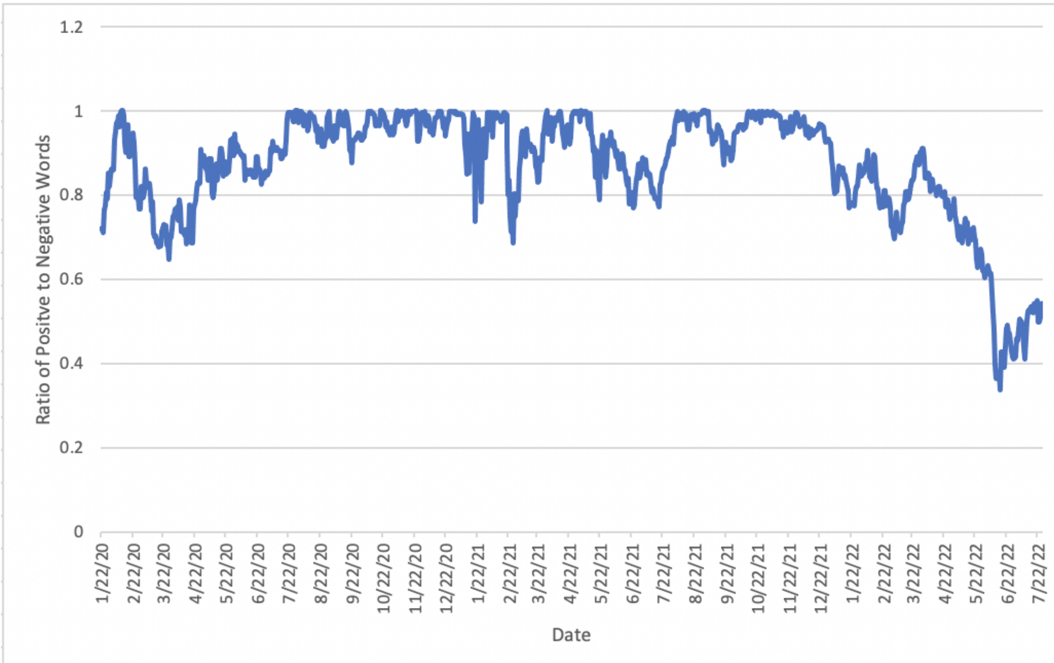


Figure 3: Time Trend of Percentage of Positive to Negative Tweets



3 Conceptual Framework

In this section, we formulate two hypotheses that can be related to the basic principles of economics. First we have Hypothesis 1.

Hypothesis 1. *Market sentiments matter for cryptocurrency price.*

Market sentiments are a way to measure the overall attitude of a certain population. Pricing is often affected by that attitude, due to the population's willingness to consume a product or spend money on something non-necessary such as cryptocurrency. As the market sentiment changes, the marginal benefit of buying cryptocurrency is lower than the marginal cost of spending money on it. This is why Hypothesis 1 holds.

Next we have Hypothesis 2.

Hypothesis 2. *Factors such as COVID deaths, oil prices, the stock market, and exchange rates are relevant interactions that may affect the effects of market sentiments on cryptocurrency price.*

Additionally, to market sentiment, there are a variety of factors that may influence that. COVID-19 had a significant impact on a variety of things ranging from commodity prices to the mood of people. Additionally, people may spend more time inside on the internet and absorbing information due to a higher risk of contracting COVID while being outside. Oil prices may increase the production cost of cryptocurrency since cryptocurrency relies on digital mining which requires significant amounts of electricity. Additionally, oil prices also may increase the cost of purchase of the electrical devices needed to produce and transport the necessary equipment to mine cryptocurrency. The overall stock market may change attitudes on the internet, due to fluctuations between bear and bull markets. Exchange rates may limit the supply of cryptocurrency available, due to the difficulty for buyers to find good opportunities to convert currency in order to purchase or sell cryptocurrency to the rest of the market. This is why Hypothesis 2 holds.

4 Data

4.1 Cryptocurrency Price

The daily price of Ethereum was collected from an online website that offers quotes for cryptocurrency prices called Polygon.io, and I also cross-referenced this data with the publicly documented closing prices on Coinbase.com. Furthermore, I also used Microsoft Excel's built-in command for searching up prices to provide another additional source of data checking to ensure the prices were accurate. The date range of this data is also January 22 of 2020 to July 28 of 2022. I define the variable price as the logarithmic value of the daily closing price of Ethereum.

4.2 Market Sentiment

The positive and negative tweets come from my own work with a Twitter scraper called SNScrape, which was distributed on Github and no longer works with the current Twitter. I installed it through GitHub and ran it through Jupyter Notebook to provide results. In order to remove bot responses and posts, I picked hashtags that were popular among cryptocurrency investors for discussion as well as picking from chat groups that cut out bot replies. I then compiled these results into a Natural Language Processor called Dovetail, which provides services online to determine sentiment with great amounts of data. I chose this Natural Language Processor specifically for its targeted use in analyzing opinions of a product, which is what a majority of the Tweets on Ethereum are from. I then uploaded the daily positive and negative word counts into Excel. The date range that this data covers is also from January 22 of 2020 to July 28 of 2022. While conducting my examination of the data, I further expanded this data into a total of five different variables. The Positive and Negative represent the logarithmic value of the daily count of positive and negative words tweeted respectively. The variable More3Positive is a dummy variable that represents whether a day has more than 3 times the positive words than negative. The variable More5Positive is a dummy variable that represents whether a day has more than 5 times the positive words than negative words. The variable More10Positive is a dummy variable that represents whether a day has more than 10 times the positive words than the negative words.

4.3 Other Variables

I obtained the daily deaths due to COVID data from an online Github database that is downloaded from the New York Times' data collection. The date range of this data is from January 22 of 2020 to July 28 of 2022. I define the variable USDeaths as the logarithmic value of the daily amount of COVID deaths. To acquire the Tweets from Trump, I went to an online database on Kaggle.com, which used two different Twitter scraping codes to collect data. This data covers the period from January 22 of 2020 to January 8 of 2021, when he was banned from Twitter. I define the variable Trcount as the logarithmic value of the daily number of Trump's tweets. The oil price and Dow Jones Index prices all came from a historical database from Investors.com, which provides accurate stock data. I define the variables Oprice and DowJones as the logarithmic values of the daily closing prices of crude oil and the Dow Jones Index respectively. The Euro to USD and RMB to USD exchange rates also came from the same database from Investors.com. The date range that this data covers is also from January 22 of 2020 to July 28 of 2022. I define the variables Europrice and Rmbprice as the logarithmic value of the daily closing exchange rate of the respective currencies with the US dollar.

5 Empirical Results

In this section, we analyze the results from the data, and create plausible explanations from the results in the tables.

5.1 Baseline Results and Robustness Checks

To begin with, we estimate the correlation between market sentiments and the price of cryptocurrency. To meet this end, we run the following specification.

$$\log Price_t = \alpha Sentiment_t + \gamma_t + u_t. \tag{1}$$

In equation 1, $\log Price_t$ is the dependent variable, the logarithm of the price of XXX at date t, $Sentiment_t$ is a measure of market sentiment, γ_t is a vector of month fixed effects, and u_t is the error term. We employ robust standard errors in all regressions.

Here are the results. Looking at the first Table 1, positive word count is positively correlated with price. Every percent increase in the positive word count leads to a 2.351 increase in the logarithmic value of price. Negative word count is also positively correlated with price. Although the negative word count is a measure of negative sentiment, it still reflects that the market is still active given a high negative word count. As a result, the price is higher.

Table 1: Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)
			log(Price)			
log(Positive WC)	2.351***		2.384***	2.932***		2.930***
	(0.0876)		(0.0912)	(0.0941)		(0.0981)
log(Negative WC)		0.0476**	0.0954***		0.148***	-0.00848
		(0.0198)	(0.0156)		(0.0265)	(0.0190)
Month Fixed Effects	N	N	N	Y	Y	Y
Observations	918	905	905	918	905	905
R-squared	0.552	0.004	0.572	0.659	0.052	0.661

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. "WC" stands for word count. * Significant at 10%, ** 5%, *** 1%.

Next, we conduct several robustness checks. In Table 2, column 1, the coefficient on the ratio of positive to negative words is positive and statistically significant. However, if we use other measures such as greater than, greater than 1.5, or greater than 2, the coefficient becomes negative. This implies that the results may not be fully robust to the choice of measurement. However, in our preferred specification, we still find that positive market sentiment, measured by the ratio of positive word count to negative word count is positively related to the price.

Table 2: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				lPrice				
Ratio Pos. to Neg.	0.551***				-0.102			
	(0.211)				(0.282)			
More3Positive		0.276***				0.155		
		(0.101)				(0.116)		
More5Positive			0.125				-0.000164	
			(0.0825)				(0.0893)	
More10Positive				0.238***				0.126
				(0.0735)				(0.0956)
Month Fixed Effects	N	N	N	N	Y	Y	Y	Y
Observations	918	918	918	918	918	918	918	918
R-squared	0.004	0.008	0.003	0.011	0.025	0.027	0.025	0.026

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. ”WC” stands for word count. ”Pos. to Neg.” stands for positive words to negative words. * Significant at 10%, ** 5%, *** 1%.

5.2 Interactions

In this subsection, we further explore several possible interactions underlying the effects of market sentiments on cryptocurrency price. To meet this end, we run the following specification.

$$\log Price_t = \alpha_1 Sentiment_t + \alpha_2 Factor_t + \alpha_3 Sentiment_t \times Factor_t + \gamma_t + u_t. \quad (2)$$

In equation 2, $\log Price_t$ is the dependent variable, the logarithm of the price of Ethereum at date t, $Sentiment_t$ is a measure of market sentiment, γ_t is a vector of month fixed effects, and u_t is the error term. $Factor_t$ represents related factors such as oil price, the state of the stock market, Donald Trump's tweets, and exchange rates. Again, we employ robust standard errors in all regressions. In Table 3, we find that the coefficient on the interaction term in column (1) is positive and statistically significant, indicating that the partial effects of positive word count are more positive when there are more COVID deaths in the United States. In column (2), the coefficient of the interaction term between negative word count and US COVID deaths is negative, which means that the partial effects of the negative word count on price are more negative when there are increasing deaths due to COVID. For the ratio of positive to negative words, the coefficient of the interaction term is positive, which means that an increase in the US COVID deaths leads to an increase in the positive-to-negative word ratio. To explain these interactions, it could be possible that US COVID deaths influence people due to the fact that they stay in more. Kaya (2020) states that during lockdowns caused by COVID-19, social media usage increased drastically. While on social media, individuals receive content that is targeted toward their interests. With the overall attitude on cryptocurrency being positive, more people who are new to social media or engage in it receive a one-dimensional view of cryptocurrency from a positive side, as Levy (2021) states. Therefore, positive words tend to rise in count as COVID deaths increase as more people engage in social media while receiving a positive outlook on cryptocurrency. This could also attribute to the negative word count decreasing while US COVID deaths rise.

Table 3: Mechanism: The Effects of COVID

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	lPrice									
log(Positive WC)	-12.79***					-7.554***				
	(0.992)					(1.428)				
log(US COVID Deaths)	-5.660***	2.105***	0.294***	0.340***	-1.549***	-3.601***	1.953***	0.319***	0.371***	-1.207***
	(0.376)	(0.150)	(0.0217)	(0.0278)	(0.284)	(0.537)	(0.155)	(0.0296)	(0.0282)	(0.328)
log(Positive WC)*log(US COVID Deaths)	1.068***					0.697***				
	(0.0700)					(0.0997)				
log(Negative WC)		5.229***					4.721***			
		(0.386)					(0.414)			
log(Negative WC)*log(US COVID Deaths)		-0.399***					-0.358***			
		(0.0296)					(0.0319)			
More5Positive			-10.92***					-10.37***		
			(3.261)					(3.721)		
More5Positive*log(US COVID Deaths)			0.845***					0.803***		
			(0.253)					(0.288)		
More10Positive				-17.32***					-17.14***	
				(0.460)					(0.547)	
More10Positive*log(US COVID Deaths)				1.347***					1.329***	
				(0.0351)					(0.0417)	
Ratio Pos. to Neg.					-33.03***					-27.87***
					(5.335)					(5.957)
Ratio Pos. to Neg.*log(US COVID Deaths)					2.533***					2.128***
					(0.402)					(0.450)
Month Fixed Effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	880	869	880	880	880	880	869	880	880	880
R-squared	0.750	0.637	0.724	0.724	0.593	0.794	0.675	0.744	0.772	0.622

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. "Pos. to Neg." stands for positive words to negative words. * Significant at 10%, ** 5%, *** 1%.

The interaction between Trump's Tweet counts and the positive and negative prices can be explained through noise. Table 4 shows that the coefficients for the interaction terms between positive and negative word count and Trump's tweets are negative and positive respectively. That means that as Trump tweets more, the magnitude of the effect of the positive word count diminishes on price, and the magnitude of the negative word count also decreases. This may be due to the fact that Trump is an influential figure on Twitter. Trump was the president of the United States, and the Tweets that he created generated a lot of attention to him, his account, and what he is saying. When he Tweets more, more attention is diverted toward Trump instead of toward cryptocurrency. Therefore, instead of positive and negative changes happening on days when he tweets a lot, individuals may choose to engage in his tweets instead. This may change how both the positive word count and the negative word count are regarded since there is more noise and diversion on the days Trump tweets more.

Table 4: Mechanism: The Effects of Donald Trump's Twitter Activities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					lPrice					
log(Positive WC)	1.952***					1.805***				
	(0.186)					(0.139)				
log(TT Count)	0.472	-0.187**	-0.0221	-0.00913	0.331*	-0.104	-0.110*	-0.106**	-0.0691**	-0.112
	(0.325)	(0.0733)	(0.0343)	(0.0283)	(0.189)	(0.200)	(0.0636)	(0.0538)	(0.0312)	(0.164)
log(Positive WC)*log(TT Count)	-0.0842					0.0159				
	(0.0565)					(0.0339)				
log(Negative WC)		-0.410***					-0.155**			
		(0.0731)					(0.0618)			
log(Negative WC)*log(TT Count)		0.0475**					0.0177			
		(0.0215)					(0.0180)			
More5Positive			1.002***					0.274		
			(0.174)					(0.198)		
More5Positive*log(TT Count)			-0.0830*					0.0599		
			(0.0502)					(0.0570)		
More10Positive				1.064***					0.297**	
				(0.167)					(0.141)	
More10Positive*log(TT Count)				-0.0998**					0.0171	
				(0.0475)					(0.0388)	
Ratio Pos. to Neg.					5.661***					2.921***
					(0.762)					(0.707)
Ratio Pos. to Neg.*log(TT Count)					-0.441**					0.0676
					(0.217)					(0.183)
Month Fixed Effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	353	344	353	353	353	353	344	353	353	353
R-squared	0.919	0.536	0.408	0.606	0.692	0.957	0.809	0.815	0.809	0.868

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. "Pos." stands for Positive. "Neg." stands for Negative. "TT" stands for Trump's tweets. "WC" stands for WC. "Pos. to Neg." stands for positive to negative word count. * Significant at 10%, ** 5%, *** 1%.

In Table 5, oil prices are another factor that is relevant. We find that the coefficient on the interaction term in column (1) is positive and statistically significant. It is also a positive value of 0.473, which means that as oil prices increase, the positive word count does too. The coefficient of the interaction term in column (2) is negative and statistically significant, which means that as oil prices increase, negative word count decreases. Similarly, all the coefficients for the interaction terms between positive word count and oil price and the ratio of positive to negative word count to oil price are statistically significant and positive. A possible theory to explain this has to do with the marginal cost and benefit of creating and selling cryptocurrency. The production of cryptocurrency depends heavily on things like graphics cards, mining rigs, and electricity, commodities that increased in price during the time period which I examined. Furthermore, as shipping prices increased during this time too, the production cost of cryptocurrency grew. As the cost increases for the producer, they must increase prices to make it profitable for them. Therefore, positive word count may increase while negative word count decreases due to the increasing price of Ethereum.

Table 5: Mechanism: The Effects of Oil Prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					lPrice					
log(Positive WC)	-0.688					1.728**				
	(0.470)					(0.708)				
log(Oil Price)	-1.293*	5.012***	1.642***	1.796***	-1.222***	1.997**	4.979***	1.742***	1.874***	-1.493***
	(0.688)	(0.159)	(0.0551)	(0.0507)	(0.241)	(0.894)	(0.150)	(0.0531)	(0.0413)	(0.267)
log(Positive WC)*log(Oil Price)	0.473***					-0.0910				
	(0.115)					(0.157)				
log(Negative WC)		2.574***					2.529***			
		(0.134)					(0.127)			
log(Negative WC)*log(Oil Price)		-0.665***					-0.646***			
		(0.0330)					(0.0314)			
More5Positive			-4.938***					-4.733***		
			(0.479)					(0.514)		
More5Positive*log(Oil Price)			1.345***					1.277***		
			(0.116)					(0.122)		
More10Positive				-7.079***					-6.836***	
				(0.332)					(0.293)	
More10Positive*log(Oil Price)				1.872***					1.776***	
				(0.0811)					(0.0712)	
Ratio Pos. to Neg.					-16.20***					-18.10***
					(1.291)					(1.493)
Ratio Pos. to Neg.*log(Oil Price)					4.409***					4.784***
					(0.295)					(0.339)
Month Fixed Effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	917	904	917	917	917	917	904	917	917	917
R-squared	0.766	0.799	0.779	0.789	0.813	0.793	0.833	0.821	0.818	0.840

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. ”WC” stands for word count. ”Pos. to Neg.” stands for positive to negative. * Significant at 10%, ** 5%, *** 1%.

Table 6 demonstrates the effects that the stock market has on cryptocurrency prices. We chose to specify and pick the Dow Jones index as an adequate measure of the overall state of the stock market. The data in column (1) shows that the coefficient for the interaction term between the positive word count and the closing price of the DOW Jones index is statistically significant and positive. Column (2) shows that the coefficient for the interaction term between the negative word count and the closing price of the DOW Jones index is statistically significant and negative. Column (3), column (4), and column (5) show similar trends with column (1) and column (2). A possible explanation for this is the overall sentiment of the market. During markets with increasing overall prices, people are more optimistic about the future and more willing to spend money on investments like cryptocurrency. With an increased demand of cryptocurrency in growing markets, prices rise, driving positive responses and decreasing negative ones.

An additional explanation for the data in Tables 5 and 6 could be that they are good representations of the current market outlook. Many investors become bullish when the market is good, spending more and investing more money into luxuries like cryptocurrencies. The Dow Jones index represents the state of the market in many ways, and oil prices could also do so too. On the contrary, bad market conditions could lead to a bearish market, withdrawing investment from excess goods. In column (1) of both tables, the coefficients of the interaction terms are positive, and in column (2), the coefficient is negative. This means that as the prices of oil or the Dow Jones index grow, the effects of the positive word count become greater on cryptocurrency price. Similarly, the effects of the negative word count on price also increase when both indicators of the market grow.

Table 6: Mechanism: The Effects of the Stock Market

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					lPrice					
log(Positive WC)	-40.98***					-34.47***				
	(2.925)					(2.720)				
log(DJI Closing Price)	-15.48***	11.43***	6.684***	7.090***	-2.563***	-11.80***	10.89***	6.910***	7.376***	-1.335*
	(1.673)	(0.300)	(0.172)	(0.143)	(0.751)	(1.533)	(0.286)	(0.178)	(0.139)	(0.764)
log(Positive WC)*log(DJI Closing Price)	3.968***					3.347***				
	(0.281)					(0.261)				
log(Negative WC)		10.25***					8.301***			
		(0.922)					(0.903)			
log(Negative WC)*log(DJI Closing Price)		-0.986***					-0.798***			
		(0.0887)					(0.0868)			
More5Positive			-21.34***					-20.01***		
			(2.104)					(2.035)		
More5Positive*log(DJI Closing Price)			2.048***					1.924***		
			(0.203)					(0.196)		
More10Positive				-25.16***					-20.84***	
				(1.970)					(1.882)	
More10Positive*log(DJI Closing Price)				2.422***					2.008***	
				(0.189)					(0.181)	
Ratio Pos. to Neg.					-128.0***					-114.8***
					(8.740)					(8.812)
Ratio Pos. to Neg.*log(DJI Closing Price)					12.33***					11.06***
					(0.843)					(0.851)
Month Fixed Effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	918	905	918	918	918	918	905	918	918	918
R-squared	0.917	0.914	0.913	0.912	0.920	0.931	0.926	0.926	0.925	0.932

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. ”WC” stands for word count. ”DJI” stands for Dow Jones Index. ”Pos. to Neg.” stands for positive to negative word count. * Significant at 10%, ** 5%, *** 1%.

Additionally, exchange rates are a relevant factor when it comes to the effects of market sentiment on cryptocurrency prices. While Ethereum is a decentralized currency that has prices independent of any major government, many currencies still require the usage of fiat currency such as the Chinese RMB in order to cash out and be able to use cryptocurrency in day-to-day exchanges. Column (1) of Table 7 shows a negative and statistically significant coefficient of the interaction term between positive word count and the exchange rate of the RMB to the USD. Column (2) shows a negative coefficient as well, while the values in columns (3), (4), and (5) are positive.

Finally, another exchange rate that we examine is the Euro. Europe engages in cryptocurrency trading by being a big seller as well as a buyer of cryptocurrency from other regions. Table 8 shows that the coefficient of the interaction term between positive word count and the exchange rate of the Euro is negative, which means that as the Euro gets more expensive, the effects of the positive word count diminish. This could be due to the fact that people, regardless of sentiment, are unwilling to trade in their coins due to unfavorable exchange rates. In order to turn Ethereum into fiat currency usable in day-to-day life, a third-party exchange is involved, which includes fees on exchange rates. With a majority of the trade coming from the US, if the dollar weakens against the Euro, the pricing and exchange rates of these third parties can become unfavorable for US customers, decreasing overall trade volume and therefore price.

Table 7: Mechanism: The Effects of the Two Exchange Rates:Rmb

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					lPrice					
log(Positive WC)	-5.658**					1.926				
	(2.467)					(2.397)				
log(ExchRMB)	42.07***	34.28***	26.42***	25.84***	16.42***	21.32***	37.74***	27.50***	27.88***	7.032**
	(7.506)	(1.661)	(0.622)	(0.321)	(3.900)	(7.205)	(1.353)	(0.412)	(0.291)	(2.759)
log(Positive WC)*log(ExchRMB)	-3.369**					0.653				
	(1.325)					(1.288)				
log(Negative WC)		-3.687***					-4.200***			
		(0.804)					(0.652)			
log(Negative WC)*log(ExchRMB)		-1.996***					-2.258***			
		(0.428)					(0.348)			
More5Positive			1.661					6.473***		
			(1.323)					(1.035)		
More5Positive*log(ExchRMB)			0.941					3.463***		
			(0.700)					(0.546)		
More10Positive				7.225***					11.00***	
				(1.293)					(1.310)	
More10Positive*log(ExchRMB)				3.887***					5.830***	
				(0.686)					(0.695)	
Ratio Pos. to Neg.					23.28***					49.65***
					(8.409)					(6.110)
Ratio Pos. to Neg.*log(ExchRMB)					12.73***					26.44***
					(4.429)					(3.219)
Month Fixed Effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	918	905	918	918	918	918	905	918	918	918
R-squared	0.857	0.853	0.836	0.840	0.846	0.913	0.910	0.901	0.905	0.907

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. "WC" stands for word count. "ExchRMB" stands for the closing exchange rate of RMB to USD. "Pos. to Neg." stands for positive to negative word count. * Significant at 10%, ** 5%, *** 1%.

Table 8: Mechanism: The Effects of the Two Exchange Rates:Euro

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					lPrice					
log(Positive WC)	5.869***					5.754***				
	(0.177)					(0.188)				
log(ExchEURO)	162.6***	23.07***	1.058	4.800***	-7.150**	136.1***	27.74***	0.738	5.109***	-3.290
	(8.585)	(2.392)	(0.839)	(0.750)	(3.412)	(9.493)	(2.423)	(0.955)	(0.782)	(3.684)
log(Positive WC)*log(ExchEURO)	-26.95***					-22.12***				
	(1.413)					(1.573)				
log(Negative WC)		0.715***					0.981***			
		(0.0775)					(0.0785)			
log(Negative WC)*log(ExchEURO)		-3.293***					-4.070***			
		(0.432)					(0.433)			
More5Positive			-1.015***					-1.123***		
			(0.234)					(0.245)		
More5Positive*log(ExchEURO)			6.998***					7.238***		
			(1.508)					(1.571)		
More10Positive				0.697*					0.662*	
				(0.356)					(0.383)	
More10Positive*log(ExchEURO)				-4.427**					-4.944**	
				(2.094)					(2.129)	
Ratio Pos. to Neg. Pos. to Neg.					-2.243***					-3.115***
					(0.445)					(0.491)
Ratio Pos. to Neg.*log(ExchEURO)					15.02***					12.19***
					(4.103)					(4.244)
Month Fixed Effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	918	905	918	918	918	918	905	918	918	918
R-squared	0.680	0.115	0.045	0.032	0.041	0.762	0.188	0.063	0.048	0.068

Notes: The sample period is 2020-2022. Robust standard error in parenthesis. "WC" stands for word count. "ExchEURO" stands for the closing exchange rate of Euro to USD. "Pos. to Neg." stands for positive to negative word count. * Significant at 10%, ** 5%, *** 1%.

6 Conclusion

Throughout the course of this paper, we study how market sentiment affects the pricing of cryptocurrency, specifically Ethereum. We find that positive market sentiments have statistically significant and positive effects on cryptocurrency prices. We also explore interactions that may influence market sentiment's effects on cryptocurrency price. We determine and examine that COVID-related deaths, Trump's tweets, oil prices, the stock market, and exchange rates of the RMB and the Euro are relevant factors at play.

For our next research steps, we aim to do the following. We should study Bitcoin, the most popular and oldest cryptocurrency available. Additionally, we should add more additional factors that may have an influence on the effects of market sentiment on cryptocurrency price. Furthermore, we could expand to just beyond the US, and include China as well as include major members of the EU to study more results from different geographical locations.

7 References

Aug 2023. URL https://en.wikipedia.org/wiki/Cryptocurrency.

H. Allcott, L. Braghieri, S. Eichmeyer, and M. Gentzkow. The welfare effects of social media. *American Economic Review*, 110(3):629–676, 2020.

M. Baker and J. Wurgler. Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2):129–151, 2007.

A. Bandopadhyaya and A. L. Jones. Measuring investor sentiment in equity markets. *Journal of Asset Management*, 7:208–215, 2006.

N. Barberis and R. Thaler. A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128, 2003.

A. Bellia. What is cryptocurrency and how does it work?, Jun 2023. URL https://usa.kaspersky.com/resource-center/definitions/what-is-cryptocurrency.

J. Benhabib, X. Liu, and P. Wang. Sentiments, financial markets, and macroeconomic fluctuations. *Journal of Financial Economics*, 120(2):420–443, 2016.

L. Braghieri, R. Levy, and A. Makarin. Social media and mental health. *American Economic Review*, 112(11):3660–3693, 2022.

L. W. Cong, Y. Li, and N. Wang. Tokenomics: Dynamic adoption and valuation. *The Review of Financial Studies*, 34(3):1105–1155, 2021.

J. M. Griffin and A. Shams. Is bitcoin really untethered? *The Journal of Finance*, 75(4):1913–1964, 2020.

D. Hirshleifer. Behavioral finance. *Annual Review of Financial Economics*, 7: 133–159, 2015.

K. John, M. O'Hara, and F. Saleh. Bitcoin and beyond. *Annual Review of Financial Economics*, 14:95–115, 2022.

T. Kaya. The changes in the effects of social media use of cypriots due to covid-19 pandemic. *Technology in society*, 63:101380, 2020.

R. Levy. Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review*, 111(3):831–870, 2021.

S. Liu. Investor sentiment and stock market liquidity. *Journal of Behavioral Finance*, 16(1):51–67, 2015.

Y. Liu and A. Tsyvinski. Risks and returns of cryptocurrency. *The Review of Financial Studies*, 34(6):2689–2727, 2021.

Y. Liu, A. Tsyvinski, and X. Wu. Common risk factors in cryptocurrency. *The Journal of Finance*, 77(2):1133–1177, 2022.

P. Maxted. A macro-finance model with sentiment. *Review of Economic Studies*, page rdad023, 2023.

E. Polyzos. Escalating tension and the war in ukraine: Evidence using impulse response functions on economic indicators and twitter sentiment. *Available at SSRN 4058364*, 2022.

T. Smith. What is market sentiment? definition, indicator types, and example, Jul 2023. URL https://www.investopedia.com/terms/m/marketsentiment.asp.

8 Acknowledgements

I am deeply indebted to my advisors Dr.Shi and Professor Blaum for guiding me throughout this project. They offered invaluable support in setting the framework of the research, computing the data, and finally teaching me the way to interpret the data. Furthermore, they both went to great lengths to teach me about STATA and its functionality, as well as introduced me to economic theories that will help on my journey forwards. Additionally, I would like to thank Nathan Godsey for providing help on the coding related to Twitter scraping. Finally, I would like to thank my parents for offering me such great opportunities to learn more about economics research.