

Influencers & Bitcoin

Welke rol spelen influencers
in de cryptomarkt?

T.M.J. Vaessen



Erasmus

Influencers & Bitcoin

Welke rol spelen influencers
in de cryptomarkt?

door

T.M.J. Vaessen

Bachelor's Thesis

BSc Economics & Business

Erasmus School Of Economics

7 augustus 2021

Auteur: T.M.J. Vaessen
Studentnummer: 436550
Thesis begeleider: Dr. J.J.G. Lemmen
Tweede lezer: Dr. A. Soebhag

The Erasmus logo, featuring the word "Erasmus" in a stylized, cursive script.

Abstract

Deze paper bestudeert de beschrijvende en voorspellende relaties tussen *influencers* en de Bitcoin marktactiviteit. De bevindingen suggereren dat er geen beschrijvende en voorspellende relaties zijn tussen het Twittersentiment van influencers en het rendement en het handelsvolume van Bitcoin. Voor de volatiliteit wordt een lage significante waarde gevonden die op een beschrijvende relatie tussen Twittersentiment en volatiliteit wijst. Een voorspellende waarde tussen deze variabelen blijft ook uit. Verder worden er geen Granger causale verbanden gevonden tussen het Twittersentiment en de marktactiviteiten van Bitcoin. De bevindingen dat er over het algemeen geen beschrijvende, voorspellende of Granger causale verbanden gelegd kunnen worden, suggereren dat de Bitcoin markt efficiënt is.

Trefwoorden: *Cryptomarkt, Bitcoin, Efficiënte Markt Hypothese, Influencers, Twitter*

Inhoudsopgave

| | |
|--|-----------|
| Lijst van tabellen | iv |
| Lijst van figuren | v |
| 1 Introductie | 1 |
| 2 Theoretisch kader & Literatuuroverzicht | 4 |
| 2.1 De Cryptomarkt | 4 |
| 2.1.1 Cryptovaluta en de onderliggende technologie Bitcoin | 4 |
| 2.1.2 Efficiënte Markt Hypothese | 5 |
| 2.2 Twitter | 5 |
| 2.2.1 Twitter en de relevantie voor de Financiële Markten | 5 |
| 2.3 Crypto-influencers | 6 |
| 2.3.1 Effect van Influencers op de Publieke Opinie | 6 |
| 2.3.2 Influencers op de Financiële Markt | 7 |
| 2.3.3 Eerder onderzoek Crypto-influencers | 7 |
| 2.4 Relevantie | 8 |
| 3 Data | 10 |
| 3.1 Twitter data: Crypto-influencers tweets | 10 |
| 3.2 Bitcoin data | 12 |
| 4 Methode | 14 |
| 4.1 Sentiment analyse (VADER) | 14 |
| 4.2 Analyse van de Variabelen | 14 |
| 4.3 Tijdreeksdata | 16 |
| 4.4 Lineaire regressiemodellen | 17 |
| 4.4.1 Beschrijvende en voorspellende regressiemodellen | 17 |
| 4.4.2 Granger Causaliteit | 18 |

| | | |
|----------|---|-----------|
| 5 | Empirische bevindingen en resultaten | 19 |
| 5.1 | Lag Bepalen en Stationariteit Testen | 19 |
| 5.1.1 | Lag bepalen | 19 |
| 5.1.2 | Augmented Dickey Fuller test | 19 |
| 5.1.3 | Phillips-Perron Test. | 20 |
| 5.2 | Relatie Twittersentiment en markt activiteit Bitcoin. | 21 |
| 5.2.1 | Beschrijvende relaties. | 21 |
| 5.2.2 | Voorspellende relaties. | 23 |
| 5.2.3 | Granger Causaliteit. | 27 |
| 6 | Conclusie | 28 |
| 6.1 | Hypotheses toetsen en het beantwoorden van de Onderzoeksvraag | 28 |
| 6.2 | Implicatie en relevantie | 29 |
| 6.3 | Limitaties en aanbevelingen voor verder onderzoek | 29 |
| A | Appendix | 35 |

Lijst van tabellen

| | | |
|-----|--|----|
| 3.1 | Variabelen | 13 |
| 4.1 | Pearson correlaties | 15 |
| 4.2 | Rendement regressiemodellen | 17 |
| 4.3 | Volatiliteit regressiemodellen | 17 |
| 4.4 | Handelsvolume regressiemodellen | 18 |
| 5.1 | Augmented Dicky Fuller test (obs. 1.659) | 20 |
| 5.2 | Phillips-Perron Test (obs. 1.659) | 20 |
| 5.3 | Rendement, beschrijvende modellen (obs. 1.659) | 21 |
| 5.4 | Volatiliteit, beschrijvende modellen (obs. 1.659) | 22 |
| 5.5 | Handelsvolume, beschrijvende modellen (obs. 1.659) | 23 |
| 5.6 | Rendement, voorspellende modellen (obs. 1.656) | 24 |
| 5.7 | Volatiliteit, voorspellende modellen (obs. 1.656) | 25 |
| 5.8 | Handelsvolume, voorspellende modellen (obs. 1.656) | 26 |
| 5.9 | Granger Causaliteitstest | 27 |
| A.1 | Voorbeeld Influencer Twitter Data | 36 |
| A.2 | Voorbeeld Tweet en Sentiment | 38 |
| A.3 | Bitcoin Data | 39 |
| A.4 | Aantal tweets per influencer | 39 |
| A.5 | VAR Diagnostiek voor Rendement | 40 |
| A.6 | VAR Diagnostiek voor Volatiliteit | 40 |
| A.7 | VAR Diagnostiek voor Handelsvolume | 40 |
| A.8 | VAR Diagnostiek voor Sentiment | 41 |

Lijst van figuren

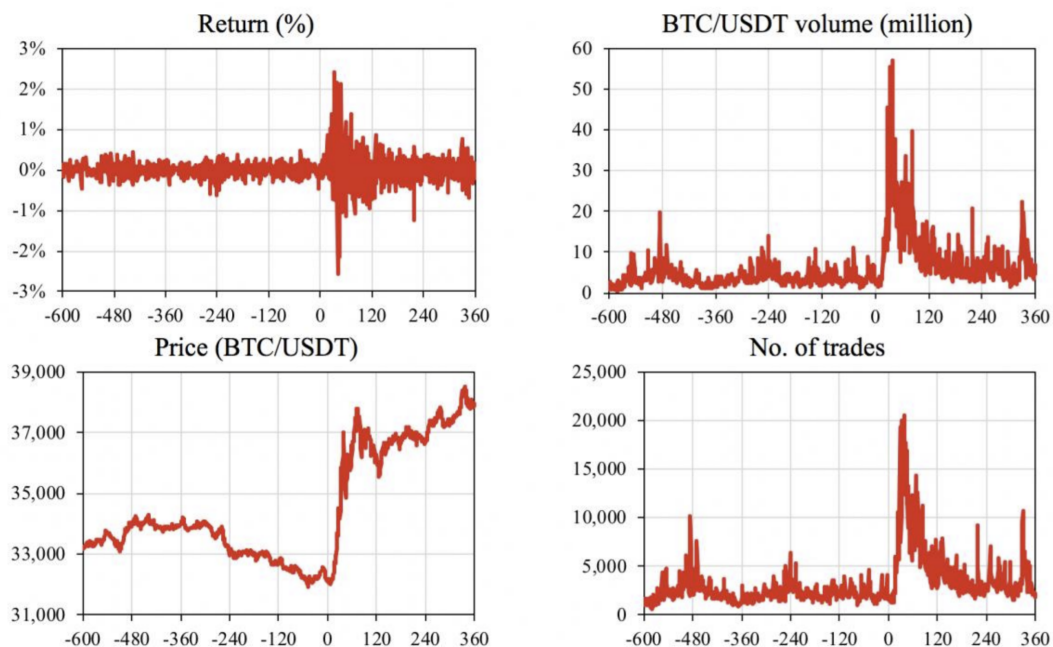
| | | |
|-----|---|----|
| 1.1 | Bitcoin marktactiviteit | 1 |
| 3.1 | Aantal tweets van geselecteerde Bitcoin influencers | 11 |

1

Introductie

Influencers nemen met een toenemende macht de financiële markt over. Elon Musk, de CEO van Tesla, voegde op 29 januari 2021 de hashtag *Bitcoin* toe aan zijn Twitterbiografie. De Bitcoin steeg binnen enkele uren met 6.000 dollar (Ante, 2021)! Het aantal investeerders dat in bezit is van een (deel van een) Bitcoin, wordt geschat op meer dan 100 miljoen (*How Many Bitcoin Users Are There? 2021*) met een totaal marktkapitaal, op het moment van schrijven, van 630 miljard dollar (*Alle cryptovaluta's g.d.*). Alle invloed op of voorspellende waarde van de cryptoprijs is daarom van groot belang voor (potentiële) investeerders, zo ook de invloed van crypto-influencers.

Figuur 1.1: Bitcoin marktactiviteit



Noot: Rendement, volume, prijs en handelsvolume per minuut rond de verandering van de Twitter biografie van Elon Musk in *Bitcoin* (Ante, 2021)

Door de opkomst van sociale media, smartphones en door de rentestand die de Centrale Banken laag houden, is er een nieuwe doelgroep tot stand gekomen voor de cryptomarkt, namelijk mensen zonder veel kennis van of affiniteit met handel in cryptomunten, maar die graag hun spaargeld willen beleggen (atNickVega, 2021). Het bewustzijn en de kennis van Bitcoin zijn in de afgelopen jaren explosief toegenomen, mede dankzij sociale media. Als de eigenaren van Bitcoin munten worden gecategoriseerd, zien we dat de grootste groep investeerders (58%) een leeftijd heeft onder de 34 jaar (*Global cryptocurrency ownership Data 2021* 2021). Dit is ook de groep die het meest actief is op sociale media (Quick en Wright, 2021). Smartphones geven een makkelijke toegang tot applicaties, zoals Bitvavo, waar men al vanaf een bedrag van vijf dollar kan investeren in een cryptomunt (*Trade the future* g.d.). Sociale media verschaffen de basisinformatie die een investeerder nodig heeft om een keuze te maken in welke munt men wil investeren en op welk moment. Maar hoe accuraat is de informatie die influencers verschaffen aan hun volgers? Hebben de tweets een voorspellende of slechts een beschrijvende waarde op de cryptoprijs? Of zijn het juist de influencers die de prijs beïnvloeden?

Al jarenlang is er veel interesse geweest vanuit de academische wereld over de rol die Twitter en influencers spelen in de financiële markt. Er bestaat nog geen eenduidig antwoord hierop vanuit de huidige literatuur. Aan de ene kant staan onderzoekers, zoals Cheuque Cerda en L. Reutter (2019), die stellen dat influencers op Twitter maar een geringe impact hebben in het voorspellen van Bitcoin prijzen en aan de andere kant zijn er academici die wel degelijk een reële significante correlatie hebben gevonden tussen de Bitcoin prijs en crypto-influencers, zoals Yamamoto e.a. (2019).

In de huidige literatuur is vaak gebruikgemaakt van een kort tijdsframe waarin tweets en Bitcoin koersen zijn geanalyseerd. In dit onderzoek zal gebruikgemaakt worden van een dataset van ruim 37.000 Bitcoin gerelateerde influencertweets over een periode van circa vier en een half jaar. In onderzoeken, zoals van Cheuque Cerda en L. Reutter (2019), bleef het effect van influencers uit door de objectieve stemming in tweets. Tweets werden gecategoriseerd in negatief, neutraal of positief. Door voor dit onderzoek een sentimentanalyse toe te passen die een continu resultaat geeft tussen -1 en 1, zal een accuratere sentimentanalyse tot stand komen. In deze studie zal naast het toetsen van de relatie tussen de Bitcoin prijs en crypto-influencers, ook dieper ingegaan worden op het soort relatie door te kijken naar beschrijvende relaties, voorspellende relaties en naar Granger causaliteit. Door een adequate sentimentanalyse en de Bitcoin koersen over een groot tijdframe tegen elkaar uit te zetten, zal de relatie tussen crypto-influencers en de Bitcoin markt beoordeeld worden om zo een antwoord te kunnen geven op de vraag of potentiële investeerders hun keuzes kunnen of zelfs zouden moeten baseren op de

informatie die crypto-influencers via Twitter verspreiden. Hieruit is de volgende onderzoeksvraag voortgevloeid:

Wat is de relatie tussen het Twittersentiment van crypto-influencers en de Bitcoin markt?

Allereerst zal de correlatie tussen de marktactiviteit variabelen van Bitcoin, rendement, volatiliteit en handelsvolume, geanalyseerd worden tezamen met de variabele van interesse, sentiment van influencer tweets. Daarna zullen de beschrijvende en voorspellende relaties voor de drie afhankelijke variabelen van de markt opgesteld worden. De bevindingen suggereren dat het Twittersentiment van influencers significant positief gecorreleerd is met handelsvolume en negatief met volatiliteit. Voor rendement is er geen significante correlatie met Twittersentiment gevonden. Verder wordt er in de beschrijvende en voorspellende regressiemodellen bijna geen significante correlatie gevonden met de variabele van interesse, Twittersentiment. Er wordt enkel voor de afhankelijke variabele volatiliteit een significante, maar lage waarde gevonden voor Twittersentiment in het beschrijvende model. Dit suggereert dat Twittersentiment in het algemeen niet gebruikt kan worden voor het verklaren of voorspellen van de Bitcoin markt.

Dit onderzoek is in een zestal hoofdstukken opgedeeld, waarvan het eerste hoofdstuk 'Introductie' zojuist behandeld is. Het tweede hoofdstuk bespreekt het theoretisch raamwerk en bespreekt relevante literatuuronderzoeken betreffende Twittersentiment en de Bitcoin markt. In het derde hoofdstuk worden de data besproken die gebruikt zullen worden om de hypothesen te testen. In hoofdstuk vier zal de methode worden beschreven en in hoofdstuk vijf zullen de resultaten van de toetsen behandeld worden. In hoofdstuk zes zullen de resultaten worden samengevat en de implicatie van de resultaten en de beperkingen van het onderzoek besproken worden.

Theoretisch kader & Literatuuroverzicht

Dit hoofdstuk vormt een theoretisch kader om de aspecten van de analyses uit dit onderzoek te kunnen begrijpen en bespreekt bestaande literatuur betreffende de onderliggende theorie. Paragraaf 2.1 bediscussieert de cryptomarkt, de onderliggende technologie Bitcoin en de Efficiënte Markt Hypothese. In paragraaf 2.2 wordt ingegaan op het socialemediaplatform Twitter en de relevantie hiervan voor de financiële markten. Paragraaf 2.3 zal dieper ingaan op het fenomeen (crypto-)influencers. Er zal besproken worden hoe zij ingezet kunnen worden als middel om de publieke opinie te reflecteren en mogelijk te beïnvloeden. In paragraaf 2.4 zal het literatuuronderzoek afgesloten worden met de belangrijkste bevindingen uit eerdere onderzoeken naar crypto-influencers, zullen de zwaktes van deze onderzoeken besproken worden en zal de relevantie van dit onderzoek op de huidige literatuur aan het licht gebracht worden.

2.1. De Cryptomarkt

2.1.1. Cryptovaluta en de onderliggende technologie Bitcoin

Een cryptomunt is een vorm van elektronisch geld die in een peer-to-peernetwerk online betalingen direct van de ene partij naar de andere partij zendt, zonder hier een financiële institutie aan te pas te laten komen (Nakamoto, 2008). De oprichter en grootste speler van deze compleet nieuwe cryptovalutamarkt, is Bitcoin. Bitcoins zijn elektronische tokens die worden gecreëerd door zogenaemde *miners*. Alle data die gepaard gaat met de transactie van deze tokens, wordt opgeslagen in een gedecentraliseerd grootboek, genaamd de Blockchain (Seybert en Katzenstein, 2018). De veelbelovende technologie achter cryptovaluta heeft ervoor gezorgd dat het Blockchainconcept verder is gegroeid dan alleen het genereren van cryptomunten en het faciliteren van transacties (Miraz en Ali, 2018). Het is een gedecentraliseerd systeem, het is beveiligd en investeerders nemen anoniem deel aan de Blockchain, omdat de technologie niet gemoeid gaat met centrale instituties, zoals banken, overheden of andere institutionele organisaties. Dit geeft de cryptomarkt de vrijheid om een monopolistisch competitief karakter aan te nemen (Böhme e.a., 2015). Inmiddels zijn er 2.539 verschillende cryptomunten actief op de

markt (waarvan de marktkapitalisatie groter dan \$0 is ten tijde van schrijven) met een totale marktkapitalisatie van \$1.281.100.822.187 (*Alle cryptovaluta's g.d.*).

Er is in de afgelopen jaren onderzoek gedaan naar de prestaties van Bitcoin en lijkt de verhandeling in Bitcoin redelijk overeen te komen met het rendement op hedgefondsen. Uit de studie van Chuen (2015) kan aangenomen worden dat met het toenemend aantal transacties van Bitcoin, ook de rendementen steeds meer zullen toenemen. Dit maakt Bitcoin een aantrekkelijke investering voor mensen die hun geld spaargeld beter willen beleggen dan het op de bank te laten staan.

2.1.2. Efficiënte Markt Hypothese

De efficiëntemarkthypothese (EMH) stelt dat een markt efficiënt is als hij alle relevante informatie weerspiegelt (Kang e.a., 2021). In de zwakke EMH wordt gesteld dat de markt niet te verslaan is door het kijken naar historische informatie. Als deze historische data niet gebruikt kan worden voor het voorspellen van toekomstige rendementen in de markt, dan is er volgens de EMH sprake van een markt die het patroon van een *random walk* volgt. Als rendementen een random walk pad volgen, kunnen ongeacht de methode, de toekomstige prijzen niet voorspeld worden. Als historische informatie wel een voorspellende waarde blijkt te hebben, dan is de cryptomarkt inefficiënt in zijn zwakke vorm.

2.2. Twitter

Twitter is een platform dat voor iedereen gratis te gebruiken is. Men kan hier in een kort bericht (max. 140 tekens), in een zogenoemde *tweet*, zijn of haar boodschap delen met een selecte groep (volgers) of openbaar maken voor het gehele publiek (Layton e.a., 2010). Twitter is een belangrijk onderdeel van dit onderzoek doordat het niet alleen informatieve updates geeft over de cryptomunten, maar ook het sentiment weergeeft van de Twittergebruikers. Tweets die geplaatst worden over de cryptomunt zijn voor het overgrote deel afkomstig van investeerders, geïnteresseerden in de cryptomunt of van influencers, die een belangrijke relatie hebben met de cryptomarkt (Cheuque Cerda en L. Reutter, 2019). In paragraaf 2.3.2 zal er verder worden ingegaan op het begrip influencers.

2.2.1. Twitter en de relevantie voor de Financiële Markten

De groei in de handel in Bitcoin heeft de afgelopen jaren gezorgd voor een stijgende media-aandacht voor de cryptomarkt. Er is een groeiende interesse in de literatuur voor de elektronische transacties en manieren om de markt te kunnen verslaan. Steeds meer academici, zoals Kumar e.a. (2014) in het boek *Twitter Data Analytics* en Mao e.a. (2011) in de studie

over het voorspellen van marktwaardes, wijzen op de grote waarde van Twitter data voor een verscheidenheid aan events, waaronder ook de waarde voor de marktwerking in de financiële wereld. De onderzoeken van Piñeiro-Chousa e.a. (2016), Sun e.a. (2016) en López-Cabarcos e.a. (2019) bevestigen deze bevinding. Ze laten een significant resultaat zien in de relatie tussen Twitter en financiële markten. Hetgeen wat Twitter onderscheidt van andere sociale media, is dat Twitter ‘real time’ is en daarmee ook invloed kan uitoefenen op de marktactiviteit van cryptovaluta (Tafti e.a., 2016).

2.3. Crypto-influencers

Influencers zijn een nieuwe soort derde partij die de mening van het publiek proberen te beïnvloeden door berichten te plaatsen op socialemediaplatformen zoals Facebook, Instagram, Youtube en Twitter. Influencers bouwen een vertrouwensrelatie op met hun volgers. Als door een gerespecteerd influencer met veel volgers een product of financieel instrument wordt aangeprezen, zijn socialemediagebruikers en volgers geneigd de influencer te geloven (Stubb e.a., 2019). Bedrijven kunnen ook als bedrijfsstrategie influencers benaderen om productinformatie aan het publiek te verschaffen of promoties te delen met hun online volgers. Dit wordt ook wel influencermarketing genoemd (Lim e.a., 2017).

2.3.1. Effect van Influencers op de Publieke Opinie

Veel studies hebben onderzoek gedaan naar het effect van influencers op de consument. In het onderzoek van Ki e.a. (2020) wordt de efficiëntie van influencermarketing aangetoond door een significant resultaat dat gevonden wordt voor het effect van socialemediainfluencers op het koopgedrag van de consument. Ook Belanche e.a. (2021) bevestigingen deze bevindingen in hun onderzoek naar de relatie tussen influencers, producten en consumenten. Ze analyseren het verband tussen een top mode-influencer en de volgers. In het onderzoek wordt geconcludeerd dat wanneer er congruentie is tussen influencer en volger, en tussen influencer en product, dit ook geldt voor volger en product. De consument (volger) zal dan eerder geneigd zijn tot het kopen van of investeren in het product. In het onderzoek van Grafström e.a. (2018) wordt de nadruk gelegd op het vertrouwen tussen de influencer en de volger. Als er weinig vertrouwen is in de influencer, wordt er een negatief effect geconstateerd. Als er echter wél vertrouwen tussen beide partijen is, dan resulteert dit in een positieve houding van de volger ten opzichte van de influencer. Als een influencer wordt gekoppeld aan een bericht van een bepaald merk, zal de koopbereidheid volgens het onderzoek van Kirkpatrick (2016) significant met 5,2 keer meer stijgen dan wanneer het gemerkte bericht niet gekoppeld is aan een influencer. De topinfluencers hebben zich ontwikkeld als opinieleiders en invloedrijke personen, die de consument ook voorzien van advies (Casaló e.a., 2020). Op deze manier kunnen zij de publieke

opinie sturen.

2.3.2. Influencers op de Financiële Markt

In de vorige paragraaf zijn studies besproken die aantonen dat influencers de publieke opinie kunnen sturen en daarmee het koopgedrag van de consument (potentiële investeerders) kunnen beïnvloeden. De invloed van influencers op de goederenmarkten is door veel onderzoeken onderbouwd. Op de financiële markten is er echter meer onduidelijk over de relatie tussen influencers en de markt, met name het effect op de cryptomarkt is nog vrij onderbelicht. Influencers zijn wel degelijk belangrijke spelers voor cryptovaluta. De cryptomarkt is namelijk een relatief nieuwe markt waardoor de traditionele nieuwsmedia niet altijd gelijk nieuws over alle cryptomunten rapporteren. Daarentegen reageren influencers op Twitter en andere socialemediaplatformen hier direct op. Echter, door een gebrekkige kennis over de cryptomarkt bij veel volgers/investeerders, is het voor hen nog lastig te beoordelen of de gedeelde informatie van de influencer correct is (Kraaijeveld en De Smedt, 2020). De influencer kan betaald worden voor influencermarketing of een belang hebben in een bepaalde cryptomunt, waardoor de informatie die verstrekt wordt aan de volgers *biased* kan zijn.

2.3.3. Eerder onderzoek Crypto-influencers

Er zijn een aantal onderzoeken met tegenstrijdige conclusies geweest naar de invloed van crypto-influencers. Yamamoto e.a. (2019) onderzochten het effect van influencer tweets op de prijzen van crypto-activa. Door middel van een *machine learning* methode wordt geprobeerd te voorspellen of de Bitcoin prijs zou stijgen of dalen door het gebruik maken van influencer tweets. In dit onderzoek worden Twitter data in twee talen (Japans en Engels) uit 2018 van 16 juli tot september 24 gebruikt als trainingdata en van 4 oktober tot 24 oktober als testdata. Om de juiste gegevens uit de tweets te halen als invoer voor de machine learning methode, zijn natuurlijk taalverwerkingstechnieken gebruikt. Tezamen met de Bitcoin prijsdata, vormt dit de basis voor het onderzoek naar het effect van influencertweets op valuta-activa (Bitcoin). Om de voorspellende waarde van de methode te testen, worden de resultaten vergeleken met een methode die alleen gebruikmaakt van de Bitcoin prijsdata en niet van de influencertweets. Uit de resultaten werd door de onderzoekers geconcludeerd dat influencertweets gebruikt kunnen worden voor het voorspellen van de Bitcoin prijs.

In het onderzoek van Cheuque Cerda en L. Reutter (2019) wordt ook gebruikgemaakt van tweets die geplaatst worden door internationaal bekende en invloedrijke influencers. Het doel is om een model te ontwerpen dat een voorspelling kan doen over de volatiliteit van Bitcoin. Hiervoor worden data gebruikt van 1 juli tot en met 30 december 2018. 135 influencers worden

geselecteerd en hiervan wordt de Twitter data gefilterd op de volgende termen: Bitcoin, btc en *cryptocurrency*. Op de dataset van uiteindelijk 9.146 tweets is een sentimentanalyse toegepast in SentiWordNet's. Dit is een taalkundige bron voor opinieonderzoek die drie sentimentcores toepast op de tekst: positiviteit, negativiteit en objectiviteit (Aesuli, g.d.). Om een juiste voorspelling van de Bitcoin prijs van de volgende dag te maken wordt gebruik gemaakt van Recurrent Neural Network in LTSM. De resultaten laten zien dat het verschil tussen het model waar alleen historische prijzen van Bitcoin worden meegenomen en het model waar ook de Twitter data in wordt geïntegreerd, verwaarloosbare verschillen laten zien. In het onderzoek wordt echter al aangehaald dat de meeste tweets over Bitcoin een relatief neutrale toon hebben, waardoor er geen positief of negatief sentiment uitgehaald kan worden, wat de minimale impact kan verklaren.

2.4. Relevantie

Aan de ene kant staan onderzoekers die beweren dat influencers alleen reageren op nieuws en het niet creëren. Aan de andere kant stellen academici zoals Yamamoto e.a. (2019), Cheuque Cerda en L. Reutter (2019) en Yang e.a. (2015) dat influencers wel degelijk impact hebben op de markt door het delen van informatie of meningen met hun volgers. Uit deze onderzoeken kan opgemaakt worden dat er een relatie is tussen influencers en de marktprijzen, alleen valt over de soort en de richting van de relatie nog te twisten. Er is tot op heden nog relatief weinig onderzoek gedaan naar de rol van influencers in de cryptomarkt. En zoals het onderzoek van Cheuque Cerda en L. Reutter (2019) en de studie van Yamamoto e.a. (2019) al laten zien, is er nog geen eenduidig antwoord op wat de relatie inhoudt en welke kant deze op gaat.

Mijn studie doet onderzoek naar de relatie tussen het Twittersentiment van de top 15 crypto-influencers en de marktactiviteit van Bitcoin. Hierbij zal ook dieper ingegaan worden op de soort relatie en welke richting deze op gaat. Om de invloed van influencers op de verschillende vormen van de marktwerking van Bitcoin te beoordelen, heeft dit onderzoek tot doel om de volgende onderzoeksvraag te beantwoorden:

Wat is de relatie tussen het Twittersentiment van influencers en de Bitcoin markt?

Zoals al eerder benoemd, zijn er studies die deze vraag in een andere hoedanigheid hebben proberen te beantwoorden (Cheuque Cerda en L. Reutter, 2019; Yamamoto e.a., 2019), maar zijn de antwoorden tegenstrijdig. De marktactiviteit van Bitcoin kan onderverdeeld worden in drie segmenten: het rendement (R), de volatiliteit (VOL) en het handelsvolume (HV). Met dit onderzoek zal aan de hand van onderstaande hypothesen getracht worden een geavanceerder

antwoord te vinden op de onderzoeksvraag door de relatie tussen het Twittersentiment (SENT) van influencers voor de verschillende marktactiviteiten separaat te onderzoeken.

*H1a: Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante **beschrijvende** relatie met het rendement op Bitcoin*

*H1b: Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante **voorspellende** relatie met het rendement op Bitcoin*

*H2a: Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante **beschrijvende** relatie met de volatiliteit op Bitcoin*

*H2b: Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante **voorspellende** relatie met de volatiliteit op Bitcoin*

*H3a: Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante **beschrijvende** relatie met het handelsvolume op Bitcoin*

*H3b: Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante **voorspellende** relatie met het handelsvolume op Bitcoin*

Het verwerpen van hypothesen (1b), (2b) en (3b) suggereert dat de influencers met het sentiment van hun tweets invloed hebben op de toekomstige marktwaarden van Bitcoin. De cryptomarkt in zijn zwakke vorm is dan inefficiënt volgens de efficiëntemarkthypothese (EMH), omdat technische analyse gebruikt kan worden om het toekomstige rendement van Bitcoin te voorspellen. Met voorgaande hypothesen zal vastgesteld worden of er sprake kan zijn van een beschrijvende of voorspellende relatie tussen de afhankelijke en onafhankelijke variabelen. Met de volgende en tevens laatste hypothese, wordt dieper ingegaan op het (Granger) oorzakelijk verband tussen de marktsegmenten van Bitcoin en het sentiment van influencer tweets.

H4: Er is geen Granger causaal verband tussen het Twittersentiment van influencer tweets ($SENT_{btc}$) en de marktsegmenten van Bitcoin ($R_{btc}, HV_{btc}, VOL_{btc}$)

3

Data

Alle verschillende databronnen die van belang zijn voor de analyses zullen in dit hoofdstuk besproken worden. De eerste bron bestaat uit de Twitter data van de crypto-influencers die gebruikt zal worden voor de sentimentanalyse voor Bitcoin. De tweede bron is de marktdata van Bitcoin opgedeeld in drie marktsegmenten. Voor elke databron zal besproken worden hoe de data is verzameld, gereedgemaakt en aangepast waar nodig.

3.1. Twitter data: Crypto-influencers tweets

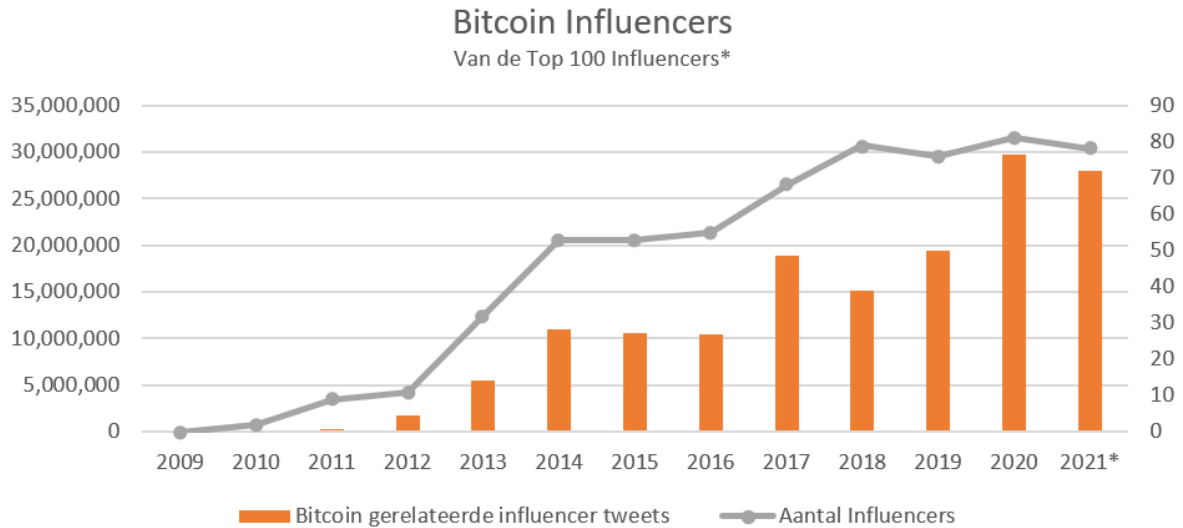
Voor het bepalen van de crypto-influencers die mee zullen worden genomen in de analyses, is de ranglijst van *The 100 Most Influential People in Crypto (2021 Edition)* van Crypto Weekly geraadpleegd (*Top 100 CRYPTOCURRENCY list of influential people in crypto* g.d.). Hierin zijn de 100 meest invloedrijke personen op het gebied van cryptovaluta gerangschikt. Voor het verzamelen van de tweets is gebruikgemaakt van de Python bibliotheek Twint (Twintproject, g.d.). Dit is een externe database waarmee handmatig tweets verzameld kunnen worden van specifieke Twittergebruikers, onderwerpen, woorden, hashtags, etc.

De code die gebruikt is om de tweets te downloaden is te vinden in appendix *listing* A.1. De variabele ‘influencers’ is in de gebruikte code de lijst van de bovengenoemde 100 influencers. De zoektermen ‘bitcoin’ en ‘btc’ geven aan dat alleen de tweets van bovengenoemde influencers verzameld worden als het woord bitcoin en/of btc er in voorkomt. Verder worden alleen tweets verzameld die tussen 1 januari 2009 en 18 juli 2021 zijn gepost. Dit resulteerde in 74.768 verschillende tweets. In Tabel A.1 van de appendix is een voorbeeld te zien van hoe deze data eruit ziet. De relevante data die verzameld is, zijn ‘date’, ‘username’, ‘tekst’ en ‘hashtags’.

De lijst van 100 influencers bestaat uit de top van de crypto-influencers anno 2021. Naast Bitcoin zijn er de laatste jaren veel munten bijgekomen die aandacht trekken van influencers. Hierdoor tweet niet elke influencer over Bitcoin. Aangezien dit onderzoek expliciet onderzoek

doet naar het effect van Twitter sentiment (SENT) op de Bitcoin marktactiviteiten, worden deze zeven influencers die niet over Bitcoin tweeten uitgesloten van het onderzoek. Voor 93 Twitter accounts was het mogelijk om tweets te verzamelen sinds het ontstaan van Bitcoin (januari 2009) tot en met 18 juli 2021.

Figuur 3.1: Aantal tweets van geselecteerde Bitcoin influencers



Noot: Het aantal Bitcoin gerelateerde tweets van de influencers van The 100 Most Influential People in Crypto (2021 Edition) van Crypto Weekly sinds het ontstaan van de Bitcoin (*Gegevens 2021 strekken tot en met 18 juli)

In Figuur 3.1 is van 2016 naar 2017 bijna een verdubbeling in het aantal Bitcoin gerelateerde tweets te zien bij de geselecteerde crypto-influencers. Het jaar 2017 wordt ook wel het jaar van Bitcoin genoemd (Stadt, 2017) waarin de prijs van €916,63 op 1 januari naar €11.800,93 op 31 december ging (*Koers, grafieken, marktkapitalisatie en andere metrics voor bitcoin (btc) g.d.*). Dit resulteerde in een explosief toenemende interesse in de munt op sociale media. Voor dit onderzoek wordt daarom de data van de top 100 influencers geëvalueerd voor een tijdsperiode van 4,5 jaar, beginnend op 1 januari 2017 tot en met 18 juli 2021. De top 100 influencers van Crypto Weekly (*Top 100 CRYPTOCURRENCY list of influential people in crypto g.d.*) is gebaseerd op het effect op de gehele cryptomarkt en niet alleen op Bitcoin. Alle influencers die in de afgelopen vijf jaar niet ten minste 52 tweets per jaar (gemiddeld 1 per week) hebben geplaatst gerelateerd aan Bitcoin, zullen worden geëlimineerd uit het onderzoek. Dit resulteert in een top 15 influencers die zullen worden gebruikt voor de analyses in dit onderzoek (appendix Tabel A.4). Over een periode van ongeveer vier en een half jaar zorgden deze influencers voor een dataset van 37.388 Bitcoin gerelateerde tweets. Als er in dit onderzoek gesproken wordt over de variabele Twittersentiment, dan heeft dit betrekking op het Twittersentiment van

Bitcoin gerelateerde tweets van de top 15 geselecteerde crypto-influencers.

3.2. Bitcoin data

De marktdata van Bitcoin van 1 januari 2017 tot en met 18 juli 2021 zijn verkregen van CoinMarketCap (*Alle cryptovaluta's* g.d.). Deze set bestaat uit dagelijkse data betreffende het handelsvolume, prijzen ten tijde van de opening, de afsluiting en de hoogst en laagst bereikte prijs van de dag (UTC-tijd) en de totale marktkapitalisatie. Alhoewel de cryptomarkt 24/7 open is, hanteert CoinMarketCap toch een sluitingsprijs. Deze prijs zal in dit onderzoek gehanteerd worden voor de analyses omtrent de rendementontwikkeling van Bitcoin. Om de marktactiviteit van Bitcoin goed in kaart te brengen, zijn er drie marktactiviteiten die worden geanalyseerd. Allereerst, het handelsvolume (HV) dat kan gevonden worden in de data van CoinMarketCap. Ten tweede, het rendement (R) van Bitcoin en ten derde de volatiliteit (VOL) van Bitcoin. Uit het onderzoek van Chuen (2015) blijkt dat het rendement op Bitcoin sterk verwant is met hedgefondsen, waarbij de opbrengst stijgt naarmate het aantal transacties toeneemt. Het handelsvolume is daarom ook een belangrijke maatstaf van de marktactiviteit van Bitcoin. Naast het handelsvolume, dat is gegeven door CoinMarketCap, moeten de andere twee marktsegmenten nog bepaald worden. Het rendement van Bitcoin wordt berekend op basis van van de dagelijkse sluitingsprijzen op de Bitcoin markt aan de hand van de volgende formule (Thorbecke, 1997):

$$R_{btc,t} = \log\left(\frac{P_t}{P_{t-1}}\right) * 100 \quad (3.1)$$

In deze formule staat R_t voor het rendement van Bitcoin op tijdstip t , en P_t en P_{t-1} zijn de Bitcoin prijzen op tijdstip t en $t - 1$. In dit onderzoek is ervoor gekozen om de meting van de volatiliteit te doen aan de hand van de studie van Garman en Klass (1980). De efficiëntie factor van dit model wordt acht maal beter geschat dan de schatter van het klassieke model. Deze studie is gewijd aan het vinden van de beste universeel beschikbare parameters voor het vaststellen van de volatiliteit. De parameters zullen gebaseerd zijn op de historische opening en sluiting, en de hoogste en laagste prijzen van Bitcoin. Alle prijzen in deze formule zijn getransformeerd in logaritmen. Dus met ‘prijs’ wordt bedoeld ‘logaritme van de prijs’ en met ‘volatiliteit’ wordt bedoeld ‘variantie van het logaritme van de prijs’. De getransformeerde prijsverandering over een bepaald tijdsinterval is normaal gedistribueerd met een gemiddelde van nul en de variantie is proportioneel met de lengte van het interval (Garman en Klass, 1980). Met deze formule kan de volatiliteit op dagelijkse basis berekend worden, in tegenstelling tot meer eenvoudige berekeningen die onder andere gebruikmaken van een *trailing window* die bestaat uit meer dan één dag. De onderzoeken van Blair e.a. (2010), Canina en Figlewski

Tabel 3.1: Variabelen

| Afhankelijke variabelen | Beschrijving |
|--------------------------|---|
| R_{btc} | Rendement van Bitcoin |
| HV_{btc} | Handelsvolume van Bitcoin |
| VOL_{btc} | Volatiliteit van Bitcoin |
| Onafhankelijke variabele | |
| $SENT_{btc}$ | Bitcoin Twittersentiment van top 15 influencers (variabele van interesse) |

Noot: De drie afhankelijke variabelen en de onafhankelijke variabele waar de hypothesen voor het beantwoorden van de onderzoeksvraag, op berusten

(1993) en Christensen en Prabhala (1998) suggereren dat het gebruik maken van dagelijkse rendementen de meest efficiënte schatter is om tot de werkelijke volatiliteit te komen. Met de methode van Garman en Klass wordt de dagelijkse historische volatiliteit berekend als de wortel van de dagelijkse variantie met de volgende formule:

$$VOL_{btc,t} = \sqrt{\frac{1}{2}(H_t - L_t)^2 - (2 \log 2 - 1)C_t^2} \quad (3.2)$$

De volatiliteit (VOL_{btc}) wordt berekend aan de hand van de volgende parameters:

$$\begin{aligned} C_t &= \log(\text{close}_t) - \log(\text{open}_t) \\ L_t &= \log(\text{low}_t) - \log(\text{open}_t) \\ H_t &= \log(\text{high}_t) - \log(\text{open}_t) \end{aligned} \quad (3.3)$$

De volatiliteit zal meegenomen worden in zowel de beschrijvende als de voorspellende regressiemodellen als controlevariabelen, omdat eerdere studies (French e.a., 1987) hebben laten zien dat volatiliteit een positieve relatie heeft met het rendement op de aandelenmarkt. Dit houdt in dat als het rendement stijgt, de volatiliteit ook toeneemt. Deze studie gaat ervan uit dat dit ook geldt voor de cryptomarkt.

4

Methode

In dit onderzoek wordt getracht antwoord te geven op de vraag of het sentiment van influencer tweets het Bitcoin rendement (R_{btc}), het handelsvolume (HV_{btc}) en/of de marktvolatiliteit (VOL_{btc}) kan verklaren, voorspellen of (Granger) veroorzaken. Allereerst zal de methode voor het verkrijgen van het sentiment uit de geselecteerde Twittertweets worden besproken. Vervolgens zal een analyse van de variabelen worden gedaan om meer inzicht te krijgen in de verhoudingen tussen de variabelen. Daarna zal getest worden op de stationariteit van de tijdreeksdata. Als laatste zal worden uitgelegd hoe de hypothesen getest zullen worden. Dit zal gedaan worden door het toetsen van beschrijvende en voorspellende regressiemodellen en met een Granger Causaliteitstoets.

4.1. Sentiment analyse (VADER)

De toon van de tweets van een influencer zal gecategoriseerd worden door middel van een sentimentscore. Deze wordt verkregen door het gebruiken van de *NLP cleaning* methode, waarbij de mening die ertoe doet, uit de tekst wordt gehaald. Vervolgens wordt er een sentiment score aan deze mening gekoppeld en wordt deze score gecategoriseerd in positief, negatief of neutraal. VADER is de afkorting voor *Valence Aware Dictionary and sEntiment Reasoner*. Dit is een veelgebruikte techniek voor het analyseren van sentiment in sociale media. Dit zijn vaak korte en cryptische teksten. In appendix listing A.2 is de code weergegeven waarmee het sentiment uit de tweets is gehaald (Band, 2021) en is er in appendix Tabel A.2 een voorbeeld weergegeven van de verkregen data uit deze analyse.

4.2. Analyse van de Variabelen

Allereerst zal de relatie tussen de verschillende variabelen geanalyseerd worden. Er is hier met name gekeken naar de correlatie tussen de afhankelijke variabelen (R, VOL en HV) en de onafhankelijke variabele van interesse (SENT). Een hoge correlatie tussen de afhankelijke en

Tabel 4.1: Pearson correlaties

| Variabelen | $SENT_{btc}$ | R_{btc} | VOL_{btc} | HV_{btc} |
|--------------|--------------|-----------|-------------|------------|
| $SENT_{btc}$ | 1,000 | | | |
| R_{btc} | 0,009 | 1,000 | | |
| VOL_{btc} | -0,014** | -0,030*** | 1,000 | |
| HV_{btc} | 0,033*** | 0,047*** | 0,292*** | 1,000 |

Noot: Pearson correlaties voor rendement (R_{btc}), volatiliteit (VOL_{btc}), handelsvolume (HV_{btc}) en sentiment ($SENT_{btc}$). * = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$

onafhankelijke variabelen zou betekenen dat de onafhankelijke variabele wellicht gebruikt zou kunnen worden om de markt te verslaan. Een lage correlatie tussen de variabelen zou duiden op een lage mate van voorspellende waarde.

Verder wordt er gezocht naar een significante correlatie tussen de onafhankelijke en afhankelijke variabelen om vast te stellen welke variabelen moeten worden toegevoegd aan het optimale regressiemodel. Daarna zal er gekeken worden naar de relatie tussen de onafhankelijke variabelen om te kijken of er sprake is van significante multicollineariteit. In dat geval zijn er twee of meer verklarende variabelen die sterk met elkaar gecorreleerd zijn, wat problemen veroorzaakt voor de betrouwbaarheid van de geschatte coëfficiënten van de parameters van het regressiemodel (*Multicollineariteit* g.d.). Ze overlappen elkaar gedeeltelijk of helemaal. Als het laatste het geval is, dan zijn de variabelen perfect multicollinear en kunnen de ‘overtollige’ variabelen verwijderd worden van het regressiemodel. Het oplossen van gedeeltelijke multicollineariteit kan door het toevoegen van meer observaties, hierdoor wordt de standaardfout kleiner. Bij structurele multicollineariteit kan het ook door het centreren van het gemiddelde of het standaardiseren van de variabelen. En bij datagebaseerde multicollineariteit kan het ook opgelost worden door op basis van de literatuur variabelen toe te voegen aan het model. Hierdoor wordt de verklarende waarde van het model vergroot en het effect zuiverder (York, 2012). De correlatie tussen de variabelen en de p-waardes zijn gebaseerd op de Pearson correlatie. De beschrijvende en voorspellende regressiemodellen zijn gebaseerd op deze correlatie voor Bitcoin.

In Tabel 4.1 is te zien dat dat bijna alle correlaties tussen de variabelen significant zijn op een niveau van ten minste 0.05, behalve de correlatie tussen het Twittersentiment van influencers en het rendement op Bitcoin. Het Bitcoin rendement is wel significant gecorreleerd met de volatiliteit en het handelsvolume. Aangezien de variabele van interesse geen significante correlatie heeft met rendement, zal worden verwacht dat de beschrijvende en voorspellende

regressies voor rendement ook zullen uitwijzen dat er geen significante relatie is tussen deze twee variabelen.

4.3. Tijdreeksdata

Allereerst zal getest worden of de tijdreeks data stationair is om onechte relaties te vermijden. Stationair houdt in dat de variabelen niet afhankelijk zijn van de tijd, oftewel er is geen sprake van een *trend of seasonality*. Hiervoor hebben Dickey en Fuller (1979) een model ontworpen dat nagaat of er sprake is van een eenheidswortel.

$$y_t = \alpha + y_{t-1} + \delta t + \varepsilon_t \quad (4.1)$$

In 1981 is dit model uitgebreid tot de Augmented Dickey Fuller test die controleert voor autocorrelatie door het toevoegen van lags van y_t .

$$\Delta y_t = \alpha + \beta t + \lambda y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \dots + \delta_p \Delta y_{t-p} + \varepsilon_t \quad (4.2)$$

Naast het toevoegen van lags, staat de ADF test ook trends of drifts toe in het random walk model. Er wordt uitgegaan dat er sprake zal zijn van autocorrelatie in y_t en daarom zal de ADF test worden toegepast in plaats van het DF-model. Parameter β in bovenstaande formule is de coëfficiënt die de trend weergeeft en p is de lag order van het autoagressieve model. Met de ADF-test wordt getoetst of λ gelijk is aan nul of niet. Dit wordt gedaan aan de hand van kritieke waardes. Als de nulhypothese, dat er sprake is van een eenheidswortel verworpen kan worden, dan is er sprake van stationaire data. Als er sprake is van niet-stationaire data (de nulhypothese kan niet worden verworpen), dan zal de data stationair gemaakt moeten worden door middel van differentiëren. Omdat er vanuit gegaan kan worden dat de data te maken zal hebben met autocorrelatie, zal er gebruik gemaakt worden van de Newey – West standard errors. Deze zijn robuust voor heteroscedasticiteit en autocorrelatie.

Als de ADF test uitwijst dat variabelen wel stationair zijn, zal om dit te bevestigen de data ook aan een andere stationariteitstoets onderworpen worden. Dit zal gedaan worden aan de hand van de Phillips-Perron unit-root test (Phillips en Perron, 1988). Dit is een gemodificeerde versie van de DF test. Voor het testen zal daarom gebruik gemaakt worden van de formule van de DF test (vergelijking 4.1). Onder de nulhypothese zal er getest worden of de data een eenheidswortel bevat. De ADF test en de Phillips-Perron test verschillen in het opzicht dat de ADF test gebruik maakt van de lags van de y-waarde en dat de Phillips-Perron test gebruik maakt van de Newey-West (Newey en West, 1986) standaardfouten om hiervoor te controleren.

4.4. Lineaire regressiemodellen

4.4.1. Beschrijvende en voorspellende regressiemodellen

Uit de significante correlaties, weergegeven in Tabel 4.1, kunnen de algemene vormen van de beschrijvende en voorspellende regressiemodellen worden opgesteld. Allereerst wordt de regressie opgesteld voor de beschrijvende relatie van de afhankelijke variabele met de eerste lag van de afhankelijke variabele. Vervolgens zal de variabele van interesse worden toegevoegd en als laatste zal ook de controlevariabele worden inbegrepen in het model. Voor het beschrijvende model zal enkel de eerste lag van de afhankelijke variabele worden meegenomen. Voor de voorspellende regressies zullen er één of meer lags van de afhankelijke variabele en de onafhankelijke variabelen worden meegenomen. Het aantal lags dat zal worden toegevoegd aan het model zal worden bepaald met de VAR Diagnostic. Deze toets zal voor verschillende criteria het optimaal aantal lags weergeven. De onderzochte informatie criteria zijn Final Prediction Error (FPE), Akaike information criterion (AIC), Hannan Quinn Information Criteria (HQIC) en Schwartz Information Criteria (SBIC). En verder zal ook de Likelihood Ratio (LR) berekend worden. Het aantal lags dat gekozen zal worden voor het regressiemodel zal worden bepaald door de meerderheid van de informatiecriteria, met een maximaal aantal lags van 4. De algemene modellen voor de beschrijvende en voorspellende regressies zijn in onderstaande tabellen 4.2, 4.3 en 4.4 voor elke afhankelijke variabele weergegeven. p staat voor het aantal lags van het rendement, q voor het aantal lags van het sentiment, r voor het aantal lags van de volatiliteit en s voor het aantal lags van het handelsvolume.

Tabel 4.2: Rendement regressiemodellen

| | |
|--------------|--|
| Beschrijvend | $R_t = \alpha_i + \beta_2 SENT_t + \beta_2 R_{t-1} + \dots + \beta_p R_{t-p} + VOL_t + \varepsilon_t$ |
| Voorspellend | $R_t = \alpha_i + \beta_2 SENT_{t-1} + \dots + \beta_2 SENT_{t-q} + \beta_2 R_{t-1} + \dots + \beta_p R_{t-p} + VOL_{t-1} + \dots + VOL_{t-r} + \varepsilon_t$ |

Noot: Het algemene beschrijvende en voorspellende regressiemodel voor de afhankelijke variabele rendement

Tabel 4.3: Volatiliteit regressiemodellen

| | |
|--------------|--|
| Beschrijvend | $VOL_t = \alpha_i + \beta_2 SENT_t + \beta_2 VOL_{t-1} + \dots + \beta_2 VOL_{t-r} + HV_t + \varepsilon_t$ |
| Voorspellend | $VOL_t = \alpha_i + \beta_2 SENT_{t-1} + \dots + \beta_2 SENT_{t-q} + \beta_2 VOL_{t-1} + \dots + \beta_2 VOL_{t-r} + HV_{t-1} + \dots + HV_{t-s} + \varepsilon_t$ |

Noot: Het algemene beschrijvende en voorspellende regressiemodel voor de afhankelijke variabele volatiliteit

Tabel 4.4: Handelsvolume regressiemodellen

| | |
|--------------|--|
| Beschrijvend | $HV_t = \alpha_i + \beta_2 SENT_t + \beta_2 HV_{t-1} + \dots + \beta_2 HV_{t-s} + VOL_t + \varepsilon_t$ |
| Voorspellend | $HV_t = \alpha_i + \beta_2 SENT_{t-1} + \dots + \beta_2 SENT_{t-q} + \beta_2 HV_{t-1} + \dots + \beta_2 HV_{t-s}$ $+ VOL_{t-1} + \dots + VOL_{t-r} + \varepsilon_t$ |

Noot: Het algemene beschrijvende en voorspellende regressiemodel voor de afhankelijke variabele handelsvolume

4.4.2. Granger Causaliteit

De vooronderstellingen voor een regressiemodel zijn dat het verband tussen de variabelen lineair en causaal is. De meest prangende vraag van dit onderzoek is of Twittersentiment van influencers bewegingen op de cryptomarkt veroorzaakt, voorspelt of reflecteert. In de voorgaande paragraaf is de methode voor het bepalen van het beste beschrijvende en voorspellende regressiemodel per afhankelijke variabele besproken. Het bepalen van de ware causaliteit is erg complex en voor dit onderzoek wordt er daarom gebruik gemaakt van het testen op Granger causaliteit. Dit wordt omschreven als ‘voorspellende causaliteit’ (Fildes e.a., 2008). Om achter de richting van het mogelijke Granger causale verband te komen, zal er gebruik gemaakt worden van de Granger Causaliteitstest (Granger, 1969). Er wordt hierbij gezocht naar een statistisch significant patroon in de vertraagde waardes van X en Y en/of er gesteld kan worden dat Y een voorspellende waarde heeft over X of andersom (Mao e.a., 2011). Voor deze toets is het een vereiste dat de data stationair is en wordt de assumptie gemaakt dat er sprake is van een lineaire relatie tussen de onderzochte variabelen. Verder zal er een maximaal aantal lags van 4 worden toegepast voor deze toets. Met behulp van de chi2-statistiek zal getest worden of er sprake is van Granger causaliteit. Het aantal lags die gevonden zijn in paragraaf 5.1.1 door middel van de VAR Diagnostic, wordt toegevoegd aan het Auto Regressief (AR) model.

$$y_t = \alpha_0 + \sum_{i=1, \dots, p} a_i y_{t-i} + \sum_{j=1, \dots, p} \beta_j x_{t-j} + \varepsilon_t \quad (4.3)$$

De nulhypothese luidt ‘geen Granger causaliteit’ en test of alle coëfficiënten $t-j$ gelijk zijn aan nul. Als dit het geval is, wordt de nulhypothese niet verworpen en kan gesteld worden dat x_{t-j} geen voorspellende waarde heeft voor y_{t-j} .

5

Empirische bevindingen en resultaten

In dit hoofdstuk worden allereerst de resultaten van de ADF en Phillips-Perron test besproken. Vervolgens zal in paragraaf 5.1 van dit hoofdstuk ook het aantal lags per variabele weergegeven worden die bepaald zijn door de VAR Diagnostic. In paragraaf 5.2 zullen de hypothesen getest worden door middel van de beschrijvende en voorspellende regressiemodellen, en een toets op Granger causaliteit. Deze modellen en toets bestuderen de relatie tussen het Twittersentiment van influencers en de marktactiviteiten van de Bitcoin. In hoofdstuk 6 zullen de interpretatie en de implicatie van de resultaten weergegeven worden.

5.1. Lag Bepalen en Stationariteit Testen

5.1.1. Lag bepalen

Om de optimale lag-lengte van de variabelen te berekenen, met een maximale lengte van 4 lags, wordt de VAR-analyse toegepast. Deze test aan de hand van verschillende informatiecriteria wat het meest adequate aantal lags is binnen de restricties. Met de meerderheid wordt uiteindelijk per variabele het aantal lags vastgesteld. In appendix Tabel A.5, A.6, A.7 en A.8 zijn de VAR Diagnostic testen voor de variabelen te vinden. Voor het testen op stationariteit en voor de voorspellende relaties, zullen voor de variabelen rendement, volatiliteit, handelsvolume en sentiment, respectievelijk de volgende lags per variabele worden meegenomen: 0, 4, 4 en 3.

5.1.2. Augmented Dickey Fuller test

Tabel 5.1 laat de test statistieken zien voor de ADF testen op de vier variabelen R_{btc} , HV_{btc} , VOL_{btc} en $SENT_{btc}$. Allereerst is de test uitgevoerd met nul lags, waarvan de resultaten te zien zijn in Tabel 5.1. Ook de kritieke waarden bij 1%, 5% en 10% significantieniveau zijn weergegeven. Als de teststatistiek in absolute termen groter is dan de absolute waarde 2,570, dan wordt de nulhypothese verworpen op een 10% significantie niveau.

Tabel 5.1: Augmented Dicky Fuller test (obs. 1.659)

| Significantie niveau | Kritieke waarde | R_{btc} | VOL_{btc} | HV_{btc} | $SENT_{btc}$ |
|----------------------|-----------------|------------|-------------|------------|--------------|
| 10% | -2,570* | | | | |
| 5% | -2,860** | -42,285*** | -24,266*** | -12,125*** | -38,351*** |
| 1% | -3,430*** | | | | |

Noot: ADF test statistieken voor rendement (R_{btc}), volatiliteit (VOL_{btc}), handelsvolume (HV_{btc}) en sentiment ($SENT_{btc}$). * = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$

De nulhypothese die voor alle vier de tijdreeksdata wordt opgesteld, luidt dat de tijdreeksdata een eenheidswortel bevat en daarom een random walk patroon volgt. De teststatistiek voor alle variabelen overstijgt de kritieke waardes in absolute zin, wat betekent dat de nulhypotheses dat de data niet stationair is, voor alle variabelen verworpen wordt. Er wordt geen autocorrelatie gevonden, dus er is geen bewijs tegen de EMH. De cryptomarkt lijkt efficiënt te zijn.

5.1.3. Phillips-Perron Test

De tweede test die nagaat of er sprake is van een eenheidswortel, is de Phillips-Perron test. Bij deze toets worden er twee statistieken berekend, namelijk de ρ en de τ . In Tabel 5.2 zijn de resultaten van de test weergegeven voor de vier variabelen (R_{btc} , HV_{btc} , VOL_{btc} en $SENT_{btc}$). Om de ADF test en de Phillips-Perron test naast elkaar te kunnen leggen, wordt de τ teststatistiek gehanteerd, omdat deze dezelfde teststatistiek gebruikt als de ADF test, zoals te zien is in Tabel 5.2. Net als bij de ADF test wordt getoetst op de variabele zonder lags.

Tabel 5.2: Phillips-Perron Test (obs. 1.659)

| Significantie niveau | Kritieke waarde | R_{btc} | VOL_{btc} | HV_{btc} | $SENT_{btc}$ |
|----------------------|-----------------|------------|-------------|------------|--------------|
| 10% | -2,570* | | | | |
| 5% | -2,860** | -42,261*** | -26,488*** | -10,777*** | -38,782*** |
| 1% | -3,430*** | | | | |

Noot: Phillips-Perron test statistieken voor rendement (R_{btc}), volatiliteit (VOL_{btc}), handelsvolume (HV_{btc}) en sentiment ($SENT_{btc}$). * = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$

Als er gekeken wordt naar de teststatistieken van deze toets, zien we dat deze niet ver van de waardes van de ADF-test affiggen. Wederom worden de nulhypotheses van niet-stationaire data verworpen. Zowel de ADF-test, als de Phillips-Perron test, wijzen uit dat de nulhypotheses van alle vier de variabelen met een relatief grote teststatistiek verworpen kunnen worden. Aan de hand van de resultaten van de testen naar stationariteit is duidelijk te zien dat alle

variabelen een stationair proces volgen, wat inhoudt dat de absolute veranderingen van de marktactiviteiten van de Bitcoin en het Twittersentiment van influencers een tijd-onafhankelijk gemiddelde en variantie hebben, oftewel ze zijn random. Volgens de zwakke vorm van de EMH, kan gesteld worden dat alle variabelen efficiënt zijn.

5.2. Relatie Twittersentiment en markt activiteit Bitcoin

5.2.1. Beschrijvende relaties

In deze paragraaf worden de resultaten van de beschrijvende modellen (Tabel 4.2, 4.3 en 4.4) bediscussieerd. Allereerst worden de resultaten van de beschrijvende regressiemodellen, die het Bitcoin rendement proberen te verklaren, besproken. In Tabel 4.1 is te zien dat er geen significante correlatie is gevonden voor rendement en Twittersentiment. Voor het meest uitgebreide beschrijvende regressiemodel, waarbij een vertraagde waarde en een controlevariabele worden toegevoegd aan het model, wordt wederom geen significante relatie gevonden. Hypothese 1a, *Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante beschrijvende relatie met het rendement van de Bitcoin*, wordt daarom aangenomen.

Tabel 5.3: Rendement, beschrijvende modellen (obs. 1.659)

| | Model 1 | Model 2 | Model 3 |
|------------------|------------------|------------------|------------------|
| $R_{btc} L1$ | - 0,038 (0,0341) | - 0,038 (0,0341) | - 0,051 (0,0374) |
| $SENT_{btc}$ | | - 0,002 (0,0024) | - 0,002 (0,0025) |
| VOL_{btc} | | | - 0,143 (0,1054) |
| F-statistiek | 1,23 | 0,85 | 1,21 |
| R^2 | 0,015 | 0,002 | 0,015 |
| Aangepaste R^2 | 0,014 | 0,000 | 0,014 |

Noot: Samengevatte resultaten van de regressie op rendement. Model 1 tot en met Model 3 rapporteren de resultaten van het beschrijvende regressiemodel. Elke kolom voegt een variabele toe aan het model. De coëfficiënt van het beschrijvende model is weergegeven, de Newey-West standaardfouten (tussen haakjes), de F-statistiek en de verklarende waarde van het model. $R_{btc} L1$ = de eerste lag van rendement, $SENT_{btc}$ = Twitter sentiment van influencers en VOL_{btc} = volatiliteit van Bitcoin (* = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$).

De resultaten voor de beschrijvende modellen voor volatiliteit worden in Tabel 5.4 weergegeven. In de eerste kolom (Model 1) wordt aangetoond dat er een significante relatie wordt gevonden voor de afhankelijke variabele volatiliteit (t) en de vertraagde waarde van volatiliteit ($t-1$). De F-statistiek in Model 3 ontbreekt, omdat er te veel variabelen zijn voor het aantal beschikbare vrijheidsgraden. Dit heeft verder geen effect op de test. De aangepaste R-kwadraatswaarde is

voor Model 3 het hoogst. Dit betekent dat Model 3 het meest betrouwbare model is van de algemene vergelijking van de beschrijvende regressie van rendement (te vinden in Tabel 4.2). Hoewel er sprake is van significantie op een 1% niveau, is de coëfficiënt van de variabele van interesse (SENT) en van de controlevariabele (HV) van lage waarde. Ze hebben zeer weinig effect op de afhankelijke variabele volatiliteit. Hypothese 2a luidt: *Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante beschrijvende relatie met de volatiliteit op de Bitcoin*. Statistisch gezien is er sprake van een beschrijvende relatie, dus wordt hypothese 2a verworpen.

Tabel 5.4: Volatiliteit, beschrijvende modellen (obs. 1.659)

| | Model 1 | Model 2 | Model 3 |
|------------------|-------------------|---------------------|--------------------------|
| VOL_{btc} L1 | 0,476*** (0,0256) | 0,474*** (0,0255) | 0,450*** (0,0264) |
| $SENT_{btc}$ | | - 0,004*** (0,0017) | - 0,005*** (0,0017) |
| HV_{btc} | | | 1,00 e-13*** (2,56 e-14) |
| F-statistiek | 345,00 | 179,85 | . |
| R^2 | 0,226 | 0,229 | 0,248 |
| Aangepaste R^2 | 0,226 | 0,228 | 0,246 |

Noot: Samengevatte resultaten van de regressie op volatiliteit. Model 1 tot en met Model 3 rapporteren de resultaten van het beschrijvende regressiemodel. Elke kolom voegt een variabele toe aan het model. De coëfficiënt van het beschrijvende model is weergegeven, de Newey-West standaardfouten (tussen haakjes), de F-statistiek en de verklarende waarde van het model. VOL_{btc} L1 = de eerste lag van volatiliteit, $SENT_{btc}$ = Twitter sentiment van influencers en HV_{btc} = handelsvolume van Bitcoin (* = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$).

Als laatste van de beschrijvende modellen, worden de modellen van de afhankelijke variabele handelsvolume onderzocht. Hierbij wordt hypothese 3a getest, *Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante beschrijvende relatie met het handelsvolume op de Bitcoin*. In Tabel 5.5 worden weer alle coëfficiënten van de modellen weergegeven. In kolom 1 (Model 1) wordt gevonden dat het handelsvolume is gecorreleerd met het handelsvolume van de vorige periode. Voor Model 2 en Model 3 wordt geen significantie gevonden voor Twittersentiment, daarom zal deze variabele worden weggelaten uit het model. Aangezien dit de variabele van interesse is in de hypothese, kunnen we hiermee gelijk hypothese 3a beantwoorden door te stellen dat de nulhypothese aangenomen zal worden.

Tabel 5.5: Handelsvolume, beschrijvende modellen (obs. 1.659)

| | Model 1 | Model 2 | Model 3 |
|------------------|-------------------|-----------------------|--------------------------|
| HV_{btc} L1 | 0,837*** (0,1087) | 0,837*** (0,1092) | 0,823*** (0,1089) |
| $SENT_{btc}$ | | 2,43 e+08 (1,77 e+09) | 1,24 e+09 (1,82 e+09) |
| VOL_{btc} | | | 1,58 e+11*** (2,90 e+10) |
| F-statistiek | 59,31 | 94,78 | 135,78 |
| R^2 | 0,701 | 0,701 | 0,714 |
| Aangepaste R^2 | 0,701 | 0,701 | 0,713 |

Noot: Samengevatte resultaten van de regressie op volume. Model 1 tot en met Model 3 rapporteren de resultaten van het beschrijvende regressiemodel. Elke kolom voegt een variabele toe aan het model. De coëfficiënt van het beschrijvende model is weergegeven, de Newey-West standaardfouten (tussen haakjes), de F-statistiek en de verklarende waarde van het model. HV_{btc} L1 = de eerste lag van handelsvolume, $SENT_{btc}$ = Twitter sentiment van influencers en VOL_{btc} = volatiliteit van Bitcoin (* = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$).

5.2.2. Voorspellende relaties

Voor de voorspellende modellen wordt voor elke variabele het aantal lags toegepast, dat is verkregen door het toepassen van de VAR Diagnostiek. Het aantal vertraagde waarden van rendement is nul waardoor er nog maar twee voorspellende modellen berekend worden. Zowel voor het beschrijvende als voorspellende model voor rendement, wordt geen significant resultaat gevonden. Bovendien liggen de coëfficiënten van de lags van SENT en de verklarende waarde van het model, erg laag. Er kan geconcludeerd worden dat er geen significant effect is van de onafhankelijke variabele SENT op de afhankelijke variabele R. Hypothese 1b, *Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante voorspellende relatie met het rendement op de Bitcoin*, kan daarom niet verworpen worden.

Tabel 5.6: Rendement, voorspellende modellen (obs. 1.656)

| | | Model 1 | Model 2 |
|------------------|----|-----------------|-----------------|
| $SENT_{btc}$ | L1 | -0,000 (0,0026) | -0,000 (0,0026) |
| | L2 | 0,001 (0,0025) | 0,001 (0,0024) |
| | L3 | -0,000 (0,0026) | -0,000 (0,0026) |
| | L4 | -0,003 (0,0024) | |
| VOL_{btc} | L1 | | 0,046 (0,0519) |
| | L2 | | -0,051 (0,0430) |
| | L3 | | 0,067* (0,0397) |
| | L4 | | -0,030 (0,0451) |
| F-statistiek | | 0,61 | 0,68 |
| R^2 | | 0,001 | 0,004 |
| Aangepaste R^2 | | -0,001 | -0,000 |

Noot: Samengevatte resultaten van de regressie op Rendement, gebaseerd op appendix 8.7.2. Model 1 en Model 2 rapporteren de resultaten van het voorspellende regressiemodel. Elke kolom voegt een variabele toe aan het model. De coëfficiënt van het voorspellende model is weergegeven, de Newey-West standaard fouten (tussen haakjes), de F-statistiek en de verklarende waarde van het model. $SENT_{btc}$ = Twitter sentiment van influencers en VOL_{btc} = volatiliteit van Bitcoin (* = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$).

Voor de afhankelijke variabele volatiliteit zijn in Tabel 5.6 ook de coëfficiënten van de modellen uitgewerkt. De hypothese die hiermee getest wordt, luidt: *Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante voorspellende relatie met de volatiliteit op de Bitcoin.* Voor Model 3 wordt een ontbrekende F-statistiek geconstateerd, doordat er wederom te veel variabelen zijn ten op zichte van het aantal beschikbare vrijheidsgraden. Voor zowel Model 2 als Model 3 geldt dat sentiment (net als in het beschrijvende model) geen significante correlatie met volatiliteit heeft in het model. Daarmee kan ook hypothese 2b niet verworpen worden.

Tabel 5.7: Volatiliteit, voorspellende modellen (obs. 1.656)

| | | Model 1 | Model 2 | Model 3 |
|------------------|----|-------------------|-------------------|--------------------------|
| VOL_{btc} | L1 | 0,347*** (0,0362) | 0,349*** (0,0360) | 0,358*** (0,0367) |
| | L2 | 0,086*** (0,0323) | 0,086*** (0,0322) | 0,083** (0,0326) |
| | L3 | 0,107*** (0,0285) | 0,105*** (0,0285) | 0,092*** (0,0293) |
| | L4 | 0,148*** (0,0292) | 0,148*** (0,0292) | 0,148*** (0,0303) |
| $SENT_{btc}$ | L1 | | 0,003 (0,0018) | 0,002 (0,0018) |
| | L2 | | -0,002 (0,0016) | -0,002 (0,0016) |
| | L3 | | -0,002 (0,0017) | -0,002 (0,0018) |
| HV_{btc} | L1 | | | -4,93 e-14** (2,33 e-14) |
| | L2 | | | 3,86 e-14 (2,71 e-14) |
| | L3 | | | 3,43 e-14 (2,27 e-14) |
| | L4 | | | -7,22 e-15 (2,28 e-14) |
| F-statistiek | | 113,54 | 65,25 | . |
| R^2 | | 0,284 | 0,285 | 0,287 |
| Aangepaste R^2 | | 0,282 | 0,282 | 0,283 |

Noot: Samengevatte resultaten van de regressie op Volatiliteit, gebaseerd op appendix 8.7.2. Model 1 tot en met Model 3 rapporteren de resultaten van het beschrijvende regressiemodel. Elke kolom voegt een variabele toe aan het model. De coëfficiënt van het voorspellende model is weergegeven, de Newey-West standaard fouten (tussen haakjes), de F-statistiek en de verklarende waarde van het model. VOL_{btc} = volatiliteit van Bitcoin, $SENT_{btc}$ = Twitter sentiment van influencers en HV_{btc} = handelsvolume van Bitcoin (* = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$).

De laatste hypothese voor de voorspellende relaties luidt: *Twittersentiment van de 15 belangrijkste Twitter influencers heeft geen statistisch significante voorspellende relatie met het handelsvolume op Bitcoin.* Ook voor deze derde categorie regressiemodellen wordt geen significantie gevonden voor de variabele van interesse (zie Tabel 5.8), maar ook niet voor de vertraagde waardes van de afhankelijke variabele zelf en de controle variabele. Daarmee kunnen we ook hypothese 3b, die stelt dat er geen statistische voorspellende relatie is tussen de afhankelijke variabele handelsvolume en de onafhankelijke variabele Twittersentiment, aannemen.

Tabel 5.8: Handelsvolume, voorspellende modellen (obs. 1.656)

| | | Model 1 | Model 2 | Model 3 |
|------------------|----|----------------|------------------------|---------------------------|
| HV_{btc} | L1 | 0,287 (0,1872) | 0,289 (0,1867) | 0,260 (0,1891) |
| | L2 | 0,189 (0,1369) | 0,189 (0,1362) | 0,217 (0,1521) |
| | L3 | 0,262 (0,1641) | 0,261 (0,1636) | 0,266 (0,1745) |
| | L4 | 0,211 (1,1321) | 0,209 (0,1322) | 0,210 (0,1456) |
| $SENT_{btc}$ | L1 | | 2,48 e+09 (1,75 e+09) | 2,67 e+09 (1,84 e+09) |
| | L2 | | -1,01 e+09 (1,25 e+09) | -1,49 e+09 (1,22 e+09) |
| | L3 | | -3,04 e+08 (1,19 e+08) | -3,73 e+08 (1,09 e+09) |
| VOL_{btc} | L1 | | | 9,21 e+10** (4,23 e+10) |
| | L2 | | | -8,19 e+10*** (3,12 e+10) |
| | L3 | | | -2,50 e+10 (4,08 e+10) |
| | L4 | | | -1,95 e+10 (3,49 e+10) |
| F-statistiek | | 190,20 | 204,63 | 176,18 |
| R^2 | | 0,789 | 0,789 | 0,794 |
| Aangepaste R^2 | | 0,788 | 0,788 | 0,792 |

Noot: Samengevatte resultaten van de regressie op Volume, gebaseerd op appendix 8.7.2. Model 1 tot en met Model 3 rapporteren de resultaten van het voorspellende regressiemodel. Elke kolom voegt een variabele toe aan het model. De coëfficiënt van het voorspellende model is weergegeven, de Newey-West standaard fouten (tussen haakjes), de F-statistiek en de verklarende waarde van het model. HV_{btc} = handelsvolume van Bitcoin, $SENT_{btc}$ = Twitter sentiment van influencers en VOL_{btc} = volatiliteit van Bitcoin (* = $p < 0,1$; ** = $p < 0,05$; *** = $p < 0,01$).

De beschrijvende resultaten van de marktactiviteiten van Bitcoin in paragraaf 5.2.1 laten zien dat er een significante negatieve relatie is tussen het Twittersentiment van influencers en de volatiliteit. Echter, deze coëfficiënt is erg klein (-0,005) en daarom zal het Twittersentiment weinig invloed hebben op de afhankelijke variabele volatiliteit. Voor de afhankelijke variabelen rendement en handelsvolume zijn geen significante modellen gevonden met de variabele van interesse, Twittersentiment. Ook voor de voorspellende relaties zijn er weinig significante coëfficiënten die een rol spelen in het voorspellen van de afhankelijk variabelen. Voor zowel de beschrijvende als de voorspellende modellen geldt dat er weinig tot geen significante invloed is van Twittersentiment op de afhankelijke variabelen.

5.2.3. Granger Causaliteit

Een belangrijke hypothese om de centrale onderzoeksvraag van dit onderzoek te kunnen beantwoorden, is of er tussen het Twittersentiment van influencers en de marktactiviteit van Bitcoin sprake is van Granger causaliteit en welke richting de voorspellende causaliteit op gaat (hypothese 4).

Tabel 5.9: Granger Causaliteitstest

| Null Hypothese | F Stat. | Prob. |
|--|---------|-------|
| Rendement voorspelt niet Twittersentiment | 5,757 | 0,218 |
| Twitter sentiment voorspelt niet rendement | 2,056 | 0,725 |
| Volatiliteit voorspelt niet Twittersentiment | 4,645 | 0,326 |
| Twitter Sentiment voorspelt niet volatiliteit | 3,239 | 0,519 |
| Handelsvolume voorspelt niet Twittersentiment | . | . |
| Twitter sentiment voorspelt niet handelsvolume | 4,465 | 0,347 |

Noot: Resultaten van de Granger Causaliteits Wald test

De resultaten van de Granger causaliteitstest bevestigen de bevindingen van de toetsen over de voorspellende waarde van de onafhankelijke variabelen op de afhankelijke variabele in paragraaf 5.2.2. De significantieniveaus van de uitgevoerde Granger causaliteitstesten liggen alle boven de waarde van 5%. Er is daarom ook in dit geval geen sprake van een significante voorspellende relatie tussen het Twittersentiment en de variabelen van de marktactiviteit van Bitcoin. De nulhypothese van ‘geen Granger causaliteit’ wordt voor alle variabelen aangenomen en daardoor kan hypothese 4, *Twittersentiment van influencer tweets* ($SENT_{btc}$) en de *marktsegmenten van Bitcoin* ($R_{btc}, HV_{btc}, VOL_{btc}$), niet verworpen worden.

6

Conclusie

Door middel van de hypothesetoetsen in dit onderzoek is getracht antwoord te geven op de onderzoeksvraag: *Wat is de relatie tussen het Twittersentiment van crypto-influencers en de Bitcoin markt?* De hypothesen hebben het verband getest tussen het Twittersentiment in influencerstweets en de marktactiviteiten van Bitcoin, en zijn dieper ingegaan op de soort relatie en de richting ervan. In dit hoofdstuk zal een kort overzicht van de resultaten van de empirische testen weergegeven worden en zullen hier conclusies en implicaties aan verbonden worden.

6.1. Hypothesen toetsen en het beantwoorden van de Onderzoeksvraag

Allereerst is er gekeken naar een algemene correlatie, waarbij een hoge mate van correlatie zou duiden op een sterk verband tussen de variabelen en een lage waarde op een zwakke relatie. Voor de volatiliteit en het handelsvolume van Bitcoin is een significante correlatie gevonden met Twittersentiment, maar voor rendement is dit niet het geval. Hoewel er sprake is van significante correlaties tussen de onafhankelijke variabele (SENT) en de afhankelijke variabelen (VOL en HV), is de correlatie van lage waarde.

De Augmented Dickey Fuller test en de Phillips-Perron test zijn uitgevoerd om na te gaan of er sprake is van een random walk. Tegen de verwachting in wezen beide testen voor alle variabelen uit dat de data stationair is. De toetsen duiden hiermee volgens de EMH op een efficiënte cryptomarkt.

De resultaten van de regressiemodellen laten zien dat er bijna zowel geen beschrijvende als voorspellende waarde is tussen Twittersentiment van influencers en de marktactiviteiten van Bitcoin. Alleen de afhankelijke variabele volatiliteit laat een significante correlatie zien in het beschrijvende model. Echter, dit betreft wel een lage waarde. Aan de hand van deze resultaten kan dus gesteld worden dat Twittersentiment alleen in een kleine mate gebruikt kan worden

in het verklaren van de volatiliteit van Bitcoin. Deze resultaten laten zien dat het rendement niet voorspeld kan worden aan de hand van historische informatie. Dit is in lijn met de zwakke EMH, wat duidt op een efficiënte markt.

De Granger causaliteitstest laat zien dat er voor geen enkele variabele een significant Granger causaal verband te vinden is. Met andere woorden, in het algemeen kan geconcludeerd worden dat het niet mogelijk is om de toekomstige bewegingen van de marktactiviteiten van Bitcoin te voorspellen door middel van het sentiment van influencer tweets. Dit sluit aan bij de eerdere bevindingen van de regressiemodellen.

Concluderend uit de resultaten kan er gesteld worden dat over het algemeen tweets van influencers geen invloed uitoefenen op de markt van Bitcoin. De markt lijkt redelijk efficiënt na het toepassen van de stationariteitstoetsen, voor zowel de Granger causaliteitstest als de Phillips-Perron test. De correlatietabel (Tabel 4.1) en de coëfficiënten van de beschrijvende en voorspellende modellen laten weinig significante resultaten zien. Ondanks de lags van zowel de onafhankelijke als de afhankelijke variabelen die aan het regressiemodel toegevoegd zijn, blijft een significant verband uit.

6.2. Implicatie en relevantie

Alle resultaten bij elkaar, kan aangenomen worden dat er geen significante relatie is tussen het sentiment van influencer tweets en de Bitcoin markt. De onderzoeksvraag die deze relatie toetst, kan hiermee beantwoord worden:

Wat is de relatie tussen het Twittersentiment van crypto-influencers en de Bitcoin markt?

Er is zowel geen beschrijvende als voorspellende relatie tussen de onafhankelijke variabele Twittersentiment en de afhankelijke variabelen rendement, volatiliteit en handelsvolume. Oftewel, potentiële investeerders dienen hun investeringsbeslissingen niet te baseren op de tweets van influencers. Het blijft zaak dat potentiële investeerders zich verdiepen in de cryptomarkt om weloverwogen investeringsbeslissingen te nemen.

6.3. Limitaties en aanbevelingen voor verder onderzoek

De resultaten en conclusies geven verschillende ideeën voor vervolgonderzoek. Een limitatie in dit onderzoek is namelijk dat het onderzoek gebaseerd is een lijst van crypto-influencers van

CryptoWeekly in 2021. Deze lijst betreft de meest invloedrijke twitteraars op het gebied van de gehele cryptomuntenmarkt en niet alleen Bitcoin. Om de beste lijst van influencers samen te stellen, kan het beste een eigen analyse gedaan worden op Bitcoin gerelateerde tweets aan de hand van het aantal volgers, het aantal likes en het aantal retweets. Optimaal zou zijn om dit voor alle onderzochte jaren separaat uit te zoeken, omdat voor een influencer die invloedrijk is in 2021, dit niet het geval hoeft te zijn voor het jaar 2017.

Een andere limitatie van dit onderzoek is dat het sentiment van een tweet is geanalyseerd wanneer het woord 'bitcoin' of de afkorting 'btc' in de tweet vermeld werd. Het kan zijn dat er sprake is van algemeen negatief sentiment in de tweet, maar dat hetgeen dat vermeld wordt over Bitcoin juist een positieve wending heeft. Dit zal relatief meevallen, omdat tweets een maximale lengte van 140 tekens hebben. Toch zou een geavanceerde sentimentanalyse toegepast kunnen worden om nog concreter het Bitcoin sentiment eruit te filteren.

Uit de bevindingen van het onderzoek van Abraham e.a. (2018) blijkt dat sentimentanalyse minder effectief is voor prijswijzigingen voor cryptovaluta wanneer er sprake is van een dalende prijs. De meeste tweets gerelateerd aan cryptomunten zijn neutraal (geen enkel sentiment) of positief van aard, ongeacht de prijsverandering. Daarom zou in een vervolgonderzoek afzonderlijk het effect van het positieve en negatieve sentiment op de marktactiviteit van Bitcoin kunnen worden geanalyseerd.

Bibliografie

- Abraham, J., Higdon, D., Nelson, J. & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Science Review*, 1(3), 1.
- Aesuli. (g.d.). Aesuli/sentiwordnet: The sentiwordnet sentiment lexicon. <https://github.com/aesuli/SentiWordNet>
- Alle cryptovaluta's. (g.d.). <https://coinmarketcap.com/nl/all/views/all/>
- Ante, L. (2021). How elon musk's twitter activity moves cryptocurrency markets. *Available at SSRN 3778844*.
- atNickVega. (2021). More than 1 in 3 cryptocurrency investors know little to nothing about it, survey finds. <https://www.cnbc.com/2021/03/04/survey-finds-one-third-of-crypto-buyers-dont-know-what-theyre-doing.html>
- Band, A. (2021). Twitter sentiment analysis using vader amp; tweepy. <https://python.plainenglish.io/twitter-sentiment-analysis-using-vader-tweepy-b2a62fba151e>
- Belanche, D., Casaló, L. V., Flavián, M. & Ibáñez-Sánchez, S. (2021). Understanding influencer marketing: The role of congruence between influencers, products and consumers. *Journal of Business Research*, 132, 186–195.
- Blair, B. J., Poon, S.-H. & Taylor, S. J. (2010). Forecasting s&p 100 volatility: The incremental information content of implied volatilities and high-frequency index returns. *Handbook of quantitative finance and risk management* (pp. 1333–1344). Springer.
- Böhme, R., Christin, N., Edelman, B. & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2), 213–38.
- Canina, L. & Figlewski, S. (1993). The informational content of implied volatility. *The Review of Financial Studies*, 6(3), 659–681.
- Casaló, L. V., Flavián, C. & Ibáñez-Sánchez, S. (2020). Influencers on instagram: Antecedents and consequences of opinion leadership. *Journal of Business Research*, 117, 510–519.
- Cheuque Cerda, G. & L. Reutter, J. (2019). Bitcoin price prediction through opinion mining. *Companion Proceedings of The 2019 World Wide Web Conference*, 755–762.
- Christensen, B. J. & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, 50(2), 125–150.
- Chuen, D. L. K. (2015). *Handbook of digital currency: Bitcoin, innovation, financial instruments, and big data*. Academic Press.

- Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427–431.
- Fildes, R. e.a. (2008). Francis x. diebold, editor, elements of forecasting , thomson, southwestern: Ohio, us (2007) isbn 978-0-324-35904-6, p. 458 hardcover. *International Journal of Forecasting*, 24(3), 552–553.
- French, K. R., Schwert, G. W. & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics*, 19(1), 3–29.
- Garman, M. B. & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of business*, 67–78.
- Global cryptocurrency ownership data 2021. (2021). <https://triple-a.io/crypto-ownership/>
- Grafström, J., Jakobsson, L. & Wiede, P. (2018). The impact of influencer marketing on consumers' attitudes.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, 424–438.
- How many bitcoin users are there? (2021). <https://www.buybitcoinworldwide.com/how-many-bitcoin-users/>
- Kang, H.-J., Lee, S.-G. & Park, S.-Y. (2021). Information efficiency in the cryptocurrency market: The efficient-market hypothesis. *Journal of Computer Information Systems*, 1–10.
- Ki, C.-W. C., Cuevas, L. M., Chong, S. M. & Lim, H. (2020). Influencer marketing: Social media influencers as human brands attaching to followers and yielding positive marketing results by fulfilling needs. *Journal of Retailing and Consumer Services*, 55, 102133.
- Kirkpatrick, D. (2016). Twitter says influencers are almost as trusted as friends. <https://www.marketingdive.com/news/twitter-says-influencers-are-almost-as-trusted-as-friends/419076/>
- Koers, grafieken, marktkapitalisatie en andere metrics voor bitcoin (btc). (g.d.). <https://coinmarketcap.com/nl/currencies/bitcoin/historical-data/>
- Kraaijeveld, O. & De Smedt, J. (2020). The predictive power of public twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65, 101188.
- Kumar, S., Morstatter, F. & Liu, H. (2014). *Twitter data analytics*. Springer.
- Layton, R., Watters, P. & Dazeley, R. (2010). Authorship attribution for twitter in 140 characters or less. *2010 Second Cybercrime and Trustworthy Computing Workshop*, 1–8.

- Lim, X. J., Radzol, A., Cheah, J. & Wong, M. W. (2017). The impact of social media influencers on purchase intention and the mediation effect of customer attitude. *Asian Journal of Business Research*, 7(2), 19–36.
- López-Cabarcos, M., Pérez-Pico, A. M. & López-Pérez, M. L. (2019). Does social network sentiment influence s&p 500 environmental & socially responsible index? *Sustainability*, 11(2), 320.
- Mao, H., Counts, S. & Bollen, J. (2011). Predicting financial markets: Comparing survey, news, twitter and search engine data. *arXiv preprint arXiv:1112.1051*.
- Miraz, M. H. & Ali, M. (2018). Applications of blockchain technology beyond cryptocurrency. *arXiv preprint arXiv:1801.03528*.
- Multicollineariteit. (g.d.). <https://nl.wikipedia.org/wiki/Multicollineariteit#:~:text=Multicollineariteit%20is%20een%20statistisch%20fenomeen,het%20model%20voorspeld%20kan%20worden>.
- Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, 21260.
- Newey, W. K. & West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.
- Phillips, P. C. & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335–346.
- Piñeiro-Chousa, J. R., López-Cabarcos, M. Á. & Pérez-Pico, A. M. (2016). Examining the influence of stock market variables on microblogging sentiment. *Journal of Business Research*, 69(6), 2087–2092.
- Quick, J. & Wright, G. (2021). Global social media statistics research summary [updated 2021]. <https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/>
- Seybert, L. A. & Katzenstein, P. J. (2018). High-tech: Power and unpredictability at the technological frontier and in bitcoin. *Protean Power: Exploring the Uncertain and Unexpected in World Politics*, 146, 124.
- Stadt, K. V. d. (2017). 2017: Het jaar waarin bitcoin bleef verbazen. https://datanews.knack.be/ict/nieuws/2017-het-jaar-waar-in-bitcoin-bleef-verbazen/article-longread-942049.html?cookie_check=1626808628
- Stubb, C., Nyström, A.-G. & Colliander, J. (2019). Influencer marketing: The impact of disclosing sponsorship compensation justification on sponsored content effectiveness. *Journal of Communication Management*.

- Sun, A., Lachanski, M. & Fabozzi, F. J. (2016). Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. *International Review of Financial Analysis*, 48, 272–281.
- Tafti, A., Zotti, R. & Jank, W. (2016). Real-time diffusion of information on twitter and the financial markets. *PloS one*, 11(8), e0159226.
- Thorbecke, W. (1997). On stock market returns and monetary policy. *The Journal of Finance*, 52(2), 635–654.
- Top 100 cryptocurrency list of influential people in crypto. (g.d.). <https://cryptoweekly.co/100>
- Trade the future. (g.d.). <https://bitvavo.com/nl>
- Twintproject. (g.d.). Twintproject/twint: An advanced twitter scraping amp; osint tool written in python that doesn't use twitter's api, allowing you to scrape a user's followers, following, tweets and more while evading most api limitations. <https://github.com/twintproject/twint>
- Yamamoto, H., Sakaji, H., Matsushima, H., Yamashita, Y., Osawa, K., Izumi, K. & Shimada, T. (2019). Forecasting crypto-asset price using influencer tweets. *International Conference on Advanced Information Networking and Applications*, 940–951.
- Yang, S. Y., Mo, S. Y. K. & Liu, A. (2015). Twitter financial community sentiment and its predictive relationship to stock market movement. *Quantitative Finance*, 15(10), 1637–1656.
- York, R. (2012). Residualization is not the answer: Rethinking how to address multicollinearity. <https://www.sciencedirect.com/science/article/pii/S0049089X12001160>

A

Appendix

```
1 import twint
2
3 influencers = ["danheld"...,"valeryvavilov"]
4 searchTerm = ["bitcoin", "btc"]
5
6 for x in influencers:
7     for y in searchTerm:
8         c = twint.Config()
9         c.Username = x
10        c.Store_csv = True
11        c.Since = "2009-01-01"
12        c.Until = "2021-07-18"
13        c.Search = y
14        # CSV Fieldnames
15        c.Custom["tweet"] = ["id", "date", "time", "username", "tweet", "hashtags"
16        , "replies_count", "retweets_count", "likes_count"]
17        # Name of the directory
18        c.Output = "twintFinal.csv"
19
20        twint.run.Search(c)
```

Listing A.1: Python code die gebruikt is om de tweets te verzamelen.

Tabel A.1: Voorbeeld Influencer Twitter Data

| date | time | username | tweet | hashtags |
|-----------|----------|----------|---|----------|
| 17-7-2021 | 23:38:31 | danheld | @ErikVoorhees Completely different scenario. Really disingenuous to compare them. Focus on addressing the Thorchain problems, not finding false equivalencies with Bitcoin. | [] |
| 17-7-2021 | 16:57:01 | danheld | @ninosgr No, Bitcoin. | [] |
| 17-7-2021 | 16:56:03 | danheld | @diamondclawz No, Bitcoin. | [] |
| 17-7-2021 | 16:54:48 | danheld | @DominicJosu Government currencies aren't even close to being the same as Bitcoin. Digital currency isn't the innovation, immutable/decentralized money is | [] |
| 17-7-2021 | 15:12:10 | danheld | Bitcoin compared to the adoption curves of other industries. It is still so early. https://t.co/xI9BkLqOIC | [] |
| ... | ... | ... | ... | ... |

Voorbeeld van de verzamelde twitter data

```

1
2 from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
3 import pandas as pd
4 import numpy as np
5 import re
6 analyzer = SentimentIntensityAnalyzer()
7
8 df = pd.read_csv('twintFinal.csv', sep=',', encoding='utf-8')
9
10 # cleaning the tweets
11 def remove_pattern(input_txt, pattern):
12     r = re.findall(pattern, input_txt)
13     for i in r:
14         input_txt = re.sub(i, '', input_txt)
15     return input_txt
16
17 def clean_tweets(tweets):
18     # remove twitter Return handles (RT @xxx:)

```

```

19 tweets = np.vectorize(remove_pattern)(tweets, "RT @[\\w]*:")
20
21 # remove twitter handles (@xxx)
22 tweets = np.vectorize(remove_pattern)(tweets, "@[\\w]*")
23
24 # remove URL links (httpxxx)
25 tweets = np.vectorize(remove_pattern)(tweets, "https?:/[A-Za-z0-9./]*")
26
27 # remove special characters, numbers, punctuations (except for #)
28 tweets = np.core.defchararray.replace(tweets, "[^a-zA-Z]", " ")
29
30 return tweets
31
32 df['text'] = clean_tweets(df['text'])
33 print(df['text'].head())
34
35 scores = []
36 # Declare variables for scores
37 compound_list = []
38 positive_list = []
39 negative_list = []
40 neutral_list = []
41 for i in range(df['text'].shape[0]):
42     # print(analyser.polarity_scores(sentiments_pd['text'][i]))
43     compound = analyzer.polarity_scores(df['text'][i])["compound"]
44     pos = analyzer.polarity_scores(df['text'][i])["pos"]
45     neu = analyzer.polarity_scores(df['text'][i])["neu"]
46     neg = analyzer.polarity_scores(df['text'][i])["neg"]
47
48     scores.append({"Compound": compound,
49                  "Positive": pos,
50                  "Negative": neg,
51                  "Neutral": neu
52                  })
53
54 sentiments_score = pd.DataFrame.from_dict(scores)
55 df = df.join(sentiments_score)
56 print(df.head())
57 df.to_csv('finalSentiment.csv', sep=',', encoding='utf-8')

```

Listing A.2: De python code die is gebruikt om het sentiment per tweet te bepalen.

Tabel A.2: Voorbeeld Tweet en Sentiment

| Tweet | Verwerkte Tweet | Gemiddeld | Positief | Negatief | Neutraal |
|--|---|------------------|-----------------|-----------------|-----------------|
| @ErikVoorhees Completely different scenario. Really disingenuous to compare them. Focus on addressing the Thorchain problems, not finding false equivalencies with Bitcoin. | Completely different scenario. Really disingenuous to compare them. Focus on addressing the Thorchain problems, not finding false equivalencies with Bitcoin. | -0.4019 | 0 | 0.124 | 0.876 |
| @ninosgr No, Bitcoin. | No, Bitcoin. | 0 | 0 | 0 | 1 |
| @diamondclawz No, Bitcoin. | No, Bitcoin. | 0 | 0 | 0 | 1 |
| @DominicJosu Government currencies aren't even close to being the same as Bitcoin. Digital currency isn't the innovation, immutable/ decentralized money is | Government currencies aren't even close to being the same as Bitcoin. Digital currency isn't the innovation, immutable/ decentralized money is | 0.3818 | 0.126 | 0 | 0.874 |
| Bitcoin compared to the adoption curves of other industries. It is still so early. | Bitcoin compared to the adoption curves of other industries. It is still so early. | 0 | 0 | 0 | 1 |
| ... | | | | | |

Noot: Voorbeeld van verwerkte tweets en hun berekende sentiment.

Tabel A.3: Bitcoin Data

| Date | Open* | High | Low | Close** | Volume |
|---------------------|-------------|-------------|-------------|-------------|------------------|
| Jul 17, 2021 | \$31,397.31 | \$31,935.94 | \$31,223.99 | \$31,533.07 | \$18,895,018,942 |
| Jul 16, 2021 | \$31,841.55 | \$32,218.41 | \$31,100.67 | \$31,421.54 | \$23,699,476,918 |
| Jul 15, 2021 | \$32,827.87 | \$33,159.64 | \$31,175.71 | \$31,780.73 | \$21,300,524,237 |
| Jul 14, 2021 | \$32,723.85 | \$33,061.40 | \$31,639.12 | \$32,822.35 | \$21,376,531,210 |
| Jul 13, 2021 | \$33,125.47 | \$33,327.10 | \$32,261.42 | \$32,702.03 | \$19,120,856,669 |

Noot: Voorbeeld van de verzamelde Bitcoin data

Tabel A.4: Aantal tweets per influencer

| Influencer | 2017 | 2018 | 2019 | 2020 | 2021* |
|----------------------|-------------|-------------|-------------|-------------|-------------|
| Meltem Demirors | 303 | 203 | 329 | 472 | 556 |
| Adam Back | 1172 | 449 | 1224 | 2388 | 1671 |
| Caitlin Long | 58 | 292 | 114 | 108 | 234 |
| Andreas Antonopoulos | 341 | 248 | 257 | 707 | 327 |
| Erik Voorhees | 584 | 280 | 337 | 475 | 505 |
| Anthony Pompliano | 118 | 864 | 1678 | 967 | 1248 |
| Jameson Lopp | 1109 | 632 | 392 | 782 | 441 |
| Vinny Lingham | 741 | 296 | 73 | 66 | 122 |
| Charlie Shrem | 398 | 137 | 454 | 435 | 204 |
| Dan Held | 129 | 638 | 1471 | 2362 | 1891 |
| Riccardo Spagni | 443 | 583 | 579 | 367 | 253 |
| Ran NeuNer | 261 | 321 | 182 | 174 | 198 |
| Emin Gün Sirer | 492 | 348 | 239 | 139 | 91 |
| Bobby Lee | 397 | 173 | 94 | 114 | 82 |
| Roger Ver | 158 | 146 | 104 | 125 | 43 |
| Totaal | 6704 | 5610 | 7527 | 9681 | 7866 |

Noot: Aantal Bitcoin gerelateerde tweets per jaar per influencer (Tot en met 18 juli 2021)

Tabel A.5: VAR Diagnostiek voor Rendement

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|-----|---------|---------|----|-------|-----------|-----------|-----------|-----------|
| 0 | 4245,26 | | | | 0,000348 | -5,12592 | -5,12471* | -5,12265* |
| 1 | 4246,52 | 2,51040 | 1 | 0,113 | 0,000348 | -5,12623 | -5,12381 | -5,11969 |
| 2 | 4248,02 | 3,01290 | 1 | 0,083 | 0,000348* | -5,12684* | -5,12321 | -5,11704 |
| 3 | 4248,03 | 0,00558 | 1 | 0,940 | 0,000348 | -5,12564 | -5,12079 | -5,11256 |
| 4 | 4248,08 | 0,09714 | 1 | 0,755 | 0,000348 | -5,12449 | -5,11843 | -5,10815 |

Noot: VAR Diagnostiek voor Rendement met de informatiecriteria FPE, AIC, HQIC, SBIC en LR met een maximum van vier lags

Tabel A.6: VAR Diagnostiek voor Volatiliteit

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|-----|---------|--------|----|-------|-----------|----------|-----------|-----------|
| 0 | 4556,26 | | | | 0,000239 | -5,50152 | -5,50031 | -5,49825 |
| 1 | 4769,43 | 426,34 | 1 | 0,000 | 0,000185 | -5,75777 | -5,75534 | -5,75123 |
| 2 | 4792,48 | 46,098 | 1 | 0,000 | 0,00018 | -5,78440 | -5,78076 | -5,77459 |
| 3 | 4814,45 | 43,95 | 1 | 0,000 | 0,000176 | -5,80973 | -5,80488 | -5,79666 |
| 4 | 4832,82 | 36,73* | 1 | 0,000 | 0,000172* | -5,8307* | -5,82464* | -5,81436* |

Noot: VAR Diagnostiek voor Volatiliteit met de informatiecriteria FPE, AIC, HQIC, SBIC en LR met een maximum van vier lags

Tabel A.7: VAR Diagnostiek voor Handelsvolume

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|-----|----------|---------|----|-------|-----------|----------|----------|----------|
| 0 | -41761,3 | | | | 4,7 e+20 | 50,4375 | 50,4387 | 50,4408 |
| 1 | -40761,7 | 1999,1 | 1 | 0,000 | 1,4 e+20 | 49,2316 | 49,2340 | 49,2381 |
| 2 | -40611,7 | 300,1 | 1 | 0,000 | 1,2 e+20 | 49,0515 | 49,0552 | 49,0614 |
| 3 | -40511,4 | 200,57 | 1 | 0,000 | 1,0 e+20 | 48,9316 | 48,9365 | 48,9447 |
| 4 | -40473,8 | 75,139* | 1 | 0,000 | 1,0 e+20* | 48,8875* | 48,8935* | 48,9038* |

Noot: VAR Diagnostiek voor Handelsvolume met de informatiecriteria FPE, AIC, HQIC, SBIC en LR met een maximum van vier lags

Tabel A.8: VAR Diagnostiek voor Sentiment

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|-----|---------|---------|----|-------|-----------|-------------|-------------|-------------|
| 0 | 562,17 | | | | 0,029729 | - 0,677741 | - 0,67653 | - 0,674473* |
| 1 | 565,137 | 5,9341 | 1 | 0,015 | 0,029659 | - 0,680117 | - 0,677694 | - 0,673581 |
| 2 | 567,815 | 5,3562 | 1 | 0,021 | 0,029599 | - 0,682144 | - 0,678509 | - 0,672339 |
| 3 | 570,928 | 6,2255* | 1 | 0,013 | 0,029523* | - 0,684695* | - 0,679849* | - 0,671622 |
| 4 | 571,288 | 0,7209 | 1 | 0,396 | 0,029546 | - 0,683923 | - 0,677865 | - 0,667582 |

Noot: VAR Diagnostiek voor Sentiment met de informatiecriteria FPE, AIC, HQIC, SBIC en LR met een maximum van vier lags