



Investor sentiment and bitcoin prices

Dimitrios Koutmos¹

I am grateful to conference participants from the 15th NYCU International Finance Conference that was held in National Yang Ming Chiao Tung University at Hsinchu, Taiwan, as well as conference participants from the 2nd Annual Bitcoin Conference sponsored by the Mays Innovation Research Center at Texas A&M University in College Station, Texas. I am also very grateful to Dr. Cheng-Few Lee (the Editor) and the anonymous referees for their comments and suggestions.

Accepted: 13 July 2022 / Published online: 22 August 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Using a rich data set of transaction-level buy and sell orders from the major digital currency exchange Coinbase, we formulate a measure for investor sentiment and shed new evidence on the sentiment-return relation for bitcoin. Using a bootstrapped quantile regression procedure we show a significant and robust relation between rising sentiment and price increases, and vice versa, across the distribution of bitcoin price changes. This relation is shown to be robust when controlling for a variety of exchange-specific and blockchain-wide variables. This relation is also robust when controlling for aggregate momentum across major cryptocurrencies. This finding is important as our data sample spans a period before and after the introduction of futures markets for bitcoin, which has arguably resulted in a regime shift in the time series behavior of its price. Taken together, our results show that bitcoin prices can undergo regime changes and that conventional regression-type models that focus on the center of the distribution of bitcoin price changes can yield misleading estimates.

Keywords Bitcoin · Bootstrap · Investor sentiment · Quantile regression · Robustness

JEL Classification C21 · C32 · G10 · G12 · G14

1 Introduction

It is yet to be definitively understood what drives cryptocurrency prices. One school of thought, inspired from the vast and very diverse literature on network theory, posits that digital currencies can be valued on the basis of actual users and their interactions together

✉ Dimitrios Koutmos
dimitrios.koutmos@tamucc.edu

¹ Texas A&M University - Corpus Christi, Corpus Christi, USA

(Alabi 2017; Economides 1993; Metcalfe 2013; Wheatley et al. 2019). Others argue that such prices may be driven by irrationality and resemble the historic bubbles that have come to pass throughout human history (Wenker 2014). From the standpoint of regulators, there are also conflicting views regarding digital currencies, their legitimacy, and what function they serve. For example, policymakers regularly issue stern warnings arguing that such currencies cannot replace, or, fundamentally serve, the function that traditional fiat currencies perform (Lo and Wang 2014). In addition, they may aid and abet criminal activity (Grinberg, 2012). Despite such warnings, however, the number of businesses accepting such digital currencies, like bitcoin, is ever-increasing.¹ Along with this growing interest, central bankers are presently discussing the feasibility of using government-backed, rather than decentralized, digital currencies in society.²

The aforementioned give only an infinitesimal and very partial glimpse of the divergences in opinion regarding cryptocurrencies. They also show the diverse and competing stakeholders that are allured into understanding more their possible uses and intrinsic properties. Perhaps most alluring to all such stakeholders is the unprecedented price volatility these digital coins exhibit. For example, in 2015, the average price of bitcoin oscillated around \$270 (in USD). Bitcoin's price first peaked December 17, 2017, where it was valued at about \$19,533 – an arithmetic return of over 7000% over its 2015 price level. After the introduction of bitcoin futures by the Chicago Board Options Exchange (CBOE) on December 10, 2017, and, subsequently, by the Chicago Mercantile Exchange (CME) on December 18, 2017, the price of bitcoin declined precipitously. It reached a low of about \$3,200 in mid-December of 2018 before rising rapidly to another peak of over \$12,600 in late June of 2019. Since the global outbreak of the severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2), bitcoin's price continues to oscillate wildly.

This type of price volatility can render conventional regression modeling approaches unstable, with coefficient estimates that lack robustness as bitcoin's price behavior undergoes regime changes across time. Lahmiri and Bekiros (2019) show strong evidence that cryptocurrency price movements are governed by regime shifts. Others show how bitcoin price regimes can be described using Markovian-type models (Koutmos 2020; Koutmos and Payne 2021; Ma et al. 2020). Yang et al. (2020) show that bitcoin miners themselves may exhibit behaviors that can be modeled using Markov models. The recurring theme in this particular literature is that bitcoin's volatility results in a significant quantity of outliers and thus thicker tails than what is commonly observed from Gaussian processes. In the words of Yermack (2015), "...bitcoin faces a number of obstacles in becoming a useful unit of account...one problem arises from its extreme volatility...the value of a bitcoin compared to other currencies changes greatly on a day-to-day basis..." (p. 38).

From a data analysis perspective, this is problematic because conclusions cannot be aptly drawn from conventional regression techniques. Catania et al. (2019) show that cryptocurrencies, like bitcoin, are subject to instabilities both in the means and at higher moments of their returns. As a result, coefficient estimates can be unstable across time, making it difficult to ascertain what factors are important in driving cryptocurrency prices and when they

¹ A partial list of major companies currently accepting bitcoins can be found here: <https://99bitcoins.com/bitcoin/who-accepts/>.

² President of the Federal Reserve Bank of Cleveland, Loretta J. Mester, has discussed the possibility of having a central bank digital currency (CBDC) at a keynote session during the 20th Anniversary Chicago Payments Symposium. This keynote talk can be found here: <https://www.clevelandfed.org/newsroom-and-events/speeches/sp-20200923-payments-and-the-pandemic>.

actually matter. In a similar vein, Atsalakis et al. (2019) show that there are strong nonlinearities inherent in cryptocurrency prices that can weaken the explanatory power of conventional regression models. In the words of Li and Wang (2017), and in view of their findings which seek to delineate the technologic and economic determinants of bitcoin, they argue, "...the relationships we identified could be subject to further changes as exchange markets develop...it will be necessary to revisit the model at some future time and consider the possibility of multiple regime changes in exchange rate dynamics..." (p. 59).

In light of the aforementioned, and focusing particularly on bitcoin, this study nests several motives and objectives together.³ First, and as mentioned, there is much debate as to what drives cryptocurrency prices and there is yet to be definitive agreement as to which variables matter and when they matter. Drawing on behavioral research, this study argues that investor sentiment is strongly linked to bitcoin price changes. Using order book data from Coinbase, which is a major bitcoin exchange, we estimate order flow imbalance across time and use this as a proxy for investor sentiment. This approach is motivated by behavioral research, such as the studies by Kumar and Lee (2006) and Chelley-Steeley et al. (2019), and references therein, which show that buy and sell orders reflect investor sentiment and can be linked to price changes. As we discuss more elaborately later on, using buy and sell order data for bitcoin can serve as a proxy for investor sentiment that is linked to shifts in bitcoin prices.

Second, several studies, such as those mentioned, find instability in coefficient estimates in models seeking to describe cryptocurrency price changes. This arises from thick tails in the distributions of their price changes. These thick tails are the result of outliers and regime changes which their prices frequently undergo, and, can be an issue when trying to obtain robust point estimates of regression coefficients. To directly address this, our study utilizes a bootstrapped quantile regression approach, motivated by Buchinsky (1995), Koenker and Basset (1978), Chakraborty (2003) and Pedersen (2015), among others, to construct robust estimates of the sentiment-return relation.

Third, this study draws on emergent research that seeks to describe bitcoin's price changes on the basis of microstructure and blockchain variables, as well as network theory. As mentioned, a segment of these studies examine whether bitcoin's value is driven by the number of users and the measurable interactions between these users. Drawing on some of these studies, our study employs a broad range of exchange-specific and blockchain-wide variables, respectively, and uses them as control variables to check the robustness of our estimated sentiment-return relation.

By combining these three motives and objectives, we depict a more complete picture of the entire distribution of bitcoin price changes. While the primary objective of our study is to quantify a robust sentiment-return relation, which we strive to achieve both empirically and theoretically, a secondary contribution that emerges is the observable heterogeneous role which microstructure and network variables play in explaining bitcoin returns across low, mid and high quantiles. This indicates that bitcoin explanatory or forecasting models are likely strongly affected by the regime, or, time sample, which they investigate. This

³ This study focuses on bitcoin, which is the largest cryptocurrency among the approximately 10,000 cryptocurrencies that are presently in circulation (see *coinmarketcap.com*). Presently, bitcoin constitutes around 50% of the market capitalization of all the cryptocurrencies in circulation. For the sample period of this study, which ranges from January 26, 2015 through October 23, 2020, the size of bitcoin's average daily trading volume was nearly \$10 billion. This amount far exceeds what is observable in other cryptocurrencies and makes bitcoin the leading digital coin for enthusiasts, merchants, speculators and investors alike.

raises questions about the robustness of their estimates and is a likely reason why aforementioned studies detect instability in their models. As mentioned by Li and Wang (2017), bitcoin prices undergo regime changes and thus display distinct statistical properties.

This study uses intraday order book data (buy and sell orders) of individual bitcoin investors to construct a proxy for investor sentiment and to quantify the sentiment-return relation across quantiles. We source our trading records from Coinbase since this is presently one of the largest digital currency exchanges in the world. Our objective, as we discuss further, is to specifically answer the following empirical question: *What is the relation between investor sentiment and price changes for bitcoin?*

With a unique data set of intraday buy and sell orders from 2015 to 2020, we construct a proxy for daily investor sentiment in order to shed light on this question and to quantify the sentiment-return relation across the conditional distribution of bitcoin price changes. Using our bootstrapped quantile regression approach to account for heteroskedasticity and ensure robustness in our results, we show reliable evidence that rising sentiment is associated with positive price changes while declining sentiment is associated with negative price changes. This relation remains robust across the distribution of bitcoin price changes, from periods of extreme negative returns (lower conditional quantiles) to periods of extreme positive returns (higher conditional quantiles). This relation also remains robust when controlling for a variety of exchange-specific and blockchain-wide variables. It also remains robust when controlling for aggregate momentum across major cryptocurrencies.

The remainder of this study is structured as follows. Section 2 discusses the data; specifically, the bitcoin prices we use from Coinbase as well as their statistical properties before and after the introduction of bitcoin futures. Section 2 also discusses the order book data and how we use it to construct a proxy of daily investor sentiment, as well as the exchange-specific and blockchain-wide variables which later serve as control variables. For Sect. 2, we also introduce a constructed index of aggregate cryptocurrency returns that we later use as yet an additional control variable to check the robustness of our sentiment-return relation. Section 3 discusses our bootstrapped quantile regression model and how it helps in achieving robust results. Sections 4 and 5 discuss the findings and conclude, respectively.

2 Data

This section discusses the sample data used in this study. Our data is sourced from the Coinbase exchange and Bitcoin's blockchain, respectively. In sub-Sects. 2.1 and 2.2 we discuss the bitcoin price data and its statistical properties before and after the introduction of futures by the CBOE. In sub-Sect. 2.3 we show how intraday order book data (buy and sell orders) can be used to construct a daily measure of investor sentiment. This measure serves as our primary regressor variable in quantifying the sentiment-return relation. Sub-Sect. 2.4 discusses the exchange-specific and blockchain-wide variables that are used as control variables. Finally, sub-Sect. 2.5 develops a cryptocurrency market index (*CRYIX*) in order to capture the effects of market-wide cryptocurrency market momentum. The construction and use of *CRYIX* serves as yet another control variable to gauge the explanatory power of the daily investor sentiment variable.

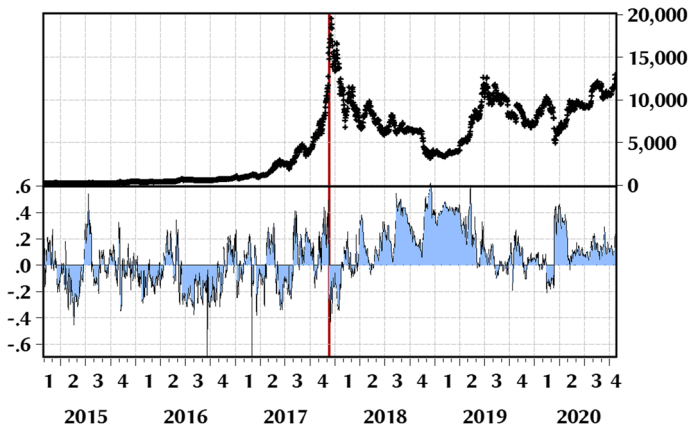


Fig 1 Time series plot of bitcoin prices and investor sentiment. This figure shows time series plots for daily bitcoin prices in USD (top and right axis) and investor sentiment (bottom and left axis), both of which are collected from the Coinbase digital currency exchange. The vertical red line that falls within the 4th quarter of the year 2017 denotes the day that the Chicago Board Options Exchange (CBOE) launched bitcoin futures products, which is December 10th. The entire sample period for this analysis is from January 26, 2015 through October 23, 2020 and the frequency of the data is daily (and includes weekend data)

2.1 Bitcoin prices

Bitcoin's explosive price volatility, relative to other asset classes, is arguably one of its major attractions for investors and speculators alike. It is also one of the reasons regulators caution that it may never serve as a medium of exchange, a unit of account, or a store of value (Lo and Wang 2014). Figure 1 shows a time series plot of bitcoin prices, as well as our measure for investor sentiment, which we discuss in sub-Sect. 2.3, over our sample period (January 26, 2015 through October 23, 2020).

For this particular sample period, bitcoin prices reached a peak of about \$19,533 (in USD) on December 17, 2017. As mentioned, the CBOE and CME introduced bitcoin futures, respectively, on December 10th and 18th of 2017.

The introduction of futures markets may be a contributing factor that led to sharp declines in bitcoin's price. By December 22, 2017 (just four days after the CME introduced futures), bitcoin's price fell to about \$13,893 (in USD) – an arithmetic price change of almost -30% from its December 17 high! From there, it declined precipitously before bottoming out on mid-December of 2018. In the words Hale et al. (2018), "...[the] one-sided speculative demand came to an end when the futures for bitcoin started trading...with the introduction of bitcoin futures, pessimists could bet on a bitcoin price decline, buying and selling contracts with a lower delivery price in the future than the spot price..." (p. 2).

2.2 Summary statistics and bitcoin risk-return metrics

Table 1 provides a glimpse into bitcoin's price behavior and its risk-return characteristics before and after the introduction of bitcoin futures on December 10, 2017. A quick inspection of the sampled minimum (min.) and maximum (max.) returns show the explosive price appreciations and declines that bitcoin has experienced. For the entire sample, the minimum daily return was about -22%, which occurred during the post-futures period, while

Table 1 Risk-return statistics of bitcoin returns

	Full sample (01/26/2015— 10/23/2020) <i>N</i> =2,098	Pre-futures (01/26/2015— 12/09/2017) <i>N</i> =1,049	Post-futures (12/10/2017— 10/23/2020) <i>N</i> =1,049
Mean	0.1825	0.3829	-0.0176
Med	0.1490	0.2255	0.0415
Min	-22.3710	-14.4627	-22.3710
Max	19.5589	19.5589	16.6754
Std. Dev	3.1230	2.8917	3.3276
Skew	-0.3547	0.0541	-0.5779
Kurt	8.8165	8.7326	8.5267
VaR	-5.9386	-5.2848	-6.5397
Modified VaR	-8.2982	-6.9451	-9.2382
Sharpe	0.0584	0.1324	-0.0053
Modified Sharpe	0.0220	0.0551	-0.0019

This table reports risk-return statistics for bitcoin returns for the full sample period (first column) and two subsample periods (second and third columns). The pre-futures period in the second column represents a time span when bitcoin futures were yet to be introduced. On December 10, 2017, the Chicago Board Options Exchange (CBOE) launched bitcoin futures. Both the pre- and post-futures subsample periods consist of 1,049 daily observations. The mean, median, minimum and maximum returns are all expressed in percentages. The value-at-risk (VaR) model and Sharpe ratio are discussed in footnote (4) while the modified VaR (MVaR) and modified Sharpe ratio are estimated using Eqs. (1) and (2), respectively. Both the VaR and MVaR are also expressed in percentages. The entire sample period for this analysis is from January 26, 2015 through October 23, 2020 and the frequency of the data is daily (and includes weekend data)

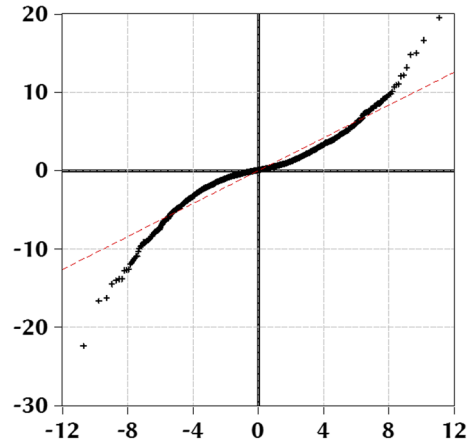
the maximum daily return was about 19%, which occurred during the pre-futures period. Across the entire sample period, as well as the two subsample periods, bitcoin returns do not follow a Gaussian distribution pattern and are instead heavy-tailed.

The lower mean returns, higher standard deviation, and heavier tails of the post-futures sample period indicate relatively heightened risks and an overall deterioration in the risk-return tradeoff. This can be observed from the calculated value-at-risk (VaR) and Sharpe ratios.⁴ To integrate higher moment risks, which can account for crash risk (skewness) and tail risk (kurtosis), we also compute the modified VaR (MVaR) and modified Sharpe ratio (Gregoriou and Gueyie 2003; Signer and Favre 2002). The MVaR can be expressed as follows (using similar notations as the VaR equation in footnote (4)):

$$\text{MVaR} = W \left[\mu - \{z_c + \frac{1}{6}(z_c^2 - 1)S + \frac{1}{24}(z_c^3 - 3z_c)K - \frac{1}{36}(2z_c^3 - 5z_c)S^2\} \sigma \right] \quad (1)$$

⁴ The Sharpe ratio for bitcoin returns, R_r , is computed as $(R_r - r_f)/\sigma$ whereby r_f denotes the daily holding period return of the risk-free rate and σ denotes the standard deviation of bitcoin returns. In our case, we use the one-month treasury, as would be used for conventional assets. Weekend data (to match bitcoin returns' 7-day-a-week frequency) is interpolated using a smoothing moving average approach that best fits the weekday data. The VaR is computed as follows: $\text{VaR} = W(\mu\Delta t - n\sigma\sqrt{\Delta t})$, whereby μ is the mean return for bitcoin; W is the value of the portfolio invested; n is the number of standard deviations depending on the confidence level; σ is the standard deviation of bitcoin returns; Δt is the time window.

Fig 2 Quantile–Quantile plot of bitcoin returns. This figure shows a quantile–quantile (Q–Q) plot for bitcoin returns. The horizontal axis (x-axis) denotes the quantiles theoretically observable in a normal distribution while the vertical axis (y-axis) denotes the quantiles of bitcoin returns. The entire sample period for this analysis is from January 26, 2015 through October 23, 2020 and the frequency of the data is daily (and includes weekend data)



whereby W is the value of the invested portfolio; z_c is the critical value for the probability $(1 - \alpha)$ and is -1.96 for a 95% probability; μ is the mean return; σ is the standard deviation of returns; S is skewness of returns and computed as $S = (1/T) \sum_{t=1}^T \left([R_t - \bar{R}] / \sigma \right)^3$; K is excess kurtosis of returns and computed as $K = (1/T) \sum_{t=1}^T \left([R_t - \bar{R}] / \sigma \right)^4 - 3$. The modified Sharpe ratio, which measures returns per unit of volatility and higher moment risks, can thus be expressed as,

$$\text{ModifiedSharpeRatio} = (R_t - r_f) / \text{MVaR} \quad (2)$$

whereby r_f serves as the risk-free rate (see footnote (4)).

Comparing between the pre- and post-futures sub-sample periods reveals distinct regimes in bitcoin's risk-return characteristics. As mentioned, bitcoin futures provided investors and speculators a vehicle in which they can bet on (or hedge against) bitcoin price declines. These distinct price regimes provide motivation for our robust quantile regression technique that we discuss in Sect. 3, which allows us to get a clear picture of bitcoin's entire return distribution. Figure 2 is a quantile–quantile plot of bitcoin returns for the entire sample period and also shows how bitcoin returns depart from a theoretical Gaussian distribution. Such departures can render instability in parameter estimates from traditional regression techniques which tend to focus on the center of variables' distributions.

2.3 Measuring investor sentiment

Coinbase is an electronic market where participants can submit buy and sell orders for bitcoin. Such orders may be executed immediately (if they are market orders) or they may be limit orders (with a limit price that may not be in close proximity to the prevailing market price) that are potentially executed at a later point in time. Taken together, these orders constitute Coinbase's *order book* and consist of outstanding orders that are awaiting

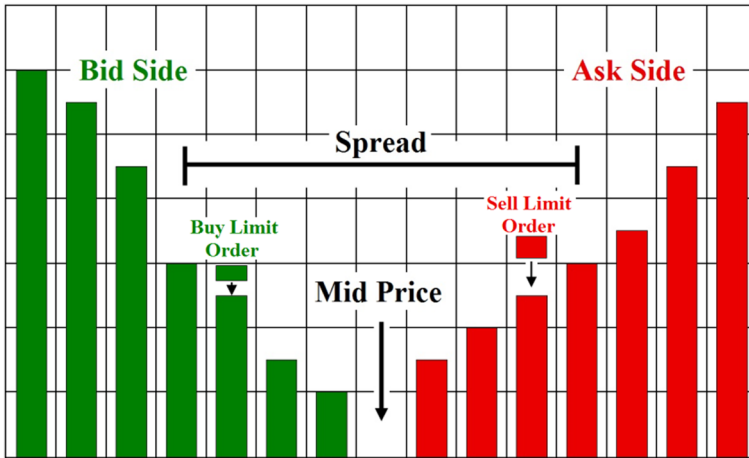


Fig 3 Example of an order book. This figure illustrates a toy example of an order book (*Source: Author's depiction*). Similar toy examples of order books for traditional asset classes are shown elsewhere in market microstructure and asset pricing literature (e.g., Bonart and Gould 2017). The green and red bars denote buy- and sell-side limit orders, respectively. The x-axis denotes possible prices and the y-axis corresponds to market depth. While differences in the bid and ask prices result in a bid-ask spread, differences in the quantity of orders on the buy- and sell-side result in an order flow imbalance

eventual execution or cancellation by the participant, subject to the rules of the exchange.⁵ The flow of such orders across time intervals is referred to as *order flow*.

Figure 3 shows a toy example of an order book that shows bid- and ask-side orders in queues at various price levels around the mid price. Fluctuations in these orders across time, and any resultant imbalances between bid- and ask-side orders, is referred to as *order flow imbalance*. It is this very imbalance that serves as a measure for investor sentiment in our study.

In discussing the incremental information content of order flow, relative to what can be gleaned by observing only trade volumes, Chordia et al. (2002) argue "...volume alone is absolutely guaranteed to conceal some important aspects of trading...consider, for example, a reported volume of one million shares...at one extreme, this might be a million shares sold...at the other extreme it could be a million shares purchased..." (p. 112). As others have noted, order flow and order flow imbalances can signal the possession of private information in the market and, when amplified, can impact liquidity conditions and drive price movements (Chelley-Steeley et al., 2019).

⁵ Coinbase's market trading rules can be found here: https://www.coinbase.com/legal/trading_rules. Care is taken in this study to exclude so-called "spoof trades," which are large non-bona fide orders used to create false appearances in the market for the purposes of manipulating prices. Such spoof trades may appear on the order book for a few minutes and are then subsequently cancelled by the spoof trader. Another characteristic they have is that their limit price can deviate substantially from the current market price. While such ostensible trades are excluded, there is no guarantee that all the trades remaining in our sample were bona fide orders. A possibly fruitful area for future research is to devise formal methods to detect such manipulative trading schemes in cryptocurrency markets.

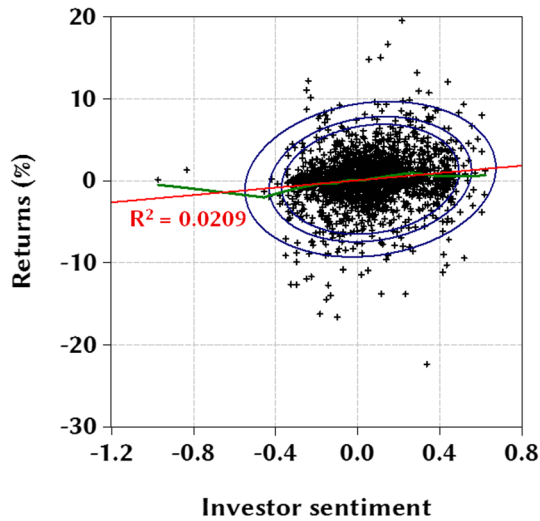


Fig 4 Scatter plot of bitcoin returns and investor sentiment. This figure shows a scatter plot of bitcoin returns in percentages (y-axis) against investor sentiment (x-axis), both of which are collected from the Coinbase digital currency exchange. While the red line is a slope estimate using ordinary least squares, which has an R^2 of 0.0209, the green line is a K-nearest neighbor fit line. The three ellipses in blue are confidence ellipses at the 0.90, 0.95 and 0.99 confidence levels, respectively. The entire sample period for this analysis is from January 26, 2015 through October 23, 2020 and the frequency of the data is daily (and includes weekend data)

In light of the aforementioned, and, using intraday bid- and ask-side orders from Coinbase across the various time intervals, we model daily order flow imbalance. This serves as our proxy for investor sentiment and is calculated as follows⁶:

$$SENT_t = \sum_1^N (Buys_t - Sells_t) / \sum_1^N (Buys_t + Sells_t) \quad (3)$$

whereby, respectively, *Buys* and *Sells* are bid- and ask-side market and limit orders (in dollar volumes) on day t , calculated from trades 1 through N on trading day t . A given trading day's *SENT* indicates whether, on aggregate, investors are optimistic (net buyers whereby $SENT > 0$) or pessimistic (net sellers whereby $SENT < 0$). This ratio serves as a direct and explicit measure of investor sentiment and it is the primary objective of this study to use this measure and quantify its linkages with bitcoin returns.

Figure 4 shows a scatter plot between our measure for investor sentiment (in Eq. 3) and bitcoin returns, on the x-axis and y-axis, respectively. The estimated R^2 from a least squares regression is approximately 0.02 while the fitted confidence ellipses at the 0.90, 0.95 and 0.99 levels, respectively, do not show signs of strong dependence patterns. However, a K-nearest neighbor fit line shows sharp nonlinearities in the sentiment-return relation—a feature in the data that cannot be robustly identified or quantified using conventional regression tools. In addition, and as shown in the scatter plot, there are a multitude of outliers, indicative of regime shifts that take place within bitcoin prices across time.

⁶ Coinbase APIs are available for market-related data, such as bid- and ask-side order book data. More information can be found here: <https://developers.coinbase.com/docs/wallet/guides/price-data>.

Table 2 Description of variables

Data source	Variable	Abbreviation	Transf	Description
Coinbase	Investor sentiment	<i>SENT</i>	lv	<i>SENT</i> serves to measure investor sentiment and is estimated as follows: $\sum_{i=1}^N (Buy_{i,t} - Sell_{i,t}) / \sum_{i=1}^N (Buy_{i,t} + Sell_{i,t})$, whereby <i>Buy_{i,t}</i> and <i>Sell_{i,t}</i> denote buyer- and seller-initiated orders, respectively. These orders are aggregated across each trading day (including on weekends). Periods of rising (declining) sentiment are associated with a positive (negative) value for this ratio
	Range volatility	<i>RV</i>	lv	<i>RV</i> estimates volatility risk and is computed as follows: $(\ln H_t - \ln L_t)^2 / (4 \ln 2)$, whereby <i>H</i> and <i>L</i> denote the intraday high and low price over day <i>t</i>
	Trade volume	<i>VOLM</i>	ln	<i>VOLM</i> is the total value of trading volume (in USD) over a given trading day. This variable serves as a proxy for information flow into the market
	Trades per minute	<i>TPM</i>	ln	<i>TPM</i> is the average number of trades per minute over a given trading day. This variable serves as a proxy for the intensity of information flow
Blockchain	Active addresses	<i>ADD</i>	$\Delta \ln$	<i>ADD</i> is the number of active bitcoin addresses. A bitcoin address is a unique identifier that serves as a destination for a bitcoin payment. Active addresses serve as a proxy for network value and activity
	Block count	<i>BCONT</i>	ln	<i>BCONT</i> is the number of blocks discovered (or “mined”) by miners. These blocks record data pertaining to bitcoin transactions. They serve as pieces of the whole ledger and, when added to the blockchain, become an immutable record within the whole ledger, which is comprised of an ever-growing number of these blocks. Blocks serve as a proxy for network value and activity
	Block size (bytes)	<i>BSIZE</i>	ln	<i>BSIZE</i> is the sum of the size (in bytes) of all the blocks created. Block sizes serve as a proxy for network value and activity
	Mining fees	<i>FEF</i>	$\Delta \ln$	<i>FEF</i> is the median fee per transaction (in USD) that miners earn when validating transactions and discovering new blocks. Mining fees serve as a proxy for network value and activity
	Hash rate	<i>HASH</i>	$\Delta \ln$	<i>HASH</i> is the hash rate (in giga hashes per second) which the Bitcoin network is performing at in order to stay operational. Hash rate is a technological unit of measurement as well as security metric. Greater hashing power reflects greater resistance to attack and enhanced security. This hashing power can be estimated from the number of blocks mined during the last 24 h (1 trading day), as well as the level of difficulty, as follows: $HASH = \lfloor 2^{32} \rfloor D/T$, whereby <i>D</i> denotes difficulty and <i>T</i> denotes the average time between mined blocks

Table 2 (continued)

Data source	Variable	Abbreviation	Transf	Description
	Issuance count	<i>ISSU</i>	$\Delta \ln$	<i>ISSU</i> is the sum of new bitcoins (in BTC) issued and measures the supply of bitcoin in circulation. Bitcoins are issued, or, “minted,” when a miner discovers a new block. Bitcoin’s total supply is set to converge to 21 million units. The number of newly issued bitcoins began at 50 and is set to decrease geometrically (halving takes place every 210,000 blocks, or, every 4 years approximately). The cumulative supply can thus be expressed as $\left\{ \sum_{i=0}^{32} 210,000 [50 * 10^8 / 2^i] \right\} / 10^8$

This table provides an explanation of all the variables used in this study. These variables are observed across a daily interval and, while *SENT*, *RV*, *VOLM* and *TPM* are collected from the Coinbase exchange, *ADD*, *BCONT*, *BSIZE*, *FEE*, *HASH* and *ISSU* are collected from bitcoin’s blockchain. While the first, second and third columns respectively list the data source, variable and abbreviation, the fourth column shows the necessary transformation applied in order to induce stationarity; *ln* denotes the natural logarithm, $\Delta \ln$ denotes the first difference of the natural logarithm, and *lv* denotes the level of the time series. Finally, the fifth column provides a brief description and some background for each of the variables. The data herein collected for this analysis is for the sample period from January 26, 2015 through October 23, 2020 (and includes weekend data)

2.4 Exchange-specific and blockchain-wide variables

To quantify the sentiment-return relation, and to obtain robust coefficient estimates, our study integrates a variety of exchange-specific and blockchain-wide variables as control variables, all of which are also summarized in Table 2. In addition, Table 2 identifies what transformations, if any, are applied to induce stationarity in their respective time series. All these variables, including our investor sentiment measure in Eq. (3), are observed, or calculated, across each trading day t .

The exchange-specific variables, obtained through Coinbase (see footnote (6)), reflect investor activity and market conditions, and are as follows:

1. *RV* Range volatility, estimated as $(\ln H_t - \ln L_t)^2 / (4 \ln 2)$, where H and L denote the intraday high and low price over day t . This is a measure of volatility risk and, in asset pricing literature, impacts risk aversion and liquidity conditions. In empirical tests of the risk-return relation in conventional assets, such as stock portfolios or currencies, there tends to be a negative relation. This phenomena is often referred to as the so-called 'volatility feedback' effect (Carr and Wu 2017).

2. *VOLM* Trade volume is the total value of all trading (in USD) over a given day t . This variable measures the flow of information to the market, and has been shown to impact asset prices and volatilities (Andersen 1996).

3. *TPM* Trades per minute is the average number of trades (on a per minute basis) over a given day t . This measure can potentially capture intensity in the information flow (Eisler and Kertesz 2006).

The blockchain-wide variables, which measure network activity and the health of the overall blockchain for bitcoin, are as follows⁷:

4. *ADD* Addresses is the number of active bitcoin addresses over day t . A bitcoin address is a unique identifier that serves as a destination for a bitcoin payment—in much the same way as an email address is needed to send or receive email messages. Active addresses serve as a proxy for network value and activity. As mentioned earlier, there is a growing body of literature on network theory that seeks to delineate the intrinsic value of bitcoin as a function of actual users and their interactions together.

5. *BCONT* Block count is the number of blocks discovered (or “mined”) by miners on day t . These blocks record data relating to bitcoin transfers or transactions among all the users. They serve as pieces of the whole ledger and, when confirmed and added to the blockchain, become an immutable record within the entire ledger. Blocks serve as a proxy for network value and activity.

6. *BSIZE* Block size is the sum of the size (measured in bytes) of all the blocks created on day t . Block sizes serve as a proxy for network value and activity.

7. *FEE* Mining fees is the median fee per transaction (in USD) that miners earn when validating transactions and discovering new blocks that are added to the blockchain. Mining fees serve as a proxy for network value and activity. The median fee on day t is observed rather than the mean fee (although in untabulated results, available upon request, our findings are robust to this choice).

⁷ One of the philosophies underlying bitcoin is that it is meant to be community-driven and transparent. Thus, there is a wealth of resources for gauging movements in blockchain variables and for obtaining time series data for forecasting or other empirical purposes. This link contains APIs for some of the variables used herein: https://www.blockchain.com/api/blockchain_api.

8. *HASH* Hash rate represents the actual hash rate (in giga hashes per second) which the Bitcoin network is performing at in order to stay operational. This is a technological unit of measurement as well as security metric. Greater hashing power reflects greater resistance to attack and enhanced security. On a given day t , the hash rate can be estimated from the number of blocks mined (in the last 24 h, or, 1 trading day), as well as the level of difficulty. Thus, it can be expressed as follows: $HASH = [2^{32}] [D/T]$, whereby D denotes network difficulty and T denotes the average time between mined blocks. Network difficulty, D , in technical terms, measures how difficult it is to find a hash below some target, and is adjusted approximately every 2 weeks (or 2,016 blocks).

9. *ISSU* Issuance count is the sum of new bitcoins (in BTC) issued on a given day t . This is a direct measure of bitcoin's money supply that is in circulation, and is analogous to the monetary base measures of 'M1' or 'M2' that are used by the Federal Reserve for US dollars. Unlike US dollars, however, bitcoins are issued, or, "minted," when a miner discovers a new block. Bitcoin's total supply is set to converge to 21 million units (government authorities behind fiat currencies, such as the US government and the dollar, do not declare a constrained supply). The number of newly issued bitcoins began at 50 and is set to decrease geometrically (halving takes place every 210,000 blocks, or, every 4 years approximately). The cumulative supply can thus be expressed as $\left\{ \sum_{i=0}^{32} 210,000 [50 * 10^8 / 2^i] \right\} / 10^8$.

2.5 Cryptocurrency market-wide Momentum

To further substantiate the robustness of our sentiment-return relation, we construct a cryptocurrency market index (*CRYIX*) in order to capture information content, investor behavior, and momentum effects arising from cryptocurrency markets on aggregate and not necessarily idiosyncratic to only bitcoin exchanges. To construct *CRYIX*, we use nine of the major cryptocurrencies that are presently in circulation. Our index is market-weighted whereby the weight of each of the respective nine constituents is determined by dividing its market capitalization by the total market capitalization of the nine cryptocurrencies. The construction of this market-weighted index is similar to, say, that of the S&P 500 index, and serves to capture aggregate market conditions. The returns for our *CRYIX* index, $R_{m,t}$, are thus constructed as

$$R_{m,t} = \sum_{i=1}^N R_{i,t} \times (MktCap_{i,t} / TotalMktCap_t) \quad (4)$$

whereby $R_{i,t}$ and $MktCap_{i,t}$ is the return and market capitalization of cryptocurrency i , respectively, while $TotalMktCap_t$ denotes the total market capitalization of all nine ($N = 9$) cryptocurrencies. The *CRYIX* index used here is constructed for each trading day t in light of the daily changes in market capitalizations of all the nine cryptocurrencies.

Selection of these nine cryptocurrencies was performed while satisfying a delicate balance between, firstly, choosing major cryptocurrencies that have some of the largest present-day market capitalizations, since they are the ones that predominantly receive attention in the financial press, and, secondly, ensuring that those selected cryptocurrencies have a sample range that availablely corresponds as close as possible with that in our present sample. This is worth noting since there are several emergent cryptocurrencies that, although may be very large in terms of current market capitalization, have only been in existence and circulation for a rather short period of time.

In light of the aforementioned, the nine cryptocurrencies used for our *CRYIX* index are as follows, respectively, along with their abbreviations: bitcoin (BTC), XRP (XRP), ethereum (ETH), stellar (XLM), litecoin (LTC), dash (DASH), monero (XMR), NEM

(XEM), and dogecoin (DOGE). In the appendix, we discuss more the various intricacies of these cryptocurrencies, including summary statistics of their price movements and average market capitalizations (in Table 5), summary statistics of their respective contributions to the *CRYIX* in terms of market capitalization (in Table six), and pairwise correlations between the nine *CRYIX* constituent cryptocurrencies (in Table 7).

Tables 5, 6 and 7 reveal that bitcoin remains the dominant cryptocurrency in terms of market capitalization among all the nine constituents, with an average contribution to the *CRYIX* of approximately 67% over the sample range. Presently, and at the time of this study, even while considering the nearly 10,000 individual cryptocurrencies in circulation, bitcoin constitutes approximately 50% of the total market capitalization of all cryptocurrencies. Thus, our *CRYIX* index is fairly representative of conditions and investor expectations in aggregate cryptocurrency markets. The average pairwise correlation among all the nine cryptocurrency constituents (shown in Table 7) is 0.4592. The lowest pairwise correlation is shown to be between Dash and XRP (at 0.3385) while the highest pairwise correlation is shown to be between bitcoin and litecoin (at 0.6697).

As is discussed in Sect. 4, *CRYIX* is used as an additional control variable that captures aggregate market momentum and expectations among cryptocurrency traders. Its inclusion as a control variable will test whether our sentiment measure, as well as our estimated sentiment-return relation, still retains explanatory power across the distribution of bitcoin price changes.

3 Robust approach to quantile modeling

This study seeks to map the linkages between investor sentiment, as defined in Eq. (3), and bitcoin price changes across their entire distribution. This serves the purpose of giving a more complete, accurate, and robust picture of the sentiment-return relation since, as shown in Sect. 2, bitcoin price changes are non-Gaussian in nature with thick tails and pronounced degrees of skewness. Because our sample range encapsulates at least two distinct regimes in bitcoin's price behaviors (pre- and post-futures), there are compelling reasons why a full description of the sentiment-return relation across the entire return distribution is fundamental to our understanding of what drives bitcoin prices.

In addition to using the bootstrapped quantile regression procedure to ensure robustness in coefficient estimates of the sentiment-return relation, this study employs a range of exchange-specific and blockchain-wide control variables, respectively. While quantifying, and, ensuring robustness in the sentiment-return relation is an important objective of this study, a secondary contribution that emerges is we identify heterogeneous linkages between these variables, which serve to capture conditions in the Coinbase exchange and blockchain network, respectively, with bitcoin returns, across the distribution. In other words, the explanatory power of these variables (and in some instances their fundamental signs, as we discuss later on) essentially varies in terms of location shifts, variance, and skewness.

In light of the aforementioned, we can begin by expressing the conditional quantile model as follows (Buchinsky 1994, 1995; Koenker and Bassett 1978):

$$y_i = x_i' \beta_\theta + u_{\theta i}, \text{with} \text{Quant}_\theta(y_i | x_i) = x_i' \beta_\theta, i = 1, \dots, n \quad (5)$$

whereby β_θ and x_i are $K \times 1$ vectors and $x_{i1} \equiv 1$. $\text{Quant}_\theta(y|x)$ represents the θ th quantile of y given x . The error term, $u_\theta \equiv y - x' \beta_\theta$ is assumed to be a continuously differentiable

cumulative distribution function, $F_{u_\theta}(\cdot|x)$, and density function $f_{u_\theta}(\cdot|x)$. From this, it follows that $\text{Quant}(u_\theta|x) = 0$ and $f_{u_\theta}(0|x) > 0$. An estimator for β_θ is obtainable as follows:

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_{\theta}(y_i - x'_i \beta) \tag{6}$$

and whereby the check function can be expressed as $\rho_{\theta}(\lambda) = (\theta - I(\lambda < 0))\lambda$ and $I(A)$ is the indicator function. As Koenker and Bassett (1978) show, Eq. (6) is a solvable linear programming problem.

As Powell (1986) demonstrates, Eq. (6) methodologically fits into the generalized method of moments (GMM) model for censored quantile regression frameworks and, under the conditions of Huber (1967), it can be shown that

$$\sqrt{N}(\hat{\beta}_{\theta} - \beta_{\theta}) \xrightarrow{L} \mathcal{N}(0, \Lambda_{\theta}) \tag{7}$$

whereby.

$$\Lambda_{\theta} = \theta(1 - \theta)(E[F_{u_{\theta}}(0|x)xx'])^{-1} E[xx'](E[F_{u_{\theta}}(0|x)xx'])^{-1} \tag{8}$$

If we have $F_{u_{\theta}}(0|x) = F_{u_{\theta}}(0)$ (i.e., the density of u_{θ} at 0 is independent of x), then Λ_{θ} can be simplified to as follows:

$$\Lambda_{\theta} = \left(\theta(1 - \theta)/F_{u_{\theta}}^2(0)\right) (E[xx'])^{-1} = \sigma_{\theta}^2 (E[xx'])^{-1} \tag{9}$$

and where.

$$\sigma_{\theta}^2 = \left(\theta(1 - \theta)/F_{u_{\theta}}^2(0)\right) \tag{10}$$

Consistency and asymptotic normality is already well-established in the literature for quantile regression modeling (Koenker 2005). Following Buchinsky (1994, 1995) and Pedersen (2015), among others, our study extracts standard errors through an (x, y) -bootstrapping procedure. This method consists of drawing B samples of (x_p, y_p) pairs each of size m and with replacement from the $N - 1$ pairs of the original sample pool, each with an equivalent probability. This method can enhance robustness since, first, it does not require identically distributed error terms and, second, can account for heteroskedasticity in a time series. As shown in Fig. 2, bitcoin price changes are heteroskedastic, exhibiting sharp regime changes at various points in time.

Thus, and in the case where $f_{u_{\theta}}(0|x) \neq f_{u_{\theta}}(0)$, the (x, y) -bootstrapping method is employed and, for each of the B samples, an estimator for $\hat{\beta}_{\theta}$ is computed in order to capture B bootstrap estimates, $\hat{\beta}_{\theta}^1, \dots, \hat{\beta}_{\theta}^B$. Therefore, an estimate for Λ_{θ} can now be expressed as.

$$\hat{\Lambda}_{\theta} = N(m/N)(1/B) \sum_{b=1}^B (\hat{\beta}_{\theta}^b - \hat{\beta}_{\theta})(\hat{\beta}_{\theta}^b - \hat{\beta}_{\theta})' \tag{11}$$

Buchinsky (1995) conducts an extensive Monte Carlo study and finds that this (x, y) -bootstrapping method can enhance robustness of coefficient estimates. In light of this, and consistent with Pedersen (2015), the size of the bootstrap samples used in this study will be equal to the original sample size.

Table 3 Quantifying the sentiment-return relation

Quantile	Center of distribution										Right tail	
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90			
<i>Panel A: Univariable modeling (no controls)</i>												
<i>SENT</i>	2.5476** (3.157)	1.5868** (3.346)	1.0209** (3.544)	0.7664** (4.290)	0.9821** (4.944)	1.3893** (5.571)	2.2434** (5.574)	3.3328** (6.447)	3.2586** (4.582)			
<i>Panel B: Multivariable modeling (exchange-specific controls)</i>												
<i>SENT</i>	1.6235** (2.850)	1.3735** (4.588)	1.3442** (3.097)	1.1713** (5.137)	1.0545** (3.841)	1.1051** (3.468)	1.1433** (3.377)	1.8260** (3.587)	1.4277** (3.420)			
<i>RV</i>	-7.6676** (-8.528)	-3.3905** (-6.513)	-3.2265** (-5.536)	-2.7997** (-6.511)	-2.1635** (-4.225)	-0.8861 (-0.880)	2.3371* (1.914)	3.8275** (6.029)	5.0618** (3.755)			
<i>VOLM</i>	-0.2566** (-1.976)	-0.2442** (-2.211)	-0.1300** (-2.210)	-0.0498 (-0.788)	0.0436 (0.657)	0.0632 (0.717)	0.0792 (0.862)	-0.0039 (-0.028)	-0.1552 (-0.943)			
<i>TPM</i>	-0.0106 (-0.042)	0.0965 (0.427)	-0.0172 (-0.140)	-0.0351 (-0.298)	-0.0907 (-0.691)	0.0041 (0.028)	0.1480 (0.876)	0.4522* (1.621)	1.0593** (3.058)			
<i>Panel C: Multivariable modeling (exchange - specific and blockchain - wide controls)</i>												
<i>SENT</i>	1.7829** (3.456)	1.6283** (5.483)	1.2279** (5.748)	1.3784** (5.721)	1.3225** (4.613)	1.3603** (4.256)	1.3534** (3.712)	1.6582** (3.367)	1.0345** (2.521)			
<i>RV</i>	-6.8269** (-7.425)	-5.0302** (-5.749)	-3.1676** (-6.756)	-2.8694** (-6.194)	-2.1392** (-3.862)	-0.9693 (-1.014)	1.9291* (1.724)	3.5652** (4.665)	4.3399** (3.969)			
<i>VOLM</i>	-0.4835** (-3.158)	-0.3936** (-3.936)	-0.2193** (-3.473)	-0.1158 (-1.442)	0.0016 (0.019)	0.0662 (0.766)	0.0962 (0.915)	0.0923 (0.658)	0.0973 (0.636)			
<i>TPM</i>	-0.0084 (-0.031)	0.0098 (0.052)	-0.1131 (-0.883)	-0.0454 (-0.354)	-0.0521 (-0.371)	0.0806 (0.556)	0.2376 (1.329)	0.3891 (1.433)	0.8879** (2.997)			
<i>ADD</i>	0.0480 (0.082)	0.0565 (0.108)	-0.4100 (-0.880)	0.0870 (0.223)	-0.1106 (-0.225)	0.6853 (1.211)	0.4415 (0.730)	1.2497* (1.802)	1.2270 (1.418)			
<i>BCONT</i>	0.6186 (0.623)	0.3738 (0.613)	0.8205 (1.896)	0.9245** (1.979)	1.2824** (2.038)	1.8047** (2.612)	2.7471** (3.385)	3.8558** (3.599)	4.9560** (5.706)			
<i>BSIZE</i>	1.6935** (4.869)	1.3540** (4.228)	1.0817** (5.293)	0.4798** (2.368)	0.1215 (0.641)	-0.1622 (-0.813)	-0.3825* (-1.689)	-0.7170** (-2.404)	-1.2162** (-3.234)			
<i>FEE</i>	0.4290 (1.260)	0.2561 (0.896)	0.8921** (3.439)	0.9697** (4.483)	1.1638** (4.097)	1.5978** (4.472)	1.9260** (5.537)	1.7623** (4.405)	1.8404** (4.488)			

Table 3 (continued)

Quantile	Center of distribution									Right tail
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	
<i>HASH</i>	1.6043 (0.556)	0.5478 (0.247)	-0.3450 (-0.216)	-0.0151 (-0.010)	0.5551 (0.330)	-1.0779 (-0.466)	-2.4144 (-0.768)	-3.5991 (-0.732)	0.4547 (0.074)	
<i>ISSU</i>	-2.6403 (-1.003)	-1.3976 (-0.670)	-0.3088 (-0.203)	-0.3488 (-0.243)	-0.8543 (-0.540)	0.3810 (0.168)	1.7805 (0.564)	2.6181 (0.529)	-0.8208 (-0.133)	

This table reports slope coefficients for each quantile $\theta = \{0.10, 0.20, \dots, 0.80, 0.90\}$ for panels A, B and C, respectively. While panel A is a univariable model, with only investor sentiment as a regressor, panels B and C are multivariable models with control variables; panel B incorporates exchange-specific control variables while panel C incorporates exchange-specific and blockchain-wide control variables. For all panels, the constant is suppressed for brevity. Coefficient estimates for *RV* are scaled by a factor of $(\times 10^7)$. Coefficient estimates for *ADD*, *FEE*, *HASH* and *ISSU*, respectively, are scaled by a factor of $(\times 10^{-2})$. This scaling is solely for illustrative purposes given the disparate units of measurement for each of the variables (see Table 2 for a discussion of each of the variables). *T*-statistics are denoted in parentheses and are estimated using heteroskedasticity-consistent and bootstrapped standard errors. Equations (5) through (11) discuss the bootstrapping procedure. Significance at the 5% and 10% levels is indicated by ** and *, respectively. The entire sample period for this analysis is from January 26, 2015 through October 23, 2020 and the frequency of the data is daily (and includes weekend data)

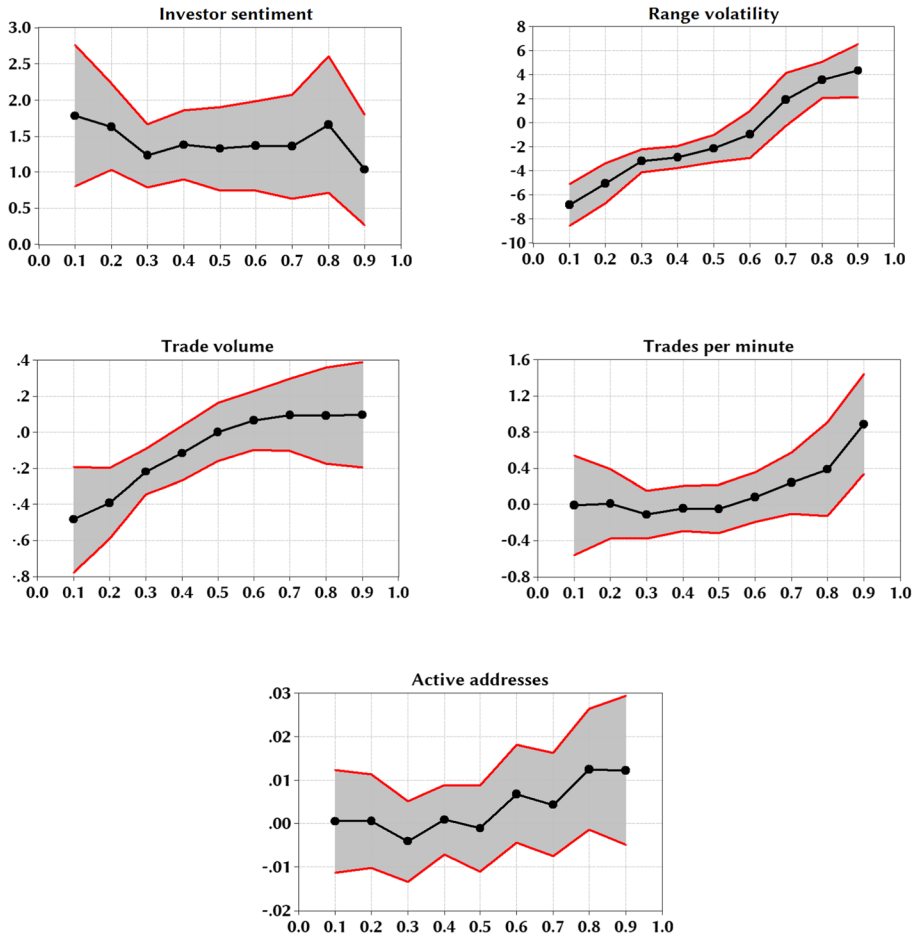


Fig 5 Slope coefficient estimates. This figure shows slope coefficient estimates for the multivariable model which includes exchange-specific and blockchain-wide control variables (panel C of Table 3) for each quantile $\alpha = \{0.10, 0.20, \dots, 0.80, 0.90\}$. The red lines give the upper and lower bounds of a 95% confidence interval size using the bootstrapped standard errors, as discussed in Eqs. (5) through (11)

4 Discussion of findings

Given the unique investor clientele and microstructure underlying the Bitcoin ecosystem, it is important to quantify the sentiment-return relation and assess whether it contrasts with the types of relations we see in conventional asset classes. As mentioned, our measure for investor sentiment in Eq. 3 is direct and explicit, and is derived from actual buying and selling behavior. To ensure robustness in our sentiment-return relation estimations, and to control for the aforementioned factors discussed in sub-Sects. 2.4 and 2.5, respectively, we show our findings in Table 3 across three panels. In panel A, we show the sentiment-return relation estimated using a univariable approach (in other words, there are no control variables and only *SENT* serves as the regressor). In panel B we employ a multivariable approach where we include exchange-specific controls (*RV*, *VOLM* and *TPM*, respectively). Finally, in panel C, we use a multivariable approach where we include both

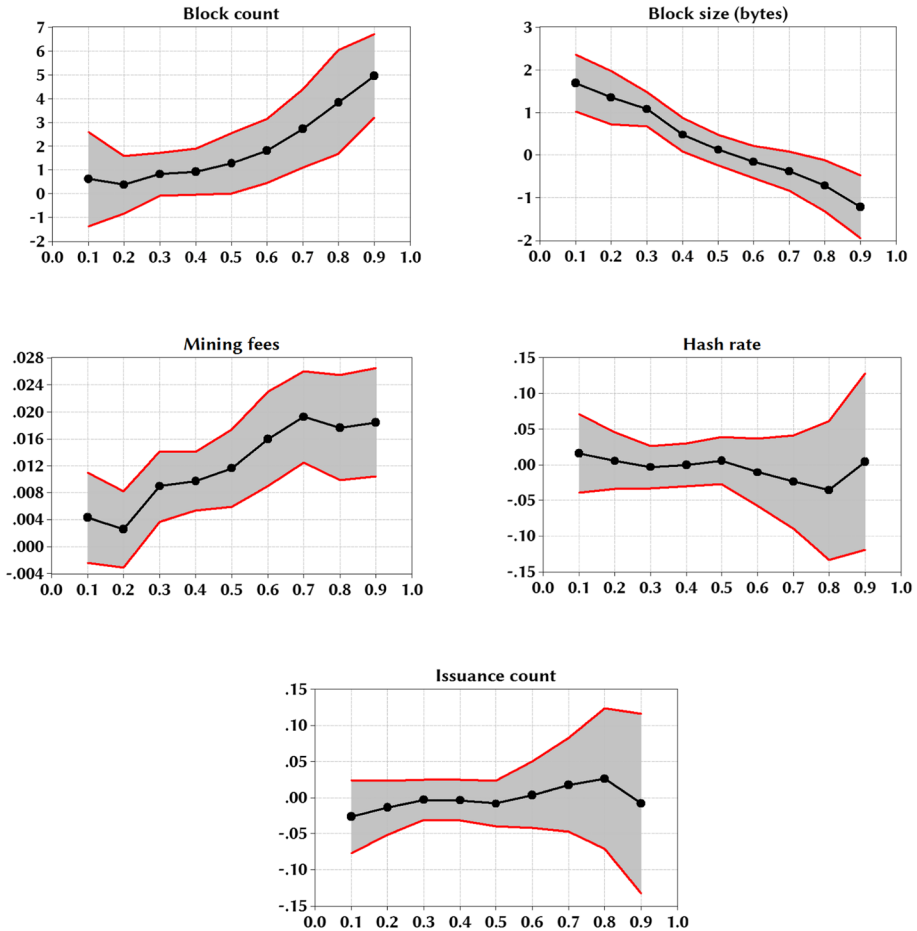


Fig 5 (continued)

exchange-specific and blockchain-wide controls (*RV*, *VOLM*, *TPM*, *ADD*, *BCONT*, *BSIZE*, *FEE*, *HASH*, and *ISSU*, respectively).⁸

⁸ The exchange-specific and blockchain-wide variables are integrated into the full multivariable model in panel C as a way to control for other variables which fundamentally reflect conditions in Bitcoin’s ecosystem. Incorporating such variables serves at least two purposes; the primary purpose, which is the focus of our study here, is to examine whether our measure for sentiment, *SENT*, maintains its statistical significance and whether we can attain a robust sign for the sentiment-return relation. A secondary purpose, which sets the foundation for future research, is to begin examining the degree to which bitcoin’s blockchain characteristics can help in explaining bitcoin’s seemingly chaotic price movements. For this secondary purpose, a natural question that arises is the degree of possible (multi-)collinearity between these predictor variables. The average pairwise correlation between all the predictors is approximately 0.18 (results not tabulated for brevity but available upon request). While some variables exhibit higher pairwise correlations with one another (such as between *ISSU* and *HASH*) several other predictors do not given the different transformations that are applied (as shown in Table 2). In untabulated results (all of which are available upon request) there is no qualitative shift in our estimated sentiment-return relation shown in Table 3 when (a) either one of the potentially collinear variables is dropped from our regression specification, or, (b) newly formed principal components are used as predictors.

The findings are extracted using the robust quantile regression approach in Eqs. (5) through (11), with the purpose of producing a higher resolution mapping of the linkages between sentiment (as well as the other control variables) and returns across the distribution of bitcoin's price changes. In panel A we show coefficient estimates of the sentiment-return relation using a univariable approach (no control variables). This serves as a baseline model to gauge the nature of the relation and to assess whether it changes qualitatively after we include controls (in panels B and C, respectively). From panel A, the coefficient for *SENT* is consistently positive and significant at the 5% level across all of the quantiles. Let us now discuss the significance and interpretation of this. As mentioned in sub-Sect. 2.3, on any given trading day t , *SENT* indicates whether investors are optimistic (net buyers, whereby $SENT > 0$) or pessimistic (net sellers, whereby $SENT < 0$). In other words, our *SENT* indicator becomes positive when bid-side orders outweigh ask-side orders and our *SENT* indicator becomes negative when ask-side orders outweigh bid-side orders. The positive coefficients across the quantiles show a positive relation, which means that if *SENT* is positive, it indicates that growing optimism is associated with positive price changes. Conversely, if *SENT* is negative, it indicates that growing pessimism is associated with negative price changes. It is important to emphasize that the sign and significance of the coefficients remain robust across bear periods (lower and left-tail quantiles) and bull periods (higher and right-tail quantiles). This finding strongly suggests that investor sentiment is an important determinant that drives bitcoin price behaviors, much in the same way as other studies show how sentiment can impact the prices of traditional asset classes, such as bonds and equities.

When integrating exchange-specific controls in panel B, we see that the significance of our coefficients for *SENT* remain qualitatively robust across quantiles. This continues to hold when we also include blockchain-wide controls, as shown in panel C. Namely, across all quantiles, we document a significant sentiment-return relation that is not diluted by the inclusion of important explanatory factors.

Inspection of the full multivariable model with all the controls (panel C), as well as the slope coefficients from this model, shown in Fig. 5, yield many interesting insights. First, there appears to be a volatility feedback effect, which posits a negative risk-return relation, in the left-tail region of bitcoin price changes (specifically, in the 0.10 through 0.50 quantiles) while, in the extreme right-tail region (the 0.80 and 0.90 quantiles), there is a positive and significant relation at the 5% level. The remainder of the quantiles show a weak or insignificant relation. This is interesting to note since the volatility feedback effect is a stylized fact in the return series of traditional asset classes, such as equities or currencies.

Second, there is partial evidence in support of a liquidity premium hypothesis, in the spirit of Amihud and Mendelson (1986), but possibly only in the left-tail quantiles (0.10, 0.20 and 0.30). One interpretation of this hypothesis, in broad terms, argues that rising illiquidity (low trading activity) should increase expected returns. We see in the aforementioned left-tail quantiles a negative relation between trade volume, which can proxy for (il-)liquidity, and bitcoin returns. However, we do not see such a relation in the other quantiles. Despite these very preliminary findings, and while not the focal point of our present study, more research into whether there exists a liquidity premium in cryptocurrency markets would be of interest since others also show that asset size, or, the hedging demand of investors can play a role in whether we see a liquidity premium and what its sign is (Llorente et al. 2002).

Third, Bitcoin microstructure variables, while they portray important characteristics underlying the blockchain, do not provide an infallible explanation as to why bitcoin prices move the way they do. For example, the coefficient for *BFSIZE* shows a sign reversal from

Table 4 Controlling for aggregate cryptocurrency market index (CRYIX) momentum

Quantile	Center of distribution									
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	Right tail
<i>SENT</i>	1.0718** (3.028)	1.0439** (4.210)	1.1229** (4.975)	1.1890** (4.613)	1.1702** (4.649)	0.9105** (3.462)	0.9500** (3.219)	0.7886** (2.476)	0.9352** (2.501)	
<i>CRYIX_{t-1}</i>	0.3604** (15.564)	0.3601** (21.961)	0.3797** (18.426)	0.3896** (16.857)	0.3929** (16.945)	0.4251** (18.824)	0.4446** (18.052)	0.4212** (20.310)	0.4427** (18.764)	
<i>CRYIX_{t-2}</i>	-0.0278** (-1.818)	-0.0201 (-1.283)	-0.0105 (-0.696)	-0.0097 (-0.531)	-0.0193 (-0.921)	-0.0113 (-0.565)	0.0025 (0.121)	-0.0080 (-0.390)	-0.0007 (-0.036)	
<i>CRYIX_{t-3}</i>	-0.0127 (-0.932)	0.0031 (0.249)	0.0118 (0.809)	0.0027 (0.156)	0.0075 (0.413)	-0.0100 (-0.581)	0.0058 (0.274)	0.0130 (0.609)	-0.0053 (-0.207)	
<i>CRYIX_{t-4}</i>	-0.0104 (-0.567)	0.0085 (0.627)	0.0123 (0.808)	0.0123 (0.705)	0.0015 (0.088)	0.0083 (0.475)	0.0155 (0.899)	0.0145 (0.679)	0.0187 (0.726)	
<i>CRYIX_{t-5}</i>	-0.0229 (-1.174)	-0.0252 (-1.554)	-0.0082 (-0.536)	0.0131 (0.727)	0.0079 (0.398)	0.0250 (1.195)	0.0325* (1.658)	0.0309* (1.804)	0.0549** (2.283)	
<i>RV</i>	-5.8648** (-7.778)	-4.3967** (-6.295)	-2.6658** (-5.458)	-1.9388** (-3.749)	-1.4274** (-3.476)	-0.5045 (-0.814)	1.4816 (1.380)	3.5662** (4.931)	5.6911** (6.671)	
<i>VOLM</i>	-0.3796** (-3.356)	-0.2995** (-3.618)	-0.2650** (-3.335)	-0.1120* (-1.693)	-0.0156 (-0.228)	0.0877 (1.228)	0.2003** (2.453)	0.1918** (2.292)	0.3007** (2.560)	
<i>TPM</i>	0.0037 (0.018)	0.0156 (0.109)	-0.0030 (-0.020)	-0.0956 (-0.774)	-0.0223 (-0.174)	0.0175 (0.127)	-0.0312 (-0.227)	0.1259 (0.818)	0.0529 (0.230)	
<i>ADD</i>	-0.5821 (-1.178)	-0.3642 (-0.851)	-0.0814 (-0.178)	0.1397 (0.310)	-0.0163 (-0.042)	0.3757 (0.772)	1.1360** (2.014)	1.0182* (1.904)	1.1293* (1.925)	
<i>BCONT</i>	-0.3725 (-0.524)	0.5730 (1.052)	0.4666 (0.888)	0.7721 (1.594)	1.3775** (2.661)	2.2394** (4.627)	2.3812** (4.160)	2.9183** (3.884)	3.4321** (4.044)	
<i>BSIZE</i>	1.0301** (3.816)	1.0282** (3.276)	0.9032** (4.161)	0.3298 (1.605)	-0.0114 (-0.074)	-0.2570 (-1.451)	-0.6413** (-2.979)	-0.7878** (-3.345)	-1.0087** (-3.702)	
<i>FEE</i>	0.8131** (2.521)	0.4413* (1.923)	0.3983** (1.972)	0.6120** (2.707)	0.9901** (4.396)	0.9801** (4.285)	1.0056** (3.681)	0.9821** (3.799)	1.0883** (3.410)	

Table 4 (continued)

Quantile	Center of distribution									
	Left tail	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	Right tail
<i>HASH</i>	2.5250 (0.762)	1.9035 (1.142)	1.3374 (0.713)	0.1959 (0.098)	0.7787 (0.476)	-0.1495 (-0.115)	-0.2413 (-0.138)	-1.3420 (-0.453)	2.0479 (0.415)	
<i>ISSU</i>	-2.4903 (-0.774)	-2.2172 (-1.345)	-1.5584 (-0.853)	-0.5663 (-0.297)	-1.2255 (-0.794)	-0.6194 (-0.506)	-0.2263 (-0.132)	0.3872 (0.132)	-2.5551 (-0.514)	

This table reports slope coefficients for each quantile $\theta = \{0.10, 0.20, \dots, 0.80, 0.90\}$ for investor sentiment and the exchange-specific and blockchain-wide control variables (as shown in Table 3). However, this table also includes lagged returns for the *CRYX* (up to five trading days) in order to control for exogenous momentum and sentiment across major cryptocurrencies. Equation (4) discusses the composition of the *CRYX* while, in the appendix, Tables 5, 6, and 7, respectively, show summary statistics of the *CRYX* constituents, their contributions to this index with respect to their market capitalizations, and the pairwise correlations between each of the constituents. Coefficient estimates for *RV* are scaled by a factor of $(\times 10^2)$. Coefficient estimates for *ADD*, *FEE*, *HASH* and *ISSU*, respectively, are scaled by a factor of $(\times 10^{-2})$. This scaling is solely for illustrative purposes given the disparate units of measurement for each of the variables (see Table 2 for a discussion of each of the variables). *T*-statistics are denoted in parentheses and are estimated using heteroskedasticity-consistent and bootstrapped standard errors. Equations (5, 6, 7, 8, 9, 10 and 11) discuss the bootstrapping procedure. Significance at the 5% and 10% levels is indicated by ** and *, respectively. The entire sample period for this analysis is from January 26, 2015 through October 23, 2020 and the frequency of the data is daily (and includes weekend data)

the left- to the right-tail quantiles, going from positive to negative. Of all the microstructure variables, *FEE* is most reliably significant across the quantiles. The sign for its coefficient is also positive, which signifies that increases in fees, arguably from a growing demand for miners' services, is linked to positive bitcoin price changes.

In Table 4, we show estimates for our sentiment-return relation when including, in addition to all the exchange-specific and blockchain-wide controls (as shown in Table 3), lagged returns of the *CRYIX* index, estimated using Eq. (4). These lags capture up to the past five trading days and, as mentioned, are designed to incorporate momentum and expectations that impact cryptocurrency markets on aggregate. As is discussed in King and Koutmos (2021), the buying and selling activities of momentum-type, or, sentiment-driven, traders are conditional upon past price changes. This is consistent with many widely-viewed cryptocurrency websites that make buy and sell recommendations that utilize some form of momentum-based strategies based on historical price data. Such websites are publicly available and can amplify momentum-type of trading behaviors.⁹

Inspection of our coefficients in Table 4 reveal the following. First, inclusion of the lagged *CRYIX* does not change the statistical significance or interpretation of our sentiment-return relation. Second, only $CRYIX_{t-1}$ is statistically significant throughout the distribution of bitcoin price changes. It is also positive in its sign, indicating that returns at time $t - 1$ are positively related to returns at time t . This provides some evidence of short-term persistence in bitcoin price changes. This is consistent with the findings of King and Koutmos (2021) who show that past prices can motivate momentum-type trading behaviors. Third, there are no major changes or reversals in signs in the remaining exchange-specific and blockchain-wide variables. One notable difference is the coefficient for *VOLM*, which in Table 3 is negative and significant only in the lower quantiles but in Table 4 is negative and significant in the lower quantiles and positive and significant in the higher quantiles. This finding, although not the focal point of our study, reveals some complexities in the volume-return relation for bitcoin and makes for interesting future research, especially from the perspective of the liquidity premium hypothesis of Amihud and Mendelson (1986).

Taken altogether, these findings provide compelling evidence that bitcoin prices undergo regime shifts across time, making it intractable to extract more precise coefficient estimates and to answer the fundamental question of what drives bitcoin prices. A broad advantage to the quantile regression approach implemented herein, apart from the bootstrapping procedure which is performed to assure robust estimates, is that quantile regressions do not assume a particular distribution for the regressand, unlike conventional regression modeling, and can overcome non-Gaussian features in data sets, such as heteroskedasticity. As shown in our findings, the ability to provide a more complete description of how sentiment (as well as other factors) is linked to returns across the distribution of bitcoin price changes, shows the future challenges cryptocurrency research must consider in terms of model building and model output interpretation.

⁹ An example of such a website is <https://www.tradingview.com/symbols/BTCUSD/ideas/>. This website gives investment ideas based on momentum-type strategies such as moving averages or the identification of "support" and "resistance" levels in bitcoin price movements. Many such websites also give users the flexibility to plot a plethora of momentum and sentiment indicators alongside bitcoin realized prices in order to form their own strategies. Despite the well intentions of such websites and interactive trader tools, they can also serve to contribute to self-reinforcing herding behaviors, as traders seek to mimic one another.

5 Concluding remarks

This study uses trading records of individual bitcoin investors to quantify the linkages between investor sentiment and bitcoin returns. Our measure for investor sentiment is estimated from Coinbase's order book and consists of bid- and ask-side limit and market orders. Using this measure, which is a direct and explicit account of investor sentiment and ensuing trading behavior, we show that rising sentiment is linked to price increases while declining sentiment is linked to price decreases.

To ensure robustness in our findings we, first, perform estimations using a bootstrapped quantile regression approach and, second, employ a host of control variables that capture conditions in the Coinbase market exchange, Bitcoin's overall blockchain, and information content pertaining to major cryptocurrencies other than just bitcoin. We show that our sentiment-return relation remains robust across quantiles and when including controls. This is an important finding since, as other studies argue, bitcoin's price changes frequently undergo regime shifts. As we show herein, after the introduction of futures by the CBOE, bitcoin experienced large declines initially and heightened volatility. As we show, this type of volatile price behavior results in many outlier observations and is difficult to model using conventional approaches which rely on using normally distributed data.

Our findings also reconcile the many observations of other studies, which argue that there is intertemporal instability in coefficient estimates in models seeking to explain bitcoin prices. This instability arises because, first, bitcoin prices undergo regime shifts and, second, conventional regression models tend to focus on the mean or the center of the distribution of bitcoin price changes. We show that such an approach, apart from possibly yielding instability in estimates, sweeps very important information under the rug regarding intervariable relationships during periods of sharp price declines (lower quantiles) versus periods of sharp price increases (higher quantiles). For example, many of the exchange-specific and blockchain-wide variables show heterogeneity across quantiles in terms of their signs and sizes in explaining bitcoin price movements. This heterogeneity is a likely reason for the lack of consensus when attempting to identify what factors are linked to cryptocurrency price movements.

In light of our findings, we recommend future research into cryptocurrency price movements proceed in at least two (interrelated) paths. The first path can address shortcomings in modeling approaches that focus exclusively on the center of the distribution of cryptocurrencies' price changes. Higher resolution mappings of intervariable relations across the distribution are needed in order to better understand the behavior of these digital coins. Second, future research can also focus on deriving more proxies for investor sentiment. While the investor clientele for cryptocurrencies may differ from that of traditional asset classes, it has been shown in studies that psychology and emotions are strong determinants of asset price movements. More proxies for sentiment can thus help in assessing to what extent their prices move due to investment behavior, or, psychology. This can then answer broader questions, such as whether the prices of these digital assets are irrational, or, whether they reflect some intrinsic value.

Appendix: Composition and description of the aggregate cryptocurrency market index (CRYIX)

Constructing an aggregate cryptocurrency market index (*CRYIX*) serves the purpose of gauging whether our sentiment-return relation remains robust across quantiles. As mentioned in Eq. (4), our index is market-weighted whereby the weight of each of the respective nine constituents is determined by dividing its market capitalization by the total market capitalization of the nine cryptocurrencies. The construction of this market-weighted index is similar to, say, that of the S&P 500 index, and serves to capture aggregate market conditions. The equation for the returns for our *CRYIX* index, $R_{m,t}$, is as follows:

$$R_{m,t} = \sum_{i=1}^N R_{i,t} \times (MktCap_{i,t} / TotalMktCap_t)$$

whereby $R_{i,t}$ and $MktCap_{i,t}$ is the return and market capitalization of cryptocurrency i , respectively, while $TotalMktCap_t$ denotes the total market capitalization of all nine of the N cryptocurrencies. Details as to the economic usefulness of this measure are discussed in sub-Sect. 2.5. The nine cryptocurrencies included in the *CRYIX* are as follows (along with their abbreviations and their sample ranges):

- Bitcoin (BTC): 1/26/2015—10/23/2020.
- XRP (XRP): 1/26/2015—10/23/2020.
- Ethereum (ETH): 8/10/2015—10/23/2020.
- Stellar (XLM): 9/30/2015—10/23/2020.
- Litecoin (LTC): 1/26/2015—10/23/2020.
- Dash (DASH): 1/26/2015—10/23/2020.
- Monero (XMR): 1/26/2015—10/23/2020.
- NEM (XEM): 4/1/2015—10/23/2020
- Dogecoin (DOGE): 1/26/2015—10/23/2020.

For those cryptocurrencies that start somewhat later (such as ETH, XLM, and XEM) than the sample period which this study examines (1/26/2015—10/23/2020) more weight is naturally given to the remaining cryptocurrencies (BTC, XRP, LTC, DASH, XMR, and DOGE). The reason some of these cryptocurrencies are observed after the beginning date of 1/26/2015 is because some of them launched relatively later on and thus had incomplete, or, irregular price data. In addition to this, so-called “stablecoins” are not included in the *CRYIX* composition since their price is explicitly “pegged” to a national currency, typically the US dollar.

Summary statistics of the price behaviors of these nine cryptocurrencies that comprise the *CRYIX* are shown in Table 5. This table shows (in USD) the average price, maximum and minimum price, and average market capitalization of the nine cryptocurrencies. The

Table 5 Summary statistics of CRYIX constituents

	Cryptocurrency	Abbrev	Avg. Price (in USD)	Max. Price (in USD)	Min. Price (in USD)	Avg. Market Cap. (in USD)
1	Bitcoin	BTC	\$4,753.3645	\$19,640.5139	\$211.9559	\$83,275,461,869
2	XRP	XRP	\$0.2352	\$2.7458	\$0.0041	\$23,515,102,703
3	Ethereum	ETH	\$210.9683	\$1,380.1445	\$0.4200	\$21,447,661,937
4	Stellar	XLM	\$0.0912	\$0.8802	\$0.0014	\$9,515,077,081
5	Litecoin	LTC	\$47.3081	\$359.0726	\$1.3443	\$2,761,356,842
6	Dash	DASH	\$130.8452	\$1,447.4691	\$1.5893	\$1,059,966,814
7	Monero	XMR	\$64.6720	\$477.1168	\$0.2317	\$1,049,930,026
8	NEM	XEM	\$0.1040	\$1.7924	\$0.0001	\$935,748,576
9	Dogecoin	DOGE	\$0.0019	\$0.0168	\$0.0001	\$221,780,057

Table 6 Contribution (%) of CRYIX constituents based on market capitalization

	Cryptocurrency	Mean	Median	Maximum	Minimum	Std. Dev	Skewness	Kurtosis
1	Bitcoin	67.3624	72.3626	88.2833	33.4940	14.3513	-0.5449	-0.9625
2	XRP	14.4971	12.6432	45.7672	2.3259	7.4112	0.5108	-0.4115
3	Ethereum	11.3880	10.3933	31.4536	0.6488	5.9192	0.4336	-0.0820
4	Stellar	4.5438	3.0865	13.2917	0.7480	3.3729	1.0051	-0.4057
5	Litecoin	1.9027	1.7564	6.1559	0.8166	0.6116	0.9731	1.4251
6	Dash	0.6636	0.4910	3.7169	0.1706	0.4868	1.9066	4.1516
7	Monero	0.5760	0.5908	1.5276	0.0320	0.3482	0.0463	-0.8420
8	NEM	0.4753	0.2917	2.5654	0.0137	0.5080	1.6809	2.1002
9	Dogecoin	0.1938	0.1730	0.6340	0.0687	0.0763	1.0196	0.8450

Table 7 Pairwise correlations between CRYIX constituent returns

	BTC	XRP	ETH	XLM	LTC	DASH	XMR	XEM	DOGE
BTC	-	0.3996	0.5422	0.4274	0.6697	0.5579	0.5636	0.4493	0.5310
XRP	0.3996	-	0.3901	0.5955	0.4612	0.3385	0.3708	0.3644	0.4927
ETH	0.5422	0.3901	-	0.3736	0.5266	0.4962	0.4954	0.3807	0.4046
XLM	0.4274	0.5955	0.3736	-	0.4553	0.3675	0.4321	0.4433	0.5005
LTC	0.6697	0.4612	0.5266	0.4553	-	0.5289	0.5067	0.4333	0.5204
DASH	0.5579	0.3385	0.4962	0.3675	0.5289	-	0.5378	0.3940	0.4117
XMR	0.5636	0.3708	0.4954	0.4321	0.5067	0.5378	-	0.3662	0.4107
XEM	0.4493	0.3644	0.3807	0.4433	0.4333	0.3940	0.3662	-	0.3927
DOGE	0.5310	0.4927	0.4046	0.5005	0.5204	0.4117	0.4107	0.3927	-

largest cryptocurrency in the sample is bitcoin, given it is the first cryptocurrency which started in January of 2009. The *CRYIX* index is thus constructed for each trading day t in light of the daily changes in market capitalizations of all the nine cryptocurrencies. Table 6 shows the summary statistics of the nine respective cryptocurrencies' contributions to the *CRYIX* index in terms of market capitalization. For example, XRP has an average contribution (i.e., $MktCap_{i,t}/TotalMktCap_t$) of 14.4971%. The maximum daily contribution was 45.7672% while the minimum daily contribution was 2.3259%.

Finally, Table 7 shows the pairwise correlations between the nine cryptocurrency constituents that make up the *CRYIX*. The lowest pairwise correlation is shown to be between Dash and XRP (at 0.3385) while the highest pairwise correlation is shown to be between bitcoin and litecoin (at 0.6697).

References

- Alabi K (2017) Digital blockchain networks appear to be following Metcalfe's Law. *Electron Commer Res Appl* 24:23–29
- Amihud Y, Mendelson H (1986) Liquidity and stock returns. *Financ Anal J* 42(3):43–48
- Andersen TG (1996) Return volatility and trading volume: an information flow interpretation of stochastic volatility. *J Finance* 51(1):169–204

- Atsalakis GS, Atsalaki IG, Pasiouras F, Zopounidis C (2019) Bitcoin price forecasting with neuro-fuzzy techniques. *Eur J Oper Res* 276(2):770–780
- Bonart J, Gould MD (2017) Latency and liquidity provision in a limit order book. *Quantitative Finance* 17(10):1601–1616
- Buchinsky M (1994) Changes in the US wage structure 1963–1987: application of quantile regression. *Econometrica* 62(2):405–458
- Buchinsky M (1995) Estimating the asymptotic covariance matrix for quantile regression models: a Monte Carlo study. *J Econ* 68(2):303–338
- Carr P, Wu L (2017) Leverage effect, volatility feedback, and self-exciting market disruptions. *J Financial Quantitative Anal* 52(5):2119
- Catania L, Grassi S, Ravazzolo F (2019) Forecasting cryptocurrencies under model and parameter instability. *Int J Forecast* 35(2):485–501
- Chakraborty B (2003) On multivariate quantile regression. *J Statistical Plan Inference* 110(1–2):109–132
- Chelley-Steeley P, Lambertides N, Savva CS (2019) Sentiment, order imbalance, and co-movement: an examination of shocks to retail and institutional trading activity. *Eur Financ Manag* 25(1):116–159
- Chordia T, Roll R, Subrahmanyam A (2002) Order imbalance, liquidity, and market returns. *J Financ Econ* 65(1):111–130
- Economides N (1993) Network economics with application to finance. *Financ Mark Inst Instrum* 2(5):89–97
- Eisler Z, Kertesz J (2006) Size matters: some stylized facts of the stock market revisited. *Eur Phys J B Condens Matter Complex Syst* 51(1):145–154
- Gregoriou GN, Gueyie JP (2003) Risk-adjusted performance of funds of hedge funds using a modified sharpe ratio. *J Wealth Manag* 6(3):77–83
- Grinberg R (2012) Bitcoin: an innovative alternative digital currency. *Hast Sci Technol Law J* 4:159–208
- Hale G, Krishnamurthy A, Kudlyak M, Shultz P (2018) How futures trading changed bitcoin prices. *FRBSF Economic Letter* 12:1–5
- Huber PJ (1967) The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, No. 1, pp. 221–233)
- King T, Koutmos D (2021) Herding and feedback trading in cryptocurrency markets. *Ann Oper Res* 300(1):79–96
- Koenker R (2005) *Quantile regression*. Cambridge University Press, Cambridge, UK
- Koenker R, Bassett G Jr (1978) Regression quantiles. *Econometrica* 46(1):33–50
- Koutmos D (2020) Market risk and bitcoin returns. *Ann Oper Res* 294(1):453–477
- Koutmos D, Payne JE (2021) Intertemporal asset pricing with bitcoin. *Rev Quant Financ Acc* 56(2):619–645
- Kumar A, Lee CM (2006) Retail investor sentiment and return comovements. *J Finance* 61(5):2451–2486
- Lahmiri S, Bekiros S (2019) Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos Solitons Fractals* 118:35–40
- Li X, Wang CA (2017) The technology and economic determinants of cryptocurrency exchange rates: the case of bitcoin. *Decis Support Syst* 95:49–60
- Llorente G, Michaely R, Saar G, Wang J (2002) Dynamic volume-return relation of individual stocks. *Rev Financial Stud* 15(4):1005–1047
- Lo S and Wang JC (2014) Bitcoin as money? current policy perspectives no. 14–4, Federal Reserve Bank of Boston
- Ma F, Liang C, Ma Y, Wahab MIM (2020) Cryptocurrency volatility forecasting: a Markov regime-switching MIDAS approach. *J Forecast* 39(8):1277–1290
- Metcalfe B (2013) Metcalfe’s law after 40 years of ethernet. *Computer* 46(12):26–31
- Pedersen TQ (2015) Predictable return distributions. *J Forecast* 34(2):114–132
- Powell JL (1986) Censored regression quantiles. *J Econom* 32(1):143–155
- Signer A, Favre L (2002) The difficulties of measuring the benefits of hedge funds. *J Altern Invest* 5(1):31–41
- Wenker N (2014) Online currencies, real-world chaos: the struggle to regulate the rise bitcoin. *Tex Rev I Pol* 19:145
- Wheatley S, Sornette D, Huber T, Reppen M, Gantner RN (2019) Are bitcoin bubbles predictable? combining a generalized Metcalfe’s law and the log-periodic power law singularity model. *R Soc Open Sci* 6(6):1–13
- Yang R, Chang X, Mišić J, Mišić VB (2020) Assessing blockchain selfish mining in an imperfect network: honest and selfish miner views. *Comput Secur* 97:1–14
- Yermack D (2015) Is Bitcoin a real currency? an economic appraisal. In *Handbook of Digital Currency* (pp 31–43). Academic Press

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.