Valuation without Cash Flows: What are Cryptoasset Fundamentals?

Nick DeRobertis*

October 5, 2020

*University of Florida, PhD Candidate, Tel: (352)392-4669, Email: Nicholas.DeRobertis@Warrington.ufl.edu

Valuation without Cash Flows: What are Cryptoasset Fundamentals?

Abstract

Cryptoassets represent a novel asset class in which tokens are generated and transacted using cryptography through blockchains. To date, few studies have attempted to derive a fundamental valuation for a cryptocurrency. I developed a model based on the Quantity Theory of Money (QTM) that informs us about fundamental value of a currency, and applied it to understand cryptocurrency valuation. For most cryptocurrencies, an expectation of future use as a currency drives the valuation. I analyzed attention, sentiment, and R&D measures as proxies that form this expectation, and found that they are all significantly related to cryptocurrency returns. A portfolio that was long high attention cryptocurrencies with weekly rebalancing would have earned a 0.58% daily alpha from mid-2017 to the end of 2019. The portfolio which is long high attention cryptocurrencies and short low attention cryptocurrencies has an even higher daily alpha of 0.72%, though it is not currently a tradeable strategy due to short-sale constraints. A portfolio formed from cryptocurrencies with high investor sentiment would have yielded a 0.33% daily alpha. R&D does not show as strong effects, but is still significantly related, and all the proxies for future usage remain significant with a variety of analyses and controls including other crypto market factors such as MKT_c, SMB_c, and UMD_c, and dual portfolio sorts on maturity, size, and momentum.

JEL Classification: E42, E44, F31, G12, G15, G41

Keywords: Cryptoassets; Cryptocurrencies; Asset returns; Currency valuation; Utility tokens; Cryptoassets; Exchange rates

I Intro

Cryptoassets are a burgeoning asset class that could soon outgrow the number of assets in all other classes. A cryptoasset is defined by an algorithm, which is inherently replicable, and most creators of cryptoassets distribute the source code freely. A new cryptoasset can be created simply by duplicating the source code that generated an earlier cryptoasset and modifying it. This has led to rapid growth in the number of cryptocurrencies: the first cryptoasset was Bitcoin (BTC) that originated in 2009, as of September of 2020, there are more than 7,000 cryptoassets tracked in the CoinMarketCap index. Meanwhile, the total cryptoasset market capitalization has grown from $1.5B in May of 2013 to $820B at the peak of the market in January 2018 (286% annualized growth) and $338B in September of 2020 (109% annualized growth) (CoinMarketCap, 2020).

[Figure 1 about here.]

As shown in Figure 1, the biggest growth in cryptoassets was driven by cryptoassets other than BTC and Etherium (ETH). From mid-2016 to end of 2019, BTC tripled and ETH doubled in value while the other cryptoassets grew 10,000x collectively. This is only possible because they represented a negligible part of the market in 2016, but is nonetheless a strong signal that the market is moving towards innovation. These numbers also understate the true growth as they exclude cryptoassets that were created and did not become popular enough to be tracked.

Cryptoassets enable payment and investment systems not reliant on fiat currency or traditional banking. The systems are powered by the source code that created the cryptoasset as well as the users of the currency, so in general as long as the cryptoasset is being used it does not require any external influence. Traditional payment and investment systems rely on counterparties and so introduce counterparty risks. Rational agents should want to minimize risk and so should be encouraged to move to cryptoassets. Even central banks are considering adopting cryptoassets either to replace or supplement fiat currency (Bech and Garratt, 2017). While there are still many challenges to overcome in the short-term, in the long-term cryptoassets are likely to play a significant role as a vehicle for payment and investment.

There are three main types of cryptoassets: cryptocurrencies, security tokens, and utility tokens, depending on the intended usage. Cryptocurrencies are those cryptoassets issued solely for use as currencies. Utility tokens are cryptocurrencies but are intended to only transact for certain goods, and is often the only direct way to purchase those goods. Security tokens are those issued to fund a project, which promise future distributions of cash.

A common criticism of current cryptoassets is their volatility relative to fiat currencies. While some newer cryptoassets such as Tether (USDT) fix their exchange rate to a fiat currency to avoid this volatility, in general this presents an issue for using cryptoassets today. Part of the reason volatility is so high is that there is no accepted model for the fundamental value of a cryptoasset, leading to widely varying estimates of value across different investors.

There is some existing literature trying to develop a valuation model for cryptocurrencies but to date, none have resulted in actionable trading strategies. Jermann (2018) argues that the value of Bitcoin is determined by its current and future usage. He developed a model based on the Quantity Theory of Money (QTM) and demand from current users as well as speculators in the currency, based on the model in Cagan (1956) and Sargent and Wallace (1973). Bolt and Oordt (2020) [BO] also argued that the value of virtual currency is determined by its current and future usage. They developed a model based on the QTM and demand from current users as well as speculators in the currency. In contrast to Jermann (2018), BO develop a more formal model based on QTM, but do not empirically test it. Both of these studies find that the future usage of the cryptocurrency is a key driver of its value, but neither suggest useful proxies for the expectation of future usage other than historical usage.[1]

[1] These are not the only existing studies modeling cryptocurrencies, but others do not focus directly on valuation. Athey, Parashekevov, Sarukkai, and Xia (2013) created a simplified model of a cryptocurrency and used Bayes' rule to update information sets about whether the cryptocurrency will survive. Fernández-Villaverde and Sanches (2019) developed a model of competing cryptocurrencies where entrepreneurs can

Numerous studies empirically examine cryptoasset valuation, but they have focused primarily on examining traditional asset factors or details about the supply-side. Further, many focus only on Bitcoin or a small subset of cryptoassets. Jerman empirically tested the model he developed using Bitcoin historical prices, transaction volumes, and supply, finding that there are two main drivers of returns: changes in transaction volume and changes in other factors, where he theorized that the other factors are driven by R&D but had no empirical test for that theory. Across various studies, size, volatility, and momentum, and investor attention factors calculated from the cryptocurrency markets have been shown to influence prices, while stock market factors tend to be unrelated.[2] Hayes (2017) focused on the supply-side and determined that the amount of competition in mining the currency, the rate of production of the currency, and the difficulty of mining the currency influenced the price of 66 studied cryptocurrencies.

The prior literature has not determined a comprehensive empirical valuation approach which is based on theory. Such an approach would inform investors how to value cryptocurrencies, leading to reduced volatility, which in turn should lead to increased usage.

In this study I have developed a model for cryptocurrency valuation based on the QTM which highlights expected future usage of the cryptocurrency as the primary driver of the current price. The model shows that investor attention, investor sentiment, and R&D specific to each cryptocurrency are part of the information set used to form this expectation.

each choose to launch their own currency by extending the Lagos and Wright (2003) model. They found multiple possible equilibria based on the assumptions, with some implying price stability and others not. Pagnotta and Buraschi (2018) asserted that the overall production of a cryptocurrency and the price of that cryptocurrency are jointly related, and developed a theoretical model of cryptocurrenies based on network effects. Easley, O'Hara, and Basu (2019) created a model explaining cryptocurrency transaction fees and mining behavior. Several papers explore the supply side of cryptoassets by developing mining theory, such as Prat and Walter (2018) and Huberman, Leshno, and Moallemi (2017).

[2]Li and Yi (2019) studied 1,656 cryptocurrencies and determined that size, volatility, and momentum factors all influence prices. Kakushadze (2018) also found that the same factors influence prices in a study of 362 cryptocurrencies. Sovbetov (2018) studied only 5 cryptocurrencies, but also confirmed a relationship of momentum and volatility with cryptocurrency prices, and further found evidence that cryptoasset index returns and volume factors can similarly drive returns. Liu and Tsyvinski (2018) found that traditional stock market factors are unrelated to cryptocurrencies, but shows that cryptocurrency momentum and investor attention are strong determinants of cryptocurrency returns, analyzing just three cryptocurrencies: Bitcoin, Etherium and Ripple. used technical analysis to show how rational learning can generate return predictability in cryptocurrencies, providing additional support for momentum in cryptocurrency returns.

Then I conducted an empirical study that provides evidence of the relationship between cryptocurrency returns and these parts of the information set.

I found that attention, sentiment, and R&D measures are all significantly related to cryptocurrency returns. A portfolio that was long high attention cryptocurrencies with weekly rebalancing would have earned a 0.58% daily alpha from mid-2017 to the end of 2019. The portfolio which is long high attention cryptocurrencies and short low attention cryptocurrencies has an even higher daily alpha of 0.72%, though it is not currently a tradeable strategy due to short-sale constraints. A portfolio formed from cryptocurrencies with high investor sentiment would have yielded a 0.33% daily alpha. R&D does not show as strong effects, but is still significantly related, and all the proxies for future usage remain significant with a variety of analyses and controls including other crypto market factors such as MKT_c, SMB_c, and UMD_c, and dual portfolio sorts on maturity, size, and momentum. I also confirmed that maturity, size and momentum are significantly related to cryptocurrency returns.

This study extends both the theoretical and empirical valuation literatures for cryptoassets. It also furthers the behavioral asset pricing literature focusing on investor attention and investor sentiment, as the two are distinct yet related and few studies have attempted to examine both simultaneously.[3] This study also introduces several novel data sources. Changes in source code have not previously been used as a measure of R&D, and Twitter, Reddit, and Facebook data have not been applied together within a Structural Equation Model (SEM) framework to create measures of investor attention and investor sentiment.

Section II develops the theoretical model to inform the empirical valuation of cryptocurrencies. Section III describes the data used in the empirical study. Section IV examines the results from the analysis, showing evidence that supports the theoretical model and

[3] There are many studies focusing on either investor attention or investor sentiment, but few that analyze both or the relationship betwen them. Da, Engelberg, and Gao (2011) price stocks using both attention and sentiment, but for their measures they find they are not very related. They only have two paragraphs on the relationship and conclude it is ambiguous. There are numerous papers that focus on one or the order such as Barber and Odean (2008) (attention), Tetlock (2007) (sentiment), Kumar and Lee (2006) (sentiment), Edmans, García, and Øyvind Norli (2007) (sentiment), Da, Engelberg, and Gao (2015) (sentiment), and Yuan (2015) (attention).

consistent patterns in prices. Section V concludes.

II Model

Like currencies, cryptocurrencies also have no intrinsic value. Currencies have three properties that differentiate them from general assets [citation needed for this], from which their value is derived. They are a store of value, unit of account, and medium of exchange. Cryptocurrencies and utility tokens in general have these properties as well, and so we can treat them as currencies for valuation purposes.

The medium of exchange function should provide the majority of the value of a cryptocurrency, at least in present day. The store of value property provided little value as there are low-risk alternatives that store value and provide returns such as treasury bills. With the exception of BTC and ETH, very few prices of assets were listed in cryptocurrencies, so the unit of account function was not providing much value as of September of 2020. The existence of transaction volumes for cryptocurrencies proves that the medium of exchange property provides value.

The value of the medium of exchange feature of a currency directly increases with the usage of that currency. For a given level of usage, those that have more desirable features such as quicker payments and lower transaction costs provide greater value, and those desirable features should increase future usage, which implies that R&D into improving desirable features should increase the valuation.

Cryptocurrencies have value as they have utility as a medium of exchange and a store of value. Based on this concept, I use the transactions form of the quantity theory of money (QTM) popularized by Persons and Fisher (1911) to form our theory for cryptoasset valuation. The QTM in its basic form states:

$$MV = PQ \tag{1}$$

Where:

- M: Money supply
- V: Money velocity
- P: Average price level
- Q: Volume of goods transacted

In a cryptoasset valuation framework, M would represent the total value of the cryptoasset, which when divided by the supply of the cryptoasset would yield the price of the cryptoasset. Therefore we can get the price of a cryptoasset by:

$$p\_c = \frac{PQ}{V * S} \tag{2}$$

Where:

- p_c: Price of cryptoasset
- S: Amount of supply of cryptoasset

In traditional economies, PQ is represented by GDP. For cryptocurrencies, we can replace the GDP with instead the share of GDP that is transacted in the cryptocurrency.

$$p\_c = \frac{GDP * u}{V * S} \tag{3}$$

Where:

- u: Percentage of GDP transacted in the cryptoasset

In this study, I did not focus on security tokens since they provide cash flows, so they can be valued as a traditional financial asset. As utility tokens are a specific type of cryptocurrency, we can think about their value in the same way. We can think of these utility tokens as enabling a miniature economy focused on particular assets rather than the economy as a whole. Within this miniature economy, the utility token is the currency, and so the determinants of their value should be similar to that of cryptocurrencies, just on a smaller scale. For utility tokens, GDP can be replaced by the total economic value of the goods or services that can be transacted with that token. If it is the only currency that can be used to transact in that miniature economy, then u would become 1 and future value of the currency would be determined by the value of these goods and services. As will later be shown, a single set of determinants can be used to form expected future usage for cryptocurrencies and also the expected future value of goods and services for utility tokens.

In thinking about cryptocurrencies, for determining the relative value of different currencies, GDP will be constant across all the currencies in a given time period, so I set it to 1 without loss of generality to simplify the model. Similarly, the velocity of money should be determined by general economic conditions and not the particulars of a currency so I also set that to 1, yielding:

$$P_c = \frac{u}{S} \tag{4}$$

With utility tokens, instead the usage will be 1 across different tokens, and the same conclusions apply about the velocity of money, leading to:

$$P_c = \frac{GDP_u}{S} \tag{5}$$

Where:

- GDP_u: Total economic value of assets traded by utility token

Because of this, the same conclusions can be drawn for utililty tokens as cryptocurrencies, just replacing u for GDP_u. Therefore the rest of the model focuses only on cryptocurrencies.

As a cryptocurrency is used for a single transaction and then the owner no longer holds it, only a single future value should be discounted. The period used should be that in which the holder of the cryptoasset transacts with it. Assuming that the holder is a rational agent, she will transact in the period that maximizes the present value of the utility of the transaction, which depends both on the individual's discount rate and the expected benefits arising from the medium of exchange and store of value properties.

$$P_{c0} = \frac{E_T[\frac{u}{S}|I_0]}{(1+r)^T} \tag{6}$$

$$T = \underset{t}{\operatorname{argmax}} \quad E[U(\frac{p_{ct}}{(1+r)^t})|I_0] \tag{7}$$

Where:

- I_t: Information set at time t
- U: Investor utility function

Over time, the growth of the cryptocurrency should slow to the point that the value calculated from the current usage is greater than the discounted value from future usage, at which point the value will be determined by the current usage.

Critically important to the valuation is the information set today, I_0. The information set contains any available information to forecast future values for u, V, and S, and determining the current fair value of the cryptoasset requires forecasting all future values of u, V, and S. Determining a concrete valuation model for a cryptoasset then requires proxying for the information set.

S at time T is affected by the supply limit L, the maximum number of units of the currency that are allowed to be created, as well as the mining rate m.

$$E[S_T|I_0] = f(L, E[\sum_{t=0}^{T} m_t|I_0]) \tag{8}$$

The mining rate is itself dependent on the price. As the price of the cryptoasset increases, benefits to mining increase and therefore so does the rate of mining. As the amount of mining increases, each unit of mining becomes less profitable. Mining is akin to a lottery system, so each additional mining unit reduces the probability of success, p_s, for all mining units, leading to an equilibrium. The costs involved in mining come from the unit cost of purchasing mining units, F, and the variable cost of electricity, e. Therefore the miner's profit, π, equation from time 0 to time T assuming a constant N is given by:

$$\pi_0^T = \sum_{t=0}^{T} (N(P_{ct} p_{st} - e_t)) - NF \tag{9}$$

$$p_s = f(N) \qquad \frac{\partial p_s}{\partial N} < 0 \tag{10}$$

The miner will choose the optimal number of mining units, N^*, to maximize expected utility of profits:

$$N^* = \underset{t}{\operatorname{argmax}} \quad E[U(\pi_0^T)|I_0] \tag{11}$$

The mining rate is set by the current number of mining units in service.

$$m_t = f(N^*) \tag{12}$$

Substituting, I found:

$$m_t = f(\underset{t}{\operatorname{argmax}} \quad E[U(\sum_{t=0}^{T} (N(P_{ct} p_{st} - e_t)) - NF)|I_t]) \tag{13}$$

$$m_t = f(\underset{t}{\operatorname{argmax}} \quad E[U(\sum_{t=0}^{T} (N(\frac{E_T[\frac{u}{S}|I_t]}{(1+r)^T} g(N) - e_t)) - NF)|I_t]) \tag{14}$$

$$m_t = E[f(U, u, S, r, F, e_t)|I_t] \tag{15}$$

Therefore, the function for S_T can be restated as:

$$E[S_T|I_0] = f(U, r, L, E[u_t, F_t, e_t|I_0]) \tag{16}$$

The other main quantity to be forecasted to determine the valuation is u. u ultimately reflects consumer preferences for which cryptoasset to use. The rational choice of a cryptoasset (or traditional currency) comes down to the three aspects that give a currency value: store of value, unit of account, and medium of exchange. The value storage benefits will be determined by the expected future price of the currency versus today $\frac{E[p_{ct}|I_0]}{p_{c0}}$. The unit of account benefit would come from a network effect when many are using the currency, so the main determinant would be $E[u_t|I_0]$. The medium of exchange benefits represents the ease of carrying out the transaction, which is a function of the technological development of the asset, D. Therefore u_t can be forecasted by:

$$E[u_T|I_0] = f(\frac{E[p_{ct}|I_0]}{p_{c0}}, E[u_T|I_0], E[D_T|I_0]) \tag{17}$$

$$E[u_T|I_0] = f(E[S_T, D_T|I_0], u_0) \tag{18}$$

However, there may be important behavioral effects in usage of the currency. Prior to the introduction of cryptoassets, the choice of which currency to transact in was narrow and well-defined. Now there are thousands of cryptoassets available, and the number will only continue to grow, and an optimal choice would require an analysis of each. Ultimately the risk-adjusted benefit for choosing the correct currency is likely less than the analysis costs for most, so most will not carry out such an analysis. As a tractable alternative, individuals may examine a subset of the available cryptoassets. Due to the network benefits of a currency, current usage may be the best predictor of future usage, and transaction volume is easily observable, so individuals may sort by transaction volume and remove coins with low volume from the analysis. Alternatively, the individual may examine only cryptoassets that are discussed in the news because they do not even know the others exist. For individuals reading the news about the cryptoassets, the polarity of the content of the articles could affect their estimates of future usage. As the individual only pays attention to these cryptoassets, their price should increase, following the logic originally proposed by Merton (1987).

For utility tokens, the same concepts apply, but as the model is based on expected value of the miniature economy, the attention and sentiment would be in regards to the assets in the miniature economy, but these values should be highly correlated with those for the utility token itself, as the utility token is the only way to transact these assets. Therefore attention and sentiment for the cryptoasset can be used to value both cryptocurrencies and utility tokens.

Considering the above, the amount of attention, a, and polarity of sentiment, s, could be important predictors for future usage.

$$E[u_T|I_0] = f(E[S_T, D_T|I_0], u_0, a_0, s_0) \tag{19}$$

This ultimately leads to a price function:

$$E[P_{cT}|I_0] = f(U, r, L, u_0, a_0, s_0, E[D_t, F_t, e_t|I_0]) \tag{20}$$

Meaning the price can be determined from the aggregate utility function, the discount rate, the supply limit, current transaction volume, attention, and sentiment, and forecasted technological development, costs of mining hardware, and costs of electricity.

To estimate returns from today to some time T in the future is still driven by the same estimate of the future price:

$$E[R_{c0T}|I_0] = \frac{E[P_{cT}|I_0] - P_{c0}}{P_{c0}} \tag{21}$$

Where:

- R_{c0T}: Cryptocurrency c return from time 0 to time T

As the current price could have been estimated by the similar information in the prior period, we can restate the former equation as:

$$E[R_{c0T}|I_0] = \frac{f(U, r, L, u_0, a_0, s_0, E[D_T, F_T, e_T|I_0]) - f(U, r, L, u_{-1}, a_{-1}, s_{-1}, D_0, F_0, e_0)}{f(U, r, L, u_{-1}, a_{-1}, s_{-1}, D_0, F_0, e_0)} \tag{22}$$

Due to the structure of this equation, we use growth rates in investor attention, investor sentiment, and R&D in the empirical analysis.

III Data

III.A Cryptoasset Exchange Rate History

I collected data on hourly and daily exchange rates between cryptoassets and fiat currencies from the CryptoCompare API.[4] This included data for 33,784 trading pairs (from symbol, to symbol, exchange) collected since 03/10/2013, totaling 300,000,000 observations. In these data, I have information on 3,404 coins and 230 exchanges. The number of days per pair varies from 1 to 2,418 with average 551 and median 499. Summary statistics on the coins are presented in Table 1 Panel A.

[Table 1 about here.]

The majority of trade volume is between cryptoassets and not between cryptoasset and fiat currency. Bitcoin is traded the most with 10,275 pairs, followed by Etherium with 8,152. Only 2,350 pairs involve the US dollar. The exchange market is highly fragmented, with the exchange with the most coins only trading 27% of coins. Bitcoin and Etherium are traded on nearly all of the exchanges, followed by Waves with 74%. The US dollar is only involved

[4] This is a process that requires using code to send multiple requests to the CryptoCompare API for each cryptoasset then aggregating the results of the queries. As part of this paper I have open-sourced a tool I built to automate this process, to aid in future research. The tool is available at https://github.com/nick-derobertis/cryptocompare-py. I request only that you cite this paper if you use this tool.

in 40% of exchanges. Summary statistics on the exchanges are presented in Table 1 Panel B.

For each coin, I calculate the Volume-Weighted Average Exchange Rate (VWAER) across the exchanges on which the coin trades, and use those for the later return analysis. After aggregating across the exchanges at a daily frequency, 754,000 price observations remain.

III.B Social History

Also from the CryptoCompare API, I have collected social information on 3,210 coins that were in the exchange rate history data. These data contain information on Facebook, Twitter, Reddit, Github, and CryptoCompare statistics about the coins, including number of subscribers and number of "likes". The social data begins on 05/25/2017, yielding 880 observations for the coins with the full time-series, but the vast majority of cryptoassets have substantial missing data. Limiting it to a subset with high data coverage, I have 79,000 observations in all. Summary statistics on the available social data are given in Table 2. Correlations of those variables are presented in Table 3

[Table 2 about here.]

[Table 3 about here.]

These data points are measures of attention, sentiment, and R&D. I define attention as the number of people thinking about a cryptoasset in a given time period, sentiment as the percentage of those people who thought about it positively, and popularity as the number of people that thought about it positively. Points such as page views, number of comments, and number of posts are purely measures of attention being paid to the cryptoasset. The Github variables other than stars are measuring the changes in the source code of the cryptoasset and so are natural proxies for R&D. Measures such as Facebook likes, Twitter followers, Reddit subscribers, and Github stars, are measuring popularity, which is a combination of attention and sentiment information. We can extract the pure sentiment information by normalizing it by the attention measures:

$$\text{Sentiment} = \frac{\text{Popularity}}{\text{Attention}} \tag{23}$$

I was concerned that the definition of sentiment being directly a function of attention could lead to strong multicollinearity and difficulty estimating the empirical models. Therefore I have proxied for attention with transaction count for the purpose of this calculation:

$$\text{Sentiment} = \frac{\text{Popularity}}{\text{Transaction Count}} \tag{24}$$

III.C Classifications

Continuing with the CryptoCompare API, I have collected classifications for each coin about their intended usage, including information on 358 of the most traded coins. From these data, I see that 84% are utility tokens, 23% are currencies, and 6% are security tokens, as there is some overlap in the classifications. I also have some information on the underlying technologies and algorithms used to build the coins.

III.D News

I also collect news data from the CryptoCompare API. These data contain the time that an article was posted, a link to the full text of the article, and the cryptoassets that were mentioned in the article. I calculated News Count as the count of articles referencing a cryptoasset within a specified time period. News Polarity was calculated following the approach from Loughran and Mcdonald (2011). First, I created a web scraper to access all the articles and download their text content. Then I classified words as positive or negative based on the Loughran-McDonald dictionary.

III.E Blockchain History

Blockchain data was also extracted from the Cryptocompare API. This includes information on daily transaction counts and supply, as well as mining difficulty.

III.F Github API and Github Archive

While there was Github data for the cyptoassets from the CryptoCompare API, it had a lot of missing data. To fill in this data, I collected information directly with the Github API.[5] In the case of Bitcoin, it had too much data to collect this way, and so I accessed the remaining Bitcoin data from the Github Archive project.[6] After combining the three sources of Github data, there is full coverage for the sample.

III.G Sample Selection

I select the coins on which I have all of the mentioned data available, yielding 53 coins and 24,000 observations.

III.H Returns Calculation

In most asset pricing studies, assets are priced in a single currency. Considering the most of the cryptoassets do not trade directly to USD, instead trading to BTC or ETH and then to USD, pricing all the assets in terms of USD may introduce BTC or ETH's volatility into

[5]This is a process that requires using code to send multiple requests to the Github API for each cryptoasset then aggregating the results of the queries. As part of this paper I have open-sourced a tool I built to automate this process, to aid in future research. The tool is available at https://github.com/nickderobertis/project-report. I request only that you cite this paper if you use this tool.

[6]This is a process that requires using code to send multiple requests to the Github Archive project for each time period, extracting the compressed files, parsing the files to extract the relevant content, then aggregating the results. As part of this paper I have open-sourced a tool I built to automate this process, to aid in future research. The tool is available at https://github.com/nickderobertis/py-gh-archive. I request only that you cite this paper if you use this tool.

the returns. Therefore we perform the analysis separately with USD, BTC, and ETH as the reference currency. In calculating returns, the cryptoasset is converted directly to the reference currency if a VWAER is available for that trading pair. If not, then the cryptoasset is first converted to BTC or ETH, then the BTC or ETH amount is converted to the final currency.

III.I Cryptoasset Market Factors

The cryptoasset market factors, MKT_c, SMB_c, and UMD_c are constructed following Fama and French (1992) but using the sample cryptoasset returns instead of stock returns. Interestingly, these factors seem to be unrelated to stock market factors, as none of the correlations between these factors and those collected from Ken French's website are significantly different than zero over the sample period. Figure 2 shows the cumulative performance of these factors over the sample period.

[Figure 2 about here.]

The market factor is the only one that finishes with a positive return. During the sample period, it seems that large cryptoassets outperformed small cryptoassets and the negative momentum performance implied some return reversal. It is important to note that BTC represents 50-80% of the entire market capitalization throughout the sample period, so it alone can drive results. In robustness checks I exclude BTC entirely from the analysis. While all of the main results are qualitatively similar, the performance of SMB becomes positive in this context, which shows that small coins have outperformed large coins, with the exception of BTC.

III.J Attention, Sentiment, and R&D Factors

Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. The results of the SEM are in Table 4.

[Table 4 about here.]

Overall, there is a good fit with the SEM. While the attention factor does not come up as significantly related to returns, this is due to the standardized variables in the model. The later analysis shows a strong relationship between attention and returns.

[Figure 3 about here.]

As Figure 3 shows, there was a distinct pattern in the average attention and sentiment for the top cryptoassets. Those that have been around longer such as BTC, Litecoin (LTC), and Ripple (XRP) had high attention but low sentiment. Those that have cutting-edge technology such as smart contracts, including Etherium (ETH) and Eosio (EOS) tended to have both high attention and high sentiment. Meanwhile, older cryptocurrencies that are not as well-known such as Stellar (XLM) and Dogecoin (DOGE) fall into the low attention, low sentiment category. Further, we can see that the growth in attention and sentiment slowed throughout 2018 and 2019, while R&D remained relatively constant over time.

IV Analysis

The analysis sought to determine whether investor attention, investor sentiment, and R&D are useful for determining the fair value of a cryptocurrency. To get at this, I explored whether growth in these measures is associated with cryptocurrency returns. I examined this in two main ways: level regressions and portfolio sorts. The portfolio sorts help deal with the non-linearity in the results.

[Table 5 about here.]

Table 5 shows OLS regressions of individual cryptocurrency returns on the attention, sentiment, and R&D factors as well as the crypto market factors and a prior return. The results were highly significant throughout the crypto market factors and the lagged return, suggesting that the traditional asset pricing factors are meaningful in the cryptoasset markets as well. Further, the attention and sentiment factors were highly significant, with the attention effect being positive and the sentiment effect being negative. It is important to note that the regression is only capturing linear effects, while the effects are actually non-linear.

[Table 6 about here.]

Table 6 presents portfolio sorts on the attention, sentiment, and R&D factors. In this analysis, the non-linearity of the effects became clear. Both attention and sentiment showed a positive effect when comparing the High portfolio to the Low portfolio, but in both cases the second portfolio had the lowest returns. While the R&D factor did not show results in the initial regression, here it showed that higher R&D growth is associated with lower returns. It is important to note that these returns were not risk-adjusted, so any effects may be due to correlations between these factors and other risk factors.

[Table 7 about here.]

Table 7 shows the result of regressions of portfolio returns on the crypto market factors. As the standard risk factors were controlled for, and the use of portfolios allows capturing non-linerity, this should be the most accurate analysis. Here the alpha represented the relative performance of the portfolio after controlling for risk factors. Attention showed a positive and significant alpha in the High portfolio, which seemed to be the main driver of the long-short portfolio's out-performance. Similarly, the High sentiment portfolio showed a positive and significant alpha, but it was not different enough from the Low portfolio for the long-short portfolio to be significant. Low R&D and High R&D both out-performed compared to medium R&D, which explains why the levels regression was not able to find significance. It is important to note that these long-short portfolios are meant to be illustrative and not a tradeable strategy. Currently there is not a robust securities lending market for cryptoassets, but it should develop over time as they become more popular.

[Figure 4 about here.]

[Table 8 about here.]

Figure 4 and Table 8 both examine the cumulative performance of portfolios over time. It is important to consider not only that effects were different by portfolio but also by time. Figure 4 shows that over time, the attention long-short portfolio performed by far the best, with a 150x return over the sample period. Meanwhile, sentiment had positive returns initially then crashed in January of 2018 and did not recover. A similar pattern is seen for R&D only it never generated a substantive positive return.

Table 8 reveals that the timing of the returns may make it more difficult to capture the results in a standard analysis. Significant and positive effects were still seen 90 days after portfolio formation in all of Attention, sentiment, and R&D. As portfolios were formed weekly, a substantial portion of the effects may be left out in other analyses.

[Table 9 about here.]

Dual-sorts among the attention, sentiment, and R&D factors are carried out in Table 9. This analysis established two pieces of evidence. One is that the factor effects were distinct, as the effect of one portfolio did not get washed out by the other. The second is that it offered potentially optimal trading strategies (once it is possible to short cryptoassets), as the dual-sort portfolio return differences were of greater magnitude. With each of the portfolio combinations, a long-short portfolio could theoretically be formed that earns high returns. For example, a portfolio that was long high attention and sentiment and short low attention and sentiment would have yielded a 2.22% daily return.

[Table 10 about here.]

[Table 11 about here.]

[Table 12 about here.]

[Table 13 about here.]

The sorts in Tables 10, 11, 12, and 13 are mainly to show that the result is not driven by some other possible risk factors. The attention, sentiment, and R&D effects are distinct from size, momentum, and maturity effects. The single sort in Table 10 also confirms the negative momentum during the sample period.

V Conclusion

Cryptoassets have seen explosive growth over the last decade and show no sign of stopping. They enable payment systems with no central authority or clearinghouse as well as decentralized applications. As these assets become mainstream and more retail investors enter the market, it is important to understand their valuation. In this study I have shown that the QTM provides a reasonable framework for cryptocurrency valuation and that investor attention, sentiment, and R&D can be useful proxies to predict the future usage that truly drives the valuation. It was surprising that R&D did not have a stronger effect, but it is perhaps that the information is not quickly absorbed by the market, or it is that the R&D as measured represented too small of changes to have a meaningful effect. It would still be useful to know what implications come from the model in a more rigorous set up with a model economy.

References

Athey, Susan, Ira Parashekevov, Vishnu Sarukkai, and Jing Xia, 2013, Bitcoin Pricing, Adoption, and Usage: Theory and Evidence, *Stanford Working Paper Series*.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Bech, Morten L., and Rodney Garratt, 2017, Central Bank Cryptocurrencies, *BIS Quarterly Review* September 2017.

Bolt, Wilko, and Maarten R.C. Van Oordt, 2020, On the Value of Virtual Currencies, *Journal of Money, Credit and Banking* 52, 835–862.

CoinMarketCap, 2020, Cryptocurrency Market Capitalizations, `https://coinmarketcap.com/`.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In Search of Attention, *Journal of Finance* 66, 1461–1499.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2015, The sum of all FEARS investor sentiment and asset prices, *Review of Financial Studies* 28, 1–32.

Easley, David, Maureen O'Hara, and Soumya Basu, 2019, From mining to markets: The evolution of bitcoin transaction fees, *Journal of Financial Economics* 134, 91–109.

Edmans, Alex, Diego García, and Øyvind Norli, 2007, Sports sentiment and stock returns, *Journal of Finance* 62, 1967–1998.

Fama, Eugene F., and Kenneth R. French, 1992, The CrossSection of Expected Stock Returns, *The Journal of Finance* 47, 427–465.

Fernández-Villaverde, Jesús, and Daniel Sanches, 2019, Can currency competition work?, *Journal of Monetary Economics* 106, 1–15.

Hayes, Adam S., 2017, Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin, *Telematics and Informatics* 34, 1308–1321.

Huberman, Gur, Jacob D. Leshno, and Ciamac Moallemi, 2017, Monopoly without a monopolist : An economic analysis of the bitcoin payment system, Research Discussion Papers 27/2017 Bank of Finland.

Jermann, Urban J., 2018, Bitcoin and Cagan's Model of Hyperinflation, *SSRN Electronic Journal*.

Kakushadze, Zura, 2018, Cryptoasset Factor Models, *SSRN Electronic Journal*.

Kumar, Alok, and Charles M.C. Lee, 2006, Retail investor sentiment and return comovements, *Journal of Finance* 61, 2451–2486.

Lagos, Ricardo, and Randall Wright, 2003, Dynamics, cycles, and sunspot equilibria in 'genuinely dynamic, fundamentally disaggregative' models of money, *Journal of Economic Theory* 109, 156–171.

Li, Jiasun, and Guanxi Yi, 2019, Toward a factor structure in crypto asset returns, *Journal of Alternative Investments* 21, 56–66.

Liu, Yukun, and Aleh Tsyvinski, 2018, Risks and Returns of Cryptocurrency, *National Bureau of Economic Research*.

Loughran, Tim, and Bill Mcdonald, 2011, When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Merton, Robert C., 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, *The Journal of Finance* 42, 483–510.

Pagnotta, Emiliano, and Andrea Buraschi, 2018, An Equilibrium Valuation of Bitcoin and Decentralized Network Assets, *SSRN Electronic Journal*.

Persons, Warren M., and Irving Fisher, 1911, Fisher's "The Purchasing Power of Money"The Purchasing Power of Money, *Publications of the American Statistical Association* 12, 818.

Prat, Julien, and Benjamin Walter, 2018, An Equilibrium Model of the Market for Bitcoin Mining, CESifo Working Paper Series 6865 CESifo Group Munich.

Sovbetov, Yhlas, 2018, Factors Influencing Cryptocurrency Prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero, *Journal of Economics and Financial Analysis* 2, 1–27.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

Yuan, Yu, 2015, Market-wide attention, trading, and stock returns, *Journal of Financial Economics* 116, 548–564.

Appendices

[Table 14 about here.]

[Table 15 about here.]

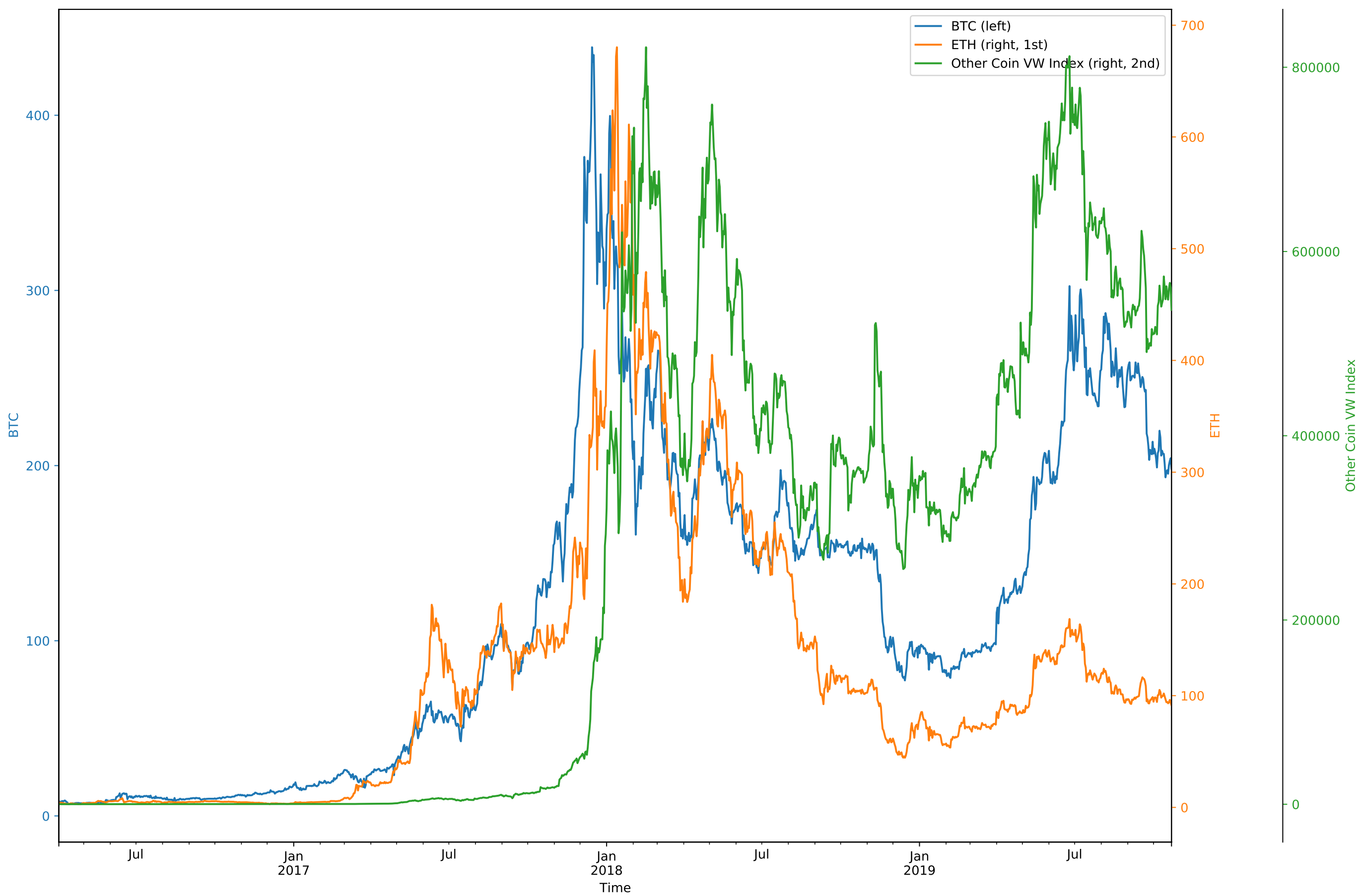


Figure 1: BTC vs Non-BTC Crypto Index. The cumulative percentage return of individual cryptoassets as well as an index of cryptoassets are presented. BTC is shown with its values based off the left axis. ETH and Other Coin VW Index are each aligned with an axis on the right side of the graph, one each and in that order. The index is formed by excluding BTC and ETH and then taking the value-weighted averages of the returns of the remaining coins in the sample each period. For all series on the graph, the cumulative buy-and-hold returns are calculated.

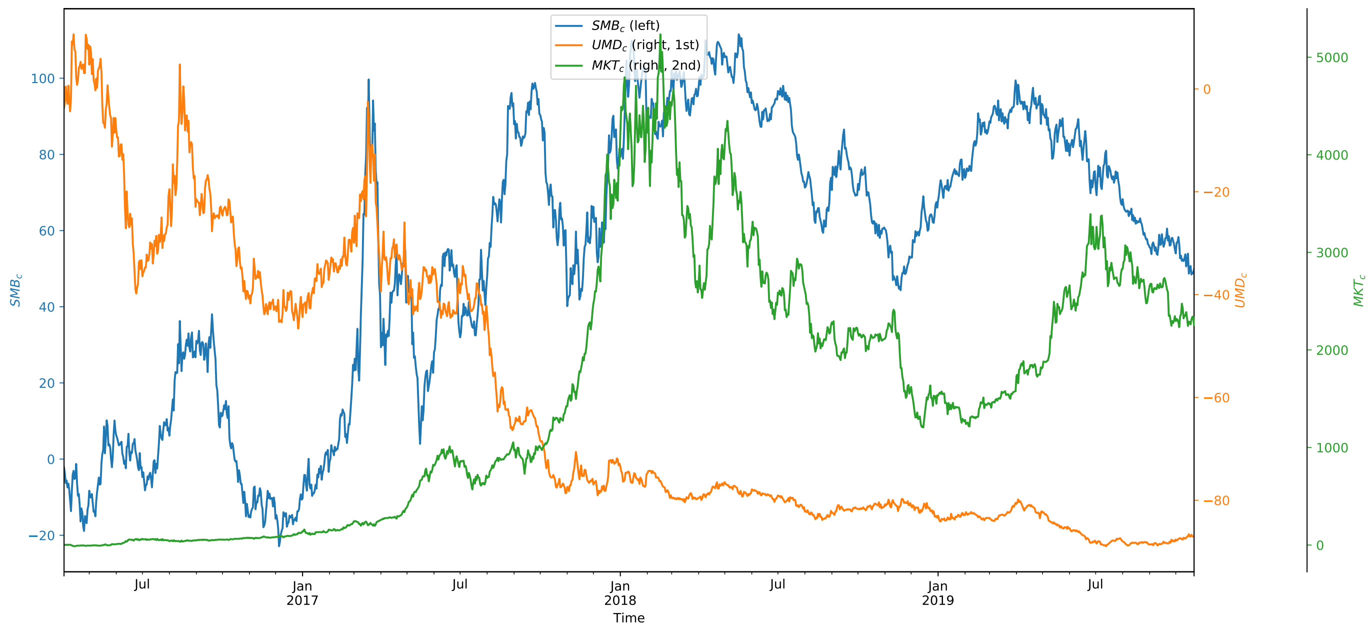
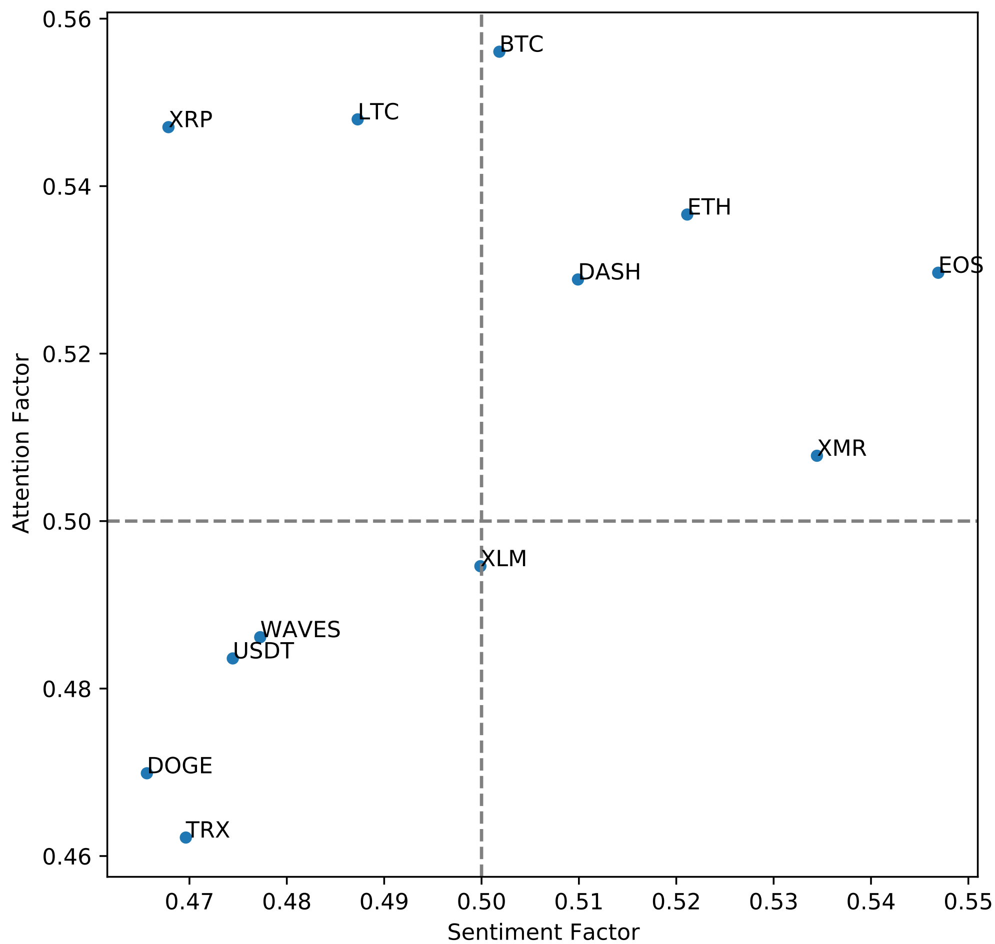
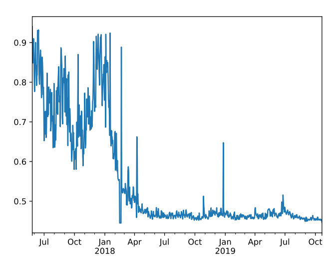


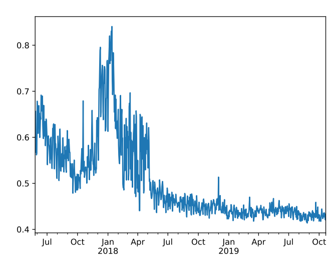
Figure 2: Cumulative Factor Performance. Cumulative buy-and-hold percentage returns of cryptoasset market factors are shown. SMB_c is shown with its values based off the left axis. UMD_c and MKT_c are each aligned with an axis on the right side of the graph, one each and in that order. The cryptoasset market factors, MKT_c, SMB_c, and UMD_c are constructed following Fama and French (1992) but using the sample cryptoasset returns instead of stock returns.



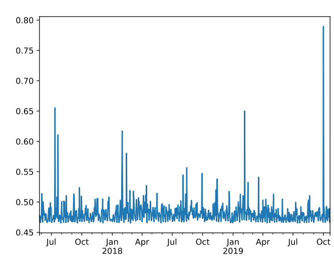
(i) Average Attention and Sentiment for Top Coins



(ii) Average Attention over Time



(iii) Average Sentiment over Time



(iv) Average R&D over Time

Figure 3: Attention, Sentiment, and R&D Overview. Subfigure (i) plots the most frequently traded cryptoassets by their average levels of attention and sentiment. Subfigures (ii), (iii), and (iv) plot value-weighted averages of attention, sentiment, and R&D, respectively, across all sample cryptoassets. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2.

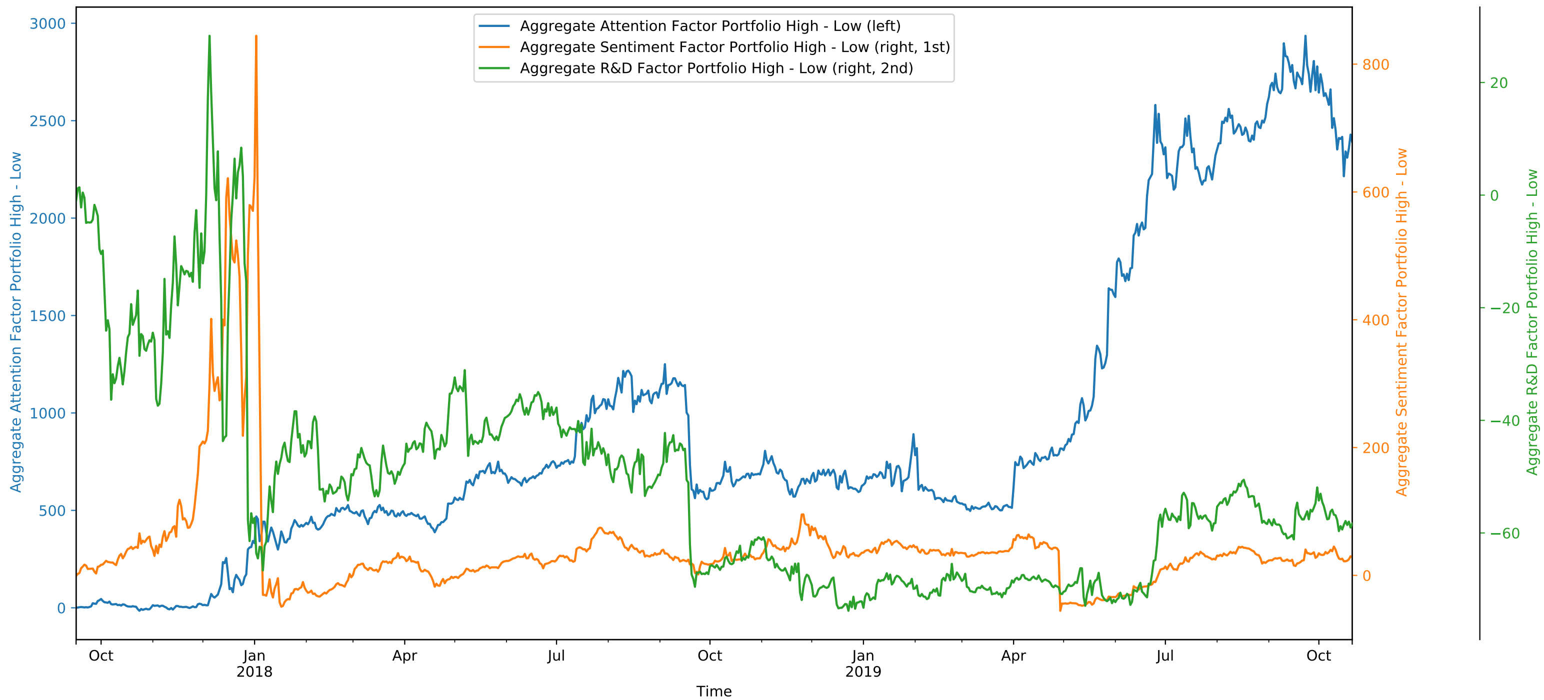


Figure 4: Performance of Long-Short Portfolios. Cumulative buy-and-hold returns of attention, sentiment, and R&D factor long-short portfolios are presented. Aggregate Attention Factor Portfolio High - Low is shown with its values based off the left axis. Aggregate Sentiment Factor Portfolio High - Low and Aggregate R&D Factor Portfolio High - Low are each aligned with an axis on the right side of the graph, one each and in that order. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, and Aggregate R&D Factor, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints. High - Low portfolios are constructed by subtracting the Low portfolio returns from the corresponding High portfolio returns in each time period.

Table 1: Cryptoasset Market Summary Statistics

Panel A: Top Cryptoassets

Name	Symbol	Count	First Listing
Bitcoin	BTC	3974	2013-03-11
Ethereum	ETH	2912	2015-08-06
Tether	USDT	1751	2015-01-08
Litecoin	LTC	573	2013-09-28
Dogecoin	DOGE	470	2014-01-31
Bitcoin Cash	BCH	387	2017-07-31
XRP	XRP	288	2014-12-01
Dash	DASH	264	2014-02-02
EOS	EOS	254	2017-06-28
Waves	WAVES	201	2016-06-19
ZCash	ZEC	184	2016-10-27
USD Coin	USDC	183	2018-10-09
TRON	TRX	177	2017-09-03
Binance Coin	BNB	176	2017-07-13
Ethereum Classic	ETC	171	2016-07-23
OMG Network	OMG	160	2017-06-19
Monero	XMR	156	2014-07-22
True USD	TUSD	155	2018-03-19
Stellar	XLM	146	2014-09-03
0x	ZRX	144	2017-08-10

Panel B: Top Exchanges

Exchange	Number of Pairs	Location	Number of Coins	First Listing
Binance	499	Malta	143	2017-07-13
Yobit	444	Russia	97	2017-05-25
OKEX	396	Malta	131	2017-11-30
HitBTC	376	United Kingdom	153	2017-05-25
HuobiPro	305	Singapore	121	2017-11-23
Cryptopia	292	New Zealand	70	2017-05-25
Kucoin	255	Hong Kong	106	2017-12-12
huobikorea	250	South Korea	106	2019-05-20
BitTrex	247	United States	128	2015-01-28
Gateio	236	Cayman Islands	112	2017-11-23
LiveCoin	222	United Kingdom	106	2017-02-21
Bitfinex	214	British Virgin Islands	82	2015-02-07
Upbit	181	South Korea	110	2018-04-26
TradeSatoshi	178	United Kingdom	36	2018-02-20
Zecoex	170	India	88	2018-08-02
CoinEx	155	Hong Kong	53	2018-04-29
Novaexchange	152	Sweden	47	2016-06-14
Liqui	143	Ukraine	51	2016-07-18
Ethfinex	133	Hong Kong	50	2019-01-21
IDEX	130	Panama	121	2018-04-17

Summary statistics for the most common cryptoassets and exchanges are presented. Panel A shows the statistics by cryptoasset, aggregated across exchanges. The Count is the sum of the number of trading pairs involving the cryptoasset across the exchanges on which that cryptoasset is listed. First listing represents the earliest date on which any exchange listed the cryptoasset. Panel B shows the statistics by exchange, aggregated across cryptoassets. Number of Pairs represents the number of unique trading pairs across the entire history of the exchange, while Number of Coins is the unique number of cryptoassets across the entire history of the exchange. Location is the most recent known location of the exchange and may be different than the founding location of the exchange. First Listing represents the earliest date on which the exchange had a cryptoasset transaction.

Table 2: Summary Statistics

	Count	Mean	Std. Dev.	Min	5%	50%	95%	Max
Panel A: Main Analysis Variables								
Price	32,131	232.42	1,266.6	0.00	0.00	0.75	390.07	19,341
Return	32,131	-0.04	6.98	-34.61	-9.99	-0.18	10.10	75.93
Attention Factor	32,131	0.49	0.10	0.39	0.44	0.45	0.75	0.99
Sentiment Factor	32,131	0.49	0.12	0.36	0.40	0.44	0.76	1.00
R&D Factor	32,131	0.49	0.08	0.42	0.47	0.47	0.66	0.98
Transaction Count	32,131	93,290	394,484	0	0	1,698	559,586	8.69M
Supply	31,505	200.04B	1,220.8B	0.09	1.90M	109.63M	104.98B	8,079.8B
Market Capitalization	31,505	6.94B	24.63B	0.38	1.89M	316.03M	36.11B	323.84B
MKT_c	32,131	0.06	3.41	-9.24	-6.06	0.09	5.83	9.70
SMB_c	32,131	-0.01	1.34	-7.23	-2.11	-0.00	2.11	7.11
UMD_c	32,131	-0.09	2.08	-8.35	-3.56	-0.17	3.18	13.02
Panel B: Social Factor Determinants								
Twitter Followers	32,131	148,931	186,096	2	2,563.5	83,358	447,820	939,841
Cc Followers	32,131	8,486.7	14,284	4	97	2,286	42,255	73,860
Cc Posts	32,131	4,861.7	13,860	1	12	329	30,975	109,240
Twitter Favorites Per Transaction	32,131	0.00	0.00	-0.00	0	0	0.01	0.03
Code Repo Stars Per Transaction	32,131	0.00	0.00	0	0	0	0.01	0.02
News Polarity	32,131	-0.09	0.30	-1.00	-0.53	-0.12	0.42	1.00
Code Repo Commits	32,131	6,677.2	6,427.9	1	23	4,581	18,510	22,699
Code Repo Loc Changed	32,131	6.77M	12.85M	147	1,269	2.93M	50.33M	68.78M
Code Repo Loc Added	32,131	1.73M	4.01M	-2.72M	459	631,339	13.88M	25.47M

Sample summary statistics are presented. Panel A contains the descriptive statistics on the variables used in the main analysis. Returns are in percentages. Panel B contains the descriptive statistics on the variables used to calculate the Attention Factor, Sentiment Factor, and R&D Factor.

Table 3: Correlations of Sample Variables

	Return	Attention Factor	Sentiment Factor	R&D Factor	Twitter Followers	Cc Followers	Cc Posts	Twitter Favourites	Code Repo Stars	News Polarity	Code Repo Commits	Code Repo Loc Changed	Code Repo Loc Added	Transaction Count
Return	1.00													
Attention Factor	0.09	1.00												
Sentiment Factor	0.02	0.41	1.00											
R&D Factor	-0.00	0.05	0.14	1.00										
Twitter Followers	-0.00	-0.07	-0.05	-0.03	1.00									
Cc Followers	0.00	0.01	-0.03	-0.09	0.85	1.00								
Cc Posts	0.00	-0.03	-0.06	-0.06	0.82	0.85	1.00							
Twitter Favourites	-0.01	-0.09	-0.05	-0.06	0.13	0.09	0.08	1.00						
Code Repo Stars	0.00	0.03	-0.01	-0.04	0.66	0.81	0.85	-0.04	1.00					
News Polarity	-0.01	0.03	0.11	-0.02	-0.16	-0.18	-0.12	-0.08	-0.10	1.00				
Code Repo Commits	0.01	0.07	0.00	-0.07	0.42	0.48	0.37	0.19	0.37	-0.16	1.00			
Code Repo Loc Changed	0.00	0.07	0.06	-0.05	0.22	0.38	0.19	0.29	0.27	-0.05	0.28	1.00		
Code Repo Loc Added	0.00	0.06	0.06	-0.04	0.15	0.31	0.14	0.29	0.20	-0.02	0.19	0.93	1.00	
Transaction Count	-0.00	0.06	0.02	0.02	0.34	0.28	0.27	0.02	0.25	-0.06	0.15	0.29	0.11	1.00

Correlations of cryptoasset returns and the variables used to construct the Attention Factor, Sentiment Factor, and R&D Factor are shown here.

Table 4: Latent Variables Structural Equation Model (SEM)

Panel A: Goodness of Fit				
Model χ^2	1322.095			
Model p-value (χ^2)	0.000			
Baseline χ^2	76463.841			
Degrees of Freedom	61			
Goodness of Fit Index (GFI)	0.983			
Adjusted-GFI (AGFI)	0.976			
Tucker-Lewis Index (TLI)	0.983			
Comparative Fit Index (CFI)	0.983			
Root Mean Square Error of Approximation (RMSEA)	0.025			
Panel B: Structural Equations				
	Return			
	Coef.	SE	Z-score	P-value
Attention Factor	-0.062	0.052	-1.193	0.233
MKT_c	0.578	0.005	127.145	0.000
Return_{t-1}	-0.030	0.004	-6.735	0.000
Sentiment Factor	1.150	0.301	3.823	0.000
SMB_c	0.147	0.005	32.442	0.000
R&D Factor	-0.034	0.009	-3.829	0.000
UMD_c	0.148	0.004	33.064	0.000
Panel C: Latent Equations				
	Attention Factor			
	Coef.	SE	Z-score	P-value
Twitter Followers	1.000			
Cc Followers	1.236	0.022	55.772	0.000
Cc Posts	1.063	0.020	52.921	0.000
	Sentiment Factor			
	Coef.	SE	Z-score	P-value
Twitter Favorites Per Transaction	1.000			
Code Repo Stars Per Transaction	3.605	0.209	17.234	0.000
News Polarity	0.349	0.057	6.104	0.000
	R&D Factor			
	Coef.	SE	Z-score	P-value
Code Repo Commits	1.000			
Code Repo Loc Added	0.768	0.087	8.874	0.000
Code Repo Loc Changed	0.877	0.091	9.600	0.000

The structure of the SEM is represented by the following equations:

$$r = AF + SF + TDF + r_{t-1} + MKT_c + SMB_c + UMD_c$$

$$AF = TF + CF + CP$$

$$SF = TFPT + CRSPT + NP$$

$$TDF = CRC + CRLC + CRLA$$

$$CF \sim CP$$

$$TF \sim TFPT$$

$$CRC \sim CRLC + CRLA + CRSPT$$

Exogenous variables were scaled to zero mean and unit variance.

Table 5: Regressions of Returns on Factors

Panel A: Levels Regressions

	OLS				
	I	II	III	IV	V
Attention Factor		4.22***		4.86***	4.86***
		(6.55)		(7.11)	(7.11)
Sentiment Factor	0.32			-1.30***	-1.24***
	(0.83)			(-3.29)	(-3.14)
R&D Factor			-0.51		-0.57
			(-1.42)		(-1.57)
Return_{t-1}	-0.02*	-0.03**	-0.02	-0.03**	-0.03**
	(-1.65)	(-2.24)	(-1.64)	(-2.28)	(-2.29)
MKT_c	1.19***	1.18***	1.19***	1.18***	1.18***
	(107.55)	(107.81)	(107.38)	(107.86)	(107.84)
SMB_c	0.79***	0.77***	0.79***	0.77***	0.77***
	(24.12)	(23.46)	(24.21)	(23.51)	(23.51)
UMD_c	0.50***	0.50***	0.50***	0.50***	0.50***
	(23.44)	(23.39)	(23.48)	(23.43)	(23.46)
Intercept	-0.21	-2.14***	0.19	-1.82***	-1.56***
	(-1.16)	(-7.01)	(1.07)	(-5.80)	(-4.61)
R-squared	0.35	0.36	0.35	0.36	0.36
N	32131	32131	32131	32131	32131

Panel B: Ln(Return) Regressions

	OLS				
	I	II	III	IV	V
Attention Factor		2.71***		3.35***	3.35***
		(4.73)		(5.55)	(5.55)
Sentiment Factor	-0.20			-1.31***	-1.27***
	(-0.56)			(-3.61)	(-3.48)
R&D Factor			-0.45		-0.41
			(-1.29)		(-1.16)
Return_{t-1}	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***
	(-2.64)	(-3.10)	(-2.65)	(-3.14)	(-3.15)
MKT_c	1.20***	1.20***	1.20***	1.20***	1.20***
	(110.60)	(110.27)	(110.45)	(110.33)	(110.31)
SMB_c	0.79***	0.78***	0.79***	0.78***	0.78***
	(25.92)	(25.31)	(25.96)	(25.36)	(25.36)
UMD_c	0.49***	0.49***	0.49***	0.49***	0.49***
	(24.41)	(24.33)	(24.43)	(24.38)	(24.40)
Intercept	-0.19	-1.62***	-0.07	-1.31***	-1.12***
	(-1.16)	(-5.99)	(-0.37)	(-4.63)	(-3.55)
R-squared	0.39	0.39	0.39	0.39	0.39
N	32131	32131	32131	32131	32131

Regressions of returns on Attention Factor, Sentiment Factor, and R&D Factor are presented. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. The cryptoasset market factors, MKT_c, SMB_c, and UMD_c are constructed following Fama and French (1992) but using the sample cryptoasset returns instead of stock returns.

Table 6: Single Portfolio Sorts

Panel A: VW Returns	Attention Factor Portfolio	Sentiment Factor Portfolio	R&D Factor Portfolio
1	0.19**	0.68***	1.26***
2	0.37***	0.12	0.63***
3	0.36***	0.92***	0.25***
4	0.76***	0.93***	-0.10

Panel B: Counts	Attention Factor Portfolio	Sentiment Factor Portfolio	R&D Factor Portfolio
1	7629	5709	7574
2	4145	5866	3107
3	5159	5200	5577
4	2993	3151	3668

Average daily returns within portfolios are presented. Panel A shows the returns while Panel B shows the count of cryptoasset/time observations in that portfolio across the sample period. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, and Aggregate R&D Factor, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. * represents significance at the 90% level. ** represents significance at the 95% level. *** represents significance at the 99% level.

Table 7: Regressions of Portfolio Returns on Factors

Portfolio	Alpha		MKT_c		SMB_c		UMD_c		N	Adj-R2
Panel A: Aggregate Attention Factor Portfolio										
1	-0.15	(-1.20)	1.17***	(24.79)	0.62***	(5.05)	0.33***	(4.14)	879	0.59
2	-0.16	(-1.10)	1.19***	(24.59)	0.84***	(4.85)	0.36***	(3.36)	858	0.51
3	-0.01	(-0.11)	1.14***	(28.18)	0.43***	(3.33)	0.05	(0.54)	879	0.63
4	0.58***	(3.22)	1.14***	(18.57)	-0.10	(-0.53)	-0.01	(-0.06)	879	0.38
High - Low	0.72***	(3.61)	-0.03	(-0.52)	-0.72***	(-3.68)	-0.34***	(-3.20)	879	0.05
Panel B: Aggregate Sentiment Factor Portfolio										
1	-0.04	(-0.21)	1.25***	(19.45)	0.22	(1.06)	0.50***	(3.62)	879	0.40
2	0.09	(0.74)	1.13***	(22.42)	0.21*	(1.69)	-0.02	(-0.28)	879	0.57
3	0.22	(1.53)	1.21***	(26.16)	0.64***	(4.04)	0.18**	(2.20)	879	0.52
4	0.33**	(2.05)	1.14***	(20.29)	0.63***	(3.72)	0.24**	(2.33)	879	0.44
High - Low	0.37	(1.52)	-0.11	(-1.57)	0.41	(1.63)	-0.26	(-1.63)	879	0.02
Panel C: Aggregate R&D Factor Portfolio										
1	0.06	(0.39)	1.17***	(21.80)	0.22	(1.32)	0.33***	(3.24)	879	0.46
2	-0.00	(-0.02)	1.15***	(25.30)	0.35**	(2.28)	0.17**	(1.97)	857	0.54
3	0.11	(0.86)	1.09***	(22.36)	0.15	(1.05)	-0.01	(-0.08)	878	0.54
4	0.07	(0.58)	1.18***	(22.70)	0.76***	(4.87)	0.24**	(2.56)	802	0.59
High - Low	0.01	(0.08)	0.01	(0.24)	0.36*	(1.80)	-0.11	(-1.14)	802	0.01

Weighted-average daily portfolio returns are regressed on cryptoasset market factors. Each row in the table represents a portfolio regression. t-statistics are in parentheses. * represents significance at the 90% level. ** represents significance at the 95% level. *** represents significance at the 99% level. The cryptoasset market factors, MKT_c, SMB_c, and UMD_c are constructed following Fama and French (1992) but using the sample cryptoasset returns instead of stock returns. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, and Aggregate R&D Factor, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints.

Table 8: Value-Weighted Cumulative Buy-and-Hold Returns

Portfolio	0	1	5	30	90	180
Panel A: Attention Portfolio						
1.0	3.24**	3.13**	-0.01	-2.06	4.50	0.57
2.0	1.50	1.66**	0.43	0.67	1.84*	0.24
3.0	2.89*	3.58***	1.54	1.82	6.73**	6.51
4.0	2.93**	3.22**	0.78	3.87	4.80	8.26
Panel B: Sentiment Portfolio						
1.0	13.28	14.09	13.12	-0.88	5.30**	12.69
2.0	3.54*	3.97**	1.56	-0.98	4.07*	6.82
3.0	2.45*	1.87**	0.10	3.33	2.85*	-0.37
4.0	2.58	3.10*	2.06	2.25	5.29	11.90
Panel C: R&D Portfolio						
1.0	-0.88	0.05	0.89	-0.92	26.97	-0.22
2.0	1.18	1.78*	-0.10	0.22	1.46	-0.47
3.0	4.14*	4.56**	2.03	2.55	4.46	9.29*
4.0	1.87	3.44	-0.47	0.93	2.68**	6.76

Cumulative value-weighted averages of returns by portfolio are presented. First, returns are averaged by portfolio and time since portfolio formation. Then the cumulative returns are calculated as product of gross returns over the time period minus one. Each column label in the table represents the number of days since portfolio formation. The returns given in the columns represent the cumulative return from the prior column's day to the current column's day. For example, the column with label 30 represents the cumulative returns from day 5-30. * represents significance at the 90% level. ** represents significance at the 95% level. *** represents significance at the 99% level. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, and Aggregate R&D Factor, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints.

Table 9: Dual Portfolio Sorts

Panel A: VW Returns

	Sentiment Factor Portfolio			
Attention Factor Portfolio	1	2	3	4
1	0.53***	-0.22	1.07***	-0.62***
2	0.16	-0.28	0.70***	0.67***
3	0.93***	0.24	0.17	0.42**
4	0.63*	0.16	2.82***	1.69***

	R&D Factor Portfolio			
Attention Factor Portfolio	1	2	3	4
1	0.89***	-0.04	-0.03	0.13
2	-1.20***	1.20***	0.28	-0.27
3	-0.04	-0.17	0.78***	-0.15
4	5.52***	1.96***	-0.19	-0.03

	R&D Factor Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	0.76***	0.29	0.88***	-0.21
2	0.92***	-0.23	0.06	0.27
3	-0.38**	1.98***	0.17	0.03
4	5.45***	0.26	0.71***	-1.41***

Panel B: Counts

	Sentiment Factor Portfolio			
Attention Factor Portfolio	1	2	3	4
1	2591	2344	2056	638
2	1098	1316	1170	561
3	1317	1588	1379	875
4	703	618	595	1077

	R&D Factor Portfolio			
Attention Factor Portfolio	1	2	3	4
1	3124	1314	1785	1406
2	1697	602	984	862
3	1689	880	1916	674
4	1064	311	892	726

	R&D Factor Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	2903	739	1197	870
2	2297	934	1820	815
3	1661	934	1512	1093
4	713	500	1048	890

Average daily returns within the intersection of portfolios are presented. Panel A shows the returns while Panel B shows the count of cryptoasset/time observations in that portfolio across the sample period. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, and Aggregate R&D Factor, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. * represents significance at the 90% level. ** represents significance at the 95% level. *** represents significance at the 99% level.

Table 10: Single Portfolio Sorts

Panel A: VW Returns	Market Capitalization Portfolio	Momentum Portfolio	Maturity Portfolio
1	-0.48***	1.12***	0.98***
2	0.59***	0.51***	1.20***
3	-0.27**	0.34***	1.31***
4	0.52***	0.15	0.30***
Panel B: Counts	Market Capitalization Portfolio	Momentum Portfolio	Maturity Portfolio
1	4761	4949	3656
2	4809	5056	3235
3	4037	5006	6789
4	6319	4915	6246

Average daily returns within portfolios are presented. Panel A shows the returns while Panel B shows the count of cryptoasset/time observations in that portfolio across the sample period. For each of Market Capitalization, Momentum, and Maturity, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints.

Table 11: Dual Portfolio Sorts with Market Capitalization

Panel A: VW Returns

	Market Capitalization Portfolio			
Attention Factor Portfolio	1	2	3	4
1	-0.10	-0.09	-0.56***	0.24
2	-0.66***	1.05***	-1.28***	0.57***
3	-0.94***	0.78***	0.20	0.37***
4	-1.26***	1.46***	0.72*	0.76***

	Market Capitalization Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	-0.54**	0.14	-0.33	0.84***
2	-0.12	0.98***	-0.30*	0.12
3	-0.67***	0.65***	0.53**	0.95***
4	-0.65**	0.65***	-1.05***	1.03***

	Market Capitalization Portfolio			
R&D Factor Portfolio	1	2	3	4
1	-0.45***	0.59***	-0.30*	1.48***
2	-0.85**	-0.77***	-0.96***	0.70***
3	-0.79***	1.10***	0.30	0.25**
4	0.16	0.82***	0.17	-0.17

Panel B: Counts

	Market Capitalization Portfolio			
Attention Factor Portfolio	1	2	3	4
1	2370	2415	1643	1201
2	795	1201	994	1155
3	797	779	874	2709
4	799	414	526	1254

	Market Capitalization Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	1210	1777	1267	1455
2	1359	1006	1245	2256
3	1376	1256	1035	1533
4	816	770	490	1075

	Market Capitalization Portfolio			
R&D Factor Portfolio	1	2	3	4
1	2276	2310	1699	1289
2	605	498	692	1312
3	1199	954	750	2674
4	681	1047	896	1044

Average daily returns within the intersection of portfolios are presented. Panel A shows the returns while Panel B shows the count of cryptoasset/time observations in that portfolio across the sample period. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, Aggregate R&D Factor, Market Capitalization, Momentum, and Maturity, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2.

Table 12: Dual Portfolio Sorts with Momentum

Panel A: VW Returns

	Momentum Portfolio			
Attention Factor Portfolio	1	2	3	4
1	1.51***	-0.05	-0.68***	0.29*
2	1.78***	1.61***	-0.66***	-1.39***
3	1.05***	0.98***	0.08	-0.55***
4	0.71**	-0.09	1.15***	1.33***

	Momentum Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	1.14***	1.47***	-0.33**	0.56**
2	0.95***	0.02	0.02	-0.21
3	1.47***	2.15***	-0.03	0.17
4	1.11***	-0.46**	2.09***	0.88***

	Momentum Portfolio			
R&D Factor Portfolio	1	2	3	4
1	1.22***	-0.43***	2.18***	2.04***
2	0.54**	0.80***	0.40*	0.77***
3	1.35***	0.65***	-0.03	-0.66***
4	1.71***	0.08	-0.65***	-0.93***

Panel B: Counts

	Momentum Portfolio			
Attention Factor Portfolio	1	2	3	4
1	2003	1876	1858	1892
2	1072	1058	1022	993
3	1207	1341	1313	1298
4	667	781	813	732

	Momentum Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	1463	1425	1441	1380
2	1413	1495	1519	1439
3	1349	1285	1215	1351
4	724	851	831	745

	Momentum Portfolio			
R&D Factor Portfolio	1	2	3	4
1	2056	1772	1789	1957
2	709	835	828	735
3	1306	1490	1430	1351
4	878	959	959	872

Average daily returns within the intersection of portfolios are presented. Panel A shows the returns while Panel B shows the count of cryptoasset/time observations in that portfolio across the sample period. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, Aggregate R&D Factor, Market Capitalization, Momentum, and Maturity, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2.

Table 13: Dual Portfolio Sorts with Maturity

Panel A: VW Returns

	Maturity Portfolio			
Attention Factor Portfolio	1	2	3	4
1	-0.71***	0.76***	0.20*	0.19
2	0.86***	0.36	1.41***	0.16
3	1.72***	0.04	0.42**	0.34***
4	1.18***	2.55***	3.87***	0.33

	Maturity Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	0.68**	0.89***	0.85***	0.62***
2	0.08	0.44	0.78***	0.06
3	2.75***	2.71***	1.37***	0.59***
4	0.98***	1.12***	2.52***	0.64***

	Maturity Portfolio			
R&D Factor Portfolio	1	2	3	4
1	0.65***	0.95***	1.44***	1.26***
2	2.74***	3.11***	1.79***	0.23
3	0.86**	0.81***	1.11***	0.13
4	0.36*	-2.55***	0.16	-0.05

Panel B: Counts

	Maturity Portfolio			
Attention Factor Portfolio	1	2	3	4
1	912	1549	3544	1624
2	638	602	1620	1285
3	726	629	1264	2540
4	1380	455	361	797

	Maturity Portfolio			
Sentiment Factor Portfolio	1	2	3	4
1	794	1183	2068	1664
2	834	712	1962	2358
3	795	873	1990	1542
4	1233	467	769	682

	Maturity Portfolio			
R&D Factor Portfolio	1	2	3	4
1	1840	1706	2618	1410
2	203	423	1254	1227
3	542	781	1799	2455
4	1071	325	1118	1154

Average daily returns within the intersection of portfolios are presented. Panel A shows the returns while Panel B shows the count of cryptoasset/time observations in that portfolio across the sample period. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, Aggregate R&D Factor, Market Capitalization, Momentum, and Maturity, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2.

Table 14: Variable Descriptions

Name	Symbol	Description
Price	Prc	The exchange rate between two cryptoassets or a cryptoasset and a fiat currency
Return	r	The percentage change in exchange rate for a cryptoasset over time
Market Capitalization	MC	The market capitalization of the cryptoasset, which is the supply multiplied by the price
Sentiment Factor	SF	A factor constructed using the SEM which represents investor sentiment
Attention Factor	AF	A factor constructed using the SEM which represents investor attention
R&D Factor	TDF	A factor constructed using the SEM which represents technological development of the cryptoasset
Cc Posts	CP	Number of posts for a cryptoasset on CryptoCompare
Cc Followers	CF	Number of followers for a cryptoasset on CryptoCompare
Twitter Followers	TF	Number of Twitter users "following" the cryptoasset
Twitter Favorites Per Transaction	$TFPT$	Number of Twitter users marking the cryptoasset as a favorite divided by transaction count
Code Repo Stars Per Transaction	$CRSPT$	Number of "stars" on the code repository divided by transaction count
Code Repo Commits	CRC	Number of discrete changes made to the code repository
Code Repo Loc Added	$CRLA$	Number of lines of code added to the code repository
Code Repo Loc Changed	$CRLC$	Number of lines of code changed in the code repository
Transaction Count	TC	Daily number of transactions with a cryptoasset
Supply	Supply	Number of units outstanding for a given cryptoasset
News Polarity	NP	A measure of how positive or negative a news article is based on the percentage of words with positive or negative connotations
SMB_c	SMB_c	Small cryptocurrency minus big cryptocurency returns
UMD_c	UMD_c	Previous winner cryptocurrency returns minus previous loser cryptocurrency returns
MKT_c	MKT_c	Value-weighted average cryptocurrency returns

Table 15: Equally-Weighted Cumulative Buy-and-Hold Returns

Portfolio	0	1	5	30	90	180
Panel A: Attention Portfolio						
1.0	5.24**	4.81**	2.08	1.38	4.47	0.49
2.0	0.88	0.75	-0.36	1.04	1.72*	-0.72
3.0	1.63	1.76*	0.49	0.81	2.02**	0.02
4.0	2.34**	2.82**	1.20	5.56	3.07	1.75
Panel B: Sentiment Portfolio						
1.0	8.09	8.19	7.68	0.58	7.24*	5.77
2.0	1.26	1.40*	-0.66	1.44	2.26**	-0.02
3.0	1.76	1.79**	-0.53	3.47	2.69**	-0.06
4.0	1.44	2.23	0.25	2.94	3.45	1.00
Panel C: R&D Portfolio						
1.0	1.38	1.61	2.24	-0.22	7.75	-2.40
2.0	1.70	2.40**	-0.26	1.23	1.63*	0.14
3.0	2.24	2.07	0.36	3.57	3.44	0.11
4.0	1.78*	1.81*	0.26	0.74	3.36	1.98

Cumulative averages of returns by portfolio are presented. First, returns are averaged by portfolio and time since portfolio formation. Then the cumulative returns are calculated as product of gross returns over the time period minus one. Each column label in the table represents the number of days since portfolio formation. The returns given in the columns represent the cumulative return from the prior column's day to the current column's day. For example, the column with label 30 represents the cumulative returns from day 5-30. * represents significance at the 90% level. ** represents significance at the 95% level. *** represents significance at the 99% level. Attention Factor, Sentiment Factor, and R&D Factor are constructed using a Latent Variables approach in a Structural Equation Model (SEM). The structural equation in the SEM relates the cryptoasset returns to the Attention Factor, Sentiment Factor, and R&D Factor. For the latent variable equations, Attention Factor was estimated using Twitter Followers, Cc Followers, and Cc Posts, Sentiment Factor was estimated using Twitter Favorites Per Transaction, Code Repo Stars Per Transaction, and News Polarity, and R&D Factor was estimated using Code Repo Commits, Code Repo Loc Changed, and Code Repo Loc Added. Finally, factors are transformed to a (0, 1) scale using a sigmoid transform, $F_t = \frac{1}{e^{-F_o k}}$ with k=0.2. For each of Aggregate Attention Factor, Aggregate Sentiment Factor, and Aggregate R&D Factor, 4 portfolios are formed by calculating the 25%, 50%, and 75% percentiles of the variables as breakpoints on a weekly basis, and sorting into portfolios based on the breakpoints.