Erasmus University Rotterdam

Erasmus School of Economics

Explaining the (cross-sectional) variation in expected cryptocurrency returns through various characteristics[1]

Master's Thesis Econometrics and Management Science

Quantitative Finance

Author:
M.H. Riemers (445591)
martijnriemers@student.eur.nl

Supervisor:
Prof. dr. P.H.B.F. Franses

Second assessor:
Prof. dr. D.J.C. van Dijk

Date final version:
23rd May 2023

Abstract

In this study, we investigate the primary drivers of cryptocurrency returns by analyzing both the cross-sectional and market returns. Our findings demonstrate that network and sentiment characteristics significantly influence market returns, while production factors play a minimal role. We also discover that a three-factor model, which includes the cryptocurrency market, size, and momentum factors, can effectively account for nearly all cross-sectional cryptocurrency returns. To achieve this, we employ a comprehensive list of market-related return predictors, complemented by projections of network, production, and sentiment characteristics. We find that using the projections of the characteristics to form long-short portfolios, did not result in significant excess returns, except for exposure to the Chinese energy market. Our results indicate that, with the exception of three return predictors, these factors capture the excess returns. Lastly, we offer an explanation for these observed effects.

[1]The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Contents

1 Introduction

Since the advent of the Bitcoin, cryptocurrencies have garnered significant attention from investors, regulators, and the general public. Some view cryptocurrencies as a way to invest in a new technology that will revolutionize the financial system, while others see them as a speculative bubble that will eventually burst. If the former is true, then cryptocurrencies are here to stay, and investors and regulators should understand the factors that drive their returns and the risks associated with them.

Current research has focused on the drivers of cryptocurrency market returns and the risk factors that explain the cross-section of cryptocurrency returns. Among the factors that have been studied are network factors, such as network size and degree of centralization, and production factors, such as mining costs. Additionally, investor sentiment has been shown to have a significant impact on cryptocurrency returns. Furthermore, the relationship between cryptocurrencies and traditional currencies, tail-risk, price inequalities, arbitrage opportunities, and potential price manipulation in the market have also been studied extensively.

This paper focuses on two primary objectives: replicating the findings of previous research and broadening the scope of those studies. Our primary research objective is to explain the key drivers of cross-sectional variations in cryptocurrency returns.

We investigate the factors driving cryptocurrency market returns and build upon Liu and Tsyvinski (2021) by examining the influence of network and production cryptocurrency characteristics. We use a more extensive dataset that includes the post-COVID-19 cryptocurrency boom to capture a broader view of market trends and dynamics.

We hypothesize that production factors, such as mining costs, will not significantly impact returns due to the maturing market, regulatory actions, rising interest rates, and the shift from Proof-of-Work to Proof-of-Stake algorithms. On the other hand, we expect network factors, which serve as proxies for adoption and utility, to have a significant impact on cryptocurrency returns.

Additionally, we explore the impact of investor sentiment on cryptocurrency market returns by employing deep-learning natural language processing techniques. We use the XLM-T model, published in Barbieri et al. (2021), to analyse an extended list of investor-centric message boards and assess the influence of investor sentiment on returns. We expect that investor sentiment, especially as seen through social media influence, will have a pronounced effect on cryptocurrency returns, as demonstrated by recent market events such as the GameStop saga and Elon Musk's Dogecoin endorsements.

Moreover, we replicate the work of Liu et al. (2022) by analysing excess returns of quintile portfolios and examining whether size and momentum factors can capture the cross-section of cryptocurrency returns. We hypothesize that our extended dataset, reflecting a more mature cryptocurrency market, will show risk premia converging more closely to those observed in equity markets, implying that the behavior of the cryptocurrency market will become increasingly similar to traditional financial markets.

We further project the characteristics of the cryptocurrency market onto individual returns to analyse the resulting return predictors. Our hypotheses include more established cryptocurrencies benefiting more from network characteristics, Proof-of-Work cryptocurrencies being sensitive to production characteristics, and “meme” cryptos being influenced by sentiment characteristics.

Our study makes several significant contributions to the existing literature. First, we broaden the scope of previous research by incorporating a much larger dataset that covers the recent post-COVID-19 cryptocurrency boom. Second, we assess the impact of investor sentiment on cryptocurrency market returns by employing a more sophisticated method for deriving sentiment from social media messages. Lastly, we extend our analysis by applying characteristic projections onto individual cryptocurrency returns and examining the resulting returns of the quintile portfolios and their respective factor loadings. This approach allows us to gain a deeper understanding of the relationships between specific cryptocurrency characteristics, their factor loadings, and their subsequent market performance. Overall, our study not only enriches the existing body of research but also provides valuable insights for investors and regulators in the ever-evolving cryptocurrency landscape. The code for this paper is available on GitHub.

This paper is structured as follows. Section 2 describes the data used in this paper. Section 3 presents the methodology used in this paper and describes the various data transformations and return predictors. Section 4 presents the results of our analysis and discusses the implications of our findings. Lastly, section 5 concludes.

1.1 Literature Review

Bitcoin, the first decentralized cryptocurrency, was introduced in 2008 when an anonymous individual or group of individuals under the pseudonym Satoshi Nakamoto published the Bitcoin whitepaper (Nakamoto, 2008). This innovative digital currency introduced the concept of a decentralized ledger that does not rely on a central authority or intermediary for reaching consensus, called a blockchain.

Blockchains consist of blocks of data that are linked together in a chain, starting from the genesis block. Each block contains three elements: the data, the hash of the block, and the hash of the previous block. The data can be any information, such as a transaction or a computational result, that is to be stored in the blockchain. The hash of the block is computed based on the data contained in the block and the hash of the previous block, forming a chain, hence the name blockchain. This mechanism ensures that the data contained in the blockchain is immutable and any tempering with the data is easily detected by validating the hash of the block. To ensure the integrity of the blockchain, the hash of the block must be computed in a way that is difficult to replicate or penalizes any bad actors from attempting to alter the data, but at the same time is easy to verify.

Bitcoin used this decentralized ledger to record transactions and verify their validity. After the Bitcoin gained popularity and widespread acceptance, other cryptocurrencies, such as Litecoin, Ripple, and Ethereum, were introduced, each with their own unique capabilities or differentiator, all centered around the same blockchain technology.

Ethereum, created by Buterin, provided an open source environment for building smart contracts and decentralized apps. It effectively serves as a Blockchain-as-a-service, enabling individuals to develop applications on top of the network by utilizing the Ethereum Virtual Machine (EVM), benefiting from its transparency and availability. Ether (ETH), the cryptocurrency embedded in Ethereum network, more commonly referred to as 'gas' in the community, is used when undertaking any transactions or computational services on the platform. It also enabled the creation of custom tokens, allowing new projects to raise capital outside government control and without the need for lengthy and costly legal scrutiny, via Initial Coin Offerings (ICOs). The Ethereum network has revolutionized the way blockchain applications have been developed.

In contrast to Ethereum and its focus on decentralized apps, Ripple prioritizes payment systems that provide real time settlement, foreign currency exchange and remittance services via its open source protocol that supports varying forms of value such as cryptocurrency commodities or fiat currency. In essence, Ripple seeks to offer secure financial transactions in near instant time and without chargebacks or other fraudulent activity, irrespective of transaction size. Ripple, powered by its native digital ledger XRP, ensures swift transaction processing through its validated consensus facilitated by a network of peer-to-peer servers (Chase & MacBrough, 2018, XRP Ledger). This approach is faster than comparable solutions offered by Bitcoin and Ethereum. Practical usage for cross-border payments or liquidity management services has been established thanks to Ripple's partnerships with major financial institutions. Blockchain technology extends its reach well beyond cryptocurrency usage alone into varied applications like supply chain management or even land registries, among others. It has significant implications for finance and online security/trust, potentially revolutionizing these fields entirely.

To ensure network safety while validating transactions, cryptocurrencies use distinct methods for reaching consensus such as Proof of Work (PoW) and Proof of Stake (PoS). The former requires miners to compute a nonce that satisfies a specific condition, generally a hash with a certain number of leading zeros, before adding the block to the blockchain. Meanwhile, PoS functions by selecting validators based on their holdings or stake of cryptocurrency, and penalizing bad actors by confiscating their stake. The energy consumption patterns between these methods are vastly different. PoW requires more energy due to miners leveraging powerful computer hardware to solve said complex math problems. Sedlmeir et al. (2020) estimates suggest that the top five cryptocurrencies by market cap consume 100 TWh annually through this method. However, newer cryptocurrencies have been embracing PoS for its lower computational power requirements resulting in less wastage of energy; an environmentally friendly approach that is also future-proof, as noted by Nguyen et al. (2019).

Furthermore, exchanges have cropped up globally where individuals can buy and sell various digital currencies since the creation of cryptocurrencies. Cryptocurrency exchanges have facilitated trading between digital assets and traditional currencies such as the US dollar or Euro. This feature has allowed individuals to invest in and speculate using cryptos similarly to fiat money. Additionally, select exchanges now offer crypto shorting and derivatives such as options (Hou et al., 2020).

Despite their popularity, cryptocurrency values remain extremely erratic with prices fluctuating significantly within brief periods of time. This unpredictability has led experts to question their potential as either a groundbreaking form of currency or simply another speculative bubble. Nevertheless, the cryptocurrency market continues its expansion with new digital currencies and innovative blockchain applications regularly emerging.

It is important to also note that while commonly referred to as "currency", these digital assets do not fit into typical currency definitions. A currency, as per economic literature (Bernstein, 2008), is defined as a medium of exchange that is widely accepted for the purchase of goods and services. It should serve as a unit of account, facilitating the comparison of the value of various goods and services. Finally, currencies act as a store of value, enabling individuals to save and use them in the future for the procurement of goods and services. For a currency to reach successful adoption worldwide, a couple conditions have to be met. It must be broadly available while maintaining stability without significant fluctuations in values. Often, this is achieved through the intervention of governments and central authorities regulating the currency's movement through oversight mechanisms preventing frauds or manipulations, ensuring credibility.

However, cryptocurrencies operate differently since they lack centralized governing bodies entirely relying on a decentralized approach, resulting in significant instability of their values, making them unpredictable and risking their reliability at being widely adopted. The lack of widespread acceptance and inconsistent regulatory status across different countries do not deter many from embracing cryptocurrencies as a potentially profitable investment choice akin to stocks and bonds. However, investing in cryptocurrencies carries significant inherent uncertainties due to a lack of proper regulations that puts investors at high risk despite the potential for high returns.

Experts and researchers employ 'risk factor analysis' to identify the underlying determinants influencing an investment its profitability against associated risks. Developed during the early twentieth century, The Capital Asset Pricing Model (CAPM) introduced the "market risk premium" as a mechanism providing an understanding for variations in returns among various stock investments (Lintner, 1975; Mossin, 1966; Sharpe, 1964; Treynor, 1961). After the introducing of the CAPM model, researchers uncovered multiple different factors to reveal variations in returns. Added factors such as size, value and momentum also had a significant influence on stock returns, producing an extensive explanation of cross-sectional variation (Carhart, 1997; Fama & French, 1993).

Furthermore, studies expanded the research beyond stocks to encompass bonds and currencies by examining the key underlying drivers that determine risk and return. For instance, credit risk and the term structure were recognized as essential for bonds while interest rates and carry trade played crucial roles in currency markets (Fama & French, 1993). Today's increased use of risk factor analysis has found its way into mainstream finance, especially portfolio management.

More recently, other asset classes such as real estate and commodities have undergone academic scrutiny in the context of the risk factor analysis. Also, with the constant increases in computational power and data processing abilities, in conjunction with more efficient and effective machine learning algorithms, new techniques such as natural language processing and computer vision have been used to extract risk factors from unstructured data. In the recent literature, these techniques have also been applied to the ever-growing market of cryptocurrencies.

One of the key challenges in conducting risk factor analysis on cryptocurrencies is the lack of historical data, the relative size compared to other equity classes and the illiquidity. This makes comparison of results with already existing literature more difficult, since findings most often do not apply to these markets. The ability to gather data for these assets is also severely hindered by the fragmented landscape of exchanges. Despite these shortcomings, several studies have attempted to identify the risk factors that drive the returns of cryptocurrencies by using various econometric methods such as principal component analysis, multivariate regression, and factor analysis.

Several risk factors have been identified as being associated with cryptocurrency returns, including investor sentiment, investor attention, the regulatory environment, and trading volume. Chen et al. (2019) constructed a sentiment index by analysing the frequency of positive and negative words within StockTwits and Reddit posts to construct an overall sentiment score. They employed a bag-of-words algorithm, which categorizes words into positive and negative sentiment categories, to identify the relevant terms. Their findings revealed that the sentiment index positively correlated with cryptocurrency returns, but only when an accurate lexicon and an investment-focused message board were utilized.

Investor attention has also been linked to cryptocurrency returns. This factor can be measured by tracking the increase in Google searches for cryptocurrency related terms like "Bitcoin". Studies have shown that increased investor attention, as indicated by a surge in relevant search queries, is positively associated with the returns of cryptocurrencies. Moreover, the actions of governments and central banks play a significant role in shaping cryptocurrency returns. Regulatory decisions and policy shifts can dramatically impact the performance of digital currencies in the market.

Another important risk factor associated with the returns of cryptocurrencies is trading volume. Cryptocurrencies are highly speculative assets and are subject to significant price volatility, making them unpopular as a general asset class and thus vulnerable to manipulation by large traders. Studies have found that the returns of cryptocurrencies are positively correlated with trading volume, or rather negatively correlated with illiquidity, which suggests that this can form an important risk factor to consider when analysing the returns of cryptocurrencies. Production factors, such as the average price of electricity and total energy consumption, have also been analysed for their contributions to the returns of cryptocurrencies (Cong, He & Li, 2021; Sockin & Xiong, 2020) due to the fact that many of the high market-cap cryptocurrencies, including Bitcoin, still use a PoW algorithm.

The influence of network factors on the risk of cryptocurrency returns has also been widely studied in recent years (Biais et al., 2018; Cong, Li et al., 2021; Pagnotta & Buraschi, 2018). One of the key network factors is the size of the network, measured by the number of nodes or participants in the network. A larger network size is generally associated with lower risk and higher liquidity, as more participants lead to greater market efficiency and reduced volatility. Another important network factor is the degree of centralization, which refers to the concentration of computational power in the hands of only a few individuals or institutions. A highly centralized network is associated with higher risk, as it is more vulnerable to manipulation and systemic failures. One such attack is called the fifty-plus-one attack where 51% of the hash-rate is in the hands of a group of individuals, giving them complete control over the blockchain. In addition, the network's level of decentralization can also impact the risk of cryptocurrency returns, with a more decentralized network being less susceptible to the influence of a few dominant participants and offering greater security and stability.

Studies have extensively explored the relationship between cryptocurrencies and other traditional currencies, including the correlation between their prices and their market dynamics (Athey et al., 2016; Jermann, 2021; Schilling & Uhlig, 2019). There has also been a significant focus on the tail-risk in the cryptocurrency markets, examining the likelihood of large and rare price movements (Borri, 2019). In recent years, researchers have looked into the presence of price inequalities and the resulting arbitrage opportunities in the cryptocurrency market (Borri & Shakhnov, 2022; Makarov & Schoar, 2020). The widespread use of Tether cryptocurrencies has sparked concerns about the potential for price manipulation within the cryptocurrency space (Griffin & Shams, 2020). Additionally, there have been studies on the relationship between Bitcoin and other cryptocurrencies, focusing on the interdependence and similarities between these digital assets (Hu et al., 2019).

2 Data

2.1 Risk-free Rate and Stock Market Returns

The risk-free rate and stock market returns are obtained from the Kenneth R. French Data Library, a repository of various factors such as SMB (Small Minus Big) and HML (High Minus Low) for different markets. The data library is maintained by Professor Kenneth R. French of the Tuck School of Business at Dartmouth College, who is a leading expert in finance and has co-authored numerous influential studies with Nobel laureate Eugene F. Fama. For the purposes of this paper, we only use data on the risk-free rate and excess stock market returns.

The dataset covers a time period from July 1, 1926, to January 31, 2023, and comprises 25,418 data points representing daily interest rates and excess stock market returns.

2.2 Cryptocurrency Market Data

The market data for cryptocurrencies is obtained from Coinmarketcap.com, a reputable source of cryptocurrency pricing and trading volume information commonly used by researchers. The website collates data from over 200 significant cryptocurrency exchanges, compiling information regarding daily opening and closing prices, highs, lows, volume, and market capitalization for over 18,000 distinct cryptocurrencies. The website includes both dedicated cryptocurrencies utilizing their own blockchain and tokens based on, for example, the Ethereum blockchain. Both active and inactive cryptocurrencies are listed on Coinmarketcap.com, alleviating concerns about survivorship bias. To be eligible for inclusion within the dataset, a cryptocurrency must sustain a minimum market capitalization of $1 million for a period of one week. Additionally, the cryptocurrency must adhere to Sections B1, B2, and C of the CoinMarketCap.com Listings Criteria.

Section B1 outlines the technical prerequisites a cryptocurrency must comply with, covering aspects such as consensus algorithm implementation and the presence of a functional block explorer. Section B2 specifies the exchange requirements necessary for a cryptocurrency to be listed on CoinMarketCap.com, including the requirement for traders to execute both purchase and sale orders, as well as the condition that the exchange must have been operational for a minimum of 60 days. Lastly, Section C elaborates on the subjective assessment policy, stating that specific attributes must be present for a cryptocurrency to be deemed "listed". These criteria encompass factors such as community engagement, institutional partnerships, and the expertise and proficiency of the project's personnel.

The dataset covers a time period from April 28, 2013, to February 16, 2023, and comprises 3,573 data points representing daily cryptocurrency closing prices, market cap, and volume.

2.3 Blockchain Data

The blockchain data is obtained from Blockchain.info, a reputable source of blockchain-related information such as hash-rate, block-size and transaction data for various cryptocurrencies. The website collates data directly from the blockchains of several coins. The website includes data on the blockchain of Bitcoin, Ethereum, Litecoin, Ripple and various other cryptocurrencies. We use data on the number of active addresses, transaction count, payment count, and hash-rate from the Bitcoin blockchain for this research.

Active addresses represent the quantity of distinct addresses created and actively engaged in Bitcoin transactions. The transaction count refers to the number of daily transactions occurring on the blockchain. It is important to note that a single transaction can encompass multiple payments. As the reward for mining a transaction—essentially confirming a transaction and recording it onto the blockchain—is equivalent to the transaction fee, exchanges employ a method known as payment batching to minimize these costs. In practice, the number of payments more accurately reflects the cryptocurrency's activity. The hash rate denotes the computational capacity available to the cryptocurrency. This metric is anticipated to increase somewhat exponentially, as the power of GPUs and CPUs consistently rises in tandem with user count growth.

The dataset covers a time period from April 28, 2013, to February 16, 2023, and comprises 3573 data points representing daily cryptocurrency network data.

2.4 Energy Market Data

The energy market data is collected from several sources. The average price of electricity in the United States, the net generation of electricity of all sectors in the United States, and the total electricity consumption of all sectors in the United States is obtained from the U.S. Energy Information Administration. This dataset encompasses a time period from January 2001 to November 2022 and comprises 263 data points representing monthly U.S. electricity data.

Owing to limitations in data availability, we utilize a proxy for the mean price of electricity in China, specifically the residential price index. The National Bureau of Statistics of China emphasizes the cost of electricity and other utilities within this index, rendering it a suitable proxy for this measurement. The residential price index for China is obtained from the National Bureau of Statistics of China.

The price-index is formulated as a growth measure relative to the previous month, where the previous month is set constant at 100. To create a cumulative monthly price index, we apply the following transformation

$$P_t^{CN} = \prod_{s=0}^{t} \frac{PI_s^{CN}}{100}, \tag{1}$$

where P_t^{CN} is the cumulative monthly price index value at time t and PI_s^{CN} is the price index value relative to the previous month (kept constant at 100) at time s.

The dataset spans from January 2003 to October 2022 and consists of 238 data points, which represent monthly electricity data for China.

2.5 Reddit Data

The Reddit post data is collected from Pushshift (Baumgartner et al., 2020). Pushshift is a data service platform that provides access to social media data, particularly focusing on Reddit data since 2015. Initially introduced by Jason Baumgartner, Pushshift later joined forces with the National Contagion Research Institute (NCRI). The data includes titles, text, links, subreddits, scores, the number of comments, and various other information related to Reddit posts accessible via the Reddit API.

For this research, we downloaded all the posts from the following subreddits: *r/CryptoCurrency*, *r/Bitcoin*, *r/Altcoin*, *r/CryptoMarkets*, *r/CryptoTechnology*, *r/CryptoMoonShots*, *r/Ethereum*, *r/Ripple*, *r/WallStreetBetsCrypto*, *r/ethtrader*, *r/NFT*, *r/bitcoinbeginners*, *r/CryptoCurrencies*, *r/CryptoCurrencyTrading*, *r/Crypto_General*, *r/Binance*, *r/CoinBase*, and *r/dogecoin*. The dataset covers a time period from January 1, 2013, to March 1, 2023, and consists of 6,912,239 data points representing individual Reddit posts.

3 Methodology

The methodology for this paper is divided into four sub-sections. Section 3.1 presents the methods used to compute the various factors employed throughout this paper. Section 3.2 outlines the various cryptocurrency-specific characteristics we test for their effect on cryptocurrency returns. Section 3.3 describes the range of return predictors we use for analysing the cross-section of cryptocurrency returns. Section 3.4 details the different factor models we estimate for the various return predictors.

3.1 Factors

Throughout this paper we use the following factors for our analysis: the market factor (MKT), the size factor (SMB), and the momentum factor (MOM). All subsequent factor models are based on these three factors.

We define the excess return of cryptocurrency s at time t as:

$$r_{s,t} = R_{s,t} - R_{f,t}, \tag{2}$$

where $R_{s,t}$ is the return of cryptocurrency s at time t, and $R_{f,t}$ is the risk-free rate at time t.

The market factor, originating from the CAPM (Capital Asset Pricing Model) published in Lintner (1975), Mossin (1966), Sharpe (1964) and Treynor (1961), represents the market risk premium or the market return. The market risk premium is the additional return that investors expect to receive for holding risky assets over risk-free assets. In essence, the market factor captures the influence of overall market movements on the returns of individual assets or portfolios. Let $t_1, t_2, \ldots, t_k$ represent the end of week indices (e.g., Sundays), then for each day t, we define:

$$\mathrm{MKT}_t = \sum_{s=1}^{n} w_{s,t'} r_{s,t}, \tag{3}$$

where t' is the most recent end of week index preceding t, i.e., $t' = \max\{t_i \mid t_i < t\}$, n denotes the total number of cryptocurrencies in the market, $w_{s,t'}$ is the weight of cryptocurrency s in the market portfolio at the end of the previous week, calculated as:

$$w_{s,t'} = \frac{MC_{s,t'}}{\sum_{j=1}^{n} MC_{j,t'}}, \tag{4}$$

where $MC_{s,t'}$ is the market cap of cryptocurrency s at the end of the previous week, and $r_{s,t}$ is the excess return of cryptocurrency s at time t.

The size factor, first introduced by Fama and French in their three-factor model (Fama & French, 1993), captures the return difference between small-cap stocks (companies with lower market capitalization) and large-cap stocks (companies with higher market capitalization). The size factor is intended to explain the observed anomaly that small-cap stocks have historically outperformed large-cap stocks on average, even after accounting for their increased risk as measured by the market factor in the CAPM model. For each day t, we define:

$$\mathrm{SMB}_t = \sum_{s \in S_{t'}} w_{s,t'} r_{s,t} - \sum_{s \in B_{t'}} w_{s,t'} r_{s,t}, \tag{5}$$

where, t' is the most recent end of week index preceding t, i.e., $t' = \max\{t_i \mid t_i < t\}$, $S_{t'}$ contains the lowest 30% of cryptocurrencies in terms of market cap at the end of the previous week, $B_{t'}$ contains the highest 70% of cryptocurrencies in terms of market cap at the end of the previous week, $w_{s,t'}$ is the weight of cryptocurrency s in the respective small-cap or large-cap value-weighted portfolio at the end of the previous week, calculated as:

$$w_{s,t'} = \frac{MC_{s,t'}}{\sum_{j \in G_{t'}} MC_{j,t'}}, \tag{6}$$

where $MC_{s,t'}$ is the market cap of cryptocurrency s at the end of the previous week, $G_{t'}$ represents the specific group, either $S_{t'}$ or $B_{t'}$, to which cryptocurrency s belongs at the end of the previous week, and $r_{s,t}$ is the excess return of cryptocurrency s at time t.

The momentum factor, first introduced by Carhart in (Carhart, 1997), captures the return difference between past winners (assets with high past returns) and past losers (assets with low past returns). The momentum factor is intended to explain the observed anomaly that assets with strong past performance continue to outperform those with weak past performance in the short term, which is not explained by either the market or size factors. For each day t, we define:

$$\mathrm{MOM}_t = \sum_{s \in W_{t'}} w_{s,t'} r_{s,t} - \sum_{s \in L_{t'}} w_{s,t'} r_{s,t}, \tag{7}$$

where t' is the most recent end of week index preceding t, i.e., $t' = \max\{t_i \mid t_i < t\}$, $W_{t'}$ contains the top 70% of cryptocurrencies in terms of 3-week aggregate returns at the end of the previous week, $L_{t'}$ contains the lowest 30% of cryptocurrencies in terms of 3-week aggregate returns at the end of the previous week, $w_{s,t'}$ is the weight of cryptocurrency s in the respective high-momentum or low-momentum value-weighted portfolio at the end of the previous week, calculated as:

$$w_{s,t'} = \frac{MC_{s,t'}}{\sum_{j \in H_{t'}} MC_{j,t'}}, \tag{8}$$

where $MC_{s,t'}$ is the market cap of cryptocurrency s at the end of the previous week, $H_{t'}$ represents the specific group, either $W_{t'}$ or $L_{t'}$, to which cryptocurrency s belongs at the end of the previous week, and $r_{s,t}$ is the excess return of cryptocurrency s at time t.

3.2 Cryptocurrency Characteristics

In the following sections, we define sets of network, production, and sentiment characteristics. Additionally, we also define the first two principal components of the network, production, and sentiment characteristics as additional characteristics.

During our analysis, we explore the relationships between these characteristics. We employ the Pearson-correlation matrix to examine the dependencies amongst the various variables. This technique also assists in comprehending the structure of the principal component and identifying which variables contribute the most variance.

To assess the explanatory power of these cryptocurrency-specific characteristics on coin market returns, we use standard linear regression. We regress the coin market returns on the characteristics.

$$\text{MKT}_t = \beta_i C_{i,t} + \varepsilon_i, \tag{9}$$

where $C_{i,t}$ denotes the value of characteristic i at time t. We report the regression coefficient β (including the significance level), the t-statistic, and the model's explanatory power, represented by the R^2 value.

Regarding the analysis of network characteristics, it is of academic interest to examine the impact of market returns on network factors. To investigate this, we perform the reverse regression, regressing the cumulative growth rates for the network characteristics on the lagged market returns and a constant.

$$C_{i,t}^{(h)} = \beta_{const} + \beta_{MKT}\text{MKT}_{t-h} + \varepsilon_i \tag{10}$$

where $C_{i,t}^{(h)}$ signifies the cumulative growth rate of characteristic i at time t and horizon h, both expressed in weeks. We conduct this regression with $1 \leq h \leq 8$.

In the analysis of sentiment characteristics, we test whether lagged values of the characteristics influence the cumulative market returns. For this investigation, we execute the following regression:

$$\text{MKT}_t^{(h)} = \beta_i C_{i,t-h} + \varepsilon_i, \tag{11}$$

where h is the horizon, expressed in weeks. We conduct this regression also with $1 \leq h \leq 8$.

3.2.1 Network Characteristics

The theoretical literature on cryptocurrency emphasizes the role of several network characteristics in determining the valuation of cryptocurrencies (Biais et al., 2018; Cong & He, 2019; Pagnotta & Buraschi, 2018; Sockin & Xiong, 2020). A critical aspect of this relationship is the network effect, which emerges from the adoption of cryptocurrencies by users. This network effect could potentially assume a central role in influencing the valuation of these digital assets.

As users adopt cryptocurrencies, they generate positive network externalities, subsequently affecting asset prices. Hence, the extent of user adoption has a direct impact on the price dynamics of cryptocurrencies. It can be inferred that fluctuations in user adoption rates within the cryptocurrency network might contribute to the observed variations in cryptocurrency prices. This is particularly relevant for cryptocurrencies with built-in anti-inflation mechanisms, which limit the number of outstanding coins, leading to increased demand and subsequent price increases.

Another potential network effect that could significantly influence cryptocurrency prices is the mining reward. This reward constitutes the compensation miners receive for verifying transactions on the blockchain. Since the mining reward differs for each cryptocurrency and is subject to varying magnitudes, the hash rate serves as a useful proxy for this effect. The hash rate represents the computational power available to miners. An increase in the hash rate results in increased difficulty, leading to a lower pay-out for mining cryptocurrency transaction blocks. This diminished pay-out may prompt miners to shift their focus to mining alternative cryptocurrencies where the pay-out might be more substantial, both in terms of the cryptocurrency itself and its associated price.

In terms of network characteristics, we consider:

- the percentage growth in active addresses,
- the percentage growth in daily transactions,
- the percentage growth in daily payments, and
- the percentage growth in hash rate.

Furthermore, we employ the first two principal components derived from these four characteristics to analyse the impact of these network characteristics-derived features on cryptocurrency prices.

3.2.2 Production Characteristics

Numerous studies have posited that mining costs are crucial for the infrastructure and security of cryptocurrencies (Abadi & Brunnermeier, 2018; Cong, He & Li, 2021; Sockin & Xiong, 2020). In particular, Sockin and Xiong (2020) demonstrate that within a general equilibrium model featuring cryptocurrency production, the prices of cryptocurrencies are closely related to the marginal cost of mining. The marginal cost of mining primarily comprises two components: the price of electricity and the price of computational power.

The price of electricity is readily available for various countries. In this paper, we construct a proxy for the price of electricity by utilizing both the production and consumption of electricity, as well as the price of electricity. Given that the majority of mining operations are geographically situated in China and the United States, we utilize the statistics from these countries for our analysis, in accordance with Liu and Tsyvinski (2021).

Finding a suitable proxy for the cost of computational power proves to be rather more challenging. Owing to the rapid development of the cryptocurrency mining industry, characterized by frequent releases of faster and more energy-efficient compute units from companies such as NVIDIA, AMD, and Intel, devising a proxy for this production effect is no easy task. Papers such as Liu and Tsyvinski (2021) employ the price of Antminer equipment as a proxy for the price of computational power. In our opinion, this does not represent a fair quantification for the cost of computational power, as this equipment varies greatly in capability between generations and is largely no longer in use by large mining operations.

Analysing the different prices across vendors and markets for mining equipment, weighted by their hash-rate, proved to be too time-consuming and beyond the scope of this paper. This relationship is better investigated by considering only the hash-rate, as discussed in the previous section.

With respect to production characteristics, we include the following variables:

- the percentage change in the monthly price of electricity in the United States,
- the percentage change in total monthly electricity generation within the United States,
- the percentage change in total monthly energy consumption in the United States,
- the percentage change in the cumulative monthly residential price index value for China, and
- the percentage change in total monthly energy production in China.

As with the network characteristics, we employ the first two principal components derived from these five production-related characteristics.

3.2.3 Sentiment Characteristics

In the case of cryptocurrencies, the absence of fundamental information, such as earnings, dividends, and other types of cash flows, renders the price discovery process exceedingly complicated. As argued by Cheah and Fry (2015), the fundamental value of Bitcoin is zero, rendering financial models based on well-defined fundamental values inapplicable to cryptocurrencies. Theoretically speaking, when the market is driven by individual investors with elevated risk preferences and in the presence of limits to arbitrage (e.g. a short-sale constraint), sentiment-driven noise traders may significantly influence the price discovery process (De Long et al., 1990). Specifically, Chen et al. (2019) demonstrate that employing a simple bag-of-words algorithm coupled with a hand-picked lexicon to ascertain the sentiment of StockTwits or Reddit posts enables the utilization of investor sentiment to predict future cryptocurrency returns. However, this approach has several drawbacks, as it cannot capture the context of the text and is limited to the English language. Moreover, it struggles to handle negations, sarcasm, and other forms of figurative language.

To enhance the accuracy of the analysed sentiment, we adopt a more sophisticated method for determining sentiment. In this paper, we make use of the XLM-T model, as presented in Barbieri et al. (2021), which is a natural language model specialized in multilingual sentiment analysis of social media data. This model is derived from the XLM-R model (Conneau et al., 2019), which itself is a variation within the BERT family of natural language models (Devlin et al., 2018). The XLM-T model is designed for improved performance in sentiment analysis tasks, particularly those involving social media data, through the incorporation of advanced pre-processing techniques and optimizations. It is particularly adept at handling challenges posed by social media posts, such as informal language, abbreviations, and code-switching.

The training of these models utilizes the masked language model (MLM) objective, which is a variant of the "Cloze Procedure" introduced by Taylor (1953). The MLM objective aims to predict masked tokens within a given input text. To achieve this, a corpus of text is first divided into (sub-)words, or tokens, using a tokenizer. Subsequently, a certain percentage of tokens, typically around 15%, are replaced with a special mask token. The model processes the masked input sequence and attempts to predict the masked tokens based on the context provided by the surrounding unmasked tokens. This results in a pre-trained model that can be further specialized for specific tasks, such as sentiment analysis.

In this paper, we use the XLM-T model to perform sentiment analysis on the Reddit post titles. We restrict ourselves to titles, and not post bodies, for two reasons. First, most Reddit posts are not text posts, but instead contain a link to an article or image. Extending the current dataset beyond just the Reddit posts to also include the content of possible backlinks proved to be too time-consuming and beyond the scope of this paper. Since the dataset contains more than 6.9 million posts and running deep learning models requires tremendous computing resources, this limitation was deemed adequate.

Secondly, running the sentiment analysis on only the posts that did have text bodies also resulted in a significant computational burden. For the computation of the sentiment based on just the title, additional computing resources already had to be acquired.

The classifier model produces the probabilities that a post is positive, neutral, and negative. To construct a sentiment score for a post, ranging from 0 to 1, we apply the following formula:

$$S_k = \frac{\mathbb{P}(S = \text{positive}) - \mathbb{P}(S = \text{negative}) + 1}{2} \tag{12}$$

where S_k is the sentiment measure for Reddit post k.

Based on the 6.9 million sentiment entries, we construct the following sentiment characteristics:

- the percentage change in the amount of posts per day,
- the percentage change in the amount of positive posts per day,
- the percentage change in the amount of negative posts per day,
- the percentage change in the amount of neutral posts per day,
- the average sentiment measure per day,

- the post score weighted average sentiment measure per day, and
- the standard deviation of the sentiment measure per day.

3.3 Return Predictors

In the subsequent sections, we describe a set of return predictors based on size, momentum, volume, volatility, and projected cryptocurrency characteristics which are utilized to construct quintile portfolios. These former four return predictors are derived from market information, as presented by Chen and Zimmermann (2020) and Feng et al. (2020). The latter is constructed based on projections of cryptocurrency market characteristics. There are two primary reasons for this approach.

First, cryptocurrencies lack inherent value, making traditional financial models based on fundamental values inapplicable. Consequently, our analysis is restricted to potential return predictors within our sole cross-sectional dataset, specifically market data.

Second, these return predictors encompass all the hypothesized excess return drivers identified in current literature that have proven significant in other asset classes. This serves as a suitable generalization for abnormal returns and effectively neutralizes potential outliers in the data.

Each week, we sort individual cryptocurrencies into quintile portfolios based on the value of a given return predictor. Subsequently, we construct the value-weighted portfolio returns for these quintiles and analyse the significance of their mean using a single-sample Student-t test. Following this, we perform factor model analysis, as explained in section 3.4, to examine the portfolio returns and their factor-loadings.

3.3.1 Size-Related Predictors

Size-related return predictors aim to capture potential return anomalies associated with the size of an asset, commonly represented by its market cap. Size factors are amongst the first risk factors proposed in the literature, dating back to Fama and French (1993). They represent the anomaly that small market cap assets tend to outperform large market assets. We define the following size-related return predictors:

Banz (1981) define the logarithm of the market cap on the last day of trading before portfolio formation as a possible return predictor:

$$\text{MCAP}_{s,t} = \log(MC_{s,t'}), \tag{13}$$

where $MC_{s,t'}$ is the market cap of cryptocurrency s at the end of the previous week.

Another considered return predictor, proposed in Miller and Scholes (1982), is the price on the last day of trading before portfolio formation:

$$\text{PRC}_{s,t} = P_{s,t'}, \tag{14}$$

where $P_{s,t'}$ is the price of cryptocurrency s at the end of the previous week.

Additionally, we consider the maximum price in the week before portfolio formation, proposed by George and Hwang (2004):

$$\mathrm{MAXDPRC}_{s,t} = \max(\mathcal{P}_{s,t'}), \tag{15}$$

where $\mathcal{P}_{s,t'} = \{P_{s,t'-h} \mid 0 \leq h \leq 6\}$ and h is expressed in days.

Finally, we also consider the number of days the cryptocurrency is listed on Coinmarketcap.com, denoted by $\mathrm{AGE}_{s,t}$, as a return predictors, put forth by Barry and Brown (1984).

3.3.2 Momentum-Related Predictors

Momentum-related return predictors generally represent compounded returns over some horizon. De Bondt and Thaler (1985) and Jegadeesh and Titman (1993) argue that buying past winners and selling past losers generates abnormal returns.

For this research, we use the past one-, two-, three-, four-, one-to-four, eight-, 16-, 50-, 100-week returns, proposed by De Bondt and Thaler (1985) and Jegadeesh and Titman (1993):

$$r_{s,t}^{(h,k)} = \left(\prod_{j=1+7k}^{7h} (1 + r_{s,t-j}) \right) - 1, \tag{16}$$

for all $(h, k) \in \{(1, 0), (2, 0), (3, 0), (4, 0), (4, 1), (8, 0), (16, 0), (50, 0), (100, 0)\}$.

3.3.3 Volume Related Predictors

Volume-related return predictors aim to capture potential return anomalies associated with the volume of an asset. These return predictors are based on the assumption that assets with higher trading volumes are more liquid and thus less risky. Another aspect of volume-related return predictors is the assumption that assets with higher trading volumes are more likely to be traded by informed investors. Volume can also be a proxy for momentum, as assets with higher trading volumes are actively traded and thus more likely to be subject to momentum trading. Finally, volume can also gauge investor attention and sentiment, as assets with higher trading volumes are more likely to be discussed in the media.

Volume is generally expressed as the number of shares traded per day. In the case for cryptocurrencies, this is expressed as the number of coins traded per day.

We use the return predictors proposed by Chordia et al. (2001). These include the logarithm of the average daily trading volume before portfolio formation, which can proxy for liquidity and investor attention:

$$\mathrm{VOL}_{s,t} = \log\left(\frac{1}{7} \sum_{j=0}^{6} V_{s,t'-j} \right). \tag{17}$$

where $V_{s,t'}$ is the trading volume of cryptocurrency s at the end of the previous week.

Additionally, we use the logarithm of the average daily trading volume times the price before portfolio formation, which can proxy for the interaction between liquidity and price movements:

$$\text{PRCVOL}_{s,t} = \log\left(\frac{1}{7}\sum_{j=0}^{6}\left(V_{s,t'-j} \cdot P_{s,t'-j}\right)\right). \tag{18}$$

Finally, we use the logarithm of the average daily trading volume times the price scaled by market capitalization before portfolio formation, which can help identify the interaction between liquidity, price movements and size:

$$\text{VOLSCALED}_{s,t} = \log\left(\frac{1}{7}\sum_{j=0}^{6}\frac{V_{s,t'-j} \cdot P_{s,t'-j}}{MC_{s,t'-j}}\right). \tag{19}$$

3.3.4 Volatility Related Predictors

Volatility-related return predictors aim to capture potential return anomalies associated with the volatility of an asset. These return predictors are based on the assumption that assets with higher volatility are more risky and thus have higher expected returns. They attempt to capture the risk associated with the price fluctuations of an asset and the impact of this risk on expected returns.

Volatility can be subdivided into four categories: historical volatility, implied volatility, idiosyncratic volatility, and realized volatility. Historical volatility is the standard deviation of the asset's returns over a given period. Implied volatility is the volatility implied by the market price of an option. It represents the market's expectation of the asset's future volatility. Idiosyncratic volatility is the volatility of an asset's returns that is not explained by the market. Realized volatility is the volatility of an asset's returns over a given period, with intraday price movements taken into account. It captures the actual risk experienced by investors over a given period.

Since only a very small subset of exchanges offer options on cryptocurrencies, implied volatility is not a viable option for this research. Furthermore, since we are using daily data, realized volatility is not a viable option either. Therefore, we only consider historical volatility and idiosyncratic volatility.

Based on the work of Ang et al. (2006) and Fama and MacBeth (1973), we employ the following model and its derived coefficients as return predictors. Firstly, we estimate the subsequent model based on the previous 365 days before portfolio formation:

$$r_{s,t} = \alpha + \beta_{MKT} MKT_t + \varepsilon_{s,t}. \tag{20}$$

We then define the following volatility return predictors:

$$\text{BETA}_{s,t} = \beta_{MKT} \tag{21}$$

$$\text{BETA2}_{s,t} = (\beta_{MKT})^2 \tag{22}$$

$$\text{IDIOVOL}_{s,t} = \text{Var}(\varepsilon_{s,t}) \tag{23}$$

Furthermore, we define the following volatility return predictors, proposed by Ang et al. (2006) and Bali et al. (2011), specifically the standard deviation and maximum of the daily returns in the week before portfolio formation:

$$\mathrm{RETVOL}_{s,t} = \sqrt{\mathrm{Var}(\mathcal{R}_{s,t'})} \tag{24}$$

$$\mathrm{MAXRET}_{s,t} = \max(\mathcal{R}_{s,t'}) \tag{25}$$

where $\mathcal{R}_{s,t} = \{r_{s,t'-j} \mid 0 \leq j \leq 6\}$.

Based on the work of Hou and Moskowitz (2005), we also consider the return predictors derived from the following model:

$$r_{s,t} = \alpha + \beta_{MKT}MKT_t + \beta_{MKT_{-1}}MKT_{t-1} + \beta_{MKT_{-2}}MKT_{t-2} + \varepsilon_{s,t}. \tag{26}$$

This model is estimated based on the previous 365 days before portfolio formation as well. We define $DELAY_{s,t}$ as the difference in R^2 between eq. (20) and eq. (26).

Finally, we examine the return predictors proposed by Amihud (2002) and Chordia et al. (2001). They define the logarithm of the standard deviation of the price multiplied by the volume and the average absolute daily return divided by the price multiplied by the volume in the week before portfolio formation as potential return predictors:

$$\mathrm{STDPRCVOL}_{s,t} = \log\left(\sqrt{\mathrm{Var}(\mathcal{PV}_{s,t'})}\right) \tag{27}$$

$$\mathrm{DAMIHUD}_{s,t} = \frac{1}{7}\sum_{j=0}^{6} \frac{\left|r_{s,t'-j}\right|}{P_{s,t'-j} \cdot V_{s,t'-j}} \tag{28}$$

where $\mathcal{PV}_{s,t'} = \{P_{s,t'-j} \cdot V_{s,t'-j} \mid 0 \leq j \leq 6\}$.

3.3.5 Projected Cryptocurrency Characteristics

We explore the use of projected cryptocurrency characteristics as return predictors, as they represent essential aspects of cryptocurrencies.

We project the characteristics of cryptocurrencies onto their individual returns using the following model:

$$r_{s,t} = \alpha + \beta_{C_i}C_{i,t} + \varepsilon_{s,t}. \tag{29}$$

where $C_{i,t}$ is the characteristic i of cryptocurrency s at time t. We then subsequently use these projection betas (β_{C_i}) as potential return predictors.

For network characteristics, we aim to create proxies for network size (based on the number of active addresses) and network activity (based on the number of transactions and payments). Given that these characteristics are derived from Bitcoin, widely considered the most used and accepted cryptocurrency, we assume that Bitcoin's network size and activity serve as valid representations of a large, active, and widely accepted cryptocurrency. These projections should, therefore, favor subsequent large, active, and well-accepted cryptocurrencies.

Regarding production characteristics, we attempt to create a proxy for the distinction between Proof-of-Work (PoW) and Proof-of-Stake (PoS) cryptocurrencies. PoW cryptocurrencies are heavily dependent on electricity prices, so they should exhibit a larger regression coefficient compared to cryptocurrencies that do not consume significant amounts of electricity.

For the sentiment characteristics, we attempt to create a proxy for "meme" coins, a term often used to describe cryptocurrencies that are exploited in pump-and-dump schemes. "Meme" coins experience rapid price growth as they gain popularity and garner significant attention, often driven by individuals with substantial social media influence.

These influencers typically accumulate a significant position in a relatively obscure cryptocurrency and subsequently promote it heavily on social media platforms, fueling the price surge. They often predict a specific date when the price will soar and sell their holdings just before that date, leaving the latecomers to bear the brunt of the price collapse. This tactic resembles a modern-day Ponzi scheme.

To capture this phenomenon, we employ sentiment and post-count as the regressors, as these metrics are closely associated with the underlying dynamics of such schemes. By analyzing sentiment and post-count data, we aim to identify and better understand the impact of "meme" coins on the cryptocurrency market.

3.4 Factor Models

Factor models play a crucial role in finance, as they help describe and predict asset return behaviors. These models break down an asset's return into systematic and idiosyncratic components, which help investors identify key drivers of asset returns. By understanding the sources of risk and return in their portfolios, investors can make more informed investment decisions.

We employ the following factor models, utilizing the factors outlined in section 3.1:

$$(1) \quad r_{p,t} = \alpha + \beta_{\mathrm{MKT}}\mathrm{MKT}_t + \varepsilon_{s,t} \tag{30}$$

$$(2) \quad r_{p,t} = \alpha + \beta_{\mathrm{MKT}}\mathrm{MKT}_t + \beta_{\mathrm{SMB}}\mathrm{SMB}_t + \varepsilon_{s,t} \tag{31}$$

$$(3) \quad r_{p,t} = \alpha + \beta_{\mathrm{MKT}}\mathrm{MKT}_t + \beta_{\mathrm{MOM}}\mathrm{MOM}_t + \varepsilon_{s,t} \tag{32}$$

$$(4) \quad r_{p,t} = \alpha + \beta_{\mathrm{MKT}}\mathrm{MKT}_t + \beta_{\mathrm{SMB}}\mathrm{SMB}_t + \beta_{\mathrm{MOM}}\mathrm{MOM}_t + \varepsilon_{s,t} \tag{33}$$

where $r_{p,t}$ represents the return of portfolio p at time t, β_j denotes the factor loading for factor j, and $\varepsilon_{s,t}$ signifies the idiosyncratic return of cryptocurrency s at time t.

The constant, α, can be interpreted as the portfolio's unique return, which is not explained by its exposure to the factors. A positive alpha indicates that the portfolio has generated returns beyond what can be explained by the factors, while a negative alpha suggests that the portfolio's returns fall short of what can be explained by the factors. This discrepancy could indicate that the factor model is missing a relevant factor that accounts for the portfolio's returns and the associated underlying risk.

The factor loadings, β_j, represent the portfolio its sensitivity to factor j. A positive β_j implies that the portfolio returns increase as factor j rises, while a negative β_j suggests that the portfolio returns decrease when factor j increases. In essence, a higher beta for a specific factor indicates that the asset's returns are more sensitive to changes in that factor, thus exposing the portfolio to the risks associated with that factor.

We utilize the framework proposed by Beck et al. (2023), for constructing the standard errors used for factor model inference. They argue that stock and factor returns exhibit time-varying volatility. The same can be argued for cryptocurrency returns, as they are often characterized by periods of high and low volatility. This time-varying volatility results in conditional heteroskedasticity of the error terms, which violates the OLS assumption of homoskedasticity. This heteroskedasticity needs to be accounted for when conducting statistical tests in order to avoid false inferences.

This phenomenon does not pose a problem for the estimation of the factor loadings, since the OLS estimator is still consistent in the presence of heteroskedasticity. However, the estimator is no longer efficient in the sense that it does not achieve the smallest asymptotic covariance matrix.

The standard approach to address this issue is to use a heteroskedasticity-consistent covariance matrix estimator (HCCME) which guarantees asymptotically valid inference under conditional heteroskedasticity of an unknown form. Examples of these estimators are the HC0 estimator, better known as the White estimator (White, 1980), the HC1-HC3 estimators (MacKinnon & White, 1985), and the HC4 estimator (Cribari-Neto, 2004). For a thorough overview of the different methods, we refer to MacKinnon (2013). For this research, as recommended by Romano and Wolf (2017), we use the HC3 estimator.

Romano and Wolf (2017) propose an alternative to OLS, suggesting the use of weighted least squares (WLS) or adaptive least squares (ALS) for estimating factor loadings. These methods employ a skedastic function, which maps the factor into the conditional variance of the error term:

$$\upsilon(x_t) = \mathbb{E}\left[\varepsilon_t^2 \mid x_t\right], \tag{34}$$

where $x_t = (1, \mathrm{MKT}_t, \mathrm{SMB}_t, \mathrm{MOM}_t)$, or a subset depending on the specific model, as specified in eqs. (30) to (33). In practice, this function is unknown, but it can be estimated using observed data, resulting in $\hat{\upsilon}(\cdot)$.

The WLS method weights the data by the reciprocal of the estimated conditional standard deviation of the error terms, $\sqrt{\hat{\upsilon}(x_t)}$, before applying OLS. This results in the following regression model:

$$\frac{r_{p,t}}{\sqrt{\hat{\upsilon}(x_t)}} = \beta' \frac{x_t}{\sqrt{\hat{\upsilon}(x_t)}} + \varepsilon_t^* \quad \text{with} \quad \varepsilon_t^* = \frac{\varepsilon_t}{\sqrt{\hat{\upsilon}(x_t)}}. \tag{35}$$

Allowing for the possibility that $\hat{\upsilon}(\cdot)$ might not be a consistent estimator of the true skedastic function, it becomes necessary to apply heteroskedasticity-consistent standard errors during model inference. This should be done in conjunction with WLS or ALS.

For this research, we use the following skedastic function:

$$\upsilon_\theta(x_t) = \exp(\nu + \gamma_2|x_{2,t}| + \cdots + \gamma_K|x_{K,t}|) \quad \text{with} \quad \theta = (\nu, \gamma_2, \cdots, \gamma_K)'. \tag{36}$$

In order to estimate θ, we use a two-step procedure. First, we estimate the standard OLS model:

$$r_{p,t} = \beta' x_t + \varepsilon_t. \tag{37}$$

Subsequently, we estimate the following model:

$$\log\left[\max\left(\delta^2, 10^q \hat{\varepsilon}_t^2\right)\right] = 10^q \nu + \gamma_2 |10^q x_{2,t}| + \cdots + \gamma_K |10^q x_{K,t}| + u_t, \tag{38}$$

where we choose $q \in [0, 1, 2, \ldots]$ as the smallest non-negative integer such that fraction of truncations on the left-hand side is at most 5% with the lower bound $\delta^2 = 0.01$.

Next, we test the joint significance of the heteroskedasticity coefficients as computed with eq. (38). We test the null hypothesis $H_0 : \gamma_2 = \cdots = \gamma_K = 0$. If this hypothesis is rejected at a significance level of 0.1, then the ALS aligns with WLS; otherwise, it aligns with standard OLS. In this research, we use the ALS method for estimating the factor loadings, as it is more robust to potential heteroskedasticity in the error terms.

4 Results

The results of this paper are organized into four sub-sections. Section 4.1 provides descriptive statistics for the market returns of cryptocurrencies, and Bitcoin, Ethereum, and Ripple (XRP). Section 4.2 details the analysis of cryptocurrency-specific characteristics. Section 4.3 offers insights into the cross-sectional examination of return predictors. Lastly, section 4.4 explores possible mechanisms that could account for the observed return predictability.

4.1 Descriptive Statistics

In this section, we investigate the time-series properties of cryptocurrency market returns. The cryptocurrency market returns are calculated using all cryptocurrencies meeting the inclusion criteria specified in section 2.2. The cryptocurrency market returns are calculated using the procedure outlined in section 3.1.

Figure 1: This figure plots the distribution of daily, weekly, and monthly cryptocurrency market returns and log returns.

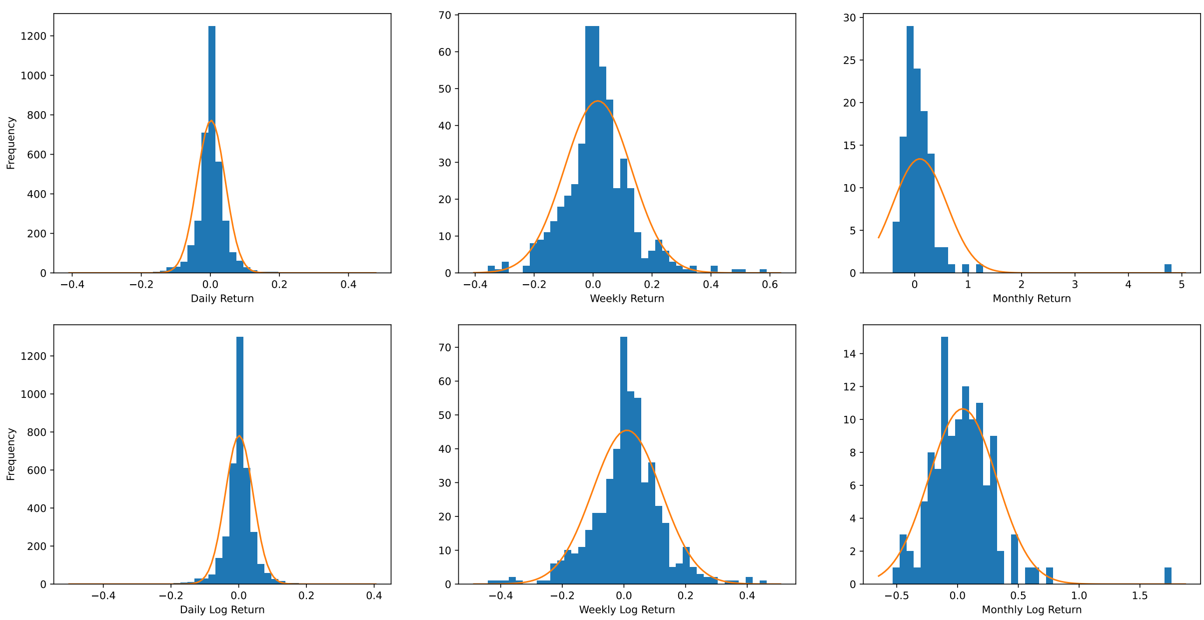

Figure 1 displays the return distribution for the cryptocurrency market index, considering both standard returns and log-returns. For comparative purposes, a normal distribution is included in each figure. Both the weekly and monthly data sets reveal significant extreme values in the tails of the distribution, with the majority of returns centered around zero.

Figure 2: This figure plots the cryptocurrency market returns, Bitcoin returns, Ethereum returns, and Ripple returns. The figure shows the value of investment over time for one dollar investment at the starting point of the graphs. The Bitcoin graph starts on the 5th of May 2013, the Ethereum graph on the 14th of August 2015, and the Ripple graph on the 11th of August 2013

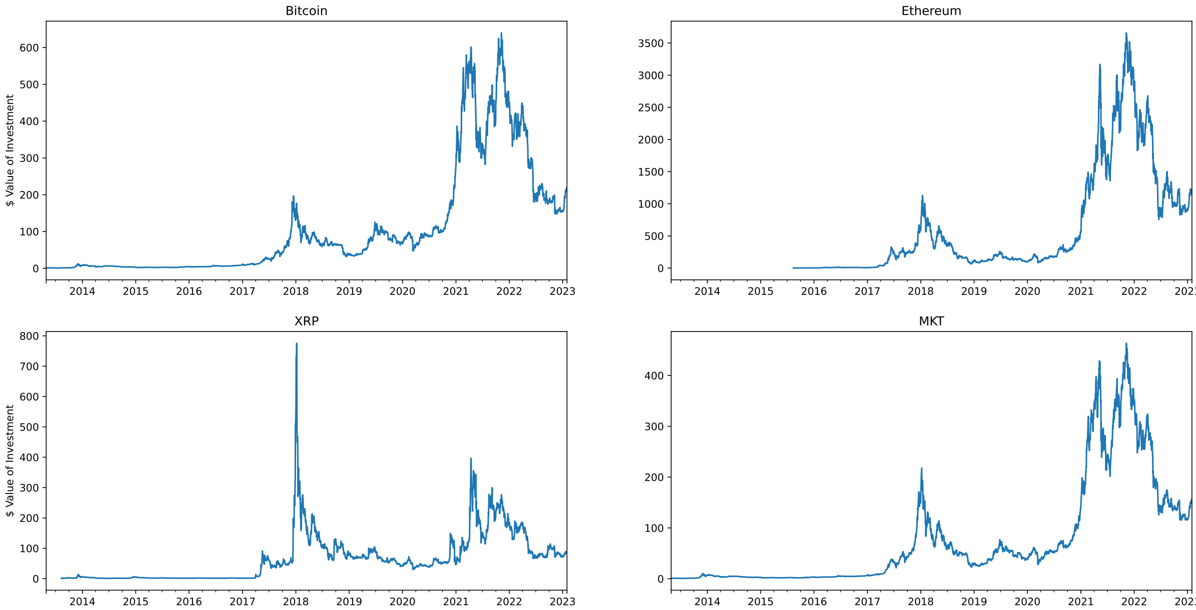

Figure 2 compares the price trajectories of Bitcoin, Ethereum, Ripple (XRP), and the cryptocurrency market index side by side. A notable degree of co-movement can be observed among these three major cryptocurrencies. Furthermore, the cryptocurrency market index has exhibited impressive rates of return over the past nine years. At the peak of the cryptocurrency "boom" in 2022, a one-dollar investment made in early 2013 in the cryptocurrency market index could have potentially appreciated to $400, significantly surpassing traditional investments such as those in the S&P500. However, this estimate does not consider potential transaction costs, the risks associated with cyberattacks on cryptocurrency exchanges, or the availability of such exchanges (liquidity). Even a modest one-dollar investment in individual cryptocurrencies like Bitcoin, Ethereum, and Ripple could have generated considerable returns.

Table 1 displays descriptive statistics for the cryptocurrency market returns, specifically Bitcoin, Ethereum, and Ripple (XRP), across daily, weekly, and monthly frequencies. A noteworthy observation is the high mean and standard deviation values associated with these cryptocurrencies. For the daily frequency, the mean return is 0.23% and the standard deviation is 4.09%; for the weekly frequency, these values are 1.62% and 11.37%; and for the monthly frequency, they are 9.68% and 50.63%. The mean stock returns for daily, weekly, and monthly frequencies are 0.05%, 0.39%, and 1.74%, respectively. Cryptocurrency market returns are significantly higher than stock returns, sometimes by nearly 500%.

The Sharpe ratios for cryptocurrency market returns are 0.06 at the daily frequency, 0.14 at the weekly frequency, and 0.19 at the monthly frequency. In contrast, the Sharpe ratios for stock

Table 1: This table documents the daily, weekly, and monthly summary statistics of the coin market returns (MKT) and compares them with the returns for Bitcoin, Ethereum, Ripple (XRP) and the stock market. The mean, standard deviation, t-statistic, Sharpe-ratio, Skewness, Kurtosis, and the percentage of positive observations are reported.

		Mean	SD	t-Stat	Sharpe	Skewness	Kurtosis	% > 0
Daily	MKT	0.23%	4.09%	3.29	0.06	-0.03	10.66	54.88
	Bitcoin	0.24%	4.11%	3.42	0.06	0.24	10.19	53.20
	Ethereum	0.43%	5.87%	4.37	0.07	0.64	6.85	38.67
	XRP	0.38%	7.68%	2.97	0.05	5.59	101.69	46.30
	Stock	0.05%	1.04%	9.49	0.05	-0.09	15.56	55.62
Weekly	MKT	1.62%	11.37%	3.22	0.14	0.57	3.05	55.58
	Bitcoin	1.61%	11.26%	3.23	0.14	0.60	2.77	55.97
	Ethereum	3.42%	18.38%	4.21	0.19	1.83	7.69	41.49
	XRP	3.75%	36.10%	2.35	0.10	9.12	116.50	45.99
	Stock	0.39%	3.57%	7.83	0.11	-0.08	6.12	60.05
Monthly	MKT	9.68%	50.63%	2.08	0.19	7.08	64.47	53.39
	Bitcoin	9.20%	47.46%	2.12	0.19	7.20	66.90	51.69
	Ethereum	16.11%	48.89%	3.60	0.33	1.88	4.81	41.53
	XRP	21.96%	108.87%	2.20	0.20	4.96	29.74	37.29
	Stock	1.74%	7.81%	7.59	0.22	0.49	6.27	62.21

returns are 0.05 for the daily frequency, 0.11 for the weekly frequency, and 0.22 for the monthly frequency. The cryptocurrency market Sharpe ratios exceed their respective stock return counterparts at the daily and weekly frequency, but are surpassed at the monthly frequency.

We now compare the Sharpe ratios of Bitcoin, Ethereum, and Ripple (XRP) to the cryptocurrency market returns. It is important to note that the return series for Bitcoin commences on the 5th of May 2013, Ethereum on the 14th of August 2015, and Ripple on the 11th of August 2013. For Bitcoin returns, the Sharpe ratios are 0.06 at the daily frequency, 0.14 at the weekly frequency, and 0.19 at the monthly frequency. For Ethereum returns, the Sharpe ratios are 0.07 at the daily frequency, 0.19 at the weekly frequency, and 0.33 at the monthly frequency. For Ripple returns, the Sharpe ratios are 0.05 at the daily frequency, 0.10 at the weekly frequency, and 0.20 at the monthly frequency. Comparing mean returns, Ripple exhibits higher mean returns across all frequencies than the cryptocurrency market index. However, its Sharpe ratio is lower than that of the cryptocurrency market index across all frequencies, as its standard deviation is substantially higher, except for the monthly frequency. The standard deviation of Ripple is approximately 80% higher than that of the cryptocurrency market returns for the daily and monthly frequencies, and 220% higher for the weekly frequency. The only cryptocurrency that outperforms the cryptocurrency market index based on Sharpe ratio is Ethereum, at all frequencies, and Ripple, at the monthly frequency.

We note that the returns of Bitcoin, Ethereum, and Ripple all exhibit positive skewness. This is in contrast to stock market returns, with the former displaying negative skewness for both the daily and weekly frequencies, and near zero skewness for the monthly frequency. The cryptocurrency market returns exhibit positive skewness at the weekly and monthly frequencies, but negative skewness at the daily frequency. The skewness of the coin market returns is -0.03, 0.57, and 7.08 for daily, weekly, and monthly frequencies, respectively. In contrast, the skewness of stock returns is -0.09, -0.08, and 0.49 for daily, weekly, and monthly frequencies, respectively.

Examination of the individual cryptocurrencies reveals that Ripple exhibits remarkably elongated tails for positive returns in comparison to other cryptocurrencies at both daily and weekly frequencies. This implies that the magnitude and frequency of positive returns for Ripple significantly surpass that of what is typically observed within the cryptocurrency market.

The kurtosis of the coin market returns is 10.66, 3.05, and 64.47 for daily, weekly, and monthly frequencies, respectively. These figures are of a similar order of magnitude to stock returns for daily and weekly frequencies, where the kurtosis measures 15.56 and 6.12. The sole striking deviation occurs at the monthly frequency, where the stock market returns only have a kurtosis of 6.27, alluding to a substantial concentration around the mean of the distribution.

Table 2: This table documents the count and percentage of extreme events for the daily coin market index returns.

Disasters	Counts	%	Miracles	Counts	%
$<$ -5 %	262	7.35%	$>$ 5 %	273	7.66%
$<$ -10 %	69	1.93%	$>$ 10 %	57	1.60%
$<$ -20 %	6	0.17%	$>$ 20 %	3	0.08%
$<$ -30 %	1	0.03%	$>$ 30 %	2	0.06%

Drawing upon the skewness and kurtosis from table 1, the results from table 2 offer additional insights into the distribution of cryptocurrency market returns. With more remarkable results at the 5% level, there are higher occurrences of adverse outcomes at the 10% and 20% thresholds. However, the 30% threshold shows a higher number of exceptional results again. With a significant concentration around the mean, a slight left skew, and elevated kurtosis, cryptocurrency market returns are more prone to exhibit extreme values compared to stock returns.

4.2 Characteristics Analysis

In this section, we examine the impact of cryptocurrency-specific characteristics on market returns, focusing on network, production, and sentiment characteristics as detailed in section 3.2.

Our results demonstrate that specific network characteristics, such as the growth rates of active addresses, transactions, and hash rate, have an impact on coin market returns. We also note that cryptocurrency market returns have a negative influence on transaction growth between the 4- to 8-week horizon. Moreover, we find that market returns positively affect hash rate growth from the 5- to 8-week horizon. Interestingly, we find that coin market returns are not affected by production factors. We also observe that coin market returns have a positive exposure to the growth rate of the number of posts, negative posts, neutral posts, and, to a smaller degree, positive posts. Lastly, we discover that the growth in social media posts, including negative, neutral, and, to a lesser extent, positive, contribute positively to long-term cryptocurrency market returns.

4.2.1 Network Characteristics

The correlation matrix for the network factors under investigation is presented in table 3. Each of the four measures exhibits a positive correlation with the others, albeit a minute positive coef-

Table 3: This table reports the correlations of the network characteristics. The data frequency is monthly.

	Δaddress	Δtransactions	Δpayments	Δhash-rate	PC_1^{network}	PC_2^{network}
Δaddress	1.00	0.48	0.11	0.32	0.18	0.08
Δtransactions		1.00	0.21	0.40	0.24	0.10
Δpayments			1.00	0.02	0.35	-0.11
Δhash-rate				1.00	0.19	0.19
PC_1^{network}					1.00	0.65
PC_2^{network}						1.00

ficient is observed between the increase in hash-rate and the increase in payments. Interestingly, the growth in transactions and the growth in payments do not exhibit a perfect correlation.

To briefly summarize, transactions and payments possess a theoretical association, wherein a single blockchain transaction consists of multiple payments. However, this batching process is only feasible for exchanges or groups of individuals managing multiple wallets, as individual users lack the capacity or coordination to combine payments. This finding is rather surprising, considering that cryptocurrencies were initially conceived as a decentralized payment method. The correlation thus indicates that the majority of transactions on the blockchain transpire via centralized exchanges. This is not entirely unexpected since these exchanges in their turn offer simplified management of wallets, a process that formed the biggest hurdle in initial adoption of cryptocurrencies.

The first principal component exhibits correlations of 0.18, 0.24, 0.35, and 0.19 with the growth rates in addresses, transactions, payments, and hash-rate, respectively. This component appears to primarily emphasize the variation in the growth rates of payments and transactions, while only partially encapsulating the variance in the growth rates of addresses and hash-rate. Conversely, the second principal component demonstrates correlations of 0.08, 0.10, -0.11, and 0.19 with the aforementioned network attributes. This component seemingly captures the relationship between the growth in address size and the growth in hash-rate, whilst adjusting for the growth in payments.

Table 4: This table reports the regression coefficients of the coin market returns on the network characteristics. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is monthly.

	(1)	(2)	(3)	(4)	(5)	(6)
Δaddress	0.93***					
	(4.62)					
Δtransactions		0.59**				
		(2.58)				
Δpayments			0.40			
			(1.04)			
Δhash-rate				0.26**		
				(2.05)		
PC_1^{network}					0.00	
					(0.01)	
PC_2^{network}						-0.08
						(-0.25)
R^2	0.15	0.05	0.01	0.03	0.00	0.00

We conduct a regression analysis of market returns on the network characteristics, with the results presented in table 4. The coin market returns exhibit positive correlations with all network characteristics, except the composite measures. The coefficient for address growth rate is statistically significant at the 1% level, whereas the coefficients for transaction and hash-rate growth rates attain significance at the 5% level. Conversely, the coefficient for payment growth rate fails to reach significance even at the 10% level, implying a weak correlation between payments and market price. This outcome is not entirely unexpected, as there is no inherent requirement for price and the volume of payments recorded on the blockchain to be correlated. Cryptocurrency transactions conducted on exchanges do not necessarily result in payments or transactions being recorded on the blockchain.

Upon examination of the R^2 measures for the factor loadings, it becomes evident that only the growth in addresses yields a measure of 15%. The other measures—5%, 1%, 3%, 0%, and 0%—corresponding to the growth rates of transactions, payments, and hash-rate, as well as the two composite measures, suggest that only user adoption is a driving force behind cryptocurrency prices and returns, when considering network characteristics.

Table 5: This table reports the regression coefficients of cumulative future coin network growth on coin market index returns. The columns represent the horizons for the cumulative network growth. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δaddress	const	0.01*	0.02***	0.02***	0.03***	0.04***	0.05***	0.06***	0.06***
		(1.95)	(3.03)	(3.73)	(4.61)	(5.05)	(5.82)	(6.39)	(6.67)
	MKT	0.01	-0.04	-0.01	-0.10	-0.09	-0.12*	-0.18**	-0.18**
		(0.32)	(-0.73)	(-0.16)	(-1.58)	(-1.34)	(-1.75)	(-2.30)	(-2.15)
	R^2	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
Δtransactions	const	0.01*	0.02***	0.02***	0.03***	0.04***	0.04***	0.05***	0.05***
		(1.89)	(2.77)	(3.07)	(3.88)	(4.22)	(4.80)	(5.20)	(5.46)
	MKT	0.02	-0.07	-0.04	-0.21***	-0.21***	-0.26***	-0.28***	-0.29***
		(0.48)	(-1.09)	(-0.56)	(-2.96)	(-2.68)	(-3.48)	(-3.65)	(-3.53)
	R^2	0.00	0.00	0.00	0.02	0.01	0.02	0.03	0.02
Δpayments	const	0.00	0.00	0.01	0.01	0.01	0.01*	0.01*	0.01**
		(0.70)	(1.20)	(1.25)	(1.47)	(1.64)	(1.71)	(1.86)	(2.06)
	MKT	0.02	-0.02	0.00	-0.01	-0.03	-0.03	-0.06	-0.09**
		(0.95)	(-0.71)	(0.14)	(-0.34)	(-0.78)	(-0.75)	(-1.31)	(-2.04)
	R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Δhash-rate	const	0.04***	0.08***	0.11***	0.15***	0.20***	0.25***	0.30***	0.36***
		(6.02)	(9.69)	(11.17)	(11.93)	(12.30)	(12.17)	(11.90)	(11.67)
	MKT	0.02	-0.07	0.11	0.14	0.28**	0.30*	0.48**	0.55**
		(0.30)	(-0.96)	(1.20)	(1.27)	(2.00)	(1.72)	(2.18)	(2.06)
	R^2	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
PC_1^{network}	const	-0.07***	-0.12***	-0.18***	-0.22***	-0.26***	-0.30***	-0.32***	-0.35***
		(-12.54)	(-17.10)	(-19.92)	(-21.27)	(-21.96)	(-22.44)	(-22.67)	(-22.79)
	MKT	0.08*	0.04	0.14*	0.15*	0.17	0.19*	0.19	0.14
		(1.80)	(0.58)	(1.86)	(1.68)	(1.63)	(1.67)	(1.52)	(1.08)
	R^2	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.00
PC_2^{network}	const	-0.02***	-0.04***	-0.06***	-0.08***	-0.10***	-0.12***	-0.14***	-0.16***
		(-5.21)	(-9.17)	(-12.37)	(-15.05)	(-17.00)	(-17.77)	(-18.57)	(-19.42)
	MKT	0.01	-0.00	0.03	0.10**	0.16***	0.17***	0.21***	0.19***
		(0.33)	(-0.06)	(0.60)	(2.11)	(3.09)	(2.80)	(3.21)	(2.73)
	R^2	0.00	0.00	0.00	0.01	0.02	0.02	0.02	0.01

We further investigate the inverse relationship, namely the influence of coin market returns on network characteristics. This mechanism forms a cornerstone of the dynamic cryptocurrency pricing model developed by Cong, Li et al. (2021). In order to accomplish this, we execute the reverse regression, as explained in section 3.2, wherein future network characteristic growth, from 1 to 8 week horizons, is regressed on coin market returns. Table 5 presents the results obtained from these regressions.

It becomes evident that short-term increases in coin market return do not significantly impact the cumulative growth in addresses. However, for more protracted periods, starting at the six-week horizon, the effect is significant at the 10% level for the 6-week durations, and at the 5% level for the 7 and 8-week durations. What is of interest and contrary to the findings of Liu and Tsyvinski (2021), is that this effect is negative. A potential rationale for this result could be that surging returns on cryptocurrencies coincide with a longer timeframe during which these digital currencies become more regulated, thereby reducing the number of exchanges trading them, which in part reduces the number of active addresses. A similar pattern is observed for the growth in transactions, where the short-term effect is not significant, but from the four-week horizon onwards, the influence becomes significant, even at the 1% level. This can be explained by individuals being less inclined to use cryptocurrencies for payments when returns are rising, opting to save them in hopes of an increase in value rather than spending them.

For the growth rate in payments, an increase in coin market return yields a virtually negligible effect, with only marginal significance observed at the 8-week horizon. The growth rate in hash-rate does show a significant effect at the 5% level for the 5, 7, and 8-week horizons, and at the 10% level for the 6-week horizon. Again, the effect of increasing coin market results on the growth rate in payments is negative, yielding a similar explanation as for the negative relationship observed for the growth in transactions. The first principal component, primarily reflecting the variation in the growth rate of transactions and payments, remains largely uninfluenced by increases in coin market returns, showing only slight significance at the 10% level for the 1, 3, 4, and 6-week horizon. Contrastingly, the second principal component exhibits an effect commencing from the 4-week horizon up to and including the 8-week horizon at the 1% level for the 5, 6, 7, and 8-week horizons and at the 5% level for the 4-week horizon.

4.2.2 Production Characteristics

Table 6: This table reports the correlations of the production characteristics. The data frequency is monthly.

	P^{US}	Gen^{US}	$Cons^{US}$	P^{CN}	Gen^{CN}	PC_1^{prod}	PC_2^{prod}
P^{US}	1.00	0.55	0.60	0.03	0.37	0.70	-0.22
Gen^{US}		1.00	0.97	-0.02	0.67	0.99	-0.07
$Cons^{US}$			1.00	-0.02	0.62	0.99	-0.16
P^{CN}				1.00	-0.15	-0.03	-0.18
Gen^{CN}					1.00	0.72	0.69
PC_1^{prod}						1.00	-0.00
PC_2^{prod}							1.00

The correlation matrix for the network factors under investigation is presented in table 6. A noteworthy result is the weak correlation between the prices of electricity in China and the United States, with a correlation coefficient of only 0.03, indicating almost no correlation at all.

This might be attributed to the significantly different main sources of electricity for China and the United States. According to the Gas Exporting Countries Forum (GECF), China mainly relies on coal and hydropower, while the US primarily depends on natural gas and coal. Another possible explanation could be the heavy influence of the Chinese government over electricity prices, compared to the United States where there exists mostly a free market.

We also observe a strong correlation between the generation and consumption of electricity in the United States, with a correlation coefficient of 0.97. This is expected, as the import and export of electricity in the United States are nearly equal. There is a correlation of 0.67 and 0.62 between the generation and consumption of electricity in the United States and the generation of electricity in China, respectively. Given China its recent rapid economic development, it is reasonable to expect that both countries will experience similar increases in electricity consumption levels.

The first principal component appears to represent the United States energy market, as evidenced by the correlation between it and the price of electricity, the generation of electricity, and the consumption of electricity in the United States, which are 0.70, 0.99, and 0.99, respectively. While it does not capture any variation in the price of electricity in China, it correlates well with the generation of electricity in China, with a coefficient of 0.72.

The second principal component seems to capture the differences between the United States and Chinese energy markets, with only negative coefficients for the former (-0.22, -0.07, and -0.16) and a substantially large positive coefficient for the generation of electricity in China, namely 0.69. However, it correlates negatively with the price of electricity in China with a coefficient of -0.18.

Table 7: This table reports the regression coefficients of the coin market returns on the production characteristics. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is monthly.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
P^{US}	-2.31 (-1.20)						
Gen^{US}		-0.01 (-0.02)					
$Cons^{US}$			-0.06 (-0.18)				
P^{CN}				54.53* (1.96)			
Gen^{CN}					0.33 (0.45)		
PC_1^{prod}						-0.05 (-0.16)	
PC_2^{prod}							-0.00 (-0.00)
R^2	0.01	0.00	0.00	0.03	0.00	0.00	0.00

We now analyse the exposures of the production characteristics with the coin market returns, presented in table 7. Surprisingly, we only observe a significant exposure of the coin market returns to the price of electricity in China. All the other factors, including the compound measures, appear to have no statistically significant exposure to the coin market returns.

When analysing the coefficient of determination (R^2), we see a similar result. Only the price of electricity in the United States and the price of electricity in China have coefficients different from 0.00.

Interestingly, the regression coefficient is largely positive, indicating a positive relationship between the return on cryptocurrencies and the price of electricity in China. This is entirely unexpected, and it is difficult to postulate a possible explanation for this finding.

4.2.3 Sentiment Characteristics

Table 8: This table reports the correlations of the sentiment characteristics. The data frequency is monthly.

	ΔN	ΔN_{pos}	ΔN_{neg}	ΔN_{neu}	ΔS_{avg}	ΔS_{wavg}	ΔS_{std}	PC_1^{sent}	PC_2^{sent}
ΔN	1.00	0.92	0.94	0.99	0.22	0.14	0.15	0.25	0.61
ΔN_{pos}		1.00	0.81	0.89	0.27	0.13	0.09	0.11	0.70
ΔN_{neg}			1.00	0.91	0.01	0.08	0.38	0.29	0.50
ΔN_{neu}				1.00	0.25	0.15	0.06	0.25	0.58
ΔS_{avg}					1.00	0.37	-0.18	0.17	0.27
ΔS_{wavg}						1.00	-0.07	0.07	0.14
ΔS_{std}							1.00	0.28	-0.06
PC_1^{sent}								1.00	-0.06
PC_2^{sent}									1.00

The correlation matrix for the sentiment characteristics under investigation is presented in table 8. We observe a strong correlation between the number of posts, positive posts, negative posts, and neutral posts, with correlation coefficients all near 0.9, except for the correlation between positive and negative posts. The correlation between positive and negative posts is 0.81, suggesting that growth rates for these posts do not move in tandem. This is expected since positive posts likely increase when cryptocurrency prices rise, while negative posts increase when prices fall.

As anticipated, the number of neutral posts correlates strongly with the total number of posts since neutral posts tend to increase when the overall post count rises. Neutral posts typically include advertisements and other content not directly related to significant positive or negative price movements.

In analyzing the mean constructed sentiment measure, it is interesting to note that it does not correlate with the growth rate in negative posts but does correlate with the growth rate in total posts, positive posts, and neutral posts. One possible explanation is that the natural language model may struggle to identify negative posts, leading to a lower sentiment measure that places more emphasis on positive and neutral posts.

The weighted average of the constructed sentiment measure exhibits a similar pattern, although this version has an even smaller correlation with the growth rates in total posts. The standard deviation of the constructed sentiment measure displays minimal correlation with the growth rates in total posts, positive posts, neutral posts, and the two constructed sentiment measures. However, it does show a substantial correlation with the growth rate in negative posts, with a coefficient of 0.38. This finding supports our previous hypothesis that the natural language model faces more difficulty identifying negative posts, resulting in greater variation in the constructed sentiment measure.

When analysing the principal components, we see that the first principal component correlates strongly with the growth rate in total posts, negative posts, and neutral posts with coefficients of 0.25, 0.29, and 0.25, respectively. This principal component captures the dependence between the growth rates in total posts, negative posts, and neutral posts. The first component also correlates with the growth rate in positive posts, albeit to a lesser extent, with a coefficient of 0.11. The first principal component also correlates well with the growth rate of the standard deviation in the constructed sentiment measure, with a coefficient of 0.28. This is not unexpected, since this measure also correlates well with the growth rate in negative posts.

The second principal component correlates strongly with the growth rate in posts and positive posts, with a coefficient of 0.61 and 0.70, respectively. It also captures some dependence between the growth rate in negative posts and the growth rate in negative and neutral posts, with coefficients of 0.50 and 0.58, respectively. To a lesser extent, it also captures a small amount of dependence between the mean constructed sentiment measure and the score weighted average of the constructed sentiment measure, with coefficients of 0.27 and 0.14, respectively.

Table 9: This table reports the regression coefficients of the coin market returns on the sentiment characteristics. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is monthly.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔN	0.12***								
	(2.66)								
ΔN_{pos}		0.05*							
		(1.90)							
ΔN_{neg}			0.14***						
			(3.23)						
ΔN_{neu}				0.14***					
				(2.81)					
ΔS_{avg}					-0.10				
					(-0.10)				
ΔS_{wavg}						0.14			
						(0.48)			
ΔS_{std}							1.07		
							(1.59)		
PC_1^{sent}								-0.09*	
								(-1.79)	
PC_2^{sent}									-0.10
									(-1.39)
R^2	0.06	0.03	0.08	0.06	0.00	0.00	0.02	0.03	0.02

We now analyse the exposures of the sentiment characteristics with the coin market returns, presented in table 9. The regression coefficients for the growth rate in posts, the growth rate in positive posts, the growth rate in negative posts, and the growth rate in neutral posts are all significant and positive, indicating a positive relationship between the return on cryptocurrencies and the growth rate in posts. The regression coefficients for the growth rate in the number of posts, negative posts, and neutral posts are significant at the 1% level. The coefficient for the growth rate in the number of positive posts is significant at the 10% level. This result is surprising and indicates that the number of positive posts does not have a substantial influence on cryptocurrency market returns. Another surprising result is that this effect is positive, meaning that more neutral and negative posts result in a higher cryptocurrency market return.

The constructed measures have no significant effect on the market return. From this, we can conclude that the relative sentiment score, which our constructed sentiment measure represents, does not have a significant effect on the cryptocurrency market returns.

When analyzing the composite measures, only the coefficient on the first principal components has slight significance and a negative coefficient. This is to be expected since this component captures the dependence structure between the post, negative post, and positive post growth rates.

Table 10: This table reports the regression coefficients of cumulative coin market return on future coin sentiment growth. The columns represent the horizons for the cumulative coin market return. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔN	0.02**	0.05***	0.09***	0.06**	0.09**	0.11***	0.11**	0.12**
	(2.47)	(3.56)	(4.17)	(2.28)	(2.56)	(2.70)	(2.26)	(2.21)
R^2	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.01
ΔN_{pos}	0.01**	0.02***	0.04***	0.03**	0.04**	0.05**	0.05*	0.04
	(2.12)	(3.37)	(4.09)	(2.21)	(2.42)	(2.22)	(1.95)	(1.64)
R^2	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.01
ΔN_{neg}	0.02**	0.04***	0.07***	0.05**	0.08***	0.10***	0.11***	0.14***
	(2.17)	(3.17)	(3.91)	(2.33)	(2.98)	(3.15)	(2.92)	(3.24)
R^2	0.01	0.02	0.03	0.01	0.02	0.02	0.02	0.02
ΔN_{neu}	0.03***	0.07***	0.11***	0.08**	0.11***	0.14***	0.14**	0.15**
	(2.75)	(3.90)	(4.37)	(2.50)	(2.63)	(2.89)	(2.40)	(2.28)
R^2	0.01	0.03	0.04	0.01	0.01	0.02	0.01	0.01
ΔS_{avg}	0.05	0.02	0.14	0.17	-0.10	-0.16	0.05	-0.45
	(0.43)	(0.11)	(0.54)	(0.51)	(-0.25)	(-0.33)	(0.09)	(-0.70)
R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ΔS_{wavg}	0.01	0.02	0.05	0.09	0.05	0.09	0.03	0.03
	(0.41)	(0.45)	(0.70)	(0.99)	(0.45)	(0.66)	(0.20)	(0.18)
R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ΔS_{std}	-0.02	-0.04	0.09	0.00	0.30	0.17	0.30	0.41
	(-0.24)	(-0.34)	(0.49)	(0.02)	(1.08)	(0.49)	(0.75)	(0.93)
R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PC_1^{sent}	-0.01*	-0.02	-0.03	-0.06***	-0.09***	-0.11***	-0.13***	-0.15***
	(-1.77)	(-1.63)	(-1.51)	(-2.76)	(-3.23)	(-3.47)	(-3.53)	(-3.62)
R^2	0.01	0.01	0.00	0.01	0.02	0.02	0.02	0.03
PC_2^{sent}	-0.02*	-0.02	-0.00	-0.00	0.00	-0.02	-0.01	-0.08
	(-1.72)	(-1.22)	(-0.10)	(-0.05)	(0.08)	(-0.27)	(-0.17)	(-1.08)
R^2	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00

We now analyse the long-term effect of the sentiment characteristics on the cryptocurrency market returns. These results are presented in table 10. We see that the growth rate in posts has a significant effect on the cryptocurrency market returns at all horizons. The effect is significant at the 5% level for the 1, 4, 5, 7, and 8-week horizons, and at the 1% level for the 2, 3, and 6-week horizons. An increase in posts has a long-lasting effect on the cumulative market returns.

The same effect is also observed for the growth rate in positive posts, albeit less significant. This effect is only significant at the 1% level for the 2 and 3-week horizons, at the 5% level for the 1, 4, 5, and 6-week horizons, and at the 10% level for the 7-week horizon.

Similar to the growth rate of the posts, the growth rates of the negative and neutral posts are significant for all horizons, for at least the 5% level. Surprisingly, all coefficients are positive, meaning an increase in negative posts results in an increase in cumulative cryptocurrency market returns. This gives rise to the phrase, "there is no such thing as bad publicity", at least not for cryptocurrencies it seems.

4.3 Cross-Sectional Analysis

In this section, we perform a cross-sectional analysis on the performance of weekly-rebalanced quintile portfolios, based on size-, momentum-, volume-, volatility-, and cryptocurrency-related return predictors presented in section 3.3. We then analyse the portfolio returns using four distinct factor models detailed in section 3.4, employing three separate factors discussed in section 3.1.

We find that of all the considered return predictors, only MCAP, PRC, MAXDPRC, 1, 2, and 3-week horizon cumulative return, PRCVOL, VOLSCALED, and STDPRCVOL form high-minus-low portfolios capable of generating significant returns. None of the return predictors based on projected cryptocurrency characteristics are able to form high-minus-low portfolios capable of generating significant returns.

After factor analysis, only MCAP, the 2-week horizon cumulative return, and the 50-week cumulative return predictors are able to generate significant excess return. The market factor, size factor, and momentum factor were able to explain the returns observed in the other return predictors.

Additionally, our findings suggest that investing in cryptocurrencies with low exposure to the growth rate of electricity generation in China can yield a modest yet significant excess return. This outcome could potentially instigate a shift in demand towards Proof-of-Stake cryptocurrencies.

4.3.1 Size-Related Return Predictors

Table 11: This table reports the weekly mean quintile portfolio returns on the size return predictors. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	$5-1$
MCAP	0.061***	0.023***	0.024***	0.034***	0.016***	-0.045***
	(5.470)	(2.934)	(2.642)	(2.856)	(3.196)	(-4.609)
PRC	0.075***	0.035***	0.012	0.018**	0.016***	-0.059**
	(2.765)	(2.888)	(1.477)	(2.080)	(3.195)	(-2.227)
MAXDPRC	0.069**	0.038***	0.011	0.021**	0.016***	-0.054**
	(2.551)	(3.031)	(1.346)	(2.304)	(3.188)	(-2.010)
AGE	0.023**	0.019**	0.024**	0.022**	0.014***	-0.009
	(2.067)	(2.073)	(2.124)	(2.454)	(2.993)	(-0.833)

We now explore the results for the size return predictors, as displayed in table 11. First, we investigate the logarithm of the market cap (MCAP) zero-investment quintile portfolios. Notably, the MCAP high-minus-low portfolio yields significant mean weekly returns of 4.5%, significant at the 1% level. However, we do not observe a consistent decrease or increase in mean return across the quintile portfolios, which could indicate that the measure may not fully capture the underlying anomaly.

Next, we examine the characteristic based on the last trading day's price (PRC) and observe similar results. The returns across the quintiles are nearly monotonically decreasing, except for the higher three quintiles. The high-minus-low portfolio based on this return predictor generates an even greater return of 5.9% per week, significant at the 1% level.

We then analyse the maximum price in the week of trading before portfolio formation (MAXDPRC), discovering a result quite similar to the previous characteristic. This high-minus-low portfolio generates an almost identical weekly return of 5.4%. Again, the return across the quintiles is nearly monotonically decreasing, except for the higher three quintiles.

Lastly, we assess the characteristic based on the cryptocurrency's age (AGE) and find no significant result. Additionally, we do not observe monotonically decreasing returns across the portfolios.

In summary, three of the four size return predictors form high-minus-low portfolios capable of generating significant returns.

We now examine the size portfolios, accounting for the coin market factor, size factor, and momentum factor, with results presented in table 12. When only adjusting for the market factor, all excess returns generated by the portfolios, except for the AGE portfolio, maintain significance, although at the 5% and 10% level for the PRC and MAXDPRC portfolios, respectively. This suggests that the portfolios created based on size-related return predictors exhibit minimal market risk. This is evident from the coefficients of the market factor in the CAPM model, which is only significant for the PRC portfolio, at the 10% level.

Correcting with the additional size factor, all portfolios, except AGE, exhibit significant loadings at the 1% level. As expected, these return predictors are intended to proxy for the return anomaly where smaller cryptocurrencies yield higher returns than larger ones, similar to the size factor. All portfolios, except for the MCAP portfolio, lose significance in their excess returns.

Considering the momentum factor, instead of the size factor, no portfolios display significant loadings. This indicates that these portfolios have no exposure to the momentum factor, which is expected. This suggests that there is no connection between the market cap of a cryptocurrency and its cumulative returns. After adjusting for the momentum factor, all portfolios, except for AGE, retain significant excess returns.

When factoring in all three factors, only the MCAP portfolio remains with significant excess returns. All other portfolio excess returns are explained by the size factor alone. The AGE portfolio is the only portfolio based on a size-related return predictor that does not generate significant excess returns, in any of the configurations.

4.3.2 Momentum-Related Return Predictors

We assess the performance of portfolios based on momentum return predictors, as shown in table 13. We notice that only the high-minus-low portfolios based on the 1, 2, and 3-week horizon return predictors generate significant mean weekly returns. These portfolios generate

Table 12: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low size return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
MCAP	(1)	-0.043***	-0.098			0.003
		(-5.419)	(-0.484)			
	(2)	-0.009***	-0.036	-1.253***		0.841
		(-3.332)	(-0.819)	(-12.920)		
	(3)	-0.051***	-0.076		0.377	0.070
		(-4.482)	(-0.421)		(1.445)	
	(4)	-0.010***	-0.035	-1.247***	0.024	0.842
		(-3.649)	(-0.804)	(-12.370)	(0.565)	
PRC	(1)	-0.051**	-0.479*			0.008
		(-1.974)	(-1.735)			
	(2)	-0.024	-0.430*	-0.971***		0.078
		(-1.046)	(-1.746)	(-3.085)		
	(3)	-0.053**	-0.471*		0.117	0.009
		(-2.185)	(-1.697)		(0.293)	
	(4)	-0.020	-0.437*	-1.014***	-0.171	0.080
		(-0.997)	(-1.769)	(-2.900)	(-0.475)	
MAXDPRC	(1)	-0.047*	-0.403			0.006
		(-1.778)	(-1.593)			
	(2)	-0.022	-0.356	-0.925***		0.067
		(-0.904)	(-1.575)	(-2.858)		
	(3)	-0.050**	-0.386		0.147	0.007
		(-2.018)	(-1.569)		(0.382)	
	(4)	-0.019	-0.361	-0.953***	-0.113	0.068
		(-0.945)	(-1.599)	(-2.728)	(-0.324)	
AGE	(1)	-0.007	-0.111			0.003
		(-0.757)	(-0.582)			
	(2)	0.011	-0.078	-0.655		0.201
		(0.899)	(-0.430)	(-1.368)		
	(3)	-0.014	-0.090		0.340	0.050
		(-1.063)	(-0.499)		(1.084)	
	(4)	0.006	-0.070	-0.613	0.166	0.211
		(0.605)	(-0.387)	(-1.385)	(0.812)	

Table 13: This table reports the weekly mean quintile portfolio returns on the momentum return predictors. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	5 − 1
$r^{(1,0)}$	0.005	0.012	0.018**	0.034***	0.027***	0.022**
	(0.668)	(1.349)	(2.087)	(3.572)	(2.999)	(2.468)
$r^{(2,0)}$	0.005	0.015	0.013**	0.028***	0.034***	0.029***
	(0.517)	(1.323)	(1.986)	(3.067)	(3.964)	(3.547)
$r^{(3,0)}$	0.010	0.003	0.011*	0.036***	0.031***	0.020***
	(1.025)	(0.393)	(1.841)	(3.663)	(3.777)	(2.598)
$r^{(4,0)}$	0.015	0.011	0.011*	0.030***	0.026***	0.011
	(1.491)	(1.370)	(1.826)	(3.206)	(3.303)	(1.449)
$r^{(4,1)}$	0.008	0.011	0.038**	0.027***	0.019**	0.011
	(0.874)	(1.523)	(2.569)	(3.362)	(2.486)	(1.366)
$r^{(8,0)}$	0.024**	0.011*	0.028**	0.024***	0.020***	-0.004
	(2.426)	(1.659)	(2.185)	(2.982)	(2.772)	(-0.444)
$r^{(16,0)}$	0.026**	0.023**	0.024**	0.019**	0.019***	-0.007
	(2.527)	(2.059)	(2.175)	(2.406)	(2.937)	(-0.710)
$r^{(50,0)}$	0.015***	0.013**	0.021**	0.016***	0.009	-0.006
	(2.606)	(2.216)	(2.397)	(2.737)	(1.626)	(-1.156)
$r^{(100,0)}$	0.033***	0.017***	0.015***	0.013**	0.012**	-0.021
	(2.605)	(2.843)	(2.614)	(2.101)	(2.338)	(-1.612)

2.2%, 2.9%, and 2.0% mean weekly return, respectively. This mean return is significant at the 5% level for the 1-week horizon and at the 1% level for the 2 and 3-week horizons. Only for the 1 and 2-week horizons are the quintile portfolio returns almost monotonically increasing. For the 3-week horizon, this effect is less pronounced. Interestingly, for the shorter maturities, from 1 up to 4-week horizons, the lower quintiles do not have statistically significant mean weekly returns. The portfolios based on the other horizons all lack significant mean weekly returns.

What is intriguing is that there appears to be a sort of reversal effect, starting at the portfolios from the 8-week horizons. From this horizon, it appears that lower momentum cryptocurrencies actually produce a higher mean return than those with high momentum.

In summary, only three momentum return predictors form high-minus-low portfolios capable of generating significant returns.

We now investigate the first set of momentum portfolios, taking into account the coin market factor, size factor, and momentum factor, with results displayed in table 14. When adjusting for just the market factor, none of the return predictors have a significant loading, even at the 10% level. Only the 1-to-4-week horizon has an insignificant excess return, even at the 10% level. The rest of the momentum-based return predictors have a significant excess return at the 1% level, except for the 4-week horizon portfolio, which is only significant at the 5% level.

In line with the results for the size-based return predictors, almost none of the momentum-based return predictors have a significant loading on the size factor. Only the 2- and 4-week horizon momentum portfolios have a significant loading on the size factor, albeit at the 10%. After correcting for the additional momentum factor, does the 2-week horizon portfolio lose its significance on the size factor, with only the 4-week horizon portfolio remaining significantly

Table 14: This table reports the results on the cryptocurrency factor adjustments of the first set of high-minus-low momentum return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
$r^{(1,0)}$	(1)	0.021***	0.031			0.000
		(2.811)	(0.168)			
	(2)	0.017	0.023	0.155		0.016
		(1.554)	(0.115)	(0.362)		
	(3)	0.012	0.055		0.417	0.101
		(1.155)	(0.330)		(1.587)	
	(4)	0.003	0.046	0.279	0.496***	0.149
		(0.367)	(0.250)	(0.611)	(2.701)	
$r^{(2,0)}$	(1)	0.031***	-0.134			0.007
		(4.317)	(-0.815)			
	(2)	0.040***	-0.118	-0.317*		0.084
		(5.113)	(-0.843)	(-1.650)		
	(3)	0.012**	-0.082		0.887***	0.549
		(2.072)	(-0.899)		(7.118)	
	(4)	0.015***	-0.079	-0.100	0.858***	0.556
		(2.608)	(-0.908)	(-1.290)	(7.146)	
$r^{(3,0)}$	(1)	0.023***	-0.155			0.010
		(3.287)	(-0.960)			
	(2)	0.031***	-0.140	-0.297		0.084
		(4.162)	(-0.992)	(-1.505)		
	(3)	0.003	-0.100		0.937***	0.667
		(0.517)	(-1.116)		(7.450)	
	(4)	0.005	-0.098	-0.065	0.918***	0.670
		(0.849)	(-1.103)	(-0.855)	(7.225)	
$r^{(4,0)}$	(1)	0.014**	-0.175			0.013
		(2.080)	(-1.099)			
	(2)	0.023***	-0.158	-0.342*		0.113
		(3.261)	(-1.183)	(-1.933)		
	(3)	-0.003	-0.127		0.820***	0.523
		(-0.597)	(-1.372)		(6.782)	
	(4)	0.001	-0.123	-0.146*	0.779***	0.540
		(0.246)	(-1.418)	(-1.894)	(6.902)	
$r^{(4,1)}$	(1)	0.011	0.006			0.000
		(1.581)	(0.040)			
	(2)	0.023**	0.028	-0.430		0.141
		(2.338)	(0.191)	(-1.090)		
	(3)	-0.000	0.037		0.536**	0.194
		(-0.034)	(0.245)		(2.315)	
	(4)	0.010	0.048	-0.318	0.445***	0.266
		(1.379)	(0.308)	(-0.960)	(2.618)	

loaded on the size factor, at 10%. This reinforces the idea that market cap and cumulative returns are not linked to one-another.

After correcting for just the market and momentum factors, none of the portfolios exhibit significant excess returns, except for the 2-week horizon portfolio. Even after adjusting for all three factors, the 2-week horizon portfolio still displays significant excess returns, whereas none of the other portfolios do.

This calls into question the choice for a 3-week horizon momentum factor. We ultimately made this decision to stay in conformity with the original research by Liu et al. (2022) and enable us to directly compare our results. In retrospect, this decision might have been misguided, as the 2-week horizon portfolio appears to be the only one with significant excess returns, even after correcting for all three factors.

Table 15: This table reports the results on the cryptocurrency factor adjustments of the second set of high-minus-low momentum return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
$r^{(8,0)}$	(1)	-0.004	-0.032			0.000
		(-0.499)	(-0.159)			
	(2)	0.012	-0.003	-0.572		0.179
		(1.069)	(-0.017)	(-1.203)		
	(3)	-0.021**	0.016		0.818***	0.327
		(-2.051)	(0.094)		(3.180)	
	(4)	-0.008	0.029	-0.393	0.706***	0.406
		(-0.984)	(0.163)	(-1.036)	(3.768)	
$r^{(16,0)}$	(1)	-0.004	-0.196			0.010
		(-0.501)	(-0.765)			
	(2)	0.015*	-0.160	-0.702*		0.280
		(1.680)	(-0.830)	(-1.872)		
	(3)	-0.020*	-0.151		0.763***	0.294
		(-1.930)	(-0.806)		(3.049)	
	(4)	-0.002	-0.133	-0.549*	0.607***	0.446
		(-0.233)	(-0.883)	(-1.928)	(4.298)	
$r^{(50,0)}$	(1)	-0.008	0.095			0.008
		(-1.513)	(1.530)			
	(2)	-0.008	0.095	-0.001		0.008
		(-1.524)	(1.527)	(-0.025)		
	(3)	-0.010*	0.101		0.104*	0.025
		(-1.888)	(1.623)		(1.812)	
	(4)	-0.011*	0.100	0.027	0.111*	0.026
		(-1.959)	(1.611)	(0.604)	(1.683)	
$r^{(100,0)}$	(1)	-0.021*	-0.001			0.000
		(-1.930)	(-0.006)			
	(2)	-0.017*	0.006	-0.151		0.007
		(-1.668)	(0.037)	(-1.264)		
	(3)	-0.021*	0.001		0.028	0.000
		(-1.719)	(0.003)		(0.213)	
	(4)	-0.016	0.006	-0.155	-0.016	0.007
		(-1.343)	(0.033)	(-1.301)	(-0.117)	

We now investigate the second set of momentum portfolios, as displayed in table 15. When adjusting for only the market factor, none of the portfolios have a significant loading on this factor, even at the 10% level. Additionally, only one of the horizons, namely the 100-week horizon, has a significant excess return, at the 10% level.

When considering only the market and size factors, we observe that only the 16-week horizon portfolio has a significant loading on this additional factor, at the 10% level. Even after considering all three factors, does the 16-week horizon portfolio remain significantly loaded on the size factor, also the 10% level. Of the 4 horizons considered, only the 100-week horizon does not have a significant loading on the momentum factor. Only after correcting for all three factors does the 50-week horizon maintain a significant excess return, at the 10% level.

4.3.3 Volume-Related Return Predictors

Table 16: This table reports the weekly mean quintile portfolio returns on the volume return predictors. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	5 − 1
VOL	0.020***	0.018***	0.019**	0.017**	0.011**	-0.009
	(2.688)	(2.849)	(2.580)	(2.074)	(2.533)	(-1.255)
PRCVOL	0.040***	0.059*	0.019***	0.020**	0.011**	-0.029**
	(2.741)	(1.801)	(2.589)	(2.025)	(2.477)	(-2.119)
VOLSCALED	0.036***	0.038**	0.014**	0.008	0.011**	-0.026**
	(2.778)	(2.372)	(2.089)	(1.180)	(2.489)	(-2.120)

We evaluate the performance of portfolios based on volume return predictors, as illustrated in table 16. Firstly, we investigate the logarithm of the average daily trading volume (VOL) zero-investment quintile portfolios. The high-minus-low portfolio produces an insignificant mean weekly return, even at the 10% level. Moreover, the mean quintile returns do not exhibit any monotonicity among them.

Next, we examine the logarithm of the average daily trading volume multiplied by the price (PRCVOL) zero-investment quintile portfolios. The high-minus-low portfolio for this return predictor yields a significant mean weekly return of 2.9%, at the 5% level. As with the previous return predictor, these quintile portfolios also do not display any monotonicity in the mean returns.

Finally, we explore the logarithm of the average daily trading volume multiplied by the price and scaled by market capitalization (VOLSCALED) zero-investment quintile portfolios. The high-minus-low portfolio for this return predictor generates a significant mean weekly return of 2.6%, at the 5% level. These quintile portfolios also do not display any monotonicity in their mean returns.

In summary, only two of the three volume return predictors form high-minus-low portfolios capable of generating significant returns.

We now analyse the volume portfolios, accounting for the coin market factor, the size factor, and the momentum factor. The results of this analysis are presented in table 17. When adjusting

Table 17: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low volume return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
VOL	(1)	-0.013	0.245*			0.031
		(-1.484)	(1.725)			
	(2)	-0.009	0.252*	-0.135*		0.051
		(-1.106)	(1.766)	(-1.955)		
	(3)	-0.012	0.244*		-0.017	0.032
		(-1.525)	(1.727)		(-0.341)	
	(4)	-0.007	0.249*	-0.150**	-0.060	0.054
		(-0.995)	(1.765)	(-2.219)	(-1.053)	
PRCVOL	(1)	-0.026**	-0.190			0.005
		(-2.062)	(-0.589)			
	(2)	-0.010	-0.161	-0.570		0.095
		(-0.868)	(-0.520)	(-1.394)		
	(3)	-0.020*	-0.205		-0.269	0.023
		(-1.896)	(-0.622)		(-1.238)	
	(4)	0.003	-0.182	-0.686*	-0.463	0.144
		(0.214)	(-0.598)	(-1.699)	(-1.558)	
VOLSCALED	(1)	-0.020**	-0.328			0.019
		(-2.027)	(-1.356)			
	(2)	-0.010	-0.309	-0.378		0.069
		(-1.049)	(-1.309)	(-1.333)		
	(3)	-0.021*	-0.326		0.029	0.019
		(-1.896)	(-1.304)		(0.127)	
	(4)	-0.008	-0.313	-0.399	-0.084	0.071
		(-0.569)	(-1.276)	(-1.305)	(-0.291)	

solely for the market factor, only the VOL portfolio has a significant loading on this factor, at the 10% level. Even when considering all different factor model configurations, does this portfolio remain significantly loaded on the market factor. This is also the only portfolio that does not have any significant excess returns, as both the PRCVOL and VOLSCALED portfolios exhibit significant excess returns, both at the 5% level.

When examining the model that includes both the market and size factors, it appears that only the VOL portfolio has a significant loading on the newly added size factor, significant at the 10% level. The remaining portfolios do not show significant loadings on this factor unless the model is supplemented by the momentum factor. With the addition of the momentum factor, the PRCVOL portfolio also demonstrates a significant loading on the size factor, again significant at the 10% level. Based on the sign of the coefficient, this demonstrates that cryptocurrencies with small volume often have small market caps, and vice versa.

None of the volume based return predictors portfolios have a significant loading on the momentum factor. This signals that investing in cryptocurrencies with high volume, often referred to as momentum by informed investors, does not lead to significant cumulative returns. However, none of the considered portfolios, after correcting for all three factors, have any significant excess return, even at the 10% level.

4.3.4 Volatility-Related Return Predictors

Table 18: This table reports the weekly mean quintile portfolio returns on the volatility return predictors. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	5 − 1
BETA	0.021**	0.028***	0.017***	0.015**	0.010	-0.012
	(2.099)	(2.778)	(2.877)	(2.197)	(1.630)	(-1.130)
BETA2	0.022**	0.028***	0.017***	0.015**	0.010	-0.012
	(2.110)	(2.759)	(2.863)	(2.206)	(1.635)	(-1.139)
IDIOVOL	0.012***	0.018***	0.019**	0.028	0.013	0.002
	(2.744)	(2.699)	(2.571)	(1.632)	(1.519)	(0.202)
RETVOL	0.024***	0.014*	0.027***	0.019**	0.020	-0.004
	(2.969)	(1.949)	(3.181)	(2.022)	(1.472)	(-0.314)
MAXRET	0.011	0.028***	0.026***	0.023***	0.024*	0.013
	(1.358)	(3.003)	(3.160)	(2.766)	(1.923)	(1.217)
DELAY	0.011***	0.010*	0.018**	0.030**	0.010*	-0.002
	(2.608)	(1.807)	(2.462)	(2.490)	(1.673)	(-0.365)
STDPRCVOL	0.038***	0.024***	0.025**	0.023**	0.011**	-0.027**
	(2.685)	(3.168)	(2.296)	(2.346)	(2.458)	(-2.056)
DAMIHUD	0.011**	0.018*	0.020**	0.041**	0.030**	0.019
	(2.388)	(1.960)	(2.250)	(2.090)	(2.333)	(1.606)

We evaluate the performance of portfolios based on volatility return predictors, as illustrated in table 18. The only high-minus-low portfolio that generates significant mean weekly return is based on the STDPRCVOL return predictor, at the 5% level. This portfolio shows a 2.7% mean weekly return. These quintiles display rough monotonicity between the different portfolios, with the three middle portfolios showing roughly the same, but significant return. In other words, going long on cryptocurrencies with low price volume standard deviation and shorting coins with high price volume volatility generates significant weekly returns.

All other return predictors show no significant mean weekly high-minus-low portfolio return. Only the quintiles scoring low for the return predictors have significant mean returns, whereas the higher quintiles all lose significance.

In summary, only one of the eight volatility based return predictors forms high-minus-low portfolios capable of generating significant returns.

We now analyse the first set of volatility portfolios, accounting for the coin market factor, the size factor, and the momentum factor. The results of this analysis are presented in table 19.

When evaluating the market factor, only the BETA and BETA2 portfolios show a significant loading on this factor, and this is at the 5% significance level. These portfolios maintain their significant loading at the 5% level, even when the market factor is considered alongside the size and/or momentum factors. Interestingly, these two portfolios are also the only ones that generate significant excess returns. However, this excess return is only significant at the 10% level and only occurs when adjustments are made for either the market factor alone or a combination of the market and size factors. None of the other considered portfolios generate a significant excess return in any configuration, even at the 10% level.

When considering the size factor in combination with the market factor, none of the considered portfolios have a significant loading on the size factor. This shows that there is no underlying connection between the returns of cryptocurrencies with high idiosyncratic volatility and size.

When considering the momentum factor in combination with the market factor, only RET-VOL and MAXRET have a significant loading on the momentum factor, at a 10% and 5% level, respectively. This shows that cryptocurrencies with high return also have high cumulative return.

We now analyse the second set of volatility portfolios, as shown in table 20. When considering only the market factor, none of the portfolios exhibit a significant loading on this factor, even at the 10% significance level. However, both the STDPRCVOL and DAMIHUD portfolios have significant excess returns in this factor model configuration, at the 5% significance level.

When adjusting for both the market and size factors, the DELAY portfolio exhibits a modest but significant loading, at the 10% level. Upon further adjustment for the momentum factor, this factor loading becomes insignificant. Yet, in this expanded model configuration, both the STDPRCVOL and DAMIHUD portfolios become significantly loaded on the size factor, also at the 10% level.

For STDPRCVOL, acting as a liquidity-based return predictor, the sign of the coefficient suggests that high liquidity is associated with a high market cap, and vice versa. For DAMIHUD, functioning as the converse of STDPRCVOL and serving as an illiquidity return predictor, the result, based on the sign of the coefficient, is comparable.

In the full factor model, only the DAMIHUD portfolio exhibits a significant loading on the momentum factor. The loading of the DAMIHUD portfolio on this factor suggests that illiquid cryptocurrencies tend to have high momentum and vice versa. Unfortunately, none of

Table 19: This table reports the results on the cryptocurrency factor adjustments of the first set of high-minus-low volatility return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
BETA	(1)	-0.016*	0.250**			0.015
		(-1.795)	(1.976)			
	(2)	-0.015*	0.251**	-0.030		0.015
		(-1.856)	(1.998)	(-0.518)		
	(3)	-0.016	0.251**		0.026	0.015
		(-1.565)	(2.016)		(0.199)	
	(4)	-0.015	0.252**	-0.025	0.018	0.016
		(-1.481)	(2.016)	(-0.366)	(0.131)	
BETA2	(1)	-0.016*	0.249**			0.015
		(-1.805)	(1.968)			
	(2)	-0.015*	0.250**	-0.030		0.015
		(-1.866)	(1.991)	(-0.521)		
	(3)	-0.016	0.250**		0.025	0.015
		(-1.571)	(2.010)		(0.194)	
	(4)	-0.015	0.251**	-0.025	0.018	0.015
		(-1.487)	(2.008)	(-0.370)	(0.128)	
IDIOVOL	(1)	-0.000	0.119			0.006
		(-0.054)	(1.062)			
	(2)	-0.004	0.111	0.152		0.026
		(-0.638)	(0.975)	(1.586)		
	(3)	0.002	0.113		-0.095	0.013
		(0.206)	(1.013)		(-0.697)	
	(4)	-0.003	0.109	0.138	-0.056	0.028
		(-0.334)	(0.936)	(1.243)	(-0.359)	
RETVOL	(1)	-0.004	0.025			0.000
		(-0.374)	(0.139)			
	(2)	-0.003	0.028	-0.046		0.001
		(-0.261)	(0.153)	(-0.352)		
	(3)	-0.011	0.045		0.336*	0.034
		(-0.979)	(0.267)		(1.724)	
	(4)	-0.013	0.044	0.042	0.348*	0.034
		(-1.133)	(0.255)	(0.469)	(1.774)	
MAXRET	(1)	0.010	0.192			0.008
		(1.035)	(1.106)			
	(2)	0.012	0.197	-0.086		0.011
		(1.297)	(1.155)	(-0.642)		
	(3)	0.003	0.211		0.323**	0.049
		(0.289)	(1.369)		(2.196)	
	(4)	0.003	0.211	-0.005	0.321**	0.049
		(0.322)	(1.361)	(-0.050)	(2.294)	

Table 20: This table reports the results on the cryptocurrency factor adjustments of the second set of high-minus-low volatility return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
DELAY	(1)	-0.001	-0.041			0.002
		(-0.237)	(-0.624)			
	(2)	-0.005	-0.048	0.138*		0.042
		(-1.074)	(-0.740)	(1.701)		
	(3)	-0.000	-0.044		-0.041	0.005
		(-0.050)	(-0.641)		(-0.602)	
	(4)	-0.005	-0.048	0.138	-0.002	0.042
		(-0.937)	(-0.713)	(1.572)	(-0.025)	
STDPRCVOL	(1)	-0.024**	-0.203			0.006
		(-1.971)	(-0.641)			
	(2)	-0.008	-0.174	-0.573		0.102
		(-0.722)	(-0.572)	(-1.406)		
	(3)	-0.018*	-0.218		-0.273	0.025
		(-1.772)	(-0.676)		(-1.287)	
	(4)	0.005	-0.195	-0.691*	-0.467	0.155
		(0.380)	(-0.655)	(-1.728)	(-1.616)	
DAMIHUD	(1)	0.018**	0.086			0.001
		(2.038)	(0.325)			
	(2)	0.004	0.060	0.507		0.095
		(0.475)	(0.240)	(1.390)		
	(3)	0.011	0.104		0.326	0.036
		(1.576)	(0.382)		(1.554)	
	(4)	-0.010	0.083	0.634*	0.505*	0.172
		(-0.965)	(0.337)	(1.779)	(1.802)	

the considered return predictors have a significant excess return when considering the full factor model.

4.3.5 Projected Network Return Predictors

We assess the performance of portfolios based on projected network characteristic return predictors, as presented in table 21. Unfortunately, none of the high-minus-low portfolios, based on the projected network characteristics, manage to generate a significant mean weekly return. Additionally, we cannot discern any visible pattern across the different quintile portfolios. It seems that no discernible return can be achieved when considering portfolios created based on exposure to network characteristics.

Considering that these network characteristics are captured only from the Bitcoin blockchain, this result is not entirely unexpected. In section 4.2, we established that only specific network characteristics had a significant influence on the cryptocurrency market returns. Given that the cryptocurrency market return mostly consists of Bitcoin returns, it is logical that these network characteristics have an influence.

When constructing the projections, all cryptocurrencies are considered, totalling more than 18,000 at the height of the boom. As Bitcoin is the largest cryptocurrency currently available, it is logical that not many cryptocurrencies have a sizeable projection with such a developed blockchain. This could be a possible reason for the lack of significant mean weekly returns for the portfolios based on projected network characteristics.

In summary, none of the four network characteristic based return predictors forms high-minus-low portfolios capable of generating significant returns.

We now examine the portfolios formed using return predictors from projected network characteristics, accounting for the coin market factor, the size factor, and the momentum factor. The results of this analysis are presented in table 22.

We observe that only the payments based portfolio exhibits a significant loading on the market factor, for all factor model configurations. Considering the size factor alongside the market factor, we find that none of the portfolios demonstrate a significant loading on the size factor. This result suggests that there is no underlying connection between exposure to a mature blockchain and the size of a cryptocurrency, which is unexpected.

When evaluating the momentum factor in conjunction with the market factor, we notice that no portfolios have a significant loading on the momentum factor. This implies that cryptocurrencies with exposure to a mature blockchain do not necessarily generate more or less cumulative returns.

Ultimately, none of the considered portfolios display a significant excess return, even at the 10% level.

4.3.6 Projected Production Return Predictors

We evaluate the performance of portfolios based on projected production characteristic return predictors, as presented in table 23. Regrettably, none of the high-minus-low portfolios, formed using projected production characteristics, manage to generate a significant mean weekly return. Additionally, we are unable to discern any visible pattern across the different quintile portfolios.

This observation suggests that the differences between Proof-of-Work (PoW) and Proof-of-Stake (PoS) cryptocurrencies do not yield a significant mean weekly return.

Interestingly, this outcome can be interpreted as positive news from an investor's perspective. The lack of significant differences in mean weekly returns between PoW and PoS cryptocurrencies implies that investors are not forced to choose power-hungry PoW cryptocurrencies for their investments. This potentially paves the way for more sustainable investment options in the cryptocurrency market, as investors may opt for environmentally friendly PoS cryptocurrencies without compromising their potential returns.

In summary, while the projected production characteristic return predictors do not display any significant differences in mean weekly returns between PoW and PoS cryptocurrencies, this result could encourage a shift towards more sustainable investment choices in the cryptocurrency market.

We now examine the portfolios formed using return predictors from projected production characteristics, accounting for the coin market factor, the size factor, and the momentum factor. The results of this analysis are presented in table 24.

Our analysis reveals that all projected characteristic portfolios have a large significant loading on the market factor, at the 1% level. This observation indicates that these portfolios are significantly exposed to market risk, further confirming our belief that the distinction between PoW and PoS cryptocurrencies should not result in worse returns.

Considering the momentum factor and the market factor, we find that only the portfolio based on the projections of the consumption of electricity in the United States exhibits a small significant loading on the momentum factor, at the 10% level. This effect is surprising and does not have an immediate explanation. It may warrant further investigation to understand the underlying factors contributing to this relationship.

When evaluating the size factor, we do not observe any significant loading for the portfolios based on projected production characteristics. This suggests that the size factor does not play a major role in the returns of these portfolios.

Ultimately, only the portfolio based on projections of electricity generation in China shows a significant excess return, albeit at the 10% level. This suggests a small difference in return between cryptocurrencies with large exposure to electricity generation in China and those with minimal exposure. The sign of this coefficient might indicate potential for greater excess return from investing in Proof-of-Stake cryptocurrencies as compared to Proof-of-Work cryptocurrencies.

4.3.7 Projected Sentiment Return Predictors

We assess the performance of portfolios based on projected sentiment characteristic return predictors, as illustrated in table 25. Interestingly, none of the high-minus-low portfolios, created using projected sentiment characteristics, manage to generate a significant mean weekly return. Moreover, we cannot discern any consistent pattern across the different quintile portfolios. This observation implies that relying on the exposure of individual cryptocurrencies to social media sentiment as a predictor of potential returns may not yield dependable outcomes.

Our analysis also reveals that cryptocurrencies with considerable exposure to social sentiment do not exhibit significantly higher returns than those with minimal exposure. This finding has positive implications for cryptocurrency investors. It suggests that including cryptocurrencies with various levels of social media sentiment exposure in an investment portfolio does not necessarily impact potential returns, providing investors with greater flexibility in their portfolio composition. However, this finding should be taken with caution, as the quintile portfolios themselves could not generate any discernible return, especially for the highly exposed quintile.

We now explore the portfolios formed using return predictors from projected sentiment characteristics, taking into account the coin market factor, the size factor, and the momentum factor. The results of this analysis are presented in tables 26 and 27. None of the portfolios have a significant loading on either the market factor, the size factor, the momentum factor or any combination of the three, even at the 10% level. Also, none of the portfolios exhibit any significant excess return. This finding indicates that there is no excess return to be gained by investing in cryptocurrencies with a high exposure to social media sentiment.

4.4 Investigating Mechanism

The results from section 4.3 largely align with the findings of Liu and Tsyvinski (2021), except the four significant excess return generating portfolios.

Similar to Liu and Tsyvinski (2021), our study also reveals that most return predictors, which initially generate excess returns, fail to do so after accounting for the market factor, the size factor, and the momentum factor. In their paper, Liu and Tsyvinski propose potential explanations for the observed size and momentum factors.

Firstly, they suggest that the size factor may serve as a proxy for an illiquidity premium. This means that cryptocurrencies with a smaller market cap tend to exhibit higher illiquidity. As a result, investors in these cryptocurrencies demand larger returns as compensation for the risk of not being able to sell their positions in a timely manner. Our analysis supports this notion, as we find that cryptocurrencies with high DAMIHUD illiquidity and low STDPRCVOL liquidity demonstrate a substantial and significant loading on the size factor.

Secondly, Liu and Tsyvinski argue that the momentum factor might be rooted in behavioural explanations, particularly psychological biases and investor over- and under-reaction. However, our analysis does not provide any evidence of an attention-based, overreaction-induced momentum effect. Surprisingly, none of our sentiment measures, including those loosely associated with investor attention, display a significant loading on the momentum factor. Additionally, the volume related return predictors, which function as a proxy for momentum by well-informed investors, also did not yield any significant loadings on the momentum factor.

5 Conclusion

This paper aimed to achieve two primary objectives: replicating the findings of previous research and expanding the scope of those studies. Our primary research objective was to identify the key drivers of cross-sectional variations in cryptocurrency returns.

Our investigation into factors driving cryptocurrency market returns built upon the work of Liu and Tsyvinski (2021) by examining the influence of network and production characteristics of cryptocurrencies. By employing an extensive dataset that includes the post-COVID-19 cryptocurrency boom, we captured a broader view of market trends and dynamics. We hypothesized that production factors would not significantly impact returns due to market maturation, regulatory actions, rising interest rates, and the shift from Proof-of-Work to Proof-of-Stake algorithms. Conversely, we expected network factors, as proxies for adoption and utility, to significantly impact cryptocurrency returns.

Additionally, we explored the impact of investor sentiment on cryptocurrency market returns using deep-learning natural language processing techniques, specifically the XLM-T model published in Barbieri et al. (2021). We expected investor sentiment, particularly social media influence, to have a pronounced effect on cryptocurrency returns, as evidenced by recent market events such as the GameStop saga and Elon Musk's Dogecoin endorsements.

Furthermore, we replicated the work of Liu et al. (2022) by analyzing excess returns of quintile portfolios and examining whether size and momentum factors capture the cross-section of cryptocurrency returns. We hypothesized that our extended dataset, reflecting a more mature cryptocurrency market, would show risk premia converging more closely to those observed in equity markets, implying that the behavior of the cryptocurrency market would become increasingly similar to traditional financial markets.

Our study yielded several significant findings. We discovered that coin market returns are influenced by certain network characteristics, such as the growth rates in active addresses, transactions, and hash rate. We also observed that market returns negatively affect the growth in transactions from the 4- to 8-week horizon and positively affect the growth in hash-rate from the 5- to 8-week horizon. Our analysis revealed that coin market returns are not exposed to production factors but exhibit positive exposure to the growth rate in the number of posts, negative posts, neutral posts, and to a lesser extent, positive posts. Lastly, we discover that the growth in social media posts, including negative, neutral, and, to a lesser extent, positive, contribute positively to long-term cryptocurrency market returns.

We found that only MCAP, PRC, MAXDPRC, 1, 2, and 3-week horizon cumulative return, PRCVOL, VOLSCALED, and STDPRCVOL form high-minus-low portfolios capable of generating significant returns, while none of the return predictors based on projected cryptocurrency characteristics were able to do so. After factor analysis, only MCAP, the 2-week horizon cumulative return, and the 50-week cumulative return predictors generated significant excess return, with the market factor, size factor, and momentum factor explaining the returns observed in the other return predictors. Additionally, we found that investing in cryptocurrencies with low exposure to the Chinese energy market generates a small but significant return, even after adjusting for all three factors.

Our findings largely align with Liu and Tsyvinski (2021), except for the three excess return generating portfolios based on return predictors. Like Liu and Tsyvinski (2021), our study reveals that most return predictors that initially generate excess returns fail to do so after accounting for the market factor, the size factor, and the momentum factor. We supported the notion that the size factor may serve as a proxy for an illiquidity premium, but our analysis did not provide any evidence of an attention-based, overreaction-induced momentum effect, as none of our sentiment measures displayed a significant loading on the momentum factor.

Our study not only enriches the existing body of research but also provides valuable insights for investors and regulators in the ever-evolving cryptocurrency landscape.

References

Abadi, J., & Brunnermeier, M. (2018, December). *Blockchain economics* (w25407). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w25407

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, *61*(1), 259–299. https://doi.org/10.1111/j.1540-6261.2006.00836.x

Athey, S., Parashkevov, I., Sarukkai, V., & Xia, J. (2016, August). *Bitcoin pricing, adoption, and usage: Theory and evidence*. Stanford University, Graduate School of Business. https://EconPapers.repec.org/RePEc:ecl:stabus:3469

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446. https://doi.org/10.1016/j.jfineco.2010.08.014

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0

Barbieri, F., Anke, L. E., & Camacho-Collados, J. (2021). XLM-t: Multilingual language models in twitter for sentiment analysis and beyond [Publisher: arXiv Version Number: 2]. https://doi.org/10.48550/ARXIV.2104.12250

Barry, C. B., & Brown, S. J. (1984). Differential information and the small firm effect. *Journal of Financial Economics*, *13*(2), 283–294. https://doi.org/10.1016/0304-405X(84)90026-6

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., & Blackburn, J. (2020). The pushshift reddit dataset [Publisher: arXiv Version Number: 1]. https://doi.org/10.48550/ARXIV.2001.08435

Beck, E., De Nard, G., & Wolf, M. (2023). Improved inference in financial factor models. *International Review of Economics & Finance*, *86*, 364–379. https://doi.org/10.1016/j.iref.2023.03.009

Bernstein, P. L. (2008). *A primer on money, banking, and gold* [OCLC: ocn233484849]. Wiley.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., & Menkveld, A. J. (2018). Equilibrium bitcoin pricing. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3261063

Borri, N. (2019). Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance*, *50*, 1–19. https://doi.org/10.1016/j.jempfin.2018.11.002

Borri, N., & Shakhnov, K. (2022). The cross-section of cryptocurrency returns (J. Pontiff, Ed.). *The Review of Asset Pricing Studies*, *12*(3), 667–705. https://doi.org/10.1093/rapstu/raac007

Buterin, V. (2013). Ethereum: A next-generation smart contract and decentralized application platform. https://ethereum.org/en/whitepaper/

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chase, B., & MacBrough, E. (2018, February 20). Analysis of the XRP ledger consensus protocol. Retrieved February 6, 2023, from http://arxiv.org/abs/1802.07242

Cheah, E.-T., & Fry, J. (2015). Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Chen, A. Y., & Zimmermann, T. (2020). Publication bias and the cross-section of stock returns (J. Pontiff, Ed.). *The Review of Asset Pricing Studies*, *10*(2), 249–289. https://doi.org/10.1093/rapstu/raz011

Chen, C. Y.-H., Despres, R., Guo, L., & Renault, T. (2019). What makes cryptocurrencies special? investor sentiment and return predictability during the bubble. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3398423

Chordia, T., Subrahmanyam, A., & Anshuman, V. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, *59*(1), 3–32. https://doi.org/10.1016/S0304-405X(00)00080-5

Cong, L. W., & He, Z. (2019). Blockchain disruption and smart contracts. *The Review of Financial Studies*, *32*(5), 1754–1797. https://doi.org/10.1093/rfs/hhz007

Cong, L. W., He, Z., & Li, J. (2021). Decentralized mining in centralized pools (W. Jiang, Ed.). *The Review of Financial Studies*, *34*(3), 1191–1235. https://doi.org/10.1093/rfs/hhaa040

Cong, L. W., Li, Y., & Wang, N. (2021). Tokenomics: Dynamic adoption and valuation (I. Goldstein, Ed.). *The Review of Financial Studies*, *34*(3), 1105–1155. https://doi.org/10.1093/rfs/hhaa089

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2019). Unsupervised cross-lingual representation learning at scale [Publisher: arXiv Version Number: 2]. https://doi.org/10.48550/ARXIV.1911.02116

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis*, *45*(2), 215–233. https://doi.org/10.1016/S0167-9473(02)00366-3

De Bondt, W. F. M., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), 793–805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, *98*(4), 703–738. https://doi.org/10.1086/261703

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding [Publisher: arXiv Version Number: 2]. https://doi.org/10.48550/ARXIV.1810.04805

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, *75*(3), 1327–1370. https://doi.org/10.1111/jofi.12883

George, T. J., & Hwang, C.-Y. (2004). The 52-week high and momentum investing. *The Journal of Finance*, *59*(5), 2145–2176. https://doi.org/10.1111/j.1540-6261.2004.00695.x

Griffin, J. M., & Shams, A. (2020). Is bitcoin really untethered? *The Journal of Finance*, *75*(4), 1913–1964. https://doi.org/10.1111/jofi.12903

Hou, A. J., Wang, W., Chen, C. Y. H., & Härdle, W. K. (2020). Pricing cryptocurrency options*. *Journal of Financial Econometrics*, *18*(2), 250–279. https://doi.org/10.1093/jjfinec/nbaa006

Hou, K., & Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, *18*(3), 981–1020. https://doi.org/10.1093/rfs/hhi023

Hu, A. S., Parlour, C. A., & Rajan, U. (2019). Cryptocurrencies: Stylized facts on a new investible instrument. *Financial Management*, *48*(4), 1049–1068. https://doi.org/10.1111/fima.12300

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Jermann, U. J. (2021). Cryptocurrencies and cagan's model of hyperinflation. *Journal of Macroeconomics*, *69*, 103340. https://doi.org/10.1016/j.jmacro.2021.103340

Lintner, J. (1975). THE VALUATION OF RISK ASSETS AND THE SELECTION OF RISKY INVESTMENTS IN STOCK PORTFOLIOS AND CAPITAL BUDGETS**. In *Stochastic optimization models in finance* (pp. 131–155). Elsevier. https://doi.org/10.1016/B978-0-12-780850-5.50018-6

Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency (I. Goldstein, Ed.). *The Review of Financial Studies*, *34*(6), 2689–2727. https://doi.org/10.1093/rfs/hhaa113

Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common risk factors in cryptocurrency. *The Journal of Finance*, *77*(2), 1133–1177. https://doi.org/10.1111/jofi.13119

MacKinnon, J. G. (2013). Thirty years of heteroskedasticity-robust inference. In X. Chen & N. R. Swanson (Eds.), *Recent advances and future directions in causality, prediction, and specification analysis* (pp. 437–461). Springer New York. https://doi.org/10.1007/978-1-4614-1653-1_17

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, *29*(3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7

Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, *135*(2), 293–319. https://doi.org/10.1016/j.jfineco.2019.07.001

Miller, M. H., & Scholes, M. S. (1982). Dividends and taxes: Some empirical evidence. *Journal of Political Economy*, *90*(6), 1118–1141. https://doi.org/10.1086/261114

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(4), 768. https://doi.org/10.2307/1910098

Nakamoto, S. (2008, October). Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf

Nguyen, C. T., Hoang, D. T., Nguyen, D. N., Niyato, D., Nguyen, H. T., & Dutkiewicz, E. (2019). Proof-of-stake consensus mechanisms for future blockchain networks: Fundamentals, applications and opportunities. *IEEE Access*, *7*, 85727–85745. https://doi.org/10.1109/ACCESS.2019.2925010

Pagnotta, E., & Buraschi, A. (2018). An equilibrium valuation of bitcoin and decentralized network assets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3142022

Romano, J. P., & Wolf, M. (2017). Resurrecting weighted least squares. *Journal of Econometrics*, *197*(1), 1–19. https://doi.org/10.1016/j.jeconom.2016.10.003

Schilling, L., & Uhlig, H. (2019). Some simple bitcoin economics. *Journal of Monetary Economics*, *106*, 16–26. https://doi.org/10.1016/j.jmoneco.2019.07.002

Sedlmeir, J., Buhl, H. U., Fridgen, G., & Keller, R. (2020). The energy consumption of blockchain technology: Beyond myth. *Business & Information Systems Engineering*, *62*(6), 599–608. https://doi.org/10.1007/s12599-020-00656-x

Sharpe, W. F. (1964). CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sockin, M., & Xiong, W. (2020, March). *A model of cryptocurrencies* (w26816). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w26816

Taylor, W. L. (1953). "cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

Treynor, J. L. (1961). Market value, time, and risk. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2600356

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817. https://doi.org/10.2307/1912934

A Projected Cryptocurrency Characteristics Results

A.1 Projected Network Return Predictors

Table 21: This table reports the weekly mean quintile portfolio returns on the projected network characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	$5-1$
Δaddress	0.012**	0.019***	0.012**	0.021***	0.009	-0.003
	(2.248)	(3.009)	(2.223)	(3.016)	(1.317)	(-0.462)
Δtransactions	0.011**	0.019***	0.013**	0.018***	0.009	-0.002
	(2.086)	(3.024)	(2.486)	(2.642)	(1.375)	(-0.243)
Δpayments	0.006	0.007	0.011*	0.012**	0.004	-0.002
	(1.101)	(1.072)	(1.834)	(2.032)	(0.670)	(-0.365)
Δhash-rate	0.013**	0.026**	0.013**	0.017***	0.010	-0.003
	(2.361)	(2.322)	(2.326)	(2.625)	(1.451)	(-0.428)

Table 22: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low projected network return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
Δaddress	(1)	-0.005	0.099			0.006
		(-0.741)	(1.301)			
	(2)	-0.007	0.095	0.075		0.013
		(-1.106)	(1.251)	(1.176)		
	(3)	-0.005	0.099		0.001	0.006
		(-0.672)	(1.260)		(0.012)	
	(4)	-0.007	0.096	0.081	0.024	0.014
		(-1.005)	(1.217)	(1.126)	(0.218)	
Δtransactions	(1)	-0.003	0.106			0.007
		(-0.535)	(1.395)			
	(2)	-0.005	0.102	0.074		0.014
		(-0.896)	(1.345)	(1.163)		
	(3)	-0.003	0.106		0.003	0.007
		(-0.493)	(1.352)		(0.033)	
	(4)	-0.006	0.103	0.081	0.026	0.015
		(-0.841)	(1.309)	(1.128)	(0.237)	
Δpayments	(1)	-0.003	0.135**			0.024
		(-0.500)	(2.522)			
	(2)	-0.002	0.137***	-0.030		0.025
		(-0.384)	(2.595)	(-0.380)		
	(3)	-0.000	0.139**		-0.108	0.036
		(-0.081)	(2.557)		(-1.285)	
	(4)	0.000	0.141***	-0.027	-0.107	0.037
		(0.059)	(2.610)	(-0.351)	(-1.274)	
Δhash-rate	(1)	-0.005	0.121			0.009
		(-0.763)	(1.594)			
	(2)	-0.007	0.118	0.071		0.015
		(-1.112)	(1.546)	(1.126)		
	(3)	-0.005	0.122		0.017	0.009
		(-0.734)	(1.552)		(0.170)	
	(4)	-0.008	0.119	0.081	0.039	0.017
		(-1.063)	(1.510)	(1.164)	(0.360)	

A.2 Projected Production Return Predictors

Table 23: This table reports the weekly mean quintile portfolio returns on the projected production characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	5 − 1
P^{US}	0.015***	0.019***	0.017***	0.024**	0.015**	-0.000
	(3.532)	(2.991)	(2.771)	(2.525)	(2.378)	(-0.069)
Gen^{US}	0.013***	0.017***	0.020***	0.021**	0.016**	0.003
	(3.201)	(2.677)	(3.174)	(2.238)	(2.572)	(0.450)
$Cons^{US}$	0.013***	0.017***	0.018***	0.022**	0.016**	0.003
	(3.182)	(2.729)	(2.844)	(2.310)	(2.570)	(0.468)
P^{CN}	0.012**	0.029**	0.013**	0.020***	0.020**	0.008
	(2.422)	(2.515)	(2.349)	(2.978)	(2.535)	(1.186)
Gen^{CN}	0.017***	0.034**	0.015***	0.021***	0.011*	-0.006
	(3.624)	(2.332)	(2.919)	(2.750)	(1.871)	(-1.174)

Table 24: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low projected production return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
P^{US}	(1)	-0.005	0.293***			0.071
		(-0.974)	(5.194)			
	(2)	-0.004	0.295***	-0.028		0.072
		(-0.841)	(5.143)	(-0.553)		
	(3)	-0.005	0.292***		-0.025	0.072
		(-0.837)	(5.073)		(-0.567)	
	(4)	-0.003	0.293***	-0.037	-0.036	0.074
		(-0.604)	(4.981)	(-0.644)	(-0.714)	
Gen^{US}	(1)	-0.002	0.302***			0.074
		(-0.442)	(5.060)			
	(2)	-0.002	0.302***	-0.015		0.075
		(-0.370)	(5.010)	(-0.287)		
	(3)	-0.001	0.297***		-0.079	0.084
		(-0.111)	(4.879)		(-1.510)	
	(4)	0.001	0.298***	-0.038	-0.090	0.086
		(0.120)	(4.834)	(-0.612)	(-1.621)	
$Cons^{US}$	(1)	-0.002	0.305***			0.075
		(-0.428)	(5.085)			
	(2)	-0.002	0.305***	-0.016		0.076
		(-0.352)	(5.035)	(-0.303)		
	(3)	-0.001	0.300***		-0.081	0.085
		(-0.093)	(4.906)		(-1.531)	
	(4)	0.001	0.301***	-0.039	-0.092*	0.087
		(0.142)	(4.857)	(-0.631)	(-1.647)	
P^{CN}	(1)	0.001	0.461***			0.109
		(0.175)	(4.343)			
	(2)	-0.001	0.458***	0.072		0.114
		(-0.154)	(4.297)	(0.649)		
	(3)	0.002	0.459***		-0.036	0.110
		(0.272)	(4.185)		(-0.287)	
	(4)	-0.000	0.457***	0.068	-0.017	0.115
		(-0.052)	(4.108)	(0.517)	(-0.114)	
Gen^{CN}	(1)	-0.010**	0.264***			0.067
		(-2.170)	(4.467)			
	(2)	-0.009**	0.265***	-0.034		0.069
		(-1.991)	(4.442)	(-0.853)		
	(3)	-0.010**	0.264***		0.001	0.067
		(-2.168)	(4.395)		(0.014)	
	(4)	-0.009*	0.264***	-0.037	-0.010	0.069
		(-1.872)	(4.335)	(-0.759)	(-0.185)	

A.3 Projected Sentiment Return Predictors

Table 25: This table reports the weekly mean quintile portfolio returns on the projected sentiment characteristics. The mean returns are the time-series averages of weekly value-weighted portfolio excess returns. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic.

	Quintiles					
	1	2	3	4	5	$5-1$
ΔN	0.010*	0.020***	0.010*	0.022***	0.008	-0.002
	(1.937)	(3.149)	(1.856)	(3.161)	(1.257)	(-0.284)
ΔN_{pos}	0.011**	0.016***	0.011*	0.021***	0.010	-0.001
	(2.212)	(2.639)	(1.956)	(3.082)	(1.433)	(-0.207)
ΔN_{neg}	0.009*	0.014**	0.010*	0.024***	0.008	-0.000
	(1.673)	(2.340)	(1.796)	(3.166)	(1.244)	(-0.078)
ΔN_{neu}	0.011**	0.019***	0.010*	0.021***	0.009	-0.002
	(2.165)	(3.002)	(1.855)	(2.981)	(1.413)	(-0.292)
ΔS_{avg}	0.012**	0.027**	0.012**	0.019***	0.009	-0.003
	(2.212)	(2.320)	(2.237)	(2.954)	(1.331)	(-0.438)
ΔS_{wavg}	0.011**	0.027**	0.015***	0.016**	0.009	-0.002
	(2.081)	(2.359)	(2.705)	(2.530)	(1.322)	(-0.337)
ΔS_{std}	0.011**	0.028**	0.009*	0.019***	0.010	-0.002
	(2.185)	(2.453)	(1.764)	(2.955)	(1.445)	(-0.260)

Table 26: This table reports the results on the cryptocurrency factor adjustments of the first set of high-minus-low projected sentiment return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
ΔN	(1)	-0.003	0.100			0.006
		(-0.561)	(1.316)			
	(2)	-0.005	0.096	0.075		0.014
		(-0.928)	(1.268)	(1.176)		
	(3)	-0.003	0.100		0.000	0.006
		(-0.503)	(1.272)		(0.004)	
	(4)	-0.006	0.097	0.081	0.023	0.014
		(-0.846)	(1.231)	(1.119)	(0.210)	
ΔN_{pos}	(1)	-0.003	0.120			0.008
		(-0.530)	(1.370)			
	(2)	-0.005	0.116	0.074		0.014
		(-0.874)	(1.324)	(1.146)		
	(3)	-0.004	0.121		0.012	0.008
		(-0.517)	(1.332)		(0.122)	
	(4)	-0.006	0.118	0.083	0.036	0.016
		(-0.863)	(1.296)	(1.168)	(0.316)	
ΔN_{neg}	(1)	-0.002	0.101			0.007
		(-0.350)	(1.344)			
	(2)	-0.004	0.098	0.054		0.011
		(-0.603)	(1.309)	(1.072)		
	(3)	-0.001	0.099		-0.030	0.008
		(-0.219)	(1.288)		(-0.325)	
	(4)	-0.003	0.097	0.050	-0.016	0.011
		(-0.435)	(1.256)	(0.838)	(-0.154)	
ΔN_{neu}	(1)	-0.003	0.094			0.006
		(-0.552)	(1.237)			
	(2)	-0.006	0.090	0.079		0.014
		(-0.937)	(1.186)	(1.209)		
	(3)	-0.003	0.094		0.002	0.006
		(-0.498)	(1.196)		(0.016)	
	(4)	-0.006	0.091	0.085	0.026	0.014
		(-0.861)	(1.154)	(1.165)	(0.231)	

Table 27: This table reports the results on the cryptocurrency factor adjustments of the second set of high-minus-low projected sentiment return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The standard errors are computed using HCCME and ALS. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
ΔS_{avg}	(1)	-0.004	0.092			0.005
		(-0.703)	(1.225)			
	(2)	-0.006	0.088	0.076		0.013
		(-1.077)	(1.174)	(1.205)		
	(3)	-0.004	0.093		0.008	0.006
		(-0.653)	(1.190)		(0.081)	
	(4)	-0.007	0.090	0.084	0.032	0.014
		(-1.004)	(1.147)	(1.193)	(0.288)	
ΔS_{wavg}	(1)	-0.004	0.086			0.005
		(-0.579)	(1.136)			
	(2)	-0.006	0.082	0.078		0.013
		(-0.959)	(1.084)	(1.234)		
	(3)	-0.004	0.086		0.000	0.005
		(-0.522)	(1.099)		(0.004)	
	(4)	-0.006	0.083	0.085	0.024	0.013
		(-0.882)	(1.056)	(1.181)	(0.220)	
ΔS_{std}	(1)	-0.004	0.121			0.009
		(-0.590)	(1.591)			
	(2)	-0.006	0.117	0.077		0.016
		(-0.959)	(1.539)	(1.178)		
	(3)	-0.004	0.121		0.013	0.009
		(-0.570)	(1.546)		(0.129)	
	(4)	-0.007	0.118	0.086	0.037	0.018
		(-0.933)	(1.501)	(1.202)	(0.335)	

B Factor Models without HCCME and ALS

B.1 Size-Related Return Predictors

Table 28: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low size return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
MCAP	(1)	-0.043***	-0.099			0.003
		(-4.399)	(-1.144)			
	(2)	-0.009**	-0.036	-1.253***		0.841
		(-2.345)	(-1.036)	(-51.850)		
	(3)	-0.052***	-0.077		0.378***	0.071
		(-5.347)	(-0.918)		(6.103)	
	(4)	-0.010**	-0.035	-1.247***	0.024	0.842
		(-2.474)	(-1.003)	(-49.700)	(0.915)	
PRC	(1)	-0.051*	-0.479**			0.008
		(-1.918)	(-2.075)			
	(2)	-0.024	-0.430*	-0.972***		0.078
		(-0.939)	(-1.930)	(-6.210)		
	(3)	-0.053**	-0.472**		0.118	0.009
		(-1.993)	(-2.043)		(0.685)	
	(4)	-0.020	-0.438*	-1.015***	-0.171	0.080
		(-0.740)	(-1.964)	(-6.248)	(-0.990)	
MAXDPRC	(1)	-0.047*	-0.403*			0.006
		(-1.750)	(-1.720)			
	(2)	-0.022	-0.356	-0.926***		0.068
		(-0.826)	(-1.569)	(-5.805)		
	(3)	-0.050*	-0.394*		0.159	0.007
		(-1.858)	(-1.679)		(0.911)	
	(4)	-0.019	-0.362	-0.954***	-0.112	0.068
		(-0.693)	(-1.590)	(-5.760)	(-0.638)	
AGE	(1)	-0.007	-0.111			0.003
		(-0.656)	(-1.196)			
	(2)	0.011	-0.078	-0.657***		0.202
		(1.131)	(-0.938)	(-11.268)		
	(3)	-0.014	-0.091		0.341***	0.051
		(-1.359)	(-1.005)		(5.060)	
	(4)	0.006	-0.070	-0.615***	0.166***	0.213
		(0.637)	(-0.851)	(-10.218)	(2.610)	

B.2 Momentum-Related Return Predictors

Table 29: This table reports the results on the cryptocurrency factor adjustments of the first set of high-minus-low momentum return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
$r^{(1,0)}$	(1)	0.021**	0.031			0.000
		(2.383)	(0.399)			
	(2)	0.017*	0.023	0.156***		0.016
		(1.896)	(0.300)	(2.863)		
	(3)	0.012	0.055		0.416***	0.101
		(1.452)	(0.746)		(7.544)	
	(4)	0.003	0.046	0.281***	0.496***	0.149
		(0.363)	(0.635)	(5.347)	(8.895)	
$r^{(2,0)}$	(1)	0.031***	-0.134*			0.007
		(3.782)	(-1.878)			
	(2)	0.040***	-0.118*	-0.318***		0.085
		(4.966)	(-1.721)	(-6.586)		
	(3)	0.012**	-0.083*		0.887***	0.549
		(2.166)	(-1.711)		(24.704)	
	(4)	0.015***	-0.079*	-0.101***	0.858***	0.556
		(2.727)	(-1.652)	(-2.884)	(23.198)	
$r^{(3,0)}$	(1)	0.023***	-0.155**			0.010
		(2.899)	(-2.263)			
	(2)	0.031***	-0.140**	-0.298***		0.084
		(4.027)	(-2.122)	(-6.427)		
	(3)	0.003	-0.100**		0.937***	0.667
		(0.606)	(-2.523)		(31.661)	
	(4)	0.005	-0.098**	-0.065**	0.918***	0.670
		(1.060)	(-2.477)	(-2.264)	(30.020)	
$r^{(4,0)}$	(1)	0.014*	-0.175**			0.013
		(1.805)	(-2.583)			
	(2)	0.023***	-0.158**	-0.343***		0.113
		(3.122)	(-2.454)	(-7.577)		
	(3)	-0.003	-0.128***		0.820***	0.523
		(-0.627)	(-2.700)		(23.323)	
	(4)	0.001	-0.123***	-0.146***	0.779***	0.540
		(0.260)	(-2.640)	(-4.313)	(21.700)	
$r^{(4,1)}$	(1)	0.011	0.006			0.000
		(1.337)	(0.084)			
	(2)	0.023***	0.028	-0.432***		0.142
		(2.931)	(0.414)	(-9.166)		
	(3)	-0.000	0.037		0.536***	0.194
		(-0.043)	(0.573)		(11.072)	
	(4)	0.010	0.048	-0.319***	0.445***	0.266
		(1.403)	(0.773)	(-7.056)	(9.276)	

Table 30: This table reports the results on the cryptocurrency factor adjustments of the second set of high-minus-low momentum return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
$r^{(8,0)}$	(1)	-0.004	-0.032			0.000
		(-0.386)	(-0.375)			
	(2)	0.012	-0.003	-0.574***		0.181
		(1.320)	(-0.041)	(-10.576)		
	(3)	-0.021***	0.016		0.819***	0.327
		(-2.620)	(0.224)		(15.719)	
	(4)	-0.008	0.029	-0.395***	0.706***	0.407
		(-1.038)	(0.441)	(-8.237)	(13.902)	
$r^{(16,0)}$	(1)	-0.004	-0.196**			0.010
		(-0.380)	(-2.313)			
	(2)	0.015*	-0.161**	-0.704***		0.281
		(1.836)	(-2.222)	(-13.835)		
	(3)	-0.020**	-0.152**		0.764***	0.294
		(-2.405)	(-2.115)		(14.297)	
	(4)	-0.002	-0.133**	-0.550***	0.608***	0.448
		(-0.222)	(-2.096)	(-11.877)	(12.374)	
$r^{(50,0)}$	(1)	-0.008	0.095**			0.008
		(-1.431)	(2.017)			
	(2)	-0.008	0.095**	-0.001		0.008
		(-1.405)	(2.014)	(-0.027)		
	(3)	-0.010*	0.101**		0.103***	0.025
		(-1.833)	(2.158)		(2.960)	
	(4)	-0.011*	0.100**	0.027	0.111***	0.026
		(-1.957)	(2.137)	(0.795)	(3.063)	
$r^{(100,0)}$	(1)	-0.021	-0.001			0.000
		(-1.592)	(-0.009)			
	(2)	-0.017	0.007	-0.151*		0.007
		(-1.261)	(0.058)	(-1.901)		
	(3)	-0.021	0.001		0.028	0.000
		(-1.620)	(0.005)		(0.329)	
	(4)	-0.016	0.006	-0.155*	-0.016	0.007
		(-1.204)	(0.051)	(-1.879)	(-0.186)	

B.3 Volume-Related Return Predictors

Table 31: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low volume return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
VOL	(1)	-0.013*	0.245***			0.031
		(-1.833)	(4.068)			
	(2)	-0.009	0.252***	-0.135***		0.051
		(-1.296)	(4.216)	(-3.216)		
	(3)	-0.012*	0.244***		-0.017	0.032
		(-1.762)	(4.044)		(-0.384)	
	(4)	-0.007	0.249***	-0.150***	-0.060	0.054
		(-1.035)	(4.171)	(-3.448)	(-1.298)	
PRCVOL	(1)	-0.026*	-0.191			0.005
		(-1.874)	(-1.594)			
	(2)	-0.010	-0.162	-0.568***		0.094
		(-0.771)	(-1.418)	(-7.083)		
	(3)	-0.020	-0.206*		-0.270***	0.023
		(-1.453)	(-1.737)		(-3.051)	
	(4)	0.003	-0.183	-0.686***	-0.464***	0.144
		(0.217)	(-1.646)	(-8.459)	(-5.404)	
VOLSCALED	(1)	-0.020*	-0.329***			0.019
		(-1.675)	(-3.118)			
	(2)	-0.010	-0.310***	-0.377***		0.069
		(-0.836)	(-3.011)	(-5.219)		
	(3)	-0.021*	-0.327***		0.029	0.019
		(-1.708)	(-3.096)		(0.369)	
	(4)	-0.008	-0.313***	-0.398***	-0.084	0.071
		(-0.627)	(-3.047)	(-5.313)	(-1.059)	

B.4 Volatility-Related Return Predictors

Table 32: This table reports the results on the cryptocurrency factor adjustments of the first set of high-minus-low volatility return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
BETA	(1)	-0.016	0.250***			0.015
		(-1.518)	(2.781)			
	(2)	-0.015	0.251***	-0.030		0.015
		(-1.420)	(2.794)	(-0.468)		
	(3)	-0.016	0.251***		0.026	0.015
		(-1.555)	(2.793)		(0.382)	
	(4)	-0.015	0.252***	-0.025	0.019	0.016
		(-1.443)	(2.799)	(-0.379)	(0.267)	
BETA2	(1)	-0.016	0.249***			0.015
		(-1.526)	(2.771)			
	(2)	-0.015	0.250***	-0.030		0.015
		(-1.427)	(2.784)	(-0.471)		
	(3)	-0.016	0.250***		0.025	0.015
		(-1.562)	(2.783)		(0.378)	
	(4)	-0.015	0.251***	-0.025	0.018	0.015
		(-1.449)	(2.789)	(-0.383)	(0.261)	
IDIOVOL	(1)	-0.000	0.119*			0.006
		(-0.049)	(1.757)			
	(2)	-0.004	0.111*	0.151***		0.026
		(-0.574)	(1.658)	(3.212)		
	(3)	0.002	0.113*		-0.095*	0.013
		(0.211)	(1.678)		(-1.890)	
	(4)	-0.003	0.109	0.137***	-0.056	0.028
		(-0.365)	(1.620)	(2.805)	(-1.082)	
RETVOL	(1)	-0.004	0.025			0.000
		(-0.343)	(0.229)			
	(2)	-0.003	0.027	-0.047		0.001
		(-0.238)	(0.251)	(-0.607)		
	(3)	-0.011	0.045		0.337***	0.034
		(-0.921)	(0.415)		(4.205)	
	(4)	-0.013	0.043	0.041	0.348***	0.034
		(-1.009)	(0.402)	(0.528)	(4.190)	
MAXRET	(1)	0.010	0.192**			0.008
		(0.921)	(2.036)			
	(2)	0.012	0.196**	-0.086		0.011
		(1.123)	(2.082)	(-1.304)		
	(3)	0.003	0.211**		0.323***	0.049
		(0.287)	(2.280)		(4.699)	
	(4)	0.003	0.211**	-0.005	0.322***	0.049
		(0.295)	(2.278)	(-0.073)	(4.503)	

Table 33: This table reports the results on the cryptocurrency factor adjustments of the second set of high-minus-low volatility return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
DELAY	(1)	-0.001	-0.041			0.002
		(-0.226)	(-0.952)			
	(2)	-0.005	-0.048	0.138***		0.042
		(-0.988)	(-1.135)	(4.642)		
	(3)	-0.000	-0.044		-0.041	0.005
		(-0.049)	(-1.007)		(-1.278)	
	(4)	-0.005	-0.048	0.138***	-0.002	0.042
		(-0.959)	(-1.136)	(4.451)	(-0.062)	
STDPRCVOL	(1)	-0.024*	-0.203*			0.006
		(-1.792)	(-1.748)			
	(2)	-0.008	-0.174	-0.572***		0.102
		(-0.646)	(-1.577)	(-7.363)		
	(3)	-0.018	-0.219*		-0.273***	0.025
		(-1.354)	(-1.900)		(-3.181)	
	(4)	0.005	-0.196*	-0.690***	-0.469***	0.155
		(0.387)	(-1.821)	(-8.822)	(-5.652)	
DAMIHUD	(1)	0.018	0.086			0.001
		(1.471)	(0.828)			
	(2)	0.004	0.061	0.506***		0.095
		(0.331)	(0.613)	(7.248)		
	(3)	0.011	0.105		0.326***	0.036
		(0.896)	(1.026)		(4.273)	
	(4)	-0.010	0.084	0.634***	0.506***	0.172
		(-0.929)	(0.882)	(9.141)	(6.888)	

B.5 Projected Network Return Predictors

Table 34: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low projected network return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
Δaddress	(1)	-0.005	0.099*			0.006
		(-0.707)	(1.759)			
	(2)	-0.007	0.095*	0.075*		0.013
		(-1.009)	(1.695)	(1.893)		
	(3)	-0.005	0.099*		0.001	0.006
		(-0.703)	(1.757)		(0.024)	
	(4)	-0.007	0.096*	0.081**	0.024	0.014
		(-1.093)	(1.712)	(1.971)	(0.552)	
Δtransactions	(1)	-0.003	0.106*			0.007
		(-0.510)	(1.899)			
	(2)	-0.005	0.102*	0.074*		0.014
		(-0.814)	(1.835)	(1.894)		
	(3)	-0.003	0.106*		0.003	0.007
		(-0.515)	(1.898)		(0.074)	
	(4)	-0.006	0.103*	0.081**	0.026	0.015
		(-0.911)	(1.854)	(1.986)	(0.604)	
Δpayments	(1)	-0.003	0.135**			0.024
		(-0.500)	(2.557)			
	(2)	-0.002	0.137**	-0.030		0.025
		(-0.344)	(2.587)	(-0.584)		
	(3)	-0.000	0.139***		-0.107*	0.036
		(-0.074)	(2.645)		(-1.797)	
	(4)	0.000	0.141***	-0.027	-0.107*	0.037
		(0.049)	(2.670)	(-0.531)	(-1.778)	
Δhash-rate	(1)	-0.005	0.121**			0.009
		(-0.728)	(2.148)			
	(2)	-0.007	0.118**	0.071*		0.015
		(-1.013)	(2.088)	(1.790)		
	(3)	-0.005	0.122**		0.016	0.009
		(-0.774)	(2.161)		(0.389)	
	(4)	-0.008	0.119**	0.081**	0.039	0.017
		(-1.162)	(2.118)	(1.966)	(0.903)	

B.6 Projected Production Return Predictors

Table 35: This table reports the results on the cryptocurrency factor adjustments of the high-minus-low projected production return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
P^{US}	(1)	-0.005	0.293***			0.071
		(-0.927)	(6.166)			
	(2)	-0.004	0.294***	-0.028		0.072
		(-0.782)	(6.184)	(-0.824)		
	(3)	-0.005	0.292***		-0.025	0.072
		(-0.822)	(6.127)		(-0.704)	
	(4)	-0.003	0.293***	-0.037	-0.035	0.074
		(-0.592)	(6.144)	(-1.054)	(-0.963)	
Gen^{US}	(1)	-0.002	0.301***			0.074
		(-0.419)	(6.339)			
	(2)	-0.002	0.302***	-0.015		0.075
		(-0.342)	(6.344)	(-0.440)		
	(3)	-0.001	0.297***		-0.079**	0.084
		(-0.109)	(6.261)		(-2.254)	
	(4)	0.001	0.298***	-0.038	-0.090**	0.086
		(0.115)	(6.280)	(-1.095)	(-2.467)	
$Cons^{US}$	(1)	-0.002	0.304***			0.075
		(-0.406)	(6.386)			
	(2)	-0.002	0.305***	-0.016		0.076
		(-0.326)	(6.391)	(-0.464)		
	(3)	-0.001	0.300***		-0.080**	0.085
		(-0.093)	(6.308)		(-2.279)	
	(4)	0.001	0.301***	-0.039	-0.092**	0.087
		(0.137)	(6.327)	(-1.127)	(-2.500)	
P^{CN}	(1)	0.001	0.461***			0.109
		(0.155)	(7.822)			
	(2)	-0.001	0.458***	0.072*		0.114
		(-0.126)	(7.787)	(1.734)		
	(3)	0.002	0.459***		-0.036	0.110
		(0.267)	(7.776)		(-0.827)	
	(4)	-0.000	0.457***	0.068	-0.017	0.115
		(-0.055)	(7.761)	(1.566)	(-0.372)	
Gen^{CN}	(1)	-0.010**	0.263***			0.067
		(-2.033)	(5.986)			
	(2)	-0.009*	0.265***	-0.034		0.069
		(-1.830)	(6.015)	(-1.096)		
	(3)	-0.010**	0.264***		0.001	0.067
		(-2.015)	(5.975)		(0.022)	
	(4)	-0.009*	0.264***	-0.037	-0.010	0.069
		(-1.745)	(5.994)	(-1.131)	(-0.285)	

B.7 Projected Sentiment Return Predictors

Table 36: This table reports the results on the cryptocurrency factor adjustments of the first set of high-minus-low projected sentiment return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
ΔN	(1)	-0.003	0.100*			0.006
		(-0.537)	(1.806)			
	(2)	-0.005	0.096*	0.075*		0.014
		(-0.849)	(1.741)	(1.941)		
	(3)	-0.003	0.100*		0.000	0.006
		(-0.533)	(1.803)		(0.007)	
	(4)	-0.006	0.097*	0.081**	0.023	0.014
		(-0.935)	(1.758)	(2.015)	(0.546)	
ΔN_{pos}	(1)	-0.003	0.120**			0.008
		(-0.494)	(2.042)			
	(2)	-0.005	0.116**	0.074*		0.014
		(-0.782)	(1.981)	(1.793)		
	(3)	-0.004	0.121**		0.012	0.008
		(-0.527)	(2.050)		(0.280)	
	(4)	-0.006	0.118**	0.083*	0.036	0.016
		(-0.914)	(2.007)	(1.938)	(0.790)	
ΔN_{neg}	(1)	-0.002	0.101*			0.007
		(-0.340)	(1.853)			
	(2)	-0.004	0.098*	0.054		0.011
		(-0.567)	(1.803)	(1.415)		
	(3)	-0.001	0.099*		-0.030	0.008
		(-0.235)	(1.818)		(-0.735)	
	(4)	-0.003	0.097*	0.050	-0.016	0.011
		(-0.489)	(1.788)	(1.263)	(-0.370)	
ΔN_{neu}	(1)	-0.003	0.094*			0.006
		(-0.530)	(1.700)			
	(2)	-0.006	0.090	0.079**		0.014
		(-0.857)	(1.632)	(2.032)		
	(3)	-0.003	0.094*		0.001	0.006
		(-0.529)	(1.698)		(0.035)	
	(4)	-0.006	0.091*	0.085**	0.026	0.014
		(-0.953)	(1.651)	(2.117)	(0.601)	

Table 37: This table reports the results on the cryptocurrency factor adjustments of the second set of high-minus-low projected sentiment return predictor portfolios. The standard t-statistic is reported in parentheses. *, **, and *** denote the significance levels at the 10%, 5%, and 1% levels based on the standard t-statistic. The data frequency is weekly.

		const	MKT	SMB	MOM	R^2
ΔS_{avg}	(1)	-0.004	0.092*			0.005
		(-0.670)	(1.667)			
	(2)	-0.006	0.088	0.076*		0.013
		(-0.984)	(1.601)	(1.962)		
	(3)	-0.004	0.093*		0.008	0.006
		(-0.689)	(1.672)		(0.190)	
	(4)	-0.007	0.090	0.084**	0.032	0.014
		(-1.104)	(1.626)	(2.088)	(0.743)	
ΔS_{wavg}	(1)	-0.004	0.086			0.005
		(-0.552)	(1.542)			
	(2)	-0.006	0.082	0.078**		0.013
		(-0.876)	(1.475)	(2.010)		
	(3)	-0.004	0.086		0.000	0.005
		(-0.547)	(1.540)		(0.007)	
	(4)	-0.006	0.083	0.085**	0.024	0.013
		(-0.964)	(1.493)	(2.087)	(0.566)	
ΔS_{std}	(1)	-0.004	0.121**			0.009
		(-0.562)	(2.147)			
	(2)	-0.006	0.117**	0.077*		0.016
		(-0.875)	(2.083)	(1.949)		
	(3)	-0.004	0.121**		0.012	0.009
		(-0.597)	(2.156)		(0.296)	
	(4)	-0.007	0.119**	0.086**	0.037	0.018
		(-1.017)	(2.111)	(2.105)	(0.850)	