Machine Forecast Disagreement in the Cryptocurrency Market*

Gang Chu†
Dehua Shen‡
Zhaobo Zhu§

June 2026

Abstract

This paper uses machine learning models to construct a crypto-level measure of simulated investor belief disagreement (MFD) for cryptocurrencies. Simulated investors use their individual models and common data information to form their individual return forecasts. Consistent with Miller (1977), we find a significantly negative relation between MFD and future cryptocurrency returns in the cross section. Past return variables are the top drivers of MFD, suggesting that disagreement is strongly associated with information contained in past returns. Moreover, the negative MFD-return relation is stronger for cryptocurrencies with larger limits-to-arbitrage or more severe overpricing, highlighting the role of mispricing and limits-to-arbitrage.

* We would like to thank an anonymous referee, Stephan Siegel (the editor), and seminar and conference participants for their helpful comments. This work is supported by the National Natural Science Foundation of China (72371138; 72401222; 72495155; 72531006), and the China Postdoctoral Science Foundation under Grant Number 2024M762452, 2025T181003, and GZB20230549. Zhu acknowledges that this study is funded by Audencia Foundation. This study is supported by Research Center for FinTech and Digital-Intelligent Management, Shenzhen University; Digital Asset Innovation Research Center, Shenzhen University; and Shenzhen Humanities & Social Sciences Key Research Bases. This paper received the Best Paper Award in 2025 Conference of Association for Big Data and Business Analytics, SMSEC. The authors contributed equally to this work. Authors are listed in alphabetical order.
† Economics and Management School, Wuhan University, Wuhan, China, Email: chugang@whu.edu.cn.
‡ School of Finance, Nankai University, Tianjin, China. Email: dhs@nankai.edu.cn.
§ Corresponding author: Shenzhen Audencia Financial Technology Institute, Shenzhen University, China; Audencia Business School, Nantes, France. Email: zb.zhu@szu.edu.cn.

I. Introduction

Cryptocurrencies have been the most popular and important emerging assets since the rapid growth in recent years, despite ongoing debate about their intrinsic values (e.g., Böhme, Christin, Edelman and Moore, 2015; Griffin and Shams, 2020; Howell, Niessner and Yermack, 2020; Sockin and Xiong, 2023; Li, Shin, and Wang, 2025).[1] Unlike traditional financial assets such as equities and bonds, cryptocurrencies do not have well-accepted fundamental collaterals. Therefore, it is challenging to precisely price cryptocurrencies. Recent studies use some common asset factors such as size and past returns to identify the relative valuation of cryptocurrencies in the cross section (e.g., Borri and Shakhnov, 2022; Liu, Tsyvinski and Wu, 2022; Fieberg et al., 2025; Lee and Wang, 2025), and some studies use cryptocurrency-specific factors (e.g., Cong, Li and Wang, 2021; Liu and Tsyvinski, 2021; Sockin and Xiong, 2023; Babiak and Bianchi, 2025).

In this paper, we examine the relative valuation of cryptocurrencies in the cross section from the perspective of heterogeneous beliefs. Specifically, we follow the approach in Bali, Kelly, Mörke and Rahman (2026) to construct a measure of simulated investor disagreement on the future values of cryptocurrencies.[2] We argue that this machine forecast disagreement (MFD) measure is particularly suitable for cryptocurrencies. There

[1] According to CoinMarketCap, the aggregate market capitalization of cryptocurrencies around the world has been $2.67 trillion until April 2025.

[2] In this paper, "investor" refers to hypothetical belief agents used to construct the MFD measure, rather than observed or identifiable cryptocurrency market participants. Consequently, investor disagreement should be interpreted as disagreement across simulated beliefs inferred from the model, rather than disagreement among actual investor in the cryptocurrency market.

is no widely accepted proxy for investor disagreement in the crypto market. Researchers could use price or earnings forecasts issued by professional analysts to construct proxies for disagreement for stocks, but analyst coverage for cryptocurrencies is limited (Cong et al., 2021). Therefore, we cannot construct a disagreement measure for most cryptocurrencies based on analyst forecast data as prior studies do for stocks (e.g., Diether, Malloy, and Scherbina, 2002).

Moreover, because of a lack of information about future cash flows and expected amounts of cryptocurrencies, it is harder for investors to precisely price cryptocurrencies than traditional assets. Investor disagreement arises because of a lack of fundamental information. In particular, because the cryptocurrency market is dominated by retail investors who have limited information processing ability and suffer from various behavioral biases (Kogan, Makarov, Niessner and Schoar, 2024), investor disagreement on cryptocurrencies' values would be large. Therefore, we need a novel approach to measure investor disagreement for cryptocurrencies. The machine forecast disagreement measure proposed by Bali et al. (2026) simulates the distribution of a large number of different beliefs by endowing each simulated investor with a (machine learning) model with random variation in model specification across investors.[3] That is, MFD directly simulates a number of investors with different beliefs in a market. As a matter of fact, Bali et al. (2026) show that their MFD is better than other proxies such as analyst forecast disagreement in

[3] We emphasize that MFD is a flexible belief-simulation framework that can accommodate various simple and sophisticated belief formation rules.

identifying the relation between belief disagreement and future stock returns in the cross section.

Following the approach in Bali et al. (2026), we construct a novel measure of belief disagreement among simulated investors in the cryptocurrency market. Our results show that MFD has a strong return predictability in the cross section for cryptocurrencies. We find that cryptos with higher MFD at the portfolio formation week tend to experience significantly lower returns in the subsequent week compared with those with lower MFD.[4] Specifically, the decile portfolio of cryptos with the highest MFD significantly underperforms the decile portfolio of cryptos with the lowest MFD by 1.79% per week in terms of value-weighted market-adjusted returns. The significant return predictability of MFD is robust after controlling for well-established factors in bivariate portfolio analyses and Fama-MacBeth regressions. Our main results hold in various robustness tests.

Moreover, the MFD approach allows us to identify the most important factors driving the disagreement among investors. Because retail investors dominate in the crypto market and they tend to extrapolate past returns (e.g., Da, Huang, and Jin, 2021; Giglio, Maggiori, Stroebel, and Utkus, 2021), we examine whether a return-based MFD that only relies on past-return variables has a strong return predictability.[5] By comparison, we also construct a non-return-based MFD that relies on other variables such as size, age, liquidity, etc. Our results show that the return-based MFD has a stronger return predictability than the non-

[4] The terms cryptocurrency and crypto are used interchangeably in this paper.
[5] We thank the anonymous referee for proposing this insightful suggestion.

return-based MFD. Moreover, a return-plus-based MFD that adds additional non-return information does not have a stronger return predictability than the return-based MFD. These results suggest that past-return information plays an important role in generating disagreement in the crypto market.

Then we explore the sources of the return predictability of MFD. The negative relation between MFD and expected crypto returns is consistent with the theoretical model of Miller (1977) that asset prices are upward biased when investors disagree with asset value and short-sale constraints exist for pessimistic investors. Consistent with the argument of Miller (1977), we find that the negative relation between MFD and future returns is significantly stronger among cryptos with higher limits-to-arbitrage. The high-low MFD portfolio has an average weekly value-weighted excess return of –0.71% (t-value is –1.29) and –3.74% (t-value is –3.05) among cryptos with low and high limits-to-arbitrage, respectively.[6]

In addition, we find that the negative MFD-return relation is more pronounced among cryptos with higher mispricing index,[7] suggesting that mispricing explains the MFD-return relation. The high-low MFD portfolio has an average weekly value-weighted excess return of –0.95% (t-value is –1.82) and –2.46% (t-value is –3.16) among low- and high-

[6] Because limits to arbitrage such as arbitrage risk and short-sale constraints in the absolute value are still high for cryptocurrencies in the low limits-to-arbitrage tercile portfolio in our analysis, the MFD-return relation is significantly negative among cryptocurrencies in the low limits-to-arbitrage tercile portfolio.

[7] Following Stambaugh, Yu and Yuan (2015), we establish cryptocurrency mispricing score (MISP) on two well-documented return predictors in the cryptocurrency literature: size and momentum (Liu et al., 2022). The mispricing measure (MISP) is defined as the arithmetic average of the ranks of the two crypto return predictors, and higher (lower) MISP indicates overvaluation (undervaluation). Detailed definition please see Section IV.D.

mispriced cryptos, respectively. Correspondingly, we also find that the negative MFD-return relation is significantly weaker following low crypto sentiment periods than following high sentiment periods. Specifically, the high-low MFD portfolio has an average weekly value-weighted excess return of –1.16% (t-value is –1.43) and –2.87% (t-value is –3.44) following low and high sentiment periods, respectively.

The negative MFD-return relation remains in various robustness tests. In particular, we compare MFD with other proxies for disagreement in the cryptocurrency market. In particular, Garfinkel, Hsiao and Hu (2025) construct a turnover-based variable to measure investor belief disagreement for cryptocurrencies. They show that their turnover measure significantly and negatively predicts crypto returns in the cross section in the daily frequency. Our results show that MFD still significantly predicts future returns even after controlling for turnover, while turnover becomes weak in predicting returns in the weekly frequency.

Our study contributes to the literature in three main ways. First, to our best knowledge, we are the first to propose a novel and valid measure of belief disagreement for cryptocurrencies, which is different from traditional proxies such as abnormal turnover in the crypto market (Garfinkel et al., 2025). The return predictability of crypto MFD remains after controlling for abnormal turnover, which does not significantly predict returns in our setting. That is, we contribute to the emerging literature on the pricing of cryptocurrencies by adding a new return predictor.[8] Even controlling for these crypto characteristics, MFD still significantly and negatively predicts future cryptocurrency returns in the cross section. Moreover, because MFD as a flexible belief-simulation framework, we provide evidence that past-return information plays an important role in generating disagreement in the crypto market.

Second, we adopt machine learning models to develop a novel and valid measure of belief disagreement for the cryptocurrency market, supporting the usefulness of machine learning models in finance, especially research in cryptocurrencies. In particular, our study complements prior studies that use machine learning models to construct a measure of belief disagreement for stocks (Bali et al., 2026), corporate culture (Li, Mai, Shen and Yan, 2021), cyber risk (Jiang, Khanna, Yang and Zhou, 2024), firm-level climate risk exposure (e.g., Li, Shan, Tang and Yao, 2024), and cryptocurrency returns (e.g., Cakici, Shahzad, Bedowska-Sojka, and Zaremba, 2024). We differ from Cakici et al. (2024) by using machine learning models to construct a new pricing factor.

Third, our study contributes to the debate on the relation between belief disagreement and future returns. Existing theoretical and empirical studies offer two opposite views. On one hand, some studies such as Merton (1987) and Johnson (2004) argue that disagreement is a source of risk and should predict positive future returns. On the other hand, Miller

[8] Prior studies document some significant cryptocurrency return predictors, including size, age, price, momentum, price volume, price volatility, failure risk, past returns of peer cryptos, network and production factors, and so on (e.g., Cong et al., 2021; Liu and Tsyvinski, 2021; Liu et al., 2022; Borri and Shakhnov, 2022; Lyandres, Palazzo and Rabetti, 2022; Zhu, Ma, and Tu, 2025).

(1977) and others argue that the market prices reflect more about the opinions of optimistic investors when short-sale constraints are binding. Prior empirical studies provide mixed evidence in the stock and bond markets. Our study provides novel evidence from the cryptocurrency market to support the argument on the negative relation between investor disagreement and future returns.

II. Data and Variables

We obtain cryptocurrency daily price and trading data from CoinMarketCap, the most reliable source for publicly available cryptocurrencies. It provides cryptocurrency daily data, including prices (open, close, high, and low), trading volume, and market capitalization. A key advantage of CoinMarketCap is its comprehensive coverage of both active and defunct cryptocurrencies (Liu et al., 2022), thus mitigating survivorship bias.

Our sample comprises 3024 cryptocurrencies from January 2013 to June 2024. Given that the majority of cryptocurrencies were created and listed after the final week of 2014, we begin our sample period in January 2015. Following common practices to address issues related to illiquidity and micro-cap bias (e.g., Bali et al., 2026), we exclude cryptocurrencies with daily price under \$0.1, daily trading volume less than \$100,000, market capitalization below \$1 million, and trading history shorter than 52 weeks.[9]

[9] The reason is that newly born cryptocurrencies may harm the overall estimation accuracy of our analyses because of the great volatility.

To construct MFD, we use 40 cryptocurrency-specific characteristics as the complete information set,[10] and randomly select a subset of 24 characteristics to represent an investor-specific incomplete information set. The MFD for week$-t$ is calculated based on characteristics during the week $t-1$, using a 104-week (2-year) rolling window to estimate the random forecast regressor. This procedure yields an MFD measure over an out-of-sample period from January 2017 to June 2024. Finally, to ensure a sufficient cross-sectional sample for asset pricing tests, we restrict our empirical analysis to the period January 2019 to June 2024.

A. MFD Construction

To model the belief formation process, we adopt the framework of Gu, Kelly and Xiu (2020) to describe the relationship between cryptocurrency expected returns and characteristics. The general prediction model is as follow:

$$r_{i,t+1} = E_t\left[r_{i,t+1}\right] + \varepsilon_{i,t+1} \tag{1}$$

where $r_{i,t+1}$ denotes the return of the cryptocurrency $i = 1, \dots, I$, in week $t = 1, \dots, T$, and $E_t\left[r_{i,t+1}\right]$ is the investor-formed expectation of the return based on available information:

$$E_t\left[r_{i,t+1}\right] = g(z_{i,t}) \tag{2}$$

Here, $z_{i,t}$ represents the information set for cryptocurrency i in the time t. $g(\cdot)$ is a

[10] We provide the detailed information of 40 characteristics in the Appendix A.

return-generating function that maps characteristics to return expectation.

Investor belief disagreement is commonly attributed to two main sources, information asymmetry (Grossman and Stiglitz, 1976; Kyle, 1985; Wang, 1994) and heterogeneous priors (Kandel and Pearson, 1995; Hong and Stein, 2007). To incorporate both channels, we propose a belief-generating model in which differences in return expectations arise from variation in investors' access to and interpretation of information.

To capture information asymmetry, we assume that investors observe different subsets of the complete information set. Each investor k has access to an incomplete information set:

$$z_{k,i,t} \in R^{d_k}, 1 \leq d_k < d \tag{3}$$

where $z_{k,i,t}$ is a subset of complete $d - dimension$ information set $z_{i,t}$, and the dimensionality d_k varies across investors. Each investor accesses the information set through a personalized selection of characteristics:

$$z_{k,i,t} = [z_{i,t}^{l_{k,1}}, \dots, z_{i,t}^{l_{k,d_k}}] \tag{4}$$

where $z_{i,t}^{l} \in R$ and denote the $l - th$ dimension of $z_{i,t}$, $l_k = \left\{l_{k,i} = \{1, \dots, d\}\right\}, and\ |l_k| = d_k$. The dimension of the investor-specific information might be different. This setup reflects heterogeneous attention, behavioral bias, and varying information processing capability across investors.

To simulate heterogeneous priors, we model each simulated investor's belief formation using a random forest regression in line with Gu et al. (2020) and Bali et al. (2026). Random forest regressions are particularly effective for capturing non-linearities and complex interactions in high-dimensional data and well-suited to represent investor-specific forecasting processes. Importantly, random forest regression simulates heterogeneous priors across simulated investors by introducing a heterogeneous layer that randomizes regression specifications across simulated investors through the use of bootstrapping and dropout.

Each investor k forecasts future returns using a random forest model trained on their unique subset $z_{k,i,t}$:

$$g_k\left(z_{k,i,t}\right) = RF_k(z_{k,i,t}) \tag{5}$$

The random forest function is defined as:

$$RF_k\left(z_{k,i,t}\right) = \frac{1}{N}\sum_{n}^{N} h_n^L(z_{k,i,t}; D^{(n)}, \Theta^{(n)}) \tag{6}$$

where N is the total number of decision trees in the forest, $h_n^L(\cdot)$ is the prediction function of the $n-th$ regression tree with the maximum depth L, $D^{(n)}$ is the bootstrap sample used to train tree n (randomly drawn with replacement from the original dataset, and $\Theta^{(n)}$ is the set of random feature selection decisions used to build tree n (i.e., which features are considered at each node split).

To implement this model empirically, we define the complete information set $z_{i,t}$ as a vector of 40 cryptocurrency characteristics. Each investor k is assigned an incomplete information set of dimensions, which are randomly selected 24 characteristics from the full set of 40 characteristics. We simulate$K = 100$ investors, each using a random forest with the following hyperparameters: the maximum tree depth is 3, the number of trees in the ensemble is 2000, the friction of features is 0.1, and the friction of the sample is 0.05.

Following Kelly, Pruitt and Su (2019) and Gu et al. (2020), we rank all characteristics cross-sectionally each week and map them to the interval [-1, +1]. When training the random forest regressors, return outliers are winsorized at 1st and 99th percentiles cross-sectionally. Each random forest model is estimated using a 104-week (2-year) rolling window and is refitted every 12 weeks to reduce computational burden. Finally, investor belief disagreement is measured by the machine forecast disagreement (MFD), defined as the cross-sectional standard deviation of the 100 investor-specific return forecasts for each cryptocurrency in a given week.

$$MFD_{i,t} = \sqrt{\frac{1}{100}\sum_{k=1}^{100}(RF_{i,k,t} - \overline{RF_{i,t}})^2} \tag{7}$$

B. Control Variables

Following the existing literature on investor disagreement and cryptocurrencies (e.g., Liu and Tsyvinski, 2021; Liu et al., 2022; Kogan et al., 2024; Bali et al., 2026; Garfinkel et al., 2025), we incorporate a broad set of characteristics that are documented to explain the cross-section of expected cryptocurrency returns.

We begin with market beta ($Beta$), a fundamental factor in asset pricing, computed as the slope coefficient from a regression of daily cryptocurrency returns on the value-weighted market returns over the prior 90 days. To control size effects, we follow Liu et al. (2022) and use the logarithm of market capitalization ($Size$) as a control variable.

To account for momentum effect, we follow Liu et al. (2022) and include the past-week return (MOM) to mitigate potential confounding influences on the MFD-return relation. Motivated by Atilgan, Bali, Demirtas and Gunaydin (2020), we also include the abnormal trading volume ($Abvol$) as control variable, defined as the difference between the logarithm of dollar trading volume in a given week and the average of the logarithm of trading volumes over the preceding 12 weeks.

We further control for several liquidity and risk related variables. We follow Amihud (2002) and define illiquidity ($ILLIQ$) as the average weekly ratio of the absolute daily returns to dollar trading volume. Following Ang, Hodrick, Xing and Zhang (2009), we define idiosyncratic volatility ($Ivol$) as the standard deviation of residuals from a regression of excess cryptocurrency returns on excess market returns over the previous 90 days. Moreover, we use the standard deviation of daily returns over the same period to control for return volatility ($Volt$).

To capture investor preferences for lottery-like payoffs, we follow Bali, Cakici and Whitelaw (2011) and use the average of the five highest daily returns over the prior five weeks to control for MAX. We also include turnover ($Turnover$), calculated as the logarithm of the ratio of average daily dollar volume to the market capitalization (Chordia, Subrahmanyam and Anshuman, 2001). Lastly, to account for asymmetry in return distributions, we use the coefficient on the squared excess market return term from a regression of daily excess cryptocurrency returns on both the linear and squared excess market returns over the past 90 days to control for co-skewness ($Coskew$).

C. Descriptive Statistics

Figure 1 presents the weekly time-series of aggregate MFD in the cryptocurrency market. The plot reveals that the cross-sectional median MFD tends to rise during periods of adverse market conditions, suggesting that disagreement intensifies in market downturns.

<Insert Figure 1 Here>

To capture the cross-sectional distribution for MFD and other characteristics of cryptocurrencies, we first compute weekly cross-sectional statistics and Table 1 reports weekly time-series averages. As Table 1 shows, MFD exhibits a time-series mean of 0.10, a median of 0.08, and a standard deviation of 0.07. The substantial cross-sectional dispersion in MFD is noteworthy, indicating considerable heterogeneity in investor disagreement across cryptocurrencies. This pronounced right-skewness in the MFD distribution aligns with the findings of Garfinkel et al. (2025).

<Insert Table 1 Here>

Table 2 reports the cross-sectional Spearman rank correlation coefficients between

MFD and other control variables. Consistent with theoretical predictions and empirical evidence (e.g., Bali et al., 2026), we observe a negative association between MFD and one-week-ahead excess returns, suggesting that higher disagreement is linked to lower future returns. Specifically, cryptocurrencies with smaller size, lower liquidity, stronger momentum, higher abnormal trading volume, lower market beta, higher idiosyncratic and total return volatility, greater lottery demand, higher turnover, and lower co-skewness tend to exhibit higher MFD values.

In particular, we find a strong negative (positive) correlation between MFD and market capitalization (idiosyncratic volatility), suggesting that MFD may capture information uncertainty in the crypto market, consistent with the arguments in Zhang (2006) and Bali et al. (2026).

<Insert Table 2 Here>

III. MFD and the Cross-Section of Crypto Returns

In this section, to investigate the return predictability of investor belief disagreement in the cryptocurrency market, we employ the univariate and bivariate portfolio analyses, and Fama-MacBeth (1973) regression analysis using MFD as the proxy for investor belief disagreement.

A. Univariate Portfolio Analysis

We begin by employing the univariate portfolio analysis to investigate the cross-sectional relationship between MFD and future cryptocurrency returns. For each week from January 2019 to June 2024, all cryptocurrencies are sorted into ten decile portfolios based on their MFD, where portfolio High contains cryptocurrencies with the highest MFD and portfolio Low contains cryptocurrencies with the lowest MFD. Then, we calculate (both equal-weighted and value-weighted) the time-series average of one-week-ahead excess return of each decile.

Besides, we compute two types of risk-adjusted returns using the CCAPM model with the cryptocurrency market factor (CMKT), [11] and Liu et al. (2022) three-factor model with CMKT, size (CSMB)[12], and momentum (CMOM)[13]. This allows us to assess whether the return spread between High-MFD and Low-MFD portfolios could be explained by standard risk factors.

Table 3 presents the results for univariate portfolio analysis. Specifically, for equal-weighted portfolios, excess returns and alphas decrease monotonically across MFD deciles. Specifically, the excess return decreases from 1.20% (t-value is 1.74) to –0.03% (t-value

---

[11] CMKT is the cryptocurrency excess market returns, defined as the difference between cryptocurrency market return and the risk-free rate. In this study, we measure the risk-free rate as the one-month Treasury bill rate.

[12] CSMB is the cryptocurrency size factor. For each week, cryptocurrencies are sorted into terciles based on market capitalization, and we form value-weighted portfolios for the bottom 30% (small-cap), the middle 40% (middle-cap), and top 30% (large-cap) groups. The CSMB is defined as the return spread between the small-cap and large-cap portfolios.

[13] CMOM is the cryptocurrency momentum factor. For each week, cryptocurrencies are sorted into terciles based on three-week momentum, and we form value-weighted portfolios for the bottom 30% (small-mom), middle 40% (middle-mom), and top 30% (large-mom) groups. The CMOM is defined as the return spread between the small-mom and large-mom portfolios.

is –0.04) when moving from the portfolio High to the portfolio Low. The long-short portfolio that short-sells cryptocurrencies in the lowest MFD decile and takes a long position in cryptocurrencies in the highest MFD decile earns an average return of 1.22% per week, and statistically significant at the 1% level with a t-value of 2.60. Notably, the magnitude and statistical significance of the return spreads on MFD-sorted portfolios does not change after controlling for well-accepted risk factors. The CCAPM and three-factor alphas for the long-short portfolio is 1.09% with a t-values of 2.54 and 1.06% with a t-value of 2.53.

A similar MFD-return pattern holds for the value-weighted portfolios. More interestingly, we find that both return and alpha spreads on the value-weighted MFD-sorted portfolios is larger in absolute magnitude and more statistically significant than that on the equal-weighted MFD-sorted portfolios. In addition, the alphas of portfolio High are significantly negative and large in absolute magnitude across all factor models, indicating that the negative MFD-return relation is primarily concentrated on the short leg of the arbitrage portfolio. That is, cryptocurrencies with high disagreement appear to be overpriced relative to those with lower disagreement.

We note that the negative MFD-return relation is stronger based on value-weighted returns than equal-weighted returns. The cryptocurrency data quality and price/return outliers are two main concerns. While our sample selection criteria could mitigate the potential data errors and extreme outliers to some extent, concerns about the potential dataset errors and extreme outliers may still exists. After addressing these extreme return outliers, the negative MFD-return relation is similar in terms of equal-weighted and value-weighted returns. We discuss this issue in detail in the Internet Appendix A.

Overall, these results are consistent with existing literature documenting a negative association between investor disagreement and future returns (e.g., Miller, 1977; Hong and Stein, 2007; Berkman, Dimitrov, Jain, Koch and Tice, 2009; Yu, 2011; Bali et al., 2026; Garfinkel et al., 2025).

<Insert Table 3 Here>

Next, we examine the cross-sectional persistence of MFD ranking across time. Following Atilgan et al. (2020), we estimate the average transition probabilities of a cryptocurrency remaining in or moving across MFD deciles over 12-week horizon. If MFD evolved randomly, the transition probabilities would be close to 10% for each decile.

Table 4 shows a strong persistence. For instance, 27.97% of cryptocurrencies in the lowest MFD decile (decile 1) remain in the same decile after 12 weeks, while 28.70% of those in the highest MFD decile (decile 10) remain in the highest MFD decile. Diagonal transition probabilities across other deciles ranges from 12.79% to 17.50%, all substantially higher than the random benchmark. Moreover, the probability of transitioning from the lowest MFD decile to the highest MFD decile or vice versa is notably lower than 10%, reinforcing the notion that MFD is a persistent characteristic in the cross-section of cryptocurrencies.

<Insert Table 4 Here>

To further investigate the persistence of MFD-based return predictability, we examine the long-term performance of MFD-based portfolios by calculating the weekly excess returns and three-factor alphas from 2 to 12 weeks after portfolio formation. As Table 5 shows, for equal-weighted portfolios, the return and three-factor alpha spreads on MFD-sorted portfolios are statistically insignificant beyond the first week. Consistent with the main results, we find a more persistent MFD-return relation in terms of value-weighted portfolio returns[14]. The significantly predictive power of MFD persists up to 5 weeks after the portfolio formation, indicating that large cryptocurrencies exhibit more persistent mispricing related to disagreement.

<Insert Table 5 Here>

In summary, the univariate portfolio analysis documents that the negative cross-sectional association between MFD and future cryptocurrency returns is economically meaningful and robust in both short term and over intermediate horizons, especially for value-weighted portfolios. These results underscore the relevance of investor disagreement as a key source of return predictability in cryptocurrency market.

B. Portfolio Characteristics

To address the possibility that the negative MFD-return relation could be driven by

[14] The long-term performance of MFD is also affected by the potential data errors and extreme outliers. We discuss it in detail in the Internet Appendix A.

other related characteristics, we control some well-known characteristics that may be useful in explaining the cross-sectional cryptocurrency returns. Specifically, Liu and Tsyvinski (2021) and Liu et al. (2022) emphasize that market return, size, and momentum are the primary driving factors for expected cryptocurrency returns. Additionally, Garfinkel et al. (2025) document a negative association between abnormal trading volume and future cryptocurrency returns.

We also include several other prominent return predictors from the equity literature, despite evidence that their predictive power is generally weaker in the cryptocurrency market (Liu et al., 2022). These variables include volatility, idiosyncratic volatility, the maximum daily returns, and liquidity.

Before proceeding with the bivariate portfolio analysis, we investigate how these crypto-specific characteristics correlate with investor disagreement, as proxied by MFD. Specifically, we sort cryptocurrencies by MFD into ten decile portfolios each week, and compute the time-series means of the cross-sectional averages of various crypto-specific characteristic within each decile portfolio, as well as the difference between High-MFD and Low-MFD portfolios. As Table 6 shows, mean MFD increases substantially across the deciles, from 0.06 in portfolio Low to 0.31 in portfolio High, consistent with Detzel, Liu, Strauss, Zhou and Zhu (2021) that cryptocurrencies are difficult to value due to the extreme heterogeneity and information frictions.

We find that, consistent with Kumar (2009) and Bali et al. (2011), cryptocurrencies with higher MFD have smaller market capitalization, lower liquidity, higher idiosyncratic volatility, and greater lottery-like features. These features have been interpreted as signals of elevated information uncertainty. In addition, we observe that cryptocurrencies with higher MFD exhibit higher turnover, consistent with Barinov (2014), who argues that turnover is positively related to uncertainty.

Moreover, High-MFD cryptocurrencies tend to exhibit abnormally low trading volume, stronger momentum, higher market beta, and higher overall return volatility. While no statistically significant difference in co-skewness is observed, it is retained as the control variables in subsequent analyses due to its theoretical relevance in capturing an asset's sensitivity to market-wide shocks and its potential influence on the pricing of investor disagreement.

<Insert Table 6 Here>

C. Bivariate Portfolio Analysis

Next, we examine robustness of the negative MFD-return relation by controlling for a broad set of well-known return predictors. These return predictors include momentum (MOM), market capitalization ($Size$), abnormal trading volume ($Abvol$), market ($Beta$), idiosyncratic volatility ($Ivol$), return volatility ($Volt$), the demand for lottery-like crypto assets (MAX), Co-skewness ($Coskew$), illiquidity ($ILLIQ$), and turnover ($Turnover$).

To implement the bivariate portfolio analysis, we conduct 5×10 dependent double sorts based on various crypto-specific characteristics and MFD. For each week, cryptocurrencies are sorted into five quintiles based on the ascending order of a given crypto-specific characteristic. Subsequently, we sort cryptocurrency within each crypto-specific characteristic quintile into ten decile portfolios based on MFD. This process generates 50 conditionally double-sorted portfolios for each crypto-specific characteristic. Due to space constraints, we do not report results for all 50 portfolios individually. Instead, we only report ten combined portfolios in Table 7, where portfolio Low (High) consists of cryptocurrencies with the lowest (highest) MFD in each crypto-specific characteristic quintile.

As Table 7 shows, the negative MFD-return relation remains economically meaningful and statistically significant after controlling for a broad set of well-established return predictors. For example, after controlling for momentum, we observe a return spreads of MFD-sorted portfolios is –1.01% (–1.58%) per week, with corresponding t-value of –2.30 (–2.71). The associated three-factor alpha is –0.82% (–1.21%) per week, with a t-value of –2.08 (–2.26).

Moreover, when we control for other crypto-specific characteristics, the return spreads of MFD-sorted portfolios range from –1.45% (t-value is –3.35) to –0.74% (t-value is –1.97) per week for equal-weighted portfolios, and from –2.09% (t-value is –3.03) to –1.31% (t-value is –2.09) per week for value-weighted portfolios. The three-factor alpha spreads on MFD-sorted portfolios remain significantly negative regardless of the crypto-specific characteristic, with the exception of the idiosyncratic volatility and return volatility in the value-weighted portfolios. These findings provide strong evidence that the MFD-based return predictability is not subsumed by other cross-sectional return predictors.

<Insert Table 7 Here>

D. Fama-MacBeth Regressions

While portfolio analysis is a widely used non-parametric approach that effectively reduces noise and highlights broad patterns, it omits large amount of cross-sectional information and cannot control for multiple factors simultaneously. To address this concern, we use the Fama and MacBeth (1973) regressions to test the MFD-based return predictability. This approach allows for a more precise estimation of the partial effect of MFD while accounting for a broad set of return predictors.

Each week, we run cross-sectional regressions (i.e., across cryptocurrencies) of one-week-ahead excess returns on MFD as well as a broad set of control variables. As discussed above, we control MOM, $Size$, $Abvol$, Beta, $Ivol$, $Volt$, MAX, $Coskew$, $ILLIQ$, and $Turnover$. We follow Newey and West (1987) and correct the standard errors of the average slope coefficients for heteroskedasticity and autocorrelation, and report the time-series average of slope coefficients in Table 8.

As Table 8 shows, the relationship between MFD and future cryptocurrency returns remains significantly negative across all specifications. The univariate regression in Column (1) shows that the time-series average slope coefficient of MFD is -7.80 and statistically significant at the 1% level with a t-value of -3.54, indicating a strong inverse relationship. In columns (2-11), we sequentially include each control variable in the cross-sectional regression. While the magnitude of the MFD coefficient varies slightly, ranging from -8.70 to -5.61, all coefficients remain significant, with t-values exceeding 2.58. Other cryptocurrency characteristics do not materially attenuate the predictive power of MFD. Overall, these results provide strong evidence that MFD contains incremental, value-relevant information about future cryptocurrency returns.

<Insert Table 8 Here>

IV. Economic Mechanisms

In this section, we further explore the underlying economic mechanisms that may drive this relation, shedding light on whether the MFD-return relation reflects mispricing or risk compensation in the crypto market.

A. The Role of Past Returns

The MFD approach allows us to identify the most important factors driving the disagreement among investors. The MFD as a flexible belief-simulation framework can accommodate various simple and sophisticated belief formation rules. Because retail investors dominate in the crypto market (Kogan et al., 2024) and they tend to extrapolate past returns (e.g., Da et al., 2021; Giglio et al., 2021; Lin et al., 2025), we examine whether a return-based MFD that relies more on past-return variables has a strong return predictability.

Figure 2 plots the average feature importance across the 100 simulated investors obtained from the random forest forecasts used to construct MFD. Characteristics related to past returns and price dynamics emerge as the most influential predictors, broadly consistent with extrapolative behavior. Meanwhile, many non-return variables also contribute to forecast variation.

To further assess the role of return extrapolation in generating belief dispersion in the cryptocurrency market, we implement three belief-formation regimes within the MFD framework that differ in investors' degree of extrapolative beliefs. First, we model expectations as being formed purely through extrapolation from past returns. That is, we construct a return-based MFD that relies only on 9 past return variables. For comparison, we construct a non-return-based MFD that relies on other variables such as size, age, liquidity, etc. We also construct a return-plus-based MFD that 9 key past return variables and 15 non-return variables.

Table 9 reports the results. Our results show that the return-based MFD has a stronger return predictability than the non-return-based MFD. The excess return spread for the return-based MFD is –1.98% with a *t*-value of –3.36, while the return spread for the non-return based MFD is only −1.09% with a t-value of −1.81, although the difference is insignificant. Moreover, a return-plus-based MFD that adds additional non-return information does not have a stronger return predictability than the return-based MFD, suggesting the incremental information of non-return variables is limited if we consider key past return variables in constructing MFD. We also confirm the similar results holds on 13 past return variables (Internet Appendix B). Overall, these results suggest that the MFD is primarily driven by past-return information, a pattern that is broadly consistent with extrapolative behavior documented in the cryptocurrency literature. We emphasize that our evidence should not be interpreted as directly identifying extrapolative beliefs. These findings also support the interpretation of MFD as a flexible belief-simulation framework that can accommodate behavioral heuristics such as extrapolation by varying the relative importance assigned to different information sources.

<Insert Figure 2 Here>

<Insert Table 9 Here>

B. MFD and Limits to Arbitrage

In this section, we investigate whether the MFD-based return predictability is driven by limits to arbitrage. Motivated by Miller (1977)'s hypothesis, we expect that return predictability of MFD should be more pronounced for cryptocurrencies with greater arbitrage frictions.

Shleifer and Vishny (1997) argue that arbitrage activity is risky and costly. Thus, mispricing may persist when the costs of arbitrage outweigh the potential gains. Consistent with Zhang (2006) and Lam and Wei (2011), we measure limits to arbitrage with three fundamental dimensions: (1) arbitrage risk, measured by idiosyncratic volatility; (2) information uncertainty, measured by cryptocurrency age and size; (3) transaction costs, measured by Amihud (2002)'s illiquidity. Cryptocurrencies with higher idiosyncratic volatility, smaller market capitalization, shorter trading history, and lower liquidity are more likely to face pronounced limits to arbitrage.

We first examine the return predictability of MFD conditional on different proxies for limit-to-arbitrage using bivariate portfolio analysis. Each week, cryptocurrencies are sorted into three terciles based on four limit-to-arbitrage proxies, idiosyncratic volatility, age, size, and illiquidity, respectively. Then, within each limit-to-arbitrage tercile, cryptocurrencies are further sorted into five quintile portfolios based on MFD.

As shown in Panel C-F of Table 10, the negative relation between MFD and future returns is significantly stronger among cryptocurrencies that are younger and exhibit higher return volatility. Moreover, the negative MFD-return relation is also more pronounced among illiquidity cryptocurrencies. These findings indicate that arbitrage frictions amplify the return predictability of MFD.

An exception arises for size, the negative MFD-return relation is weak and statistically insignificant among smallest cryptocurrencies. One plausible explanation is that a subset of "small and high-MFD" cryptocurrencies experience extreme positive returns, which may obscure the underlying relation[15]. To address the influence of extreme outliers, we winsorize weekly returns at the 1st and 99th percentiles and repeat the bivariate portfolio analysis (see Internet Appendix C).

Consistent with expectations, after winsorization, the negative MFD-return relation is strongest among youngest, most volatile and illiquid cryptocurrencies. More importantly, the negative MFD-return relation among smallest cryptocurrencies also becomes stronger and statistically significant. Taken together, these findings are consistent with the predictions of Miller (1977).

To provide a comprehensive measure for limits-to-arbitrage, we construct a composite index using these four aforementioned proxies. Following Atilgan et al. (2020), each week, we rank cryptocurrencies in ascending order according to their idiosyncratic volatility and illiquidity, and in descending order according to their size and age. Then, cryptocurrencies are given a score of their decile rank for each limits-to-arbitrage proxy. The crypto-specific arbitrage index (ARB) is calculated as the sum of these four scores, ranging from 4 to 40, where higher values indicate greater arbitrage costs for a cryptocurrency.

We first calculate the time-series average of the cross-sectional ARB index for cryptocurrencies in MFD quintile portfolios and reported in Panel A of Table 10. We observe that the average ARB index increases sharply across MFD deciles. The difference

[15] We discuss this issue in Internet Appendix A.

in ARB index between High-MFD and Low-MFD portfolios is 3.70 with a t-value of 21.78, indicating that high-MFD cryptocurrencies are more difficult to arbitrage.

Next, to explore the interaction between arbitrage costs and MFD-based return predictability, we conduct dependent double sorts on ARB and MFD. Specifically, we sort cryptocurrencies into three tercile portfolios based on their ARB index each week. Subsequently, cryptocurrencies are sorted into five MFD-based quintiles within each ARB tercile. We calculate return and three-factor alpha spreads between High-MFD and Low-MFD portfolios within each ARB tercile.

As Panel B of Table 10 shows, the return and three-factor alpha spreads on value-weighted MFD-based portfolios increase monotonically in absolute magnitude moving from low-ARB to high-ARB terciles. Specifically, the return and three-factor alpha spreads on MFD-sorted portfolios are –3.74% per week with t-value of –3.05 and –3.29% with a t-value of –3.12 for cryptocurrencies with high arbitrage cost, i.e., in the high ARB tercile. In contrast, these spreads are smaller and often insignificant in the low ARB tercile. More importantly, the difference in the return (three-factor alpha) spread between high ARB and low ARB terciles is –3.03% (–3.11%) with a t-value of –2.41 (–2.83). This result suggests that the negative cross-sectional MFD-return relation is more pronounced for cryptocurrencies with high arbitrage costs, consistent with the role of short-sale constraints in Miller (1977).

<Insert Table 10 Here>

C. MFD and Investor Sentiment

Lee and So (2015) conceptually argue that investor sentiment could better explain the origin of mispricing and limits to arbitrage could better explain the persistence of mispricing. Stambaugh, Yu, and Yuan (2012) show that stock market anomalies are more pronounced following high sentiment. Cen, Lu, and Yang (2013) examine the breadth–return relation by jointly considering sentiment and disagreement. In particular, Kim, Ryu, and Seo (2014) show that the negative relationship between investor disagreement and future returns is stronger during high sentiment in the stock market. Di Francesco and Hommes (2025) examine the sentiment-driven speculation in the cryptocurrency market in a framework of heterogenous beliefs. In this section, we test how the predictability of MFD varies conditional on different investor sentiment.

We use the *CMC Crypto Fear and Greed Index* from CoinMarketCap as the measure of weekly investor sentiment in the cryptocurrency market.[16] Following standard practice, we classify each week as either a high-sentiment or low-sentiment period based on whether the sentiment index is above or below the annual median. After that, we repeat the univariate MFD portfolio analysis separately for high-sentiment and low-sentiment periods.

As Table 11 shows, the negative MFD-return relation is more pronounced following

---

[16] The CMC Fear and Greed Index, developed by coinmarketcap.com, serves as a widely recognized analytical tool for assessing prevailing sentiment in the cryptocurrency market. This index ranges from 1 to 100, where lower scores denote extreme fear and higher scores indicate extreme greed. The dynamics of the sentiment level help us capture the emotion state of market participants. Notably, this index also offers potential insights into whether the cryptocurrency market may be overvalued (extreme greed) or undervalued (extreme fear).

high-sentiment periods. The return spread on the value-weighted (equal-weighted) MFD-sorted portfolios is –2.87% (–2.18%) per week following high sentiment periods, and statistically significant at the 1% level with a t-value of –3.44 (–2.92). Moreover, the three-factor alpha spread is statistically and economically significant.

In contrast, the return and three-factor alpha spreads on MFD-sorted portfolios are statistically insignificant following low sentiment periods across most specifications. Notably, the return (three-factor alpha) difference between high- and low-sentiment periods are statistically significant on the equal-weighted portfolios. The return spread is economically significant based on value-weighted returns. These findings suggest that the overpricing of high MFD cryptocurrencies is concentrated during high-sentiment periods and that the cryptocurrency market exhibits behavioral dynamics similar to those documented in the stock market.

<Insert Table 11 Here>

D. Mispricing versus Risk

The negative cross-sectional MFD-return relation challenges the interpretation of disagreement as a proxy for uncertainty or risk (e.g., Merton and Others, 1987; Johnson, 2004). According to Miller (1977), when investor disagreement is high, short-sale constraints impede pessimistic views to be fully incorporated into asset prices, leading to overpricing and subsequently lower future returns. If the MFD-based return predictability is indeed driven by mispricing, then the MFD-return relation should be more pronounced for overpriced cryptocurrencies.

To test this mispricing argument, we construct a mispricing index (MISP) by adapting the approach in Stambaugh, Yu and Yuan (2015). Given the relative immaturity of the cryptocurrency market and the limited number of established anomalies compared to traditional asset classes, we base our mispricing measure on two well-documented return predictors in the cryptocurrency literature: size and momentum (Liu et al., 2022). Each week, cryptocurrencies are independently ranked according to these two characteristics, with higher ranks indicating higher future returns, and thus, greater underpricing. The mispricing index (MISP) for each cryptocurrency is computed as the arithmetic average of its two ranks, where a higher MISP value indicates greater overpricing.

We first report the time-series average of cross-sectional MISP score for cryptocurrencies in MFD quintile portfolios. Panel A of Table 12 shows that cryptocurrencies with high MFD indeed have higher average MISP score than cryptocurrencies with low MFD. The high-minus-low difference in average MISP score is 0.34, and statistically significant at the 1% level. This provides evidence that cryptocurrencies with high disagreement is more likely to be overpriced.

To further explore whether the return predictability of MFD is concentrated in overpriced cryptocurrencies, we perform bivariate portfolio sorts. Specifically, we sort cryptocurrencies into three terciles based on MISP score each week. Subsequently, cryptocurrencies are sorted into five quintiles based on MFD within each MISP decile.

As shown in Panel B of Table 12, the return and three-factor alpha spreads on value-weighted MFD-sorted portfolios are significantly negative and large in the absolute magnitude for overvalued cryptocurrencies, compared to the return and three-factor alpha spreads on value-weighted MFD-sorted portfolios for undervalued cryptocurrencies. In the high-MISP tercile (i.e., overvalued cryptocurrencies), the return spread MFD-sorted portfolios is −2.46% with a t-value of −3.16, and the corresponding three-factor alpha spread on MFD-sorted portfolios is −1.87% with a t-value of −2.75 per week. However, in the low-MISP tercile (i.e., undervalued cryptocurrencies), the return and three-factor alpha spreads are insignificant. Notably, the difference in return spreads is −1.51% with t-value of −1.84, indicating that the difference of the return spreads of the overvalued cryptocurrencies (High-MISP tercile) vs. undervalued cryptocurrencies (Low-MISP tercile) also generate difference in MFD premium.

Taken together, these results demonstrate that the negative MFD-return relationship is concentrated for the most overpriced cryptocurrencies, providing evidence consistent with a mispricing interpretation.

<Insert Table 12 Here>

V. Additional Analyses

In this section, we provide additional evidence that MFD captures investor disagreement and have robust predictive power on cryptocurrency future returns. First, we compare MFD to turnover-based disagreement proxy in return predictability. Then, we examine the predictive power of MFD for trading volume and volatility. Finally, we conduct several robustness tests, including transaction costs, alternative construction of MFD, extensive sample periods, and alternative sample selection criteria.

A. Comparing MFD to Turnover-Based Disagreement

Existing literature employs a variety of proxies to measure investor disagreement. Trading turnover has been widely used as a disagreement proxy in both equity and cryptocurrency markets. For example, Boehme, Danielsen and Sorescu (2006) use turnover directly, while Garfinkel (2009) construct changes in market-adjusted turnover to isolate disagreement-driven trading from liquidity effects and market-wide volume trends. More recently, Garfinkel et al. (2025) employ abnormal trading volume as a proxy for investor disagreement in the cryptocurrency market.

In this section, we compare the return predictability of turnover-based disagreement measures with that of MFD. Following Garfinkel (2009), we construct the change in turnover (ΔTO) as an alternative proxy for investor belief disagreement.[17] We first examine the correlation between ΔTO and MFD using portfolio sorts. Panel A of Table 13 shows a monotonic increase in ΔTO across MFD quintiles, with a statistically significant spread of 0.206 between the highest and lowest quintiles, indicating a strong positive

[17] The detail construction information presented in Appendix B.

correlation between ΔTO and MFD.

To further assess the relative strength of the cross-sectional return predictability of MFD and ΔTO, we conduct a bivariate portfolio analysis. Each week, cryptocurrencies are first sorted into three terciles based on ΔTO. Within each ΔTO tercile, cryptocurrencies are then sorted into 5 quintiles based on MFD. Panel B of Table 13 reports the results. The value-weighted return and three-factor alpha spreads of MFD-sorted portfolios are statistically significant across all ΔTO terciles, with the exception of the three-factor alpha in the middle ΔTO tercile. In the highest ΔTO tercile, the return and three-factor alpha spreads are –2.11% (with a t-value of –2.86) and –1.65% (with a t-value of –2.75), respectively. However, the impact of belief disagreement is not significantly amplified when turnover-based disagreement is more severe. Although the spreads are larger in absolute magnitude in the highest ΔTO tercile than in the lowest ΔTO tercile, the differences are not statistically significant.

We next revisit the predictive power of ΔTO at different frequencies. Garfinkel et al. (2025) document that trading-based disagreement predicts daily cryptocurrency returns. Consistent with prior findings, the average daily return and alpha spreads between high-ΔTO and low-ΔTO decile portfolios are significantly negative for both equal- and value-weighted portfolios. In contrast, the average weekly return and alpha spreads are statistically insignificant for the both weighting schemes. Thus, while ΔTO exhibits predictive power at the daily frequency, its ability to predict returns dissipates at the weekly horizon (see Internet Appendix D).

Finally, we compare the relative performance of MFD and ΔTO using cryptocurrency-level Fama-MacBeth regressions that control for other well-known return predictors. Table 14 shows that ΔTO does not exhibit predictive power for future cryptocurrency returns. More importantly, including ΔTO does not affect the statistical or economic significance of MFD.

Taken together, these findings suggest that MFD contains incremental predictive information beyond that captured by ΔTO at a weekly frequency. Turnover-based disagreement proxies, which are derived from trading activity, capture short-lived fluctuations in disagreement that tend to resolve within few days. By contrast, MFD, constructed using a 2-year rolling window and a rich set of characteristics, appears to capture a more persistent and fundamental component of belief disagreement. Moreover, because MFD is model-based and inherently forward-looking, it may better reflect differences in expectations embedded in investor beliefs. We argue that both measures are complementary in different frequencies.

<Insert Table 13 Here>

<Insert Table 14 Here>

B. Explanatory Power for Trading Volume and Volatility

A large body of literature documents that greater belief disagreement is associated with higher trading activity (Garfinkel, 2009; Atmaz and Basak, 2018; Cookson and Niessner, 2020; Cookson, Fos, and Niessner, 2021; and Bali et al., 2026). Atmaz and Basak (2018) develop a theoretical model in which heterogenous investors become more active as belief dispersion increase, thereby generating higher trading volume. Bali et al. (2026) provide complementary empirical evidence, showing that a machine-learning-based disagreement is an important determinant of trading volume. In addition, both theoretical and empirical studies suggest that belief disagreement is positively related to market volatility (Atmaz and Basak, 2018; Bali et al., 2026).

Building on these findings from the equity market, we examine the impact of investor belief disagreement on trading volume and volatility in the cryptocurrency market. Thus, we estimate the following panel regression models.

$$y_{i,t} = \alpha_i + \gamma_t + \beta_1 \times MFD_{i,t} + \beta_2 \times y_{i,t-1} + \gamma \times Controls_{i,t} + \epsilon_{i,t} \quad (8)$$

where $y_{i,t}$ is the weekly turnover or historical volatility of cryptocurrency i in week t. To account for persistence in turnover and volatility, we include the first lag of our dependent variable as controls. We additionally control for short-term reversal (RET), one-week momentum (MOM), cryptocurrency capitulation ($Size$), trading volume (Vol), cryptocurrency beta ($Beta$), and cryptocurrency lottery-like features (MAX). Standard errors are double-clustered by week and crypto. Notably, the effect of investor belief disagreement is captured by the coefficient β_1.

Table 15 reports the panel regression results based on Eq. (8). In columns (1) and (3), the coefficients on MFD are 3.53 and 15.13 and statistically significant at the 1% level, respectively, indicating a positive and economically meaningful relation between MFD and contemporaneous turnover and historical volatility. More importantly, MFD also exhibits strong predictive power for future market activity. As shown in columns (2) and (4), the coefficient on MFD is 4.42 for one-week-ahead volatility and 14.38 for one-week-ahead turnover, both statistically significant at the 1% level.

Overall, these results are consistent with the existing disagreement literature (Atmaz and Basak, 2018; Bali et al., 2026), and suggest that belief disagreement, as captured by MFD, is strongly linked not only to future returns but also to higher-order moments of the cryptocurrency market. The evidence provides evidence consistent with the view that MFD captures economically meaningful variation in investor disagreement that manifests in both trading volume and volatility.

<Insert Table 15 Here>

C. Robustness Tests

The strong negative MFD-return relation is robust to a comprehensive set of tests, which we detail in the Internet Appendix E. In traditional asset classes, many return anomalies become unprofitable once transaction costs are taken into account (Novy-Marx and Velikow, 2016), and even the most effective mitigation techniques often preserve only a small fraction of gross returns (Patton and Weller, 2020). Accordingly, we evaluate the net profitability of MFD-based on long-short strategy. As detailed in Internet Appendix E1, the strategy remains economically profitable after accounting for conservative estimates of transaction costs.

Additionally, our results are robust to alternative specifications of the incomplete information set and alternative machine learning architectures used to construct MFD (Internet Appendix E2). To assess the temporal stability of the results, we show that the negative MFD-return relation persists across two equally split subsamples and an extended sample period (Internet Appendix E3). Finally, the MFD premium remains robust under alternative size-based screening criteria and after excluding stablecoins from the sample. (Internet Appendix E4).

VI. Conclusion

Because of the lack of a reasonable proxy for investor disagreement in the cryptocurrency market, we follow the approach in Bali et al. (2026) to construct a crypto-level measure of investor belief disagreement for cryptocurrencies. We assume that crypto investors use their individual models and common data information to form their individual return forecasts for cryptocurrencies. We find a significantly negative relation between MFD and future cryptocurrency returns in the cross section. The negative MFD-return relation holds in various robustness tests. In particular, MFD significantly predicts future returns after controlling for other disagreement proxies such as turnover.

Moreover, we find that past return variables are among the top drivers of disagreement, suggesting that MFD is broadly consistent with disagreement associated with extrapolative behavior. In addition, the negative relation is stronger for cryptocurrencies with larger limits-to-arbitrage or more severe overpricing. The negative relation is also stronger following high crypto investor sentiment periods. Our study sheds novel light on the relation between investor disagreement and future returns from the cryptocurrency market.

Reference

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *The Journal of Finance* 61, 259-299.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2009, High idiosyncratic volatility and low returns: International and further US evidence, *Journal of Financial Economics* 91, 1-23.

Atilgan, Yigit, Turan G. Bali, K. Ozgur Demirtas, and A. Doruk Gunaydin, 2020, Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns, *Journal of Financial Economics* 135, 725-753.

Atmaz, Adem, and Suleyman Basak, 2018, Belief dispersion in the stock market, *The Journal of Finance* 73(3), 1225-1279.

Bianchi, Daniele, Mykola Babiak, and Alexander Dickerson, 2022, Trading volume and liquidity provision in cryptocurrency markets, *Journal of Banking & Finance* 142, 106547.

Babiak, Mykola, and Daniele Bianchi, 2025, Mispricing and risk compensation in cryptocurrency returns, *Journal of Financial and Quantitative Analysis,* 1-42.

Bali, Turan G., Bryan T. Kelly, Mathis Mörke, and Jamil Rahman, 2026, Machine forecast disagreement, *The Review of Financial Studies, forthcoming*.

Bali, Turan G., and Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427-446.

Barinov, Alexander, 2014, Turnover: liquidity or uncertainty? *Management Science* 60, 2478-2495.

Berkman, Henk, Valentin Dimitrov, Prem C. Jain, Paul D. Koch, and Sheri Tice, 2009, Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements, *Journal of Financial Economics* 92, 376-399.

Boehme, Rodney D., and Bartley R. Danielsen, and Sorin M. Sorescu, 2006, Short-sale constraints, differences of opinion, and overvaluation, *Journal of Financial and Quantitative Analysis* 41, 455-487.

Böhme, Rainer, Nicolas Christin, Benjamin Edelman, and Tyler Moore, 2015, Bitcoin: Economics, technology, and governance, *Journal of Economic Perspectives* 29, 213-238.

Borri, Nicola, and Kirill Shakhnov, 2022, The cross-section of cryptocurrency returns, *The Review of Asset Pricing Studies* 12, 667-705.

Cakici, Nusret, Syed Jawad Hussain Shahzad, Barbara Bedowska-Sojka, and Adam Zaremba, 2024, Machine learning and the cross-section of cryptocurrency returns, *International Review of Financial Analysis* 94, 103244.

Cen, Ling, Hai Lu, Liyan Yang, 2013, Investor Sentiment, Disagreement, and the Breadth – Return Relationship, *Management Science* 59, 1076-1091.

Chordia, Tarun, and Avanidhar Subrahmanyam, and V. Ravi Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3-32.

Cong, Lin William, and Ye Li, and Neng Wang, 2021, Tokenomics: Dynamic adoption and valuation, *The Review of Financial Studies* 34, 1105-1155.

Cookson, J. Anthony, and Marina Niessner, 2020, Why don't we agree? Evidence from a social network of investors, *The Journal of Finance* 75(1), 173-228.

Cookson, J. Anthony, Vyacheslav Fos, and Marina Niessner, 2022, Does disagreement facilitate informed trading? *Journal of Financial and Quantitative Analysis,* 1-28.

Da, Zhi, Xing Huang, and Lawrence J. Jin, 2021, Extrapolative Beliefs in the Cross-Section: What Can We Learn from the Crowds? *Journal of Financial Economics* 140, 175-196.

Detzel, Andrew, Hong Liu, Jack Strauss, Guofu Zhou, and Yingzi Zhu, 2021, Learning and predictability via technical analysis: Evidence from bitcoin and stocks with hard-to-value fundamentals, *Financial Management* 50, 107-137.

Di Francesco, Tommaso, and Cars Hommes, 2025, Sentiment-driven speculation in financial markets with heterogeneous beliefs: A machine learning approach, *Journal of Economic Dynamics & Control* 175, 105092.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113 – 2141.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Fieberg, Christian, Gerrit Liedtke, Thorsten Poddig, Thomas Walker, and Adam Zaremba, 2025, A Trend Factor for the Cross-Section of Cryptocurrency Returns, *Journal of Financial and Quantitative Analysis* 60(7), 3116-3153.

Garfinkel, Jon A., 2009, Measuring investors' opinion divergence, *Journal of Accounting Research* 47, 1317-1348.

Garfinkel, Jon A., and Lawrence Hsiao, and Danqi Hu, 2025, Disagreement and returns: The case of cryptocurrencies, *Financial Management*.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021, Five facts about beliefs and portfolios, *American Economic Review* 111(5), 1481-1522.

Griffin, John M., and Amin Shams, 2020, Is Bitcoin really untethered? *The Journal of Finance* 75, 1913-1964.

Grossman, Sanford J., and Joseph E. Stiglitz, 1976, Information and competitive price systems, *The American Economic Review* 66, 246-253.

Gu, Shihao, and Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *The Review of Financial Studies* 33, 2223-2273.

Hong, Harrison, and Jeremy C. Stein, 2007, Disagreement and the stock market, *Journal* *of Economic Perspectives* 21, 109-128.

Howell, Sabrina T., and Marina Niessner, and David Yermack, 2020, Initial coin offerings: Financing growth with cryptocurrency token sales, *The Review of Financial Studies* 33, 3925-3974.

Jiang, Hao, Naveen Khanna, Qian Yang, and Jiayu Zhou, 2024, The cyber risk premium, *Management Science* 70, 8791-8817.

Johnson, Timothy C., 2004, Forecast dispersion and the cross section of expected returns, *The Journal of Finance* 59, 1957-1978.

Kandel, Eugene, and Neil D. Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831-872.

Kelly, Bryan T., and Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501-524.

Kim, Jun Sik, and Doojin Ryu, and Sung Won Seo, 2014, Investor sentiment and return predictability of disagreement, *Journal of Banking & Finance* 42, 166-178.

Kogan, Shimon, Igor Makarov, Marina Niessner, and Antoinette Schoar, 2024, Are cryptos different? evidence from retail trading, *Journal of Financial Economics* 159, 103897.

Kumar, Alok, 2009, Who gambles in the stock market? *The Journal of Finance* 64, 1889-1933.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* 1315-1335.

Lam, FY Eric C., and KC John Wei, 2011, Limits-to-arbitrage, investment frictions, and the asset growth anomaly, *Journal of Financial Economics* 102, 127-149.

Lee, Charles M. C., and Eric C. So, 2015, Alphanomics: The Informational Underpinnings of Market Efficiency, Foundations and Trends in Accounting 9, 59-258.

Li, Kai, Feng Mai, Rui Shen, and Xinyan Yan, 2021, Measuring corporate culture using machine learning, *The Review of Financial Studies* 34, 3265-3315.

Li, Qing, Hongyu Shan, Yuehua Tang, and Vincent Yao, 2024, Corporate climate risk: Measurements and responses, *The Review of Financial Studies* 37, 1778-1830.

Lee, Suzanne S., and Minho Wang, 2024, Variance Decomposition and Cryptocurrency Return Prediction, *Journal of Financial and Quantitative Analysis*, Forthcoming.

Li, Tao, Donghwa Shin, and Baolian Wang, 2025, Cryptocurrency Pump-and-Dump Schemes, *Journal of Financial and Quantitative Analysis* 60, 3622-3659.

Liu, Yukun, and Aleh Tsyvinski, 2021, Risks and returns of cryptocurrency, *The Review of Financial Studies* 34, 2689--2727.

Liu, Yukun, and Aleh Tsyvinski, and Xi Wu, 2022, Common risk factors in cryptocurrency, *The Journal of Finance* 77, 1133-1177.

Lyandres, Evgeny, and Berardino Palazzo, and Daniel Rabetti, 2022, Initial coin offering

(ICO) success and post-ICO performance, *Management Science* 68, 8658-8679.

Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *The Journal of Finance* 42, 483-510.

Miller, Edward M., 1977, Risk, uncertainty, and divergence of opinion, *The Journal of Finance* 32, 1151-1168.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation, *Econometrica* 55, 703-708.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *The Review of Financial Studies* 29(1), 104-147.

Patton, Andrew J., and Brian M. Weller, 2020, What you see is not what you get: The costs of trading market anomalies, *Journal of Financial Economics* 137(2), 515-549.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35-55.

Sockin, Michael, and Wei Xiong, 2023, A model of cryptocurrencies, *Management Science* 69, 6684-6707.

Stambaugh, Robert F., and Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288-302.

Stambaugh, Robert F., and Jianfeng Yu, and Yu Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *The Journal of Finance* 70, 1903-1948.

Wang, Jiang, 1994, A model of competitive stock trading volume, *Journal of Political Economy* 102, 127-168.

Yu, Jialin, 2011, Disagreement and return predictability of stock portfolios, *Journal of Financial Economics* 99, 162-183.

Yu, Jianfeng, and Yu Yuan, 2011, Investor sentiment and the mean--variance relation, *Journal of Financial Economics* 100, 367-381.

Zhang, X. Frank, 2006, Information uncertainty and stock returns, *The Journal of Finance* 61, 105-137.

Zhu, Zhaobo, Donglian Ma, and Jun Tu, 2025, In search of cryptocurrency failure, Working paper.

Figure 1: Time Series of MFD

The figure presents the weekly time-series plot of the median cryptocurrency-level MFD. The dashed line presents the time series of the interquartile range. The sample period is from January 2019 to June 2024.

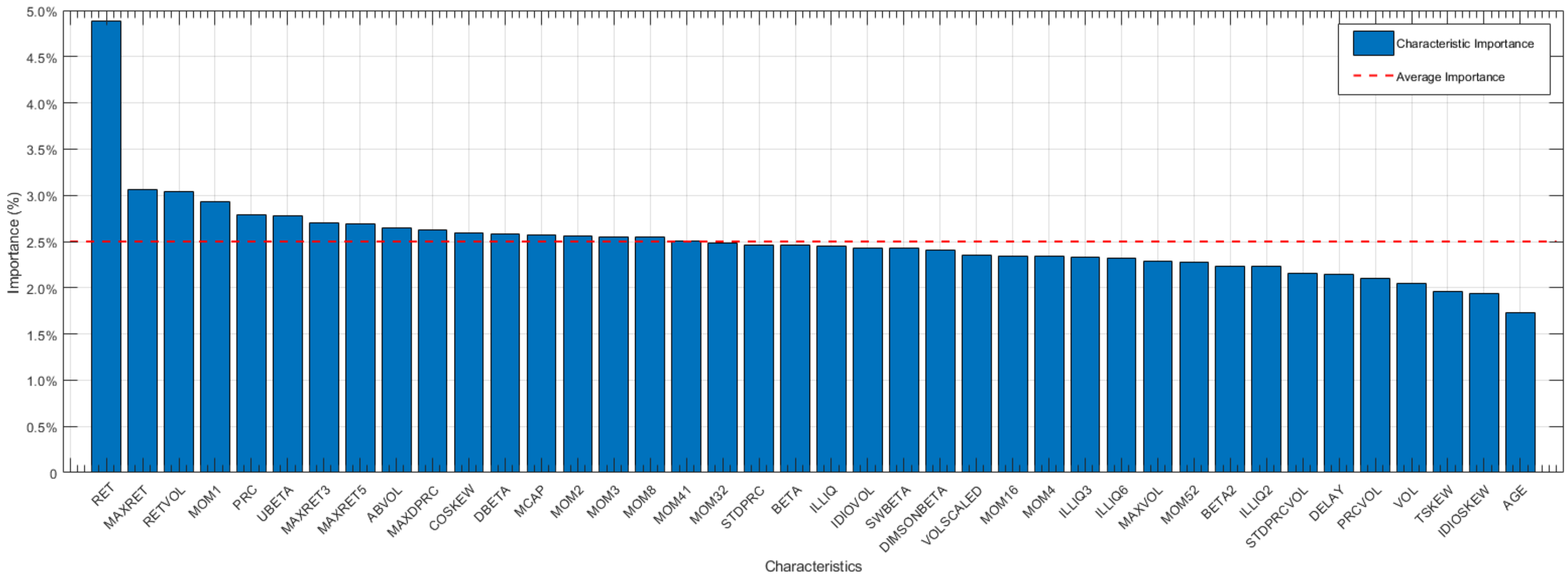

Figure 2: Feature Importance of 40 Cryptocurrency Characteristics in Random Forest Regression Prediction

This figure presents the average feature importance of 40 characteristics across 100 investors, which measures how much feature contributes to random forest regression prediction accuracy. The red dashed line represents the baseline feature importance of all 40 characteristics.

Table 1: Descriptive Statistics

This table reports the summary statistics for the cross-sectional variables. Cryptocurrencies with market capitalization below $1 billion dollars, trading volume below $10 thousand, price below 0.1 dollars, and trading history shorter than 52 weeks are excluded from the analysis. Ret_1 is the one-week ahead returns in excess of the risk-free rate of individual cryptocurrency (in %). MFD is the machine forecast disagreement variable (scaled by 100). $Size$ is the logarithm of cryptocurrency market value. MOM is the cryptocurrency return in the previous week (in %). $Abvol$ is the natural logarithm of the average trading volume for a cryptocurrency in the portfolio formation week subtracted by the natural logarithm of the average dollar trading volume of the past 12 weeks. $Beta$ is calculated by regressing the daily excess returns of each cryptocurrency on the value-weighted market excess returns during the past 90 days. $Ivol$ is calculated as the standard deviation of residuals from regressing excess cryptocurrency returns on excess market returns during the past 90 days (in %). $Volt$ is calculated as the standard deviation of daily return in the portfolio formation week (in %). MAX is the maximum daily return of the portfolio formation week (in %). $CosKew$ is the coefficient of the squared excess market return term when regressing daily excess cryptocurrency returns on the daily excess market returns and the square daily excess market returns in the past 90 days. $ILLIQ$ is calculated as the average absolute daily return divided by price volume during the portfolio formation week (scaled by 10^7). $Turnover$ is calculated as the logarithm of the average daily volume times price scaled by market capitalization in the portfolio formation week. ARB is the limit-to-arbitrage index (see Section IV.B). $MISP$ is the mispricing index (see Section IV.D). ΔTO is the change in turnover (an alternative proxy for investor belief disagreement, see section V.A). $AlterRF\_MFD$ is the MFD constructed by random forest regression using an alternative investor–specific incomplete information set, and $XGBM\_MFD$ is the MFD constructed by gradient boosted regression trees (see Section V.C). $PastRet\_MFD$ is the return-based MFD, $NonPastRet\_MFD$ is the Non-return-based MFD, and $PastRetPlus\_MFD$ is the return-plus-based MFD (see Section IV.A). All non-return variables are winsorized at the 1% and 99% levels. The mean, standard deviation (Std), minimum (Min), 10^{th} percentile (10^{th}), 25^{th} percentile (25^{th}), 50^{th} percentile (median), 75^{th} percentile (75^{th}), 90^{th} percentile (90^{th}), and maximum (Max) are reported. The sample is from January 2019 to June 2024.

	Mean	Std	Min	10^{th}	25^{th}	Median	75^{th}	90^{th}	Max
Ret_1	0.931	15.638	–44.658	–11.026	–6.094	–1.130	4.921	14.091	134.016
MFD	0.098	0.066	0.060	0.066	0.070	0.077	0.094	0.145	0.511
MOM	1.590	13.496	–30.559	–10.364	–5.657	–0.664	5.615	15.637	68.641
$Size$	18.374	2.037	14.407	16.012	16.980	18.103	19.537	21.084	24.655
$Abvol$	0.017	0.486	–1.223	–0.518	–0.248	–0.017	0.229	0.583	1.772
$Beta$	0.923	0.403	–0.037	0.341	0.718	0.963	1.175	1.377	1.970
$Ivol$	5.357	3.606	0.203	2.171	3.319	4.583	6.428	9.216	23.366
$Volt$	5.249	3.508	0.133	2.084	3.287	4.532	6.289	9.019	21.992
MAX	8.175	7.598	0.113	2.009	3.816	6.071	9.866	16.389	46.918
$CosKew$	–2.570	3.679	–14.799	–6.841	–4.771	–2.642	–0.382	1.505	9.121
$ILLIQ$	0.355	0.644	0.000	0.001	0.013	0.068	0.328	1.022	3.713
$Turnover$	–2.572	1.409	–6.493	–4.383	–3.468	–2.519	–1.607	–0.811	0.774
ARB	11.493	4.366	4.000	6.183	8.477	11.493	14.718	17.695	23.525
$MISP$	7.757	2.606	3.035	4.410	5.734	7.621	9.695	11.259	13.771
ΔTO	–0.006	0.242	–0.906	–0.191	–0.065	–0.007	0.043	0.171	1.095
$AlterRF\_MFD$	0.115	0.079	0.070	0.077	0.081	0.089	0.110	0.171	0.616
$XGBM\_MFD$	0.183	0.119	0.081	0.100	0.116	0.144	0.199	0.306	0.823
$PastRet\_MFD$	0.027	0.025	0.013	0.015	0.016	0.018	0.026	0.046	0.180
$NonPastRet\_MFD$	0.190	0.080	0.115	0.135	0.146	0.165	0.201	0.268	0.617
$PastRetPlus\_MFD$	0.082	0.045	0.053	0.059	0.061	0.067	0.081	0.122	0.338

Table 2: Cross-Sectional Correlations to MFD

This table reports the summary statistics on the correlations of various cryptocurrency characteristics with MFD. Correlation is measured using Spearman's ρ. The cryptocurrency characteristics are defined in Table 1. The mean, standard deviation (Std), 10^{th} percentile (10^{th}), 25^{th} percentile (25^{th}), 50^{th} percentile (median), 75^{th} percentile (75^{th}), and 90^{th} percentile (90^{th}) are reported. The sample is from January 2019 to June 2024.

	Mean	Std	10^{th}	25^{th}	Median	75^{th}	90^{th}
Ret_1	–0.046	0.139	–0.203	–0.123	–0.055	0.032	0.130
MOM	0.106	0.147	–0.070	0.015	0.091	0.200	0.297
$Size$	–0.103	0.079	–0.193	–0.158	–0.111	–0.058	–0.007
$Abvol$	0.105	0.101	–0.022	0.036	0.101	0.176	0.236
$Beta$	–0.019	0.159	–0.228	–0.114	–0.010	0.097	0.178
$Ivol$	0.456	0.129	0.280	0.380	0.479	0.547	0.599
$Volt$	0.291	0.156	0.082	0.186	0.307	0.394	0.487
MAX	0.227	0.160	0.017	0.107	0.227	0.337	0.436
$CosKew$	–0.006	0.165	–0.188	–0.117	0.005	0.110	0.177
$ILLIQ$	0.182	0.121	0.039	0.101	0.170	0.265	0.338
$Turnover$	0.067	0.133	–0.076	–0.023	0.053	0.150	0.238

Table 3: Univariate Portfolio Sorts on MFD

This table reports excess returns and alphas on different cryptocurrency deciles sorted based on MFD and formed on a weekly basis between January 2019 and June 2024. Each week, cryptocurrencies are sorted into decile portfolios by MFD. Portfolio Low is the portfolio of cryptocurrencies with the lowest MFD, while cryptocurrencies with the highest MFD comprise portfolio High. All equal-weighted and value-weighted excess returns and alphas are a week ahead of the portfolio formation period. Excess return is the return in excess of the risk-free rate. Alpha is the intercept from a time-series regression of weekly excess returns on the factors of alternative models: CAPM and three-factor models proposed by Liu, Tsyvinski, and Wu (2022). This table also reports differences between portfolios High and Low. Newey-West (1987) adjusted t statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal-Weighted			Value-Weighted		
	Excess Return	CAPM	3-Factor	Excess Return	CAPM	3-Factor
Low	1.196*	0.059	0.037	1.456**	0.318	0.395
	(1.74)	(0.16)	(0.10)	(2.14)	(0.91)	(1.12)
2	1.197*	0.075	0.073	0.763	–0.466	–0.514*
	(1.84)	(0.21)	(0.21)	(1.10)	(–1.47)	(–1.71)
3	1.401**	0.197	0.140	2.454***	1.072**	0.955*
	(2.13)	(0.49)	(0.38)	(2.56)	(2.08)	(1.89)
4	1.079	–0.055	–0.046	1.136	–0.152	–0.197
	(1.57)	(–0.15)	(–0.13)	(1.60)	(–0.41)	(–0.58)
5	0.847	–0.203	–0.243	1.326**	0.172	0.155
	(1.31)	(–0.57)	(–0.73)	(1.99)	(0.48)	(0.44)
6	0.917	–0.257	–0.200	1.032*	–0.0720	0.049
	(1.35)	(–0.80)	(–0.62)	(1.76)	(–0.23)	(0.16)
7	0.933	–0.158	–0.084	0.160	–0.950***	–0.643*
	(1.33)	(–0.40)	(–0.23)	(0.25)	(–2.64)	(–1.70)
8	0.920	–0.141	–0.245	0.404	–0.508	–0.487
	(1.45)	(–0.43)	(–0.78)	(0.68)	(–1.28)	(–1.26)
9	0.483	–0.484	–0.441	0.532	–0.159	–0.200
	(0.79)	(–1.31)	(–1.23)	(1.13)	(–0.48)	(–0.63)
High	–0.027	–1.028**	–1.025**	–0.565	–1.468***	–1.433***
	(–0.04)	(–2.33)	(–2.34)	(–0.90)	(–3.24)	(–3.43)
High-Low	–1.223***	–1.087**	–1.063**	–2.022***	–1.786***	–1.828***
	(–2.60)	(–2.54)	(–2.53)	(–3.29)	(–3.22)	(–3.54)

Table 4: Transition Matrix

This table reports transition probabilities for MFD at a lag of 12 weeks from January 2019 to June 2024. For each week, we sort all cryptocurrencies in our sample into deciles based on the ascending order of MFD. We then repeat this process in the week after 12 weeks. Cryptocurrencies with the lowest MFD are categorized in portfolio Low. Cryptocurrencies with the highest MFD are categorized in portfolio High. For each MFD decile, we then calculate the percentage of cryptocurrencies that fall into each of the MFD decile in the week after 12 weeks and present time-series averages of these transition probabilities in this table. Each row corresponds to a decile portfolio in a week, and each column corresponds to a decile portfolio in the week after 12 weeks.

	Low	2	3	4	5	6	7	8	9	High
Low	27.968	16.979	11.395	8.913	7.171	5.930	5.131	5.345	5.536	5.632
2	17.086	17.502	14.691	11.281	9.338	7.309	6.734	5.842	5.194	5.023
3	12.040	14.081	14.530	13.364	11.153	9.100	7.678	6.390	6.075	5.589
4	9.120	10.787	13.435	13.288	12.736	11.290	8.789	7.514	6.693	6.350
5	7.583	9.930	11.626	12.228	12.793	12.265	10.532	8.590	7.693	6.759
6	7.067	8.652	9.318	11.605	12.076	13.069	12.379	9.765	8.846	7.224
7	6.236	7.100	8.489	9.134	11.046	12.605	14.066	12.617	10.194	8.513
8	5.820	6.053	7.386	7.642	9.098	11.739	13.402	15.627	13.426	9.807
9	4.894	4.639	5.294	7.262	7.820	10.516	12.410	15.301	17.438	14.426
High	3.905	3.817	4.118	4.718	6.020	7.009	9.111	12.929	19.675	28.698

Table 5: Long-Term Portfolio Returns

This table reports the long-term predictive power of MFD. For each week $t+n$, where $n \in \{2, 3, \ldots, 12\}$, individual cryptocurrencies are sorted into decile portfolios based on $week - t$ MFD. The Table reports the average weekly return and alpha (adjusted with three-factor model proposed by Liu, Tsyvinski, and Wu (2022)) of the equal-weighted and value-weighted MFD-sorted high-minus-low portfolios. Newey-West (1987) adjusted t statistics are presented in parentheses. $*$, $**$, and $***$ indicate the significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 2019 and June 2024.

	Equal-Weighted		Value-Weighted	
	High–Low	Alpha	High–Low	Alpha
$t+2$	–0.084	–0.235	–1.555**	–1.184**
	(–0.15)	(–0.51)	(–2.29)	(–2.34)
$t+3$	0.125	–0.079	–1.273**	–0.908*
	(0.17)	(–0.12)	(–2.14)	(–1.86)
$t+4$	–0.301	–0.130	–2.057***	–1.421**
	(–0.67)	(–0.31)	(–3.03)	(–2.45)
$t+5$	–0.252	–0.095	–0.980	–0.728
	(–0.61)	(–0.24)	(–1.61)	(–1.30)
$t+6$	1.016	0.950	0.850	1.128
	(1.15)	(1.18)	(0.50)	(0.65)
$t+7$	1.552	1.485*	–0.703	–0.167
	(1.62)	(1.79)	(–0.87)	(–0.23)
$t+8$	0.589	0.528	–1.939***	–1.454***
	(0.66)	(0.69)	(–3.41)	(–2.88)
$t+9$	1.225	1.209	–0.894	–0.479
	(1.12)	(1.25)	(–1.58)	(–0.93)
$t+10$	0.490	0.427	–0.870	–0.398
	(0.42)	(0.41)	(–1.38)	(–0.75)
$t+11$	6.995	6.943	0.355	0.917
	(1.13)	(1.11)	(0.40)	(1.00)
$t+12$	8.575	8.657	0.085	0.103
	(1.18)	(1.17)	(0.11)	(0.13)

Table 6: Average Cryptocurrency Characteristics of MFD-sorted Portfolio

This table reports the time-series average for MFD (Scaled by 100) and other cryptocurrency characteristics for cryptocurrency deciles sorted based on MFD and formed on a weekly basis from January 2019 to June 2024. Other cryptocurrency characteristics include MOM (in %), $Size$, $Abvol$, $Beta$, $Ivol$ (in %), $Volt$ (in %), MAX (in %), $CosKew$, $ILLIQ$ (Scaled by 10^7), Turnover. Portfolio Low is the portfolio of cryptocurrencies with the lowest MFD, while Portfolio High is the portfolio of cryptocurrencies with the highest MFD. The last two rows report the difference between the High and Low (High-Low) and the associated Newey-West (1987) adjusted t statistics (t-stats). $*$, $**$, and $***$ indicate the significance at the 10%, 5%, and 1% levels, respectively.

	MFD	MOM	$Size$	$Abvol$	$Beta$	$Ivol$	$Volt$	MAX	$CosKew$	$ILLIQ$	$Turnover$
Low	0.064	0.488	18.444	-0.044	0.891	3.986	4.549	7.487	–2.791	0.248	–2.703
2	0.068	0.590	18.649	-0.035	0.920	4.068	4.588	7.072	–2.747	0.219	–2.690
3	0.070	0.889	18.602	-0.010	0.936	4.182	4.585	6.823	–2.735	0.226	–2.638
4	0.072	0.841	18.510	-0.001	0.942	4.419	4.733	6.936	–2.710	0.250	–2.628
5	0.075	1.226	18.535	0.001	0.940	4.566	4.801	7.119	–2.542	0.282	–2.651
6	0.079	1.013	18.503	0.011	0.932	4.867	4.873	7.231	–2.515	0.291	–2.595
7	0.084	1.017	18.281	0.028	0.912	5.292	5.090	7.676	–2.499	0.365	–2.579
8	0.094	1.523	18.262	0.024	0.907	5.728	5.331	8.274	–2.308	0.384	–2.502
9	0.118	2.746	18.140	0.048	0.924	6.740	6.037	9.848	–2.206	0.446	–2.426
High	0.307	5.682	17.843	0.146	0.921	9.621	7.868	13.272	–2.639	0.633	–2.334
High–Low	0.243^{***}	5.194^{***}	$–0.6014^{***}$	0.190^{***}	0.030	5.634^{***}	3.319^{***}	5.785^{***}	0.152	0.385^{***}	0.369^{***}
	(22.17)	(10.37)	(–10.18)	(16.47)	(1.18)	(28.02)	(26.63)	(14.77)	(0.87)	(16.11)	(4.56)

Table 7: Bivariate Portfolio Analysis

This table reports the results from bivariate portfolio analyses based on dependent double sorts of various cryptocurrency-specific characteristics and MFD. We first form quintile portfolios every week based on a given cryptocurrency characteristic. Then, we additionally form quintile portfolios based on MFD in each cryptocurrency characteristic quintile. Portfolio Low is the combined portfolio of cryptocurrencies with the Lowest MFD in each cryptocurrency characteristic quintile, while portfolio High is the combined portfolio of cryptocurrencies with the highest MFD in each cryptocurrency characteristic quintile. The characteristics are described in Table 1. The Table show the excess return and alphas (adjusted with three-factor model proposed by Liu, Tsyvinski, and Wu (2022)) for each of the equal-weighted and value-weighted MFD-sorted high-minus-low portfolios. All Newey-West (1987) adjusted t statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal-Weighted		Value-Weighted	
	High–Low	Alpha	High–Low	Alpha
MOM	–1.011**	–0.823**	–1.576***	–1.210**
	(–2.30)	(–2.08)	(–2.71)	(–2.26)
$Size$	–1.414***	–1.134***	–1.960***	–1.407***
	(–3.31)	(–2.89)	(–2.95)	(–2.70)
$Abvol$	–1.254***	–1.105***	–2.094***	–1.530***
	(–3.19)	(–3.06)	(–3.03)	(–2.68)
$Beta$	–1.434***	–1.168***	–1.666***	–1.356**
	(–3.74)	(–3.40)	(–2.62)	(–2.34)
$Ivol$	–0.743**	–0.568*	–1.475**	–0.716
	(–1.97)	(–1.69)	(–2.15)	(–1.37)
$Volt$	–0.818*	–0.715*	–1.305**	–0.697
	(–1.90)	(–1.77)	(–2.09)	(–1.26)
MAX	–1.016**	–0.875**	–1.411**	–0.925*
	(–2.49)	(–2.30)	(–2.28)	(–1.65)
$CosKew$	–1.020***	–0.858**	–1.991***	–1.358**
	(–2.63)	(–2.52)	(–3.15)	(–2.50)
$ILLIQ$	–1.452***	–1.268***	–1.706**	–0.976*
	(–3.35)	(–3.20)	(–2.46)	(–1.88)
$Turnover$	–1.113***	–1.007**	–1.391**	–1.338**
	(–2.70)	(–2.54)	(–2.19)	(–2.37)

Table 8: Fama-MacBeth Cross-Sectional Regression

This table reports Fama-MacBeth cross-sectional regressions of one-week-ahead excess returns on MFD and a series of other cryptocurrency characteristics as control variables from January 2019 to June 2024. Coefficients and adjusted R2 are the time-series average from weekly Fama and MacBeth (1973) regressions. And the associated t-statists in parentheses are adjusted with the Newey-West (1987) procedure. The control variables are described in Table 1. All non-return variables are winsorized at the 1% and 99% levels. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MFD	–7.795***	–8.698***	–8.659***	–8.621***	–8.051***	–6.829***	–5.971***	–5.606***	–5.985***	–6.201***	–6.443***
	(–3.54)	(–4.25)	(–4.21)	(–4.19)	(–3.98)	(–3.08)	(–2.75)	(–2.58)	(–2.76)	(–2.88)	(–2.91)
MOM		0.020*	0.020*	0.022**	0.013	0.014	0.014	0.016	0.014	0.016	0.016
		(1.76)	(1.78)	(1.98)	(1.23)	(1.30)	(1.20)	(1.30)	(1.17)	(1.31)	(1.36)
$Size$			–0.001	–0.001	–0.001	–0.001**	–0.001**	–0.001**	–0.001**	–0.001*	–0.001*
			(–1.40)	(–1.60)	(–1.50)	(–2.09)	(–2.22)	(–2.27)	(–2.31)	(–1.77)	(–1.95)
$Abvol$				–0.004*	–0.004**	–0.004*	–0.002	–0.002	–0.002	–0.002	–0.001
				(–1.91)	(–1.98)	(–1.90)	(–0.76)	(–0.93)	(–1.04)	(–0.93)	(–0.68)
$Beta$					0.008	0.009	0.010*	0.010	0.010	0.010	0.010
					(1.22)	(1.39)	(1.69)	(1.53)	(1.64)	(1.57)	(1.56)
$Ivol$						–0.049	–0.027	–0.031	–0.022	–0.019	–0.016
						(–1.13)	(–0.62)	(–0.70)	(–0.45)	(–0.39)	(–0.33)
$Volt$							–0.079**	–0.145	–0.134	–0.129	–0.119
							(–1.98)	(–1.43)	(–1.29)	(–1.16)	(–1.08)
MAX								0.030	0.028	0.020	0.017
								(0.65)	(0.60)	(0.41)	(0.35)
$CosKew$									0.001	0.000	0.001
									(1.05)	(0.75)	(0.83)
$ILLIQ$										0.014	0.001

										(0.72)	(0.38)
$Turnover$											–0.001
											(–0.97)
$Constant$	0.017**	0.017**	0.031**	0.034**	0.029**	0.036***	0.037***	0.036***	0.036***	0.033**	0.033**
	(2.25)	(2.38)	(2.26)	(2.38)	(2.10)	(2.67)	(2.84)	(2.80)	(2.90)	(2.39)	(2.44)
$Adj. R^2$	0.021	0.038	0.047	0.055	0.087	0.097	0.109	0.122	0.130	0.140	0.146
N	81068	81068	81068	81068	81068	81068	81068	81068	81068	81068	81068

Table 9: The Role of Past Returns

This table reports the average weekly return spread between the highest and lowest decile portfolios based on return-based MFD, non-return-based MFD, and return-plus-based MFD, respectively. Return information set consists of 13 characteristics related to past returns, such as past one–week return (MOM_1), past two–week return (MOM_2), past three–week return (MOM_3), past four–week return (MOM_4), past one–to–four–week return ($MOM_{4\to1}$), past eight–week return (MOM_8), past sixteen–week return (MOM_{16}), past thirty–two–week return (MOM_{32}), past fifty two–week return (MOM_{52}), maximum daily return in portfolio–formation week (MAX_1), average maximum daily return of the previous three weeks before the portfolio–formation week (MAX_3), average maximum daily return of the previous five weeks before the portfolio–formation week (MAX_5), and return on portfolio–formation week (RET). Non-return information set consists of other 27 characteristics that are not directly relative to past returns. Return-based MFD is constructed by 9 key return characteristics selected from return information set, which play more significant role in capturing extrapolative expectations in investor belief formation, including MOM_1, MOM_2, MOM_3, MOM_4, $MOM_{4\to1}$, MAX_1, MAX_3, MAX_5, and RET. Non-return-based MFD is constructed by 9 characteristics randomly selected from non-return information set. Return-plus-based MFD is constructed by mixed characteristics, which comprise 9 key return characteristics and 15 characteristics randomly selected from non-return information set. Each week, cryptocurrencies are sorted into decile portfolios by MFD. Portfolio Low is the portfolio of cryptocurrencies with the lowest MFD, while cryptocurrencies with the highest MFD comprise portfolio High. All value–weighted excess returns and alphas are a week ahead of the portfolio formation period. Excess return is the return in excess of the risk–free rate. Alpha is the intercept from a time–series regression of weekly excess returns on the three–factor models proposed by Liu, Tsyvinski, and Wu (2022). This table also reports differences between portfolios High and Low. Besides, this table reports the return difference between return-based MFD and non-return-based MFD, and between return-plus-based MFD and non-return-based MFD. All measures are reported in percentage terms. Newey–West (1987) adjusted t statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Return-based MFD (A)		Non-return-based MFD (B)		Return-plus-based MFD (C)		Diff (A-B)		Diff (C-B)	
	Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha
Low	0.984	–0.169	1.328*	0.098	1.615**	0.336	–0.344	–0.267	0.287	0.238
	(1.48)	(–0.55)	(1.73)	(0.23)	(2.01)	(0.74)	(–0.70)	(–0.61)	(0.75)	(0.64)
2	1.310*	0.013	1.667**	0.427	1.354*	0.064	–0.0357	–0.414	–0.313	–0.363
	(1.70)	(0.04)	(2.01)	(0.86)	(1.86)	(0.19)	(–0.61)	(–0.78)	(–0.64)	(–0.77)
3	1.088	–0.111	1.133	–0.247	1.011	–0.178	–0.045	0.136	–0.122	0.069
	(1.51)	(–0.32)	(1.46)	(–0.65)	(1.51)	(–0.64)	(–0.09)	(0.29)	(–0.28)	(0.18)
4	1.319**	0.107)	1.194	–0.083	1.576**	0.353	0.125	0.190	0.382	0.436
	(1.99)	(0.39)	(1.52)	(–0.23)	(2.28)	(1.06)	(0.27)	(0.48)	(0.86)	(1.13)
5	0.499	–0.550	1.463*	0.211	1.506**	0.105	–0.965*	–0.761*	0.043	–0.107
	(0.78)	(–1.58)	(1.85)	(0.55)	(2.01)	(0.31)	(–1.85)	(–1.65)	(0.11)	(–0.27)
6	0.692	–0.314	0.845	–0.254	0.796	–0.368	–0.153	–0.061	–0.049	–0.115
	(1.01)	(–0.96)	(1.38)	(–0.84)	(1.05)	(–0.86)	(–0.34)	(–0.16)	(–0.10)	(–0.25)
7	2.390**	1.186*	0.473	–0.627*	0.491	–0.401	1.917**	1.813**	0.018	0.226
	(2.35)	(1.68)	(0.69)	(–1.79)	(0.92)	(–1.40)	(2.20)	(2.41)	(0.04)	(0.65)
8	0.760	–0.220	0.748	–0.173	0.746	–0.145	0.012	–0.046	–0.002	0.028
	(1.11)	(–0.44)	(1.27)	(–0.50)	(1.11)	(–0.37)	(0.02)	(–0.08)	(–0.00)	(0.07)
9	0.421	–0.520	0.043	–0.523	0.047	–0.604	0.378	0.002	0.004	–0.081
	(0.74)	(–1.27)	(0.08)	(–1.39)	(0.08)	(–1.64)	(0.67)	(0.00)	(0.01)	(–0.17)
High	–0.999**	–1.481***	0.236	–0.682**	–0.475	–1.016***	–1.236**	0.799	–0.711*	–0.334
	(–2.06)	(–3.56)	(0.42)	(–2.00)	(–1.05)	(–2.84)	(–2.26)	(–1.49)	(–1.75)	(–0.90)
High–Low	–1.984***	–1.312***	–1.092*	–0.780	–2.090***	–1.352**	–0.892	–0.532	–0.998*	–0.572
	(–3.36)	(–2.68)	(–1.81)	(–1.43)	(–3.06)	(–2.39)	(–1.28)	(–0.79)	(–1.85)	(–1.10)

Table 10: MFD and Limits-to-Arbitrage

This table reports results from the value–weighted bivariate portfolios based on dependent double sorts of limit–to–arbitrage measures and MFD from January 2019 to June 2024. The cryptocurrency-level limits-to-arbitrage score (ARB) is constructed using cryptocurrency size ($Size$), age (Age), idiosyncratic volatility ($Ivol$), and illiquidity ($ILLIQ$). A low (high) cryptocurrency size, low (high) cryptocurrency age, high (low) idiosyncratic volatility, and high (low) illiquidity indicate a high (low) arbitrage cost. To calculate it, cryptocurrencies are sorted each week into deciles according to their size, age, idiosyncratic volatility, and illiquidity. The decile ranks are attributed to each cryptocurrency increasing in their idiosyncratic volatility and illiquidity, and decreasing in their size and age. Each cryptocurrency is given the corresponding score of its decile rank for each variable. Finally, the limit-to-arbitrage score on the cryptocurrency-level is the sum of the three scores such that it ranges from 4 to 40. A low (high) ARB indicates a lower (higher) arbitrage cost. Panel A reports the time-series averages of the weekly cross-sectional average of a limit-to-arbitrage score (ARB) for MFD-sorted univariate quintile portfolios. Panel B reports excess returns of 3×5 dependent bivariate value-weighted portfolio sorts. Tercile portfolios are formed every week using ARB, and then quintile portfolios are formed based on MFD within each cryptocurrency-specific ARB tercile. Panel C reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of $Size$ and MFD. Tercile portfolios are formed every week using $Size$, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific $Size$ tercile. Panel D reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of Age and MFD. Tercile portfolios are formed every week using Age, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific Age tercile. Panel E reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of $Ivol$ and MFD. Tercile portfolios are formed every week using $Ivol$, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific $Ivol$ tercile. Panel F reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of $ILLIQ$ and MFD. Tercile portfolios are formed every week using $ILLIQ$, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific $ILLIQ$ tercile. The last column reports the alpha adjusted by the three–factor model proposed by Liu, Tsyvinski, and Wu (2022) for each of the MFD–sorted high–minus–low portfolios. All alphas are reported in percentage terms. Newey and West's (1987) adjusted t–values are reported in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Average ARB in MFD Decile Portfolio

	Low	2	3	4	High	H–L	t–stat
ARB	10.584	10.460	10.940	12.159	14.281	3.696***	21.78

Panel B: Bivariate Portfolio Sort on ARB

	Low	2	3	4	High	H–L	Alpha
ARB Low	1.209*	1.574**	1.353**	0.739	0.499	–0.711	–0.178
	(1.82)	(2.10)	(1.99)	(1.37)	(1.05)	(–1.29)	(–0.44)
ARB 2	1.581**	1.012	1.153	0.807	–0.105	–1.686***	–1.534***
	(2.17)	(1.42)	(1.64)	(1.12)	(–0.16)	(–2.99)	(–3.00)
ARB High	2.216*	1.174	0.619	1.052	–1.520**	–3.736***	–3.291***
	(1.81)	(1.45)	(0.70)	(1.07)	(–2.13)	(–3.05)	(–3.12)
ARB High–Low	1.007	–0.401	–0.734	0.313	–2.019***	–3.025**	–3.113***
	(1.03)	(–0.71)	(–1.21)	(0.37)	(–3.17)	(–2.41)	(–2.83)

Panel C: Bivariate Portfolio Sort on Size ($Size$)

	Low	2	3	4	High	H–L	Alpha
Low	1.246	1.375*	0.855	1.119	0.631	–0.615	–0.547
	(1.47)	(1.71)	(1.16)	(1.62)	(0.79)	(–0.85)	(–0.80)
Median	1.261*	0.995	0.866	1.037	–0.249	–1.510***	–1.294***
	(1.75)	(1.44)	(1.19)	(1.46)	(–0.38)	(–2.91)	(–2.63)
High	1.279*	2.386***	1.160*	0.208	0.334	–0.945	–0.320
	(1.85)	(2.80)	(1.80)	(0.41)	(0.77)	(–1.63)	(–0.77)
High–Low	0.033	1.011*	0.305	–0.911*	–0.297	–0.330	0.227
	(0.05)	(1.65)	(0.67)	(–1.80)	(–0.41)	(–0.34)	(0.28)

Panel D: Bivariate Portfolio Sort on Age (Age)

	Low	2	3	4	High	H–L	Alpha
Low	1.252	1.968*	0.772	0.114	–1.117*	–2.369***	–1.916***
	(1.64)	(1.69)	(1.02)	(0.16)	(–1.90)	(–3.50)	(–3.34)
Median	0.794	0.945	1.466**	0.758	0.569	–0.226	–0.081
	(1.08)	(1.11)	(2.04)	(1.04)	(0.75)	(–0.37)	(–0.14)
High	1.188*	1.723**	0.752	0.635	0.465	–0.723	–0.202
	(1.87)	(2.40)	(1.14)	(1.13)	(1.16)	(–1.34)	(–0.54)
High–Low	–0.064	–0.244	–0.020	0.522	1.582***	1.646**	1.714***
	(–0.14)	(–0.25)	(–0.03)	(0.88)	(2.88)	(2.31)	(2.72)

Panel E: Bivariate Portfolio Sort on Idiosyncratic Volatility ($Ivol$)

	Low	2	3	4	High	H–L	Alpha
Low	1.336*	1.582**	1.179*	0.579	0.348	–0.988	–0.333
	(1.88)	(2.14)	(1.72)	(1.01)	(0.98)	(–1.61)	(–0.75)
Median	0.813	1.191*	2.111*	0.880	0.628	–0.185	–0.280
	(1.07)	(1.54)	(1.80)	(1.05)	(0.70)	(–0.25)	(–0.42)
High	0.958	0.426	–0.122	0.563	–1.604**	–2.563***	–2.493***
	(1.20)	(0.52)	(–0.17)	(0.63)	(–1.98)	(–2.79)	(–2.97)
High–Low	–0.378	–1.156*	–1.301***	–0.016	–1.952***	–1.574	–2.160**
	(–0.68)	(–1.72)	(–2.66)	(–0.02)	(–2.64)	(–1.50)	(–2.45)

Panel F: Bivariate Portfolio Sort on Illiquidity ($ILIQ$)

	Low	2	3	4	High	H–L	Alpha
Low	1.275*	2.013**	1.207*	0.483	–0.134	–1.410**	–0.715*
	(1.92)	(2.33)	(1.94)	(0.90)	(–0.33)	(–2.56)	(–1.85)
Median	1.537**	1.407*	1.210*	0.745	–0.375	–1.912***	–1.529***
	(2.00)	(1.77)	(1.74)	(0.99)	(–0.56)	(–2.93)	(–2.63)
High	1.509*	0.513*	0.427	1.118	–0.843	–2.352***	–1.661**
	(1.93)	(0.71)	(0.64)	(1.07)	(–1.17)	(–2.76)	(–2.00)
High–Low	0.234	–1.500**	–0.780*	0.635	–0.709	–0.942	–0.946
	(0.47)	(–2.55)	(–1.84)	(0.63)	(–0.97)	(–1.00)	(–0.97)

Table 11: MFD and Investor Sentiment

This table reports the one-week-ahead excess-return of portfolios with the highest MFD and lowest MFD as well as their differences when investor sentiment is high or low. Equal-weighted and value-weighted one-week-ahead excess returns of different cryptocurrency terciles are sorted based on MFD and formed on a weekly basis. This table also reports the alphas (adjusted with the three-factor model proposed by Liu, Tsyvinski, and Wu (2022)) for each of the MFD-sorted high-minus-low portfolios. We define high (low) sentiment periods as weeks whose CMC Crypto Fear and Greed Index is higher (or not higher) than the median of each year. We also report the return spread difference between the low- and high-sentiment groups. Newey-West (1987) adjusted t statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal-Weighted		Value-Weighted	
	Low Sentiment (A)	High Sentiment (B)	Low Sentiment (A)	High Sentiment (B)
Low	0.081	2.295**	0.328	2.569***
	(0.10)	(2.46)	(0.40)	(2.73)
2	0.334	2.048**	0.107	1.410
	(0.42)	(2.32)	(0.12)	(1.57)
3	0.565	2.226**	2.171	2.733**
	(0.66)	(2.23)	(1.53)	(2.55)
4	0.658	1.495	0.046	2.212**
	(0.80)	(1.61)	(0.05)	(1.97)
5	0.513	1.175	1.045	1.604*
	(0.66)	(1.42)	(1.18)	(1.74)
6	0.328	1.498	0.645	1.414*
	(0.41)	(1.63)	(0.80)	(1.66)
7	0.438	1.421	0.003	0.315
	(0.56)	(1.53)	(0.00)	(0.31)
8	0.640	1.195	0.225	0.581
	(0.82)	(1.46)	(0.31)	(0.68)
9	0.325	0.639	–0.071	1.126
	(–0.43)	(0.78)	(–0.12)	(1.48)
High	–0.174	0.118	–0.836	–0.298
	(–0.20)	(0.12)	(–0.88)	(–0.35)
High-Low	–0.255	–2.177***	–1.164	–2.867***
	(–0.41)	(–2.92)	(–1.43)	(–3.44)
Alpha	–0.348	–1.683**	–1.327*	–2.061**
	(–0.58)	(–2.25)	(–1.89)	(–2.51)
Excess Return Diff. (A-B)	1.923**		1.703	
Alpha Diff. (A-B)	1.335*		0.734	

Table 12: MFD and Mispricing

This table reports results from the value-weighted bivariate portfolios based on dependent double sorts of the mispricing score (MISP) and MFD from January 2019 to June 2024. Following Stambaugh et al. (2015), we construct the cryptocurrency-level mispricing score (MISP) based on cryptocurrency size and momentum anomalies (Liu, Tsyvinski, and Wu, 2022). To calculate it, cryptocurrencies are sorted each week into deciles according to their size, and momentum. The decile ranks are attributed to each cryptocurrency increasing in their momentum, and decreasing in their size. Each cryptocurrency is given the corresponding score of its decile rank for each variable. Finally, the mispricing score on the cryptocurrency-level is the average of the two scores such that it ranges from 1 to 10. Low (high) MISP indicates undervaluation (overvaluation). Panel A reports the time-series averages of the weekly cross-sectional median of a mispricing score (MISP) for MFD-sorted univariate quintile portfolios. Panel B reports excess returns of 3×5 dependent bivariate value-weighted portfolio sorts. Tercile portfolios are formed every week using MISP, and then quintile portfolios are formed based on MFD within each cryptocurrency-specific MISP tercile. The last column reports the alpha (adjusted by the three-factor model proposed by Liu, Tsyvinski, and Wu (2022)) for each of the MFD-sorted high-minus-low portfolios. Newey and West's (1987) adjusted t-values are reported in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Average MISP in MFD Decile Portfolios

	Low	2	3	4	High	High–Low	t–stat
MISP	7.714	7.623	7.651	7.727	8.054	0.340***	5.82

Panel B: Bivariate Portfolio Sort on MISP

	Low	2	3	4	High	H–L	Alpha
MISP Low	0.833	1.157	0.616	0.343	–0.114	–0.947*	–0.721
	(1.17)	(1.52)	(0.86)	(0.55)	(–0.19)	(–1.82)	(–1.53)
MISP 2	1.180	1.872**	1.006	0.964	0.398	–0.782	–0.547
	(1.64)	(2.47)	(1.58)	(1.40)	(0.62)	(–1.24)	(–0.99)
MISP High	2.043**	2.322**	1.863**	0.666	–0.414	–2.457***	–1.866***
	(2.51)	(2.35)	(2.14)	(0.96)	(–0.60)	(–3.16)	(–2.75)
MISP High–Low	1.210**	1.166	1.246*	0.321	–0.300	–1.510*	–1.145
	(2.03)	(1.45)	(1.76)	(0.46)	(–0.47)	(–1.84)	(–1.45)

Table 13: MFD versus Change in Turnover

Panel A reports the average change in turnover (ΔTO) of MFD-sorted univariate quintile portfolios. Low (high) ΔTO indicates a lower (higher) average change in turnover. Panel B reports 3×5 dependent bivariate value-weighted portfolio sorts. First, quintile portfolios are formed every week using ΔTO. Next decile portfolios are formed based on MFD within each cryptocurrency-specific ΔTO quintile. Newey and West (1987) adjusted t-values are reported in parentheses. *, **, *** denote statistical significance at the 10%,5%, and 1% level, respectively. The sample period is from January 2019 to June 2024.

Panel A: Average ΔTO in MFD Decile Portfolio

	Low	2	3	4	High	High–Low	t–stat
ΔTO	–0.059	–0.034	–0.026	0.038	0.147	0.206***	4.04

Panel B: Bivariate Portfolio Sort on ΔTO

	Low	2	3	4	High	H–L	Alpha
ΔTO Low	1.630**	0.829	0.500	–0.067	–0.015	–1.645***	–1.045**
	(2.38)	(1.22)	(0.78)	(–0.11)	(–0.03)	(–2.70)	(–2.11)
ΔTO 2	1.049	1.701**	1.235*	1.555*	0.071	–0.978	–0.529
	(1.48)	(2.36)	(1.81)	(1.81)	(0.12)	(–1.61)	(–1.13)
ΔTO High	1.120	1.769*	1.107	0.246	–0.986	–2.106***	–1.651***
	(1.43)	(1.67)	(1.53)	(0.37)	(–1.34)	(–2.86)	(–2.75)
ΔTO High–Low	–0.510	0.941	0.607	0.313	–0.972	–0.462	–0.607
	(–0.91)	(1.10)	(1.08)	(0.51)	(–1.29)	(–0.50)	(–0.72)

Table 14: Fama-MacBeth Cross-Sectional Regressions on MFD and Turnover-Based Disagreement

This table reports Fama–MacBeth cross–sectional regression for MFD while additionally controlling for turnover–based disagreement measure. MFD, ΔTO, and other control variables in week $\boldsymbol{t}$ are matched to cryptocurrency returns in week $\boldsymbol{t+1}$. The dependent variable is the cryptocurrencies' future return in the first row. The control variables are described in Table 1. Winsorized at 1%. Newey and West (1987) adjusted t−statistics are reported in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 2019 to June 2024.

	Excess Return	Excess Return	Excess Return	Excess Return
MFD			–7.005***	–5.345**
			(–2.99)	(–2.33)
ΔTO	–0.005	–0.004	–0.001	–0.002
	(–1.18)	(–0.97)	(–0.13)	(–0.50)
MOM		0.011		0.014
		(0.87)		(1.16)
$Size$		–0.001		–0.001*
		(–1.87)		(–1.81)
$Abvol$		–0.001		–0.001
		(–0.63)		(–0.39)
$Beta$		0.0104		0.009
		(1.62)		(1.45)
$Ivol$		–0.065		–0.012
		(–1.41)		(–0.23)
$Volt$		–0.121		–0.110
		(–1.06)		(–0.95)
MAX		0.0141		0.013
		(0.29)		(0.27)
$CosKew$		–0.000		0.000
		(–0.62)		(0.40)
$ILLIQ$		0.001		0.001
		(0.68)		(0.45)
$Turnover$		–0.001		–0.001
		(–0.83)		(–0.96)
$Constant$	0.009	0.029	0.016**	0.031**
	(1.41)	(2.00)	(2.19)	(2.19)
$Adj.R^2$	0.008	0.139	0.029	0.153
N	79174	79174	79174	79174

Table 15: Predicting weekly Turnover and Historical Volatility

Panel A reports results for panel regression of trading volume and cryptocurrency volatility on MFD. For trading volume, we use weekly turnover as a proxy. For volatility, we use the standard deviation of daily cryptocurrencies returns in a week. We estimate the following regression $y_{i,t} = \alpha_i + \gamma_t + \beta_1 \times MFD_{i,t} + \beta_2 \times y_{i,t-1} + \gamma \times Controls_{i,t} + \epsilon_{i,t}$, where $y_{i,t}$ is either weekly turnover or the historical volatility of cryptocurrency i in week t. We also include the first lag of our dependent variable to account for persistence in volume and volatility, respectively, and add week (γ_t) and cryptocurrency (α_i) fixed effects. Additional controls ($Controls_{i,t}$) consists of the short–term reversal (Ret), momentum (MOM), cryptocurrency market capitalization (Size), trading volume (Volume), cryptocurrency beta (Beta), and lottery–like (MAX) variables. Standard errors are double–clustered by week and cryptocurrency. The dependent and independent variables are winsorized at 1% in both tails. *, **, *** denote statistical significance at the 10%,5%, and 1% level, respectively. The sample period is from January 2019 to June 2024.

	$Volt_t$	$Volt_{t+1}$	$Turnover_t$	$Turnover_{t+1}$
MFD_t	3.532***	4.415***	15.133***	14.380***
	(15.88)	(12.33)	(3.09)	(3.05)
$Volt_{t-1}$	0.037***			
	(9.71)			
$Volt_t$		0.340***		
		(11.22)		
$Turnover_{t-1}$			0.281***	
			(13.21)	
$Turnover_t$				0.828***
				(49.89)
$Controls$	YES	YES	YES	YES
$Entity\_fixed\ effect$	YES	YES	YES	YES
$Week\_fixed\ effect$	YES	YES	YES	YES
Obs.	77,506	77,506	76,057	76,057
Adj R^2	0.933	0.435	0.968	0.889
$Cluster\ by\ entity\ and\ week$	YES	YES	YES	YES

Appendix A: Variables for the Construction of MFD

No.	Category	predictor	Reference	Definition
1	Size	MCAP	Banz (1981)	Log last–day market capitalization in the portfolio formation week.
2	Size	PRC	Miller and Scholes (1982)	Log last–day price in the portfolio formation week.
3	Size	MAXDPRC	George and Hwang (2004)	Maximum price of the portfolio formation week.
4	Size	AGE	Barry and Brown (1984)	Log number of days listed onCoinmarketcap.com
5	Mom	R1.0	Jegadeesh and Titman (1993)	Past one–week return
6	Mom	R2.0	Jegadeesh and Titman (1993)	Past two–week return
7	Mom	R3.0	Jegadeesh and Titman (1993)	Past three–week return
8	Mom	R4.0	Jegadeesh and Titman (1993)	Past four–week return
9	Mom	R4.1	Jegadeesh and Titman (1993)	Past one–to–four–week return
10	Mom	R8.0	Jegadeesh and Titman (1996)	Past eight–week return
11	Mom	R16.0	Jegadeesh and Titman (1993)	Past 16–week return
12	Mom	R32.0	De Bondt and Thaler (1985)	Past 32–week return
13	Mom	R52.0	De Bondt and Thaler (1986)	Past 52–week return
14	Volume	VOL	Chordia, Subrahmanyam, and Anshuman (2001)	Log average daily volume in the portfolio formation week.
15	Volume	ABVOL	Chen, Jian, et al. (2022)	Log average daily volume for a cryptocurrency in the portfolio formation week is subtracted by the log average daily volume of the past 12 weeks.
16	Volume	PRCVOL	Chordia, Subrahmanyam, and Anshuman (2001)	Log average daily volume time price in the portfolio formation week.

17	Volume	VOLSCALED	Chordia, Subrahmanyam, and Anshuman (2001)	Log average daily volume times price scaled by market capitalization in the portfolio formation week.
18	Vol	BETA	Fama and MacBeth (1973)	The regression coefficient β_{CMKT}^{i} in $R_i - R_f = \alpha^i + \beta_{CMKT}^{i}(CMKT_t - R_f) + \varepsilon_i$. The model is estimated using daily returns of the previous 90 days before the portfolio formation week.
19	Vol	BETA2	Fama and MacBeth (1973)	Beta square.
20	Vol	IDIOVOL	Ang et al. (2006)	Idiosyncratic volatility, measured as the standard deviation of the residual after estimating $R_{i,t} - R_{f,t} = \alpha^i + \beta_{CMKT}^{i}(CMKT_t - R_f) + \varepsilon_i$. The model is estimated using daily returns of the previous 90 days before the portfolio formation week.
21	Vol	DBETA	Ang et al. (2006)	The market beta during the downside period: $DBETA_i = \frac{cov(R_{i,t}, CMKT_t \mid CMKT_t < \overline{CMKT})}{var(, CMKT_t \mid CMKT_t < \overline{CMKT})}$, where $\overline{CMKT}$ is the average market returns. The model is estimated using daily returns of the previous 90 days before the portfolio formation week.
22	Vol	UBETA	Scholes and Williams (1977)	The market beta during the upside period: $UBETA_i = \frac{cov(R_{i,t}, CMKT_t \mid CMKT_t > \overline{CMKT})}{var(, CMKT_t \mid CMKT_t > \overline{CMKT})}$, where $\overline{CMKT}$ is the average market returns. The model is estimated using daily returns of the previous 90 days before the portfolio formation week.

23	Vol	SWBETA	Scholes and Williams (1977)	$SWBETA_i = \frac{DBETA_i + BETA_i + UBETA_i}{1+2\rho}$, where ρ is the first–order serial correlation of the market portfolio's excess return.
24	Vol	DIMSONBETA	Dimson (1979)	$\text{DIMSONBETA}_i = \sum_{k=-5}^{k=5} \hat{\beta}_k^i$, where the $\hat{\beta}_k^i$ are the estimated slope coefficients from the regression model: $R_i - R_f = \alpha^i + \sum_{k=-5}^{k=5} \beta_k^i (CMKT_{t-k} - R_f) + \varepsilon_i$.
25	Vol	RETVOL	Ang et al. (2006)	The standard deviation of daily return in the portfolio formation week.
26	Vol	MAXRET	Bali Cakici and Whitelaw (2011)	Maximum daily return of the portfolio formation week.
27	Vol	MAXRET3	Bali Cakici and Whitelaw (2011)	Average maximum daily return of the previous three weeks before the portfolio formation week.
28	Vol	MAXRET5	Bali Cakici and Whitelaw (2011)	Average maximum daily return of the previous five weeks before the portfolio formation week.
29	Vol	DELAY	Hou and Moskowitz (2005)	The improvement of R2 in $R_{i,t} - R_{f,t} = \alpha^i + \beta_{CMKT}^i CMKT_t + \beta_{CMKT-1}^i CMKT_{t-1} + \beta_{CMKT-2}^i CMKT_{t-2} + \varepsilon_i$, where $CMKT_{t-1}$ and $CMKT_{t-2}$ are the lagged one– and two–day cryptocurrency market index excess returns, compared to using only current cryptocurrency market excess returns. The model is estimated using daily returns of the previous 90 days before the formation week.
30	Vol	STDPRCVOL	Chordiam Subrahmanyam and Anshuman (2001)	Log the standard deviation of price volume in the portfolio formation week.
31	Vol	STDPRC	Chordiam Subrahmanyam and Anshuman (2001)	Log the standard deviation of price in the portfolio formation week.

32	Vol	TSKEW	Bali Engle and Murry (2016)	The skewness of historical realized cryptocurrency returns. $TSKEW_i = \frac{\frac{1}{n}\sum_{t=1}^{n}(R_{i,t}-\overline{R_i})^3}{(\frac{1}{n}\sum_{t=1}^{n}(R_{i,t}-\overline{R_i})^2)^{3/2}}$, where $\overline{R_i}$ is the average periodic return of cryptocurrency i. The model is estimated using daily returns of the previous 90 days before the formation week.
33	Vol	IDIOSKEW	Boyer Mitton and Vorkink (2010)	The skewness of the residuals CAPM model regression: $IDIOSKEW_i = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_{i,t}{}^3}{(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_{i,t}{}^3)^{3/2}}$, where $\varepsilon_{i,t}$ are the residuals from the regression model: $R_{i,t} - R_{f,t} = \alpha^i + \beta_{CMKT}^i CMKT_t + \varepsilon_i$. The model is estimated using daily returns of the previous 90 days before the portfolio formation week.
34	Vol	COSKEW	Harvey and Siddique (2000)	The slope coefficient on squared excess market return from a regression of excess cryptocurrency returns on the excess return of the market portfolio and the squared excess market return: $R_{i,t} - R_{f,t} = \alpha^i + \beta_{CMKT}^i CMKT_t + COSKEW_i CMKT_t^2 + \varepsilon_i$, The model is estimated using daily returns of the previous 90 days before the portfolio formation week.
35	Vol	MAXVOL	Assness, Frazzini, Gormsen, Pedersen (2020)	Maximum daily return scaled by the daily return volatility of the portfolio formation week.
36	Liquidity	ILLIQ1	Amihud (2002), Bali Engle and Murry (2016)	Average absolute daily return divided by price volume during the portfolio formation week

37	Liquidity	ILLIQ3	Amihud (2002)	Average absolute daily return divided by price volume during the past 3 weeks.
38	Liquidity	ILLIQ6	Bali Engle and Murry (2016)	Average absolute daily return divided by price volume during the past 6 weeks.
39	Liquidity	ILLIQ12	Amihud (2002)	Average absolute daily return divided by price volume during the past 12 weeks.
40	Reversal	RET	Bali Engle and Murry (2016)	The return during the portfolio formation week.

Appendix B: Change in Turnover

Following Garfinkel et al. (2006; 2009), we use the change in turnover (ΔTO) as the proxy for investor disagreement. We begin by calculating weekly turnover for each cryptocurrency each week as the average cryptocurrency daily trading volume scaled by its market capitalization. However, some cryptocurrencies may exhibit more trading in some weeks because a macroeconomic event creates significant trading in the whole market. We therefore, subtract market–wide turnover calculated the same way, but across all cryptocurrencies in the cryptocurrency market. This creates our weekly market–adjusted turnover measure.

Then, cryptocurrencies with relatively higher turnover at some weeks may reasonably be the same cryptocurrencies with relatively higher turnover overall. Thus, market–adjusted turnover may capture more than just volume attributable to investor disagreements; it can also include liquidity trading. We therefore market–adjusted turnover is subtracted by market–adjusted turnover averaged over the past 48 weeks.

$$\Delta TO = \left(TO_{i,t} - TO_{MKT,t}\right) - \frac{1}{48}\sum_{j=t-1}^{t-48}\left(TO_{i,t} - TO_{MKT,t}\right) \tag{B.1}$$

Internet Appendix

Machine Forecast Disagreement in the Cryptocurrency Market

Table of contents:

Internet Appendix A. Influence of extreme price/return outliers

In Section III.A, we document that cryptocurrencies with high MFD tend to be smaller on average. Nevertheless, the negative relation between MFD and future returns is stronger for value–weighted portfolios and remains robust when the sample is restricted to larger cryptocurrencies. Moreover, value–weighted predictability persists for up to five weeks, whereas equal–weighted predictability appears considerably more volatile. In this section we provide an economic explanation for these patterns.

A key challenge in cryptocurrency research concerns data quality and the prevalence of extreme price/return outliers. Figure IA1 illustrates the distribution of weekly maximum returns from January 2019 to June 2024. In the raw data, we observe extreme daily returns, sometimes exceeding 10000%, that often reverse sharply the following day, occasionally resulting in losses greater than 99%. Although we impose multiple sample filters to mitigate the influence of such extreme observations, extreme returns remain present in the data. As shown in Figure IA1, approximately three–quarters of weekly maximum returns exceed 50%, and nearly one–quarter exceed 150%, with some surpassing 300%.

We next examine the distribution of extreme returns across cryptocurrencies with different market capitalization and MFD levels. Specifically, we independently sort cryptocurrencies each week into size and MFD deciles and compute, for each decile pair, the share of observations that fall into the top 1% of one–week–ahead returns. Figure IA2 presents the resulting distribution. Approximately 3.54% of the top 1% extreme return observations are concentrated in the portfolio with the highest market capitalization and highest MFD, substantially exceeding the 1% benchmark. These findings provide a potential explanation for the empirical patterns documented in Section III.A.

To account for extreme outliers and hedge against potential database errors, we winsorize weekly cryptocurrency returns at the 1st and 99th percentiles. Table IA1 replicates the baseline results in Table 3 using winsorized returns. After winsorization, equal–weighted returns become comparable in magnitude to value–weighted returns. This evidence indicates that extreme return outliers weaken the return predictability of MFD when portfolios are equal–weighted. In particular, extreme positive returns outliers lead to higher (or less negative) future equal–weighted returns for the high–MFD portfolio. The high–MFD portfolio has an equal–weighted excess return of –0.86% with winsorized returns, compared with –0.027% without winsorization. By contrast, the low–MFD portfolio exhibits similar future returns with and without winsorization.

Finally, we assess the impact of extreme return outliers on the long–term predictive power of MFD by replicating the analysis in Table 5 using winsorized weekly returns. Table IA2 reports excess returns and three–factor alphas for MFD–sorted portfolios from 2 to 12 weeks after portfolio formation. Consistent with the short–horizon evidence, extreme outliers attenuate the long–term predictive power of MFD for equal–weighted portfolios. Specifically, extreme positive return outliers lead to higher (or less negative) future equal–weighted returns for the high–MFD portfolio. Although equal–weighted returns and three–factor spreads remain significantly negative up to 12 weeks, they exhibit some volatility across horizons. By contrast, the value–weighted return predictability of MFD persists only for 10 weeks.

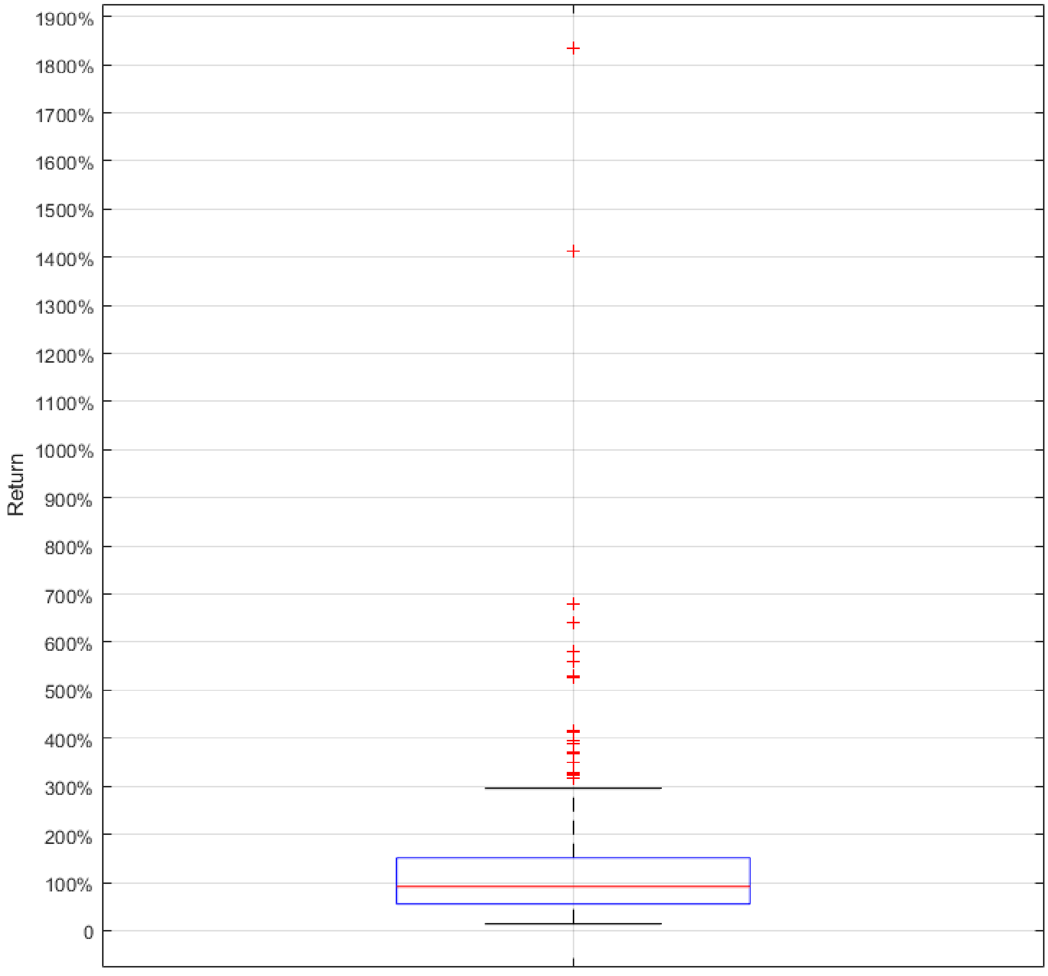

Figure IA1. The Distribution of the Maximum Return of Each Week

This figure presents the distribution of the maximum return of each week.

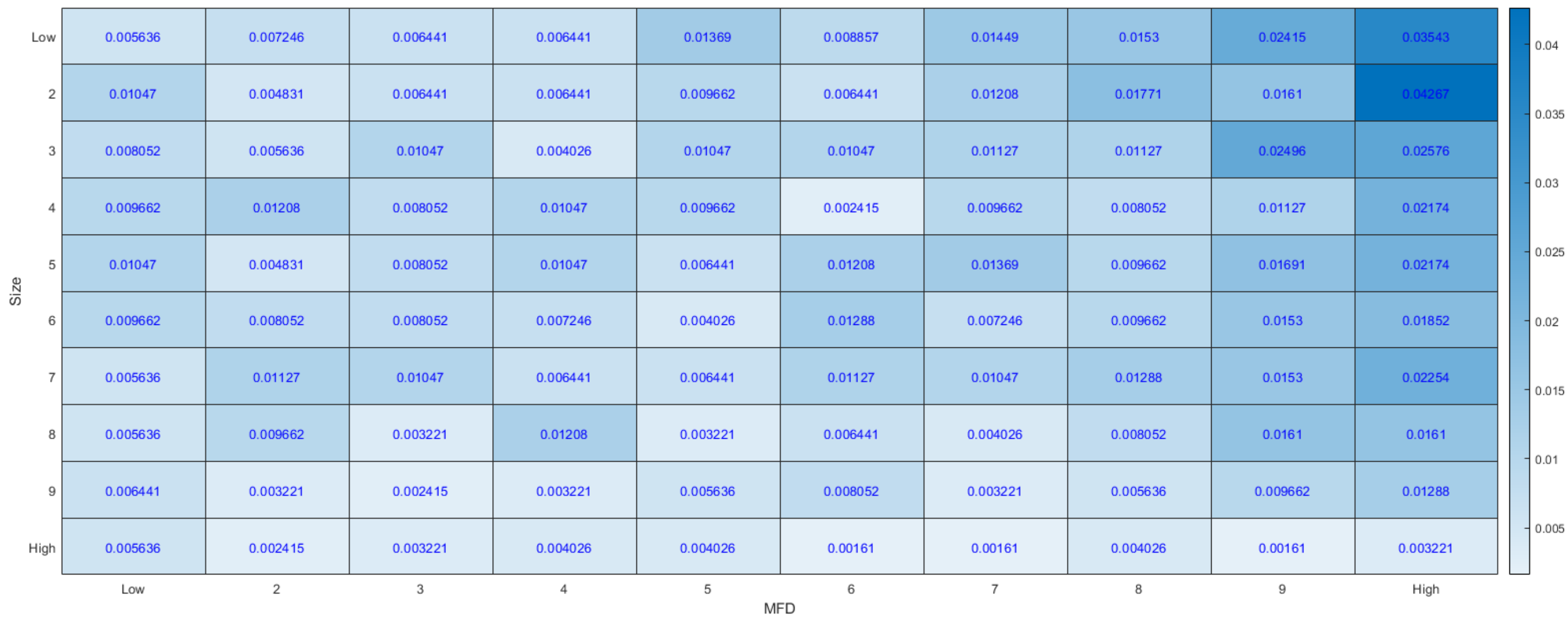

Figure IA2. The Distribution of Extreme High Value of One–week–ahead Return on Size and MFD Portfolios

This figure presents the distribution of top 1% extreme value of one–week–ahead return on size and MFD portfolios. First, all cryptocurrencies are sorted each week into deciles according to their size and MFD. Then, each week, we count the number of cryptocurrencies with top 1% one–week–ahead return in each size and MFD decile. Finally, we calculate the percentage of top 1% extreme value of one–week–ahead return in each size and MFD decile. The sample is from January 2019 to June 2024.

Table IA1. Univariate Portfolio Sort on MFD: Winsorize at 1% level

This table reports excess returns and alphas on different cryptocurrency deciles sorted based on MFD and formed on weekly basis between January 2019 and June 2024. To further account for extreme outliers and hedge against potential database errors, we winsorize returns at the 1% and 99% percentiles. Each week, cryptocurrencies are sorted into decile portfolios by MFD. Portfolio Low is the portfolio of cryptocurrencies with the lowest MFD, while cryptocurrencies with the highest MFD comprise portfolio High. All equal–weighted and value–weighted excess returns and alphas are a week ahead of the portfolio formation period. Excess return is the return in excess of the risk–free rate. Alpha is the intercept from a time–series regression of weekly excess returns on the factors of alternative models: CCAPM and three–factor models proposed by Liu, Tsyvinski, and Wu (2022). This table also reports differences between portfolios High and Low. All measures are reported in percentage terms. Newey–West (1987) adjusted t statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal–Weighted			Value–Weighted		
	Excess Return	CCAPM	3–factor	Excess Return	CCAPM	3–factor
Low	1.034	–0.102	–0.123	1.350**	0.220	0.296
	(1.45)	(–0.27)	(–0.35)	(2.01)	(0.65)	(0.87)
2	0.974	–0.156	–0.168	0.738	–0.487	–0.535*
	(1.44)	(–0.45)	(–0.53)	(1.06)	(–1.55)	(–1.80)
3	1.059	–0.120	–0.157	2.176***	0.846**	0.735**
	(1.49)	(–0.34)	(–0.48)	(2.65)	(2.33)	(2.08)
4	0.799	–0.332	–0.316	0.908	–0.338	–0.350
	(1.19)	(–0.98)	(–0.99)	(1.36)	(–1.01)	(–1.09)
5	0.618	–0.441	–0.487	1.274*	0.120	0.101
	(0.96)	(–1.28)	(–1.52)	(1.91)	(0.34)	(0.29)
6	0.627	–0.514*	–0.459	0.981*	–0.117	0.004
	(0.96)	(–1.67)	(–1.50)	(1.67)	(–0.38)	(0.01)
7	0.498	–0.563*	–0.467	0.189	–0.900***	–0.595*
	(0.79)	(–1.69)	(–1.49)	(0.31)	(–2.81)	(–1.72)

8	0.518	–0.524*	–0.597**	0.263	–0.636*	–0.627*
	(0.84)	(–1.67)	(–2.01)	(0.45)	(–1.65)	(–1.70)
9	–0.120	–1.075***	–1.077***	0.471	–0.222	–0.265
	(–0.21)	(–3.19)	(–3.36)	(1.01)	(–0.68)	(–0.84)
High	–0.860	–1.838***	–1.834***	–0.578	–1.478***	–1.432***
	(–1.50)	(–5.69)	(–5.82)	(–0.95)	(–3.47)	(–3.64)
High–Low	–1.894***	–1.736***	–1.711***	–1.928***	–1.699***	–1.728***
	(–5.03)	(–5.16)	(–5.33)	(–3.29)	(–3.20)	(–3.49)

Table IA2. Long–Term Portfolio Returns: Winsorize at 1% level

This table reports the long-term predictive power of MFD. For each week $t+n$, where $n \in \{2, 3, \ldots, 12\}$, individual cryptocurrencies are sorted into decile portfolios based on $week - t$ MFD. Each week, returns are winsorized at 1st and 99th percentiles. The Table reports the average weekly return and alpha (adjusted with three-factor model proposed by Liu, Tsyvinski, and Wu (2022)) of the equal-weighted and value-weighted MFD-sorted high-minus-low portfolios. Newey-West (1987) adjusted t statistics are presented in parentheses. $*$, $**$, and $***$ indicate the significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 2019 and June 2024.

	Equal-Weighted		Value-Weighted	
	High–Low	Alpha	High–Low	Alpha
$t+2$	–1.205***	–1.141***	–1.563**	–1.156**
	(–3.41)	(–3.70)	(–2.32)	(–2.38)
$t+3$	–1.386***	–1.260***	–1.540***	–1.133**
	(–3.54)	(–4.02)	(–2.61)	(–2.37)
$t+4$	–1.241***	–1.088***	–2.170***	–1.546***
	(–3.95)	(–3.93)	(–3.78)	(–3.57)
$t+5$	–1.058***	–0.915***	–1.065*	–0.799
	(–3.68)	(–3.32)	(–1.79)	(–1.49)
$t+6$	–1.054***	–0.968***	–1.119**	–0.871**
	(–3.57)	(–3.53)	(–2.19)	(–2.09)
$t+7$	–0.404	–0.269	–1.063	–0.560
	(–1.20)	(–0.85)	(–1.54)	(–0.91)
$t+8$	–1.042***	–0.939***	–2.088***	–1.604***
	(–3.46)	(–3.24)	(–3.82)	(–3.35)
$t+9$	–0.846***	–0.729**	–1.079**	–0.678
	(–2.64)	(–2.48)	(–2.05)	(–1.43)
$t+10$	–0.625*	–0.509*	–1.042*	–0.579

	(–1.88)	(–1.74)	(–1.73)	(–1.16)
$t + 11$	–0.445*	–0.359	–0.612	–0.136
	(–1.86)	(–1.38)	(–1.21)	(–0.29)
$t + 12$	–0.569**	–0.482*	–0.564	–0.596
	(–2.10)	(–1.70)	(–1.09)	(–1.28)

Internet Appendix B. Past-Return Information and MFD

Section IV.A examines the role of past returns. In this section, we provide additional evidence using alternative sets of key characteristics to further assess the robustness of our conclusions. Past returns and price dynamics may contribute to disagreement among investors when forming expectations about future outcomes. To further explore whether information related to past returns contributes disproportionately to MFD, we classify cryptocurrency characteristics into two groups. **The return information set** consists of 13 characteristics related to past returns and price dynamics; whereas the **non–return information set** comprises 27 characteristics that are not directly related to past returns or price dynamics.

We extend the baseline analysis by assuming that all 13 return characteristics to be potentially informative about disagreement related to past-return information. Using all 13 return characteristics, we construct a return–based MFD. By comparison, we construct a non–return–based MFD using 13 characteristics randomly selected from the non–return information set. In addition, we construct a return–plus–based MFD using a mixed information set that combines the 13 return characteristics with 11 characteristics randomly selected from the non–return information set.

Table IB1 replicates the results of Table 9 using these alternative sets of 13return characteristics. The value–weighted return and three–factor alpha spreads based on the return–based MFD are –2.11% and –1.41%, with t–values of –3.09 and –2.62, respectively. By contrast the return and three–factor alpha spreads based on the non–return–based MFD are –1.34% and –1.00%, with t-values of –1.80 and –1.64, respectively, indicating that the predictive power of the non-return-based MFD is much weaker than that of return-based MFD, although the difference is insignificant. Moreover, the return-plus–based MFD does not have stronger return predictability than the return-based MFD, suggesting that non-return characteristics only provide limited incremental information in the formation of belief disagreement. These results are consistent with the view that MFD is more strongly associated with disagreement arising from past-return-related information.

Table IB1. The Role of Past Return: 13 Return Characteristics

This table reports the average weekly spread return between the highest and lowest decile of univariate portfolios based on return–based MFD, non–return–based MFD, and return–plus–based MFD, respectively. The return information set consists of 13 characteristics related to past returns and price dynamics, such as past one–week return (MOM_1), past two–week return (MOM_2), past three–week return (MOM_3), past four–week return (MOM_4), past one–to–four–week return ($MOM_{4\to1}$), past eight–week return (MOM_8), past sixteen–week return (MOM_{16}), past thirty–two–week return (MOM_{32}), past fifty-two–week return (MOM_{52}), maximum daily return in the portfolio–formation week (MAX_1), average maximum daily return over the previous three weeks before the portfolio–formation week (MAX_3), average maximum daily return over the previous five weeks before the portfolio–formation week (MAX_5), and return on the portfolio–formation week (RET). The non–return information set consists of the remaining 27 characteristics that are not directly related to past returns and price dynamics. Return–based MFD is constructed using all 13 return characteristics selected from return information set. Non–return–based MFD is constructed using 13 characteristics randomly selected from the non–return information set. Return–plus–based MFD is constructed using a mixed information set comprising 13 return characteristics and 11 characteristics randomly selected from non–return information set. Each week, cryptocurrencies are sorted into decile portfolios by MFD. Portfolio Low is the portfolio of cryptocurrencies with the lowest MFD, while cryptocurrencies with the highest MFD comprise portfolio High. All value–weighted excess returns and alphas are measured one week ahead of the portfolio formation period. Excess return is the return in excess of the risk–free rate. Alpha is the intercept from a time–series regression of weekly excess returns on the three–factor models proposed by Liu, Tsyvinski, and Wu (2022). This table also reports differences between portfolios High and Low. Besides, this table reports the return difference between return-based MFD and non-return-based MFD, and between return-plus-based MFD and non-return-based MFD. All measures are reported in percentage terms. Newey–West (1987) adjusted t-statistics are presented in parentheses. $*$, $**$, and $***$ indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Return–based MFD (A)		Non–return–based MFD (B)		Return–plus–based MFD (C)		Diff (A-B)		Diff (C-B)	
	Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha
Low	1.407*	0.176	1.584*	0.383	1.891**	0.583*	–0.177	–0.208	0.307	0.199
	(1.86)	(0.52)	(1.85)	(0.80)	(2.51)	(1.76)	(–0.32)	(–0.43)	(0.70)	(0.46)
2	1.340*	0.183	1.310	0.005	1.495*	0.184	0.030	0.178	0.184	0.178
	(1.87)	(0.57)	(1.63)	(0.01)	(1.81)	(0.38)	(0.06)	(0.38)	(0.48)	(0.48)
3	1.138*	–0.096	0.853	–0.388	1.491**	0.170	0.285	0.293	0.638*	0.558

	(1.83)	(–0.35)	(1.17)	(–1.07)	(1.96)	(0.49)	(0.71)	(0.79)	(1.75)	(1.58)
4	1.069*	–0.207	1.323*	0.003	1.264*	–0.022	–0.254	–0.210	–0.059	–0.025
	(1.78)	(–0.90)	(1.68)	(0.01)	(1.72)	(–0.06)	(–0.55)	(–0.48)	(–0.11)	(–0.05)
5	0.373	–0.617**	1.761**	0.573	1.111*	–0.064	–1.388***	–1.190***	–0.650	–0.637
	(0.60)	(–2.13)	(2.53)	(1.56)	(1.65)	(–0.16)	(–3.08)	(–2.78)	(–1.52)	(–1.51)
6	0.928	–0.167	1.575**	0.243	0.531	–0.413	–0.647*	–0.410	–1.045**	–0.656*
	(1.35)	(–0.43)	(2.22)	(0.76)	(0.84)	(–1.19)	(–1.67)	(–1.12)	(–2.50)	(–1.74)
7	2.373**	1.201	0.713	–0.353	0.797	–0.019	1.660**	1.554**	0.084	0.334
	(2.22)	(1.55)	(1.12)	(–1.21)	(1.26)	(–0.05)	(1.98)	(2.15)	(0.19)	(0.82)
8	1.215	0.033	0.192	–0.871***	0.423	–0.362	1.023*	0.905*	0.230	0.509
	(1.54)	(0.07)	(0.30)	(–2.93)	(0.66)	(–0.90)	(1.77)	(1.72)	(0.49)	(1.20)
9	–0.101	–0.751*	0.495	–0.142	0.086	–0.711*	–0.596	–0.609	–0.409	–0.569
	(–0.17)	(–1.73)	(0.83)	(–0.38)	(0.16)	(–1.80)	(–0.97)	(–1.05)	(–0.78)	(–1.11)
High	–0.699	–1.238***	0.242	–0.615*	–0.244	–1.002***	–0.941*	–0.623	–0.487	–0.387
	(–1.29)	(–2.71)	(0.45)	(–1.76)	(–0.44)	(–2.71)	(–1.77)	(–1.18)	(–1.39)	(–1.18)

High–Low	–2.106***	–1.414***	–1.341*	–0.998	–2.135***	–1.584***	–0.765	–0.415	–0.794	–0.586
	(–3.09)	(–2.62)	(–1.80)	(–1.64)	(–3.38)	(–3.42)	(–0.99)	(–0.57)	(–1.29)	(–1.00)

Internet Appendix C. Limits–to–Arbitrage and MFD

In Section IV.B, we investigate the return predictability of MFD conditional on specific proxies for limits–to–arbitrage using bivariate portfolio analysis. We find that the negative relation between MFD and future returns is strongest among the youngest, most volatile and illiquid cryptocurrencies. However, in contrast to the theory of Miller (1977), the negative MFD–return relation is weak and insignificant among the smallest cryptocurrencies. As shown in Figure IA2, some small and high MFD cryptocurrencies experience extreme positive returns. To mitigate the influence of extreme outliers, we winsorize returns at the 1^{st} and 99^{th} percentiles. We then re–examine the impact of four limit–to–arbitrage proxies on the MFD–return relation using bivariate portfolio analysis.

Table IC1 reports the results. Panel A shows the results for the value–weighted bivariate portfolios using size as the proxy for limits–to–arbitrage. For the smallest size tercile, the return and three–factor alpha spreads are –1.13% and –1.11%, with t–values of –2.16 and –2.30, respectively. As expected, both the return and three–factor alpha spreads are smaller in absolute magnitude for the largest size tercile portfolio, with neither being statistically significant.

Panels B–D present the results from the value–weighted bivariate portfolio using age, idiosyncratic volatility, and illiquidity as proxies for limits–to–arbitrage, respectively. In all three panels, the return and three–factor alpha spreads in the low limits–to–arbitrage tercile portfolios are significantly larger in absolute magnitude than those in the high limits–to–arbitrage tercile portfolios. More importantly, the difference in return and three–factor alpha spreads between low and high limits–to–arbitrage tercile portfolios increase in absolute magnitude after excluding extreme outliers.

Overall, Table IC1 shows that the negative cross–sectional MFD–return relation is more pronounced for cryptocurrencies with smaller size, youngest, higher volatility, and lower liquidity, consistent with the view that arbitrage frictions amplify the return predictability associated with MFD.

Table IC1. MFD and Four Limits–to–Arbitrage Measures: Winsorizing Returns

This table reports results from the value–weighted bivariate portfolios based on independent double sorts of four limits–to–arbitrage measures and MFD from January 2019 to June 2024. The cryptocurrency–level limits–to–arbitrage measures include cryptocurrency size ($Size$), age (Age), idiosyncratic volatility ($IVOL$), and illiquidity ($ILLIQ$). A low (high) cryptocurrency size, low (high) cryptocurrency age, high (low) idiosyncratic volatility, and high (low) illiquidity indicate a high (low) arbitrage cost. Panel A reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of $Size$ and MFD. Tercile portfolios are formed every week using $Size$, and then quintile portfolios are formed based on MFD within each crypto–specific $Size$ tercile. Panel B reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of Age and MFD. Tercile portfolios are formed every week using Age, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific Age tercile. Panel C reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of $IDIO$ and MFD. Tercile portfolios are formed every week using $IDIO$, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific $IDIO$ tercile. Panel D reports excess returns of 3×5 dependent bivariate value–weighted portfolio sorts of $ILLIQ$ and MFD. Tercile portfolios are formed every week using $ILLIQ$, and then quintile portfolios are formed based on MFD within each cryptocurrency–specific $ILLIQ$ tercile. The last column reports the alpha adjusted by the three–factor model proposed by Liu, Tsyvinski, and Wu (2022) for each of the MFD–sorted high–minus–low portfolios. To deal with extreme outliers and hedge against potential database errors, we winsorize returns at the 1st and 99th percentiles. All alphas are reported in percentage terms. Newey and West's (1987) adjusted t–statistics are reported in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Bivariate Portfolio Sort on Size (Size)

	Low	2	3	4	High	H–L	Alpha
Low	0.900	0.729	0.396	0.451	–0.226	–1.125**	–1.108**
	(1.14)	(1.03)	(0.58)	(0.70)	(–0.32)	(–2.16)	(–2.30)
Median	1.107	0.810	0.565	0.623	–0.975*	–2.082***	–1.874***
	(1.55)	(1.17)	(0.83)	(0.93)	(–1.71)	(–5.15)	(–5.12)
High	1.253*	2.138***	1.087*	0.223	0.317	–0.936	–0.324
	(1.83)	(2.73)	(1.70)	(0.44)	(0.73)	(–1.64)	(–0.80)
High–Low	0.354	1.409***	0.690*	–0.228	0.543	0.189	0.784
	(0.58)	(2.98)	(1.66)	(–0.52)	(0.88)	(0.23)	(1.18)

Panel B: Bivariate Portfolio Sort on Age (Age)

	Low	2	3	4	High	H–L	Alpha
Low	1.167	1.390	0.676	–0.063	–1.219**	–2.386***	–1.925***
	(1.55)	(1.59)	(0.90)	(–0.09)	(–2.19)	(–3.70)	(–3.55)
Median	0.741	0.823	1.390*	0.643	0.405	–0.336	–0.195
	(1.02)	(0.99)	(1.92)	(0.91)	(0.56)	(–0.63)	(–0.39)
High	1.175*	1.678**	0.713	0.622	0.430	–0.745	–0.222
	(1.85)	(2.37)	(1.09)	(1.11)	(1.08)	(–1.39)	(–0.60)
High–Low	0.008	0.288	0.037	0.686	1.650***	1.642**	1.703***
	(0.02)	(0.43)	(0.06)	(1.18)	(3.14)	(2.40)	(2.83)

Panel C: Bivariate Portfolio Sort on Idiosyncratic Volatility (IVOL)

	Low	2	3	4	High	H–L	Alpha
Low	1.314*	1.467**	1.170*	0.559	0.340	–0.974	–0.333
	(1.87)	(2.05)	(1.71)	(0.98)	(0.96)	(–1.61)	(–0.76)
Median	0.722	1.099	1.638*	0.685	0.560	–0.162	–0.266
	(0.95)	(1.43)	(1.82)	(0.87)	(0.63)	(–0.23)	(–0.42)
High	0.716	0.248	–0.311	–0.158	–1.879***	–2.594***	–2.497***
	(0.89)	(0.32)	(–0.42)	(–0.20)	(–2.61)	(–3.27)	(–3.60)
High–Low	–0.598	–1.219**	–1.480***	–0.717	–2.219***	–1.621*	–2.164***
	(–1.11)	(–1.99)	(–3.03)	(–1.23)	(–3.44)	(–1.73)	(–2.86)

Panel D: Bivariate Portfolio Sort on Illiquidity (IILIQ)

	Low	2	3	4	High	H–L	Alpha
Low	1.269*	1.800**	1.140*	0.492	–0.151	–1.419***	–0.729*
	(1.92)	(2.28)	(1.85)	(0.92)	(–0.37)	(–2.59)	(–1.91)
Median	1.400*	1.230	0.899	0.583	–0.637	–2.037***	–1.615***
	(1.84)	(1.58)	(1.34)	(0.79)	(–1.03)	(–3.41)	(–3.08)
High	1.273*	0.337	0.191	–0.217	–1.373**	–2.646***	–2.119***
	(1.66)	(0.47)	(0.30)	(–0.35)	(–2.32)	(–3.84)	(–3.44)
High–Low	0.005	–1.463***	–0.948**	–0.708	–1.222**	–1.227*	–1.391**
	(0.01)	(–3.00)	(–2.33)	(–1.58)	(–2.18)	(–1.67)	(–1.98)

Internet Appendix D. Turnover–Based Disagreement

In this section, we examine the relation between the turnover–based disagreement proxy and future cryptocurrency returns. We begin by constructing the daily turnover–based disagreement measure, ΔTO, over the period from January 2019 to June 2024. Each day, cryptocurrencies are sorted into decile portfolios based on ΔTO, and we compute the one–day–ahead equal–weighted and value–weighted average excess returns for each decile. To assess the cross–sectional relation between ΔTO and future returns, we form a long–short portfolio that takes a long position in the lowest decile of ΔTO and a short position in the highest decile of ΔTO.

Table ID1 reports the average daily excess returns for each ΔTO–sorted decile portfolio, as well as for the long–short portfolio. Table ID1 also reports abnormal returns (alphas) estimated using the cryptocurrency capital asset pricing model (CCAPM) and the three–factor model. As shown in Table ID1, for equal–weighted (value–weighted) portfolios, the excess return spread between the highest and lowest ΔTO deciles is –0.36% (–0.31%), with a t–value of –4.74 (–3.42). After controlling for the three well–known pricing factors, the long–short portfolio that shorts cryptocurrencies in the highest ΔTO decile and goes long cryptocurrencies in the lowest ΔTO decile earns an average daily abnormal return of 0.40% (0.33%) on an equal–weighted (value–weighted) basis, with a t–value of 6.22 (3.20). These results indicate that turnover–based disagreement exhibits strong predictive power for daily cryptocurrency returns, consistent with the findings of Garfinkel, Hsiao, and Hu (2025).

We next examine whether turnover–based disagreement predicts returns at longer horizons. To this end, we construct a weekly ΔTO measure over the same sample period and conduct a univariate portfolio analysis. Table ID2 reports the results. In contrast to the daily evidence, both the return and alpha spreads for weekly ΔTO–sorted portfolios are statistically insignificant, indicating that turnover–based disagreement does not exhibit predictive power for cryptocurrency returns at the weekly frequency.

Table ID1. Change in Turnover and Cryptocurrency Returns: Daily Frequency

This table reports excess returns and alphas on different cryptocurrency deciles sorted based on turnover–based disagreement measure ΔTO and formed on a daily basis between January 2019 and June 2024. Each day, cryptocurrencies are sorted into decile portfolios by ΔTO. Portfolio Low is the portfolio of cryptocurrencies with the lowest ΔTO, while cryptocurrencies with the highest ΔTO comprise portfolio High. All equal–weighted and value–weighted excess returns and alphas are measured one day ahead of the portfolio formation period. Excess return is the return in excess of the risk–free rate. Alpha is the intercept from a time–series regression of daily excess returns on the factors of alternative models: the CCAPM and the three–factor model proposed by Liu, Tsyvinski, and Wu (2022). This table also reports differences between portfolios High and Low. All excess returns and alphas are reported in percentage terms. Newey–West (1987) adjusted t-statistics are presented in parentheses. $*$, $**$, and $***$ indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal–Weighted			Value–Weighted		
	Excess Return	CCAPM	3–factor	Excess Return	CCAPM	3–factor
Low	0.126	–0.126**	–0.273***	0.081	–0.136*	–0.184**
	(1.52)	(–2.32)	(–3.87)	(0.90)	(–1.88)	(–2.06)
2	0.080	–0.217***	–0.348***	0.073	–0.232***	–0.326***
	(0.92)	(–4.39)	(–5.77)	(0.82)	(–4.49)	(–5.18)
3	0.200**	–0.095*	–0.212***	0.084	–0.225***	–0.299***
	(2.28)	(–1.78)	(–3.19)	(0.95)	(–4.40)	(–4.77)
4	0.126	–0.162***	–0.300***	0.136*	–0.168***	–0.262***
	(1.54)	(–3.63)	(–5.18)	(1.65)	(–3.95)	(–4.48)
5	0.165**	–0.114***	–0.264***	0.179**	–0.111***	–0.195***
	(2.11)	(–2.57)	(–4.81)	(2.25)	(–2.56)	(–3.25)
6	0.128*	–0.143***	–0.282***	0.111	–0.173***	–0.237***
	(1.67)	(–3.43)	(–5.20)	(1.43)	(–4.39)	(–4.90)
7	0.192**	–0.081*	–0.218***	0.149*	–0.132***	–0.235***
	(2.47)	(–1.72)	(–3.60)	(1.86)	(–2.81)	(–3.97)

8	0.280***	0.007	–0.152**	0.135	–0.151***	–0.220***
	(3.24)	(0.13)	(–2.10)	(1.58)	(–2.87)	(–3.24)
9	0.305***	0.009	–0.146**	0.274**	–0.047	–0.141
	(3.47)	(0.16)	(–2.20)	(2.51)	(–0.71)	(–1.55)
High	–0.234**	–0.504***	–0.677***	–0.228***	–0.456***	–0.518***
	(–2.33)	(–6.66)	(–8.66)	(–2.81)	(–7.46)	(–6.91)
High–Low	–0.361***	–0.377***	–0.404***	–0.309***	–0.320***	–0.334***
	(–4.74)	(–5.06)	(–6.22)	(–3.42)	(–3.53)	(–3.20)

Table ID2. Change in Turnover and Cryptocurrency Returns: Weekly Frequency

This table reports the average weekly excess returns and alphas on univariate portfolios of cryptocurrencies sorted by the change in turnover (ΔTO). Each week, cryptocurrencies are sorted into decile portfolios based on the change in turnover (ΔTO), we then calculate the excess return in the subsequent week. This table reports equal–weighted and value–weighted portfolio sorts. Excess return is the return in excess of the risk–free–rate. Alpha is the intercept from a time–series regression of weekly excess returns on the three–factor model (Liu et al., 2022). Newey and West (1987) adjusted t–statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%,5%, and 1% level, respectively. The sample period is from January 2019 to June 2024.

	Equal–Weighted			Value–Weighted		
	Excess Return	CCAPM	3–factor	Excess Return	CCAPM	3–factor
Low	0.521	–0.436	–0.433	0.108	–0.566*	–0.478
	(0.85)	(–1.18)	(–1.27)	(0.25)	(–1.68)	(–1.42)
2	1.040	–0.096	–0.099	1.059	–0.060	–0.075
	(1.55)	(–0.27)	(–0.29)	(1.49)	(–0.15)	(–0.20)
3	0.897	–0.191	–0.164	0.736	–0.435	–0.439
	(1.34)	(0.50)	(–0.46)	(1.14)	(–1.31)	(–1.38)
4	1.063	–0.087	–0.077	0.759	–0.266	–0.206
	(1.56)	(–0.24)	(–0.22)	(1.23)	(–0.79)	(–0.63)
5	0.823	–0.279	–0.267	1.350*	0.128	0.109
	(1.21)	(–0.77)	(–0.79)	(1.73)	(0.34)	(0.33)
6	1.306*	0.151	0.161	1.231*	0.038	0.228
	(1.71)	(0.39)	(0.44)	(1.94)	(0.14)	(0.75)
7	1.338**	0.307	0.293	1.073*	–0.023	0.059
	(2.00)	(0.77)	(0.78)	(1.72)	(–0.07)	(0.17)
8	0.739	–0.323	–0.320	1.138	0.072	0.125
	(1.15)	(–1.00)	(–1.01)	(1.60)	(0.16)	(0.27)

9	1.317*	0.091	0.216	0.838	–0.439	–0.410
	(1.80)	(0.22)	(0.53)	(1.13)	(–1.07)	(–1.03)
High	0.140	–0.909**	–1.083***	0.283	–0.785	–1.030**
	(0.22)	(–2.21)	(–2.88)	(0.40)	(–1.56)	(–2.29)
High–Low	–0.381	–0.473	–0.649*	0.175	–0.219	–0.553
	(–1.00)	(–1.28)	(–1.82)	(0.26)	(0.35)	(–0.94)

Internet Appendix E. Robustness Tests

IE1. Transaction Costs

Although the long-short strategy based on MFD earns statistically significant abnormal returns relative to the CCAPM and the three-factor model, transaction costs may materially reduce its real-world profitability. Prior studies document that, in traditional asset classes, many return anomalies become unprofitable once transaction costs are taken into account (Novy-Marx and Velikow, 2016), and even the most effective mitigation techniques often preserve only a small fraction of gross returns (Patton and Weller, 2020). To assess whether the MFD-return relation in the cryptocurrency market survives realistic implementation frictions, we evaluate the net profitability of the MFD-based long-short strategy. The net return of the long-short strategy in week t is calculated as follows:

$$r_t^{net} = \left(\sum\nolimits_{i \in L} w_{i,t}\, r_{i,t} - tc^l \sum\nolimits_{i \in L} \left|w_{i,t} - w_{i,t-1}\right|\right) - \left(\sum\nolimits_{j \in S} w_{j,t}\, r_{j,t} - tc^s \sum\nolimits_{j \in S} \left|w_{j,t} - w_{j,t-1}\right|\right) \tag{IF.1}$$

where $i \in L$ and $j \in S$ denote that the cryptocurrency belongs to the long or short leg, respectively. tc^l and tc^s indicate the transaction cost rates applied to the long and short positions, respectively. Bianchi, Babiak and Dickerson (2022) adopt conservative transaction cost assumptions of 30 basis points for long positions and 40 basis points for short positions. However, their estimates are based on cryptocurrencies covered by CryptoCompare.com, which primarily includes relatively large and liquid cryptocurrencies. Given that our sample includes smaller and potentially less-liquid cryptocurrencies, we consider higher transaction cost scenarios, setting transaction costs to 30 (40), 40 (50), and 50 (60) basis points for the long (short) portfolios, respectively.

Table IE1 reports the results. For equal-weighted (value-weighted) portfolios, the weekly alpha declines from 1.063% (1.828%) before transaction costs to a range of 0.926% (1.690%) to 0.908% (1.680%), depending on the assumed transaction cost level. Importantly, both excess returns and three-factor alphas remain statistically significant across all transaction cost rate assumptions.

Overall, these findings suggest that MFD-based trading strategies are relatively robust to transaction costs and retain economically meaningful abnormal returns even under conservative assumptions about trading frictions.

Table IE1. Transaction Cost

This table reports the impact of transaction costs on the MFD–return performance of cryptocurrency. This table reports the excess return (in %) and the alpha (in %) adjusted by the three–factor model proposed by Liu, Tsyvinski, and Wu (2022)) for the MFD–sorted high–minus–low portfolios. The portfolios are value–weighted and rebalanced weekly. The transaction cost rates are assumed to be constant, that is, (I) 30 bps and 40 bps, (II) 40 bps and 50 bps, and (III) 50 bps and 60 bps for the long and short leg, respectively. Excess returns and alphas are all reported in percentage terms. Newey and West's (1987) adjusted t–statistics are reported in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 2019 to June 2024.

	Excess Return	Alpha	I		II		III	
			Excess Return	Alpha	Excess Return	Alpha	Excess Return	Alpha
Equal-weighted	–1.223***	–1.063**	–1.085**	–0.926**	–1.075**	–0.917**	–1.066**	–0.908**
	(–2.60)	(–2.53)	(–2.31)	(–2.21)	(–2.29)	(–2.19)	(–2.27)	(–2.17)
Value-weighted	–2.022***	–1.828***	–1.879***	–1.690***	–1.874***	–1.685***	–1.868***	–1.680***
	(–3.29)	(–3.54)	(–3.03)	(–3.24)	(–3.02)	(–3.23)	(–3.00)	(–3.21)

IE2. Alternative Measures of MFD

To address concerns about the potential sensitivity of our results to the specific dimensions of the incomplete information set and the machine learning architecture choices used in MFD construction, we first perform a robustness check by altering the number of characteristics in the incomplete information set to 20, while keeping all other aspects of the MFD construction unchanged. Then, we examine whether our findings remain robust when using alternative gradient boosted tree model (XGBM).

We replicate the univariate portfolio analysis using the alternative MFD measures. As Table IE2 shows, the return and three-factor alpha spreads on MFD-sorted portfolios remain economically and statistically significant. Specifically, for the alternative MFD with a 20-dimensional incomplete information set, the return spread on equal-weighted MFD-sorted portfolios is –1.11% per week, and statistically significant at the 5% level with a t-value of –2.30. The three-factor alpha spread on equal-weighted MFD-sorted portfolios is –0.95%, and statistically significant at the 5% level with a t-value of –2.17. Notably, the results are even more pronounced for value-weighted portfolios, where the return (three-factor alpha) spread on value-weighted MFD-sorted portfolios is –1.48% (–1.05%), and statistically significant at the 5% level. Similarly, the negative relation between MFD and future returns is quite clear from both excess returns and alpha differences between portfolios High and Low when MFD is constructed by XGBM.

Table IE2. Alternative MFD measure

This table reports the average weekly spread return between the highest and lowest decile of univariate portfolios based on MFD for alternative investor–specific incomplete information set for random forest regression (RF) and gradient boosted regression trees (XGBM). Each week, cryptocurrencies are grouped into deciles based on their MFD in the past week. Portfolio Low is the portfolio of cryptocurrencies with the lowest alternative MFD measure, while cryptocurrencies with the highest alternative MFD measure comprise portfolio High. All equal–weighted and value–weighted excess returns are one week ahead of the portfolio formation period. Besides, this table also reports alphas (adjusted using the three–factor model proposed by Liu, Tsyvinski, and Wu (2022)) for each of the MFD–sorted high–minus–low portfolios. For RF, we configured a 20–dimensional investor–specific incomplete information set while keeping all other hyperparameters unchanged. For XGBM, we use the following hyperparameters: L1/L2 regularization (0.03/0.02); the learning rate (0.01); feature sample by node/by level/by tree (0.25/0.5/0.25); minimum child weight (5); number of investors (100), maximum depth (4), number of tree (1000), subsample (0.25), and investor–specific incomplete information set (24 dimensions). Newey–West (1987) adjusted t–statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal–Weighted		Value–Weighted	
	Alternative RF	XGBM	Alternative RF	XGBM
Low	1.070	1.113	0.979	1.343*
	(1.48)	(1.48)	(1.31)	(1.74)
2	1.601**	1.179	2.010**	1.359*
	(2.17)	(1.62)	(2.25)	(1.77)
3	1.168*	1.283*	1.378**	0.815
	(1.67)	(1.83)	(2.08)	(1.31)
4	0.931	0.873	1.368**	1.136*
	(1.32)	(1.24)	(2.01)	(1.77)
5	1.304*	0.831	1.491**	0.975
	(1.90)	(1.18)	(1.96)	(1.41)
6	0.696	1.360**	0.892	1.309*
	(1.05)	(1.98)	(1.29)	(1.77)

7	0.855	0.945	0.878	0.566
	(1.27)	(1.42)	(1.22)	(0.84)
8	0.793	0.874	0.435	0.508
	(1.29)	(1.47)	(0.72)	(0.92)
9	0.572	0.597	0.207	0.858
	(0.95)	(1.00)	(0.39)	(1.43)
High	–0.043	–0.126	–0.501	–0.513
	(–0.07)	(–0.19)	(–0.80)	(–0.76)
High–Low	–1.113**	–1.239**	–1.480**	–1.856***
	(–2.30)	(–2.43)	(–2.24)	(–2.64)
Alpha	–0.952**	–1.144**	–1.053**	–1.408**
	(–2.17)	(–2.52)	(–2.06)	(–2.43)

IE3. Subsamples

A key challenge in cryptocurrency research is the relatively short sample period of available data, as noted by Liu and Tsyvinski (2021). This limitation is particularly relevant in our study, where the main out-of-sample analysis period spans from January 2019 to June 2024. While this window helps reduce concerns related to market immaturity and investor learning, the short-lived nature of certain cryptocurrencies may still raise questions regarding the generalizability of our findings.

Following Liu and Tsyvinski (2021), we split our sample period into two equal subsamples to examine the temporal stability of our results. Table IE3 reports the results. We find that the MFD-return relation is consistent across the two subperiods, suggesting that the MFD-return relation is not driven by a specific market regime or time-specific effect. Furthermore, we extend the sample period from 2019-2024 to 2017-2024 to re-examine the robustness of our findings. Table IE3 shows that cryptocurrencies with high MFD continue to earn significantly lower future returns, consistent with our main results.

Table IE3. Different Sample Periods

This table shows the results from the univariate portfolio analysis after extending the sample and splitting the sample into halves. First, the sample period is extended to 2017–2024. Then, the whole sample period is split into 2019.01–2021.12 and 2022.01–2024.06. Equal–weighted and value–weighted one–week–ahead excess returns of different cryptocurrency deciles are sorted based on MFD and formed on a weekly basis. This table also reports the alphas (adjusted using the three–factor model proposed by Liu, Tsyvinski, and Wu (2022)) for each of the MFD–sorted high–minus–low portfolios. Newey–West (1987) adjusted t-statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal–Weighted			Value–Weighted		
	Sample Extension	Sample Split		Sample Extension	Sample Split	
	2017–2024	2019–2021	2022–2024	2017–2024	2019–2021	2022–2024
Low	1.744**	2.207**	–0.036	1.902**	2.540**	0.136
	(2.10)	(1.99)	(–0.05)	(2.42)	(2.47)	(0.18)
2	2.109**	2.157**	0.027	1.627**	2.141**	–0.917
	(2.39)	(2.08)	(0.03)	(2.07)	(2.05)	(–1.19)
3	2.060**	2.510**	0.050	2.850***	3.882**	0.714
	(2.56)	(2.15)	(0.06)	(3.08)	(2.51)	(0.79)
4	1.341**	1.961*	0.005	1.242	1.886*	0.223
	(2.01)	(1.86)	(0.01)	(1.64)	(1.72)	(0.27)
5	1.498**	1.538	0.004	1.664**	2.040**	0.457
	(2.25)	(1.64)	(0.00)	(2.29)	(2.13)	(0.52)
6	1.509**	1.770*	–0.123	1.557**	1.603*	0.337
	(1.98)	(1.75)	(–0.15)	(2.15)	(1.76)	(0.51)
7	1.588**	2.008*	–0.377	1.298	0.921	–0.768
	(2.11)	(1.87)	(–0.48)	(1.29)	(1.01)	(–0.94)
8	1.056	2.037**	–0.442	0.631	1.041	–0.372

	(1.44)	(2.32)	(–0.52)	(0.85)	(1.13)	(–0.55)
9	0.626	1.289	–0.4992	0.554	0.745	0.273
	(0.88)	(1.60)	(–0.56)	(0.84)	(0.94)	(0.72)
High	0.802	0.852	–1.098	–0.076	–0.128	–1.098
	(0.94)	(0.99)	(–1.16)	(–0.10)	(–0.18)	(–1.04)
High–Low	–0.942**	–1.355*	–1.062**	–1.978***	–2.668***	–1.234*
	(–1.98)	(–1.72)	(–2.56)	(–2.85)	(–2.83)	(–1.72)
Alpha	–1.152**	–0.754	–1.220***	–1.656***	–1.825**	–1.741***
	(–2.33)	(–1.11)	(–3.25)	(–2.68)	(–2.15)	(–3.18)

IE4. Alternative Screens

Our empirical analyses in the above sections restrict the sample to cryptocurrencies with a market capitalization larger than \$1 million. To examine the robustness of our findings to alternative size-based screens, we re-estimate our analysis using two additional thresholds: cryptocurrencies with market capitalization above \$0.5 million and above \$5 million. We keep all other criteria unchanged.

Table IE4 reports the results. Columns (1) and (4) report the results for the \$0.5-million screen. We find that the return spread for MFD-sorted portfolios remains significantly negative, particularly for the value-weighted portfolios. However, both the return and alpha spreads are somewhat attenuated when compared to the baseline results in Table 3.

Columns (2) and (5) report the results using a \$5 million size screen. In this more restrictive sample, the return and alpha differences become larger in magnitude compared to the \$0.5 million screen and remain statistically significant. These findings suggest that, while firm size does influence the strength of the MFD-return relationship, the negative MFD-return relation is robust across different size thresholds. More interestingly, the negative MFD-return relation is stronger for larger cryptocurrencies[18].

Considering that stablecoins may exhibit different disagreement dynamics, we

[18] One plausible explanation is that smaller cryptocurrencies are more prone to extreme price fluctuations. We have discussed it in the Internet Appendix A.

perform a robustness test by excluding stablecoins. Columns (3) and (6) in Table IE4 reports the results using sample without stablecoins. The excess return and alpha differences between portfolio High and Low increase slightly, and all remain statistically significant. These findings suggest that stablecoins in our sample have a negligible effect on the relation between MFD and cryptocurrency returns.

Table IE4. Alternative Screens

This table reports the results of the univariate portfolio analysis with alternative screens. We first limit our sample to cryptocurrencies with a capitalization larger than $0.5 million at the end of each week. Then, we limit our sample to cryptocurrencies with a capitalization larger than $5 million at the end of each week. Finally, we exclude stablecoins and keep all cryptocurrencies in our sample with a capitalization larger than 1 million at the end of each week. To ensure adequate liquidity, we also exclude cryptocurrencies whose price is lower than $0.1, trading volume is lower than $0.1 million, and that do not have at least 52 weeks of trading history. We sort cryptocurrencies into decile portfolios based on MFD on a weekly basis. Cryptocurrencies with the lowest MFD are categorized into portfolio Low. Cryptocurrencies with the highest MFD are categorized in portfolio High. This table also reports the alphas (adjusted using the three–factor model proposed by Liu, Tsyvinski, and Wu (2022)) for each of the MFD–sorted high–minus–low portfolios. Newey–West (1987) adjusted t-statistics are presented in parentheses. *, **, and *** indicate the significance at the 10%, 5%, and 1% levels, respectively.

	Equal–Weighted			Value–Weighted		
	Cap>0.5 M	Cap>5 M	Exclude Stablecoins	Cap>0.5 M	Cap>5 M	Exclude Stablecoins
Low	1.221*	1.313*	1.240*	1.424**	1.554**	1.460**
	(1.72)	(1.86)	(1.73)	(2.10)	(2.29)	(2.14)
2	1.185*	1.215*	1.310*	0.748	0.666	0.812
	(1.72)	(1.72)	(1.81)	(1.07)	(0.95)	(1.15)
3	1.391*	1.263*	1.197	2.454**	1.965***	1.813**
	(1.81)	(1.74)	(1.60)	(2.55)	(2.61)	(2.40)
4	1.109	1.097	1.118	1.084	1.768	2.360**
	(1.64)	(1.54)	(1.55)	(1.57)	(1.62)	(1.96)
5	0.861	0.946	0.882	1.205*	1.306*	1.217*
	(1.31)	(1.42)	(1.28)	(1.81)	(1.95)	(1.72)
6	0.850	0.987	1.106	0.995*	1.177**	1.346*
	(1.25)	(1.42)	(1.50)	(1.68)	(1.97)	(1.91)

7	0.924	0.844	0.952	0.244	0.477	0.238
	(1.33)	(1.21)	(1.27)	(0.39)	(0.75)	(0.36)
8	0.771	0.812	0.916	0.382	0.446	0.183
	(1.23)	(1.25)	(1.34)	(0.65)	(0.79)	(0.28)
9	0.570	0.547	0.878	0.510	0.328	0.716
	(0.93)	(0.92)	(1.30)	(1.07)	(0.66)	(0.94)
High	–0.002	0.081	0.008	–0.601	–0.642	–0.733
	(–0.00)	(0.13)	(0.01)	(–0.96)	(–1.05)	(–0.97)
High–Low	-1.223^{***}	-1.232^{***}	-1.232^{***}	-2.025^{***}	-2.196^{***}	-2.194^{***}
	(–2.60)	(–2.64)	(–2.72)	(–3.30)	(–3.60)	(–3.53)
Alpha	-1.062^{**}	-1.099^{**}	-1.147^{***}	-1.794^{***}	-2.003^{***}	-2.349^{***}
	(–2.54)	(–2.57)	(–2.77)	(–3.52)	(–4.10)	(–4.46)